

**Visokošolski učbenik**

**OPTIMIZACIJA LOGISTIČNIH PROCESOV**

**Dejan Dragan**

Fakulteta za logistiko

Celje, 2010

**Naslov:**

Optimizacija logističnih procesov

**Izdajatelj:**

Fakulteta za logistiko Univerza v Mariboru

**Avtor:**

dr. Dejan Dragan

**Pomoč pri urejanju:**

Tea Vizinger

**Recenzenta:**

Doc.dr. Damir Vrančič

Doc. dr. Đani Juričič

# KAZALO

1 UVOD V OPTIMIZACIJO .....	7
2 PRIMERI OPTIMIZACIJE.....	11
2.1 Nekaj splošnih primerov iz optimizacije.....	11
2.2 Nekateri primeri optimizacije v logistiki in transportu .....	14
2.2.1 <i>Transportni problem</i> .....	14
2.2.2 <i>Problem načrtovanja logističnih sistemov</i> .....	18
2.2.3 <i>Ostali možni problemi optimizacije logističnih sistemov</i> .....	21
3 TEMELJNI KONCEPTI OPTIMIRANJA .....	23
3.1 Modeliranje optimizacijskih modelov.....	24
3.2 Ekstremi funkcij ene in več spremenljivk.....	27
3.2.1 <i>Ekstremi funkcij ene spremenljivke</i> .....	29
3.2.2 <i>Ekstremi funkcij več spremenljivk</i> .....	30
3.3 Lokalni in globalni minimumi ciljnih funkcij.....	32
4 PREGLED GLAVNIH OPTIMIZACIJSKIH METOD .....	33
4.1 Razdelitev in lastnosti optimizacijskih metod.....	33
4.2 Linearno programiranje.....	37
4.3 Celoštevilo programiranje .....	41
4.3.1 <i>Modeliranje celoštevilskih problemov</i> .....	43
4.3.2 <i>Reševanje celoštevilskih problemov</i> .....	48
4.4 Dinamično programiranje .....	48
4.5 Mrežna optimizacija.....	49
5 NEKATERE MATEMATIČNE OSNOVE .....	53
5.1 Kvadratne forme matrik .....	53
5.2 Odvodi funkcij, gradient in Hessejeva matrika.....	54
5.3 Konveksnost funkcij.....	56
5.4 Klasična analitična optimizacija in analitičen pristop iskanja ekstremov.....	60
5.4.1 <i>Stacionarne točke</i> .....	60
5.4.2 <i>Razlika med analitičnim in numeričnim iskanjem ekstrema</i> .....	63
5.4.3 <i>Pretvorba na problem brez omejitev z eliminacijo spremenljivk</i> .....	64
5.4.4 <i>Pretvorba na problem brez omejitev z Lagrangeovimi multiplikatorji</i> .....	65
6 NELINEARNA NUMERIČNA OPTIMIZACIJA BREZ OMEJITEV.....	70
6.1 Modeliranje nelinearnih problemov .....	74
6.2 Enodimenzionalna optimizacija .....	79
6.2.1 <i>Lastnosti optimizacijskih algoritmov</i> .....	80
6.2.2 <i>Metoda zlatega reza</i> .....	81
6.2.3 <i>Newtonova metoda za eno spremenljivko</i> .....	88
6.2.4 <i>Sekantna metoda</i> .....	94
6.3 Multidimenzionalna optimizacija.....	97
6.3.1 <i>Osnovne metode</i> .....	97

6.3.1.1	Direktna Hook-Jeevesova metoda vzorčnih premikov .....	98
6.3.1.2	Direktna Nelder-Meadova simpleksna metoda .....	110
6.3.1.3	Cauchyeva metoda najstrmejšega spusta (gradientna metoda I. reda).....	127
6.3.1.4	Newtonova metoda (gradientna metoda II. reda) .....	137
6.3.1.5	Marquardtova metoda.....	142
6.3.2	<i>Sofisticirane metode</i> .....	143
6.3.2.1	Powellova metoda konjugiranih smeri .....	143
6.3.2.2	Metoda konjugiranih gradientov .....	161
6.3.2.3	Quasi-Newtonove metode.....	172
7	GEOMETRIJSKO PROGRAMIRANJE .....	186
7.1	Geometrijsko programiranje brez omejitev.....	187
7.1.1	<i>Izpeljava pristopa z diferencialnim računom</i> .....	189
7.1.2	<i>Postopek reševanja</i> .....	193
7.1.3	<i>Ozadje Cauchyeve neenakosti in reševanje problemov višje stopnje</i> .....	207
7.1.3.1	Pristop k reševanju problemov višje stopnje z odvajanjem logaritmske dualne ciljne funkcije .....	209
7.1.3.2	Pristop k reševanju problemov višje stopnje z odvajanjem dualnih odločitvenih spremenljivk .....	216
7.2	Geometrijsko programiranje z omejitvami .....	223
7.2.1	<i>Aplikativni primer GP z omejitvami</i> .....	231
8	ZBIRKA REŠENIH NALOG .....	244
8.1	Primeri analitične optimizacije.....	244
8.1.1	<i>Klasičen način reševanja</i> .....	245
8.1.2	<i>Reševanje z Lagrangeovimi multiplikatorji za probleme z omejitvami</i> .....	272
8.2	Primeri enodimenzionalne numerične optimizacije .....	284
8.2.1	<i>Metoda zlatega reza</i> .....	284
8.2.2	<i>Newtonova metoda za eno spremenljivko</i> .....	292
8.2.3	<i>Sekantna metoda</i> .....	299
8.3	Primeri multidimenzionalne numerične optimizacije z osnovnimi metodami.....	301
8.3.1	<i>Direktna Hook-Jeevesova metoda vzorčnih premikov</i> .....	301
8.3.2	<i>Cauchyeva metoda najstrmejšega spusta (gradientna metoda I. reda)</i> .....	313
8.3.3	<i>Newtonova multidimenzionalna metoda</i> .....	329
8.3.4	<i>Marquardtova metoda</i> .....	339
8.4	Primeri multidimenzionalne numerične optimizacije s sofisticiranimi metodami.	344
8.4.1	<i>Powellova metoda</i> .....	344
8.4.2	<i>Metoda konjugiranih gradientov</i> .....	360
8.4.3	<i>Quasi-Newtonove metode</i> .....	369
8.5	Primeri geometrijskega programiranja.....	380
8.5.1	<i>Geometrijsko programiranje brez omejitev</i> .....	380
8.5.2	<i>Geometrijsko programiranje z omejitvami</i> .....	402
9	PRILOGE .....	407
9.1	Programi v Matlabu ali Scilabu.....	407
9.1.1	<i>Algoritem zlatega reza (Scilab)</i> .....	409
9.1.2	<i>Algoritem Newtonove metode za eno spremenljivko (Scilab)</i> .....	410



9.1.3 Algoritem sekantne metode (Scilab).....	410
9.1.4 Algoritem Hooke-Jeevesove metode vzorčnih premikov (Matlab).....	411
9.1.5 Algoritem Nelder-Meadove simpleksne metode (Matlab).....	412
9.1.6 Algoritem metode najstrmejšega spusta brez uporabe Symbolic toolboxa (Matlab).....	415
9.1.7 Algoritem metode najstrmejšega spusta s pomočjo dodatne uporabe Symbolic toolboxa (Matlab).....	420
9.1.8 Algoritem Newtonove metode za več spremenljivk (Matlab).....	423
9.1.9 Algoritem DFP QuasiNewtonove metode za več spremenljivk s pomožno analitično enodimenzionalno optimizacijo v Symbolic toolboxu (Matlab).....	426
LITERATURA.....	428

## ***Spremna beseda***

To delo je zasnovano kot začasna skripta v pripravi, ki obravnava snov predmeta »Optimizacija logističnih procesov«, ki se predava na podiplomskem študiju Fakultete za logistiko.

Delo pokriva kratke teoretične osnove, zbirko analitično rešenih nalog, ter primere uporabe nekaterih računalniških programov. Mišljeno je kot začasno gradivo, ki naj bi ga študentje uporabljali poleg ostalih gradiv (prosojnic, itn) pri študiju tega predmeta, dokler ne bo v celoti izdelan vsestransko obsegajoč in še nekoliko bolj razširjen učbenik iz tega predmeta.

Pri delu s študenti je bilo ugotovljeno, da potrebujejo veliko računskih vaj pri osvajanju snovi s tega področja. To pomeni, da morajo preizkusiti metode za reševanje problemov na kar največjem naboru nalog, da so res učinkovito kos izpitnim vprašanjem in sposobni kvalitetnega razumevanja snovi.

V okviru tega se pričakuje vsaj osnovno predznanje študentov s področja visokošolske matematike, ponekod pa tudi vsaj osnovno razumevanje iz osnov programiranja.

V gradivu je takorekoč vsaka naloga opremljena ne le s končno rešitvijo, pač pa tudi s celotnim postopkom reševanja. Pri tem je nazorno prikazan prav vsak korak računanja, z jasnim ciljem, da lahko sleherni študent ujame ritem razlage in slednjemu brez težav sledi do konca izračunov.

Poudarimo še, da je to delo šele prva verzija skripte v pripravi, zato ni izključena možnost določenih tiskarskih in podobnih napak. Za morebitne napake se bralcu že vnaprej opravičujemo in bomo hvaležni za vsak kritičen komentar.

# 1 UVOD V OPTIMIZACIJO

V življenju je potrebno nenehno sprejemati odločitve. Seveda si želimo, da bi bili rezultati sprejetih odločitev čim boljši. Pri tem imamo ponavadi na izbiro več možnosti, ki vodijo do zastavljenega cilja, med njimi pa bi hoteli izbrati najbolj optimalno možnost.

Pravimo, da se v družbi nenehno odvija določena optimizacija. Tako potujemo po najkrajši poti, letalska podjetja razporejajo osebje in letala na takšen način, da se minimizirajo stroški, v industriji se proizvodni procesi oblikujejo tako, da se maksimizira produktivnost, doseči želimo optimalno proizvodnjo, ki bo usklajena s povpraševanjem, naložiti želimo tovarnjak, da bomo nanj ob omejitvah glede mase in prostornine spravili čim več blaga, itn.

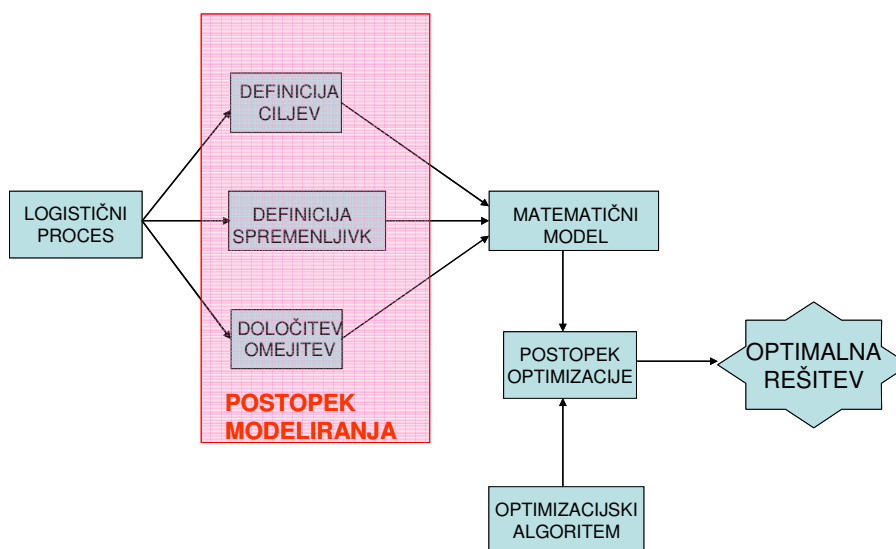
Tudi v naravi se nenehno odvija optimizacija. Tako fizikalni sistemi težijo k stanju najmanjše energije. Molekule v izoliranih kemijskih sistemih medsebojno delujejo toliko časa, dokler skupna potencialna energija njihovih elektronov ne doseže minimuma. Tudi svetlobni žarki se širijo v prostor na takšen način, da se minimizira čas širjenja žarka, oziroma le-ta potuje po najkrajši poti.

Postopki optimiranja so temeljno orodje v znanstvenem odločanju in v analizi fizikalnih sistemov. Da pa bi se sploh lahko govorilo o optimiranju, je najprej potrebno identificirati cilj, oziroma kvantitativno mero obnašanja proučevanega sistema. Cilj je lahko zaslužek, čas, potencialna energija ali neka druga veličina ali kombinacija veličin, katere se lahko predstavi z numerično vrednostjo. Cilj zavisi od določenih karakteristik sistema, imenovanih spremenljivke. Cilj postopka optimiranja je določiti kombinacijo vrednosti spremenljivke, ki bo najboljše izpolnila oz. dosegla cilj.

Pogosto so vrednosti spremenljivk na nek način omejene, saj realne fizikalne veličine, ki jih spremenljivke predstavljajo, lahko zavzemajo le vrednosti na določenem omejenem intervalu. Npr., spremenljivka, ki predstavlja zakasnitev paketa v računalniški mreži, ne more zavzeti negativne vrednosti. Poleg fizikalnih limitacij, ki predstavljajo omejitve, so slednje lahko tudi funkcionalne, če predstavljajo skrajne limitacije obnašanja sistema, ki ga želimo optimirati. Lahko pa so omejitve tudi posledica skrajne razpoložljivosti različnih kapacitet, bodisi proizvodnih, transportnih, itn.

Proces definiranja cilja, spremenljivk in morebitnih omejitev za dani problem se imenuje postopek modeliranja (osenčeno polje na sliki 1) [4, 25]. Velja namreč, da logistični sistem v celoti, ali tudi njegove podsisteme proučujemo tako, da kreiramo modele (pod)sistemov in nato na modelih preizkušamo reševanje problemov. Model nam pri tem pomeni poenostavljeno predstavo realnosti. Realni logistični sistemi so izjemno

kompleksni in s tem zahtevni za proučevanje. Pri izgradnji modela, ki je bolj ali manj uspešna predstavitev realnosti, stremimo za tem, da naj bo čim enostavnejši in naj vključuje le zares pomembne podrobnosti. V modelu nikoli ne moremo upoštevati prav vseh možnosti, spremenljivk in njihovih povezav, zato je model vedno le približek realnega problema, včasih bolj, včasih pa manj uspešen.



Slika 1: Postopek optimiranja

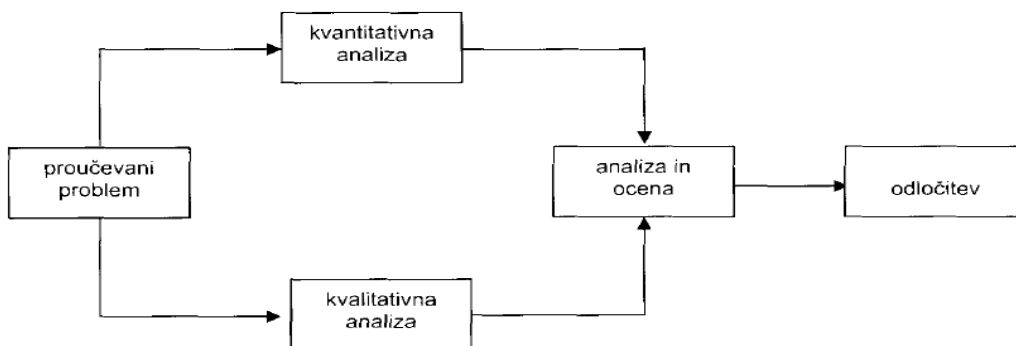
Oblikovanje dobrega modela je prvi korak, pogosto pa tudi najpomembnejši, v procesu optimiranja [25]. Če je model preveč enostaven, potem ne bo dobro opisoval sistema, ki ga želimo optimirati. Če pa je preveč kompliciran, pa lahko postane postopek optimiranja pretežek za reševanje.

Po opravljeni definiciji modela se lahko uporabijo postopki optimizacije z namenom iskanja optimalne rešitve problema (glej sliko 1). Slednji se v praksi realizirajo z implementacijo algoritma optimiranja na računalniku. Žal ne obstaja enoten algoritem optimiranja, ki bi bil primeren za vse probleme. Tako v praksi obstaja veliko število algoritmov, izmed katerih pa je za vsak primeren le za določeno skupino optimizacijskih problemov.

V procesu odločanja imamo na razpolago več ali manj podatkov, ki so lahko kvalitativni ali kvantitativni (slika 2) [25]. Pri tem pod pojmom “kvantitativni” razumemo vse tisto, kar lahko merimo in nato izrazimo s konkretnimi števili. Kvantitativne metode optimiranja so torej takšne metode, s katerimi lahko preračunavamo merljive podatke in izračunavamo možnosti, spet izražene s števili. Te metode uporabljajo različne matematične operacije od najenostavnejših (seštevanje, množenje, itn) do izjemno zahtevnih (verjetnostni račun, diferencialni račun, itn). Rezultat optimiranja je izražen v

numerični obliki in s tem predan odločevalcu v uporabo. V konkretnih praktičnih primerih imamo opraviti z veliko spremenljivkami z mnogo različnimi podatki, zato bi bilo “ročno” računanje zahtevno, dolgotrajno in mukotrpno. Z uporabo računalnikov pa so postali tudi numerični izračuni neprimerljivo hitrejši, zato je uporaba kvantitativnih metod optimiranja vedno bolj smiselna in učinkovita.

Marsikje pa kvantitativnih metod ne moremo uporabiti, npr. v humanističnih znanostih, kadrovskih vprašanjih, medsebojnih odnosih itn. V vsakem primeru je pri proučevanju pojavov in njihovem analiziranju potrebno kombinirati različne metode, za odločevalca pa je poleg različnih informacij potrebna tudi izkušnost in intuicija.



Slika 2: Kvalitativni in kvantitativni vidik sprejemanja odločitev

V tem delu bomo spoznali nekatere metode optimiranja, ter tudi njihovo uporabo v konkretni logistični praksi. Pri tem bo glavni poudarek knjige namenjen obravnavi postopkov nelinearne optimizacije. V praksi so namreč mnogi problemi optimizacije po naravi nelinearni. To pomeni, da pride v postopku modeliranja bodisi do nelinearne ciljne funkcije, bodisi do nelinearnih omejitev, ali celo do obojega. Seveda postane problem nelinearen že pri eni sami nelinearni omejitvi.

Za razliko od problemov linearne optimizacije, kjer obstajajo univerzalne metode za reševanje linearnih programov (npr. metoda simpleksov [21, 25]), je pri nelinearni optimizaciji problematika reševanja nelinearnih problemov mnogo bolj zapletena. Tako ne obstaja neka univerzalna numerična metoda, ki bi učinkovito reševala vse probleme nelinearne optimizacije. Zato reševanje tovrstnih problemov zahteva poglobljeno razumevanje teoretičnih rezultatov in izbranih numeričnih metod s tega področja [4].

Tipični nelinearni problemi, ki se lahko pojavijo v praksi, so: nelinearni transportni problem, nelinearni problem optimalne razporeditve virov, nelinearni problem

optimalnega izbora asortimana proizvodnje, nelinearni problemi pri optimizaciji proizvodnje, itn [17, 18].

Knjiga je zgrajena tako, da najprej v 2. poglavju spoznamo nekatere tipične primere optimizacije, v 3. poglavju temeljne koncepte optimiranja, ter nato v 4. poglavju izvedemo kratek pregled glavnih optimizacijskih metod.

V 5. poglavju sledi obravnava nekaterih matematičnih osnov, kot so kvadratne forme matrik, odvodi funkcij, konveksnost funkcij, ter temeljni principi klasične analitične optimizacije. Te osnove so nujno potrebne za boljše razumevanje problematike nelinearne (numerične) optimizacije, ki je snov naslednjega, 6. poglavja.

Marsikaterega optimizacijskega problema se namreč ne da rešiti s klasičnim analitičnim reševanjem in tedaj se izkaže uporaba numeričnih optimizacijskih tehnik kot najbolj ustrežna. Reševanje problemov nelinearne optimizacije večkrat imenujemo tudi z imenom »Nelinearno programiranje« [17]. Tej problematiki, ki se osredotoči na obravnavo nelinearnih problemov numerične optimizacije brez omejitev, za eno ali več spremenljivk, je namenjeno jedro tega dela v 6. poglavju.

Naslednje pomembno poglavje, 7. po vrsti, pa obravnava problematiko takoimenovanega geometrijskega programiranja [16, 17, 18, 19]. Gre za metodologijo reševanja posebnega razreda nelinearnih optimizacijskih problemov, ki se izkaže kot izredno koristno pri reševanju določene skupine praktičnih logističnih optimizacijskih problemov.

8. poglavje vključuje zbirko rešenih nalog iz različnih področij nelinearne optimizacije, izmed katerih so bile mnoge zbrane tudi iz preteklih kolokvijev in izpitov. Gre za sklope nalog s področja analitične optimizacije, numerične optimizacije brez omejitev za eno ali več spremenljivk, ter naloge s področja geometrijskega programiranja.

V zadnjem, 9. poglavju, so podane nekatere priloge, s poudarkom na zbirki obravnavanih algoritmov optimiranja, napisanih v programskih jezikih Matlab ali Scilab.

## 2 PRIMERI OPTIMIZACIJE

Naštejmo nekaj tipičnih aplikacij iz različnih inženirskih področij, kjer lahko s pridom uporabimo optimizacijo [21]:

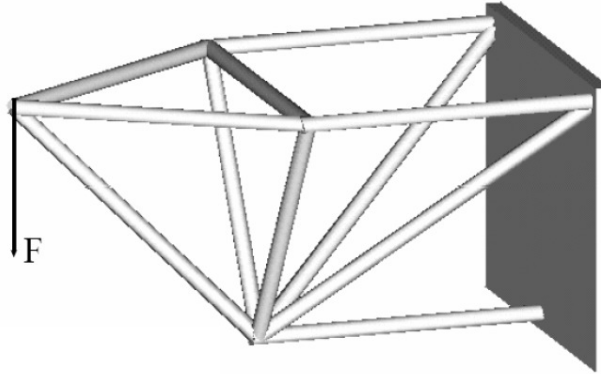
- Načrtovanje struktur zračnih plovil, kjer bo dosežena minimalna teža.
- Iskanje optimalnih trajektorij vesoljskih plovil.
- Načrtovanje struktur v gradbeništvu, kot npr. mostov, dimnikov, predorov, itn, kjer bodo doseženi minimalni stroški.
- Načrtovanje oskrbnih sistemov z vodo, kjer bodo dosežene maksimalne koristi.
- Optimalno načrtovanje mehanskih komponent in strojev.
- Načrtovanje črpalk, turbin in opreme za prenos toplote, kjer bo dosežena maksimalna učinkovitost.
- Optimalno načrtovanje električnih strojev, kot npr. motorjev, generatorjev in transformatorjev.
- Doseganje najkrajše poti trgovskih potnikov pri obiskovanju različnih mest tekom enega potovanja.
- Optimalno planiranje in vodenje proizvodnje.
- Optimalno razvrščanje opravil.
- Optimalno planiranje vzdrževanja in servisiranja opreme z namenom redukcije operativnih stroškov.
- Optimalno upravljanje zalog.
- Optimalna alokacija resursov med različnimi aktivnostmi z namenom maksimizacije koristi in minimizacije stroškov.
- Optimalen nadzor in vodenje čakalnih vrst v produkcijskih linijah z namenom zmanjšanja stroškov.
- Planiranje najboljše poslovne strategije z namenom doseganja maksimalnega profita v prisotnosti konkurence. Itn.

### 2.1 Nekaj splošnih primerov iz optimizacije

V nadaljevanju na kratko ilustrirajmo 4 primere optimizacije: v gradbeništvu, pri reševanju težkih fizikalnih problemov, pri apliciranju visokih tehnologij, ter pri arhitekturi.

## A. Primer optimizacije v gradbeništvu

Pri kompliciranih gradbenih elementih bi lahko prišlo do velikega števila vrednosti spremenljivk in posledično do tako obsežnega računanja optimalnih vrednosti, da brez optimizacije ne bi nikamor prišli (slika 3).

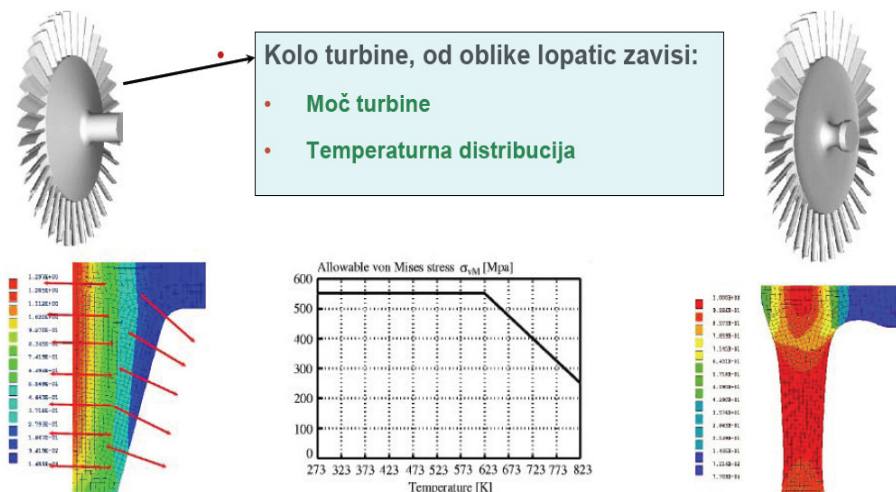


- 23 spremenljivk (11 palic z debelino, 4 stičišča s tremi koordinatami)
- 3 vrednosti za vsako:  $3^{23} \rightarrow 9.4 \cdot 10^{23}$  kombinacij
- Če 1 kombinacija/sekundo  $\rightarrow$  3000 let za računanje

Slika 3: Primer optimizacije v gradbeništvu

## B. Primer optimizacije pri reševanju težkih fizikalnih problemov

Tudi v takšnih primerih je optimizacija pogostokrat nepogrešljiv pripomoček, saj je razumevanje fizike problemov velikokrat tako težavno, da brez optimizacije optimalno in pravilno načrtovanje enostavno ne bi bilo mogoče (slika 4).



Slika4: Primer optimizacije pri reševanju težkih fizikalnih problemov



### C. Primer optimizacije pri apliciranju visokih tehnologij

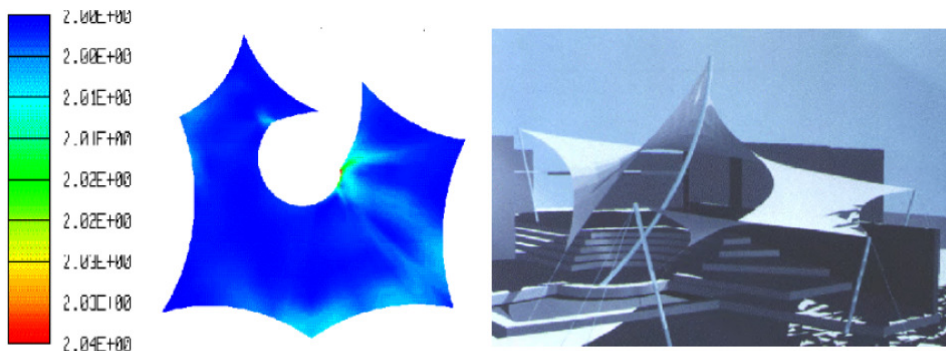


- V izostrenih konkurenčnih razmerah se ne da preživeti, če ni vse res odlično narejeno

- Pri astronautih pa gre res dobesedno za preživetje, če ni vse brezhibno narejeno!!

Slika 5: Primer optimizacije pri apliciranju visokih tehnologij

### D. Primer optimizacije pri arhitekturi



- Prostor načrtovanja: vsi možni robovi membrane
- Cilj:
  - Tehničen: Minimalni stroški in materiali
  - Dejanski: Lep izgled
- Omejitve: Strukturna integriteta

Slika 6: Primer optimizacije v arhitekturi

Kot vidimo na sliki 5, pri apliciranju visokih tehnologij stremimo k popolnosti, brezhibnosti in odličnosti izdelave, kar je mogoče le z uporabo optimizacije. Slednja je nepogrešljiva tudi v arhitekturi, kjer lahko bistveno pripomore k minimizaciji stroškov in materialov, ter lepem izgledu arhitekturnega objekta (slika 6).

## 2.2 Nekateri primeri optimizacije v logistiki in transportu

V nadaljevanju si bomo pogledali še nekatere primere optimizacije v logistiki in transportu. Omejili se bomo na dva tipična problema: transportni problem in problem optimalnega načrtovanja logističnih sistemov. Nato pa se bomo seznanili še z ostalimi možnimi problemi, ki se lahko pojavijo pri optimizaciji logističnih sistemov.

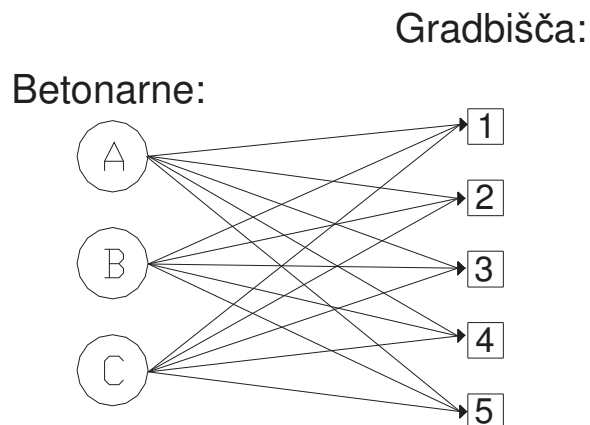
### 2.2.1 Transportni problem

Transportni problem je izjemno pomemben pri organiziranju transportnih storitev [25, 26]. V splošnem imamo opravka z  $m$  izvori (npr. tovarne), odkoder izvira blago, ki se porablja v  $n$  ponorih (npr. potrošna središča). Od izvorov se blago prenaša do ponorov, pri čemer nastajajo prevozni stroški.

Gre za problem, kako naj se dirigira prevoz blaga od izvorov do ponorov, da bodo skupni transportni stroški najmanjši (iskanje optimalne organizacije transporta blaga). Seveda so pri tem iz različnih razlogov, kot so različne oddaljenosti, različni načini prevoza, itn, stroški prevoza na enoto proizvoda odvisni tako od posameznega izvora, iz katerega se vozi blago, kot tudi od posameznega ponora, v katerega se blago vozi.

Problematiko transportnega problema si pogledjmo na naslednjem primeru.

Imamo 3 betonarne ( $m = 3$ ) in 5 gradbišč ( $n = 5$ ), katerih povezave so predstavljene na sliki 7. Zanima nas takšna organizacija transporta, pri kateri bodo cene prevozov betona najmanjše.



Slika 7: Primer transportnega problema

V skladu s sliko 1 najprej opravimo postopek modeliranja, pri čemer dobimo kot rezultat matematični model transportnega procesa.

V prvem koraku postopka modeliranja najprej definiramo vse potrebne veličine:

$c_{ij}$ .....cene prevozov na enoto betona iz betonarne  $i$  na gradbišče  $j$

$x_{ij}$ .....količine betona, pripeljanega iz betonarne  $i$  na gradbišče  $j$

$a_j$ .....količina, ki jo potrebuje gradbišče  $j$

$b_i$ .....količina, ki jo izdelamo v betonarni  $i$

$i = 1, 2, 3, \quad j = 1, 2, 3, 4, 5$

kjer nam količine  $x_{ij}, i = 1, \dots, 3; j = 1, \dots, 5$  predstavljajo takoimenovane **odločitvene spremenljivke**.

Nato predstavimo problem še v obliki naslednje pregledne tabele (tabela 1):

	Gradb.1	Gradb.2	Gradb.3	Gradb.4	Gradb.5	Kapaciteta betonarn
Bet. 1	$x_{11}/c_{11}$	$x_{12}/c_{12}$	...			$b_1$
Bet.2	...					$b_2$
Bet.3						$b_3$
Potrebe gradbišča	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	

Tabela 1: Podatki za primer transportnega problema

V tabeli 1 so vpisane transportne tarife ( $c_{ij}$ ) za vsako relacijo ( $i, j$ ), ter neznane količine betona  $x_{ij}$ , ki naj se transportirajo na teh relacijah. Podane so tudi potrebe posameznih gradbišč  $a_j$ , ter kapacitete posameznih betonarn  $b_i$ .

Torej, določiti je potrebno naslednje neznane količine, pripeljane pri transportu:

$$x_{ij}, i = 1, \dots, 3; j = 1, \dots, 5, \quad x_{ij} \geq 0$$

kjer smo z omejitvami  $x_{ij} \geq 0$  upoštevali dejstvo, da prepeljane količine ne morejo biti negativne.

Kot nam narekuje postopek modeliranja na sliki 1, je potrebno v drugem koraku določiti še vse ostale omejitve v procesu, ki ustrezajo potrebam gradbišč in kapacitet betonarn:

$$\begin{array}{r}
 x_{11} + x_{21} + x_{31} \geq a_1 \\
 x_{12} + x_{22} + x_{32} \geq a_2 \\
 \dots \\
 x_{15} + x_{25} + x_{35} \geq a_5 \\
 \hline
 x_{11} + x_{12} + x_{13} + x_{14} + x_{15} \leq b_1 \\
 \dots \\
 x_{31} + x_{32} + x_{33} + x_{34} + x_{35} \leq b_3 \\
 \\
 x_{ij} \geq 0
 \end{array} \tag{2.1.}$$

Količine  $x_{ij}$ ,  $i=1,\dots,3$ ;  $j=1,\dots,5$  je potrebno določiti na takšen način, da bo kriterijska (ciljna) funkcija stroškov:

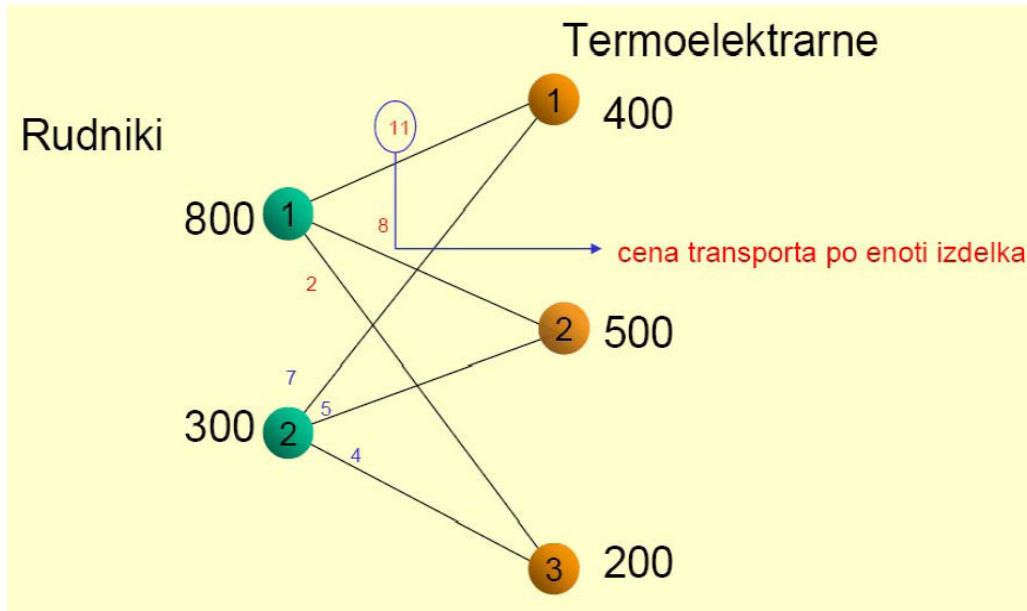
$$f(X) = \sum_{i=1}^3 \sum_{j=1}^5 x_{ij} c_{ij} \tag{2.2.}$$

minimalna. Tako smo opredelili še cilj, ki ga želimo doseči z optimizacijo (glej sliko 1).

Dobljena matematična formulacija, izražena v izrazih (2.1.) in (2.2.), vsebuje tako cilj, kot tudi vse spremenljivke in omejitve, ter nam predstavlja matematični model transportnega procesa. V takšni obliki je primerna za uporabo postopka za iskanje optimalne rešitve (glej sliko 1), to je takšnih prepeljanih količin  $x_{ij}$ ,  $i=1,\dots,3$ ;  $j=1,\dots,5$ , pri katerih bo ciljna funkcija (2.2.) minimalna.

V skladu s sliko 1 si je potrebno izbrati še primeren optimizacijski algoritem. Kot se izkaže, je za rešitev tega problema najprimernejši postopek **linearne programiranja (LP)** [21, 25], kjer uporabimo simpleksno metodo [21, 25]. Razlog za takšno izbiro tiči v dejstvu, da so vsi izrazi v (2.1.) in (2.2.) linearne narave.

V nadaljevanju si pogledjmo še konkreten primer transportnega problema, ilustriran na sliki 8 [4].



Slika 8: Konkreten primer transportnega problema

Iz slike 8 je razvidno, da imamo opravka z dvema rudnikoma s kapaciteto 800 ton in 300 ton premoga, ter tremi termoelektrarnami s potrebami 400, 500 in 200 ton premoga. Označene so tudi cene transporta na enoto premoga v določeni denarni enoti. Kot vidimo, gre za uravnotežen problem transporta, saj je vsota proizvedene tonaže premoga v rudnikih (1100 ton) enaka vsoti potreb termoelektrarn.

Kot v prejšnjem primeru, lahko problem predstavimo tudi v obliki naslednje tabele (tabela 2):

	TC 1	TC 2	TC 3	
Rudnik1	$c_{11}=11$	$c_{12}=8$	$c_{13}=2$	800
Rudnik 2	$c_{21}=7$	$c_{22}=5$	$c_{23}=4$	300
	400	500	200	

Kapaciteta

Povpraševanje

Tabela 2: Podatki za primer konkretnega transportnega problema

kjer so  $c_{ij}$  cene transporta na enoto premoga.

Če označimo odločitvene spremenljivke (neznane količine) kot:

$s_{ij}$ .....količina premoga iz rudnika  $i$  h termoelektrarni  $j$ ,  
 $i = 1, 2, j = 1, 2, 3$

potem lahko matematični model formuliramo v naslednji obliki:

$$\begin{aligned} \text{minimiziraj } f(s_{ij}) &= 11s_{11} + 8s_{12} + 2s_{13} + 7s_{21} + 5s_{22} + 4s_{23} \\ \text{omejitve:} & \\ s_{11} + s_{12} + s_{13} &= 800 \\ s_{21} + s_{22} + s_{23} &= 300 \\ s_{11} + s_{21} &= 400 \\ s_{12} + s_{22} &= 500 \\ s_{13} + s_{23} &= 200 \\ s_{ij} &\geq 0, \quad i = 1, 2, \quad j = 1, 2, 3 \end{aligned} \tag{2.3.}$$

Dobljena matematična formulacija, izražena v izrazih (2.3.), vsebuje tako ciljno funkcijo  $f(s_{ij})$ , kot tudi vse spremenljivke  $s_{ij}$  in omejitve (ravnotežje dobave in povpraševanja), ter nam predstavlja matematični model transportnega procesa. V takšni obliki je primerna za uporabo postopka za iskanje optimalne rešitve (glej sliko 1), to je takšnih prepeljanjih količin premoga  $s_{ij}$ ,  $i = 1, 2; j = 1, 2, 3$ , pri katerih bo ciljna funkcija  $f(s_{ij})$  minimalna.

### 2.2.2 Problem načrtovanja logističnih sistemov

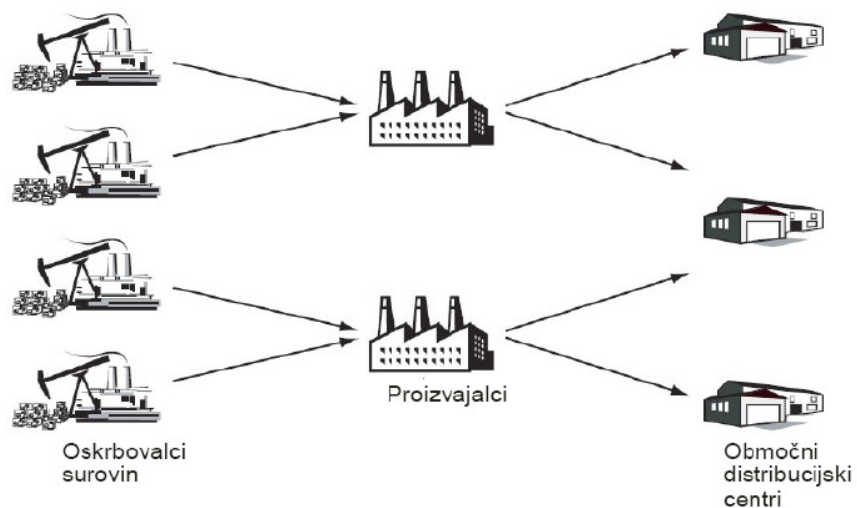
Tudi problem optimalnega načrtovanja logističnih sistemov je izredno pomemben pri reševanju optimizacijskih problemov v logistiki.

Za ta problem so značilne naslednje lastnosti [7]:

- Gre za načrtovanje lokacij logističnih virov in njihovih porazdelitev.
- Problem je potrebno rešiti že v fazi načrtovanja sistema.
- Odločitve o lokaciji lahko ključno vplivajo na povpraševanje.
- Zelo je pomembna optimalna topologija (centralna ali območna skladišča, itn).
- Zelo je pomembno določiti optimalne pretoke blaga (le ena vrsta ali več vrst blaga, itn).

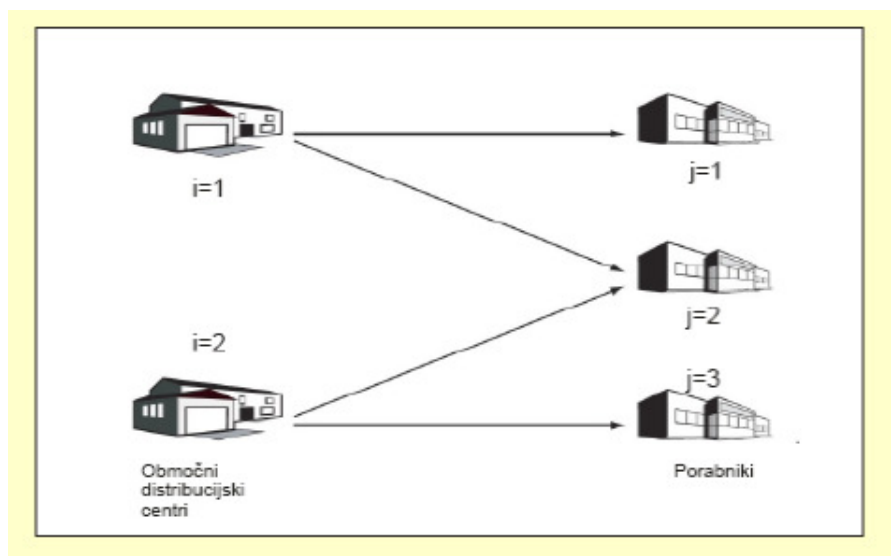
**Ključnega pomena pri tej problematiki je npr. določitev, koliko skladišč vzeti, kje jih postaviti, koliko vrst blaga prenašati in na katerih lokacijah to izvajati.**

Na sliki 9 si lahko ogledamo, kakšen bi bil tipičen primer izgleda dvonivojskega logističnega sistema [7].



Slika 9: Primer izgleda dvonivojskega logističnega sistema

Tipičen primer enonivojskega logističnega sistema pa je prikazan na sliki 10 [7].



Slika 10: Primer izgleda enonivojskega logističnega sistema

Na sliki 10 imamo opravka z določenim številom območnih distribucijskih centrov (ODC), ki oskrbujejo z blagom določeno število porabnikov. Konkretno je narisana primer, če bi imeli dva ODC in tri porabnike.

Naloga optimalnega načrtovanja tovrstnega logističnega sistema je zlasti ugotoviti dvoje [7]:

- Ugotoviti optimalno število ODC-jev, ter
- Ugotoviti optimalni transport (količine) blaga med ODC-ji in porabniki.

Za dotični (enonivojski) logistični sistem poskušajmo sedaj nakazati, kako bi se lotili postopka optimiranja.

V skladu s sliko 1 najprej opravimo postopek modeliranja, pri čemer dobimo kot rezultat matematični model enonivojskega logističnega sistema.

V postopku modeliranja najprej definiramo vse potrebne veličine:

$s_{ij}$  → količina blaga

$d_j$  → povpraševanje  $j$  – tega uporabnika

$u_i$  → trenutna kapaciteta

$C_{ij}(s_{ij})$  → cena transporta v odvisnosti od količine blaga

$F_i(u_i)$  → obratovni stroški  $i$  – tega oskrbovalnega centra

$q_i$  → maksimalna kapaciteta

kjer nam količine blaga  $s_{ij}$  predstavljajo odločitvene spremenljivke.

Matematični model bi potem lahko postavili v naslednji obliki [7]:

$$\begin{aligned} \text{minimiziraj } f_{KRIT} &= \sum_{i \in V_1} \sum_{j \in V_2} C_{ij}(s_{ij}) + \sum_{i \in V_1} F_i(u_i) \\ \text{omejitve:} & \\ \sum_{j \in V_2} s_{ij} &= u_i, \quad i \in V_1 \\ \sum_{i \in V_1} s_{ij} &= d_j, \quad j \in V_2 \\ u_i &\leq q_i, \quad i \in V_1 \\ s_{ij} &\geq 0, \quad i \in V_1, \quad j \in V_2 \\ u_i &\geq 0, \quad i \in V_1 \end{aligned} \tag{2.4.}$$

kjer je  $f_{KRIT}$  kriterijska funkcija, ki jo želimo minimizirati. Model (2.4.) še ni izražen v obliki, primerni za postopek optimiranja, saj v njem še ni definirana narava funkcij  $C_{ij}(s_{ij})$  in  $F_i(u_i)$ . Če je katera od njiju nelinearna funkcija, potem imamo opravka s problemom **nelinearnega programiranja (NLP)**. Tedaj moramo za algoritem



optimizacije izbrati primeren nelinearni optimizacijski algoritem, da lahko poiščemo optimalno rešitev.

V nadaljevanju si pogledjmo, v kakšno obliko preide model (2.4.), če imamo opravka z enostavnim LP problemom.

Prvo pa predpostavimo še naslednje:

- Topologija sistema je fiksirana.
- Lahko zanemarimo obratovalne stroške, to je:  $F_i(u_i) \approx 0$
- Stroški transporta so proporcionalni količini blaga, ki se transportira:  $C_{ij}(s_{ij}) = \bar{c}_{ij}s_{ij}$ , kjer je  $\bar{c}_{ij}$  cena na enoto izdelka.

Model (2.4.) pod temi predpostavkami preide v obliko [7]:

$$\begin{aligned} \text{minimiziraj } f_{KRT} &= \sum_{i \in V_1} \sum_{j \in V_2} \bar{c}_{ij} s_{ij} \\ \text{omejitve:} \\ \sum_{j=1}^n s_{ij} &= u_i, \quad i = 1, \dots, m \\ \sum_{i=1}^m s_{ij} &= d_j, \quad j = 1, \dots, n \\ s_{ij} &\geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n \\ 0 &\leq u_i \leq q_i \end{aligned} \tag{2.5.}$$

primerno za nadaljnje izvajanje postopka optimizacije.

### 2.2.3 Ostali možni problemi optimizacije logističnih sistemov

Naštejmo še nekaj ostalih možnih problemov optimizacije logističnih sistemov [17, 18]:

- Nelinearni transportni problem.
- Alokacija virov in njihova optimalna porazdelitev.
- Nelinearni problemi optimalne razporeditve istorodnih in raznorodnih resursov.
- Problem optimalnega železniškega in zračnega transporta, ter transporta v proizvodnji.
- Problem optimalnega razvrščanja opravil, transportnih sredstev in ljudi.
- Problem optimalnega razpošiljanja in dostave blaga.
- Problem minimalne poti in maksimalnega pretoka v logistiki.

- Problem optimalne izgradnje ali razširitve energetskega sistema.
- Problem optimalnega upravljanja zalog.
- Problem optimalnega izbora asortimana proizvodnje. Itn.

Tudi v naštetih primerih je potrebno v skladu s sliko 1 najprej določiti matematični model procesa, ter nato s pomočjo primernega optimizacijskega algoritma poiskati optimalno rešitev [17, 18].

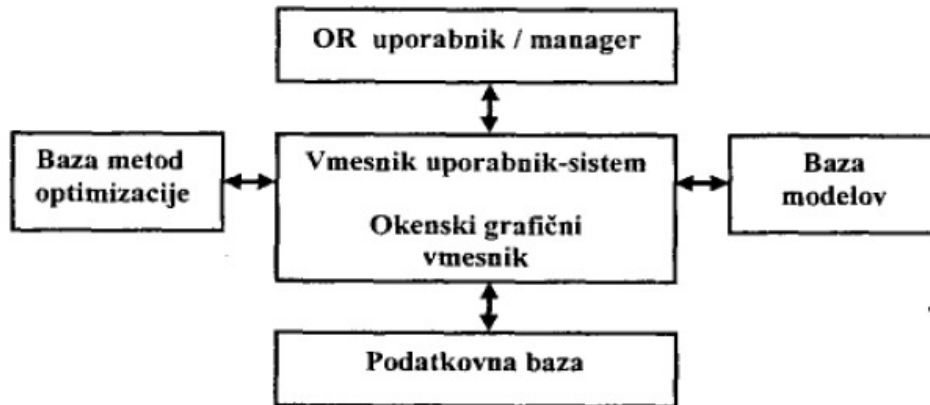
### 3 TEMELJNI KONCEPTI OPTIMIRANJA

Kot je razvidno iz slike 1 v 1. poglavju, se postopek optimiranja sestoji iz dveh temeljnih korakov:

- Postavitev strukture matematičnega modela logističnega procesa s pomočjo postopka modeliranja, ter
- Uporaba najbolj primerne algoritma optimizacije pri iskanju optimalne rešitve za dani matematični model.

V praksi so optimizacijski problemi zelo zapleteni in se večinoma rešujejo s pomočjo uporabe računalnikov. Pri tem žal ne obstaja enoten algoritem optimizacije, primeren za reševanje vseh problemov. To pomeni, da v praksi sicer obstaja veliko algoritmov, vendar so ti primerni le za reševanje določene skupine problemov.

Na sliki 11 si lahko ogledamo strukturo sistema za podporo optimalnemu odločanju [3, 11, 12].



Slika 11: Struktura sistema za podporo optimalnemu odločanju

Kot je razvidno iz slike 11, ima uporabnik (npr. manager) na voljo neko bazo metod optimizacije, bazo modelov, ter bazo podatkov iz logističnega procesa. Na osnovi teh baz se mora pri reševanju določenega logističnega problema odločiti, katera naj bi bila optimalna kombinacija za rešitev problema. Torej, da na osnovi znanj iz procesa izbere najprimernejši matematični model in zanj uporabi najprimernejšo optimizacijsko metodo pri iskanju optimalne rešitve problema.

V nadaljevanju si bomo najprej ogledali še nekaj značilnosti matematičnega modeliranja, kot sestavnega dela postopka optimiranja (slika 1!). Nato pa bomo nekaj spregovorili še o ekstremih funkcij ene ali več spremenljivk, katerih razumevanje je ključnega pomena pri postopku optimizacije. Poglavje bomo sklenili z lastnostmi lokalnih in globalnih minimumov ciljnih funkcij.

### 3.1 Modeliranje optimizacijskih modelov

Kot je prikazano na sliki 1, je matematično modeliranje logističnih procesov izredno pomembna faza v postopku optimiranja.

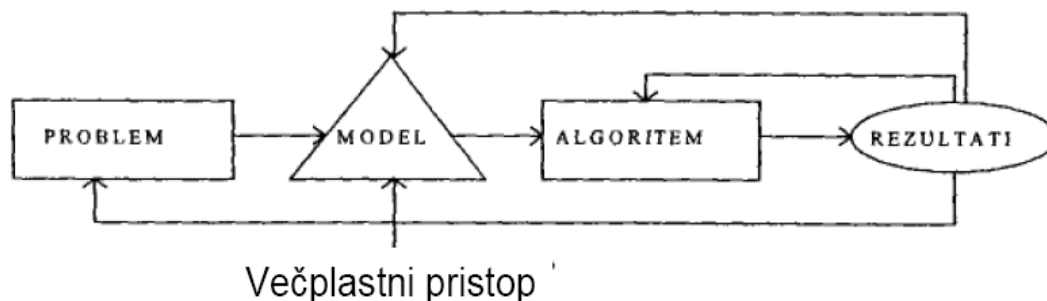
Naštejmo nekaj lastnosti matematičnih modelov [4]:

- Upeljemo jih, saj so realni sistemi ponavadi zelo zapleteni (kvantitativni in kvalitativni, deterministični, naključni, hevristični in drugi dejavniki).
- Model je poenostavitev realnosti (približek realnega problema).
- Postavitev dobrega modela je ključna za proces optimiranja.
- Če je model preveč enostaven → opis sistema ne bo dober .
- Če je model preveč kompliciran → postopek optimiranja bo pretežek.
- Model naj bo čim enostavnejši, a vseeno z vsemi zares potrebnimi podrobnostmi.
- Izpeljani model je osnova za uporabo postopkov optimiranja.

Glavne faze oblikovanja modelov pa so [4]:

- Postavitev naloge (formulacija problema),
- Zapis modela (npr. enačb),
- Iskanje rešitve (metode, algoritmi),
- Preverjanje modela in ocena rešitve,
- Uvajanje modela v prakso in kontrola njegove pravilnosti.

Pogosto se pri modeliranju zaradi zapletenosti problemov poslužujemo tako imenovanega večplastnega (čebulnega) pristopa, ki ga lahko ilustriramo z naslednjo sliko (slika 12) [3, 11, 12]:



Slika 12: Večplastni pristop k modeliranju

Pri njem najprej postavimo zelo enostaven model, ki vsebuje le najpomembnejše parametre. Nato pa ta model, brž ko ugotovimo, da je primeren, razširimo še z dodajanjem novih parametrov. Postopek ponavljamo, dokler model v celoti ne zadovolji vseh ciljev v postopku optimiranja.

Poudarimo še, da so izbor vrste modela, oblikovanje modela, obdelava eksperimentalnih podatkov, filtracije motenj, ocenjevanje napak, opredelitve karakteristik ter pomoč računalnika, usodno odvisni od raziskovalčevega poznavanja ustreznih matematičnih metod s področja operacijskih raziskav. Pri tem vse procedure in vmesne rešitve v procesu oblikovanja modela združujejo v sebi potrebo po jasnem razumevanju problema, ter vsebinski utemeljitvi, analizi in kritičnemu ocenjevanju uporabljenih modelov.

V nadaljevanju si pogledjmo še splošen matematičen zapis oz. formulacijo matematičnega modela, ter določitev strukture modela ponazorimo na preprostem primeru.

Pri postopku modeliranja običajno dobimo nek vektor odločitvenih spremenljivk, ki ga v primeru  $n$  spremenljivk v splošnem zapišemo na naslednji način:

$$x = (x_1, x_2, \dots, x_n)^T \quad (3.1.)$$

Tedaj kriterijska (ciljna) funkcija zavzame obliko:

$$f(x) = f\left[(x_1, x_2, \dots, x_n)^T\right] \quad (3.2.)$$

pri čemer je potrebno upoštevati še določene omejitve, ki jih v splošnem zapišemo v obliki nekih funkcij, izraženih z enačbami in/ali neenačbami:

$$\begin{aligned}g_i(x) &\leq 0, & i = 1, 2, \dots, p \\h_j(x) &= 0, & j = 1, 2, \dots, q\end{aligned}\tag{3.3.}$$

Glavna naloga postopka optimizacije je potemtakem poiskati takšen vektor (3.1.), pri katerem bo ciljna funkcija (3.2) zavzela optimalno vrednost (bodisi minimum ali maksimum), upoštevajoč seveda omejitve (3.3.).

V odvisnosti od prisotnosti ali odsotnosti omejitev (3.3.) pri problemu, ki ga rešujemo, ločimo dva možna primera:

- Problem optimizacije brez omejitev, ter
- Problem optimizacije z omejitvami.

V logistični praksi imamo ponavadi vedno opravka s primeri problemov, kjer so prisotne številne omejitve. Se pa v nekaterih primerih lahko tudi zgodi, da omejitev ni, kar vsekakor v mnogočem olajša postopek optimiranja.

Določitev izrazov (3.1.), (3.2.) in (3.3.) lahko ponazorimo na naslednjem preprostem primeru.

### **Problem**

*Grosist želi izprazniti skladišče z 10 tonami robe. Na razpolago ima skladišča dveh lokalnih odjemalcev. Prvi lahko sprejme največ 6 ton, drugi pa največ 5 ton robe. Strošek transporta med grosistom in prvim odjemalcem je 100 EUR/tono, med grosistom in drugim odjemalcem pa 150 EUR/tono. Določi optimalno razporeditev blaga (minimalni stroški!).*

Najprej definirajmo odločitveni spremenljivki:

$x_1$ .....količina, ki jo prenesemo v 1. skladišče  
 $x_2$ .....količina, ki jo prenesemo v 2. skladišče

pri čemer vektor odločitvenih spremenljivk (3.1.) zavzame dvokomponentno obliko:

$$x = (x_1, x_2)^T \quad (3.4.)$$

Kriterijska (stroškovna) oz. ciljna funkcija (3.2.) je glede na postavljeni problem in obliko vektorja (3.4.) naslednja:

$$f(x) = f[(x_1, x_2)^T] = 100 \cdot x_1 + 150 \cdot x_2 \quad (3.5.)$$

pri čemer smo stroška transporta (na tono) med grosistom in prvim oz. drugim odjemalcem pomnožili s prvo oz. drugo odločitveno spremenljivko.

Podajmo še vse omejitve:

$$\begin{array}{ll} x_1 + x_2 = 10 & \text{bilanca na strani grosista} \\ x_1 \leq 6 & \text{omejitve kapacitete 1. skladišča} \\ x_2 \leq 5 & \text{omejitev kapacitete 2. skladišča} \end{array} \quad (3.6.)$$

pri čemer izraza (3.3.) na osnovi (3.6.) preideta v obliko:

$$\begin{array}{l} h_1(x_1, x_2) = x_1 + x_2 - 10 = 0 \\ g_1(x_1) = x_1 - 6 \leq 0 \\ g_2(x_2) = x_2 - 5 \leq 0 \end{array} \quad (3.7.)$$

Dobljena matematična formulacija, izražena v izrazih (3.6.) in (3.7.), nam predstavlja matematični model za podani problem. V takšni obliki je primerna za uporabo postopka za iskanje optimalne rešitve (glej sliko 1), to je takšnih prenešenih količin  $x_1, x_2$ , pri katerih bo ciljna funkcija (3.5.) minimalna. Kot se izkaže, je tudi za rešitev tega problema najprimernejši postopek linearnega programiranja, kjer lahko zaradi le dveh odločitvenih spremenljivk uporabimo grafično metodo [21, 25].

## 3.2 Ekstremi funkcij ene in več spremenljivk

V postopku optimizacije, kjer za dani matematični model z uporabo določenega optimizacijskega algoritma iščemo optimalno rešitev, imamo opravka z iskanjem ekstremov funkcij ene ali več spremenljivk.

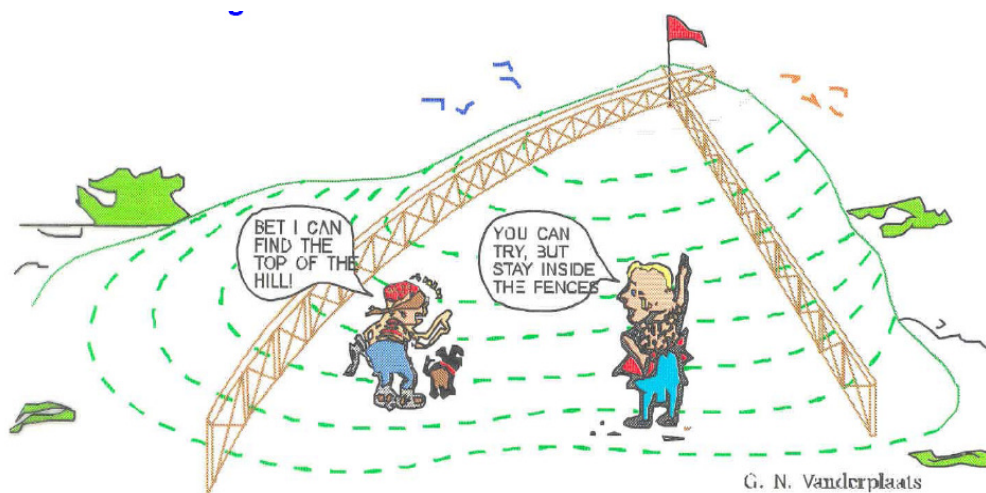
Pri tem ločimo dva pristopa:

- Pristop z analitičnim iskanjem ekstremov (analitično računanje), ter

- Pristop z numeričnim iskanjem ekstremov (numerične metode).

Poudariti je potrebno, da je možno prvi pristop uporabiti le v najenostavnejših primerih, sicer pa se v logistični praksi običajno vedno poslužujemo uporabe numeričnih optimizacijskih metod. Razlogi za to bodo podani oz. razvidni v naslednjih poglavjih, kjer se bomo osredotočili predvsem na uporabo numeričnih metod, analitični pristop iskanja ekstremov pa bomo le občasno in bežno obravnavali.

V primeru iskanja maksimuma si lahko postopek optimizacije ponazorimo tudi z naslednjo šaljivo sliko, ki predstavlja ilustrativni zgled iz narave (slika 13):



Slika 13: Ilustrativni zgled iz narave: Iskanje maksimuma

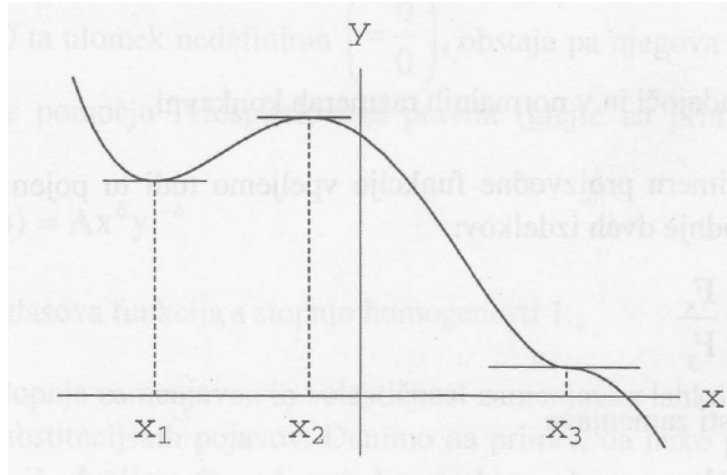
Kot je razvidno iz slike 13, velja naslednje:

- Cilj: Najti najvišjo točko hriba (maksimum).
- Spremenljivke: Prostorske koordinate.
- Omejitve: Ostati znotraj ograj.

Ker imamo na sliki 13 opravka s tremi odločitvenimi spremenljivkami (prostorskimi koordinatami  $x, y, z$ ) in je ciljna funkcija - funkcija treh spremenljivk  $f(x, y, z)$ , gre v tem primeru za problematiko iskanja ekstrema (maksimuma) funkcije večih (treh) spremenljivk.







Slika 15: Primer funkcije z enim minimumom, enim maksimumom, ter eno prevojno točko

Za dvakrat zvezno odvedljive funkcije ene spremenljivke ( $y = f(x)$ ) izračunamo ekstreme tako, da najprej z rešitvijo enačbe [17]:

$$f'(x) = \frac{df}{dx} = 0 \quad (3.8.)$$

poiščemo takoimenovane **stacionarne točke** (v katerih je strmina funkcije enaka 0 oz. je tangenta v njih vzporedna x osi), nato pa z drugim odvodom ugotovimo, v kateri stacionarni točki ima funkcija maksimum oz. minimum. Če je v stacionarni točki drugi odvod enak 0, o nastopu ekstrema odločajo višji odvodi.

Za funkcijo na sliki 15 velja naslednje [17]:

$$\begin{aligned} x = x_1 &\rightarrow \text{minimum, saj je } f'(x_1) = 0, \text{ ter } f''(x_1) > 0 \\ x = x_2 &\rightarrow \text{maksimum, saj je } f'(x_2) = 0, \text{ ter } f''(x_2) < 0 \\ x = x_3 &\rightarrow \text{prevoj, saj je } f'(x_3) = 0, f''(x_3) = 0, \text{ ter } f'''(x_3) \neq 0 \end{aligned} \quad (3.9.)$$

Na kratko si pogledjmo še primer, ko imamo opravka s funkcijami več spremenljivk.

### 3.2.2 Ekstremi funkcij več spremenljivk

Na tem mestu se bomo dotaknili le problematike iskanja ekstremov funkcij dveh spremenljivk. Na podoben način bi namreč lahko obravnavo posplošili tudi na funkcije treh in več spremenljivk.

Če imamo opravka s funkcijo dveh spremenljivk:  $z = f(x, y)$ , izračunamo stacionarne točke tako, da rešimo naslednji sistem enačb [17]:

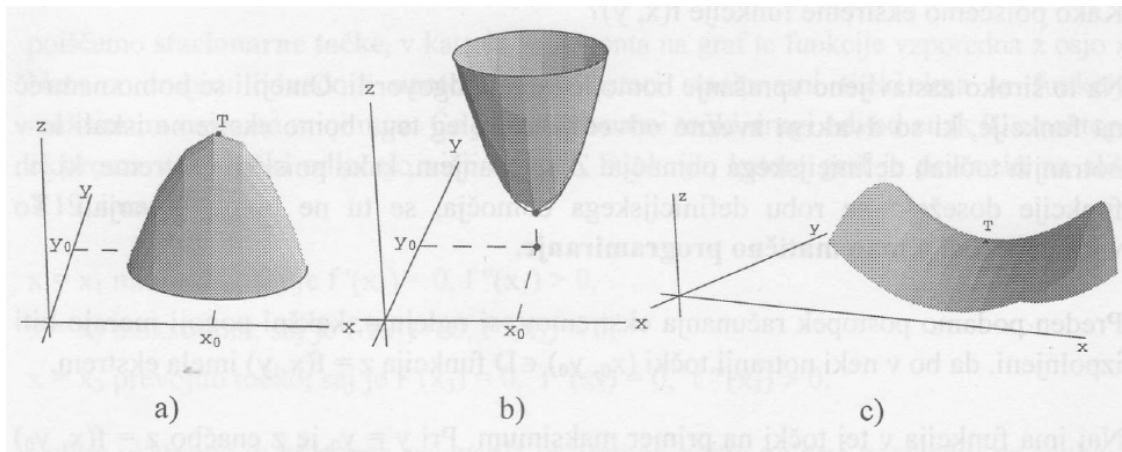
$$\nabla z = \nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} = [0 \quad 0] \quad (3.10.)$$

oziroma:

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} = 0$$

kjer je  $\nabla z = \nabla f(x, y)$  **gradientni vektor parcialnih odvodov** funkcije  $z = f(x, y)$ . Pri tem je potrebno poudariti, da na takšen način dobimo šele potrební pogoj za nastop ekstrema funkcije v neki notranji točki definicijskega območja.

V stacionarnih točkah obravnavane funkcije je tangentsna ravnina ploskve, ki je določena s to funkcijo, vzporedna x-y ravnini. Slika 16 prikazuje tri primere funkcij dveh spremenljivk, za katere je točka  $(x_0, y_0)$  stacionarna točka.



Slika 16: Primeri stacionarnih točk funkcij dveh spremenljivk: a) maksimum; b) minimum; c) sedlasta točka

Primer c) na sliki 16 nazorno kaže, da izpolnitev sistema enačb (3.10.) še ni zadosten pogoj za to, da ima funkcija v stacionarni točki res ekstrem (bodisi minimum ali maksimum).

Zadosten pogoj, da ima funkcija v stacionarni točki  $(x_0, y_0)$  ekstrem, je nekoliko bolj zapleteno preverjati, kot je bilo to pri funkciji ene spremenljivke (glej izraze (3.9)).

Bralec si tovrstno problematiko lahko поблиže pogleda v literaturi [17], poleg tega bo o tem več govora v naslednjih poglavjih (zlasti v poglavju 5). Tam bodo podrobneje obravnavani tudi pojmi, kot so gradient funkcije, Hessova matrika, itn, ki so ključnega pomena pri iskanju ekstremov in ugotavljanju tipov stacionarnih točk za funkcije več spremenljivk.

V nadaljevanju se na kratko seznanimo še s problematiko ugotavljanja lokalnih in globalnih minimumov funkcij, s čimer bomo zaključili obravnavo 3. poglavja.

### **3.3 Lokalni in globalni minimumi ciljnih funkcij**

Poudariti je potrebno, da večina numeričnih optimizacijskih postopkov, opisanih v literaturi, pripelje le do lokalnega minimuma kriterijske funkcije, ki ni nujno tudi globalen (glej sliko 14). Teža tega problema je odvisna od tipa funkcije. Če ima funkcija en sam minimum, ta problem sploh izgine. Kadar vemo, da ima kriterijska funkcija le nekaj minimumov in vsaj približno poznamo njihovo število, ni težko s ponavljanjem numerične optimizacije pri različnih začetnih pogojih poiskati vseh in z medsebojno primerjavo najti globalnega. Pogosto niti ni potrebno najti globalnega minimuma, ampak zadošča zadovoljiv lokalni minimum. V dvome pa nas spravi lokalni minimum, ki ne daje zadovoljivega rezultata. Takrat ne vemo, ali lahko z drugačnimi začetnimi pogoji optimizacijskega postopka dosežemo boljši minimum, ali pa se je potrebno problema lotiti na drugačen način. Več o tej problematiki si lahko bralec поблиže pogleda v literaturi [17].

## 4 PREGLED GLAVNIH OPTIMIZACIJSKIH METOD

V tem poglavju bomo opravili krajši pregled nekaterih glavnih optimizacijskih metod, opisali razlike med njimi, ter podali njihove glavne značilnosti.

### 4.1 Razdelitev in lastnosti optimizacijskih metod

Metode iskanja optimuma (metode numerične optimizacije) so poznane tudi pod imenom *Metode matematičnega programiranja* in so v splošnem obravnavane kot področje oz. del operacijskih raziskav [21]. Slednje predstavljajo posebno vejo v matematiki, ki zadeva apliciranje znanstvenih metod in disciplin na probleme odločanja (angl. *Decision making problems*), kjer se iščejo najboljše oz. najbolj optimalne rešitve.

Metode matematičnega programiranja bi lahko razvrstili na naslednji način [21]:

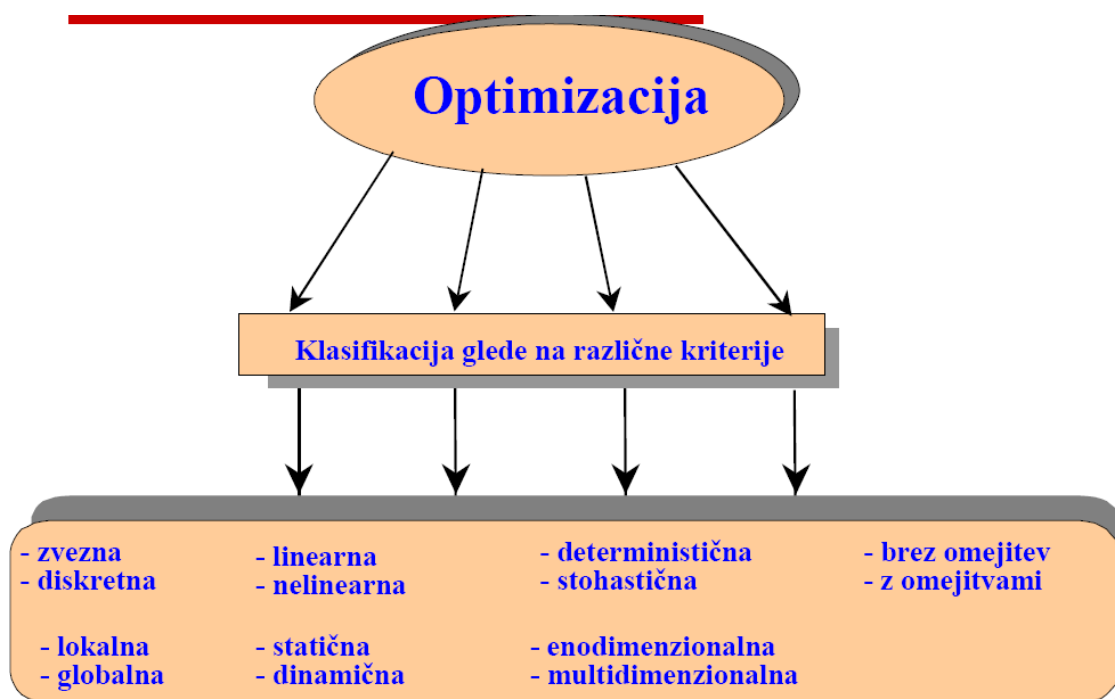
- Nelinearno programiranje (NLP),
- Geometrično oz. geometrijsko programiranje (GP),
- Kvadratno programiranje,
- Linearno programiranje (LP) \* ,
- Dinamično programiranje (DP) \* ,
- Celoštevilsko linearno programiranje (CLP) \* ,
- Mešano celoštevilsko linearno programiranje (MCLP),
- Stohastično programiranje,
- Separabilno programiranje,
- Večkriterijsko programiranje,
- Mrežna optimizacija (MO) \* ,
- Teorija iger,
- Simulirano ohlajanje,
- Genetski algoritmi, ter
- Nevronske mreže.

Pri pregledu metod se bomo v tem poglavju dotaknili le metod, označenih z \* , medtem

ko bomo nelinearno programiranje in geometrijsko programiranje temeljiteje obdelali v naslednjih poglavjih. Kar se pa tiče ostalih metod, pa si bralec lahko поблиže pogleda njihovo obravnavo v literaturi [21]. Večina metod v zgornji razvrstitvi je primerna zlasti za iskanje lokanih ekstremov. Po drugi strani pa se kot učinkoviti pristopi k iskanju globalnih ekstremov še posebej izkažejo postopki simuliranega ohlajanja, genetskih algoritmov in nevronske mreže [21, 25].

Pri obravnavi problemov optimiranja in njihove razvrstitve velikokrat upoštevamo predvsem dva kriterija. Po prvem kriteriju dobimo razvrstitev problemov glede na uporabljeno metodo matematičnega programiranja. Po tem kriteriju razlikujemo sledeče tipe problemov matematičnega programiranja: probleme linearnega programiranja, probleme nelinearnega programiranja, probleme celoštevilčnega (linearnega) programiranja, probleme mrežne optimizacije, itn. Po drugem kriteriju pa lahko razvrstimo probleme glede na njihovo vsebino, kot npr. proizvodni problemi, transportni problemi, problemi zmesi, problemi zalog, problemi čakalnih vrst, itd [22].

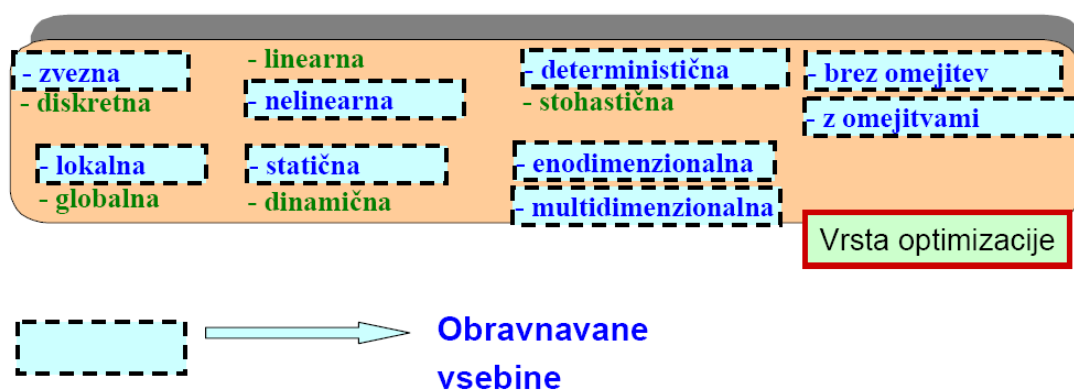
V splošnem pa se optimizacijske metode in problemi lahko klasificirajo še v skladu z naslednjimi tipičnimi delitvami, ki jih prikazuje slika 17 [4]:



Slika 17: Klasifikacija optimizacijskih metod in problemov

Kot je razvidno na sliki 17, se lahko klasifikacija optimizacijskih metod in problemov izvrši glede na to, a so prisotne omejitve v problemu ali ne. Klasifikacijo lahko izvedemo tudi glede na to, a so parametri v izrazih matematičnega modela fiksni (statični), ali pa se spreminjajo s časom (so dinamični). Naslednji pomemben vidik klasifikacije temelji na (linearni ali nelinearni) naravi izrazov (ciljne funkcije in omejitev) matematičnega modela. Če so vsi izrazi linearni, imamo opravka z linearno optimizacijo. Že npr. ena sama nelinearna omejitev pa povzroči, da moramo problem obravnavati s stališča nelinearne optimizacije. Optimizacijo lahko razvrstimo še na lokalno/globalno, glede na naravo ekstrema, ki ga iščemo (lokalni ali globalni), na enodimenzionalno/večdimenzionalno, glede na število neodvisnih odločitvenih spremenljivk (ena ali več spremenljivk), ter na deterministično/stohastično, glede na deterministično ali stohastično naravo spremenljivk in parametrov v matematičnem modelu. Če se v problemu pojavijo naključne spremenljivke, potem ga moramo obravnavati kot stohastični problem. Nenazadnje bi bila klasifikacija seveda lahko izvršena tudi glede na dovoljene vrednosti iz definicijskega območja odločitvenih spremenljivk. Če pridejo v poštev vsa realna števila iz tega območja, potem imamo opravka z zvezno optimizacijo. Če pa določene odločitvene spremenljivke lahko zavzamejo le celoštevilčne vrednosti, pa gre za problem diskretne optimizacije.

Slika 18 prikazuje področja optimizacije, ki bodo glavna domena obravnave v tem delu.



Slika 18: Področja optimizacije, ki bodo glavna domena obravnave

Torej se bomo v tem delu osredotočili predvsem na obravnavo optimizacije, ki ima zvezen, lokalni, nelinearen, statičen in determinističen značaj. Pri tem bomo obravnavali tako primere z eno ali več spremenljivkami, kot tudi primere brez omejitev ali z omejitvami. Kar se tiče primerov z omejitvami, bodo le-ti podani predvsem v sklopu obravnave analitične optimizacije in geometrijskega programiranja, kar bo razvidno v naslednjih poglavjih.

Nekatere izmed nepoudarjenih vsebin iz slike 18 (npr. linearna, diskretna, dinamična optimizacija, itn) bomo v tem poglavju le na kratko preleteli. Natančnejšo obravnavo vseh snovi, podanih na sliki 18, pa si bralec lahko поблиže pogleda v literaturi [7, 21].

V nadaljevanju tega poglavja bomo torej na kratko pogledali opis naslednjih sklopov optimizacijskih metod [7, 21]:

- Linearno programiranje.
- Celoštevilčno programiranje.
- Dinamično programiranje.
- Mrežna optimizacija.

Te metode namreč poleg metod nelinearnega in geometrijskega programiranja, katerim bo posvečeno največ pozornosti v naslednjih poglavjih, tudi pokrivajo marsikatera temeljna področja optimizacije, tako pri obravnavi splošnih problemov, kot tudi problemov v logistiki in transportu.

Zato jih je smiselno na kratko predstaviti, tudi zato, da bralec dobi občutek, v čem se razlikujejo od glavnih vsebin tega dela, to je nelinearnega in geometrijskega programiranja.



## 4.2 Linearno programiranje

Linearnega programiranja smo se nekoliko že dotaknili v dveh primerih v poglavjih 2.2.1. in 2.2.2. (pri poenostavljenem modelu). Na tem mestu si bomo linearno programiranje še nekoliko pobliže ogledali oz. podali kratek opis tovrstne metodologije. Več o tovrstni metodologiji pa si bralec lahko pogleda v literaturi [7, 21, 25].

Mnogi problemi, s katerimi se srečujemo v gospodarstvu in tehniki, so povezani z optimalno uporabo ali razporeditvijo kakršnihkoli sredstev, ki so na razpolago le v omejenih količinah.

Tudi v logistiki je veliko problemov povezanih z razporejanjem dobrin (materiala, blaga, proizvodov). Takšno (pre)razporejanje je v konkretnih praktičnih problemih možno izvesti na več različnih načinov. Med njimi vedno skušamo izbrati najugodnejšo možnost, ki predstavlja optimalno rešitev. Npr se poskušamo odločiti, kako:

- določiti časovno najhitrejši prevoz blaga.
- imeti najmanjše stroške pri proizvodnji določenih artiklov.
- imeti največji dobiček pri prodaji itn.

Pri tem pa pogostokrat naletimo na številne omejitve, kot npr.:

- omejen čas dobave.
- omejene kapacitete proizvodnje.
- omejene kapacitete prevozov.
- omejena količina materiala.
- omejena trajnost blaga itn.

V takšnih primerih seveda govorimo o optimizaciji z omejitvami. Probleme takšne narave lahko rešujemo na različne načine, ena od metod, s katerimi je mogoče dobiti optimalne rešitve pri določenih omejitvah, pa je linearno programiranje [21, 25].

Pri tem moramo poudariti, da v tej besedni zvezi pojem »programiranje« nima nobene zveze s programiranjem računalniških programov, pač pa je iz zgodovinskih razlogov pod tem pojmom mišljeno ustrezno načrtovanje matematičnega modela v primerno obliko za optimizacijo.

Linearno programiranje uporabljamo v primerih, kjer se lahko predvidi ukrepanje, ki naj bi bilo s kakega vidika optimalno. Zaradi tega se lahko uporabi npr. v gospodarstvu pri planiranju gospodarskih dejavnosti, v industriji pri organiziranju najbolj ekonomičnih tehnoloških procesov, v prometu pri najbolj ekonomičnem izkoriščanju obstoječih prometnih sredstev itn. Naštejmo nekaj posebno značilnih in pogostih možnosti uporabe linearnega programiranja [26]:

- **Proizvodni problem:** Proizvodno podjetje ima na razpolago omejene količine surovin, ki jih lahko uporabi pri proizvodnji raznih izdelkov. Za vsak izdelek je določeno, kolikšen delež surovine je zanj potreben in koliko znaša njegova tržna cena. Pri teh podatkih se postavlja v podjetju problem, kako naj planira svojo proizvodnjo izdelkov, da bo imelo od prodaje izdelkov čim večji dohodek.
- **Transportni problem:** Več tovarn, katerih proizvodne kapacitete so znane, oskrbuje s kakim blagom več potrošnih središč, katerih potrebe po tem blagu so tudi znane. Znani so tudi transportni stroški za enoto blaga od vsake tovarne do vsakega potrošnega središča. Pri teh podatkih se postavlja problem, kako naj se načrtuje prevoz blaga od tovarn do potrošnih središč, da bodo skupni transportni stroški najmanjši.
- **Problem omrežja:** Dve mesti sta povezani s prometnim omrežjem, v katerem je prepustnost (kapaciteta) vsake prometne veje določena, pri čemer teče promet po omrežju od prvega do drugega mesta. Pri teh podatkih se postavlja problem, kako naj se načrtuje promet, da bo prepustnost celotnega omrežja čim večja.

Z matematičnega vidika predstavlja linearni program problem vezanega ekstrema (Jamnik, 1981), pri katerem iščemo ekstrem linearne ciljne funkcije na konveksni množici, ki je določena s sistemom linearnih neenačb in pogojem nenegativnosti spremenljivk.

Matematično osnovo linearnega programiranja sestavljata teorija linearnih neenačb in enačb in teorija konveksnih poliedrov, torej dve panogi teoretične matematike, ki sta za potrebe linearnega programiranja dovolj razviti [26]. Probleme linearnega programiranja rešujemo v splošnem z algebrskimi metodami. Matematična sredstva, ki jih pri tem uporabljamo, so v splošnem odvisna od števila nastopajočih odločitvenih spremenljivk in od števila neenačb, ki se pojavljajo v omejitvah.

Značilna načina reševanja problemov linearnega programiranja sta [25]:

- Grafična metoda (če imamo opravka le z dvema odločitvenima spremenljivkama), ter
- Metoda simpleksov (če imamo opravka z več kot dvema odločitvenima spremenljivkama).

Torej je rešitve problemov linearnega programiranja mogoče dobiti v grafični obliki, če imamo le dve odločitveni spremenljivki in je število računskih operacij majhno, sicer pa je potreben računski postopek. Pri tem velja, da je metoda simpleksov splošna računška metoda za reševanje linearnega programa, ki jo je leta 1948 razvil ameriški matematik G.B. Dantzig. Nadvse priporočljivo jo je uporabiti, ko imamo opravka z več odločitvenimi spremenljivkami, saj tedaj število numeričnih operacij strmo narašča in bi bili brez uporabe računalnika dokaj nemočni.

V samem postopku reševanja problemov linearnega programiranja ločimo tri pomembne korake [25]:

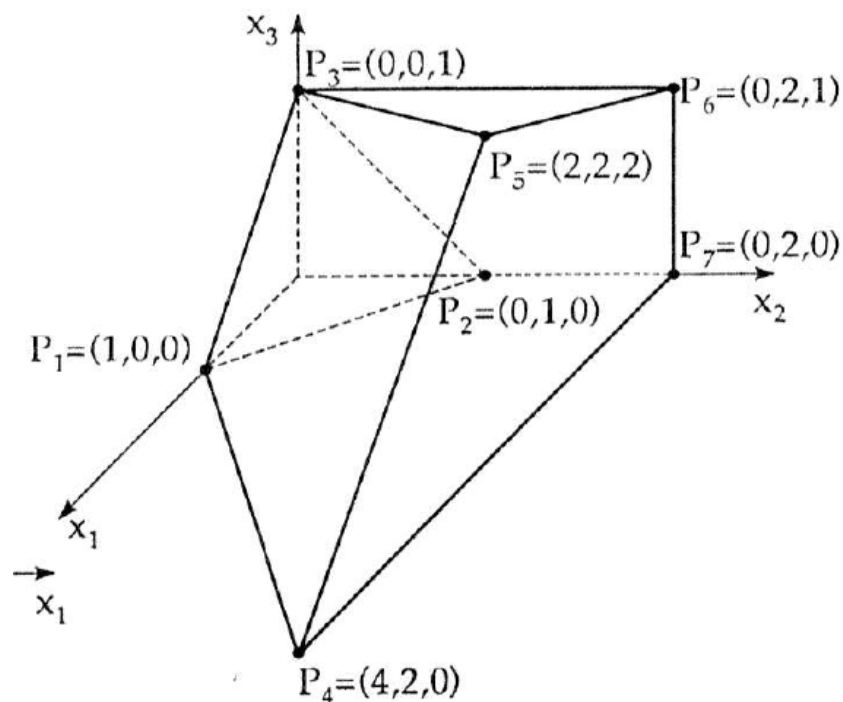
- Formulacija problema, to je postavitve problema v pravilni obliki,
- Rešitev, to je izračun optimalne možnosti.
- Senzitivnostna analiza, to je ugotovitev, kaj bi se zgodilo, če bi se pogoji našega zastavljenega problema za zelo malo spremenili.

Za probleme linearnega programiranja velja, da neenačbe, ki predstavljajo omejitve, tvorijo takoimenovan konveksni večkotnik (polieder) (glej sliko 19). Vse točke znotraj tega lika in na njegovih straneh ustrezajo vsem neenačbam istočasno. Konveksni polieder ima zanimivo lastnost: če katerikoli dve točki v notranjosti povežemo z daljico, leži ta daljica v celoti znotraj tega lika [25].

Ciljna funkcija lahko v posameznih točkah konveksnega poliedra (bodisi znotraj ali na njegovem robu) zavzema različne vrednosti. Vsem tem točkam pravimo možne rešitve linearnega programa, saj je le tu dovoljeno iskati optimalno rešitev (v področju neenačb oz. omejitev).

***Pri tem se da dokazati, da najdemo optimalno rešitev (če obstaja) vedno le v enem od oglišč konveksnega poliedra! [25]. Tam bo torej ciljna funkcija gotovo zavzela ekstremno vrednost, rešitev pa bo optimalna.***

Konveksni polieder pri treh odločitvenih spremenljivkah  $(x_1, x_2, x_3)$  prikazuje slika 19.



Slika 19: Konveksni polieder pri treh odločitvenih spremenljivkah

Poudariti velja, da ima pri reševanju problemov linearnega programiranja z veliko odločitvenimi spremenljivkami pripadajoči konveksni polieder tudi veliko oglišč, med katerimi iščemo optimalno rešitev. To pomeni, da je optimalno rešitev potrebno iskati karseda sistematično. In ravno metoda simpleksov je postopek, ki to omogoča.

Postavimo sedaj splošno formulacijo problema linearnega programiranja. V splošnem lahko zapišemo LP program na naslednji način [25]:

$$\begin{aligned}
 & \text{optimiraj } f(x_1, x_2, \dots, x_n) = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (\text{ciljna funkcija}) \\
 & \text{pri omejitvah:} \\
 & a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \leq b_1 \\
 & a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \leq b_2 \\
 & \dots \\
 & a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n \leq b_m \\
 & x_k \geq 0, \quad k = 1, 2, \dots, n
 \end{aligned} \tag{4.1}$$

kjer optimum pomeni skupni izraz za maksimum v nekaterih primerih ali pa za minimum v nekaterih drugih situacijah.

Izraz (4.1.) lahko zapišemo tudi v bolj pregledni vektorko-matrični obliki na naslednji način [25]:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}, \quad (4.2.)$$

$$\mathbf{C} = [c_1 \quad c_2 \quad \dots \quad c_n], \quad \mathbf{x} = (x_1, x_2, \dots, x_n)^T$$

$$f = \mathbf{C} \cdot \mathbf{x}$$

$$\mathbf{A} \cdot \mathbf{x} \leq \mathbf{B}$$

$$\mathbf{x} \geq 0$$

Seveda je glavni cilj pri linearnem programiranju poiskati takšen optimalen vektor odločitvenih spremenljivk  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ , pri katerem se bo dosegla optimalna (minimalna ali maksimalna) vrednost ciljne funkcije  $f(x_1, x_2, \dots, x_n)$ .

Poudarimo še, da je le redkokdaj mogoče poiskati optimalno rešitev brez uporabe računalnikov oz. računalniških programov, kar še posebej velja pri velikem številu odločitvenih spremenljivk.

Praktične primere uporabe linearnega programiranja za reševanje problemov z dvema odločitvenima spremenljivkama, kjer je možno uporabiti grafično metodo, lahko bralec zasledi v literaturi [5, 7, 25]. Prav tako je mogoče v literaturi zaslediti tudi primere z več odločitvenimi spremenljivkami, kjer je potrebno uporabiti metodo simpleksov [21, 25].

### 4.3 Celoštevilčno programiranje

V mnogih primerih je nemogoče imeti odločitvene spremenljivke, ki bi lahko zavzele vse realne vrednosti. Naštejmo nekaj primerov:

- Nemogoče je uporabiti 1,6 ventila v termalni močnostni postaji,
- Nemogoče je uporabiti 1,9 delavca pri določenem opravilu,

- Nemogoče je uporabiti 2,76 zidaka v določenem objektu,
- Nemogoče je zgraditi 1.37 šole,
- Nemogoče je izdelati 11.74 letal, itn.

V takšnih primerih je priporočljivo vpeljati celoštevilске odločitvene spremenljivke, oz. je smiselno za tovrstne spremenljivke vpeljati cela števila.

Spremenljivki, ki ima samo celoštevilске vrednosti (0, 1, 2, ...), torej rečemo celoštevilska spremenljivka. Poseben primer celoštevilskih spremenljivk so spremenljivke, ki so lahko enake le 0 ali 1. Zanje je v rabi izraz 0-1 oz. binarne spremenljivke. Z njimi je v modelu mogoče upoštevati, ali nek pojav nastopi ali ne. Če pojav nastopi, ima celoštevilska spremenljivka vrednost 1, če ne nastopi, pa ima vrednost 0.

Celoštevilске spremenljivke so pomembne tudi zato, ker je z njimi mogoče izražati tudi negladke funkcije. To so funkcije, ki niso povsod zvezne, ali pa niso povsod odvedljive. Pri tem so celoštevilске spremenljivke še posebej primerne v primeru odsekoma linearnih funkcij, s katerimi se pogostokrat srečujemo.

Če v optimizacijskem modelu nastopajo samo celoštevilске spremenljivke, imamo celoštevilski model. Če poleg celoštevilskih nastopajo tudi zvezne spremenljivke iz nabora vseh realnih števil, imamo mešani celoštevilski model. Če so pri tem vse omejitve in ciljna funkcija linearne, imamo linearni celoštevilski ali linearni mešani celoštevilski optimizacijski model.

V splošnem lahko zapišemo linearni celoštevilski problem (CLP) na naslednji način.

Ciljna funkcija  $n$  odločitvenih spremenljivk  $(x_1, x_2, \dots, x_n)$  ima naslednjo linearno obliko:

$$f(x_1, x_2, \dots, x_n) = c_1x_1 + c_2x_2 + \dots + c_nx_n \quad (4.3)$$

pri čemer veljajo naslednje linearne omejitve:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &\leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &\leq b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &\leq b_m \end{aligned} \quad (4.4)$$

in velja še:

$$x_1, x_2, \dots, x_n \in \mathbb{Z} \quad (4.5)$$

torej lahko odločitvene spremenljivke zavzemajo vrednosti le iz nabora celih števil.

Podani matematični model se na prvi pogled ne razlikuje dosti od klasičnega linearnega programa (LP) (primerjaj z izrazom (4.1.)), razen seveda izraza (4.5.), ki ga pri linearnem programu ni.

Ker so v tem primeru vse odločitvene spremenljivke celoštevilčne, seveda govorimo o "čistem" celoštevilskem linearnem programu.

### ***4.3.1 Modeliranje celoštevilskih problemov***

V nadaljevanju si pogledjmo dva značilna primera modeliranja celoštevilčnih problemov, kjer se pojavijo celoštevilške odločitvene spremenljivke. V enem primeru bomo imeli opravka s situacijo, kjer so odločitvene spremenljivke lahko le binarne, torej zgolj z vrednostjo 0 ali 1. V drugem primeru pa lahko celoštevilške spremenljivke zavzamejo vrednosti iz celega nabora celih števil.

#### **Primer 1:**

Pet različno uspešnih delavcev hočemo razporediti na pet različno zahtevnih delovnih mest. Ker želimo, da bi bil skupen učinek vseh delavcev čim boljši, izmerimo uspešnost vsakega delavca na vsakem delovnem mestu. Mera uspešnosti naj bo čas, potreben za to, da delavec opravi opravilo na posameznem delovnem mestu. Ker so delavci različno usposobljeni in različno spretni, so ti časi za vsakega delavca in za vsako delovno mesto drugačni. Rezultati preizkusa delavcev so podani s tabelo 3. Želimo poiskati takšen razpored delavcev na delovna mesta, da bo vsota porabljenih časov za vsa opravila minimalna.

	M <sub>1</sub>	M <sub>2</sub>	M <sub>3</sub>	M <sub>4</sub>	M <sub>5</sub>
D <sub>1</sub>	25	31	40	29	27
D <sub>2</sub>	36	32	32	27	40
D <sub>3</sub>	28	29	37	30	41
D <sub>4</sub>	35	42	36	28	33
D <sub>5</sub>	40	39	41	42	45

Tabela 3: Rezultati uspešnosti delavcev na posameznih delovnih mestih, merjeni s časi, potrebnimi za dokončanje opravila na določenem delovnem mestu (D<sub>i</sub> – delavci, M<sub>i</sub> – delovna mesta)

### Nastavitev modela:

Očitno gre za celoštevilski problem, kjer moramo uporabiti binarne celoštevilске spremenljivke. Zato matematični model definiramo na naslednji način.

Najprej vpeljemo naslednje binarne celoštevilске spremenljivke:

$u_{ij}$  ... celoštevilška spremenljivka za  $i$ -tega delavca in  $j$ -to del.mesto

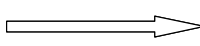
ki jih definiramo na naslednji način:

$$u_{ij} = \begin{cases} 1, & \text{če } i\text{-tega delavca razporedimo na } j\text{-to del.mesto} \\ 0, & \text{če } i\text{-tega delavca ne razporedimo na } j\text{-to del.mesto} \end{cases} \quad (4.6)$$

$i = 1, 2, \dots, 5, \quad j = 1, 2, \dots, 5$

Nato definiramo naslednjo kriterijsko funkcijo, ki predstavlja seštevek vseh potrebnih časov vseh delavcev na vseh delovnih mestih (glej tabelo 3.):

$$f = 25u_{11} + 31u_{12} + 40u_{13} + 29u_{14} + 27u_{15} + \\ + 36u_{21} + 32u_{22} + 32u_{23} + 27u_{24} + 40u_{25} + \\ + 28u_{31} + 29u_{32} + 37u_{33} + 30u_{34} + 41u_{35} + \\ + 35u_{41} + 42u_{42} + 36u_{43} + 28u_{44} + 33u_{45} + \\ + 40u_{51} + 39u_{52} + 41u_{53} + 42u_{54} + 45u_{55}$$

 Ciljna funkcija  
(iščemo minimum)

(4.7.)

Definirati moramo tudi omejitve, ki upoštevajo dejstvo, da gre lahko vsak delavec le na eno delovno mesto, oziroma, da je vsako delovno mesto lahko dodeljeno le enemu delavcu. Tako dobimo:



$$\begin{array}{l}
 u_{i1} + u_{i2} + u_{i3} + u_{i4} + u_{i5} = 1, \quad i = 1, 2, \dots, 5 \\
 u_{1j} + u_{2j} + u_{3j} + u_{4j} + u_{5j} = 1, \quad j = 1, 2, \dots, 5
 \end{array}
 \begin{array}{l}
 \Rightarrow \text{i-ti delavec gre lahko največ na} \\
 \text{1 del. mesto} \\
 \Rightarrow \text{j-to del.mesto je lahko dodeljeno} \\
 \text{največ 1 delavcu}
 \end{array}$$

(4.8.)

Z definiranjem izrazov (4.6.), (4.7.) in (4.8.) smo postavili ustrezno strukturo matematičnega modela, sestavljeno iz kriterijske funkcije in omejitev. Kriterijska funkcija lahko pri različnih razporeditvah enk in ničel v binarnih spremenljivkah zavzema različne vrednosti. Seveda iščemo takšno razporeditev enk in ničel, kjer bo kriterijska funkcija dosegla najmanjšo vrednost.

Denimo razporedimo i-tega delavca na i-to del. mesto:

$$u_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \quad i = 1, 2, \dots, 5, \quad j = 1, 2, \dots, 5 \quad (4.9.)$$

Seštevek vseh potrebnih časov vseh delavcev na vseh delovnih mestih je tedaj naslednji:

$$\begin{array}{l}
 f = 25.1 + 31.0 + 40.0 + 29.0 + 27.0 + \\
 + 36.0 + 32.1 + 32.0 + 27.0 + 40.0 + \\
 + 28.0 + 29.0 + 37.1 + 30.0 + 41.0 + \\
 + 35.0 + 42.0 + 36.0 + 28.1 + 33.0 + \\
 + 40.0 + 39.0 + 41.0 + 42.0 + 45.1
 \end{array}
 \begin{array}{l}
 \Rightarrow \text{Boljši razpored} \\
 \text{f} = 25.1 + 32.1 + 37.1 + 28.1 + 45.1 = 167
 \end{array}$$

(4.10.)

pri čemer je prvi delavec razporejen na prvo delovno mesto, drugi delavec na drugo delovno mesto, itn, ..., zadnji delavec na zadnje delovno mesto. Zato so enke postavljene le v "diagonali", ničle pa povsod drugod. Pri takšni razporeditvi enk in ničel kriterijska funkcija zavzame skupno vrednost 167 porabljenih časovnih enot, da se opravijo vsa opravila delavcev na delovnih mestih.

Če pa naredimo majhno modifikacijo izraza (4.10.) in 1. delavca damo na 2. delovno mesto, 2. delavca pa na 1. delovno mesto, se kriterijska funkcija spremeni v naslednjo obliko:

$$\begin{aligned}
 f = & 25.0 + 31.1 + 40.0 + 29.0 + 27.0 + \\
 & + 36.1 + 32.0 + 32.0 + 27.0 + 40.0 + \\
 & + 28.0 + 29.0 + 37.1 + 30.0 + 41.0 + \\
 & + 35.0 + 42.0 + 36.0 + 28.1 + 33.0 + \\
 & + 40.0 + 39.0 + 41.0 + 42.0 + 45.1
 \end{aligned}
 \Rightarrow
 \begin{aligned}
 & f = 31.1 + 36.1 + 37.1 + 28.1 + 45.1 = 177
 \end{aligned}$$

(4.11.)

Slabši raspored

kjer sta se v 1. in 2. vrstici glede na izraz (4.10.) “zamenjali” ničla in enica.

Na osnovi primerjave izrazov (4.10.) in (4.11.) opazimo, da je v primeru izraza (4.10.) gotovo boljši raspored delavcev, saj zavzame kriterijska funkcija manjšo vrednost 167 porabljenih časovnih enot, v primeru izraza (4.11.) pa slabši raspored delavcev, saj zavzame kriterijska funkcija večjo vrednost 177 porabljenih časovnih enot.

Kako torej rešimo problem optimalnega razporejanja delavcev na delavna mesta?

Seveda ne bomo šli “na roke” poskušati vse možne kombinacije različnih razporeditev enk in ničel in opazovati, kakšno vrednost zavzema kriterijska funkcija, pač pa bomo to raje prepustili računalniškemu programu.

Zato sledi za rešitev problema, definirane v matematičnem modelu (4.6.), (4.7.), (4.8.):

**Najugodnejši raspored delavcev z najmanj porabljenimi časovnimi enotami za realizacijo vseh opravil je mogoče poiskati z enim izmed računalniških programov za celoštevilsko optimizacijo!**

### Primer 2:

Podjetje izdeluje mize in stole. Za mizo je potrebnih 1 ura dela in 9 kvadratnih enot lesa. Za stol je potrebnih 1 ura dela in 5 kvadratnih enot lesa. V danem trenutku je na razpolago 6 ur dela in 45 kvadratnih enot lesa. Vsaka miza doprinese 8 EUR profita in vsak stol 5 EUR profita. Formulirajte problem na takšen način, da se doseže maksimalen profit podjetja (problem celoštevilskega linearnega programiranja – nastavite matematični model).

### Nastavitev modela:

Definirajmo odločitvene spremenljivke:

$x_1$  = število izdelanih miz  
 $x_2$  = število izdelanih stolov

Matematični model je potem naslednji:

$$\begin{aligned} \max f(x_1, x_2) &= 8x_1 + 5x_2 \\ \text{omejitve:} \\ x_1(1 \text{ ura dela}) + x_2(1 \text{ ura dela}) &\leq 6(\text{ur dela}) \quad (\text{omejitev za ure dela}) \\ 9(\text{enot lesa})x_1 + 5(\text{enot lesa})x_2 &\leq 45(\text{enot lesa}) \quad (\text{omejitev za les}) \\ x_1, x_2 &\geq 0, x_1, x_2 \in Z \end{aligned} \quad (4.12.)$$

kjer lahko celoštevilске spremenljivke za razliko od prejšnjega primera zavzamejo vrednosti iz celega nabora celih števil (seveda v skladu z omejitvami).

Dobimo iz izraza (4.12), ko pokrajšamo enote:

$$\begin{aligned} \max f(x_1, x_2) &= 8x_1 + 5x_2 \\ \text{omejitve:} \\ x_1 + x_2 &\leq 6 \quad (\text{omejitev za ure dela}) \\ 9x_1 + 5x_2 &\leq 45 \quad (\text{omejitev za les}) \\ x_1, x_2 &\geq 0, x_1, x_2 \in Z \end{aligned} \quad (4.13.)$$

Izraz (4.13.) zopet predstavlja nek celoštevilski program, ki ga je potrebno rešiti z enim izmed računalniških programov za celoštevilsko optimizacijo.

Kot se izkaže, je optimalna rešitev tega problema z uporabo „razveji in omeji” metode naslednja:

$$\begin{aligned} x_1^* &= 5 \text{ miz} \\ x_2^* &= 0 \text{ stolov} \\ f(x_1^*, x_2^*) &= f_{\max} = 40 \text{ EUR} \end{aligned} \quad (4.14.)$$

Maksimalen profit 40 EUR se torej v danem trenutku doseže, če podjetje izdela 5 miz in nič stolov.

### **4.3.2 Reševanje celoštevilskih problemov**

Najbolj značilna načina reševanja celoštevilčnih problemov sta naslednja [7, 10, 21, 23]:

- Branch and Bound (razveji in omeji) metoda
- Gomory Cut algoritem (metoda sekajočih ravnin).

V praksi je še posebej priljubljen prvi način, torej “razveji in omeji” metoda. Ta pristop odpravi nekatere pomankljivosti naslednjih dveh pristopov [7]:

- Pristop z “relaksiranim” LP problemom in zaokroževanjem,
- Pristop z izračuni vseh možnosti,

Prvi pristop je namreč nekoliko naiven in nas lahko hitro pripelje do napačnih rešitev. Drugi pristop pa je v praksi preveč potraten, saj je potrebno izračunavati kriterijsko funkcijo za vse kombinacije celih števil danih odločitvenih spremenljivk, ki se nahajajo v območju dopustnih rešitev.

Več o teoriji celoštevilskega programiranja in praktičnih primerih njegove uporabe za reševanje celoštevilčnih problemov s pomočjo »Razveji in omeji« metode lahko bralec zasledi v literaturi [7, 10, 21, 23].

## **4.4 Dinamično programiranje**

Dinamično programiranje je posebna panoga uporabne matematike in je namenjeno za obravnavo mnogofaznih procesov. To so procesi, kjer rezultati prejšnjih faz učinkujejo na naslednje faze. Značilen primer reševanja problemov s pomočjo dinamičnega programiranja je npr. s področja kmetijske proizvodnje, kjer npr. [27]:

- V prvi fazi pridelamo krmo,
- V drugi fazi pa nato pokrmimo živino.

Pri dinamičnem programiranju ne določimo vseh odločitev hkrati, ampak jih določujemo zapovrstjo drugo za drugo. Pri tem pa je potrebno poiskati takšno zaporedje odločitev, s katerim dosežemo optimalen učinek z vidika izbranega kriterija optimalnosti (optimalen plan vodenja procesa)!

Naštejmo nekaj značilnih primerov problemov, ki se lahko rešujejo s pomočjo dinamičnega programiranja [27]:

- Računanje najkrajše (najhitrejše) poti,
- Računanje optimalnega plana proizvodnje,
- Računanje optimalne investicijske politike,
- Računanje optimalnega vlaganja proizvodnih faktorjev,
- Računanje optimalnega plana nadomeščanja strojev,
- Računanje optimalnega plana zalog, itn.

V splošnem lahko glede na naravo problema in njegovih spremenljivk ločimo med determinističnim in stohastičnim dinamičnim programiranjem. Druga ločitev pa lahko poteka tudi glede na metode, ki rešujejo probleme z diskretnimi spremenljivkami, ter tiste, ki rešujejo probleme z zveznimi spremenljivkami.

Za dinamično programiranje velja, da običajno terja dokaj zahtevno numerično računanje, pojmi in matematična simbolika pa so precej komplicirani. Več o teoriji dinamičnega programiranja in praktičnih primerih njegove uporabe lahko bralec zasledi v literaturi [27].

## **4.5 Mrežna optimizacija**

Pri tovrstni problematiki gre za takoimenovano optimizacijo mrežnih struktur [1, 30]. Mreže igrajo pomembno vlogo na področju transporta, komunikacij, produkcije, distribucije, projektnega planiranja, itn. Z njimi lahko na eleganten način vizuelno in konceptualno ilustriramo in reprezentiramo medsebojne odnose in povezave, ki vladajo med posameznimi komponentami nekega sistema.

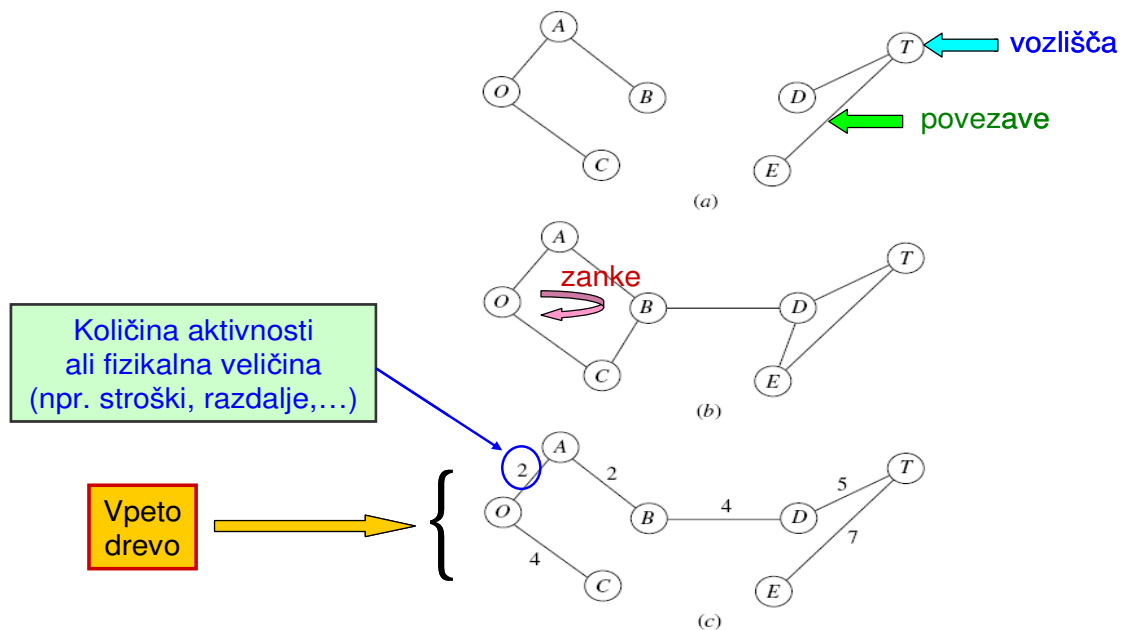
Poglejmo si nekaj osnovnih definicij, ki izhajajo iz topologije mrež in grafov [1, 30]:

- Vozlišča,
- Povezave,
- Mreža ali graf (skupek vozlišč in povezav),
- Drevo (povezan podgraf brez ciklov (zank), ki vsebuje nekatera izmed vozlišč grafa).
- Vpeto drevo (drevo, ki vsebuje vsa vozlišča grafa).
- Minimalno vpeto drevo (vpeto drevo, pri katerem je skupna dolžina povezav med posameznimi pari vozlišč najmanjša).
- Hamiltonov cikel (cikel, ki obiše vsako vozlišče grafa natanko enkrat).

Slika 20 ilustrira nekaj primerov mrež oz. grafov, kjer so npr. razvidna vozlišča, povezave, zanke in vpeto drevo.

Tukaj imamo opravka s takoimenovanim mrežnim modeliranjem logističnih procesov, katerega glavne faze so naslednje:

- Če želimo procesu prirediti graf, ga moramo najprej razčleniti na elementarne aktivnosti.
- Vsakemu njegovemu elementu nato priredimo vozlišče v obliki kroga.
- Poleg tega definiramo tudi medsebojne povezave in merske enote za aktivnosti.



Slika 20: Ilustracija primerov mrež oz. grafov, kjer so razvidna vozlišča, povezave, zanke in vpeto drevo.

Značilni problemi pri mrežni optimizaciji so naslednji [1, 7, 30]:

**1.) Problem minimalno vpetega drevesa (MVD):**

- Želimo realizirati takšno omrežje, ki bo povezal vsa vozlišča grafa brez tvorjenja zank, pri čemer bo vsota povezav najkrajša/najcenejša.

**2.) Problem trgovskega potnika (TP):**

- Želimo povezati vozlišča grafa tako, da obiščemo vsa vozlišča natanko enkrat in se vrnemo v izhodiščno vozlišče po najkrajši skupni poti.

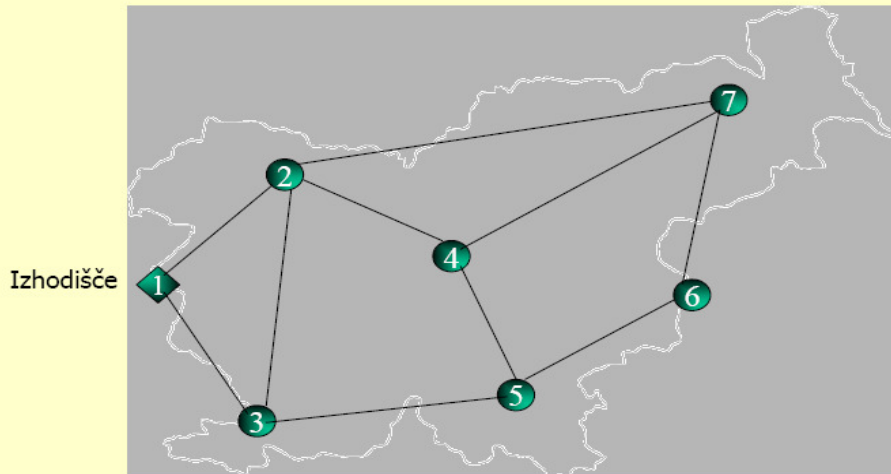
**3.) Problem maksimalnega pretoka:**

- Želimo določiti takšen razpored razpošiljanja materiala med vhodnim in izhodnim vozliščem, pri katerem se bo prenesla maksimalna količina tega materiala (maksimalen promet skozi omrežje med vhom in izhodom).

Kar se tiče problematike minimalno vpetih dreves, lahko naštejemo nekaj značilnih primerov iz logistične prakse [7]:

- Načrtovanje sistema transportnih poti, ki bo povezal vse lokacije (vozlišča) s čim manjšimi stroški (stroški izgradnje in vzdrževanja poti in vmesnih vozlišč ter pripadajoče infrastrukture bodo najmanjši).
- Načrtovanje sistema kanalov za izsuševanje, ki se bo povezal z vsemi lokacijami (vozlišči) pri čim manjših stroških (stroški izgradnje in vzdrževanja kanalov bodo najmanjši).
- Načrtovanje sistema optičnih kablov, ki bo povezal vsa vozlišča z izhodiščem, pri čemer bo skupna dolžina kablov najmanjša (slika 21).

## Minimalno vpeto drevo



Problem: npr. kako položiti minimalno dolžino optičnega kabla tako, da so vsa vozlišča povezana z izhodiščem?

56

Slika 21: Ilustracija problema načrtovanja sistema optičnih kablov, ki bo povezal vsa vozlišča z izhodiščem, pri čemer bo skupna dolžina kablov najmanjša.

Obstajajo različne metode, kako reševati probleme minimalno vpetega drevesa, maksimalnega pretoka in trgovskega potnika pri problematiki mrežne optimizacije.

Podrobnosti o tovrstnih metodah in praktičnih primerih reševanja problematike mrežne optimizacije lahko bralec zasledi v literaturi [1, 7, 30].

S predstavitvijo mrežne optimizacije smo končali kratek pregled najbolj značilnih optimizacijskih metod v logistiki. V naslednjih poglavjih pa bomo svojo pozornost usmerili le še na obravnavo nelinearnega in geometrijskega programiranja, ki predstavljata glavno snov tega dela.

Da pa bi sploh lahko začeli kvalitetno obravnavati tovrstne vsebine, moramo najprej ponoviti nekatere matematične osnove, ki so ključne pri razumevanju obravnave teh vsebin.



## 5 NEKATERE MATEMATIČNE OSNOVE

V tem poglavju se bomo seznanili s pregledom elementarnih matematičnih pojmov in rezultatov, ki so bistveni za razumevanje problematike nelinearnega programiranja.

Pri obravnavi bomo naleteli tudi na vektorje in matrike. Pri tem moramo opozoriti, da se prve v literaturi velikokrat označuje z malimi črkami in dodatno podčrtano oznako »-«, druge pa z velikimi črkami in dvema podčrtanima črtama, da se jih ustrezno loči od skalarjev. Zaradi poenostavljenosti obravnave podčrtovanja sicer večinoma ne bomo uporabljali, vendar pa moramo imeti to dejstvo vedno na umu in ustrezno ločevati vektorje oz. matrike od skalarjev.

### 5.1 Kvadratne forme matrik

Denimo imamo opravka z naslednjima dvema vektorjema:

$$\begin{aligned}x &= (x_i), \dots, i = 1, 2, \dots, n \\y &= (y_i), \dots, i = 1, 2, \dots, n\end{aligned}\tag{5.1.}$$

Njun skalarni produkt potem lahko zapišemo v naslednji obliki:

$$(x, y) = \sum_{i=1}^n x_i y_i\tag{5.2.}$$

Če velja relacija  $y = \mathbf{A} \cdot x$ , potem lahko skalarni produkt zapišemo na naslednji način:

$$\begin{aligned}f(x) &= (x, Ax), \text{ kjer } A = (a_{ij})_{n \times n} \\f(x) &= x^T (Ax) = \sum_{i=1}^n x_i (Ax)_i = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j\end{aligned}\tag{5.3.}$$

Funkcija  $f(x)$  se v tem primeru imenuje takoiimenovana kvadratna forma matrike  $A$  in spada med najenostavnejše nelinearne funkcije [17].

Kvadratne forme se glede na predznak delijo v naslednje štiri skupine [17]:

1.  $(x, Ax) \geq 0, \forall x$ .....*nenegativna forma*;
  2.  $(x, Ax) > 0, \forall x \neq 0$ .....*pozitivna forma*;
  3.  $(x, Ax) \leq 0, \forall x$ .....*nepozitivna forma*;
  4.  $(x, Ax) < 0, \forall x \neq 0$ .....*negativna forma*;
- (5.4.)

Pri tej delitvi za matriko A velja [17]:

1. A.....*pozitivno semidefinitna matrika*;
  2. A.....*pozitivno definitna matrika*;
  3. A.....*negativno semidefinitna matrika*;
  4. A.....*negativno definitna matrika*;
- (5.5.)

Kot bomo kasneje videli, je ugotavljanje tipa definitnosti določene matrike zelo pomembno pri ugotavljanju tipa oz. narave ekstreme nelinearne funkcije. Pri tem je še posebej pomembno ugotoviti, kdaj je obravnavana matrika  $A = (a_{ij})$  pozitivno definitna.

Dokazati se da, da je pogoj za pozitivno definitnost matrike A izpolnjen tedaj, ko velja takoimenovana Sylvestrova lema, ki se glasi [17]:

$$a_{11} > 0, \quad \det \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} > 0, \quad \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} > 0, \dots, \quad \det A > 0 \quad (5.6.)$$

Torej, simetrična matrika n-tega reda  $A = (a_{ij})$  je pozitivno definitna le v primeru, ko so vsi glavni minorji matrike A pozitivni.

## 5.2 Odvodi funkcij, gradient in Hessejeva matrika

Operatorji odvajanja funkcij so zaradi več razlogov izredno pomembni pri nelinearnem programiranju. Odvodi namreč opisujejo obnašanje funkcij in jih na ta način klasificirajo v linearne, konveksne, konkavne, itn.

Poleg tega nam služijo tudi za formuliranje pogojev optimalnosti, kar pomeni preverjanje, če je neka dopustna rešitev tudi optimalna ali pa ne. Če neka dopustna rešitev ni optimalna, pa nam koristijo tudi za formuliranje numeričnih metod, ki nam pomagajo poiskati boljše dopustne rešitve in morebiti tudi optimalno rešitev.

Običajno imamo pri nelinearnem programiranju opravka le z eno skalarno ciljno (ponavadi nelinearno) funkcijo:

$$f(x) = f\left[(x_1, x_2, \dots, x_n)^T\right] \quad (5.7.)$$

kjer je  $x$  vektor  $n$  odločitvenih spremenljivk.

Če funkcijo (5.7.) parcialno odvajamo po vseh spremenljivkah, pridemo do pojma gradienta (gradientnega vektorja), ki je sestavljen iz parcialnih odvodov in se glasi [17]:

$$\nabla f(x) = \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \quad \dots \quad \frac{\partial f(x)}{\partial x_n} \right]^T \quad (5.8.)$$

Velikokrat bomo pri obravnavi nelinearnega programiranja uporabljali tudi takoimenovano Hessovo matriko drugih parcialnih odvodov  $H$ , ki jo dobimo s ponovnim odvajanjem gradienta (5.8.) in se glasi [17]:

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}. \quad (5.9.)$$

### **Primer:**

Dano imamo skalarno ciljno funkcijo treh odločitvenih spremenljivk:

$$f(x_1, x_2, x_3) = 4x_1^2 + 6x_2^2 + 6x_3^2 + 6x_1x_2 - 4x_1x_3 + 4x_2x_3 \quad (5.10.)$$

Poiščite njen gradient in Hessovo matriko!

Gradientni vektor se glasi:

$$\begin{aligned} \nabla f(x) &= \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \quad \frac{\partial f(x)}{\partial x_3} \right] = \\ &= [8x_1 + 6x_2 - 4x_3 \quad 12x_2 + 6x_1 + 4x_3 \quad 12x_3 - 4x_1 + 4x_2] \end{aligned} \quad (5.11.)$$

Hessova matrika pa se glasi:

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f(x)}{\partial x_3 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_3 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 8 & 6 & -4 \\ 6 & 12 & 4 \\ -4 & 4 & 12 \end{bmatrix} \quad (5.12.)$$

Poznavanje prvih in drugih odvodov funkcij lahko znatno olajša obravnavo problema pri nelinearnem programiranju. Za boljše razumevanje slednjega pa si v nadaljevanju pogledimo še pojem konveksnosti funkcij.

### 5.3 Konveksnost funkcij

Konveksnost je pomemben pojem v optimizaciji, Če je namreč neka funkcija ali množica konveksna, je mogoče to lastnost koristno uporabiti v algoritmi optimiranja.

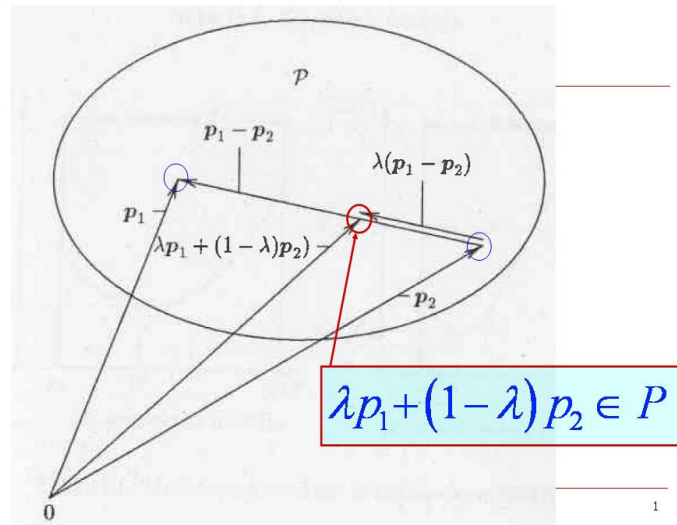
Neka **množica** je konveksna, če velja trditev [13]:

*Množica  $P$  je konveksna, če velja za katerikoli njeni točki*

*$p_1$  in  $p_2 \in P$  in za vsak skalar  $\lambda \in [0,1]$  trditev:*

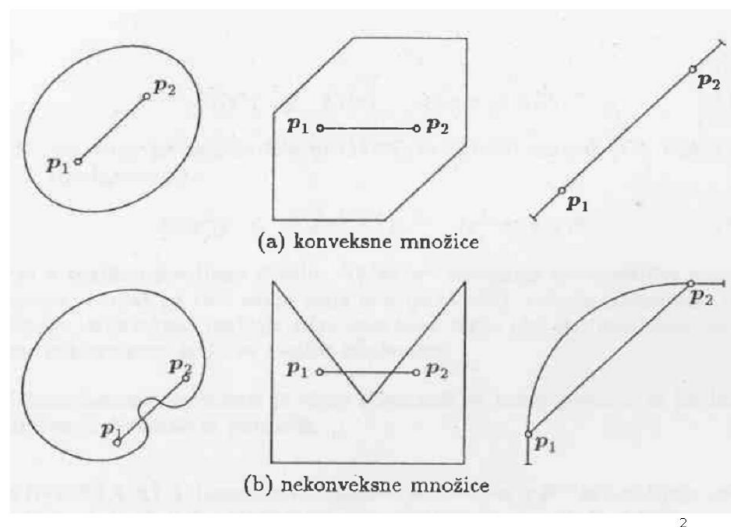
$$\lambda p_1 + (1 - \lambda) p_2 \in P$$

To pomeni, da je množica konveksna, če je zveznica, ki povezuje dve točki množice, v celoti vsebovana v tej množici! Konveksna množica je ilustriрана na sliki 22 [13].



Slika 22: Ilustracija konveksne množice

Zgledi nekaterih konveksnih in nekonveksnih množic pa so podani na sliki 23 [13].



Slika 23: Zgledi konveksnih in nekonveksnih množic

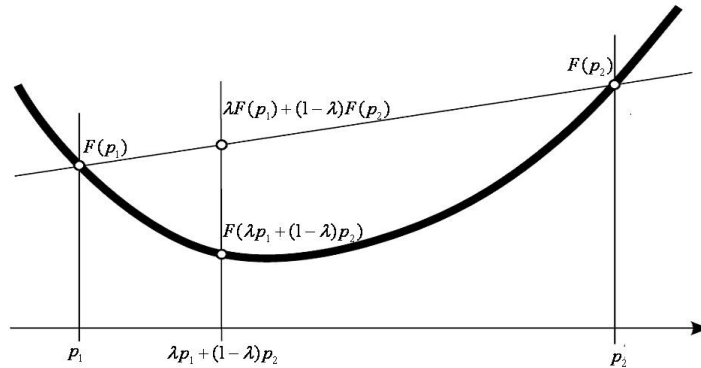
Podobna definicija kot pri konveksnih množicah velja tudi za konveksnost funkcij.

Neka **funkcija** je konveksna, če velja trditev [13]:

*Funkcija  $F(p)$  je konveksna, če velja za katerikoli njeni točki  $p_1$  in  $p_2 \in P$  in za vsak skalar  $\lambda \in [0,1]$  relacija:*  

$$F(\lambda p_1 + (1-\lambda)p_2) \leq \lambda \cdot F(p_1) + (1-\lambda) \cdot F(p_2)$$

Če v gornji trditvi ni enačaja, pač pa le neenačaj, imamo opravka s strogo konveksno funkcijo. Konveksna funkcija je ilustrirana na sliki 24.



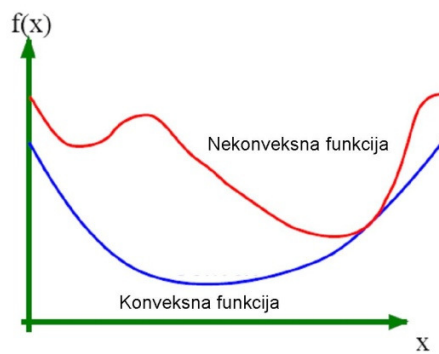
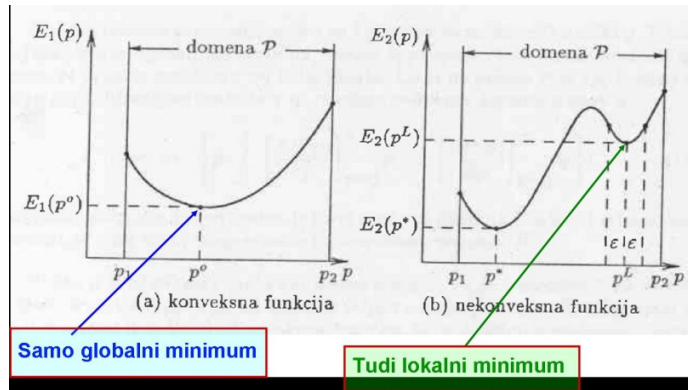
Slika 24: Konveksna funkcija

Za konveksne funkcije  $f$  veljajo nekatere pomembne lastnosti [17]:

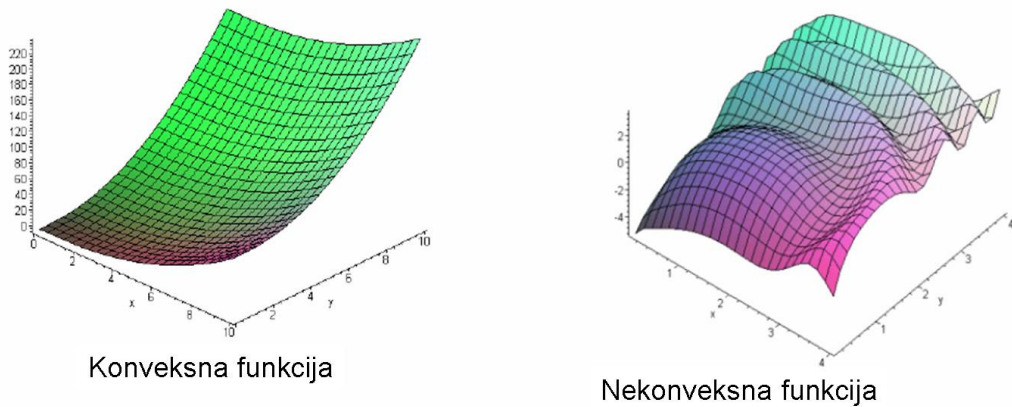
- Če je pri njih v neki točki gradient  $\nabla f = 0$ , potem je tam gotovo globalni minimum (izpolnjena sta potreben in zadosten pogoj).
- Pri poljubnih funkcijah pa  $\nabla f = 0$  pomeni zgolj potreben pogoj (stacionarno točko), ne pa tudi zadosten pogoj za globalni minimum.
- Konveksne funkcije imajo samo globalni minimum, nekonveksne funkcije pa imajo lahko še lokalne minimume!
- Pri konveksnih funkcijah je stacionarna točka gotovo tudi (globalni) minimum!

Podobna analogija seveda velja tudi za obravnavo maksimumov in konkavne funkcije. Bistvene razlike med konveksno in nekonveksno funkcijo v enodimenzionalnem prostoru ilustrira slika 25, v dvodimenzionalnem prostoru pa slika 26.

V splošnem velja, da moramo poiskati Hessovo matriko drugih odvodov obravnavane funkcije, da bi lahko ugotovili naravo stacionarnih točk. Za to matriko moramo preveriti tudi njeno definitnost (npr. s Sylvestrovo lemo, glej izraze (5.6.)) in če gre npr. za pozitivno definitnost, potem je v stacionarni točki gotovo lokalni minimum.



Slika 25: Razlike med konveksno in nekonveksno funkcijo v enodimenzionalnem prostoru



Slika 26: Razlike med konveksno in nekonveksno funkcijo v dvodimenzionalnem prostoru

## 5.4 Klasična analitična optimizacija in analitičen pristop iskanja ekstremov

V tem poglavju je podan kratek pregled najpomembnejših klasičnih rezultatov pri analitičnem iskanju optimumov funkcije več spremenljivk v okviru teorije klasične optimizacije.

### 5.4.1 Stacionarne točke

Denimo imamo opravka s funkcijo  $f(x) = f\left[(x_1, x_2, \dots, x_n)^T\right]$ , ki je **odvedljiva** in definirana na celotnem prostoru realnih števil. Potem za potreben pogoj za nastop lokalnega optimuma funkcije velja naslednje:

Če je vektor  $x^*$  optimalen vektor odločitvenih spremenljivk (lokalni minimum ali maksimum), tedaj velja naslednji izraz [17]:

$$\nabla f(x^*) = 0 \tag{5.13.}$$

torej je  $x^*$  gotovo stacionarna točka (seveda v obratni smeri za to trditev ni gotovo, da velja). S pomočjo izraza (5.13.) na analitičen način poiščemo stacionarne točke. Pri tem velja, da so le-te za konveksne funkcije gotovo tudi ekstremi (bodisi minimumi ali maksimumi), v primeru nekonveksnih funkcij pa ne moremo kar tako sklepati na naravo stacionarnih točk, pač pa moramo preveriti definitnost Hessove matrike.

Pri tem velja naslednje:

Denimo imamo opravka s funkcijo  $f(x) = f\left[(x_1, x_2, \dots, x_n)^T\right]$ , ki je **dvakrat odvedljiva**, zvezna in definirana na celotnem prostoru realnih števil. Potem za potreben in zadosten pogoj za nastop lokalnega optimuma funkcije velja naslednje [17]:

- Če je vektor  $x^*$  optimalen vektor odločitvenih spremenljivk (lokalni minimum), tedaj velja izraz (5.13.) in je Hessova matrika  $\nabla^2 f(x^*)$  pozitivno semidefinitna.
- Velja tudi obratno: Če velja izraz (5.13.) in je Hessova matrika  $\nabla^2 f(x^*)$  pozitivno semidefinitna za vsak  $x$  v neki okolici  $N(x^*)$ , potem je vektor  $x^*$  optimalen vektor odločitvenih spremenljivk (lokalni minimum).



Zgornji trditvi seveda veljata tudi, če obravnavamo lokalne maksimume in imamo opravka z negativno semidefinitno Hessovo matriko.

Matematično najbolj korektna pot za ugotavljanje definitnosti Hessove matrike bi bila s pomočjo Sylvestrove leme (glej izraze (5.6.)).

V praksi pa se pogostokrat poslužujemo nekoliko enostavnejšega načina za preverjanje narave stacionarne točke, ki se glasi [17]:

$$\begin{aligned}
 \det H(x^*) < 0 & \Rightarrow x^* \text{ je sedlo} \\
 \det H(x^*) > 0 \text{ in } \frac{\partial^2 f}{\partial x_1^2}(x^*) > 0 & \Rightarrow x^* \text{ je minimum} \\
 \det H(x^*) > 0 \text{ in } \frac{\partial^2 f}{\partial x_1^2}(x^*) < 0 & \Rightarrow x^* \text{ je maksimum}
 \end{aligned}
 \tag{5.14.}$$

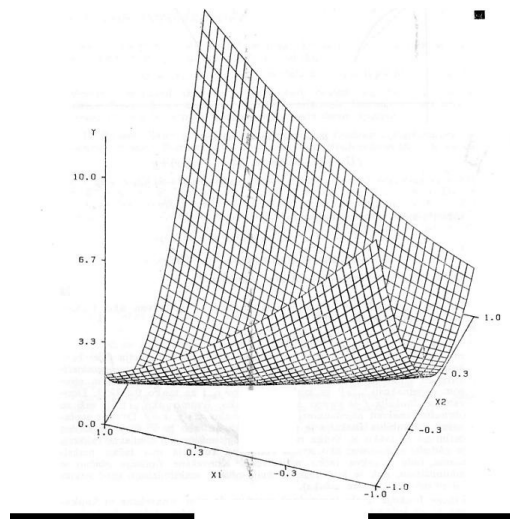
V nadaljevanju si pogledjmo dva primera, kjer je razvidno, kako poskušamo na konveksni in nekonveksni funkciji analitično poiskati stacionarno točko.

### **1. Primer konveksne funkcije:**

Dano imamo naslednjo konveksno nelinearno funkcijo:

$$f(x_1, x_2) = x_1^2 + 5x_2^2 + 4x_1x_2 \tag{5.15.}$$

ki je ponazorjena na sliki 27. Za dano funkcijo poiščite stacionarno točko!



Slika 27: Primer konveksne funkcije

Gradient te funkcije se glasi:

$$\nabla f(x) = \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \right] = [2x_1 + 4x_2 \quad 10x_2 + 4x_1] \quad (5.16.)$$

Stacionarno točko na osnovi (5.13.) dobimo tako, da izraz (5.16.) enačimo z ničelnim vektorjem:

$$\begin{aligned} \nabla f(x) = 0 &\Rightarrow 2x_1 + 4x_2 = 0; \\ &10x_2 + 4x_1 = 0 \end{aligned} \quad (5.17.)$$

Ko rešimo sistem (5.17.), za rešitev dobimo naslednjo stacionarno točko:

$$x_1^* = 0; \quad x_2^* = 0 \quad (5.18.)$$

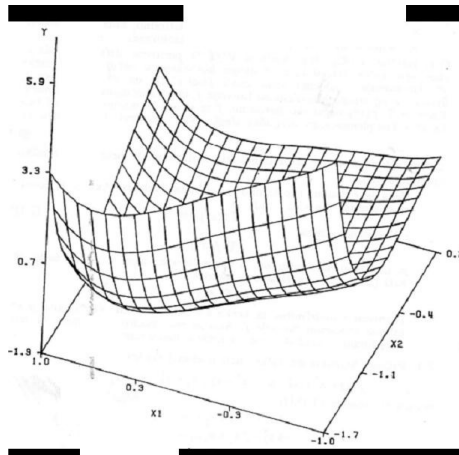
ki je hkrati tudi globalni minimum, saj gre za konveksno funkcijo.

## **2. Primer nekonveksne funkcije:**

Dano imamo naslednjo nekonveksno nelinearno funkcijo:

$$f(x_1, x_2, x_3) = x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3 \quad (5.19.)$$

ki je ponazorjena na sliki 28. Za dano funkcijo poiščite stacionarno točko!



Slika 28: Primer nekonveksne funkcije

Gradient te funkcije se glasi:

$$\begin{aligned}\nabla f(x) &= \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \quad \frac{\partial f(x)}{\partial x_3} \right] = \\ &= \left[ 4x_1^3 + 3x_1^2 - 1 + x_2x_3 \quad 4x_2^3 - 2x_2 + 1 + x_1x_3 \quad 2x_3 - 1 + x_1x_2 \right]\end{aligned}\tag{5.20.}$$

Stacionarno točko na osnovi (5.13.) dobimo tako, da izraz (5.20.) enačimo z ničelnim vektorjem:

$$\begin{aligned}\nabla f(x) = 0 \quad \Rightarrow \quad & 4x_1^3 + 3x_1^2 - 1 + x_2x_3 = 0 \\ & 4x_2^3 - 2x_2 + 1 + x_1x_3 = 0 \\ & 2x_3 - 1 + x_1x_2 = 0\end{aligned}\tag{5.21.}$$

Ko rešimo sistem (5.21.), za rešitev dobimo naslednjo stacionarno točko:

$$x_1^* = 0.571; \quad x_2^* = -0.940; \quad x_3^* = 0.768\tag{5.22.}$$

Ker imamo tokrat opravka z nekonveksno funkcijo, ne vemo, kakšna je narava stacionarne točke. Da bi slednjo ugotovili, bi morali najprej poiskati Hessovo matriko in nato npr. preveriti izraze (5.14.)

#### **5.4.2 Razlika med analitičnim in numeričnim iskanjem ekstrema**

Na kratko podajmo še opis bistvenih razlik med analitičnim in numeričnim iskanjem ekstremov. Običajno se v praksi ekstremi ne iščejo na način, kot ga opisuje izraz (5.13.).

Razlog za to bi lahko našli zlasti v naslednjih dveh dejstvih:

- Ker imamo pri tvorjenju izraza (5.13.) ponavadi opravka z zahtevnimi nelinearnimi sistemi enačb, ki jih moramo rešiti, ter lahko naletimo na nerešljive težave pri računanju.
- Ker bi bile lahko mnoge od teh rešitev stacionarne točke, ki pa ne bi bile hkrati tudi lokalni optimumi in bi bil velik del napora pri reševanju zaman.

Zaradi naštetih razlogov se torej v praksi raje poslužujemo numeričnih metod, kar velja tudi v primeru konveksnih funkcij.

Vendar pa, kot bomo videli v poglavju 8, obstajajo tudi določene situacije v praksi, ko z analitičnim pristopom vseeno uspemo poiskati ekstreme funkcij.

V nadaljevanju si bomo pogledali še dva pristopa, ki jih uporabljamo pri analitični optimizaciji, če imamo opravka tudi z določenimi omejitvami.

### 5.4.3 Pretvorba na problem brez omejitev z eliminacijo spremenljivk

Denimo imamo opravka z naslednjim optimizacijskim problemom z omejitvami:

$$\begin{array}{l}
 \text{PROBLEM } P \\
 \text{opt } f(x) \dots \dots \dots \text{funkcija cilja} \\
 \text{p.o. (pri omejitvah)} \\
 \left. \begin{array}{l}
 h^1(x) = 0 \\
 h^2(x) = 0 \\
 \dots \\
 h^m(x) = 0
 \end{array} \right\} \dots \dots \dots \text{funkcije omejitev}
 \end{array} \tag{5.23.}$$

kjer ima vektor  $x$   $n$  komponent:  $x = (x_1, x_2, \dots, x_n)^T$ .

Očitno je takšen problem zaradi prisotnosti omejitev mnogo težje rešiti kot v primeru, ko omejitev ne bi bilo. Ideja pri pristopu z eliminacijo spremenljivk je v tem, da bi poskušali omejitve nekako odpraviti, s tem, da bi jih uvrstili v ciljno funkcijo. Seveda je najbolj idealna varianta ta, da bi nam na takšen način uspelo odpraviti vse omejitve in problem (5.23.) pretvoriti na problem brez vsakršnih omejitev.

Za ilustracijo te ideje si pogledjmo naslednji primer.

#### Primer:

Poiskati želimo optimum naslednje funkcije:

$$f(x_1, x_2) = 4x_1^2 + 5x_2^2 \tag{5.24.}$$

ki ima naslednjo omejitev:

$$h(x) = 2x_1 + 3x_2 - 6 = 0 \tag{5.25.}$$

Rešitev: Najprej izrazimo eno spremenljivko v izrazu (5.25.) z drugo na naslednji način:

$$x_1 = \frac{3}{2}(2 - x_2) \tag{5.26.}$$

Izraz (5.26.) nato vstavimo v funkcijo (5.24.), ki preide v obliko:

$$f(x_2) = 4 \left[ \frac{3}{2}(2 - x_2) \right]^2 + 5x_2^2 = 9(4 - 4x_2 + x_2^2) + 5x_2^2 \quad (5.27.)$$

pri čemer dobimo:

$$f(x_2) = 14x_2^2 - 36x_2 + 36 \quad (5.28.)$$

Ker smo se omejitve (5.25.) uspeli v izrazu (5.28.) znebiti in dobiti le funkcijo ene spremenljivke, smo očitno optimizacijski problem z omejitvijo prevedli na problem brez omejitve. Tega nato rešujemo po ustaljeni poti kot je to običajno za probleme brez omejitev.

Pri tem dobimo:

$$\frac{df(x_2)}{dx_2} = 28x_2 - 36 = 0 \Rightarrow x_2^* = \frac{9}{7} \quad (5.29.)$$

$$x_1^* = \frac{3}{2} \left( 2 - \frac{9}{7} \right) = 3 - \frac{27}{14} = \frac{15}{14}$$

Ker je drugi odvod funkcije (5.28.) pozitiven, je dobljena rešitev:

$$x_1^* = \frac{15}{14}, \quad x_2^* = \frac{9}{7} \quad (5.28.)$$

očitno minimum.

Žal pa v mnogo primerih problema z omejitvami ni tako enostavno prevesti na problem brez omejitev, kot je bilo prikazano v tem poglavju. To še posebej velja pri nelinearnih funkcijah omejitev, ko se je potrebno poslužiti kakšnega drugega, bolj sofisticiranega načina za reševanje problemov z omejitvami. O tem bo govora v naslednjem poglavju.

#### **5.4.4 Pretvorba na problem brez omejitev z Lagrangeovimi multiplikatorji**

Denimo imamo zopet opravka z optimizacijskim problemom z omejitvami, ki ga opisujejo izrazi (5.23.), kjer je  $x = (x_1, x_2, \dots, x_n)^T$  vektor  $n$  odločitvenih spremenljivk.

Vpeljimo takoimenovano Lagrangeovo funkcijo, ki jo definiramo na naslednji način [17]:

$$L(x, \mu) = f(x) + \sum_{k=1}^m \mu_k h^k(x) \quad (5.29.)$$

kjer so  $\mu = (\mu_1, \mu_2, \dots, \mu_m)^T$  .....Lagrangeovi multiplikatorji . Potrebni pogoj za lokalni optimum se potem lahko izrazi na naslednji način:

Če je vektor  $x^*$  kandidat za optimalen vektor odločitvenih spremenljivk (lokalni minimum ali maksimum), potem obstaja takšen vektor  $\mu^* = (\mu_1^*, \mu_2^*, \dots, \mu_m^*)^T$  , da velja naslednji izraz [17]:

$$\nabla L(x^*, \mu^*) = 0, \text{ pri čemer tvorimo } \frac{\partial}{\partial}() \text{ nad } L \text{ za vse } x_1, \dots, x_n, \mu_1, \dots, \mu_m \quad (5.30.)$$

oziroma:

$$\nabla L(x^*, \mu^*) = \left[ \begin{array}{ccccccc} \frac{\partial L(x, \mu)}{\partial x_1} & \frac{\partial L(x, \mu)}{\partial x_2} & \dots & \frac{\partial L(x, \mu)}{\partial x_n} & \frac{\partial L(x, \mu)}{\partial \mu_1} & \frac{\partial L(x, \mu)}{\partial \mu_2} & \dots & \frac{\partial L(x, \mu)}{\partial \mu_m} \end{array} \right]_{x^*, \mu^*} = 0 \quad (5.31.)$$

Tvorimo torej sistem enačb na osnovi izraza (5.30.) in ga poskušamo rešiti, pri čemer dobimo rešitev  $(x^*, \mu^*)$ . Ta rešitev je stacionarna točka Lagrangeove funkcije (5.29.), pri čemer je izpolnjen potreben pogoj za nastop lokalnega optimuma v originalnem problemu (5.23.). Seveda je ta rešitev samo kandidat za lokalni optimum. Da pa se prepričamo, da je to res lokalni optimum, bi pa bilo potrebno preveriti še kakšnega od zadostnih pogojev za nastop optimuma.

Pokazati se da, da je za izpolnitev zadostnega pogoja za nastop lokalnega optimuma potrebno preveriti definitnost naslednje matrike [17]:

$$\nabla_x^2 L(x^*, \mu^*) = \left[ \begin{array}{cccc} \frac{\partial^2 L(x, \mu)}{\partial x_1^2} & \frac{\partial^2 L(x, \mu)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 L(x, \mu)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 L(x, \mu)}{\partial x_2 \partial x_1} & \frac{\partial^2 L(x, \mu)}{\partial x_2^2} & \dots & \frac{\partial^2 L(x, \mu)}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 L(x, \mu)}{\partial x_n \partial x_1} & \frac{\partial^2 L(x, \mu)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 L(x, \mu)}{\partial x_n^2} \end{array} \right]_{x = x^*, \mu = \mu^*} \quad (5.32.)$$

### Primer:

Potrebno je najti ekstrem naslednje funkcije z omejitvami (originalni problem P):

$$\begin{aligned}
f(x) &= 7x_1 - 6x_2 + 4x_3 \\
p.o. & \\
h^1(x) &= x_1^2 + 2x_2^2 + 3x_3^2 - 1 = 0 \\
h^2(x) &= 5x_1 + 5x_2 - 3x_3 - 6 = 0
\end{aligned}
\tag{5.33.}$$

Lagrangeova funkcija se glasi:

$$\begin{aligned}
L(x, \mu) &= f(x) + \mu_1 h^1(x) + \mu_2 h^2(x) \\
L(x, \mu) &= 7x_1 - 6x_2 + 4x_3 + \mu_1 (x_1^2 + 2x_2^2 + 3x_3^2 - 1) + \\
&\quad + \mu_2 (5x_1 + 5x_2 - 3x_3 - 6)
\end{aligned}
\tag{5.34.}$$

Nastavimo potreben pogoj za nastop optimuma originalnega problema P na osnovi izrazov (5.30.) oz. (5.31.). Dobimo:

$$\begin{aligned}
\frac{\partial}{\partial x_1} L(x, \mu) &= 7 + 2\mu_1 x_1 + 5\mu_2 = 0 \\
\frac{\partial}{\partial x_2} L(x, \mu) &= -6 + 4\mu_1 x_2 + 5\mu_2 = 0 \\
\frac{\partial}{\partial x_3} L(x, \mu) &= 4 + 6\mu_1 x_3 - 3\mu_2 = 0 \\
\frac{\partial}{\partial \mu_1} L(x, \mu) &= h^1(x) = x_1^2 + 2x_2^2 + 3x_3^2 - 1 = 0 \\
\frac{\partial}{\partial \mu_2} L(x, \mu) &= h^2(x) = 5x_1 + 5x_2 - 3x_3 - 6 = 0
\end{aligned}
\tag{5.35.}$$

Dobimo torej naslednji sistem enačb:

$$\begin{bmatrix} 2x_1^* & 5 \\ 4x_2^* & 5 \\ 6x_3^* & -3 \end{bmatrix} \begin{bmatrix} \mu_1^* \\ \mu_2^* \end{bmatrix} = \begin{bmatrix} -7 \\ 6 \\ -4 \end{bmatrix}$$

$$(x_1^*)^2 + 2(x_2^*)^2 + 3(x_3^*)^2 - 1 = 0$$

$$\tag{5.36.}$$

$$5(x_1^*) + 5(x_2^*) - 3(x_3^*) - 6 = 0$$

katerega iskanje rešitev terja precej naporov.

Po daljšem izračunu dobimo naslednji dve rešitvi:

$$\begin{aligned} 1. \text{ rešitev: } x_p^* &= \begin{bmatrix} x_{p1}^* \\ x_{p2}^* \\ x_{p3}^* \end{bmatrix} = \begin{bmatrix} 0.535 \\ 0.535 \\ -0.219 \end{bmatrix}, \mu_p^* = \begin{bmatrix} \mu_{p1}^* \\ \mu_{p2}^* \end{bmatrix} = \begin{bmatrix} -12.187 \\ 4.006 \end{bmatrix} \\ 2. \text{ rešitev: } x_d^* &= \begin{bmatrix} x_{d1}^* \\ x_{d2}^* \\ x_{d3}^* \end{bmatrix} = \begin{bmatrix} 0.947 \\ 0.207 \\ -0.077 \end{bmatrix}, \mu_d^* = \begin{bmatrix} \mu_{d1}^* \\ \mu_{d2}^* \end{bmatrix} = \begin{bmatrix} 12.187 \\ -3.216 \end{bmatrix} \end{aligned} \quad (5.37.)$$

Preveriti moramo še zadosten pogoj za nastop ekstrema. V ta namen moramo preveriti definitnost matrike (5.32.). Prvo izračunamo vektor prvih odvodov:

$$\begin{aligned} \nabla_x L(x, \mu) &= \left[ \frac{\partial L(x, \mu)}{\partial x_1} \quad \frac{\partial L(x, \mu)}{\partial x_2} \quad \frac{\partial L(x, \mu)}{\partial x_3} \right] = \\ &= [7 + 2\mu_1 x_1 + 5\mu_2 \quad -6 + 4\mu_1 x_2 + 5\mu_2 \quad 4 + 6\mu_1 x_3 - 3\mu_2] \end{aligned} \quad (5.38.)$$

Nato pa matriko drugih odvodov:

$$\nabla_x^2 L(x, \mu) = \left[ \frac{\partial \nabla_x L(x, \mu)}{\partial x_1} \quad \frac{\partial \nabla_x L(x, \mu)}{\partial x_2} \quad \frac{\partial \nabla_x L(x, \mu)}{\partial x_3} \right] = \begin{bmatrix} 2\mu_1 & 0 & 0 \\ 0 & 4\mu_1 & 0 \\ 0 & 0 & 6\mu_1 \end{bmatrix}. \quad (5.39.)$$

Če vstavimo v matriko (5.39.) 1. rešitev iz (5.37.), dobimo:

$$\nabla_x^2 L(x_p^*, \mu_p^*) = \begin{bmatrix} 2\mu_{p1}^* & 0 & 0 \\ 0 & 4\mu_{p1}^* & 0 \\ 0 & 0 & 6\mu_{p1}^* \end{bmatrix} = \begin{bmatrix} -24.374 & 0 & 0 \\ 0 & -48.748 & 0 \\ 0 & 0 & -73.122 \end{bmatrix}. \quad (5.40.)$$



Odtod sledi za 1. rešitev (na osnovi Sylvestrove leme (5.6.)):

$$\nabla_x^2 L(x_p^*, \mu_p^*) \rightarrow \textit{negativno definitna} \Rightarrow x_p^* = \begin{bmatrix} 0.535 \\ 0.535 \\ -0.219 \end{bmatrix} = \textit{max rešitev (P)} \quad (5.41.)$$

Če vstavimo v matriko (5.39.) 2. rešitev iz (5.37.), dobimo:

$$\nabla_x^2 L(x_d^*, \mu_d^*) = \begin{bmatrix} 2\mu_{d1}^* & 0 & 0 \\ 0 & 4\mu_{d1}^* & 0 \\ 0 & 0 & 6\mu_{d1}^* \end{bmatrix} = \begin{bmatrix} 24.374 & 0 & 0 \\ 0 & 48.748 & 0 \\ 0 & 0 & 73.122 \end{bmatrix} \quad (5.42.)$$

Odtod sledi za 2. rešitev (na osnovi Sylvestrove leme (5.6.)):

$$\nabla_x^2 L(x_d^*, \mu_d^*) \rightarrow \textit{pozitivno definitna} \Rightarrow x_d^* = \begin{bmatrix} 0.947 \\ 0.207 \\ -0.077 \end{bmatrix} = \textit{min rešitev (P)} \quad (5.43.)$$

Končen sklep: Originalni problem (5.33.) ima 2 optimuma, ki nastopita v točkah (5.41.) in (5.43.), pri čemer je prva točka lokalni maksimum, druga točka pa lokalni minimum.

## 6 NELINEARNA NUMERIČNA OPTIMIZACIJA BREZ OMEJITEV

Mnogi problemi optimizacije so po naravi nelinearni. To pomeni, da so bodisi ciljna funkcija, bodisi omejitve, ali celo oboje, nelinearni. Problem je nelinearen že tedaj, če je ena sama omejitev nelinearna.

Pri linearnem programiranju (LP) je simpleksna metoda univerzalna metoda za reševanje linearnih programov. Žal pri nelinearnem programiranju (NLP) ni tako in ni univerzalne metode za reševanje nelinearnih programov. Zato je zaželjeno, da se tovrstna tematika поблиže spozna, vsaj kar se tiče osnovnih teoretičnih rezultatov in izbranih numeričnih metod.

Nekateri klasični problemi NLP so:

- Nelinearni transportni problem (npr. stroški so korenska funkcija količine).
- Nelinearni problem optimalne razporeditve virov.
- Reševanje nelinearnih regresijskih problemov z metodami nelinearne optimizacije za potrebe načrtovanja bodočih aktivnosti .
- Nelinearni problem izbora ponudbe.
- Nelinearni problemi v proizvodnji.
- Itn...

Nekatere splošne lastnosti NLP pa so:

- NLP spada v skupino determinističnih metod za reševanje statičnih problemov.
- Velikokrat je potrebno za določen problem, glede na nelinearno naravo, prilagoditi katero izmed metod optimiranja ali jo celo na novo razviti!
- Še vedno obstaja cela vrsta težko rešljivih ali celo nerešljivih nelinearnih problemov, saj še niso razviti učinkoviti algoritmi.
- Zaradi naštetih razlogov se reševanje večine NLP problemov obravnava kot raziskovalna in ne rutinska naloga!

Še nekatere značilnosti obravnave NLP v tem delu:

- Obravnavali bomo le zvezne realne spremenljivke.
- Prostor obravnave bo lahko enodimenzionalen ali multidimenzionalen.
- V obravnavi nelinearne numerične optimizacije se bomo omejili zgolj na probleme optimizacije brez omejitev. Probleme optimizacije z omejitvami pa si bralec lahko natančneje pogleda v literaturi [8, 15, 17, 21, 28].

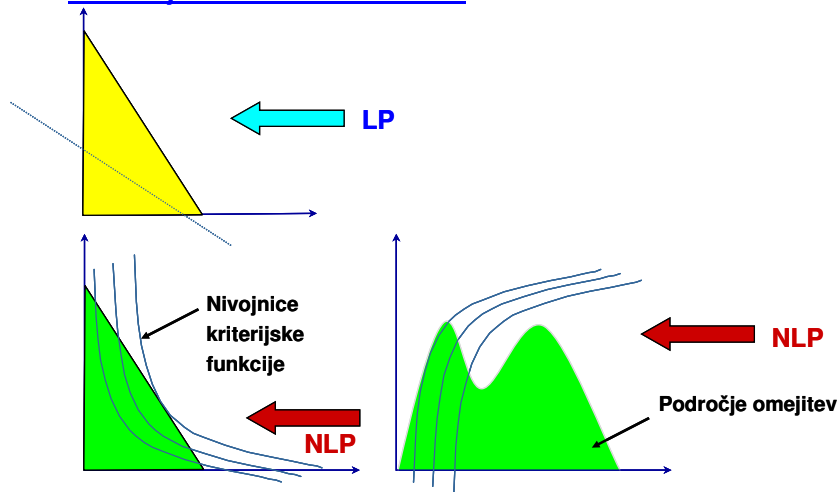
Slika 29 prikazuje nekatere značilne primere ciljnih funkcij s pripadajočimi omejitvami pri nelinearnem programiranju. Kot je razvidno, se v teh primerih nelinearnosti (potence in koreni, itn.) pojavljajo bodisi v ciljni funkciji, bodisi v omejitvah, ali celo tako pri funkciji kot tudi omejitvah.

- **Nekaj primerov NLP:**
- **1. primer:**  
 $f(x_1, x_2) = 3x_1 + 2x_2^4$  → Ciljna funkcija  
 $x_1 + x_2^2 \leq 1, x_1 \geq 0, x_2 \text{ neomejen}$  → Omejitvi
- **2. primer:**  
 $f(x) = 4x_1^2 + 5x_2^2$  → Ciljna funkcija  
 $h(x) = 2x_1 + 3x_2 - 6 = 0$  → Omejitev
- **3. primer:**  
 $f(x) = 7x_1 - 6x_2 + 4x_3$  → Ciljna funkcija  
 $h^1(x) = x_1^2 + 2x_2^2 + 3x_3^2 - 1 = 0$   
 $h^2(x) = 5x_1 + 5x_2 - 3x_3 - 6 = 0$  } → Omejitvi
- **4. primer:**  
 $f(x, y) = \sqrt{(x-14)^2 + (y-15)^2}$  → Ciljna funkcija  
 $(x-8)^2 + (y-9)^2 \leq 49$   
 $x \geq 2, x \leq 13, x + y \leq 24$  } → Omejitve

Slika 29: Nekateri značilni primeri ciljnih funkcij s pripadajočimi omejitvami pri NLP

Slika 30 pa prikazuje ilustracijo razlike med linearnim in nelinearnim programiranjem v smislu nivojnic ciljne funkcije in področja omejitev (dopustnih rešitev).

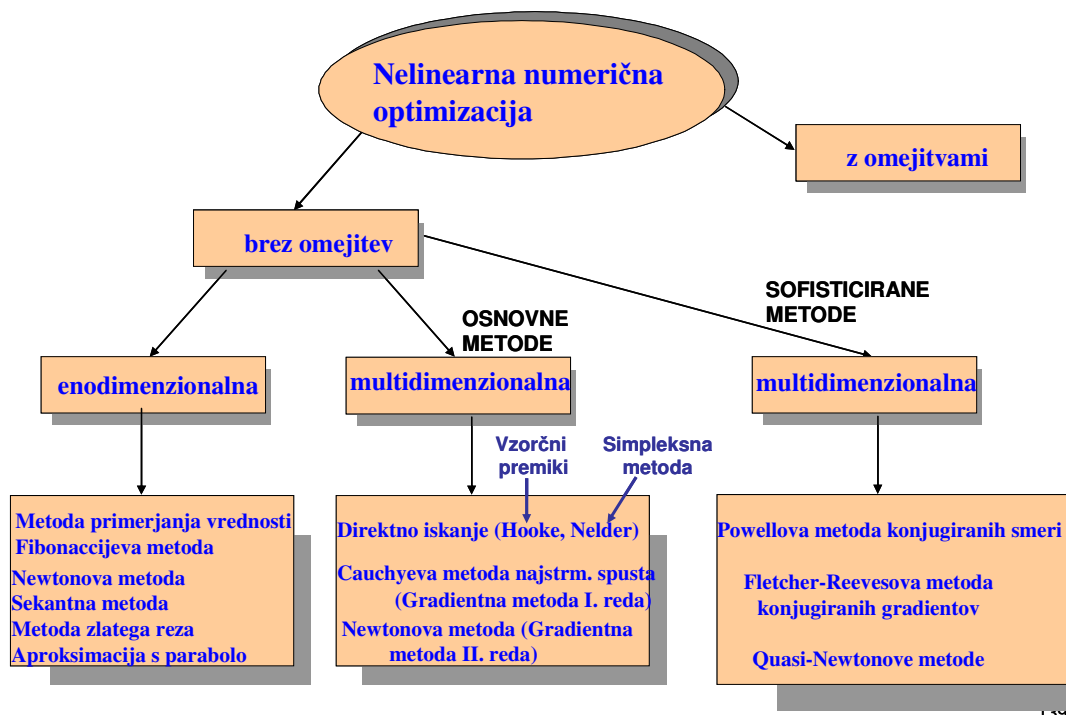
- **Ilustracija razlike med LP in NLP:**



Slika 30: Ilustracija razlike med linearnim in nelinearnim programiranjem v smislu nivojnic ciljne funkcije in področja omejitev (dopustnih rešitev).

Seveda gre v primeru LP za ravne črte (daljice oz. premice), v primeru NLP pa nivojnice ciljne funkcije in meje področja omejitev zavzamejo poljubno obliko krivulje.

Pri problematiki nelinearne numerične optimizacije oz. nelinearnega programiranja imamo opravka z metodologijo oz. skupino numeričnih metod, ki nam pomagajo na numerični način poiskati optimume ciljne funkcije. Slika 31 prikazuje eno izmed možnih delitev teh metod [4, 14, 17, 21].



Slika 31: Ena izmed možnih delitev numeričnih metod pri nelinearni numerični optimizaciji

Kot je razvidno iz slike 31, v splošnem numerične metode lahko razdelimo na tiste, ki obravnavajo probleme brez omejitev, ter tiste pri problemih z omejitvami. Naslednja pomembna delitev je lahko delitev na metode, ki iščejo optimume funkcije ene spremenljivke, ter tiste, ki iščejo optimume funkcije več spremenljivk. Pri slednjih pa lahko ločimo med tistimi metodami, ki so bolj osnovnega značaja, ter tiste, ki optimume iščejo na bolj sofisticirane načine. Seveda se bomo pri nadaljnji obravnavi omejili le na nekatere najbolj značilne metode, prikazane na sliki 31.

Preden pa se lotimo natančnejše obravnave posameznih metod za iskanje optimumov, pa si v naslednjem poglavju še pogledjmo na nekaj primerih, kako sploh poteka modeliranje nelinearnih problemov. Kot je znano, moramo v postopku optimizacije z modeliranjem najprej sploh določiti strukturo ciljne funkcije, da bi potem lahko uporabili katero izmed primernih numeričnih metod za iskanje optimumov te funkcije (glej sliko 1!).

## 6.1 Modeliranje nelinearnih problemov

V tem poglavju bomo na nekaj primerih pogledali, kako poteka modeliranje nelinearnih problemov. Omejili se bomo na naslednje vrste problemov:

- Nelinearni transportni problem,
- Nelinearni problem optimalne razporeditve virov, ter
- Nelinearni problem optimalnega načrtovanja transportnih zabožnikov.

### Primer 1:

*Pogoji transporta neke vrste blaga iz treh skladišč do treh trgovskih centrov so takšni, da vsi stroški niso linearno odvisni od količine blaga, pač pa odvisnosti prikazuje tabela 4. Nastavite matematični model za dani problem!*

	TC.1	TC.2	TC.3	Kapacit. Skladišč
Skl. 1	$\sqrt{x_{11}} \cdot c_{11}$	$\sqrt{x_{12}} \cdot c_{12}$	$x_{13} \cdot c_{13}$	$a_1$
Skl.2	$\sqrt{x_{21}} \cdot c_{21}$	$\sqrt{x_{22}} \cdot c_{22}$	$\sqrt{x_{23}} \cdot c_{23}$	$a_2$
Skl.3	$x_{31} \cdot c_{31}$	$\sqrt{x_{32}} \cdot c_{32}$	$x_{33} \cdot c_{33}$	$a_3$
Potrebe trgovskih centrov	$b_1$	$b_2$	$b_3$	

Tabela 4: Nelinearni transportni problem

### Rešitev:

Matematični model lahko zapišemo v naslednji obliki.

Ciljna funkcija:

$$f(x) = c_{11}\sqrt{x_{11}} + c_{12}\sqrt{x_{12}} + c_{13}x_{13} + c_{21}\sqrt{x_{21}} + c_{22}\sqrt{x_{22}} + c_{23}\sqrt{x_{23}} + c_{31}x_{31} + c_{32}\sqrt{x_{32}} + c_{33}x_{33} \quad (6.1.)$$

Omejitve:

$$x_{11} + x_{21} + x_{31} = b_1$$

$$x_{12} + x_{22} + x_{32} = b_2$$

$$x_{13} + x_{23} + x_{33} = b_3$$

$$x_{11} + x_{12} + x_{13} = a_1$$

$$x_{21} + x_{22} + x_{23} = a_2$$

$$x_{31} + x_{32} + x_{33} = a_3$$

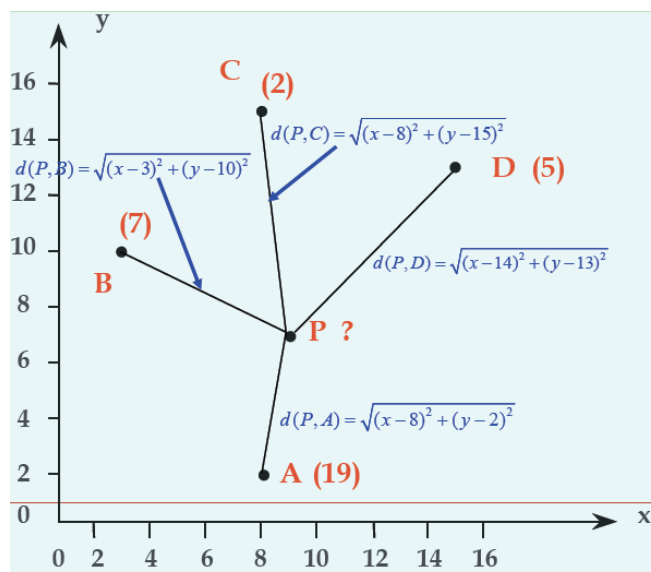
(6.2.)

$$x_{ij} \geq 0, \quad i = 1, 2, 3, \quad j = 1, 2, 3$$

Kot je razvidno iz izrazov (6.1.) in (6.2.), imamo opravka z nelinearno ciljno funkcijo 9 spremenljivk in linearnimi omejitvami. Za ta model je nato potrebno izbrati eno izmed primernih numeričnih metod, ki rešujejo probleme z omejitvami, ter poiskati optimalno rešitev problema. Seveda je v tem primeru optimalna rešitev tista, pri kateri ciljna funkcija (6.1.) doseže minimalno vrednost.

### **Primer 2:**

*Potrebno je določiti lokacijo skladišča, ki bo z minimalnimi stroški oskrbovalo štiri trgovine A, B, C in D. Trgovina A se nahaja na lokaciji  $(x_A, y_A) = (8, 2)$ , trgovina B na lokaciji  $(x_B, y_B) = (3, 10)$ , trgovina C na lokaciji  $(x_C, y_C) = (8, 15)$ , ter trgovina D na lokaciji  $(x_D, y_D) = (14, 13)$ . Dnevne potrebe trgovin so: A (19 enot blaga), B (7 enot blaga), C (2 enoti blaga) in D (5 enot blaga). Problem je predstavljen na sliki 32. Nastavite matematični model za dani problem!*



Slika 32: Problem načrtovanja optimalnih transportnih poti med neznano lokacijo skladišča P in znanimi lokacijami trgovin A, B, C in D (primer brez omejitev)

**Rešitev:**

Če točke A, B, C in D povežemo z iskano neznano (optimalno) lokacijo skladišča  $P(x,y)$ , dobimo na osnovi Pitagorovega izreka naslednje evklidske razdalje med lokacijami trgovin in skladišča:

$$\begin{aligned}
 d(P, A) &= \sqrt{(x-8)^2 + (y-2)^2} \\
 d(P, B) &= \sqrt{(x-3)^2 + (y-10)^2} \\
 d(P, C) &= \sqrt{(x-8)^2 + (y-15)^2} \\
 d(P, D) &= \sqrt{(x-14)^2 + (y-13)^2}
 \end{aligned}
 \tag{6.3.}$$

V skladu z dnevnimi potrebami trgovin tvorimo ciljno funkcijo na takšen način, da pomnožimo posamezne razdalje od skladišča do trgovin s pripadajočimi dnevnimi potrebami trgovin:

$$\begin{aligned}
 f(x, y) &= 19\sqrt{(x-8)^2 + (y-2)^2} + 7\sqrt{(x-3)^2 + (y-10)^2} + \\
 &+ 2\sqrt{(x-8)^2 + (y-15)^2} + 5\sqrt{(x-14)^2 + (y-13)^2}
 \end{aligned}
 \tag{6.4.}$$

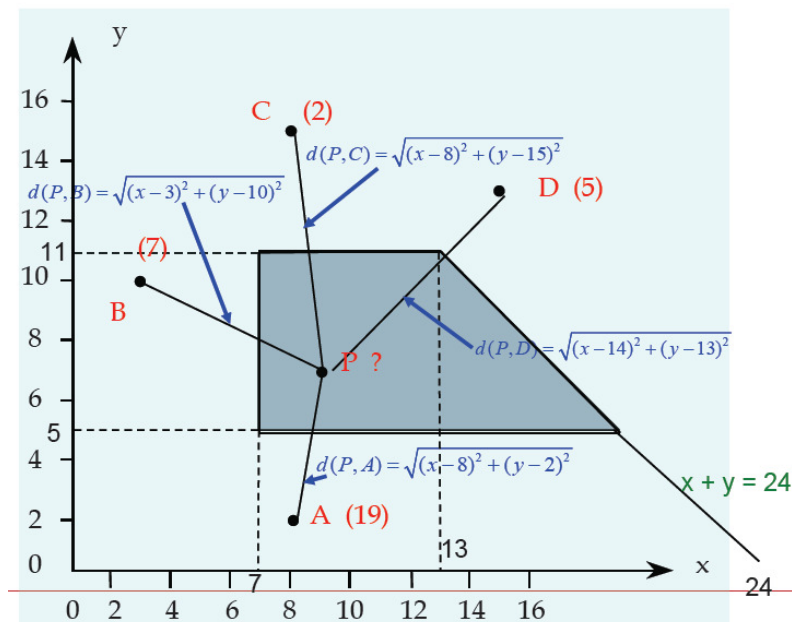
Dobljeni izraz predstavlja matematični model nelinearne ciljne funkcije dveh spremenljivk:  $x$  in  $y$ . Z optimizacijo je nato potrebno poiskati takšni optimalni koordinati skladišča  $(x_p, y_p) = (x^*, y^*)$ , pri katerih bo ciljna funkcija minimalna:



$f(x^*, y^*) = \min f(x, y)$  Tedaj bo skladišče gotovo imelo optimalno lokacijo, saj bo skupna vsota razdalj, pomnožena s potrebami, najmanjša, kar pa pomeni tudi minimalne stroške (saj so ti ob fiksnih dnevnih potrebah trgovin proporcionalni z razdaljami).

Kot je razvidno, se v matematičnem modelu za ciljno funkcijo (6.4.) ne pojavi nobena omejitev, saj lokacija skladišča  $P(x, y)$  lahko leži kjerkoli v x-y ravnini. V tem primeru gre torej očitno za problem optimizacije brez omejitev.

Obravnavajmo še popolnoma enak primer prejšnjemu, le da so dane še prostorske omejitve, prikazane na sliki 33.



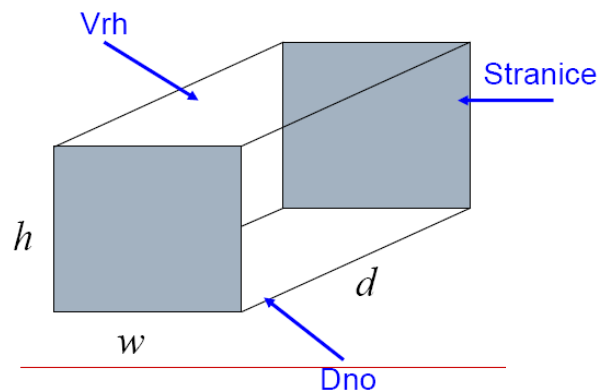
Slika 33: Problem načrtovanja optimalnih transportnih poti med neznano lokacijo skladišča P in znanimi lokacijami trgovin A, B, C in D (primer z omejitvami)

Izpeljava matematičnega modela poteka na popolnoma enak način kot v prejšnjem primeru, pri čemer dobimo enako ciljno funkcijo (6.4.). Seveda pa v tem primeru nastopijo še naslednje omejitve (glej sliko 33!):

$$\begin{aligned} x &\geq 7 \\ 5 &\leq y \leq 11 \\ x + y &\leq 24 \end{aligned} \tag{6.5.}$$

### Primer 3:

Potrebno je načrtati transportne zabojnike pri minimalnih stroških. Kvadratni meter materiala, iz katerega sta zgrajena vrh in dno zabojnika, stane 325 EUR. Kvadratni meter materiala za stranice pa stane 175 EUR. Prostornina zabojnika ne sme presegati 100 m<sup>3</sup>, višina ne sme presegati 3m, površina stranic pa ne sme presegati 75 m<sup>2</sup>. Problem je predstavljen na sliki 34. Nastavite matematični model nelinearne ciljne funkcije.



Slika 34: Problem načrtovanja transportnega zabojnika pri minimalnih stroških

### Rešitev:

Površino zabojnika lahko izrazimo na naslednji način:

$$S = \underbrace{2 \cdot d \cdot w}_{S_{VRH,DNO}} + \underbrace{2 \cdot h \cdot w + 2 \cdot d \cdot h}_{S_{STRANICE}} \quad (6.6.)$$

Ciljna funkcija ima naslednjo obliko:

$$\begin{aligned} f_{STROŠKI} &= S_{VRH,DNO} \cdot 325 + S_{STRANICE} \cdot 175 = \\ &= \underbrace{2 \cdot d \cdot w}_{S_{VRH,DNO}} \cdot 325 + \underbrace{(2 \cdot h \cdot w + 2 \cdot d \cdot h)}_{S_{STRANICE}} \cdot 175 = \\ &= 650 \cdot d \cdot w + 350(h \cdot w + d \cdot h) = 650 \cdot d \cdot w + 350(w + d) \cdot h \end{aligned} \quad (6.7.)$$

Če upoštevamo še omejitve, dobimo naslednji matematični model:

$$\begin{aligned} \min_{d,w,h} \quad & f(w,d,h) = 650dw + 350(d+w)h \\ & 2(dh + wh) \leq 75 \\ & h \leq 3 \\ & wd h \leq 100 \\ & w, d, h \geq 0 \end{aligned} \quad (6.8.)$$

Kot je razvidno iz izrazov (6.8.), imamo opravka z nelinearno ciljno funkcijo 3 spremenljivk  $w, d, h$  in z nelinearnimi omejitvami. Za ta model je nato potrebno izbrati eno izmed primernih numeričnih metod, ki rešujejo probleme z omejitvami, ter poiskati optimalno rešitev problema. Seveda je v tem primeru optimalna rešitev tista, pri kateri ciljna funkcija  $f(d,w,h)$  v (6.8.) doseže minimalno vrednost.

Na podoben način, kot smo ponazorili modeliranje nelinearnih problemov v zgornjih treh primerih tega poglavja, je možno pri reševanju nelinearnih logističnih problemov formulirati še številne druge probleme v obliki matematičnih modelov.

Ko pa so problemi v obliki modelov ustrezno izpeljani, pa je seveda potrebno izbrati primerno numerično optimizacijsko metodo, ter s postopkom optimizacije poiskati optimalno rešitev problema.

V nadaljevanju se bomo posvetili obravnavi optimizacijskih metod, pri čemer se bomo omejili le na metode, ki rešujejo probleme brez omejitev. Podobni principi delovanja veljajo tudi za metode, primerne za probleme z omejitvami. Seveda pa so te še nekoliko zahtevnejše in jih bralec lahko zasledi v literaturi [8, 15, 17, 21, 28].

## 6.2 Enodimenzionalna optimizacija

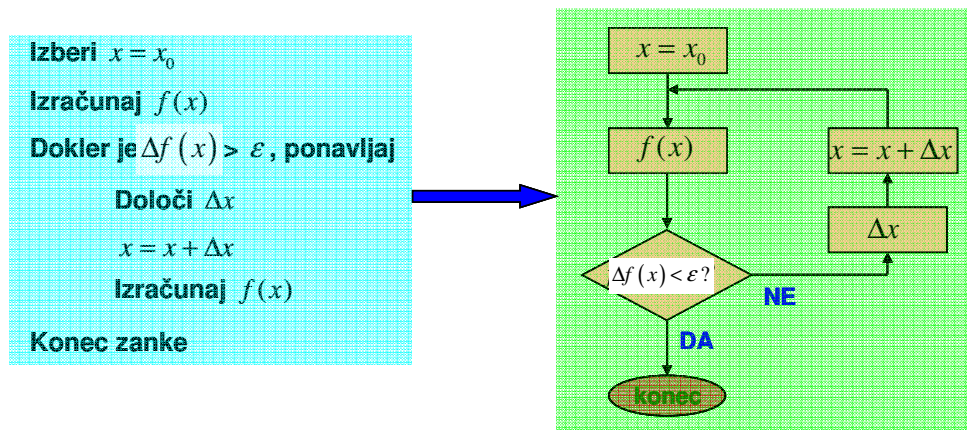
V tem poglavju si bomo pogledali 3 numerične metode, ki so značilne pri reševanju problemov, kjer imamo opravka s ciljno funkcijo le ene same spremenljivke. Kot bo nato razvidno v nadaljevanju, pa so te metode ponekod primerne tudi kot pomožno sredstvo

oz. mehanizem v okviru reševanja problemov s pomočjo nekaterih multidimenzionalnih numeričnih metod.

Preden se lotimo opisa enodimenzionalnih optimizacijskih metod, pa spregovorimo še nekaj besed o lastnostih optimizacijskih algoritmov, ki v principu veljajo prav za vse postopke optimiranja.

### 6.2.1 Lastnosti optimizacijskih algoritmov

Vsak optimizacijski postopek lahko strnemo v naslednji simbolični algoritem, kot je prikazan na sliki 35 [13].



Slika 35: Simbolični algoritem postopka optimiranja

Kot je razvidno iz slike 35, moramo vektor odločitvenih spremenljivk  $x$  spreminjati toliko časa, dokler se vrednost ciljne funkcije  $f$  ne spreminja več, kot je pa predpisano s toleranco  $\varepsilon$ . Pri tem mora vsaka sprememba vektorja odločitvenih spremenljivk  $\Delta x$  voditi k zmanjšanju ciljne funkcije  $f$  (če iščemo minimum, seveda), kar naj bi se zgodilo v sleherni iteraciji optimizacijskega postopka.

Torej, kriterija pri določanju vektorja spremembe  $\Delta x$  sta predvsem naslednja dva [13]:

- Zagotavljati mora manjšo vrednost ciljne funkcije v sledeči iteraciji,
- Pripeljati mora do končne rešitve v čim manj iteracijah.

Če sta zgornja kriterija izpolnjena, bosta s tem zagotovljeni konvergenca postopka in njegova kar največja učinkovitost.

Postopki optimiranja se razlikujejo v strategiji določevanja vektorja  $\Delta x$  in jih delimo na [13]:

- Neposredne (direktne) postopke (kjer se izračunava le vrednost ciljne funkcije)
- Gradientne postopke (1. ali 2. reda) (kjer se izračunavajo tudi 1. oz. 2. odvodi).

Naštejmo nekaj prednosti in slabosti posameznih postopkov [13]:

- Prednosti direktnih postopkov: Uporabni so tedaj, ko gradienta ni moč izračunati, ali le-ta ne obstaja (npr. pri nezveznih ciljnih funkcijah).
- Prednosti gradientnih postopkov: Hitrejša konvergenca v bližini minimuma.
- Slabosti gradientnih postopkov: Računanje gradienta je lahko težavno.

Gornje navedbe kažejo na večjo počasnost optimizacijskih metod z neposrednim iskanjem in na hitrejšo, a manj zanesljivo delovanje gradientnih metod. Pri tem večina postopkov pripelje le do lokalnega minimuma ciljne funkcije, ki pa ni nujno tudi globalen.

V nadaljevanju si bomo pogledali nekaj enodimenzionalnih metod optimizacije. Pri njih seveda  $x$  na sliki 35 ne predstavlja vektorja, pač pa eno samo neodvisno spremenljivko v obliki skalarja.

### **6.2.2 Metoda zlatega reza**

Predpostavimo, da je  $f(x)$  realna funkcija spremenljivke  $x$ . Naj bo  $x^*$  (neznani globalni) minimum, ki leži na intervalu  $[a_{k-1}, b_{k-1}]$ , kjer je  $k$  indeks iteracije. Vrednosti funkcije bomo izračunali v vmesnih točkah  $y_k$  in  $z_k$  (slika 36). Za ti dve točki zahtevamo:

1.) enako oddaljenost od robov intervala  $[a_{k-1}, b_{k-1}]$ :

$$y_k - a_{k-1} = b_{k-1} - z_k \quad (6.9.)$$

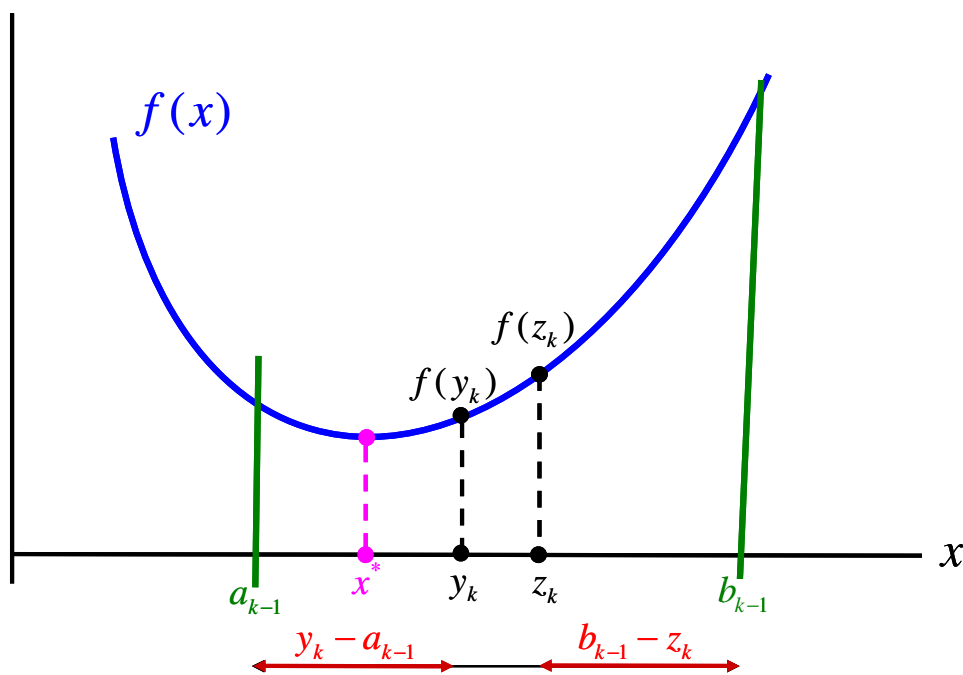
2.) da je  $y_k$  na levi strani,  $z_k$  pa na desni strani znotraj intervala  $[a_{k-1}, b_{k-1}]$ :

$$y_k = a_{k-1} + c_{k-1}(b_{k-1} - a_{k-1}) \quad (6.10.)$$

$$z_k = b_{k-1} - c_{k-1}(b_{k-1} - a_{k-1})$$

pri čemer je  $c_{k-1}$  neko število, za katerega mora veljati:

$$0 < c_{k-1} < \frac{1}{2} \quad (6.11.)$$



Slika 36: Tvorjenje vmesnih točk  $y_k$  in  $z_k$  pri metodi zlatega reza

Točko  $z_k$  lahko izrazimo tudi nekoliko drugače:

$$\begin{aligned} z_k &= b_{k-1} - a_{k-1} + a_{k-1} - c_{k-1}(b_{k-1} - a_{k-1}) = \\ &= a_{k-1} + (b_{k-1} - a_{k-1}) - c_{k-1}(b_{k-1} - a_{k-1}) = \\ &= a_{k-1} + (1 - c_{k-1})(b_{k-1} - a_{k-1}) \end{aligned} \quad (6.12.)$$

Če primerjamo vrednosti funkcije  $f(y_k)$  in  $f(z_k)$  na sliki 36, imamo opravka z dvema možnostima. Poglejmo si najprej 1. možnost.

**1. možnost:**

Iz primerjave vrednosti funkcije  $f(y_k)$  in  $f(z_k)$  sklepamo o optimumu naslednje:

$$f(y_k) \leq f(z_k) \quad \Rightarrow \quad a_{k-1} \leq x^* \leq z_k \quad (6.13.)$$

zato tedaj izberemo:

$$\begin{aligned} a_k &= a_{k-1}, \\ b_k &= z_k \end{aligned} \quad (6.14.)$$

torej bomo desno krajišče intervala premaknili na levo v točko  $z_k$ , pod domnevo, da se optimum nahaja znotraj intervala, ki ga opisuje izraz (6.13.).

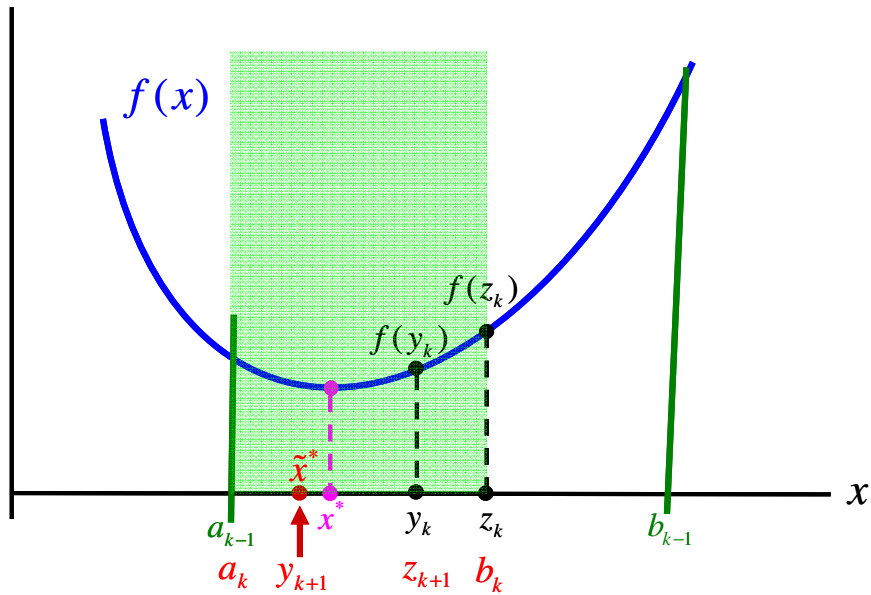
Postavimo tudi (glej (6.10.)):

$$\begin{aligned} y_{k+1} &= a_k + c_k(b_k - a_k) \\ z_{k+1} &= y_k \end{aligned} \quad (6.15.)$$

In za približek rešitve (optimuma) v trenutni iteraciji vzamemo:

$$\tilde{x}^* = x_{k+1} = y_{k+1} \quad (6.16.)$$

Pri 1. možnosti torej dobimo situacijo, kot jo npr. ilustrira slika 37.



Slika 37: 1. možnost za postavitev novega intervala pri metodi zlatega reza

**2. možnost:**

Poglejmo si še 2. možnost, ki jo ilustrira slika 38. Tedaj glede na primerjavo vrednosti funkcije  $f(y_k)$  in  $f(z_k)$  o optimumu sklepamo naslednje:

$$f(y_k) > f(z_k) \quad \Rightarrow \quad y_k \leq x^* \leq b_{k-1} \quad (6.17.)$$

zato tedaj izberemo:

$$\begin{aligned} b_k &= b_{k-1}, \\ a_k &= y_k \end{aligned} \quad (6.18.)$$

torej bomo levo krajišče intervala premaknili na desno v točko  $y_k$ , pod domnevo, da se optimum nahaja znotraj intervala, ki ga opisuje izraz (6.17.).

Postavimo tudi (glej (6.12.)):

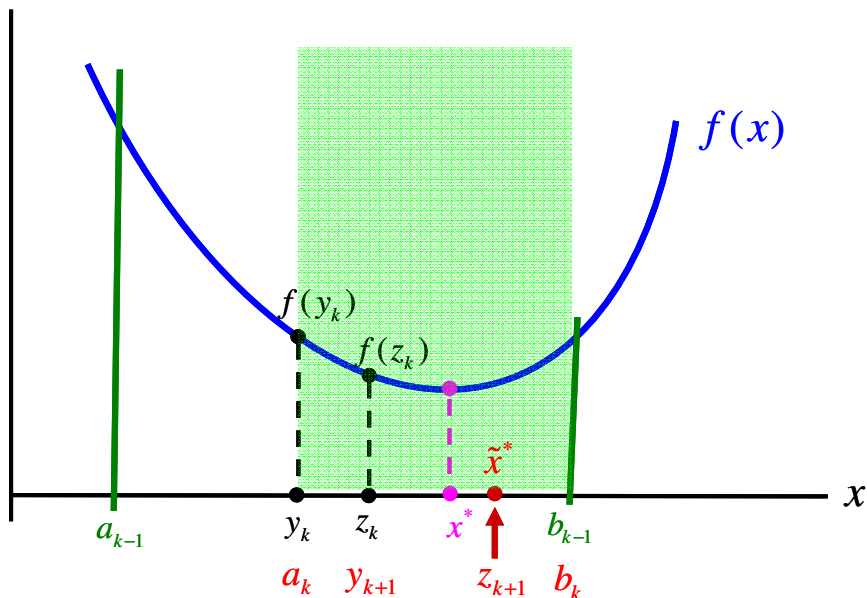
$$\begin{aligned} z_{k+1} &= b_k - c_k(b_k - a_k) = a_k + (1 - c_k)(b_k - a_k) \\ y_{k+1} &= z_k \end{aligned} \quad (6.19.)$$



In za približek rešitve (optimuma) v trenutni iteraciji vzamemo:

$$\tilde{x}^* = x_{k+1} = z_{k+1} \quad (6.20.)$$

Pri 2. možnosti torej dobimo situacijo, kot jo npr. ilustrira slika 38.



Slika 38: 2. možnost za postavitev novega intervala pri metodi zlatega reza

Kot je razvidno iz slik 37 in 38, smo v  $k$ .ti iteraciji postopka interval prejšnje iteracije zožali, ter tako dobili zožan interval (osenčeno območje na slikah 37 oz. 38), na katerem se gotovo nahaja pravi optimum. Če postopek nato v naslednjih iteracijah na podoben način ponavljamo, so novi intervali vedno ožji in ožji, ter vedno bolj »oklepajo« pravi optimum. Ko konvergenca postopka doseže neko predpisano toleranco  $\varepsilon$ , se postopek konča, iz izračunanih vrednosti zadnje iteracije pa sklepamo o optimumu funkcije.

Preostane nam še naloga, kako določiti takšen ustrezen faktor  $c_k$ , pri katerem bodo v vsaki iteraciji izpolnjeni pogoji (6.9.), (6.10.) in (6.11.). Pokazati se da, da bodo ti pogoji izpolnjeni, če velja naslednja relacija [17]:

$$c_k = \frac{1 - 2c_{k-1}}{(1 - c_{k-1})}; \quad k = 1, 2, \dots \quad (6.21.)$$

Če želimo ohraniti v postopku enaka razmerja pri zožanju intervala, vzamemo [17]:

$$c_k = c_{k-1}; \quad k = 1, 2, \dots \quad (6.22.)$$

$$c_k = \frac{1-2c_k}{(1-c_k)} \Rightarrow c_k(1-c_k) = 1-2c_k; \quad k = 1, 2, \dots$$

Iz (6.22.) dobimo naslednjo kvadratno enačbo:

$$c_k - c_k^2 = 1 - 2c_k \rightarrow c_k^2 - 3c_k + 1 = 0 \quad (6.23.)$$

katere rešitev pri upoštevanju izraza (6.11.) je:

$$c_k = c_{zr} = \frac{3-\sqrt{5}}{2} = 0.38197 \quad (6.24.)$$

Ker velja:

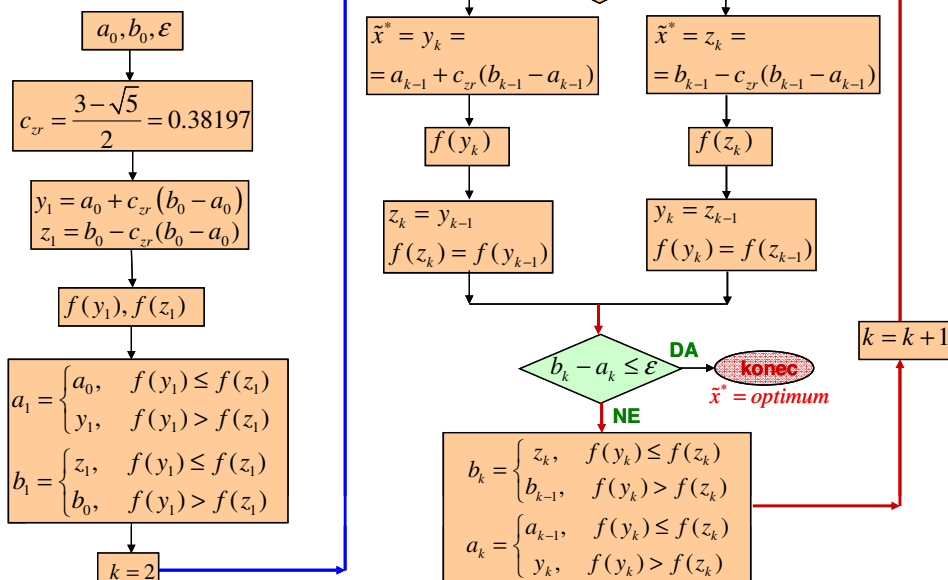
$$1 + (1 - c_{zr}) = 1 + (1 - 0.38197) = 1.6103 \quad (6.25.)$$

kar je značilna vrednost pri zlatem rezu, je metoda na osnovi tega dejstva dobila svoje ime.

Slika 39 prikazuje celoten postopek (algoritem), ki velja za metodo zlatega reza, kjer je  $[a_0, b_0]$  začetni opazovani interval v 0. iteraciji, ki ga bomo v naslednjih iteracijah ožali in na takšen način poiskali optimum funkcije  $f$  [4, 5, 13, 17].

Poudarimo še, da metoda zlatega reza, ki je direktna metoda, spada v kategorijo takoimenovanih metod primerjanja vrednosti funkcije. V to kategorijo npr. spada tudi znamenita Fibonaccijeva metoda iskanja ekstrema in dokazati se da, da obe metodi pri določenih pogojih preideta v enako metodo oz. dajeta enake rezultate.

## Metoda Zlatega reza



Slika 39: Postopek (algoritem) pri metodi zlatega reza

### Primer:

S pomočjo metode zlatega reza poiščite ekstrem naslednje funkcije:

$$f(x) = 2x^4 - 3x \quad (6.26.)$$

kjer so podani naslednji predpisani parametri optimizacije:  $a_0 = 0$ ,  $b_0 = 1$ ,  $\varepsilon = 10^{-5}$ .

### Rešitev:

Na analitičen način bi ekstrem funkcije (6.26.) poiskali na naslednji način:

$$\frac{df(x)}{dx} = 8x^3 - 3 = 0 \quad (6.27.)$$

$$x^* = \sqrt[3]{\frac{3}{8}} = 0.7211$$

Pri numeričnem načinu s pomočjo metode zlatega reza seveda ne iščemo ekstrema na takšen način, pač pa uporabimo računalniški program, ki sledi postopku na sliki 39.

Pri tem so izračunani rezultati tega postopka v posameznih iteracijah podani v tabeli, prikazani v obliki slike 40.

**Tabela rezultatov konvergence :**

$k$	$a_{k-1}$	$y_k$	$z_k$	$b_{k-1}$	$f(y_k)$	$f(z_k)$
1	0	0,38197	0,61803	1	-1,1033	-1,5623
2	0,38197	0,61803	0,76393	1	-1,5623	-1,6106
3	0,61803	0,76393	0,8541	1	-1,6106	-1,498
4	0,61803	0,7082	0,76393	0,8541	-1,6215	-1,6106
5	0,61803	0,67376	0,7082	0,76393	-1,6091	-1,6215
6	0,67376	0,7082	0,72949	0,76393	-1,6215	-1,6221
7	0,7082	0,72949	0,74265	0,76393	-1,6221	-1,6196
8	0,7082	0,72136	0,72949	0,74265	-1,6225	-1,6221
9	0,7082	0,71633	0,72136	0,72949	-1,6224	-1,6225
10	0,71633	0,72136	0,72447	0,72949	-1,6225	-1,6225
11	0,71633	0,71944	0,72136	0,72447	-1,6225	-1,6225
12	0,71944	0,72136	0,72255	0,72447	-1,6225	-1,6225
13	0,71944	0,72063	0,72136	0,72255	-1,6225	-1,6225
14	0,72063	0,72136	0,72181	0,72255	-1,6225	-1,6225
15	0,72063	0,72108	0,72136	0,72181	-1,6225	-1,6225
16	0,72063	0,72091	0,72108	0,72136	-1,6225	-1,6225
17	0,72091	0,72108	0,72119	0,72136	-1,6225	-1,6225
18	0,72091	0,72101	0,72108	0,72119	-1,6225	-1,6225
19	0,72101	0,72108	0,72112	0,72119	-1,6225	-1,6225
20	0,72108	0,72112	0,72115	0,72119	-1,6225	-1,6225
21	0,72108	0,7211	0,72112	0,72115	-1,6225	-1,6225
22	0,7211	0,72112	0,72113	0,72115	-1,6225	-1,6225
23	0,72112	0,72113	0,72114	0,72115	-1,6225	-1,6225
24	0,72112	0,72113	0,72113	0,72114	-1,6225	-1,6225
25	0,72112	0,72112	0,72113	0,72113	-1,6225	-1,6225
26	0,72112	0,72113	0,72112	0,72113	-1,6225	-1,6225

**Izračunan minimum :**  
 $x^* = x_{\text{min}} =$   
 $= x_{26} = 0.7211$

Slika 40: Rezultati pri iskanju optimuma funkcije (6.26.) s pomočjo metode zlatega reza

Kot je razvidno iz slike 40, je bila konvergenca postopka dosežena po 26 iteracijah, dobljeni numerični rezultat pa se očitno ujema z rezultatom (6.27.), ki bi ga dobili na analitičen način (numerični izračun je bil izveden v programskem jeziku Matlabu).

### 6.2.3 Newtonova metoda za eno spremenljivko

Newtonova metoda temelji na dejstvu, da v stacionarni točki dane funkcije  $f$  velja [17]:

$$\frac{df(x^*)}{dx} = f'(x^*) = 0 \quad (6.28.)$$

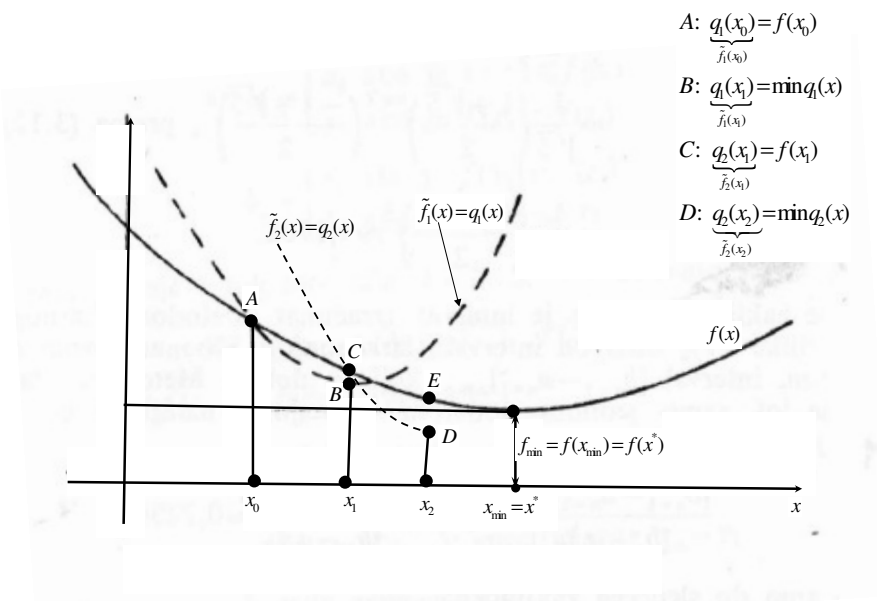
Če je funkcija konveksna, je ta stacionarna točka seveda tudi globalni minimum.

Ideja Newtonove metode je dokaj enostavna in temelji na sledečih predpostavkah. Če z  $x_0$  označimo neko začetno aproksimacijo stacionarne točke  $x^*$ , potem se lahko vrednosti funkcije  $f$  v okolici točke  $x_0$  aproksimirajo s prvimi 3 členi Taylorjeve vrste (parabolo) na naslednji način [17]:

$$\begin{aligned}
\tilde{f}_1(x) &\approx q_1(x) = \\
&= \underbrace{f(x_0)}_{K_1} + \underbrace{f'(x_0)}_{K_2}(x-x_0) + \underbrace{\frac{1}{2}f''(x_0)}_{K_3}(x-x_0)^2 = \\
&= K_1 + K_2(x-x_0) + K_3(x-x_0)^2
\end{aligned}
\tag{6.29}$$

Nato poiščemo minimum parabole (6.29.) in njena ekstremna točka  $x_1$  se vzame za novo aproksimacijo dejanske stacionarne točke  $x^*$  (glej sliko 41), kar se zgodi v 1. iteraciji postopka:

$$\begin{aligned}
\frac{d}{dx} \tilde{f}_1(x) &= \frac{d}{dx} q_1(x) = K_2 + 2K_3(x-x_0) = 0 \rightarrow x_{eks} = x_1 = x_0 - \frac{K_2}{2K_3} \rightarrow \\
x_1 &= x_0 - \frac{f'(x_0)}{2\frac{1}{2}f''(x_0)} = x_0 - \frac{f'(x_0)}{f''(x_0)}
\end{aligned}
\tag{6.30}$$



Slika 41: Ideja Newtonove metode

Če se sedaj vrednosti funkcije  $f$  v okolici točke  $x_1$  aproksimirajo s prvimi 3 členi Taylorjeve vrste (parabolo), dobimo:

$$\begin{aligned}
\tilde{f}_2(x) &\approx q_2(x) = \\
&= \underbrace{f(x_1)}_{K_1} + \underbrace{f'(x_1)}_{K_2}(x-x_1) + \underbrace{\frac{1}{2}f''(x_1)}_{K_3}(x-x_1)^2 = \\
&= K_1 + K_2(x-x_1) + K_3(x-x_1)^2
\end{aligned} \tag{6.31.}$$

Nato poiščemo minimum parabole (6.31.) in njena ekstremna točka  $x_2$  se vzame za novo aproksimacijo dejanske stacionarne točke  $x^*$  (glej sliko 41), kar se zgodi v 2. iteraciji postopka:

$$\begin{aligned}
\frac{d}{dx} \tilde{f}_2(x) &= \frac{d}{dx} q_2(x) = K_2 + 2K_3(x-x_1) = 0 \rightarrow x_{eks} = x_2 = x_1 - \frac{K_2}{2K_3} \rightarrow \\
x_2 &= x_1 - \frac{f'(x_1)}{\frac{1}{2}f''(x_1)} = x_1 - \frac{f'(x_1)}{f''(x_1)}
\end{aligned} \tag{6.32.}$$

Če postopek ponavljamo za naslednje iteracije, potem lahko za  $k$ -to iteracijo zapišemo sledečo Newtonovo iteracijsko enačbo (za funkcijo ene spremenljivke) [13, 17, 21]:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, \quad k = 0, 1, 2, \dots \tag{6.33.}$$

Ta enačba nas načeloma gotovo vodi do stacionarne točke funkcije  $f$  in nam tako pomaga poiskati pravi optimum funkcije.

Če je funkcija  $f$  dovolj gladka (ima zvezno definirane odvode) in če je začetna aproksimacija  $x_0$  relativno dovolj blizu točke  $x^*$ , potem niz aproksimacij  $x_1, x_2, x_3, \dots$  dokaj hitro (s kvadratno hitrostjo konvergence) konvergira k točki  $x^*$ .

### **Primer:**

S pomočjo Newtonove metode poiščite ekstrem naslednje funkcije,

$$f(x) = 2x^4 - 3x \tag{6.34.}$$

pri čemer vzemite začetno aproksimacijo  $x_0 = 0.5$ .

**Rešitev:**

V skladu z enačbo (6.33.) najprej izračunamo vrednosti 1. in 2. odvoda funkcije, saj ju bomo potrebovali v postopku:

$$f'(x) = 8x^3 - 3 \quad (6.35.)$$

$$f''(x) = 24x^2$$

Odvoda (6.35.) nato vstavimo v enačbo (6.33.) in dobimo:

$$x_{k+1} = x_k - \frac{8x_k^3 - 3}{24x_k^2}, \quad k = 0, 1, 2, \dots \quad (6.36.)$$

Če bi želeli enačbo (6.36.) ročno izračunavati za posamezne iteracije, bi dobili pri  $x_0 = 0.5$ :

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} = x_0 - \frac{8x_0^3 - 3}{24x_0^2} = 0.5 - \frac{8(0.5)^3 - 3}{24(0.5)^2} = 0.5 - \frac{-2}{6} = 0.8333 \quad (6.37.)$$

$$x_2 = x_1 - \frac{f'(x_1)}{f''(x_1)} = x_1 - \frac{8x_1^3 - 3}{24x_1^2} = 0.8333 - \frac{8(0.8333)^3 - 3}{24(0.8333)^2} = 0.7355$$

Itn...

Natanko tako, kot je prikazano v izrazih (6.37.), išče ekstrem funkcije tudi računalniški program. Pri tem bi bile poiskane vrednosti aproksimacij v posameznih iteracijah takšne, kot jih prikazuje slika 42 (pri predpisani toleranci oz. točnosti postopka).

optimum je:

X =

0.5

0.8333333

0.7355556

0.7214060



**Vidimo, da izredno hitro konvergira (glede na predpisano točnost)**

Slika 42: Rezultati pri iskanju optimuma funkcije (6.34.) s pomočjo Newtonove metode,

če je  $x_0 = 0.5$ .

Dobljene rezultate bi lahko strnili tudi v pregledno tabelo 5.

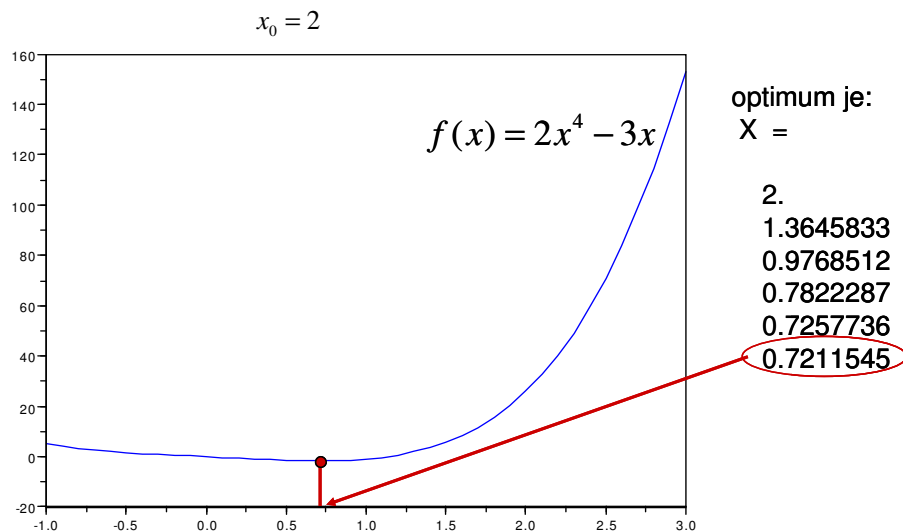
$x_0$	$k$	$x_{k+1}$	$\tilde{f}_{k+1}(x_{k+1})$	$x_k$	$f(x_k)$
0,5	0	0,83333	-1,7083	0,5	-1,375
0,5	1	0,73556	-1,6152	0,83333	-1,5355
0,5	2	0,72141	-1,6225	0,73556	-1,6212
0,5	3	0,72112	-1,6225	0,72141	-1,6225
0,5	4	0,72112	-1,6225	0,72112	-1,6225

Tabela 5: Pregledna tabela rezultatov pri iskanju optimuma funkcije (6.34.) s pomočjo Newtonove metode

Kot je razvidno iz tabele 5, metoda konvergira v ekstrem z vrednostjo 0.7211, kar je seveda enak rezultat, kot ga je dala za isti primer metoda zlatega reza (glej izraz (6.27.) in sliko 40). Vendar pa je bila Newtonova metoda bistveno hitrejša, saj je ekstrem našla že po 4 iteracijah, metoda zlatega reza pa je za to potrebovala kar 26 iteracij.

Iz tabele 5 so razvidne tudi vrednosti funkcije  $f$  in njenih aproksimacijskih parabol  $\tilde{f}$  v posameznih točkah neodvisne spremenljivke  $x_k$ . Seveda so si te vrednosti tem bolj podobne, čim bolj smo blizu ekstrema funkcije.

Poglejmo si še, kakšen bi bil potek konvergence postopka, če bi si izbrali začetno vrednost  $x_0 = 2$ . Rezultati so v tem primeru prikazani na sliki 43.

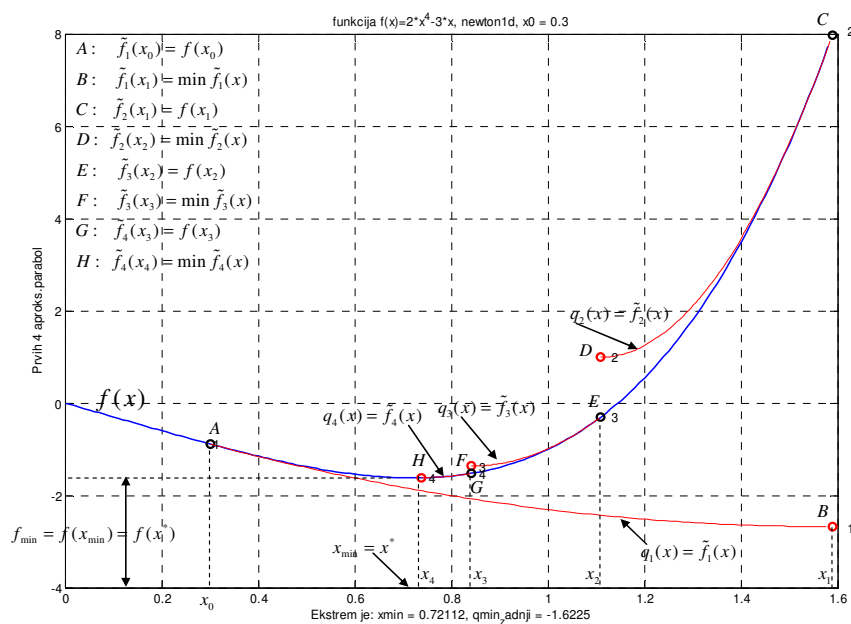


Slika 43: Rezultati pri iskanju optimuma funkcije (6.34.) s pomočjo Newtonove metode, če je  $x_0 = 2$ .



Na koncu še poudarimo, da ni nujno, da bi vrednost izračunanih aproksimacij  $x_k$  z naraščanjem iteracij nenehno le zvezno padala, kot je to na primeru na sliki 43 oz. osnovni sliki 41, ki ilustrira idejo Newtonove metode. Pride lahko tudi do “preskoka” na drugo stran krivulje, kar bomo opisali v nadaljevanju.

Če izberemo za primer funkcije (6.34.) začetno aproksimacijo  $x_0 = 0.3$ , dobimo potek konvergence postopka oz. potek posameznih aproksimacijskih parabol, kot je to prikazano na sliki 44.



Slika 44: Rezultati pri iskanju optimuma funkcije (6.34.) s pomočjo Newtonove metode, če je  $x_0 = 0.3$ .

Kot je razvidno iz slike 44, začnemo postopek v točki A. Poiščemo nato prvo aproksimacijsko parabelo (v 1. iteraciji) ter njen minimum (točko B), kar pa nas posledično prestavi v točko C na desni strani (preskok!). To je izhodišče nove aproksimacijske parabole (v 2. iteraciji), katere minimum (točka D) nas premakne v točko E in postopek se nadaljuje. Torej se šele od točke C naprej postopek zvezno približuje pravemu ekstremu funkcije 0.7211, predtem pa je prišlo tudi do preskoka.

#### 6.2.4 Sekantna metoda

Sekantno metodo lahko izpeljemo iz Newtonove metode, če drugi odvod v izrazu (6.33.) nadomestimo z naslednjo aproksimacijo [17]:

$$f''(x_k) \approx \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}} \quad (6.38.)$$

Potem izraz (6.33.) preide v naslednjo obliko [17]:

$$x_{k+1} = x_k - \frac{f'(x_k)}{\frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}} = x_k - \frac{f'(x_k)(x_k - x_{k-1})}{f'(x_k) - f'(x_{k-1})}, \quad k = 1, 2, \dots \quad (6.39.)$$

kar je iteracijska enačba za sekantno metodo.

Iteracijski postopek potem poteka na naslednji način:

$$\begin{aligned} x_2 &= x_1 - \frac{f'(x_1)(x_1 - x_0)}{f'(x_1) - f'(x_0)}, \quad x_0 \text{ in } x_1 \text{ izberemo} \\ x_3 &= x_2 - \frac{f'(x_2)(x_2 - x_1)}{f'(x_2) - f'(x_1)} \\ x_4 &= x_3 - \frac{f'(x_3)(x_3 - x_2)}{f'(x_3) - f'(x_2)} \\ &\dots \end{aligned} \quad (6.40.)$$

pri čemer si moramo tokrat, za razliko od Newtonove metode, izbrati dve začetni točki  $x_0, x_1$ .

Prikaz delovanja sekantne metode bomo osvetlili na naslednjem primeru:

#### **Primer:**

S pomočjo sekantne metode poiščite ekstrem naslednje funkcije,

$$f(x) = 2x^4 - 3x \quad (6.41.)$$

pri čemer vzemite začetni točki  $x_0 = 0, x_1 = 0.5$ .

**Rešitev:**

Najprej izračunajmo vrednosti 1. odvoda funkcije, saj ga bomo potrebovali v postopku:

$$f'(x) = 8x^3 - 3 \quad (6.42.)$$

Odvod (6.42.) nato vstavimo v iteracijsko enačbo (6.39.), pri čemer dobimo:

$$\begin{aligned} x_{k+1} &= x_k - \frac{(8x_k^3 - 3)(x_k - x_{k-1})}{(8x_k^3 - 3) - (8x_{k-1}^3 - 3)} = \\ &= x_k - \frac{(8x_k^3 - 3)(x_k - x_{k-1})}{8(x_k^3 - x_{k-1}^3)} = \\ &= x_k - \frac{8\left(x_k^3 - \frac{3}{8}\right)(x_k - x_{k-1})}{8(x_k - x_{k-1})(x_k^2 + x_k x_{k-1} + x_{k-1}^2)} = \\ &= x_k - \frac{\left(x_k^3 - \frac{3}{8}\right)}{(x_k^2 + x_k x_{k-1} + x_{k-1}^2)}, \quad k = 1, 2, \dots \end{aligned} \quad (6.43.)$$

Če bi želeli enačbo (6.43.) ročno izračunavati za posamezne iteracije, bi dobili pri  $x_0 = 0, x_1 = 0.5$  naslednji rezultat za prvo iteracijo:

$$\begin{aligned} x_2 &= x_1 - \frac{f'(x_1)(x_1 - x_0)}{f'(x_1) - f'(x_0)} = x_1 - \frac{(8x_1^3 - 3)(x_1 - x_0)}{8x_1^3 - 3 - (8x_0^3 - 3)} = \\ &= x_1 - \frac{(8x_1^3 - 3)(x_1 - x_0)}{8(x_1^3 - x_0^3)} = x_1 - \frac{8\left(x_1^3 - \frac{3}{8}\right)(x_1 - x_0)}{8(x_1 - x_0)(x_1^2 + x_1 x_0 + x_0^2)} = \\ &= x_1 - \frac{\left(x_1^3 - \frac{3}{8}\right)}{(x_1^2 + x_1 x_0 + x_0^2)} = 0.5 - \frac{\left(0.5^3 - \frac{3}{8}\right)}{(0.5^2 + 0.5 \cdot 0 + 0^2)} = \\ &= 0.5 - \frac{\left(0.5^3 - \frac{3}{8}\right)}{(0.5^2)} = 1.5 \end{aligned} \quad (6.44.)$$

Za 2. iteracijo pa bi dobili rezultat:

$$\begin{aligned}
 x_3 &= x_2 - \frac{f'(x_2)(x_2 - x_1)}{f'(x_2) - f'(x_1)} = x_2 - \frac{(8x_2^3 - 3)(x_2 - x_1)}{8x_2^3 - 3 - (8x_1^3 - 3)} = \dots = \\
 &= x_2 - \frac{\left(x_2^3 - \frac{3}{8}\right)}{\left(x_2^2 + x_2x_1 + x_1^2\right)} = 1.5 - \frac{\left(1.5^3 - \frac{3}{8}\right)}{\left(1.5^2 + 1.5 \cdot 0.5 + 0.5^2\right)} = 0.5769
 \end{aligned}
 \tag{6.45.}$$

Podobno bi nato računali za vse naslednje iteracije.

Natanko tako, kot je prikazano v izrazih (6.44.) in (6.45.), išče ekstrem funkcije tudi računalniški program. Pri tem bi bile poiskane vrednosti aproksimacij v posameznih iteracijah takšne, kot jih prikazuje slika 45 (pri predpisani toleranci oz. točnosti postopka).

optimum je:  
X =

0.  
0.5  
1.5  
0.5769231  
0.6299871  
0.7443037  
0.7180070  
0.7210264

Slika 45: Rezultati pri iskanju optimuma funkcije (6.41.) s pomočjo sekantne metode, če je  $x_0 = 0, x_1 = 0.5$ .

Dobljene rezultate bi lahko strnili tudi v pregledno tabelo 6.

$k$	$x_{k+1}$	$x_k$	$x_{k-1}$	$f'(x_k)$	$f'(x_{k-1})$
1	1,5	0,5	0	-2	-3
2	0,57692	1,5	0,5	24	-2
3	0,62999	0,57692	1,5	-1,4638	24
4	0,7443	0,62999	0,57692	-0,99975	-1,4638
5	0,71801	0,7443	0,62999	0,29868	-0,99975
6	0,72103	0,71801	0,7443	-0,038743	0,29868
7	0,72113	0,72103	0,71801	-0,0012276	-0,038743
8	0,72112	0,72113	0,72103	5,32E-06	-0,0012276

Tabela 6: Pregledna tabela rezultatov pri iskanju optimuma funkcije (6.41.) s pomočjo sekantne metode

Kot je razvidno iz tabele 6, metoda konvergira v ekstrem z vrednostjo 0.7211, kar je seveda enak rezultat, kot sta ga dali za isti primer metoda zlatega reza (glej izraz (6.27.) in sliko 40) in Newtonova metoda (glej tabelo 5). Iz primerjave rezultatov konvergence vseh treh metod (Newtonova metoda – 4 iteracije, Sekantna metoda – 8 iteracij, metoda zlatega reza – 26 iteracij) lahko sklepamo, da je sekantna metoda nekoliko počasnejša od Newtonove, vendar še vedno občutno hitrejša od metode zlatega reza. Razlog je v tem, da nosi manj informacije o obnašanju funkcije kot Newtonova metoda (saj ni 2. odvoda), vendar še vedno več kot pri metodi zlatega reza (opazujemo namreč tudi 1. odvod funkcije, ki ga pri metodi zlatega reza nismo).

Na koncu tega poglavja še poudarimo, da smo obravnavali le tri tipične metode, metodo zlatega reza, Newtonovo in sekantno metodo. Seveda obstaja še več enodimenzionalnih metod, kot npr. Fibonaccijeva metoda, metoda aproksimacije s parabolami, itn, katere si lahko bralec bolj natančno pogleda v literaturi [14, 17, 21].

### **6.3 Multidimenzionalna optimizacija**

V nadaljevanju si bomo pogledali optimizacijske metode za nelinearne primere, ko imamo opravka z več odločitvenimi spremenljivkami. Najprej si bomo pogledali bolj preproste, osnovne metode. Nato pa tudi nekoliko bolj zahtevne, sofisticirane metode.

Kot bomo videli, bodo nekatere izmed obravnavanih metod imele direkten značaj (opazujemo le funkcijo), druge pa gradienten značaj (opazujemo tudi parcialne odvode funkcije).

#### **6.3.1 Osnovne metode**

V tem poglavju si bomo pogledali osnovne optimizacijske metode nelinearnega programiranja. Opravka bomo imeli z dvema direktnima metodama, nato pa še z dvema gradientnima metodama.

### 6.3.1.1 Direktna Hook-Jeevesova metoda vzorčnih premikov

Metoda Hookea in Jeevesa [4, 6, 14, 21] spada med neposredne (direktne) metode in temelji le na računanju ciljne funkcije v različnih točkah prostora odločitvenih spremenljivk. To pomeni, da pri tej metodi izračunavanje odvodov ciljne funkcije ni potrebno, kar pa nekoliko upočasni hitrost konvergence. Metoda je enostavna in zanesljiva in zato zelo popularna v praksi.

Gre za dvokoračno metodo, kjer se v 1. koraku izvaja premikanje proti najbolj ugodni smeri (kjer se ciljna funkcija najbolj pomanjša), v 2. koraku pa se izvede še morebiten dodaten skok v izbrani (najbolj ugodni) smeri.

Metoda temelji na naslednji ideji. Pri vsaki aproksimaciji  $x_p(k) = (x_{p1}(k), x_{p2}(k), \dots, x_{pn}(k))^T$  optimalnega vektorja odločitvenih spremenljivk  $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$  se določijo smeri, pri katerih ima ciljna funkcija »hribe« (naraščanje) in »doline« (padanje), ter se nato »doline« sledijo do nove aproksimacije  $x_p(k+1)$ . To pomeni, da se vsaka iteracija sestoji iz dveh delov: iskanja najbolj ugodne smeri (kjer se ciljna funkcija najbolj pomanjša), ter nato premika v izbrani smeri.

Osnovni princip delovanja metode je predstavljen na slikah 46 in 47 [17]. V prvem koraku vsake iteracije  $k$  se najprej poišče najbolj ugodna smer za vektor  $x_p(k) = x_p^k$ , ki se na začetku postopka nahaja v začetnem »baznem« vektorju ( $x_b$ ), ki je hkrati tudi začetni vektor  $x_0$ . Tekom iskanja najbolj ugodne smeri gre postopek preko vseh dimenzij (koordinatnih osi) prostora odločitvenih spremenljivk (slika 46) in vzdolž vsake dimenzije  $i$  ( $i=1,2,\dots,n$ ) napravi pozitiven ali negativen »korak«  $\Delta x_i(k) = d_i(k) = d_i^k$ :

$$x_p^k \pm \Delta x_i^k = x_p^k \pm d_i^k, \quad i = 1, 2, \dots, n \quad (6.46.)$$

kjer so »koraki«:

$$d_1^k = \begin{bmatrix} \delta(k) \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \quad d_2^k = \begin{bmatrix} 0 \\ \delta(k) \\ 0 \\ \dots \\ 0 \end{bmatrix}, \quad \dots, \quad d_n^k = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dots \\ \delta(k) \end{bmatrix} \quad (6.47.)$$

in je  $\delta(k)$  izbrano pozitivno število (ki se po potrebi lahko še (večkrat) prepolovi).

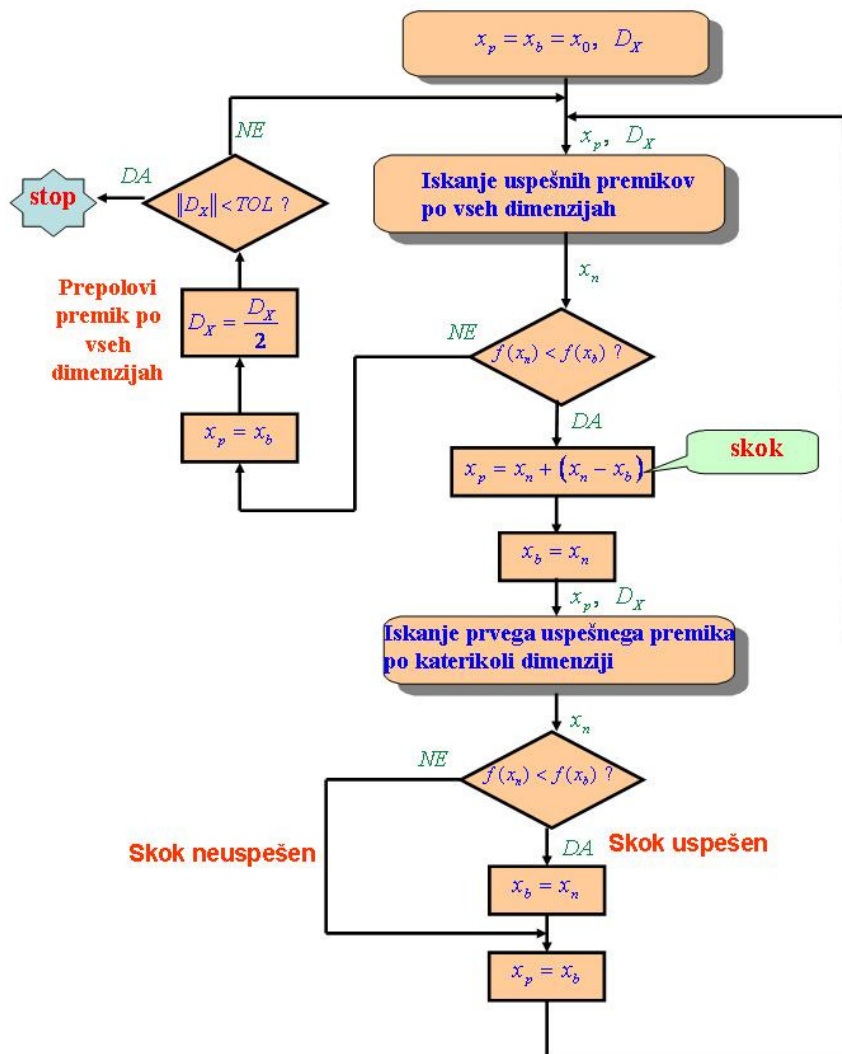
Kot je razvidno iz postopka premikanja vzdolž vseh dimenzij na sliki 47, se pri  $i$ -ti dimenziji vrednost vektorja  $x$  poveča ali pomanjša za korak  $\Delta x$ , če se vrednost funkcije v smeri tega premika pomanjša ( $f(x \pm \Delta x) < f(x)$ ), sicer pa ostane nespremenjena.

Ko je iskanje najbolj ugodne smeri po vseh dimenzijah uspešno končano (je  $f(x_n) < f(x_b)$ ) in se dobi za rezultat vektor  $x_n$ , se izvrši še »skok« v tej smeri (glej sliko 46):

$$x_p = x_n + (x_n - x_b) \quad (6.48.)$$

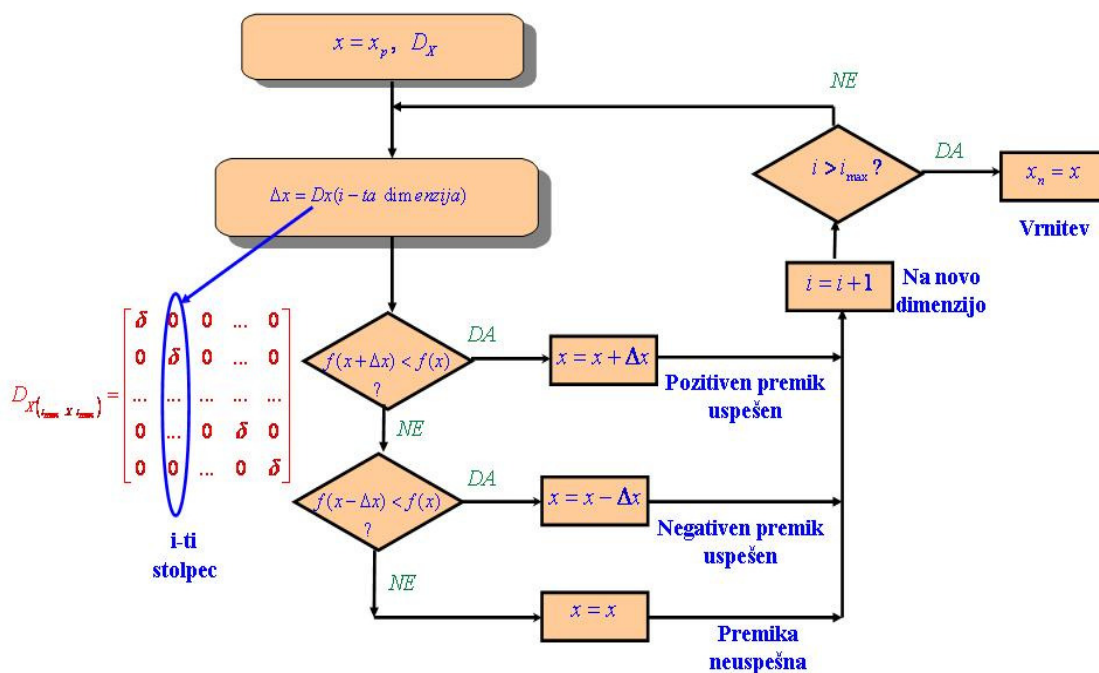
kjer je  $x_b$  star bazni vektor (na začetku hkrati enak tudi  $x_0$ ). Skok se izvrši pod predpostavko, da se »dolina« še naprej nadaljuje s spuščanjem (manjšanjem) ciljne funkcije v dotični, najbolj ugodni smeri.

Poleg tega si postopek tudi zapomne (glej sliko 46) nov bazni vektor:  $x_b = x_n$ , saj je  $x_n$ , ki ga je izračunal algoritem za uspešno iskanje najbolj ugodne smeri, vodil do pomanjšanja ciljne funkcije (kar je predpogoj za postavitev novega, boljšega baznega vektorja).



Slika 46: Osnovni princip delovanja Hook-Jeevesove metode (osnovna varianta)





Slika 47: Algoritem za iskanje uspešnih premikov po vseh dimenzijah (iskanje najbolj ugodne smeri)

Sledi preverjanje uspešnosti skoka (glej sliko 46), kar pomeni, da se zopet uporabi postopek na sliki 47, z edino razliko, da se poišče le prvi uspešen premik (po skoku) po katerikoli dimenziji (kar da za rezultat nov vektor  $x_n$ ). Če je skok uspešen (se ciljna funkcija pomanjša še bolj glede na prej), predstavljata skok+prvi uspešen premik po skoku dodatno korekcijo baznega vektorja. Če temu ni tako in je skok neuspešen, pa bazni vektor ostane nespremenjen glede na situacijo pred preverjanjem uspešnosti skoka.

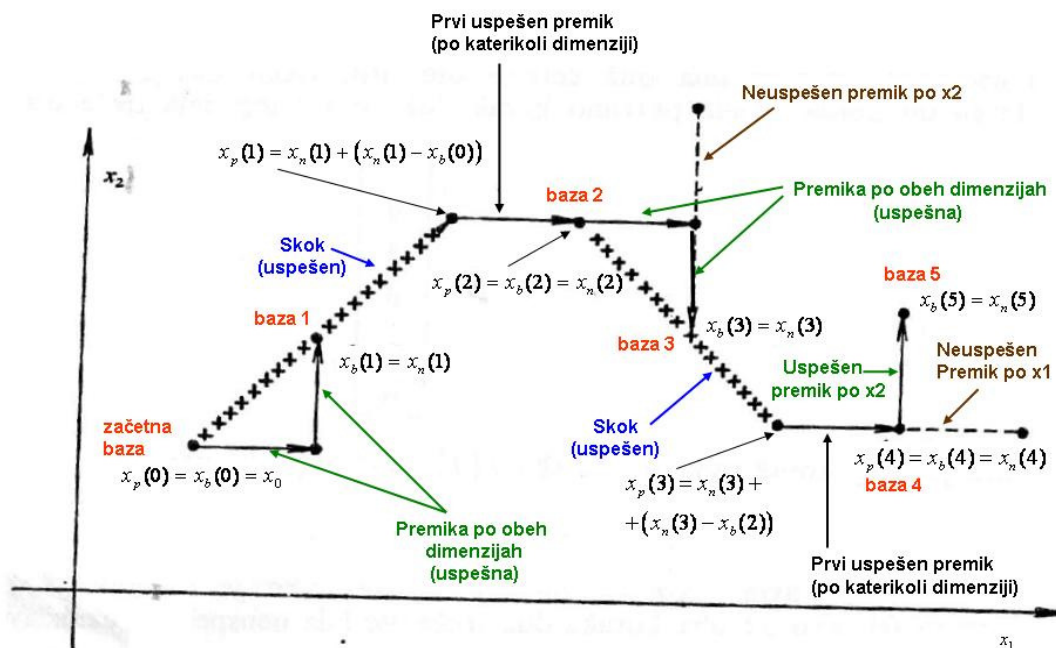
Ko algoritem preveri uspešnost ali neuspešnost skoka, aproksimacija  $x_p(k)$  zavzame vrednost dokončno izračunanega baznega vektorja (v trenutni iteraciji), nakar se postopek ponovi za naslednjo iteracijo.

Če se v določeni iteraciji zgodi (glej sliko 46), da iskanje najbolj ugodne smeri (uspešnih premikov po vseh dimenzijah) ne da več uporabnih rezultatov (se ciljna funkcija več ne pomanjša), to pomeni, da se je postopek optimalni rešitvi že dokaj približal. Če se želijo rezultati še bolj natančno določiti, se mora v ta namen prepoloviti premike (»korake«) po

vseh dimenzijah (prepoloviti  $\delta(k)$  v »korakih« izrazov (6.47.)), ter nato nadaljevati postopek. Tovrstna prepolovitev premikov se izvrši vsakič, ko znova pride do takšne situacije. Postopek se konča, ko postane norma matrike »korakov« v izrazu (6.47.)  $\|D_X(k)\| = \left\| \begin{bmatrix} d_1^k & d_2^k & \dots & d_n^k \end{bmatrix} \right\|$  manjša od predpisane tolerance  $TOL$  (ki je neko predpisano število). Tedaj zadnji bazni vektor, ki ga doseže postopek, predstavlja končno (željeno) aproksimacijo optimalne rešitve.

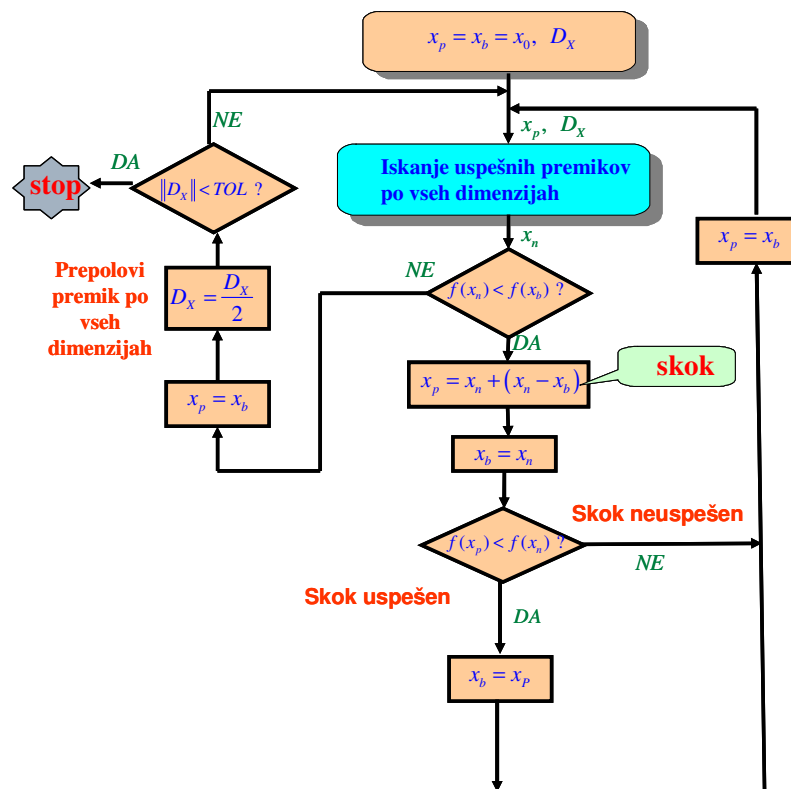
Hooke-Jeevesova direktna optimizacijska metoda je ilustrirana z dvodimenzionalnim zgledom na sliki 48 [17].

Prikazanih je nekoliko začetnih iteracij. Iz slike je razvidno, da se vedno izvede dvokoračni postopek. Najprej se poskušata izvesti premika po obeh dimenzijah, nato pa skok in prvi premik po skoku. V primeru uspešnosti teh premikov se seveda ciljna funkcija nenehno zmanjšuje, ko postopek napreduje z aproksimacijo  $x_p(k)$  preko baznih vektorjev 1, 2, 3, 4 in 5.



Slika 48: Ilustracija Hooke-Jeevesove metode na dvodimenzionalnem zgledu

Pri Hooke-Jeevesovi metodi moramo še poudariti, da obstaja tudi takoimenovana poenostavljena varianta metode, prikazana na sliki 49 [17].



Slika 49: Osnovni princip delovanja Hook-Jeevesove metode (poenostavljena varianta)

Kot je razvidno iz primerjave slik 46 in 49, poenostavljena varianta ne vsebuje bloka za iskanje prvega uspešnega premika (po katerikoli dimenziji) po opravljenem skoku, torej se v tem primeru dodatno ne verificira uspešnosti skoka.

V nadaljevanju si še na kratko pogledjmo tri primere, ki nam bodo pomagali dodatno osvetliti delovanje Hooke-Jeevesove metode.

### 1. primer:

S pomočjo Hooke-Jeevesove metode vzorčnih premikov določite optimalno točko za funkcijo treh odločitvenih spremenljivk  $x_1, x_2, x_3$ :

$$f(x_1, x_2, x_3) = x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3 \quad (6.49.)$$

kjer je začetni vektor odločitvenih spremenljivk enak:  $x(0) = x_0 = [x_1(0) \ x_2(0) \ x_3(0)] = [0 \ 0 \ 0]^T$  in je korak vzdolž dimenzij enak:  $\delta = 0.1$ .

**Rešitev:**

Pri iskanju optimalnih rezultatov si pomagamo z algoritmoma na slikah 46 in 47. Če ta dva algoritma apliciramo v obliki računalniškega programa, nam slednji da rezultate, prikazane na sliki 50.

<b>1.stolp: iteracija (1. in morebitni 2. korak)</b>	<b>2.-4.stolp: komponente baznega vektorja</b>	<b>5.stolp: vrednost funkcije</b>	<b>6.stolp: vrednost koraka</b>	<b>Število izračunavanj 1. oz. 2. koraka: 30</b>
0	0	0	0	0
1	0,1	-0,1	0,1	-0,2998
1	0,3	-0,2	0,2	-0,6813
2	0,4	-0,3	0,3	-0,9383
2	0,5	-0,5	0,4	-1,34
3	0,5	-0,6	0,5	-1,5429
3	0,5	-0,8	0,6	-1,8229
4	0,5	-0,9	0,7	-1,8914
4	0,6	-1	0,8	-1,8944
5	0,6	-0,9	0,8	-1,9003
6	0,55	-0,95	0,75	-1,9095
6	0,55	-1	0,7	-1,8871
7	0,55	-0,95	0,75	-1,9095
7	0,55	-0,95	0,8	-1,9081
8	0,6	-0,95	0,8	-1,9084
9	0,575	-0,95	0,775	-1,9113
9	0,575	-0,95	0,75	-1,9108
...	0,575	-0,95	0,775	-1,9113
...	0,575	-0,9375	0,775	-1,9117
...	0,56875	-0,9375	0,76875	-1,9117
...	0,56875	-0,9375	0,7625	-1,9117
...	0,56875	-0,9375	0,76875	-1,9117
...	0,57188	-0,94063	0,76875	-1,9118
...	0,57188	-0,94375	0,76875	-1,9117
...	0,57188	-0,94063	0,76875	-1,9118
...	0,57031	-0,93906	0,76719	-1,9118
...	0,57031	-0,9375	0,76563	-1,9118
...	0,57031	-0,93906	0,76719	-1,9118
...	0,57031	-0,93906	0,76875	-1,9118

**Izračunan minimum :**

$$\tilde{x}^* = [\tilde{x}_1^* \quad \tilde{x}_2^* \quad \tilde{x}_3^*]^T$$

Slika 50: Rezultati pri iskanju optimuma funkcije (6.49.) s pomočjo Hooke-Jeevesove metode

Kot je razvidno iz slike 50, je bilo izvedenih 30 izračunavanj 1. oz. 2. koraka postopka. Pri tem se je korak vzdolž dimenzij tudi večkrat prepolovil, kar je razvidno iz 6. stolpca na sliki 50.

Končni rezultat za aproksimacijski vektor odločitvenih spremenljivk se torej glasi (numerični izračun je bil izveden v programskem jeziku Matlabu):

$$x_p(N) = \tilde{x}^* = [\tilde{x}_1^* \quad \tilde{x}_2^* \quad \tilde{x}_3^*]^T = [0,57031 \quad -0,93906 \quad 0,76875]^T \quad (6.50.)$$

v katerem funkcija (6.49.) zavzame naslednjo minimalno vrednost:

$$f(\tilde{x}_1^*, \tilde{x}_2^*, \tilde{x}_3^*) = \min f(x_1, x_2, x_3) = -1,9118 \quad (6.51.)$$

## 2. primer:

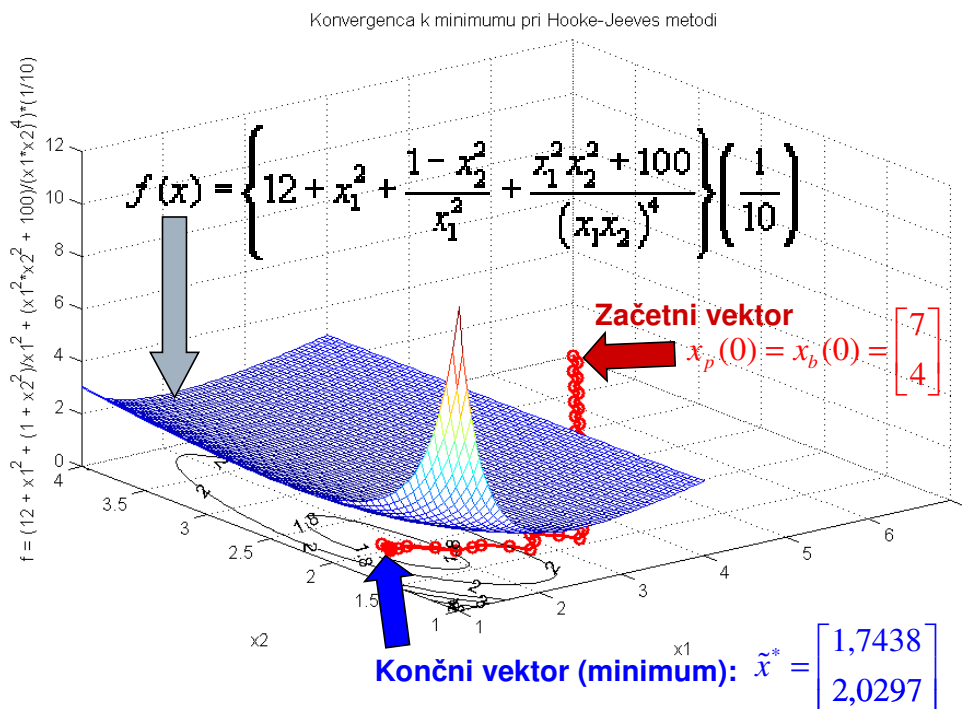
S pomočjo Hooke-Jeevesove metode vzorčnih premikov določite optimalno točko za naslednjo funkcijo dveh odločitvenih spremenljivk  $x_1, x_2$ :

$$f(x) = \left\{ 12 + x_1^2 + \frac{1 + x_2^2}{x_1^2} + \frac{x_1^2 x_2^2 + 100}{(x_1 x_2)^4} \right\} \begin{pmatrix} 1 \\ 10 \end{pmatrix} \quad (6.52.)$$

kjer je začetni vektor odločitvenih spremenljivk enak:  $x(0) = x_0 = [x_1(0) \ x_2(0)]^T = [7 \ 4]^T$  in je korak vzdolž dimenzij enak:  $\delta = 0.1$  ( $TOL = 10^{-3}$ ).

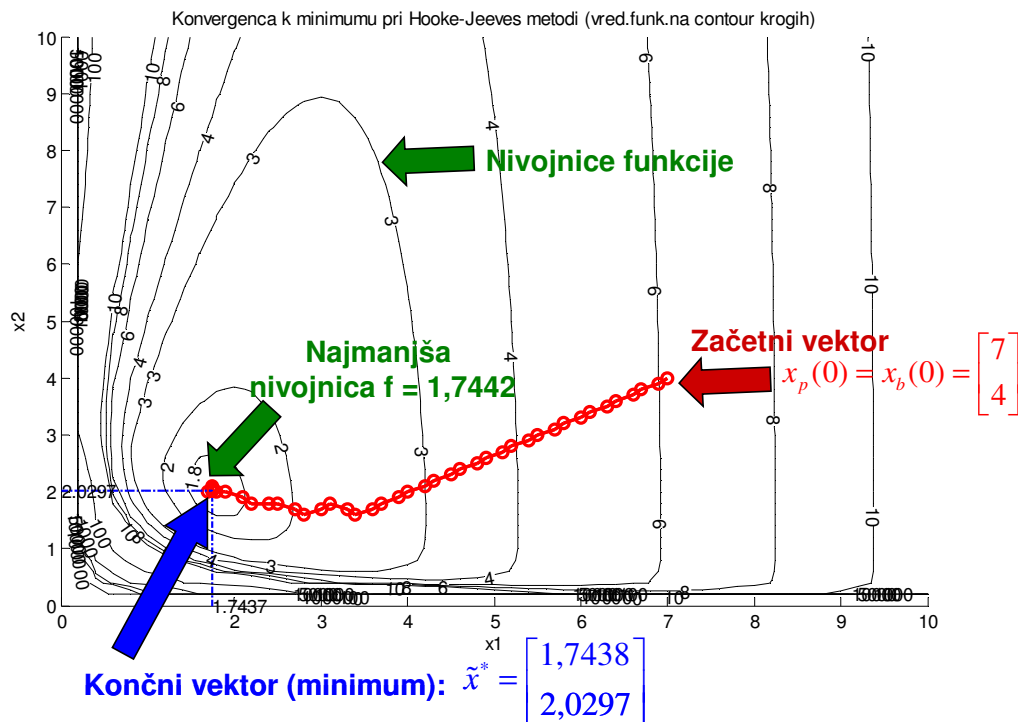
## Rešitev:

Pri iskanju optimalnih rezultatov si ponovno pomagamo z algoritmoma na slikah 46 in 47. Če ta dva algoritma apliciramo v obliki računalniškega programa, nam slednji da dinamiko konvergence rezultatov, prikazano na sliki 51.



Slika 51: Dinamika konvergence rezultatov pri iskanju optimuma funkcije (6.52.) s pomočjo Hooke-Jeevesove metode, opazovana v prostoru  $(x_1, x_2, f(x_1, x_2))$  (slika A)

Dinamiko konvergence rezultatov, prikazano na sliki 51, si lahko še natančneje ogledamo na sliki 52, kjer so podane tudi nivojnice funkcije (6.52.), preslikane na  $x_1 - x_2$  ravnino.



Slika 52: Dinamika konvergence rezultatov pri iskanju optima funkcije (6.52.) s pomočjo Hooke-Jeevesove metode, opazovana na  $x_1 - x_2$  ravnini. (slika B)

Iz slik 51 in 52 je razvidno, kako v posameznih iteracijah vrednosti aproksimacij optimalnih odločitvenih spremenljivk konsistentno napredujejo od začetne točke  $(x_{p_1}(0), x_{p_2}(0)) = (7, 4)$  do končne (optimalne) točke  $(x_1^*, x_2^*) = (1.748, 2.0297)$ , ki se doseže po 45 iteracijah.

Tovrstno napredovanje (konvergiranje k optimalni rešitvi) je na slikah 51, 52 označeno s povezanimi krogi (o), označene pa so tudi nivojnice ciljne funkcije (10, 8, 6, 4, 3, 2, ...). Te pri konvergiranju postopka postopoma padajo od začetne vrednosti  $f(7, 4)$  do končne (minimalne) vrednosti 1.7442.

Na sliki 51 spodaj je tudi razvidno, da funkcija  $f(x_1, x_2)$  dejansko doseže najnižjo točko (minimum) v prostoru  $(x_1, x_2, f(x_1, x_2))$ , ko postopek konvergira k optimalni rešitvi  $(x_1^*, x_2^*) = (1.748, 2.0297)$ .

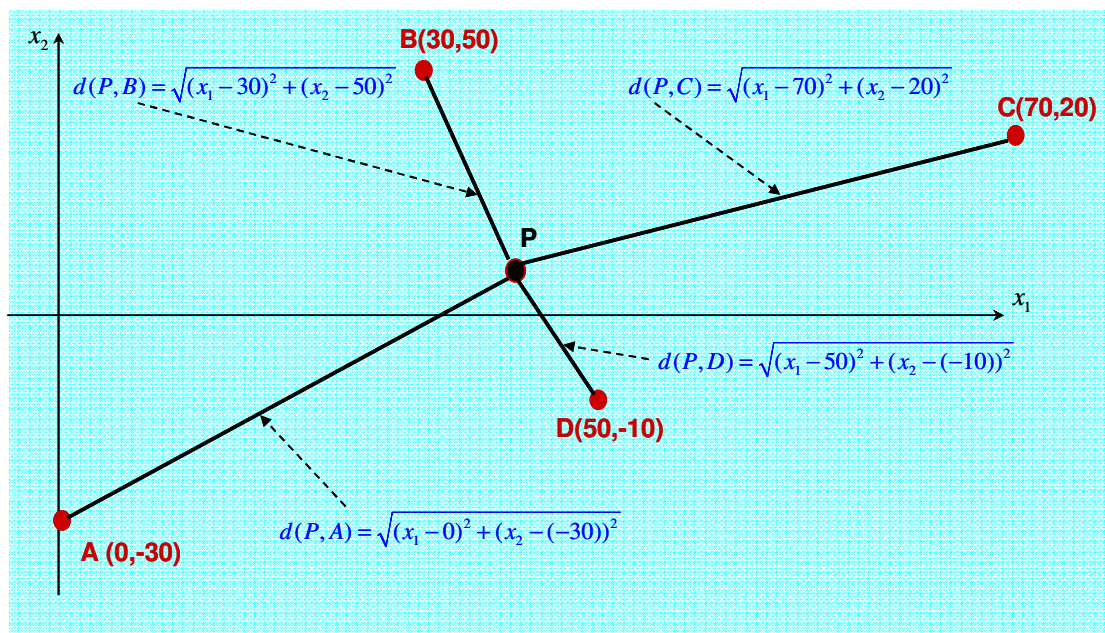
Končni optimalni rezultati torej so:

$$\tilde{x}^* = \begin{bmatrix} \tilde{x}_1^* & \tilde{x}_2^* \end{bmatrix}^T = \begin{bmatrix} 1.7438 & 2.0297 \end{bmatrix}^T \quad (6.53.)$$

$$f(\tilde{x}^*) = \min f(x) = 1.7442$$

### 3. primer:

Potrebno je določiti lokacijo oskrbovalnega centra  $P$ , od koder se bodo s helikopterji oskrbovala težko dostopna črpališča nafte  $A$ ,  $B$ ,  $C$  in  $D$ . Podjetje za črpanje nafte želi določiti to lokacijo na takšen način, da bo vsota prevoženih razdalj helikopterjev od oskrbovalnega centra do črpališč minimalna. Pri tem imajo črpališča prostorske koordinate  $(x,y) = (0,-30)$ ,  $(30,50)$ ,  $(70,20)$  in  $(50,-10)$ . Problem je predstavljen na sliki 53. Poiščite optimalno rešitev s pomočjo Hooke-Jeevesove metode, pri čemer vzemite:  $(x_{p1}(0), x_{p2}(0)) = (10,10)$ ,  $\delta(0) = 1$ , ter  $TOL = 10^{-3}$ .



Slika 53: Problem načrtovanja optimalnih helikopterskih poti za potrebe oskrbe črpališč nafte

### Rešitev:

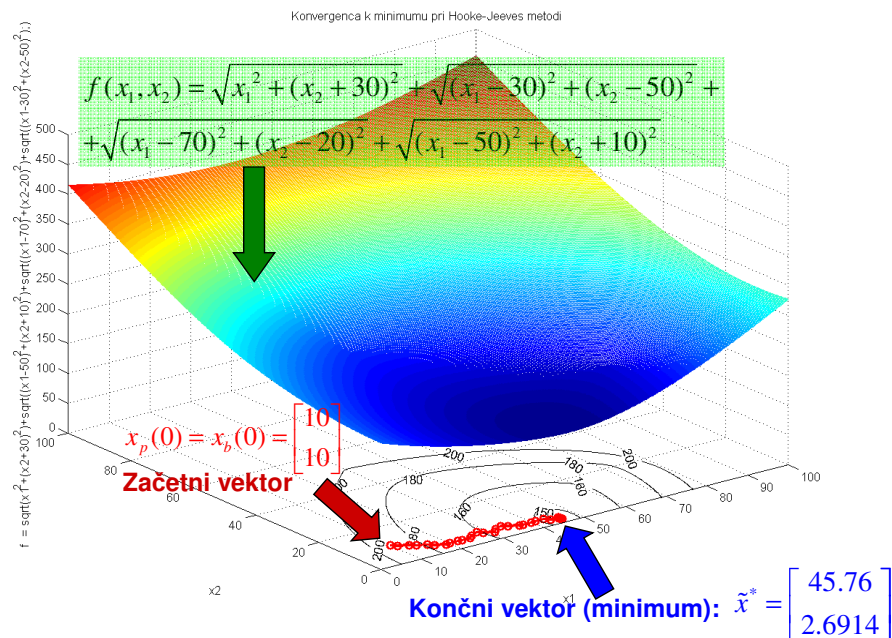
Iz slike 53 je razvidno, da imajo izrazi za posamezne razdalje ( $d(P,A)$ ,  $d(P,B)$ ,  $d(P,C)$  in  $d(P,D)$ ) nelinearno obliko kvadratnega korena. Če definiramo ciljno funkcijo kot vsoto teh razdalj, dobimo naslednji nelinearni izraz:

$$f(x_1, x_2) = \sqrt{x_1^2 + (x_2 + 30)^2} + \sqrt{(x_1 - 30)^2 + (x_2 - 50)^2} + \sqrt{(x_1 - 70)^2 + (x_2 - 20)^2} + \sqrt{(x_1 - 50)^2 + (x_2 + 10)^2} \quad (6.54.)$$

Lokacijo oskrbovalnega centra P je potrebno postaviti pri tistih optimalnih prostorskih koordinatah  $(x_1^*, x_2^*)$ , pri katerih bo ciljna funkcija dosegla minimum:  $f(x_1^*, x_2^*) = \min f(x_1, x_2)$ .

Kot je razvidno iz predstavitve problema, niso v njem predpisane nikakršne omejitve. Torej gre za problem optimizacije brez omejitev. Ker se pri problemu pojavita dve neodvisni spremenljivki,  $(x_1, x_2)$ , gre očitno za problem večdimenzionalne optimizacije.

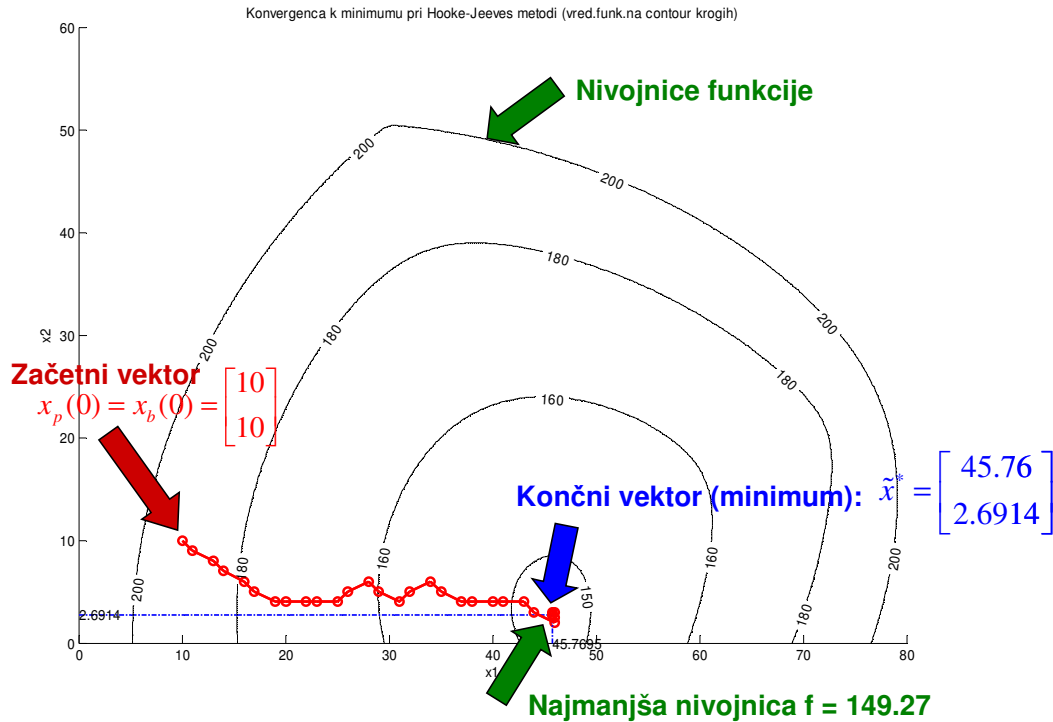
Pri iskanju optimalnih rezultatov si ponovno pomagamo z algoritmoma na slikah 46 in 47. Če ta dva algoritma apliciramo v obliki računalniškega programa, nam slednji da dinamiko konvergence rezultatov, prikazano na sliki 54.



Slika 54: Dinamika konvergence rezultatov pri iskanju optimuma funkcije (6.54.) s pomočjo Hooke-Jeevesove metode, opazovana v prostoru  $(x_1, x_2, f(x_1, x_2))$  (slika A)



Dinamiko konvergence rezultatov, prikazano na sliki 54, si lahko še natančneje ogledamo na sliki 55, kjer so podane tudi nivojnice funkcije (6.54.), preslikane na  $x_1 - x_2$  ravnino.



Slika 55: Dinamika konvergence rezultatov pri iskanju optimuma funkcije (6.54.) s pomočjo Hooke-Jeevesove metode, opazovana na  $x_1 - x_2$  ravnini. (slika B)

Iz slik 54 in 55 je razvidno, kako v posameznih iteracijah vrednosti aproksimacij optimalnih odločitvenih spremenljivk konsistentno napredujejo od začetne točke  $(x_{p1}(0), x_{p2}(0)) = (10, 10)$  do končne (optimalne) točke  $(x_{p1}(36), x_{p2}(36)) = (45.76, 2.6914)$ . Pri tem je bilo izvedenih 36 izračunavanj 1. oz. 2. koraka postopka.

Tovrstno napredovanje (konvergiranje k optimalni rešitvi) je na slikah 54, 55 označeno s povezanimi krogi (o), označene pa so tudi nivojnice ciljne funkcije (z resolucijo 20: 200, 180, 160, ...). Te pri konvergiranju postopoma padajo od začetne vrednosti  $f(x_{p1}(0), x_{p2}(0)) = f(10, 10) = 191.5$  do končne (minimalne) vrednosti 149.27.

Na sliki 54 spodaj je tudi razvidno, da funkcija  $f(x_1, x_2)$  dejansko doseže najnižjo točko (minimum) v prostoru  $(x_1, x_2, f(x_1, x_2))$ , ko postopek konvergira k optimalni rešitvi  $(x_{p1}(36), x_{p2}(36)) = (45.76, 2.6914)$ .

Končni optimalni rezultati torej so:

$$\tilde{x}^* = [\tilde{x}_1^* \quad \tilde{x}_2^*] ^T = [45.76 \quad 2.6914] ^T \quad (6.55.)$$

$$f(\tilde{x}^*) = \min f(x) = 149.27$$

Odgovor na rešitev problema bi se torej glasil:

**Lokacija oskrbovalnega centra P mora imeti koordinate (45.76,2.6914), da bo vsota prevoženih razdalj helikopterjev od oskrbovalnega centra do črpališč minimalna (enaka 149.27).**

### 6.3.1.2 Direktna Nelder-Meadova simpleksna metoda

Simpleksna metoda je verjetno najbolj uporabljena multidimenzionalna metoda v praksi. Uvrščamo jo med direktne metode, torej računamo le vrednosti ciljne funkcije brez odvodov. Metoda temelji na izračunavanju simpleksov in nas zaradi izjemne prožnosti oblike in velikosti simpleksa skoraj zagotovo pripelje do minimuma [13, 14, 17, 21].

Pojem simpleks izhaja iz topološke posplošitve trikotnika. Tako kot tvori trikotnik množica treh točk v dvodimenzionalni ravnini, tvori simpleks množica  $n+1$  temenskih točk v  $n$ -dimenzionalnem prostoru. Če so vse temenske točke medseboj enako oddaljene, govorimo o *pravilnem simpleksu*. Torej je v dvodimenzionalni ravnini pravilen simpleks enakostraničen trikotnik [13].

Osnovna ideja optimiranja ciljne funkcije  $f(x)$  z  $n$ -odločitvenimi spremenljivkami s simpleksno metodo je preprosta. Denimo iščemo minimum ciljne funkcije. Potem najprej v  $n$ -dimenzionalnem prostoru odločitvenih spremenljivk izberemo  $n+1$  temenskih točk, ki sestavljajo začetni simpleks okoli točke začetnega vektorja odločitvenih spremenljivk. Nato iz tega simpleksa izločimo temensko točko, pri kateri je vrednost ciljne funkcije največja, ter jo zamenjamo z novo »boljšo« točko. Postopek nato ponavljamo z novimi simpleksi. Ko v neki iteraciji ugotovimo, da leži optimalna točka znotraj določenega simpleksa, ga skrčimo in nadaljujemo postopek. Optimizacijski postopek se ustavi, ko so vse točke nekega simpleksa dovolj blizu skupaj, iz njih pa se da razbrati optimalna rešitev [13].

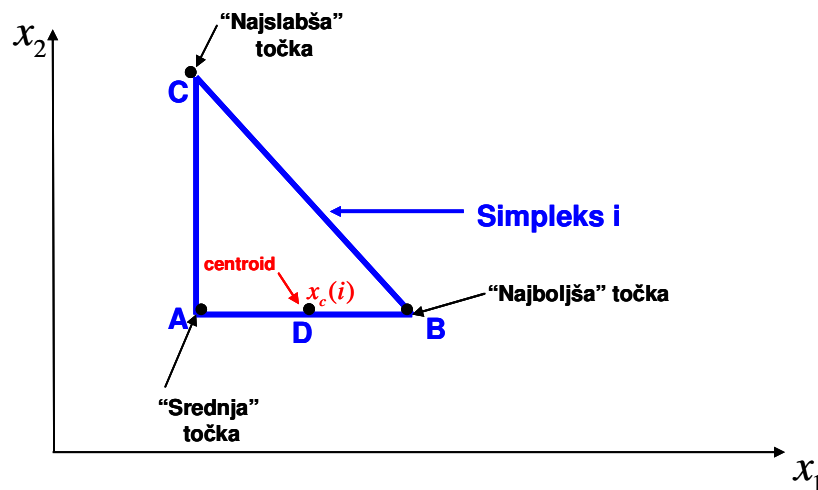
#### **Potek postopka tvorjenja novega simpleksa:**

1. ) **Klasifikacija temenskih točk:**

Pri tvorjenju novega simpleksa izvajamo različne operacije. Najprej med vsemi temenskimi točkami simpleksa  $x_1, x_2, \dots, x_{n+1}$  poiščemo vrednosti ciljne funkcije  $f(x_1, x_2, \dots, x_{n+1})$  za posamezne točke.

Če imamo opravka z dvema odločitvenima spremenljivkama  $x_1, x_2$  in dvodimenzionalnim prostorom  $x_1, x_2, f(x_1, x_2)$ , to pomeni, da glede na vrednost ciljne funkcije (pri iskanju minimuma) lahko ločimo med (glej sliko 56):

1. točko  $x_h$ , v kateri je vrednost ciljne funkcije največja (točka C).
2. točko  $x_s$ , v kateri ima ciljna funkcija drugo največjo vrednost (točka A).
3. točko  $x_l$ , v kateri je vrednost ciljne funkcije najmanjša (točka B).



Slika 56: Simpleks v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije

## 2.) Izračun centroida :

Po opravljeni klasifikaciji temenskih točk izračunamo tudi točko  $x_c$ , ki je centroid vseh temenskih točk  $x_i$  ( $i = 1, 2, \dots, n+1$ ), razen najslabše točke  $x_h$ . Za n-dimenzionalen prostor tako dobimo [13]:

$$x_c = \frac{1}{n} \sum_{\substack{i=1 \\ i \neq h}}^{n+1} x_i \quad (6.56.)$$

Pri dvodimenzionalnem prostoru pa ima centroid vrednost (glej sliko 56):

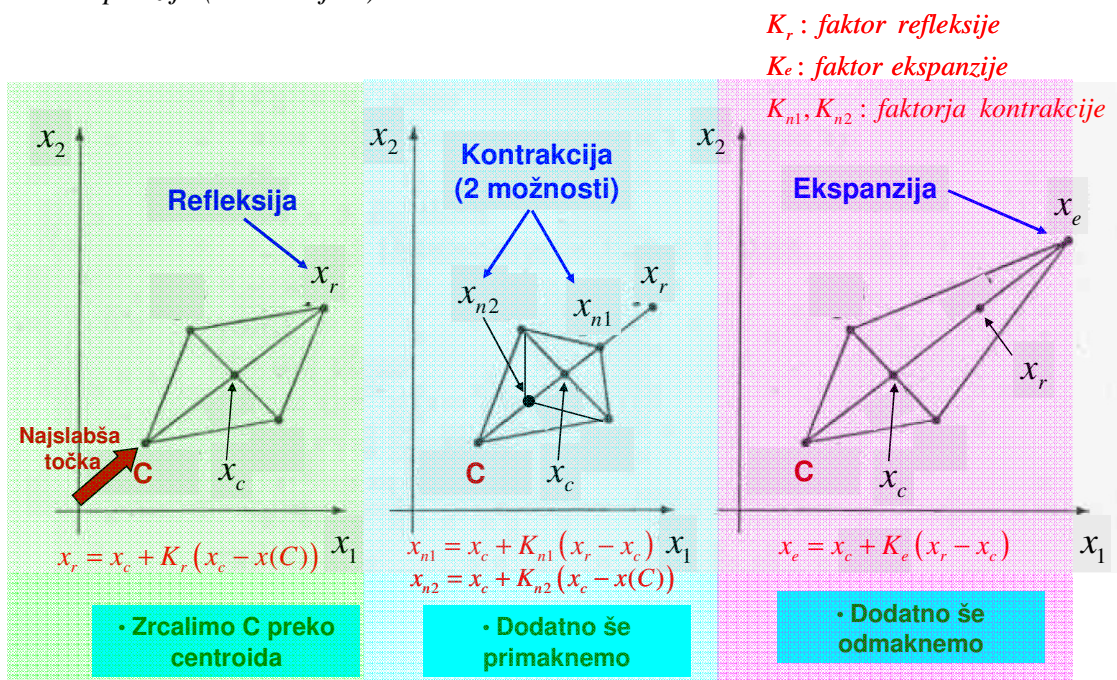
$$x_c(i) = \frac{1}{n} \sum_{\substack{k=1 \\ k \neq \text{indeks za } C}}^{n+1} x(k) = \frac{x(A) + x(B)}{2} = x(D) \quad (6.57.)$$

### 3.) Premik najslabše točke preko centroida, ter njena nadomestitev z boljšo točko:

Ko je centroid izračunan, ga lahko uporabimo za premik najslabše točke preko njega. Tako se lahko najslabše točke znebimo in jo nadomestimo z novo, boljšo točko.

Novo točko simpleksa lahko določimo s tremi operacijami (glej sliko 57) [13]:

- z *refleksijo* (zrcaljenjem),
- s *kontrakcijo* (primikanjem) in z
- *ekspanzijo* (odmikanjem).

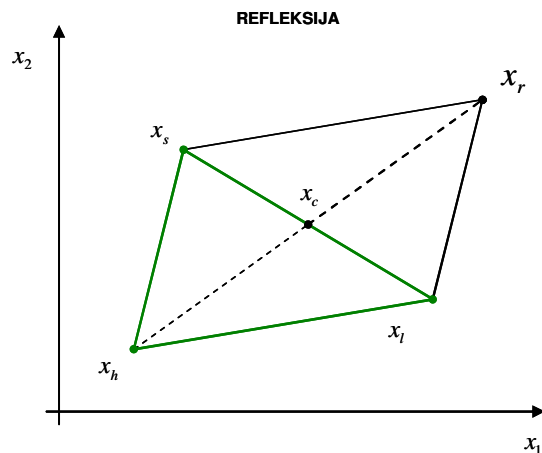


Slika 57: Tri možne operacije za določitev nove točke simpleksa v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije

Pri postopku refleksije točko, v kateri ima ciljna funkcija »najslabšo vrednost«  $x_h$ , zrcalimo preko centroidne točke  $x_c$  v točko  $x_r$  po relaciji [13] (glej sliko 58):

$$x_r = x_c + K_r(x_c - x_h) \quad (6.58.)$$

kjer je  $K_r$  koeficient zrcaljenja in podaja dolžino zrcaljenja. Slednjega določimo s poskušanjem, ponavadi pa uporabimo vrednost  $K_r = 1$  [13].



Slika 58: Prikaz postopka refleksije simpleksa v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije

Postopek refleksije (zrcaljenja) najslabše točke  $x_h$  preko centroida načeloma opravimo v vsakem primeru, če le nam da neko boljšo točko  $x_r$  (kjer je ciljna funkcija manjša).

Včasih pa je možno po opravljeni refleksiji dobljeno točko  $x_r$  še izboljšati, oz. jo nadomestiti s še boljšo točko (če se seveda ciljna funkcija v njej še bolj zmanjša). V ta namen vsakič testiramo, ali bi bilo smotno opraviti še dodatno kontrakcijo ali ekspanzijo.

#### Kontrakcija:

V primeru kontrakcije imamo 2 možnosti (glej sliko 59). Pri prvi možnosti lahko točko  $x_r$  še nekoliko približamo točki  $x_c$  in pridemo v točko  $x_{n1} = x_{k1}$ , pri čemer uporabimo izraz [13]:

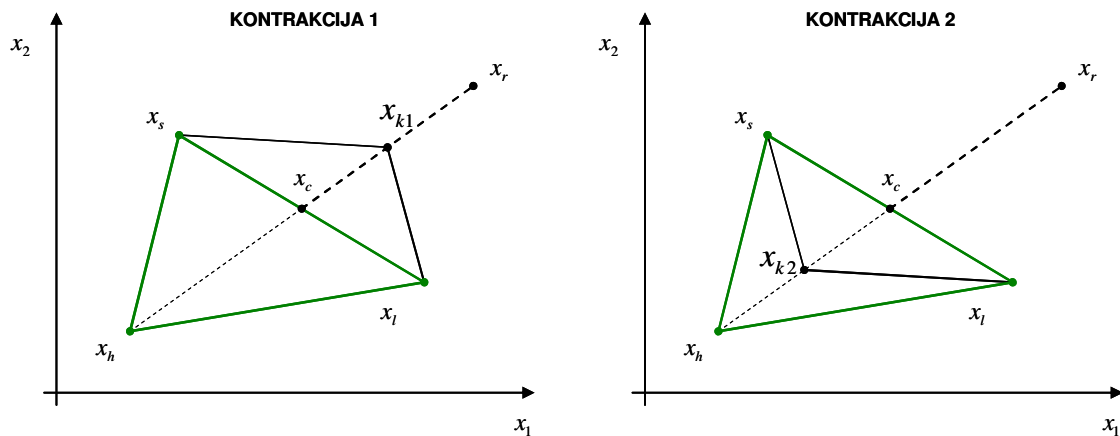
$$x_{n1} = x_c + K_{n1}(x_r - x_c) \quad (6.59.)$$

kjer je  $K_{n1}$  koeficient primika, manjši od ena (običajno vzamemo vrednost 0.5 [13]).

Pri drugi možnosti pa lahko točko  $x_r$  celo pomaknemo znotraj simpleksa in pridemo v točko  $x_{n2} = x_{k2}$ , pri čemer uporabimo izraz [13]:

$$x_{n2} = x_c + K_{n2}(x_c - x_h) \quad (6.60.)$$

kjer je  $K_{n2}$  koeficient primika, manjši od ena (običajno vzamemo vrednost 0.5 [13]).



Slika 59: Prikaz postopka kontrakcije simpleksa v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije (2 možnosti)

Na sliki 59 sta prikazani obe možnosti kontrakcije v dvodimenzionalni ravnini. Na levi sliki točko  $x_r$  nekoliko približamo točki  $x_c$ . To storimo tako, da jo primaknemo v točko  $x_{k1}$ , ki se nahaja zunaj simpleksa. Na desni sliki pa je prikazano, kako primaknemo točko  $x_r$  preko točke  $x_c$  v točko  $x_{k2}$ , ki se nahaja znotraj simpleksa.

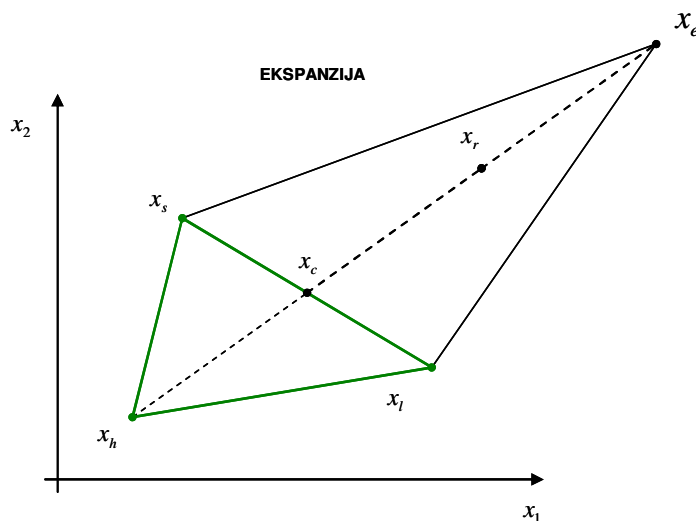
#### Ekspanzija:

Namesto kontrakcije pa v nekaterih primerih lahko točko  $x_r$  še dodatno odmaknemo od točke  $x_c$  s postopkom ekspanzije, pri čemer dobimo točko  $x_e$  na osnovi izraza (glej sliko 60) [13]:

$$x_e = x_c + K_e(x_r - x_c) \quad (6.61.)$$

kjer je  $K_e$  koeficient odmika, običajno ima vrednost  $K_e = 2$ .

Na sliki 60 je prikazan postopek ekspanzije v dvodimenzionalni ravnini, kjer točko  $x_r$  še bolj odmaknemo od točke  $x_c$  v  $x_e$ .



Slika 60: Prikaz postopka ekspanzije simpleksa v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije

### **Potek Nelder-Meadovega postopka:**

Nelder-meadov simpleksni postopek optimizacije se izvaja tako, da najprej izberemo začetni simpleks ter izračunamo vrednost ciljne funkcije v vseh temenskih točkah, poiščemo temena  $x_h$ ,  $x_s$  in  $x_l$  ter izračunamo  $x_c$  po enačbi (6.56). V nadaljevanju temensko točko  $x_h$  zrcalimo preko  $x_c$  v točko  $x_r$ . Če velja  $f(x_l) \leq f(x_r) \leq f(x_s)$ , zamenjamo  $x_h$  z  $x_r$  in dobimo nov simpleks, s katerim ponovimo postopek od začetka [13].

Če je  $f(x_r) \leq f(x_l)$ , ima zrcaljena točka nižjo vrednost ciljne funkcije od vseh simpleksnih točk. V tem primeru se pomaknemo še naprej v isti ugodni smeri v točko ekspanzije  $x_e$ . Nov simpleks sedaj dobimo tako, da točko  $x_h$  zamenjamo s tisto izmed točk  $x_r$  ali  $x_e$ , ki daje manjšo vrednost ciljne funkcije. Spet ponovimo postopek od začetka z novim simpleksom [13].

Če je  $f(x_s) \leq f(x_r) \leq f(x_h)$ , je ciljna funkcija v zrcaljeni točki nižja le od vrednosti v »najslabši« temenski točki, kar pomeni le rahlo izboljšanje. Da bi ugotovili, če smo boljše točko v tej smeri že preskočili, primaknemo  $x_r$  s postopkom kontrakcije v točko  $x_{k1}$ . Točko  $x_h$  zamenjamo s tisto od točk  $x_{k1}$  ali  $x_r$ , ki ima nižjo vrednost ciljne funkcije, ter se z novim simpleksom lotimo postopka od začetka [13].

Če je  $f(x_h) \leq f(x_r)$  ima zrcaljena točka višjo vrednost ciljne funkcije od vseh temenskih točk. V tem primeru sklepamo, da se že nahajamo v okolici optimalne točke in s postopkom kontrakcije točko  $x_r$  premaknemo v točko  $x_{k2}$ . Točko  $x_h$  zamenjamo s točko

$x_{k2}$ , če ima slednja nižjo vrednost ciljne funkcije. Nato se z novim simpleksom lotimo postopka od začetka [13].

V najslabšem primeru velja, da nobeden izmed postopkov refleksije, kontrakcije oziroma ekspanzije ne doseže nobenega izboljšanja. Ob tej ugotovitvi sklepamo, da leži minimum verjetno znotraj simpleksa, zato ga skrčimo okoli točke z najnižjo vrednostjo ciljne funkcije  $x_l$  tako, da spremenimo vse druge temenske točke simpleksa po naslednjem pravilu [13]:

$$x_i = \frac{1}{2}(x_i + x_l), \quad i = 1, 2, \dots, n + 1 \quad (6.62.)$$

Postopek se ustavi, ko je izpolnjen pogoj [13]:

$$\sqrt{\frac{1}{n} \sum_{\substack{i=1 \\ i \neq h}}^{n+1} [f(x_i) - f(x_l)]^2} \leq \varepsilon \quad (6.63.)$$

kjer je  $\varepsilon$  vnaprej podano majhno število.

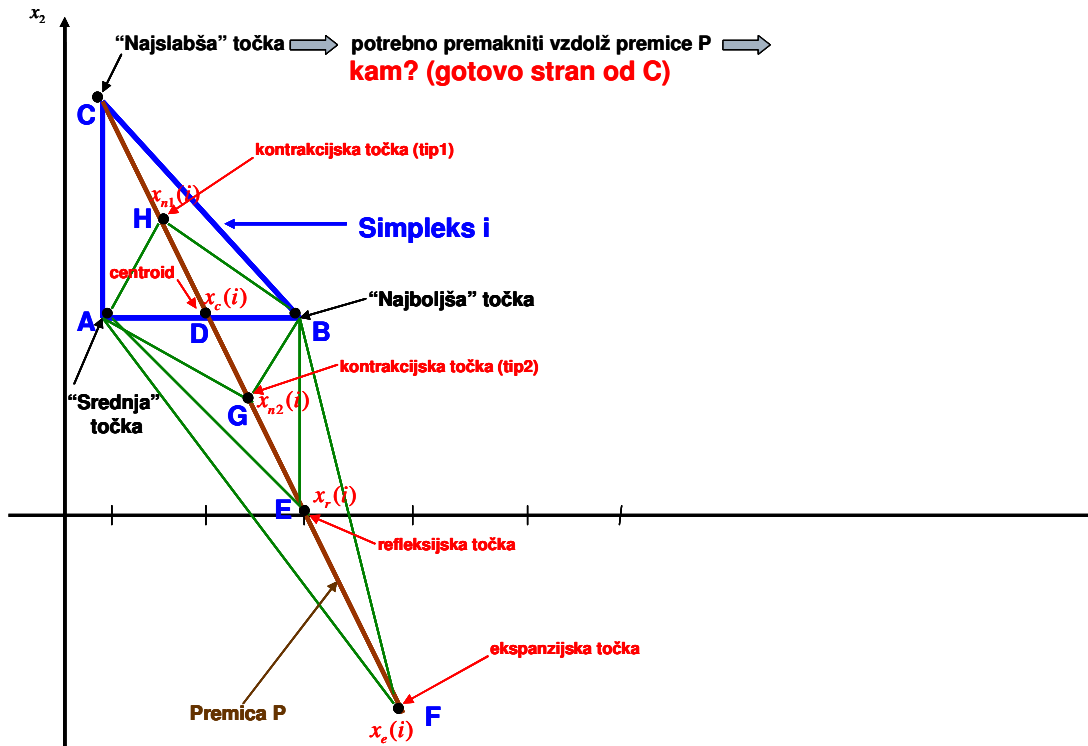
Strategijo premika »najslabše« točke simpleksa  $i$  v dvodimenzionalnem prostoru ( $i=2$ ) nam nazorno prikazuje slika 61 [4]. Glede na Nelder-Meadov algoritem najslabšo točko  $x_h$  premaknemo v eno izmed točk  $x_{n1}$ ,  $x_{n2}$ ,  $x_r$  ali  $x_e$ . Pri tem simpleks  $i$  zavzame novo obliko enega izmed trikotnikov, ki nastane, ko povežemo novonastalo točko (H, G, E ali F) s točkama starega simpleksa A in B.

Če pa ne dosežemo s premikom najslabše točke nobene izboljšave, leži verjetno optimalna točka znotraj starega simpleksa. Zato le tega skrčimo okrog »najboljše točke«  $x_l$  (točka B), kot je prikazano na sliki 62 [13].

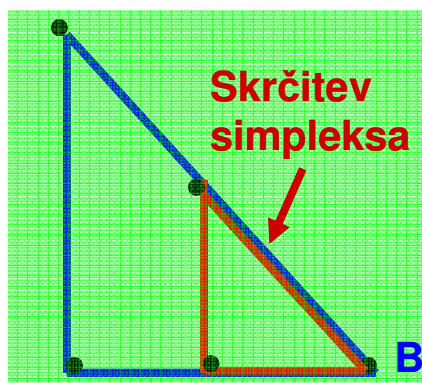
Slika 63 pa prikazuje ilustracijo principa delovanja simpleksne metode še pri konkretnih numeričnih vrednostih [4]. Začetni simpleks 0 (ogljjišča A, B, C) se v 1. iteraciji spremeni v simpleks 1 (ogljjišča A, B, F), saj se izkaže, da je najbolj primerno premakniti najslabšo točko C s pomočjo ekspanzije v točko F. Nato se za simpleks 1 ugotovi najslabša točka (točka A), ki jo je zopet potrebno nekam premakniti. Kot se izkaže, je ekspanzija ponovno najbolj primerna operacija, pri čemer se simpleks 1 spremeni v simpleks 2.



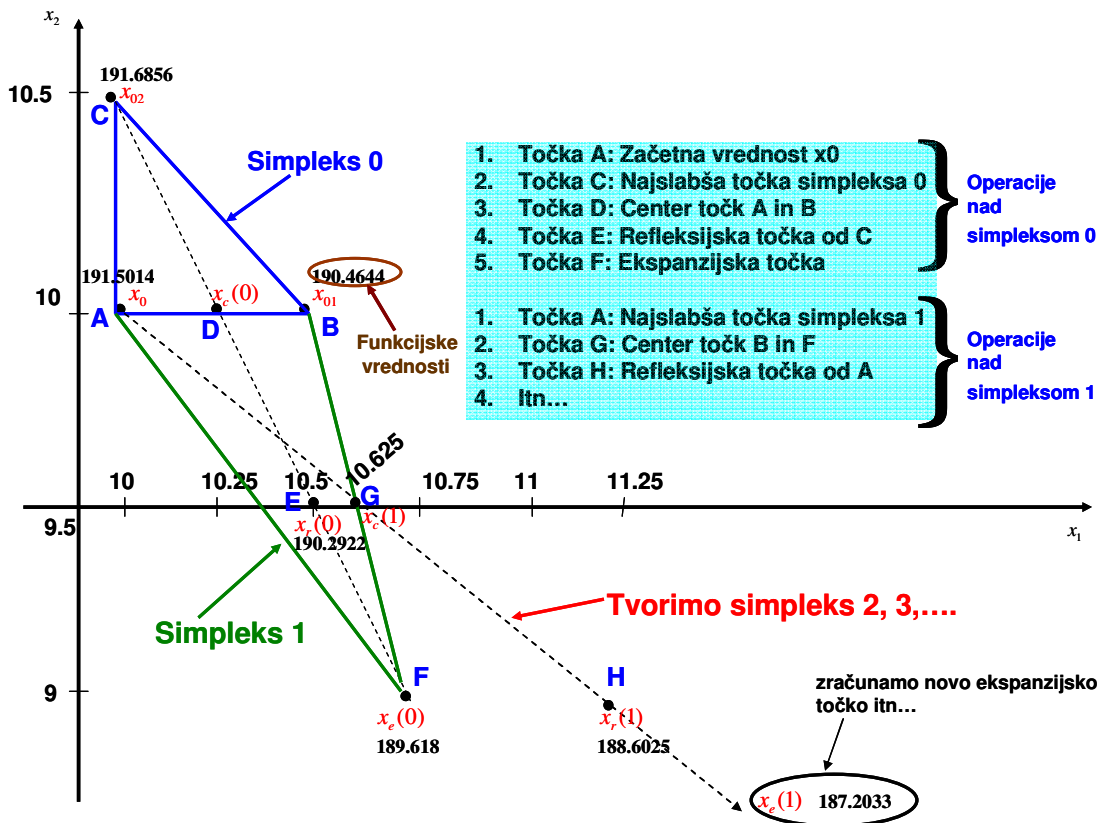
slednjega dobimo, če povežemo točke B, F in  $x_e(1)$ . Postopek se nato z opisanim principom nadaljuje, pri čemer se generirajo simpleksi 3, 4, itn,..., dokler se ne doseže optimalna rešitev.



Slika 61: Prikaz strategije premika najslabše točke simpleksa v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije

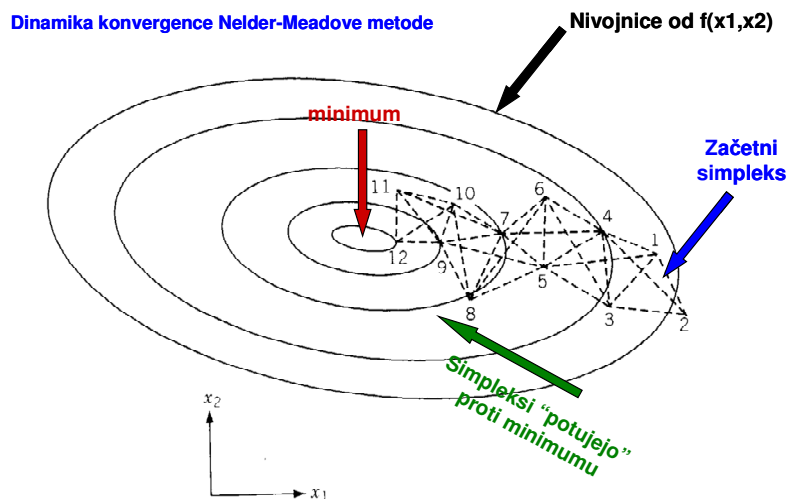


Slika 62: Prikaz postopka skrčitve simpleksa okrog »najboljše točke« v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije



Slika 63: Ilustracija principa delovanja simpleksne metode v dvodimenzionalnem prostoru  $x_1, x_2, f(x_1, x_2)$ , kjer iščemo minimum funkcije

Slika 64 prikazuje dinamiko konvergence Nelder-Meadove metode, ki se odvija od začetnega simpleksa, pa preko vmesnih simpleksov vse do iskanega minimuma ciljne funkcije.



Slika 64: Primer dinamike konvergence Nelder-Meadove metode

Poglejmo si še simboličen potek Nelder-Meadovega algoritma, ki je prikazan na sliki 65 [4].

```

// Vhodne veličine: x0, Kr, Ke, Kn1, Kn2
Na osnovi x0 izračunaj n+1 točk začetnega simpleksa x[i], i = 1..n+1;
ponavljaj
{ sortiraj po velikosti funkcijske vrednosti v točkah trenutnega simpleksa x[i], i = 1..n+1
  sortiraj točke trenutnega simpleksa od najboljše x[l] do najslabše x[h]
  // seveda velja: F(x[h]) = max, F(x[l]) = min;
  določi centroid xc;
  xr = Refleksija();
  če F(xr)<F(x[l]) // če refleksijska točka boljša od najboljše
  { xe = Ekspanzija();
    če F(xe)<F(x[l]) // če ekspanzijska točka boljša od najboljše
    x[h] = xe; // najslabšo točko nadomesti z ekspanzijsko
    sicer
    x[h] = xr; // najslabšo točko nadomesti z refleksijsko
  }
  sicer // sicer je refleksijska točka slabša od najboljše
  { če F(xr)>F(x[j]) za vsak j=1..n+1, j != h // če refleksijska točka slabša od vseh ostalih brez najslabše
    { če F(xr)<F(x[h]) // če je refleksijska točka boljša od najslabše
      x[h] = xr; // najslabšo točko nadomesti z refleksijsko
      xn = Kontrakcija();
      če F(xn)<F(x[h]) // če kontrakcijska boljša od najslabše
      x[h] = xn; // najslabšo točko nadomesti z kontrakcijsko
      sicer // kontrakcijska slabša od najslabše (ni napredka)
      pomakni vse točke proti x[l]; // skrči vse točke simpleksa proti najboljši
    }
    sicer // refleksijska točka boljša od katere izmed ostalih brez najslabše
    x[h] = xr; // najslabšo točko nadomesti z refleksijsko
  }
}
} dokler ni izpolnjen pogoj za zaustavitev;

```

Slika 65: Simboličen potek Nelder-Meadovega algoritma

Seveda se algoritem na sliki 65 obnaša v skladu s principi, ki smo jih razložili v tem poglavju.

V nadaljevanju si še na kratko pogledimo tri primere, ki nam bodo pomagali dodatno osvetliti delovanje Nelder-Meadove metode.

### 1. primer:

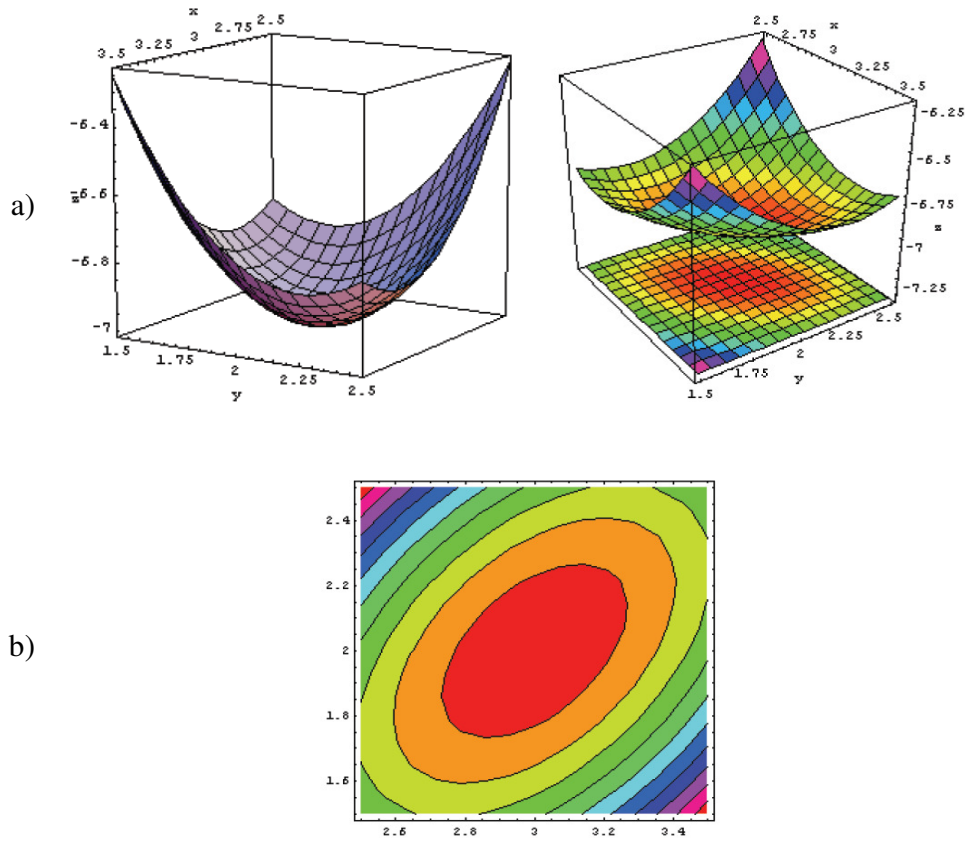
Z uporabo Nelder-Meadove metode poiščite minimum naslednje funkcije:

$$z = f(x, y) = x^2 - 4 \cdot x + y^2 - y - x \cdot y \quad (6.64.)$$

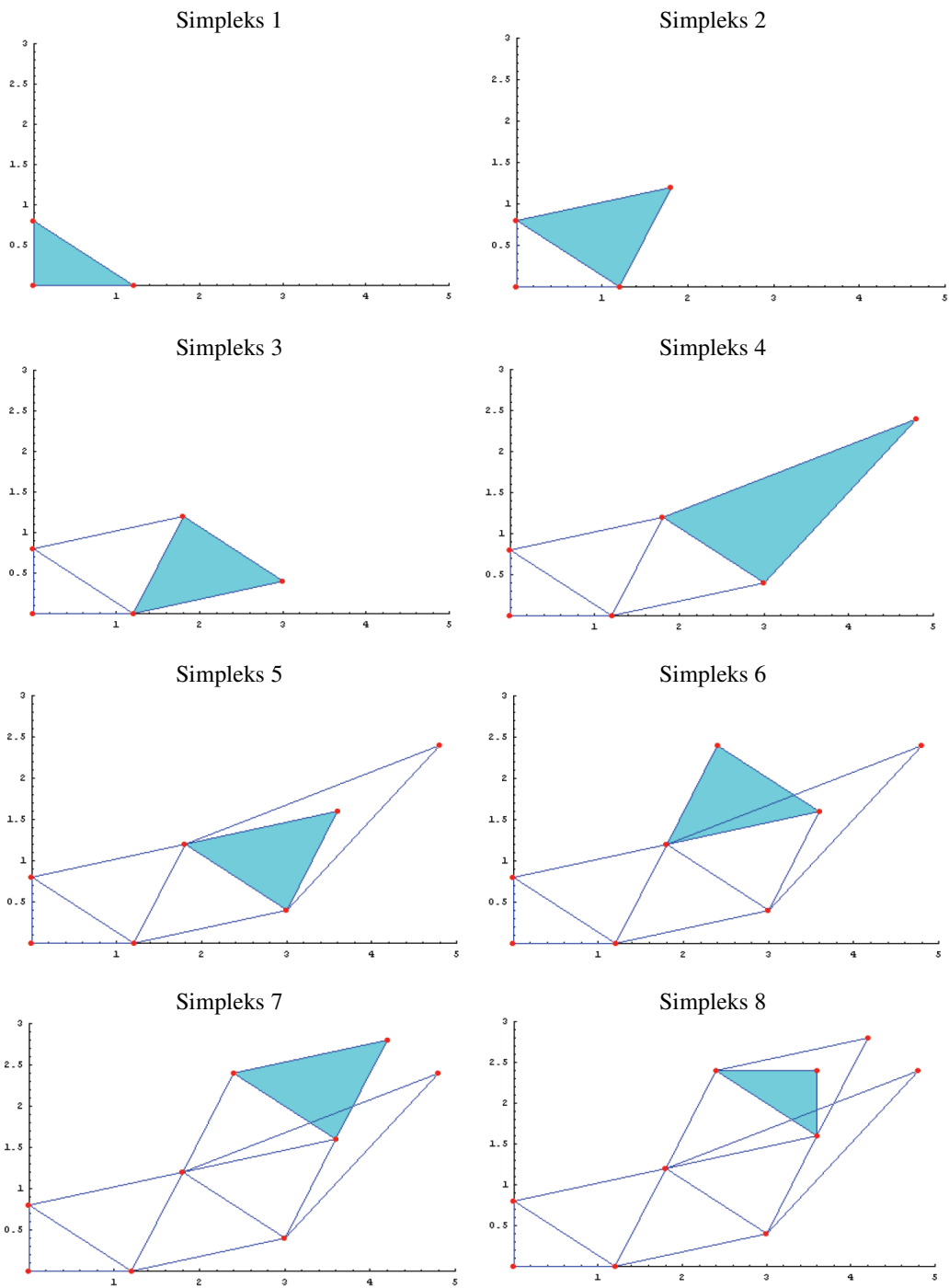
Slika 66 prikazuje funkcijo (6.64.) iz dveh različnih zornih kotov, ter njen minimum, ki ima koordinate  $(x^*, y^*) = (3, 2)$  [4].

Sliki 67 in 68 pa prikazujeta dinamiko tvorjenja sekvence prvih oz. drugih osmih simpleksov, ki so potrebni, da dosežemo minimum funkcije. Kot je razvidno iz simpleksa

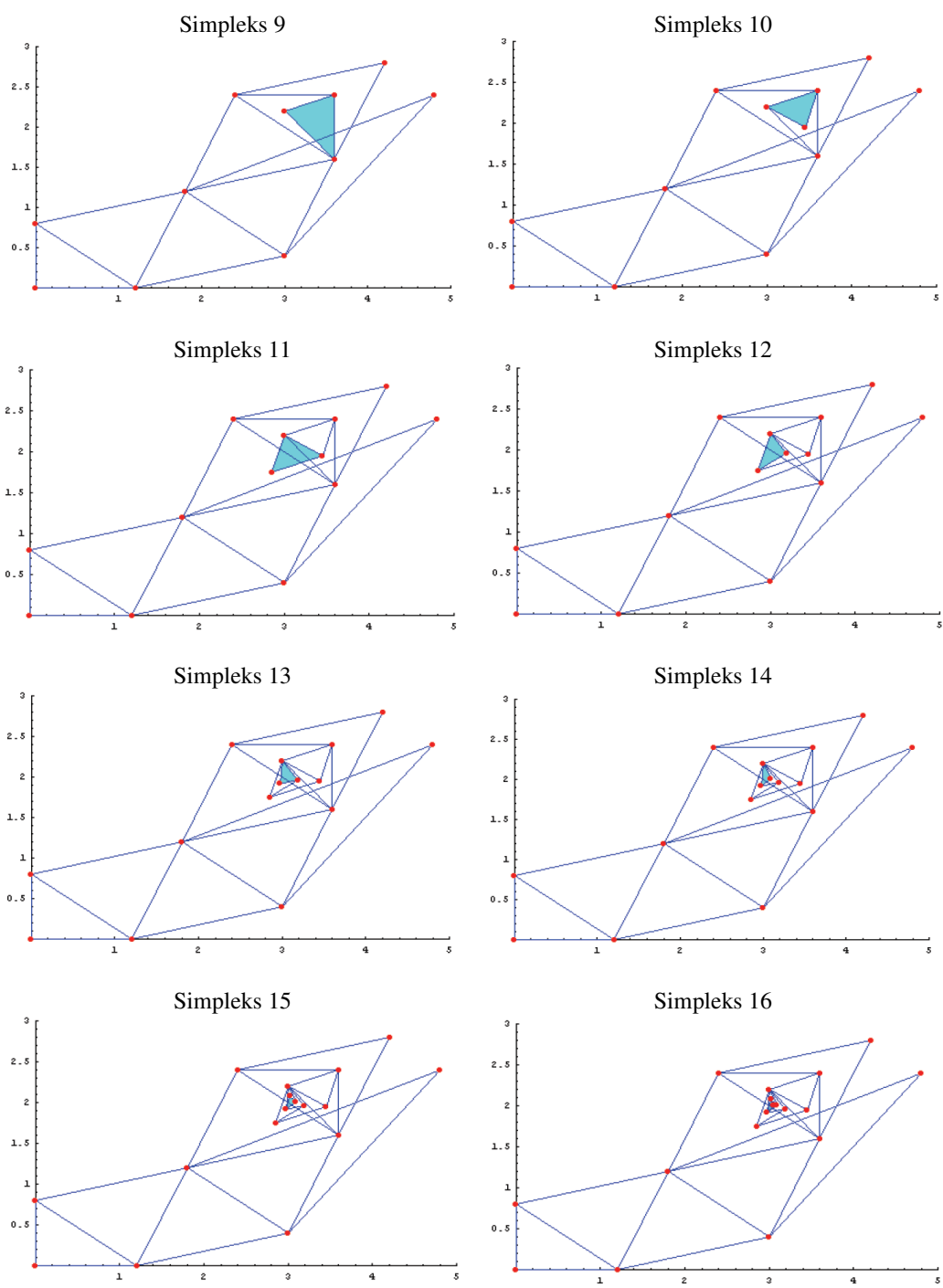
16 na slici 68, metoda dejansko konvergira k pravi vrednosti minimuma funkcije  $f(x, y)$ , to je  $(x^*, y^*) = (3, 2)$ .



Slika 66: a) Prikaz dvodimenzionalne funkcije (6.64.) iz dveh različnih zornih kotov;  
 b) Prikaz minimuma  $(x^*, y^*) = (3, 2)$  funkcije (6.64.) s pomočjo njenih nivojnic.



Slika 67: Potek dinamike tvorjenja sekvence prvih osmih simpleksov za primer funkcije (6.64.)



Slika 68: Potek dinamike tvorjenja sekvence drugih osmih simpleksov za primer funkcije (6.64.)

## 2. primer:

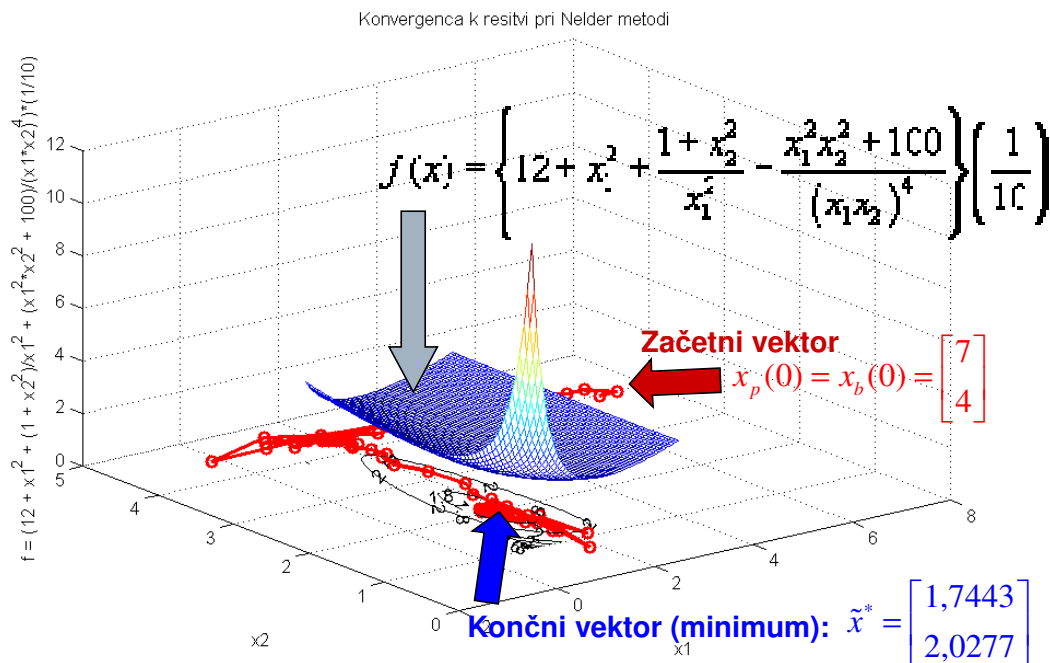
S pomočjo Nelder-Meadove metode določite optimalno točko za naslednjo funkcijo dveh odločitvenih spremenljivk  $x_1, x_2$ :

$$f(x) = \left\{ 12 + x_1^2 + \frac{1 + x_2^2}{x_1^2} + \frac{x_1^2 x_2^2 + 100}{(x_1 x_2)^4} \right\} \begin{pmatrix} 1 \\ 10 \end{pmatrix} \quad (6.65.)$$

kjer je začetni vektor odločitvenih spremenljivk enak:  $x(0) = x_0 = [x_1(0) \quad x_2(0)] = [7 \quad 4]^T$ .

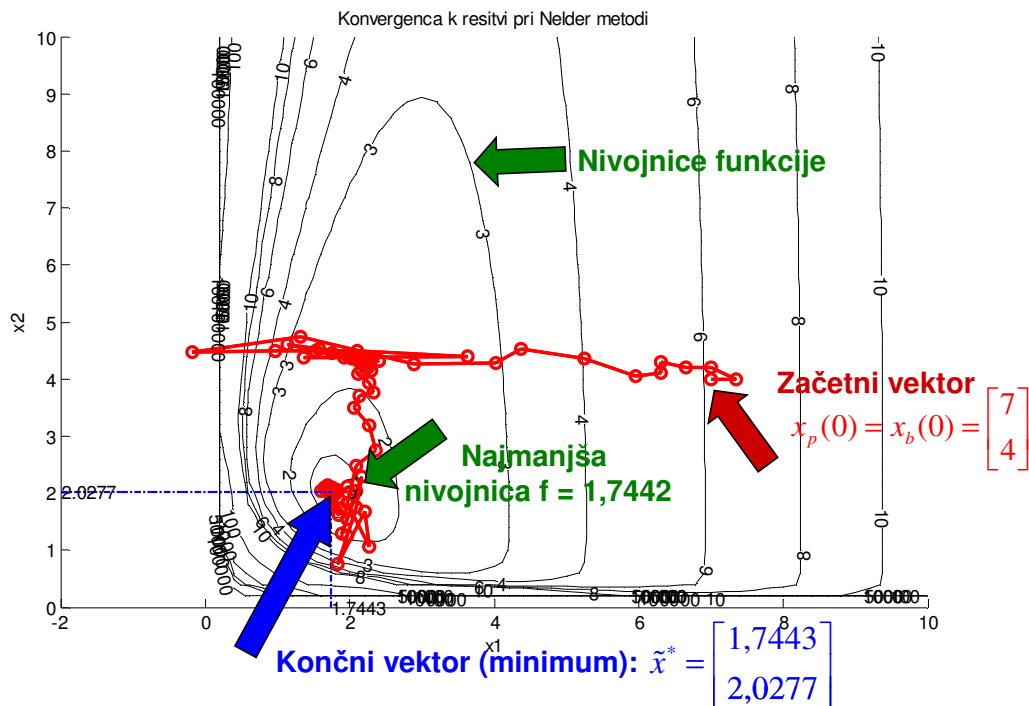
### Rešitev:

Pri iskanju optimalnih rezultatov si seveda pomagamo z Nelder-Meadovim algoritmom, ki je bil prej opisan v tem poglavju. Pri tem dobimo dinamiko konvergence rezultatov, prikazano na sliki 69.



Slika 69: Dinamika konvergence rezultatov pri iskanju optimuma funkcije (6.65.) s pomočjo Nelder-Meadove metode, opazovana v prostoru  $(x_1, x_2, f(x_1, x_2))$  (slika A)

Dinamiko konvergence rezultatov, prikazano na sliki 69, si lahko še natančneje ogledamo na sliki 70, kjer so podane tudi nivojnice funkcije (6.65.), preslikane na  $x_1 - x_2$  ravnino.



Slika 70: Dinamika konvergence rezultatov pri iskanju optimuma funkcije (6.65.) s pomočjo Nelder-Meadove metode, opazovana na  $x_1 - x_2$  ravnini. (slika B)

Iz slik 69 in 70 je razvidno, kako v posameznih iteracijah vrednosti aproksimacij optimalnih odločitvenih spremenljivk konsistentno napredujejo od začetne točke  $(x_{p1}(0), x_{p2}(0)) = (7, 4)$  do končne (optimalne) točke  $(x_1^*, x_2^*) = (1.7443, 2.0277)$ .

Tovrstno napredovanje (konvergiranje k optimalni rešitvi) je na slikah 69, 70 označeno s povezanimi krogi (o), označene pa so tudi nivojnice ciljne funkcije (10, 8, 6, 4, 3, 2, ...). Te pri konvergiranju postopka postopoma padajo od začetne vrednosti  $f(7, 4)$  do končne (minimalne) vrednosti 1.7442.

Na sliki 69 spodaj je tudi razvidno, da funkcija  $f(x_1, x_2)$  dejansko doseže najnižjo točko (minimum) v prostoru  $(x_1, x_2, f(x_1, x_2))$ , ko postopek konvergira k optimalni rešitvi  $(x_1^*, x_2^*) = (1.7443, 2.0277)$ .



Končni optimalni rezultati torej so:

$$\tilde{x}^* = \begin{bmatrix} \tilde{x}_1^* & \tilde{x}_2^* \end{bmatrix}^T = \begin{bmatrix} 1.7443 & 2.0277 \end{bmatrix}^T \quad (6.66.)$$

$$f(\tilde{x}^*) = \min f(x) = 1.7442$$

in so seveda zelo podobni rezultatom, ki jih je dosegla Hooke-Jeevesova metoda (primerjaj z izrazom (6.53.)).

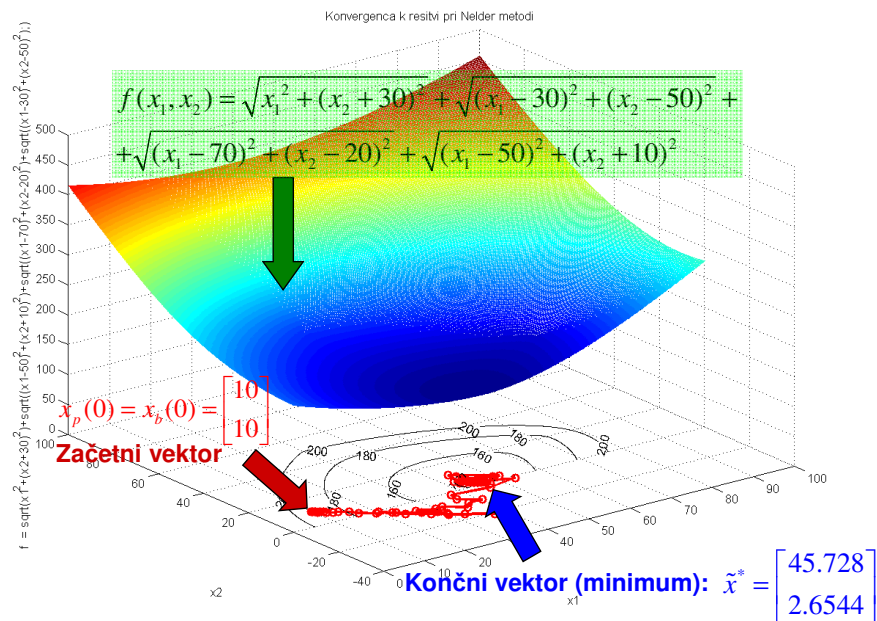
Kot je razvidno iz slik 69 in 70, metoda s precejšnjimi nihanji konvergira k pravi rešitvi. V primerjavi s Hooke-Jeevesovo metodo vzorčnih premikov je tudi precej počasnejša, saj potrebuje za konvergenco k rešitvi 95 iteracij, Hooke-Jeevesova metoda pa je za konvergenco potrebovala znatno manj, to je 45 iteracij (glej slike 51 in 52).

### **3. Primer:**

Gre za enak primer, kot smo ga obravnavali že pri Hooke-Jeevesovi metodi (glej sliko 53 in izraz (6.54.)), pri čemer zopet vzamemo začetni vektor odločitvenih spremenljivk  $(x_1(0), x_2(0)) = (10, 10)$ .

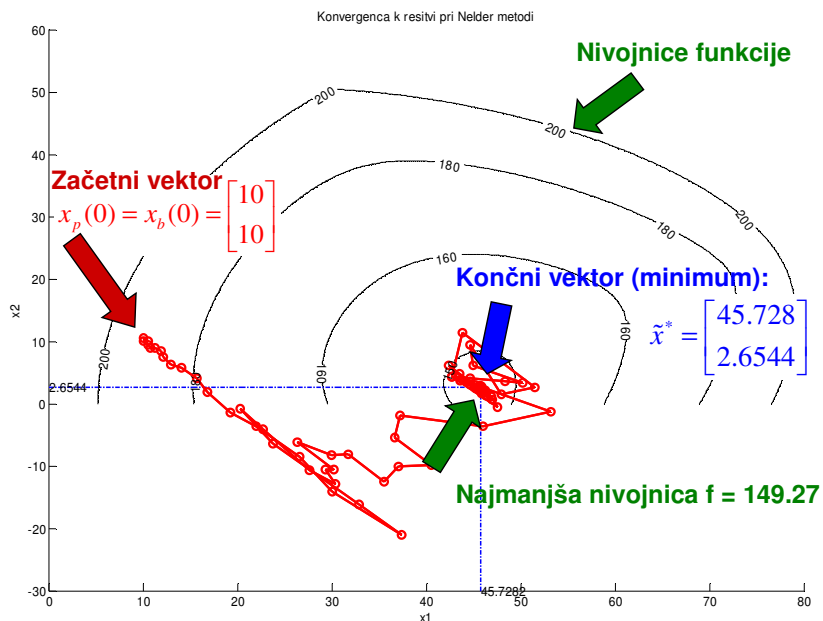
Pri iskanju optimalnih rezultatov si seveda zopet pomagamo z Nelder-Meadovim algoritmom, ki je bil prej opisan v tem poglavju. Pri tem dobimo dinamiko konvergence rezultatov, prikazano na sliki 71.

.



Slika 71: Dinamika konvergence rezultatov pri iskanju optimuma funkcije (6.54.) s pomočjo Nelder-Meadove metode, opazovana v prostoru  $(x_1, x_2, f(x_1, x_2))$  (slika A)

Dinamiko konvergence rezultatov, prikazano na sliki 71, si lahko še natančneje ogledamo na sliki 72, kjer so podane tudi nivojnice funkcije (6.54.), preslikane na  $x_1 - x_2$  ravnino.



Slika 72: Dinamika konvergence rezultatov pri iskanju optimuma funkcije (6.54.) s

pomočjo Nelder-Meadove metode, opazovana na  $x_1 - x_2$  ravnini. (slika B)

Iz slik 71 in 72 je razvidno, kako v posameznih iteracijah vrednosti aproksimacij optimalnih odločitvenih spremenljivk konsistentno napredujejo od začetne točke  $(x_1(0), x_2(0)) = (10, 10)$  do končne (optimalne) točke  $(x_1^*, x_2^*) = (45.728, 2.6544)$ .

Tovrstno napredovanje (konvergiranje k optimalni rešitvi) je na slikah 71, 72 označeno s povezanimi krogi (o), označene pa so tudi nivojnice ciljne funkcije (z resolucijo 20: 200, 180, 160, ...). Te pri konvergiranju postopka postopoma padajo od začetne vrednosti  $f(x_1(0), x_2(0)) = f(10, 10) = 191.5$  do končne (minimalne) vrednosti 149.27.

Na sliki 71 spodaj je tudi razvidno, da funkcija  $f(x_1, x_2)$  dejansko doseže najnižjo točko (minimum) v prostoru  $(x_1, x_2, f(x_1, x_2))$ , ko postopek konvergira k optimalni rešitvi  $(x_1^*, x_2^*) = (45.728, 2.6544)$ .

Končni optimalni rezultati torej so:

$$\tilde{x}^* = \begin{bmatrix} \tilde{x}_1^* & \tilde{x}_2^* \end{bmatrix}^T = \begin{bmatrix} 45.728 & 2.6544 \end{bmatrix}^T \quad (6.67.)$$

$$f(\tilde{x}^*) = \min f(x) = 149.27$$

in so seveda zelo podobni rezultatom, ki jih je dosegla Hooke-Jeevesova metoda (primerjaj z izrazom (6.55.)).

Tudi v tem primeru je razvidno iz slik 71 in 72, da metoda s precejšnjimi nihanji konvergira k pravi rešitvi. To je tudi razlog, da ta metoda ponovno v primerjavi s Hooke-Jeevesovo metodo precej počasneje konvergira k pravi rešitvi.

### 6.3.1.3 Cauchyeva metoda najstrmejšega spusta (gradientna metoda I. reda)

Cauchyovo metodo najstrmejšega spusta običajno imenujemo kar Metoda najstrmejšega spusta. Uvrščamo jo med tipične gradientne metode prvega reda, saj moramo poleg računanja vrednosti ciljne funkcije izračunati tudi njen prvi odvod, kar pomeni, da je uporabna samo za odvedljive funkcije. Metoda ima zelo zanesljivo konvergenco, tudi v primerih, ko je začetna aproksimacija precej oddaljena, po drugi strani pa ima dokaj počasno konvergenco v bližini optimalne točke [13].

Metoda najstrmejšega spusta je temeljna metoda in zato silno pomembna, saj številne druge, bolj sofisticirane metode tega tipa, temeljijo prav na njej ter predstavljajo zgolj določeno modifikacijo te metode, z namenom izboljšave konvergenčnih lastnosti.

V nadaljevanju bomo izpeljali to metodo, pri čemer bomo predpostavili, da iščemo minimum funkcije (podobne ugotovitve seveda veljajo tudi za maksimum).

Pri izpeljavi metode bomo morali izpeljati tudi vektor optimalnih sprememb smeri po vseh dimenzijah  $d_k$  [12], kjer je  $k$  indeks iteracije. V ta namen lahko najprej izrazimo funkcijo vektorske spremenljivke v obliki Taylorjeve vrste, pri čemer zanemarimo višje člene od drugega naprej. Dobimo naslednji izraz [13, 14, 17, 21]:

$$f(x_{k+1}) = f(x_k + \Delta x_k) \approx f(x_k) + [\nabla f(x_k)]^T \cdot \Delta x_k + \frac{1}{2} \Delta x_k^T \cdot \nabla^2 f(x_k) \cdot \Delta x_k \quad (6.68.)$$

kjer je  $\nabla f(x_k)$  gradient (vektor 1. odvodov funkcije),  $\nabla^2 f(x_k)$  je Hessova matrika 2. odvodov funkcije,  $x_{k+1}$  je vektor odločitvenih spremenljivk v naslednji iteraciji in  $x_k$  vektor odločitvenih spremenljivk v trenutni iteraciji postopka.  $\Delta x_k$  je vektor spremembe odločitvenih spremenljivk in ga določimo kot produkt velikosti  $\alpha_k$  in smeri  $d_k$ :

$$\Delta x_k = \alpha_k \cdot d_k \quad (6.69.)$$

kjer  $d_k = [d_{k1} \ d_{k2} \ \dots \ d_{kn}]^T$  pomeni vektor optimalnih sprememb smeri po vseh dimenzijah ( $n$  jih je) in predstavlja smer najhitrejšega padanja funkcije  $f$ .  $\alpha_k$  pa predstavlja dolžino koraka sprememb smeri in določa velikost najstrmejšega padca funkcije  $f$ , ter ga v vsaki iteraciji določimo z enodimenzionalno optimizacijo (glej poglavje 6.2.).

Zaradi nadaljnje čim večje poenostavitve računanja zanemarimo še drugi člen pri izrazu (6.68.) in iz njega dobimo izraz [17]:

$$f(x_{k+1}) \approx f(x_k) + [\nabla f(x_k)]^T \cdot \Delta x_k \quad (6.70.)$$

Če v izrazu (6.70.) upoštevamo izraz (6.69.), dobimo [13, 17]:

$$f(x_{k+1}) \approx f(x_k) + [\nabla f(x_k)]^T \cdot \underbrace{\alpha_k \cdot d_k}_{\Delta x_k} = f(x_k) + \alpha_k [\nabla f(x_k)]^T d_k \quad (6.71.)$$

Za izraz (6.71.) nato postavimo pogoj, da se mora funkcija v vsaki naslednji iteraciji  $k + 1$  vedno zmanjšati, pri čemer dobimo:

$$f(x_{k+1}) \approx f(x_k) + \alpha_k [\nabla f(x_k)]^T d_k < f(x_k) \quad (6.72.)$$

Iz tega sledi izraz:

$$\alpha_k [\nabla f(x_k)]^T d_k < 0 \quad (6.73.)$$

ki zaradi pozitivne dolžine koraka  $\alpha_k$  preide v obliko:

$$[\nabla f(x_k)]^T d_k < 0 \quad (6.74.)$$

Za izraz (6.74.) lahko sedaj uporabimo takoimenovano *Schwartzovo neenakost* [13, 14, 17, 21], ki pri  $\|d_k\| \leq 1$  preide v obliko:

$$-\|\nabla f(x_k)\| \leq -\|\nabla f(x_k)\| \cdot \|d_k\| \leq [\nabla f(x_k)]^T \cdot d_k < 0 \quad (6.75.)$$

Enakost levega dela tega izraza velja (v levem neenačaju), če vzamemo [17]:

$$d_k = -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|} \quad (6.76.)$$

pri čemer se da dokazati, da bo pri veljavnosti izraza (6.76.) funkcija  $f$  doživela najstrmejši spust (padec).

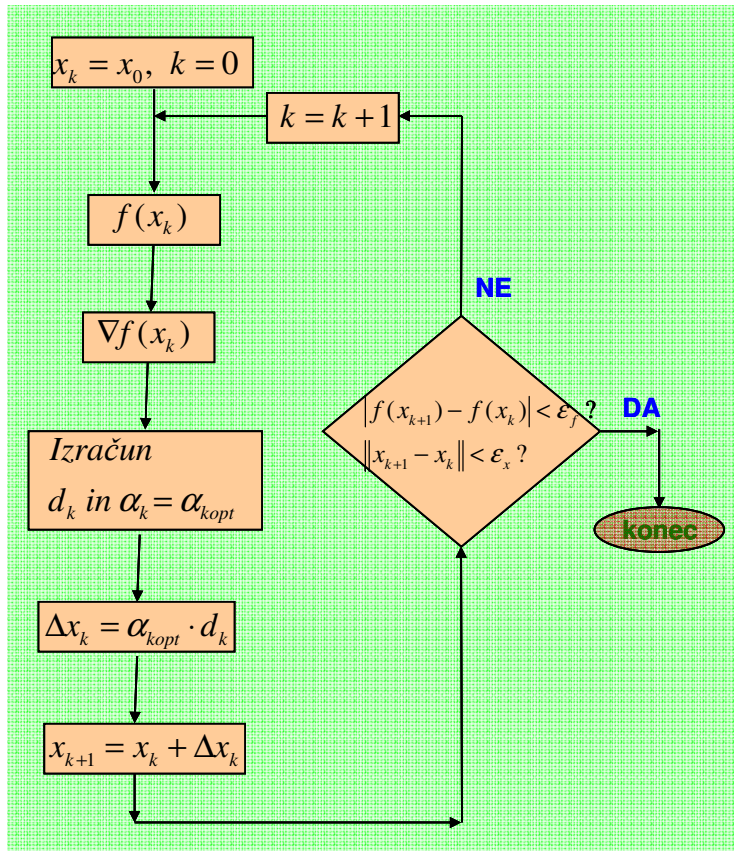
Ker vemo, da je  $x_{k+1} = x_k + \Delta x_k$ , lahko torej za vektor odločitvenih spremenljivk v naslednji iteraciji z upoštevanjem izrazov (6.69.) in (6.76.) zapišemo naslednjo iteracijsko enačbo metode najstrmejšega spusta [13, 17, 21]:

$$x_{k+1} = x_k + \Delta x_k = x_k + \alpha_k \cdot d_k = x_k - \alpha_k \cdot \frac{\nabla f(x_k)}{\|\nabla f(x_k)\|} \quad (6.77.)$$

Potemtakem vektor odločitvenih spremenljivk v naslednji iteraciji dobimo iz vektorja odločitvenih spremenljivk prejšnje iteracije  $x_k$ , od katerega še odštejemo vektor

$$\alpha_k \cdot \frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}.$$

Diagram poteka Cauchyve metode najstrmejšega spusta prikazuje slika 73 [13, 17, 21].

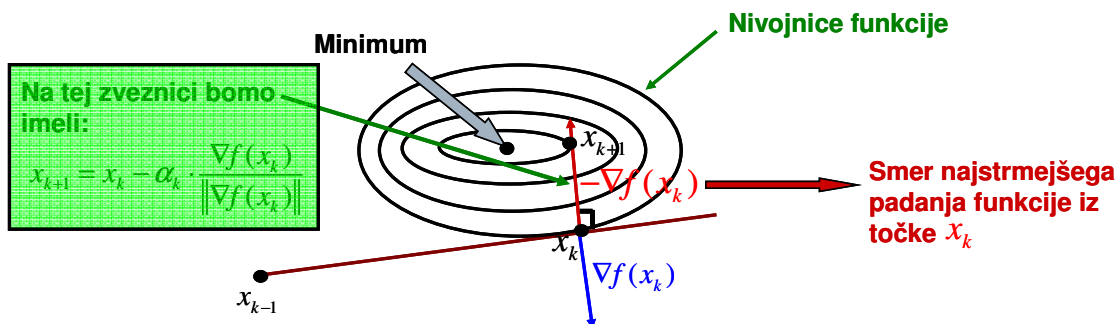


Slika 73: Diagram poteka Cauchyve metode najstrmejšega spusta

V vsaki iteraciji izračunamo gradient funkcije, ki ga, da se čim bolj stabilizirano bližamo k optimumu, še dodatno normiramo (glej izraz (6.76.)), nato pa vzdolž optimalne smeri  $d_k$  izračunamo optimalno velikost dolžine koraka  $\alpha_k$  s pomočjo katere od enodimenzionalnih optimizacijskih metod [13, 17]. Dobljena rezultata nato zmnožimo in odštejemo od  $x_k$  ter se na ta način »najstrmeje spuščamo« proti optimalni vrednosti vektorja  $x_{k+1}$ . Ta vektor nato uporabimo za izračun funkcije in njenega gradienta v naslednji iteraciji ter nato opravimo še vse nadaljnje izračune, potrebne v tej iteraciji. Postopek ponavljamo toliko časa, dokler niso izpolnjeni konvergenčni pogoji, torej tisti, pri katerih se vrednost funkcije oz. vektorja odločitvenih spremenljivk ne spreminja več bolj od predpisane tolerance.

Za pozitivni gradient funkcije  $+\nabla f$  velja, da vedno kaže v smeri, kjer funkcija najhitreje narašča. Potemtakem lahko sklepamo, da negativni gradient funkcije  $-\nabla f$  vedno kaže v smeri, kjer funkcija najhitreje pada. Ker pa je gradient  $\nabla f$  še dodatno normiran (glej izraza (6.76.) in (6.77.)), je potemtakem tudi stabilnost metode večja.

Geometrijsko interpretacijo Cauchyve metode najstrmejšega spusta za iteracijo  $k+1$  nam prikazuje slika 74 [4].



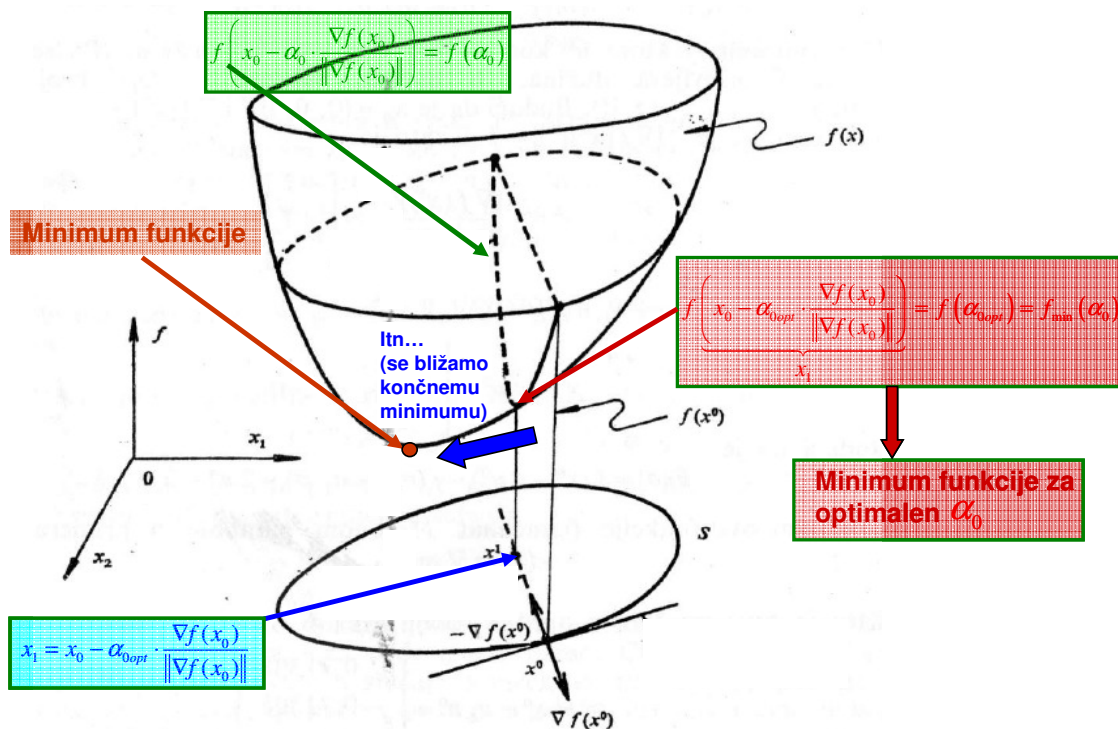
Slika 74: Geometrijska interpretacija Cauchyve metode najstrmejšega spusta za iteracijo  $k+1$

Na sliki 74 je s pomočjo nivojnic funkcije geometrijsko prikazano delovanje Cauchyve metode najstrmejšega spusta, pri čemer središčna točka (najmanjša nivojnica) predstavlja minimum funkcije. Smer najstrmejšega padanja funkcije iz točke  $x_k$  predstavlja negativni gradient  $-\nabla f(x_k)$ .

To pomeni, da metoda izračuna takšen vektor odločitvenih spremenljivk v naslednji iteraciji  $x_{k+1}$ , ki se bo premaknil najbližje (najhitreje) dejanskemu minimumu, kar pa je seveda pravokoten premik v smeri  $-\nabla f(x_k)$  glede na veznico  $(x_{k-1}, x_k)$ .

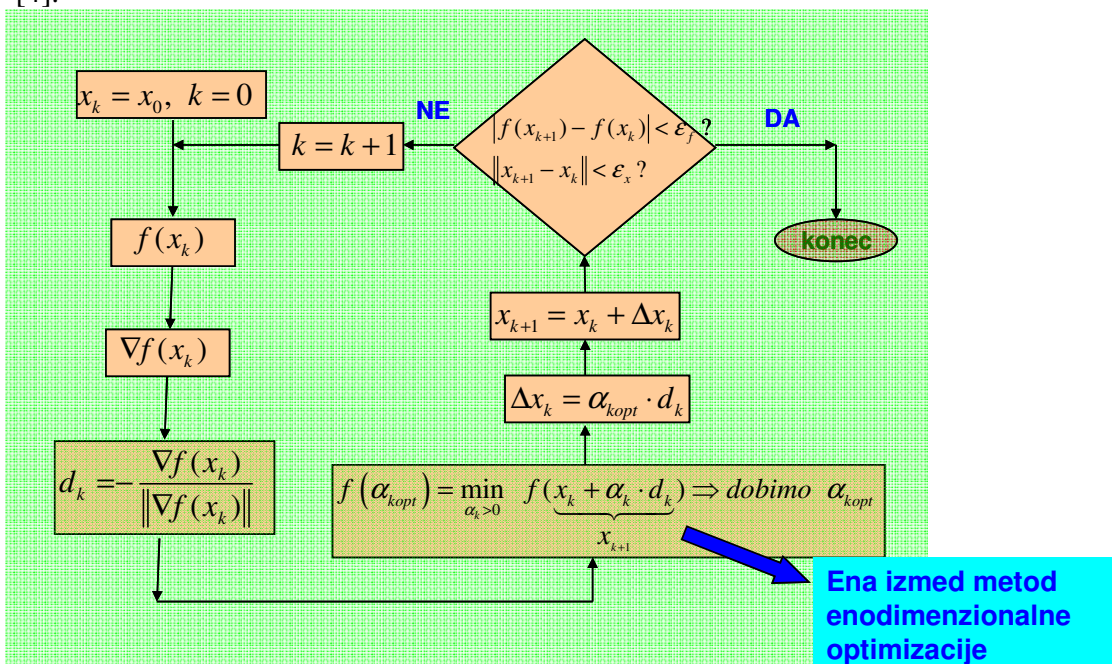
Na sliki 75 si lahko ogledamo še geometrijsko interpretacijo Cauchyve metode najstrmejšega spusta, gledano iz prostorske perspektive [17].

$x_0$  je začetni vektor odločitvenih spremenljivk,  $x_1$  pa vektor odločitvenih spremenljivk v 1. iteraciji. Slednjega določimo tako, da se iz  $x_0$  pomaknemo v smeri negativnega gradienta  $-\nabla f(x_0)$ , kar je gotovo optimalna smer premika. Optimalno velikost premika pa dobimo tako, da s katero od enodimensionalnih metod optimizacije določimo takšen  $\alpha_0$ , ki bo minimiziral  $f(x_0 + \alpha_0 \cdot (-\nabla f(x_0)))$ . Postopek nato ponavljamo s premiki v  $x_2, x_3, \dots$ , toliko časa, dokler ni dosežen minimum funkcije  $f(x_n)$  v  $n$ -ti iteraciji. Seveda se pri premikih proti optimumu tudi korak  $\alpha_k$  nenehno spreminja, torej zavzame vrednosti  $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_{n-1}$ , spreminja pa se tudi gradient  $-\nabla f(x_k)$ ,  $k = 0, 1, 2, \dots, n-1$ .



Slika 75: Geometrijska interpretacija Cauchyve metode najstrmejšega spusta za iteracijo  $k+1$

Na sliki 76 je prikazan celoten algoritem Cauchyve metode najstrmejšega spusta, ki ga dobimo iz slike 73 tako, da vnesemo še nekatere dodatne detajle glede izračunov metode [4].



Slika 76: Celoten algoritem Cauchyve metode najstrmejšega spusta





Slika 77: a) 1. primer dinamike konvergence za konveksno funkcijo; b) 2. primer dinamike konvergence konveksne funkcije; c) Primer dinamike konvergence za nekonveksno funkcijo

Osvetlino delovanje metode najstrmejšega spusta še na naslednjem primeru.

**Primer:**

Gre za enak primer, kot smo ga obravnavali že pri Hooke-Jeevesovi metodi (glej 1. primer in izraz (6.49.)), pri čemer zopet vzamemo začetni vektor odločitvenih spremenljivk  $x(0) = x_0 = [x_1(0) \ x_2(0) \ x_3(0)] = [0 \ 0 \ 0]^T$ .

Pri iskanju optimalnih rezultatov si seveda pomagamo z algoritmom najstrmejšega spusta, ki je bil opisan v tem poglavju.

**Rešitev:**

Najprej zgolj informativno preverimo, kakšna je vrednost funkcije (6.49.) v začetni točki:

$$f(0,0,0) = 0^4 + 0^3 - 0 + 0^4 - 0^2 + 0 + 0^2 - 0 + 0 \cdot 0 \cdot 0 = 0 \quad (6.78.)$$

Iteracijska enačba (6.77.) bo v 1. iteraciji zavzela vrednost:

$$x_1 = x_0 + \Delta x_0 = x_0 + \alpha_0 \cdot d_0 = x_0 - \alpha_0 \cdot \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} \quad (6.79.)$$

V enačbi (6.79.) očitno potrebujemo tudi gradient v začetni točki. Vendar, da bi slednjega lahko izračunali, moramo najprej izračunati gradient v splošni točki, ki se glasi:

$$\begin{aligned} \nabla f(x) &= \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \quad \frac{\partial f(x)}{\partial x_3} \right]^T = \\ &= \left[ 4x_1^3 + 3x_1^2 - 1 + x_2x_3 \quad 4x_2^3 - 2x_2 + 1 + x_1x_3 \quad 2x_3 - 1 + x_1x_2 \right] \end{aligned} \quad (6.80.)$$

Če v gradient (6.80.) vstavimo začetno točko, dobimo:

$$\begin{aligned} \nabla f(x_0) &= \left[ 4 \cdot 0^3 + 3 \cdot 0^2 - 1 + 0 \cdot 0 \quad 4 \cdot 0^3 - 2 \cdot 0 + 1 + 0 \cdot 0 \quad 2 \cdot 0 - 1 + 0 \cdot 0 \right]^T = \\ &= \left[ -1 \quad 1 \quad -1 \right]^T \end{aligned} \quad (6.81.)$$

Chebisevljeva norma [17] tega gradienta je enaka:

$$\|\nabla f(x_0)\| = \|[ -1 \ 1 \ -1 ]\| = \max\{|-1| \ |1| \ |-1|\} = 1 \quad (6.82.)$$

pri čemer bi lahko za izračun norme vektorja uporabili tudi kakšen drug tip norme [17].

Na osnovi izrazov (6.81.) in (6.82.) lahko izračunamo smerni vektor  $d_0$ :

$$d_0 = -\frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} = -\frac{[-1 \ 1 \ -1]^T}{1} = -[-1 \ 1 \ -1]^T = [1 \ -1 \ 1]^T \quad (6.83.)$$

Če vstavimo izraz (6.83.) v enačbo (6.79.), dobimo pri upoštevanju začetne točke  $x(0) = x_0 = [x_1(0) \ x_2(0) \ x_3(0)] = [0 \ 0 \ 0]^T$  naslednji izraz:

$$x_1 = x_0 + \alpha_0 \cdot d_0 = [0 \ 0 \ 0]^T + \alpha_0 \cdot [1 \ -1 \ 1]^T = [\alpha_0 \ -\alpha_0 \ \alpha_0]^T \quad (6.84.)$$

Vstavimo sedaj točko (6.84.) v funkcijo (6.49.), pri čemer dobimo:

$$\begin{aligned} f([\alpha_0 \ -\alpha_0 \ \alpha_0]^T) &= \\ &= (\alpha_0^4 + \alpha_0^3 - \alpha_0 + (-\alpha_0)^4 - (-\alpha_0)^2 + (-\alpha_0) + \alpha_0^2 - \alpha_0 + \alpha_0(-\alpha_0)\alpha_0) = \\ &= (2\alpha_0^4 - 3\alpha_0) \end{aligned} \quad (6.85.)$$

Očitno bo najbolj ugodna velikost  $\alpha_0$  premika v smeri  $d_0$  tista, pri kateri bo dosežen minimum funkcije (6.85.). Rešiti torej moramo problem:

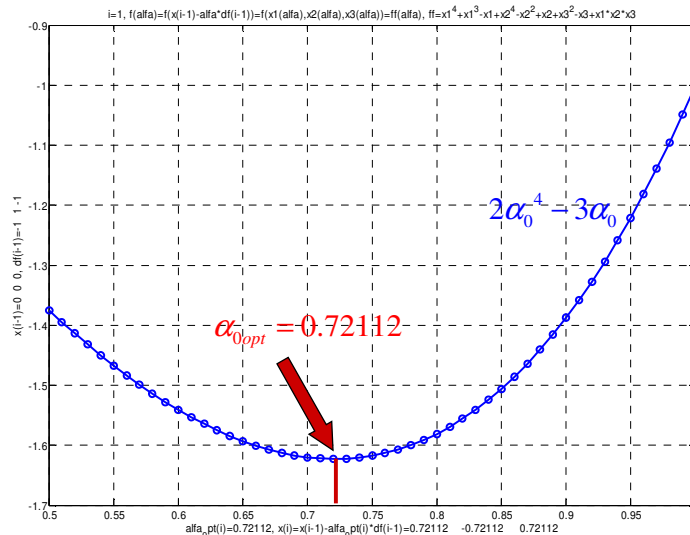
$$\min_{\alpha_0 > 0} f([\alpha_0 \ -\alpha_0 \ \alpha_0]^T) = \min_{\alpha_0 > 0} (2\alpha_0^4 - 3\alpha_0) \quad (6.86.)$$

Vsekakor gre pri izrazu (6.86.) za problem enodimenzionalne optimizacije. Na analitičen način bi  $\alpha_{0opt}$  poiskali z naslednjim izrazom:

$$\frac{d}{d\alpha_0} (2\alpha_0^4 - 3\alpha_0) = 8\alpha_0^3 - 3 = 0 \quad \Rightarrow \alpha_{0opt} = 0.7212 \quad (6.87.)$$

Vendar računalniški program običajno nima vgrajene funkcije za analitično računanje. Zato moramo  $\alpha_{0opt}$  poiskati z numerično enodimenzionalno optimizacijo z eno izmed metod iz poglavja 6.2.

Potek funkcije (6.85.) ima obliko, prikazano na sliki 78.



Slika 78: Potek funkcije (6.85.) in njen minimum

Če bi poiskali minimum te funkcije z eno izmed numeričnih enodimenzionalnih metod, bi dobili rezultat:

$$\alpha_{0opt} = 0.7212 \quad (6.88.)$$

ki je seveda enak rezultatu v izrazu (6.87.). Če rezultat (6.88.) vstavimo v izraz (6.84.), dobimo:

$$x_1 = [0.72112 \quad -0.72112 \quad 0.72112]^T \quad (6.89.)$$

kar je izraz za optimalni vektor odločitvenih spremenljivk po 1. iteraciji.

Zopet informativno preverimo, kakšna je vrednost funkcije (6.49.) v točki (6.89.):

$$\begin{aligned} f(0.72112, -0.72112, 0.72112) &= \\ &= 0.72112^4 + 0.72112^3 - 0.72112 + \\ &+ (-0.72112)^4 - (-0.72112)^2 + (-0.72112) + 0.72112^2 - 0.72112 + \\ &+ 0.72112 \cdot (-0.72112) \cdot 0.72112 = -1.6225 \end{aligned} \quad (6.90.)$$

Očitno se je vrednost funkcije že precej pomanjšala glede na izraz (6.78.).

Če postopek na osnovi izraza (6.77.) ter slike 76 ponavljamo do izpolnitve konvergenčnih kriterijev, dobimo za 2. in naslednje iteracije rezultate, prikazane v tabeli na sliki 79.

k	x1(k)	x2(k)	x3(k)	$\alpha(k-1)$	f(k)
1	0,72112	-0,72112	0,72112	0,72112	-1,6225
2	0,5288	-0,90373	0,73083	0,12488	-1,9021
3	0,56307	-0,9397	0,73325	0,14894	-1,9106
4	0,5699	-0,93199	0,75121	0,28684	-1,9113
5	0,56785	-0,93955	0,75523	0,14011	-1,9116
6	0,57064	-0,93681	0,7618	0,28489	-1,9117
7	0,56974	-0,93955	0,76333	0,13925	-1,9117
8	0,57077	-0,93853	0,76579	0,2843	-1,9118
9	0,57044	-0,93956	0,76636	0,13896	-1,9118
10	0,57083	-0,93917	0,76728	0,28405	-1,9118
11	0,5707	-0,93956	0,76749	0,13886	-1,9118

Izračunan minimum :  
 $\vec{x}^* = [\vec{x}_1^* \quad \vec{x}_2^* \quad \vec{x}_3^*]^T$

$\alpha_{kopt}$   
se tako spreminja!

Slika 79: Doseženi rezultati pri iskanju minimuma funkcije (6.49.) s pomočjo metode najstrmejšega spusta.

Če končen rezultat na sliki 79 primerjamo z rezultatom (6.50.) (glej tudi sliko 50), doseženim pri Hooke-Jeevesovi metodi, seveda dobimo enak rezultat. Vendar je Hooke-Jeevesova metoda potrebovala precej več iteracij (30 izračunavanj) kot pa Metoda najstrmejšega spusta (11 iteracij). Razlog je seveda v tem, da je slednja uporabljala pri izračunih tudi odvode funkcije, kar je zaradi dodatnih informacij dokaj pospešilo konvergenco metode.

#### 6.3.1.4 Newtonova metoda (gradientna metoda II. reda)

Newtonova multidimenzionalna optimizacijska metoda predstavlja posplošitev enodimenzionalne Newtonove metode. Njena uporaba je pogojena z dejstvom, da mora biti funkcija vsaj dvakrat odvedljiva, saj uporablja Hessovo matriko 2. odvodov funkcije. Kadar funkcija konvergira, se metoda izkaže za izredno hitro, večje težave pa se lahko pojavijo pri komplicirani naravi oz. obliki funkcije, kar lahko vodi celo do divergence postopka.

Pri izračunavanju uporabljamo inverzno Hessovo matriko, zaradi česar pa se lahko pojavijo težave v primeru, če je Hessova matrika singularna oz. skoraj singularna [4]. V primeru večjega števila spremenljivk in kompliciranih funkcij, lahko metoda postane časovno potratna ter zahteva veliko računskih operacij.

Podobno kot pri Cauchyevi metodi najstrmejšega spusta, tudi tu izrazimo funkcijo vektorske spremenljivke v obliki Taylorjeve vrste, kjer zanemarimo višje člene [4, 13, 14, 17, 21]:

$$f(x_{k+1}) = f(x_k + \Delta x_k) \approx f(x_k) + [\nabla f(x_k)]^T \cdot \Delta x_k + \frac{1}{2} \Delta x_k^T \cdot \nabla^2 f(x_k) \cdot \Delta x_k \quad (6.91.)$$

kjer je  $\nabla f(x_k)$  gradient (vektor 1. odvodov funkcije),  $\nabla^2 f(x_k)$  je Hessova matrika 2. odvodov funkcije,  $x_{k+1}$  je vektor odločitvenih spremenljivk v naslednji iteraciji in  $x_k$  vektor odločitvenih spremenljivk v trenutni iteraciji postopka.

Ker želimo poiskati minimum funkcije  $f$ , izraz (6.91) odvajamo in enačimo z nič [13, 17]:

$$\begin{aligned} \frac{\partial}{\partial \Delta x_k} f(x_{k+1}) &= \frac{\partial}{\partial \Delta x_k} f(x_k + \Delta x_k) \approx \\ &\approx 0 + [\nabla f(x_k)]^T \cdot 1 + \frac{1}{2} \cdot \frac{\partial}{\partial \Delta x_k} (\Delta x_k^T \cdot \nabla^2 f(x_k) \cdot \Delta x_k) = \\ &= [\nabla f(x_k)]^T + \frac{1}{2} \cdot 2 \cdot \nabla^2 f(x_k) \cdot \Delta x_k = [\nabla f(x_k)]^T + \nabla^2 f(x_k) \cdot \Delta x_k = 0, \end{aligned} \quad (6.92.)$$

pri čemer pa smo kot spremenljivko vzeli le  $\Delta x_k$ , ostale izraze pa smo fiksirali. Tako dobimo:

$$[\nabla f(x_k)]^T + \nabla^2 f(x_k) \cdot \Delta x_k = 0 \quad (6.93.)$$

oziroma:

$$\Delta x_k = -[\nabla^2 f(x_k)]^{-1} \cdot [\nabla f(x_k)]^T = x_{k+1} - x_k \quad (6.94.)$$

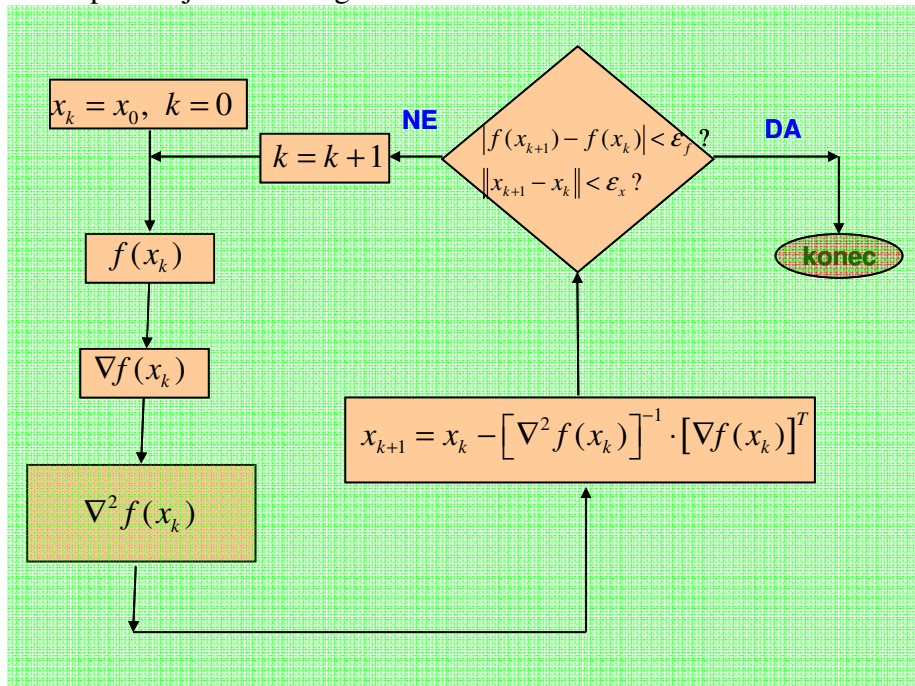
Izraz (6.94.) nas pripelje do iteracijske enačbe Newtonovega postopka [13, 17, 21]:

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \cdot [\nabla f(x_k)]^T, \quad k = 0, 1, 2, \dots \quad (6.95.)$$

kjer je  $k$  indeks iteracij.

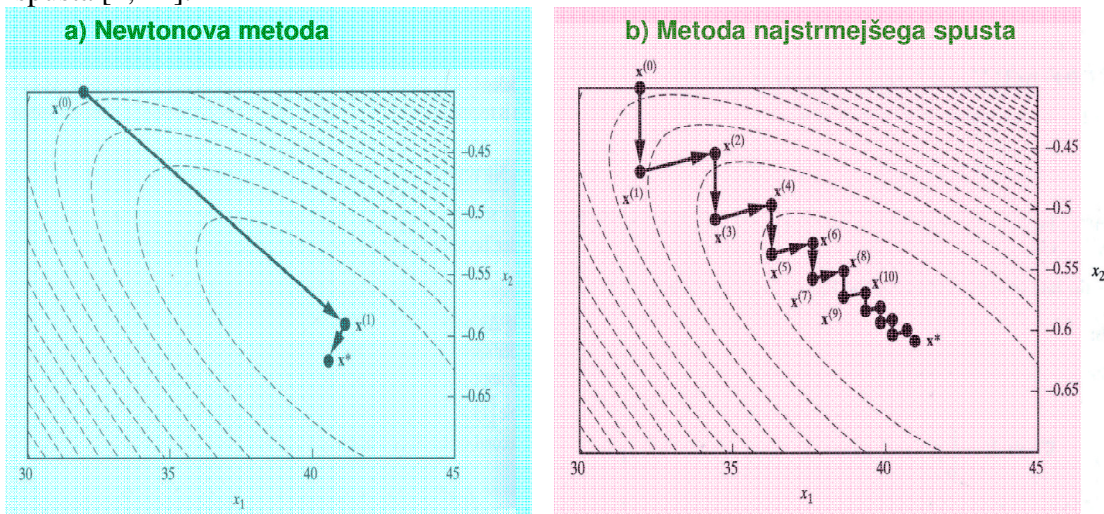
Opazimo lahko, da pride do invertiranja Hessove matrike, kar zna biti problematično s stališča morebitne singularnosti ali potratnega računanja.

Slika 80 nam prikazuje celoten algoritem Newtonove multidimenzionalne metode [4].



Slika 80: Algoritem Newtonove multidimenzionalne metode.

Na sliki 81 lahko vidimo primerjavo dinamike konvergence za Newtonovo multidimenzionalno metodo in dinamiko konvergence Cauchyve metode najstrmejšega spusta [4, 22].



Slika 81: Primerjava dinamike konvergence za Newtonovo multidimenzionalno metodo in Cauchyvo metodo najstrmejšega spusta

Kot je razvidno iz slike 81, je Newtonova multidimenzionalna metoda v primeru konvergence izredno hitra, saj doseže optimalno rešitev v minimalnem številu iteracij,

medtem ko je Cauchyeva metoda najstrmejšega spusta običajno nekoliko počasnejša. Vendar pa se lahko pojavi težava pri funkcijah kompleksne narave oziroma kompleksne oblike, ki niso dvakrat odvedljive, saj lahko tedaj Newtonova multidimenzionalna metoda divergira, ali pa se bistveno upočasni njena divergenca.

Osvetlimo delovanje Newtonove metode še na naslednjem primeru.

**Primer:**

Gre za enak primer, kot smo ga obravnavali že pri Hooke-Jeevesovi metodi (glej 1. primer in izraz (6.49.)) in Metodi najstrmejšega spusta (edini primer), pri čemer tokrat vzamemo začetni vektor odločitvenih spremenljivk  $x_0^T = [1 \quad -1 \quad 1]$ .

Pri iskanju optimalnih rezultatov si seveda pomagamo z algoritmom Newtonove metode, ki je bil opisan v tem poglavju.

**Rešitev:**

Najprej zgolj informativno preverimo, kakšna je vrednost funkcije (6.49.) v začetni točki:

$$f(1, -1, 1) = 1^4 + 1^3 - 1 + (-1)^4 - (-1)^2 + (-1) + 1^2 - 1 + 1 \cdot (-1) \cdot 1 = -1 \quad (6.96.)$$

Gradient in Hessova matrika funkcije (6.49.) v poljubni točki sta enaka:

$$\begin{aligned} \nabla f(x) &= \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \quad \frac{\partial f(x)}{\partial x_3} \right]^T = \\ &= \left[ 4x_1^3 + 3x_1^2 - 1 + x_2x_3 \quad 4x_2^3 - 2x_2 + 1 + x_1x_3 \quad 2x_3 - 1 + x_1x_2 \right]^T \\ \nabla^2 f(x) &= \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f(x)}{\partial x_3 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_3 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 12x_1^2 + 6x_1 & x_3 & x_2 \\ x_3 & 12x_2^2 - 2 & x_1 \\ x_2 & x_1 & 2 \end{bmatrix} \end{aligned} \quad (6.97.)$$



Iteracijska enačba (6.95.) bo v 1. iteraciji zavzela vrednost:

$$x_1 = x_0 - [\nabla^2 f(x_0)]^{-1} \cdot [\nabla f(x_0)]^T \quad (6.98.)$$

Gradient in hessova matrika izraza (6.97.) sta v začetni točki enaka:

$$\begin{aligned} \nabla f(x_0) &= \\ &= [4 \cdot 1^3 + 3 \cdot 1^2 - 1 + (-1) \cdot 1 \quad 4 \cdot (-1)^3 - 2 \cdot (-1) + 1 + 1 \cdot 1 \quad 2 \cdot 1 - 1 + 1 \cdot (-1)]^T = \\ &= [5 \quad 0 \quad 0]^T \end{aligned} \quad (6.99.)$$

$$\nabla^2 f(x_0) = \begin{bmatrix} 12 \cdot 1^2 + 6 \cdot 1 & 1 & -1 \\ 1 & 12(-1)^2 - 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 18 & 1 & -1 \\ 1 & 10 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

Če vstavimo izraza (6.99.) v izraz (6.98.), dobimo za vektor odločitvenih spremenljivk v 1. iteraciji naslednjo rešitev:

$$\begin{aligned} x_1 &= x_0 - [\nabla^2 f(x_0)]^{-1} \cdot [\nabla f(x_0)]^T = \\ &= [1 \quad -1 \quad 1]^T - \begin{bmatrix} 18 & 1 & -1 \\ 1 & 10 & 1 \\ -1 & 1 & 2 \end{bmatrix}^{-1} \cdot [5 \quad 0 \quad 0]^T = \\ &= \begin{bmatrix} 0.7103 \\ -0.9542 \\ 0.8323 \end{bmatrix} \end{aligned} \quad (6.100.)$$

Zopet informativno preverimo, kakšna je vrednost funkcije (6.49.) v točki (6.100.):

$$f(0.7103, -0.9542, 0.8323) = \dots = -1.8367 \quad (6.101.)$$

Očitno se je vrednost funkcije že občutno pomanjšala glede na izraz (6.96.).

Če postopek na osnovi izraza (6.95.) ter slike 80 ponavljamo do izpolnitve konvergenčnih kriterijev, dobimo za 2. in naslednje iteracije rezultate, prikazane v tabeli na sliki 82.

k	x1(k)	x2(k)	x3(k)	f(k)
0	1	-1	1	-1
1	0,71037	-0,95427	0,83232	-1,8367
2	0,59255	-0,94165	0,77825	-1,9101
3	0,5715	-0,93962	0,76847	-1,9118
4.0	0,57086	-0,93956	0,76818	-1,9118
5.0	0,57086	-0,93956	0,76818	-1,9118

Izračunan minimum :  $\vec{x}^* = [\vec{x}_1^* \ \vec{x}_2^* \ \vec{x}_3^*]^T$

Slika 82: Doseženi rezultati pri iskanju minimuma funkcije (6.49.) s pomočjo Newtonove multidimenzionalne metode

Če končen rezultat na sliki 82 primerjamo z rezultatom (6.50.) (glej tudi sliko 50), doseženim pri Hooke-Jeevesovi metodi, ter rezultatom za Metodo najstrmejšega spusta iz slike 79, dobimo seveda enak rezultat. Vendar je Hooke-Jeevesova metoda potrebovala precej več iteracij (30 izračunavanj), pa tudi Metoda najstrmejšega spusta jih je potrebovala občutno več (11 iteracij), medtem ko Newtonova metoda potrebuje le 5 iteracij za dosego minimuma. Razlog je seveda v tem, da je slednja uporabljala pri izračunih tudi druge odvode funkcije poleg prvih, kar je zaradi dodatnih informacij še pospešilo konvergenco metode v primerjavi z metodo najstrmejšega spusta, da ne govorimo o Hooke-Jeevesovi metodi.

### 6.3.1.5 Marquardtova metoda

Včasih se pri Newtonovi multidimenzionalni metodi pojavijo problemi, povezani s Hessovo matriko drugih odvodov, saj lahko invertiranje te matrike pripelje do njene singularnosti oz. slabo pogojenega sistema enačb pri računanju inverza. To pa lahko vodi do divergence postopka, ali pa se njegova konvergenca bistveno upočasni.

Da bi premostili te probleme, je v takšnih primerih smiselno vpeljati tako imenovano Marquardtovo metodo, ki sta jo razvila Levenberg in Marquardt [17, 21] in predstavlja izboljšano različico Newtonove multidimenzionalne metode.

Metoda se od Newtonove multidimenzionalne metode razlikuje v tem, da je Hessova matrika modificirana na naslednji način [17, 21]:

$$\tilde{H} = H + \lambda \cdot I \quad (6.102.)$$

kjer je  $H(x_k) = [\nabla^2 f(x_k)]$ ,  $k = 0, 1, 2, \dots$  originalna Hessova matrika,  $I$  je enotina matrika,  $\lambda$  pa je dodatni skalarni faktor, ki ga je potrebno ustrezno nastaviti. Podrobnosti o tem lahko bralec zasledi v literaturi [17, 21].

Če upoštevamo izraz (6.102), bo Newtonova iteracijska enačba (6.95) zavzela naslednjo spremenjeno obliko [17, 21]:

$$x_{k+1} = x_k - \tilde{H}^{-1} \cdot [\nabla f(x_k)]^T, \quad k = 0, 1, 2, \dots \quad (6.103.)$$

ki pomaga odpraviti morebitne težave Newtonove multidimenzionalne metode.

Tako lahko Marquardtova metoda še posebej učinkovito nadomesti Newtonovo multidimenzionalno metodo, kadar se pojavi katera od naslednjih značilnih težav pri iskanju minimuma [17, 21]:

- Newtonova multidimenzionalna metoda lahko včasih divergira oz. konvergira k sedelni točki ali lokalnemu maksimumu, namesto da bi konvergirala k lokalnemu minimumu.
- Newtonova multidimenzionalna metoda lahko včasih, kljub dodatni informaciji, ki jo nosi drugi odvod, prepočasi konvergira k lokalnemu minimumu.

V takšnih situacijah se uporaba Marquardtove metode izkaže za primerno nadomestilo, ki učinkovito reši omenjene probleme. Več o njeni uporabi in ustreznih primerih lahko bralec zasledi v literaturi [17, 21].

Z opisom Marquardtove metode smo končali pregled osnovnih optimizacijskih multidimenzionalnih metod. V nadaljevanju si bomo pogledali še nekaj bolj sofisticiranih metod, ki se v praksi tudi pogostokrat uporabljajo zaradi dobrih lastnosti, ki jih imajo.

### **6.3.2 Sofisticirane metode**

Pri obravnavi sofisticiranih multidimenzionalnih optimizacijskih metod se bomo omejili na tri tipe metod: Powellovo metodo konjugiranih smeri, Metodo konjugiranih gradientov, ter takoimenovane Quasi-Newtonove metode.

#### **6.3.2.1 Powellova metoda konjugiranih smeri**

Metoda spada v razred metod konjugiranih smeri. Razvita je bila za tip konveksnih kvadratnih funkcij in se jo da tudi najlažje pojasniti za ta tip funkcij. Seveda se jo pa da uporabiti tudi za druge tipe funkcij. Vsekakor je v praksi zelo popularna in uporabljana metoda [13, 14, 17, 21].

Powellova metoda temelji na predpostavki, da imamo opravka z naslednjo kvadratno funkcijo oz. kvadratnim problemom [13, 14, 17, 21]:

$$\min_x f(\underline{x}) = \min \left\{ \frac{1}{2} \underline{x}^T \cdot \underline{A} \cdot \underline{x} + \underline{b}^T \cdot \underline{x} + a \right\}, \quad (6.104)$$

kjer je  $\underline{A}$  simetrična pozitivno definitna matrika koeficientov kvadratne forme,  $\underline{b}$  vektor koeficientov linearnega člena in  $a$  konstanta.

V nadaljevanju bomo videli bistvo te metode, in sicer, da temelji na konjugiranih vektorjih (smereh) matrike  $\underline{A}$  !!!

### **Definicija konjugiranih smeri matrike $\underline{A}$ :**

Če imamo pozitivno definitno simetrično matriko  $\underline{A}$  z  $n$  različnimi lastnimi vrednostmi, je možno poiskati  $n$  neodvisnih smeri, ki se skladajo s smermi lastnih vektorjev.

Za konjugirane vektorje  $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n$  veljajo naslednji izrazi [13, 14, 17, 21]:

$$\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_j = 0; \quad i \neq j; \quad i, j = 0, 1, \dots, n-1 \quad (6.105)$$

Če je denimo  $\underline{A} = \underline{I}$ , sledi:

$$\underline{u}_i^T \cdot \underline{I} \cdot \underline{u}_j = \underline{u}_i^T \cdot \underline{u}_j = 0 \quad (6.106)$$

kar pa je lastnost ORTOGONALNOSTI MED VEKTORJEMA  $\underline{u}_i$  in  $\underline{u}_j$ . Torej je ortogonalnost vektorjev poseben primer konjugiranosti vektorjev. Konjugirani vektorji  $\underline{u}_i$  določajo konjugirane smeri glede na matriko  $\underline{A}$  in so linearno neodvisni. Pri tem velja naslednja trditev: Set lastnih vektorjev  $\underline{u}_i$  ustvari set konjugiranih smeri glede na matriko  $\underline{A}$  !!!

### **Teorem 1:**

Če so  $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n$  konjugirane smeri matrike  $\underline{A}$ , potem je mogoče določiti minimum kvadratne funkcije  $f(\underline{x})$  iz izraza (6.104.) s pomočjo enkratnega in samo enkratnega iskanja minimuma vzdolž vsake osi. Pri tem je  $n$  – dimenzija prostora (enaka številu

odločitvenih spremenljivk). Rešitev se torej da poiskati v  $n$  ali manj tskoimenovanih Powellovih ciklih, pri čemer imamo kvadratno konvergenco [13].

**Dokaz teorema 1:**

$\underline{x}_0$  – začetni vektor (zberemo!)

a.) Denimo gremo vzdolž prve konjugirane smeri  $\underline{u}_1$  in poiščemo minimum funkcije  $f(\underline{x})$  v tej smeri. Ta minimum leži v točki:  $\underline{x}_1 = \underline{x}_0 + \Theta_1 \cdot \underline{u}_1$ , kjer smo  $\Theta_1$  poiskali z enodimenzionalno optimizacijo.

b.) Nato se iz  $\underline{x}_1$  pomaknemo vzdolž druge konjugirane smeri  $\underline{u}_2$  in poiščemo spet minimum  $f(\underline{x})$  v tej smeri. Ta minimum leži v točki:  $\underline{x}_2 = \underline{x}_1 + \Theta_2 \cdot \underline{u}_2$ , kjer smo  $\Theta_2$  poiskali z enodimenzionalno optimizacijo.

↓ ITN,..., postopek nadaljujemo, pri čemer dobimo naslednji izraz:

$$\underline{x}_n = \underline{x}_{n-1} + \Theta_n \cdot \underline{u}_n \tag{6.107}$$

Za vektor (6.107.) lahko na osnovi a.), b.), itn..., zapišemo:

$$\begin{aligned} \underline{x}_n &= \left( \underbrace{\underline{x}_{n-2} + \Theta_{n-1} \cdot \underline{u}_{n-1}}_{\underline{x}_{n-1}} \right) + \Theta_n \cdot \underline{u}_n = \underbrace{\underline{x}_{n-3} + \Theta_{n-2} \cdot \underline{u}_{n-2}}_{\underline{x}_{n-2}} + \Theta_{n-1} \cdot \underline{u}_{n-1} + \Theta_n \cdot \underline{u}_n = \dots \\ &= \dots = \underbrace{\underline{x}_0 + \Theta_1 \cdot \underline{u}_1}_{\underline{x}_1} + \Theta_2 \cdot \underline{u}_2 + \Theta_3 \cdot \underline{u}_3 + \dots + \Theta_n \cdot \underline{u}_n \end{aligned} \tag{6.108}$$

Izraz (6.108.) lahko zapišemo tudi kot [13]:

$$\underline{x}_n = \underline{x}_0 + \left[ \Theta_1 \cdot \underline{u}_1 + \Theta_2 \cdot \underline{u}_2 + \dots + \Theta_n \cdot \underline{u}_n \right] \tag{6.109}$$

$$\boxed{\underline{x}_n = \underline{x}_0 + \sum_{i=1}^n \Theta_i \cdot \underline{u}_i}$$

Torej lahko končni vektor po  $n$  ciklih izrazimo kot začetni vektor in linearno kombinacijo  $n$  konjugiranih vektorjev, pri čemer smo  $\Theta_i$ ,  $i = 1, \dots, n$  parametrov poiskali z enodimenzionalnimi optimizacijami po  $n$  konjugiranih smereh. Oziroma, z drugimi besedami, če nam je poznanih  $n$  linearno neodvisnih konjugiranih vektorjev matrike  $\underline{A}$ , se problem iskanja optimalne točke  $\underline{x}_n = \underline{x}^*$  zreducira na reševanje  $n$  enodimenzionalnih problemov, s katerimi poiščemo parametre  $\Theta_i$ ,  $i = 1, \dots, n$  [13].

Seveda pa je potrebno še dokazati, če je  $\underline{x}_n = \underline{x}_0 + \sum_{i=1}^n \Theta_i \cdot \underline{u}_i$  res optimalna točka  $\underline{x}^*$ .

V ta namen najprej razvijmo funkcijo  $f(\underline{x})$  izraza (6.104.) v Taylorjevo vrsto v okolici točke  $\underline{x}_i$ , pri čemer upoštevamo le prve 3 člene [13]:

$$f(\underline{x}_i + \Delta \underline{x}_i) \doteq f(\underline{x}_i) + [\nabla f(\underline{x}_i)]^T \cdot \Delta \underline{x} + \frac{1}{2} \cdot \Delta \underline{x}^T \cdot \underline{\underline{H}} \cdot [f(\underline{x}_i)] \cdot \Delta \underline{x} \quad (6.110)$$

torej se v izrazu pojavita gradient in Hessova matrika funkcije.

Ker je  $f(\underline{x}_i) = \frac{1}{2} \cdot \underline{x}_i^T \cdot \underline{\underline{A}} \cdot \underline{x}_i + \underline{b}^T \cdot \underline{x}_i + a$ . (glej izraz (6.104.)), za gradient sledi [13]:

$$\nabla f(\underline{x}_i) = \left[ \frac{\partial f(\underline{x}_i)}{\partial \underline{x}_i} \right] = \frac{1}{2} \cdot 2 \cdot \underline{\underline{A}} \cdot \underline{x}_i + \underline{b} = \underline{\underline{A}} \cdot \underline{x}_i + \underline{b} \quad (6.111.)$$

pri čemer smo uporabili pravila za vektorsko odvajanje [13].

Za Hessovo matriko pa velja, če upoštevamo izraz (6.111.):

$$H[f(\underline{x}_i)] = \nabla^2 f(\underline{x}_i) = \nabla [\nabla f(\underline{x}_i)] = \frac{\partial}{\partial \underline{x}_i} [\underline{\underline{A}} \cdot \underline{x}_i + \underline{b}] = \underline{\underline{A}} \quad (6.112.)$$

V splošnem velja tudi izraz (glej (6.107.)) [13]:

$$\underline{x}_i = \underline{x}_{i-1} + \Theta_i \cdot \underline{u}_i = \underline{x}_{i-1} + \Delta \underline{x}_i \quad \text{oz.} \quad (6.113.)$$

$$\Delta \underline{x}_i = \Theta_i \cdot \underline{u}_i$$

Če izraze (6.111.), (6.112.) in (6.113.) upoštevamo v Taylorjevi vrsti (6.110.), dobimo:

$$\begin{aligned} f(\underline{x}_i + \Theta_i \cdot \underline{u}_i) &\doteq f(\underline{x}_i) + (\underline{\underline{A}} \cdot \underline{x}_i + \underline{b})^T \cdot \Theta_i \cdot \underline{u}_i + \frac{1}{2} \cdot (\Theta_i \cdot \underline{u}_i)^T \cdot \underline{\underline{A}} \cdot (\Theta_i \cdot \underline{u}_i) = \\ &= f(\underline{x}_i) + (\underline{\underline{A}} \cdot \underline{x}_i + \underline{b})^T \cdot \Theta_i \cdot \underline{u}_i + \frac{1}{2} \cdot \Theta_i^2 \cdot \underline{u}_i^T \cdot \underline{\underline{A}} \cdot \underline{u}_i \end{aligned} \quad (6.114.)$$

Če izraz (6.114.) odvajamo po  $\Theta_i$  in enačimo z 0, dobimo:

$$\begin{aligned} \frac{\partial}{\partial \Theta_i} [f(\underline{x}_i + \Theta_i \cdot \underline{u}_i)] &= \frac{\partial}{\partial \Theta_i} \left[ f(\underline{x}_i) + (\underline{\underline{A}} \cdot \underline{x}_i + \underline{b})^T \cdot \Theta_i \cdot \underline{u}_i + \frac{\Theta_i^2}{2} \cdot \underline{u}_i^T \cdot \underline{\underline{A}} \cdot \underline{u}_i \right] = \\ &= (\underline{\underline{A}} \cdot \underline{x}_i + \underline{b})^T \cdot \underline{u}_i + \frac{2\Theta_i}{2} \cdot \underline{u}_i^T \cdot \underline{\underline{A}} \cdot \underline{u}_i = 0 \end{aligned} \quad (6.115.)$$

Iz izraza (6.115.) torej sledi [13]:

$$\Theta_i = \frac{-\left(\underline{A} \cdot \underline{x}_i + \underline{b}\right)^T \cdot \underline{u}_i}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \quad (6.116.)$$

kar je pomembna relacija med parametri enodimenzionalnih optimizacij  $\Theta_i$  in pripadajočimi konjugiranimi vektorji  $\underline{u}_i$ .

Na osnovi (6.111.) lahko zapišemo naslednji izraz, ki velja v optimumu [13]:

$$\nabla f(\underline{x}_n) = \underline{A} \cdot \underline{x}_n + \underline{b} = \underset{\text{V OPTIMUMU}}{\underline{0}} \quad \text{oz.} \quad (6.117.)$$

$$\boxed{\underline{x}_n = \underbrace{-\underline{A}^{-1} \cdot \underline{b}}_{\text{V OPTIMUMU}}}$$

Če je  $\underline{x}_n$  res optimum  $\underline{x}^*$ , potem bi se moral izraz (6.109.) v optimumu izenačiti z izrazom (6.117.), kar bomo skušali dokazati v nadaljevanju.

V nadaljevanju izhajamo iz izraza (6.116.), pri čemer upoštevamo KOMUTATIVNOST. Na osnovi izraza (6.109.) potem lahko zapišemo:

$$\begin{aligned} \Theta_i &= \frac{-\underline{u}_i^T \cdot (\underline{A} \cdot \underline{x}_i + \underline{b})}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} = \frac{-\underline{u}_i^T}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \left[ \underline{A} \cdot \left[ \underline{x}_0 + \sum_{j=0}^{i-1} \Theta_j \cdot \underline{u}_j \right] + \underline{b} \right] = \\ &= \frac{-1}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \left[ \underline{u}_i^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b}) + \sum_{j=0}^{i-1} \underline{u}_i^T \cdot \underline{A} \cdot \Theta_j \cdot \underline{u}_j \right] \end{aligned} \quad (6.118.)$$

pri čemer smo vektor po  $i$  ciklih izrazili kot začetni vektor in linearno kombinacijo  $i$  konjugiranih vektorjev.

Z upoštevanjem izraza (6.105.) izraz (6.118.) preide v obliko:

$$\Theta_i = \frac{-\underline{u}_i^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b})}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \quad (6.119.)$$

Če izraz (6.119.) vstavimo v izraz (6.109.), le-ta preide v obliko:

$$\underline{x}_n = \underline{x}_0 + \sum_{i=1}^n \left[ \frac{-\underline{u}_i^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b})}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \right] \cdot \underline{u}_i \quad (6.120.)$$

Sedaj moramo pokazati, da izraz (6.120.) v optimumu postane enak izrazu (6.117.), to je izrazu  $[-\underline{A}^{-1} \cdot \underline{b}]$ . V ta namen izvedemo naslednje operacije nad izrazom (6.120.):

$$\begin{aligned} \underline{x}_n - \underline{x}_0 &= -\sum_{i=1}^n \frac{\underline{u}_i^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b}) \cdot \underline{u}_i}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \Big/ \Big/^T \\ (\underline{x}_n - \underline{x}_0)^T &= -\sum_{i=1}^n \frac{\underline{u}_i^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b}) \cdot \underline{u}_i^T}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \Big/ \cdot \underline{A} \cdot \underline{u}_j \end{aligned} \quad (6.121.)$$

$$\begin{aligned} (\underline{x}_n - \underline{x}_0)^T \cdot \underline{A} \cdot \underline{u}_j &= \\ &= -\sum_{\substack{i=1 \\ i \neq j}}^n \underbrace{\frac{\underline{u}_i^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b}) \cdot \underline{u}_i^T}{\underline{u}_i^T \cdot \underline{A} \cdot \underline{u}_i} \cdot \underline{A} \cdot \underline{u}_j}_{\text{to je 0 zaradi konjugiranih smeri!}} - \underbrace{\frac{\underline{u}_j^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b}) \cdot \underline{u}_j^T}{\underline{u}_j^T \cdot \underline{A} \cdot \underline{u}_j} \cdot \underline{A} \cdot \underline{u}_j}_{\text{Člen pri } i=j \text{ smo dali posebej}} \end{aligned}$$

pri čemer smo upoštevali tudi izraz (6.105.). Iz izrazov (6.121.) je razvidno, da ponekod operacije transponiranja niso imele vpliva, saj je šlo za skalarje.

Na osnovi izrazov (6.121.) torej dobimo:

$$(\underline{x}_n - \underline{x}_0)^T \cdot \underline{A} \cdot \underline{u}_j = -\underline{u}_j^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b}) \quad (6.122.)$$

Izraz (6.122.) lahko zaradi skalarja na levi strani in simetričnosti matrice  $\underline{A} = \underline{A}^T$  zapišemo tudi kot [13]:

$$\begin{aligned} \underline{u}_j^T \cdot \underline{A} \cdot (\underline{x}_n - \underline{x}_0) &= -\underline{u}_j^T \cdot (\underline{A} \cdot \underline{x}_0 + \underline{b}) \\ \underline{A} \cdot (\underline{x}_n - \underline{x}_0) &= -(\underline{A} \cdot \underline{x}_0 + \underline{b}) \quad \Rightarrow \quad \underline{A} \cdot \underline{x}_n = -\underline{b} \quad \Rightarrow \quad \boxed{\underline{x}_n = -\underline{A}^{-1} \cdot \underline{b}} \end{aligned} \quad (6.123.)$$

Torej smo dokazali, da izraz (6.109.) po  $n$  ciklih res zavzame vrednost izraza (6.117.), kar pomeni, da gre pri izrazu (6.109.) res za optimum. S tem smo tudi uspešno končali dokaz teorema 1.

V nadaljevanju si bomo pogledali primer, kako bi poiskali optimum dane funkcije s pomočjo uporabe konjugiranih smeri, pod predpostavko, **da bi konjugirane vektorje poznali vnaprej.**

### **Primer:**

Spomočjo uporabe konjugiranih smeri določite minimum naslednje funkcije:



$$f(x_1, x_2) = \frac{1}{2} \cdot (2x_1^2 + 2x_1 \cdot x_2 + x_2^2) + x_1 + 10x_2 \quad (6.124.)$$

Funkcijo (6.124.) najprej poskušajmo prevesti v obliko, opredeljeno v izrazu (6.104), torej:

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2} \cdot \underline{x}^T \cdot \underline{A} \cdot \underline{x} + \underline{b}^T \cdot \underline{x} + a = \\ &= \frac{1}{2} \cdot \begin{bmatrix} x_1 & x_2 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + a \end{aligned} \quad (6.125.)$$

Izraz (6.125.) lahko pretvorimo v obliko:

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2} \cdot \begin{bmatrix} x_1 & x_2 \end{bmatrix} \cdot \begin{bmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 \end{bmatrix} + b_1 \cdot x_1 + b_2 \cdot x_2 + a = \\ &= \frac{1}{2} \cdot \left[ x_1 \cdot (a_{11} \cdot x_1 + a_{12} \cdot x_2) + x_2 \cdot (a_{21} \cdot x_1 + a_{22} \cdot x_2) \right] + b_1 \cdot x_1 + b_2 \cdot x_2 + a = \\ &= \frac{1}{2} \cdot \left[ a_{11} \cdot x_1^2 + a_{12} \cdot x_1 \cdot x_2 + a_{21} \cdot x_1 \cdot x_2 + a_{22} \cdot x_2^2 \right] + b_1 \cdot x_1 + b_2 \cdot x_2 + a \end{aligned} \quad (6.126.)$$

Če sedaj primerjamo istoležne člene izrazov (6.124.) in (6.126.), pri čemer upoštevamo simetričnost matrike  $\underline{A}$ , dobimo:

$$\begin{aligned} 2 &= a_{11} \\ a_{12} + a_{21} &= 2; \\ \text{ker pa je } a_{12} &= a_{21}, \text{ sledi:} \\ a_{12} + a_{12} &= 2 \Rightarrow a_{12} = 1 \\ a_{22} &= 1 \end{aligned} \quad (6.127.)$$

Na osnovi primerjave istoležnih členov izrazov (6.124.) in (6.126.), ter izpeljanih izrazov v (6.127.), torej velja:

$$\underline{A} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad b_1 = 1, \quad b_2 = 10, \quad a = 0 \quad (6.128.)$$

Funkcijo (6.124.) torej lahko na osnovi izrazov (6.125.) in (6.128.) izrazimo v obliki:

$$f(x_1, x_2) = \frac{1}{2} \cdot \underbrace{\begin{bmatrix} x_1 & x_2 \end{bmatrix}}_{\underline{x}^T} \cdot \underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{\underline{A}} \cdot \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\underline{x}} + \underbrace{\begin{bmatrix} 1 & 10 \end{bmatrix}}_{\underline{b}^T} \cdot \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\underline{x}} + 0 \quad (6.129.)$$

V nadaljevanju predpostavimo, da konjugirana vektorja poznamo vnaprej in se glasita:

$$\underline{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad \underline{u}_2 = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \quad (6.130.)$$

Preverimo, če sta vektorja v (6.130.) res konjugirana:

$$\underline{u}_1^T \cdot \underline{A} \cdot \underline{u}_2 = [1 \quad 1] \cdot \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -3 \end{bmatrix} = [1 \quad 1] \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \underline{0} \quad (6.131.)$$

DA, RES STA!

Na osnovi teorema 1 sklepamo, da se da minimum poiskati v  $n = 2$  ciklih, kar na osnovi izraza (6.109.) pomeni:

$$\underline{x}_n = \underline{x}_2 = \underline{x}_0 + \sum_{i=1}^2 \Theta_i \cdot \underline{u}_i = \underline{x}_0 + \Theta_1 \cdot \underline{u}_1 + \Theta_2 \cdot \underline{u}_2 \quad (6.132.)$$

Če je začetni vektor npr:  $\underline{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , potem iz (6.130.) in (6.132.) sledi:

$$\underline{x}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \Theta_1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \Theta_2 \cdot \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \underline{x}^* \quad (6.133.)$$

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} \Theta_1 + 2\Theta_2 \\ \Theta_1 - 3\Theta_2 \end{bmatrix}$$

Optimalni vektor v (6.133.) lahko torej izračunamo, če izračunamo še  $\Theta_1$  in  $\Theta_2$ . Pri tem pa si pomagamo z relacijo (6.116.), ki jo uporabimo na naslednji način (glej tudi izraza (6.129.) in (6.130.)):

$$\Theta_1 = \frac{-\underline{u}_1^T \cdot \left( \overbrace{\underline{A} \cdot \underline{x}_0}^{=0} + \underline{b} \right)}{\underline{u}_1^T \cdot \underline{A} \cdot \underline{u}_1} = \frac{-[1 \quad 1] \cdot \begin{bmatrix} 1 \\ 10 \end{bmatrix}}{[1 \quad 1] \cdot \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{-11}{[1 \quad 1] \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix}} = \frac{-11}{5} \quad (6.134.)$$

$$\Theta_2 = \frac{-\underline{u}_2^T \cdot \left( \overbrace{\underline{A} \cdot \underline{x}_0}^{=0} + \underline{b} \right)}{\underline{u}_2^T \cdot \underline{A} \cdot \underline{u}_2} = \frac{-[2 \quad -3] \cdot \begin{bmatrix} 1 \\ 10 \end{bmatrix}}{[2 \quad -3] \cdot \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -3 \end{bmatrix}} = \frac{-(-28)}{[2 \quad -3] \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}} = \frac{28}{5}$$

Če izraza (6.134.) vstavimo v izraz (6.133.), dobimo končen rezultat za optimalni vektor odločitvenih spremenljivk:

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} \frac{-11}{5} + 2 \cdot \frac{28}{5} \\ \frac{-11}{5} - 3 \cdot \frac{28}{5} \end{bmatrix} = \begin{bmatrix} +\frac{45}{5} \\ -\frac{95}{5} \end{bmatrix} = \begin{bmatrix} +9 \\ -19 \end{bmatrix} \rightarrow \text{OPTIMUM} \quad (6.135.)$$

Naredimo še analitični preizkus, če je rezultat (6.135.) res pravilen. V ta namen gradient funkcije (6.124.) enačimo z 0, pri čemer dobimo naslednji sistem enačb:

$$\left. \begin{array}{l} \frac{\partial f}{\partial x_1} = 2x_1 + x_2 + 1 = 0 \\ \frac{\partial f}{\partial x_2} = x_1 + x_2 + 10 = 0 \quad / \cdot 2 \end{array} \right\} - \quad (6.136.)$$

Če sistem enačb (6.136.) rešimo, dobimo:

$$\begin{array}{ll} -x_2^* + 1 - 20 = 0 & \Rightarrow x_2^* = -19 \\ 2x_1^* - 19 + 1 = 0 & \Rightarrow x_1^* = 9 \end{array} \quad (6.137.)$$

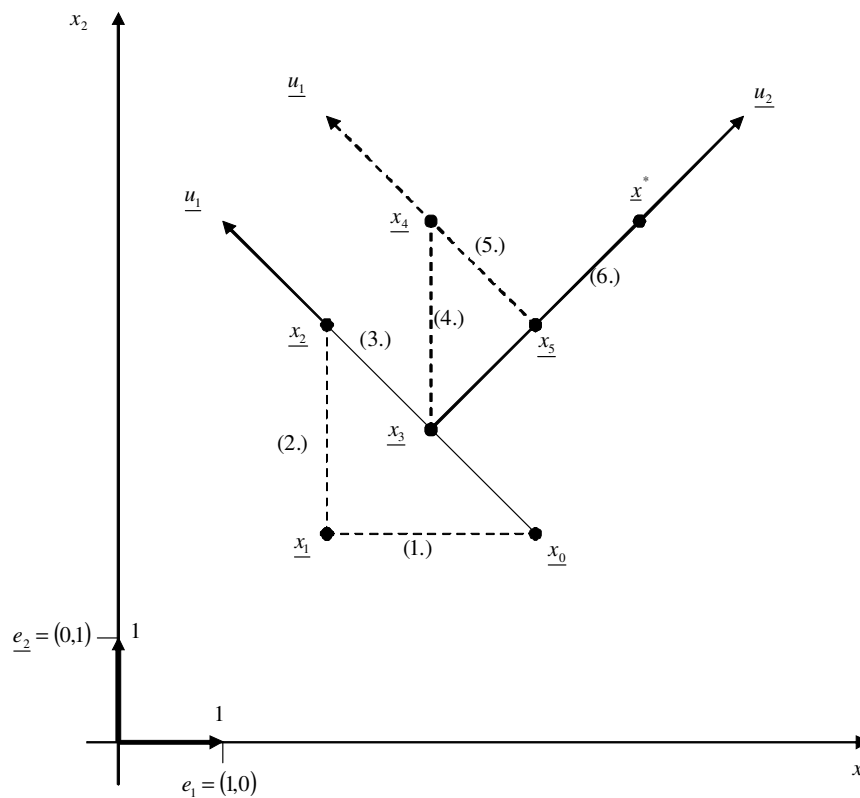
torej je dobljeni rezultat (6.35.) res pravilen.

Obravnavani primer nas je prepričal, kako lahko dokaj hitro in preprosto pridemo do optimuma dane funkcije s pomočjo konjugiranih smeri, če le poznamo konjugirana vektorja  $\underline{u}_1$  in  $\underline{u}_2$ . Žal pa se v praksi velikokrat zgodi, da konjugiranih vektorjev matrike  $\underline{A}$  ne poznamo vnaprej. Tedaj se zdi uporaba takoimenovane Powellove metode primerna rešitev, kako rešiti nastali problem [13].

### **Glavna ideja Powellove metode**

Za to metodo je značilno, da hkrati (sočasno) računa tako koeficiente enodimenzionalnih (1D) optimizacij  $\Theta_i$ , kot tudi konjugirane vektorje  $\underline{u}_i$  [13, 17].

Idejo metode najlažje ilustriramo na primeru 2 dimenzij, kjer sta  $\underline{e}_1 = (1, 0)$  in  $\underline{e}_2 = (0, 1)$  enotska vektorja v smeri  $x_1$  oz.  $x_2$  [13, 17].



Slika 83: Ilustracija ideje Powellove metode

Kot je razvidno iz slike 83, najprej opravimo iz začetnega vektorja premik v smeri enotskega vektorja  $\underline{e}_1$ . Pri tem pridemo v točko  $\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1$ , kjer optimalen  $\tilde{\Theta}_1$  določimo z enodimenzionalno optimizacijo. To pomeni, da točko  $\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1$  vstavimo v dano funkcijo, s čimer le-ta postane funkcija ene same spremenljivke  $\tilde{\Theta}_1$ . Seveda je pri tem najbolj ugoden premik pri tistem  $\tilde{\Theta}_1$ , pri katerem  $f(\tilde{\Theta}_1)$  doseže minimum. Nato postopek ponovimo s premikom v smeri enotskega vektorja  $\underline{e}_2$ . Pri tem pridemo v točko  $\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2$ , kjer optimalen  $\tilde{\Theta}_2$  ponovno določimo z enodimenzionalno optimizacijo. To pomeni, da točko  $\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2$  kot prej vstavimo v dano funkcijo, s čimer le-ta postane funkcija ene same spremenljivke  $\tilde{\Theta}_2$ . Seveda je pri tem najbolj ugoden premik pri tistem  $\tilde{\Theta}_2$ , pri katerem  $f(\tilde{\Theta}_2)$  doseže minimum. Pokazati se da, da ima vektor  $\underline{x}_2 - \underline{x}_0$  enako smer kot prvi konjugiran vektor, to je  $\underline{u}_1$ . V nadaljevanju se moramo v smeri slednjega optimalno premakniti, kar pomeni premik v  $\underline{x}_3 = \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0)$ . To zopet pomeni, da točko  $\underline{x}_3 = \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0)$  vstavimo v dano funkcijo, s čimer le-ta postane funkcija ene same spremenljivke  $\tilde{\Theta}_3$ . In zopet velja, da je najbolj ugoden premik pri tistem  $\tilde{\Theta}_3$ , pri katerem  $f(\tilde{\Theta}_3)$  doseže minimum. Ko

dosežemo optimalno točko  $\underline{x}_3$ , je 1. iteracija končana. Nato postopek ponovimo s ponovnim premikom v smeri enotskega vektorja  $\underline{e}_2$ . Pri tem pridemo v točko  $\underline{x}_4 = \underline{x}_3 + \tilde{\Theta}_4 \cdot \underline{e}_2$ , kjer optimalen  $\tilde{\Theta}_4$  ponovno določimo z enodimenzionalno optimizacijo. To pomeni, da točko  $\underline{x}_4 = \underline{x}_3 + \tilde{\Theta}_4 \cdot \underline{e}_2$  vstavimo v dano funkcijo, s čimer le-ta postane funkcija ene same spremenljivke  $\tilde{\Theta}_4$ . Seveda je pri tem najbolj ugoden premik pri tistem  $\tilde{\Theta}_4$ , pri katerem  $f(\tilde{\Theta}_4)$  doseže minimum. Iz  $\underline{x}_4$  se moramo nato optimalno premakniti, ponovno v smeri 1. konjugiranega vektorja. Pri tem pridemo v točko  $\underline{x}_5 = \underline{x}_4 + \tilde{\Theta}_5 \cdot (\underline{x}_3 - \underline{x}_0)$ , kjer optimalen  $\tilde{\Theta}_5$  ponovno določimo z enodimenzionalno optimizacijo. To pomeni, da točko  $\underline{x}_5 = \underline{x}_4 + \tilde{\Theta}_5 \cdot (\underline{x}_3 - \underline{x}_0)$  vstavimo v dano funkcijo, s čimer le-ta postane funkcija ene same spremenljivke  $\tilde{\Theta}_5$ . Seveda je pri tem najbolj ugoden premik pri tistem  $\tilde{\Theta}_5$ , pri katerem  $f(\tilde{\Theta}_5)$  doseže minimum. Pokazati se da, da ima vektor  $\underline{x}_5 - \underline{x}_3$  enako smer kot drugi konjugiran vektor, to je  $\underline{u}_2$ . V nadaljevanju se moramo v smeri slednjega optimalno premakniti, kar pomeni premik v  $\underline{x}^* = \underline{x}_3 + \tilde{\Theta}_6 \cdot (\underline{x}_5 - \underline{x}_3)$ . To zopet pomeni, da točko  $\underline{x}^* = \underline{x}_3 + \tilde{\Theta}_6 \cdot (\underline{x}_5 - \underline{x}_3)$  vstavimo v dano funkcijo, s čimer le-ta postane funkcija ene same spremenljivke  $\tilde{\Theta}_6$ . In zopet velja, da je najbolj ugoden premik pri tistem  $\tilde{\Theta}_6$ , pri katerem  $f(\tilde{\Theta}_6)$  doseže minimum. Ko dosežemo optimalno točko  $\underline{x}^*$ , je 2. iteracija končana.

Če analiziramo pravkar opisani postopek, lahko vidimo, da se očitno sočasno računajo tako koeficienti enodimenzionalnih optimizacij  $\Theta_i$ , kot tudi konjugirani vektorji  $\underline{u}_i$  [13, 17]. Seveda je v skladu z ugotovitvami predhodne obravnave po 2 iteracijah oz. 2 Povellovih ciklih gotovo dosežen optimum v  $\underline{x}^*$ , saj je  $n = 2$  (dimenzija prostora) v opisanem primeru.

Strnimo torej pravkar opisan postopek v bolj kompaktno obliko opisa [13, 14, 17].

PRVA ITERACIJA:

$$\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1 \quad (6.138.)$$

$$\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2 \quad (6.139.)$$

$$\underline{x}_3 = \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0) \quad (6.140.)$$

DRUGA ITERACIJA:

$$\underline{x}_4 = \underline{x}_3 + \tilde{\Theta}_4 \cdot \underline{e}_2 \quad (6.141.)$$

$$\underline{x}_5 = \underline{x}_4 + \tilde{\Theta}_5 \cdot (\underline{x}_3 - \underline{x}_0) \quad (6.142.)$$

$$\underline{x}^* = \underline{x}_3 + \tilde{\Theta}_6 \cdot (\underline{x}_5 - \underline{x}_3) \quad (6.143.)$$

1. ITERACIJA:

- Gremo iz  $\underline{x}_0$  v  $\underline{x}_1$  v smeri  $\underline{e}_1$ , z 1D optimizacijo. Pri tem izračunamo  $\tilde{\Theta}_1$ .
- Gremo nato iz  $\underline{x}_1$  v  $\underline{x}_2$  v smeri  $\underline{e}_2$ , z 1D optimizacijo. Pri tem izračunamo  $\tilde{\Theta}_2$ .
- Dokazati se da, da je  $\underline{u}_1 = \underline{x}_2 - \underline{x}_0 \rightarrow$  1. konjugiran vektor.
- Gremo nato iz  $\underline{x}_2$  v  $\underline{x}_3$  v smeri  $\underline{u}_1 = \underline{x}_2 - \underline{x}_0$ , pri čemer z 1D optimizacijo izračunamo tudi  $\tilde{\Theta}_3$ .

2. ITERACIJA:

- Gremo nato iz  $\underline{x}_3$  v  $\underline{x}_4$  v smeri enotskega vektorja  $\underline{e}_2$ , pri čemer z 1D optimizacijo izračunamo tudi  $\tilde{\Theta}_4$ .
- Gremo nato iz  $\underline{x}_4$  v  $\underline{x}_5$ , zopet v smeri  $\underline{u}_1$  (saj je  $\underline{x}_5 - \underline{x}_4$  paralelen z  $\underline{u}_1$ ), pri čemer z 1D optimizacijo izračunamo tudi  $\tilde{\Theta}_5$ .
- Dokazati se da, da smo sedaj dobili še 2. konjugiran vektor, ki je  $\underline{u}_2 = \underline{x}_5 - \underline{x}_3$ .
- Končno gremo v  $\underline{x}_3$  ter v velikem skoku iz njega skočimo v optimalno rešitev  $\underline{x}^*$ , pri čemer storimo premik v smeri 2. konjugiranega vektorja  $\underline{u}_2$ . Pri tem se z 1D optimizacijo hkrati izračuna  $\tilde{\Theta}_6$ .

Izvršili smo 2 iteraciji, v vsaki pa smo izračunali tudi konjugiran vektor (torej  $\underline{u}_1$  in  $\underline{u}_2$ ). Vsaka iteracija je imela 3 podkorake, saj smo morali zaradi hkratnega računanja konjugiranih smeri  $\underline{u}_1$ ,  $\underline{u}_2$ , namesto parametrov  $\tilde{\Theta}_1$  in  $\tilde{\Theta}_2$ , s 6 vmesnimi optimizacijami računati tudi vmesne parametre  $\tilde{\Theta}_i$ ,  $i = 1, \dots, 6$  (pomožni parametri).

Kot je razvidno iz slike 83, se je smer premikov spreminjala v skladu z naslednjo sekvenco:  $(\underline{e}_1, \underline{e}_2, \underline{u}_1, \underline{e}_2, \underline{u}_1, \underline{u}_2)$ . Iz te sekvence je razvidno, da pri 2 dimenzijah pare smeri premika v takoimenovani smerni matriki  $\underline{S}$  organiziramo tako, da enotske vektorje postopno nadomestimo s konjugiranimi vektorji [13]:

$$\begin{aligned}
1. \quad \underline{\underline{S}} &= \begin{bmatrix} \underline{e}_1 & \underline{e}_2 \end{bmatrix} \\
2. \quad \underline{\underline{S}} &= \begin{bmatrix} \underline{e}_2 & \underline{u}_1 \end{bmatrix} \\
3. \quad \underline{\underline{S}} &= \begin{bmatrix} \underline{u}_1 & \underline{u}_2 \end{bmatrix}
\end{aligned} \tag{6.144.}$$

Popolnoma enaka filozofija velja tudi za višje dimenzije, kjer prav tako enotske vektorje postopno nadomestimo s konjugiranimi vektorji. Npr. pri 3 dimenzijah bi se smerna matrika obnašala na naslednji način [13]:

$$\begin{aligned}
1. \quad \underline{\underline{S}} &= \begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \end{bmatrix} \\
2. \quad \underline{\underline{S}} &= \begin{bmatrix} \underline{e}_2 & \underline{e}_3 & \underline{u}_1 \end{bmatrix} \\
3. \quad \underline{\underline{S}} &= \begin{bmatrix} \underline{e}_3 & \underline{u}_1 & \underline{u}_2 \end{bmatrix} \\
4. \quad \underline{\underline{S}} &= \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \underline{u}_3 \end{bmatrix}
\end{aligned} \tag{6.145.}$$

Seveda postane metoda pri višjih dimenzijah že dokaj komplicirana in težje razumljiva, čeprav deluje po popolnoma enakem principu kot pri dveh dimenzijah.

Osvetlino delovanje Powellove metode še na naslednjem primeru, ki predstavlja problem v dveh dimenzijah.

### **Primer:**

S Powellovo metodo določite minimum naslednje kvadratne funkcije:

$$f(x_1, x_2) = x_1^2 + x_1 \cdot x_2 + \frac{1}{2}x_2^2 + x_1 + 10x_2 \tag{6.146.}$$

kjer je  $\underline{x}_0^T = [0 \ 0]$  in sta enotska vektorja  $\underline{e}_1^T = [1 \ 0]$ ,  $\underline{e}_2^T = [0 \ 1]$ .

### **Rešitev:**

Pri računanju optimalne rešitve bomo uporabili izraze (6.138.) do (6.143.). Najprej uporabimo izraz (6.138.), pri čemer dobimo:

$$\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \tilde{\Theta}_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{\Theta}_1 \\ 0 \end{bmatrix} \tag{6.147.}$$

Točko (6.147.) nato vstavimo v funkcijo (6.146.), iz česar sledi:

$$f(\underline{x}_1) = f\left(\begin{bmatrix} \tilde{\Theta}_1 \\ 0 \end{bmatrix}\right) = \tilde{\Theta}_1^2 + \tilde{\Theta}_1 \cdot 0 + \frac{1}{2} \cdot 0^2 + \tilde{\Theta}_1 + 10 \cdot 0 = \tilde{\Theta}_1^2 + \tilde{\Theta}_1 = f(\tilde{\Theta}_1) \quad (6.148.)$$

$$\frac{df(\underline{x}_1)}{d\tilde{\Theta}_1} = 2\tilde{\Theta}_1 + 1 = 0 \quad \Rightarrow \quad \tilde{\Theta}_1^* = -\frac{1}{2} \quad \Rightarrow \quad \underline{x}_1^* = \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}$$

kjer smo ekstrem funkcije  $f(\tilde{\Theta}_1)$  poiskali na analitičen način. V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije.

Dobljeno točko  $\underline{x}_1^*$  nato vstavimo v izraz (6.139.), pri čemer dobimo:

$$\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2 = \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix} + \tilde{\Theta}_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \tilde{\Theta}_2 \end{bmatrix} \quad (6.149.)$$

Točko (6.149.) nato vstavimo v funkcijo (6.146.), iz česar sledi:

$$\begin{aligned} f(\underline{x}_2) &= f\left(\begin{bmatrix} -\frac{1}{2} \\ \tilde{\Theta}_2 \end{bmatrix}\right) = \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right) \cdot \tilde{\Theta}_2 + \frac{1}{2} \cdot (\tilde{\Theta}_2)^2 + \left(-\frac{1}{2}\right) + 10\tilde{\Theta}_2 = \\ &= -\frac{1}{2} + \frac{19}{2}\tilde{\Theta}_2 + \frac{1}{2}\tilde{\Theta}_2^2 \end{aligned} \quad (6.150.)$$

$$\frac{df(\underline{x}_2)}{d\tilde{\Theta}_2} = \frac{19}{2} + \tilde{\Theta}_2 = 0 \quad \Rightarrow \quad \tilde{\Theta}_2^* = -\frac{19}{2} \quad \Rightarrow \quad \underline{x}_2^* = \begin{bmatrix} -\frac{1}{2} \\ -\frac{19}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -9.5 \end{bmatrix}$$

kjer smo ekstrem funkcije  $f(\tilde{\Theta}_2)$  poiskali na analitičen način. V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije.

Dobljeno točko  $\underline{x}_2^*$  nato vstavimo v izraz (6.140.), pri čemer dobimo:

$$\underline{x}_3 = \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0) = \begin{bmatrix} -\frac{1}{2} \\ \frac{19}{2} \\ -\frac{19}{2} \end{bmatrix} + \tilde{\Theta}_3 \cdot \left( \begin{bmatrix} -\frac{1}{2} \\ \frac{19}{2} \\ -\frac{19}{2} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} -\frac{1}{2} - \frac{1}{2}\tilde{\Theta}_3 \\ \frac{19}{2} + \frac{19}{2}\tilde{\Theta}_3 \\ -\frac{19}{2} - \frac{19}{2}\tilde{\Theta}_3 \end{bmatrix} \quad (6.151.)$$

Točko (6.151.) nato vstavimo v funkcijo (6.146.), iz česar sledi:



$$\begin{aligned}
f(\underline{x}_3) &= \left(-\frac{1}{2} - \frac{1}{2}\tilde{\Theta}_3\right)^2 + \left(-\frac{1}{2} - \frac{1}{2}\tilde{\Theta}_3\right) \cdot \left(-\frac{19}{2} - \frac{19}{2}\tilde{\Theta}_3\right) + \frac{1}{2} \cdot \left(-\frac{19}{2} - \frac{19}{2}\tilde{\Theta}_3\right)^2 + \\
&+ \left(-\frac{1}{2} - \frac{1}{2}\tilde{\Theta}_3\right) + 10 \cdot \left(-\frac{19}{2} - \frac{19}{2}\tilde{\Theta}_3\right) = \\
&= \frac{1}{4} \cdot (1 + \tilde{\Theta}_3)^2 - \frac{1}{2} \cdot (1 + \tilde{\Theta}_3) \cdot \left(-\frac{19}{2}\right) \cdot (1 + \tilde{\Theta}_3) + \frac{1}{2} \cdot \frac{19^2}{4} \cdot (1 + \tilde{\Theta}_3)^2 - \frac{1}{2} \cdot (1 + \tilde{\Theta}_3) - \quad (6.152.) \\
&- \frac{190}{2} \cdot (1 + \tilde{\Theta}_3) = \left[\frac{1}{4} + \frac{19}{4} + \frac{19^2}{2 \cdot 4}\right] \cdot (1 + \tilde{\Theta}_3)^2 - \left[\frac{1}{2} + \frac{190}{2}\right] \cdot (1 + \tilde{\Theta}_3) = \\
&= 50.125 \cdot (1 + \tilde{\Theta}_3)^2 - \frac{191}{2} \cdot (1 + \tilde{\Theta}_3)
\end{aligned}$$

Ekstrem funkcije  $f(\tilde{\Theta}_3)$  zopet poiščemo na analitičen način. V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije. Velja torej:

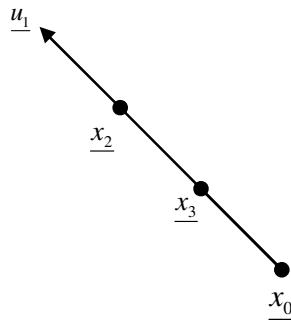
$$\begin{aligned}
\frac{df(\underline{x}_3)}{d\tilde{\Theta}_3} &= 50.125 \cdot 2 \cdot (1 + \tilde{\Theta}_3) - \frac{191}{2} = 0 \quad \Rightarrow \\
\Rightarrow \tilde{\Theta}_3^* &= \frac{191}{2 \cdot 2 \cdot 50.125} - 1 = 0.9526 - 1 = -0.047 \quad \Rightarrow \quad (6.153.) \\
\Rightarrow \underline{x}_3^* &= \begin{bmatrix} -\frac{1}{2} \cdot (1 + \tilde{\Theta}_3) \\ -\frac{19}{2} \cdot (1 + \tilde{\Theta}_3) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \cdot (1 - 0.047) \\ -\frac{19}{2} \cdot (1 - 0.047) \end{bmatrix} = \begin{bmatrix} -0.4763 \\ -9.0535 \end{bmatrix}
\end{aligned}$$

Pri izračunu točke  $\underline{x}_3^*$  in koeficienta  $\tilde{\Theta}_3$  omenimo, da obstaja tudi nekoliko drugačen način za izračun te točke, ki se v nekaterih primerih lahko izkaže celo kot lažji.

Če sliko 83 nekoliko natančneje opazujemo, lahko iz nje izluščimo točke, ki ležijo na 1. konjugiranem vektorju. Očitno lahko vektor  $\underline{x}_3^*$ , ki smo ga računali z izrazom (6.140.) oz (6.151.), izračunamo tudi kot vektor:

$$\underline{x}_3 = \underline{x}_0 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \tilde{\Theta}_3 \cdot \left( \begin{bmatrix} -\frac{1}{2} \\ -\frac{19}{2} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} -\frac{\tilde{\Theta}_3}{2} \\ -\frac{19}{2}\tilde{\Theta}_3 \end{bmatrix} \quad (6.154.)$$

Ta izraz bi nato vstavili v funkcijo (6.146.), dobili  $f(\tilde{\Theta}_3)$ , izvedli  $\frac{df(\tilde{\Theta}_3)}{d\tilde{\Theta}_3} = 0$ , vstavili dobljeni koeficient  $\tilde{\Theta}_3 = 0.9526$  v izraz (6.154.), ter tako dobili enak rezultat za  $\underline{x}_3$  kot smo ga dobili v izrazu (6.153.).



Slika 84: Točke na 1. konjugiranem vektorju iz slike 83

V nadaljevanju dobljeno točko  $\underline{x}_3^*$  iz izraza (6.153.) vstavimo v izraz (6.141.), pri čemer dobimo:

$$\underline{x}_4 = \underline{x}_3 + \tilde{\Theta}_4 \cdot \underline{e}_2 = \begin{bmatrix} -0.4763 \\ -9.0535 \end{bmatrix} + \tilde{\Theta}_4 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.4763 \\ -9.0535 + \tilde{\Theta}_4 \end{bmatrix} \quad (6.155.)$$

Točko (6.155.) nato vstavimo v funkcijo (6.146.), iz česar sledi:

$$f(\underline{x}_4) = f\left(\begin{bmatrix} -0.4763 \\ -9.0535 + \tilde{\Theta}_4 \end{bmatrix}\right) = (-0.4763)^2 + (-0.4763) \cdot (-9.0535 + \tilde{\Theta}_4) + \frac{1}{2} \cdot (-9.0535 + \tilde{\Theta}_4)^2 + (-0.4763) + 10 \cdot (-9.0535 + \tilde{\Theta}_4) \quad (6.156.)$$

Ekstrem funkcije  $f(\tilde{\Theta}_4)$  zopet poiščemo na analitičen način. V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije. Velja torej:

$$\begin{aligned} \frac{df(\underline{x}_4)}{d\tilde{\Theta}_4} &= -0.4763 + (-9.0535 + \tilde{\Theta}_4) + 10 = 0 \quad \Rightarrow \\ \Rightarrow \tilde{\Theta}_4^* &= 0.4763 + 9.0535 - 10 = -0.4702 \end{aligned} \quad (6.157.)$$

Dobljeni rezultat vstavimo v točko (6.155.) in dobimo:

$$\underline{x}_4^* = \begin{bmatrix} -0.4763 \\ -9.0535 - 0.4702 \end{bmatrix} = \begin{bmatrix} -0.4763 \\ -9.5237 \end{bmatrix} \quad (6.158.)$$

V nadaljevanju dobljeno točko  $\underline{x}_4^*$  iz izraza (6.158.) vstavimo v izraz (6.142.), pri čemer dobimo:

$$\begin{aligned}\underline{x}_5 &= \underline{x}_4 + \tilde{\Theta}_5 \cdot (\underline{x}_3 - \underline{x}_0) = \\ &= \begin{bmatrix} -0.4763 \\ -9.5237 \end{bmatrix} + \tilde{\Theta}_5 \cdot \left\{ \begin{bmatrix} -0.4763 \\ -9.0535 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} = \begin{bmatrix} -0.4763 - 0.4763\tilde{\Theta}_5 \\ -9.5237 - 9.0535\tilde{\Theta}_5 \end{bmatrix} = \\ &= \begin{bmatrix} -0.4763 \cdot (1 + \tilde{\Theta}_5) \\ -9.5237 - 9.0535\tilde{\Theta}_5 \end{bmatrix}\end{aligned}\quad (6.159.)$$

Točko (6.159.) nato vstavimo v funkcijo (6.146.), iz česar sledi:

$$\begin{aligned}f(\underline{x}_5) &= (-0.4763)^2 \cdot (1 + \tilde{\Theta}_5)^2 + 0.4763 \cdot (1 + \tilde{\Theta}_5) \cdot (9.5237 + 9.0535\tilde{\Theta}_5) + \\ &+ \frac{1}{2} \cdot (-9.5237 - 9.0535\tilde{\Theta}_5)^2 - 0.4763 \cdot (1 + \tilde{\Theta}_5) - 10 \cdot (9.5237 + 9.0535\tilde{\Theta}_5)\end{aligned}\quad (6.160.)$$

Ekstrem funkcije  $f(\tilde{\Theta}_5)$  zopet poiščemo na analitičen način. V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije. Velja torej:

$$\begin{aligned}\frac{df(\underline{x}_5)}{d\tilde{\Theta}_5} &= \\ &= (0.4763)^2 \cdot 2 \cdot (1 + \tilde{\Theta}_5) + 0.4763 \cdot \{9.5237 + 9.0535\tilde{\Theta}_5 + (1 + \tilde{\Theta}_5) \cdot 9.0535\} + \\ &+ \frac{2}{2} \cdot (-9.5237 - 9.0535\tilde{\Theta}_5) \cdot (-9.0535) - 0.4763 - 10 \cdot 9.0535 = \\ &= \tilde{\Theta}_5 \cdot [2 \cdot 0.4763^2 + 0.4763 \cdot (9.0535 \cdot 2) + (9.0535)^2] + \\ &+ [2 \cdot 0.4763^2 + 0.4763 \cdot \{9.5237 + 9.0535\} + 9.5237 \cdot 9.0535 - 0.4763 - 10 \cdot 9.0535] = \\ &= \tilde{\Theta}_5 \cdot 91.0439 + 4.5135 = 0 \quad \Rightarrow \tilde{\Theta}_5^* = \frac{-4.5135}{91.0439} = -0.0495 \\ &\Rightarrow \underline{x}_5^* = \begin{bmatrix} -0.4763 \cdot (1 - 0.0495) \\ -9.5237 - 9.0535 \cdot (-0.0495) \end{bmatrix} = \begin{bmatrix} -0.4527 \\ -9.0755 \end{bmatrix}\end{aligned}\quad (6.161.)$$

V nadaljevanju dobljeno točko  $\underline{x}_5^*$  iz izraza (6.161.) vstavimo v izraz (6.143.), pri čemer dobimo:

$$\begin{aligned}
\underline{x}^* = \underline{x}_3 + \tilde{\Theta}_6 \cdot (\underline{x}_5 - \underline{x}_3) &= \begin{bmatrix} -0.4763 \\ -9.0535 \end{bmatrix} + \tilde{\Theta}_6 \cdot \left\{ \begin{bmatrix} -0.4527 \\ -9.0755 \end{bmatrix} - \begin{bmatrix} -0.4763 \\ -9.0535 \end{bmatrix} \right\} = \\
&= \begin{bmatrix} -0.4763 + \tilde{\Theta}_6 \cdot (0.4763 - 0.4527) \\ -9.0535 + \tilde{\Theta}_6 \cdot (9.0535 - 9.0755) \end{bmatrix} = \\
&= \begin{bmatrix} -0.4763 + \tilde{\Theta}_6 \cdot 0.0236 \\ -9.0535 - \tilde{\Theta}_6 \cdot 0.022 \end{bmatrix} = \underline{x}_6
\end{aligned} \tag{6.162.}$$

Točko (6.162.) nato vstavimo v funkcijo (6.146.), iz česar sledi:

$$\begin{aligned}
f(\underline{x}_6) &= (-0.4763 + \tilde{\Theta}_6 \cdot 0.0236)^2 + (-0.4763 + \tilde{\Theta}_6 \cdot 0.0236) \cdot (-9.0535 - \tilde{\Theta}_6 \cdot 0.022) + \\
&+ \frac{1}{2} \cdot (-9.0535 - \tilde{\Theta}_6 \cdot 0.022)^2 + (-0.4763 + \tilde{\Theta}_6 \cdot 0.0236) + 10 \cdot (-9.0535 - \tilde{\Theta}_6 \cdot 0.022)
\end{aligned} \tag{6.163.}$$

Ekstrem funkcije  $f(\tilde{\Theta}_6)$  zopet poiščemo na analitičen način. V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije. Velja torej:

$$\begin{aligned}
\frac{df(\underline{x}_6)}{d\tilde{\Theta}_6} &= 2 \cdot (0.0236\tilde{\Theta}_6 - 0.4763) \cdot 0.0236 - 0.0236 \cdot (9.0535 + 0.022\tilde{\Theta}_6) + \\
&+ (0.0236\tilde{\Theta}_6 - 0.4763) \cdot (-0.022) + (-9.0535 - 0.022\tilde{\Theta}_6) \cdot (-0.022) + \\
&+ 0.0236 - 10 \cdot 0.022 = \\
&= \tilde{\Theta}_6 \cdot [2 \cdot 0.0236^2 - 2 \cdot (0.0236 \cdot 0.022) + 0.022^2] + \\
&+ [-2 \cdot 0.4763 \cdot 0.0236 - 0.0236 \cdot 9.0535 + 0.4763 \cdot 0.022 + \\
&+ 9.0535 \cdot 0.022 + 0.0236 - 10 \cdot 0.022] = \\
&= \tilde{\Theta}_6 \cdot 5.595 \cdot 10^{-4} - 0.22288 = 0 \quad \Rightarrow \quad \tilde{\Theta}_6^* = 398.37 \\
\Rightarrow \underline{x}^* &= \begin{bmatrix} -0.4763 + 398.37 \cdot 0.0236 \\ -9.0535 - 398.37 \cdot 0.022 \end{bmatrix} = \begin{bmatrix} 8.925 \\ -17.817 \end{bmatrix}
\end{aligned} \tag{6.164.}$$

Kot se izkaže, je pravi rezultat naslednji:

$$\underline{x}^* = \begin{bmatrix} 9 \\ -19 \end{bmatrix} \tag{6.165.}$$

do razlike v rezultatu pa je prišlo zaradi numeričnih napak pri zaokroževanju, ki smo jih delali pri »ročnem računanju«. Seveda bi bile numerične napake zaradi občutno večje natančnosti računanja pri izračunu z računalniškim programom bistveno manjše.

### 6.3.2.2 Metoda konjugiranih gradientov

Prvotno je bila uvedena za reševanje velikih linearnih sistemov enačb (metoda linearnih konjugiranih gradientov), kasneje pa se je izkazala kot učinkovita tudi pri reševanju nelinearnih optimizacijskih problemov (metoda nelinearnih konjugiranih gradientov) [32].

Najprej si bomo ogledali Metodo linearnih konjugiranih gradientov, nato pa še modifikacijo te metode za uporabo na nelinearnih sistemih.

#### I.) METODA LINEARNIH KONJUGIRANIH GRADIENTOV ZA REŠEVANJE SISTEMOV ENAČB TIPA $\underline{\underline{A}} \cdot \underline{x} = \underline{b}$

Imamo dan naslednji sistem linearnih enačb:

$$\begin{aligned}\underline{\underline{A}} \cdot \underline{x} &= \underline{b} \\ \underline{\underline{A}} \cdot \underline{x} - \underline{b} &= 0\end{aligned}\tag{6.166.}$$

kjer je matrika  $\underline{\underline{A}}$  simetrična in pozitivno definitna matrika dimenzije  $n \times n$ .

Izberimo si naslednjo funkcijo:

$$\Phi(\underline{x}) = \frac{1}{2} \underline{x}^T \cdot \underline{\underline{A}} \cdot \underline{x} - \underline{x}^T \cdot \underline{b}\tag{6.167.}$$

in dokažimo, da gradient te funkcije ustreza enačbi (6.166.):

$$\nabla \Phi(\underline{x}) = \frac{\partial}{\partial \underline{x}} \left[ \frac{1}{2} \underline{x}^T \cdot \underline{\underline{A}} \cdot \underline{x} - \underline{x}^T \cdot \underline{b} \right] = \underline{\underline{A}} \cdot \underline{x} - \underline{b} = \underbrace{\underline{r}(\underline{x})}_{\text{RESIDUAL}}\tag{6.168.}$$

Očitno je torej, da je iskanje vektorja  $\underline{x}$  pri sistemu  $\underline{\underline{A}} \cdot \underline{x} = \underline{b}$  enaka operacija, kot če bi minimizirali funkcijo  $\Phi(\underline{x})$  glede na  $\underline{x}$ . Rešitev bi namreč obakrat prišla enaka, saj izenačitev gradienta funkcije (6.167.) z 0 da enak izraz, kot je zapisan v enačbi (6.166). Kot je razvidno iz izraza (6.168.), smo gradient poimenovali residual, saj velja, da je  $\nabla \Phi(\underline{x})$  enak residualu sistema  $\underline{\underline{A}} \cdot \underline{x} = \underline{b}$ .

V nadaljevanju predpostavimo, da imamo dan tudi nek set konjugiranih vektorjev  $\{\underline{u}_1, \dots, \underline{u}_n\}$ . Zanje tako kot pri Powellovi metodi velja (glej izraz (6.105)):

$$\underline{u}_i^T \cdot \underline{\underline{A}} \cdot \underline{u}_j = 0, \quad i \neq j, \quad i, j = 1, \dots, n\tag{6.169.}$$

Denimo lahko rešitev  $\underline{x}_k$  v  $k$ -ti iteraciji podobno kot pri Powellovi metodi (glej izraza (6.107.) in (6.113.)) izrazimo v obliki naslednjega izraza:

$$\underline{x}_{k+1} = \underline{x}_k + \alpha_k \cdot \underline{u}_k \quad (6.170.)$$

kjer je  $\underline{x}_{k+1}$  rešitev za  $k+1$ . iteracijo,  $\underline{x}_k$  je rešitev za  $k$ . iteracijo, ter je  $\underline{u}_k$  konjugiran vektor, izračunan v  $k$ -ti iteraciji.

Poiščimo sedaj na analitičen način optimalen koeficient  $\alpha_k$ . Funkcija (6.167.) v  $k+1$ . iteraciji zavzame na osnovi (6.170.) naslednjo obliko:

$$\Phi(\underline{x}_{k+1}) = \Phi(\underline{x}_k + \alpha_k \cdot \underline{u}_k) \quad (6.171.)$$

Če jo odvajamo po  $\alpha_k$  in enačimo z 0, dobimo:

$$\frac{\partial \Phi(\underline{x}_{k+1})}{\partial \alpha} = \underbrace{\frac{\partial \Phi(\underline{x}_k + \alpha_k \cdot \underline{u}_k)}{\partial \underline{x}_{k+1}}}_{\text{VERIŽNO ODVAJANJE!}} \cdot \underbrace{\frac{\partial \underline{x}_{k+1}}{\partial \alpha_k}}_{\underline{u}_k} = \nabla \Phi^T(\underline{x}_{k+1}) \cdot \underline{u}_k = 0 \quad (6.172.)$$

Torej mora biti v optimumu  $\alpha^*$  takšen, da sta vektorja  $\nabla \Phi^T(\underline{x}_{k+1})$  in  $\underline{u}_k$  ortogonalna med seboj!

Na osnovi izrazov (6.168.) in (6.172.) dobimo [32]:

$$\nabla \Phi^T(\underline{x}_{k+1}) \cdot \underline{u}_k = 0 \quad (6.173.)$$

ozi.

$$\underline{r}_{k+1}^T \cdot \underline{u}_k = 0$$

kjer smo poenostavljeno označili:  $\underline{r}(\underline{x}_{k+1}) = \underline{r}_{k+1}$ . Ker pa je na osnovi izraza (6.168.)  $\underline{r}_{k+1} = \underline{A} \cdot \underline{x}_{k+1} - \underline{b}$ , sledi:

$$(\underline{A} \cdot \underline{x}_{k+1} - \underline{b})^T \cdot \underline{u}_k = 0 \quad (6.174.)$$

Če v izrazu (6.174.) upoštevamo izraz (6.170.), dobimo:

$$\begin{aligned} & [\underline{A} \cdot (\underline{x}_k + \alpha_k \cdot \underline{u}_k) - \underline{b}]^T \cdot \underline{u}_k = 0 \\ & [\underline{A} \cdot \underline{x}_k + \alpha_k \cdot \underline{A} \cdot \underline{u}_k - \underline{b}]^T \cdot \underline{u}_k = 0 \\ & \left[ \underbrace{\underline{A} \cdot \underline{x}_k - \underline{b}}_{\underline{r}_k} + \alpha_k \cdot \underline{A} \cdot \underline{u}_k \right]^T \cdot \underline{u}_k = 0 \end{aligned} \quad (6.175.)$$

$$\begin{aligned} [r_k + \alpha_k \cdot \underline{\underline{A}} \cdot \underline{\underline{u}}_k]^T \cdot \underline{\underline{u}}_k &= 0 \\ r_k^T \cdot \underline{\underline{u}}_k + \alpha_k \cdot \underline{\underline{u}}_k^T \cdot \underbrace{\underline{\underline{A}}^T}_{\underline{\underline{A}}} \cdot \underline{\underline{u}}_k &= 0 \end{aligned}$$

Tako dobimo optimalen koeficient, ki se glasi [32]:

$$\alpha_k^* = \frac{-r_k^T \cdot \underline{\underline{u}}_k}{\underline{\underline{u}}_k^T \cdot \underline{\underline{A}} \cdot \underline{\underline{u}}_k} \quad (6.176.)$$

Torej lahko analitično izrazimo optimalen koeficient (v optimumu) kot funkcijo residuala in konjugiranega vektorja.

Če izraz (6.176.) vstavimo v izraz (6.170.), dobimo:

$$\underline{\underline{x}}_{k+1} = \underline{\underline{x}}_k + \frac{-r_k^T \cdot \underline{\underline{u}}_k}{\underbrace{\underline{\underline{u}}_k^T \cdot \underline{\underline{A}} \cdot \underline{\underline{u}}_k}_{\alpha_k^*}} \cdot \underline{\underline{u}}_k$$

kjer je  $\underline{\underline{r}}_k = \underline{\underline{A}} \cdot \underline{\underline{x}}_k - \underline{\underline{b}}$  na osnovi izraza (6.168.). Na takšen način izražen vektor  $\underline{\underline{x}}_{k+1}$  bo gotovo šel proti pravi (optimalni rešitvi) sistema (6.166.) oz. pravemu optimumu funkcije (6.167), saj smo uporabili optimalno obliko koeficienta (6.176.).

Poglejmo si še, kako se določajo konjugirani vektorji  $\underline{\underline{u}}_k$ . Postavili jih bomo na naslednji način [32]:

$$\underline{\underline{u}}_k = -r_k + \beta_k \underline{\underline{u}}_{k-1} \quad (6.177.)$$

kjer je  $\underline{\underline{u}}_k$  konjugiran vektor v  $k$ -ti iteraciji,  $r_k$  je residual v  $k$ -ti iteraciji,  $\underline{\underline{u}}_{k-1}$  je konjugiran vektor v  $k-1$ . iteraciji,  $\beta_k$  pa je koeficient iteracijske enačbe konjugiranih vektorjev.

V skladu z izrazom (6.169.) gotovo velja naslednje:

$$\underline{\underline{u}}_{k-1}^T \cdot \underline{\underline{A}} \cdot \underline{\underline{u}}_k = 0 \quad (6.178.)$$

torej je vektor  $\underline{\underline{u}}_k$  konjugiran glede na vektor  $\underline{\underline{u}}_{k-1}$ .

Množimo (6.177) z leve z  $\underline{\underline{u}}_{k-1}^T \cdot \underline{\underline{A}}$  in na osnovi (6.178.) dobimo:

$$\underbrace{\underline{\underline{u}}_{k-1}^T \underline{\underline{A}} \underline{\underline{u}}_k}_0 = -\underline{\underline{u}}_{k-1}^T \underline{\underline{A}} r_k + \beta_k \underline{\underline{u}}_{k-1}^T \underline{\underline{A}} \underline{\underline{u}}_{k-1} \quad (6.179.)$$

Sledi:

$$\underline{u}_{k-1}^T \underline{A} \underline{r}_k = \beta_k \underline{u}_{k-1}^T \underline{A} \underline{u}_{k-1} \quad (6.180.)$$

Odtod pa lahko izrazimo koeficient iteracijske enačbe konjugiranih vektorjev [32]:

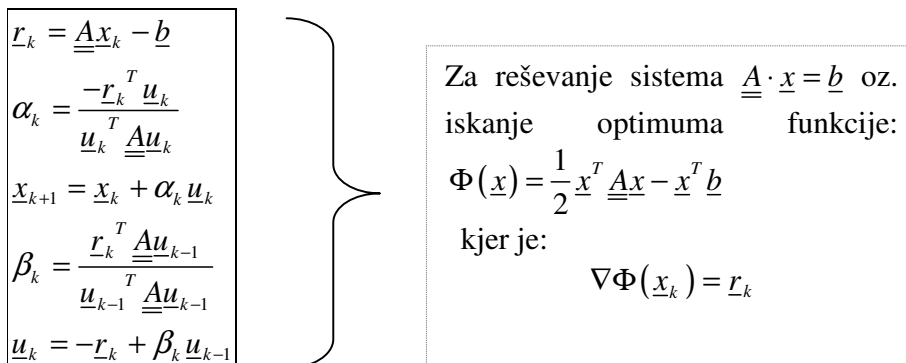
$$\beta_k = \frac{\underline{u}_{k-1}^T \underline{A} \underline{r}_k}{\underline{u}_{k-1}^T \underline{A} \underline{u}_{k-1}} = \frac{\underline{r}_k^T \underline{A} \underline{u}_{k-1}}{\underline{u}_{k-1}^T \underline{A} \underline{u}_{k-1}} \quad (6.181.)$$

kjer smo upoštevali, da imamo tako v števcu kot tudi imenovalcu skalar, ter je  $\underline{A}$  simetrična matrika.

Če vstavimo izraz (6.181.) v izraz (6.177.), dobimo iteracijsko enačbo za konjugirane vektorje v celoti izraženo z znanimi veličinami:

$$\underline{u}_k = -\underline{r}_k + \overbrace{\frac{\underline{r}_k^T \underline{A} \underline{u}_{k-1}}{\underline{u}_{k-1}^T \underline{A} \underline{u}_{k-1}}}^{\beta_k} \underline{u}_{k-1} \quad (6.182.)$$

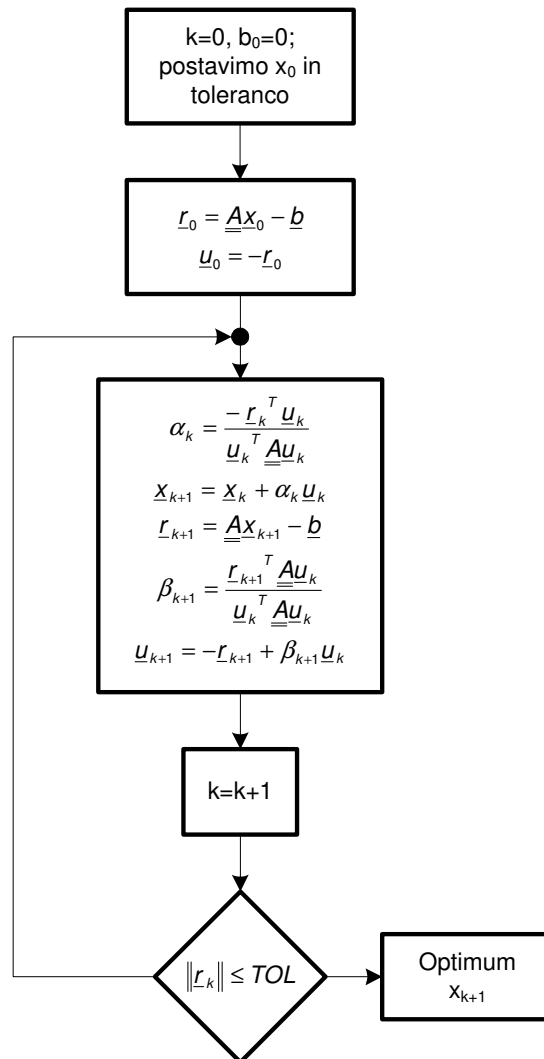
Slika 85 prikazuje sklop vseh glavnih rezultatov, dosedaj izpeljanih v tem poglavju [32].



Slika 85: Sklop nekaterih glavnih rezultatov pri metodi linearnih konjugiranih gradientov.

Slika 86 pa prikazuje dosedaj izpeljane glavne rezultate v obliki simboličnega algoritma oz. diagrama poteka [32].





Slika 86: Osnovni simbolični algoritem metode linearnih konjugiranih gradientov.

**Predelava algoritma v bolj praktično (ekonomično) obliko:**

V nadaljevanju bomo izpeljani algoritem pretvorili še v nekoliko bolj praktično oz. ekonomično obliko za uporabo na računalniku.

Na osnovi izrazov (6.168.), (6.170.) in (6.173.) lahko zapišemo [32]:

$$\begin{aligned}
 \underline{r}_{k+1} &= \underline{A}\underline{x}_{k+1} - \underline{b} = \underline{A}[\underline{x}_k + \alpha_k \underline{u}_k] - \underline{b} = \\
 &= \underbrace{\underline{A}\underline{x}_k - \underline{b}}_{\underline{r}_k} + \alpha_k \underline{A}\underline{u}_k
 \end{aligned}
 \tag{6.183.}$$

$$\underline{r}_{k+1} = \underline{r}_k + \alpha_k \underline{A}\underline{u}_k$$

S čimer smo uspeli dobiti bolj ekonomično obliko za residual, izražen v obliki iteracijske enačbe.

V nadaljevanju izpeljimo bolj ekonomično obliko še za  $\alpha_k$  in  $\beta_{k+1}$ :

Če izraz (6.177.) množimo z leve z  $\underline{r}_k^T$ , na osnovi izraza (6.173.) dobimo:

$$\begin{aligned} \underline{u}_k &= -\underline{r}_k + \beta_k \underline{u}_{k-1} \quad \left| \cdot \underline{r}_k^T \text{ z leve} \right. \\ \underline{r}_k^T \underline{u}_k &= -\underline{r}_k^T \underline{r}_k + \beta_k \underbrace{\underline{r}_k^T \underline{u}_{k-1}}_0 \\ \underline{r}_k^T \underline{u}_k &= -\underline{r}_k^T \underline{r}_k \end{aligned} \tag{6.184.}$$

Izraz (6.184.) upoštevamo v izrazu (6.176.) in dobimo [32]:

$$\alpha_k = \frac{-\underline{r}_k^T \underline{u}_k}{\underline{u}_k^T \underline{A} \underline{u}_k} = \frac{\underline{r}_k^T \underline{r}_k}{\underline{u}_k^T \underline{A} \underline{u}_k} \tag{6.185.}$$

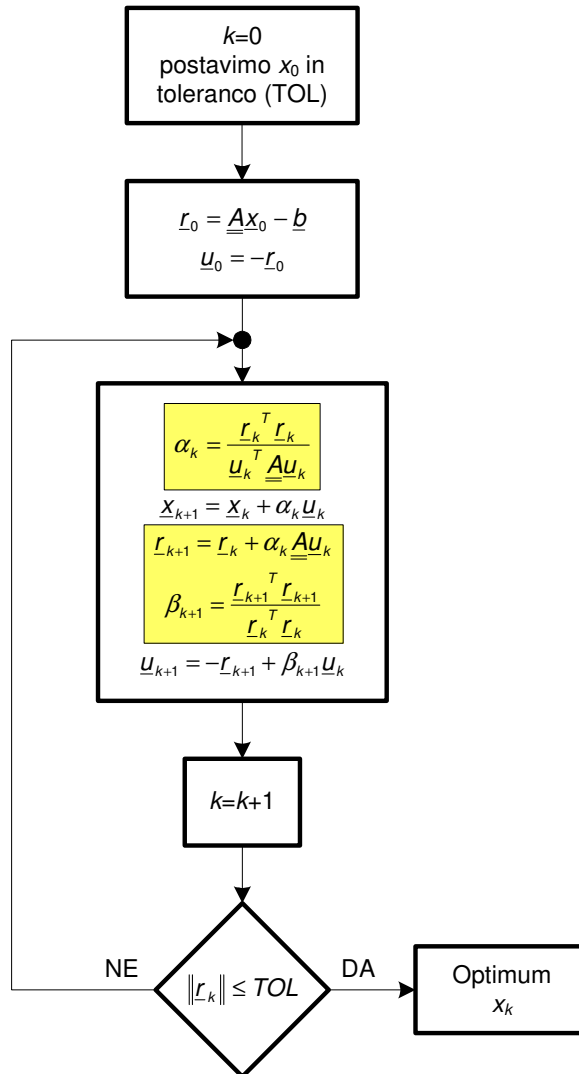
Predelajmo še koeficient  $\beta_k$  v bolj ekonomično obliko. Iz izraza (6.183.) na osnovi izrazov (6.173.) in (6.184.) sledi:

$$\begin{aligned} \underline{r}_{k+1} - \underline{r}_k &= \alpha_k \underline{A} \underline{u}_k \quad \Rightarrow \quad \underline{A} \underline{u}_k = \frac{\underline{r}_{k+1} - \underline{r}_k}{\alpha_k} \quad \left| \cdot \underline{r}_{k+1}^T \text{ z leve} \right. \\ \underline{r}_{k+1}^T \underline{A} \underline{u}_k &= \frac{\underline{r}_{k+1}^T \underline{r}_{k+1} - \underline{r}_{k+1}^T \underline{r}_k}{\alpha_k} \\ \underline{r}_{k+1}^T \underline{A} \underline{u}_k &= \frac{\underline{r}_{k+1}^T \underline{r}_{k+1}}{\alpha_k} \end{aligned} \tag{6.186.}$$

Če upoštevamo izraz (6.186.) v izrazu (6.181.), dobimo pri indeksu k+1 z upoštevanjem izraza (6.185.) naslednji rezultat [32]:

$$\begin{aligned} \beta_{k+1} &= \frac{\underline{r}_{k+1}^T \underline{r}_{k+1}}{\underbrace{\alpha_k \underline{u}_k^T \underline{A} \underline{u}_k}_{\underline{r}_k^T \underline{r}_k}} \\ \beta_{k+1} &= \frac{\underline{r}_{k+1}^T \underline{r}_{k+1}}{\underline{r}_k^T \underline{r}_k} \end{aligned} \tag{6.187.}$$

Ko smo uspeli predelati koeficienta  $\alpha_k$  in  $\beta_{k+1}$  v bolj ekonomično obliko, pri čemer smo dobili izraza (6.185.) in (6.187.), lahko slednja skupaj z bolj ekonomično obliko za residual v izrazu (6.183.) upoštevamo v osnovnem simboličnem algoritmu na sliki 86. S tem dobimo bolj ekonomičen simboličen algoritem, prikazan na sliki 87 [32].



Slika 87: Bolj ekonomičen simbolični algoritem metode linearnih konjugiranih gradientov.

## II.) METODA KONJUGIRANIH GRADIENTOV ZA REŠEVANJE NELINEARNIH OPTIMIZACIJSKIH PROBLEMOV

Kot se izkaže, je metoda konjugiranih gradientov izredno učinkovita tudi pri reševanju nelinearnih optimizacijskih problemov. Tedaj govorimo o metodi nelinearnih

konjugiranih gradientov, ki ji večkrat rečemo tudi Fletcher – Reevesova metoda nelinearnih konjugiranih gradientov [32].

### **FLETCHER – REEVESOVA METODA NELINEARNIH KONJUGIRANIH GRADIENTOV**

Fletcher in Reeves sta pokazala, da je algoritem metode linearnih konjugiranih gradientov z manjšimi modifikacijami mogoče uporabiti tudi za iskanje optimumov nelinearnih konveksnih kvadratnih funkcij, generalnih konveksnih funkcij, ali celo generalnih splošnih (poljubnih) nelinearnih funkcij  $f(\underline{x})$ . Pri tem je treba vpeljati 2 preprosti modifikaciji glede na algoritem metode linearnih konjugiranih gradientov na sliki 87 [32]:

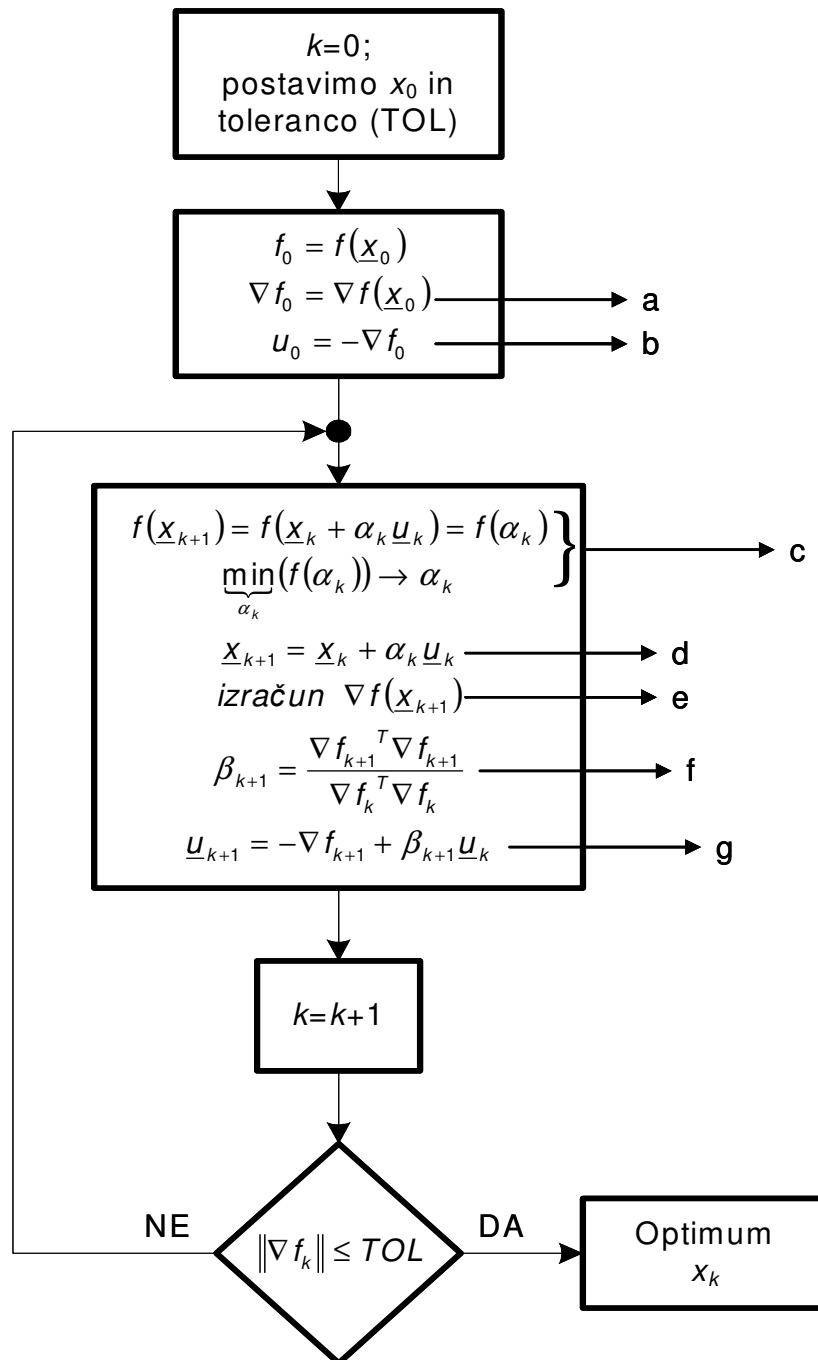
a.)  $\underline{r}_k$  se nadomesti z gradientom  $\nabla f_k = \nabla f(\underline{x}_k)$ .

b.) Namesto izraza  $\alpha_k = \frac{\underline{r}_k^T \cdot \underline{r}_k}{\underline{u}_k^T \cdot \underline{A} \cdot \underline{u}_k}$  se  $\alpha_k$  v vsaki iteraciji k računa z enodimenzionalno optimizacijo vzdolž smeri  $\underline{u}_k$ .

Tako dobimo Fletcher-Reevesov algoritem metode nelinearnih konjugiranih gradientov, ki je prikazan na sliki 88 [32].

Če primerjamo metodo najstrmejšega spusta s Fletcher - Reevesovo metodo, se izkaže, da lahko s slednjo znatno izboljšamo konvergenčne karakteristike (hitrost konvergence) metode najstrmejšega spusta.

Metoda konjugiranih gradientov pa se od metode konjugiranih smeri (Powellove metode) razlikuje v tem, da so pri prvi tudi gradienti vpleteni v postopku, delita pa si dobre konvergenčne lastnosti, saj vemo, da Powellova metoda potrebuje le  $n$  iteracij za dosego optimuma (pri nelinearnih kvadratnih funkcijah). Torej, v splošnem je Fletcher-Reevesova metoda vsekakor mnogo hitrejša od metode najstrmejšega spusta, saj gotovo doseže optimum po  $n$  iteracijah, kjer je  $n$  dimenzija prostora (število odločitvenih spremenljivk).



Slika 88: Fletcher-Reevesov algoritem metode nelinearnih konjugiranih gradientov.

**Primer:**

Minimizirajte naslednjo funkcijo z uporabo Fletcher-Reeves metode nelinearnih konjugiranih gradientov:

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1 \cdot x_2 + x_2^2, \quad \underline{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (6.188.)$$

Analitični izračun pravega minimuma se glasi:

$$\nabla \underline{f}(x_1, x_2) = 0$$

sledi:

$$\left. \begin{array}{l} 1 + 4x_1 + 2x_2 = 0 \\ -1 + 2x_1 + 2x_2 = 0 \end{array} \right\} -$$

sledi:

$$2x_1 + 2 = 0 \quad \Rightarrow x_{1A}^* = -1 \quad (6.189.)$$

sledi:

$$1 + 4 \cdot (-1) + 2x_2 = 0 \quad \Rightarrow x_{2A}^* = \frac{3}{2}$$

sledi:

$$\underline{x}_A^* = (x_{1A}^*, x_{2A}^*) = \left( -1, \frac{3}{2} \right)$$

**Rešitev s Fletcher-Reevesovo metodo** pa je naslednja (glej bločno shemo na sliki 88):

$k = 0, 1$ . ITERACIJA:

Gradient funkcije (6.188.) v poljubni točki je:

$$\nabla \underline{f}(x_1, x_2) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \quad (6.190.)$$

Realizacija izraza a na sliki 88:

$$\nabla \underline{f}(x_0) = \nabla \underline{f}(0, 0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \nabla \underline{f}_0 \quad (6.191.)$$

Realizacija izraza b na sliki 88:

$$\underline{u}_0 = -\nabla \underline{f}_0 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (6.192.)$$

Realizacija izraza c na sliki 88:

$$\underline{x}_1 = \underline{x}_0 + \alpha_0 \cdot \underline{u}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \alpha_0 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\alpha_0 \\ \alpha_0 \end{bmatrix} \quad (6.193.)$$

$$f(\underline{x}_1) = f(-\alpha_0, \alpha_0) = (-\alpha_0) - \alpha_0 + 2 \cdot (-\alpha_0)^2 + 2 \cdot (-\alpha_0) \cdot \alpha_0 + \alpha_0^2 = -2 \cdot \alpha_0 + \alpha_0^2$$

$$f(\alpha_0) = \alpha_0^2 - 2 \cdot \alpha_0$$

$$\frac{df}{d\alpha_0} = 2\alpha_0 - 2 = 0 \quad \Rightarrow \quad \alpha_0^* = 1$$

Realizacija izraza d na slici 88:

$$\underline{x}_1 = \begin{bmatrix} -\alpha_0^* \\ \alpha_0^* \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (6.194.)$$

Realizacija izraza e na slici 88:

$$\nabla f(\underline{x}_1) = \nabla f(-1, 1) = \begin{bmatrix} 1 + 4 \cdot (-1) + 2 \cdot 1 \\ -1 + 2 \cdot (-1) + 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \nabla f_1 \quad (6.195.)$$

Realizacija izraza f na slici 88:

$$\beta_1 = \frac{\nabla f_1^T \cdot \nabla f_1}{\nabla f_0^T \cdot \nabla f_0} = \frac{\begin{bmatrix} -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix}}{\begin{bmatrix} 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}} = \frac{2}{2} = 1 \quad (6.196.)$$

Realizacija izraza g na slici 88:

$$\underline{u}_1 = -\nabla f_1 + \beta_1 \cdot \underline{u}_0 = -\begin{bmatrix} -1 \\ -1 \end{bmatrix} + 1 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad (6.197.)$$

1. iteracija je končana, gremo na 2.iteraciju.

$k = 1$ , 2. ITERACIJA:

Realizacija izraza c na slici 88:

$$\underline{x}_2 = \underline{x}_1 + \alpha_1 \cdot \underline{u}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \alpha_1 \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 + 2 \cdot \alpha_1 \end{bmatrix}$$

$$\begin{aligned} f(\underline{x}_2) &= (-1) - (1 + 2\alpha_1) + 2 \cdot (-1)^2 + 2 \cdot (-1) \cdot (1 + 2\alpha_1) + (1 + 2\alpha_1)^2 = \\ &= -2 - 2\alpha_1 + 2 - 2 - 4\alpha_1 + (1 + 2\alpha_1)^2 = \\ &= -2 - 6\alpha_1 + (1 + 2\alpha_1)^2 = f(\alpha_1) \end{aligned} \quad (6.198.)$$

$$\frac{df(\alpha_1)}{d\alpha_1} = -6 + 2 \cdot (1 + 2\alpha_1^*) \cdot 2 = 0 \quad \Rightarrow \quad \alpha_1^* = \left(\frac{6}{4} - 1\right) \cdot \frac{1}{2} = \frac{2}{4} \cdot \frac{1}{2} \quad \Rightarrow$$

$$\alpha_1^* = \frac{1}{4}$$

Realizacija izraza d na sliki 88:

$$\underline{x}_2 = \begin{bmatrix} -1 \\ 1 + 2 \cdot \frac{1}{4} \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} \quad (6.199.)$$

Če primerjamo rezultat (6.199.) z rezultatom (6.189.), vidimo, da smo po  $n=2$  iteracijah (dimenzija prostora!) že prišli v optimum, kar je bilo tudi za pričakovati, glede na lastnosti Fletcher-Reevesove metode in kvadratično naravo nelinearne funkcije (6.188.).

V nadaljevanju vseeno še preverimo, če smo res že prišli v optimum.

Realizacija izraza e na sliki 88:

$$\nabla f(\underline{x}_2) = \nabla f\left(-1, \frac{3}{2}\right) = \begin{bmatrix} 1 + 4 \cdot (-1) + 2 \cdot \frac{3}{2} \\ -1 + 2 \cdot (-1) + 2 \cdot \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (6.200.)$$

Izstopni pogoj je očitno izpolnjen, saj  $\nabla f(\underline{x}_2) = 0 \leq \text{TOL}$ , končni rezultat pa je:

$$\underline{x}^* = \underline{x}_2 = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} \quad (6.201)$$

Na koncu tega primera še opozorimo, da smo smo ekstrem funkcije  $f(\alpha_k), k=0,1$  poiskali na analitičen način. V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije.

### 6.3.2.3 Quasi-Newtonove metode

V Newtonovi metodi smo določili smer spreminjanja vektorja odločitvenih spremenljivk (smer novega koraka) tako, da smo množili inverzno Hessejevo matriko in gradient funkcije. Pri metodi najstrmejšega spusta si lahko predstavljamo, da smo v ta namen množili enotino matriko in gradient funkcije. V splošnem pa bi lahko smer spreminjanja vektorja odločitvenih spremenljivk določili z množenjem gradienta funkcije s kako drugo



matriko, ki jo imenujemo takoimenovana metrična matrika. Slednja je običajno simetrična in, če je tudi pozitivno definitna, načeloma postopek konvergira.

Za invertiranje Hessove matrike pri Newtonovi metodi se včasih izkaže, da je dokaj delikatna in računsko potratna operacija. Zato se v takšnih primerih zdi smiselno, da bi se invertiranju Hessove matrike izognili. Tedaj se zdi uporaba Quasi – Newtonovih metod smiselna alternativna rešitev, saj le-te, kot bomo videli, invertiranja matrike sploh ne uporabljajo, pač pa uporabljajo mehanizem metričnih matrik.

V splošnem lahko v  $(i+1.)$  iteraciji spremembo vektorja odločitvenih spremenljivk  $\Delta \underline{x}_i = \underline{x}_{i+1} - \underline{x}_i$  določimo z izrazom [13]:

$$\Delta \underline{x}_i = -\alpha_i \cdot \underline{G}_i \cdot \nabla f(\underline{x}_i) \quad (6.202.)$$

kjer je  $\alpha_i$  neko število, pri katerem je funkcija  $f(\underline{x}_i)$  v izbrani smeri minimalna,  $\underline{G}_i$  pa je metrična matrika. Ker se slednja od iteracije do iteracije spreminja, pravimo takšnim postopkom spremenljive metrične metode.

Izraz (6.202.) bomo kasneje uporabili v postopku, še prej pa moramo izpeljati matriko  $\Delta \underline{G}_i$ .

V ta namen najprej razvijmo dano funkcijo v Taylorjevo vrsto (3 členi!) v okolici točke  $\underline{x}_0$ :

$$f(\underline{x}_{i+1}) = f(\underline{x}_i) + [\nabla f(\underline{x}_i)]^T \Delta \underline{x}_i + \frac{1}{2} (\Delta \underline{x}_i)^T \underline{H}[f(\underline{x}_i)] \Delta \underline{x}_i \quad (6.203.)$$

Izrazimo gradient Taylorjeve vrste (6.203.):

$$\begin{aligned} \nabla f(\underline{x}_{i+1}) &= \frac{\partial}{\partial \Delta \underline{x}_i} \left[ f(\underline{x}_i) + [\nabla f(\underline{x}_i)]^T \Delta \underline{x}_i + \frac{1}{2} (\Delta \underline{x}_i)^T \underline{H}[f(\underline{x}_i)] \Delta \underline{x}_i \right] = \\ &= 0 + \nabla f(\underline{x}_i) + \frac{2}{2} \underline{H}[f(\underline{x}_i)] \Delta \underline{x}_i \end{aligned} \quad (6.204.)$$

Dobimo torej relacijo:

$$\begin{aligned} \nabla f(\underline{x}_{i+1}) &= \nabla f(\underline{x}_i) + \underline{H}[f(\underline{x}_i)] \Delta \underline{x}_i \\ \underbrace{\nabla f(\underline{x}_{i+1}) - \nabla f(\underline{x}_i)}_{\Delta g(\underline{x}_i)} &= +\underline{H}[f(\underline{x}_i)] \Delta \underline{x}_i \end{aligned} \quad (6.205.)$$

$$\Delta \underline{x}_i = \underline{\underline{H}}^{-1} [f(\underline{x}_i)] \Delta g(\underline{x}_i)$$

kjer je  $g(\underline{x}_i)$  gradient,  $\Delta g(\underline{x}_i)$  je sprememba gradienta,  $\Delta \underline{x}_i$  pa sprememba vektorja odločitvenih spremenljivk (sprememba rešitve), pri čemer veljajo tudi relacije:

$$\Delta g(\underline{x}_i) = \nabla f(\underline{x}_{i+1}) - \nabla f(\underline{x}_i) \quad (6.206.)$$

$$\nabla f(\underline{x}_{i+1}) = \underbrace{\nabla f(\underline{x}_i)}_{g(\underline{x}_i)} + \Delta g(\underline{x}_i)$$

Nadomestimo sedaj v izrazu (6.205.) inverzno Hessovo matriko  $\underline{\underline{H}}^{-1}$  z neko metrično matriko :

$$\underline{\underline{G}}_{i+1} = \underline{\underline{G}}_i + \Delta \underline{\underline{G}}_i \quad (6.207.)$$

kar nam da izraz [13]:

$$\Delta \underline{x}_i = \underline{\underline{G}}_{i+1} \cdot \Delta g(\underline{x}_i) = [\underline{\underline{G}}_i + \Delta \underline{\underline{G}}_i] \Delta g(\underline{x}_i) \quad (6.208.)$$

$$\Delta \underline{x}_i = \underline{\underline{G}}_i \cdot \Delta g(\underline{x}_i) + \Delta \underline{\underline{G}}_i \Delta g(\underline{x}_i)$$

$$\Delta \underline{\underline{G}}_i \Delta g(\underline{x}_i) = \Delta \underline{x}_i - \underline{\underline{G}}_i \cdot \Delta g(\underline{x}_i)$$

V nadaljevanju bi želeli v izrazu (6.208.) izraziti le  $\Delta \underline{\underline{G}}_i$ . Torej se moramo vektorja  $\Delta g(\underline{x}_i)$  na levi strani izraza nekako znebiti, vendar se ne da kar deliti z njim, pač pa moramo izraz ustrezno še prevesti na skalarje.

Da se dokazati, da lahko preidemo v naslednjo obliko (če se znebimo  $\Delta g(\underline{x}_i)$  na levi strani izraza (6.208.)) [13]:

$$\Delta \underline{\underline{G}}_i = \frac{\Delta \underline{x}_i \cdot \underline{y}^T}{\underline{y}^T \cdot \Delta \underline{g}_i} - \frac{\underline{\underline{G}}_i \cdot \Delta \underline{g}_i \cdot \underline{z}^T}{\underline{z}^T \cdot \Delta \underline{g}_i} \quad (6.209.)$$

kar je takoimenovan **Davidonov izraz**, pri čemer je  $\Delta \underline{g}_i = \Delta g(\underline{x}_i)$  in sta  $\underline{y}$  in  $\underline{z}$  takšna vektorja, ki izpolnjujeta naslednji pogoj [13]:

$$\begin{aligned} \underline{y}^T \Delta \underline{g}_i &= a \\ \underline{z}^T \Delta \underline{g}_i &= b \end{aligned} \quad (6.210.)$$

kjer sta a in b poljubni realni konstanti.

Z različno izbiro vektorjev  $\underline{y}$  in  $\underline{z}$  lahko dobimo različne tipe metričnih (poimenovanih tudi **Quasi - Newtonovih**) metod!

V nadaljevanju si bomo pogledali dve takšni metodi:

- Davidon – Fletcher – Powellovo (**DFP**) Quasi – Newtonovo metodo, ter
- Broyden – Fletcher – Goldfarb – Shannovo (**BFGS**) Quasi – Newtonovo metodo.

### DAVIDON – FLETCHER – POWELLOVA METODA (DFP)

Fletcher in Powell sta v splošen DAVIDONOV IZRAZ (6.209.) vstavila naslednja vektorja [13]:

$$\begin{aligned}\underline{y} &= \Delta \underline{x}_i \\ \underline{z} &= \underline{G}_i \Delta \underline{g}_i\end{aligned}\quad (6.211.)$$

S tem izraz (6.209) preide v naslednjo obliko [13]:

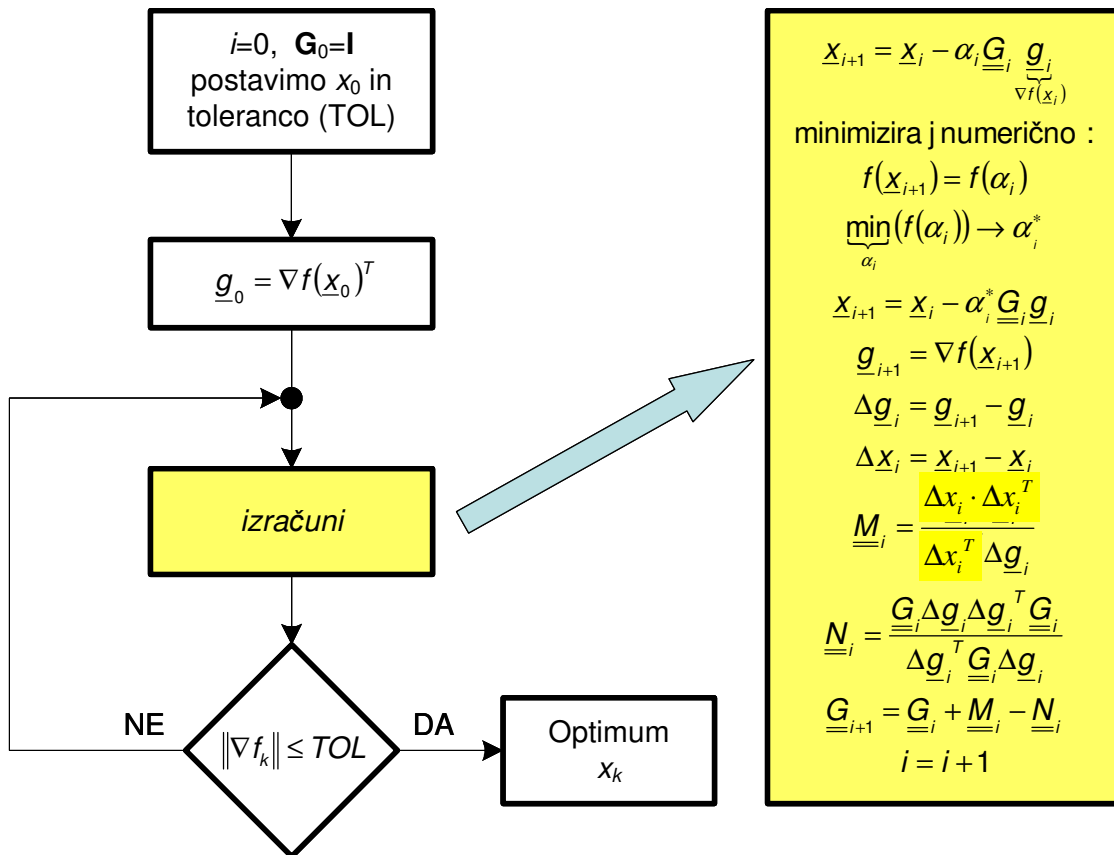
$$\Delta \underline{G}_i = \frac{\Delta \underline{x}_i \Delta \underline{x}_i^T}{\Delta \underline{x}_i^T \Delta \underline{g}_i} - \frac{\underline{G}_i \Delta \underline{g}_i (\underline{G}_i \Delta \underline{g}_i)^T}{(\underline{G}_i \Delta \underline{g}_i)^T \Delta \underline{g}_i} = \frac{\Delta \underline{x}_i \Delta \underline{x}_i^T}{\Delta \underline{x}_i^T \Delta \underline{g}_i} - \frac{\underline{G}_i \Delta \underline{g}_i \Delta \underline{g}_i^T \overbrace{\underline{G}_i^T}^{\underline{G}_i}}{\Delta \underline{g}_i^T \overbrace{\underline{G}_i}^{\underline{G}_i} \Delta \underline{g}_i}\quad (6.212.)$$

$$\Delta \underline{G}_i = \underbrace{\frac{\Delta \underline{x}_i \Delta \underline{x}_i^T}{\Delta \underline{x}_i^T \Delta \underline{g}_i}}_{\underline{M}_i} - \underbrace{\frac{\underline{G}_i \Delta \underline{g}_i \Delta \underline{g}_i^T \underline{G}_i}{\Delta \underline{g}_i^T \underline{G}_i \Delta \underline{g}_i}}_{\underline{N}_i} = \underline{M}_i - \underline{N}_i$$

kjer smo tudi upoštevali, da je  $\underline{G}_i$  simetrična matrika. S tem smo določili spremembo metrične matrike  $\underline{G}_i$  v  $i+1$ .iteraciji. Matrika  $\underline{G}_{i+1}$  v  $i+1$ .iteraciji pa se na osnovi (6.207.) in (6.212.) glasi [13]:

$$\begin{aligned}\underline{G}_{i+1} &= \underline{G}_i + \Delta \underline{G}_i = \underline{G}_i + \underline{M}_i - \underline{N}_i, \quad \text{kjer sta} \\ \underline{M}_i &= \frac{\Delta \underline{x}_i \Delta \underline{x}_i^T}{\Delta \underline{x}_i^T \Delta \underline{g}_i}, \quad \underline{N}_i = \frac{\underline{G}_i \Delta \underline{g}_i \Delta \underline{g}_i^T \underline{G}_i}{\Delta \underline{g}_i^T \underline{G}_i \Delta \underline{g}_i}\end{aligned}\quad (6.213.)$$

Slika 89 prikazuje simbolični algoritem DFP metode, ki temelji na vseh izpeljanih izrazih od (6.202.) do (6.213.) [13].



Slika 89: Simbolični algoritem Quasi-Newtonove DFP metode

Princip delovanja DFP metode bomo bolj osvetlili na naslednjem primeru.

**Primer:**

S pomočjo DFP metode poiščite minimum naslednje funkcije:

$$f(x_1, x_2) = x_1^2 - 2x_1x_2 + 2x_2^2 - x_1 - 2x_2, \quad (6.214.)$$

pri čemer je  $\underline{x}_0 = [1 \quad 1.5]^T$ .

**Rešitev:**

V nadaljevanju bomo natančno sledili posameznim blokom simboličnega algoritma na sliki 89.

Najprej izberemo začetno metrično matriko:

$$\underline{G}_0 = \underline{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (6.215.)$$

Gremo na 1. iteracijo ( $i = 0$ ).

Najprej je potrebno izračunati gradient v začetni točki:

$$\underline{g}_0 = \nabla f(\underline{x}_0) = \left. \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} \right|_{\underline{x}=\underline{x}_0} = \left. \begin{bmatrix} 2x_1 - 2x_2 - 1 \\ -2x_1 + 4x_2 - 2 \end{bmatrix} \right|_{\underline{x}=\underline{x}_0} = \begin{bmatrix} 2 \cdot 1 - 2 \cdot 1.5 - 1 \\ -2 \cdot 1 + 4 \cdot 1.5 - 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \quad (6.216.)$$

Nato je potrebno izračunati nov vektor odločitvenih spremenljivk:

$$\underline{x}_1 = \underline{x}_0 - \alpha_0 \cdot \underline{G}_0 \cdot \underline{g}_0 = \begin{bmatrix} 1 \\ 1.5 \end{bmatrix} - \alpha_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 + 2\alpha_0 \\ 1.5 - 2\alpha_0 \end{bmatrix} \quad (6.217.)$$

V točki (6.217.) funkcija (6.214.) zavzame vrednost:

$$f(\underline{x}_1) = (1 + 2\alpha_0)^2 - 2(1 + 2\alpha_0)(1.5 - 2\alpha_0) + 2(1.5 - 2\alpha_0)^2 - (1 + 2\alpha_0) - 2(1.5 - 2\alpha_0)$$

$$\begin{aligned} f(\underline{x}_1) &= (1 + 2\alpha_0)^2 - 2[1.5 - 2\alpha_0 + 3\alpha_0 - 4\alpha_0^2] + 2(1.5 - 2\alpha_0)^2 - 1 - 2\alpha_0 - 3 + 4\alpha_0 = \quad (6.218.) \\ &= (1 + 2\alpha_0)^2 + 2(1.5 - 2\alpha_0)^2 + 8\alpha_0^2 - 2\alpha_0 - 3 - 4 + 2\alpha_0 = \\ &= (1 + 2\alpha_0)^2 + 2(1.5 - 2\alpha_0)^2 + 8\alpha_0^2 - 7 \end{aligned}$$

in postane seveda funkcija ene spremenljivke – koeficienta  $\alpha_0$ . Slednji je optimalen tedaj, ko funkcija (6.218.) zavzame pri njem minimalno vrednost. Poiščimo  $\alpha_0^*$  na analitičen način:

$$\begin{aligned} \frac{df(\underline{x}_1)}{d\alpha_0} &= 2(1 + 2\alpha_0) \cdot 2 + 4(1.5 - 2\alpha_0)(-2) + 16\alpha_0 = 4 + 8\alpha_0 - 8(1.5 - 2\alpha_0) + 16\alpha_0 = \\ &= -8 + 40\alpha_0 = 0 \quad \Rightarrow \alpha_0^* = \frac{8}{40} = \frac{1}{5} = 0.2 \quad (6.219.) \end{aligned}$$

V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije.

Če izraz (6.219.) vstavimo v izraz (6.217.), dobimo:

$$\underline{x}_1 = \begin{bmatrix} 1 + 2 \cdot 0.2 \\ 1.5 - 2 \cdot 0.2 \end{bmatrix} = \begin{bmatrix} 1.4 \\ 1.1 \end{bmatrix} \quad (6.220.)$$

Gradient funkcije (6.214.) v točki (6.220.) se glasi:

$$\underline{g}_1 = \nabla f(\underline{x}_1) = \left[ \begin{array}{c} 2x_1 - 2x_2 - 1 \\ -2x_1 + 4x_2 - 2 \end{array} \right] \Big|_{\underline{x}=\underline{x}_1} = \left[ \begin{array}{c} 2 \cdot 1.4 - 2 \cdot 1.1 - 1 \\ -2 \cdot 1.4 + 4 \cdot 1.1 - 2 \end{array} \right] = \left[ \begin{array}{c} -0.4 \\ -0.4 \end{array} \right] \quad (6.221.)$$

Spremembo gradienta pa dobimo, če tvorimo razliko med izrazoma (6.221.) in (6.216.):

$$\Delta \underline{g}_0 = \underline{g}_1 - \underline{g}_0 = \left[ \begin{array}{c} -0.4 \\ -0.4 \end{array} \right] - \left[ \begin{array}{c} -2 \\ 2 \end{array} \right] = \left[ \begin{array}{c} 1.6 \\ -2.4 \end{array} \right] \quad (6.222.)$$

V nadaljevanju je potrebno v skladu s sliko 89 izračunati:

$$\underline{\underline{M}}_0 = \frac{\Delta \underline{x}_0 \Delta \underline{x}_0^T}{\Delta \underline{x}_0^T \Delta \underline{g}_0} \quad (6.223.)$$

Seveda pa je potrebno najprej izračunati vektor  $\Delta \underline{x}_0$ , pri čemer od vektorja (6.220.) odštejemo začetno točko:

$$\Delta \underline{x}_0 = \underline{x}_1 - \underline{x}_0 = \left[ \begin{array}{c} 1.4 \\ 1.1 \end{array} \right] - \left[ \begin{array}{c} 1 \\ 1.5 \end{array} \right] = \left[ \begin{array}{c} 0.4 \\ -0.4 \end{array} \right] \quad (6.224.)$$

Če izraza (6.224.) in (6.222.) vstavimo v izraz (6.223.), dobimo:

$$\underline{\underline{M}}_0 = \frac{\left[ \begin{array}{c} 0.4 \\ -0.4 \end{array} \right] \left[ \begin{array}{cc} 0.4 & -0.4 \end{array} \right]}{\left[ \begin{array}{cc} 0.4 & -0.4 \end{array} \right] \left[ \begin{array}{c} 1.6 \\ -2.4 \end{array} \right]} = \frac{\left[ \begin{array}{cc} 0.4^2 & -0.4^2 \\ -0.4^2 & 0.4^2 \end{array} \right]}{0.4 \cdot 1.6 + 0.4 \cdot 2.4} = \frac{0.4^2 \left[ \begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array} \right]}{1.6} = \left[ \begin{array}{cc} 0.1 & -0.1 \\ -0.1 & 0.1 \end{array} \right] \quad (6.225.)$$

V nadaljevanju je potrebno v skladu s sliko 89 izračunati:

$$\begin{aligned} \underline{\underline{N}}_0 &= \frac{\overset{I}{\underline{\underline{G}}}_0 \Delta \underline{g}_0 \Delta \underline{g}_0^T \overset{I}{\underline{\underline{G}}}_0}{\Delta \underline{g}_0^T \underline{\underline{G}}_0 \Delta \underline{g}_0} = \frac{\left[ \begin{array}{c} 1.6 \\ -2.4 \end{array} \right] \left[ \begin{array}{cc} 1.6 & -2.4 \end{array} \right]}{\left[ \begin{array}{cc} 1.6 & -2.4 \end{array} \right] \cdot \underline{\underline{I}} \cdot \left[ \begin{array}{c} 1.6 \\ -2.4 \end{array} \right]} = \frac{\left[ \begin{array}{cc} 1.6^2 & -1.6 \cdot 2.4 \\ -2.4 \cdot 1.6 & 2.4^2 \end{array} \right]}{1.6^2 + 2.4^2} = \\ &= \left[ \begin{array}{cc} 2.56 & -3.84 \\ -3.84 & 5.76 \end{array} \right] \frac{1}{8.32} = \left[ \begin{array}{cc} 0.307 & -0.461 \\ -0.461 & 0.692 \end{array} \right] \end{aligned} \quad (6.226.)$$

kjer smo uporabili tudi izraza (6.215.) in (6.222.).

V nadaljevanju je potrebno v skladu s sliko 89 na osnovi izrazov (6.215.), (6.225.) in (6.226.) izračunati še:

$$\begin{aligned} \underline{\underline{G}}_1 &= \underline{\underline{G}}_0 + \underline{\underline{M}}_0 - \underline{\underline{N}}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.1 & -0.1 \\ -0.1 & 0.1 \end{bmatrix} - \begin{bmatrix} 0.307 & -0.461 \\ -0.461 & 0.692 \end{bmatrix} = \\ &= \begin{bmatrix} 0.793 & 0.361 \\ 0.361 & 0.408 \end{bmatrix} \end{aligned} \quad (6.227.)$$

Na tej točki je algoritem končal s 1. iteracijo. Zato gremo na 2. iteracijo ( $i = 1$ ):

Najprej je potrebno izračunati nov vektor odločitvenih spremenljivk:

$$\begin{aligned} \underline{x}_2 &= \underline{x}_1 - \alpha_1 \underline{\underline{G}}_1 \underline{g}_1 = \begin{bmatrix} 1.4 \\ 1.1 \end{bmatrix} - \alpha_1 \begin{bmatrix} 0.793 & 0.361 \\ 0.361 & 0.408 \end{bmatrix} \begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix} = \begin{bmatrix} 1.4 \\ 1.1 \end{bmatrix} - \alpha_1 \begin{bmatrix} -0.462 \\ -0.3076 \end{bmatrix} = \\ &= \begin{bmatrix} 1.4 + 0.462\alpha_1 \\ 1.1 + 0.3076\alpha_1 \end{bmatrix} \end{aligned} \quad (6.228.)$$

V točki (6.228.) funkcija (6.214.) zavzame vrednost:

$$\begin{aligned} f(\underline{x}_2) &= f\left(\begin{bmatrix} 1.4 + 0.462\alpha_1 \\ 1.1 + 0.3076\alpha_1 \end{bmatrix}\right) = \\ &= (1.4 + 0.462\alpha_1)^2 - 2(1.4 + 0.462\alpha_1)(1.1 + 0.3076\alpha_1) + \\ &\quad + 2(1.1 + 0.3076\alpha_1)^2 - (1.4 + 0.462\alpha_1) - 2(1.1 + 0.3076\alpha_1) \end{aligned} \quad (6.229.)$$

in postane seveda funkcija ene spremenljivke - koeficienta  $\alpha_1$ . Slednji je optimalen tedaj, ko funkcija (6.229.) zavzame pri njem minimalno vrednost. Poiščimo  $\alpha_1^*$  na analitičen način:

$$\begin{aligned} \frac{df(\underline{x}_2)}{d\alpha_1} &= 2(1.4 + 0.462\alpha_1) \cdot 0.462 - 2[0.462(1.1 + 0.3076\alpha_1) + (1.4 + 0.462\alpha_1) \cdot 0.3076] + \\ &\quad + 4(1.1 + 0.3076\alpha_1) \cdot 0.3076 - 0.462 - 2 \cdot 0.3076 = \\ &= 1.2936 + 0.426\alpha_1 - 2[0.506 + 0.142\alpha_1 + 0.4306 + 0.142\alpha_1] + 1.353 + 0.378\alpha_1 - \\ &\quad - 1.075 = \\ &= -0.3016 + 0.236\alpha_1 = 0 \quad \Rightarrow \alpha_1^* = 1.277 \end{aligned} \quad (6.230.)$$

V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije.

Če izraz (6.230.) vstavimo v izraz (6.228.), dobimo:

$$\underline{x}_2 = \begin{bmatrix} 1.4 + 0.462 \cdot 1.277 \\ 1.1 + 0.3076 \cdot 1.277 \end{bmatrix} = \begin{bmatrix} 1.99 \\ 1.492 \end{bmatrix} \approx \begin{bmatrix} 2 \\ 1.5 \end{bmatrix} \quad (6.231.)$$

Gradient funkcije (6.214.) v točki (6.231.) se glasi:

$$\underline{g}_2 = \nabla f(\underline{x}_2) = \left. \begin{bmatrix} 2x_1 - 2x_2 - 1 \\ -2x_1 + 4x_2 - 2 \end{bmatrix} \right|_{\underline{x}=\underline{x}_2} = \begin{bmatrix} 2 \cdot 2 - 2 \cdot 1.5 - 1 \\ -2 \cdot 2 + 4 \cdot 1.5 - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (6.232.)$$

Ker pride gradient enak 0, končamo s postopkom. Končna rešitev torej je:

$$\underline{x}^* = \underline{x}_2 = \begin{bmatrix} 2 \\ 1.5 \end{bmatrix} \quad (6.233.)$$

Kot je razvidno iz postopka na tem primeru, je bila očitno konvergenca izredno hitra!

### **BROYDEN, FLETCHER, GOLDFARB, SHANNO METODA (BFGS)**

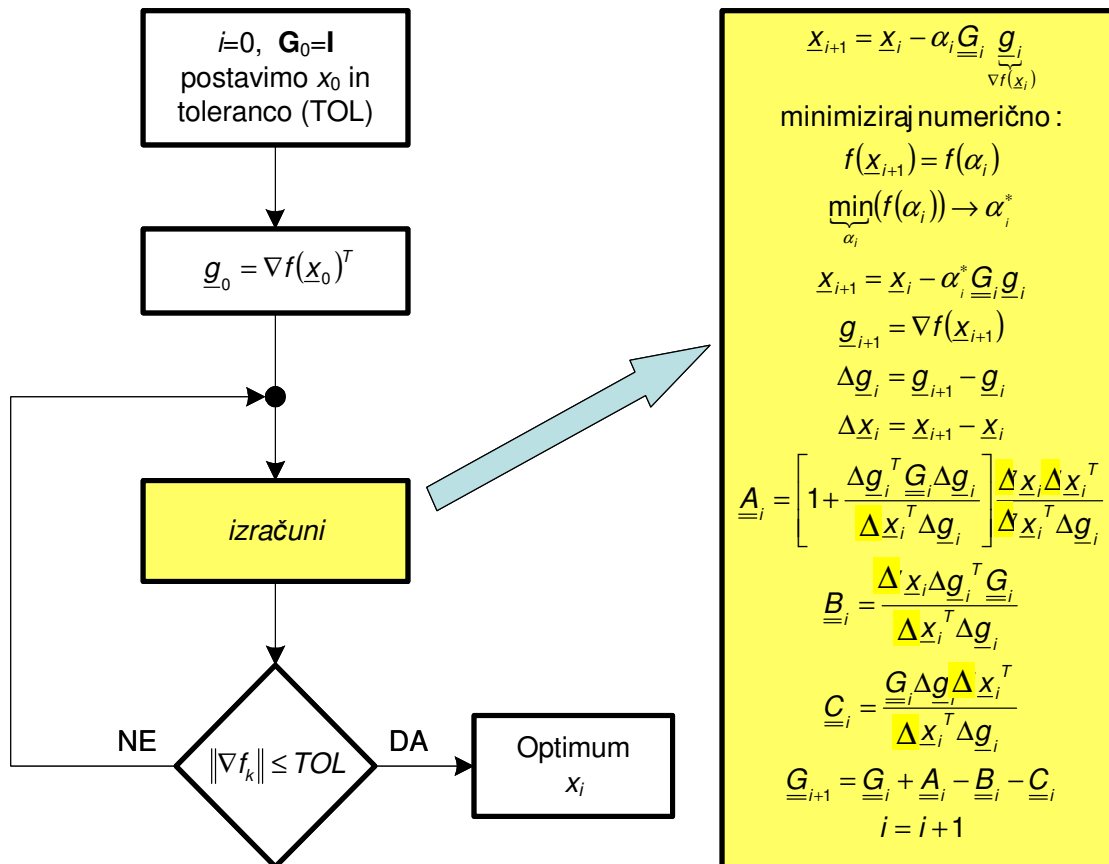
Je alternativa DFP metodi in je trenutno med najbolj popularnimi Quasi-Newtonovimi metodami. Izpeljava metode z vsemi ustreznimi dokazi je zelo zahtevna, zato je na tem mestu ne bomo podali, jo pa bralec lahko zasledi v literaturi [2, 14, 17, 21].

Matrika  $\underline{G}_{i+1}$  se v  $i+1$ . iteraciji izračuna na naslednji način [21]:

$$\underline{G}_{i+1} = \underline{G}_i + \underbrace{\left( 1 + \frac{\Delta \underline{g}_i^T \cdot \underline{G}_i \cdot \Delta \underline{g}_i}{\Delta \underline{x}_i^T \cdot \Delta \underline{g}_i} \right)}_A \cdot \underbrace{\frac{\Delta \underline{x}_i \cdot \Delta \underline{x}_i^T}{\Delta \underline{x}_i^T \cdot \Delta \underline{g}_i}}_B - \underbrace{\frac{\Delta \underline{x}_i \cdot \Delta \underline{g}_i^T \cdot \underline{G}_i}{\Delta \underline{x}_i^T \cdot \Delta \underline{g}_i}}_B - \underbrace{\frac{\underline{G}_i \cdot \Delta \underline{g}_i \cdot \Delta \underline{x}_i^T}{\Delta \underline{x}_i^T \cdot \Delta \underline{g}_i}}_C \quad (6.234.)$$

Algoritem je po osnovni strukturi enak kot pri DFP metodi (glej sliko 89), le da se  $\underline{G}_{i+1}$  v skladu z izrazom (6.234.) računa drugače. Slika 90 prikazuje simbolični algoritem BFGS metode [21].





Slika 90: Simbolični algoritem Quasi-Newtonove BFGS metode

Tudi delovanje BFGS metode bomo natančneje osvetlili na naslednjem primeru.

**Primer:**

S pomočjo DFP metode poiščite minimum naslednje funkcije:

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2, \quad (6.235.)$$

pri čemer je  $\underline{x}_0 = [0 \ 0]^T$ .

**Rešitev:**

V nadaljevanju bomo natančno sledili posameznim blokom simboličnega algoritma na sliki 90.

Najprej izberemo začetno metrično matriko:

$$\underline{\mathbf{G}}_0 = \underline{\mathbf{I}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (6.236.)$$

Gremo na 1. iteracijo ( $i = 0$ ).

Najprej je potrebno izračunati gradient v začetni točki:

$$\underline{g}_0 = \nabla f(\underline{x}_0)^T = \left[ \begin{array}{c} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{array} \right] \bigg|_{\underline{x}=\underline{x}_0} = \left[ \begin{array}{c} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{array} \right] \bigg|_{\underline{x}=\underline{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (6.237.)$$

Nato je potrebno izračunati nov vektor odločitvenih spremenljivk:

$$\underline{x}_1 = \underline{x}_0 - \alpha_0 \cdot \underline{G}_0 \cdot \underline{g}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \alpha_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -\alpha_0 \\ \alpha_0 \end{bmatrix} \quad (6.238.)$$

V točki (6.238.) funkcija (6.235.) zavzame vrednost:

$$f(\underline{x}_1) = (-\alpha_0) - \alpha_0 + 2(-\alpha_0)^2 + 2(-\alpha_0)\alpha_0 + \alpha_0^2 = -2\alpha_0 + \alpha_0^2 \quad (6.239.)$$

in postane seveda funkcija ene spremenljivke – koeficienta  $\alpha_0$ . Slednji je optimalen tedaj, ko funkcija (6.239.) zavzame pri njem minimalno vrednost. Poiščimo  $\alpha_0^*$  na analitičen način:

$$\frac{df(\underline{x}_1)}{d\alpha_0} = 2\alpha_0 - 2 = 0 \quad \Rightarrow \quad \alpha_0^* = 1 \quad (6.240.)$$

V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije.

Če izraz (6.240.) vstavimo v izraz (6.238.), dobimo:

$$\underline{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (6.241.)$$

Gradient funkcije (6.235.) v točki (6.241.) se glasi:

$$\underline{g}_1 = \nabla f(\underline{x}_1) = \left[ \begin{array}{c} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{array} \right] \bigg|_{\underline{x}=\underline{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \left[ \begin{array}{c} 1 - 4 + 2 \\ -1 - 2 + 2 \end{array} \right] = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad (6.242.)$$

Spremembo gradienta pa dobimo, če tvorimo razliko med izrazoma (6.242.) in (6.237.):

$$\Delta \underline{g}_0 = \underline{g}_1 - \underline{g}_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \quad (6.243.)$$

V nadaljevanju je potrebno izračunati tudi vektor  $\Delta \underline{x}_0$ , pri čemer od vektorja (6.241.) odštejemo začetno točko:

$$\Delta \underline{x}_0 = \underline{x}_1 - \underline{x}_0 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (6.244.)$$

V nadaljevanju je potrebno v skladu s sliko 90 izračunati naslednje matrike:

$$\begin{aligned} \underline{\underline{A}}_0 &= \left[ 1 + \frac{\Delta \underline{g}_0^T \underline{\underline{G}}_0 \Delta \underline{g}_0}{\Delta \underline{x}_0^T \Delta \underline{g}_0} \right] \frac{\Delta \underline{x}_0 \Delta \underline{x}_0^T}{\Delta \underline{x}_0^T \Delta \underline{g}_0} = \left[ 1 + \frac{\begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}}{\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}} \right] \cdot \frac{\begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix}}{\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}} = \\ &= \underbrace{\left[ 1 + \frac{4+0}{2+0} \right]}_3 \cdot \frac{\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}}{2+0} = \frac{3}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{aligned} \quad (6.245.)$$

$$\underline{\underline{B}}_0 = \frac{\Delta \underline{x}_0 \Delta \underline{g}_0^T \overset{I}{\underline{\underline{G}}}_0}{\Delta \underline{x}_0^T \Delta \underline{g}_0} = \frac{\begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \end{bmatrix}}{\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}} = \frac{\begin{bmatrix} 2 & 0 \\ -2 & 0 \end{bmatrix}}{2} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \quad (6.246.)$$

$$\underline{\underline{C}}_0 = \frac{\overset{I}{\underline{\underline{G}}}_0 \Delta \underline{g}_0 \Delta \underline{x}_0^T}{\Delta \underline{x}_0^T \Delta \underline{g}_0} = \frac{\begin{bmatrix} -2 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix}}{\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}} = \frac{\begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix}}{2} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \quad (6.247.)$$

Nato pa je potrebno v skladu s sliko 90 izračunati še:

$$\begin{aligned} \underline{\underline{G}}_1 &= \underline{\underline{G}}_0 + \underline{\underline{A}}_0 - \underline{\underline{B}}_0 - \underline{\underline{C}}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{3}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \frac{3}{2} & -\frac{3}{2} \\ -\frac{3}{2} & \frac{3}{2} \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{2} \end{bmatrix} \end{aligned} \quad (6.248.)$$

Na tej točki je algoritem končal s 1. iteracijo. Zato gremo na 2. iteracijo ( $i = 1$ ):

Najprej je potrebno izračunati nov vektor odločitvenih spremenljivk:

$$\underline{x}_2 = \underline{x}_1 - \alpha_1 \underline{G}_1 \underline{g}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \alpha_1 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{2} \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \alpha_1 \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 + 2\alpha_1 \end{bmatrix} \quad (6.249.)$$

V točki (6.249.) funkcija (6.235.) zavzame vrednost:

$$\begin{aligned} f(\underline{x}_2) &= (-1) - (1 + 2\alpha_1) + 2(-1)^2 + 2(-1)(1 + 2\alpha_1) + (1 + 2\alpha_1)^2 = \\ &= -1 - 1 - 2\alpha_1 + 2 - 2 - 4\alpha_1 + (1 + 4\alpha_1 + 4\alpha_1^2) = -1 - 2\alpha_1 + 4\alpha_1^2 \end{aligned} \quad (6.250.)$$

in postane seveda funkcija ene spremenljivke - koeficienta  $\alpha_1$ . Slednji je optimalen tedaj, ko funkcija (6.250.) zavzame pri njem minimalno vrednost. Poiščimo  $\alpha_1^*$  na analitičen način:

$$\frac{df(\underline{x}_2)}{d\alpha_1} = -2 + 8\alpha_1 = 0 \quad \Rightarrow \quad \alpha_1^* = \frac{1}{4} \quad (6.251.)$$

V praksi bi seveda računalniški program to storil z eno izmed metod enodimenzionalne numerične optimizacije.

Če izraz (6.251.) vstavimo v izraz (6.249.), dobimo:

$$\underline{x}_2 = \begin{bmatrix} -1 \\ 1 + 2\alpha_1^* \end{bmatrix} = \begin{bmatrix} -1 \\ 1 + 2\frac{1}{4} \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} \quad (6.252.)$$

Gradient funkcije (6.235.) v točki (6.252.) se glasi:

$$\underline{g}_2 = \nabla f(\underline{x}_2) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \Bigg|_{\underline{x}=\underline{x}_2 = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}} = \begin{bmatrix} 1 - 4 + 3 \\ -1 - 2 + 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (6.253.)$$

Ker pride gradient enak 0, končamo s postopkom. Končna rešitev torej je:

$$\underline{x}^* = \underline{x}_2 = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix} \quad (6.254.)$$

Kot je razvidno iz postopka na tem primeru, je bila očitno konvergenca izredno hitra!

## 7 GEOMETRIJSKO PROGRAMIRANJE

Geometrijsko programiranje (v nadaljevanju GP) je metodologija reševanja posebnega razreda nelinearnih optimizacijskih problemov, pri katerih so ciljna funkcija in omejitve nelinearne funkcije posebnih oblik. Zato se geometrijsko programiranje lahko priznava kot poseben primer nelinearnega programiranja [17].

V splošnem poznamo tri osnovne pristope, na osnovi katerih gradimo teorijo geometrijskega programiranja [17]:

- pristop z DIFERENCIALNIM RAČUNOM,
- pristop z uporabo CAUCHYJEVE neenakosti, ki daje povezavo med geometrijsko in njej pripadajočo aritmetično sredino (odtod tudi ime GP), in
- pristop z uporabo LEGANDREOVE transformacije (LT).

Slednji je primeren za uporabo zlasti tedaj, kadar je primarni matematični model zelo kompliciran. Se pa v veliko primerih tega da prevesti v takoimenovani dualni model s pomočjo LT, ki je konveksen in ima bistveno lažjo obliko za reševanje. V nadaljevanju se bomo omejili zgolj na pristop z diferencialnim računom in rezultate, ki temeljijo na izpeljavi tega pristopa, podrobnosti o drugih dveh pristopih pa si bralec lahko pogleda v literaturi [17].

Ena izmed ključnih značilnosti GP je v tem, da s to metodologijo najprej lahko poiščemo optimalno vrednost kriterijske funkcije  $f^*$ , šele nato pa, če je sploh potrebno (včasih se tudi to lahko zgodi!), izračunamo še optimalno vrednost odločitvenih spremenljivk  $x_i^*$ .

Značilnost GP je tudi v tem, da vedno zahteva ciljno funkcijo in omejitve zapisane v obliki tako imenovanih **POSINOMOV**, katerih podrobnosti bodo razvidne v nadaljevanju. Zato na tem mestu omenimo le, da posinome štejem v posebno kategorijo generaliziranih polinomov, kjer eksponenti niso nujno pozitivna cela števila.

Pri GP ločimo naslednja dva tipa problemov:

- probleme brez omejitev, ter
- probleme z omejitvami.

V nadaljevanju si bomo pogledali tako prvo kategorijo problemov, kot tudi drugo kategorijo problemov.

Kot bo razvidno, se da veliko nelinearnih problemov, ki so tipično logistične ali transportne narave, učinkovito reševati z uporabo GP. Še celo več, veliko se jih da rešiti celo zgolj z uporabo analitičnega računanja, brez uporabe numeričnih metod, ali pa slednje uporabimo le za izračun nekaterih pomožnih rezultatov.

Relativno enostavno analitično računanje pa se da uporabiti tudi za določene probleme z omejitvami, zato si bomo ogledali tudi nekaj primerov tovrstnega računanja.

Na koncu poudarimo še, da pri reševanju problemov geometrijskega programiranja obstajajo tako analitične kot numerične metode. Pri tem načeloma velja, da probleme rešujemo zgolj z analitičnim pristopom tedaj, kadar imamo opravka z malo odločitvenimi spremenljivkami in omejitvami. Kadar pa so problemi bolj kompleksni, bodisi imajo veliko odločitvenih spremenljivk in omejitev, bodisi imajo višje stopnje težavnosti reševanja problemov ali bolj zapletene strukture modelov, pa se poslužujemo tudi pomoči numeričnih metod.

## 7.1 Geometrijsko programiranje brez omejitev

Ciljno funkcijo lahko pri GP izrazimo na naslednji način:

$$f(\underline{x}) = U_1(\underline{x}) + U_2(\underline{x}) + \dots + U_N(\underline{x}) = \sum_{i=1}^N U_i(\underline{x}) \quad (7.1)$$

kjer so:

$f(\underline{x})$  - celotni stroški

$U_1(\underline{x})$  - delni stroški 1

$U_2(\underline{x})$  - delni stroški 2

...

$U_N(\underline{x})$  - delni stroški  $N$

V mnogo primerih se dajo komponente  $U_i$  izraziti kot takoimenovane močnostne funkcije, to je:

$$U_i = C_i \cdot x_1^{a_{i1}} \cdot x_2^{a_{i2}} \cdot \dots \cdot x_n^{a_{in}}, \quad i = 1, 2, \dots, N \quad (7.2)$$

oziroma:

$$\begin{aligned} U_1 &= C_1 \cdot x_1^{a_{11}} \cdot x_2^{a_{12}} \cdot \dots \cdot x_n^{a_{1n}} \\ U_2 &= C_2 \cdot x_1^{a_{21}} \cdot x_2^{a_{22}} \cdot \dots \cdot x_n^{a_{2n}} \\ &\dots \\ U_N &= C_N \cdot x_1^{a_{N1}} \cdot x_2^{a_{N2}} \cdot \dots \cdot x_n^{a_{Nn}} \end{aligned} \quad (7.3)$$

kjer velja:

$C_i \geq 0$  - pozitivne konstante

$a_{ij} \in \mathfrak{R}$  - realne konstante

$x_i \geq 0$  - pozitivne odločitvene spremenljivke

Če vse naštetu velja, potem lahko funkcijo  $f(\underline{x})$  poimenujemo tudi kot tako imenovano POSINOMSKO funkcijo (ali POSINOM). Če bi veljalo:  $a_{ij} \geq 0 \Rightarrow$  potem bi bila  $f(\underline{x})$  klasična POLINOMSKA funkcija, tu pa so lahko koeficienti  $a_{ij}$  tudi:  $a_{ij} < 0$ .

Polinomsko funkcijo lahko na primer zapišemo na naslednji način:

$$\begin{aligned} f(x_1, x_2, x_3) &= 6 + 3x_1 - 8x_2 + 7x_3 + 2x_1x_2 - 3x_1x_3 + \frac{4}{3}x_2x_3 + \frac{8}{7}x_1^2 - \\ &\quad - 9x_2^2 + x_3^2 \end{aligned} \quad (7.4)$$

Posinom pa npr. lahko zapišemo takole:

$$g(x_1, x_2, x_3) = x_1x_2x_3 + x_1^2x_2 + 4x_3 + \frac{2}{x_1x_2} + 5x_3^{-1/2} \quad (7.5)$$

$U_i$  iz izrazov (7.2.) oz. (7.3.) in  $f(\underline{x})$  iz izraza (7.1.) lahko zapišemo tudi v kompaktni obliki:



$$\begin{aligned}
 U_i &= \prod_{i=1}^n C_j \cdot x_i^{a_{ij}} = C_j \cdot \prod_{i=1}^n x_i^{a_{ij}} \\
 f(\underline{x}) &= \sum_{j=1}^N U_j(\underline{x}) = \sum_{j=1}^N C_j \cdot \prod_{i=1}^n x_i^{a_{ij}}
 \end{aligned}
 \tag{7.6}$$

V nadaljevanju si bomo ogledali izpeljavo pristopa k GP s pomočjo diferencialnega računa. Ta pristop bo, kot rečeno, edini, na katerega se bomo omejili v tem delu. Pri obravnavi se bomo omejili tudi zgolj na obravnavo optimumov v smislu minimumov funkcij, podobna logika pa seveda velja tudi za maksimume funkcij.

### 7.1.1 Izpeljava pristopa z diferencialnim računom

Potreben pogoj, da funkcija  $f(\underline{x}^*)$  iz izraza (7.6.) doseže minimum v točki  $\underline{x}^* = [x_1^*, \dots, x_n^*]^T$ , je naslednji:

$$\frac{\partial f(\underline{x})}{\partial x_k} = 0, \quad k = 1, \dots, n
 \tag{7.7}$$

iz česar sledi:

$$\frac{\partial f(\underline{x})}{\partial x_k} = \frac{\partial}{\partial x_k} \left[ \sum_{j=1}^N C_j \cdot \prod_{i=1}^n x_i^{a_{ij}} \right] = \frac{\partial}{\partial x_k} \left[ \sum_{j=1}^N C_j [x_1^{a_{1j}} \cdot x_2^{a_{2j}} \cdot \dots \cdot x_n^{a_{nj}}] \right]
 \tag{7.8}$$

Če funkcijo  $f(\underline{x})$  odvajamo po prvi odločitveni spremenljivki  $x_1$ , dobimo:

$$\begin{aligned}
 \frac{\partial f}{\partial x_1} &= \frac{\partial}{\partial x_1} \sum_{j=1}^N C_j [x_1^{a_{1j}} \cdot x_2^{a_{2j}} \cdot \dots \cdot x_n^{a_{nj}}] = \\
 &= \frac{\partial}{\partial x_1} [C_1 x_1^{a_{11}} x_2^{a_{21}} \dots x_n^{a_{n1}} + C_2 x_1^{a_{12}} x_2^{a_{22}} \dots x_n^{a_{n2}} + \dots + C_N x_1^{a_{1N}} x_2^{a_{2N}} \dots x_n^{a_{nN}}] = \\
 &= C_1 a_{11} x_1^{a_{11}-1} x_2^{a_{21}} \dots x_n^{a_{n1}} + C_2 a_{12} x_1^{a_{12}-1} x_2^{a_{22}} \dots x_n^{a_{n2}} + \dots + \\
 &\quad + C_N a_{1N} x_1^{a_{1N}-1} x_2^{a_{2N}} \dots x_n^{a_{nN}} = \\
 &= \sum_{j=1}^N C_j \cdot a_{1j} \cdot x_1^{a_{1j}-1} \cdot \prod_{i=2}^n x_i^{a_{ij}}
 \end{aligned}
 \tag{7.9}$$

Podobno dobimo (za odvod po 2., 3., ..., n. odločitveni spremenljivki):

$$\begin{aligned}\frac{\partial f}{\partial x_2} &= \sum_{j=1}^N C_j a_{2j} x_2^{a_{2j}-1} \prod_{\substack{i=1 \\ i \neq 2}}^n x_i^{a_{ij}} \\ \frac{\partial f}{\partial x_3} &= \sum_{j=1}^N C_j a_{3j} x_3^{a_{3j}-1} \prod_{\substack{i=1 \\ i \neq 3}}^n x_i^{a_{ij}} \\ &\dots \text{ ITN.} \\ \frac{\partial f}{\partial x_n} &= \sum_{j=1}^N C_j a_{nj} x_n^{a_{nj}-1} \prod_{\substack{i=1 \\ i \neq n}}^n x_i^{a_{ij}}\end{aligned}\tag{7.10}$$

V splošnem torej velja:

$$\frac{\partial f(\underline{x})}{\partial x_k} = \sum_{j=1}^N C_j a_{kj} x_k^{a_{kj}-1} \prod_{\substack{i=1 \\ i \neq k}}^n x_i^{a_{ij}}, \quad k = 1, \dots, n\tag{7.11}$$

Izraz (7.11.) še malo preuredimo in dobimo:

$$\begin{aligned}\frac{\partial f(\underline{x})}{\partial x_k} &= \frac{x_k}{x_k} \sum_{j=1}^N C_j a_{kj} x_k^{a_{kj}-1} \prod_{\substack{i=1 \\ i \neq k}}^n x_i^{a_{ij}} = \frac{1}{x_k} \sum_{j=1}^N C_j a_{kj} x_k^{a_{kj}} \prod_{\substack{i=1 \\ i \neq k}}^n x_i^{a_{ij}} = \\ &= \frac{1}{x_k} \sum_{j=1}^N C_j a_{kj} \underbrace{\prod_{\substack{i=1 \\ i \neq k}}^n x_i^{a_{ij}}}_{\prod_{i=1}^n x_i^{a_{ij}}} x_k^{a_{kj}}\end{aligned}\tag{7.12}$$

oziroma:

$$\frac{\partial f(\underline{x})}{\partial x_k} = \frac{1}{x_k} \sum_{j=1}^N C_j \cdot a_{kj} \cdot \prod_{i=1}^n x_i^{a_{ij}} = \frac{1}{x_k} \sum_{j=1}^N a_{kj} \underbrace{\prod_{i=1}^n C_j \cdot x_i^{a_{ij}}}_{U_j}\tag{7.13}$$

V minimumu funkcije  $f(\underline{x})$  torej sledi:

$$\frac{\partial f(\underline{x})}{\partial x_k} = \frac{1}{x_k} \sum_{j=1}^N a_{kj} \cdot U_j = 0, \quad k = 1, \dots, n\tag{7.14}$$

kar nam po krajšanju z odločitveno spremenljivko da rezultat:

$$\sum_{j=1}^N a_{kj} \cdot U_j^* = 0, \quad k = 1, \dots, n \quad \Big| \div f^*\tag{7.15}$$

Izraz (7.15) še delimo z optimalno ciljno funkcijo  $f^*$  in dobimo:

$$\sum_{j=1}^N a_{kj} \cdot \underbrace{\frac{U_j^*}{f^*}}_{y_j^*} = 0, \quad k = 1, \dots, n \quad (7.16)$$

Kot je razvidno iz izraza (7.16.), smo v optimumu upeljali tudi nove spremenljivke  $y_j^*$ , iz česar sledi:

$$\sum_{j=1}^N a_{kj} \cdot y_j^* = 0, \quad k = 1, \dots, n, \quad (7.17)$$

Za spremenljivke  $y_j^*$  v izrazu (7.17) mora veljati tudi naslednje:

$$y_j^* > 0 \quad \text{ter} \quad \sum_{j=1}^N y_j^* = 1 \quad (7.18)$$

saj na osnovi definicije teh spremenljivk (glej izraz (7.16.)) in izraza (7.1.) očitno sledi:

$$\sum_{j=1}^N y_j^* = \sum_{j=1}^N \frac{U_j^*}{f^*} = \frac{U_1^* + \dots + U_N^*}{f^*} = \frac{f^*}{f^*} = 1 \quad (7.19)$$

Torej, če strnemo potrebne pogoje za nastop minimuma v  $f(x)$  v pregledno obliko, dobimo naslednje izraze:

1. Pogoj **ortogonalnosti** takoimenovanega dualnega vektorja  $\underline{y}^*$ :

$$\sum_{j=1}^N a_{kj} \cdot y_j^* = 0, \quad k = 1, \dots, n \quad (7.20)$$

kjer je dualni vektor (glej izraz (7.16.)) definiran takole:

$$y_j^* = \frac{U_j^*}{f^*}, \quad j = 1, \dots, N \quad (7.21)$$

2. Pogoj **normiranosti** dualnega vektorja  $\underline{y}^*$ :

$$\sum_{j=1}^N y_j^* = 1 \quad (7.22)$$

3. Pogoj **stroge pozitivnosti** komponent dualnega vektorja  $\underline{y}^*$ :

$$y_j^* > 0, \quad j = 1, \dots, N \quad (7.23)$$

pri čemer za spremenljivke v izrazih (7.20) do (7.23) velja naslednje:

$$\begin{aligned} \underline{y}_j^* &= [y_1^*, \dots, y_j^*, \dots, y_N^*]^T \rightarrow \text{Vektor dualnih odločitvenih spremenljivk} \\ \underline{a}_i &= [a_{i1}, \dots, a_{ij}, \dots, a_{iN}]^T \rightarrow \text{Vektor eksponentov } (i = 1, \dots, n) \end{aligned} \quad (7.24)$$

pri čemer je seveda:

$$\underline{x}^* = [x_1^*, \dots, x_n^*]^T \rightarrow \text{Vektor primarnih odločitvenih spremenljivk}$$

Iz pogoja ortogonalnosti torej sledi, da je v minimumu optimalni dualni vektor  $\underline{y}^*$  ortogonalen na vsakega od vektorjev eksponentov  $\underline{a}_i$ , ki pripadajo vektorju primarnih odločitvenih spremenljivk  $\underline{x}^*$ .

Izraz (7.20) lahko zapišemo tudi kot:

$$\left. \begin{aligned} \sum_{j=1}^N a_{1j} \cdot y_j^* &= 0 \\ \sum_{j=1}^N a_{2j} \cdot y_j^* &= 0 \\ \dots \\ \sum_{j=1}^N a_{nj} \cdot y_j^* &= 0 \end{aligned} \right\} \Rightarrow \begin{aligned} a_{11} \cdot y_1^* + a_{12} \cdot y_2^* + \dots + a_{1N} \cdot y_N^* &= 0 \\ a_{21} \cdot y_1^* + a_{22} \cdot y_2^* + \dots + a_{2N} \cdot y_N^* &= 0 \\ \dots \\ a_{n1} \cdot y_1^* + a_{n2} \cdot y_2^* + \dots + a_{nN} \cdot y_N^* &= 0 \end{aligned} \quad (7.25)$$

oziroma:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nN} \end{bmatrix}}_{\underline{A}} \cdot \underbrace{\begin{bmatrix} y_1^* \\ y_2^* \\ \dots \\ y_N^* \end{bmatrix}}_{\underline{y}^*} = \underbrace{\begin{bmatrix} 0 \\ \dots \\ \dots \\ 0 \end{bmatrix}}_{\underline{0}} \quad (7.26)$$

Iz izraza (7.26) torej sledi:

$$\underline{A} \cdot \underline{y}^* = \underline{0} \quad (7.27)$$

kjer smo vpeljali matriko eksponentov  $\underline{A}$ , kjer je:

$n$  – število enačb

$N$  – število dualnih neznank  $y_i^*$ .

Definiramo lahko tudi takoimenovano stopnjo težavnosti reševanja problemov v GP na naslednji način:

$$T = N - n - 1 \quad (7.28)$$

Če je  $N - n - 1 = 0$  oz.  $N = n + 1$ , potem ima problem 0. stopnjo težavnosti in se neznanke  $y_i^*$  lahko določijo enolično iz pogojev ortogonalnosti in normiranosti. Kadar pa je  $N > n + 1$ , je neznank  $y_i^*$  več kot enačb v izrazih (7.25.) do (7.27.). Stopnja težavnosti  $T$  je tedaj pozitivna, reševanje pa zahteva bolj specialne prijeme.

V dosedaj razloženem postopku smo pokazali, kako lahko na čimbolj racionalen način izračunamo takoimenovane dualne odločitvene spremenljivke  $y_i^*$ . Ko pa so slednje izračunane, je možno z določenim postopkom izračunati tudi optimalno ciljno funkcijo (7.1.), ter nato optimalne primarne odločitvene spremenljivke  $\underline{x}^* = [x_1^*, \dots, x_n^*]^T$ . Izvedba tega postopka bo razvidna v nadaljevanju.

### **7.1.2 Postopek reševanja**

V nadaljevanju bomo opisali ključne korake, po katerih poteka trikoračni postopek iskanja optimalne rešitve s pomočjo geometrijskega programiranja na osnovi rezultatov z diferencialnim računom.

#### **1. KORAK (izračun $y_i^*$ ):**

Iz pogoja ortogonalnosti  $\underline{A} \cdot \underline{y}^* = \underline{0}$  (izraza (7.26.) oz. (7.27.)) izračunamo optimalne dualne spremenljivke  $y_i^*$ ,  $i = 1, \dots, N$ . Pri reševanju dotičnega (linearnega) sistema enačb uporabimo npr. Gaussovo metodo eliminacije.

#### **2. KORAK (izračun $f^*$ ):**

Nato lahko izračunamo optimalno ciljno funkcijo  $f^*$ . Na osnovi pogoja normiranosti (7.22.) namreč velja:

$$f^* = (f^*)^1 = (f^*)^{\sum_{j=1}^N y_j^*} = (f^*)^{y_1^* + y_2^* + \dots + y_N^*} = f^{*y_1^*} \cdot f^{*y_2^*} \cdot \dots \cdot f^{*y_N^*} \quad (7.29)$$

Velja pa tudi (glej izraz (7.21)):

$$f^* = \frac{U_1^*}{y_1^*} = \frac{U_2^*}{y_2^*} = \dots = \frac{U_N^*}{y_N^*} \quad (7.30)$$

Če izraz (7.30.) uporabimo v izrazu (7.29.), dobimo:

$$f^* = \left( \frac{U_1^*}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{U_2^*}{y_2^*} \right)^{y_2^*} \cdot \dots \cdot \left( \frac{U_N^*}{y_N^*} \right)^{y_N^*} \quad (7.31)$$

Seveda iz izraza (7.31.) ne moremo izračunati optimalne ciljne funkcije, saj delne funkcije  $U_j^*$  iz izraza (7.1.) niso znane. Ker pa je  $U_j^* = \prod_{i=1}^n C_j \cdot (x_i^*)^{a_{ij}}$  (glej izraz (7.6)), lahko izraz (7.31.) preoblikujemo v naslednjo obliko:

$$\begin{aligned} f^* &= \left[ \frac{\prod_{i=1}^n C_1 \cdot (x_1^*)^{a_{i1}}}{y_1^*} \right]^{y_1^*} \cdot \left[ \frac{\prod_{i=1}^n C_2 \cdot (x_2^*)^{a_{i2}}}{y_2^*} \right]^{y_2^*} \cdot \dots \cdot \left[ \frac{\prod_{i=1}^n C_N \cdot (x_N^*)^{a_{iN}}}{y_N^*} \right]^{y_N^*} \\ f^* &= \left[ \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \left\{ \prod_{i=1}^n (x_1^*)^{a_{i1}} \right\}^{y_1^*} \right] \cdot \left[ \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \left\{ \prod_{i=1}^n (x_2^*)^{a_{i2}} \right\}^{y_2^*} \right] \cdot \dots \\ &\quad \cdot \left[ \left( \frac{C_N}{y_N^*} \right)^{y_N^*} \left\{ \prod_{i=1}^n (x_N^*)^{a_{iN}} \right\}^{y_N^*} \right] = \\ &= \underbrace{\left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \dots \cdot \left( \frac{C_N}{y_N^*} \right)^{y_N^*}}_{\prod_{j=1}^N \left( \frac{C_j}{y_j^*} \right)^{y_j^*}} \cdot \underbrace{\left\{ \prod_{i=1}^n (x_1^*)^{a_{i1}} \right\}^{y_1^*} \cdot \dots \cdot \left\{ \prod_{i=1}^n (x_N^*)^{a_{iN}} \right\}^{y_N^*}}_{\prod_{j=1}^N \left[ \prod_{i=1}^n (x_i^*)^{a_{ij}} \right]^{y_j^*}} \end{aligned} \quad (7.32)$$

Dokazati se da (glej [17]), da izraz  $\prod_{j=1}^N \left[ \prod_{i=1}^n (x_i^*)^{a_{ij}} \right]^{y_j^*}$  v izrazu (7.32.) pride enak 1. Na tej osnovi pa dobimo:

$$f^* = \prod_{j=1}^N \left( \frac{C_j}{y_j^*} \right)^{y_j^*} \quad (7.33)$$

in tako je torej v 2. koraku možno izračunati tudi ciljno funkcijo  $f^*$ , saj so vse veličine v izrazu (7.33.) poznane.

### **3. KORAK (izračun odločitvenih spremenljivk $x_i^*$ ):**

Ko so dualne spremenljivke  $y_i^*$  in ciljna funkcija  $f^*$  v prvih dveh korakih izračunane, lahko izračunamo tudi (primarne) odločitvene spremenljivke  $x_i^*$ .

Spomnimo se namreč, da na osnovi izrazov (7.6.) in (7.21.) veljata naslednji relaciji:

$$y_j^* = \frac{U_j^*}{f^*}, \quad j = 1, \dots, N \quad (7.34)$$

ter

$$U_j^* = \prod_{i=1}^n C_j \cdot (x_i^*)^{a_{ij}}, \quad j = 1, \dots, N \quad (7.35)$$

Če izrazimo  $U_j^*$  v izrazu (7.34.) z  $y_j^*$  ter dobljeni izraz enačimo z izrazom (7.35.), dobimo:

$$U_j^* = f^* \cdot y_j^* = \prod_{i=1}^n C_j \cdot (x_i^*)^{a_{ij}}, \quad j = 1, \dots, N \quad (7.36)$$

Izraz (7.36.) lahko zapišemo tudi kot:

$$\frac{f^* \cdot y_j^*}{C_j} = \prod_{i=1}^n (x_i^*)^{a_{ij}} = (x_1^*)^{a_{1j}} \cdot (x_2^*)^{a_{2j}} \cdot \dots \cdot (x_n^*)^{a_{nj}}, \quad j = 1, \dots, N \quad (7.37)$$

Logaritmirajmo obe strani v izrazu (7.37.) in dobimo:

$$\ln \left( \frac{f^* y_j^*}{C_j} \right) = a_{1j} \ln x_1^* + a_{2j} \ln x_2^* + \dots + a_{nj} \ln x_n^*, \quad j = 1, \dots, N \quad (7.38)$$

Nato lahko vpeljemo nove spremenljivke:

$$w_i = \ln x_i^*, \quad i = 1, \dots, n \quad (7.39)$$

Sistem (7.38.) lahko z upoštevanjem (7.39.) zapišemo tudi kot:

$$\begin{aligned}
 \ln\left(\frac{f^* y_1^*}{C_1}\right) &= a_{11}w_1 + a_{21}w_2 + \dots + a_{n1}w_n \\
 \ln\left(\frac{f^* y_2^*}{C_2}\right) &= a_{12}w_1 + a_{22}w_2 + \dots + a_{n2}w_n \\
 &\dots \\
 \ln\left(\frac{f^* y_N^*}{C_N}\right) &= a_{1N}w_1 + a_{2N}w_2 + \dots + a_{nN}w_n
 \end{aligned} \tag{7.40}$$

iz česar sledi:

$$\underbrace{\begin{bmatrix} \ln\left(\frac{f^* y_1^*}{C_1}\right) \\ \ln\left(\frac{f^* y_2^*}{C_2}\right) \\ \dots \\ \ln\left(\frac{f^* y_N^*}{C_N}\right) \end{bmatrix}}_{\underline{K}} = \underbrace{\begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1N} & a_{2N} & \dots & a_{nN} \end{bmatrix}}_{\underline{A}^T} \cdot \underbrace{\begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_n \end{bmatrix}}_{\underline{w}} \tag{7.41}$$

Dobljeni sistem (7.41.) smo zapisali v kompaktni vektorsko-matrični obliki, kjer smo vpeljali tudi nova vektorja  $\underline{K}$  in  $\underline{w}$ . Sistem (7.41.) lahko zapišemo v še bolj pregledni obliki:

$$\underline{K} = \underline{A}^T \cdot \underline{w} \tag{7.42}$$

Ker so vsi koeficienti v vektorju  $\underline{K}$  in matriki  $\underline{A}$  znani, lahko iz sistema (7.42.) izračunamo tudi vektor neznank  $\underline{w}$ .

V primeru 0. stopnje težavnosti problema (ko je  $N = n + 1$ , glej izraz (7.28.)) se lahko vektor  $\underline{w}$  določi enolično iz izraza (7.42.). Ko so spremenljivke  $w_1, \dots, w_n$  izračunane, pa lahko izračunamo tudi primarne odločitvene spremenljivke v optimumu. V ta namen naredimo naslednjo operacijo nad izrazom (7.39.):

$$w_i = \ln x_i^*, \quad i = 1, \dots, n \quad |^e \tag{7.43}$$



pri čemer dobimo:

$$e^{w_i} = \underbrace{e^{\ln x_i^*}}_{x_i^*} \quad (7.44)$$

Sledi torej:

$$x_i^* = e^{w_i}, \quad i = 1, \dots, n \quad (7.45)$$

s čimer so izračunane tudi primarne odločitvene spremenljivke.

Če je stopnja težavnosti problema višja od ničte ( $N > n + 1$ ), si pa izberemo  $n$  linearno neodvisnih enačb v sistemu (7.42.) in na takšen način pridemo do optimalne rešitve za primarne odločitvene spremenljivke.

Morda se zdi procedura za izračun optimalnih primarnih odločitvenih spremenljivk, opisana v 3. koraku, nekoliko komplicirana. Kot bodo pokazali primeri, se dejansko da s spretnim »manevriranjem« izrazov (7.34.) in (7.35.), pri upoštevanju strukture delnih funkcij  $U_j^*$  iz izraza (7.1.), po nekoliko lažji (krajši) poti izračunati spremenljivke  $x_i^*$ .

### **Primer (0. stopnja težavnosti):**

Odločeno je bilo, da je potrebno transportirati določen material iz skladišča v tovarno. Material naj se prenese v odprti škatli dolžine  $x_1$  metrov, širine  $x_2$  metrov in višine  $x_3$  metrov. Izdelava dna škatle stane  $80 \frac{\text{€}}{\text{m}^2}$ , stranic stane  $10 \frac{\text{€}}{\text{m}^2}$  in konca škatle stane  $20 \frac{\text{€}}{\text{m}^2}$ . Za vsak transport škatle so stroški 1€. Poiščite minimalne skupne stroške izdelave škatle in transporta  $80\text{m}^3$  materiala. Poiščite tudi optimalne dimenzije škatle, pri katerih se dosežejo minimalni skupni stroški.

### **Rešitev:**

Skupno stroškovno funkcijo lahko definiramo kot:

$$\underbrace{\text{SKUPNI STROŠKI}}_f = \underbrace{\text{STROŠKI ŠKATLE}}_{f_1} + \underbrace{\text{STROŠKI TRANSPORTA}}_{f_2} \quad (7.46)$$

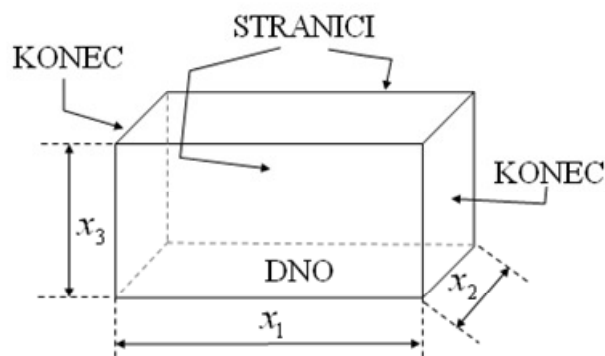
Pri tem so stroški izdelave škatle enaki:

$$f_1 = \underbrace{\text{STROŠKI STRANIC}}_{s_1} + \underbrace{\text{STROŠKI DNA}}_{s_2} + \underbrace{\text{STROŠKI KONCA}}_{s_3} \quad (7.47)$$

Stroški transporta škatle pa so enaki:

$$f_2 = \underbrace{\text{ŠTEVILO TRANSPORTOV}}_{s_4} \cdot \underbrace{\text{STROŠEK ZA VSAK TRANSPORT}}_{s_5} \quad (7.48)$$

Geometrijske veličine škatle so ustrezno predstavljene na sliki 91.



Slika 91: Geometrijske veličine škatle

Delne stroške izdelave škatle iz izraza (7.47.) lahko na osnovi slike 91 in podatkov izrazimo na naslednji način (površina – število kvadratnih metrov krat cena za en kvadratni meter):

$$\begin{aligned} S_1 &= (2x_1x_3) \cdot 10 \\ S_2 &= (x_1x_2) \cdot 80 \\ S_3 &= (2x_2x_3) \cdot 20 \end{aligned} \quad (7.49)$$

Na osnovi izrazov (7.49.) izraz (7.47.) preide v obliko:

$$f_1 = (2x_1x_3) \cdot 10 + (x_1x_2) \cdot 80 + (2x_2x_3) \cdot 20 \quad (7.50)$$

Število transportov iz izraza (7.48.) lahko izrazimo na naslednji način (volumen, ki ga je potrebno prepeljati, deljeno z volumnom škatle):

$$S_4 = \frac{80m^3}{x_1m \cdot x_2m \cdot x_3m} = \frac{80}{\underbrace{x_1x_2x_3}_v} \quad (7.51)$$

kjer je  $V$  volumen škatle.

Če bi npr. bil volumen škatle  $V = x_1 x_2 x_3 = 80 m^3$ , potem bi bil  $S_4 = \frac{80}{80} = 1$  in bi bilo treba opraviti le en transport. Če bi bil pa  $V = 40 m^3$ , bi bil  $S_4 = \frac{80}{40} = 2$  in bi bilo potrebno opraviti dva transporta, itn. Podobno sklepamo naprej  $\rightarrow$  manjši volumen  $\rightarrow$  več transportov.

Ker je  $S_5 = 1$  (saj je strošek 1€ za vsak transport), lahko na osnovi izraza (7.51.) zapišemo izraz (7.48.) v naslednji obliki:

$$f_2 = S_4 \cdot S_5 = \frac{80}{x_1 x_2 x_3} \cdot 1 = \frac{80}{x_1 x_2 x_3} \quad (7.52)$$

Če upoštevamo izraza (7.50.) in (7.52.) v izrazu (7.46.), dobimo skupno stroškovno (ciljno) funkcijo, ki se glasi:

$$\begin{aligned} f(x_1, x_2, x_3) &= 20x_1 x_3 + 80x_1 x_2 + 40x_2 x_3 + \frac{80}{x_1 x_2 x_3} = \\ &= 20x_1^1 x_2^0 x_3^1 + 80x_1^1 x_2^1 x_3^0 + 40x_1^0 x_2^1 x_3^1 + 80x_1^{-1} x_2^{-1} x_3^{-1} \end{aligned} \quad (7.53)$$

Kot je razvidno iz izraza (7.53.), smo skupno stroškovno funkcijo zapisali tudi v obliki, primerni za reševanje z geometrijskim programiranjem.

Naš optimizacijski (geometrijski) problem brez dodatnih omejitev se torej glasi:

$$\min_{x_1, x_2, x_3} f(x_1, x_2, x_3) \quad (7.54)$$

Potrebno je torej poiskati takšne dimenzije škatle (odločitvene spremenljivke  $x_1, x_2, x_3$ ), pri katerih bo skupna stroškovna funkcija (7.54.) dosegla minimalno vrednost.

V splošnem lahko ciljno funkcijo (7.54.) na osnovi izraza (7.6.) zapišemo tudi na naslednji način (upoštevamo 3 odločitvene spremenljivke in strukturo funkcije  $f$  iz izraza (7.53.), ki je sestavljena iz 4 členov – delnih funkcij):

$$\begin{aligned} f(\underline{x}) &= \sum_{j=1}^N C_j \prod_{i=1}^n x_i^{a_{ij}} = \sum_{j=1}^4 C_j \prod_{i=1}^3 x_i^{a_{ij}} = \sum_{j=1}^4 C_j [x_1^{a_{1j}} \cdot x_2^{a_{2j}} \cdot x_3^{a_{3j}}] = \\ &= C_1 x_1^{a_{11}} x_2^{a_{21}} x_3^{a_{31}} + C_2 x_1^{a_{12}} x_2^{a_{22}} x_3^{a_{32}} + C_3 x_1^{a_{13}} x_2^{a_{23}} x_3^{a_{33}} + C_4 x_1^{a_{14}} x_2^{a_{24}} x_3^{a_{34}} \end{aligned} \quad (7.55)$$

Izraz (7.55.) bomo primerjali z izrazom (7.53.), ki ga v ta namen še nekoliko preuredimo:

$$f(x_1, x_2, x_3) = \underbrace{80x_1^1 x_2^1 x_3^0}_{u_1(x)} + \underbrace{40x_1^0 x_2^1 x_3^1}_{u_2(x)} + \underbrace{20x_1^1 x_2^0 x_3^1}_{u_3(x)} + \underbrace{80x_1^{-1} x_2^{-1} x_3^{-1}}_{u_4(x)} \quad (7.56)$$

Če primerjamo istoležne člene v izrazih (7.53) in (7.56), dobimo:

$$C_1 = 80, C_2 = 40, C_3 = 20, C_4 = 80 \quad (7.57)$$

ter

$$\begin{aligned} a_{11} &= 1, & a_{21} &= 1, & a_{31} &= 0 \\ a_{12} &= 0, & a_{22} &= 1, & a_{32} &= 1 \\ a_{13} &= 1, & a_{23} &= 0, & a_{33} &= 1 \\ a_{14} &= -1, & a_{24} &= -1, & a_{34} &= -1 \end{aligned} \quad (7.58)$$

Na osnovi koeficientov v izrazu (7.58) pa lahko sestavimo matriko  $\underline{\underline{A}}$ , ki smo jo obravnavali v izrazih (7.26.) oz. (7.27.):

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \end{bmatrix} \quad (7.59)$$

Na osnovi pogoja ortogonalnosti (glej izraz (7.27.)) sledi:

$$\underline{\underline{A}} \cdot \underline{\underline{y}}^* = \underbrace{\begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \end{bmatrix}}_{\underline{\underline{A}}} \cdot \underbrace{\begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix}}_{\underline{\underline{y}}^*} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{\underline{\underline{0}}} \quad (7.60)$$

Na osnovi pogoja normiranosti (glej izraz (7.22.)) pa sledi:

$$\sum_{j=1}^N y_j^* = 1 \Rightarrow \sum_{j=1}^4 y_j^* = 1 \Rightarrow y_1^* + y_2^* + y_3^* + y_4^* = 1 \quad (7.61)$$

Preverimo, s kakšno stopnjo težavnosti T reševanja problema imamo opravka (glej izraz (7.28.)):

$$\left. \begin{aligned} N &= 4 \\ n+1 &= 3+1=4 \end{aligned} \right\} \Rightarrow N = n+1 \Rightarrow 0. \text{ STOPNJA TEŽAVNOSTI!} \quad (7.62)$$

torej imamo opravka z ničto stopnjo težavnosti, kar pomeni relativno enostavno reševanje problema.

Če izraza (7.60) in (7.61) združimo, dobimo:

$$\begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (7.63)$$

Sistem v izrazu (7.63) ima 4 enačbe in 4 neznanke. Če ga bomo rešili, bomo s tem poiskali neznanke, to so optimalne dualne odločitvene spremenljivke  $y_1^*$ ,  $y_2^*$ ,  $y_3^*$  in  $y_4^*$ . Na osnovi izračunanih dualnih spremenljivk pa bo nato možno izračunati tudi optimalno ciljno funkcijo ter optimalne primarne odločitvene spremenljivke.

V nadaljevanju torej izvedimo trikoračni postopek, opisan prej v tem poglavju.

### 1. KORAK (izračun $y_j^*$ )

Sistem (7.63.) lahko zapišemo tudi v obliki naslednjih enačb:

$$\begin{aligned} E_1: y_1^* + y_3^* - y_4^* &= 0 & \Rightarrow y_4^* &= y_1^* + y_3^* \\ E_2: y_1^* + y_2^* - y_4^* &= 0 & \Rightarrow y_4^* &= y_1^* + y_2^* \\ E_3: y_2^* + y_3^* - y_4^* &= 0 & \Rightarrow y_4^* &= y_2^* + y_3^* \\ E_4: y_1^* + y_2^* + y_3^* + y_4^* &= 1 \end{aligned} \quad (7.64)$$

Izenačimo enačbi  $E_2$  in  $E_3$  v izrazu (7.64.), pri čemer dobimo:

$$y_1^* + y_2^* = y_2^* + y_3^* \Rightarrow y_1^* = y_3^* \quad (7.65)$$

Če izenačimo še  $E_1$  in  $E_2$ , dobimo:

$$y_4^* = y_1^* + y_3^* = y_1^* + y_2^* \Rightarrow y_2^* = y_3^* \quad (7.66)$$

Na podlagi izrazov (7.65) in (7.66) torej velja:

$$y_1^* = y_2^* = y_3^* \quad (7.67)$$

Upoštevajmo izraz (7.67) v  $E_4$ , pri čemer dobimo:

$$y_1^* + y_1^* + y_1^* + y_4^* = 1 \Rightarrow y_4^* = -3y_1^* + 1 \quad (7.68)$$

Nato vstavimo izraz (7.68) v enačbo  $E_1$ , ki ob upoštevanju izraza (7.67.) preide v naslednjo optimalno rešitev za prvo dualno spremenljivko:

$$1 - 3y_1^* = y_1^* + y_1^* \Rightarrow y_1^* = \frac{1}{5} \quad (7.69)$$

Na osnovi rešitve (7.69.) in izrazov (7.67.) in (7.68.) lahko zapišemo optimalne rešitve še za preostale tri dualne spremenljivke:

$$y_2^* = \frac{1}{5}, \quad y_3^* = \frac{1}{5} \quad \text{in} \quad y_4^* = 1 - 3y_1^* = 1 - 3 \cdot \frac{1}{5} = \frac{2}{5} \quad (7.70)$$

V nadaljevanju se lotimo 2. koraka postopka reševanja, ki nam bo dal optimalno ciljno funkcijo.

## 2. KORAK (izračun $f^*$ )

Na osnovi izrazov (7.33.) in (7.62.) lahko zapišemo za optimalno ciljno funkcijo:

$$f^* = \prod_{j=1}^4 \left( \frac{C_j}{y_j^*} \right)^{y_j^*} = \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4^*} \right)^{y_4^*} \quad (7.71)$$

Če upoštevamo izraz (7.57.) ( $C_1 = 80$ ,  $C_2 = 40$ ,  $C_3 = 20$ ,  $C_4 = 80$ ) ter rešitve v izrazih (7.69.), (7.70.) ( $y_1^* = y_2^* = y_3^* = \frac{1}{5}$ ,  $y_4^* = \frac{2}{5}$ ) v izrazu (7.71.), dobimo naslednji rezultat za optimalno ciljno funkcijo:

$$\begin{aligned} f^* &= \left( \frac{80}{\frac{1}{5}} \right)^{\frac{1}{5}} \cdot \left( \frac{40}{\frac{1}{5}} \right)^{\frac{1}{5}} \cdot \left( \frac{20}{\frac{1}{5}} \right)^{\frac{1}{5}} \cdot \left( \frac{80}{\frac{2}{5}} \right)^{\frac{2}{5}} = (400)^{\frac{1}{5}} \cdot (200)^{\frac{1}{5}} \cdot (100)^{\frac{1}{5}} \cdot (200)^{\frac{2}{5}} \\ &= (4 \cdot 10^2 \cdot 2 \cdot 10^2 \cdot 10^2)^{\frac{1}{5}} \cdot (200)^{\frac{2}{5}} = (8 \cdot 10^6)^{\frac{1}{5}} \cdot (200)^{\frac{2}{5}} = \\ &= 24.0224 \cdot 8.3255 = 199.999 \approx 200\text{€} \end{aligned} \quad (7.72)$$

Torej minimalni možni skupni stroški transporta in izdelave škatle dosegajo vrednost v višini 200€.

Izračunajmo še primarne odločitvene spremenljivke, ki predstavljajo seveda optimalne dimenzije škatle.

### 3. KORAK (izračun $x_i^*$ )

Primarne odločitvene spremenljivke bomo izračunali na dva načina. Pri prvem načinu bomo uporabili pristop s spretnim »manevriranjem« izrazov (7.34.) in (7.35.), kjer bomo seveda poštevali tudi strukture delnih funkcij  $U_j^*$  iz izraza (7.56.). Pri drugem načinu pa bomo računali na daljši način, kot je to opisano v izrazih (7.34.) do (7.45.).

#### 1. način:

Delne funkcije v izrazu (7.56) imajo obliko:

$$\begin{aligned} U_1(\underline{x}^*) &= 80x_1^* x_2^* \\ U_2(\underline{x}^*) &= 40x_2^* x_3^* \\ U_3(\underline{x}^*) &= 20x_1^* x_3^* \\ U_4(\underline{x}^*) &= 80 \cdot \frac{1}{x_1^* x_2^* x_3^*} \end{aligned} \quad (7.73)$$

Na osnovi izraza (7.36) pa lahko zapišemo z upoštevanjem izrazov (7.69.), (7.70.) in (7.72.) tudi naslednje izraze:

$$\begin{aligned} U_j^* &= f^* \cdot y_j^*, \quad j=1,2,3,4 \\ U_1^* &= f^* \cdot y_1^* = 200 \cdot \frac{1}{5} = 40 \\ U_2^* &= f^* \cdot y_2^* = 200 \cdot \frac{1}{5} = 40 \\ U_3^* &= f^* \cdot y_3^* = 200 \cdot \frac{1}{5} = 40 \\ U_4^* &= f^* \cdot y_4^* = 200 \cdot \frac{2}{5} = 80 \end{aligned} \quad (7.74)$$

Če izenačimo delne funkcije v izrazih (7.73) in (7.74), dobimo:

$$\begin{aligned}
F_1: U_1(\underline{x}^*) &= 80x_1^* x_2^* = 40 \Rightarrow x_2^* = \frac{1}{2x_1^*} \\
F_2: U_2(\underline{x}^*) &= 40x_2^* x_3^* = 40 \Rightarrow x_2^* = \frac{1}{x_3^*} \\
F_3: U_3(\underline{x}^*) &= 20x_1^* x_3^* = 40 \Rightarrow x_1^* = \frac{2}{x_3^*} \\
F_4: U_4(\underline{x}^*) &= \frac{80}{x_1^* x_2^* x_3^*} = 80 \Rightarrow x_1^* x_2^* x_3^* = 1
\end{aligned} \tag{7.75}$$

Če nato izenačimo izraza  $F_1$  in  $F_2$  v (7.75.), dobimo:

$$x_2^* = \frac{1}{2x_1^*} = \frac{1}{x_3^*} \Rightarrow x_3^* = 2x_1^* \tag{7.76}$$

Če upoštevamo rezultata iz izraza (7.76), ter izraza  $F_3$  v (7.75.):

$$x_3^* = 2x_1^*, \quad x_1^* x_3^* = 2 \tag{7.77}$$

dobimo naslednji optimalen rezultat za prvo primarno odločitveno spremenljivko:

$$x_1^* \cdot 2x_1^* = 2 \Rightarrow x_1^* = 1m \tag{7.78}$$

Vstavimo rezultat (7.78.) v izraz  $F_1$  iz (7.75.), pri čemer dobimo optimalen rezultat za drugo primarno odločitveno spremenljivko:

$$x_2^* = \frac{1}{2x_1^*} = \frac{1}{2 \cdot 1} = \frac{1}{2}m \tag{7.79}$$

Vstavimo rezultat (7.79.) v izraz  $F_2$  iz (7.75.), pri čemer dobimo optimalen rezultat še za tretjo primarno odločitveno spremenljivko:

$$x_2^* x_3^* = 1 \Rightarrow \frac{1}{2}x_3^* = 1 \Rightarrow x_3^* = 2m \tag{7.80}$$

Na osnovi izrazov (7.78), (7.79) in (7.80) sledi, da je volumen optimalne škatle enak:

$$V^* = x_1^* x_2^* x_3^* = 1m \cdot \frac{1}{2}m \cdot 2m = 1m^3 \tag{7.81}$$

Odgovor na vprašanje, kakšne morajo biti optimalne dimenzije škatle, da bodo doseženi minimalni skupni stroški njene izdelave in predpisanega transporta le-te, se torej glasi:



Če bi bile dimenzije škatle optimalne s stranicami  $(x_1^*, x_2^*, x_3^*) = \left(1m, \frac{1}{2}m, 2m\right)$ , bi lahko dosegli minimalne skupne stroške izdelave škatle in njenega transporta v višini 200€, če bi hoteli transportirati  $80m^3$  materiala. Pri tem bi opravili  $S_4^* = \frac{80}{(x_1^* x_2^* x_3^*)} = \frac{80}{V^*} = \frac{80}{1} = 80$  transportov.

Na osnovi izrazov (7.75.) lahko podamo še distribucijo optimalnih (minimalnih) doseženih stroškov glede na izdelavo in transport škatle, ki se glasi:

$$f^*(x) = f(x^*) = U_1^* + U_2^* + U_3^* + U_4^* = \underbrace{40 + 40 + 40}_{120\text{€}} + \underbrace{80}_{\substack{\text{STROŠKI} \\ \text{TRANSPORTA}}} = \underbrace{120\text{€}}_{\substack{\text{STROŠKI} \\ \text{IZDELAVE} \\ \text{ŠKATLE}}} + \underbrace{80}_{\substack{\text{STROŠKI} \\ \text{TRANSPORTA}}} = 200\text{€} \quad (7.82)$$

$\downarrow$   
 80 TRANSPORTOV  $\times$   
 1€ ZA EN TRANSPORT

V nadaljevanju si pogledjmo še drugi način, kako izračunati optimalne primarne odločitvene spremenljivke.

## 2. način:

Najprej tvorimo sistem (7.41.), pri čemer upoštevamo izraz (7.57.) ( $C_1 = 80, C_2 = 40, C_3 = 20, C_4 = 80$ ), matriko  $\underline{A}^T$  v izrazu (7.59.), izračunane optimalne dualne spremenljivke v izrazih (7.69.) in (7.70.), ter izračunano optimalno ciljno funkcijo v izrazu (7.72.). Tako dobimo:

$$\underbrace{\begin{bmatrix} \ln\left(\frac{f^* y_1^*}{C_1}\right) \\ \ln\left(\frac{f^* y_2^*}{C_2}\right) \\ \ln\left(\frac{f^* y_3^*}{C_3}\right) \\ \ln\left(\frac{f^* y_4^*}{C_4}\right) \end{bmatrix}}_K = \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}}_{\underline{A}^T} \cdot \underbrace{\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}}_w \quad (7.83)$$

kjer lahko vektor  $\underline{K}$  izrazimo kot:

$$\underline{K} = \begin{bmatrix} \ln\left(\frac{200 \cdot \frac{1}{5}}{80}\right) \\ \ln\left(\frac{200 \cdot \frac{1}{5}}{40}\right) \\ \ln\left(\frac{200 \cdot \frac{1}{5}}{20}\right) \\ \ln\left(\frac{200 \cdot \frac{2}{5}}{80}\right) \end{bmatrix} = \begin{bmatrix} \ln\left(\frac{200}{400}\right) \\ \ln\left(\frac{200}{200}\right) \\ \ln\left(\frac{200}{100}\right) \\ \ln\left(\frac{400}{400}\right) \end{bmatrix} = \begin{bmatrix} \ln \frac{1}{2} \\ \ln 1 \\ \ln 2 \\ \ln 1 \end{bmatrix} \quad (7.84)$$

Imamo torej sistem:

$$\begin{aligned} G_1: w_1 + w_2 &= \ln \frac{1}{2} \\ G_2: w_2 + w_3 &= \ln 1 \\ G_3: w_1 + w_3 &= \ln 2 \\ G_4: -w_1 - w_2 - w_3 &= \ln 1 \end{aligned} \quad (7.85)$$

Če seštejemo  $G_1$ ,  $G_2$  in  $G_4$ , dobimo:

$$w_1 + w_2 + w_2 + w_3 - w_1 - w_2 - w_3 = \ln \frac{1}{2} + \ln 1 + \ln 1 \quad (7.86)$$

Odtod sledi:

$$w_2^* = \ln\left(\frac{1}{2} \cdot 1 \cdot 1\right) = \ln \frac{1}{2} \quad (7.87)$$

Vstavimo izraz (7.87.) v  $G_1$  izraza (7.85.), pri čemer dobimo:

$$w_1^* = \ln \frac{1}{2} - w_2^* = \ln \frac{1}{2} - \ln \frac{1}{2} = 0 \quad (7.88)$$

Vstavimo še izraz (7.87.) v  $G_2$  izraza (7.85.), pri čemer dobimo:

$$w_3^* = \ln 1 - w_2^* = \ln 1 - \ln \frac{1}{2} = \ln \left( \frac{1}{\frac{1}{2}} \right) = \ln 2 \quad (7.89)$$

Dobimo torej naslednje rezultate za koeficiente vektorja  $\underline{w}$ :

$$w_2^* = \ln \frac{1}{2}, \quad w_1^* = 0, \quad w_3^* = \ln 2 \quad (7.90)$$

Seveda nas zanimajo optimalne primarne odločitvene spremenljivke, ki jih lahko izračunamo, če uporabimo izraz (7.45), to je:

$$x_i^* = e^{w_i^*}, \quad i = 1, 2, 3 \quad (7.91)$$

Odtod dobimo naslednje rezultate za optimalne dimenzije škatle:

$$\begin{aligned} x_1^* &= e^{w_1^*} = e^0 = 1m \\ x_2^* &= e^{w_2^*} = e^{\ln \frac{1}{2}} = \frac{1}{2}m \\ x_3^* &= e^{w_3^*} = e^{\ln 2} = 2m \end{aligned} \quad (7.92)$$

Seveda dobimo enake rezultate kot pri 1. načinu v izrazih (7.78.) do (7.80.).

V nadaljevanju si bomo pogledali ozadje takoimenovane Cauchyve neenakosti, ki nam še bolj osvetli matematične relacije med primarno in dualno ciljno funkcijo. Razumevanje tovrstne problematike pa nam bo pomagalo tudi pri reševanju problemov višje stopnje.

### ***7.1.3 Ozadje Cauchyve neenakosti in reševanje problemov višje stopnje***

Imamo naslednjo primarno ciljno funkcijo, ki jo želimo minimizirati, ki se je pojavila že v izrazu (7.6.):

$$f(\underline{x}) = \sum_{j=1}^N U_j(\underline{x}) = \sum_{j=1}^N C_j \cdot \prod_{i=1}^n x_i^{a_{ij}} \quad (7.93)$$

V prejšnjih poglavjih smo izpeljali tudi dualno ciljno funkcijo (glej izraz (7.33)):

$$f^*(\underline{y}) = \prod_{j=1}^N \left( \frac{C_j}{y_j^*} \right)^{y_j^*} \quad (7.94)$$

pri čemer smo  $y_j^*$  izračunali na osnovi pogojev ortogonalnosti in normiranosti.

Bistvo Cauchyve neenakosti (kar zaradi zapletenosti ne bomo dokazovali) je, da velja naslednja relacija med primarno in dualno ciljno funkcijo:

$$\min_{\underline{x}} f(\underline{x}) = \max_{\underline{y}} f^*(\underline{y}) \quad (7.95)$$

Torej, če minimiziramo primarno funkcijo  $f(\underline{x})$ , dobimo enak rezultat, kot če bi maksimizirali dualno funkcijo  $f^*(\underline{y})$ .

Velja še naslednja povezava, ki se jo da dokazati:

$$\max_{\underline{y}} f^*(\underline{y}) = \max_{\underline{y}} \ln f^*(\underline{y}) \quad (7.96)$$

Torej, če nam kdaj bolj odgovarja računati s funkcijo  $\ln f^*(\underline{y})$  namesto s funkcijo  $f^*(\underline{y})$ , tedaj bomo to storili, saj pridemo do enakega rezultata za optimum.

Seveda veljajo tu še omejitve, iz katerih izračunamo  $y_j^*$  in smo jih srečali že v izrazih (7.22.) in (7.27.), ki se glasijo:

$$\underline{A} \cdot \underline{y}^* = 0, \quad \text{ter} \quad \sum_{j=1}^N y_j^* = 1 \quad (7.97)$$

V nadaljevanju si bomo pogledali primer, kjer pride stopnja težavnosti višja od 0. Že intuitivno lahko slutimo, da je v takšnih primerih reševanje težje, kot pa je bilo v primerih 0. stopnje težavnosti. Vendar bomo pokazali, da si lahko na osnovi sklepov Cauchyve neenakosti dokaj olajšamo delo, torej če upoštevamo relaciji v izrazih (7.95.) in (7.96.).

Dani primer bomo rešili na dva načina:

- Način z odvajanjem logaritmske dualne ciljne funkcije, ter
- Poenostavljen način z odvajanjem dualnih odločitvenih spremenljivk

### 7.1.3.1 Pristop k reševanju problemov višje stopnje z odvajanjem logaritske dualne ciljne funkcije

Najprej si pogledjmo pristop k reševanju problemov višje stopnje z odvajanjem logaritske dualne ciljne funkcije. Metodologijo tega pristopa bomo osvetlili na naslednjem primeru.

#### Primer (1. stopnja težavnosti):

Opravka imamo z inštalacijo črpalke nekega rezervoarja [21]. Strošek instalacije cevi je dan z odvisnostjo:

$$S_1(D) = 100D + 50D^2 \quad (7.98)$$

kjer je  $D$  premer cevi v  $cm$ .

Strošek, povezan z rezervoarjem, je dan z odvisnostjo:

$$S_2(Q) = \frac{20}{Q} \quad (7.99)$$

kjer je  $Q$  pretok v  $\frac{m^3}{s}$ .

Strošek črpanja pa je dan z odvisnostjo:

$$S_3(D, Q) = \frac{3000Q^2}{D^5} \quad (7.100)$$

Poiščite takšni optimalni veličini  $D^*$  in  $Q^*$ , pri katerih bodo skupni stroški instalacije in uporabe črpalke rezervoarja minimalni.

Skupna stroškovna (ciljna) funkcija se glasi (vsota vseh treh delnih stroškov):

$$S(D, Q) = S_1(D) + S_2(Q) + S_3(D, Q) = 100D + 50D^2 + \frac{20}{Q} + \frac{3000Q^2}{D^5} \quad (7.101)$$

Zapišimo jo v primerni obliki za GP:

$$S(D, Q) = \underbrace{100D^1 Q^0}_{u_1} + \underbrace{50D^2 Q^0}_{u_2} + \underbrace{20D^0 Q^{-1}}_{u_3} + \underbrace{300D^{-5} Q^2}_{u_4} \quad (7.102)$$

Originalen (primaren) optimizacijski problem se glasi:

$$\min_{D,Q} S(D,Q) \quad (7.103)$$

V splošnem lahko funkcijo  $S$  iz izraza (7.102.) na osnovi izraza (7.6.) zapišemo tudi na naslednji način (2 odločitveni spremenljivki in 4 delne funkcije v  $S$ ):

$$S\left(\underbrace{D,Q}_x\right) = \sum_{j=1}^4 C_j \prod_{i=1}^2 x_i^{a_{ij}}; \quad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} D \\ Q \end{bmatrix} \quad (7.104)$$

Imamo torej:

$N = 4 \rightarrow$  4 členi funkcije  $S$  v izrazu (7.102.)

$n = 2 \rightarrow$  2 odločitveni spremenljivki ( $D, Q$ ), ki se pojavita v izrazu (7.102.).

Izraz (7.104.) lahko zapišemo tudi v obliki:

$$S(\underline{x}) = \sum_{j=1}^4 [x_1^{a_{1j}} \quad x_2^{a_{2j}}] \cdot C_j = x_1^{a_{11}} x_2^{a_{21}} C_1 + x_1^{a_{12}} x_2^{a_{22}} C_2 + x_1^{a_{13}} x_2^{a_{23}} C_3 + x_1^{a_{14}} x_2^{a_{24}} C_4 \quad (7.105)$$

oziroma, če upoštevamo  $x_1 = D, \quad x_2 = Q$ , v obliki:

$$S(\underline{x}) = D^{a_{11}} Q^{a_{21}} C_1 + D^{a_{12}} Q^{a_{22}} C_2 + D^{a_{13}} Q^{a_{23}} C_3 + D^{a_{14}} Q^{a_{24}} C_4 \quad (7.106)$$

Če primerjamo istoležne člene izrazov (7.102) in (7.106), dobimo:

$$\begin{aligned} C_1 &= 100, \quad C_2 = 50, \quad C_3 = 20, \quad C_4 = 300 \\ a_{11} &= 1, \quad a_{21} = 0, \quad a_{12} = 2, \quad a_{22} = 0, \quad a_{13} = 0, \quad a_{23} = -1, \quad a_{14} = -5, \quad a_{24} = 2 \end{aligned} \quad (7.107)$$

Na osnovi koeficientov v izrazu (7.107.) lahko sestavimo matriko  $\underline{\underline{A}}$ , ki smo jo obravnavali v izrazih (7.26.) oz. (7.27.):

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & -1 & 2 \end{bmatrix} \quad (7.108)$$

Na osnovi pogoja ortogonalnosti (glej izraz (7.27.)) sledi:

$$\underline{\underline{A}} \cdot \underline{y}^* = \begin{bmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (7.109)$$

Na osnovi pogoja normiranosti (glej izraz (7.22.)) pa sledi:

$$\sum_{j=1}^4 y_j^* = 1 \Rightarrow y_1^* + y_2^* + y_3^* + y_4^* = 1 \quad (7.110)$$

Preverimo, s kakšno stopnjo težavnosti  $T$  reševanja problema imamo opravka (glej izraz (7.28.)):

$$\left. \begin{array}{l} N = 4 \\ n + 1 = 2 + 1 = 3 \end{array} \right\} \Rightarrow N - (n + 1) = 4 - 3 = T = 1 \quad (7.111)$$

Torej imamo opravka z 1. stopnjo težavnosti GP problema.

Če izraza (7.109) in (7.110) združimo, dobimo:

$$\begin{bmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & -1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (7.112)$$

Ta sistem moramo sedaj rešiti, da dobimo  $y_i^*$ . Vendar pa ima sistem v izrazu (7.112.) 3 enačbe in 4 neznanke. Torej, ker je stopnja težavnosti enaka  $T=1$ , direktne rešitve za  $y_i^*$  ne moremo dobiti. Zato bomo morali ta sistem rešiti na drugačen način.

V ta namen najprej spremenljivke  $y_1^*, y_2^*$  in  $y_3^*$  v sistemu (7.112.) izrazimo s spremenljivko  $y_4^*$ . Pri tem dobimo:

$$\begin{aligned} E_1 : y_1^* + 2y_2^* - 5y_4^* &= 0 \\ E_2 : -y_3^* + 2y_4^* &= 0 \Rightarrow y_3^* = 2y_4^* \\ E_3 : y_1^* + y_2^* + y_3^* + y_4^* &= 1 \end{aligned} \quad (7.113)$$

Če enačbo  $E_2$  vstavimo v  $E_3$ , dobimo:

$$y_1^* + y_2^* + 2y_4^* + y_4^* = 1 \text{ oziroma } y_1^* + y_2^* + 3y_4^* = 1 \quad (7.114)$$

Ker se enačba  $E_1$  glasi:

$$y_1^* + 2y_2^* - 5y_4^* = 0 \quad (7.115)$$

lahko odštejemo (7.114) in (7.115) ter dobimo:

$$\begin{aligned} y_1^* + y_2^* + 3y_4^* - (y_1^* + 2y_2^* - 5y_4^*) &= 1 - 0 \\ -y_2^* + 8y_4^* &= 1 \Rightarrow y_2^* = 8y_4^* - 1 \end{aligned} \quad (7.116)$$

Dobljeno spremenljivko  $y_2^*$  nato vstavimo v  $E_1$  in dobimo:

$$\begin{aligned} y_1^* + 2(8y_4^* - 1) - 5y_4^* &= 0 \\ y_1^* + 11y_4^* &= 2 \Rightarrow y_1^* = 2 - 11y_4^* \end{aligned} \quad (7.117)$$

$y_1^*$ ,  $y_2^*$  in  $y_3^*$  torej lahko izrazimo s spremenljivko  $y_4^*$  na osnovi izrazov (7.113.), (7.116.) in (7.117.) na naslednji način:

$$\begin{aligned} y_1^* &= 2 - 11y_4^* \\ y_2^* &= 8y_4^* - 1 \\ y_3^* &= 2y_4^* \end{aligned} \quad (7.118)$$

Dualna ciljna funkcija (7.33.) dobi na osnovi izrazov (7.107.) in (7.118.) naslednjo obliko:

$$\begin{aligned} f^*(\underline{y}) &= \prod_{j=1}^4 \left( \frac{C_j}{y_j^*} \right)^{y_j^*} = \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4^*} \right)^{y_4^*} = \\ &= \left( \frac{100}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{50}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{20}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{300}{y_4^*} \right)^{y_4^*} = \\ &= (\text{izrazimo le } z \ y_4^*) = \\ &= \left( \frac{100}{2 - 11y_4^*} \right)^{2 - 11y_4^*} \cdot \left( \frac{50}{8y_4^* - 1} \right)^{8y_4^* - 1} \cdot \left( \frac{20}{2y_4^*} \right)^{2y_4^*} \cdot \left( \frac{300}{y_4^*} \right)^{y_4^*} = \\ &= g_1(y_4^*) \end{aligned} \quad (7.119)$$

pri čemer smo vstavili številčne vrednosti za  $C_1, C_2, C_3, C_4$  ter  $y_1^*, y_2^*$  in  $y_3^*$  izražene z  $y_4^*$ . Očitno je funkcija (7.119.) odvisna od spremenljivke  $y_4^*$ . Problem je torej sedaj poiskati takšno optimalno spremenljivko  $y_4^*$ , ki bo dala najbolj ustrezne rezultate.

Spomnimo se sedaj sklepa Cauchyve neenakosti, ki temelji na izrazih (7.95.) oz. (7.96.), kar pomeni:

$$\min_{\underline{x}} f(\underline{x}) = \max_{\underline{y}} f^*(\underline{y}) = \max_{\underline{y}} \ln f^*(\underline{y}) \quad (7.120)$$



Torej, če bomo maksimizirali  $f^*$  oz.  $\ln f^*$ , glede na optimalno spremenljivko  $y_{4\text{OPT}}^*$ , bomo hkrati dosegli tudi minimum primarne funkcije  $f(x)$  (ki je v našem primeru določena z izrazom (7.102.)). Torej bo optimalna spremenljivka  $y_4^*$  tista, pri kateri bo dosežen izraz (7.120.). To pa z drugimi besedami pomeni, da je potrebno optimirati funkcijo (7.119.) po spremenljivki  $y_4^*$ , tovrstna optimizacija pa bo gotovo dala najbolj ustrezen rezultat.

Kot se izkaže, je boljše (lažje) poiskati  $y_{4\text{OPT}}^*$  pri opazovanju funkcije  $\ln f^* = g_2(y_4^*)$ , kot pa funkcije  $f^* = g_1(y_4^*)$  v izrazu (7.119.). Tvorimo torej:

$$\begin{aligned} \ln f^* &= \ln\left(\frac{100}{2-11y_4^*}\right)^{2-11y_4^*} + \ln\left(\frac{50}{8y_4^*-1}\right)^{8y_4^*-1} + \ln\left(\frac{20}{2y_4^*}\right)^{2y_4^*} + \ln\left(\frac{300}{y_4^*}\right)^{y_4^*} = \\ &= (2-11y_4^*) \cdot [\ln 100 - \ln(2-11y_4^*)] + (8y_4^*-1) \cdot [\ln 50 - \ln(8y_4^*-1)] + \\ &\quad + (2y_4^*) \cdot [\ln 10 - \ln y_4^*] + y_4^* \cdot [\ln 300 - \ln y_4^*] \end{aligned} \quad (7.121)$$

Optimalno vrednost spremenljivke  $y_{4\text{OPT}}^*$  poiščemo z enačenjem odvoda izraza (7.121.) z ničlo:

$$\frac{\partial}{\partial y_4^*}(\ln f^*) = \frac{d}{dy_4^*}(\ln f^*) = 0 \quad (7.122)$$

Sledi:

$$\begin{aligned} \frac{d(\ln f^*)}{dy_4^*} &= -11 \cdot [\ln 100 - \ln(2-11y_4^*)] + (2-11y_4^*) \cdot \left[ -\frac{-11}{2-11y_4^*} \right] + \\ &\quad + 8 \cdot [\ln 50 - \ln(8y_4^*-1)] + (8y_4^*-1) \cdot \left[ \frac{-8}{8y_4^*-1} \right] + \\ &\quad + 2 \cdot [\ln 10 - \ln y_4^*] + 2y_4^* \cdot \frac{-1}{y_4^*} + [\ln 300 - \ln y_4^*] + y_4^* \cdot \frac{-1}{y_4^*} = \end{aligned} \quad (7.123)$$

$$\begin{aligned}
&= -11 \cdot \ln \frac{100}{2-11y_4^*} + 11 + 8 \cdot \ln \frac{50}{8y_4^* - 1} - 8 + 2 \ln \frac{10}{y_4^*} - 2 + \ln \frac{300}{y_4^*} - 1 = \\
&= \left( \ln \frac{100}{2-11y_4^*} \right)^{-11} + \left( \ln \frac{50}{8y_4^* - 1} \right)^8 + \left( \ln \frac{10}{y_4^*} \right)^2 + \ln \frac{300}{y_4^*} = \\
&= \ln \left[ \frac{(2-11y_4^*)^{11}}{100^{11}} \cdot \frac{50^8}{(8y_4^* - 1)^8} \cdot \frac{100}{(y_4^*)^2} \cdot \frac{300}{y_4^*} \right] = \\
&= \ln \left[ \frac{(2-11y_4^*)^{11}}{(8y_4^* - 1)^8 \cdot (y_4^*)^2 \cdot y_4^*} \cdot \frac{50^8 \cdot 100 \cdot 300}{100^{11}} \right] = \\
&= \ln \left[ \frac{(2-11y_4^*)^{11} \cdot 4}{(8y_4^* - 1)^8 \cdot (2y_4^*)^2 \cdot y_4^*} \cdot \left( \frac{100^{11}}{50^8 \cdot 100 \cdot 300} \right)^{-1} \right] = \\
&= \ln \left[ \frac{(2-11y_4^*)^{11}}{(8y_4^* - 1)^8 \cdot (2y_4^*)^2 \cdot y_4^*} \right] + \ln \left[ \left( \frac{1}{4} \right)^{-1} \cdot \left( \frac{100^{11}}{50^8 \cdot 100 \cdot 300} \right)^{-1} \right] = \\
&= \ln \left[ \frac{(2-11y_4^*)^{11}}{(8y_4^* - 1)^8 \cdot (2y_4^*)^2 \cdot y_4^*} \right] - \ln \left[ \frac{100^{11}}{4 \cdot 50^8 \cdot 100 \cdot 300} \right] = \\
&= \ln \left[ \frac{(2-11y_4^*)^{11}}{(8y_4^* - 1)^8 \cdot (2y_4^*)^2 \cdot y_4^*} \right] - \ln \underbrace{\left[ \frac{100^{11}}{50^8 \cdot (20)^2 \cdot 300} \right]}_{2133.3} = 0
\end{aligned}$$

Dobimo torej naslednjo enačbo:

$$\frac{(2-11y_4^*)^{11}}{(8y_4^* - 1)^8 \cdot (2y_4^*)^2 \cdot y_4^*} = 2133.3 \quad (7.124)$$

Rešiti enačbo (7.124) ni enostavno. Kot pa se izkaže, lahko s pomočjo tako imenovanega »TRIAL AND ERROR« postopka ([21]) dobimo naslednji rezultat:

$$y_{4\text{OPT}}^* \approx 0.147 \quad (7.125)$$

Dobljena spremenljivka (7.125.) bo gotovo optimalna, saj bo optimirala funkcijo (7.121.), kar pa zaradi izraza (7.120.) pomeni tudi optimiranje dualne funkcije (7.119.) oz. primarne ciljne funkcije (7.102.).

Na osnovi izrazov (7.118.) sedaj lahko izrazimo še preostale optimalne dualne spremenljivke  $y_{1\text{OPT}}^*$ ,  $y_{2\text{OPT}}^*$ ,  $y_{3\text{OPT}}^*$ :

$$\begin{aligned} y_{1\text{OPT}}^* &= 2 - 11y_{4\text{OPT}}^* = 2 - 11 \cdot 0.147 = 0.383 \\ y_{2\text{OPT}}^* &= 8y_{4\text{OPT}}^* - 1 = 8 \cdot 0.147 - 1 = 0.176 \\ y_{3\text{OPT}}^* &= 2y_{4\text{OPT}}^* = 2 \cdot 0.147 = 0.294 \end{aligned} \quad (7.126)$$

Če optimalne dualne spremenljivke v izrazih (7.125.) in (7.126.) upoštevamo v dualni funkciji (7.119.), dobimo njeno optimalno vrednost:

$$\begin{aligned} f_{\text{OPT}}^* &= \left(\frac{100}{0.383}\right)^{0.383} \cdot \left(\frac{50}{0.176}\right)^{0.176} \cdot \left(\frac{20}{0.294}\right)^{0.294} \cdot \left(\frac{300}{0.147}\right)^{0.147} = \\ &= 8.426 \cdot 2.702 \cdot 3.457 \cdot 3.06 = 241.294 \end{aligned} \quad (7.127)$$

Vsekakor smo pri tem hkrati dobili tudi optimalno vrednost primarne ciljne funkcije  $f$ , ki je v našem primeru določena z izrazom (7.102). Seveda bi do enakega rezultata prišli tudi, če bi vstavili  $y_{4\text{OPT}}^* = 0.147$  v izraz (7.119), kjer je dana le odvisnost od spremenljivke  $y_4^*$ .

Izračunati moramo še optimalni primarni odločitveni spremenljivki  $x_1 = D$  in  $x_2 = Q$ . Denimo se odločimo, da ju bomo izpeljali na poenostavljen način preko »manevriranja« izrazov, ki tvorijo delne funkcije ciljne funkcije (7.102.).

Na osnovi izraza (7.102) lahko zapišemo za delne funkcije:

$$\begin{aligned} U_1^* &= 100 \cdot (D^*)^1 \cdot (Q^*)^0 = 100D^* \\ U_2^* &= 50 \cdot (D^*)^2 \cdot (Q^*)^0 = 50D^{*2} \\ U_3^* &= 20 \cdot (D^*)^0 \cdot (Q^*)^{-1} = \frac{20}{Q^*} \\ U_4^* &= 300 \cdot (D^*)^{-5} \cdot (Q^*)^2 = 300 \frac{Q^{*2}}{D^{*5}} \end{aligned} \quad (7.128)$$

Hkrati pa vemo, da velja tudi relacija na osnovi izraza (7.36.):

$$U_j^* = y_j^* \cdot f^*, \quad j = 1, \dots, N = 4 \quad (7.129)$$

kar lahko z upoštevanjem izrazov (7.125.), (7.126.) in (7.127.) zapišemo tudi kot:

$$\begin{aligned}
 U_1^* &= y_1^* \cdot f^* = y_{1\text{OPT}}^* \cdot f^* = 0.383 \cdot 241.294 = 92.41 \\
 U_2^* &= y_2^* \cdot f^* = y_{2\text{OPT}}^* \cdot f^* = 0.176 \cdot 241.294 = 42.467 \\
 U_3^* &= y_3^* \cdot f^* = y_{3\text{OPT}}^* \cdot f^* = 0.294 \cdot 241.294 = 70.94 \\
 U_4^* &= y_4^* \cdot f^* = y_{4\text{OPT}}^* \cdot f^* = 0.147 \cdot 241.294 = 35.47
 \end{aligned}
 \tag{7.130}$$

Če izenačimo tiste izraze v (7.128) in (7.130), ki jih potrebujemo za določitev optimalnih primarnih spremenljivk, dobimo:

$$\begin{aligned}
 U_1^* = 100D^* = 92.41 &\Rightarrow D^* = 0.9241\text{cm} \\
 U_3^* = \frac{20}{Q^*} = 70.94 &\Rightarrow Q^* = 0.281 \frac{\text{m}^3}{\text{s}}
 \end{aligned}
 \tag{7.131}$$

Odgovor na vprašanje, kakšna optimalna premer cevi in pretok izbrati, da bodo skupni stroški minimalni, bi se torej glasil:

Če bo premer cevi  $D^*$  enak 0.9241 cm in pretok  $Q^*$  enak 0.281 m<sup>3</sup>/s , bodo skupni stroški instalacije in uporabe črpalke rezervoarja najmanjši.

### 7.1.3.2 Pristop k reševanju problemov višje stopnje z odvajanjem dualnih odločitvenih spremenljivk

Poglejmo si še poenostavljen pristop k reševanju problemov višje stopnje z odvajanjem dualnih odločitvenih spremenljivk. Metodologijo tega pristopa bomo osvetlili na istem primeru, ki smo ga ravnokar obravnavali (glej izraze (7.98.) do (7.102.)). Tudi izpeljava vse do izraza (7.118.) poteka na popolnoma enak način kot prej.

Pri tem pristopu gre za to, kako lažje priti do izraza (7.124) ter se izogniti kompliciranemu postopku odvajanja logaritma, to je izpeljavi enačbe (7.122.), ki jo zaradi preglednosti napišimo še enkrat:

$$\frac{d \ln f^*}{dy_4^*} = 0
 \tag{7.132}$$

Zapišimo še enkrat povezave, ki smo jih dobili v izrazu (7.118), na naslednji način:

$$\begin{aligned}
y_1^* &= 2 - 11y_4^* \\
y_2^* &= 8y_4^* - 1 \\
y_3^* &= 2y_4^* \\
y_4^* &= y_4^*
\end{aligned}
\tag{7.133}$$

Namesto, da bi tvorili izraze (7.121.) do (7.123.), se bomo izpeljave izraza (7.124.) raje lotili tako, da bomo izvedli operacijo odvajanja dualnih odločitvenih spremenljivk ter nato z uporabo posebnega izraza (glej [17]) prišli po krajši poti do izraza (7.124.).

V ta namen najprej izvedimo operacijo odvajanja vseh spremenljivk v izrazu (7.133.) po spremenljivki  $y_4^*$ , pri čemer dobimo:

$$\begin{aligned}
\frac{dy_1^*}{dy_4^*} &= -11 \\
\frac{dy_2^*}{dy_4^*} &= 8 \\
\frac{dy_3^*}{dy_4^*} &= 2 \\
\frac{dy_4^*}{dy_4^*} &= 1
\end{aligned}
\tag{7.134}$$

Spomnimo se, da smo imeli v primeru dane naslednje koeficiente primarne ciljne funkcije, zapisane že v izrazu (7.107.):

$$C_1 = 100, \quad C_2 = 50, \quad C_3 = 20, \quad C_4 = 300 \tag{7.135}$$

Dokazati se da (glej [17]), da velja naslednja relacija, ki povezuje dualne odločitvene spremenljivke (7.133.), odvode (7.134.) in koeficiente (7.135.):

$$\left(y_1^*\right)^{\frac{dy_1^*}{dy_4^*}} \cdot \left(y_2^*\right)^{\frac{dy_2^*}{dy_4^*}} \cdot \left(y_3^*\right)^{\frac{dy_3^*}{dy_4^*}} \cdot \left(y_4^*\right)^{\frac{dy_4^*}{dy_4^*}} = \left(C_1^*\right)^{\frac{dy_1^*}{dy_4^*}} \cdot \left(C_2^*\right)^{\frac{dy_2^*}{dy_4^*}} \cdot \left(C_3^*\right)^{\frac{dy_3^*}{dy_4^*}} \cdot \left(C_4^*\right)^{\frac{dy_4^*}{dy_4^*}} \tag{7.136}$$

ki nam da naslednji izraz:

$$\left(2 - 11y_4^*\right)^{-11} \cdot \left(8y_4^* - 1\right)^8 \cdot \left(2y_4^*\right)^2 \cdot \left(y_4^*\right)^1 = 100^{-11} \cdot 50^8 \cdot 20^2 \cdot 300^1 \tag{7.137}$$

kar lahko zapišemo tudi kot:

$$\frac{(2-11y_4^*)^{11}}{(8y_4^*-1)^8 \cdot (2y_4^*)^2 \cdot (y_4^*)^1} = \frac{100^{11}}{\underbrace{50^8 \cdot 20^2 \cdot 300^1}_{2^{133.3}}} \quad (7.138)$$

Torej očitno dobimo popolnoma enak rezultat kot v izrazu (7.124), le da na dokaj krajši in lažji način. Nadaljnji postopek nato seveda poteka popolnoma enako kot prej (glej izraze (7.125.) do (7.131.)).

Uporabo dotičnega pristopa z odvajanjem dualnih odločitvenih spremenljivk si pogledjmo še na enem primeru, kar nam bo pomagalo okrepiti razumevanje tega pristopa.

### **Primer:**

Poiščite minimum za naslednjo funkcijo:

$$\begin{aligned} f(x_1, x_2, x_3) &= 2x_1x_2 + \frac{5}{x_1^2x_3} + 4x_2x_3 + \frac{x_1}{x_2} + x_3^2 = \\ &= \underbrace{2x_1^1x_2^1x_3^0}_{u_1} + \underbrace{5x_1^{-2}x_2^0x_3^{-1}}_{u_2} + \underbrace{4x_1^0x_2^1x_3^1}_{u_3} + \underbrace{x_1^1x_2^{-1}x_3^0}_{u_4} + \underbrace{x_1^0x_2^0x_3^2}_{u_5} \end{aligned} \quad (7.139)$$

Koeficienti  $C_j$  so enaki:

$$C_1 = 2, C_2 = 5, C_3 = 4, C_4 = 1, C_5 = 1 \quad (7.140)$$

Matriko  $\underline{\underline{A}}$  lahko neposredno zapišemo, če koeficiente v delnih funkcijah funkcije  $f$  gledamo kot stolpce:

$$\underline{\underline{A}} = \begin{bmatrix} 1 & -2 & 0 & 1 & 0 \\ 1 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 2 \end{bmatrix} \quad (7.141)$$

kar je krajši način tvorjenja te matrike glede na prejšnje primere.

Na osnovi pogoja ortogonalnosti (glej izraz (7.27.)) sledi:

$$\underline{A} \cdot \underline{y}^* = \begin{bmatrix} 1 & -2 & 0 & 1 & 0 \\ 1 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (7.142)$$

Na osnovi pogoja normiranosti (glej izraz (7.22.)) pa sledi:

$$\sum_{i=1}^{N=5} y_i^* = 1 \Rightarrow y_1^* + y_2^* + y_3^* + y_4^* + y_5^* = 1 \quad (7.143)$$

Preverimo, s kakšno stopnjo težavnosti T reševanja problema imamo opravka (glej izraz (7.28.)):

$$\left. \begin{array}{l} N = 5 \\ n + 1 = 3 + 1 = 4 \end{array} \right\} \Rightarrow T = N - (n + 1) = 5 - 4 = 1 \quad (7.144)$$

Opravka imamo z 1. stopnjo težavnosti GP problema.

Če združimo pogoja ortogonalnosti (7.142.) in normiranosti (7.143.), dobimo:

$$\begin{bmatrix} 1 & -2 & 0 & 1 & 0 \\ 1 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 2 \\ -1 & 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (7.145)$$

Očitno imamo v sistemu (7.145.) opravka 4. enačbami in 5. neznankami. Podobno kot v prejšnjem primeru, moramo izraziti z eno dualno spremenljivko ostale dualne spremenljivke. Denimo si izberemo, da bomo spremenljivke  $y_1^*$ ,  $y_2^*$ ,  $y_3^*$  in  $y_4^*$  izrazili s spremenljivko  $y_5^*$ . V ta namen najprej izraz (7.145.) zapišimo v naslednji obliki:

$$\begin{aligned} E_1 : y_1^* - 2y_2^* + y_4^* &= 0 & \Rightarrow y_1^* &= 2y_2^* - y_4^* \\ E_2 : y_1^* + y_3^* - y_4^* &= 0 & \Rightarrow y_1^* &= -y_3^* + y_4^* \\ E_3 : -y_2^* + y_3^* + 2y_5^* &= 0 & \Rightarrow y_5^* &= \frac{y_2^* - y_3^*}{2} \\ E_4 : y_1^* + y_2^* + y_3^* + y_4^* + y_5^* &= 1 \end{aligned} \quad (7.146)$$

Nato izenačimo enačbi  $E_1$  in  $E_2$  v izrazu (7.46.):

$$y_1^* = 2y_2^* - y_4^* = -y_3^* + y_4^* \quad \Rightarrow \quad y_4^* = \frac{2y_2^* + y_3^*}{2} \quad (7.147)$$

Vstavimo rezultat v izrazu (7.147) nazaj v enačbo  $E_1$ , pri čemer dobimo:

$$y_1^* = 2y_2^* - \frac{2y_2^* + y_3^*}{2} = y_2^* - \frac{y_3^*}{2} \quad (7.148)$$

Vstavimo izraza (7.148), (7.147) in rezultat za enačbo  $E_3$  v enačbo  $E_4$ . Dobimo:

$$\begin{aligned} \left( y_2^* - \frac{y_3^*}{2} \right) + y_2^* + y_3^* + \frac{2y_2^* + y_3^*}{2} + \frac{y_2^* - y_3^*}{2} &= 1 \quad | \cdot 2 \\ 2y_2^* - y_3^* + 2y_2^* + 2y_3^* + 2y_2^* + y_3^* + y_2^* - y_3^* &= 2 \\ 7y_2^* + y_3^* = 2 \quad \Rightarrow \quad y_3^* &= 2 - 7y_2^* \end{aligned} \quad (7.149)$$

Vstavimo dobljeni rezultat za  $y_3^*$  v enačbo  $E_3$  in dobimo:

$$y_5^* = \frac{y_2^* - (2 - 7y_2^*)}{2} \Rightarrow 2y_5^* = 8y_2^* - 2 \Rightarrow y_2^* = \frac{2 + 2y_5^*}{8} = \frac{1 + y_5^*}{4} \quad (7.150)$$

Tako smo uspeli izraziti prvo dualno spremenljivko z  $y_5^*$ .

V nadaljevanju vstavimo rezultat za izraz (7.150.) v rezultat za izraz (7.149.), pri čemer dobimo:

$$y_3^* = 2 - 7 \left( \frac{1 + y_5^*}{4} \right) = \frac{1}{4} - \frac{7}{4} y_5^* \Rightarrow \frac{1 - 7y_5^*}{4} = y_3^* \quad (7.151)$$

Tako smo uspeli izraziti drugo dualno spremenljivko z  $y_5^*$ .

Vstavimo izraza (7.150) in (7.151) v izraz (7.147), pri čemer dobimo:

$$y_4^* = \frac{1 + y_5^*}{4} + \frac{1}{2} \left( \frac{1 - 7y_5^*}{4} \right) = \frac{2 + 2y_5^*}{8} + \frac{1 - 7y_5^*}{8} = \frac{3 - 5y_5^*}{8} = y_4^* \quad (7.152)$$

Tako smo uspeli izraziti tretjo dualno spremenljivko z  $y_5^*$ .

Če vstavimo izraza (7.150) in (7.152) v enačbo  $E_1$  izraza (7.146.), dobimo:



$$y_1^* = 2 \cdot \frac{1+y_5^*}{4} - \frac{3-5y_5^*}{8} = \frac{4+4y_5^* - 3+5y_5^*}{8} = \frac{1+9y_5^*}{8} = y_1^* \quad (7.153)$$

Tako smo uspeli izraziti še četrto dualno spremenljivko z  $y_5^*$ .

Preizkusimo, če se nismo kje zmotili (izraze (7.150), (7.151), (7.152) in (7.153) vstavimo v enačbo  $E_4$  izraza (7.146.) in mora priti rezultat enak 1):

$$\begin{aligned} & \underbrace{\frac{1+9y_5^*}{8}}_{y_1^*} + \underbrace{\frac{1+y_5^*}{4}}_{y_2^*} + \underbrace{\frac{1-7y_5^*}{4}}_{y_3^*} + \underbrace{\frac{3-5y_5^*}{8}}_{y_4^*} + y_5^* = \\ & = \frac{1+9y_5^* + 2+2y_5^* - 2-14y_5^* + 3-5y_5^* + 8y_5^*}{8} = \frac{8+0}{8} = 1 \end{aligned} \quad (7.154)$$

Do tega trenutka smo uspeli vse dualne spremenljivke izraziti s prosto dualno spremenljivko  $y_5^*$ . V nadaljevanju moramo poiskati optimalno vrednost dualne ciljne funkcije. V ta namen bomo uporabili pristop z odvajanjem dualnih odločitvenih spremenljivk, kot smo to storili v prejšnjem primeru (glej prejšnjo nalogo). S tem se bomo torej izognili dokaj kompliciranemu postopku odvajanja logaritma, to je operaciji

$$\frac{d \ln f^*}{dy_5^*} = 0 \quad (\text{o čemer govori poglavje 7.1.3.1.}).$$

Če odvajamo rezultate v izrazih (7.150.) do (7.153.) po spremenljivki  $y_5^*$ , dobimo:

$$\begin{array}{l} y_1^* = \frac{1+9y_5^*}{8} \\ y_2^* = \frac{1+y_5^*}{4} \\ y_3^* = \frac{1-7y_5^*}{4} \\ y_4^* = \frac{3-5y_5^*}{8} \\ y_5^* = y_5^* \end{array} \Rightarrow \begin{array}{l} \text{Če to odvajamo} \\ \text{po } y_5^*, \text{ dobimo} \end{array} \Rightarrow \begin{array}{l} \frac{dy_1^*}{dy_5^*} = \frac{9}{8} = P_1 \\ \frac{dy_2^*}{dy_5^*} = \frac{1}{4} = P_2 \\ \frac{dy_3^*}{dy_5^*} = -\frac{7}{4} = P_3 \\ \frac{dy_4^*}{dy_5^*} = -\frac{5}{8} = P_4 \\ \frac{dy_5^*}{dy_5^*} = 1 = P_5 \end{array} \quad (7.155)$$

Podobno kot v prejšnjem primeru (prejšnji nalogi) se da dokazati, da velja naslednja relacija, ki povezuje dualne odločitvene spremenljivke in odvode iz izrazov (7.155.), ter koeficiente (7.140.):

$$(y_1^*)^{P_1} \cdot (y_2^*)^{P_2} \cdot (y_3^*)^{P_3} \cdot (y_4^*)^{P_4} \cdot (y_5^*)^{P_5} = C_1^{P_1} \cdot C_2^{P_2} \cdot C_3^{P_3} \cdot C_4^{P_4} \cdot C_5^{P_5} \quad (7.156)$$

Tako dobimo:

$$\begin{aligned} & \left( \frac{1+9y_5^*}{8} \right)^{\frac{9}{8}} \cdot \left( \frac{1+y_5^*}{4} \right)^{\frac{1}{4}} \cdot \left( \frac{1-7y_5^*}{4} \right)^{\frac{7}{4}} \cdot \left( \frac{3-5y_5^*}{8} \right)^{\frac{5}{8}} \cdot y_5^{*1} = \\ & = 2^{\frac{9}{8}} \cdot 5^{\frac{1}{4}} \cdot 4^{\frac{7}{4}} \cdot 1^{\frac{5}{8}} \cdot 1^1 \end{aligned} \quad (7.157)$$

V nadaljevanju moramo poiskati iz izraza (7.157.) pravo rešitev za spremenljivko  $y_5^*$ . Kot v prejšnjem primeru se izkaže, da je to vse prej kot preprosto opravilo. Tako moramo tudi tukaj za pomoč uporabiti numerične metode, pri čemer dobimo naslednji rezultat [17]:

$$y_{5_{\text{OPT}}}^* = 0.053423 \quad (7.158)$$

Če nato ta rezultat vstavimo v izraze (7.150.) do (7.153.) (oz. v dualne spremenljivke v (7.155.)), dobimo optimalno rešitev še za preostale dualne spremenljivke:

$$\begin{aligned} y_{1_{\text{OPT}}}^* &= \frac{1+9y_{5_{\text{OPT}}}^*}{8} = 0.1851 \\ y_{2_{\text{OPT}}}^* &= \frac{1+y_{5_{\text{OPT}}}^*}{4} = 0.2633 \\ y_{3_{\text{OPT}}}^* &= \frac{1-7y_{5_{\text{OPT}}}^*}{4} = 0.1565 \\ y_{4_{\text{OPT}}}^* &= \frac{3-5y_{5_{\text{OPT}}}^*}{8} = 0.3416 \end{aligned} \quad (7.159)$$

Če nato optimalne rešitve za dualne spremenljivke v (7.158.), (7.159.), ter koeficiente (7.140.) vstavimo v izraz za ciljno funkcijo (7.33.), dobimo za slednjo naslednjo optimalno rešitev:

$$f_{\text{OPT}}^* = \prod_{j=1}^5 \left( \frac{C_j}{y_{j\text{OPT}}^*} \right)^{y_{j\text{OPT}}^*} = \dots = 9.4539 \quad (7.160)$$

Nato pa je možno (npr. z »manipuliranjem« preko delnih funkcij  $U_j$  funkcije (7.139.) – glej prejšnja poglavja) dobiti še naslednje optimalne primarne odločitvene spremenljivke:

$$\begin{aligned} x_1^* &= 1.681 \\ x_2^* &= 0.520 \\ x_3^* &= 0.710 \end{aligned} \quad (7.161)$$

S tem primerom smo zaključili obravnavo problemov geometrijskega programiranja brez omejitev. V nadaljevanju pa si na kratko pogledjmo še problematiko nekaterih osnovnih problemov geometrijskega programiranja z omejitvami.

## 7.2 Geometrijsko programiranje z omejitvami

V tem poglavju si bomo najprej pogledali kratke teoretične osnove problematike geometrijskega programiranja z omejitvami. Nato pa bomo problematiko dodatno osvetlili na dveh primerih, od katerih bo en primer tudi aplikativno obarvan.

Splošna formulacija problemov geometrijskega programiranja z omejitvami se podobno kot pri problemih brez omejitev (glej strukturo primarne ciljne funkcije v izrazu (7.6.)) glasi:

$$\min f(\underline{x}) = \sum_{j=1}^N C_j \prod_{i=1}^n x_i^{a_{ij}}, \quad C_j > 0, \quad a_{ij} \in \mathfrak{R}, \quad \underline{x} = \{x_i\} > 0 \quad (7.162)$$

P.O.

$$g_K(\underline{x}) \leq 1, \quad K = 1, \dots, P \quad (7.163)$$

Torej želimo poiskati optimum (npr. minimum) ciljne funkcije s strukturo, kot je zapisana v izrazu (7.162.), pri čemer pa imamo opravka tudi s  $P$  omejitvami, podanimi v izrazu (7.163.). Le-te lahko zapišemo tudi na naslednji pregleden način:

$$\begin{aligned}
g_1(\underline{x}) &\leq 1 \\
g_2(\underline{x}) &\leq 1 \\
&\dots \\
g_p(\underline{x}) &\leq 1
\end{aligned}
\tag{7.164}$$

Kot pri problematiki brez omejitev (glej izraz (7.1.)) je ciljna funkcija  $f(\underline{x})$  lahko podana tudi v obliki strukture iz  $N$  členov delnih funkcij  $U_j$ , to je:

$$f(\underline{x}) = U_1(\underline{x}) + U_2(\underline{x}) + \dots + U_N(\underline{x}) = \sum_{j=1}^N U_j(\underline{x}) \tag{7.165}$$

Tudi omejitve v izrazu (7.164.) lahko zapišemo v obliki delnih funkcij. Tako je npr. prva omejitev sestavljena iz  $K_1$  členov (delnih funkcij):

$$g_1(\underline{x}) = U_{N+1}(\underline{x}) + U_{N+2}(\underline{x}) + \dots + U_{N+K_1}(\underline{x}) \tag{7.166}$$

Druga omejitev je npr. sestavljena iz  $K_2$  členov (delnih funkcij):

$$\begin{aligned}
g_2(\underline{x}) &= U_{N+K_1+1}(\underline{x}) + U_{N+K_1+2}(\underline{x}) + \dots + U_{N+K_1+K_2}(\underline{x}) \\
&\dots
\end{aligned}
\tag{7.167}$$

Itn...

$P$ -ta omejitev pa je npr. sestavljena iz  $K_p$  členov (delnih funkcij):

$$g_p(\underline{x}) = U_{N+K_1+K_2+\dots+K_{p-1}+1}(\underline{x}) + \dots + U_{N+K_1+K_2+\dots+K_{p-1}+K_p}(\underline{x}) \tag{7.168}$$

Dokazati se da, da za optimalno dualno ciljno funkcijo velja izraz, katerega del strukture je zelo podoben kot pri problematiki brez omejitev (glej izraz (7.33.)), del strukture pa se pojavi dodatno kot posledica omejitev [17]:

$$\begin{aligned}
f^*(\underline{y}^*) &= \prod_{j=1}^{\tilde{N}=N+K_1+K_2+\dots+K_p} \left( \frac{C_j}{y_j^*} \right)^{y_j^*} \cdot \prod_{K=1}^P (\lambda_K^*)^{\lambda_K^*} = \\
&= \underbrace{\left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \dots \cdot \left( \frac{C_{\tilde{N}}}{y_{\tilde{N}}^*} \right)^{y_{\tilde{N}}^*}}_{\text{PODOBNO KOT BREZ OMEJITEV}} \cdot \underbrace{(\lambda_1^*)^{\lambda_1^*} \cdot (\lambda_2^*)^{\lambda_2^*} \cdot \dots \cdot (\lambda_p^*)^{\lambda_p^*}}_{\text{NOVO ZARADI OMEJITEV}},
\end{aligned}
\tag{7.169}$$

kjer za parametre  $\lambda_i$ , ki jih pri problematiki brez omejitev ni bilo, velja:

$$\begin{aligned}
 \lambda_1^* &= y_{N+1}^* + y_{N+2}^* + \dots + y_{N+K_1}^* \rightarrow \text{Za 1. omejitev} \\
 \lambda_2^* &= y_{N+K_1+1}^* + y_{N+K_1+2}^* + \dots + y_{N+K_1+K_2}^* \rightarrow \text{Za 2. omejitev} \\
 &\dots \\
 \lambda_P^* &= y_{N+K_1+K_2+\dots+K_{P-1}+1}^* + y_{N+K_1+K_2+\dots+K_{P-1}+2}^* + \dots + y_{N+K_1+K_2+\dots+K_{P-1}+K_P}^* \rightarrow \text{Za } P. \\
 &\hspace{15em} \text{omejitev.}
 \end{aligned} \tag{7.170}$$

torej jih tvorimo kot določene kombinacije dualnih odločitvenih spremenljivk (več o tem bo razvidno v kasnejših primerih).

Kot je razvidno iz izraza (7.169.), se v njem pojavi koeficient  $\tilde{N} = N + K_1 + K_2 + \dots + K_P$ , kjer se upošteva vsota členov tako delnih funkcij funkcije (7.165.), kot tudi vsota vseh členov, ki tvorijo vse omejitve (7.164.).

Primarne odločitvene spremenljivke dobimo podobno kot pri situaciji brez omejitev (glej izraz (7.36.)), s pomočjo manipuliranja z izrazi:

$$y_j^* = \frac{U_j^*(\underline{x})}{f^*} \quad \text{oziroma} \quad U_j^*(\underline{x}) = f^* \cdot y_j^*, \quad j = 1, \dots, N \tag{7.171}$$

pri čemer pa se pojavijo še dodatne relacije zaradi omejitev. Tako za 1. omejitev velja:

$$\begin{aligned}
 U_{N+1}^*(\underline{x}) &= \frac{y_{N+1}^*}{\lambda_1^*} \\
 U_{N+2}^*(\underline{x}) &= \frac{y_{N+2}^*}{\lambda_1^*} \\
 &\dots \\
 U_{N+K_1}^*(\underline{x}) &= \frac{y_{N+K_1}^*}{\lambda_1^*}
 \end{aligned} \tag{7.172}$$

Za 2. omejitev velja:

$$\begin{aligned}
U_{N+K_1+1}^*(\underline{x}) &= \frac{y_{N+K_1+1}^*}{\lambda_2^*} \\
U_{N+K_1+2}^*(\underline{x}) &= \frac{y_{N+K_1+2}^*}{\lambda_2^*} \\
&\dots \\
U_{N+K_1+K_2}^*(\underline{x}) &= \frac{y_{N+K_1+K_2}^*}{\lambda_2^*}
\end{aligned} \tag{7.173}$$

Itn...

Za  $P$ -to omejitev pa velja:

$$\begin{aligned}
U_{N+K_1+K_2+\dots+K_{p-1}+1}^*(\underline{x}) &= \frac{y_{N+K_1+K_2+\dots+K_{p-1}+1}^*}{\lambda_p^*} \\
U_{N+K_1+K_2+\dots+K_{p-1}+2}^*(\underline{x}) &= \frac{y_{N+K_1+K_2+\dots+K_{p-1}+2}^*}{\lambda_p^*} \\
&\dots \\
U_{N+K_1+K_2+\dots+K_{p-1}+K_p}^*(\underline{x}) &= \frac{y_{N+K_1+K_2+\dots+K_{p-1}+K_p}^*}{\lambda_p^*}
\end{aligned} \tag{7.174}$$

Ko rešimo sistem, ki ga definirajo izrazi (7.171), (7.172), (7.173) in (7.174), dobimo optimalen vektor odločitvenih spremenljivk  $\underline{x} = \{x_1, \dots, x_n\}$ .

Sicer pa ostali atributi postopka reševanja (pogoja ortogonalnosti in normiranosti, stopnja težavnosti, problematika reševanja pri višjih stopnjah težavnosti z odvajanjem logaritmske dualne ciljne funkcije ali odvajanjem dualnih odločitvenih spremenljivk, itn.) potekajo podobno kot pri problematiki geometrijskega programiranja brez omejitev (glej poglavje 7.1.), več bo razvidno na primerih.

Omenimo še, da pri izpeljavi izrazov (7.169.) do (7.174.) ključno vlogo igrajo upoštevanje Cauchyve neenakosti ter metodologija z Lagrangeovimi multiplikatorji. Več o tovrstni problematiki si lahko bralec pogleda v literaturi [17].

V nadaljevanju si bomo pogledali dva primera, ki bosta še dodatno osvetlila razumevanje problematike geometrijskega programiranja z omejitvami.

**Primer:**

Minimizirajte naslednjo funkcijo:

$$f(\underline{x}) = \frac{x_1^2 x_2}{3} + \frac{x_3^2}{x_2} \quad (7.175)$$

P.O.

$$g(\underline{x}) = \frac{3x_2}{4x_1 x_3} + \frac{2}{x_3} \leq 1, \quad x_i > 0 \quad (7.176)$$

Očitno imamo opravka v izrazu (7.163.) z  $P=1$ , torej eno samo omejitvijo. Zapišimo dano funkcijo (7.175.) in njeno omejitvev (7.176.) v obliki delnih funkcij, primerni za reševanje z geometrijskim programiranjem.

Tako dobimo (primerjaj z izrazoma (7.165.) in (7.166.)):

$$f(\underline{x}) = \underbrace{\frac{1}{3} \cdot x_1^2 x_2^1 x_3^0}_{u_1} + \underbrace{1 \cdot x_1^0 x_2^{-1} x_3^2}_{u_2} \quad \left. \vphantom{f(\underline{x})} \right\} N = 2 \text{ (2 člena)} \quad (7.177)$$

$$g(\underline{x}) = \underbrace{\frac{3}{4} \cdot x_1^{-1} x_2^1 x_3^{-1}}_{u_3} + \underbrace{2 \cdot x_1^0 x_2^0 x_3^{-1}}_{u_4} \quad \left. \vphantom{g(\underline{x})} \right\} K_1 = 2 \text{ (2 člena)} \quad (7.178)$$

Matriko  $\underline{\underline{A}}$  lahko neposredno zapišemo, če koeficiente v delnih funkcijah funkcije  $f$  in omejitve  $g$  gledamo kot stolpce, pri čemer dobimo naslednji izraz za pogoj ortogonalnosti (7.27.):

$$\underline{\underline{A}} = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & -1 \end{bmatrix}; \quad \underline{\underline{A}} \cdot \underline{y}^* = \underline{0} \quad (7.179)$$

Očitno je bistvena razlika glede na problematiko brez omejitvev v tem, da smo v matriki  $A$  dodali poleg stolpcev, ki pripadajo ciljni funkciji, tudi stolpca, ki pripadata njeni omejitvi.

Za pogoj normiranosti (7.22.) velja, da ga vedno pri tovrstni problematiki pišemo le za tisti indeks vsote ( $N=2$ ), ki se tiče ciljne funkcije (torej omejitvev tukaj nikoli ne upoštevamo). Tako dobimo:

$$\sum_{j=1}^{N=2} y_j^* = 1 \quad \Rightarrow \quad y_1^* + y_2^* = 1 \quad (7.180)$$

Izraz (7.179.) lahko zapišemo tudi v obliki:

$$\begin{bmatrix} 2 & 0 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (7.181)$$

oziroma:

$$\begin{aligned} E_1: \quad 2y_1^* - y_3^* &= 0 \quad \Rightarrow \quad y_1^* = \frac{y_3^*}{2} \\ E_2: \quad y_1^* - y_2^* + y_3^* &= 0 \quad \Rightarrow \quad y_1^* = y_2^* - y_3^* \\ E_3: \quad 2y_2^* - y_3^* - y_4^* &= 0 \end{aligned} \quad (7.182)$$

Enačimo  $E_1$  in  $E_2$  in dobimo:

$$\frac{y_3^*}{2} = y_2^* - y_3^* \quad \Rightarrow \quad y_3^* = \frac{2}{3} y_2^* \quad (7.183)$$

Iz izraza  $E_1$  (ali  $E_2$ ) in izraza (7.183) sledi:

$$y_1^* = \frac{1}{3} y_2^* \quad (7.184)$$

Iz  $E_3$  in izraza (7.183) sledi:

$$2y_2^* - \frac{2}{3} y_2^* = y_4^* \quad \Rightarrow \quad y_4^* = \frac{4}{3} y_2^* \quad (7.185)$$

Upoštevajmo izraz (7.184.) v pogoju normiranosti (7.180.), pri čemer dobimo:

$$y_1^* + y_2^* = 1 \quad \Rightarrow \quad \underbrace{\frac{1}{3} y_2^*}_{\text{glej 7.184}} + y_2^* = 1 \quad \Rightarrow \quad y_2^* = \frac{3}{4} \quad (7.186)$$



Izraz (7.186.) upoštevamo v izrazih (7.183), (7.184), (7.185), pri čemer dobimo še preostale optimalne dualne spremenljivke:

$$y_1^* = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}; \quad y_2^* = \frac{3}{4}; \quad y_3^* = \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{2}; \quad y_4^* = \frac{4}{3} \cdot \frac{3}{4} = 1 \quad (7.187)$$

Optimalna rešitev za vse dualne spremenljivke se torej glasi:

$$(y_1^*, y_2^*, y_3^*, y_4^*) = \left( \frac{1}{4}, \frac{3}{4}, \frac{1}{2}, 1 \right) \quad (7.188)$$

Optimalna dualna funkcija (7.169.) zavzame naslednjo obliko:

$$\begin{aligned} f^*(\underline{y}^*) &= \prod_{j=1}^{N+K_1=2+2} \left( \frac{C_j}{y_j} \right)^{y_j^*} \cdot \prod_{K=1}^{P=1} (\lambda_K^*)^{\lambda_K^*} = \\ &= \left( \frac{C_1}{y_1} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4} \right)^{y_4^*} \cdot (\lambda_1^*)^{\lambda_1^*}, \end{aligned} \quad (7.189)$$

kjer je  $\tilde{N} = N + K_1 = 2 + 2 = 4$  in za (edini) parameter  $\lambda_i$  ( $i = 1$ ) iz izraza (7.170.) velja:

$$\lambda_1^* = y_{\underbrace{N+1}_{=2}}^* + y_{\underbrace{N+K_1}_{=2}}^* = y_3^* + y_4^* \quad (7.190)$$

Če upoštevamo koeficiente  $C_i$  izrazov (7.177.) in (7.178.) ter optimalno rešitev za dualne spremenljivke (7.188.) v izrazih (7.189.) in (7.190.), dobimo naslednjo optimalno dualno ciljno funkcijo:

$$\begin{aligned} f^*(\underline{y}^*) &= \left( \frac{1}{\frac{3}{4}} \right)^{\frac{1}{4}} \cdot \left( \frac{1}{\frac{3}{4}} \right)^{\frac{3}{4}} \cdot \left( \frac{3}{\frac{1}{2}} \right)^{\frac{1}{2}} \cdot \left( \frac{2}{1} \right)^1 \cdot \left( \frac{1}{2} + 1 \right)^{\frac{1}{2}+1} = \\ &= \left( \frac{4}{3} \right)^{0.25} \cdot \left( \frac{4}{3} \right)^{0.75} \cdot \left( \frac{3}{2} \right)^{0.5} \cdot 2 \cdot (1.5)^{1.5} = \left( \frac{4}{3} \right)^1 \cdot 1.5^2 \cdot 2 = 6 \end{aligned} \quad (7.191)$$

Pri določitvi optimalnih primarnih odločitvenih spremenljivk upoštevamo izraze (7.171.), (7.177.), (7.188.) in (7.191.), pri čemer dobimo:

$$U_1^* = f^* \cdot y_1^* \Rightarrow \frac{x_1^{*2} \cdot x_2^*}{3} = 6 \cdot \frac{1}{4} \Rightarrow x_1^{*2} \cdot x_2^* = 4.5 \quad (7.192)$$

$$U_2^* = f^* \cdot y_2^* \Rightarrow \frac{x_3^{*2}}{x_2^*} = 6 \cdot \frac{3}{4} \Rightarrow x_3^{*2} = 4.5 \cdot x_2^* \quad (7.193)$$

Pojavijo se še dodatne relacije zaradi omejitev, pri čemer izrazi za edino omejitev (7.172.) z upoštevanjem izrazov (7.178.), (7.188.) in (7.190.), preidejo v obliko:

$$U_{N+1}^* = \frac{y_{N+1}^*}{\lambda_1^*} \Rightarrow U_3^* = \frac{3x_2^*}{4x_1^* x_3^*} = \frac{y_3^*}{y_3^* + y_4^*} = \frac{\frac{1}{2}}{\frac{1}{2} + 1} = \frac{1}{3} \quad (7.194)$$

$$U_{N+2}^* = \frac{y_{N+2}^*}{\lambda_1^*} \Rightarrow U_4^* = \frac{2}{x_3^*} = \frac{y_4^*}{y_3^* + y_4^*} = \frac{1}{\frac{1}{2} + 1} = \frac{2}{3} \quad (7.195)$$

Iz izraza (7.195) sledi:

$$x_3^* = 3 \quad (7.196)$$

Iz izraza (7.193) na osnovi izraza (7.196.) sledi:

$$3^2 = 4.5x_2^* \Rightarrow x_2^* = 2 \quad (7.197)$$

Iz izraza (7.192) na osnovi izraza (7.197.) pa sledi:

$$x_1^{*2} \cdot 2 = 4.5 \Rightarrow x_1^* = 1.5 \quad (7.198)$$

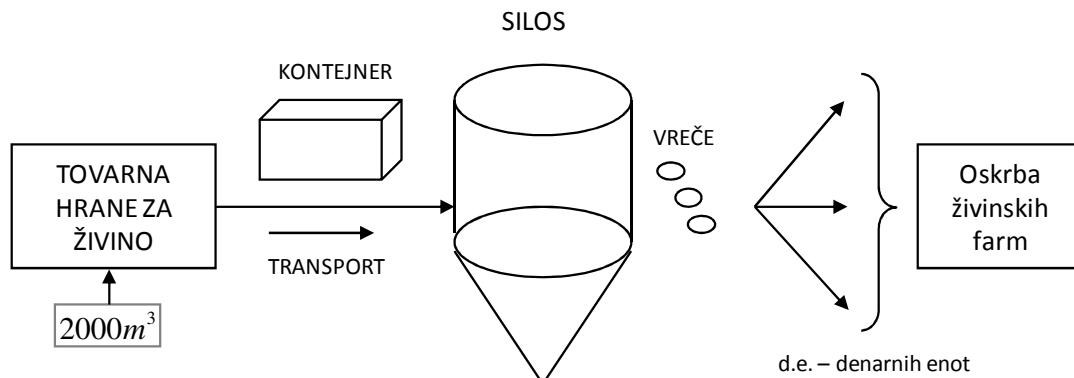
Optimalni vektor odločitvenih spremenljivk je torej na osnovi izrazov (7.196.) do (7.198.) naslednji:

$$\underline{x}^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} 1.5 \\ 2 \\ 3 \end{bmatrix} \quad (7.199)$$

V nadaljevanju si bomo pogledali še en primer, ki se tiče problematike geometrijskega programiranja z omejitvami. Za ta primer je značilno, da je aplikativno obarvan, saj izhaja iz čisto realnega problema iz logistične transportne prakse.

### 7.2.1 Aplikativni primer GP z omejitvami

Organizirati je potrebno transport  $2000m^3$  koncentrata za prehrano živine, ki ga bo tovarna hrane za živino proizvedla v določenem času. V ta namen je potrebno tudi zgraditi odprt kvadratast kontejner za transport hrane, ter odprt silos za polnjenje vreč s hrano (glej sliko 92).



Slika 92: Transport koncentrata hrane za živino

Pri tem se mora upoštevati naslednje:

- 1.) Stroški prevoza kontejnerja so 10 d.e./kontejner.
- 2.) Stroški izdelave kontejnerja so  $10 \frac{\text{d.e.}}{m^2}$ .
- 3.) Volumen kontejnerja ne sme biti večji od  $10m^3$ , razmerje širine proti višini pa ne sme biti večje od 5:4.
- 4.) Razmerje dolžine proti širini kontejnerja ne sme biti manjše od 4:1.
- 5.) Volumen cilindričnega dela silosa mora biti vsaj tako velik kot volumen kontejnerja.
- 6.) Kot konusnega dela silosa ne sme biti večji od  $60^\circ$ . Stroški izdelave cilindričnega dela so  $15 \frac{\text{d.e.}}{m^2}$ , konusa pa  $20 \frac{\text{d.e.}}{m^2}$ .
- 7.) Organizator tega posla razpolaga s 750 d.e. za izdelavo kontejnerja in silosa.

Določite takšne optimalne dimenzije kontejnerja in silosa, da bodo skupni stroški transporta in izgradnje najmanjši možni.

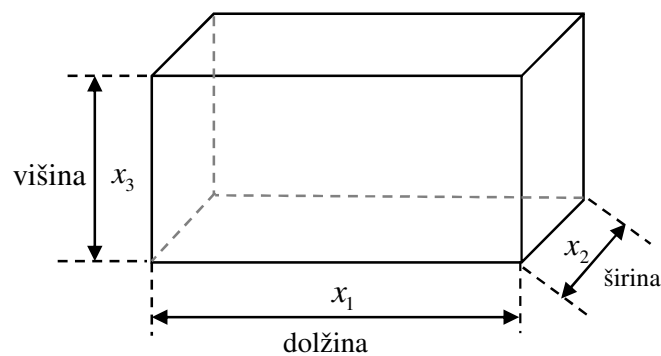
#### Rešitev:

Rešitev problema bomo iskali v 2 fazah, pri čemer bomo:

- 1.) najprej poiskali optimalne dimenzije kontejnerja,
- 2.) nato pa poiskali še optimalne dimenzije silosa.

### I.) Optimizacija kontejnerja

V prvi fazi bomo poiskali optimalne dimenzije kontejnerja, katerega geometrijske veličine so predstavljene na sliki 93. V ta namen pa je potrebno najprej seveda nastaviti matematični model stroškovne ciljne funkcije (glej podoben primer v poglavju 7.1.2., le da tedaj omejitev ni bilo).



Slika 93: Geometrijske veličine kontejnerja

Skupno stroškovno funkcijo lahko definiramo kot:

$$\underbrace{\text{Skupni stroški}}_f = \underbrace{\text{Stroški škatle}}_{f_1} + \underbrace{\text{Stroški transporta}}_{f_2} \quad (7.200)$$

Pri tem so stroški izdelave kontejnerja (škatle) enaki:

$$f_1 = 10 \cdot S_{\text{kontejnerja}} = 10 \cdot [x_1 x_2 + 2x_1 x_3 + 2x_2 x_3] \quad (7.201)$$

kjer je  $S_{\text{kontejnerja}}$  skupna površina kontejnerja, katerega izdelava kvadratnega metra stane 10 d.e. (glej podatke naloge!).

Stroški transporta kontejnerja so enaki:

$$f_2 = \underbrace{\text{Število transportov}}_{\frac{2000}{x_1 x_2 x_3}} \cdot \underbrace{\text{Strošek za vsak transport}}_{10} \quad (7.202)$$

Če upoštevamo izraza (7.201.) in (7.202.) v izrazu (7.200.), dobimo skupno stroškovno (ciljno) funkcijo, ki se glasi:

$$f = 10 \cdot [x_1 x_2 + 2x_1 x_3 + 2x_2 x_3] + 10 \cdot \frac{2000}{x_1 x_2 x_3} \quad (7.203)$$

Zapišimo še vse omejitve, ki se na osnovi podatkov naloge glasijo:

$$\begin{aligned} g_1: & x_1 x_2 x_3 \leq 10 \\ g_2: & \frac{x_2}{x_3} \leq \frac{5}{4} \\ g_3: & \frac{x_1}{x_2} \geq \frac{4}{1} \\ & x_i > 0 \end{aligned} \quad (7.204)$$

Če zapišemo model (7.203.), (7.204.) v obliki, primerni za reševanje GP problemov, dobimo naslednji zapis:

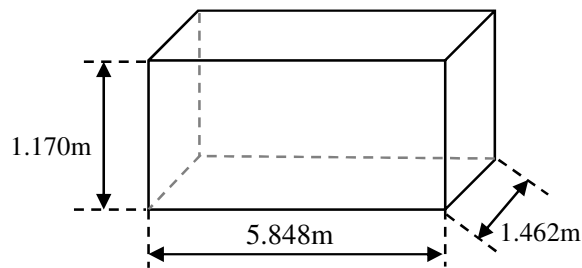
$$\begin{aligned} f(x_1, x_2, x_3) &= 20000x_1^{-1}x_2^{-1}x_3^{-1} + 10x_1^1x_2^1x_3^0 + 20x_1^1x_2^0x_3^1 + 20x_1^0x_2^1x_3^1; \\ g_1(x_1, x_2, x_3) &= \frac{1}{10}x_1^1x_2^1x_3^1 \leq 1, \\ g_2(x_1, x_2, x_3) &= \frac{4}{5}x_1^0x_2^1x_3^{-1} \leq 1, \\ g_3(x_1, x_2, x_3) &= 4x_1^{-1}x_2^1x_3^0 \leq 1, \\ & x_i > 0 \end{aligned} \quad (7.205)$$

Kot se izkaže, gre za dokaj zapleten problem GP z omejitvami. Če problem rešimo (glej [17]), se optimalna rešitev za dimenzije kontejnerja glasi:

$$\underline{x}^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} 5.848m \\ 1.462m \\ 1.170m \end{bmatrix} \quad (7.206)$$

Torej bodo pri teh dimenzijah kontejnerja skupni stroški izdelave in transporta kontejnerja najmanjši možni.

Optimalni kontejner je ilustriran na sliki 94, kjer upoštevamo dimenzije v izrazu (7.206.)



Slika 94: Optimalni kontejner (z optimalnimi dimenzijami)

Torej so optimalni (minimalni) stroški izdelave kontejnerja enaki (vstavimo izraz (7.206.) v izraz (7.201)):

$$f_1^* = 10 \cdot [x_1^* x_2^* + 2x_1^* x_3^* + 2x_2^* x_3^*] = 10 \cdot [5.848 \cdot 1.462 + 2 \cdot 5.848 \cdot 1.170 + 2 \cdot 1.462 \cdot 1.170] = 256.5 \text{ d.e.} \quad (7.207)$$

Za izdelavo silosa nam torej glede na razpoložljiva denarna sredstva (glej podatke naloge) ostane  $750 - 256.5 = 493.5 \text{ d.e.}$

Minimalni stroški transporta kontejnerja so enaki (vstavimo izraz (7.206.) v izraz (7.202.)):

$$f_2^* = \frac{2000}{x_1^* x_2^* x_3^*} \cdot 10 = \frac{2000}{\underbrace{5.848 \cdot 1.462 \cdot 1.170}_{10.003 \rightarrow \text{Volumen kontejnerja pride cca. } 10m^3}} \cdot 10 \approx 2000 \text{ d.e.} \quad (7.208)$$

pri čemer je razvidno tudi, da pride optimalni volumen kontejnerja enak cca. 10 kubičnih metrov. Iz izraza (7.208.) je razvidno, da bodo stroški transporta kontejnerja enaki cca. 2000 d.e, ki ga bomo transportirali  $\frac{2000}{10.003}$  krat, to je cca. 200 krat.

Skupne minimalne stroške transporta in izdelave kontejnerja dobimo, če na osnovi izraza (7.200.) seštejemo izraza (7.207.) in (7.208.):

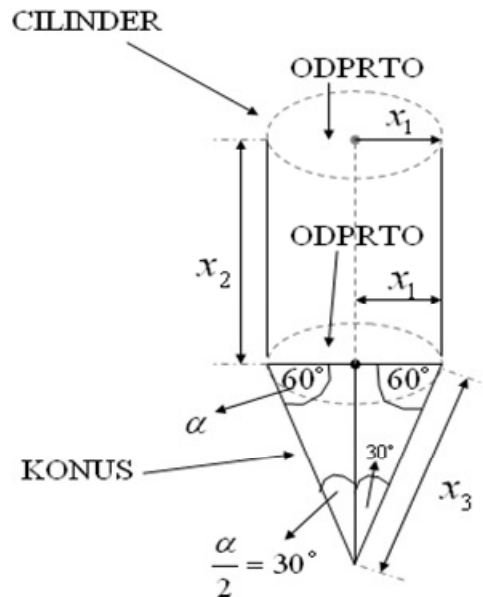
$$f^* = f_1^* + f_2^* = 2256.5 \text{ d.e.} \quad (7.209)$$

Skupni minimalni stroški transporta in izdelave kontejnerja bodo torej enaki 2256.5 d.e.

V nadaljevanju bomo poskušali poiskati še optimalne dimenzije silosa.

## II.) Optimizacija silosa

Silos na sliki 92 je sestavljen iz cilindričnega in konusnega dela. Geometrijske veličine, ki se pri njem pojavljajo, natančneje prikazuje slika 95.



Slika 95: Geometrijske veličine silosa

Volumen cilindričnega dela silosa je enak:

$$V_{cilindra} = \pi \cdot x_1^2 x_2 \quad (7.210)$$

Površini cilindričnega in konusnega dela silosa pa sta enaki:

$$\begin{aligned} S_{cilindra} &= 2\pi x_1 x_2 \\ S_{konusa} &= \underbrace{\pi x_1 x_3}_{\text{za stožec}} \end{aligned} \quad (7.211)$$

V izrazu (7.211) očitno ne upoštevamo površini osnovnih ploskev, saj je na teh mestih silos odprt.

Zaradi zahtev za konusni del ( $\alpha \leq 60^\circ$ ) vzamemo kot izrez stožca enakostranični trikotnik, kjer so mejne vrednosti vseh treh kotov enake  $60^\circ$ .

$$\alpha \leq 60^\circ \quad \Rightarrow \quad \frac{\alpha}{2} \leq 30^\circ \quad \Rightarrow \quad \sin \frac{\alpha}{2} \leq \sin 30^\circ = \frac{1}{2} \quad (7.212)$$

Po drugi strani pa velja:

$$\sin \frac{\alpha}{2} = \frac{x_1}{x_3} = \frac{\text{nasprotna kateta}}{\text{hipotenuza}} \quad (7.213)$$

Na osnovi izrazov (7.212.) in (7.213.) lahko zapišemo naslednjo omejitev:

$$\frac{x_1}{x_3} \leq \frac{1}{2} \quad \Rightarrow \quad 2x_1 \leq x_3 \quad \rightarrow \text{omejitev} \quad (7.214)$$

Stroškovna ciljna funkcija za izdelavo konusa je na osnovi izraza (7.211.) enaka (glej tudi podatke naloge):

$$\begin{aligned} f_S(\underline{x}) &= 15 \cdot S_{cilindra} + 20 \cdot S_{konusa} = 15 \cdot 2\pi \cdot x_1 x_2 + 20\pi \cdot x_1 x_3 = \\ &= 30\pi \cdot x_1 x_2 + 20\pi \cdot x_1 x_3 \end{aligned} \quad (7.215)$$

Na osnovi podatkov naloge moramo upoštevati še eno omejitev, ki se na osnovi izračunanega volumna kontejnerja (sestavni del izraza (7.208.)) glasi:

$$V_{cilindra} \geq V_{kontejnerja} \approx 10m^3 \quad (7.216)$$

Odtod z upoštevanjem izraza (7.210.) dobimo:

$$\pi \cdot x_1^2 x_2 \geq 10m^3 \quad (7.217)$$

Združimo izraze (7.215.), (7.214.) in (7.217.) v enoten matematični model:

$$\begin{aligned} f_S(x) &= 30\pi \cdot x_1 x_2 + 20\pi \cdot x_1 x_3 \\ \text{P.O.:} \\ g_1: \quad &\pi \cdot x_1^2 x_2 \geq 10 \\ g_2: \quad &2x_1 \leq x_3 \\ &x_i > 0 \end{aligned} \quad (7.218)$$

Prevedimo izraze (7.218.) v primerno obliko za reševanje GP problema z omejitvami, pri čemer upoštevamo, da imamo v izrazu (7.163.) opravka z  $P = 2$ , torej dvema omejitvama. Seveda moramo v ta namen dano ciljno funkcijo in njeni omejitvi iz izraza (7.218.) zapisati v obliki delnih funkcij, primerni za reševanje z geometrijskim programiranjem.



Tako dobimo naslednje izraze:

$$\left. \begin{aligned}
 f_S(\underline{x}) &= \overbrace{30\pi \cdot x_1^1 x_2^1 x_3^0}^{U_1} + \overbrace{20\pi \cdot x_1^1 x_2^0 x_3^1}^{U_2} \\
 g_1(\underline{x}) &= \frac{\overbrace{10}^{C_3}}{\pi} x_1^{-2} x_2^{-1} x_3^0 \leq 1 \\
 g_2(\underline{x}) &= \frac{\overbrace{2}^{C_4}}{2} x_1^1 x_2^0 x_3^{-1} \leq 1
 \end{aligned} \right\} \begin{aligned}
 &N = 2 \text{ (2 člena)} \\
 &K_1 = 1 \text{ (1 člen)} \\
 &K_2 = 1 \text{ (1 člen)}
 \end{aligned} \quad (7.219)$$

$x_i > 0$

Matriko  $\underline{A}$  lahko neposredno zapišemo, če koeficiente v delnih funkcijah ciljne funkcije in obeh omejitvah gledamo kot stolpce, pri čemer dobimo naslednji izraz za pogoj ortogonalnosti (7.27.):

$$\underline{A} = \begin{bmatrix} 1 & 1 & -2 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}; \quad \underbrace{\underline{A} \cdot \underline{y}^* = \underline{0}}_{\text{ORTOGONALNOST}} \quad (7.220)$$

Tudi tukaj smo podobno, kot v prejšnjem primeru, v matriki A dodali poleg stolpcev, ki pripadajo ciljni funkciji, tudi stolpca, ki pripadata njenima omejitvama.

Za pogoj normiranosti (7.22.) velja, da ga podobno kot v prejšnjem primeru pišemo le za tisti indeks vsote (N=2), ki se tiče ciljne funkcije (torej omejitvah tukaj ne upoštevamo).

Tako dobimo:

$$\sum_{j=1}^{N=2} y_j^* = 1; \quad \underbrace{y_1^* + y_2^* = 1}_{\text{NORMIRANOST}} \quad (7.221)$$

Izraz (7.220.) lahko zapišemo tudi v obliki:

$$\begin{bmatrix} 1 & 1 & -2 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (7.222)$$

oziroma:

$$\begin{aligned} E_1: & \quad \overbrace{y_1^* + y_2^*}^{=1 \text{ (zaradi normiranosti)}} - 2y_3^* + y_4^* = 0 \\ E_2: & \quad y_1^* - y_3^* = 0 \quad \Rightarrow y_1^* = y_3^* \\ E_3: & \quad y_2^* - y_4^* = 0 \quad \Rightarrow y_2^* = y_4^* \end{aligned} \quad (7.223)$$

Odtod na osnovi pogoja (7.221.) dobimo:

$$\begin{aligned} E_1: & \quad 1 - 2y_3^* + y_4^* = 0 \\ E_2: & \quad y_1^* = y_3^* \\ E_3: & \quad y_2^* = y_4^* \end{aligned} \quad (7.224)$$

Na osnovi pogoja (7.221.) in izraza (7.224.) velja:  $y_1^* + y_2^* = y_3^* + y_4^* = 1$ . Če to upoštevamo v enačbi  $E_1$  izraza (7.224.), dobimo:

$$\begin{aligned} y_3^* + y_4^* - 2y_3^* + y_4^* &= 0 \\ 2y_4^* = y_3^* &\quad \Rightarrow y_4^* = \frac{y_3^*}{2} \end{aligned} \quad (7.225)$$

Vstavimo izraz (7.225) v  $E_1$  in dobimo:

$$\begin{aligned} 1 - 2y_3^* + \frac{y_3^*}{2} &= 0 \\ 1 = 2y_3^* - \frac{y_3^*}{2} &\quad \Rightarrow y_3^* = \frac{2}{3} \end{aligned} \quad (7.226)$$

Če upoštevamo izraz (7.226.) v izrazih (7.224.) in (7.225), dobimo še preostale optimalne dualne spremenljivke:

$$\begin{aligned}
y_4^* &= \frac{y_3^*}{2} = \frac{1}{3} \\
y_1^* &= y_3^* = \frac{2}{3} \\
y_2^* &= y_4^* = \frac{1}{3}
\end{aligned}
\tag{7.227}$$

Na osnovi izrazov (7.226.) in (7.227.) dobimo torej naslednji set optimalnih dualnih odločitvenih spremenljivk:

$$(y_1^*, y_2^*, y_3^*, y_4^*) = \left( \frac{2}{3}, \frac{1}{3}, \frac{2}{3}, \frac{1}{3} \right)
\tag{7.228}$$

Optimalna dualna funkcija (7.169.) zavzame naslednjo obliko:

$$\begin{aligned}
f_S^*(\underline{y}^*) &= \prod_{j=1}^{N+K_1+K_2=4} \left( \frac{C_j}{y_j^*} \right)^{y_j^*} \cdot \prod_{K=1}^{P=2} (\lambda_K^*)^{\lambda_K^*} = \\
&= \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4^*} \right)^{y_4^*} \cdot (\lambda_1^*)^{\lambda_1^*} \cdot (\lambda_2^*)^{\lambda_2^*},
\end{aligned}
\tag{7.229}$$

kjer je  $\tilde{N} = N + K_1 + K_2 = 4$  in za parametra  $\lambda_i$  ( $i = 1, 2$ ) iz izraza (7.170.) velja:

$$\begin{aligned}
\lambda_1^* &= \underbrace{y_{N+1}^*}_{3} + y_{N+2}^* + \dots + \underbrace{y_{N+K_1}^*}_{3} = y_3^* \\
\lambda_2^* &= \underbrace{y_{N+K_1+1}^*}_{4} + \dots + \underbrace{y_{N+K_1+K_2}^*}_{4} = y_4^*
\end{aligned}
\tag{7.230}$$

Če upoštevamo koeficiente  $C_i$  iz izrazov (7.219.) ter optimalno rešitev za dualne spremenljivke (7.228.) v izrazih (7.229.) in (7.230.), dobimo naslednjo optimalno dualno ciljno funkcijo:

$$\begin{aligned}
f_s^* &= \left(\frac{30\pi}{\frac{2}{3}}\right)^{\frac{2}{3}} \cdot \left(\frac{20\pi}{\frac{1}{3}}\right)^{\frac{1}{3}} \cdot \left(\frac{10}{\frac{\pi}{2}}\right)^{\frac{2}{3}} \cdot \left(\frac{2}{\frac{1}{3}}\right)^{\frac{1}{3}} \cdot \left(\frac{2}{3}\right)^{\frac{2}{3}} \cdot \left(\frac{1}{3}\right)^{\frac{1}{3}} = \\
&= (45\pi)^{\frac{2}{3}} \cdot (60\pi)^{\frac{1}{3}} \cdot \left(\frac{15}{\pi}\right)^{\frac{2}{3}} \cdot (6)^{\frac{1}{3}} \cdot \left(\frac{2}{3}\right)^{\frac{2}{3}} \cdot \left(\frac{1}{3}\right)^{\frac{1}{3}} = \\
&= \left[\frac{45\pi \cdot 15 \cdot 2}{\pi \cdot 3}\right]^{\frac{2}{3}} \cdot \left[\frac{60\pi \cdot 6}{3}\right]^{\frac{1}{3}} = 450^{\frac{2}{3}} \cdot (120\pi)^{\frac{1}{3}} = \\
&= 58.7230 \cdot 7.2239 = 424.2142
\end{aligned} \tag{7.231}$$

Pri določitvi optimalnih primarnih odločitvenih spremenljivk upoštevamo izraze (7.171.), (7.219.), (7.228.) in (7.231.), pri čemer dobimo:

$$\begin{aligned}
U_1^* = f^* \cdot y_1^* &\Rightarrow 30\pi x_1^* x_2^* = 424.2142 \cdot \frac{2}{3} \\
U_2^* = f^* \cdot y_2^* &\Rightarrow 20\pi x_1^* x_3^* = 424.2142 \cdot \frac{1}{3}
\end{aligned} \tag{7.232}$$

Odtod sledi:

$$x_1^* x_2^* = 3.000 \tag{7.233}$$

$$x_1^* x_3^* = 2.2505 \tag{7.234}$$

Pojavijo se še dodatne relacije zaradi omejitev, pri čemer izrazi za omejitvev (7.172.) z upoštevanjem izrazov (7.219.), (7.228.) in (7.230.) preidejo v naslednjo obliko za 1. omejitev:

$$U_{\underbrace{N+K_1}_3}^*(x) = \frac{\underbrace{y_{N+K_1}}_3^*}{\lambda_1^*} = \frac{y_3^*}{\lambda_1^*} = \frac{y_3^*}{y_3^*} = 1 \tag{7.235}$$

kar da po izenačitvi z delno funkcijo v (7.219.):

$$\frac{10}{\pi} \cdot \frac{1}{x_1^{*2} x_2^*} = 1 \Rightarrow x_1^{*2} x_2^* = 3.1830 \tag{7.236}$$

Izrazi za omejitvev (7.173.) pa z upoštevanjem izrazov (7.219.), (7.228.) in (7.230.) preidejo v naslednjo obliko za 2. omejitev:

$$U_{\frac{N+K_1+K_2}{4}}^*(\underline{x}) = \frac{y_{\frac{N+K_1+K_2}{4}}^*}{\lambda_2^*} = \frac{y_4^*}{\lambda_2^*} = \frac{y_4^*}{y_4^*} = 1 \quad (7.237)$$

kar da po izenačitvi z delno funkcijo v (7.219.):

$$2 \cdot \frac{x_1^*}{x_3^*} = 1 \quad \Rightarrow \quad x_3^* = 2x_1^* \quad (7.238)$$

Na osnovi izrazov (7.234) in (7.238) sledi:

$$x_1^* \cdot 2x_1^* = 2.505 \quad \Rightarrow \quad x_1^* = \sqrt{\frac{2.2505}{2}} = 1.0607m \quad (7.239)$$

Če izraz (7.239.) vstavimo v izraz (7.233.), dobimo:

$$1.0607 \cdot x_2^* = 3.000 \quad \Rightarrow \quad x_2^* = 2.8281m \quad (7.240)$$

Če vstavimo izraz (7.239) v izraz (7.238.), pa dobimo:

$$x_3^* = 2 \cdot 1.0607 = 2.1214m \quad (7.241)$$

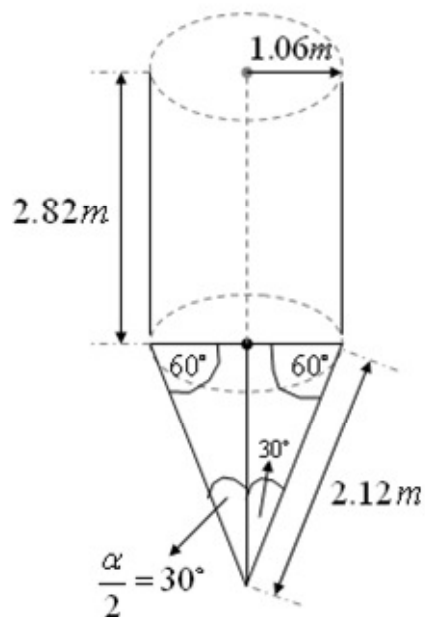
Optimalni vektor primarnih odločitvenih spremenljivk za silos se torej na osnovi izrazov (7.239.) do (7.241.) glasi:

$$\underline{x}^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} 1.0607m \\ 2.8281m \\ 2.1214m \end{bmatrix} \quad (7.242)$$

kar so optimalne dimenzije silosa. Poglejmo še, kakšen pride optimalni konusni kot silosa, ki ga dobimo na osnovi izrazov (7.213.) in (7.242.):

$$\sin \frac{\alpha^*}{2} = \frac{1.0607}{2.1214} = \frac{1}{2} \quad \Rightarrow \quad \frac{\alpha^*}{2} = 30^\circ \quad \Rightarrow \quad \alpha^* = 60^\circ \quad (7.243)$$

Optimalni silos je ilustriran na sliki 96, kjer upoštevamo dimenzije v izrazu (7.242.) in kot v izrazu (7.243.).



Slika 96: Optimalni silos (z optimalnimi dimenzijami in optimalnim konusnim kotom)

Preverimo še, če bi pri optimalno izdelanih dimenzijah kontejnerja (glej izraz (7.206.)) in silosa (glej izraza (7.242.), (7.243.)) ostali v okviru predvidenih maksimalnih dovoljenih stroškov za izdelavo kontejnerja in silosa skupaj (750 d.e., glej podatke naloge).

Tako na osnovi izrazov (7.207.) in (7.231.) sledi za optimalne (minimalne) stroške izdelave kontejnerja oz. silosa:

$$\begin{aligned}
 f_1^* &= 256.5 \text{ d.e.} \} \text{ Za kontejner} \\
 f_s^* &= 424.2142 \text{ d.e.} \} \text{ Za silos} \\
 f_s^* + f_1^* &= 680.71 \text{ d.e.} \} \text{ Stroški izdelave kontejnerja in silosa}
 \end{aligned}
 \tag{7.244}$$

Ker je  $680.71 \text{ d.e.} < 750 \text{ d.e.}$ , smo očitno uspeli ostati v okviru predvidenih maksimalnih dovoljenih stroškov za izdelavo kontejnerja in silosa skupaj.

Preverimo še, kolikšni bodo vsi stroški (v optimumu), ki jih bomo imeli z izdelavo silosa in izdelavo ter transporta kontejnerja. V ta namen seštejemo izraza (7.209.) in (7.231.), pri čemer dobimo:

$$f_{skupni}^* = f^* + f_S^* = 2256.5 + 424.2142 = 2680.71 \text{ d.e.} \quad (7.245)$$

Skupni minimalni stroški (v optimumu!) izdelave silosa in izdelave ter transporta kontejnerja torej so 2680.71 d.e. To so torej najmanjši možni skupni stroški, ki jih bomo potrebovali za optimalno rešitev danega problema.

S tem primerom smo končali teoretično obravnavo geometrijskega programiranja z omejitvami in geometrijskega programiranja nasploh. Prav tako smo na tej točki končali s teoretičnimi osnovami, katerih razumevanje je zelo dobrodošlo pri reševanju raznoraznih problemov s področja nelinearne optimizacije logističnih procesov.

V naslednjem poglavju pa se bomo soočili še z dokaj obširno zbirko rešenih nalog. Namen tega poglavja je, preko reševanja različnih primerov utrditi razumevanje problematike reševanja nelinearnih logističnih problemov s pomočjo optimizacije.

Zbirka je zasnovana tako, da po strukturi bolj ali manj sledi strukturi teoretičnih osnov v prejšnjih poglavjih. Kjer je potrebno, pa so na začetku posameznih primerov ali začetku celotnega paketa primerov navedene tudi usmeritve, kako čim lažje poiskati povezavo med posameznim primerom (paketom primerov) in pa teorijo prejšnjih poglavij, ki se tiče dotičnega primera (paketa primerov).

## 8 ZBIRKA REŠENIH NALOG

Zbirka obsega naslednje primere:

- primere analitične optimizacije,
- primere enodimenzionalne numerične optimizacije,
- primere multidimenzionalne numerične optimizacije z osnovnimi metodami,
- primere multidimenzionalne numerične optimizacije s sofisticiranimi metodami, ter
- primere geometrijskega programiranja.

Teoretične osnove, ki se tičejo posameznih primerov, so bile razložene v prejšnjih poglavjih tega dela.

V nadaljevanju si najprej pogledajmo primere optimizacije, ki jih lahko rešimo z analitičnim izračunom.

### 8.1 Primeri analitične optimizacije

Primere, ki jih bomo reševali, lahko razdelimo na 2 kategoriji:

- Primeri, ki jih lahko rešujemo s klasičnim analitičnim računanjem. Večinoma gre za primere z omejitvami, ki jih lahko z eliminacijo spremenljivk prevedemo na enodimenzionalne ali dvodimenzionalne analitične probleme brez omejitev, kjer pri iskanju ekstrema enačimo odvod funkcije z 0 (ali gradient funkcije z ničelnim vektorjem).
- Primeri z omejitvami, ki jih rešujemo z analitičnim pristopom z Lagrangeovimi multiplikatorji.

Opomnimo še, da bo v nekaterih primerih potrebno določiti tudi naravo oz. tip ekstremov, to je ustrezno klasificirati stacionarne točke.

V nadaljevanju si bomo najprej pogledali prvo kategorijo primerov.



### 8.1.1 Klasičen način reševanja

V okviru tega poglavja si najprej pogledajmo primere, ki jih rešujemo s klasičnim načinom reševanja, kjer pri iskanju ekstrema enačimo odvod funkcije z 0 oziroma gradient funkcije z ničelnim vektorjem.

#### Primer 1:

Delite število 120 na dva dela tako, da je produkt  $P$  enega dela in kvadrata drugega dela maximum.

$$\left. \begin{array}{l} \text{del 1} = x \\ \text{del 2} = 120 - x \end{array} \right\} + = 120 \quad (8.1)$$

Problem je sledeč:

$$\begin{aligned} P &= x^2 \cdot (120 - x) = P(x) \\ P &= 120x^2 - x^3, \quad 0 \leq x \leq 120 \end{aligned} \quad (8.2)$$

Poiščemo ekstrem in dobimo stacionarne točke:

$$\begin{aligned} \frac{dP}{dx} &= 240x - 3x^2 = 0 \\ 3x^2 - 240x &= 0 \\ x^2 - 80x &= 0 \\ x \cdot (x - 80) &= 0 \\ x_1^* &= 0, \quad x_2^* = 80 \end{aligned} \quad (8.3)$$

Optimum je sledeč:

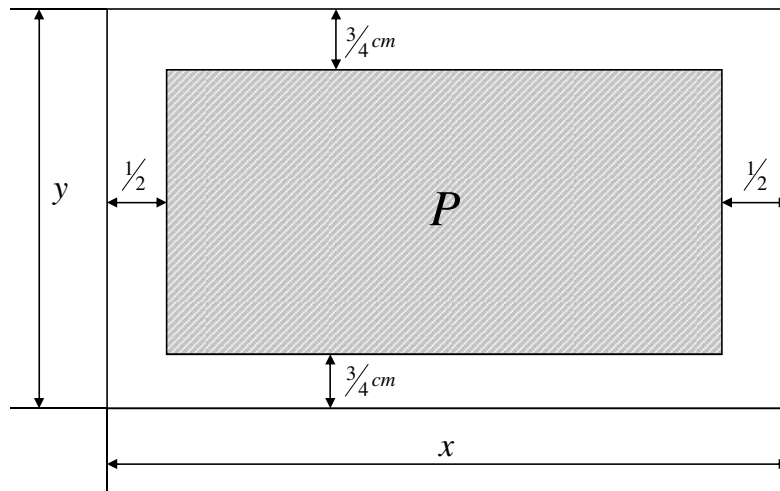
$$P(x_1^*) = 0; \quad P(x_2^*) = 120 \cdot 80^2 - 80^3 = 256000 \rightarrow \text{MAX} \quad (8.4)$$

Sledi:

$$120 = 80 + 40 \quad (8.5)$$

### Primer 2:

List papirja za poster naj obsega  $18\text{cm}^2$  (glej sliko 97). Na sliki so označeni robni pogoji za tiskano površino. Kolikšno mora biti razmerje strani lista, da bo nanj natiskana največja možna površina?



Slika 97: List papirja

Problem optimizacije celotnega lista papirja je naslednji:

$$P_{CEL} = x \cdot y \rightarrow \text{Prevedimo le na eno spremenljivko} \quad (8.6)$$

Problem optimizacije tiskane površine pa je naslednji:

$$P(x, y) = \left(x - \frac{1}{2} - \frac{1}{2}\right) \cdot \left(y - \frac{3}{4} - \frac{3}{4}\right) = (x - 1) \cdot \left(y - \frac{3}{2}\right) \quad (8.7)$$

Eliminiramo eno spremenljivko na podlagi izraza (8.6) in dobimo:

$$x \cdot y = 18 \Rightarrow y = \frac{18}{x} \quad (8.8)$$

Naredimo odvod funkcije (8.7) in ga enačimo z 0:

$$\begin{aligned} P(x) &= (x - 1) \cdot \left(\frac{18}{x} - \frac{3}{2}\right) \Rightarrow \frac{dP}{dx} = 1 \cdot \left(\frac{18}{x} - \frac{3}{2}\right) + (x - 1) \cdot \left(-\frac{18}{x^2}\right) \\ \frac{dP}{dx} &= \left(\frac{18}{x} - \frac{3}{2}\right) - (x - 1) \cdot \frac{18}{x^2} = \left(\frac{18}{x} - \frac{3}{2}\right) - \frac{18}{x} + \frac{18}{x^2} = \frac{18}{x^2} - \frac{3}{2} = 0 \end{aligned} \quad (8.9)$$

In tako dobimo stacionarne točke:

$$18 - \frac{3}{2}x^2 = 0 \quad \Rightarrow \quad x^2 = \frac{36}{3} = 12 \quad \Rightarrow \quad x^* = \sqrt{12} = \sqrt{4 \cdot 3} = 2\sqrt{3}$$

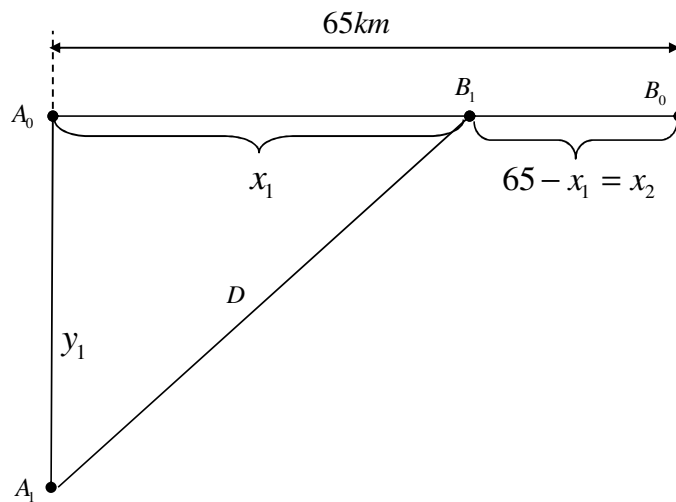
$$y^* = \frac{18}{x^*} = \frac{18}{2\sqrt{3}} = \frac{9\sqrt{3}}{\sqrt{3} \cdot \sqrt{3}} = \frac{9\sqrt{3}}{3} = 3\sqrt{3} \quad (8.10)$$

List moramo torej deliti v razmerju  $2\sqrt{3}$  in  $3\sqrt{3}$ , da dobimo maksimalno tiskano površino.

### Primer 3:

Ob 9. uri je ladja  $B$   $65\text{km}$  vzhodno od ladje  $A$ . Ladja  $B$  potem potuje zahodno s hitrostjo  $10 \frac{\text{km}}{\text{h}}$ ,  $A$  pa proti jugu s  $15 \frac{\text{km}}{\text{h}}$ . Če nadaljujeta enakomerno v predpisani smeri, kdaj

bosta najbližje druga drugi, in kako blizu?



Slika 98: Ladja

Problem je podan z izrazoma (8.11) in (8.12):

$$v_1 = \frac{dy}{dt} = \frac{y_1}{t} \quad \Rightarrow \quad y_1 = v_1 \cdot t = 15t \quad (8.11)$$

$$v_2 = \frac{dx}{dt} = \frac{x_2}{t} \quad \Rightarrow \quad x_2 = v_2 \cdot t = 10t$$

$$x_1 = 65 - x_2 = 65 - 10t$$

$$y_1 = 15t \quad (8.12)$$

Razdalja med dvema točkama je sledeča:

$$D(x_1, y_1) = \sqrt{x_1^2 + y_1^2} \Rightarrow D^2(x_1, y_1) = x_1^2 + y_1^2 = (65 - 10t)^2 + (15t)^2 = f(t) \quad (8.13)$$

Kdaj je  $D = D_{MIN}$ ? Ko je  $\frac{df}{dt} = 0$ . Naredimo torej odvod po času in dobimo:

$$\begin{aligned} \frac{df}{dt} &= 2 \cdot (65 - 10t) \cdot (-10) + 15^2 \cdot 2t = 0 \Rightarrow \\ &\Rightarrow 15^2 \cdot 2t = 20 \cdot (65 - 10t) \Rightarrow \\ &\Rightarrow 15 \cdot 15 \cdot t = 5 \cdot 2 \cdot (65 - 10t) \Rightarrow \\ &\Rightarrow 3 \cdot 15 \cdot t = 2 \cdot (65 - 10t) \Rightarrow \\ &\Rightarrow 45t = 130 - 20t \Rightarrow \\ &\Rightarrow 65t = 130 \Rightarrow \quad t^* = \frac{130}{65} = 2h \end{aligned} \quad (8.14)$$

Razdalja med ladjama pa je po dveh urah sledeča:

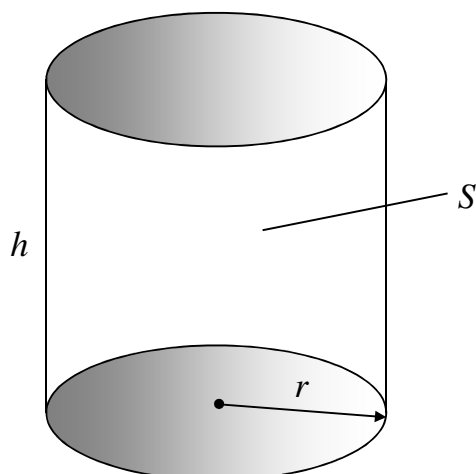
$$D(t^*) = \sqrt{f(t^*)} = \sqrt{(65 - 10 \cdot 2)^2 + (15 \cdot 2)^2} = \sqrt{45^2 + 30^2} = 54km \quad (8.15)$$

Torej ob 11. uri sta ladji najbližje, to je 54km narazen.

#### Primer 4:

Cilindrični kontejner mora držati  $64cm^3$ . Poiščite njegove dimenzije na takšen način, da bo uporabljena vsebina (površina) materiala v njem, najmanjša možna – torej, da bomo uporabili najmanj materiala oz., da bodo stroški najmanjši. Pri tem sta možni dve varianti:

- a.) cilinder je zgoraj odprt in
- b.) cilinder je zgoraj zaprt.



Slika 99: Cilinder

Rešitev za a:

$$V = \pi \cdot r^2 \cdot h = 64 \quad \Rightarrow h = \frac{64}{\pi \cdot r^2} \quad (8.16)$$

Površina materiala, ki ga je potrebno optimizirati je sledeča:

$$S = \pi \cdot r^2 + 2\pi \cdot r \cdot h = S(r, h) \quad (8.17)$$

Prevedimo na eno spremenljivko:

$$S = \pi \cdot r^2 + 2\pi \cdot r \cdot \frac{64}{\pi \cdot r^2} = \pi \cdot r^2 + \frac{128}{r} = S(r) \quad (8.18)$$

Iščemo tak  $r$ , da bo  $S(r) = S_{MIN}$ . Sledi:

$$\frac{dS}{dr} = 2\pi \cdot r - \frac{128}{r^2} = 0 \quad \Rightarrow 2\pi \cdot r^3 - 128 = 0 \quad (8.19)$$

Sledi:

$$r^3 = \frac{128}{2\pi} = \frac{64}{\pi} \quad \Rightarrow r^* = \sqrt[3]{\frac{64}{\pi}} = \frac{4}{\sqrt[3]{\pi}} \quad (8.20)$$

$$h^* = \frac{64}{\pi \cdot r^{*2}} = \frac{64}{\pi \cdot \left(\frac{4}{\sqrt[3]{\pi}}\right)^2} = \frac{64}{\pi \cdot \frac{16}{\pi^{\frac{2}{3}}}} = \frac{4}{\pi \cdot \frac{1}{\pi^{\frac{2}{3}}}} = 4 \cdot \frac{\pi^{\frac{2}{3}}}{\pi^{\frac{2}{3}}} = \frac{4}{\sqrt[3]{\pi}} \quad (8.21)$$

Sklep:

$$h^* = r^* \quad (8.22)$$

Rešitev za b:

$$V = \pi \cdot r^2 \cdot h = 64 \quad \Rightarrow h = \frac{64}{\pi \cdot r^2} \quad (8.23)$$

Površina materiala, ki ga je potrebno optimizirati je sledeča:

$$S = \pi \cdot r^2 + \pi \cdot r^2 + 2\pi \cdot r \cdot h = 2\pi \cdot r^2 + 2\pi \cdot r \cdot h = 2\pi \cdot r \cdot (r + h) = S(r, h) \quad (8.24)$$

$$S = 2\pi \cdot r \cdot \left( r + \frac{64}{\pi \cdot r^2} \right) = S(r) = 2\pi \cdot r^2 + \frac{2\pi \cdot r \cdot 64}{\pi \cdot r^2} = 2\pi \cdot r^2 + \frac{128}{r} \quad (8.25)$$

Naredimo odvod, da dobimo optimalne dimenzije:

$$\frac{dS}{dr} = 4\pi \cdot r - \frac{128}{r^2} = 0 \quad \Rightarrow 4\pi \cdot r^3 - 128 = 0 \quad \Rightarrow r^3 = \frac{128}{4\pi} = \frac{32}{\pi} \quad \Rightarrow$$

$$r^* = \sqrt[3]{\frac{32}{\pi}} = \sqrt[3]{\frac{8 \cdot 4}{\pi}} = 2 \cdot \sqrt[3]{\frac{4}{\pi}};$$

$$h^* = \frac{64}{\pi \cdot r^{*2}} = \frac{64}{\pi \cdot \left( 2 \cdot \sqrt[3]{\frac{4}{\pi}} \right)^2} = \frac{64}{\pi \cdot \left( 4 \cdot \left( \frac{4}{\pi} \right)^{\frac{2}{3}} \right)} = \frac{16}{\pi \cdot \left( \frac{4}{\pi} \right)^{\frac{2}{3}}} = \frac{4 \cdot 4}{\pi^{\frac{3}{3}} \cdot \frac{4^{\frac{2}{3}}}{\pi^{\frac{2}{3}}}} = \frac{4 \cdot 4^{\frac{3-2}{3}}}{\pi^{\frac{1}{3}}} =$$

$$= 4 \cdot \sqrt[3]{\frac{4}{\pi}} = 2r^*$$

Sklep:

$$h^* = 2r^* \quad (8.27)$$

### Primer 5:

Celotna stroškovna funkcija za  $x$  izdelanih radio setov se glasi:  $\frac{x^2}{4} + 35x + 25$ . Prodajna

funkcija dohodka, po kateri lahko prodajo en set na  $x$  izdelanih radiiov, pa se glasi:

$\left( 50 - \frac{1}{2}x \right)$ . Kakšen naj bo dnevni proizvodni plan, da bo dosežen maksimalen profit?

Problem je naslednji:

$$\begin{aligned}
C(x) &= \frac{x^2}{4} + 35x + 25 && \text{(STROŠKI)} \\
D(x) &= \left[ 50 - \frac{1}{2}x \right] \cdot x && \text{(DOHODKI)} \rightarrow \text{moramo še množiti z } x \text{ (z št. setov)} \\
P(x) &= D(x) - C(x) && \text{(PROFIT)} \\
P(x) &= \left[ 50 - \frac{1}{2}x \right] \cdot x - \left( \frac{x^2}{4} + 35x + 25 \right) = 50x - \frac{x^2}{2} - \frac{x^2}{4} - 35x - 25 = \\
&= 15x - \frac{3x^2}{4} - 25
\end{aligned} \tag{8.28}$$

Naredimo odvod po  $x$  izdelanih radio setov, da dosežemo maksimalen profit:

$$\begin{aligned}
\frac{dP}{dx} &= 15 - \frac{6x}{4} = 0 && \Rightarrow x^* = \frac{60}{6} = 10 \\
\frac{d^2P}{dx^2} &= -\frac{6}{4} < 0 && \Rightarrow \text{v } x^*, P(x^*) \text{ doseže maksimum}
\end{aligned} \tag{8.29}$$

Torej, če bomo izdelali 10 radio setov, bomo dosegli maksimalen profit.

### Primer 6:

Poiščite ekstrem funkcije dane v izrazu (8.30) in poiščite vse možne stacionarne točke.

Problem je podan z naslednjim izrazom:

$$\begin{aligned}
R(w) &= \frac{w^2 + 1}{w^2 - w - 6} \\
\frac{dR}{dw} &= \frac{2w \cdot (w^2 - w - 6) - (w^2 + 1) \cdot (2w - 1)}{(w^2 - w - 6)^2} = \\
&= \frac{2w^3 - 2w^2 - 12w - \{2w^3 - w^2 + 2w - 1\}}{(w^2 - w - 6)^2} = \\
&= \frac{-w^2 - 14w + 1}{(w^2 - w - 6)^2} = \frac{w^2 + 14w - 1}{(w^2 - w - 6)^2}
\end{aligned} \tag{8.30}$$

$$w^2 - w - 6 = 0 \quad \Rightarrow (w - 3) \cdot (w + 2) = 0 \quad \Rightarrow R(w) \tag{8.31}$$

Torej  $R(w)$  ne obstaja v  $w = 3, -2$ , prav tako ne  $R'(w)$ .

Dobljen odvod enačimo z 0 in dobimo:

$$\frac{dR}{dw} = 0 \Rightarrow w^2 + 14w - 1 = 0 \quad (8.32)$$

Sledi:

$$w_{1,2} = \frac{-14 \pm \sqrt{14^2 + 4}}{2} = -7 \pm \frac{1}{2} \sqrt{200} = -7 \pm \frac{1}{2} \cdot 10 \cdot \sqrt{2} = -7 \pm 5 \cdot \sqrt{2} \quad (8.33)$$

Stacionarni točki sta sledeči:

$$w_1 = -7 + 5 \cdot \sqrt{2}, \quad w_2 = -7 - 5 \cdot \sqrt{2} \quad (8.34)$$

### Primer 7:

Poiščite ekstrem funkcije dane v izrazu (8.35) in poiščite vse možne stacionarne točke.

Problem je podan z naslednjim izrazom:

$$h(t) = 10 \cdot t \cdot e^{3-t^2} \quad (8.35)$$

Naredimo odvod in ga enačimo z 0:

$$\frac{dh}{dt} = 10 \cdot (e^{3-t^2} + t \cdot e^{3-t^2} \cdot (-2t)) = 0 \quad / \quad \div e^{3-t^2} \cdot 10 \quad (8.36)$$

Sledi:

$$1 - 2t^2 = 0 \Rightarrow t_{1,2} = \pm \sqrt{\frac{1}{2}} \quad (8.37)$$

Stacionarni točki sta sledeči:

$$t_1 = \sqrt{\frac{1}{2}}, \quad t_2 = -\sqrt{\frac{1}{2}} \quad (8.38)$$

### Primer 8:

Poiščite ekstrem funkcije dane v izrazu (8.39) in poiščite vse možne stacionarne točke.

Problem je podan z naslednjim izrazom:



$$f(x) = x^2 \cdot \ln 3x + 6 \quad (8.39)$$

Naredimo odvod in ga enačimo z 0:

$$\frac{df}{dx} = 2x \cdot \ln 3x + x^2 \cdot \frac{1}{3x} \cdot 3 = x \cdot (2 \cdot \ln 3x + 1) = 0 \quad / \div x \quad (8.40)$$

$x = 0$  ni stacionarna točka, saj  $f(x)$  in  $f'(x)$  sploh ne obstajata v njej.

Sledi:

$$\begin{aligned} 2 \cdot \ln 3x + 1 &= 0 \quad / e \\ e^{2 \cdot \ln 3x + 1} &= e^0 \\ e^{2 \cdot \ln 3x} \cdot e^1 &= 1 \\ e^{2 \cdot \ln 3x} &= e^{-1} \\ e^{(\ln 3x)^2} &= e^{-1} \\ (3x)^2 &= e^{-1} \quad / \sqrt{\phantom{x}} \\ 3x &= e^{-\frac{1}{2}} \end{aligned} \quad (8.41)$$

Stacionarna točka je sledeča:

$$x = \frac{e^{-\frac{1}{2}}}{3} = \frac{1}{3\sqrt{e}} \quad (8.42)$$

### Primer 9:

Denimo, da se znesek na bančnem računu giblje po naslednji zakonitosti:

$$A(t) = 2000 - t \cdot e^{5 - \frac{t^2}{8}} \cdot 10. \text{ Ugotovite največji in najmanjši znesek v prvih 10 letih.}$$

Problem je podan z naslednjim izrazom:

$$A(t) = 2000 - 10t \cdot e^{5 - \frac{t^2}{8}} \quad (8.43)$$

Naredimo odvod po času in ga enačimo z 0:

$$\frac{dA}{dt} = -10 \cdot e^{5 - \frac{t^2}{8}} - 10t \cdot e^{5 - \frac{t^2}{8}} \cdot \left(-\frac{2t}{8}\right) = -10 \cdot e^{5 - \frac{t^2}{8}} \cdot \left(1 - \frac{t^2}{4}\right) = 0 \quad (8.44)$$

Sledi:

$$1 - \frac{t^2}{4} = 0 \quad \Rightarrow t_{1,2} = \pm 2 \quad (8.45)$$

$t^* = 2$ , saj gledamo  $0 \leq t \leq 10$ .

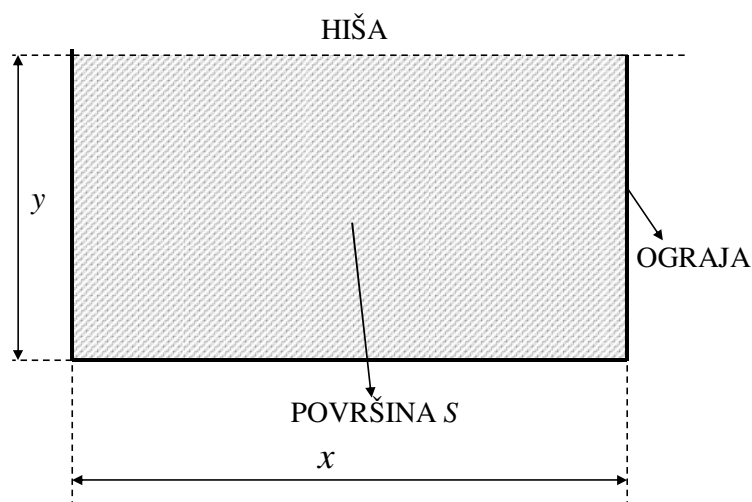
Poglejmo  $A(t)$  v mejah in v stacionarni točki:

$$A(0) = 2000 - 10 \cdot 0 \cdot e^{\frac{5 \cdot 0^2}{8}} = 2000 \text{ €}$$
$$A(2) = 2000 - 10 \cdot 2 \cdot e^{\frac{5 \cdot 2^2}{8}} = 2000 - 20 \cdot e^{\frac{9}{2}} = 20 \cdot \left(100 - e^{\frac{9}{2}}\right) = 199.65 \text{ €} \quad (8.46)$$
$$A(10) = 2000 - 10 \cdot 10 \cdot e^{\frac{5 \cdot 10^2}{8}} = 2000 - 100 \cdot e^{\frac{5 \cdot 100}{8}} = 1999.94 \text{ €}$$

Torej, maksimalni znesek bo 2000€ ob  $t = 0$ , minimalni znesek pa 199.65€ ob  $t = 2$  leti.

### Primer 10:

Z ograjo je treba zagraditi veliko hišo (stavbo), pri čemer imamo na voljo 500m materiala. Ugotovite dimenzije zagrajenega polja, ki bodo obsegale največjo možno površino.



Slika 100: Ograja

Optimalni problem z omejitvijo je naslednji:

$$\begin{aligned} \max S(x, y) &= x \cdot y \\ x + y + y &= 500 \end{aligned} \quad (8.47)$$

Rešimo z eliminacijo ene spremenljivke:

$$x = 500 - 2y \quad (8.48)$$

Površina, ki jo je potrebno optimizirati je naslednja:

$$S(y) = (500 - 2y) \cdot y = 500y - 2y^2 \quad (8.49)$$

Naredimo odvod, ga enačimo z 0 in dobimo optimalne dolžine stranic:

$$\frac{dS}{dy} = 500 - 4y = 0 \quad \Rightarrow \quad y^* = \frac{500}{4} = 125m \quad (8.50)$$

$$x^* = 500 - 2y^* = 500 - 2 \cdot 125m = 250m \quad (8.51)$$

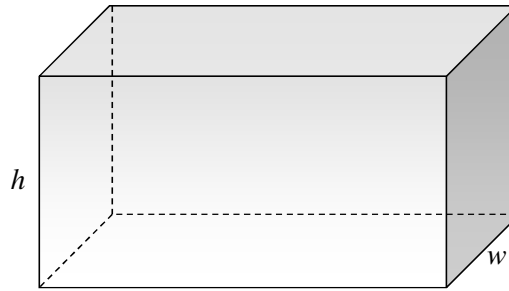
Največja možna površina je sledeča:

$$S^* = S(x^*, y^*) = x^* \cdot y^* = 250 \cdot 125 = 31250m^2 \quad (8.52)$$

Torej moramo ograjo zasnovati tako, da je ena stranica dolga  $125m$ , druga pa  $250m$ , da dobimo največjo možno površino  $31250m^2$ .

### Primer 11:

Zasnovati želimo škatlo, pri kateri je dolžina  $3 \times$  večja od širine. Material, potreben za vrh in dno, stane  $10 \frac{\text{€}}{\text{cm}^2}$ , material za stranice pa stane  $6 \frac{\text{€}}{\text{cm}^2}$ . Določite stranice škatle, pri katerih bodo stroški minimalni, pri čemer mora volumen biti  $50\text{cm}^3$ .



$l = 3w$   
Slika 101: Škatla

Površina stranic, vrha in dna je dana s sledečima relacijama:

$$\begin{aligned} S_{VRH,DNO} &= 3w \cdot w + 3w \cdot w = 6w^2 \\ S_{STRANICE} &= 2h \cdot l + 2h \cdot w = 2h \cdot (l + w) = 2h(3w + w) = 2h \cdot 4w = 8h \cdot w \end{aligned} \quad (8.53)$$

Funkcija stroškov je dana z izrazom (8.54):

$$C = 10 \cdot S_{VRH,DNO} + 6 \cdot S_{STRANICE} = 10 \cdot 6w^2 + 6 \cdot 8h \cdot w = 60w^2 + 48h \cdot w = C(h, w) \quad (8.54)$$

Volumen škatle je sledeč:

$$V = 50 \text{ cm}^3 = h \cdot l \cdot w = h \cdot 3w \cdot w = 50 \quad \Rightarrow h = \frac{50}{3w^2} \quad (8.55)$$

Funkcija, ki jo je potrebno optimizirati, da dosežemo minimalne stroške, pa je sledeča:

$$C(w) = 60w^2 + 48 \cdot \frac{50}{3w^2} \cdot w = 60w^2 + 16 \cdot \frac{50}{w} = 60w^2 + \frac{800}{w} \quad (8.56)$$

Naredimo odvod, ga enačimo z 0 in dobimo:

$$\begin{aligned} \frac{dC}{dw} &= 120w - \frac{800}{w^2} = 0 \quad / \cdot w^2 \\ 120w^3 - 800 &= 0 \quad \Rightarrow w^3 = \frac{800}{120} = \frac{20}{3} \quad \Rightarrow w^* = \sqrt[3]{\frac{20}{3}} = 1.88 \end{aligned} \quad (8.57)$$

Optimalne stranice torej so:

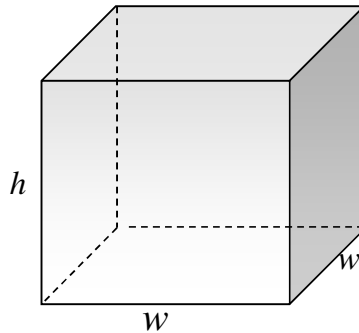
$$\begin{aligned} w^* &= 1.88 \text{ cm} \\ l^* &= 3w^* = 5.64 \text{ cm} \\ h^* &= \frac{50}{3w^{*2}} = \frac{50}{3 \cdot (1.88)^2} = 4.71 \text{ cm} \end{aligned} \quad (8.58)$$

Minimalni stroški pa so sledeči:

$$C(w^*) = C^* = 60w^{*2} + \frac{800}{w^*} = 60 \cdot 1.88^2 + \frac{800}{1.88} = 637.5 \text{ €} \quad (8.59)$$

**Primer 12:**

Zasnovati želimo škatlo, ki ima kvadratno dno, pri čemer imamo na razpolago  $10m^2$  materiala. Kolikšen maksimalen volumen škatle lahko naredimo?



Slika 102: Škatla

Volumen škatle, ki ga je potrebno maksimizirati je naslednji:

$$\begin{aligned} V &= h \cdot w \cdot w = h \cdot w^2 \\ 10 &= 2 \cdot w \cdot w + 4 \cdot h \cdot w \end{aligned} \quad (8.60)$$

Optimalni problem je:

$$\begin{aligned} \max V(h, w) &= h \cdot w^2 \\ 10 &= 2w \cdot (w + 2h) \end{aligned} \quad (8.61)$$

Izrazimo  $h$ :

$$h = \left( \frac{10}{2w} - w \right) \cdot \frac{1}{2} = \frac{10 - 2w^2}{4w} = \frac{5 - w^2}{2w} \quad (8.62)$$

$h$  vstavimo v enačbo (8.61) in dobimo:

$$V(w) = \frac{5 - w^2}{2w} \cdot w^2 = \frac{5 - w^2}{2} \cdot w \quad (8.63)$$

Dobljen volumen odvajamo po stranici  $w$  in ga nato enačimo z 0:

$$\frac{dV}{dw} = \frac{1}{2} \cdot [5 - 3w^2] = 0 \quad \Rightarrow w^* = \sqrt{\frac{5}{3}}$$

$$h^* = \frac{5 - w^{*2}}{2w^*} = \frac{5 - \frac{5}{3}}{2 \cdot \sqrt{\frac{5}{3}}} = \frac{\frac{10}{3}}{2 \cdot \sqrt{\frac{5}{3}}} = \frac{\sqrt{\frac{10^2}{3^2}}}{2 \cdot \sqrt{\frac{5}{3}}} = \frac{1}{2} \cdot \sqrt{\frac{10^2}{3^2} \cdot \frac{3}{5}} = \frac{1}{2} \cdot \sqrt{\frac{20}{3}} = \sqrt{\frac{5}{3}} \quad (8.64)$$

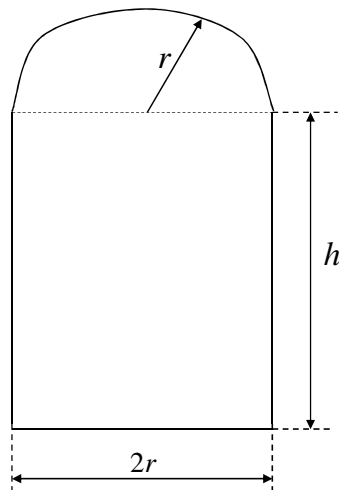
Optimalni volumen je sledeč:

$$V^* = h^* \cdot w^{*2} = \sqrt{\frac{5}{3}} \cdot \left(\frac{5}{3}\right) = \left(\frac{5}{3}\right)^{\frac{3}{2}} = \sqrt{\left(\frac{5}{3}\right)^3} = 2.15m^2 \quad (8.65)$$

Torej, če bomo uporabili kocko s stranico  $\sqrt{\frac{5}{3}} = 1.29m$ , bomo dosegli največji možni volumen  $2.15m^2$ , pri danih  $10m^2$  materiala.

### Primer 13:

Za okno v obliki pravokotnika in polkroga imamo na razpolago  $12m$  materiala. Kakšne naj bodo dimenzije okna, da bo prepuščalo največ svetlobe?



Slika 103: Okno

Imamo optimalni problem z omejitvijo:

$$\begin{aligned} \max S(r, h) &= 2r \cdot h + \frac{\pi \cdot r^2}{2} \\ 12 &= h + 2r + h + \frac{2\pi \cdot r}{2} \end{aligned} \quad (8.66)$$

Dobimo:

$$12 = 2r + 2h + \pi \cdot r \quad \Rightarrow \quad h = \frac{12 - (\pi + 2) \cdot r}{2} = 6 - \frac{\pi \cdot r}{2} - r \quad (8.67)$$

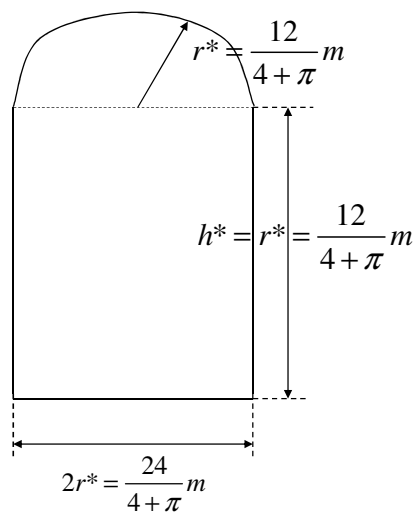
Površina okna, ki jo je potrebno maksimizirati je naslednja:

$$\begin{aligned} S(r) &= 2r \cdot \left(6 - \frac{\pi \cdot r}{2} - r\right) + \frac{\pi \cdot r^2}{2} = 12r - \pi \cdot r^2 - 2r^2 + \frac{\pi \cdot r^2}{2} = \\ &= 12r - \frac{\pi \cdot r^2}{2} - 2r^2 \end{aligned} \quad (8.68)$$

Naredimo odvod, ga enačimo z 0 in dobimo optimalne dimenzije okna:

$$\begin{aligned} \frac{dS}{dr} &= 12 - \frac{2\pi \cdot r}{2} - 4r = 0 \quad \Rightarrow \\ \Rightarrow 12 &= (\pi + 4) \cdot r \quad \Rightarrow \quad r^* = \frac{12}{\pi + 4} = 1.68m \\ h^* &= 6 - \frac{\pi \cdot r^*}{2} - r^* = 6 - \frac{\pi}{2} \cdot \frac{12}{\pi + 4} - \frac{12}{\pi + 4} = \frac{12 \cdot (\pi + 4) - 12\pi - 12 \cdot 2}{(\pi + 4) \cdot 2} = \\ &= \frac{24}{(\pi + 4) \cdot 2} = \frac{12}{\pi + 4} = r^* \end{aligned} \quad (8.69)$$

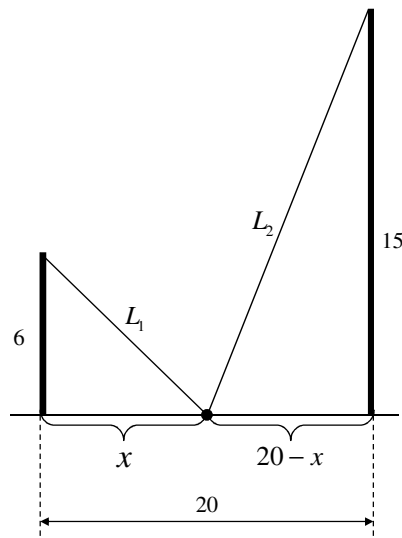
Torej, more okno imeti naslednjo strukturo stranic:



Slika 104: Okno

**Primer 14:**

Dana imamo dva stebra, visoka 6 in 15 m, ki sta 20 m narazen. Nanju (na vrhe stebrov) je privezana žica, ki se nekje med njima stika tal. Kje naj se žica stika tal, da bo zanjo uporabljenega najmanj materiala?



Slika 105: Žica

Optimalni problem je naslednji:

$$\min \{L = L_1 + L_2 = L(x)\}; \quad 0 \leq x \leq 20 \quad (8.70)$$

Razdalja med žicama je naslednja:

$$L(x) = \sqrt{6^2 + x^2} + \sqrt{15^2 + (20 - x)^2} \quad (8.71)$$

Naredimo odvod dobljene razdalje v izrazu (8.71) in ga enačimo z 0:

$$\begin{aligned} \frac{dL}{dx} &= \frac{1}{\sqrt{6^2 + x^2}} \cdot (2x) + \frac{1}{\sqrt{15^2 + (20 - x)^2}} \cdot (2 \cdot (20 - x) \cdot (-1)) = 0 \\ \frac{2x}{\sqrt{6^2 + x^2}} &= \frac{2 \cdot (20 - x)}{\sqrt{15^2 + (20 - x)^2}} \quad (8.72) \\ x \cdot \sqrt{15^2 + (20 - x)^2} &= (20 - x) \cdot \sqrt{6^2 + x^2} \quad /^2 \\ x^2 \cdot (15^2 + (20 - x)^2) &= (20 - x)^2 \cdot (6^2 + x^2) \end{aligned}$$



$$\begin{aligned}
15^2 \cdot x^2 + \cancel{(20-x)^2 \cdot x^2} &= 36 \cdot (20-x)^2 + \cancel{x^2 \cdot (20-x)^2} \\
225x^2 &= 36 \cdot [400 - 40x + x^2] \\
225x^2 &= 36 \cdot 400 - 36 \cdot 40x + 36x^2 \\
189x^2 + 36 \cdot 40x - 36 \cdot 400 &= 0 \\
21x^2 + 160x - 1600 &= 0
\end{aligned}$$

Dobimo razdaljo žice od manjšega stebra:

$$\begin{aligned}
x_{1,2} &= \frac{-160 \pm \sqrt{160^2 + 4 \cdot 21 \cdot 1600}}{42} = \frac{-160 \pm 400}{42} \\
x_1^* &= \frac{-160 - 400}{42} = -\frac{560}{42}, & x_2^* &= \frac{240}{42} = \frac{40}{7} = 5.71m
\end{aligned} \tag{8.73}$$

NE PRIDE  
V POŠTEV

Optimalna dolžina žice je naslednja:

$$L^*(x) = \underbrace{\sqrt{6^2 + \frac{40^2}{49}}}_{8.28m} + \underbrace{\sqrt{15^2 + \left(20 - \frac{40}{7}\right)^2}}_{20.71m} = 29m \tag{8.74}$$

Torej bo minimalna dolžina žice 29m porabljena, če se žica stika tal 5.71m stran od manjšega stebra.

### Primer 15:

Stanovanjsko naselje ima 250 apartmajev. Če lastniki izposodijo  $x$  apartmajev na mesec, se izkaže, da je funkcija profita dana kot:  $P(x) = -8x^2 + 3200x - 80000$ . Koliko apartmajev morajo oddati, da bo profit največji?

Imamo optimalni problem z omejitvijo:

$$\begin{aligned}
\max P(x) \\
0 \leq x \leq 250
\end{aligned} \tag{8.75}$$

Naredimo odvod po številu apartmajev in dobimo:

$$\frac{dP}{dx} = -16x + 3200 = 0 \quad \Rightarrow \quad x^* = \frac{3200}{16} = 200 \tag{8.76}$$

Optimalna funkcija profita pa je sledeča:

$$P(x^*) = -8 \cdot 200^2 + 3200 \cdot 200 - 80000 = 240000 \text{ €} \quad (8.77)$$

Torej je bolje oddati 200 apartmajev, namesto vseh 250, saj je profit večji.

Dokaz:

$$P(250) = -8 \cdot 250^2 + 3200 \cdot 250 - 80000 = 220000 \text{ €} \quad (8.78)$$

### Primer 16:

Podjetje je zmožno izdelati 60 000 kosov na dan, pri čemer je stroškovna funkcija za  $x$  kosov/dan dana kot:  $C(x) = 250000 + 0.08x + \frac{200000.000}{x}$ . Koliko kosov naj podjetje izdela na dan, da bodo stroški najmanjši?

Imamo optimalni problem z omejitvijo:

$$\begin{aligned} \min C(x) \\ 0 \leq x \leq 60000 \end{aligned} \quad (8.79)$$

Naredimo odvod po številu kosov in dobimo:

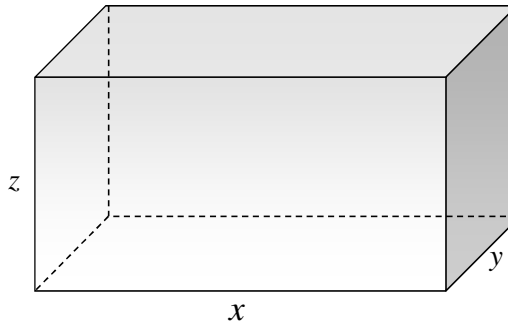
$$\begin{aligned} \frac{dC}{dx} = 0.08 - \frac{200000.000}{x^2} = 0 &\Rightarrow 0.08x^2 = 200000.000 \Rightarrow \\ &\Rightarrow x^* = \sqrt{\frac{200000.000}{0.08}} = 50000 \end{aligned} \quad (8.80)$$

Torej naj podjetje izdela 50 000 kosov/dan, da bodo stroški najmanjši.

Opomba: Za razumevanje teoretičnega ozadja v primerih 17 do 21, ki sledijo v nadaljevanju, glej poglavje 5.4.1, izraze (5.13) do (5.14).

### Primer 17:

Zgraditi želimo velik kovinski kontejner, odprt na vrhu, ki mora držati  $256m^3$ . Kakšne naj bodo dimenzije, da bodo stroški izgradnje čim manjši?



Slika 106: Kontejner

Optimalni problem z omejitvijo:

$$\begin{aligned} \min S &= x \cdot y + 2 \cdot y \cdot z + 2 \cdot x \cdot z \\ V &= 256 = x \cdot y \cdot z \end{aligned} \quad (8.81)$$

Izrazimo spremenljivko  $z$ :

$$z = \frac{256}{x \cdot y} \quad (8.82)$$

Površina kontejnerja je sledeča:

$$\begin{aligned} S(x, y) &= x \cdot y + 2 \cdot \cancel{x} \cdot \frac{256}{x \cdot \cancel{x}} + 2 \cdot \cancel{y} \cdot \frac{256}{\cancel{y} \cdot y} \\ S(x, y) &= x \cdot y + \frac{512}{x} + \frac{512}{y} \quad \Rightarrow \text{POZOR - FUNKCIJA 2. SPREMENLJIVK!} \end{aligned} \quad (8.83)$$

Gradient funkcije površine kontejnerja enačimo z ničelnim vektorjem, da dobimo stacionarne točke:

$$\nabla S(x, y) = \begin{bmatrix} \frac{\partial S}{\partial x} \\ \frac{\partial S}{\partial y} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.84)$$

Gradient funkcije je sledeč:

$$\frac{\partial S}{\partial x} = y - \frac{512}{x^2}; \quad \frac{\partial S}{\partial y} = x - \frac{512}{y^2} \quad (8.85)$$

Sledi:

$$\begin{bmatrix} \frac{\partial S}{\partial x} \\ \frac{\partial S}{\partial y} \end{bmatrix} = \begin{bmatrix} y - \frac{512}{x^2} \\ x - \frac{512}{y^2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.86)$$

Dobimo:

$$\left. \begin{array}{l} y = \frac{512}{x^2} \\ x = \frac{512}{y^2} \end{array} \right\} \Rightarrow x = \sqrt{\frac{512}{y}} \Rightarrow \sqrt{\frac{512}{y}} = \frac{512}{y^2} \Rightarrow \frac{y^2}{\sqrt{y}} = \frac{512}{\sqrt{512}} \Rightarrow y^{\frac{3}{2}} = 512^{\frac{1}{2}}$$

$$\Rightarrow y^* = (\sqrt{512})^{\frac{2}{3}} = \sqrt[3]{512} = 8 \Rightarrow x^* = \sqrt{\frac{512}{8}} = 8 \quad (8.87)$$

Torej, stacionarna točka je:

$$(x^*, y^*) = (8, 8) \quad (8.88)$$

Preverimo, če je to res minimum:

$$\underline{\underline{H}}(x, y) = \nabla^2 S(x, y) = \begin{bmatrix} \frac{\partial^2 S}{\partial x^2} & \frac{\partial^2 S}{\partial x \partial y} \\ \frac{\partial^2 S}{\partial y \partial x} & \frac{\partial^2 S}{\partial y^2} \end{bmatrix} = \begin{bmatrix} \frac{512 \cdot 2}{x^3} & 1 \\ 1 & \frac{512 \cdot 2}{y^3} \end{bmatrix} = \begin{bmatrix} \frac{1024}{x^3} & 1 \\ 1 & \frac{1024}{y^3} \end{bmatrix} \quad (8.89)$$

Vstavimo stacionarno točko, dobljeno v izrazu (8.88) in dobimo:

$$\underline{\underline{H}}(8, 8) = \begin{bmatrix} \frac{1024}{8^3} & 1 \\ 1 & \frac{1024}{8^3} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad (8.90)$$

$\det \underline{\underline{H}}(8, 8) = 2 \cdot 2 - 1 \cdot 1 > 0$  in hkrati  $\frac{\partial^2 S}{\partial x^2}(8, 8) = 2 > 0$ , zato je to sigurno minimum.

Izrazimo še optimalno dolžino stranice  $z^*$ :

$$z^* = \frac{256}{x^* \cdot y^*} = \frac{256}{8 \cdot 8} = \frac{256}{64} = 4 \quad (8.91)$$

Torej, dimenzije morajo biti  $x^* \times y^* \times z^* = 8 \times 8 \times 4$ , da bodo stroški izgradnje kontejnerja najmanjši.

**Primer 18:**

Podjetje proizvaja 2 proizvoda po cenah  $p_1$  in  $p_2$ . Izkaže se, da sta funkciji povpraševanja zanj naslednji:

$$\begin{aligned} Q_1(p_1, p_2) &= 40 - 2p_1 - p_2 \\ Q_2(p_1, p_2) &= 35 - p_1 - p_2 \end{aligned} \quad (8.92)$$

kjer sta količini  $Q_1$  in  $Q_2$ , odvisni od cen  $p_1$  in  $p_2$ , ki morata biti narejeni, da zadovoljita povpraševanje.

Funkcija stroškov pa je naslednja:  $S(p_1, p_2) = Q_1^2 + 2Q_2^2 + 10$ . Kolikšni morata biti ceni  $p_1$  in  $p_2$ , da bo profit maksimalen?

Funkcija prihodkov je dana z naslednjo relacijo:

$$P(p_1, p_2) = p_1 \cdot Q_1(p_1, p_2) + p_2 \cdot Q_2(p_1, p_2) \quad (8.93)$$

Vstavimo količini  $Q_1$  in  $Q_2$  ter dobimo:

$$\begin{aligned} P(p_1, p_2) &= p_1 \cdot (40 - 2p_1 - p_2) + p_2 \cdot (35 - p_1 - p_2) = \\ &= 40p_1 - 2p_1^2 - p_1 \cdot p_2 + 35p_2 - p_1 \cdot p_2 - p_2^2 = \\ &= 40p_1 - 2p_1^2 - 2p_1 \cdot p_2 + 35p_2 - p_2^2 \end{aligned} \quad (8.94)$$

Funkcija profita je dana z enačbo (8.95):

$$\underbrace{PR(p_1, p_2)}_{\text{PROFIT}} = \underbrace{P(p_1, p_2)}_{\text{PRIHODKI}} - \underbrace{S(p_1, p_2)}_{\text{STROŠKI}} \quad (8.95)$$

Vstavimo dobljene prihodke in dane stroške ter dobimo:

$$\begin{aligned} PR(p_1, p_2) &= 40p_1 - 2p_1^2 - 2p_1 \cdot p_2 + 35p_2 - p_2^2 - \\ &\quad - \left[ (40 - 2p_1 - p_2)^2 + 2 \cdot (35 - p_1 - p_2)^2 + 10 \right] \end{aligned} \quad (8.96)$$

Sledi:

$$\begin{aligned} (40 - (2p_1 + p_2))^2 &= 40^2 - 2 \cdot 40 \cdot (2p_1 + p_2) + (2p_1 + p_2)^2 = \\ &= 1600 - 160p_1 - 80p_2 + 4p_1^2 + 4p_1p_2 + p_2^2 \end{aligned} \quad (8.97)$$

Sledi:

$$(35 - (p_1 + p_2))^2 = 35^2 - 2 \cdot 35 \cdot (p_1 + p_2) + p_1^2 + 2p_1 \cdot p_2 + p_2^2 \quad (8.98)$$

Funkcija profita se za dani primer torej glasi:

$$\begin{aligned} PR(p_1, p_2) &= 40p_1 - 2p_1^2 - 2p_1 \cdot p_2 + 35p_2 - p_2^2 - \{1600 - 160p_1 - 80p_2 + \\ &\quad + 4p_1^2 + 4p_1p_2 + p_2^2 + 2 \cdot 35^2 - 4 \cdot 35 \cdot (p_1 + p_2) + 2p_1^2 + \\ &\quad + 4p_1 \cdot p_2 + 2p_2^2 + 10\} = \\ &= 340p_1 - 8p_1^2 - 10p_1 \cdot p_2 + 255p_2 - 4p_2^2 - 4060 \end{aligned} \quad (8.99)$$

Gradient dane funkcije profita enačimo z ničelnim vektorjem in dobimo:

$$\begin{aligned} \frac{\partial PR}{\partial p_1} &= 340 - 16p_1 - 10p_2 = 0 && \Rightarrow 340 = 16p_1 + 10p_2 \\ \frac{\partial PR}{\partial p_2} &= -10p_1 + 255 - 8p_2 = 0 && \Rightarrow 255 = 10p_1 + 8p_2 \end{aligned} \quad (8.100)$$

Zapišimo v vektorsko matrični obliki:

$$\begin{bmatrix} 16 & 10 \\ 10 & 8 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 340 \\ 255 \end{bmatrix} \quad (8.101)$$

Sledi:

$$\begin{aligned} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}^* &= \begin{bmatrix} 16 & 10 \\ 10 & 8 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 340 \\ 255 \end{bmatrix} = \frac{\begin{bmatrix} 8 & -10 \\ -10 & 16 \end{bmatrix}}{16 \cdot 8 - 100} \cdot \begin{bmatrix} 340 \\ 255 \end{bmatrix} = \\ &= \frac{1}{28} \cdot \begin{bmatrix} 8 \cdot 340 - 10 \cdot 255 \\ -10 \cdot 340 + 16 \cdot 255 \end{bmatrix} = \frac{1}{28} \cdot \begin{bmatrix} 170 \\ 680 \end{bmatrix} = \begin{bmatrix} 6.07 \\ 24.28 \end{bmatrix} = \begin{bmatrix} p_1^* \\ p_2^* \end{bmatrix} \end{aligned} \quad (8.102)$$

Hessova matrika je sledeča:

$$\underline{\underline{H}}(p_1, p_1) = \begin{bmatrix} \frac{\partial^2 PR}{\partial p_1^2} & \frac{\partial^2 PR}{\partial p_1 \partial p_2} \\ \frac{\partial^2 PR}{\partial p_2 \partial p_1} & \frac{\partial^2 PR}{\partial p_2^2} \end{bmatrix} = \begin{bmatrix} -16 & -10 \\ -10 & -8 \end{bmatrix} \quad (8.103)$$

Preverimo, če je to res maksimum:

$$\left. \begin{aligned} \det \underline{H}(p_1, p_2) &= (-16) \cdot (-8) - (-10)^2 = \\ &= 28 > 0 \\ \frac{\partial^2 PR}{\partial p_1^2} &< 0 \end{aligned} \right\} \begin{aligned} &\text{Zato je } v(p_1^*, p_2^*) = (6.07, 24.28) \\ &\text{MAKSIMUM funkcije } PR \end{aligned} \quad (8.104)$$

Maksimalen profit torej je:

$$\begin{aligned} PR(p_1^*, p_2^*) &= 340 \cdot 6.07 - 8 \cdot 6.07^2 - 10 \cdot 6.07 \cdot 24.28 + 255 \cdot 24.28 - \\ &- 4 \cdot 24.28^2 - 4060 = 68.57 \end{aligned} \quad (8.105)$$

### Primer 19:

Poišči in določi tip stacionarne točke za:  $f(x, y) = x \cdot y - x^4 - y^2 + 2$

Gradient dane funkcije enačimo z ničelnim vektorjem:

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} y - 4x^3 \\ x - 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.106)$$

Sledi:

$$\left. \begin{aligned} y - 4x^3 &= 0 & \Rightarrow y &= 4x^3 \\ x - 2y &= 0 & \Rightarrow y &= \frac{x}{2} \end{aligned} \right\} \Rightarrow 4x^3 = \frac{x}{2} \Rightarrow 4x^3 - \frac{x}{2} = 0 / \div 4 \Rightarrow$$

$$\Rightarrow x^3 - \frac{x}{8} = 0 \quad (8.107)$$

$$\Rightarrow x \cdot \left( x^2 - \frac{1}{8} \right) = 0 \Rightarrow$$

$$\Rightarrow x \cdot \left( x - \frac{1}{\sqrt{8}} \right) \cdot \left( x + \frac{1}{\sqrt{8}} \right) = 0$$

Sledi:

$$x_1^* = 0, \quad x_2^* = \frac{1}{\sqrt{8}}, \quad x_3^* = -\frac{1}{\sqrt{8}} \quad (8.108)$$

Stacionarne točke so sledeče:

$$\begin{aligned}
 \text{Če je } x_1^* = 0 & \Rightarrow y_1^* = \frac{0}{2} = 0 & \Rightarrow T_1^*(0,0) \\
 \text{Če je } x_2^* = \frac{1}{\sqrt{8}} & \Rightarrow y_2^* = \frac{1}{2 \cdot \sqrt{8}} & \Rightarrow T_2^*\left(\frac{1}{\sqrt{8}}, \frac{1}{2 \cdot \sqrt{8}}\right) \\
 \text{Če je } x_3^* = -\frac{1}{\sqrt{8}} & \Rightarrow y_3^* = -\frac{1}{2 \cdot \sqrt{8}} & \Rightarrow T_3^*\left(-\frac{1}{\sqrt{8}}, -\frac{1}{2 \cdot \sqrt{8}}\right)
 \end{aligned}
 \tag{8.109}$$

Imamo 3 stacionarne točke.

Preverimo v katerih izmed dobljenih stacionarnih točk smo dobili optimum:

$$\nabla^2 f(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} -12x^2 & 1 \\ 1 & -2 \end{bmatrix}
 \tag{8.110}$$

a.) Za 1. stacionarno točko:

$$\begin{aligned}
 \nabla^2 f(0,0) &= \begin{bmatrix} -12 \cdot 0^2 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} = H(0,0) \\
 \det \underline{\underline{H}}(0,0) &= 0 \cdot (-2) - 1 \cdot 1 = -1 < 0 \quad \Rightarrow T_1^*(0,0) \text{ je sedelna točka}
 \end{aligned}
 \tag{8.111}$$

b.) Za 2. stacionarno točko:

$$\begin{aligned}
 \underline{\underline{H}}\left(\frac{1}{\sqrt{8}}, \frac{1}{2 \cdot \sqrt{8}}\right) &= \begin{bmatrix} -12 \cdot \left(\frac{1}{\sqrt{8}}\right)^2 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} -12 \cdot \frac{1}{8} & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & 1 \\ 1 & -2 \end{bmatrix} \\
 \det \underline{\underline{H}}\left(\frac{1}{\sqrt{8}}, \frac{1}{2 \cdot \sqrt{8}}\right) &= \left(-\frac{3}{2}\right) \cdot (-2) - 1 \cdot 1 = 3 - 1 = 2 > 0 \text{ in ker je} \\
 \frac{\partial^2 f}{\partial x^2}\left(\frac{1}{\sqrt{8}}, \frac{1}{2 \cdot \sqrt{8}}\right) &= -\frac{3}{2} < 0, \text{ sledi, da je } T_2^*\left(\frac{1}{\sqrt{8}}, \frac{1}{2 \cdot \sqrt{8}}\right) \text{ maksimum.}
 \end{aligned}
 \tag{8.112}$$

c.) Za 3. stacionarno točko:

$$\underline{\underline{H}}\left(-\frac{1}{\sqrt{8}}, -\frac{1}{2 \cdot \sqrt{8}}\right) = \begin{bmatrix} -12 \cdot \left(-\frac{1}{\sqrt{8}}\right)^2 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & 1 \\ 1 & -2 \end{bmatrix}
 \tag{8.113}$$



Isto kot pri b.), zato je  $T_3^* \left( -\frac{1}{\sqrt{8}}, -\frac{1}{2 \cdot \sqrt{8}} \right)$  tudi maksimum.

**Primer 20:**

Klasificiraj stacionarne točke za:  $f(x, y) = x^3 + y^3 - 6xy$

Gradient dane funkcije enačimo z ničelnim vektorjem:

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 3x^2 - 6y \\ 3y^2 - 6x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.114)$$

Sledi:

$$\left. \begin{array}{l} 3x^2 - 6y = 0 / \cdot x \\ 3y^2 - 6x = 0 / \cdot y \end{array} \right\} - \quad (8.115)$$

Sledi:

$$\begin{aligned} 3x^3 - 6xy - \{3y^3 - 6xy\} &= 0 \\ x^3 = y^3 &\Rightarrow x = y \end{aligned} \quad (8.116)$$

Iz 1. velja:

$$\begin{aligned} 3x^2 - 6x = 0 &\Rightarrow x^2 - 2x = 0 &\Rightarrow x \cdot (x - 2) = 0 &\Rightarrow \\ x_1^* = 0, & \quad x_2^* = 2 & & \end{aligned} \quad (8.117)$$

Zaradi (8.116) je:

$$y_1^* = 0, \quad y_2^* = 2 \quad (8.118)$$

Stacionarni točki sta sledeči:

$$T_1^* = (0, 0), \quad T_2^* = (2, 2) \quad (8.119)$$

Preverimo v katerih izmed dobljenih stacionarnih točk smo dobili optimum:

$$\underline{\underline{H}}(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 6x & -6 \\ -6 & 6y \end{bmatrix} \quad (8.120)$$

1.) Za 1. stacionarno točko velja:

$$\underline{\underline{H}}(0,0) = \begin{bmatrix} 6 \cdot 0 & -6 \\ -6 & 6 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 & -6 \\ -6 & 0 \end{bmatrix} \quad (8.121)$$

$$\det \underline{\underline{H}}(0,0) = 0 \cdot 0 - (-6)^2 = -36 < 0 \quad \Rightarrow T_1^* \text{ je sedlo}$$

2.) Za 2. stacionarno točko velja:

$$\underline{\underline{H}}(2,2) = \begin{bmatrix} 6 \cdot 2 & -6 \\ -6 & 6 \cdot 2 \end{bmatrix} = \begin{bmatrix} 12 & -6 \\ -6 & 12 \end{bmatrix}$$

$$\det \underline{\underline{H}}(2,2) = 12^2 - (-6)^2 = 144 - 36 = 108 > 0 \quad (8.122)$$

ker velja tudi  $\frac{\partial^2 f}{\partial x^2}(2,2) = 12 > 0$ , sledi, da je  $T_2^*$  minimum.

### Primer 21:

Poišči in klasificiraj vse stacionarne točke od funkcije  $f(x, y) = e^{-(x^2+y^2-4y)}$

Gradient dane funkcije enačimo z ničelnim vektorjem:

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} e^{-(x^2+y^2-4y)} \cdot (-2x) \\ e^{-(x^2+y^2-4y)} \cdot (-2y+4) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.123)$$

Sledi:

$$\begin{aligned} e^{-(x^2+y^2-4y)} \cdot (-2x) &= 0 & \Rightarrow x &= 0 \\ \underbrace{-e^{-(x^2+y^2-4y)} \cdot (2y-4)}_{\text{to je 0 šele v } \infty} &= 0 & \Rightarrow y &= 2 \end{aligned} \quad (8.124)$$

Edina stacionarna točka je:  $T^* = (0, 2)$ . Preverimo, če se v njej res skriva optimum:

$$\underline{\underline{H}}(x, y) = \underline{\underline{\nabla^2 f}}(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \quad (8.125)$$

Sledi:

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x} \cdot \left[ -2x \cdot e^{-(x^2+y^2-4y)} \right] = -2 \cdot \left[ e^{-(x^2+y^2-4y)} + x \cdot e^{-(x^2+y^2-4y)} \cdot (-2x) \right] = \\ &= -2 \cdot e^{-(x^2+y^2-4y)} \cdot \{1 - 2x^2\} \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial y} \cdot \left[ -2x \cdot e^{-(x^2+y^2-4y)} \right] = -2x \cdot e^{-(x^2+y^2-4y)} \cdot (-2y + 4) = \\ &= e^{-(x^2+y^2-4y)} \cdot (4xy - 8x) \\ \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial x} \cdot \left[ -e^{-(x^2+y^2-4y)} \cdot (4 - 2y) \right] = (4 - 2y) \cdot e^{-(x^2+y^2-4y)} \cdot (-2x) = \\ &= e^{-(x^2+y^2-4y)} \cdot (4xy - 8x) \\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y} \cdot \left[ e^{-(x^2+y^2-4y)} \cdot (4 - 2y) \right] = \\ &= (-2) \cdot e^{-(x^2+y^2-4y)} + (4 - 2y) \cdot e^{-(x^2+y^2-4y)} \cdot (-2) = \\ &= e^{-(x^2+y^2-4y)} \cdot \left[ (4 - 2y)^2 - 2 \right] \end{aligned} \quad (8.126)$$

Sledi:

$$\begin{aligned} \underline{\underline{H}}(x, y) &= e^{-(x^2+y^2-4y)} \cdot \begin{bmatrix} -2 \cdot (1 - 2x^2) & 4xy - 8x \\ 4xy - 8x & (4 - 2y)^2 - 2 \end{bmatrix} \\ \underline{\underline{H}}(0, 2) &= e^{-(0^2+2^2-4 \cdot 2)} \cdot \begin{bmatrix} -2 \cdot (1 - 2 \cdot 0^2) & 4 \cdot 0 \cdot 2 - 8 \cdot 0 \\ 4 \cdot 0 \cdot 2 - 8 \cdot 0 & (4 - 2 \cdot 2)^2 - 2 \end{bmatrix} = \\ &= e^{+4} \cdot \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} -2e^4 & 0 \\ 0 & -2e^4 \end{bmatrix} \end{aligned} \quad (8.127)$$

Dobimo:

$$\det \underline{\underline{H}}(0, 2) = (-2e^4)^2 - 0^2 = 4 \cdot e^8 > 0 \quad (8.128)$$

Ker je  $\frac{\partial^2 f}{\partial x^2}(0,2) = -2e^4 < 0$ , sledi, da je  $T^*$  lokalni maksimum.

### 8.1.2 Reševanje z Lagrangeovimi multiplikatorji za probleme z omejitvami

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 5.4.4, izraze (5.29) do (5.43).

#### Primer 1:

Denimo, da proizvajalec na dan izdelava  $x$  kosov ene dobrine in  $y$  kosov druge dobrine. Funkcija profita je dana v obliki (normirani!):

$$U(x, y) = (x+1)(y+1) = xy + x + y + 1 \quad (8.129)$$

Cena prve dobrine na kos je 1, cena druge dobrine na kos pa 2 denarni enoti. Koliko kosov prve in druge dobrine mora proizvajalec izdelati na dan, da bo  $U(x,y)$  maksimalen, pri čemer je dnevni proračun 30 denarnih enot. Gre za optimizacijski problem, kjer iščemo maksimum kriterijske funkcije z omejitvami:

$$\begin{aligned} \text{kriterijska funkcija : } & \max U(x, y) \\ \text{omejitev : } & 1 \cdot x + 2 \cdot y = 30 \end{aligned} \quad (8.130)$$

I. način reševanja (substitucija oz. eliminacija ene spremenljivke):

$$x = 30 - 2y$$

$$U(y) = (30 - 2y)y + 30 - 2y + y + 1 = 30y - 2y^2 + 31 - y = -2y^2 + 29y + 31$$

$$\frac{dU}{dy} = -4y + 29 = 0 \Rightarrow y^* = \frac{29}{4}$$

$$x^* = 30 - 2y^* = 30 - 2 \cdot \frac{29}{4} = \frac{31}{2} \quad (8.131)$$

$$(x^*, y^*) = \left( \frac{31}{2}, \frac{29}{4} \right)$$

$$\frac{d^2U}{dy^2} = -4 < 0 \Rightarrow \forall y^* = \frac{29}{4} \text{ doseže } U(y^*) \text{ maksimum}$$

II. način reševanja (Lagrangeovi multiplikatorji):

$$U(x, y) = xy + x + y + 1$$

$$x + 2y = 30 \Rightarrow x + 2y - 30 = h(x, y)$$

$$L(x, y, \lambda) = U(x, y) + \lambda \cdot h(x, y) = xy + x + y + 1 + \lambda(x + 2y - 30) \quad (8.132)$$

$$\frac{\partial L}{\partial x} = y + 1 + \lambda = 0 \Rightarrow y = -1 - \lambda$$

$$\frac{\partial L}{\partial y} = x + 1 + 2\lambda = 0 \Rightarrow x = -1 - 2\lambda$$

Iz (8.132) sledi:

$$\frac{\partial L}{\partial \lambda} = x + 2y - 30 = 0 \quad (8.133)$$

Sledi:

$$(-1 - 2\lambda) + 2(-1 - \lambda) - 30 = 0 \quad (8.134)$$

Dobimo:

$$-3 - 4\lambda - 30 = 0 \Rightarrow \lambda^* = -\frac{33}{4} \quad (8.135)$$

$$x^* = -1 - 2\left(-\frac{33}{4}\right) = -1 + \frac{33}{2} = \frac{31}{2}$$

Iz (8.135) sledi:

$$y^* = -1 - \left(-\frac{33}{4}\right) = -1 + \frac{33}{4} = \frac{29}{4} \quad (8.136)$$

$$(x^*, y^*) = \left(\frac{31}{2}, \frac{29}{4}\right)$$

Dobimo seveda isti rezultat kot prej. Preverimo tudi lahko, če v  $(x^*, y^*, \lambda^*)$  doseže  $L(x^*, y^*, \lambda^*)$  res maksimum (tedaj v  $x^*, y^*$  doseže  $U(x^*, y^*)$  tudi maksimum).

Denimo izberemo:

$$x^* = \frac{30}{2} = 15, \quad y^* = \frac{28}{4} = 7 \quad (8.137)$$

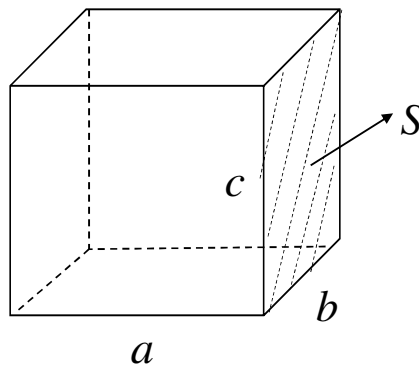
Potem velja:

$$x^* \cdot 2y^* = 15 + 2 \cdot 7 = 29 < 30 \quad \text{OK} \quad (8.138)$$

Torej, proizvajalec mora izdelati 15 kosov prve dobrine in 7 kosov druge dobrine na dan, da bo imel največji prihodek.

**Primer 2:**

Podjetje izdeluje kovinske škatle brez pokrova, ki morajo biti kvadri volumna  $200dm^3$ . Kakšne naj bodo dimenzije škatel, da bodo stroški najmanjši (rešujte s pomočjo metode Lagrangeovih multiplikatorjev!).



Slika 107: Kovinska škatla

Imamo optimizacijski problem z omejitvijo:

$$\begin{aligned} \min S &= ab + 2ac + 2bc \\ V &= abc = 200 \end{aligned} \tag{8.139}$$

Površina stranice je sledeča:

$$S(a, b, c) = ab + 2ac + 2bc \tag{8.140}$$

Lagrangeova funkcija je naslednja:

$$L(a, b, c, \lambda) = S(a, b, c) + \lambda(200 - abc) = ab + 2ac + 2bc + \lambda(200 - abc) \tag{8.141}$$

Funkcijo (8.141) odvajamo po dotičnih spremenljivkah in dobljene relacije enačimo z 0:

$$\begin{aligned}
\frac{\partial L}{\partial a} &= b + 2c - \lambda bc = 0 \\
\frac{\partial L}{\partial b} &= a + 2c - \lambda ac = 0 \\
\frac{\partial L}{\partial c} &= 2a + 2b - \lambda ab = 0 \\
\frac{\partial L}{\partial \lambda} &= 200 - abc = 0
\end{aligned}
\tag{8.142}$$

Sledi:

$$\begin{aligned}
b + 2c - \lambda bc = 0 &\Rightarrow c = \frac{-b}{2 - \lambda b} \\
a + 2c - \lambda ac = 0 &\Rightarrow c = \frac{-a}{2 - \lambda a}
\end{aligned}
\tag{8.143}$$

Enačimo izraza v (8.143) in dobimo:

$$\begin{aligned}
\frac{b}{2 - \lambda b} &= \frac{a}{2 - \lambda a} \\
&\downarrow \\
b(2 - \lambda a) &= a(2 - \lambda b) \\
2b - \lambda ab &= 2a - \lambda ab
\end{aligned}
\tag{8.144}$$

Sledi:

$$a^* = b^* \tag{8.145}$$

Sledi:

$$\begin{aligned}
2a + 2b - \lambda ab &= 0 \\
2a + 2a - \lambda aa &= 0 \\
4a - \lambda a^2 = 0 &\Rightarrow \lambda = \frac{4a}{a^2} = \frac{4}{a}
\end{aligned}
\tag{8.146}$$

Iz (8.146) sledi:

$$c^* = \frac{-a}{2 - \lambda a} = \frac{-a}{2 - \frac{4}{a}a} = \frac{-a}{-2} = \frac{a^*}{2} \tag{8.147}$$

Dobimo:

$$\begin{aligned}
200 - abc &= 0 \\
200 &= a^* b^* c^* \\
200 &= a^* a^* \frac{a^*}{2} \Rightarrow a^{*3} = 400 \Rightarrow a^* = \sqrt[3]{400} = 7.36dm \\
c^* &= \frac{a^*}{2} = 3.68dm
\end{aligned} \tag{8.148}$$

Optimalne dimenzije škatle (kjer bodo stroški njene izdelave najmanjši) torej so:

$$\begin{aligned}
a^* &= b^* = 7.36dm \\
c^* &= 3.68dm
\end{aligned} \tag{8.149}$$

Površina pa je sledeča:

$$\begin{aligned}
S^* &= a^* b^* + 2a^* c^* + 2b^* c^* = a^{*2} + 2a^* \frac{a^*}{2} + 2a^* \frac{a^*}{2} = \\
&= a^{*2} + a^{*2} + a^{*2} = 3a^{*2} = 3 \cdot 7.36^2 = 162.5dm^2
\end{aligned} \tag{8.150}$$

Za izdelavo ene optimalne škatle bomo porabili  $162.5dm^2$  materiala.

### Primer 3:

Količina proizvodnje neke tovarne je določena z COBB-DOUGLASOVO funkcijo proizvodnje:

$$Q(L, K) = 12 \cdot \sqrt[3]{L} \cdot \sqrt[3]{K^2} = 12 \cdot \sqrt[3]{LK^2} \tag{8.151}$$

Cene proizvodnih faktorjev dela in kapitala so 3 in 6 enot (na vloženo enoto dela oz. kapitala). Kakšen je minimalni proračun, potreben za proizvodnjo 1200 produktov?

Optimizacijski problem se glasi:

$$\begin{aligned}
\min B(L, K) &= 3L + 6K \\
Q(L, K) &= 12 \cdot \sqrt[3]{LK^2} = 1200 \Rightarrow \sqrt[3]{LK^2} = 100
\end{aligned} \tag{8.152}$$

I. način:

$$\sqrt[3]{LK^2} = 100 \Rightarrow LK^2 = 100^3 \Rightarrow L = \frac{100^3}{K^2} \tag{8.153}$$

Sledi:



$$\begin{aligned}
B(K) &= 3 \cdot \frac{100^3}{K^2} + 6K \\
\frac{dB}{dK} &= -3 \cdot 2 \cdot \frac{100^3}{K^3} + 6 = 0 \Rightarrow K^* = 100 \\
L^* &= \frac{100^3}{K^{*2}} = \frac{100^3}{100^2} = 100 \\
B(L^*, K^*) &= 3 \cdot 100 + 6 \cdot 100 = 900 \text{ enot}
\end{aligned}
\tag{8.154}$$

Torej, če angažiramo 100 enot dela, 100 enot kapitala, ter višino proračuna (budžet) 900 enot, lahko z minimalnim možnim budžetom proizvedemo 1200 produktov.

II. način:

$$\begin{aligned}
F(L, K, \lambda) &= \overbrace{B(L, K)}^{3L+6K} + \lambda \left( 12 \cdot \sqrt[3]{LK^2} - 1200 \right) \\
\frac{\partial F}{\partial L} &= 3 + \lambda \cdot \frac{1}{2} \cdot (LK^2)^{-\frac{1}{2}} \cdot K^2 = 0 \\
\frac{\partial F}{\partial K} &= 6 + \lambda \cdot \frac{1}{2} \cdot (LK^2)^{-\frac{1}{2}} \cdot L \cdot 2K = 0 \\
\frac{\partial F}{\partial \lambda} &= 12 \cdot \sqrt[3]{LK^2} - 1200 = 0
\end{aligned}
\tag{8.155}$$

Sledi:

$$\left. \begin{aligned}
3 + \frac{\lambda K^2}{2\sqrt{LK^2}} &= 0 \\
6 + \frac{\lambda LK}{\sqrt{LK^2}} &= 0 \quad \left| \cdot \frac{K}{2L} \right.
\end{aligned} \right\} -
\tag{8.156}$$

Odštejemo izraza v (8.156) in dobimo:

$$3 - \frac{6K}{2L} + \frac{\lambda K^2}{2\sqrt{LK^2}} - \frac{\lambda LK^2}{\sqrt{LK^2} \cdot 2L} = 0
\tag{8.157}$$

Sledi:

$$\begin{aligned}
3 - \frac{6K}{2L} &= 0 \\
3 - \frac{3K}{L} &= 0 \quad \Big| \cdot L \\
3L - 3K &= 0 \Rightarrow L^* = K^*
\end{aligned}
\tag{8.158}$$

Sledi:

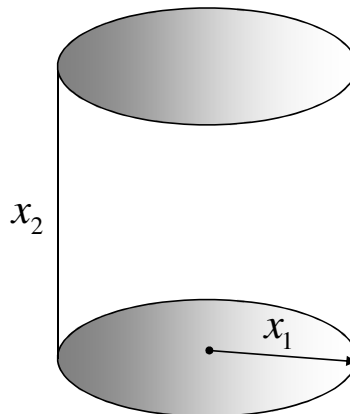
$$\begin{aligned}
12 \cdot \sqrt[3]{L \cdot L^2} &= 1200 \\
\sqrt[3]{L^3} &= 100 \Rightarrow L^* = 100 \Rightarrow K^* = 100
\end{aligned}
\tag{8.159}$$

Optimalni parametri so sledeči:

$$\begin{aligned}
L^* &= 100 \\
K^* &= 100 \\
B(L^*, K^*) &= 3 \cdot 100 + 6 \cdot 100 = 900
\end{aligned}
\tag{8.160}$$

#### Primer 4:

Poiščite optimalne dimenzije cilindričnega kontejnerja (s pokrovom), ki bodo dosegle njegov maksimalen volumen, pri čemer naj bo celotna površina enaka  $24\pi$ .



Slika 108: Cilindrični kontejner

Optimizacijski problem:

$$\begin{aligned}
\max V &= \pi x_1^2 \cdot x_2 = V(x_1, x_2) \\
S &= 2\pi x_1^2 + 2\pi x_1 x_2 = 24\pi
\end{aligned}
\tag{8.161}$$

Lagrangeova funkcija:

$$L(x_1, x_2, \lambda) = \underbrace{V(x_1, x_2)}_{\pi x_1^2 x_2} + \lambda(24\pi - 2\pi x_1^2 - 2\pi x_1 x_2) \quad (8.162)$$

Sledi:

$$\begin{aligned} 1. \quad \frac{\partial L}{\partial x_1} &= 2\pi x_1 x_2 - \lambda 4\pi x_1 - \lambda 2\pi x_2 = 0 & \Rightarrow \lambda &= \frac{x_1 x_2 \cdot 2}{4x_1 + 2x_2} = \frac{x_1 x_2}{2x_1 + x_2} \\ 2. \quad \frac{\partial L}{\partial x_2} &= \pi x_1^2 - \lambda 2\pi x_1 = 0 & \Rightarrow \lambda &= \frac{x_1}{2} \\ 3. \quad \frac{\partial L}{\partial \lambda} &= 24\pi - 2\pi x_1^2 - 2\pi x_1 x_2 = 0 & \Rightarrow 12 &= x_1^2 + x_1 x_2 \end{aligned} \quad (8.163)$$

Izenačimo 1. in 2.:

$$\frac{x_1 x_2}{2x_1 + x_2} = \frac{x_1}{2} = \lambda \quad (8.164)$$

Sledi:

$$\begin{aligned} 2x_1 x_2 &= x_1(2x_1 + x_2) \\ 2x_2 &= 2x_1 + x_2 \\ x_2 &= 2x_1 \end{aligned} \quad (8.165)$$

Vstavimo izraz (8.165) v enačbo 3. in dobimo:

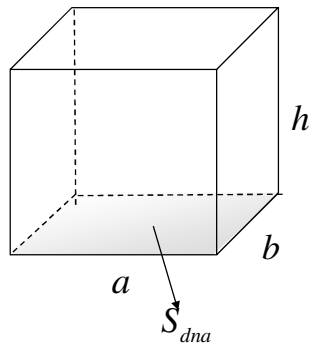
$$\begin{aligned} 12 &= x_1^2 + x_1 \cdot 2x_1 = 3x_1^2 & \Rightarrow x_1 &\pm 2; & \text{Vzamemo } x_1^* &= 2. \\ x_2^* &= 2x_1^* = 2 \cdot 2 = 4 \end{aligned} \quad (8.166)$$

Stacionarna točka torej je:  $(x_1^*, x_2^*) = (2, 4)$ . Dokazati bi še morali, da je tu res maksimum.

Torej, če bo radij enak 2, višina pa  $2 \times$  večja od radija, bomo dosegli maksimalen volumen pri  $S = 24\pi$ .

**Primer 5:**

Želimo izdelati določeno škatlo, pri čemer je strošek izdelave dna  $3\times$  večji od stroška izdelave stranic in vrha. Najdite optimalne dimenzije škatle, ki bodo minimizirale strošek izdelave pri danem volumnu  $V$ .



Slika 109: Škatla

Površine posameznih stranic škatle so naslednje:

$$\begin{aligned} S_{dna} &= ab \\ S_{ostalo} &= ab + 2bh + 2ah \end{aligned} \quad (8.167)$$

Strošek izdelave pa se glasi:

$$C = 3c \cdot S_{dna} + c \cdot S_{ostalo} \quad (8.168)$$

kjer je  $c =$  strošek izdelave stranic in vrha /  $m^2$ .

Optimizacijski problem:

$$\begin{aligned} C &= 3c \cdot a \cdot b + c \cdot (ab + 2bh + 2ah) = C(a, b, h) \\ V &= a \cdot b \cdot h \end{aligned} \quad (8.169)$$

Lagrangeova funkcija:

$$\begin{aligned} L(a, b, h) &= C(a, b, h) + \lambda(V - abh) \\ L(a, b, h) &= 3abc + c(ab + 2bh + 2ah) + \lambda(V - abh) \end{aligned} \quad (8.170)$$

Sledi:

$$\begin{aligned}
1. \quad \frac{\partial L}{\partial a} &= 3bc + cb + 2ch - \lambda bh = 0 & \Rightarrow 4bc + 2ch &= \lambda bh \\
2. \quad \frac{\partial L}{\partial b} &= 3ac + ca + 2ch - \lambda ah = 0 & \Rightarrow 4ac + 2ch &= \lambda ah \\
3. \quad \frac{\partial L}{\partial h} &= 2bc + 2ac - \lambda ab = 0 & \Rightarrow 2ac + 2bc &= \lambda ab \\
4. \quad \frac{\partial L}{\partial \lambda} &= V - abh = 0 & \Rightarrow V &= abh
\end{aligned}
\tag{8.171}$$

Izrazimo  $h$  iz 1.:

$$h = \frac{4bc}{\lambda b - 2c} \tag{8.172}$$

Izrazimo  $h$  iz 2.:

$$h = \frac{4ac}{\lambda a - 2c} \tag{8.173}$$

Izenačimo (8.172) in (8.173) ter dobimo:

$$\begin{aligned}
\frac{4bc}{\lambda b - 2c} &= \frac{4ac}{\lambda a - 2c} \\
b(\lambda a - 2c) &= a(\lambda b - 2c) \\
-2bc &= -2ac \quad \Rightarrow \quad a^* = b^*
\end{aligned}
\tag{8.174}$$

Izraz (8.174) vstavimo v enačbo 3. in dobimo:

$$\begin{aligned}
2a^*c + 2a^*c &= \lambda^* a^* a^* \\
4a^*c &= \lambda^* a^{*2} \quad \Rightarrow \quad \lambda^* = \frac{4c}{a^*}
\end{aligned}
\tag{8.175}$$

Dobimo povezavo med stranicama  $h$  in  $a$  :

$$h^* = \frac{4a^*c}{\lambda^* a^* - 2c} = \frac{4a^*c}{\frac{4c}{a^*} a^* - 2c} = \frac{4a^*c}{2c} = 2a^* \tag{8.176}$$

Dobimo povezavo med volumnom in stranico  $a$ :

$$V = a^* \cdot a^* \cdot h^* = a^* \cdot a^* \cdot 2a^* = 2a^{*3} \quad \Rightarrow \quad a^* = \sqrt[3]{\frac{V}{2}} \tag{8.177}$$

Dobimo povezavo med volumnom in stranico  $b$  :

$$b^* = \sqrt[3]{\frac{V}{2}} \quad (8.178)$$

Dobimo povezavo med volumnom in stranico  $h$ :

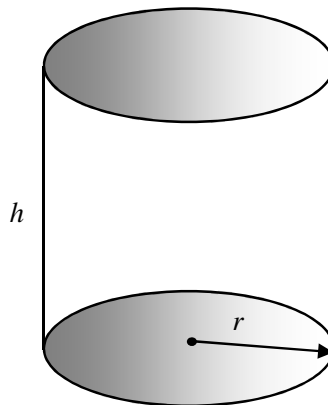
$$h^* = 2a^* = 2 \cdot \sqrt[3]{\frac{V}{2}} = \sqrt[3]{8 \cdot \frac{V}{2}} = \sqrt[3]{4V} \quad (8.179)$$

Torej, če sta stranici škatle  $\sqrt[3]{\frac{V}{2}}$ , ter je višina  $\sqrt[3]{4V}$ , bodo stroški izdelave škatle minimalni pri danih pogojih.

### Primer 6:

Stroški izdelave kovinskega cilindričnega kontejnerja, ki nosi 300l, so  $8 \frac{\text{€}}{\text{dm}^2}$  za dno,  $6 \frac{\text{€}}{\text{dm}^2}$  za pokrov, ter  $4 \frac{\text{€}}{\text{dm}^2}$  za plašč.

Določite optimalne dimenzije, pri katerih so stroški izdelave minimalni. Rešujte z metodo Lagrangeovih multiplikatorjev.



Slika 110: Cilindrični kontejner

Površine posameznih stranic cilindričnega kontejnerja so naslednje:

$$\begin{aligned} S_{dno} &= \pi r^2 \\ S_{vrh} &= \pi r^2 \\ S_{plašč} &= 2\pi r h \end{aligned} \quad (8.180)$$

Strošek izdelave pa se glasi:

$$\begin{aligned} C &= 8 \cdot S_{dno} + 6 \cdot S_{vrh} + 4 \cdot S_{plášč} \\ C(r, h) &= 8 \cdot \pi r^2 + 6 \cdot \pi r^2 + 4 \cdot 2\pi rh \end{aligned} \quad (8.181)$$

Optimizacijski problem:

$$\begin{aligned} C(r, h) &= 14\pi r^2 + 8\pi rh \\ 300 &= \pi r^2 h \end{aligned} \quad (8.182)$$

Lagrangeova funkcija:

$$L(r, h, \lambda) = C(r, h) + \lambda(300 - \pi r^2 h) = 14\pi r^2 + 8\pi rh + \lambda(300 - \pi r^2 h) \quad (8.183)$$

Sledi:

$$\begin{aligned} 1. \quad \frac{\partial L}{\partial r} &= 28\pi r + 8\pi h - 2\lambda\pi rh = 0 & \Rightarrow \lambda &= \frac{14r + 4h}{rh} \\ 2. \quad \frac{\partial L}{\partial h} &= 8\pi r - \lambda\pi r^2 = 0 & \Rightarrow \lambda &= \frac{8}{r} \\ 3. \quad \frac{\partial L}{\partial \lambda} &= 300 - \pi r^2 h = 0 \end{aligned} \quad (8.184)$$

Izenačimo 1. in 2. in dobimo:

$$\begin{aligned} \frac{14r + 4h}{rh} &= \frac{8}{r} \\ 8rh &= r(14r + 4h) \\ 8h &= 14r + 4h \\ 14r &= 4h & \Rightarrow h^* &= \frac{7}{2}r^* \end{aligned} \quad (8.185)$$

Iz enačbe 3. sledi:

$$300 = \pi r^{*2} h^* = \pi r^{*2} \frac{7}{2} r^* = \frac{7\pi}{2} r^{*3} \quad (8.186)$$

Optimalne dimenzije, pri katerih bodo stroški izdelave kontejnerja najmanjši so:

$$\begin{aligned} r^* &= \sqrt[3]{\frac{2 \cdot 300}{7\pi}} = \sqrt[3]{\frac{600}{7\pi}} = 3.01 \text{ dm} \\ h^* &= \frac{7}{2} \cdot r^* = \frac{7}{2} \cdot 3.01 = 10.53 \text{ dm} \end{aligned} \quad (8.187)$$

## 8.2 Primeri enodimenzionalne numerične optimizacije

V nadaljevanju si bomo pogledali nekaj primerov uporabe metod enodimenzionalne numerične optimizacije, pri čemer se bomo omejili na:

- Metodo zlatega reza,
- Newtonovo 1D metodo, ter
- Sekantno metodo.

### 8.2.1 Metoda zlatega reza

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.2.2, izraze (6.9) do (6.27), ter zlasti sliko 39.

#### Primer 1:

S pomočjo metode zlatega reza poiščite ekstrem naslednje funkcije (glej sliko 39):

$$f(x) = 2x^4 - 3x \quad (8.188)$$

kjer sta podana naslednja predpisana parametra optimizacije:

$$a_0 = 0, \quad b_0 = 1 \quad (8.189)$$

in je  $[a_0, b_0]$  začetni interval. Izračunajte prvi dve iteraciji.

Pri reševanju si pomagamo s sliko 39 (Diagram poteka za metodo zlatega reza).

Najprej v skladu s sliko 39 izračunamo koeficient zlatega reza:

$$c_{zr} = \frac{3 - \sqrt{5}}{2} = 0.38197 \quad (8.190)$$

Nato za orientacijo analitično preverimo, kje se nahaja pravi optimum, pri čemer odvod funkcije (8.188) enačimo z 0:



$$\frac{df}{dx} = 8x^3 - 3 = 0 \quad \Rightarrow \quad x^* = \sqrt[3]{\frac{3}{8}} = 0.7211 \quad (8.191)$$

Nato v skladu s sliko 39 opravimo izračune 1. iteracije ( $k = 1$ ). Notranji točki danega začetnega intervala sta:

$$\begin{aligned} y_1 &= a_0 + c_{xr}(b_0 - a_0) = 0 + 0.38197(1 - 0) = 0.38197 \\ z_1 &= b_0 - c_{xr}(b_0 - a_0) = 1 - 0.38197(1 - 0) = 0.61803 \end{aligned} \quad (8.192)$$

Funkcijski vrednosti funkcije (8.188) v teh dveh točkah sta:

$$\begin{aligned} f(y_1) &= 2(y_1)^4 - 3y_1 = 2(0.38197)^4 - 3 \cdot 0.38197 = -1.1031 \\ f(z_1) &= 2(z_1)^4 - 3z_1 = 2(0.61803)^4 - 3 \cdot 0.61803 = -1.56231 \end{aligned} \quad (8.193)$$

V splošnem za novi meji intervala (v 1. iteraciji) velja (glej diagram poteka na sliki 39!):

$$\begin{aligned} a_1 &= \begin{cases} a_0, & f(y_1) \leq f(z_1) \\ y_1, & f(y_1) > f(z_1) \end{cases} \\ b_1 &= \begin{cases} z_1, & f(y_1) \leq f(z_1) \\ b_0, & f(y_1) > f(z_1) \end{cases} \end{aligned} \quad (8.194)$$

Ker je  $f(y_1) > f(z_1)$ , sledi iz (8.194):

$$\begin{aligned} a_1 &= y_1 = 0.38197 \\ b_1 &= b_0 = 1 \\ [a_1, b_1] &= [0.38197, 1] \end{aligned} \quad (8.195)$$

kar je nov, zožan interval 1. iteracije.

Opravimo še izračune 2. iteracije ( $k = 2$ ), pri čemer v diagramu poteka na sliki 39 upoštevamo, da je  $f(y_1) > f(z_1)$ . Tako pridemo do novih notranjih točk intervala, tokrat tistega, ki smo ga izračunali v 1. iteraciji. Ti dve točki ter funkcijski vrednosti funkcije (8.188) v teh dveh točkah se glasita:

$$\begin{aligned} z_2 &= b_1 - c_{xr}(b_1 - a_1) = 1 - 0.38197(1 - 0.38197) = 0.7639 \\ f(z_2) &= 2(z_2)^4 - 3z_2 = 2(0.7639)^4 - 3 \cdot 0.7639 = -1.6076 \end{aligned} \quad (8.196)$$

in:

$$y_2 = z_1 = 0.61803, \quad f(y_2) = f(z_1) = -1.5623 \quad (8.197)$$

V splošnem za novi meji intervala (v 2. iteraciji, kjer je  $k = 2$ ) velja (glej diagram poteka na sliki 39!):

$$a_k = \begin{cases} a_{k-1}, & f(y_k) \leq f(z_k) \\ y_k, & f(y_k) > f(z_k) \end{cases} \quad (8.198)$$

$$b_k = \begin{cases} z_k, & f(y_k) \leq f(z_k) \\ b_{k-1}, & f(y_k) > f(z_k) \end{cases}$$

Ker je  $f(y_2) > f(z_2)$ , sledi iz (8.198) pri upoštevanju  $k = 2$ :

$$\begin{aligned} a_2 &= y_2 = 0.61803 \\ b_2 &= b_1 = 1 \\ [a_2, b_2] &= [0.61803, 1] \end{aligned} \quad (8.199)$$

kar je rešitev za interval po 2. iteracijah (nov, zožan interval 2. iteracije). Seveda se pravi, analitično izračunani optimum 0.7211 (glej izraz (8.191)) nahaja znotraj tega intervala.

### Primer 2:

S pomočjo metode zlatega reza določite optimalno točko za funkcijo (glej sliko 111):

$$f(x) = 2 - 4x + e^x \quad (8.200)$$

kjer je začetni interval  $[a_0, b_0] = [0.5, 2.61]$ . Izračunajte prvi dve iteraciji. Pri reševanju si pomagamo s sliko 39 (Diagram poteka za metodo zlatega reza).

Najprej v skladu s sliko 39 izračunamo koeficient zlatega reza:

$$c_{zr} = \frac{3 - \sqrt{5}}{2} = 0.38197 \quad (8.201)$$

Nato za orientacijo analitično preverimo, kje se nahaja pravi optimum, pri čemer odvod funkcije (8.200) enačimo z 0:

$$\begin{aligned}\frac{df}{dx} &= -4 + e^x = 0 \\ 4 &= e^x \\ \ln 4 &= \ln e^x = x \ln e = x \\ \Rightarrow x^* &= \ln 4 = 1.38629\end{aligned}\tag{8.202}$$

Nato v skladu s sliko 39 opravimo izračune 1. iteracije ( $k = 1$ ). Notranji točki danega začetnega intervala sta:

$$\begin{aligned}y_1 &= a_0 + c_{zr}(b_0 - a_0) = 0.5 + 0.38197(2.61 - 0.5) = 1.3059 \\ z_1 &= b_0 - c_{zr}(b_0 - a_0) = 2.61 - 0.38197(2.61 - 0.5) = 1.80404\end{aligned}\tag{8.203}$$

Funkcijski vrednosti funkcije (8.200) v teh dveh točkah sta:

$$\begin{aligned}f(y_1) &= 2 - 4 \cdot 1.3059 + e^{1.3059} = 0.4674 \\ f(z_1) &= 2 - 4 \cdot 1.80404 + e^{1.80404} = 0.8579\end{aligned}\tag{8.204}$$

V splošnem za novi meji intervala (v 1. iteraciji) velja (glej diagram poteka na sliki 39!):

$$\begin{aligned}a_1 &= \begin{cases} a_0, & f(y_1) \leq f(z_1) \\ y_1, & f(y_1) > f(z_1) \end{cases} \\ b_1 &= \begin{cases} z_1, & f(y_1) \leq f(z_1) \\ b_0, & f(y_1) > f(z_1) \end{cases}\end{aligned}\tag{8.205}$$

Ker je  $f(y_1) < f(z_1)$ , sledi iz (8.205):

$$\begin{aligned}a_1 &= a_0 = 0.5 \\ b_1 &= z_1 = 1.80404 \\ [a_1, b_1] &= [0.5, 1.80404]\end{aligned}\tag{8.206}$$

kar je nov, zožan interval 1. iteracije.

Opravimo še izračune 2. iteracije ( $k = 2$ ), pri čemer v diagramu poteka na sliki 39 upoštevamo, da je  $f(y_1) < f(z_1)$ . Tako pridemo do novih notranjih točk intervala, tokrat tistega, ki smo ga izračunali v 1. iteraciji. Ti dve točki ter funkcijski vrednosti funkcije (8.200) v teh dveh točkah se glasita:

$$\begin{aligned}
 y_2 &= a_1 + c_{zr}(b_1 - a_1) = 0.5 + 0.38197(1.80404 - 0.5) = 0.9981 \\
 f(y_2) &= 2 - 4.0.9981 + e^{0.9981} = 0.7207
 \end{aligned}
 \tag{8.207}$$

in:

$$\begin{aligned}
 z_2 &= y_1 = 1.3059 \\
 f(z_2) &= f(y_1) = 0.4674
 \end{aligned}
 \tag{8.208}$$

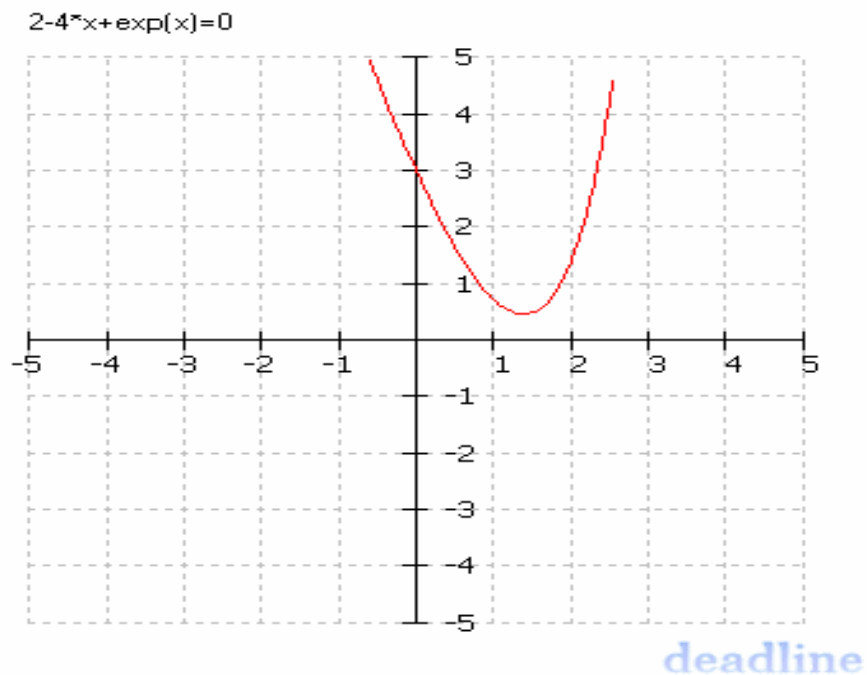
V splošnem za novi meji intervala (v 2. iteraciji, kjer je  $k = 2$ ) velja (glej diagram poteka na sliki 39!):

$$\begin{aligned}
 a_k &= \begin{cases} a_{k-1}, & f(y_k) \leq f(z_k) \\ y_k, & f(y_k) > f(z_k) \end{cases} \\
 b_k &= \begin{cases} z_k, & f(y_k) \leq f(z_k) \\ b_{k-1}, & f(y_k) > f(z_k) \end{cases}
 \end{aligned}
 \tag{8.209}$$

Ker je  $f(y_2) > f(z_2)$ , sledi iz (8.209) pri upoštevanju  $k = 2$ :

$$\begin{aligned}
 a_2 &= y_2 = 0.9981 \\
 b_2 &= b_1 = 1.80404 \\
 [a_2, b_2] &= [0.9981, 1.80404]
 \end{aligned}
 \tag{8.210}$$

kar je rešitev za interval po 2. iteracijah (nov, zožan interval 2. iteracije). Seveda se pravi, analitično izračunani optimum 1.38629 (glej izraz (8.202)) nahaja znotraj tega intervala.



Slika 111: Funkcija  $f(x) = 2 - 4x + e^x$ ,  $(x^*, f(x^*)) = (1.38629, 0.454822)$

### Primer 3:

S pomočjo metode zlatega reza določite optimalno točko za funkcijo (glej sliko 112):

$$f(x) = -x(5\pi - x) \quad (8.211)$$

kjer je začetni interval  $[a_0, b_0] = [0, 20]$ . Izračunajte prvi dve iteraciji. Pri reševanju si pomagamo s sliko 39 (Diagram poteka za metodo zlatega reza).

Najprej v skladu s sliko 39 izračunamo koeficient zlatega reza:

$$c_{zr} = \frac{3 - \sqrt{5}}{2} = 0.38197 \quad (8.212)$$

Nato za orientacijo analitično preverimo, kje se nahaja pravi optimum, pri čemer odvod funkcije (8.211) enačimo z 0:

$$\frac{df}{dx} = -5\pi + 2x = 0 \quad \Rightarrow \quad x^* = \frac{5\pi}{2} = 7.85 \quad (8.213)$$

Nato v skladu s sliko 39 opravimo izračune 1. iteracije ( $k = 1$ ). Notranji točki danega začetnega intervala sta:

$$\begin{aligned} y_1 &= a_0 + c_{zr}(b_0 - a_0) = 0 + 0.38197(20 - 0) = 7.6394 \\ z_1 &= b_0 - c_{zr}(b_0 - a_0) = 20 - 0.38197(20 - 0) = 12.3606 \end{aligned} \quad (8.214)$$

Funkcijski vrednosti funkcije (8.211) v teh dveh točkah sta:

$$\begin{aligned} f(y_1) &= -7.6394(5\pi - 7.6394) = -61.6389 \\ f(z_1) &= -12.3606(5\pi - 12.3606) = -41.3754 \end{aligned} \quad (8.215)$$

V splošnem za novi meji intervala (v 1. iteraciji) velja (glej diagram poteka na sliki 39!):

$$\begin{aligned} a_1 &= \begin{cases} a_0, & f(y_1) \leq f(z_1) \\ y_1, & f(y_1) > f(z_1) \end{cases} \\ b_1 &= \begin{cases} z_1, & f(y_1) \leq f(z_1) \\ b_0, & f(y_1) > f(z_1) \end{cases} \end{aligned} \quad (8.216)$$

Ker je  $f(y_1) < f(z_1)$ , sledi iz (8.216):

$$\begin{aligned} a_1 &= a_0 = 0 \\ b_1 &= z_1 = 12.36 \\ [a_1, b_1] &= [0, 12.36] \end{aligned} \quad (8.217)$$

kar je nov, zožan interval 1. iteracije.

Opravimo še izračune 2. iteracije ( $k = 2$ ), pri čemer v diagramu poteka na sliki 39 upoštevamo, da je  $f(y_1) < f(z_1)$ . Tako pridemo do novih notranjih točk intervala, tokrat tistega, ki smo ga izračunali v 1. iteraciji. Ti dve točki ter funkcijski vrednosti funkcije (8.211) v teh dveh točkah se glasita:

$$\begin{aligned} y_2 &= a_1 + c_{zr}(b_1 - a_1) = 0 + 0.38197(12.36 - 0) = 4.7211 \\ f(y_2) &= -4.7211(5\pi - 4.7211) = -51.87 \end{aligned} \quad (8.218)$$

in:

$$\begin{aligned} z_2 &= y_1 = 7.6394 \\ f(z_2) &= f(y_1) = -61.6389 \end{aligned} \quad (8.219)$$

V splošnem za novi meji intervala (v 2. iteraciji, kjer je  $k = 2$ ) velja (glej diagram poteka na sliki 39!):

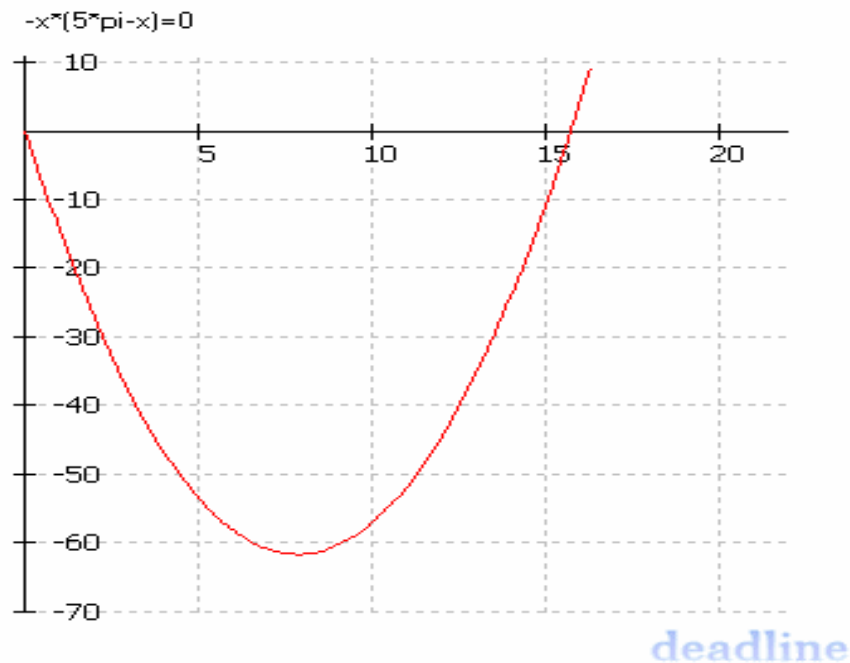
$$a_k = \begin{cases} a_{k-1}, & f(y_k) \leq f(z_k) \\ y_k, & f(y_k) > f(z_k) \end{cases} \quad (8.220)$$

$$b_k = \begin{cases} z_k, & f(y_k) \leq f(z_k) \\ b_{k-1}, & f(y_k) > f(z_k) \end{cases}$$

Ker je  $f(y_2) > f(z_2)$ , sledi iz (8.220) pri upoštevanju  $k = 2$ :

$$\begin{aligned} a_2 &= y_2 = 4.7211 \\ b_2 &= b_1 = 12.36 \\ [a_2, b_2] &= [4.7211, 12.36] \end{aligned} \quad (8.221)$$

kar je rešitev za interval po 2. iteracijah (nov, zožan interval 2. iteracije). Seveda se pravi, analitično izračunani optimum 7.85 (glej izraz (8.213)) nahaja znotraj tega intervala.



Slika 112: Funkcija  $f(x) = -x(5\pi - x)$ ,  $(x^*, f(x^*)) = (7.85398, -61.685027)$

### 8.2.2 Newtonova metoda za eno spremenljivko

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.2.3 in izraze (6.28) do (6.37).

#### Primer 1:

S pomočjo Newtonove metode poiščite maksimum  $(x^*, f(x^*))$  naslednje funkcije:

$$f(x) = 720 - \frac{12}{x} - 108x \quad (8.222)$$

Izračunajte prve tri iteracije ( $k + 1 = 1, 2, 3$ ), pri čemer vzemite začetno vrednost  $x(0) = x_0 = 0.25$ . Dobljeni numerični rezultat za ekstrem  $x_3 = x^*$  primerjajte z analitično izračunanim rezultatom za ekstrem  $(x_A^*, f(x_A^*))$ . A tri iteracije zadoščajo za dovolj dober rezultat, če želimo absolutno natančnost numeričnega rezultata  $|x^* - x_A^*| \leq 10^{-2}$ ?

Najprej izračunajmo vrednosti 1. in 2. odvoda funkcije (8.222), saj ju bomo potrebovali v postopku:

$$\begin{aligned} f'(x) &= \frac{12}{x^2} - 108 \\ f''(x) &= -\frac{24}{x^3} \end{aligned} \quad (8.223)$$

Iteracijska enačba za Newtonovo metodo je:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, \quad k = 0, 1, 2, \dots \quad (8.224)$$

ki z upoštevanjem odvodov (8.223) preide v obliko:



$$\begin{aligned}
x_{k+1} &= x_k - \frac{\frac{12}{x_k^2} - 108}{-\frac{24}{x_k^3}}, \quad k = 0, 1, 2, \dots \\
x_{k+1} &= x_k - \frac{12 - 108x_k^2}{-24x_k}, \quad k = 0, 1, 2, \dots \\
x_{k+1} &= x_k + \frac{12x_k - 108x_k^3}{24}, \quad k = 0, 1, 2, \dots \\
x_{k+1} &= \frac{36x_k - 108x_k^3}{24}, \quad k = 0, 1, 2, \dots \\
x_{k+1} &= \frac{3x_k - 9x_k^3}{2}, \quad k = 0, 1, 2, \dots \\
x_{k+1} &= \frac{3x_k(1 - 3x_k^2)}{2}, \quad k = 0, 1, 2, \dots
\end{aligned} \tag{8.225}$$

Za 1. iteracijo velja z upoštevanjem (8.225) pri  $k + 1 = 1$ :

$$x_1 = \frac{3x_0(1 - 3x_0^2)}{2} = \frac{3 \cdot 0.25(1 - 3 \cdot 0.25^2)}{2} = 0.3046 \tag{8.226}$$

Za 2. iteracijo velja z upoštevanjem (8.225) pri  $k + 1 = 2$  in z upoštevanjem (8.226):

$$x_2 = \frac{3x_1(1 - 3x_1^2)}{2} = \frac{3 \cdot 0.3046(1 - 3 \cdot 0.3046^2)}{2} = 0.3297 \tag{8.227}$$

Za 3. iteracijo velja z upoštevanjem (8.225) pri  $k + 1 = 3$  in z upoštevanjem (8.227):

$$x^* = x_3 = \frac{3x_2(1 - 3x_2^2)}{2} = \frac{3 \cdot 0.3297(1 - 3 \cdot 0.3297^2)}{2} = 0.3332 \tag{8.228}$$

Analitično določimo ekstrem tako, da prvi odvod funkcije (8.222) enačimo z 0. Slednjega smo izračunali že v izrazu (8.223) in potemtakem dobimo:

$$\begin{aligned}
f'(x_A^*) &= \frac{12}{x_A^{*2}} - 108 = 0 \\
12 - 108x_A^{*2} &= 0 \\
1 - 9x_A^{*2} &= 0 \\
x_A^* &= \sqrt{\frac{1}{9}} = \frac{1}{3}
\end{aligned} \tag{8.229}$$

Izračunajmo absolutno natančnost numeričnega rezultata, to je absolutno vrednost razlike med analitično izračunanim ekstremom v (8.229) in numerično izračunanim ekstremom v 3. iteraciji (8.228):

$$|x^* - x_A^*| = \left| 0.33332 - \frac{1}{3} \right| = \left| 1.3 \cdot 10^{-4} \right| \leq 10^{-2} \quad (8.230)$$

Torej očitno prve tri iteracije ( $k + 1 = 1, 2, 3$ ) zadoščajo za dovolj dober rezultat, saj je razlika  $|x^* - x_A^*|$  znotraj predpisane vrednosti  $10^{-2}$ .

### Primer 2:

S pomočjo Newtonove metode poiščite optimum  $(x^*, f(x^*))$  naslednje funkcije (glej sliko 113):

$$f(x) = \frac{(x^2 - 4)^2}{8} - 1; \quad 1 < x \leq 3 \quad (8.231)$$

Izračunajte prve tri iteracije ( $k + 1 = 1, 2, 3$ ), pri čemer vzemite začetno vrednost  $x(0) = x_0 = 1.5$ . Dobljeni numerični rezultat za ekstrem  $x_3 = x^*$  primerjajte z analitično izračunanim rezultatom za ekstrem  $(x_A^*, f(x_A^*))$ . A tri iteracije zadoščajo za dovolj dober rezultat, če želimo absolutno natančnost numeričnega rezultata  $|x^* - x_A^*| \leq 10^{-2}$ ?

Najprej izračunajmo vrednosti 1. in 2. odvoda funkcije (8.231), saj ju bomo potrebovali v postopku:

$$f'(x) = \frac{2(x^2 - 4)2x}{8} = \frac{(x^2 - 4)x}{2} = \frac{(x - 2)(x + 2)x}{2} \quad (8.232)$$

$$f''(x) = \frac{3x^2 - 4}{2}$$

Iteracijska enačba za Newtonovo metodo je:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, \quad k = 0, 1, 2, \dots \quad (8.233)$$

ki z upoštevanjem odvodov (8.232) preide v obliko:

$$\begin{aligned} x_{k+1} &= x_k - \frac{(x_k^2 - 4)x_k}{\frac{3x_k^2 - 4}{2}}, \quad k = 0, 1, 2, \dots \\ x_{k+1} &= x_k - \frac{(x_k^2 - 4)x_k}{3x_k^2 - 4}, \quad k = 0, 1, 2, \dots \\ x_{k+1} &= \frac{(3x_k^2 - 4)x_k - (x_k^2 - 4)x_k}{3x_k^2 - 4}, \quad k = 0, 1, 2, \dots \\ x_{k+1} &= \frac{3x_k^3 - 4x_k - x_k^3 + 4x_k}{3x_k^2 - 4}, \quad k = 0, 1, 2, \dots \\ x_{k+1} &= \frac{2x_k^3}{3x_k^2 - 4}, \quad k = 0, 1, 2, \dots \end{aligned} \quad (8.234)$$

Za 1. iteracijo velja z upoštevanjem (8.234) pri  $k + 1 = 1$ :

$$x_1 = \frac{2x_0^3}{3x_0^2 - 4} = \frac{2 \cdot 1.5^3}{3 \cdot 1.5^2 - 4} = 2.4545 \quad (8.235)$$

Za 2. iteracijo velja z upoštevanjem (8.234) pri  $k + 1 = 2$  in z upoštevanjem (8.235):

$$x_2 = \frac{2x_1^3}{3x_1^2 - 4} = \frac{2 \cdot 2.4545^3}{3 \cdot 2.4545^2 - 4} = 2.1014 \quad (8.236)$$

Za 3. iteracijo velja z upoštevanjem (8.234) pri  $k + 1 = 3$  in z upoštevanjem (8.236):

$$x^* = x_3 = \frac{2x_2^3}{3x_2^2 - 4} = \frac{2 \cdot 2.1014^3}{3 \cdot 2.1014^2 - 4} = 2.006 \quad (8.237)$$

Analitično določimo ekstrem tako, da prvi odvod funkcije (8.231) enačimo z 0. Slednjega smo izračunali že v izrazu (8.232) in potemtakem dobimo:

$$f'(x_A^*) = \frac{2(x_A^{*2} - 4)2x_A^*}{8} = \frac{(x_A^{*2} - 4)x_A^*}{2} = \frac{(x_A^* - 2)(x_A^* + 2)x_A^*}{2} = 0 \quad (8.238)$$

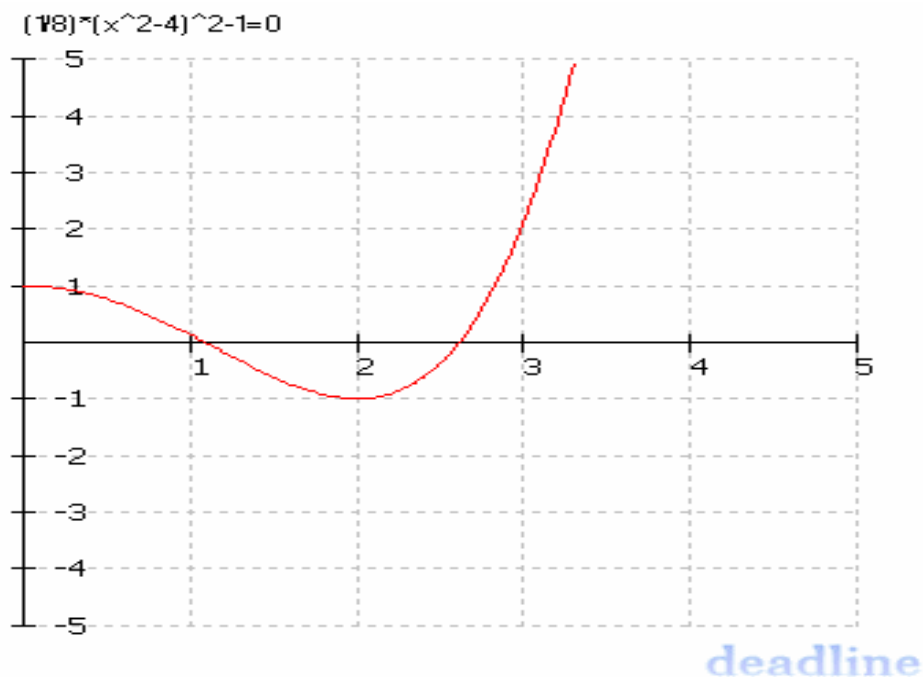
Ker dana funkcija (8.231) lahko zavzema vrednost le za  $1 < x \leq 3$ , je analitična rešitev za ekstrem naslednja:

$$x_A^* = 2 \quad (8.239)$$

Izračunajmo absolutno natančnost numeričnega rezultata, to je absolutno vrednost razlike med analitično izračunanim ekstremom v (8.239) in numerično izračunanim ekstremom v 3. iteraciji (8.237):

$$|x^* - x_A^*| = |2.006 - 2| = 0.006 = 6 \cdot 10^{-3} = 0.6 \cdot 10^{-2} \leq 10^{-2} \quad (8.240)$$

Torej očitno prve tri iteracije ( $k + 1 = 1, 2, 3$ ) zadoščajo za dovolj dober rezultat, saj je razlika  $|x^* - x_A^*|$  znotraj predpisane vrednosti  $10^{-2}$ .



Slika 113: Funkcija  $f(x) = \frac{(x^2 - 4)^2}{8} - 1$ ;  $1 < x \leq 3$ ,  $(x^*, f(x^*)) = (2, -1)$

### Primer 3:

S pomočjo Newtonove metode poiščite optimum  $(x^*, f(x^*))$  naslednje funkcije (glej sliko 114):

$$f(x) = x^2 - \sin x \quad (8.241)$$

Izračunajte prvi dve iteraciji ( $k + 1 = 1, 2$ ), pri čemer vzemite začetno vrednost  $x(0) = x_0 = 0$ .

Najprej izračunajmo vrednosti 1. in 2. odvoda funkcije (8.241), saj ju bomo potrebovali v postopku:

$$\begin{aligned} f'(x) &= 2x - \cos x \\ f''(x) &= 2 + \sin x \end{aligned} \quad (8.242)$$

Iteracijska enačba za Newtonovo metodo je:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, \quad k = 0, 1, 2, \dots \quad (8.243)$$

ki z upoštevanjem odvodov (8.242) preide v obliko:

$$\begin{aligned} x_{k+1} &= x_k - \frac{2x_k - \cos x_k}{2 + \sin x_k}, \quad k = 0, 1, 2, \dots \\ x_{k+1} &= \frac{(2 + \sin x_k)x_k - 2x_k + \cos x_k}{2 + \sin x_k}, \quad k = 0, 1, 2, \dots \\ x_{k+1} &= \frac{x_k \sin x_k + \cos x_k}{2 + \sin x_k}, \quad k = 0, 1, 2, \dots \end{aligned} \quad (8.244)$$

Za 1. iteracijo velja z upoštevanjem (8.244) pri  $k + 1 = 1$ :

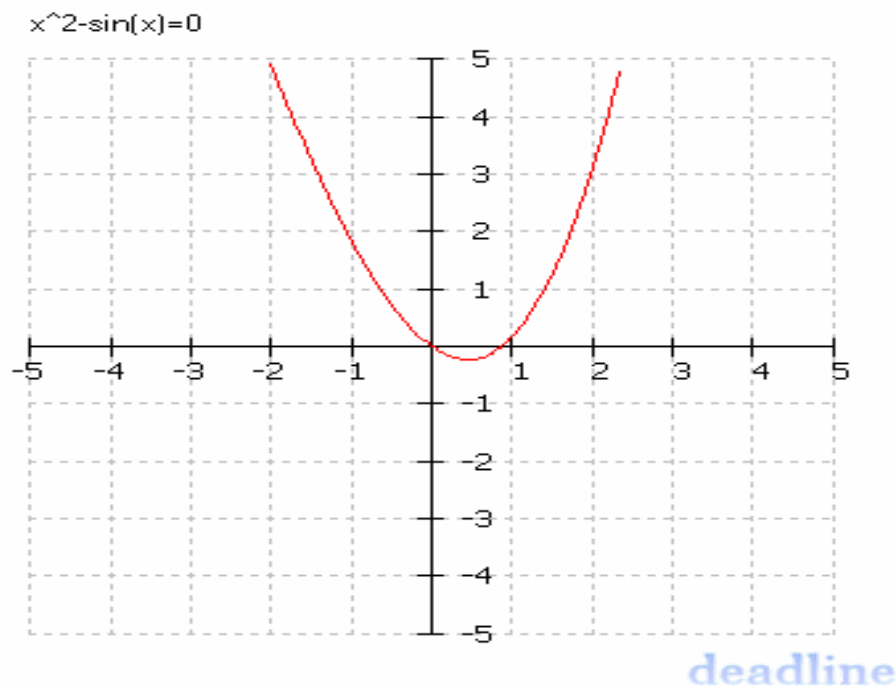
$$\begin{aligned} x_1 &= \frac{x_0 \sin x_0 + \cos x_0}{2 + \sin x_0} \\ x_1 &= \frac{0 \cdot \sin 0 + \cos 0}{2 + \sin 0} = \frac{1}{2} \end{aligned} \quad (8.245)$$

Za 2. iteracijo velja z upoštevanjem (8.244) pri  $k + 1 = 2$  in z upoštevanjem (8.245):

$$\begin{aligned}
 x_2 &= \frac{x_1 \sin x_1 + \cos x_1}{2 + \sin x_1} \\
 x_2 &= \frac{\frac{1}{2} \cdot \sin\left(\frac{1}{2} \text{ rad}\right) + \cos\left(\frac{1}{2} \text{ rad}\right)}{2 + \sin\left(\frac{1}{2} \text{ rad}\right)} = \frac{\frac{1}{2} \cdot \sin\left(\frac{1}{2} \cdot \frac{180^\circ}{\pi}\right) + \cos\left(\frac{1}{2} \cdot \frac{180^\circ}{\pi}\right)}{2 + \sin\left(\frac{1}{2} \cdot \frac{180^\circ}{\pi}\right)} = \\
 &= \frac{\frac{1}{2} \cdot \sin\left(\frac{90^\circ}{\pi}\right) + \cos\left(\frac{90^\circ}{\pi}\right)}{2 + \sin\left(\frac{90^\circ}{\pi}\right)} = \frac{\frac{1}{2} \cdot 0.4794 + 0.8775}{2 + 0.4794} = \frac{0.2397 + 0.8775}{2.4794} = 0.45059
 \end{aligned}
 \tag{8.246}$$

pri čemer smo v argumentih obeh trigonometričnih funkcij še množili s faktorjem  $\frac{180^\circ}{\pi}$ , da smo lahko izvedli pretvorbo iz radianov v stopinje.

Numerični rezultat po dveh iteracijah torej je:  $x^* = x_2 = 0.45059$ . Če ga primerjamo s pravim optimumom 0.45018 na sliki 114, lahko vidimo, da smo že po dveh iteracijah dobili zelo dober rezultat.



Slika 114: Funkcija  $f(x) = x^2 - \sin x$ ,  $(x^*, f(x^*)) = (0.45018, -0.23246)$

### 8.2.3 Sekantna metoda

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.2.4 in izraze (6.38) do (6.45).

#### Primer:

S pomočjo sekantne metode poiščite optimum  $(x^*, f(x^*))$  naslednje funkcije (glej sliko 114):

$$f(x) = x^2 - \sin x \quad (8.247)$$

Izračunajte prvi dve iteraciji ( $k + 1 = 2, 3$ ), pri čemer vzemite začetni vrednosti  $x(0) = x_0 = 0$ ,  $x(1) = x_1 = 1$ .

Najprej izračunajmo vrednost 1. odvoda funkcije (8.247), saj jo bomo potrebovali v postopku:

$$f'(x) = 2x - \cos x \quad (8.248)$$

Iteracijska enačba za sekantno metodo je:

$$x_{k+1} = x_k - \frac{f'(x_k)(x_k - x_{k-1})}{f'(x_k) - f'(x_{k-1})}, \quad k = 1, 2, \dots \quad (8.249)$$

ki z upoštevanjem odvoda (8.248) preide v obliko:

$$\begin{aligned} x_{k+1} &= x_k - \frac{(2x_k - \cos x_k)(x_k - x_{k-1})}{(2x_k - \cos x_k) - (2x_{k-1} - \cos x_{k-1})}, \quad k = 1, 2, \dots \\ x_{k+1} &= x_k - \frac{(2x_k - \cos x_k)(x_k - x_{k-1})}{2(x_k - x_{k-1}) - (\cos x_k - \cos x_{k-1})}, \quad k = 1, 2, \dots \\ x_{k+1} &= x_k - \frac{2x_k - \cos x_k}{2 - \frac{\cos x_k - \cos x_{k-1}}{x_k - x_{k-1}}}, \quad k = 1, 2, \dots \end{aligned} \quad (8.250)$$

Za 1. iteracijo velja z upoštevanjem (8.250) pri  $k + 1 = 2$ :

$$\begin{aligned}
x_2 &= x_1 - \frac{2x_1 - \cos x_1}{2 - \frac{\cos x_1 - \cos x_0}{x_1 - x_0}} \\
x_2 &= 1 - \frac{2.1 - \cos(1rad)}{2 - \frac{\cos(1rad) - \cos 0}{1 - 0}} \\
x_2 &= 1 - \frac{2.1 - \cos\left(1 \cdot \frac{180^0}{\pi}\right)}{2 - \frac{\cos\left(1 \cdot \frac{180^0}{\pi}\right) - \cos 0}{1 - 0}} = 1 - \frac{2 - 0.5403}{2 - \frac{0.5403 - 1}{1}} = \\
&= 1 - \frac{1.4597}{2 + 0.4597} = 1 - \frac{1.4597}{2.4597} = 0.406
\end{aligned} \tag{8.251}$$

Za 2. iteracijo velja z upoštevanjem (8.250) pri  $k + 1 = 3$  in z upoštevanjem (8.251):

$$\begin{aligned}
x_3 &= x_2 - \frac{2x_2 - \cos x_2}{2 - \frac{\cos x_2 - \cos x_1}{x_2 - x_1}} \\
x_3 &= 0.406 - \frac{2 \cdot 0.406 - \cos(0.406rad)}{2 - \frac{\cos(0.406rad) - \cos(1rad)}{0.406 - 1}} \\
x_3 &= 0.406 - \frac{2 \cdot 0.406 - \cos\left(0.406 \frac{180^0}{\pi}\right)}{2 - \frac{\cos\left(0.406 \frac{180^0}{\pi}\right) - \cos\left(1 \frac{180^0}{\pi}\right)}{0.406 - 1}} \\
x_3 &= 0.406 - \frac{0.812 - 0.9187}{2 - \frac{0.9187 - 0.5403}{-0.594}} = 0.406 - \frac{-0.1067}{2 - \frac{0.3784}{-0.594}} = \\
&= 0.406 + \frac{0.1067}{2 + 0.637} = 0.446
\end{aligned} \tag{8.252}$$

V izrazih (8.251) in (8.252) smo v argumentu trigonometrične funkcije cosinus še množili s faktorjem  $\frac{180^0}{\pi}$ , da smo lahko izvedli pretvorbo iz radianov v stopinje.



Numerični rezultat po dveh iteracijah torej je:  $x^* = x_3 = 0.446$ . Če ga primerjamo s pravim optimumom 0.45018 na sliki 114, lahko vidimo, da smo že po dveh iteracijah dobili dokaj dober rezultat, vendar za malenkost slabši kot v primeru z Newtonovo metodo. Razlog za nekoliko hitrejšo konvergenco Newtonove metode glede na sekantno metodo je v tem, da prva metoda vsebuje tudi 2. odvod, ki podaja dodatno informacijo za čim hitrejšo napredovanje k optimumu.

### 8.3 Primeri multidimenzionalne numerične optimizacije z osnovnimi metodami

V nadaljevanju si bomo pogledali nekaj primerov uporabe metod multidimenzionalne numerične optimizacije, pri čemer se bomo omejili na:

- Direktno Hook-Jeevesovo metodo vzorčnih premikov,
- Cauchyovo metodo najstrmejšega spusta,
- Newtonovo multidimenzionalno metodo, ter
- Marquardtovo metodo.

#### 8.3.1 Direktna Hook-Jeevesova metoda vzorčnih premikov

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.3.1.1, izraze (6.46) do (6.55), ter zlasti slike 46 do 49.

#### Primer 1:

S pomočjo Hook-Jeevesove metode vzorčnih premikov določite optimalno točko za funkcijo treh odločitvenih spremenljivk  $x_1, x_2, x_3$ :

$$f(x_1, x_2, x_3) = x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3 \quad (8.253)$$

kjer je začetni vektor odločitvenih spremenljivk enak:  $x(0) = x_0 = [x_1(0) \ x_2(0) \ x_3(0)] = [0 \ 0 \ 0]^T$  in je korak vzdolž dimenzij enak:  $\delta = 0.1$ .

Izračunajte prvo iteracijo aproksimacijskega vektorja  $x_p(1) = [x_{1p}(1) \ x_{2p}(1) \ x_{3p}(1)]$ , kjer številka »1« pomeni indeks prve iteracije.

Pri reševanju si pomagamo s sliko 46 (diagram poteka za Hook-Jeevesovo metodo) in s sliko 47 (Diagram poteka za postopek premikanja proti najbolj ugodni smeri). V slednjem je natančneje prikazana struktura bloka »iskanje uspešnih premikov po vseh dimenzijah«, ki jo kliče algoritem iz slike 46. Za boljšo orientacijo poimenujmo ta blok z  $B_1$ .

Najprej v skladu z algoritmom na sliki 46 nastavimo vrednosti začetnega aproksimacijskega vektorja  $x_p(0)$  in začetnega baznega vektorja  $x_b(0)$  na vrednost  $x_0$ :

$$x_p(0) = x_b(0) = x_0 = [x_1(0) \ x_2(0) \ x_3(0)] = [0 \ 0 \ 0]^T.$$

Nato v skladu z blokom  $B_1$  (glej sliki 46 in 47) najprej izvedemo za 1. iteracijo 1. korak, to je iskanje uspešnih premikov po vseh (treh) dimenzijah.

V ta namen vpeljemo vmesno spremenljivko, to je vektor  $x$ , ki ima na začetku vrednost:  $x(0) = x_p(0) = x_0$ , s katerim operiramo, dokler se izvaja algoritem bloka  $B_1$  (glej sliko 47!).

V tem algoritmu najprej izvedemo premik  $\Delta x_1$  po prvi dimenziji  $x_1$ , izhajajoč iz začetnega vektorja  $x(0) = [0 \ 0 \ 0]^T$ :

$$\Delta x_1 = \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} \tag{8.254}$$

Sledi:

$$x(0) + \Delta x_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} \quad (8.255)$$

Funkcija (8.253) ima v vektorju (8.255) vrednost:

$$f(0.1, 0, 0) = 0.1^4 + 0.1^3 - 0.1 + 0^4 - 0^2 + 0 + 0^2 - 0 + 0.1 \cdot 0 \cdot 0 = -0.0989 \quad (8.256)$$

Ker je  $f(0.1, 0, 0) = -0.0989 < f(0, 0, 0) = 0^4 + 0^3 - 0 + 0^4 - 0^2 + 0 + 0^2 - 0 + 0 \cdot 0 \cdot 0 = 0$ , je bil premik uspešen, saj smo zmanjšali vrednost funkcije. Sledi, da nov vektor  $x(1)$  nastavimo na vrednost (8.255) (glej sliko 47!):

$$x(1) = \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} \quad (8.257)$$

Odtod nato izvedemo premik  $\Delta x_2 = [0 \ 0.1 \ 0]^T$  po drugi dimenziji  $x_2$ :

$$x(1) + \Delta x_2 = \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0.1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0.1 \\ 0 \end{bmatrix} \quad (8.258)$$

Funkcija (8.253) ima v vektorju (8.258) vrednost:

$$\begin{aligned} f(0.1, 0.1, 0) &= 0.1^4 + 0.1^3 - 0.1 + 0.1^4 - 0.1^2 + 0.1 + 0^2 - 0 + 0.1 \cdot 0.1 \cdot 0 = \\ &= -0.0089 \end{aligned} \quad (8.259)$$

Ker je  $f(0.1, 0.1, 0) = -0.0089 > f(0.1, 0, 0) = -0.0989$ , je bil premik neuspešen, saj smo namesto, da bi zmanjšali, povečali vrednost funkcije. Zato se je potrebno premakniti še v nasprotni smeri  $-\Delta x_2$ :

$$x(1) - \Delta x_2 = \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0.1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.1 \\ -0.1 \\ 0 \end{bmatrix} \quad (8.260)$$

Funkcija (8.253) ima v vektorju (8.260) vrednost:

$$\begin{aligned}
f(0.1, -0.1, 0) &= \\
&= 0.1^4 + 0.1^3 - 0.1 + (-0.1)^4 - (-0.1)^2 + (-0.1) + 0^2 - 0 + 0.1 \cdot (-0.1) \cdot 0 = \quad (8.261) \\
&= -0.2088
\end{aligned}$$

Ker je  $f(0.1, -0.1, 0) = -0.2088 < f(0.1, 0, 0) = -0.0989$ , je bil premik uspešen, saj smo zmanjšali vrednost funkcije. Sledi, da nov vektor  $x(1)$  nastavimo na vrednost (8.260) (glej sliko 47!):

$$x(1) = \begin{bmatrix} 0.1 \\ -0.1 \\ 0 \end{bmatrix} \quad (8.262)$$

Odtod nato izvedemo premik  $\Delta x_3 = [0 \ 0 \ 0.1]^T$  po tretji dimenziji  $x_3$ :

$$x(1) + \Delta x_3 = \begin{bmatrix} 0.1 \\ -0.1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.1 \\ -0.1 \\ 0.1 \end{bmatrix} \quad (8.263)$$

Funkcija (8.253) ima v vektorju (8.263) vrednost:

$$\begin{aligned}
f(0.1, -0.1, 0.1) &= \\
&= 0.1^4 + 0.1^3 - 0.1 + (-0.1)^4 - (-0.1)^2 + (-0.1) + 0.1^2 - 0.1 + 0.1 \cdot (-0.1) \cdot 0.1 = \quad (8.264) \\
&= -0.2998
\end{aligned}$$

Ker je  $f(0.1, -0.1, 0.1) = -0.2998 < f(0.1, -0.1, 0) = -0.2088$ , je bil premik uspešen, saj smo zmanjšali vrednost funkcije. Sledi, da nov vektor  $x(1)$  nastavimo na vrednost (8.263) (glej sliko 47!):

$$x(1) = x_b(1) = x_n(1) = \begin{bmatrix} 0.1 \\ -0.1 \\ 0.1 \end{bmatrix} \quad (8.265)$$

pri čemer smo v skladu z algoritmom na sliki 39 to vrednost dodelili tudi vektorju  $x_n$ . S tem se je postopek v bloku  $B_1$ , prikazanem na sliki 47, končal. Zato se vrnemo v glavni algoritem metode na sliki 46. Kot je razvidno iz izraza (8.265), smo prenavajali tudi bazni vektor, ki je zavzel novo vrednost:  $x_b(1) = [0.1 \ -0.1 \ 0.1]^T$ . Slednjega namreč

prenastavimo vedno, ko smo prepričani o izboljšanju rezultata. To je sicer v algoritmu na sliki 46 iz razlogov programiranja ponazorjeno nekoliko kasneje, izza bloka za skok.

In dejansko je do izboljšanja prišlo, saj je vrednost funkcije (8.253) v vektorju  $x_n(1)$  gotovo manjša kot v začetnem baznem vektorju  $x_b(0)$ , velja namreč:  $f(0.1, -0.1, 0.1) = -0.2998 < f(0, 0, 0) = 0$ . To pomeni, da nas je blok  $B_1$  na začetku postopka (tekem 1. koraka 1. iteracije) gotovo premaknil bliže optimumu, saj je zmanjšal vrednost funkcije (8.253). Seveda v kasnejših iteracijah ni več gotovo, da bo temu res tako. Tedaj pa je potrebno prepoloviti premik po vseh dimenzijah.

Ker je bil 1. korak postopka uspešen, gremo sedaj lahko na 2. korak 1. iteracije, to je morebiten dodaten skok v izbrani (najbolj ugodni) smeri (glej sliko 46). V ta namen postavimo aproksimacijski vektor na novo vrednost:

$$x_p(1) = x_n(1) + (x_n(1) - x_b(0)) = 2x_n(1) - x_b(0) = 2 \begin{bmatrix} 0.1 \\ -0.1 \\ 0.1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.2 \\ -0.2 \\ 0.2 \end{bmatrix} \quad (8.266)$$

Ta skok od izboljšanega novega vektorja  $x_b(1) = x_n(1)$  v smeri razlike novega in starega vektorja  $x_n(1) - x_b(0)$  se je izvršil pod predpostavko, da se »dolina« (padanje funkcije) še naprej nadaljuje s spuščanjem (manjšanjem) ciljne funkcije (8.253) v dotični, najbolj ugodni smeri.

Funkcija (8.253) ima v vektorju (8.266) vrednost:

$$f(0.2, -0.2, 0.2) = 0.2^4 + 0.2^3 - 0.2 + (-0.2)^4 - (-0.2)^2 + (-0.2) + 0.2^2 - 0.2 + 0.2 \cdot (-0.2) \cdot 0.2 = -0.5968 \quad (8.267)$$

Ker je  $f(0.2, -0.2, 0.2) = -0.5968 < f(0.1, -0.1, 0.1) = -0.2998$ , je skok zaenkrat uspešen, saj smo ponovno zmanjšali vrednost funkcije.

Zato v nadaljevanju zopet preidemo v algoritem bloka  $B_1$  na sliki 47, pri čemer tudi ponovno vpeljemo vmesno spremenljivko, to je vektor  $x$ , ki ga izenačimo z vektorjem po skoku v izrazu (8.266):

$$x(1) = \begin{bmatrix} 0.2 \\ -0.2 \\ 0.2 \end{bmatrix} \quad (8.268)$$

Seveda bomo tokrat izvedli le iskanje prvega uspešnega premika po katerikoli dimenziji, kar je potrebno zaradi izvedbe končne verifikacije uspešnosti skoka.

V ta namen denimo izvedemo premik  $\Delta x_1$  po prvi dimenziji  $x_1$ , izhajajoč iz vektorja  $x(1) = [0.2 \ -0.2 \ 0.2]^T$ :

$$x(1) + \Delta x_1 = \begin{bmatrix} 0.2 \\ -0.2 \\ 0.2 \end{bmatrix} + \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.3 \\ -0.2 \\ 0.2 \end{bmatrix} \quad (8.269)$$

Funkcija (8.253) ima v vektorju (8.269) vrednost:

$$\begin{aligned} f(0.3, -0.2, 0.2) &= \\ &= 0.3^4 + 0.3^3 - 0.3 + (-0.2)^4 - (-0.2)^2 + (-0.2) + \\ &\quad + 0.2^2 - 0.2 + 0.3 \cdot (-0.2) \cdot 0.2 = \\ &= -0.6813 \end{aligned} \quad (8.270)$$

Ker je  $f(0.3, -0.2, 0.2) = -0.6813 < f(0.2, -0.2, 0.2) = -0.5968$ , je skok dokončno uspešen, saj smo spet zmanjšali vrednost funkcije.

Sledi, da nov vektor  $x(1)$  nastavimo na vrednost (8.269) (glej sliko 47!):

$$x_p(1) = x_b(1) = x_n(1) = x(1) = \begin{bmatrix} 0.3 \\ -0.2 \\ 0.2 \end{bmatrix} \quad (8.271)$$

S tem postopek v bloku  $B_1$ , prikazanem na sliki 47, takoj končamo, saj smo iskali le prvi uspešen premik, ter se ponovno vrnemo v glavni algoritem metode na sliki 46.

Kot je razvidno iz izraza (8.271), smo vrednost  $[0.3 \ -0.2 \ 0.2]^T$  dodelili tudi vektorjem,  $x_n(1)$ ,  $x_b(1)$  in  $x_p(1)$ . Prvemu zato, ker je to izhodna spremenljivka iz algoritma na sliki 47. Drugemu in tretjemu pa zato, ker je to tudi končni rezultat za bazni vektor 1. iteracije, ter hkrati končni rezultat za aproksimacijski vektor 1. iteracije:

$$x_p(1) = [x_{1p}(1) \ x_{2p}(1) \ x_{3p}(1)].$$

Kot je razvidno iz slike 46, nato lahko nadaljujemo z naslednjimi iteracijami, to je 2., 3., ..., N. iteracijo. Postopek končamo, ko smo se dovolj približali optimalni rešitvi oz. so izpolnjeni pogoji za končanje postopka.

Končni rezultat za aproksimacijski vektor odločitvenih spremenljivk se glasi (numerični izračun v programskem jeziku Matlab):

$$x_p(N) = \tilde{x}^* = [\tilde{x}_1^* \ \tilde{x}_2^* \ \tilde{x}_3^*]^T = [0.57031 \ -0.93906 \ 0.76875]^T$$

v katerem zavzame funkcija (8.253) minimalno vrednost:

$$f(\tilde{x}_1^*, \tilde{x}_2^*, \tilde{x}_3^*) = \min f(x_1, x_2, x_3) = -1.9118$$

### Primer 2:

S pomočjo Hook-Jeevesove metode vzorčnih premikov določite optimalno točko za funkcijo dveh odločitvenih spremenljivk  $x_1, x_2$ :

$$f(x_1, x_2) = (x_1 - \sqrt{5})^2 + (x_2 - \pi)^2 + 10 \quad (8.272)$$

kjer je začetni vektor odločitvenih spremenljivk enak:

$$x(0) = x_0 = [x_1(0) \ x_2(0)] = [0 \ 0]^T \text{ in je korak vzdolž dimenzij enak: } \delta = 0.5.$$

Izračunajte prvo iteracijo aproksimacijskega vektorja:  $x_p(1) = [x_{1p}(1) \ x_{2p}(1)]$ .

Podrobnosti reševanja in natančnejše komentarje si lahko bralec pogleda v prejšnjem primeru.

Pri reševanju si pomagamo s slikama 46 (diagram poteka za Hook-Jeevesovo metodo), ter 47 (Diagram poteka za postopek premikanja proti najbolj ugodni smeri). V slednjem je natančneje prikazana struktura bloka »iskanje uspešnih premikov po vseh dimenzijah«, ki jo kliče algoritem iz slike 46. Za boljšo orientacijo poimenujmo ta blok z  $B_1$ .

Najprej v skladu z algoritmom na sliki 46 nastavimo vrednosti začetnega aproksimacijskega vektorja  $x_p(0)$  in začetnega baznega vektorja  $x_b(0)$  na vrednost  $x_0$ :

$$x_p(0) = x_b(0) = x_0 = [x_1(0) \quad x_2(0)] = [0 \quad 0]^T.$$

Nato v skladu z blokom  $B_1$  (glej sliki 46 in 47) najprej izvedemo za 1. iteracijo 1. korak, to je iskanje uspešnih premikov po vseh (dveh) dimenzijah.

V ta namen vpeljemo vmesno spremenljivko, to je vektor  $x$ , ki ima na začetku vrednost:  $x(0) = x_p(0) = x_0$ , s katerim operiramo, dokler se izvaja algoritem bloka  $B_1$  (glej sliko 47!).

V tem algoritmu najprej izvedemo premik  $\Delta x_1$  po prvi dimenziji  $x_1$ , izhajajoč iz začetnega vektorja  $x(0) = [0 \quad 0]^T$ :

$$\Delta x_1 = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \quad (8.273)$$

Sledi:

$$x(0) + \Delta x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \quad (8.274)$$

Funkcija (8.272) ima v vektorju (8.274) vrednost:



$$f(0.5,0) = (0.5 - \sqrt{5})^2 + (0 - \pi)^2 + 10 = 22.8835 \quad (8.275)$$

Ker je  $f(0.5,0) < f(0,0) = (0 - \sqrt{5})^2 + (0 - \pi)^2 + 10 = 24.86$ , je bil premik uspešen, saj smo zmanjšali vrednost funkcije. Sledi, da nov vektor  $x(1)$  nastavimo na vrednost (8.274) (glej sliko 47!):

$$x(1) = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \quad (8.276)$$

Odtod nato izvedemo premik  $\Delta x_2 = [0 \ 0.5]^T$  po drugi dimenziji  $x_2$ :

$$x(1) + \Delta x_2 = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \quad (8.277)$$

Funkcija (8.272) ima v vektorju (8.277) vrednost:

$$f(0.5,0.5) = (0.5 - \sqrt{5})^2 + (0.5 - \pi)^2 + 10 = 19.99 \quad (8.278)$$

Ker je  $f(0.5,0.5) < f(0.5,0)$ , je bil premik uspešen, saj smo zmanjšali vrednost funkcije. Sledi, da nov vektor  $x(1)$  nastavimo na vrednost (8.277) (glej sliko 47!):

$$x(1) = x_b(1) = x_n(1) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \quad (8.279)$$

pri čemer smo v skladu z algoritmom na sliki 47 to vrednost dodelili tudi vektorju  $x_n$ . S tem se je postopek v bloku  $B_1$ , prikazanem na sliki 47, končal. Zato se vrnemo v glavni algoritem metode na sliki 46. Kot je razvidno iz izraza (8.275), smo prenestavili tudi bazni vektor, ki je zavzel novo vrednost:  $x_b(1) = [0.5 \ 0.5]^T$ .

Ker je bil 1. korak postopka uspešen, gremo sedaj lahko na 2. korak 1. iteracije, to je morebiten dodaten skok v izbrani (najbolj ugodni) smeri (glej sliko 46). V ta namen postavimo aproksimacijski vektor na novo vrednost:

$$x_p(1) = x_n(1) + (x_n(1) - x_b(0)) = 2x_n(1) - x_b(0) = 2 \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (8.280)$$

Funkcija (8.272) ima v vektorju (8.280) vrednost:

$$f(1,1) = (1 - \sqrt{5})^2 + (1 - \pi)^2 + 10 = 16.1143 \quad (8.281)$$

Ker je  $f(1,1) < f(0.5,0.5)$ , je skok zaenkrat uspešen, saj smo ponovno zmanjšali vrednost funkcije.

Zato v nadaljevanju zopet preidemo v algoritem bloka  $B_1$  na sliki 47, pri čemer tudi ponovno vpeljemo vmesno spremenljivko, to je vektor  $x$ , ki ga izenačimo z vektorjem po skoku v izrazu (8.280):

$$x(1) = x_p(1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (8.282)$$

Seveda bomo tokrat izvedli le iskanje prvega uspešnega premika po katerikoli dimenziji, kar je potrebno zaradi izvedbe končne verifikacije uspešnosti skoka.

V ta namen denimo izvedemo premik  $\Delta x_1$  po prvi dimenziji  $x_1$ , izhajajoč iz vektorja  $x(1) = [1 \ 1]^T$ :

$$x(1) + \Delta x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} \quad (8.283)$$

Funkcija (8.272) ima v vektorju (8.283) vrednost:

$$f(1.5,1) = (1.5 - \sqrt{5})^2 + (1 - \pi)^2 + 10 = 15.12 \quad (8.284)$$

Ker je  $f(1.5,1) < f(1,1)$ , je skok dokončno uspešen, saj smo spet zmanjšali vrednost funkcije.

Sledi, da nov vektor  $x(1)$  nastavimo na vrednost (8.283) (glej sliko 47!):

$$x_p(1) = x_b(1) = x_n(1) = x(1) = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} \quad (8.285)$$

S tem postopek v bloku  $B_1$ , prikazanem na sliki 47, takoj končamo, saj smo iskali le prvi uspešen premik, ter se ponovno vrnemo v glavni algoritem metode na sliki 46.

Kot je razvidno iz izraza (8.285), smo vrednost  $[1.5 \ 1]^T$  dodelili tudi vektorjem  $x_n(1)$ ,  $x_b(1)$  in  $x_p(1)$ . Prvemu zato, ker je to izhodna spremenljivka iz algoritma na sliki 47. Drugemu in tretjemu pa zato, ker je to tudi končni rezultat za bazni vektor 1. iteracije, ter hkrati končni rezultat za aproksimacijski vektor 1. iteracije:  $x_p(1) = [x_{1p}(1) \ x_{2p}(1)]$ .

Kot je razvidno iz slike 46, nato lahko nadaljujemo z naslednjimi iteracijami, to je 2., 3., ..., N. iteracijo. Postopek končamo, ko smo se dovolj približali optimalni rešitvi oz. so izpolnjeni pogoji za končanje postopka.

Prikaz dinamike konvergence Hook-Jeevesove metode si lahko ogledamo na sliki 115, kjer je narisana tudi potek ciljne funkcije  $f(x_1, x_2)$  iz izraza (8.272) pri ekvidistančnih vrednostih za  $(x_1, x_2)$  (slika 115 spodaj). Iz slike je razvidno, kako v posameznih iteracijah vrednosti aproksimacij optimalnih odločitvenih spremenljivk konsistentno napredujejo od začetne točke  $x_p(0) = [x_{1p}(0) \ x_{2p}(0)] = [0 \ 0]^T$  do končne (optimalne) točke:

$$x_p(N) = [x_{1p}(N) \ x_{2p}(N)] = [2.2363 \ 3.1406]^T.$$

Tovrstno napredovanje (konvergiranje k optimalni rešitvi) je na sliki 115 označeno s povezanimi krogi (o), označene pa so tudi nivojnice ciljne funkcije (8.272) v  $(x_1, x_2)$  ravnini (z resolucijo 1: 19, 18, 17, ...). Te pri konvergiranju postopka postopoma padajo od začetne vrednosti  $f(0,0) = (0 - \sqrt{5})^2 + (0 - \pi)^2 + 10 = 24.86$  do končne (minimalne) vrednosti  $f(2.2363, 3.1406) = (2.2363 - \sqrt{5})^2 + (3.1406 - \pi)^2 + 10 \approx 0 + 0 + 10 = 10$ .

Na sliki 115 spodaj je tudi razvidno, da funkcija  $f(x_1, x_2)$  dejansko doseže najnižjo točko (minimum) v prostoru  $(x_1, x_2, f(x_1, x_2))$ , ko postopek konvergira k optimalni rešitvi  $[2.2363 \ 3.1406]^T$ .

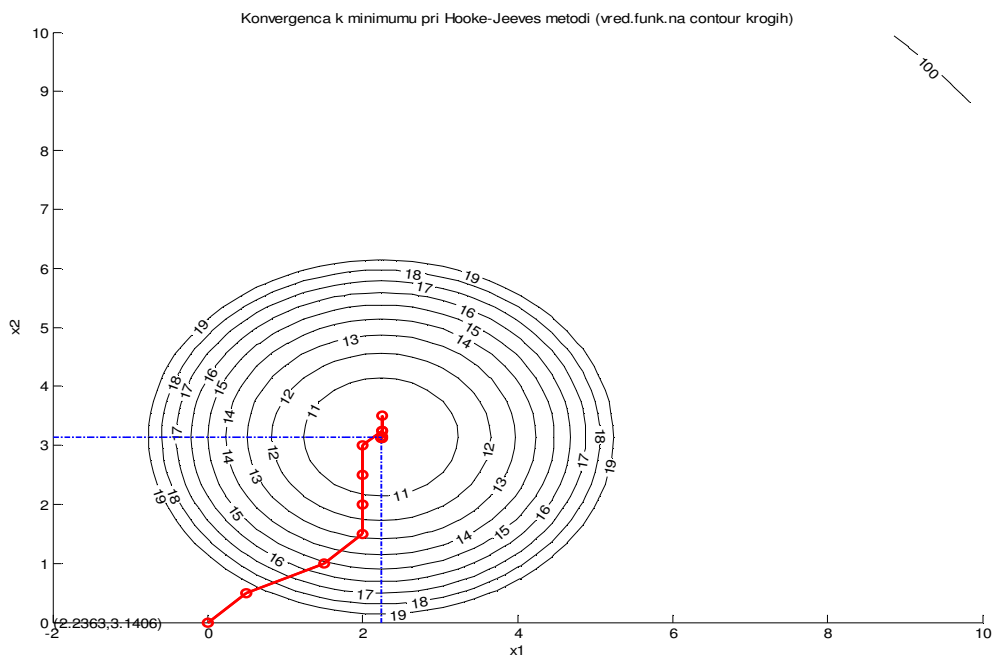
Končni rezultat za aproksimacijski vektor odločitvenih spremenljivk se torej glasi (numerični izračun v programskem jeziku Matlab):

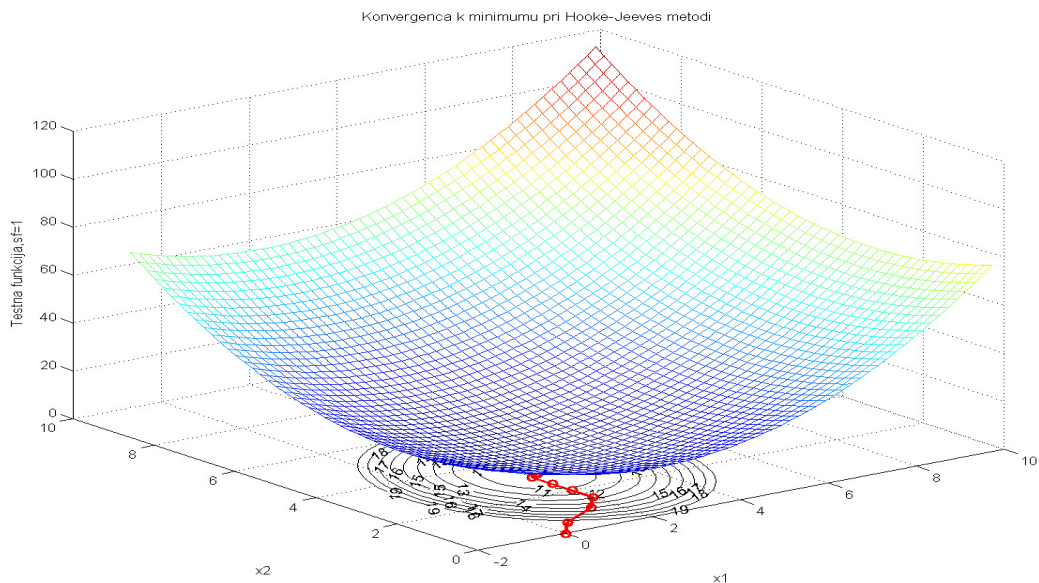
$$x_p(N) = \tilde{x}^* = [\tilde{x}_1^* \ \tilde{x}_2^*]^T = [2.2363 \ 3.1406]^T$$

in je bil izračunan po 16 iteracijah. Tedaj funkcija (8.272) zavzame minimalno vrednost:

$$f(\tilde{x}_1^*, \tilde{x}_2^*) = \min f(x_1, x_2) = 10$$

Ta rezultat za optimum tudi lahko primerjamo z analitično dobljenim rezultatom za ekstrem, prikazanim v naslednji nalogi (izraz 8.297), da vidimo, kako smo se dejansko zelo natančno približali pravi optimalni rešitvi.





Slika 115: Prikaz dinamike konvergence Hook-Jeevesove metode: Zgoraj - Prikaz konvergiranja k optimalni rešitvi (**povezani krogi**) v prostoru odločitvenih spremenljivk  $(x_1, x_2)$ ; Spodaj – Prikaz ciljne funkcije  $f(x_1, x_2)$  iz izraza (8.272) v odvisnosti od  $(x_1, x_2)$ ; **Nivojnice so prikazane z resolucijo 1.**

### 8.3.2 Cauchyeva metoda najstrmejšega spusta (gradientna metoda I. reda)

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.3.1.3, izraze (6.68) do (6.90), ter zlasti sliki 73 in 76.

#### Primer 1:

S pomočjo metode najstrmejšega spusta določite optimalno točko za skalarno funkcijo vektorske odločitvene spremenljivke (brez omejitev):

$$f(\underline{x}) = f([x_1 \quad x_2]^T) = f(x_1, x_2) = (x_1 - \sqrt{5})^2 + (x_2 - \pi)^2 + 10 \quad (8.286)$$

kjer je začetni vektor odločitvenih spremenljivk enak:  $\underline{x}(0) = \underline{x}_0 = [x_1(0) \quad x_2(0)]^T = [0 \quad 0]^T$ . Prečna črta spodaj (zaradi preglednosti!) označuje vektor. Izračunajte prvo iteracijo vektorja odločitvenih spremenljivk  $\underline{x}(1) = \underline{x}_1 = [x_1(1) \quad x_2(1)]^T$ . Uporabite izraz brez norme. Izračunajte tudi analitično

rešitev  $\underline{x}_A$  za vektor ekstremov. A smo se po 1. iteraciji že dovolj približali pravi rešitvi, če mora biti  $\|\underline{x}_A - \underline{x}(1)\| \leq \varepsilon = 10^{-2}$ ? Če je odgovor pritrديلen, pokažite in razložite, zakaj ni potrebno računati parametra  $\alpha_{opt}$  v naslednji, to je 2. iteraciji.

Gradient funkcije (8.286) v poljubni točki tridimenzionalnega prostora  $[x_1, x_2, f(x_1, x_2)]$  je enak:

$$\nabla_{\underline{f}}(\underline{x}) = \nabla_{\underline{f}}([x_1 \quad x_2]^T) = \nabla_{\underline{f}}(x_1, x_2) = \left[ \frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \right]^T = \left[ 2(x_1 - \sqrt{5}) \quad 2(x_2 - \pi) \right]^T \quad (8.287)$$

Gradient funkcije (8.286) v začetni točki  $\underline{x}(0) = \underline{x}_0$  je v skladu z (8.287) enak:

$$\begin{aligned} \nabla_{\underline{f}}(\underline{x}_0) &= \nabla_{\underline{f}}([x_1(0) \quad x_2(0)]^T) = \nabla_{\underline{f}}([0 \quad 0]^T) = \nabla_{\underline{f}}(0, 0) = \\ &= \left[ 2(0 - \sqrt{5}) \quad 2(0 - \pi) \right]^T = \left[ 2(-\sqrt{5}) \quad 2(-\pi) \right]^T = \left[ -2\sqrt{5} \quad -2\pi \right]^T \end{aligned} \quad (8.288)$$

Pri metodi najstrmejšega spusta velja za vektor odločitvenih spremenljivk naslednja relacija v  $k+1$ . iteraciji metode:

$$\underline{x}_{k+1} = \underline{x}_k + \Delta \underline{x}_k = \underline{x}_k + \alpha_k \cdot \underline{d}_k = \underline{x}_k - \alpha_k \cdot \frac{\nabla_{\underline{f}}(\underline{x}_k)}{\|\nabla_{\underline{f}}(\underline{x}_k)\|}, \quad k = 0, 1, 2, \dots \quad (8.289)$$

oziroma, če norme ne upoštevamo:

$$\underline{x}_{k+1} = \underline{x}_k + \Delta \underline{x}_k = \underline{x}_k + \alpha_k \cdot \underline{d}_k = \underline{x}_k - \alpha_k \cdot \nabla_{\underline{f}}(\underline{x}_k), \quad k = 0, 1, 2, \dots \quad (8.290)$$

Za prvo iteracijo na osnovi (8.290) torej velja ( $k + 1 = 1$ ):

$$\underline{x}_1 = \underline{x}(1) = [x_1(1) \quad x_2(1)] = \underline{x}_0 - \alpha_0 \cdot \nabla_{\underline{f}}(\underline{x}_0) \quad (8.291)$$

kjer moramo parameter  $\alpha_0$  nastaviti tako, da bo funkcija (8.286) v vektorju  $\underline{x}_1$ , to je  $f(\underline{x}_1) = f(\underline{x}_0 - \alpha_0 \cdot \nabla_{\underline{f}}(\underline{x}_0)) = f(\alpha_0)$ , najmanjša.

Očitno gre pri iskanju optimalnega parametra  $\alpha_{0opt}$  za problem enodimenzionalne optimizacije, ki ga z upoštevanjem izrazov (8.286), (8.288) in (8.291) formuliramo na naslednji način:

$$\begin{aligned}
 f(\alpha_{0opt}) &= \min_{\alpha_0 > 0} f(\underline{x}_1) = \min_{\alpha_0 > 0} f(\underline{x}_0 - \alpha_0 \cdot \underline{\nabla} f(\underline{x}_0)) = \\
 &= \min_{\alpha_0 > 0} f([0 \ 0]^T - \alpha_0 \cdot [-2\sqrt{5} \ -2\pi]^T) = \\
 &= \min_{\alpha_0 > 0} f([2\alpha_0\sqrt{5} \ 2\alpha_0\pi]^T) = \min_{\alpha_0 > 0} f(2\alpha_0\sqrt{5}, 2\alpha_0\pi) = \\
 &= \min_{\alpha_0 > 0} \left( (2\alpha_0\sqrt{5} - \sqrt{5})^2 + (2\alpha_0\pi - \pi)^2 + 10 \right) = \\
 &= \min_{\alpha_0 > 0} \left( 5(2\alpha_0 - 1)^2 + \pi^2(2\alpha_0 - 1)^2 + 10 \right) = \min_{\alpha_0 > 0} \underbrace{\left( (2\alpha_0 - 1)^2(5 + \pi^2) + 10 \right)}_{f(\alpha_0)}
 \end{aligned} \tag{8.292}$$

V praksi rešimo ta problem na takšen način, da računalnik z eno od numeričnih metod enodimenzionalne optimizacije, ki smo jih obravnavali v 6. poglavju, poišče optimum za  $\alpha_0$ . Seveda je to potrebno storiti ne le v prvi, pač pa tudi v vsaki naslednji iteraciji. Za potrebe razumevanja in utrjevanja snovi tega predmeta pa bomo poiskali dotični optimum s pomočjo analitičnega pristopa.

Funkcija  $f(\alpha_0)$ , podana na koncu izraza (8.292), doseže minimalno vrednost tedaj, ko velja:

$$\begin{aligned}
 \frac{df(\alpha_0)}{d\alpha_0} &= 0 \Rightarrow \\
 \frac{d\left((2\alpha_0 - 1)^2(5 + \pi^2) + 10\right)}{d\alpha_0} &= 0 \Rightarrow \\
 (5 + \pi^2) \cdot 2 \cdot (2\alpha_0 - 1) \cdot 2 &= 0 \Rightarrow \\
 \alpha_{0OPT} &= \frac{1}{2}
 \end{aligned} \tag{8.293}$$

Nov vektor odločitvenih spremenljivk (v 1. iteraciji), z upoštevanjem izrazov (8.288), (8.291) in (8.293), potemtakem zavzame naslednjo vrednost:

$$\underline{x}_1 = \underline{x}_0 - \alpha_{0OPT} \cdot \nabla_{\underline{f}}(x_0) = [0 \ 0]^T - \frac{1}{2} \cdot [-2\sqrt{5} \ -2\pi]^T = [\sqrt{5} \ \pi]^T \quad (8.294)$$

Tako smo dobili rezultat za vektor odločitvenih spremenljivk po 1. iteraciji metode najstrmejšega spusta, pri čemer smo si za izračun parametra  $\alpha_0$ , ki določa optimalno (najstrmejšo) smer tega spusta, pomagali z analitičnim izračunom.

Seveda je rezultat (8.294) le vmesni rezultat metode, ki z napredovanjem nadaljnjih iteracij pomika vektor odločitvenih spremenljivk k numerično izračunanemu vektorju ekstremov funkcije odločitvenih spremenljivk (8.286).

Primerjajmo sedaj ta rezultat (8.294), ki torej predstavlja numerično izračunan vektor odločitvenih spremenljivk ekstremov po 1. iteraciji, z analitično izračunanim vektorjem ekstremov funkcije (8.286). Analitično izračunamo vektor ekstremov funkcije (8.286) na takšen način, da njen gradient v splošni točki prostora  $[x_1, x_2, f(x_1, x_2)]$ , ki smo ga izračunali že v izrazu (8.287), izenačimo z ničelnim vektorjem. Tako dobimo potreben pogoj za obstoj ekstremne točke, ki se glasi:

$$\begin{aligned} \nabla_{\underline{f}}(\underline{x}) &= \nabla_{\underline{f}}([x_1 \ x_2]^T) = \nabla_{\underline{f}}(x_1, x_2) = \left[ \frac{\partial f}{\partial x_1} \ \frac{\partial f}{\partial x_2} \right]^T = \\ &= [2(x_1 - \sqrt{5}) \ 2(x_2 - \pi)]^T = [0 \ 0] \end{aligned} \quad (8.295)$$

oziroma:

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 2(x_1 - \sqrt{5}) = 0 \\ \frac{\partial f}{\partial x_2} &= 2(x_2 - \pi) = 0 \end{aligned} \quad (8.296)$$

Analitično izračunan vektor ekstremov na osnovi izrazov (8.296) torej zavzame vrednost:

$$\left. \begin{array}{l} x_1^* = \sqrt{5} \\ x_2^* = \pi \end{array} \right\} \Rightarrow \underline{x}_A = \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} \sqrt{5} \\ \pi \end{bmatrix} \quad (8.297)$$



V naslednjem koraku želimo primerjati rezultata (8.294) in (8.297). Čeprav sta v tem primeru izjemoma rezultata že na prvi pogled enaka med seboj, vseeno podajmo bolj splošen način za primerjavo dveh vektorjev.

To storimo tako, da izračunamo normo razlike med obema vektorjema ekstremov, to je med analitično izračunanim vektorjem ekstremov (8.297) in numerično izračunanim vektorjem ekstremov v 1. iteraciji (8.294):

$$\|\underline{x}_A - \underline{x}_1\| = \|\underline{x}_A - \underline{x}(1)\| = \left\| \begin{bmatrix} \sqrt{5} \\ \pi \end{bmatrix} - \begin{bmatrix} \sqrt{5} \\ \pi \end{bmatrix} \right\| = \left\| \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\| = 0 \leq \varepsilon = 10^{-2} \quad (8.298)$$

Torej očitno že 1. iteracija zadošča za dovolj dober rezultat, saj se že pri njej numeričen rezultat popolnoma ujema z analitičnim.

To je razvidno tudi, če bi poskušali na osnovi (8.290) izračunati še drugo iteracijo numeričnega postopka ( $k + 1 = 2$ ).

Imeli bi:

$$\underline{x}_2 = \underline{x}(2) = [x_1(2) \quad x_2(2)] = \underline{x}_1 - \alpha_1 \cdot \nabla f(\underline{x}_1) \quad (8.299)$$

Gradient funkcije (8.286) v vektorju  $\underline{x}(1) = \underline{x}_1$  iz (8.294) je z upoštevanjem (8.287) enak:

$$\begin{aligned} \nabla f(\underline{x}_1) &= \nabla f([x_1(1) \quad x_2(1)]^T) = \nabla f\left(\begin{bmatrix} \sqrt{5} \\ \pi \end{bmatrix}\right) = \nabla f(\sqrt{5}, \pi) = \\ &= \begin{bmatrix} 2(\sqrt{5} - \sqrt{5}) & 2(\pi - \pi) \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \end{bmatrix}^T = \underline{0} \end{aligned} \quad (8.300)$$

Čim pa je gradient enak ničelnemu vektorju, pa ni potrebno računati parametra  $\alpha_{1opt}$  v naslednji, to je 2. iteraciji. Z iteracijskim postopkom smo namreč ne glede na  $\alpha_{1opt}$  že na koncu, saj na osnovi (8.300) velja:

$$\underline{x}_2 = \underline{x}(2) = [x_1(2) \quad x_2(2)] = \underline{x}_1 - \alpha_1 \cdot \underline{0} = \underline{x}_1 = \begin{bmatrix} \sqrt{5} \\ \pi \end{bmatrix} \quad (8.301)$$

Torej se rezultat 2. iteracije (in vseh naslednjih) glede na 1. iteracijo ne spremeni več. Razlog za to leži v dejstvu, da je gradient funkcije (8.286) v naslednjih iteracijah vseskozi enak  $\underline{0}$ , kot se je to zgodilo pri izrazu (8.300).

Velja torej:

$$\begin{aligned} \nabla f(\underline{x}_{k-1}) &= \underline{0}, \quad k = 2, 3, \dots, N \\ \underline{x}_k = \underline{x}(k) &= [x_1(k) \quad x_2(k)] = \underline{x}_{k-1} - \alpha_{k-1} \cdot \underline{0} = \underline{x}_{k-1} = \begin{bmatrix} \sqrt{5} \\ \pi \end{bmatrix}, \quad k = 2, 3, \dots, N \end{aligned} \quad (8.302)$$

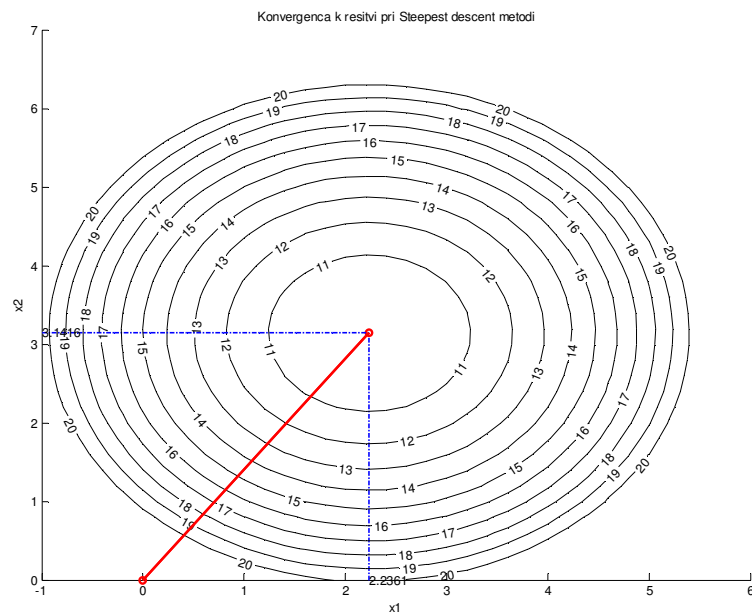
pri čemer je rešitev v izrazu (8.302), to je  $\begin{bmatrix} \sqrt{5} & \pi \end{bmatrix}^T$ , numerično izračunan vektor ekstremov funkcije odločitvenih spremenljivk (8.286).

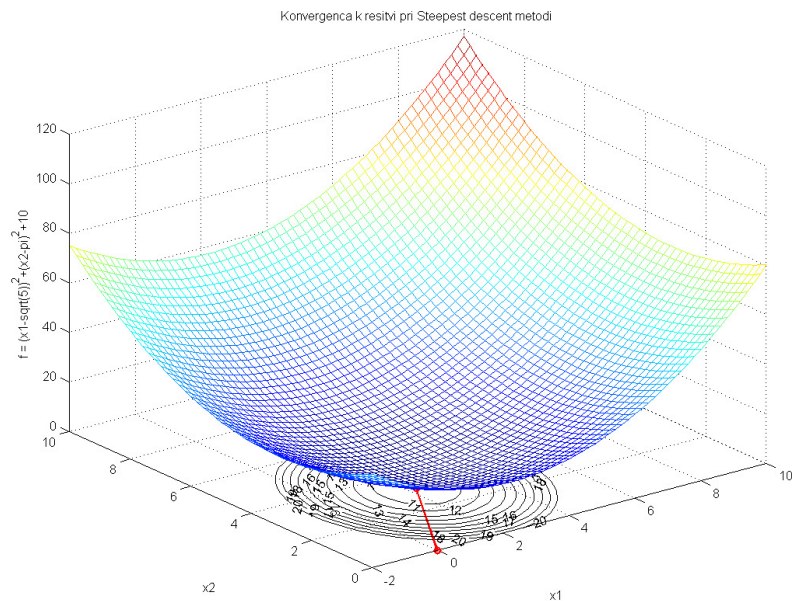
Prikaz dinamike konvergence metode najstrmejšega spusta si lahko ogledamo na sliki 116, kjer je narisani tudi potek ciljne funkcije  $f(x_1, x_2)$  iz izraza (8.286) pri ekvidistančnih vrednostih za  $(x_1, x_2)$  (slika 116 spodaj). Iz slike je razvidno, kako že v eni iteraciji vrednosti aproksimacij optimalnih odločitvenih spremenljivk skočijo od začetne točke  $\underline{x}(0) = \underline{x}_0 = [x_1(0) \quad x_2(0)]^T = [0 \quad 0]^T$  do končne (optimalne) točke:  $\underline{x}(N) = \underline{x}(1) = \underline{x}_1 = [x_1(1) \quad x_2(1)]^T = [2.2361 \quad 3.1416]^T \approx \begin{bmatrix} \sqrt{5} & \pi \end{bmatrix}^T$ .

Tovrsten skok je na sliki 116 označen s povezanima krogcema (o), označene pa so tudi nivojnice ciljne funkcije (8.286) v  $(x_1, x_2)$  ravnini (z resolucijo 1: 20, 19, 18, 17, ...). Pri tem skoku funkcija (8.286) v trenutku pade od začetne vrednosti  $f(0, 0) = (0 - \sqrt{5})^2 + (0 - \pi)^2 + 10 = 24.86$  do končne (minimalne) vrednosti  $f(2.2361, 3.1416) = (2.2361 - \sqrt{5})^2 + (3.1416 - \pi)^2 + 10 \approx 0 + 0 + 10 = 10$ .

Na sliki 116 spodaj je tudi razvidno, da funkcija  $f(x_1, x_2)$  dejansko doseže najnižjo točko (minimum) v prostoru  $(x_1, x_2, f(x_1, x_2))$ , ko postopek konvergira (skoči) k optimalni rešitvi  $[2.2361 \ 3.1416]^T$ .

Omenimo še, da so bili numerični izračuni opravljeni v programskem jeziku Matlab. Za potrebe numeričnega izračuna optimuma  $\alpha_{opt}$  je bila namesto analitičnega izračuna (8.293) uporabljena enodimenzionalna optimizacijska metoda aproksimacije s parabolo.





Slika 116: Prikaz dinamike konvergence metode najstrmejšega spusta: Zgoraj - Prikaz konvergiranja k optimalni rešitvi (**povezani krogi**) v prostoru odločitvenih spremenljivk  $(x_1, x_2)$ ; Spodaj – Prikaz ciljne funkcije  $f(x_1, x_2)$  iz izraza (8.286) v odvisnosti od  $(x_1, x_2)$ ; **Nivojnice so prikazane z resolucijo 1.**

Primerjava slik 115 in 116 pokaže očitne prednosti metode najstrmejšega spusta glede na Hook-Jeevesovo metodo. Tako je prva metoda konvergirala k optimalni rešitvi že po eni iteraciji, druga pa šele po 16 iteracijah.

### Primer 2:

S pomočjo metode najstrmejšega spusta določite optimalno točko za skalarno funkcijo vektorske odločitvene spremenljivke (brez omejitev):

$$\begin{aligned}
 f(\underline{x}) &= f([x_1 \quad x_2 \quad x_3]^T) = f(x_1, x_2, x_3) = \\
 &= x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3
 \end{aligned}
 \tag{8.303}$$

kjer je začetni vektor odločitvenih spremenljivk enak:  
 $\underline{x}(0) = \underline{x}_0 = [x_1(0) \quad x_2(0) \quad x_3(0)]^T = [0 \quad 0 \quad 0]^T$ . Prečna črta spodaj (zaradi

preglednosti!) označuje vektor. Izračunajte prvo iteracijo vektorja odločitvenih spremenljivk  $\underline{x}(1) = \underline{x}_1 = [x_1(1) \ x_2(1) \ x_3(1)]^T$ , pri čemer uporabite izraz brez norme.

Gradient funkcije (8.303) v poljubni točki štiridimenzionalnega prostora  $[x_1, x_2, x_3, f(x_1, x_2, x_3)]$  je enak:

$$\begin{aligned} \nabla f(\underline{x}) &= \nabla f([x_1 \ x_2 \ x_3]^T) = \nabla f(x_1, x_2, x_3) = \left[ \frac{\partial f(\underline{x})}{\partial x_1} \quad \frac{\partial f(\underline{x})}{\partial x_2} \quad \frac{\partial f(\underline{x})}{\partial x_3} \right]^T = \\ &= [4x_1^3 + 3x_1^2 - 1 + x_2x_3 \quad 4x_2^3 - 2x_2 + 1 + x_1x_3 \quad 2x_3 - 1 + x_1x_2]^T \end{aligned} \quad (8.304)$$

Gradient funkcije (8.303) v začetni točki  $\underline{x}(0) = \underline{x}_0$  je v skladu z (8.304) enak:

$$\begin{aligned} \nabla f(\underline{x}_0) &= \nabla f([x_1(0) \ x_2(0) \ x_3(0)]^T) = \nabla f([0 \ 0 \ 0]^T) = \nabla f(0, 0, 0) = \\ &= [4 \cdot 0^3 + 3 \cdot 0^2 - 1 + 0 \cdot 0 \quad 4 \cdot 0^3 - 2 \cdot 0 + 1 + 0 \cdot 0 \quad 2 \cdot 0 - 1 + 0 \cdot 0]^T = \\ &= [-1 \ 1 \ -1]^T \end{aligned} \quad (8.305)$$

Pri metodi najstrmejšega spusta velja za vektor odločitvenih spremenljivk naslednja relacija v  $k+1$ . iteraciji metode:

$$\underline{x}_{k+1} = \underline{x}_k + \Delta \underline{x}_k = \underline{x}_k + \alpha_k \cdot \underline{d}_k = \underline{x}_k - \alpha_k \cdot \frac{\nabla f(\underline{x}_k)}{\|\nabla f(\underline{x}_k)\|}, \quad k = 0, 1, 2, \dots \quad (8.306)$$

oziroma, če norme ne upoštevamo:

$$\underline{x}_{k+1} = \underline{x}_k + \Delta \underline{x}_k = \underline{x}_k + \alpha_k \cdot \underline{d}_k = \underline{x}_k - \alpha_k \cdot \nabla f(\underline{x}_k), \quad k = 0, 1, 2, \dots \quad (8.307)$$

Za prvo iteracijo na osnovi (8.307) velja ( $k + 1 = 1$ ):

$$\underline{x}_1 = \underline{x}(1) = [x_1(1) \ x_2(1) \ x_3(1)] = \underline{x}_0 - \alpha_0 \cdot \nabla f(\underline{x}_0) \quad (8.308)$$

kjer moramo parameter  $\alpha_0$  nastaviti tako, da bo funkcija (8.303) v vektorju  $\underline{x}_1$ , to je  $f(\underline{x}_1) = f(\underline{x}_0 - \alpha_0 \cdot \nabla f(\underline{x}_0)) = f(\alpha_0)$ , najmanjša.

Očitno gre pri iskanju optimalnega parametra  $\alpha_{0opt}$  za problem enodimenzionalne optimizacije, ki ga z upoštevanjem izrazov (8.303), (8.305) in (8.308) formuliramo na naslednji način:

$$\begin{aligned}
 f(\alpha_{0opt}) &= \min_{\alpha_0 > 0} f(\underline{x}_1) = \min_{\alpha_0 > 0} f(\underline{x}_0 - \alpha_0 \cdot \nabla f(\underline{x}_0)) = \\
 &= \min_{\alpha_0 > 0} f([0 \ 0 \ 0]^T - \alpha_0 \cdot [-1 \ 1 \ -1]^T) = \\
 &= \min_{\alpha_0 > 0} f([\alpha_0 \ -\alpha_0 \ \alpha_0]^T) = \min_{\alpha_0 > 0} f(\alpha_0, -\alpha_0, \alpha_0) = \\
 &= \min_{\alpha_0 > 0} (\alpha_0^4 + \alpha_0^3 - \alpha_0 + (-\alpha_0)^4 - (-\alpha_0)^2 + (-\alpha_0) + \alpha_0^2 - \alpha_0 + \alpha_0(-\alpha_0)\alpha_0) = \\
 &= \min_{\alpha_0 > 0} \underbrace{(2\alpha_0^4 - 3\alpha_0)}_{f(\alpha_0)}
 \end{aligned} \tag{8.309}$$

V praksi rešimo ta problem na takšen način, da računalnik z eno od numeričnih metod enodimenzionalne optimizacije, ki smo jih obravnavali v 6. poglavju, poišče optimum za  $\alpha_0$ . Seveda je to potrebno storiti ne le v prvi, pač pa tudi v vsaki naslednji iteraciji. Za potrebe razumevanja in utrjevanja snovi tega predmeta pa bomo poiskali dotični optimum s pomočjo analitičnega pristopa.

Funkcija  $f(\alpha_0)$ , podana na koncu izraza (8.309), doseže minimalno vrednost tedaj, ko velja:

$$\begin{aligned}
 \frac{df(\alpha_0)}{d\alpha_0} &= 0 \Rightarrow \\
 \frac{d(2\alpha_0^4 - 3\alpha_0)}{d\alpha_0} &= 0 \Rightarrow \\
 8\alpha_0^3 - 3 &= 0 \Rightarrow \\
 \alpha_{0OPT} &= \sqrt[3]{\frac{3}{8}} = 0.7211
 \end{aligned} \tag{8.310}$$

Nov vektor odločitvenih spremenljivk (v 1. iteraciji), z upoštevanjem izrazov (8.305), (8.308) in (8.310), potemtakem zavzame naslednjo vrednost:

$$\begin{aligned} \underline{x}_1 &= \underline{x}_0 - \alpha_{OPT} \cdot \nabla f(\underline{x}_0) = [0 \ 0 \ 0]^T - 0.7211 \cdot [-1 \ 1 \ -1]^T = \\ &= [0.72112 \ -0.72112 \ 0.72112]^T \end{aligned} \quad (8.311)$$

Tako smo dobili rezultat za vektor odločitvenih spremenljivk po 1. iteraciji metode najstrmejšega spusta, pri čemer smo si za izračun parametra  $\alpha_0$ , ki določa optimalno (najstrmejšo) smer tega spusta, pomagali z analitičnim izračunom.

Seveda je rezultat (8.311) le vmesni rezultat metode, ki z napredovanjem nadaljnjih iteracij pomika vektor odločitvenih spremenljivk k numerično izračunanemu vektorju ekstremov funkcije odločitvenih spremenljivk (8.303). Zato lahko rečemo, da dani rezultat predstavlja numerično izračunan vektor odločitvenih spremenljivk ekstremov funkcije (8.303) po 1. iteraciji.

Postopek se nadaljuje z naslednjimi iteracijami, to je 2., 3., ..., N. iteracijo. Končamo ga, ko smo se dovolj približali optimalni rešitvi oz. so izpolnjeni pogoji za končanje postopka.

Končni rezultat za aproksimacijski vektor odločitvenih spremenljivk se glasi (numerični izračun v programskem jeziku Matlab):

$$x_p(N) = \tilde{x}^* = [\tilde{x}_1^* \ \tilde{x}_2^* \ \tilde{x}_3^*]^T = [0.5707 \ -0.93956 \ 0.76749]^T$$

v katerem zavzame funkcija (8.303) minimalno vrednost:

$$f(\tilde{x}_1^*, \tilde{x}_2^*, \tilde{x}_3^*) = \min f(x_1, x_2, x_3) = -1.9118$$

Če ta rezultat primerjamo z rezultatom, ki smo ga dobili pri Hook-Jeevesovi metodi, lahko vidimo, da smo dobili zelo podobno optimalno rešitev. Razlika je v tem, da smo pri metodi najstrmejšega spusta izračunali le 11 iteracij, pri Hook-Jeevesovi metodi pa 30 iteracij.

### Primer 3:

S pomočjo metode najstrmejšega spusta določite optimalno točko za funkcijo:

$$f(\underline{x}) = f([x_1 \ x_2]^T) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 \quad (8.312)$$

kjer je začetni vektor odločitvenih spremenljivk enak:

$$\underline{x}(0) = \underline{x}_0 = [x_1(0) \ x_2(0)]^T = [0 \ 0]^T. \text{ Prečna črta spodaj (zaradi preglednosti!) označuje}$$

vektor. Izračunajte prvo iteracijo vektorja odločitvenih spremenljivk:

$$\underline{x}(1) = \underline{x}_1 = [x_1(1) \ x_2(1)]^T, \text{ pri čemer uporabite izraz brez norme.}$$

Gradient funkcije (8.312) v poljubni točki tridimenzionalnega prostora  $[x_1, x_2, f(x_1, x_2)]$

je enak:

$$\begin{aligned} \nabla \underline{f}(\underline{x}) &= \nabla \underline{f}([x_1 \ x_2]^T) = \nabla \underline{f}(x_1, x_2) = \left[ \frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \right]^T = \\ &= [1 + 4x_1 + 2x_2 \quad -1 + 2x_1 + 2x_2]^T \end{aligned} \quad (8.313)$$

Gradient funkcije (8.312) v začetni točki  $\underline{x}(0) = \underline{x}_0$  je v skladu z (8.313) enak:

$$\nabla \underline{f}(\underline{x}_0) = \nabla \underline{f}([x_1(0) \ x_2(0)]^T) = \nabla \underline{f}([0 \ 0]^T) = \nabla \underline{f}(0,0) = [1 \ -1]^T \quad (8.314)$$

Pri metodi najstrmejšega spusta velja za vektor odločitvenih spremenljivk naslednja relacija v  $k+1$ . iteraciji metode:

$$\underline{x}_{k+1} = \underline{x}_k + \Delta \underline{x}_k = \underline{x}_k + \alpha_k \cdot \underline{d}_k = \underline{x}_k - \alpha_k \cdot \frac{\nabla \underline{f}(\underline{x}_k)}{\|\nabla \underline{f}(\underline{x}_k)\|}, \quad k = 0, 1, 2, \dots \quad (8.315)$$

oziroma, če norme ne upoštevamo:

$$\underline{x}_{k+1} = \underline{x}_k + \Delta \underline{x}_k = \underline{x}_k + \alpha_k \cdot \underline{d}_k = \underline{x}_k - \alpha_k \cdot \nabla \underline{f}(\underline{x}_k), \quad k = 0, 1, 2, \dots \quad (8.316)$$

Za prvo iteracijo na osnovi (8.316) torej velja ( $k + 1 = 1$ ):

$$\underline{x}_1 = \underline{x}(1) = [x_1(1) \ x_2(1)] = \underline{x}_0 - \alpha_0 \cdot \nabla \underline{f}(\underline{x}_0) \quad (8.317)$$



kjer moramo parameter  $\alpha_0$  nastaviti tako, da bo funkcija (8.312) v vektorju  $\underline{x}_1$ , to je  $f(\underline{x}_1) = f(\underline{x}_0 - \alpha_0 \cdot \nabla f(\underline{x}_0)) = f(\alpha_0)$ , najmanjša.

Očitno gre pri iskanju optimalnega parametra  $\alpha_{0opt}$  za problem enodimenzionalne optimizacije, ki ga z upoštevanjem izrazov (8.312), (8.314) in (8.317) formuliramo na naslednji način:

$$\begin{aligned}
 f(\alpha_{0opt}) &= \min_{\alpha_0 > 0} f(\underline{x}_1) = \min_{\alpha_0 > 0} f(\underline{x}_0 - \alpha_0 \cdot \nabla f(\underline{x}_0)) = \\
 &= \min_{\alpha_0 > 0} f([0 \ 0]^T - \alpha_0 \cdot [1 \ -1]^T) = \\
 &= \min_{\alpha_0 > 0} f([- \alpha_0 \ \alpha_0]^T) = \\
 &= \min_{\alpha_0 > 0} (-\alpha_0 - \alpha_0 + 2(-\alpha_0)^2 + 2(-\alpha_0)\alpha_0 + \alpha_0^2) = \\
 &= \min_{\alpha_0 > 0} \underbrace{(-2\alpha_0 + \alpha_0^2)}_{f(\alpha_0)}
 \end{aligned} \tag{8.318}$$

V praksi rešimo ta problem na takšen način, da računalnik z eno od numeričnih metod enodimenzionalne optimizacije, ki smo jih obravnavali v 6. poglavju, poišče optimum za  $\alpha_0$ . Seveda je to potrebno storiti ne le v prvi, pač pa tudi v vsaki naslednji iteraciji. Za potrebe razumevanja in utrjevanja snovi tega predmeta pa bomo poiskali dotični optimum s pomočjo analitičnega pristopa.

Funkcija  $f(\alpha_0)$ , podana na koncu izraza (8.318), doseže minimalno vrednost tedaj, ko velja:

$$\begin{aligned}
 \frac{df(\alpha_0)}{d\alpha_0} &= 0 \Rightarrow \\
 \frac{d(-2\alpha_0 + \alpha_0^2)}{d\alpha_0} &= 0 \Rightarrow \\
 -2 + 2\alpha_0 &= 0 \Rightarrow \alpha_{0OPT} = 1
 \end{aligned} \tag{8.319}$$

Nov vektor odločitvenih spremenljivk (v 1. iteraciji), z upoštevanjem izrazov (8.314), (8.317) in (8.319), potemtakem zavzame naslednjo vrednost:

$$\underline{x}_1 = \underline{x}_0 - \alpha_{0OPT} \cdot \underline{\nabla} f(\underline{x}_0) = [0 \ 0]^T - 1 \cdot [1 \ -1]^T = [-1 \ 1]^T \quad (8.320)$$

Tako smo dobili rezultat za vektor odločitvenih spremenljivk po 1. iteraciji metode najstrmejšega spusta, pri čemer smo si za izračun parametra  $\alpha_0$ , ki določa optimalno (najstrmejšo) smer tega spusta, pomagali z analitičnim izračunom.

Seveda je rezultat (8.320) le vmesni rezultat metode, ki z napredovanjem nadaljnjih iteracij pomika vektor odločitvenih spremenljivk k numerično izračunanemu vektorju ekstremov funkcije odločitvenih spremenljivk (8.312). Zato lahko rečemo, da dani rezultat predstavlja numerično izračunan vektor odločitvenih spremenljivk ekstremov funkcije (8.312) po 1. iteraciji.

Postopek se nadaljuje z naslednjimi iteracijami, to je 2., 3., ..., N. iteracijo. Končamo ga, ko smo se dovolj približali optimalni rešitvi oz. so izpolnjeni pogoji za končanje postopka.

#### Primer 4:

Poiščite minimum funkcije s pomočjo metode najstrmejšega spusta:

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 \quad (8.384)$$

Pri tem vzemite  $\underline{x}_0 = [0 \ 0]^T$ . Izračunajte prve 3 iteracije.

Gradient funkcije je:

$$\underline{\nabla} f(x_1, x_2) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \quad (8.385)$$

Iteracijska enačba je (če je norma = 1):

$$\underline{x}_{k+1} = \underline{x}_k - \alpha_k \cdot \underline{\nabla} f(\underline{x}_k), \quad k = 0, 1, 2, \dots \quad (8.386)$$

Sledi (1. iteracija):

$$\underline{x}_1 = \underline{x}_0 - \alpha_0 \cdot \nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \alpha_0 \cdot \begin{bmatrix} 1 + 4 \cdot 0 + 2 \cdot 0 \\ -1 + 2 \cdot 0 + 2 \cdot 0 \end{bmatrix} = \begin{bmatrix} -\alpha_0 \\ \alpha_0 \end{bmatrix} \quad (8.387)$$

Funkcija v  $\underline{x}_1$  zavzame vrednost:

$$\begin{aligned} f(\underline{x}_1) &= f(-\alpha_0, \alpha_0) = (-\alpha_0) - \alpha_0 + 2 \cdot (-\alpha_0)^2 + 2 \cdot (-\alpha_0) \cdot \alpha_0 + \alpha_0^2 = \\ &= -\alpha_0 - \alpha_0 + 2\alpha_0^2 - 2\alpha_0^2 + \alpha_0^2 = -2\alpha_0 + \alpha_0^2 = \alpha_0(\alpha_0 - 2) = f(\alpha_0) \end{aligned} \quad (8.388)$$

Odvajamo po  $\alpha_0$  in dobimo:

$$\frac{df}{d\alpha_0} = 2\alpha_0 - 2 = 0 \Rightarrow \alpha_0^* = 1 \quad (8.389)$$

Sledi:

$$\underline{x}_1 = \begin{bmatrix} -\alpha_0^* \\ \alpha_0^* \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (8.390)$$

2. iteracija:

$$\begin{aligned} \underline{x}_2 &= \underline{x}_1 - \alpha_1 \cdot \nabla \underline{f}(\underline{x}_1) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} 1 + 4 \cdot (-1) + 2 \cdot 1 \\ -1 + 2 \cdot (-1) + 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \\ &= \begin{bmatrix} -1 + \alpha_1 \\ 1 + \alpha_1 \end{bmatrix} \end{aligned} \quad (8.391)$$

Funkcija v  $\underline{x}_2$  zavzame vrednost:

$$\begin{aligned} f(\underline{x}_2) &= (-1 + \alpha_1) - (1 + \alpha_1) + 2 \cdot (-1 + \alpha_1)^2 + 2 \cdot (-1 + \alpha_1) \cdot (1 + \alpha_1) + (1 + \alpha_1)^2 = \\ &= -2 + 2(\alpha_1^2 - 2\alpha_1 + 1) + 2(\alpha_1^2 - 1) + (1 + 2\alpha_1 + \alpha_1^2) = \\ &= -2 + 2\alpha_1^2 - 4\alpha_1 + 2 + 2\alpha_1^2 - 2 + 1 + 2\alpha_1 + \alpha_1^2 = 5\alpha_1^2 - 2\alpha_1 - 1 = f(\alpha_1) \end{aligned} \quad (8.392)$$

Odvajamo po  $\alpha_1$  in dobimo:

$$\frac{df}{d\alpha_1} = 10\alpha_1 - 2 = 0 \Rightarrow \alpha_1^* = \frac{2}{10} = \frac{1}{5} = 0.2 \quad (8.393)$$

Sledi:

$$\underline{x}_2 = \begin{bmatrix} -1 + \alpha_1^* \\ 1 + \alpha_1^* \end{bmatrix} = \begin{bmatrix} -1 + 0.2 \\ 1 + 0.2 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} \quad (8.394)$$

3. iteracija:

$$\begin{aligned} \underline{x}_3 &= \underline{x}_2 - \alpha_2 \cdot \nabla f(\underline{x}_2) = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} - \alpha_2 \cdot \begin{bmatrix} 1 + 4 \cdot (-0.8) + 2 \cdot 1.2 \\ -1 + 2 \cdot (-0.8) + 2 \cdot 1.2 \end{bmatrix} = \\ &= \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} - \alpha_2 \cdot \begin{bmatrix} 0.2 \\ -0.2 \end{bmatrix} = \begin{bmatrix} -0.8 - 0.2\alpha_2 \\ 1.2 + 0.2\alpha_2 \end{bmatrix} \end{aligned} \quad (8.395)$$

Funkcija v  $\underline{x}_3$  zavzame vrednost:

$$\begin{aligned} f(\underline{x}_3) &= f(-0.8 - 0.2\alpha_2; 1.2 + 0.2\alpha_2) = (-0.8 - 0.2\alpha_2) - (1.2 + 0.2\alpha_2) + \\ &+ 2 \cdot (-0.8 - 0.2\alpha_2)^2 + 2 \cdot (-0.8 - 0.2\alpha_2) \cdot (1.2 + 0.2\alpha_2) + \\ &+ (1.2 + 0.2\alpha_2)^2 = \\ &= -2 - 0.4\alpha_2 + 2(0.2\alpha_2 + 0.8)^2 + \\ &+ 2 \cdot (-0.8 \cdot 1.2 - 0.8 \cdot 0.2\alpha_2 - 0.2 \cdot 1.2\alpha_2 - 0.2^2 \alpha_2^2) + \\ &+ (1.2 + 0.2\alpha_2)^2 \quad (8.396) \\ &= -2 - 0.4\alpha_2 + 2(0.2\alpha_2 + 0.8)^2 - 1.92 - 0.8\alpha_2 - 0.08\alpha_2^2 + \\ &+ (1.2 + 0.2\alpha_2)^2 = \\ &= -3.92 - 1.2\alpha_2 + 2(0.2\alpha_2 + 0.8)^2 - 0.08\alpha_2^2 + (1.2 + 0.2\alpha_2)^2 = f(\alpha_2) \end{aligned}$$

Odvajamo po  $\alpha_2$  in dobimo:

$$\begin{aligned} \frac{df(\alpha_2)}{d\alpha_2} &= -1.2 + 4(0.2\alpha_2 + 0.8) \cdot 0.2 - 0.16\alpha_2 + 2(1.2 + 0.2\alpha_2) \cdot 0.2 = \\ &= -1.2 + 0.64 + 0.16\alpha_2 - 0.16\alpha_2 + 0.48 + 0.08\alpha_2 = \\ &= -0.08 + 0.08\alpha_2 \Rightarrow \alpha_2^* = 1 \end{aligned} \quad (8.397)$$

Sledi:

$$\underline{x}_3 = \begin{bmatrix} -0.8 - 0.2\alpha_2^* \\ 1.2 + 0.2\alpha_2^* \end{bmatrix} = \begin{bmatrix} -0.8 - 0.2 \cdot 1 \\ 1.2 + 0.2 \cdot 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1.4 \end{bmatrix} \quad (8.398)$$

### 8.3.3 Newtonova multidimenzionalna metoda

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.3.1.4, izraze (6.91) do (6.101), ter zlasti sliko 80.

#### Primer 1:

Minimizirajte naslednjo funkcijo s pomočjo Newtonove metode:

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 \quad (8.321)$$

pri danem začetnem vektorju:

$$\underline{x}(0) = [\underline{x}_1(0) \quad \underline{x}_2(0)]^T = [0 \quad 0]^T \quad (8.322)$$

Gradient funkcije (8.321) je:

$$\nabla \underline{f}(\underline{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \quad (8.323)$$

Hessova matrika 2. odvodov je:

$$\nabla^2 \underline{f}(\underline{x}) = \underline{\underline{H}}(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \quad (8.324)$$

Inverzna Hessova matrika je:

$$\underline{\underline{H}}^{-1}(\underline{x}) = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}^{-1} = \frac{\text{adj} \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}}{\det \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}} = \frac{\begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}}{4 \cdot 2 - 2 \cdot 2} = \frac{1}{4} \cdot 2 \cdot \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \quad (8.325)$$

$$= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$$

Inverzna Hessova matrika je vedno enaka, ne glede na  $\underline{x}$ !

Iteracijska formula se glasi:

$$\underline{x}_{k+1} = \underline{x}_k - \underline{\underline{H}}^{-1}[\underline{x}_k] \cdot \nabla \underline{f}(\underline{x}_k); \quad k = 0, 1, 2, \dots \quad (8.326)$$

Za 1. iteracijo velja:

$$\begin{aligned} \underline{x}_1 &= \underline{x}_0 - \underline{\underline{H}}^{-1}(\underline{x}_0) \cdot \nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} \underline{x}_1(1) \\ \underline{x}_2(1) \end{bmatrix} = \\ &= \begin{bmatrix} \underline{x}_1(0) \\ \underline{x}_2(0) \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{bmatrix} \Bigg|_{\underline{x}=\underline{x}_0} = \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1+4 \cdot 0+2 \cdot 0 \\ -1+2 \cdot 0+2 \cdot 0 \end{bmatrix} = \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} \end{aligned} \quad (8.327)$$

Za 2. iteracijo velja:

$$\begin{aligned} \underline{x}_2 &= \underline{x}_1 - \underline{\underline{H}}^{-1}(\underline{x}_1) \cdot \nabla \underline{f}(\underline{x}_1) = \begin{bmatrix} \underline{x}_1(2) \\ \underline{x}_2(2) \end{bmatrix} = \\ &= \begin{bmatrix} \underline{x}_1(1) \\ \underline{x}_2(1) \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{bmatrix} \Bigg|_{\underline{x}=\underline{x}_1} = \end{aligned} \quad (8.328)$$

$$\begin{aligned}
&= \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 + 4 \cdot (-1) + 2 \cdot \frac{3}{2} \\ -1 + 2 \cdot (-1) + 2 \cdot \frac{3}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \\
&= \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} \rightarrow \text{OPTIMUM}
\end{aligned}$$

Torej, ker je gradient  $\nabla f(\underline{x}_1)$  enak  $\underline{0}$ , smo očitno že dosegli optimum, kar se je zgodilo že v 1. iteraciji, potrdilo pa v 2. iteraciji.

### Primer 2:

Minimizirajte naslednjo funkcijo s pomočjo Newtonove metode, ( $\underline{x}(0) = [0 \ 0]^T$ ):

$$f(x_1, x_2) = x_1 - x_2 + x_1^2 - x_1 x_2 \quad (8.329)$$

Gradient je enak:

$$\nabla f(\underline{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 + 2x_1 - x_2 \\ -1 - x_1 \end{bmatrix} \quad (8.330)$$

Hessova matrika je enaka:

$$\underline{\underline{H}}(\underline{x}) = \nabla^2 f(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}}_{\text{JE KONSTANTNA}} \quad (8.331)$$

Njen inverz je enak (ne glede na  $\underline{x}$ ):

$$\underline{\underline{H}}^{-1}(\underline{x}) = \frac{\text{adj}\underline{\underline{H}}(\underline{x})}{\det \underline{\underline{H}}(\underline{x})} = \frac{\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}}{2 \cdot 0 - 1 \cdot 1} = \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix} \quad (8.332)$$

Iteracijska formula je:

$$\underline{x}_{k+1} = \underline{x}_k - \underline{H}^{-1}(\underline{x}_k) \cdot \nabla \underline{f}(\underline{x}_k); \quad k = 0, 1, 2, \dots \quad (8.333)$$

1. iteracija:

$$\begin{aligned} \underline{x}_1 &= \underline{x}_0 - \underline{H}^{-1}(\underline{x}_0) \cdot \nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ -1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1+2x_1-x_2 \\ -1-x_1 \end{bmatrix} \Bigg|_{\underline{x}=\underline{x}_0} = \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1+2 \cdot 0 - 0 \\ -1 - 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \end{aligned} \quad (8.334)$$

2. iteracija:

$$\begin{aligned} \underline{x}_2 &= \underline{x}_1 - \underline{H}^{-1}(\underline{x}_1) \cdot \nabla \underline{f}(\underline{x}_1) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ -1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1+2 \cdot (-1) - (-1) \\ -1 - (-1) \end{bmatrix} = \\ &= \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\nabla f(\underline{x}_1)=0!} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \Rightarrow \text{OPTIMUM} \end{aligned} \quad (8.335)$$

Preizkus (analitični):

$$\nabla \underline{f}(\underline{x}) = \begin{bmatrix} 1+2x_1-x_2 \\ -1-x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.336)$$

Sledi:

$$\left. \begin{array}{l} x_2 = 1 + 2x_1 \\ x_1 = -1 \end{array} \right\} \Rightarrow x_2 = 1 + 2 \cdot (-1) = 1 - 2 = -1 \quad (8.337)$$

Torej, analitični preizkus da  $(\underline{x}_{1A}^*, \underline{x}_{2A}^*) = (-1, -1)$ , kar potrди, da se je pravilno izračunal rezultat  $(-1, -1)$  z Newtonovo metodo (že v 1. iteraciji).

### Primer 3:

Minimizirajte naslednjo funkcijo z Newtonovo metodo,  $(\underline{x}(0) = [0 \ 0]^T)$ :



$$f(x_1, x_2) = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2 \quad (8.338)$$

Gradient je enak:

$$\underline{\nabla} f(\underline{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4 - 4x_1 - 2x_2 \\ 6 - 2x_1 - 4x_2 \end{bmatrix} \quad (8.339)$$

Hessova matrika je enaka:

$$\underline{\underline{H}}(\underline{x}) = \nabla^2 f(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} -4 & -2 \\ -2 & -4 \end{bmatrix} \quad (8.340)$$

JE KONSTANTNA

Njen inverz je enak (ne glede na  $\underline{x}$ ):

$$\underline{\underline{H}}^{-1}(\underline{x}) = \frac{\text{adj}\underline{\underline{H}}(\underline{x})}{\det\underline{\underline{H}}(\underline{x})} = \frac{\begin{bmatrix} -4 & 2 \\ 2 & -4 \end{bmatrix}}{4 \cdot 4 - 2 \cdot 2} = \frac{1}{12} \cdot \begin{bmatrix} -4 & 2 \\ 2 & -4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{3} \end{bmatrix} \quad (8.341)$$

Iteracijska formula je:

$$\underline{x}_{k+1} = \underline{x}_k - \underline{\underline{H}}^{-1}(\underline{x}_k) \cdot \underline{\nabla} f(\underline{x}_k); \quad k = 0, 1, 2, \dots \quad (8.342)$$

1. iteracija:

$$\begin{aligned} \underline{x}_1 &= \underline{x}_0 - \underline{\underline{H}}^{-1}(\underline{x}_0) \cdot \underline{\nabla} f(\underline{x}_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -\frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 4 - 4x_1 - 2x_2 \\ 6 - 2x_1 - 2x_2 \end{bmatrix} \Bigg|_{\underline{x}=\underline{x}_0} = \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 4 - 0 - 0 \\ 6 - 0 - 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} - 1 \\ -\frac{4}{6} + 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{4}{3} \end{bmatrix} \end{aligned} \quad (8.343)$$

2. iteracija:

$$\begin{aligned}
\underline{x}_2 = \underline{x}_1 - \underline{H}^{-1}(\underline{x}_1) \cdot \nabla \underline{f}(\underline{x}_1) &= \begin{bmatrix} \frac{1}{3} \\ \frac{4}{3} \\ \frac{4}{3} \end{bmatrix} - \begin{bmatrix} -\frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 4 - 4 \cdot \frac{1}{3} - 2 \cdot \frac{4}{3} \\ 6 - 2 \cdot \frac{1}{3} - 4 \cdot \frac{4}{3} \end{bmatrix} = \\
&= \begin{bmatrix} \frac{1}{3} \\ \frac{4}{3} \\ \frac{4}{3} \end{bmatrix} + \begin{bmatrix} \frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{3} \end{bmatrix} \cdot \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\nabla f(\underline{x}_1)=0!} = \begin{bmatrix} \frac{1}{3} \\ \frac{4}{3} \\ \frac{4}{3} \end{bmatrix} \Rightarrow \text{OPTIMUM}
\end{aligned} \tag{8.344}$$

Preizkus (analitični):

$$\nabla \underline{f}(\underline{x}) = \begin{bmatrix} 4 - 4x_1 - 2x_2 \\ 6 - 2x_1 - 4x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{8.345}$$

$$\left. \begin{aligned} 4 = 4x_1 + 2x_2 &\Rightarrow 2 = 2x_1 + x_2 \\ 6 = 2x_1 + 4x_2 &\Rightarrow 3 = x_1 + 2x_2 \quad | \cdot 2 \end{aligned} \right\} \tag{8.346}$$

Sledi:

$$\begin{aligned} 2 - 6 &= 2x_1 + x_2 - 2 \cdot (x_1 + 2x_2) \\ -4 = x_2 - 4x_2 &= -3x_2 \Rightarrow x_{2A}^* = \frac{4}{3} \end{aligned} \tag{8.347}$$

Zato sledi:

$$2 = 2x_1 + \frac{4}{3} \Rightarrow x_{1A}^* = \frac{2 - \frac{4}{3}}{2} = \frac{\frac{2}{3}}{2} = \frac{1}{3} \tag{8.348}$$

Analitični preizkus da  $(\underline{x}_{1A}^*, \underline{x}_{2A}^*) = \left(\frac{1}{3}, \frac{4}{3}\right)$ , kar potrdi, da se je pravilno izračunal

rezultat  $\left(\frac{1}{3}, \frac{4}{3}\right)$  z Newtonovo metodo (že v 1. iteraciji).

#### Primer 4:

Uporabite Newtonovo metodo, da maksimizirate funkcijo:

$$z(x_1, x_2) = -(x_1 - \sqrt{5})^2 - (x_2 - \pi)^2 - 10 \quad (8.349)$$

pri čemer vzemite  $\underline{x}_0 = [6.59 ; 5.89]$ . A dobimo po 1. iteraciji že pravi optimum?

Gradient je:

$$\nabla_{\underline{z}}(x_1, x_2) = \begin{bmatrix} \frac{\partial z}{\partial x_1} \\ \frac{\partial z}{\partial x_2} \end{bmatrix} = \begin{bmatrix} -2 \cdot (x_1 - \sqrt{5}) \\ -2 \cdot (x_2 - \pi) \end{bmatrix} \quad (8.350)$$

Hessova matrika je:

$$\underline{\underline{H}}(x_1, x_2) = \begin{bmatrix} \frac{\partial^2 z}{\partial x_1^2} & \frac{\partial^2 z}{\partial x_2 \partial x_1} \\ \frac{\partial^2 z}{\partial x_1 \partial x_2} & \frac{\partial^2 z}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \underline{\underline{H}}_{konst} \quad (8.351)$$

Njen inverz je enak:

$$\underline{\underline{H}}_{konst}^{-1} = \frac{adj \underline{\underline{H}}_{konst}}{\det \underline{\underline{H}}_{konst}} = \frac{\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}}{(-2)^2 - 0^2} = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \quad (8.352)$$

Iteracijska formula je:

$$\underline{x}_{k+1} = \underline{x}_k - \underline{\underline{H}}^{-1}(\underline{x}_k) \cdot \nabla_{\underline{z}}(\underline{x}_k); \quad k = 0, 1, 2, \dots \quad (8.353)$$

1. iteracija:

$$\underline{x}_1 = \underline{x}_0 - \underline{\underline{H}}_{konst}^{-1} \cdot \nabla_{\underline{z}}(\underline{x}_0) = \begin{bmatrix} 6.59 \\ 5.89 \end{bmatrix} - \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} -2 \cdot (x_1 - \sqrt{5}) \\ -2 \cdot (x_2 - \pi) \end{bmatrix} \Bigg|_{\substack{x_1=6.59 \\ x_2=5.89}} = \quad (8.354)$$

$$= \begin{bmatrix} 6.59 \\ 5.89 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} -2 \cdot (6.59 - \sqrt{5}) \\ -2 \cdot (5.89 - \pi) \end{bmatrix} = \begin{bmatrix} 6.59 \\ 5.89 \end{bmatrix} + \begin{bmatrix} -(6.59 - \sqrt{5}) \\ -(5.89 - \pi) \end{bmatrix} = \begin{bmatrix} \sqrt{5} \\ \pi \end{bmatrix}$$

A smo v (8.354) že dobili pravi optimum? Preverimo z  $\nabla_{\underline{z}}(\underline{x}_1)$ , ki je enak:

$$\nabla_{\underline{z}}(\underline{x}_1) = \begin{bmatrix} -2 \cdot (x_1 - \sqrt{5}) \\ -2 \cdot (x_2 - \pi) \end{bmatrix} \Bigg|_{\substack{x_1 = \sqrt{5} \\ x_2 = \pi}} = \begin{bmatrix} -2 \cdot (\sqrt{5} - \sqrt{5}) \\ -2 \cdot (\pi - \pi) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.355)$$

Torej dobimo že po 1. iteraciji pravi OPTIMUM, ker  $\nabla_{\underline{z}}(\underline{x}_1) = \underline{0}$ .

### Primer 5:

Uporabite Newtonovo metodo, da poiščete maksimum funkcije:

$$z(x_1, x_2) = -\sin x_1 x_2 + \cos(x_1 - x_2) \quad (8.356)$$

pri čemer vzemite  $\underline{x}_0 = [0.04 ; 0.08]$ . A dobimo po 1. iteraciji že pravi optimum?

Gradient je enak:

$$\begin{aligned} \nabla_{\underline{z}}(x_1, x_2) &= \begin{bmatrix} \frac{\partial z}{\partial x_1} \\ \frac{\partial z}{\partial x_2} \end{bmatrix} = \begin{bmatrix} -(\cos x_1 x_2) \cdot x_2 - \sin(x_1 - x_2) \cdot 1 \\ -(\cos x_1 x_2) \cdot x_1 - \sin(x_1 - x_2) \cdot (-1) \end{bmatrix} = \\ &= \begin{bmatrix} -x_2 \cdot \cos x_1 x_2 - \sin(x_1 - x_2) \\ -x_1 \cdot \cos x_1 x_2 + \sin(x_1 - x_2) \end{bmatrix} \end{aligned} \quad (8.357)$$

Hessejeva matrika je:

$$\underline{\underline{H}}(x_1, x_2) = \begin{bmatrix} \frac{\partial^2 z}{\partial x_1^2} & \frac{\partial^2 z}{\partial x_2 \partial x_1} \\ \frac{\partial^2 z}{\partial x_1 \partial x_2} & \frac{\partial^2 z}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (8.358)$$

$$\begin{aligned}
a_{11} &= -x_2 \cdot (-\sin x_1 x_2) \cdot x_2 - \cos(x_1 - x_2) \cdot 1 = x_2^2 \cdot \sin x_1 x_2 - \cos(x_1 - x_2) \\
a_{12} &= -1 \cdot \cos x_1 x_2 - x_2 \cdot (-\sin x_1 x_2) \cdot x_1 - \cos(x_1 - x_2) \cdot (-1) = \\
&= -\cos x_1 x_2 + x_1 x_2 \cdot \sin x_1 x_2 + \cos(x_1 - x_2) \\
a_{21} &= -1 \cdot \cos x_1 x_2 - x_1 \cdot (-\sin x_1 x_2) \cdot x_2 + \cos(x_1 - x_2) \cdot 1 = \\
&= -\cos x_1 x_2 + x_1 x_2 \cdot \sin x_1 x_2 + \cos(x_1 - x_2) \\
a_{22} &= -x_1 \cdot (-\sin x_1 x_2) \cdot x_1 + \cos(x_1 - x_2) \cdot (-1) = x_1^2 \cdot \sin x_1 x_2 - \cos(x_1 - x_2)
\end{aligned}$$

Sledi torej:

$$\begin{aligned}
\underline{\underline{H}}(x_1, x_2) &= \\
&= \begin{bmatrix} x_2^2 \cdot \sin x_1 x_2 - \cos(x_1 - x_2) & -\cos x_1 x_2 + x_1 x_2 \cdot \sin x_1 x_2 + \cos(x_1 - x_2) \\ -\cos x_1 x_2 + x_1 x_2 \cdot \sin x_1 x_2 + \cos(x_1 - x_2) & x_1^2 \cdot \sin x_1 x_2 - \cos(x_1 - x_2) \end{bmatrix} \quad (8.359)
\end{aligned}$$

Njena determinanta je sledeča:

$$\begin{aligned}
\det \underline{\underline{H}}(x_1, x_2) &= (x_2^2 \cdot \sin x_1 x_2 - \cos(x_1 - x_2)) \cdot (x_1^2 \cdot \sin x_1 x_2 - \cos(x_1 - x_2)) - \\
&\quad - (-\cos x_1 x_2 + x_1 x_2 \cdot \sin x_1 x_2 + \cos(x_1 - x_2)) \cdot \\
&\quad \cdot (-\cos x_1 x_2 + x_1 x_2 \cdot \sin x_1 x_2 + \cos(x_1 - x_2)) = \\
&= (x_1 x_2)^2 \cdot \sin^2 x_1 x_2 - x_2^2 \cdot \sin x_1 x_2 \cdot \cos(x_1 - x_2) - \\
&\quad - x_1^2 \cdot \sin x_1 x_2 \cdot \cos(x_1 - x_2) + \cos^2(x_1 - x_2) - \\
&\quad - \left[ \cos^2 x_1 x_2 - x_1 x_2 \cdot \sin x_1 x_2 \cdot \cos x_1 x_2 - \cos x_1 x_2 \cdot \cos(x_1 - x_2) - \right. \\
&\quad \left. - x_1 x_2 \cdot \sin x_1 x_2 \cdot \cos x_1 x_2 + (x_1 x_2)^2 \cdot \sin^2 x_1 x_2 + \right. \\
&\quad \left. + x_1 x_2 \cdot \sin x_1 x_2 \cdot \cos(x_1 - x_2) - \cos(x_1 - x_2) \cdot \cos x_1 x_2 + \right. \\
&\quad \left. + x_1 x_2 \cdot \sin x_1 x_2 \cdot \cos(x_1 - x_2) + \cos^2(x_1 - x_2) \right] \quad (8.360)
\end{aligned}$$

Očitno bo splošno obliko za  $\underline{\underline{H}}^{-1}(x_1, x_2)$ , zelo težko dobiti in se bomo temu izognili!

Iteracijska enačba je:

$$\underline{x}_{k+1} = \underline{x}_k - \underline{\underline{H}}^{-1}(\underline{x}_k) \cdot \nabla \underline{z}(\underline{x}_k); \quad k = 0, 1, 2, \dots \quad (8.361)$$

1. iteracija:

$$\underline{x}_1 = \underline{x}_0 - \underline{\underline{H}}^{-1}(\underline{x}_0) \cdot \nabla \underline{z}(\underline{x}_0) \quad (8.362)$$

$$\begin{aligned}
\nabla_{\underline{x}}(\underline{x}_0) &= \left[ \begin{array}{c} -x_2 \cdot \cos x_1 x_2 - \sin(x_1 - x_2) \\ -x_1 \cdot \cos x_1 x_2 + \sin(x_1 - x_2) \end{array} \right] \Bigg|_{\substack{x_1=0.04 \\ x_2=0.08}} = \\
&= \left[ \begin{array}{c} \overbrace{-0.08 \cdot \cos \left[ \overbrace{0.04 \cdot 0.08 \cdot \frac{180^\circ}{\pi}}^{0.9999} \right]}^{0.9999} - \overbrace{\sin \left[ \overbrace{(0.04 - 0.08) \cdot \frac{180^\circ}{\pi}}^{-0.04} \right]}^{-0.04} \\ \overbrace{-0.04 \cdot \cos \left[ \overbrace{0.04 \cdot 0.08 \cdot \frac{180^\circ}{\pi}}^{0.9999} \right]}^{0.9999} + \overbrace{\sin \left[ \overbrace{(0.04 - 0.08) \cdot \frac{180^\circ}{\pi}}^{-0.04} \right]}^{-0.04} \end{array} \right] = \\
&= \begin{bmatrix} -0.08 + 0.04 \\ -0.04 - 0.04 \end{bmatrix} = \begin{bmatrix} -0.04 \\ -0.08 \end{bmatrix}
\end{aligned} \tag{8.363}$$

Poiščimo še  $\underline{\underline{H}}(x_1, x_2)$  v točki  $\underline{x}_0$ . Sledi:

$$\begin{aligned}
\underline{\underline{H}}(\underline{x}_0) &= \underline{\underline{H}}\left(\underbrace{0.04}_{x_1}, \underbrace{0.08}_{x_2}\right) = \\
&= \begin{bmatrix} \overbrace{0.08^2 \cdot \sin \left[ 0.04 \cdot 0.08 \cdot \frac{180^\circ}{\pi} \right]}^{\tilde{a}_{11}} - \overbrace{\cos \left[ (0.04 - 0.08) \cdot \frac{180^\circ}{\pi} \right]}^{\tilde{a}_{12}} \\ \underbrace{0.04^2 \cdot \sin \left[ 0.04 \cdot 0.08 \cdot \frac{180^\circ}{\pi} \right]}_{\tilde{a}_{22}} - \overbrace{\cos \left[ (0.04 - 0.08) \cdot \frac{180^\circ}{\pi} \right]}^{\tilde{a}_{12}} \end{bmatrix}
\end{aligned} \tag{8.364}$$

$$\tilde{a}_{11} = 0.08^2 \cdot 3.2 \cdot 10^{-3} - 0.9991 \doteq -0.9991$$

$$\tilde{a}_{12} = -\cos \left[ 0.04 \cdot 0.08 \cdot \frac{180^\circ}{\pi} \right] + \underbrace{0.04 \cdot 0.08 \cdot \sin \left[ 0.04 \cdot 0.08 \cdot \frac{180^\circ}{\pi} \right]}_{3.2 \cdot 10^{-3}} +$$

$$+ \underbrace{\cos \left[ (0.04 - 0.08) \cdot \frac{180^\circ}{\pi} \right]}_{0.9999} \doteq -0.9999 + 0 + 0.9999 \approx 0$$

$$\tilde{a}_{22} = \underbrace{0.04^2 \cdot 3.2 \cdot 10^{-3}}_0 - 0.9999 \doteq -0.9999$$

Sledi torej:

$$\underline{\underline{H}}(\underline{x}_0) \doteq \begin{bmatrix} -0.9991 & 0 \\ 0 & -0.9999 \end{bmatrix} \quad (8.365)$$

Inverz pa je (sedaj pa bo možno izračunati):

$$\underline{\underline{H}}^{-1}(\underline{x}_0) = \frac{\text{adj}\underline{\underline{H}}(\underline{x}_0)}{\det\underline{\underline{H}}(\underline{x}_0)} = \frac{\begin{bmatrix} -0.9991 & 0 \\ 0 & -0.9999 \end{bmatrix}}{\underbrace{(-0.9991) \cdot (-0.9999) - 0^2}_1} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad (8.366)$$

Sledi:

$$\begin{aligned} \underline{x}_1 &= \underbrace{\begin{bmatrix} 0.04 \\ 0.08 \end{bmatrix}}_{\underline{x}_0} - \underbrace{\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}}_{\underline{\underline{H}}^{-1}(\underline{x}_0)} \cdot \underbrace{\begin{bmatrix} -0.04 \\ -0.08 \end{bmatrix}}_{\nabla_{\underline{z}}(\underline{x}_0)} = \begin{bmatrix} 0.04 \\ 0.08 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -0.04 \\ -0.08 \end{bmatrix} = \\ &= \begin{bmatrix} 0.04 \\ 0.08 \end{bmatrix} + \begin{bmatrix} -0.04 \\ -0.08 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned} \quad (8.367)$$

Če je to že pravi optimum, preverimo z  $\nabla_{\underline{z}}(\underline{x}_1)$ :

$$\nabla_{\underline{z}}(\underline{x}_1) = \begin{bmatrix} -x_2 \cdot \cos x_1 x_2 - \sin(x_1 - x_2) \\ -x_1 \cdot \cos x_1 x_2 - \sin(x_1 - x_2) \end{bmatrix} \Bigg|_{\substack{x_1=0 \\ x_2=0}} = \begin{bmatrix} -0 \cdot \cos 0 - \sin 0 \\ -0 \cdot \cos 0 + \sin 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.368)$$

Torej dobimo že po 1. iteraciji pravi OPTIMUM, saj je  $\nabla_{\underline{z}}(\underline{x}_1) = \underline{0}$ .

### 8.3.4 Marquardtova metoda

Algoritem združuje dobre lastnosti Newtonove metode in metode najstrmejšega spusta. Hessova matrika Newtonove metode je pri tem nekoliko modificirana, in sicer:

$$\underline{\underline{\tilde{H}}}(\underline{x}_k) = \underline{\underline{H}}(\underline{x}_k) + \alpha_k \cdot \underline{\underline{I}}, \quad (8.369)$$

kjer je  $\underline{\underline{I}}$  enotna matrika,  $\alpha_k$  pa pozitivna konstanta, ki zagotovi pozitivno definitnost matrike  $\underline{\underline{\tilde{H}}}$ , ko  $\underline{\underline{H}}$  ni pozitivno definitna.

Ko je  $\alpha_k$  dovolj velik, (razreda  $10^4$ ), velja naslednje:

$\alpha_k \cdot \underline{I} \gg \underline{H}(\underline{x}_k)$ , torej člen  $\alpha_k \cdot \underline{I}$  dominira glede na člen  $\underline{H}(\underline{x}_k)$ , diagonalni elementi od  $\alpha_k \cdot \underline{I}$  pa pridejo izrazito do izraza. Tedaj velja:

$$\underline{\tilde{H}}(\underline{x}_k) \approx \alpha_k \cdot \underline{I} \Rightarrow \underline{\tilde{H}}^{-1}(\underline{x}_k) = [\underline{H}(\underline{x}_k) + \alpha_k \cdot \underline{I}]^{-1} \approx [\alpha_k \cdot \underline{I}]^{-1} = \frac{1}{\alpha_k} \cdot \underline{I} \quad (8.370)$$

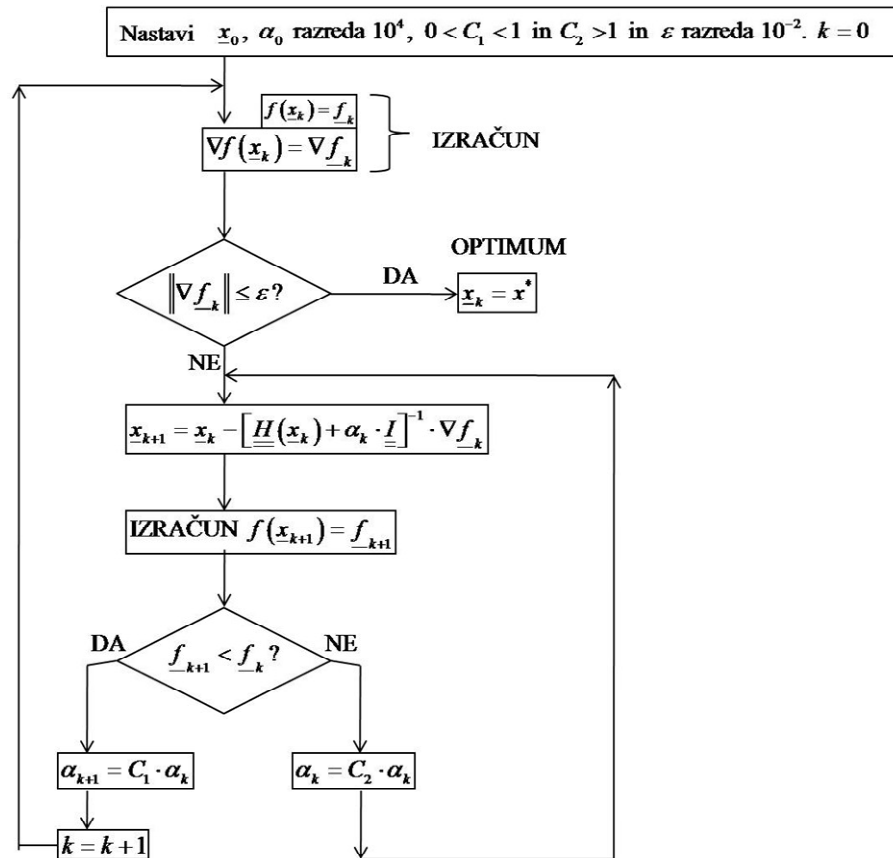
Pri Marquardtovi metodi je priporočljivo, da so vrednosti  $\alpha_k$  velike na začetku iterativnega procesa, potem pa vedno manjše z vsako naslednjo iteracijo.

Na začetku torej prevlada člen  $\alpha_k \cdot \underline{I}$ , proti koncu pa člen  $\underline{H}(\underline{x}_k)$  (saj postane  $\alpha_k \rightarrow 0$ ).

Na začetku se torej metoda obnaša podobno kot metoda najstrmejšega spusta, proti koncu pa kot Newtonova metoda (saj le-ta pospeši konvergenco, ko smo blizu optimuma).

Algoritem je naslednji:





Slika 117: Algoritem Marquardtove metode

Opomba: Za razumevanje teoretičnega ozadja glej tudi poglavje 6.3.1.5, izraze (6.102) do (6.103) ter zlasti sliko 117.

### Primer 1:

Določite minimum za funkcijo:

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 \quad (8.371)$$

pri naslednjih začetnih pogojih;  $\underline{x}_0 = [0 \ 0]^T$ ,  $\alpha_0 = 10^4$ ,  $C_1 = \frac{1}{4}$ ,  $C_2 = 2$ ,  $\varepsilon = 10^{-2}$ .

Uporabite Marquardtovo metodo, izračunajte 2 iteraciji.

Gradient je enak:

$$\nabla \underline{f}(\underline{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \quad (8.372)$$

Hessova matrika:

$$\underline{\underline{H}}(\underline{x}) = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} = \underline{\underline{H}} \quad (8.373)$$

1. iteracija:

Vrednost funkcije v začetni točki:

$$f(\underline{x}_0) = 0 - 0 + 2 \cdot 0^2 + 2 \cdot 0 \cdot 0 + 0^2 = 0 \quad (8.374)$$

Gradient funkcije v začetni točki pa je sledeč:

$$\nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} 1 + 4 \cdot 0 + 2 \cdot 0 \\ -1 + 2 \cdot 0 + 2 \cdot 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (8.375)$$

Sledi:

$$\|\nabla \underline{f}(\underline{x}_0)\| = \sqrt{1^2 + (-1)^2} = \sqrt{2} > \varepsilon \quad (8.376)$$

Dobimo:

$$\begin{aligned} \underline{x}_1 &= \underline{x}_0 - [\underline{\underline{H}} + \alpha_0 \cdot \underline{\underline{I}}]^{-1} \cdot \nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \left[ \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} + 10^4 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right]^{-1} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 10004 & 2 \\ 2 & 10002 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 10002 & -2 \\ -2 & 10004 \end{bmatrix}}{10004 \cdot 10002 - 2 \cdot 2} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} \overbrace{10^4+2}^{\approx 10^4} & -2 \\ -2 & \underbrace{10^4+4}_{\approx 10^4} \end{bmatrix}}{\underbrace{(10^4+4) \cdot (10^4+2) - 2^2}_{\approx 10^8}} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \frac{1}{10^8} \cdot \begin{bmatrix} 10^4 & -2 \\ -2 & 10^4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \end{aligned} \quad (8.377)$$

$$= -\frac{1}{10^8} \cdot \begin{bmatrix} 10^4 + 2 \\ -2 - 10^4 \end{bmatrix} \approx -\frac{1}{10^8} \cdot \begin{bmatrix} 10^4 \\ -10^4 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \cdot 10^{-4} = \underbrace{\begin{bmatrix} -10^{-4} \\ 10^{-4} \end{bmatrix}}_{\underline{x}_1}$$

Funkcija v  $\underline{x}_1$  zavzame vrednost:

$$\begin{aligned} f(\underline{x}_1) &= (-10^{-4}) - 10^{-4} + 2 \cdot (-10^{-4})^2 + 2 \cdot (-10^{-4}) \cdot 10^{-4} + 10^{-8} = \\ &= -2 \cdot 10^{-4} + 2 \cdot 10^{-8} - 2 \cdot 10^{-8} + 10^{-8} \approx -2 \cdot 10^{-4} \end{aligned} \quad (8.378)$$

Ker je  $f(\underline{x}_1) < f(\underline{x}_0)$ , sledi:

$$\alpha_1 = C_1 \cdot \alpha_0 = \frac{1}{4} \cdot 10^4 = 2500 \quad (8.379)$$

2. iteracija;  $k = 1$ :

$$f(\underline{x}_1) = -2 \cdot 10^{-4} \quad (8.380)$$

Gradient je enak:

$$\nabla \underline{f}(\underline{x}_1) = \begin{bmatrix} 1 + 4 \cdot (-10^{-4}) + 2 \cdot 10^{-4} \\ -1 + 2 \cdot (-10^{-4}) + 2 \cdot 10^{-4} \end{bmatrix} = \begin{bmatrix} 1 - 2 \cdot 10^{-4} \\ -1 \end{bmatrix} = \begin{bmatrix} 0.9998 \\ -1 \end{bmatrix} \quad (8.381)$$

Sledi:

$$\|\nabla \underline{f}(\underline{x}_1)\| = \sqrt{0.9998^2 + (-1)^2} = 1.4141 > \varepsilon \quad (8.382)$$

Dobimo:

$$\begin{aligned} \underline{x}_2 &= \underline{x}_1 - [\underline{H} + \alpha_1 \cdot \underline{I}]^{-1} \cdot \nabla \underline{f}(\underline{x}_1) = \\ &= \begin{bmatrix} -10^{-4} \\ 10^{-4} \end{bmatrix} - \left[ \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} + 2500 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right]^{-1} \cdot \begin{bmatrix} 0.9998 \\ -1 \end{bmatrix} = \\ &= \begin{bmatrix} -10^{-4} \\ 10^{-4} \end{bmatrix} - \begin{bmatrix} 2504 & 2 \\ 2 & 2502 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0.9998 \\ -1 \end{bmatrix} = \\ &= \begin{bmatrix} -10^{-4} \\ 10^{-4} \end{bmatrix} - \frac{1}{2504 \cdot 2502 - 2 \cdot 2} \cdot \begin{bmatrix} 2502 & -2 \\ -2 & 2504 \end{bmatrix} \cdot \begin{bmatrix} 0.9998 \\ -1 \end{bmatrix} = \end{aligned} \quad (8.383)$$

$$\begin{aligned}
&= \begin{bmatrix} -10^{-4} \\ 10^{-4} \end{bmatrix} - \frac{1}{6265004} \cdot \begin{bmatrix} 2502 \cdot 0.9998 + 2 \\ -2 \cdot 0.9998 - 2504 \end{bmatrix} = \\
&= \begin{bmatrix} -10^{-4} \\ 10^{-4} \end{bmatrix} - \frac{1}{6265004} \cdot \begin{bmatrix} 2503.4996 \\ -2505.999 \end{bmatrix} = \\
&= \begin{bmatrix} -10^{-4} \\ 10^{-4} \end{bmatrix} - \underbrace{\begin{bmatrix} 4 \cdot 10^{-4} \\ -4 \cdot 10^{-4} \end{bmatrix}}_{\hat{x}_2} = \begin{bmatrix} -5 \cdot 10^{-4} \\ 5 \cdot 10^{-4} \end{bmatrix}
\end{aligned}$$

Tako smo izračunali rezultat po 2. iteracijah. Algoritem pa bi postopek nadaljeval.

Opomba: Sicer v tem primeru bistveno počasnejša konvergenca kot pri Newtonovi metodi (glej Primer 1). Vendar se v praksi pri čudnih funkcijah mnogokrat izkaže, da Newtonova metoda divergira, Maquardtova pa lepo konvergira.

## 8.4 Primeri multidimenzionalne numerične optimizacije s sofisticiranimi metodami

V nadaljevanju si bomo pogledali nekaj primerov uporabe sofisticiranih metod multidimenzionalne numerične optimizacije, pri čemer se bomo omejili na:

- Powellovo metodo,
- Metodo konjugiranih gradientov, ter
- Quasi-Newtonove metode.

### 8.4.1 Powellova metoda

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.3.2.1 ter izraze (6.104) do (6.165).

#### Primer 1:

S pomočjo Powellove metode določite minimum za funkcijo:

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 \quad (8.399)$$

pri tem vzemite  $\underline{x}_0 = [0 \ 0]^T$ .

1.) Najprej opravimo premik iz začetnega vektorja v smeri enotskega vektorja  $\underline{e}_1$ :

$$\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \tilde{\Theta}_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{\Theta}_1 \\ 0 \end{bmatrix} \quad (8.400)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_1) &= \tilde{\Theta}_1 - 0 + 2 \cdot \tilde{\Theta}_1^2 + 0^2 + 0^2 = \tilde{\Theta}_1 + 2\tilde{\Theta}_1^2 = f(\tilde{\Theta}_1) \\ \frac{df(\underline{x}_1)}{d\tilde{\Theta}_1} &= 1 + 4\tilde{\Theta}_1 = 0 \quad \Rightarrow \tilde{\Theta}_1^* = -\frac{1}{4} \quad \Rightarrow \underline{x}_1^* = \begin{bmatrix} -\frac{1}{4} \\ 0 \end{bmatrix} \end{aligned} \quad (8.401)$$

2.) Postopek ponovimo s premikom v smeri enotskega vektorja  $\underline{e}_2$ :

$$\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2 = \begin{bmatrix} -\frac{1}{4} \\ 0 \end{bmatrix} + \tilde{\Theta}_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} \\ \tilde{\Theta}_2 \end{bmatrix} \quad (8.402)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_2) &= \left(-\frac{1}{4}\right) - \tilde{\Theta}_2 + 2 \cdot \left(-\frac{1}{4}\right)^2 + 2 \cdot \left(-\frac{1}{4}\right) \cdot \tilde{\Theta}_2 + \tilde{\Theta}_2^2 = \\ &= \tilde{\Theta}_2^2 - \tilde{\Theta}_2 - \frac{1}{2}\tilde{\Theta}_2 - \frac{1}{4} + \frac{1}{8} = \tilde{\Theta}_2^2 - \frac{3}{2}\tilde{\Theta}_2 - \frac{1}{8} \\ \frac{df}{d\tilde{\Theta}_2} &= 2\tilde{\Theta}_2 - \frac{3}{2} = 0 \quad \Rightarrow \tilde{\Theta}_2^* = \frac{3}{4} \quad \Rightarrow \underline{x}_2^* = \begin{bmatrix} -\frac{1}{4} \\ \frac{3}{4} \end{bmatrix} \end{aligned} \quad (8.403)$$

3.) Sledi premik v smeri  $\underline{x}_2 - \underline{x}_0$ :

$$\begin{aligned} \underline{x}_3 &= \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0) = \begin{bmatrix} -\frac{1}{4} \\ \frac{3}{4} \end{bmatrix} + \tilde{\Theta}_3 \cdot \left( \begin{bmatrix} -\frac{1}{4} \\ \frac{3}{4} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} -\frac{1}{4} - \frac{1}{4}\tilde{\Theta}_3 \\ \frac{3}{4} + \frac{3}{4}\tilde{\Theta}_3 \end{bmatrix} = \\ &= \begin{bmatrix} -\frac{1}{4} \cdot (1 + \tilde{\Theta}_3) \\ \frac{3}{4} \cdot (1 + \tilde{\Theta}_3) \end{bmatrix} \end{aligned} \quad (8.404)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_3) &= -\frac{1}{4} \cdot (1 + \tilde{\Theta}_3) - \frac{3}{4} \cdot (1 + \tilde{\Theta}_3) + 2 \cdot \left(-\frac{1}{4}\right)^2 \cdot (1 + \tilde{\Theta}_3)^2 + \\ &+ 2 \cdot \left(-\frac{1}{4}\right) \cdot \frac{3}{4} \cdot (1 + \tilde{\Theta}_3)^2 + \left(\frac{3}{4}\right)^2 \cdot (1 + \tilde{\Theta}_3)^2 = \\ &= -(1 + \tilde{\Theta}_3) + \frac{1}{8} \cdot (1 + \tilde{\Theta}_3)^2 - \frac{3}{8} \cdot (1 + \tilde{\Theta}_3)^2 + \frac{9}{16} \cdot (1 + \tilde{\Theta}_3)^2 = \\ &= -1 - \tilde{\Theta}_3 + \frac{5}{16} \cdot (1 + \tilde{\Theta}_3)^2 \end{aligned} \quad (8.405)$$

Odvajamo po  $\tilde{\Theta}_3$  in dobimo:

$$\begin{aligned} \frac{df}{d\tilde{\Theta}_3} &= -1 + \frac{5}{16} \cdot 2 \cdot (1 + \tilde{\Theta}_3) = -1 + \frac{5}{8} \cdot (1 + \tilde{\Theta}_3) = -\frac{3}{8} + \frac{5}{8}\tilde{\Theta}_3 = 0 \\ \Rightarrow \tilde{\Theta}_3^* &= \frac{3}{5} \quad \Rightarrow \underline{x}_3^* = \begin{bmatrix} -\frac{1}{4} \cdot \left(1 + \frac{3}{5}\right) \\ \frac{3}{4} \cdot \left(1 + \frac{3}{5}\right) \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \\ \frac{6}{5} \end{bmatrix} = \begin{bmatrix} -0.4 \\ 1.2 \end{bmatrix} \end{aligned} \quad (8.406)$$

4.) Sledi ponovni premik v smeri enotskega vektorja  $\underline{e}_2$ :

$$\underline{x}_4 = \underline{x}_3 + \tilde{\Theta}_4 \cdot \underline{e}_2 = \begin{bmatrix} -0.4 \\ 1.2 \end{bmatrix} + \tilde{\Theta}_4 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.4 \\ 1.2 + \tilde{\Theta}_4 \end{bmatrix} \quad (8.407)$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_4) &= -0.4 - (1.2 + \tilde{\Theta}_4) + 2 \cdot (-0.4)^2 + 2 \cdot (-0.4) \cdot (1.2 + \tilde{\Theta}_4) + (1.2 + \tilde{\Theta}_4)^2 = \\
&= -0.4 - 1.2 - \tilde{\Theta}_4 + 0.64 - 0.96 - 0.8\tilde{\Theta}_4 + (1.2 + \tilde{\Theta}_4)^2 = \\
&= -1.92 - 1.8\tilde{\Theta}_4 + (1.2 + \tilde{\Theta}_4)^2
\end{aligned} \tag{8.408}$$

Odvajamo po  $\tilde{\Theta}_4$  in dobimo:

$$\begin{aligned}
\frac{df}{d\tilde{\Theta}_4} &= -1.8 + 2 \cdot (1.2 + \tilde{\Theta}_4) = -1.8 + 2.4 + 2\tilde{\Theta}_4 = 0 \\
\Rightarrow 2\tilde{\Theta}_4 &= -0.6 \quad \Rightarrow \tilde{\Theta}_4^* = -0.3 \\
\Rightarrow \underline{x}_4^* &= \begin{bmatrix} -0.4 \\ 0.9 \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \\ \frac{9}{10} \end{bmatrix}
\end{aligned} \tag{8.409}$$

5.) Sledi premik v smeri  $\underline{x}_3 - \underline{x}_0$ :

$$\underline{x}_5 = \underline{x}_4 + \tilde{\Theta}_5 \cdot (\underline{x}_3 - \underline{x}_0) = \begin{bmatrix} -\frac{2}{5} \\ \frac{9}{10} \end{bmatrix} + \tilde{\Theta}_5 \cdot \begin{bmatrix} -\frac{2}{5} \\ \frac{6}{5} \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} - \frac{2}{5}\tilde{\Theta}_5 \\ \frac{9}{10} + \frac{6}{5}\tilde{\Theta}_5 \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \cdot (1 + \tilde{\Theta}_5) \\ \frac{3}{10} \cdot (3 + 4\tilde{\Theta}_5) \end{bmatrix} \tag{8.410}$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_5) &= -\frac{2}{5} \cdot (1 + \tilde{\Theta}_5) - \frac{3}{10} \cdot (3 + 4\tilde{\Theta}_5) + 2 \cdot \left( -\frac{2}{5} \cdot (1 + \tilde{\Theta}_5) \right)^2 + \\
&\quad + 2 \cdot \left( -\frac{2}{5} \cdot (1 + \tilde{\Theta}_5) \right) \cdot \frac{3}{10} \cdot (3 + 4\tilde{\Theta}_5) + \left( \frac{3}{10} \cdot (3 + 4\tilde{\Theta}_5) \right)^2 = \\
&= -\frac{2}{5} - \frac{2}{5}\tilde{\Theta}_5 - \frac{9}{10} - \frac{12}{10}\tilde{\Theta}_5 + \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + 2 \cdot \left( -\frac{2}{5} - \frac{2}{5}\tilde{\Theta}_5 \right) \cdot \left( \frac{9}{10} + \frac{12}{10}\tilde{\Theta}_5 \right) + \\
&\quad + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 = \\
&= -\frac{13}{10} - \frac{8}{5}\tilde{\Theta}_5 + 2 \cdot \left( -\frac{18}{50} - \frac{24}{50}\tilde{\Theta}_5 - \frac{18}{50}\tilde{\Theta}_5 - \frac{24}{50}\tilde{\Theta}_5^2 \right) + \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \\
&\quad + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 =
\end{aligned} \tag{8.411}$$

$$\begin{aligned}
&= -\frac{13}{10} - \frac{8}{5} \tilde{\Theta}_5 + 2 \cdot \left( -\frac{9}{25} - \frac{21}{25} \tilde{\Theta}_5 - \frac{12}{25} \tilde{\Theta}_5^2 \right) + \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 = \\
&= \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 - \frac{8}{5} \tilde{\Theta}_5 - \frac{13}{10} - \frac{24}{25} \tilde{\Theta}_5^2 - \frac{42}{25} \tilde{\Theta}_5 - \frac{18}{25} = \\
&= \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 - \frac{24}{25} \tilde{\Theta}_5^2 - \frac{82}{25} \tilde{\Theta}_5 - \frac{101}{50} = f(\tilde{\Theta}_5)
\end{aligned}$$

Sledi:

$$\begin{aligned}
f(x_5) &= -\frac{2}{5} \cdot (1 + \tilde{\Theta}_5) - \frac{3}{10} \cdot (3 + 4\tilde{\Theta}_5) + 2 \cdot \left( -\frac{2}{5} \cdot (1 + \tilde{\Theta}_5) \right)^2 + \\
&\quad + 2 \cdot \left( -\frac{2}{5} \cdot (1 + \tilde{\Theta}_5) \right) \cdot \frac{3}{10} \cdot (3 + 4\tilde{\Theta}_5) + \left( \frac{3}{10} \cdot (3 + 4\tilde{\Theta}_5) \right)^2 = \\
&= -\frac{2}{5} - \frac{2}{5} \tilde{\Theta}_5 - \frac{9}{10} - \frac{12}{10} \tilde{\Theta}_5 + \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \\
&\quad + 2 \cdot \left( -\frac{2}{5} - \frac{2}{5} \tilde{\Theta}_5 \right) \cdot \left( \frac{9}{10} + \frac{12}{10} \tilde{\Theta}_5 \right) + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 = \\
&= -\frac{13}{10} - \frac{8}{5} \tilde{\Theta}_5 + 2 \cdot \left( -\frac{18}{50} - \frac{24}{50} \tilde{\Theta}_5 - \frac{18}{50} \tilde{\Theta}_5 - \frac{24}{50} \tilde{\Theta}_5^2 \right) + \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \\
&\quad + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 = \\
&= -\frac{13}{10} - \frac{8}{5} \tilde{\Theta}_5 + 2 \cdot \left( -\frac{9}{25} - \frac{21}{25} \tilde{\Theta}_5 - \frac{12}{25} \tilde{\Theta}_5^2 \right) + \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \\
&\quad + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 = \\
&= \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 - \frac{8}{5} \tilde{\Theta}_5 - \frac{13}{10} - \frac{24}{25} \tilde{\Theta}_5^2 - \frac{42}{25} \tilde{\Theta}_5 - \frac{18}{25} = \\
&= \frac{8}{25} \cdot (1 + \tilde{\Theta}_5)^2 + \frac{9}{100} \cdot (3 + 4\tilde{\Theta}_5)^2 - \frac{24}{25} \tilde{\Theta}_5^2 - \frac{82}{25} \tilde{\Theta}_5 - \frac{101}{50} = f(\tilde{\Theta}_5)
\end{aligned} \tag{8.412}$$

Odvajamo po  $\tilde{\Theta}_5$  in dobimo:

$$\begin{aligned}
\frac{df}{d\tilde{\Theta}_5} &= \frac{8}{25} \cdot 2 \cdot (1 + \tilde{\Theta}_5) + \frac{9}{100} \cdot 2 \cdot 4 \cdot (3 + 4\tilde{\Theta}_5) - 2 \cdot \frac{24}{25} \tilde{\Theta}_5 - \frac{82}{25} = \\
&= \frac{16}{25} \cdot (1 + \tilde{\Theta}_5) + \frac{18}{25} \cdot (3 + 4\tilde{\Theta}_5) - \frac{48}{25} \tilde{\Theta}_5 - \frac{82}{25} = \\
&= \frac{16}{25} + \frac{16}{25} \tilde{\Theta}_5 + \frac{54}{25} + \frac{72}{25} \tilde{\Theta}_5 - \frac{48}{25} \tilde{\Theta}_5 - \frac{82}{25} =
\end{aligned} \tag{8.413}$$



$$\begin{aligned}
&= \frac{40}{25} \tilde{\Theta}_5 - \frac{12}{25} = 0 \quad | \cdot 25 \Rightarrow 40\tilde{\Theta}_5 - 12 = 0 \Rightarrow \tilde{\Theta}_5^* = \frac{12}{40} = \frac{3}{10} = 0.3 \\
\Rightarrow \underline{x_5}^* &= \begin{bmatrix} -\frac{2}{5} \cdot \left(1 + \frac{3}{10}\right) \\ \frac{3}{10} \cdot \left(3 + 4 \cdot \frac{3}{10}\right) \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} - \frac{6}{50} \\ \frac{9}{10} + \frac{36}{100} \end{bmatrix} = \begin{bmatrix} -\frac{10}{25} - \frac{3}{25} \\ \frac{45}{50} + \frac{18}{50} \end{bmatrix} = \begin{bmatrix} -\frac{13}{25} \\ \frac{63}{50} \end{bmatrix}
\end{aligned}$$

6.) Sledi premik v smeri  $\underline{x_5} - \underline{x_3}$ :

$$\begin{aligned}
\underline{x_6} = \underline{x}^* &= \underline{x_3} + \tilde{\Theta}_6 \cdot (\underline{x_5} - \underline{x_3}) = \begin{bmatrix} -\frac{2}{5} \\ \frac{6}{5} \end{bmatrix} + \tilde{\Theta}_6 \cdot \begin{bmatrix} -\frac{13}{25} + \frac{2}{5} \\ \frac{63}{50} - \frac{6}{5} \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} - \frac{3}{25} \tilde{\Theta}_6 \\ \frac{6}{5} + \frac{3}{50} \tilde{\Theta}_6 \end{bmatrix} \\
&= \begin{bmatrix} -\frac{1}{25} \cdot (10 + 3\tilde{\Theta}_6) \\ \frac{3}{50} \cdot (20 + \tilde{\Theta}_6) \end{bmatrix}
\end{aligned} \tag{8.414}$$

Dobimo:

$$\begin{aligned}
f(\underline{x}^*) &= -\frac{1}{25} \cdot (10 + 3\tilde{\Theta}_6) - \frac{3}{50} \cdot (20 + \tilde{\Theta}_6) + 2 \cdot \left(-\frac{1}{25} \cdot (10 + 3\tilde{\Theta}_6)\right)^2 + \\
&\quad + 2 \cdot \left(-\frac{1}{25} \cdot (10 + 3\tilde{\Theta}_6)\right) \cdot \left(\frac{3}{50} \cdot (20 + \tilde{\Theta}_6)\right) + \left(\frac{3}{50} \cdot (20 + \tilde{\Theta}_6)\right)^2 = \\
&= -\frac{10}{25} - \frac{3}{25} \tilde{\Theta}_6 - \frac{60}{50} - \frac{3}{50} \tilde{\Theta}_6 + 2 \cdot \frac{1}{625} \cdot (10 + 3\tilde{\Theta}_6)^2 + \\
&\quad + 2 \cdot \left(-\frac{10}{25} - \frac{3}{25} \tilde{\Theta}_6\right) \cdot \left(\frac{60}{50} + \frac{3}{50} \tilde{\Theta}_6\right) + \frac{9}{2500} \cdot (20 + \tilde{\Theta}_6)^2 = \\
&= -\frac{80}{50} - \frac{9}{50} \tilde{\Theta}_6 + \frac{2}{625} \cdot (10 + 3\tilde{\Theta}_6)^2 + \frac{9}{2500} \cdot (20 + \tilde{\Theta}_6)^2 + \\
&\quad + 2 \cdot \left(-\frac{600}{1250} - \frac{30}{1250} \tilde{\Theta}_6 - \frac{180}{1250} \tilde{\Theta}_6 - \frac{9}{1250} \tilde{\Theta}_6^2\right) = \\
&= -\frac{80}{50} - \frac{9}{50} \tilde{\Theta}_6 + \frac{2}{625} \cdot (10 + 3\tilde{\Theta}_6)^2 + \frac{9}{2500} \cdot (20 + \tilde{\Theta}_6)^2 - \\
&\quad - \frac{120}{125} - \frac{42}{125} \tilde{\Theta}_6 - \frac{18}{1250} \tilde{\Theta}_6^2 = \\
&= \frac{2}{625} \cdot (10 + 3\tilde{\Theta}_6)^2 + \frac{9}{2500} \cdot (20 + \tilde{\Theta}_6)^2 - \frac{18}{1250} \tilde{\Theta}_6^2 - \frac{129}{250} \tilde{\Theta}_6 - \frac{640}{250} = f(\tilde{\Theta}_6)
\end{aligned} \tag{8.415}$$

Odvajamo po  $\tilde{\Theta}_6$ :

$$\begin{aligned}
 \frac{df}{d\tilde{\Theta}_6} &= \frac{2}{625} \cdot 2 \cdot 3 \cdot (10 + 3\tilde{\Theta}_6) + \frac{9}{2500} \cdot 2 \cdot (20 + \tilde{\Theta}_6) - \frac{18 \cdot 2}{1250} \tilde{\Theta}_6 - \frac{129}{250} = \\
 &= \frac{120}{625} + \frac{36}{625} \tilde{\Theta}_6 + \frac{360}{2500} + \frac{18}{2500} \tilde{\Theta}_6 - \frac{36}{1250} \tilde{\Theta}_6 - \frac{129}{250} = \\
 &= \frac{45}{1250} \tilde{\Theta}_6 - \frac{45}{250} = 0 \quad \Bigg| \cdot 1250 \\
 &\Rightarrow 45\tilde{\Theta}_6 - 225 = 0 \quad \Rightarrow \tilde{\Theta}_6^* = 5 \\
 &\Rightarrow \underline{x}^* = \underline{x}_6^* = \begin{bmatrix} -\frac{1}{25} \cdot (10 + 3 \cdot 5) \\ \frac{3}{50} \cdot (20 + 5) \end{bmatrix} = \begin{bmatrix} -\frac{1}{25} \cdot 25 \\ \frac{75}{50} \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}
 \end{aligned} \tag{8.416}$$

Torej v 6. koraku dobimo minimum v točki  $\underline{x}^* = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}^T$ .

### Primer 2:

S pomočjo Powellove metode poiščite minimum za naslednjo funkcijo:

$$f(x_1, x_2) = 3 + (x_1 - 1.5x_2)^2 + (x_2 - 2)^2 \tag{8.417}$$

pri čemer je  $\underline{x}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$ .

1.) Opravimo premik iz začetnega vektorja v smeri enotskega vektorja  $\underline{e}_1$ :

$$\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} + \tilde{\Theta}_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 + \tilde{\Theta}_1 \\ 0.5 \end{bmatrix} \tag{8.418}$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_1) &= 3 + (0.5 + \tilde{\Theta}_1 - 1.5 \cdot 0.5)^2 + (0.5 - 2)^2 \\
\frac{df}{d\tilde{\Theta}_1} &= 2 \cdot (0.5 + \tilde{\Theta}_1 - 1.5 \cdot 0.5) = 0 \quad \Rightarrow \quad \tilde{\Theta}_1^* = 1.5 \cdot 0.5 - 0.5 = \\
&= 0.5 \cdot (1.5 - 1) = 0.5^2 = \frac{1}{4} \quad (8.419) \\
\Rightarrow \underline{x}_1^* &= \begin{bmatrix} 0.5 + \frac{1}{4} \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.5 \end{bmatrix}
\end{aligned}$$

2.) Postopek ponovimo s premikom v smeri enotskega vektorja  $\underline{e}_2$ :

$$\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2 = \begin{bmatrix} 0.75 \\ 0.5 \end{bmatrix} + \tilde{\Theta}_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.5 + \tilde{\Theta}_2 \end{bmatrix} \quad (8.420)$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_2) &= 3 + \{0.75 - 1.5 \cdot (0.5 + \tilde{\Theta}_2)\}^2 + \{0.5 + \tilde{\Theta}_2 - 2\}^2 \\
\frac{df}{d\tilde{\Theta}_2} &= 2 \cdot \{0.75 - 1.5 \cdot (0.5 + \tilde{\Theta}_2)\} \cdot (-1.5) + 2 \cdot \{0.5 + \tilde{\Theta}_2 - 2\} = \\
&= -3 \cdot \{-1.5\tilde{\Theta}_2\} + 2 \cdot \{\tilde{\Theta}_2 - 1.5\} = 0 \quad (8.421) \\
\Rightarrow 4.5\tilde{\Theta}_2 + 2\tilde{\Theta}_2 &= 3 \quad \Rightarrow \quad \tilde{\Theta}_2^* = \frac{3}{6.5} = 0.4615 \\
\Rightarrow \underline{x}_2^* &= \begin{bmatrix} 0.75 \\ 0.5 + 0.4615 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.9615 \end{bmatrix}
\end{aligned}$$

3.) Sledi premik v smeri  $\underline{x}_2 - \underline{x}_0$ :

$$\begin{aligned}
\underline{x}_3 &= \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0) = \begin{bmatrix} 0.75 \\ 0.9615 \end{bmatrix} + \tilde{\Theta}_3 \cdot \left( \begin{bmatrix} 0.75 \\ 0.9615 \end{bmatrix} - \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \right) = \\
&= \begin{bmatrix} 0.75 + 0.25\tilde{\Theta}_3 \\ 0.9615 + 0.4615\tilde{\Theta}_3 \end{bmatrix} \quad (8.422)
\end{aligned}$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_3) &= 3 + \left(0.75 + 0.25\tilde{\Theta}_3 - 1.5 \cdot \{0.9615 + 0.4615\tilde{\Theta}_3\}\right)^2 + \\
&\quad + \left(0.9615 + 0.4615\tilde{\Theta}_3 - 2\right)^2 = \\
&= 3 + \left(0.75 + 0.25\tilde{\Theta}_3 - 1.44225 - 0.69225\tilde{\Theta}_3\right)^2 + \left(0.4615\tilde{\Theta}_3 - 1.0385\right)^2 = \quad (8.423) \\
&= 3 + \left(-0.69225 - 0.44225\tilde{\Theta}_3\right)^2 + \left(0.4615\tilde{\Theta}_3 - 1.0385\right)^2 = \\
&= 3 + \left(0.69225 + 0.44225\tilde{\Theta}_3\right)^2 + \left(0.4615\tilde{\Theta}_3 - 1.0385\right)^2
\end{aligned}$$

Odvajamo po  $\tilde{\Theta}_3$ :

$$\begin{aligned}
\frac{df}{d\tilde{\Theta}_3} &= 2 \cdot \left(0.69225 + 0.44225\tilde{\Theta}_3\right) \cdot 0.44225 + \\
&\quad + 2 \cdot \left(0.4615\tilde{\Theta}_3 - 1.0385\right) \cdot 0.4615 = 0 \quad (8.424)
\end{aligned}$$

Sledi:

$$\begin{aligned}
0.3061 + 0.1955\tilde{\Theta}_3 + 0.2129\tilde{\Theta}_3 - 0.4792 &= 0 \\
0.4084\tilde{\Theta}_3 = 0.1731 &\Rightarrow \tilde{\Theta}_3^* = 0.4238 \quad (8.425)
\end{aligned}$$

Dobimo:

$$\underline{x}_3^* = \begin{bmatrix} 0.75 + 0.4238 \cdot 0.25 \\ 0.9615 + 0.4238 \cdot 0.4615 \end{bmatrix} = \begin{bmatrix} 0.8559 \\ 1.1570 \end{bmatrix} \quad (8.426)$$

4.) Sledi ponovni premik v smeri enotskega vektorja  $\underline{e}_2$ :

$$\underline{x}_4 = \underline{x}_3 + \tilde{\Theta}_4 \cdot \underline{e}_2 = \begin{bmatrix} 0.8559 \\ 1.1570 \end{bmatrix} + \tilde{\Theta}_4 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.8559 \\ 1.1570 + \tilde{\Theta}_4 \end{bmatrix} \quad (8.427)$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_4) &= 3 + \left(0.8559 - 1.5 \cdot (1.1570 + \tilde{\Theta}_4)\right)^2 + \left(1.1570 + \tilde{\Theta}_4 - 2\right)^2 = \\
&= 3 + \left(-0.8796 - 1.5\tilde{\Theta}_4\right)^2 + \left(-0.843 + \tilde{\Theta}_4\right)^2 = \quad (8.428) \\
&= 3 + \left(0.8796 + 1.5\tilde{\Theta}_4\right)^2 + \left(\tilde{\Theta}_4 - 0.843\right)^2
\end{aligned}$$

Odvajamo po  $\tilde{\Theta}_4$ :

$$\begin{aligned}
\frac{df}{d\tilde{\Theta}_4} &= 2 \cdot (0.8796 + 1.5\tilde{\Theta}_4) \cdot 1.5 + 2 \cdot (\tilde{\Theta}_4 - 0.843) = 0 \\
1.3194 + 2.25\tilde{\Theta}_4 + \tilde{\Theta}_4 - 0.843 &= 0 \\
3.25\tilde{\Theta}_4 &= -0.4764 \\
\Rightarrow \tilde{\Theta}_4^* &= -0.1465 \\
\Rightarrow \underline{x}_4 &= \begin{bmatrix} 0.8559 \\ 1.1570 - 0.1465 \end{bmatrix} = \begin{bmatrix} 0.8559 \\ 1.0104 \end{bmatrix}
\end{aligned} \tag{8.429}$$

5.) Sledi premik v smeri  $\underline{x}_3 - \underline{x}_0$ :

$$\begin{aligned}
\underline{x}_5 &= \underline{x}_4 + \tilde{\Theta}_5 \cdot (\underline{x}_3 - \underline{x}_0) = \begin{bmatrix} 0.8559 \\ 1.0104 \end{bmatrix} + \tilde{\Theta}_5 \cdot \left\{ \begin{bmatrix} 0.8559 \\ 1.1570 \end{bmatrix} - \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \right\} = \\
&= \begin{bmatrix} 0.8559 + 0.3559\tilde{\Theta}_5 \\ 1.0104 + 0.6570\tilde{\Theta}_5 \end{bmatrix}
\end{aligned} \tag{8.430}$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_5) &= 3 + (0.8559 + 0.3559\tilde{\Theta}_5 - 1.5 \cdot (1.0104 + 0.6570\tilde{\Theta}_5))^2 + \\
&\quad + (1.0104 + 0.6570\tilde{\Theta}_5 - 2)^2 = \\
&= 3 + (-0.6597 - 0.6296\tilde{\Theta}_5)^2 + (0.6570\tilde{\Theta}_5 - 0.9896)^2 = \\
&= 3 + (0.6597 + 0.6296\tilde{\Theta}_5)^2 + (0.6570\tilde{\Theta}_5 - 0.9896)^2
\end{aligned} \tag{8.431}$$

Odvajamo po  $\tilde{\Theta}_5$ :

$$\begin{aligned}
\frac{df}{d\tilde{\Theta}_5} &= 2 \cdot (0.6597 + 0.6296\tilde{\Theta}_5) \cdot 0.6296 + 2 \cdot (0.6570\tilde{\Theta}_5 - 0.9896) \cdot 0.6570 = 0 \\
0.4153 + 0.3963\tilde{\Theta}_5 + 0.4316\tilde{\Theta}_5 - 0.6501 &= 0 \\
0.8279\tilde{\Theta}_5 &= 0.2348 \\
\Rightarrow \tilde{\Theta}_5^* &= \frac{0.2348}{0.8279} = 0.2836 \\
\Rightarrow \underline{x}_5^* &= \begin{bmatrix} 0.8559 + 0.2836 \cdot 0.3559 \\ 1.0104 + 0.2836 \cdot 0.6570 \end{bmatrix} = \begin{bmatrix} 0.9568 \\ 1.196 \end{bmatrix}
\end{aligned} \tag{8.432}$$

6.) Sledi premik v smeri  $\underline{x}_5 - \underline{x}_3$ :

$$\begin{aligned}\underline{x}_6 = \underline{x}^* &= \underline{x}_3 + \tilde{\Theta}_6 \cdot (\underline{x}_5 - \underline{x}_3) = \begin{bmatrix} 0.8559 \\ 1.1570 \end{bmatrix} + \tilde{\Theta}_6 \cdot \left\{ \begin{bmatrix} 0.9568 \\ 1.196 \end{bmatrix} - \begin{bmatrix} 0.8559 \\ 1.1570 \end{bmatrix} \right\} = \\ &= \begin{bmatrix} 0.8559 + 0.1009\tilde{\Theta}_6 \\ 1.1570 + 0.039\tilde{\Theta}_6 \end{bmatrix}\end{aligned}\quad (8.433)$$

Dobimo:

$$\begin{aligned}f(\underline{x}^*) &= 3 + (0.8559 + 0.1009\tilde{\Theta}_6 - 1.5 \cdot (1.15 + 0.039\tilde{\Theta}_6))^2 + \\ &\quad + (1.1570 + 0.039\tilde{\Theta}_6 - 2)^2 = \\ &= 3 + (-0.8691 + 0.0424\tilde{\Theta}_6)^2 + (0.039\tilde{\Theta}_6 - 0.843)^2\end{aligned}\quad (8.434)$$

Odvajamo po  $\tilde{\Theta}_6$ :

$$\begin{aligned}\frac{df}{d\tilde{\Theta}_6} &= 2 \cdot (-0.8691 + 0.0424\tilde{\Theta}_6) \cdot 0.0424 + 2 \cdot (0.039\tilde{\Theta}_6 - 0.843) \cdot 0.039 = 0 \\ &- 0.0697 + 1.7977 \cdot 10^{-3} \tilde{\Theta}_6 + 1.521 \cdot 10^{-3} \tilde{\Theta}_6 = 0 \\ \Rightarrow \tilde{\Theta}_6^* &= \frac{0.0697}{3.3187} \cdot 10^3 = 21.002 \\ \Rightarrow \underline{x}^* &= \begin{bmatrix} 0.8559 + 21.002 \cdot 0.1009 \\ 1.1570 + 21.002 \cdot 0.039 \end{bmatrix} = \begin{bmatrix} 2.9750 \\ 1.9760 \end{bmatrix}\end{aligned}\quad (8.435)$$

Pravi rezultat  $\underline{x}^* = [3 \quad 2]^T$ . Do odstopanja je prišlo zaradi napak pri zaokroževanju.

### Primer 3:

S pomočjo Powellove metode določite minimum za funkcijo:

$$f(x_1, x_2) = x_1^2 + x_2^2 - x_1 - x_2 \quad (8.436)$$

kjer je  $\underline{x}_0 = [0 \quad 0]^T$ .

1.) Najprej opravimo premik iz začetnega vektorja v smeri enotskega vektorja  $\underline{e}_1$ :

$$\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \tilde{\Theta}_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{\Theta}_1 \\ 0 \end{bmatrix} \quad (8.437)$$

Dobimo:

$$f(\underline{x}_1) = \tilde{\Theta}_1^2 + 0 - \tilde{\Theta}_1 - 0$$
$$\frac{df}{d\tilde{\Theta}_1} = 2\tilde{\Theta}_1 - 1 = 0 \quad \Rightarrow \quad \tilde{\Theta}_1^* = \frac{1}{2} \quad \Rightarrow \quad \underline{x}_1^* = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} \quad (8.438)$$

2.) Postopek ponovimo s premikom v smeri enotskega vektorja  $\underline{e}_2$ :

$$\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2 = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} + \tilde{\Theta}_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \tilde{\Theta}_2 \end{bmatrix} \quad (8.439)$$

Dobimo:

$$f(\underline{x}_2) = \frac{1}{4} + \tilde{\Theta}_2^2 - \frac{1}{2} - \tilde{\Theta}_2$$
$$\frac{df}{d\tilde{\Theta}_2} = 2\tilde{\Theta}_2 - 1 = 0 \quad \Rightarrow \quad \tilde{\Theta}_2^* = \frac{1}{2} \quad \Rightarrow \quad \underline{x}_2^* = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad (8.440)$$

3.) Sledi premik v smeri  $\underline{x}_2 - \underline{x}_0$ :

$$\underline{x}_3 = \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0) = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} + \tilde{\Theta}_3 \cdot \left( \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}\tilde{\Theta}_3 \\ \frac{1}{2} + \frac{1}{2}\tilde{\Theta}_3 \end{bmatrix} \quad (8.441)$$

Dobimo:

$$\begin{aligned}
f(\underline{x}_3) &= \left(\frac{1}{2} + \frac{1}{2}\tilde{\Theta}_3\right)^2 + \left(\frac{1}{2} + \frac{1}{2}\tilde{\Theta}_3\right)^2 - \left(\frac{1}{2} + \frac{\tilde{\Theta}_3}{2}\right) - \left(\frac{1}{2} + \frac{\tilde{\Theta}_3}{2}\right) = \\
&= 2 \cdot \left(\frac{1}{2} + \frac{1}{2}\tilde{\Theta}_3\right)^2 - 2 \cdot \left(\frac{1}{2} + \frac{\tilde{\Theta}_3}{2}\right) = \\
&= 2 \cdot \left[\frac{1}{4} + 2 \cdot \frac{1}{2} \cdot \frac{\tilde{\Theta}_3}{2} + \frac{\tilde{\Theta}_3^2}{4} - \frac{1}{2} - \frac{\tilde{\Theta}_3}{2}\right] = \\
&= 2 \cdot \left[\frac{\tilde{\Theta}_3^2}{4} - \frac{1}{4}\right] = \frac{\tilde{\Theta}_3^2}{2} - \frac{1}{2}
\end{aligned} \tag{8.442}$$

Odvajamo po  $\tilde{\Theta}_3$  in dobimo:

$$\frac{df}{d\tilde{\Theta}_3} = \frac{2\tilde{\Theta}_3}{2} = 0 \quad \Rightarrow \quad \tilde{\Theta}_3^* = 0 \quad \Rightarrow \quad \underline{x}_3^* = \begin{bmatrix} \frac{1}{2} + 0 \\ \frac{1}{2} + 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \underline{x}_2^* = \underline{x}^* \tag{8.443}$$

Torej smo že v optimumu, ker se vektor ne spreminja več.

Analitični preizkus:

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 - 1 \\ 2x_2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad x_{1A} = \frac{1}{2}, \quad x_{2A} = \frac{1}{2} \tag{8.444}$$

Torej smo s Powellovo metodo dobili dejanski ekstrem.

#### Primer 4:

S pomočjo Powellove metode poiščite ekstrem naslednje funkcije:

$$f(x_1, x_2) = (x_1 - 3x_2)^2 + (x_2 - 1)^2 \tag{8.445}$$

Vzemite  $\underline{x}_0 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ .



1.) Opravimo premik iz začetnega vektorja v smeri enotskega vektorja  $\underline{e}_1$  :

$$\underline{x}_1 = \underline{x}_0 + \tilde{\Theta}_1 \cdot \underline{e}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \tilde{\Theta}_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 + \tilde{\Theta}_1 \\ 2 \end{bmatrix} \quad (8.446)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_1) &= (2 + \tilde{\Theta}_1 - 6)^2 + (2 - 1)^2 = (\tilde{\Theta}_1 - 4)^2 + 1 \\ \frac{df}{d\tilde{\Theta}_1} &= 2 \cdot (\tilde{\Theta}_1 - 4) = 0 \quad \Rightarrow \tilde{\Theta}_1^* = 4 \\ \Rightarrow \underline{x}_1^* &= \begin{bmatrix} 2 + 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} \end{aligned} \quad (8.447)$$

2.) Postopek ponovimo s premikom v smeri enotskega vektorja  $\underline{e}_2$  :

$$\underline{x}_2 = \underline{x}_1 + \tilde{\Theta}_2 \cdot \underline{e}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix} + \tilde{\Theta}_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 + \tilde{\Theta}_2 \end{bmatrix} \quad (8.448)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_2) &= (6 - 3 \cdot (2 + \tilde{\Theta}_2))^2 + (2 + \tilde{\Theta}_2 - 1)^2 = (6 - 6 - 3\tilde{\Theta}_2)^2 + (\tilde{\Theta}_2 + 1)^2 = \\ &= 9\tilde{\Theta}_2^2 + (\tilde{\Theta}_2 + 1)^2 \\ \frac{df}{d\tilde{\Theta}_2} &= 18\tilde{\Theta}_2 + 2 \cdot (\tilde{\Theta}_2 + 1) = 20\tilde{\Theta}_2 + 2 = 0 \\ \Rightarrow \tilde{\Theta}_2^* &= -\frac{1}{10} \quad \Rightarrow \underline{x}_2^* = \begin{bmatrix} 6 \\ 2 - \frac{1}{10} \end{bmatrix} = \begin{bmatrix} 6 \\ \frac{19}{10} \end{bmatrix} \end{aligned} \quad (8.449)$$

3.) Sledi premik v smeri  $\underline{x}_2 - \underline{x}_0$  :

$$\begin{aligned} \underline{x}_3 &= \underline{x}_2 + \tilde{\Theta}_3 \cdot (\underline{x}_2 - \underline{x}_0) = \begin{bmatrix} 6 \\ \frac{19}{10} \end{bmatrix} + \tilde{\Theta}_3 \cdot \left( \begin{bmatrix} 6 \\ \frac{19}{10} \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right) = \begin{bmatrix} 6 \\ \frac{19}{10} \end{bmatrix} + \tilde{\Theta}_3 \cdot \begin{bmatrix} 4 \\ -\frac{1}{10} \end{bmatrix} = \\ &= \begin{bmatrix} 6 + 4\tilde{\Theta}_3 \\ \frac{19}{10} - \frac{\tilde{\Theta}_3}{10} \end{bmatrix} \end{aligned} \quad (8.450)$$

Dobimo:

$$\begin{aligned}
 f(\underline{x}_3) &= \left\{ 6 + 4\tilde{\Theta}_3 - 3 \cdot \left( \frac{19}{10} - \frac{\tilde{\Theta}_3}{10} \right) \right\}^2 + \left( \frac{19}{10} - \frac{\tilde{\Theta}_3}{10} - 1 \right)^2 = \\
 &= \left\{ \frac{3}{10} + \frac{43}{10} \tilde{\Theta}_3 \right\}^2 + \left( \frac{9}{10} - \frac{\tilde{\Theta}_3}{10} \right)^2
 \end{aligned} \tag{8.451}$$

Odvajamo po  $\tilde{\Theta}_3$ :

$$\begin{aligned}
 \frac{df}{d\tilde{\Theta}_3} &= 2 \cdot \left\{ \frac{3}{10} + \frac{43}{10} \tilde{\Theta}_3 \right\} \cdot \frac{43}{10} + 2 \cdot \left( \frac{9}{10} - \frac{\tilde{\Theta}_3}{10} \right) \cdot \left( -\frac{1}{10} \right) = 0 \\
 \left\{ 3 + 43\tilde{\Theta}_3 \right\} \cdot \frac{43}{10^2} - \left\{ 9 - \tilde{\Theta}_3 \right\} \cdot \frac{1}{10^2} &= 0 \\
 3 \cdot 43 + 43^2 \cdot \tilde{\Theta}_3 - 9 + \tilde{\Theta}_3 &= 0 \\
 120 + 1850\tilde{\Theta}_3 = 0 \quad \Rightarrow \quad \tilde{\Theta}_3^* &= -\frac{12}{185}
 \end{aligned} \tag{8.452}$$

Dobimo:

$$\underline{x}_3^* = \begin{bmatrix} 6 - \frac{48}{185} \\ \frac{19}{10} + \frac{12}{1850} \end{bmatrix} = \begin{bmatrix} 5.7405 \\ 1.9064 \end{bmatrix} \tag{8.453}$$

4.) Sledi ponovni premik v smeri enotskega vektorja  $\underline{e}_2$ :

$$\underline{x}_4 = \underline{x}_3 + \tilde{\Theta}_4 \cdot \underline{e}_2 = \begin{bmatrix} 5.7405 \\ 1.9064 \end{bmatrix} + \tilde{\Theta}_4 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5.7405 \\ 1.9064 + \tilde{\Theta}_4 \end{bmatrix} \tag{8.454}$$

Dobimo:

$$\begin{aligned}
 f(\underline{x}_4) &= \left\{ 5.7405 - 3 \cdot (1.9064 + \tilde{\Theta}_4) \right\}^2 + (1.9064 + \tilde{\Theta}_4 - 1)^2 = \\
 &= \left\{ 0.0213 - 3\tilde{\Theta}_4 \right\}^2 + (0.9064 + \tilde{\Theta}_4)^2
 \end{aligned} \tag{8.455}$$

Odvajamo po  $\tilde{\Theta}_4$ :

$$\begin{aligned}\frac{df}{d\tilde{\Theta}_4} &= 2 \cdot (0.0213 - 3\tilde{\Theta}_4) \cdot (-3) + 2 \cdot (0.9064 + \tilde{\Theta}_4) = 0 \\ &-0.0639 + 9\tilde{\Theta}_4 + 0.9064 + \tilde{\Theta}_4 = 0 \\ 10\tilde{\Theta}_4 + 0.8425 &= 0 \quad \Rightarrow \tilde{\Theta}_4^* = -0.08425\end{aligned}\tag{8.456}$$

Dobimo:

$$\underline{x}_4^* = \begin{bmatrix} 5.7405 \\ 1.9064 - 0.08425 \end{bmatrix} = \begin{bmatrix} 5.7405 \\ 1.82215 \end{bmatrix}\tag{8.457}$$

5.) Sledi premik v smeri  $\underline{x}_3 - \underline{x}_0$ :

$$\begin{aligned}\underline{x}_5 &= \underline{x}_4 + \tilde{\Theta}_5 \cdot (\underline{x}_3 - \underline{x}_0) = \begin{bmatrix} 5.7405 \\ 1.82215 \end{bmatrix} + \tilde{\Theta}_5 \cdot \left\{ \begin{bmatrix} 5.7405 \\ 1.9064 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\} = \\ &= \begin{bmatrix} 5.7405 + 3.7405\tilde{\Theta}_5 \\ 1.82215 - 0.0936\tilde{\Theta}_5 \end{bmatrix}\end{aligned}\tag{8.458}$$

Dobimo:

$$\begin{aligned}f(\underline{x}_5) &= (5.7405 + 3.7405\tilde{\Theta}_5 - 3 \cdot (1.82215 - 0.0936\tilde{\Theta}_5))^2 + \\ &+ (1.82215 - 0.0936\tilde{\Theta}_5 - 1)^2 = \\ &= (0.2740 + 4.0213\tilde{\Theta}_5)^2 + (0.82215 - 0.0936\tilde{\Theta}_5)^2\end{aligned}\tag{8.459}$$

Odvajamo po  $\tilde{\Theta}_5$  in dobimo:

$$\begin{aligned}\frac{df}{d\tilde{\Theta}_5} &= 2 \cdot (0.2740 + 4.0213\tilde{\Theta}_5) \cdot 4.0213 - 2 \cdot (0.82215 - 0.0936\tilde{\Theta}_5) \cdot 0.0936 = 0 \\ 1.1018 + 16.1708\tilde{\Theta}_5 - 0.0769 + 8.76 \cdot 10^{-3} \cdot \tilde{\Theta}_5 &= 0 \\ 16.1795\tilde{\Theta}_5 + 1.0249 &= 0 \quad \Rightarrow \tilde{\Theta}_5^* = -0.0633 \\ &\Rightarrow \underline{x}_5^* = \begin{bmatrix} 5.503 \\ 1.828 \end{bmatrix}\end{aligned}\tag{8.460}$$

6.) Sledi premik v smeri  $\underline{x}_5 - \underline{x}_3$ :

$$\begin{aligned}\underline{x}_6 = \underline{x}^* &= \underline{x}_3 + \tilde{\Theta}_6 \cdot (\underline{x}_5 - \underline{x}_3) = \begin{bmatrix} 5.7405 \\ 1.9064 \end{bmatrix} + \tilde{\Theta}_6 \cdot \left\{ \begin{bmatrix} 5.503 \\ 1.828 \end{bmatrix} - \begin{bmatrix} 5.7405 \\ 1.9064 \end{bmatrix} \right\} = \\ &= \begin{bmatrix} 5.7405 - 0.2375\tilde{\Theta}_6 \\ 1.9064 - 0.0784\tilde{\Theta}_6 \end{bmatrix}\end{aligned}\quad (8.461)$$

Dobimo:

$$\begin{aligned}f(\underline{x}^*) &= \left\{ (5.7405 - 0.2375\tilde{\Theta}_6) - 3 \cdot (1.9064 - 0.0784\tilde{\Theta}_6) \right\}^2 + \\ &+ (1.9064 - 0.0784\tilde{\Theta}_6 - 1)^2 = \\ &= (0.0213 - 2.3 \cdot 10^{-3} \cdot \tilde{\Theta}_6)^2 + (0.9064 - 0.0784\tilde{\Theta}_6)^2\end{aligned}\quad (8.462)$$

Odvajamo po  $\tilde{\Theta}_6$ :

$$\begin{aligned}\frac{df}{d\tilde{\Theta}_6} &= 2 \cdot (0.0213 - 2.3 \cdot 10^{-3} \cdot \tilde{\Theta}_6) \cdot (-2.3 \cdot 10^{-3}) + \\ &+ 2 \cdot (0.9064 - 0.0784\tilde{\Theta}_6) \cdot (-0.0784) = 0 \\ -4.899 \cdot 10^{-5} &+ 5.29 \cdot 10^{-6} \cdot \tilde{\Theta}_6 - 0.0710 + 6.1465 \cdot 10^{-3} \tilde{\Theta}_6 = 0 \\ 6.15185 \cdot 10^{-3} \cdot \tilde{\Theta}_6 &= 0.07104 \quad \Rightarrow \tilde{\Theta}_6^* = 11.5492\end{aligned}\quad (8.463)$$

Dobimo:

$$\underline{x}^* = \left. \begin{bmatrix} 5.7405 - 0.2375 \cdot 11.5492 \\ 1.9064 - 0.0784 \cdot 11.5492 \end{bmatrix} = \begin{bmatrix} 2.9975 \\ 1.0009 \end{bmatrix} \right\} \text{ Rešitev} \quad (8.464)$$

Torej v 6. koraku dobimo minimum v točki  $\underline{x}^* = [2.9975 \quad 1.0009]^T$ .

### 8.4.2 Metoda konjugiranih gradientov

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.3.2.2, izraze (6.166) do (6.201) ter zlasti sliko 88.

#### Primer 1:

Poiščite minimum naslednje funkcije:

$$f(x_1, x_2) = x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 + x_1 + 10x_2, \quad \underline{x}_0 = [0 \ 0]^T \quad (8.465)$$

Uporabite Fletcher – Reevesovo metodo konjugiranih gradientov (glejte sliko 88).

Gradient funkcije (8.465) je:

$$\nabla \underline{f}(x_1, x_2) = \begin{bmatrix} 2x_1 + x_2 + 1 \\ x_1 + x_2 + 10 \end{bmatrix} \quad (8.466)$$

Pravi (analitični) ekstrem je:

$$\begin{aligned} \nabla \underline{f}(x_1, x_2) &= 0 \\ \left. \begin{aligned} 2x_1 + x_2 + 1 &= 0 \\ x_1 + x_2 + 10 &= 0 \end{aligned} \right\} \cdot 2 & - \\ \left. \begin{aligned} -x_2 - 19 &= 0 \\ 2x_{1A}^* - 19 + 1 &= 0 \end{aligned} \right\} \Rightarrow \begin{aligned} x_{2A}^* &= -19 \\ x_{1A}^* &= 9 \end{aligned} & \left. \begin{aligned} \left[ \begin{array}{c} x_{1A}^* \\ x_{2A}^* \end{array} \right] &= \left[ \begin{array}{c} 9 \\ -19 \end{array} \right] \end{aligned} \right\} \end{aligned} \quad (8.467)$$

1. iteracija:

Najprej realiziramo izraz a na sliki 88:

$$\nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} 2 \cdot 0 + 0 + 1 \\ 0 + 0 + 10 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix} = \nabla \underline{f}_0 \quad (8.468)$$

Nato realiziramo izraz b na sliki 88:

$$\underline{u}_0 = -\nabla \underline{f}_0 = \begin{bmatrix} -1 \\ -10 \end{bmatrix} \quad (8.469)$$

Sledi realizacija izraza c na sliki 88:

$$\underline{x}_1 = \underline{x}_0 + \alpha_0 \cdot \underline{u}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \alpha_0 \cdot \begin{bmatrix} -1 \\ -10 \end{bmatrix} = \begin{bmatrix} -\alpha_0 \\ -10\alpha_0 \end{bmatrix} \quad (8.470)$$

Dobimo:

$$\begin{aligned}
 f(\underline{x}_1) &= (-\alpha_0)^2 + (-\alpha_0) \cdot (-10\alpha_0) + \frac{1}{2} \cdot (-10\alpha_0)^2 + (-\alpha_0) + 10 \cdot (-10\alpha_0) = \\
 &= \alpha_0^2 + 10\alpha_0^2 + 50\alpha_0^2 - 101\alpha_0 = 61\alpha_0^2 - 101\alpha_0
 \end{aligned} \tag{8.471}$$

Sledi:

$$\frac{df}{d\alpha_0} = 122\alpha_0 - 101 = 0 \quad \Rightarrow \quad \alpha_0^* = \frac{101}{122} = 0.8278 \tag{8.472}$$

Realizacija izraza d na slici 88:

$$\underline{x}_1 = \begin{bmatrix} -0.8278 \\ -8.2786 \end{bmatrix} \tag{8.473}$$

Realizacija izraza e na slici 88:

$$\nabla \underline{f}(\underline{x}_1) = \begin{bmatrix} 2 \cdot (-0.8278) - 8.2786 + 1 \\ -0.8278 - 8.2786 + 10 \end{bmatrix} = \begin{bmatrix} -8.9334 \\ 0.8936 \end{bmatrix} = \nabla \underline{f}_1 \tag{8.474}$$

Realizacija izraza f na slici 88:

$$\beta_1 = \frac{\nabla \underline{f}_1^T \cdot \nabla \underline{f}_1}{\nabla \underline{f}_0^T \cdot \nabla \underline{f}_0} = \frac{\begin{bmatrix} -8.9334 & 0.8936 \end{bmatrix} \cdot \begin{bmatrix} -8.9334 \\ 0.8936 \end{bmatrix}}{\begin{bmatrix} 1 & 10 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 10 \end{bmatrix}} = \frac{80.6041}{101} = 0.798 \tag{8.475}$$

Realizacija izraza g na slici 88:

$$\begin{aligned}
 \underline{u}_1 &= -\nabla \underline{f}_1 + \beta_1 \cdot \underline{u}_0 = -\begin{bmatrix} -8.9334 \\ 0.8936 \end{bmatrix} + 0.798 \cdot \begin{bmatrix} -1 \\ -10 \end{bmatrix} = \begin{bmatrix} 8.9334 - 0.798 \\ -0.8936 - 0.798 \cdot 10 \end{bmatrix} = \\
 &= \begin{bmatrix} 8.1354 \\ -8.8736 \end{bmatrix}
 \end{aligned} \tag{8.476}$$

$k = 1$ , 2. iteracija:

Realizacija izraza c na slici 88:

$$\underline{x}_2 = \underline{x}_1 + \alpha_1 \cdot \underline{u}_1 = \begin{bmatrix} -0.8278 \\ -8.2786 \end{bmatrix} + \alpha_1 \cdot \begin{bmatrix} 8.1354 \\ -8.8736 \end{bmatrix} = \begin{bmatrix} -0.8278 + 8.1354\alpha_1 \\ -8.2786 - 8.8736\alpha_1 \end{bmatrix} \tag{8.477}$$

Dobimo:

$$\begin{aligned}
 f(\underline{x}_2) &= (-0.8278 + 8.1354\alpha_1)^2 + \\
 &+ (-0.8278 + 8.1354\alpha_1) \cdot (-8.2786 - 8.8736\alpha_1) + \\
 &+ \frac{1}{2} \cdot (-8.2786 - 8.8736\alpha_1)^2 + (-0.8278 + 8.1354\alpha_1) + \\
 &+ 10 \cdot (-8.2786 - 8.8736\alpha_1) =
 \end{aligned} \tag{8.478}$$

Sledi:

$$\begin{aligned}
 \frac{df}{d\alpha_1} &= 2 \cdot (8.1354\alpha_1 - 0.8278) \cdot 8.1354 + 8.1354 \cdot (-8.2786 - 8.8736\alpha_1) + \\
 &+ (-0.8278 + 8.1354\alpha_1) \cdot (-8.8736) + \\
 &+ (-8.2786 - 8.8736\alpha_1) \cdot (-8.8736) + 8.1354 - 10 \cdot 8.8736 = \\
 &= 132.369\alpha_1 - 13.4689 - 67.3497 - 72.1902\alpha_1 + 7.3455 - 72.190\alpha_1 + \\
 &+ 73.4609 + 78.740\alpha_1 - 80.6006 = \\
 &= 66.7288\alpha_1 - 80.6128 = 0 \quad \Rightarrow \alpha_1^* = 1.2080
 \end{aligned} \tag{8.479}$$

Realizacija izraza d na sliki 88:

$$\underline{x}_2 = \begin{bmatrix} -0.8278 + 1.2080 \cdot 8.1354 \\ -8.2786 - 1.2080 \cdot 8.8736 \end{bmatrix} = \begin{bmatrix} 8.9997 \\ -18.9979 \end{bmatrix} \tag{8.480}$$

Če rezultat, dobljen v izrazu (8.480) primerjamo s pravim analitičnim ekstremom, vidimo, da smo konvergirali v optimum – že po 2. iteracijah.

### Primer 2:

Poiščite minimum naslednje funkcije:

$$f(x_1, x_2) = 2x_1^2 + x_2^2 - x_1x_2, \quad \underline{x}_0 = [2 \quad 2]^T \tag{8.481}$$

Uporabite Fletcher – Reevesovo metodo konjugiranih gradientov (glejte sliko 88).

Gradient funkcije (8.481) je:

$$\nabla \underline{f}(x_1, x_2) = \begin{bmatrix} 4x_1 - x_2 \\ 2x_2 - x_1 \end{bmatrix} \tag{8.482}$$

Pravi (analitični) ekstrem je:

$$\begin{aligned}
 \nabla \underline{f}(x_1, x_2) &= 0 \\
 4x_1 - x_2 &= 0 \quad \Rightarrow x_2 = 4x_1 \\
 2x_2 - x_1 &= 0 \quad \Rightarrow 2 \cdot 4x_1 - x_1 = 0 \quad \Rightarrow 7x_1 = 0 \\
 \Rightarrow x_{1A}^* &= 0, \quad x_{2A}^* = 0 \quad \Rightarrow \begin{bmatrix} x_{1A}^* \\ x_{2A}^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
 \end{aligned} \tag{8.483}$$

1. iteracija:

Realizacija izraza a na sliki 88:

$$\nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} 4 \cdot 2 - 2 \\ 2 \cdot 2 - 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \nabla \underline{f}_0 \tag{8.484}$$

Realizacija izraza b na sliki 88:

$$\underline{u}_0 = -\nabla \underline{f}_0 = \begin{bmatrix} -6 \\ -2 \end{bmatrix} \tag{8.485}$$

Realizacija izraza c na sliki 88:

$$\underline{x}_1 = \underline{x}_0 + \alpha_0 \cdot \underline{u}_0 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \alpha_0 \cdot \begin{bmatrix} -6 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 - 6\alpha_0 \\ 2 - 2\alpha_0 \end{bmatrix} \tag{8.486}$$

Dobimo:

$$f(\underline{x}_1) = 2 \cdot (2 - 6\alpha_0)^2 + (2 - 2\alpha_0)^2 - \underbrace{(2 - 6\alpha_0) \cdot (2 - 2\alpha_0)}_{(4 - 16\alpha_0 + 12\alpha_0^2)} \tag{8.487}$$

Sledi:

$$\begin{aligned}
 \frac{df}{d\alpha_0} &= 4 \cdot (2 - 6\alpha_0) \cdot (-6) + 2 \cdot (2 - 2\alpha_0) \cdot (-2) + 16 - 24\alpha_0 = \\
 &= -48 + 144\alpha_0 - 8 + 8\alpha_0 + 16 - 24\alpha_0 = -40 + 128\alpha_0 = 0 \\
 \Rightarrow \alpha_0^* &= \frac{40}{128} = 0.3125
 \end{aligned} \tag{8.488}$$

Realizacija izraza d na sliki 88:



$$\underline{x}_1 = \begin{bmatrix} 2 - 6 \cdot 0.3125 \\ 2 - 2 \cdot 0.3125 \end{bmatrix} = \begin{bmatrix} 0.125 \\ 1.375 \end{bmatrix} \quad (8.489)$$

Realizacija izraza e na slici 88:

$$\nabla \underline{f}(\underline{x}_1) = \begin{bmatrix} 4 \cdot 0.125 - 1.375 \\ 2 \cdot 1.375 - 0.125 \end{bmatrix} = \begin{bmatrix} -0.875 \\ 2.625 \end{bmatrix} = \nabla \underline{f}_1 \quad (8.490)$$

Realizacija izraza f na slici 88:

$$\beta_1 = \frac{\nabla \underline{f}_1^T \cdot \nabla \underline{f}_1}{\nabla \underline{f}_0^T \cdot \nabla \underline{f}_0} = \frac{\begin{bmatrix} -0.875 & 2.625 \end{bmatrix} \cdot \begin{bmatrix} -0.875 \\ 2.625 \end{bmatrix}}{\begin{bmatrix} 6 & 2 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 2 \end{bmatrix}} = \frac{7.6562}{40} = 0.1914 \quad (8.491)$$

Realizacija izraza g na slici 88:

$$\underline{u}_1 = -\nabla \underline{f}_1 + \beta_1 \cdot \underline{u}_0 = \begin{bmatrix} 0.875 \\ -2.625 \end{bmatrix} + \begin{bmatrix} 0.1914 \cdot (-6) \\ 0.1914 \cdot (-2) \end{bmatrix} = \begin{bmatrix} -0.2734 \\ -3.0078 \end{bmatrix} \quad (8.492)$$

$k = 1, 2$ . iteracija:

Realizacija izraza c na slici 88:

$$\underline{x}_2 = \underline{x}_1 + \alpha_1 \cdot \underline{u}_1 = \begin{bmatrix} 0.125 \\ 1.375 \end{bmatrix} + \alpha_1 \cdot \begin{bmatrix} -0.2734 \\ -3.0078 \end{bmatrix} = \begin{bmatrix} 0.125 - 0.2734\alpha_1 \\ 1.375 - 3.0078\alpha_1 \end{bmatrix} \quad (8.493)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_2) &= 2 \cdot (0.125 - 0.2734\alpha_1)^2 + (1.375 - 3.0078\alpha_1)^2 - \\ &\quad - (0.125 - 0.2734\alpha_1) \cdot (1.375 - 3.0078\alpha_1) = \\ &= 2 \cdot (0.125 - 0.2734\alpha_1)^2 + (1.375 - 3.0078\alpha_1)^2 - \\ &\quad - \{0.125 \cdot 1.375 - 0.125 \cdot 3.0078\alpha_1 - 0.2734 \cdot 1.375\alpha_1 + 0.2734 \cdot 3.0078\alpha_1^2\} \end{aligned} \quad (8.494)$$

Sledi:

$$\begin{aligned} \frac{df}{d\alpha_1} &= -4 \cdot (0.125 - 0.2734\alpha_1) \cdot 0.2734 - 2 \cdot (1.375 - 3.0078\alpha_1) \cdot 3.0078 + \\ &\quad + 0.125 \cdot 3.0078 + 0.2734 \cdot 1.375 - 2 \cdot 0.2734 \cdot 3.0078\alpha_1 = \\ &= -7.65625 + 16.7480\alpha_1 = 0 \quad \Rightarrow \alpha_1^* = 0.4571 \end{aligned} \quad (8.495)$$

Realizacija izraza d na sliki 88:

$$\underline{x}_2 = \begin{bmatrix} 0.125 - 0.2734 \cdot 0.4571 \\ 1.375 - 3.0078 \cdot 0.4571 \end{bmatrix} = \begin{bmatrix} 2.88 \cdot 10^{-5} \\ 1.34 \cdot 10^{-4} \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.496)$$

Če rezultat, dobljen v izrazu (8.496) primerjamo s pravim analitičnim ekstremom, vidimo, da smo ponovno konvergirali v optimum – že po 2. iteracijah.

### Primer 3:

S pomočjo Fletcher – Reevesove metode konjugiranih gradientov poiščite minimum za naslednjo funkcijo:

$$f(x_1, x_2) = 3 + (x_1 - 1.5x_2)^2 + (x_2 - 2)^2, \quad \text{pri čemer je } \underline{x}_0 = [0.5 \quad 0.5]^T. \quad (8.497)$$

Gradient funkcije (8.497) je:

$$\begin{aligned} \nabla \underline{f}(x_1, x_2) &= \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2 \cdot (x_1 - 1.5x_2) \\ 2 \cdot (x_1 - 1.5x_2) \cdot (-1.5) + 2 \cdot (x_2 - 2) \end{bmatrix} = \\ &= \begin{bmatrix} 2x_1 - 3x_2 \\ -3x_1 + 4.5x_2 + 2x_2 - 4 \end{bmatrix} = \begin{bmatrix} 2x_1 - 3x_2 \\ -3x_1 + 6.5x_2 - 4 \end{bmatrix} \end{aligned} \quad (8.498)$$

Pravi (analitični) ekstrem je:

$$\begin{aligned} \nabla \underline{f}(x_1, x_2) = 0 &\Rightarrow 2x_1 - 3x_2 = 0 &\Rightarrow x_1 = \frac{3}{2}x_2 \\ &\Rightarrow -3x_1 + 6.5x_2 - 4 = 0 &\Rightarrow -3 \cdot \frac{3}{2}x_2 + 6.5x_2 = 4 \\ &\Rightarrow 2x_2 = 4 &\Rightarrow x_{2A}^* = 2, \quad x_{1A}^* = \frac{3}{2} \cdot 2 = 3 \\ &\Rightarrow \begin{bmatrix} x_{1A}^* \\ x_{2A}^* \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \end{aligned} \quad (8.499)$$

1. iteracija:

Realizacija izraza a na sliki 88:

$$\nabla \underline{f}(\underline{x}_0) = \begin{bmatrix} 2 \cdot 0.5 - 3 \cdot 0.5 \\ -3 \cdot 0.5 + 6.5 \cdot 0.5 - 4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -2.25 \end{bmatrix} = \nabla \underline{f}_0 \quad (8.500)$$

Realizacija izraza b na slici 88:

$$\underline{u}_0 = -\nabla \underline{f}_0 = \begin{bmatrix} \frac{1}{2} \\ 2.25 \end{bmatrix} \quad (8.501)$$

Realizacija izraza c na slici 88:

$$\underline{x}_1 = \underline{x}_0 + \alpha_0 \cdot \underline{u}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} + \alpha_0 \cdot \begin{bmatrix} \frac{1}{2} \\ 2.25 \end{bmatrix} = \begin{bmatrix} 0.5 + 0.5\alpha_0 \\ 0.5 + 2.25\alpha_0 \end{bmatrix} \quad (8.502)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_1) &= 3 + (0.5 + 0.5\alpha_0 - 1.5 \cdot (0.5 + 2.25\alpha_0))^2 + (0.5 + 2.25\alpha_0 - 2)^2 = \\ &= 3 + (-0.25 - 2.875\alpha_0)^2 + (-1.5 + 2.25\alpha_0)^2 = \\ &= 3 + (0.25 + 2.875\alpha_0)^2 + (-1.5 + 2.25\alpha_0)^2 \end{aligned} \quad (8.503)$$

Sledi:

$$\begin{aligned} \frac{df}{d\alpha_0} &= 2 \cdot (0.25 + 2.875\alpha_0) \cdot 2.875 + 2 \cdot (2.25\alpha_0 - 1.5) \cdot 2.25 = \\ &= 1.4375 + 16.5312\alpha_0 + 10.125\alpha_0 - 6.75 = 0 \\ &\Rightarrow \alpha_0^* = \frac{5.3125}{26.6562} = 0.1992 \end{aligned} \quad (8.504)$$

Realizacija izraza d na slici 88:

$$\underline{x}_1 = \begin{bmatrix} 0.5 + 0.1992 \cdot 0.5 \\ 0.5 + 0.1992 \cdot 2.25 \end{bmatrix} = \begin{bmatrix} 0.5996 \\ 0.9482 \end{bmatrix} \quad (8.505)$$

Realizacija izraza e na slici 88:

$$\nabla \underline{f}(\underline{x}_1) = \begin{bmatrix} 2 \cdot 0.5996 - 3 \cdot 0.9482 \\ -3 \cdot 0.5996 + 6.5 \cdot 0.9482 - 4 \end{bmatrix} = \begin{bmatrix} -1.6454 \\ 0.3645 \end{bmatrix} = \nabla \underline{f}_1 \quad (8.506)$$

Realizacija izraza f na sliki 88:

$$\beta_1 = \frac{\nabla \underline{f}_1^T \cdot \nabla \underline{f}_1}{\nabla \underline{f}_0^T \cdot \nabla \underline{f}_0} = \frac{[-1.6454 \quad 0.3645] \cdot \begin{bmatrix} -1.6454 \\ 0.3645 \end{bmatrix}}{\begin{bmatrix} -\frac{1}{2} & -2.25 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{2} \\ -2.25 \end{bmatrix}} = \frac{2.8402}{5.3125} = 0.5346 \quad (8.507)$$

Realizacija izraza g na sliki 88:

$$\underline{u}_1 = -\nabla \underline{f}_1 + \beta_1 \cdot \underline{u}_0 = \begin{bmatrix} 1.6454 \\ -0.3645 \end{bmatrix} + 0.5346 \cdot \begin{bmatrix} \frac{1}{2} \\ 2.25 \end{bmatrix} = \begin{bmatrix} 1.9127 \\ 0.8383 \end{bmatrix} \quad (8.508)$$

$k = 1$ , 2. iteracija:

Realizacija izraza c na sliki 88:

$$\underline{x}_2 = \underline{x}_1 + \alpha_1 \cdot \underline{u}_1 = \begin{bmatrix} 0.5996 \\ 0.9482 \end{bmatrix} + \alpha_1 \cdot \begin{bmatrix} 1.9127 \\ 0.8383 \end{bmatrix} = \begin{bmatrix} 0.5996 + 1.9127\alpha_1 \\ 0.9482 + 0.8383\alpha_1 \end{bmatrix} \quad (8.509)$$

Dobimo:

$$\begin{aligned} f(\underline{x}_2) &= 3 + (0.5996 + 1.9127\alpha_1 - 1.5 \cdot (0.9482 + 0.8383\alpha_1))^2 + \\ &\quad + (0.9482 + 0.8383\alpha_1 - 2)^2 = \\ &= 3 + (-0.8227 + 0.6552\alpha_1)^2 + (-1.0518 + 0.8383\alpha_1)^2 \end{aligned} \quad (8.510)$$

Sledi:

$$\begin{aligned} \frac{df}{d\alpha_1} &= 2 \cdot (0.6552\alpha_1 - 0.8227) \cdot 0.6552 + 2 \cdot (0.8383\alpha_1 - 1.0518) \cdot 0.8383 = 0 \\ 1.1320\alpha_1 - 1.4207 &= 0 \quad \Rightarrow \alpha_1^* = 1.2550 \end{aligned} \quad (8.511)$$

Realizacija izraza d na sliki 88:

$$\underline{x}_2 = \begin{bmatrix} 0.5996 + 1.2550 \cdot 1.9127 \\ 0.9482 + 1.2550 \cdot 0.8383 \end{bmatrix} = \begin{bmatrix} 3.0000 \\ 2.0002 \end{bmatrix} \quad (8.512)$$

Torej smo v pravi optimum zopet konvergirali že po 2. iteracijah.

### 8.4.3 Quasi-Newtonove metode

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 6.3.2.3, izraze (6.202) do (6.254) ter zlasti sliki 89 in 90.

#### Primer 1:

S pomočjo DFP metode, poiščite minimum funkcije:

$$f(x_1, x_2) = x_1^2 + 4x_2^2 - 4x_1 \quad (8.513)$$

Vzemite  $\underline{x}_0 = [1 \ 1]^T$ . (Glejte sliko 89!)

Začetna metrična matrika se glasi:

$$\underline{G}_0 = \underline{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (8.514)$$

$i = 0$ , 1. iteracija:

Izračunajmo gradient v začetni točki:

$$\underline{g}_0 = \nabla f(\underline{x}_0) = \left. \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} \right|_{\underline{x}=\underline{x}_0} = \left. \begin{bmatrix} 2x_1 - 4 \\ 8x_2 \end{bmatrix} \right|_{\substack{x_1=1 \\ x_2=1}} = \begin{bmatrix} -2 \\ 8 \end{bmatrix} \quad (8.515)$$

Izračunajmo še nov vektor odločitvenih spremenljivk:

$$\underline{x}_1 = \underline{x}_0 - \alpha_0 \cdot \underline{G}_0 \cdot \underline{g}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \alpha_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 + 2\alpha_0 \\ 1 - 8\alpha_0 \end{bmatrix} \quad (8.516)$$

V točki (8.516) funkcija (8.513) zavzame vrednost:

$$f(\underline{x}_1) = (1 + 2\alpha_0)^2 + 4 \cdot (1 - 8\alpha_0)^2 - 4 \cdot (1 + 2\alpha_0) \quad (8.517)$$

$\alpha_0^*$  poiščemo na analitičen način:

$$\begin{aligned} \frac{df}{d\alpha_0} &= 2 \cdot (1 + 2\alpha_0) \cdot 2 + 4 \cdot 2 \cdot (1 - 8\alpha_0) \cdot (-8) - 8 = 4 + 8\alpha_0 - 64 \cdot (1 - 8\alpha_0) - 8 = \\ &= 4 + 8\alpha_0 - 64 + 512\alpha_0 - 8 = -68 + 520\alpha_0 = 0 \quad \Rightarrow \alpha_0^* = \frac{68}{520} = 0.1307 \end{aligned} \quad (8.518)$$

Če izraz (8.518) vstavimo v izraz (8.516), dobimo:

$$\underline{x}_1 = \begin{bmatrix} 1 + 2 \cdot 0.1307 \\ 1 - 8 \cdot 0.1307 \end{bmatrix} = \begin{bmatrix} 1.2615 \\ -0.0456 \end{bmatrix} \quad (8.519)$$

Gradient funkcije (8.513) v točki (8.519) se glasi:

$$\underline{g}_1 = \nabla f(\underline{x}_1) = \begin{bmatrix} 2x_1 - 4 \\ 8x_2 \end{bmatrix} \bigg|_{\substack{x_1=1.2615 \\ x_2=-0.0456}} = \begin{bmatrix} -1.477 \\ -0.3648 \end{bmatrix} \quad (8.520)$$

Spremembo gradienta pa dobimo, če tvorimo razliko med izrazoma (8.520) in (8.515):

$$\Delta \underline{g}_0 = \underline{g}_1 - \underline{g}_0 = \begin{bmatrix} -1.477 \\ -0.3648 \end{bmatrix} - \begin{bmatrix} -2 \\ 8 \end{bmatrix} = \begin{bmatrix} 0.523 \\ -8.3648 \end{bmatrix} \quad (8.521)$$

V nadaljevanju je potrebno v skladu s sliko 89 izračunati:

$$\underline{\underline{M}}_0 = \frac{\Delta \underline{x}_0 \cdot \Delta \underline{x}_0^T}{\Delta \underline{x}_0^T \cdot \Delta \underline{g}_0} \quad (8.522)$$

Seveda pa je potrebno najprej izračunati vektor  $\Delta \underline{x}_0$ , pri čemer od vektorja (8.519) odštejemo začetno točko:

$$\Delta \underline{x}_0 = \underline{x}_1 - \underline{x}_0 = \begin{bmatrix} 1.2615 \\ -0.0456 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.2615 \\ -1.0456 \end{bmatrix} \quad (8.523)$$

Če izraza (8.523) in (8.521) vstavimo v izraz (8.522), dobimo:

$$\begin{aligned} \underline{\underline{M}}_0 &= \frac{\begin{bmatrix} 0.2615 \\ -1.0456 \end{bmatrix} \cdot \begin{bmatrix} 0.2615 & -1.0456 \end{bmatrix}}{\begin{bmatrix} 0.2615 & -1.0456 \end{bmatrix} \cdot \begin{bmatrix} 0.523 \\ -8.3648 \end{bmatrix}} = \frac{\begin{bmatrix} 0.0683 & -0.2734 \\ -0.2734 & 1.0932 \end{bmatrix}}{8.8829} = \\ &= \begin{bmatrix} 7.688 \cdot 10^{-3} & -0.0307 \\ -0.0307 & 0.1230 \end{bmatrix} \end{aligned} \quad (8.524)$$

V nadaljevanju je potrebno v skladu s sliko 89 izračunati:

$$\begin{aligned} \underline{\underline{N}}_0 &= \frac{\frac{\underline{\underline{I}}}{\underline{\underline{G}}_0} \cdot \Delta \underline{\underline{g}}_0 \cdot \Delta \underline{\underline{g}}_0^T \cdot \frac{\underline{\underline{I}}}{\underline{\underline{G}}_0}}{\Delta \underline{\underline{g}}_0^T \cdot \underline{\underline{G}}_0 \cdot \Delta \underline{\underline{g}}_0} = \frac{\begin{bmatrix} 0.523 \\ -8.3648 \end{bmatrix} \cdot \begin{bmatrix} 0.523 & -8.3648 \end{bmatrix}}{\begin{bmatrix} 0.523 & -8.3648 \end{bmatrix} \cdot \underline{\underline{I}} \cdot \begin{bmatrix} 0.523 \\ -8.3648 \end{bmatrix}} = \\ &= \frac{\begin{bmatrix} 0.2735 & -4.3747 \\ -4.3747 & 69.9698 \end{bmatrix}}{70.2434} = \begin{bmatrix} 3.8936 \cdot 10^{-3} & -0.0622 \\ -0.0622 & 0.9961 \end{bmatrix} \end{aligned} \quad (8.525)$$

Na osnovi izrazov (8.514), (8.524) in (8.525) izračunamo še (glej sliko 89):

$$\begin{aligned} \underline{\underline{G}}_1 &= \underline{\underline{G}}_0 + \underline{\underline{M}}_0 - \underline{\underline{N}}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 7.688 \cdot 10^{-3} & -0.0307 \\ -0.0307 & 0.1230 \end{bmatrix} - \\ &\quad - \begin{bmatrix} 3.8936 \cdot 10^{-3} & -0.0622 \\ -0.0622 & 0.9961 \end{bmatrix} = \\ &= \begin{bmatrix} 1.0037 & 0.0315 \\ 0.0315 & 0.1269 \end{bmatrix} \end{aligned} \quad (8.526)$$

$i = 1, 2.$  iteracija:

V 2. iteraciji najprej izračunamo nov vektor odločitvenih spremenljivk:

$$\begin{aligned} \underline{\underline{x}}_2 &= \underline{\underline{x}}_1 - \alpha_1 \cdot \underline{\underline{G}}_1 \cdot \underline{\underline{g}}_1 = \begin{bmatrix} 1.2615 \\ -0.0456 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} 1.0037 & 0.0315 \\ 0.0315 & 0.1269 \end{bmatrix} \cdot \begin{bmatrix} -1.477 \\ -0.3648 \end{bmatrix} = \\ &= \begin{bmatrix} 1.2615 \\ -0.0456 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} -1.4939 \\ -0.0928 \end{bmatrix} = \begin{bmatrix} 1.2615 + 1.4939\alpha_1 \\ -0.0456 + 0.0928\alpha_1 \end{bmatrix} \end{aligned} \quad (8.527)$$

V točki (8.527) funkcija (8.513) zavzame vrednost:

$$f(\underline{x}_2) = (1.2615 + 1.4939\alpha_1)^2 + 4 \cdot (-0.0456 + 0.0928\alpha_1)^2 - 4 \cdot (1.2615 + 1.4939\alpha_1) \quad (8.528)$$

$\alpha_1^*$  poiščemo na analitičen način:

$$\begin{aligned} \frac{df}{d\alpha_1} &= 2 \cdot (1.2615 + 1.4939\alpha_1) \cdot 1.4939 + 4 \cdot 2 \cdot (-0.0456 + 0.0928\alpha_1) \cdot 0.0928 - \\ &- 4 \cdot 1.4939 = 0 \quad (8.529) \\ 1.8845 + 2.2317\alpha_1 - \underbrace{4.2316 \cdot 10^{-3}}_{\text{ZANEMARIMO}} + \underbrace{8.6118 \cdot 10^{-3}\alpha_1}_{\text{ZANEMARIMO}} - 2.9878 &= 0 \end{aligned}$$

Sledi:

$$\alpha_1^* = \frac{2.9878 - 1.8845}{2.2317} = 0.4943 \quad (8.530)$$

Če izraz (8.530) vstavimo v izraz (8.527), dobimo:

$$\underline{x}_2 = \begin{bmatrix} 1.2615 + 0.4943 \cdot 1.4939 \\ -0.0456 + 0.4943 \cdot 0.0928 \end{bmatrix} = \begin{bmatrix} 2.0000 \\ 2.71 \cdot 10^{-4} \end{bmatrix} \approx \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad (8.531)$$

Gradient funkcije (8.513) v točki (8.531) se glasi:

$$\underline{g}_2 = \nabla \underline{f}(\underline{x}_2) = \begin{bmatrix} 2x_1 - 4 \\ 8x_2 \end{bmatrix} \Bigg|_{\substack{x_1=2 \\ x_2=0}} = \begin{bmatrix} 2 \cdot 2 - 4 \\ 8 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.532)$$

Ker pride gradient v izrazu (8.532) enak 0, končamo s postopkom. Končna rešitev torej je:

$$\underline{x}^* = \underline{x}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad (8.533)$$

Konvergenca je bila zelo hitra.

Analitični preizkus:

$$\nabla \underline{f}(\underline{x}) = 0 \quad \Rightarrow \quad \left. \begin{array}{l} 2x_1 - 4 = 0 \\ 8x_2 = 0 \end{array} \right\} \begin{array}{l} x_{1A} = 2 \\ x_{2A} = 0 \end{array} \quad \Rightarrow \quad \underline{x}_A = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad (8.534)$$



**Primer 2:**

Poiščite minimum naslednje funkcije:

$$f(x_1, x_2) = x_1^2 + x_1 x_2 + \frac{1}{2} x_2^2 + x_1 + 10x_2, \quad \underline{x}_0 = [0 \ 0]^T \quad (8.535)$$

Uporabite DFP Quasi - Newtonovo metodo.

Začetna metrična matrika se glasi:

$$\underline{G}_0 = \underline{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (8.536)$$

$i = 0, 1$ . iteracija:

Izračunajmo gradient v začetni točki:

$$\underline{g}_0 = \nabla f(\underline{x}_0) = \left. \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} \right|_{\underline{x}=\underline{x}_0} = \left. \begin{bmatrix} 2x_1 + x_2 + 1 \\ x_1 + x_2 + 10 \end{bmatrix} \right|_{\substack{x_1=0 \\ x_2=0}} = \begin{bmatrix} 1 \\ 10 \end{bmatrix} \quad (8.537)$$

Izračunajmo še nov vektor odločitvenih spremenljivk:

$$\underline{x}_1 = \underline{x}_0 - \alpha_0 \cdot \underline{G}_0 \cdot \underline{g}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \alpha_0 \cdot \underline{I} \cdot \begin{bmatrix} 1 \\ 10 \end{bmatrix} = \begin{bmatrix} -\alpha_0 \\ -10\alpha_0 \end{bmatrix} \quad (8.538)$$

V točki (8.538) funkcija (8.535) zavzame vrednost:

$$\begin{aligned} f(\underline{x}_1) &= (-\alpha_0)^2 + (-\alpha_0) \cdot (-10\alpha_0) + \frac{1}{2} (-10\alpha_0)^2 + (-\alpha_0) + 10 \cdot (-10\alpha_0) = \\ &= \alpha_0^2 + 10\alpha_0^2 + 50\alpha_0^2 - \alpha_0 - 100\alpha_0 = 61\alpha_0^2 - 101\alpha_0 \end{aligned} \quad (8.539)$$

$\alpha_0^*$  poiščemo na analitičen način:

$$\frac{df}{d\alpha_0} = 122\alpha_0 - 101 = 0 \quad \Rightarrow \alpha_0^* = \frac{101}{122} = 0.8278 \quad (8.540)$$

Če izraz (8.540) vstavimo v izraz (8.538), dobimo:

$$\underline{x}_1 = \begin{bmatrix} -0.8278 \\ -8.278 \end{bmatrix} \quad (8.541)$$

Gradient funkcije (8.535) v točki (8.541) se glasi:

$$\underline{g}_1 = \nabla f(\underline{x}_1) = \begin{bmatrix} 2x_1 + x_2 + 1 \\ x_1 + x_2 + 10 \end{bmatrix} \Bigg|_{\substack{x_1 = -0.8278 \\ x_2 = -8.278}} = \begin{bmatrix} 2 \cdot (-0.8278) - 8.278 + 1 \\ -0.8278 - 8.278 + 10 \end{bmatrix} = \begin{bmatrix} -8.9336 \\ 0.8942 \end{bmatrix} \quad (8.542)$$

Spremembo gradienta pa dobimo, če tvorimo razliko med izrazoma (8.542) in (8.537):

$$\Delta \underline{g}_0 = \underline{g}_1 - \underline{g}_0 = \begin{bmatrix} -8.9336 \\ 0.8942 \end{bmatrix} - \begin{bmatrix} 1 \\ 10 \end{bmatrix} = \begin{bmatrix} -9.9336 \\ -9.1058 \end{bmatrix} \quad (8.543)$$

Izračunajmo še vektor  $\Delta \underline{x}_0$ , pri čemer od vektorja (8.541) odštejemo začetno točko:

$$\Delta \underline{x}_0 = \underline{x}_1 - \underline{x}_0 = \begin{bmatrix} -0.8278 \\ -8.278 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.8278 \\ -8.278 \end{bmatrix} \quad (8.544)$$

Če izraza (8.544) in (8.543) vstavimo v izraz (8.522), dobimo (glej sliko 89):

$$\begin{aligned} \underline{\underline{M}}_0 &= \frac{\Delta \underline{x}_0 \cdot \Delta \underline{x}_0^T}{\Delta \underline{x}_0^T \cdot \Delta \underline{g}_0} = \frac{\begin{bmatrix} -0.8278 \\ -8.278 \end{bmatrix} \cdot \begin{bmatrix} -0.8278 & -8.278 \end{bmatrix}}{\begin{bmatrix} -0.8278 & -8.278 \end{bmatrix} \cdot \begin{bmatrix} -9.9336 \\ -9.1058 \end{bmatrix}} = \frac{\begin{bmatrix} 0.6852 & 6.8525 \\ 6.8525 & 68.5252 \end{bmatrix}}{83.6008} = \\ &= \begin{bmatrix} 8.1960 \cdot 10^{-3} & 0.0819 \\ 0.0819 & 0.8196 \end{bmatrix} \end{aligned} \quad (8.545)$$

V nadaljevanju je potrebno v skladu s sliko 89 izračunati:

$$\begin{aligned} \underline{\underline{N}}_0 &= \frac{\underline{\underline{G}}_0 \cdot \Delta \underline{g}_0 \cdot \Delta \underline{g}_0^T \cdot \underline{\underline{G}}_0}{\Delta \underline{g}_0^T \cdot \underline{\underline{G}}_0 \cdot \Delta \underline{g}_0} = \frac{\begin{bmatrix} -9.9336 \\ -9.1058 \end{bmatrix} \cdot \begin{bmatrix} -9.9336 & -9.1058 \end{bmatrix}}{\begin{bmatrix} -9.9336 & -9.1058 \end{bmatrix} \cdot \begin{bmatrix} -9.9336 \\ -9.1058 \end{bmatrix}} = \\ &= \frac{\begin{bmatrix} 98.676 & 90.4533 \\ 90.4533 & 82.9155 \end{bmatrix}}{181.5920} = \begin{bmatrix} 0.5433 & 0.4981 \\ 0.4981 & 0.4566 \end{bmatrix} \end{aligned} \quad (8.546)$$

Na osnovi izrazov (8.536), (8.545) in (8.546) izračunamo še (glej sliko 89):

$$\begin{aligned}\underline{\underline{G}}_1 &= \underline{\underline{G}}_0 + \underline{\underline{M}}_0 - \underline{\underline{N}}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 8.1960 \cdot 10^{-3} & 0.0819 \\ 0.0819 & 0.8196 \end{bmatrix} - \begin{bmatrix} 0.5433 & 0.4981 \\ 0.4981 & 0.4566 \end{bmatrix} = \\ &= \begin{bmatrix} 0.4648 & -0.4162 \\ -0.4162 & 1.363 \end{bmatrix}\end{aligned}\quad (8.547)$$

$i = 1$ , 2. iteracija:

V 2. iteraciji najprej izračunamo nov vektor odločitvenih spremenljivk:

$$\begin{aligned}\underline{x}_2 &= \underline{x}_1 - \alpha_1 \cdot \underline{\underline{G}}_1 \cdot \underline{g}_1 = \begin{bmatrix} -0.8278 \\ -8.278 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} 0.4648 & -0.4162 \\ -0.4162 & 1.363 \end{bmatrix} \cdot \begin{bmatrix} -8.9336 \\ 0.8942 \end{bmatrix} = \\ &= \begin{bmatrix} -0.8278 \\ -8.278 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} -4.5245 \\ 4.9369 \end{bmatrix} = \begin{bmatrix} -0.8278 + 4.5245\alpha_1 \\ -8.278 - 4.9369\alpha_1 \end{bmatrix}\end{aligned}\quad (8.548)$$

V točki (8.548) funkcija (8.535) zavzame vrednost:

$$\begin{aligned}f(\underline{x}_2) &= (-0.8278 + 4.5245\alpha_1)^2 + (-0.8278 + 4.5245\alpha_1) \cdot \\ &\quad \cdot (-8.278 - 4.9369\alpha_1) + \frac{1}{2} \cdot (-8.278 - 4.9369\alpha_1)^2 + \\ &\quad + (-0.8278 + 4.5245\alpha_1) + 10 \cdot (-8.278 - 4.9369\alpha_1) = \\ &= (-0.8278 + 4.5245\alpha_1)^2 + \\ &\quad + (6.8525 + 4.0867\alpha_1 - 37.4538\alpha_1 - 22.337\alpha_1^2) + \\ &\quad + \frac{1}{2} \cdot (-8.278 - 4.9369\alpha_1)^2 + (-0.8278 + 4.5245\alpha_1) - \\ &\quad - 82.780 - 49.369\alpha_1 =\end{aligned}\quad (8.549)$$

$\alpha_1^*$  poiščemo na analitičen način:

$$\begin{aligned}\frac{df}{d\alpha_1} &= 2 \cdot (-0.8278 + 4.5245\alpha_1) \cdot 4.5245 + 4.0867 - 37.4538 - 44.674\alpha_1 + \\ &\quad + (-8.278 - 4.9369\alpha_1) \cdot (-4.9369) + 4.5245 - 49.369 = \\ &= -44.8347 + 20.6411\alpha_1 = 0 \quad \Rightarrow \alpha_1^* = \frac{44.8347}{20.6411} = 2.1720\end{aligned}\quad (8.550)$$

Če izraz (8.550) vstavimo v izraz (8.548), dobimo:

$$\underline{x}_2 = \begin{bmatrix} -0.8278 + 2.1720 \cdot 4.5245 \\ -8.278 - 2.1720 \cdot 4.9369 \end{bmatrix} = \begin{bmatrix} 8.9994 \\ -19.0009 \end{bmatrix} \approx \begin{bmatrix} 9 \\ -19 \end{bmatrix} \quad (8.551)$$

Gradient funkcije (8.535) v točki (8.551) se glasi:

$$\underline{g}_2 = \nabla f(\underline{x}_2) = \begin{bmatrix} 2x_1 + x_2 + 1 \\ x_1 + x_2 + 10 \end{bmatrix} \Bigg|_{\substack{x_1=9 \\ x_2=-19}} = \begin{bmatrix} 2 \cdot 9 - 19 + 1 \\ 9 - 19 + 10 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.552)$$

Ker pride gradient v izrazu (8.552) enak 0, končamo s postopkom.

Končna rešitev torej je:

$$\underline{x}^* = \underline{x}_2 = \begin{bmatrix} 9 \\ -19 \end{bmatrix} \quad (8.553)$$

Konvergenca je bila ponovno zelo hitra.

### Primer 3:

Poiščite minimum funkcije:

$$f(x_1, x_2) = 2x_1^2 + x_2^2 - x_1 x_2, \quad \underline{x}_0 = [2 \quad 2]^T \quad (8.554)$$

Uporabite BFGS Quasi – Newtonovo metodo (glejte sliko 90!).

Najprej izberemo začetno metrično matriko:

$$\underline{G}_0 = \underline{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (8.555)$$

$i = 0$ , 1. iteracija:

Na začetku 1. iteracije najprej izračunamo gradient v začetni točki:

$$\underline{g}_0 = \nabla f(\underline{x}_0) = \left. \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} \right|_{\underline{x}=\underline{x}_0} = \left. \begin{bmatrix} 4x_1 - x_2 \\ 2x_2 - x_1 \end{bmatrix} \right|_{\substack{x_1=2 \\ x_2=2}} = \begin{bmatrix} 4 \cdot 2 - 2 \\ 2 \cdot 2 - 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} \quad (8.556)$$

Nato izračunamo nov vektor odločitvenih spremenljivk:

$$\underline{x}_1 = \underline{x}_0 - \alpha_0 \cdot \underline{G}_0 \cdot \underline{g}_0 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \alpha_0 \cdot \underline{I} \cdot \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 - 6\alpha_0 \\ 2 - 2\alpha_0 \end{bmatrix} \quad (8.557)$$

V točki (8.557) funkcija (8.554) zavzame vrednost:

$$\begin{aligned} f(\underline{x}_1) &= 2 \cdot (2 - 6\alpha_0)^2 + (2 - 2\alpha_0)^2 - (2 - 6\alpha_0) \cdot (2 - 2\alpha_0) = \\ &= 2 \cdot (2 - 6\alpha_0)^2 + (2 - 2\alpha_0)^2 - (4 - 4\alpha_0 - 12\alpha_0 + 12\alpha_0^2) = \\ &= 2 \cdot (2 - 6\alpha_0)^2 + 4 \cdot (1 - \alpha_0)^2 - 4 + 16\alpha_0 - 12\alpha_0^2 \end{aligned} \quad (8.558)$$

$\alpha_0^*$  poiščemo na analitičen način:

$$\begin{aligned} \frac{df}{d\alpha_0} &= 4 \cdot (2 - 6\alpha_0) \cdot (-6) + 8 \cdot (1 - \alpha_0) \cdot (-1) + 16 - 24\alpha_0 = 0 \\ &-12 + 36\alpha_0 - 2 + 2\alpha_0 + 4 - 6\alpha_0 = 0 \\ &-10 + 32\alpha_0 = 0 \quad \Rightarrow \alpha_0^* = \frac{10}{32} = 0.3125 \end{aligned} \quad (8.559)$$

Če izraz (8.559) vstavimo v izraz (8.557), dobimo:

$$\underline{x}_1 = \begin{bmatrix} 2 - 6 \cdot 0.3125 \\ 2 - 2 \cdot 0.3125 \end{bmatrix} = \begin{bmatrix} 0.125 \\ 1.375 \end{bmatrix} \quad (8.560)$$

Gradient funkcije (8.554) v točki (8.560) se glasi:

$$\underline{g}_1 = \nabla f(\underline{x}_1) = \left. \begin{bmatrix} 4x_1 - x_2 \\ 2x_2 - x_1 \end{bmatrix} \right|_{\substack{x_1=0.125 \\ x_2=1.375}} = \begin{bmatrix} 4 \cdot 0.125 - 1.375 \\ 2 \cdot 1.375 - 0.125 \end{bmatrix} = \begin{bmatrix} -0.875 \\ 2.625 \end{bmatrix} \quad (8.561)$$

Spremembo gradienta pa dobimo, če tvorimo razliko med izrazoma (8.561) in (8.556):

$$\Delta \underline{g}_0 = \underline{g}_1 - \underline{g}_0 = \begin{bmatrix} -0.875 \\ 2.625 \end{bmatrix} - \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} -6.875 \\ 0.625 \end{bmatrix} \quad (8.562)$$

Nato izračunamo vektor  $\Delta \underline{x}_0$ , pri čemer od vektorja (8.560) odštejemo začetno točko:

$$\Delta \underline{x}_0 = \underline{x}_1 - \underline{x}_0 = \begin{bmatrix} 0.125 \\ 1.375 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -1.875 \\ -0.625 \end{bmatrix} \quad (8.563)$$

V nadaljevanju je potrebno v skladu s sliko 90 izračunati naslednje matrike:

$$\begin{aligned} \underline{\underline{A}}_0 &= \left[ 1 + \frac{\Delta \underline{g}_0^T \cdot \underline{\underline{G}}_0 \cdot \Delta \underline{g}_0}{\Delta \underline{x}_0^T \cdot \Delta \underline{g}_0} \right] \cdot \frac{\Delta \underline{x}_0 \cdot \Delta \underline{x}_0^T}{\Delta \underline{x}_0^T \cdot \Delta \underline{g}_0} = \\ &= \left[ 1 + \frac{\begin{bmatrix} -6.875 & 0.625 \end{bmatrix} \cdot \begin{bmatrix} -6.875 \\ 0.625 \end{bmatrix}}{\begin{bmatrix} -1.875 & -0.625 \end{bmatrix} \cdot \begin{bmatrix} -6.875 \\ 0.625 \end{bmatrix}} \right] \cdot \frac{\begin{bmatrix} -1.875 \\ -0.625 \end{bmatrix} \cdot \begin{bmatrix} -1.875 & -0.625 \end{bmatrix}}{\begin{bmatrix} -1.875 & -0.625 \end{bmatrix} \cdot \begin{bmatrix} -6.875 \\ 0.625 \end{bmatrix}} = \\ &= \left[ 1 + \frac{47.6562}{12.3312} \right] \cdot \frac{\begin{bmatrix} 3.5156 & 1.1718 \\ 1.1718 & 0.3906 \end{bmatrix}}{12.5} = 0.3891 \cdot \begin{bmatrix} 3.5156 & 1.1718 \\ 1.1718 & 0.3906 \end{bmatrix} = \\ &= \begin{bmatrix} 1.3681 & 0.4559 \\ 0.4559 & 0.1519 \end{bmatrix} \end{aligned} \quad (8.564)$$

$$\begin{aligned} \underline{\underline{B}}_0 &= \frac{\Delta \underline{x}_0 \cdot \Delta \underline{g}_0^T \cdot \underline{\underline{G}}_0}{\Delta \underline{x}_0^T \cdot \Delta \underline{g}_0} = \frac{\begin{bmatrix} -1.875 \\ -0.625 \end{bmatrix} \cdot \begin{bmatrix} -6.875 & 0.625 \end{bmatrix}}{\begin{bmatrix} -1.875 & -0.625 \end{bmatrix} \cdot \begin{bmatrix} -6.875 \\ 0.625 \end{bmatrix}} = \frac{\begin{bmatrix} 12.8906 & -1.1718 \\ 4.2968 & -0.3906 \end{bmatrix}}{12.5} = \\ &= \begin{bmatrix} 1.0312 & -0.0937 \\ 0.3437 & -0.0312 \end{bmatrix} \end{aligned} \quad (8.565)$$

$$\begin{aligned} \underline{\underline{C}}_0 &= \frac{\underline{\underline{G}}_0 \cdot \Delta \underline{g}_0 \cdot \Delta \underline{x}_0^T}{\Delta \underline{x}_0^T \cdot \Delta \underline{g}_0} = \frac{\begin{bmatrix} -6.875 \\ 0.625 \end{bmatrix} \cdot \begin{bmatrix} -1.875 & -0.625 \end{bmatrix}}{\begin{bmatrix} -1.875 & -0.625 \end{bmatrix} \cdot \begin{bmatrix} -6.875 \\ 0.625 \end{bmatrix}} = \frac{\begin{bmatrix} 12.8906 & 4.2968 \\ -1.1718 & -0.3906 \end{bmatrix}}{12.5} = \\ &= \begin{bmatrix} 1.0312 & 0.3437 \\ -0.0937 & -0.0312 \end{bmatrix} \end{aligned} \quad (8.566)$$

Nato pa je potrebno v skladu s sliko 90 izračunati še:

$$\begin{aligned} \underline{\underline{G}}_1 &= \underline{\underline{G}}_0 + \underline{\underline{A}}_0 - \underline{\underline{B}}_0 - \underline{\underline{C}}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1.3681 & 0.4559 \\ 0.4559 & 0.1519 \end{bmatrix} - \begin{bmatrix} 1.0312 & -0.0937 \\ 0.3437 & -0.0312 \end{bmatrix} - \\ & - \begin{bmatrix} 1.0312 & 0.3437 \\ -0.0937 & -0.0312 \end{bmatrix} = \begin{bmatrix} 0.3057 & 0.2059 \\ 0.2059 & 1.2143 \end{bmatrix} \end{aligned} \quad (8.567)$$

$i = i + 1$ , 2. iteracija:

Na začetku 2. iteracije najprej izračunamo nov vektor odločitvenih spremenljivk:

$$\begin{aligned} \underline{x}_2 &= \underline{x}_1 - \alpha_1 \cdot \underline{\underline{G}}_1 \cdot \underline{g}_1 = \begin{bmatrix} 0.125 \\ 1.375 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} 0.3057 & 0.2059 \\ 0.2059 & 1.2143 \end{bmatrix} \cdot \begin{bmatrix} -0.875 \\ 2.625 \end{bmatrix} = \\ & = \begin{bmatrix} 0.125 \\ 1.375 \end{bmatrix} - \alpha_1 \cdot \begin{bmatrix} 0.273 \\ 3.007 \end{bmatrix} = \begin{bmatrix} 0.125 - 0.273\alpha_1 \\ 1.375 - 3.007\alpha_1 \end{bmatrix} \end{aligned} \quad (8.568)$$

V točki (8.568) funkcija (8.554) zavzame vrednost:

$$\begin{aligned} f(\underline{x}_2) &= 2 \cdot (0.125 - 0.273\alpha_1)^2 + (1.375 - 3.007\alpha_1)^2 - \\ & - (0.125 - 0.273\alpha_1) \cdot (1.375 - 3.007\alpha_1) = \\ & = 2 \cdot (0.125 - 0.273\alpha_1)^2 + (1.375 - 3.007\alpha_1)^2 - \\ & - \{0.1718 - 0.7512\alpha_1 + 0.8209\alpha_1^2\} \end{aligned} \quad (8.569)$$

$\alpha_1^*$  poiščemo na analitičen način:

$$\begin{aligned} \frac{df}{d\alpha_1} &= 4 \cdot (0.125 - 0.273\alpha_1) \cdot (-0.273) + 2 \cdot (1.375 - 3.007\alpha_1) \cdot (-3.007) + \\ & + 0.7512 - 2 \cdot 0.8209\alpha_1 = \\ & = -7.6545 + 16.7404\alpha_1 = 0 \quad \Rightarrow \alpha_1^* = \frac{7.6545}{16.7404} = 0.4572 \end{aligned} \quad (8.570)$$

Če izraz (8.570) vstavimo v izraz (8.568), dobimo:

$$\underline{x}_2 = \begin{bmatrix} 0.125 - 0.4572 \cdot 0.273 \\ 1.375 - 0.4572 \cdot 3.007 \end{bmatrix} = \begin{bmatrix} 1.84 \cdot 10^{-4} \\ 1.99 \cdot 10^{-4} \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.571)$$

Gradient funkcije (8.554) v točki (8.571) se glasi:

$$\underline{g}_2 = \nabla \underline{f}(\underline{x}_2) = \begin{bmatrix} 4x_1 - x_2 \\ 2x_2 - x_1 \end{bmatrix} \Bigg|_{\substack{x_1=0 \\ x_2=0}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.572)$$

Ker pride gradient v izrazu (8.572) enak 0, končamo s postopkom. Zelo hitra konvergenca!

## 8.5 Primeri geometrijskega programiranja

V nadaljevanju si bomo pogledali nekaj primerov geometrijskega programiranja, pri čemer se bomo omejili na:

- Geometrijsko programiranje brez omejitev, ter
- Geometrijsko programiranje z omejitvami.

### 8.5.1 Geometrijsko programiranje brez omejitev

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 7.1, ter izraze (7.1) do (7.161).

#### Primer 1:

Opravka imamo z minimizacijo skupnih stroškov proizvodnje in skladiščenja določenega proizvoda, pri čemer se izkaže, da ima stroškovna funkcija naslednjo obliko:

$$f(x_1, x_2) = C_1 \cdot x_1^{-3} x_2^{-2} + C_2 \cdot x_1^3 x_2^1 + C_3 \cdot x_1^{-3} x_2^3 \quad (8.573)$$

kjer je:

$x_1$  - število ton proizvoda v določenem časovnem obdobju,

$x_2$  - število ton proizvoda, ki gre v skladišče v tem časovnem obdobju, ter

$(C_1, C_2, C_3) = (50, 30, 10) \rightarrow$  stroškovni koeficienti, odvisni od časovne periode in pogojev produkcije.



Stroškovno funkcijo zapišimo v obliki primerni za GP:

$$f(x_1, x_2) = \underbrace{50 \cdot x_1^{-3} x_2^{-2}}_{U_1} + \underbrace{30 \cdot x_1^3 x_2^1}_{U_2} + \underbrace{10 \cdot x_1^{-3} x_2^3}_{U_3} \quad (8.574)$$

kjer je:

$U_1$  - stroški nabave materiala in priprave proizvodnje,

$U_2$  - fiksni stroški proizvodnje in izgub časa, ter

$U_3$  - stroški skladiščenja in rokovanja s proizvodom.

Matriko  $\underline{A}$  lahko neposredno zapišemo, če koeficiente v  $f$  gledamo kot stolpce:

$$\underline{A} = \begin{bmatrix} -3 & 3 & -3 \\ -2 & 1 & 3 \end{bmatrix} \quad (8.575)$$

Stopnja težavnosti reševanja problema je sledeča:

$$\left. \begin{array}{l} N = 3 \\ n + 1 = 2 + 1 = 3 \end{array} \right\} \Rightarrow T = N - (n + 1) = 3 - 3 = 0 \quad (8.576)$$

Iz izraza (8.576) sledi, da imamo opravka z 0. stopnjo težavnosti problema.

Pogoj ortogonalnosti da:

$$\underline{A} \cdot \underline{y}^* = \begin{bmatrix} -3 & 3 & -3 \\ -2 & 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.577)$$

Na osnovi pogoja normiranosti sledi:

$$\sum_{i=1}^{N=3} y_i^* = 1 \Rightarrow y_1^* + y_2^* + y_3^* = 1 \quad (8.578)$$

Sledi:

$$\begin{bmatrix} -3 & 3 & -3 \\ -2 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (8.579)$$

Sledi:

$$\left. \begin{array}{l} E_1: -3y_1^* + 3y_2^* - 3y_3^* = 0 \\ E_2: -2y_1^* + y_2^* + 3y_3^* = 0 \end{array} \right\} + \Rightarrow -5y_1^* + 4y_2^* = 0 \Rightarrow \underbrace{y_1^* = \frac{4}{5}y_2^*}_{E_{1,2}} \quad (8.580)$$

$$E_3: y_1^* + y_2^* + y_3^* = 1$$

Vstavimo  $E_{1,2}$  v  $E_2$  in dobimo:

$$\begin{aligned} -2 \cdot \frac{4}{5} \cdot y_2^* + y_2^* + 3y_3^* &= 0 \\ -\frac{3}{5}y_2^* + 3y_3^* &= 0 \Rightarrow y_3^* = \frac{1}{5}y_2^* \end{aligned} \quad (8.581)$$

Vstavimo  $E_{1,2}$  in izraz (8.581) v enačbo  $E_3$  in dobimo:

$$\begin{aligned} \frac{4}{5}y_2^* + y_2^* + \frac{1}{5}y_2^* &= 1 \Rightarrow 2y_2^* = 1 \Rightarrow y_2^* = \frac{1}{2} \\ y_1^* &= \frac{4}{5} \cdot \frac{1}{2} = \frac{2}{5}, \quad y_3^* = \frac{1}{5} \cdot \frac{1}{2} = \frac{1}{10} \end{aligned} \quad (8.582)$$

Optimalna kriterijska funkcija je:

$$\begin{aligned} f^*(x_1, x_2) &= \prod_{j=1}^3 \left( \frac{C_j}{y_j^*} \right)^{y_j^*} = \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} = \\ &= \left( \frac{50}{\frac{2}{5}} \right)^{\frac{2}{5}} \cdot \left( \frac{30}{\frac{1}{2}} \right)^{\frac{1}{2}} \cdot \left( \frac{10}{\frac{1}{10}} \right)^{\frac{1}{10}} = (125)^{\frac{2}{5}} \cdot (60)^{\frac{1}{2}} \cdot (100)^{\frac{1}{10}} = 84.69 \end{aligned} \quad (8.583)$$

Izračunajmo še optimalne odločitvene spremenljivke:

$$\begin{aligned} U_1^* &= \frac{50}{x_1^{*3} \cdot x_2^{*2}} = f^* \cdot y_1^* = 84.69 \cdot \frac{2}{5} \\ U_2^* &= \frac{30 \cdot x_1^{*3} \cdot x_2^*}{1} = f^* \cdot y_2^* = 84.69 \cdot \frac{1}{2} \\ U_3^* &= \frac{10}{x_1^{*3} \cdot x_2^{*3}} = f^* \cdot y_3^* = 84.69 \cdot \frac{1}{10} \end{aligned} \quad (8.584)$$

Sledi:

$$\left. \begin{array}{l} E_1 : x_1^{*3} \cdot x_2^{*2} = 1.4759 \\ E_2 : x_1^{*3} \cdot x_2^* = 1.4115 \\ E_3 : x_1^{*3} \cdot x_2^{*3} = 1.1807 \end{array} \right\} \div \Rightarrow \frac{x_1^{*3} \cdot x_2^{*2}}{x_1^{*3} \cdot x_2^*} = \frac{1.4759}{1.4115} \Rightarrow x_2^* = 1.0456 \quad (8.585)$$

Dobimo:

$$x_1^* = \sqrt[3]{\frac{1.4759}{1.0456^2}} = 1.1052 \quad (8.586)$$

Torej, če bomo proizvedli 1.1052 tone proizvoda in ga skladiščili 1.0456 tone v danem časovnem obdobju, bomo dosegli minimalne skupne stroške proizvodnje in skladiščenja.

### Primer 2:

Opravka imamo z obdelavo odpadne snovi. Celotni strošek se sestoji iz kemične obdelave te snovi, njenega razredčenja, zahtev glede črpanja snovi ter stroškov nabave ustreznih cevi. Kot se izkaže, je skupna stroškovna funkcija dana z:

$$C(D, Q) = 150D + \frac{972000Q^2}{D^5} + \frac{432}{Q} \quad (8.587)$$

kjer se  $C$  meri v €,  $D$  v  $cm$  in  $Q$  v  $m^3 / sec$  ( $D$  je premer pipe,  $Q$  pa pretok). S pomočjo geometrijskega programiranja poiščite takšna optimalna  $D^*$  in  $Q^*$ , da bodo celotni stroški minimalni.

Zapišimo v primerni obliki:

$$C = \underbrace{150D^1 Q^0}_{v_1} + \underbrace{972000D^{-5} Q^2}_{v_2} + \underbrace{432D^0 Q^{-1}}_{v_3}; \quad (8.588)$$

$$C_1 = 150, \quad C_2 = 972000, \quad C_3 = 432$$

Zapišimo matriko  $\underline{\underline{A}}$ :

$$\underline{\underline{A}} = \begin{bmatrix} 1 & -5 & 0 \\ 0 & 2 & -1 \end{bmatrix} \quad (8.589)$$

Stopnja težavnosti problema GP je sledeča:

$$\left. \begin{array}{l} N = 3 \\ n + 1 = 2 + 1 = 3 \end{array} \right\} \Rightarrow T = N - (n + 1) = 3 - 3 = 0 \quad (8.590)$$

Iz izraza (8.590) sledi, da imamo ponovno opravka z 0. stopnjo težavnosti problema.

Pogoj ortogonalnosti da:

$$\underline{A} \cdot \underline{y}^* = \begin{bmatrix} 1 & -5 & 0 \\ 0 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8.591)$$

Pogoj normiranosti:

$$\sum_{j=1}^{N=3} y_j^* = 1 \Rightarrow y_1^* + y_2^* + y_3^* = 1 \quad (8.592)$$

Sledi:

$$\begin{bmatrix} 1 & -5 & 0 \\ 0 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (8.593)$$

Sledi:

$$\left. \begin{array}{l} E_1: y_1^* - 5y_2^* = 0 \Rightarrow y_1^* = 5y_2^* \\ E_2: 2y_2^* - y_3^* = 0 \Rightarrow y_3^* = 2y_2^* \end{array} \right\} E_{1,2} \quad (8.594)$$

$$E_3: y_1^* + y_2^* + y_3^* = 1$$

Vstavimo  $E_{1,2}$  v enačbo  $E_3$  in dobimo:

$$\begin{aligned} 5y_2^* + y_2^* + 2y_2^* = 1 & \Rightarrow y_2^* = \frac{1}{8} \\ y_1^* = 5 \cdot \frac{1}{8} = \frac{5}{8}, \quad y_3^* = 2 \cdot \frac{1}{8} = \frac{1}{4} \end{aligned} \quad (8.595)$$

Optimalna kriterijska funkcija je:

$$\begin{aligned}
C^*(D, Q) &= \prod_{j=1}^3 \left( \frac{C_j}{y_j^*} \right)^{y_j^*} = \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} = \\
&= \left( \frac{150}{\frac{5}{8}} \right)^{\frac{5}{8}} \cdot \left( \frac{972000}{\frac{1}{8}} \right)^{\frac{1}{8}} \cdot \left( \frac{432}{\frac{1}{4}} \right)^{\frac{1}{4}} = \\
&= (240)^{\frac{5}{8}} \cdot (7776000)^{\frac{1}{8}} \cdot (1728)^{\frac{1}{4}} = 1440 \text{ €}
\end{aligned} \tag{8.596}$$

Izračunajmo še optimalni odločitveni spremenljivki:

$$\begin{aligned}
U_1^* = 150D^* = C^* \cdot y_1^* = 1440 \cdot \frac{5}{8} &\Rightarrow D^* = \frac{1440 \cdot 5}{150 \cdot 8} = 6 \text{ cm} \\
U_2^* = \dots \rightarrow \text{raje bomo } Q^* \text{ izrazili iz } U_3^* & \\
U_3^* = \frac{432}{Q^*} = C^* \cdot y_3^* = 1440 \cdot \frac{1}{4} &\Rightarrow Q^* = \frac{4 \cdot 432}{1440} = 1.2 \frac{\text{m}^3}{\text{s}}
\end{aligned} \tag{8.597}$$

Torej, če bomo načrtovali pipe s premerom 6cm ter skozi njih spuščali pretok 1.2 m<sup>3</sup> / s, bomo dosegli minimalne skupne stroške 1440€.

### Primer 3:

Naftno podjetje želi načrtati plinski transportni cevovod in ugotovi, da se skupni letni stroški gibljejo po naslednji zakonitosti:

$$C(L, F, D) = \overbrace{4.55 \cdot 10^5}^{C_1} \frac{L^{\frac{1}{2}}}{F^{0.387} \cdot D^{\frac{2}{3}}} + \overbrace{3.69 \cdot 10^4}^{C_2} D + \overbrace{6.57 \cdot 10^6}^{C_3} \frac{1}{L} + \overbrace{7.72 \cdot 10^8}^{C_4} \frac{F}{L} \tag{8.598}$$

kjer je  $L$  skupna dolžina cevi v  $m$ ,  $D$  je premer cevi v  $cm$ , ter je  $F = r^{0.219} - 1$ ,  $r$  pa je razmerje med notranjim in zunanjim tlakom. Poiščite takšne optimalne  $L^*$ ,  $D^*$ ,  $F^*$  in  $r^*$ , da bomo dosegli minimalne skupne stroške.

Funkcijo, dano v izrazu (8.598) zapišimo v obliki primerni za GP:

$$C(L, F, D) = \underbrace{4.55 \cdot 10^5 \cdot L^{0.5} \cdot F^{-0.3871} \cdot D^{\frac{2}{3}}}_{u_1} + \underbrace{3.69 \cdot 10^4 \cdot L^0 \cdot F^0 \cdot D^1}_{u_2} + \underbrace{6.57 \cdot 10^6 \cdot L^{-1} \cdot F^0 \cdot D^0}_{u_3} + \underbrace{7.72 \cdot 10^8 \cdot L^{-1} \cdot F^1 \cdot D^0}_{u_4} \quad (8.599)$$

Matrika  $\underline{\underline{A}}$  se glasi:

$$\underline{\underline{A}} = \begin{bmatrix} 0.5 & 0 & -1 & -1 \\ -0.387 & 0 & 0 & 1 \\ -\frac{2}{3} & 1 & 0 & 0 \end{bmatrix} \quad (8.600)$$

Stopnja težavnosti problema GP:

$$\left. \begin{array}{l} N = 4 \\ \underbrace{n}_{\substack{\text{saj} \\ L, F, D}} + 1 = 3 + 1 = 4 \end{array} \right\} \Rightarrow T = N - (n + 1) = 4 - 4 = 0 \Rightarrow \quad (8.601)$$

Iz izraza (8.601) sledi, da imamo ponovno opravka z 0. stopnjo težavnosti GP problema.

Pogoj ortogonalnosti:

$$\underline{\underline{A}} \cdot \underline{y}^* = \underline{0} \quad (8.602)$$

Pogoj normiranosti:

$$\sum_{j=1}^{N=4} y_j^* = 1 \quad (8.603)$$

Združimo in dobimo:

$$\begin{bmatrix} 0.5 & 0 & -1 & -1 \\ -0.387 & 0 & 0 & 1 \\ -\frac{2}{3} & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (8.604)$$

Sledi:

$$\begin{aligned}
 E_1: & 0.5y_1^* - y_3^* - y_4^* = 0 \\
 E_2: & -0.387y_1^* + y_4^* = 0 \quad \Rightarrow \quad y_4^* = 0.387y_1^* \\
 E_3: & -\frac{2}{3}y_1^* + y_2^* = 0 \quad \Rightarrow \quad y_2^* = \frac{2}{3}y_1^* \\
 E_4: & y_1^* + y_2^* + y_3^* + y_4^* = 1
 \end{aligned} \tag{8.605}$$

Vstavimo  $E_2$  v enačbo  $E_1$ :

$$0.5y_1^* - y_3^* - 0.387y_1^* = 0 \quad \Rightarrow \quad y_3^* = 0.113y_1^* \tag{8.606}$$

Vstavimo  $E_2$ ,  $E_3$  in izraz (8.606) v enačbo  $E_4$  in dobimo:

$$\begin{aligned}
 y_1^* + \frac{2}{3}y_1^* + 0.113y_1^* + 0.387y_1^* &= 1 \quad \Rightarrow \quad y_1^* = \frac{1}{2.166} = 0.4615 \\
 y_2^* &= \frac{2}{3} \cdot 0.4615 = 0.3076 \\
 y_3^* &= 0.113 \cdot 0.4615 = 0.0521 \\
 y_4^* &= 0.387 \cdot 0.4615 = 0.1786
 \end{aligned} \tag{8.607}$$

Optimalna kriterijska funkcija je:

$$\begin{aligned}
 C^*(F, L, D) &= \prod_{j=1}^4 \left( \frac{C_j}{y_j^*} \right)^{y_j^*} = \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4^*} \right)^{y_4^*} = \\
 &= \left( \frac{4.55 \cdot 10^5}{0.4615} \right)^{0.4615} \cdot \left( \frac{3.69 \cdot 10^4}{0.3076} \right)^{0.3076} \cdot \left( \frac{6.57 \cdot 10^6}{0.0521} \right)^{0.0521} \cdot \\
 &\quad \cdot \left( \frac{7.72 \cdot 10^8}{0.1786} \right)^{0.1786} = \\
 &= 583.656 \cdot 36.501 \cdot 2.6427 \cdot 52.5947 = 2.961 \cdot 10^6 \frac{\text{€}}{\text{leto}}
 \end{aligned} \tag{8.608}$$

Določimo še optimalne odločitvene spremenljivke:

$$U_1^* = 4.55 \cdot 10^5 \frac{(L^*)^{\frac{1}{2}}}{(F^*)^{0.387} \cdot (D^*)^{\frac{2}{3}}} = C^* \cdot y_1^* = 2.961 \cdot 10^6 \cdot 0.4615 \quad (M_1)$$

$$U_2^* = 3.69 \cdot 10^4 D^* = C^* \cdot y_2^* = 2.961 \cdot 10^6 \cdot 0.3076 \quad (M_2) \quad (8.609)$$

$$U_3^* = \frac{6.57 \cdot 10^6}{L^*} = C^* \cdot y_3^* = 2.961 \cdot 10^6 \cdot 0.0521 \quad (M_3)$$

$$U_4^* = 7.72 \cdot 10^8 \frac{F^*}{L^*} = C^* \cdot y_4^* = 2.961 \cdot 10^6 \cdot 0.1786 \quad (M_4)$$

Iz  $M_2$  sledi:

$$D^* = \frac{2.961 \cdot 10^6 \cdot 0.3076}{3.69 \cdot 10^4} = 24.683 \text{ cm} \quad (8.610)$$

Iz  $M_3$  sledi:

$$L^* = \frac{6.57 \cdot 10^6}{2.961 \cdot 10^6 \cdot 0.0521} = 42.588 \text{ m} \quad (8.611)$$

Iz  $M_4$  sledi:

$$F^* = \frac{2.961 \cdot 10^6 \cdot 0.1786 \cdot 42.588}{7.72 \cdot 10^8} = 0.029 = (r^*)^{0.219} - 1 \quad (8.612)$$

Sledi:

$$1.029 = (r^*)^{0.219} \quad \Bigg| \ln$$

$$\ln 1.029 = 0.219 \cdot \ln r^* \quad (8.613)$$

$$\ln r^* = \frac{0.02858}{0.219} = 0.1305$$

$$r^* = e^{0.1305} = 1.1393$$

Torej, če bomo vpeljali skupno dolžino cevi 42.588 m, premer cevi 24.683 cm, ter razmerje med notranjim in zunanjim tlakom 1.1393, bomo dosegli minimalne skupne letne stroške v višini 2.961 milijona €.



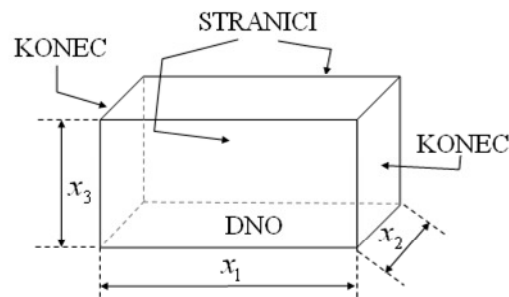
**Primer 4:**

$400m^3$  nekega materiala je potrebno transportirati čez reko. Material naj se prenese v odprti škatli dolžine  $x_1$  metrov, širine  $x_2$  metrov in višine  $x_3$  metrov (glej sliko 118).

Izdelava dna škatle stane  $20 \frac{\text{€}}{m^2}$ , izdelava stranic stane  $5 \frac{\text{€}}{m^2}$ , izdelava koncev škatle pa

stane  $20 \frac{\text{€}}{m^2}$ . Vsak prevoz na trajektu stane 0.1€. Poiščite takšne optimalne dimenzije

škatle, pri katerih se dosežejo minimalni skupni stroški izdelave in transporta škatle.



Slika 118: Škatla

Problem je sledeč:

$$\underbrace{\text{SKUPNI STROŠKI}}_f = \underbrace{\text{STROŠKI ŠKATLE}}_{f_1} + \underbrace{\text{STROŠKI PREVOZA ČEZ REKO}}_{f_2} \quad (8.614)$$

kjer je  $f_1$  definiran z naslednjim izrazom:

$$f_1 = \underbrace{\text{STROŠKI STRANIC}}_{s_1} + \underbrace{\text{STROŠKI DNA}}_{s_2} + \underbrace{\text{STROŠKI KONCEV}}_{s_3} \quad (8.615)$$

$f_2$  pa z izrazom (8.616):

$$f_2 = \underbrace{\text{ŠTEVILO TRANSPORTOV}}_{s_4} \cdot \underbrace{\text{STROŠEK ZA VSAK TRANSPORT}}_{s_5} \quad (8.616)$$

Stroški izdelave škatle so definirani z relacijami, podanih v izrazu (8.617):

$$\begin{aligned}
S_1 &= 2x_1x_3 \cdot 5 = 10x_1x_3 \\
S_2 &= x_1x_2 \cdot 20 = 20x_1x_2 \\
S_3 &= 2x_2x_3 \cdot 20 = 40x_2x_3
\end{aligned}
\tag{8.617}$$

Stroški transporta škatle pa so definirani na sledeč način:

$$\begin{aligned}
S_4 &= \frac{400m^3}{x_1x_2x_3} = \frac{400}{V} \\
S_5 &= 0.1
\end{aligned}
\tag{8.618}$$

Če v izrazu (8.615) upoštevamo relacije, podane v (8.617), dobimo:

$$f_1 = 10x_1x_3 + 20x_1x_2 + 40x_2x_3 \tag{8.619}$$

Če v izrazu (8.616) upoštevamo relaciji, podani v (8.618), dobimo:

$$f_2 = \frac{400}{x_1x_2x_3} \cdot 0.1 = \frac{40}{x_1x_2x_3} \tag{8.620}$$

Stroškovna funkcija je torej naslednja:

$$f(x_1, x_2, x_3) = \underbrace{10}_{c_1} x_1x_3 + \underbrace{20}_{c_2} x_1x_2 + \underbrace{40}_{c_3} x_2x_3 + \frac{\underbrace{40}_{c_4}}{x_1x_2x_3} \tag{8.621}$$

Če funkcijo (8.621) prevedemo v obliko, primerno za GP, dobimo:

$$f(x_1, x_2, x_3) = \underbrace{10x_1^1x_2^0x_3^1}_{u_1} + \underbrace{20x_1^1x_2^1x_3^0}_{u_2} + \underbrace{40x_1^0x_2^1x_3^1}_{u_3} + \underbrace{40x_1^{-1}x_2^{-1}x_3^{-1}}_{u_4} \tag{8.622}$$

Zapišimo matriko  $\underline{\underline{A}}$ :

$$\underline{\underline{A}} = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \end{bmatrix} \tag{8.623}$$

Ortogonalnost:

$$\underline{\underline{A}} \cdot \underline{\underline{y}}^* = \underline{\underline{0}} \tag{8.624}$$

Normiranost:

$$\sum_{j=1}^{N=4} y_j^* = 1 \quad (8.625)$$

Stopnja težavnosti:

$$T = N - (n+1) = 4 - (3+1) = 0 \quad (8.626)$$

Ponovno imamo opravka z 0. stopnjo težavnosti GP problema!

Sledi (ortogonalnost + normiranost):

$$\begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (8.627)$$

Sledi:

$$\begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (8.628)$$

Dobimo:

$$\begin{aligned} E_1 : y_1^* + y_2^* - y_4^* &= 0 & \Rightarrow y_4^* &= y_1^* + y_2^* \\ E_2 : y_2^* + y_3^* - y_4^* &= 0 & \Rightarrow y_4^* &= y_2^* + y_3^* \\ E_3 : y_1^* + y_3^* - y_4^* &= 0 & \Rightarrow y_4^* &= y_1^* + y_3^* \\ E_4 : y_1^* + y_2^* + y_3^* + y_4^* &= 1 \end{aligned} \quad (8.629)$$

Izenačimo  $E_1$  in  $E_2$  ter dobimo:

$$y_1^* + y_2^* = y_2^* + y_3^* \Rightarrow y_1^* = y_3^* \quad (8.630)$$

Izenačimo  $E_1$  in  $E_3$ , dobimo:

$$y_1^* + y_2^* = y_1^* + y_3^* \Rightarrow y_2^* = y_3^* \quad (8.631)$$

Vstavimo (8.630) in (8.631) v  $E_4$  in dobimo:

$$y_1^* + y_1^* + y_1^* + y_4^* = 1 \Rightarrow y_4^* = 1 - 3y_1^* \quad (8.632)$$

Sedaj pa vstavimo (8.632) v  $E_1$  in dobimo:

$$\begin{aligned} 1 - 3y_1^* &= y_1^* + y_1^* \Rightarrow y_1^* = \frac{1}{5} = y_2^* = y_3^* \\ y_4^* &= 1 - \frac{3}{5} = \frac{2}{5} \end{aligned} \quad (8.633)$$

Sledi:

$$(y_1^*, y_2^*, y_3^*, y_4^*) = \left( \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{2}{5} \right) \quad (8.634)$$

Optimalna stroškovna funkcija je:

$$\begin{aligned} f^* &= \prod_{j=1}^4 \left( \frac{C_j}{y_j^*} \right)^{y_j^*} = \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4^*} \right)^{y_4^*} \\ &= \left( \frac{10}{\frac{1}{5}} \right)^{\frac{1}{5}} \cdot \left( \frac{20}{\frac{1}{5}} \right)^{\frac{1}{5}} \cdot \left( \frac{40}{\frac{1}{5}} \right)^{\frac{1}{5}} \cdot \left( \frac{40}{\frac{2}{5}} \right)^{\frac{2}{5}} = 50^{\frac{1}{5}} \cdot 100^{\frac{1}{5}} \cdot 200^{\frac{1}{5}} \cdot 100^{\frac{2}{5}} \\ &= (50 \cdot 100 \cdot 200)^{\frac{1}{5}} \cdot 100^{\frac{2}{5}} = 100\text{€} \end{aligned} \quad (8.635)$$

Določimo še optimalne dimenzije škatle:

$$\begin{aligned} U_1^* &= 10x_1^* x_3^* = f^* \cdot y_1^* = 100 \cdot \frac{1}{5} = 20 \\ U_2^* &= 20x_1^* x_2^* = f^* \cdot y_2^* = 100 \cdot \frac{1}{5} = 20 \\ U_3^* &= 40x_2^* x_3^* = f^* \cdot y_3^* = 100 \cdot \frac{1}{5} = 20 \\ U_4^* &= \frac{40}{x_1^* x_2^* x_3^*} = f^* \cdot y_4^* = 100 \cdot \frac{2}{5} = 40 \end{aligned} \quad (8.636)$$

Sledi:

$$\begin{aligned}
E_1: & \quad x_1^* x_3^* = 2 \\
E_2: & \quad x_1^* x_2^* = 1 \\
E_3: & \quad x_2^* x_3^* = \frac{1}{2} \\
E_4: & \quad x_1^* x_2^* x_3^* = 1
\end{aligned}
\tag{8.637}$$

Delimo  $E_1$  in  $E_4$  in dobimo:

$$\begin{aligned}
\frac{x_1^* x_3^*}{x_1^* x_2^* x_3^*} &= \frac{2}{1} \Rightarrow x_2^* = \frac{1}{2} \\
x_1^* &= \frac{1}{x_2^*} = \frac{1}{\frac{1}{2}} = 2 \\
x_3^* &= \frac{1}{x_1^* x_2^*} = \frac{1}{2 \cdot \frac{1}{2}} = 1
\end{aligned}
\tag{8.638}$$

Torej, če bomo zgradili škatlo dimenzij  $2m \times \frac{1}{2}m \times 1m$ , bomo dosegli minimalne skupne stroške izdelave in transporta škatle čez reko v višini 100€.

### Primer 5:

Poiščite minimum za naslednjo funkcijo:

$$f(\underline{x}) = 5 \frac{x_1}{x_2} + 2 \frac{x_2}{x_1} + 5x_1 + \frac{1}{x_2}.
\tag{8.639}$$

Uporabite postopek geometrijskega programiranja. Če je stopnja težavnosti večja od 0, uporabite postopek odvajanja logaritma. Rezultat potem spravite v polinomske oblike. Kakšna je potem optimalna rešitev, če je edina prava rešitev polinoma enaka 0.16274 (preverite to rešitev polinoma!).

Prevedimo funkcijo, dano v izrazu (8.639), v obliko primerno za GP:

$$f(\underline{x}) = \underbrace{5 x_1^1 x_2^{-1}}_{u_1} + \underbrace{2 x_1^{-1} x_2^1}_{u_2} + \underbrace{5 x_1^1 x_2^0}_{u_3} + \underbrace{1 x_1^0 x_2^{-1}}_{u_4}
\tag{8.640}$$

Stopnja težavnosti:

$$\left. \begin{array}{l} N = 4 \\ n + 1 = 2 + 1 = 3 \end{array} \right\} T = N - (n + 1) = 1 \quad (8.641)$$

Iz izraza (8.641) sledi, da imamo opravka z 1. stopnjo GP problema.

Zapišimo matriko  $\underline{\underline{A}}$ :

$$\underline{\underline{A}} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & -1 \end{bmatrix} \quad (8.642)$$

Spomnimo se še enkrat funkcije, dane v izrazu (8.639), ki jo je potrebno minimizirati:

$$f(\underline{x}) = 5 \frac{x_1}{x_2} + 2 \frac{x_2}{x_1} + 5x_1 + \frac{1}{x_2} \quad (8.643)$$

Ortogonalnost:

$$\underline{\underline{A}} \cdot \underline{y}^* = \underline{0} \quad (8.644)$$

Normiranost:

$$\sum_{j=1}^4 y_j^* = y_1^* + y_2^* + y_3^* + y_4^* = 1 \quad (8.645)$$

Ortogonalnost + normiranost nam da:

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (8.646)$$

Sledi:

$$\left. \begin{array}{l} E_1: y_1^* - y_2^* + y_3^* = 0 \\ E_2: -y_1^* + y_2^* - y_4^* = 0 \end{array} \right\} + \Rightarrow y_3^* = y_4^* \quad (8.647)$$

$$E_3: y_1^* + y_2^* + y_3^* + y_4^* = 1$$

Iz  $E_3$  sledi:

$$y_1^* + y_2^* + y_3^* + y_4^* = 1 \Rightarrow y_1^* + y_2^* + 2y_3^* = 1 \Rightarrow y_1^* = 1 - y_2^* - 2y_3^* \quad (8.648)$$

Iz  $E_1$  sledi:

$$y_1^* = y_2^* - y_3^* \quad (8.649)$$

Izenačimo (8.648) in (8.649):

$$y_2^* - y_3^* = 1 - y_2^* - 2y_3^* \Rightarrow 2y_2^* = 1 - y_3^* \Rightarrow y_2^* = \frac{1 - y_3^*}{2} \quad (8.650)$$

Iz  $E_1$  sledi (vstavimo (8.649)):

$$y_1^* - \frac{1 - y_3^*}{2} + y_3^* = 0 \Rightarrow y_1^* = \frac{1}{2} - \frac{3y_3^*}{2} = \frac{1}{2} \cdot (1 - 3y_3^*) \quad (8.651)$$

Dualna funkcija ima obliko (upoštevamo izraze (8.647), (8.651) in (8.650)).

$$\begin{aligned} f^*(\underline{y}) &= \prod_{j=1}^4 \left( \frac{C_j}{y_j^*} \right)^{y_j^*} = \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4^*} \right)^{y_4^*} \\ &= \left[ \frac{5}{\frac{1}{2} \cdot (1 - 3y_3^*)} \right]^{\frac{1}{2}(1 - 3y_3^*)} \cdot \left[ \frac{2}{\frac{1 - y_3^*}{2}} \right]^{\frac{1 - y_3^*}{2}} \cdot \left[ \frac{5}{y_3^*} \right]^{y_3^*} \cdot \left[ \frac{1}{y_3^*} \right]^{y_3^*} = \\ &= \left( \frac{10}{1 - 3y_3^*} \right)^{\frac{1 - 3y_3^*}{2}} \cdot \left( \frac{4}{1 - y_3^*} \right)^{\frac{1 - y_3^*}{2}} \cdot \left( \frac{5}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{1}{y_3^*} \right)^{y_3^*} \end{aligned} \quad (8.652)$$

Logaritem te funkcije ima obliko:

$$\begin{aligned} \ln f^* &= \frac{1 - 3y_3^*}{2} \cdot \ln \frac{10}{1 - 3y_3^*} + \frac{1 - y_3^*}{2} \cdot \ln \frac{4}{1 - y_3^*} + y_3^* \cdot \ln \frac{5}{y_3^*} + y_3^* \cdot \ln \frac{1}{y_3^*} = \\ &= \frac{1 - 3y_3^*}{2} \cdot [\ln 10 - \ln(1 - 3y_3^*)] + \frac{1 - y_3^*}{2} \cdot [\ln 4 - \ln(1 - y_3^*)] + \\ &\quad + y_3^* \cdot [\ln 5 - \ln y_3^*] + y_3^* \cdot \left[ \frac{\ln 1}{0} - \ln y_3^* \right] \end{aligned} \quad (8.653)$$

$$\ln f^* = \frac{1-3y_3^*}{2} \cdot [\ln 10 - \ln(1-3y_3^*)] + \frac{1-y_3^*}{2} \cdot [\ln 4 - \ln(1-y_3^*)] + y_3^* \cdot [\ln 5 - 2 \cdot \ln y_3^*]$$

Logaritem funkcije odvajamo po  $y_3^*$ :

$$\begin{aligned} \frac{d \ln f^*}{dy_3^*} &= -\frac{3}{2} \cdot [\ln 10 - \ln(1-3y_3^*)] + \frac{1-3y_3^*}{2} \cdot \left[ -\frac{1}{1-3y_3^*} \cdot (-3) \right] - \\ &\quad - \frac{1}{2} \cdot [\ln 4 - \ln(1-y_3^*)] + \\ &\quad + \frac{1-y_3^*}{2} \cdot \left[ -\frac{1}{1-y_3^*} \cdot (-1) \right] + \ln 5 - 2 \cdot \ln y_3^* + y_3^* \cdot \left[ \frac{-2}{y_3^*} \right] = \\ &= \left[ -\frac{3}{2} \cdot \ln 10 + \frac{3}{2} - \frac{1}{2} \cdot \ln 4 + \frac{1}{2} + \ln 5 - 2 \right] + \frac{3}{2} \cdot \ln(1-3y_3^*) + \\ &\quad + \frac{1}{2} \cdot \ln(1-y_3^*) - 2 \cdot \ln y_3^* = \\ &= \left[ -\ln 10^{\frac{3}{2}} - \ln 4^{\frac{1}{2}} - \ln 5^{-1} \right] + \left[ \ln(1-3y_3^*)^{\frac{3}{2}} + \ln(1-y_3^*)^{\frac{1}{2}} + \ln(y_3^*)^{-2} \right] = \\ &= \ln \left[ (1-3y_3^*)^{\frac{3}{2}} \cdot (1-y_3^*)^{\frac{1}{2}} \cdot (y_3^*)^{-2} \right] - \ln \left[ 10^{\frac{3}{2}} \cdot 4^{\frac{1}{2}} \cdot 5^{-1} \right] \end{aligned} \quad (8.654)$$

Odvod enačimo z 0 in dobimo:

$$\frac{d \ln f^*}{dy_3^*} = 0 \quad \Rightarrow \ln \left[ (1-3y_3^*)^{\frac{3}{2}} \cdot (1-y_3^*)^{\frac{1}{2}} \cdot (y_3^*)^{-2} \right] = \ln \left[ 10^{\frac{3}{2}} \cdot 4^{\frac{1}{2}} \cdot 5^{-1} \right] \quad (8.655)$$

Sledi:

$$\begin{aligned} (1-3y_3^*)^{\frac{3}{2}} \cdot (1-y_3^*)^{\frac{1}{2}} \cdot (y_3^*)^{-2} &= 10^{\frac{3}{2}} \cdot 4^{\frac{1}{2}} \cdot 5^{-1} \quad \Big| \quad ^2 \\ \frac{(1-3y_3^*)^3 \cdot (1-y_3^*)}{(y_3^*)^4} &= 160 \end{aligned} \quad (8.656)$$

$$\begin{aligned} (1-3y_3^*)^3 \cdot (1-y_3^*) &= 160 \cdot (y_3^*)^4 \\ (1-3y_3^*) \cdot (1-6y_3^* + 9y_3^{*2}) \cdot (1-y_3^*) &= 160y_3^{*4} \\ (1-6y_3^* + 9y_3^{*2})(1-4y_3^* + 3y_3^{*2}) &= 160y_3^{*4} \end{aligned}$$



$$\begin{aligned}
1 - 4y_3^* + 3y_3^{*2} - 6y_3^* + 24y_3^{*2} - 18y_3^{*3} + 9y_3^{*2} - 36y_3^{*3} + 27y_3^{*4} &= 160y_3^{*4} \\
1 - 10y_3^* + 36y_3^{*2} - 54y_3^{*3} &= 133y_3^{*4}
\end{aligned}$$

Sledi:

$$133y_3^{*4} + 54y_3^{*3} - 36y_3^{*2} + 10y_3^* - 1 = 0 \quad (8.657)$$

In smo v izrazu (8.657) prišli do polinomske oblike.

Preverimo rešitev polinoma 0.16274:

$$\begin{aligned}
133 \cdot 0.16274^4 + 54 \cdot 0.16274^3 - 36 \cdot 0.16274^2 + 10 \cdot 0.16274 - 1 &= \\
= -3.3 \cdot 10^{-6} \approx 0 \quad \text{OK} & \quad (8.658)
\end{aligned}$$

Torej imamo (glej izraze (8.647), (8.651) in (8.650)):

$$\begin{aligned}
y_3^* &= 0.16274 \\
y_1^* &= \frac{1}{2} \cdot (1 - 3 \cdot 0.16274) = 0.2558 \\
y_2^* &= \frac{1 - 0.16274}{2} = 0.4186 \\
y_4^* &= 0.16274
\end{aligned} \quad (8.659)$$

V izrazu (8.659) pa pridemo do dualnih rešitev.

Dualna funkcija zavzame optimalno vrednost:

$$\begin{aligned}
f^*(\underline{y}) &= \left(\frac{5}{0.2558}\right)^{0.2558} \cdot \left(\frac{2}{0.4186}\right)^{0.4186} \cdot \left(\frac{5}{0.16274}\right)^{0.16274} \cdot \left(\frac{1}{0.16274}\right)^{0.16274} = \\
&= 2.1392 \cdot 1.9245 \cdot 1.7461 \cdot 1.3437 = 9.6596
\end{aligned} \quad (8.660)$$

Poiščimo še optimalne odločitvene spremenljivke:

$$\begin{aligned}
U_1^* &= \frac{5x_1^*}{x_2^*} = f^* \cdot y_1^* = 9.6596 \cdot 0.2558 \\
U_2^* &= \frac{2x_2^*}{x_1^*} = f^* \cdot y_2^* = 9.6596 \cdot 0.4186
\end{aligned} \quad (8.661)$$

$$U_3^* = 5x_1^* = f^* \cdot y_3^* = 9.6596 \cdot 0.16274$$

$$U_4^* = \frac{1}{x_2^*} = f^* \cdot y_4^* = 9.6596 \cdot 0.16274$$

Uporabimo  $U_3$  in  $U_4$ , ter dobimo:

$$\begin{aligned} x_1^* &= 0.3144 \\ x_2^* &= 0.6361 \end{aligned} \quad (8.662)$$

### Primer 6:

Poiščite minimum naslednje funkcije:

$$f(x_1, x_2) = x_1^2 + 2x_2 + \frac{3}{x_1 x_2} + 2x_1 x_2 + x_2^2, \quad x_1 > 0, x_2 > 0. \quad (8.663)$$

Uporabite postopek GP, 2. način (brez odvajanja logaritma). Kolikšna je optimalna rešitev, če je rešitev dualnega sistema enačb enaka:  $y_3^* = 0.4589$ ,  $y_4^* = 0.1794$  (preverite!).

Funkcija, dana v izrazu (8.663), primerna za GP se glasi:

$$f(x_1, x_2) = \underbrace{1 x_1^2 x_2^0}_{u_1} + \underbrace{2 x_1^0 x_2^1}_{u_2} + \underbrace{3 x_1^{-1} x_2^{-1}}_{u_3} + \underbrace{2 x_1^1 x_2^1}_{u_4} + \underbrace{1 x_1^0 x_2^2}_{u_5} \quad (8.664)$$

Stopnja težavnosti:

$$\left. \begin{aligned} N &= 5 \\ n+1 &= 3 \end{aligned} \right\} T = N - (n+1) = 2 \quad (8.665)$$

Glede na izraz (8.665) imamo opravka z 2. stopnjo težavnosti. Zato bosta 2 spremenljivki prosti, 3 pa odvisne.

Zapišimo matriko  $\underline{A}$ :

$$\underline{A} = \begin{bmatrix} 2 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 & 2 \end{bmatrix} \quad (8.666)$$

Ortogonalnost:

$$\underline{\underline{A}} \cdot \underline{y}^* = \underline{0} \quad (8.667)$$

Normiranost:

$$\sum_{j=1}^{N=5} y_j^* = 1 \quad (8.668)$$

Ortogonalnost + normiranost nam da:

$$\begin{bmatrix} 2 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (8.669)$$

Sledi:

$$\begin{aligned} E_1: 2y_1^* - y_3^* + y_4^* &= 0 & \Rightarrow y_1^* &= \frac{y_3^* - y_4^*}{2} \\ E_2: y_2^* - y_3^* + y_4^* + 2y_5^* &= 0 & \Rightarrow y_2^* &= y_3^* - y_4^* - 2y_5^* \\ E_3: y_1^* + y_2^* + y_3^* + y_4^* + y_5^* &= 1 \end{aligned} \quad (8.670)$$

Izraze  $y_1^*$  in  $y_2^*$  izpeljane v obrazcu (8.670) vstavimo v enačbo  $E_3$  in dobimo:

$$\begin{aligned} \left( \frac{y_3^* - y_4^*}{2} \right) + (y_3^* - y_4^* - 2y_5^*) + y_3^* + y_4^* + y_5^* &= 1 \\ \frac{y_3^* - y_4^*}{2} + 2y_3^* - y_5^* &= 1 \quad \Rightarrow y_5^* = \frac{y_3^* - y_4^*}{2} + 2y_3^* - 1 \end{aligned} \quad (8.671)$$

Izraz  $y_5^*$  izpeljan v obrazcu (8.671) pa vstavimo v enačbo  $E_2$  in dobimo:

$$\begin{aligned} y_2^* &= y_3^* - y_4^* - 2 \cdot \left[ \frac{y_3^* - y_4^*}{2} + 2y_3^* - 1 \right] = y_3^* - y_4^* - y_3^* + y_4^* - 4y_3^* + 2 = \\ &= 2 - 4y_3^* = y_2^* \end{aligned} \quad (8.672)$$

Preizkus (Damo izpeljane izraze (8.670), (8.671) in (8.672) v enačbo  $E_3$  in dobimo):

$$\frac{y_3^* - y_4^*}{2} + 2 - 4y_3^* + y_3^* + y_4^* + \frac{y_3^* - y_4^*}{2} + 2y_3^* - 1 = 2 - 1 = 1 \quad \text{OK} \quad (8.673)$$

Torej sta  $y_3^*$ ,  $y_4^*$  prosti,  $y_1^*$ ,  $y_2^*$  in  $y_5^*$  pa odvisne od njiju:

Parcialni odvodi so:

$$\begin{aligned} y_1^* &= \frac{y_3^* - y_4^*}{2} & P_1 &= \frac{\partial y_1^*}{\partial y_3^*} = \frac{1}{2}; & \frac{\partial y_1^*}{\partial y_4^*} &= -\frac{1}{2} = q_1 \\ y_2^* &= 2 - 4y_3^* & \Rightarrow P_2 &= \frac{\partial y_2^*}{\partial y_3^*} = -4; & \frac{\partial y_2^*}{\partial y_4^*} &= 0 = q_2 \\ y_5^* &= \frac{y_3^* - y_4^*}{2} + 2y_3^* - 1 & P_5 &= \frac{\partial y_5^*}{\partial y_3^*} = \frac{5}{2}; & \frac{\partial y_5^*}{\partial y_4^*} &= -\frac{1}{2} = q_5 \\ & & P_3 &= \frac{\partial y_3^*}{\partial y_3^*} = 1; & \frac{\partial y_3^*}{\partial y_4^*} &= 0 = q_3 \\ & & P_4 &= \frac{\partial y_4^*}{\partial y_3^*} = 0; & \frac{\partial y_4^*}{\partial y_4^*} &= 1 = q_4 \end{aligned} \quad (8.674)$$

Veljata naslednji relaciji:

$$\begin{aligned} (y_1^*)^{P_1} \cdot (y_2^*)^{P_2} \cdot (y_3^*)^{P_3} \cdot (y_4^*)^{P_4} \cdot (y_5^*)^{P_5} &= C_1^{P_1} \cdot C_2^{P_2} \cdot C_3^{P_3} \cdot C_4^{P_4} \cdot C_5^{P_5} \\ (y_1^*)^{q_1} \cdot (y_2^*)^{q_2} \cdot (y_3^*)^{q_3} \cdot (y_4^*)^{q_4} \cdot (y_5^*)^{q_5} &= C_1^{q_1} \cdot C_2^{q_2} \cdot C_3^{q_3} \cdot C_4^{q_4} \cdot C_5^{q_5} \end{aligned} \quad (8.675)$$

Sledi:

$$\begin{aligned} \left(\frac{y_3^* - y_4^*}{2}\right)^{\frac{1}{2}} \cdot (2 - 4y_3^*)^{-4} \cdot y_3^{*1} \cdot y_4^{*0} \cdot \left(\frac{y_3^* - y_4^*}{2} + 2y_3^* - 1\right)^{\frac{5}{2}} &= \\ = 1^{\frac{1}{2}} \cdot 2^{-4} \cdot 3^1 \cdot 2^0 \cdot 1^{\frac{5}{2}} & \\ \left(\frac{y_3^* - y_4^*}{2}\right)^{-\frac{1}{2}} \cdot (2 - 4y_3^*)^0 \cdot y_3^{*0} \cdot y_4^{*1} \cdot \left(\frac{y_3^* - y_4^*}{2} + 2y_3^* - 1\right)^{-\frac{1}{2}} &= \\ = 1^{-\frac{1}{2}} \cdot 2^0 \cdot 3^0 \cdot 2^1 \cdot 1^{-\frac{1}{2}} & \end{aligned} \quad (8.676)$$

Sledi:

$$\begin{aligned}
& \left( \frac{y_3^* - y_4^*}{2} \right)^{\frac{1}{2}} \cdot (2 - 4y_3^*)^{-4} \cdot y_3^* \cdot \left( \frac{5y_3^*}{2} - \frac{y_4^*}{2} - 1 \right)^{\frac{5}{2}} = \frac{3}{2^4} \quad \Bigg| \cdot 2 \\
& \left( \frac{y_3^* - y_4^*}{2} \right)^{\frac{1}{2}} \cdot y_4^* \cdot \left( \frac{5y_3^*}{2} - \frac{y_4^*}{2} - 1 \right)^{\frac{1}{2}} = 2 \quad \Bigg| \cdot 2 \\
& \frac{y_3^* - y_4^*}{2} \cdot \frac{1}{(2 - 4y_3^*)^8} \cdot y_3^{*2} \cdot \left( \frac{5y_3^*}{2} - \frac{y_4^*}{2} - 1 \right)^5 = \frac{9}{2^8} \\
& \frac{2}{y_3^* - y_4^*} \cdot y_4^{*2} \cdot \left( \frac{5y_3^*}{2} - \frac{y_4^*}{2} - 1 \right)^{-1} = 4
\end{aligned} \tag{8.677}$$

Sledi:

$$\begin{aligned}
& \frac{y_3^* - y_4^*}{2} \cdot \frac{1}{2^8 \cdot (1 - 2y_3^*)^8} \cdot y_3^{*2} \cdot \frac{1}{2^5} \cdot (5y_3^* - y_4^* - 1)^5 = \frac{9}{2^8} \\
& \frac{2}{y_3^* - y_4^*} \cdot y_4^{*2} \cdot (5y_3^* - y_4^* - 2)^{-1} \cdot 2 = 4
\end{aligned} \tag{8.678}$$

Sledi:

$$\begin{aligned}
& \frac{(y_3^* - y_4^*) \cdot y_3^{*2} \cdot (5y_3^* - y_4^* - 2)^5}{2^{14} \cdot (1 - 2y_3^*)^8} = \frac{9}{\underbrace{2^8}_{\approx 0.034}} \\
& \frac{2^2 \cdot y_4^{*2}}{(y_3^* - y_4^*) \cdot (5y_3^* - y_4^* - 2)} = 2^2
\end{aligned} \tag{8.679}$$

Dobili smo torej sistem 2 enačb z 2 neznankama  $y_3^*$  in  $y_4^*$ .

Če je  $y_3^* = 0.4589$  in  $y_4^* = 0.1794$ , potem sledi (preverimo):

$$\begin{aligned}
& \frac{(0.4589 - 0.1794) \cdot 0.4589^2 \cdot (5 \cdot 0.4589 - 0.1794 - 2)^5}{2^{14} \cdot (1 - 2 \cdot 0.4589)^8} = \frac{1.18903 \cdot 10^{-6}}{3.4150 \cdot 10^{-5}} = \\
& \qquad \qquad \qquad = 0.0348 \quad \text{OK} \tag{8.680} \\
& \frac{0.1794^2}{(0.4589 - 0.1794) \cdot (5 \cdot 0.4589 - 0.1794 - 2)} = 1.0004 \quad \text{OK}
\end{aligned}$$

Torej sta to res pravi rešitvi. Sledi (glej (8.670), (8.671) in (8.672)):

$$\begin{aligned} y_1^* &= \frac{0.4589 - 0.1794}{2} = 0.1397 \\ y_2^* &= 2 - 4 \cdot 0.4589 = 0.1644 \\ y_5^* &= \frac{0.4589 - 0.1794}{2} + 2 \cdot 0.4589 - 1 = 0.0575 \end{aligned} \quad (8.681)$$

Dualne rešitve torej so:

$$\begin{aligned} y_1^* &= 0.1397 \\ y_2^* &= 0.1644 \\ y_3^* &= 0.4589 \\ y_4^* &= 0.1794 \\ y_5^* &= 0.0575 \end{aligned} \quad (8.682)$$

Optimalna dualna funkcija je:

$$\begin{aligned} f^*(\underline{y}) &= \left(\frac{1}{0.1397}\right)^{0.1397} \cdot \left(\frac{2}{0.1644}\right)^{0.1644} \cdot \left(\frac{3}{0.4589}\right)^{0.4589} \cdot \left(\frac{2}{0.1794}\right)^{0.1794} \cdot \\ &\quad \cdot \left(\frac{1}{0.0575}\right)^{0.0575} = \\ &= 1.3164 \cdot 1.5079 \cdot 2.3669 \cdot 1.5412 \cdot 1.1784 = 8.5333 \end{aligned} \quad (8.683)$$

Optimalni odločitveni spremenljivki pa sta:

$$\begin{aligned} U_1^* = x_1^{*2} = f^* \cdot y_1^* &= 8.5333 \cdot 0.1397 \Rightarrow x_1^* = 1.0918 \\ U_5^* = x_2^{*2} = f^* \cdot y_5^* &= 8.5333 \cdot 0.0575 \Rightarrow x_2^* = 0.7004 \end{aligned} \quad (8.684)$$

### 8.5.2 Geometrijsko programiranje z omejitvami

Opomba: Za razumevanje teoretičnega ozadja glej poglavje 7.2, ter izraze (7.162) do (7.245).

#### Primer 7:

Minimizirajte naslednjo funkcijo:

$$f(\underline{x}) = 20x_1x_3 + 40x_2x_3 + 80x_1x_2 \quad (8.685)$$

pri omejitvi:

$$g_1 : \frac{80}{x_1x_2x_3} \leq 10 \Rightarrow \frac{8}{x_1x_2x_3} \leq 1. \quad (8.686)$$

Funkcijo, dano v izrazu (8.685), prevedimo v obliko primerno za GP:

$$f(\underline{x}) = \underbrace{\overbrace{20}^{c_1} x_1^1 x_2^0 x_3^1}_{u_1} + \underbrace{\overbrace{40}^{c_2} x_1^0 x_2^1 x_3^1}_{u_2} + \underbrace{\overbrace{80}^{c_3} x_1^1 x_2^1 x_3^0}_{u_3} \quad \left. \vphantom{f(\underline{x})} \right\} N = 3 \text{ (3 členi)} \quad (8.687)$$

Omejitev, primerna za reševanje problema GP, pa se glasi:

$$g_1(\underline{x}) = \underbrace{\overbrace{8}^{c_4} x_1^{-1} x_2^{-1} x_3^{-1}}_{u_4} \leq 1 \quad \left\{ \begin{array}{l} P = 1 \text{ (1 omejitev)} \\ K_1 = 1 \text{ (1 člen)} \end{array} \right. \quad (8.688)$$

Zapišimo matriko  $\underline{\underline{A}}$ :

$$\underline{\underline{A}} = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 0 & -1 \end{bmatrix} \quad (8.689)$$

Ortogonalnost:

$$\underline{\underline{A}} \cdot \underline{y}^* = \underline{0} \quad (8.690)$$

Normiranost:

$$\sum_{j=1}^{N=3} y_j^* = 1 \Rightarrow \underbrace{y_1^* + y_2^* + y_3^*}_{E_4} = 1 \quad (8.691)$$

Če združimo ortogonalnost in normiranost, dobimo:

$$\begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (8.692)$$

Sledi:

$$\begin{aligned} E_1: y_1^* + y_3^* - y_4^* &= 0 \Rightarrow y_4^* = y_1^* + y_3^* \\ E_2: y_2^* + y_3^* - y_4^* &= 0 \Rightarrow y_4^* = y_2^* + y_3^* \\ E_3: y_1^* + y_2^* - y_4^* &= 0 \Rightarrow y_4^* = y_1^* + y_2^* \end{aligned} \quad (8.693)$$

Izenačimo  $E_1$  in  $E_2$  in dobimo:

$$y_1^* + y_3^* = y_2^* + y_3^* \Rightarrow y_1^* = y_2^* \quad (8.694)$$

Izenačimo  $E_1$  in  $E_3$ , sledi:

$$y_1^* + y_3^* = y_1^* + y_2^* \Rightarrow y_2^* = y_3^* \quad (8.695)$$

Upoštevamo (8.694) in (8.695) v enačbi  $E_4$ , sledi:

$$y_1^* + y_1^* + y_1^* = 1 \Rightarrow y_1^* = \frac{1}{3} = y_2^* = y_3^* \quad (8.696)$$

Na osnovi  $E_1$  in (8.696) dobimo:

$$y_1^* + y_1^* + y_1^* = 1 \Rightarrow y_1^* = \frac{1}{3} = y_2^* = y_3^* \quad (8.697)$$

Sledi:

$$y_4^* = y_1^* + y_3^* = \frac{1}{3} + \frac{1}{3} = \frac{2}{3} \quad (8.698)$$

Dobimo torej:

$$(y_1^*, y_2^*, y_3^*, y_4^*) = \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right) \quad (8.699)$$

Dualna funkcija se glasi:



$$\begin{aligned}
 f^*(\underline{y}^*) &= \prod_{j=1}^{N+K_1=3+1} \left( \frac{C_j}{y_j^*} \right)^{y_j^*} \cdot \prod_{K=1}^{P=1} (\lambda_K^*)^{\lambda_K^*} = \\
 &= \left( \frac{C_1}{y_1^*} \right)^{y_1^*} \cdot \left( \frac{C_2}{y_2^*} \right)^{y_2^*} \cdot \left( \frac{C_3}{y_3^*} \right)^{y_3^*} \cdot \left( \frac{C_4}{y_4^*} \right)^{y_4^*} \cdot (\lambda_1^*)^{\lambda_1^*}
 \end{aligned}
 \tag{8.700}$$

kjer je  $\lambda_1^* = y_{\frac{N+K_1}{3} \frac{1}{1}}^* = y_4^*$ .

Torej:

$$\begin{aligned}
 f^* &= \left( \frac{20}{\frac{1}{3}} \right)^{\frac{1}{3}} \cdot \left( \frac{40}{\frac{1}{3}} \right)^{\frac{1}{3}} \cdot \left( \frac{80}{\frac{1}{3}} \right)^{\frac{1}{3}} \cdot \left( \frac{8}{\frac{2}{3}} \right)^{\frac{2}{3}} \cdot \left( \frac{2}{3} \right)^{\frac{2}{3}} = 60^{\frac{1}{3}} \cdot 120^{\frac{1}{3}} \cdot 240^{\frac{1}{3}} \cdot 12^{\frac{2}{3}} \cdot \left( \frac{2}{3} \right)^{\frac{2}{3}} = \\
 &= (60 \cdot 120 \cdot 240)^{\frac{1}{3}} \cdot \left( \frac{12 \cdot 2}{3} \right)^{\frac{2}{3}} = (1728000)^{\frac{1}{3}} \cdot 8^{\frac{2}{3}} = 120 \cdot 4 = 480
 \end{aligned}
 \tag{8.701}$$

Določitev optimalnih odločitvenih spremenljivk:

$$\begin{aligned}
 U_1^* = f^* \cdot y_1^* &\Rightarrow 20x_1^* x_3^* = 480 \cdot \frac{1}{3} = 160 \\
 U_2^* = f^* \cdot y_2^* &\Rightarrow 40x_2^* x_3^* = 480 \cdot \frac{1}{3} = 160 \\
 U_3^* = f^* \cdot y_3^* &\Rightarrow 80x_1^* x_2^* = 480 \cdot \frac{1}{3} = 160
 \end{aligned}
 \tag{8.702}$$

Sledi:

$$\begin{aligned}
 x_1^* x_3^* &= 8 \quad (M_1) \\
 x_2^* x_3^* &= 4 \quad (M_2) \\
 x_1^* x_2^* &= 2 \quad (M_3)
 \end{aligned}
 \tag{8.703}$$

Za omejitev pa velja:

$$U_{\underbrace{N+K_1}_4}^* = \frac{y_{N+K_1}^*}{\lambda_1^*} \Rightarrow U_4^* = \frac{y_4^*}{\lambda_1^*} = \frac{y_4^*}{y_4^*} = 1
 \tag{8.704}$$

Sledi:

$$\frac{8}{x_1^* x_2^* x_3^*} = 1 \quad (M_4) \quad (8.705)$$

Iz  $M_1$  in  $M_4$  sledi:

$$\frac{8}{8x_2^*} = 1 \quad \Rightarrow x_2^* = 1 \quad (8.706)$$

Iz  $M_3$  zato sledi:

$$x_2^* \cdot 1 = 2 \quad \Rightarrow x_1^* = 2 \quad (8.707)$$

Iz  $M_2$  pa sledi:

$$1 \cdot x_3^* = 4 \quad \Rightarrow x_3^* = 4 \quad (8.708)$$

Optimalni vektor odločitvenih spremenljivk torej je:

$$\underline{x}^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} \quad (8.709)$$

## 9 PRILOGE

V tem poglavju so predstavljeni primeri izpisov nekaterih računalniških programov za numerično (nelinearno) optimizacijo. Programi so bili večinoma napisani v programskem paketu Matlab, nekateri pa tudi v programskem paketu Scilab. Za slednjega velja, da ima zelo podobno strukturo programskih vrstic kot Matlab, saj je le njegov odprtokodni »klon«.

V programih so razvidni primeri različnih nelinearnih ciljnih funkcij, ki so bile optimirane, prikazi dinamike konvergence tekom postopka njihove optimizacije pa ilustrirani v okviru slik prejšnjih poglavij (glej predvsem poglavje 6). Tudi teoretične osnove, ki stojijo v ozadju mehanizmov delovanja posameznih programov, si bralec lahko pogleda v poglavju 6.

Programi so tudi opremljeni z ustreznimi komentarji, kjer je to potrebno. Kjer so v programih vgrajeni tudi podprogrami, je to jasno razvidno. To pomeni, najprej je podan izpis osnovnega programa, nato pa še vseh podprogramov.

### 9.1 Programi v Matlabu ali Scilabu

V nadaljevanju je predstavljen izpis nekaterih programov za:

- Enodimenzionalno numerično nelinearno optimizacijo, ter
- Multidimenzionalno numerično nelinearno optimizacijo.

Pri enodimenzionalni numerični optimizaciji smo podali primere izpisov programov za naslednje metode:

- Metodo zlatega reza,
- Newtonovo metodo za eno spremenljivko, ter
- Sekantno metodo.

Pri multidimenzionalni numerični optimizaciji pa smo podali primere izpisov programov za naslednje osnovne metode:

- Hooke-Jeevesovo metodo vzorčnih premikov,
- Nelder-Meadovo simpleksno metodo,
- Metodo najstrmejšega spusta, ter
- Newtonovo multidimenzionalno metodo.

Prav tako smo pri multidimenzionalni numerični optimizaciji podali tudi primer izpisa programa za naslednjo sofisticirano metodo:

- DFP QuasiNewtonovo metodo.

Poudarimo še, da smo pri metodi najstrmejšega spusta podali tako primer izpisa programa, če si ne pomagamo s Symbolic toolboxom, kot tudi primer izpisa programa, kjer si pomagamo s Symbolic toolboxom. Gre za poseben pripomoček v sklopu Matlaba, ki nam je omogočil na določenih točkah v programu tudi simbolično (analitično) računanje, kar je olajšalo oz. poenostavilo določene izračune.

Omenimo še, da smo tovrsten pripomoček (Symbolic toolbox) uporabili tudi pri DFP QuasiNewtonovi metodi. Pri tem optimalnega parametra  $\alpha_k$  v vsaki iteraciji ne iščemo s katero izmed numeričnih enodimenzionalnih metod optimizacije, pač pa ga iščemo na

analitičen način na osnovi izraza  $\frac{df(\alpha_k)}{d\alpha_k} = 0$ . Pri tem je  $f$  multidimenzionalna ciljna

funkcija, ki jo zreduciramo na enodimenzionalno odvisnost od  $\alpha_k$  ( $k$  je indeks iteracije).

Seveda ta pristop zaradi kompleksnosti ni več možen pri bolj kompliciranih ciljnih funkcijah, kjer moramo namesto tega uporabiti v vsaki iteraciji katero izmed numeričnih enodimenzionalnih metod optimizacije.

Poglejmo si torej izpise programov za navedene metode.

### 9.1.1 Algoritem zlatega reza (Scilab)

```
// golden_outfun

function REZ = golden_outfun()

clc
clear

f = input('vnesi funkcijo v pravilni sintaksi (npr. x^2+2*x):','string');

a0 = input('a0= (npr.-3)');
b0 = input('b0= (npr. 5)');
eps = 1e-4;

n = 100;

cZR = (3 - sqrt(5))/2; // drugace kot fibonacci

y1 = a0 + cZR*(b0-a0);
z1 = b0 - cZR*(b0-a0);

x=y1;fy1 = evstr(f);
x=z1;fz1 = evstr(f);

ak(1) = a0;
bk(1) = b0;
yk(1) = y1;
zk(1) = z1;
fyk(1) = fy1;
fzk(1) = fz1;

for k = 2:n
    if fyk(k-1) <= fzk(k-1)
        ak(k) = ak(k-1);
        bk(k) = zk(k-1);
        yk(k) = ak(k) + cZR*(bk(k) - ak(k)); // drugace kot fibonacci
        x = yk(k);
        fyk(k) = evstr(f);
        zk(k) = yk(k-1);
        x = zk(k);
        fzk(k) = evstr(f);
    else
        ak(k) = yk(k-1);
        bk(k) = bk(k-1);
        zk(k) = bk(k) - cZR*(bk(k) - ak(k)); // drugace kot fibonacci
        x = zk(k);
        fzk(k) = evstr(f);
        yk(k) = zk(k-1);
        x = yk(k);
        fyk(k) = evstr(f);
    end
    if abs(bk(k)-ak(k))<eps
        break
    end
end

REZ = [[0:1:k-1]' ak yk zk bk fyk fzk];

endfunction
```

### 9.1.2 Algoritem Newtonove metode za eno spremenljivko (Scilab)

// newton\_outfun.sce za funkcijo: npr.  $f(x) = 2*x^4 - 3*x$

```
function X = newton_outfun()

clc
clear

f = '2*x0^4 - 3*x0'; // funkcija
df = '8*x0^3-3'; // sem vnesemo 1. odvod
d2f = '24*x0^2' // sem vnesemo 2. odvod

x0 = input('x0='); // zacetni pogoj

N = 100;
eps = 1e-3;

X = x0;

for k = 1:N
    fx = evstr(f);
    dfx = evstr(df);
    d2fx = evstr(d2f);
    x1 = x0 - dfx/d2fx;
    e = abs((x1-x0)/x1);
    if e < eps
        break
    end
    x0 = x1;
    X = [X;x0];
end

disp('optimum je:')
x1

endfunction

// klic funkcije je: X = newton_outfun()
```

### 9.1.3 Algoritem sekantne metode (Scilab)

// sekant\_outfun.sce za funkcijo: npr.  $f(x) = 2*x^4 - 3*x$

```
function X = sekant_outfun()

clc
clear

f0 = '2*x0^4 - 3*x0'; // funkcija v koraku k-1
df0 = '8*x0^3-3'; // odvod funkcije v koraku k-1
df1 = '8*x1^3-3'; // odvod funkcije v koraku k

x0 = input('x0='); // zacetni pogoj 1
x1 = input('x1='); // zacetni pogoj 2

N = 100;
eps = 1e-3;

X = [x0;x1];

for k = 1:N
    dfx0 = evstr(df0);
    dfx1 = evstr(df1);
    x2 = x1 - dfx1*(x1-x0)/(dfx1-dfx0); // izracun za k+1
```

```

e = abs((x2-x1)/x2);
if e < eps
    break
end
x0 = x1;
x1 = x2;
X = [X;x2];
end

disp('optimum je:')
x2

endfunction

// klic funkcije je: X = sekant_outfun()

```

### 9.1.4 Algoritem Hooke-Jeevesove metode vzorčnih premikov (Matlab)

```

%-----
% ols_hook
%
% klic je npr. TOCKE=ols_hook(2,1,[7 4]); ali TOCKE=ols_hook(2,1,[0.1 0.1]); ali TOCKE=ols_hook(3,2,[0 0 0]);
%-----
function TOCKE = ols_hook(dim,sf,x0)

% dim = dimenzija problema, sf je številka testne funkcije

close all
clc
disp('Hooke')

Dxmejni = 1e-3;      % toleranca za normo od Dx
itmejni = 100;      % max. stevilo iteracij
delta = 0.1;      % vrednost premika po dimenziji
%x0 = zeros(dim,1); % začetni vektor neodv. sprem.
xp = x0;          % začetni vektor novega premika
xb = x0;          % začetni bazni vektor
Dx = delta*eye(dim); % matrika vektorjev premikov po dimenzijah

xn = -100*xb;      % začetni vektor uspešnega premika

TOCKE = [xb F(xb,sf) norm(Dx)]; % sem shranjujemo kronologijo premikov

F(xn,sf)
F(xb,sf)

it = 0;
while (it < itmejni)
    [xn,uspesno] = ols_hook_poisc(xp,Dx,0,sf) % nov premik, gledamo vse smeri
    F(xn,sf)
    F(xb,sf)
    if (F(xn,sf) < F(xb,sf)) & (uspesno > 0) % boljši od stare baze in nov premik uspešen (vsaj v eno smer)- xn boljši od xp
        xp = xn + (xn - xb) % nov premik je nov skok
        xb = xn % bazo postavimo na premik pred skokom
        it = it + 1, disp('*****'), TOCKE = [TOCKE; [xb F(xb,sf) norm(Dx)]]; % kronol.le pri uspešni spr.baze
        [xn,uspesno] = ols_hook_poisc(xp,Dx,1,sf) % nov premik, gledamo le prvo uspešno smer
        if (uspesno == 1) & (norm(xb-xn) > 1e-3) % če ena od smeri uspešna in nov premik drugačen od stare baze
            xb = xn % potem je skok uspešen in bazo postavimo na nov skok+prvo uspešnih smeri
            it = it + 1, disp('*****'), TOCKE = [TOCKE; [xb F(xb,sf) norm(Dx)]]; % kronol.le pri uspešni spr.baze
            xp = xb % nov premik v naslednji iteraciji enak spremenjeni bazi
        else % nobena od smeri uspešna ali nov premik enak stari bazi
            xp = xb % nov premik v naslednji iteraciji enak nespremenjeni bazi
        end
    else % ni boljši od stare baze ali nov premik neuspešen (v nobeno smer)
        Dx = Dx/2 % zato prepolovimo vrednost premika po vseh dimenzijah
        xp = xb % nov premik enak kot prej (enak stari bazi), le da Dx prepolovljen
    end
end

```

```

    if norm(Dx) < Dxmejni % ce Dx zmanjsamo do vrednosti tolerance, koncamo (ali pa, ce itmejni dosežemo)
        break
    end
end
end
end

```

```

disp('koncni rezultat:')
x = xb

```

```

if dim == 2 % pri 2D problemu lahko se sliko za konvergiranje k resitvi narisemo
    ols_spaceplotter(TOCKE(:,1:end-1'))
end

```

```

return

```

```

%-----

```

```

% ols_hook_poisci

```

```

%-----
function [x,uspesno] = ols_hook_poisci(xp,Dx,flag,sf)

```

```

x = xp
uspesno = 0;

```

```

for i = 1:length(x) % gremo preko vseh dimenzij
    P = F(x,sf)
    x = x + Dx(i,:); % prvo pozitiven premik
    N = F(x,sf)
    if N > P % pozitiven premik ni uspesen
        x = x - 2*Dx(i,:); % zato nazaj in negativen premik
        N = F(x,sf)
        if N > P % negativen premik tudi ni uspesen
            x = x + Dx(i,:); % zato nazaj v prvotno stanje in na novo dimenzijo
        else % negativen premik je uspesen, gremo na novo dim.
            uspesno = uspesno + 1;
            if flag == 1;return;end % ce gledamo le prvo uspesno smer, vrnitev nazaj
        end
    else % pozitiven premik je uspesen, gremo na novo dim.
        uspesno = uspesno + 1;
        if flag == 1;return;end % ce gledamo le prvo uspesno smer, vrnitev nazaj
    end
end
end
return

```

```

%-----

```

```

% F - nabor nasih testnih funkcij

```

```

%-----
function f = F(x,sf)

```

```

if sf == 1
    x1 = x(1);
    x2 = x(2);
    f = (12 + x1^2 + (1 + x2^2)/x1^2 + (x1^2*x2^2 + 100)/(x1*x2)^4)*(1/10);
else
    x1 = x(1);
    x2 = x(2);
    x3 = x(3);
    f = x1^4+x1^3-x1+x2^4-x2^2+x2+x3^2-x3+x1*x2*x3;
end

```

### 9.1.5 Algoritem Nelder-Meadove simpleksne metode (Matlab)

```

%-----

```

```

% ols_nelder

```



```

%-----
function [X,Y] = ols_nelder(dim,sf,x0)

% dim = dimenzija problema, sf je številka testne funkcije, x0 je začetni vektor
% klic je npr.: [X,Y]=ols_nelder(2,1,[7;4]); ali [X,Y]=ols_nelder(2,1,[0.1;0.1]); ali [X,Y]=ols_nelder(3,2,[0.1;0.1;0.1]);

close all
clc

if sf == 1
    diary ols_nelder_rez1.m
else
    diary ols_nelder_rez2.m
end

if sf == 1 % Pri Fentonovi funkciji ne smemo iti prevec v singularnost.
    if (x0(1)<0.1)
        x0(1) = 0.1
        disp('Pozor: Pri Fentonovi funkciji si sel prevec v singularnost')
    end
    if (x0(2)<0.1)
        x0(2) = 0.1
        disp('Pozor: Pri Fentonovi funkciji si sel prevec v singularnost')
    end
end

disp('Nelder')

itmejni = 300; % max. stevilo iteracij

eps = 0.1; % faktor razseznosti simplexa
Rho = 1; % faktor refleksijske tocke
Chi = 2; % faktor ekspanzijske tocke
Gamma = 0.5; % faktor skrcitvene tocke (tip1 oziroma tip2)
Sigma = 0.5; % faktor pomanjsanja simplexa
TolX = 1e-6; % toleranca konvergence za vektor x
TolFun = 1e-6; % toleranca konvergence za funkcijo

n = length(x0); % dimenzija prostora
n1 = n+1; % stevilo ogljisc simplexa

x = [x0 x0(:,ones(n,1))+diag(0.05*x0+0.00025*(abs(x0)<eps))]' % generacija tock zacetnega simplexa

y = [];
for i = 1:n1
    y = [y;F(x(i,:),sf)]; % generacija vrednosti funkcije v zacetnem simplexu
end
y

X = x % matrika zgodovine vektorja x (ki je trenutna domnevno izboljsana tocka simpleksa)
Y = y % vektor zgodovine funkcije v vektorju x
st = n1; % stevec sprememb vektorja x

for i = 2:itmejni
    [y,ind] = sort(y) % sortira funkcijo od najmanjse do največje vrednosti
    x = x(ind,:); % sortira tocke simplexa od najboljše do najslabše
    centroid = mean(x(1:end-1,:)) % težišče točk simplexa razen največje (najslabše)
    x_r = centroid + Rho*(centroid - x(end,:)) % največja točka se prezrcali v refleksijsko točko
    y_r = F(x_r,sf) % vrednost funkcije v njej
    st = st + 1;
    X = [X ; x_r]; % zapomnimo si domnevno izboljsano točko simplexa
    Y = [Y ; y_r]; % in funkcijo v njej
    if y_r >= y(1) & y_r < y(end-1) % če refleksijska točka med najmanjšo in drugo največjo točko
        x(end,:) = x_r % največjo točko nadomestimo z refleksijsko (dejansko izboljsamo)
        y(end) = y_r % nadomestimo tudi funkcijo v tej točki
    elseif y_r < y(1) % če refleksijska točka manjša od najmanjše, potem ima najnižjo vred. od vseh simplex. točk.
        x_e = centroid + Chi*(x_r - centroid) % potem ugodna smer in refleksij.točko pomaknemo v ekspanzijsko
        y_e = F(x_e,sf) % vrednost funkcije v njej
        st = st + 1
        X = [X ; x_e]; % zapomnimo si domnevno izboljsano točko simplexa
    end
end

```

```

Y = [Y ; y_e]; % in funkcijo v njej
if y_e < y_r % ce ekspanzijska točka manjša (boljša) od refleksijske
    x(end,:) = x_e % največjo točko nadomestimo z ekspanzijsko (dejansko izboljšamo - kar 2x)
    y(end) = y_e % nadomestimo tudi funkcijo v tej točki
else
    x(end,:) = x_r % sicer pa največjo točko nadomestimo z refleksijsko (dejansko izboljšamo, a le 1x)
    y(end) = y_r % nadomestimo tudi funkcijo v tej točki
end
else % potem je refleksijska točka večja od druge največje točke in moramo skrciti
    shrink = 0 % zastavica za pomanjšanje simplexa je na začetku ne
    if y(end-1) <= y_r & y_r < y(end) % ce je refleksijska točka med največjo in drugo največjo
        x_c = centroid + Gamma*(x_r - centroid) % skrcitvena točka med centroidom in refleksijsko (tip1)
        y_c = F(x_c,sf)
        st = st + 1
        X = [X ; x_c]; % zapomnimo si domnevno izboljšano točko simplexa
        Y = [Y ; y_c]; % in funkcijo v njej
        if y_c <= y_r % ce je skrcitvena točka tipa1 manjša (boljša) od refleksijske
            x(end,:) = x_c % največjo točko nadomestimo s skrcitveno tipa1 (dejansko izboljšamo)
            y(end) = y_c % nadomestimo tudi funkcijo v tej točki
        else % ni koristi od zrcaljenja, simplex moramo zmanjšati
            shrink = 1
        end
    else % refleksijska točka večja od največje
        x_c = centroid + Gamma*(centroid - x(end,:)) % skrcitvena točka med največjo (najslabšo) in centroidom (tip2)
        y_c = F(x_c,sf)
        st = st + 1
        X = [X ; x_c]; % zapomnimo si domnevno izboljšano točko simplexa
        Y = [Y ; y_c]; % in funkcijo v njej
        if y_c <= y(end) % ce je skrcitvena točka tipa2 boljša od največje (najslabše)
            x(end,:) = x_c % največjo točko nadomestimo s skrcitveno tipa2 (dejansko izboljšamo)
            y(end) = y_c % nadomestimo tudi funkcijo v tej točki
        else % ni koristi od zrcaljenja, simplex moramo zmanjšati
            shrink = 1
        end
    end
end
if shrink == 1 % ce ni nobene koristi od zrcaljenja, moramo simplex zmanjšati
    for j=2:n+1
        x(j,:) = x(1,:) + Sigma*(x(j,:) - x(1,:)) % vsako oglišče simplexa zmanjšamo
        y(j) = F(x(j,:),sf) % in vrednost funkcije v vsakem oglišču
        st = st + 1
        X = [X ; x(j,:)]; % zapomnimo si zmanjšana oglišča simplexa
        Y = [Y ; y(j)]; % in vrednost funkcije v vsakem oglišču
    end
end
end
if max(abs(min(x) - max(x))) < TolX
    disp('optimizacija končana uspešno (TolX kriterij)'); % ce izpolnjen kriterij konvergence za x, stop
    break;
end
if abs(max(y) - min(y))/max(abs(y)) < TolFun
    disp('optimizacija končana uspešno (TolFun kriterij)'); % ce izpolnjen kriterij konvergence za funkcijo, stop
    break
end
end
end

if i == itmejni % postopek ni konvergiraj
    disp('Pozor : maksimalno stevilo iteracij preseženo');
end

diary off

if dim == 2 % pri 2D problemu lahko se sliko za konvergiranje k rešitvi narisemo
    ols_spaceplotter_nelder([X Y])
    figure(1)
    title('Konvergenca k rešitvi pri Nelder metodi')
    figure(2)
    title('Konvergenca k rešitvi pri Nelder metodi')
end

return

```

```

%-----
% F - nabor nasih testnih funkcij
%-----
function f = F(x,sf)

if sf == 1
    x1 = x(1);
    x2 = x(2);
    f = (12 + x1^2 + (1 + x2^2)/x1^2 + (x1^2*x2^2 + 100)/(x1*x2)^4)*(1/10);
else
    x1 = x(1);
    x2 = x(2);
    x3 = x(3);
    f = x1^4+x1^3-x1+x2^4-x2^2+x2+x3^2-x3+x1*x2*x3;
end

```

### 9.1.6 Algoritem metode najstrmejšega spusta brez uporabe Symbolic toolboxa (Matlab)

```

%-----
% ols_steepest
%
% dim = dimenzija problema, sf je številka testne funkcije
%
% Pri močno zgresenih začetnih (ogromnih) vrednostih metoda s parabolo divergira, s sekanto pa ne.
% Sicer pa metoda s parabolo nekoliko hitreje konvergira od metode s sekanto.
%-----
function [X,Y,ALFA,GRAD] = ols_steepest(dim,sf,x0,tip_met,tip_kor)

close all
clc
disp('Steepest descent')

if tip_met == 1
    alfa0 = 10
else
    alfa0 = 1e-5;
end

if tip_kor == 2 % fiksni korak
    alfa_fix = 2/13; % zberemo neko, naceleoma manjša mora biti od 2/max_eig(Hesse)=2/10 - glej Gershgorin.kroge
    alfa0 = alfa_fix;
end

kmax = 100
MaxIter = 200
TolX = 1e-4
TolFun = 1e-4

x = x0
fx0 = F(x0,sf,1)
fx = fx0
alfa = alfa0

X = x; Y = fx; ALFA = alfa; GRAD = [];

warning = 0

for k = 1:MaxIter
    g = F(x,sf,2);
    if tip_kor == 1
        g = g/norm(g)
        alfa = alfa*2;
        if tip_met == 1

```

```

        [alfa,k1] = ols_steep_parab(x,sf,alfa,kmax)
    else
        [alfa,k1] = ols_steep_sekant(x,sf,alfa,kmax)
        % Tu bi lahko klicali se druge metode (armijo, backtracking, itn...)
    end
else
    alfa = alfa_fix;
    k1 = 0;
end
GRAD = [GRAD;g];
ALFA = [ALFA;alfa];
x = x - alfa*g
fx = F(x,sf,1)
if k1 >= kmax
    warning = warning + 1;
else
    warning = 0;
end
X = [X;x];Y = [Y;fx];
if (warning >= 2)|(norm(x-x0)<TolX) & (abs(fx-fx0)<TolFun)
    break
end
x0 = x
fx0 = fx
end

g = F(x,sf,2);
if tip_kor == 1
    g = g/norm(g)
end
GRAD = [GRAD;g];

if warning == 2
    disp('Pozor: 2x zaporedoma je k1 presegel kmax!!!')
end

if tip_met == 1
    disp('koncni rezultat (linesearch = parab):')
else
    disp('koncni rezultat (linesearch = sekant):')
end

xfinal = x
fxfinal = fx

if dim == 2
    % pri 2D problemu lahko se sliko za konvergiranje k resitvi narisemo
    ols_spaceplotter_steep([X Y]);
    figure(1)
    title('Konvergenca k resitvi pri Steepest descent metodi')
    figure(2)
    title('Konvergenca k resitvi pri Steepest descent metodi')
end

return

%-----
%                PODPROGRAMI
%-----

%-----
% ols_steep_parab.m - metoda aproksimacije s parabolo
%
% klic funkcije: [alfa,k1] = ols_steep_parab(x,sf,alfa,kmax)
%
% Dokazat se da, da pri ekvidistanci h velja (spelji v wordu, glej knjigo Applied numer. meth. using Matlab,
% formula 7.1.5):
% xap = 0.5*(f2*(x1^2-x3^2)+f1*(x3^2-x2^2)+f3*(x2^2-x1^2))/(f2*(x1-x3)+f1*(x3-x2)+f3*(x2-x1)) =

```

```

%  $x_1 = x_0 + h \cdot (3 \cdot f_1 - 4 \cdot f_2 + f_3) / (-2 \cdot f_1 + 4 \cdot f_2 - 2 \cdot f_3)$ , ce je  $x_1 = x_0 + h$  in  $x_2 = x_1 + h$ ;
%
% To pa lahko se poenostavimo, ce  $x_1$  in  $h$  skrijemo v alfa:
%  $\alpha = \alpha \cdot (3 \cdot f_1 - 4 \cdot f_2 + f_3) / (-2 \cdot f_1 + 4 \cdot f_2 - 2 \cdot f_3)$ 
%-----
function [alfa,k1] = ols_steep_parab(x0,sf,alfa,kmax)

TolFun = 1e-8;

f1 = F(x0,sf,1);
g = F(x0,sf,2);
g = g/norm(g);

f2 = F(x0-2*alfa*g,sf,1); % funkcija v zacetnih treh tockah

for k1 = 1:kmax % blizaj se minimumu, dokler ni dosezena predp.tocnost (se srednja tocka ne sprem. vec)
    f3 = f2;
    f2 = F(x0-alfa*g,sf,1);
    if (f1>f2+TolFun)&(f2+TolFun<f3)
        alfa = alfa*(3*f1-4*f2+f3)/(-2*f1+4*f2-2*f3)
        alfa = - alfa;
        return
    else
        alfa = alfa/2;
    end
end

return

%-----
% ols_steep_sekant.m - Sekantna metoda
%-----
function [alfa,k] = ols_steep_sekant(x0,sf,alfa,kmax)

% resolucija
Tolalfa = 1e-5;

alfa_2 = 0;

g = F(x0,sf,2);
g = g/norm(g);

df1 = g;
df2 = df1;

% glavna zanka: poteka, dokler  $\text{abs}(x_k(k) - x_k(k-1)) > \text{eps}$  (isto kot pri Newtonovi metodi)

for k = 1:kmax
    alfa_1 = alfa_2;
    alfa_2 = alfa;
    df1 = df2;
    df2 = F(x0+alfa_2*g,sf,2);
    alfa = (df2*alfa_1-df1*alfa_2)/(df2-df1); % Izpeljemo iz  $\alpha_2 - \text{df}_2 \cdot (\alpha_2 - \alpha_1) / (\text{df}_2 - \text{df}_1)$ 
    if alfa < 0
        alfa = - alfa;
    end
    if abs(alfa_2-alfa) < Tolalfa
        break
    end
end

return

%-----
% ols_spaceplotter_steep.m: Prikaz konvergence (steepest) pri 2D problemu
%-----

```

```

function ols_spaceplotter_steep(pts)

% Inputs:
% Pts - 2D data array of x,y,z points
%
% Outputs:
% x1 - vector of x1 axis values
% x2 - vector of x2 axis values
% f - 2D array of function values

% Close all open figures
close all

%-----
% generacija prve slike
%-----

% Define Plot Limits
d = 0.2;
x1_min = -3;
x1_max = 5;
x2_min = -3;
x2_max = 5;

% Define contour lines
c = [-2 -1 -0.2 -0.1 -0.05 0 1 1.8 2 3 4 6 8 10];

% Define x1 and x2
x1 = x1_min:d:x1_max;
x2 = x2_min:d:x2_max;

warning off

% Compute function values
for i = 1:length(x1)
    for j = 1:length(x2)
        f(j,i) = F2d([x1(i),x2(j)]);
    end
end

warning on

% Display Plots
% Contour Plot
figure
hold on
[C,h] = contour(x1,x2,f,c,'-k');
clabel(C,h);

% Plot function only if no point data supplied
if nargin == 1
    plot3(pts(1,:),pts(2,:),pts(3,:), 'r', 'LineWidth', 2)
    plot3(pts(1,:),pts(2,:),pts(3,:), '-ro', 'LineWidth', 2)
end

show_opt(pts)

hold off
xlabel('x1');
ylabel('x2');
title('Konvergenca k minimumu pri steepest metodi (vred.funk.na contour krogih)')

%-----
% generacija druge slike
%-----

d = 0.05;
x1_min = -0.5;
x1_max = 2;
x2_min = -0.5;

```

```

x2_max = 2;

% Define x1 and x2
x1 = x1_min:d:x1_max;
x2 = x2_min:d:x2_max;

warning off

% Compute function values
for i = 1:length(x1)
    for j = 1:length(x2)
        f(j,i) = F2d([x1(i),x2(j)]);
    end
end

warning on

figure
mesh(x1,x2,f)
hold on
[C,h] = contour(x1,x2,f,c,'-k');
clabel(C,h);

% Plot function only if no point data supplied
if nargin == 1
    plot(pts(1,:),pts(2,:),'r','LineWidth',2)
    plot(pts(1,:),pts(2,:),'-ro','LineWidth',2)
end

hold off
xlabel('x1');
ylabel('x2');
zlabel('f = 4*x1^2 - sin(x1+x2)+ 2*x2^2')
title('Konvergenca k minimumu pri steepest metodi')

return

%-----
% show_opt(pts)
%-----
function show_opt(pts)

d = axis;

x1opt=pts(1,end);
x2opt=pts(2,end);
X2OPT=linspace(d(3),x2opt,10);
X2X = x1opt*ones(length(X2OPT),1);
plot(X2X,X2OPT,'-', 'LineWidth',1.5)
X1OPT=linspace(d(1),x1opt,10);
X1X = x2opt*ones(length(X1OPT),1);
plot(X1OPT,X1X,'-', 'LineWidth',1.5)
%plot(x1opt,x2opt,'p', 'LineWidth',4)
text(x1opt,d(3),num2str(x1opt))
text(d(1),x2opt,num2str(x2opt))
return

%-----
% Nasa funkcija (pri 2D prikazu ols_space_plotter_steep)
% (mora biti seveda ista kot v naslednjem podprogramu)
%-----
function f = F2d(x)

x1 = x(1);
x2 = x(2);
f = 4*x1^2 - sin(x1+x2)+ 2*x2^2;

```

```

return

%-----
% F - nabor nasih testnih funkcij
%-----
function objekt1 = F(x,sf,objekt)

if sf == 1
    x1 = x(1);
    x2 = x(2);
    f = 4*x1^2 - sin(x1+x2) + 2*x2^2;
    df = [8*x1 - cos(x1+x2) 4*x2 - cos(x1+x2)];
    d2f = [8 + sin(x1+x2) sin(x1+x2); sin(x1+x2) 4 + sin(x1+x2)];
elseif sf == 2
    x1 = x(1);
    x2 = x(2);
    x3 = x(3);
    f = x1^4 + x1^3 - x1 + x2^4 - x2^2 + x2 + x3^2 - x3 + x1 * x2 * x3;
    df = [4*x1^3 + 3*x1^2 - 1 + x2*x3 4*x2^3 - 2*x2 + x1*x3 2*x3 - 1 + x1*x2];
    d2f = [12*x1^2 + 6*x1 x3 x2; x3 12*x2^2 - 2 x1; x2 x1 2];
end

if objekt == 1
    objekt1 = f;
elseif objekt == 2
    objekt1 = df;
else
    objekt1 = d2f;
end

return

```

### 9.1.7 Algoritem metode najstrmejšega spusta s pomočjo dodatne uporabe Symbolic toolboxa (Matlab)

```

%-----
% ols_steep_simbol.m
%-----
function REZ = ols_steep_simbol(tip_koraka) % tip_koraka = 1 (variabilen), 2(fiksen)

close all
clc

disp('ols_steep_simbol')

if tip_koraka == 2 % konst. korak
    alfa_fix = 2/20; % zberemo neko, naceloma manjsa mora bit od 2/max_eig(Hesse) - glej Gershgorin.kroge
end

maxiter = 100;
TolX = 1e-3;

syms x1s x2s x3s alfas
fs = x1s^4 + x1s^3 - x1s + x2s^4 - x2s^2 + x2s + x3s^2 - x3s + x1s * x2s * x3s

dfx1s = diff(fs,x1s)
dfx2s = diff(fs,x2s)
dfx3s = diff(fs,x3s)
dfs = [dfx1s dfx2s dfx3s]

x0 = [0 0 0]
f0 = subs(fs, {x1s,x2s,x3s}, {x0(1),x0(2),x0(3)})

```



```

X = x0; ALFA = [NaN]; DFX = []; FX = f0;

for i = 1:maxiter

clear df0 xns x1s x2s x3s falfas f0
syms x1s x2s x3s

df0 = subs(dfs, {x1s,x2s,x3s}, {x0(1),x0(2),x0(3)})
DFX = [DFX;df0];
%df0 = df0/norm(df0)

xns = x0 - alfas*df0;
disp('xns = x0 - alfas*df0')

x1s = xns(1);
x2s = xns(2);
x3s = xns(3);

falfas=subs(fs, {x1s,x2s,x3s}, {'x1s','x2s','x3s'});

disp('')
disp('falfas=fs(xns)')

if tip_koraka == 1
    alfa_opt = ols_steep_par_ena_ko_ls(falfas);
else
    alfa_opt = alfa_fix
end

ALFA = [ALFA;alfa_opt];

if i > 1
    al_pl = 0:0.01:0.5;
else
    al_pl = 0.5:0.01:1;
end

fal_pl = [];
for ii = al_pl;
    fal_pl = [fal_pl;subs(falfas,alfas,ii)];
end

if tip_koraka == 1
    figure
    plot(al_pl,fal_pl,'LineWidth',2)
    hold on,grid
    plot(al_pl,fal_pl,'o','LineWidth',2)
    fdisp1 = ' f(alfa)=f(x(i-1)-alfa*df(i-1))=f(x1(alfa),x2(alfa),x3(alfa))=ff(alfa)';
    fdisp2 = ' ff=x1^4+x1^3-x1+x2^4-x2^2+x2+x3^2-x3+x1*x2*x3';
    fdisp3 = ['alfa_opt(i)= num2str(alfa_opt)];
    title(['i=' num2str(i) fdisp1 fdisp2])
end

xold = x0;
x0 = x0 - alfa_opt*df0

if tip_koraka == 1
    fdisp4 = ['x(i-1)= num2str(xold)];
    fdisp5 = [' x(i)=x(i-1)-alfa_opt(i)*df(i-1)= num2str(x0)];
    fdisp6 = [' df(i-1)= num2str(df0)];
    ylabel([fdisp4 fdisp6])
    xlabel([fdisp3 fdisp5])
end

X =[X;x0];
x1 = x0(1);x2 = x0(2);x3 = x0(3);
f0 = x1^4+x1^3-x1+x2^4-x2^2+x2+x3^2-x3+x1*x2*x3
FX =[FX;f0];

if norm(xold-x0)<ToIX

```

```

        break
    end

end

disp('-----')
disp('KONCEN REZULTAT')
disp('REZ = [[0:1:i] X ALFA [DFX;NaN NaN NaN] FX]')

REZ = [[0:1:i]' X ALFA [DFX;NaN NaN NaN] FX];
disp('xopt(it)=xopt(it-1)-alfa_opt(it)*dfopt(it-1) pri it>0')

return

%-----
%                PODPROGRAMI
%-----

%-----
% ols_steep_par_ena_k_ls.m - metoda aproksimacije s parabolom
%
% klic funkcije: alfa_opt = ols_steep_par_ena_k_ls(falfas);
%-----
function alfa_opt = ols_steep_par_ena_k_ls(falfas)

x1 = 0;
x2 = 1;          % zacetne 3 tocke
x3 = 2;

a0 = 0; b0 = 1; % interval za prikaz

f1 = funkc_a_p(x1,falfas);
f2 = funkc_a_p(x2,falfas); % funkcija v zacetnih treh tockah
f3 = funkc_a_p(x3,falfas);

if ~((f1>f2)&&(f2<f3)) % test pravilnosti zacetnih tock
    REZ = [];
    disp('Nisi dobro izbral zacetnih tock!!!')
    return
end

eps = 1e-5; % resolucija
dx = 0.05; % korak pri prikazu funkcije

xkold = x2; % zacetna stara vrednost srednje tocke
xknew = x1; % nekaj

k = 0;
while abs(xknew-xkold)>eps % blizaj se minimumu, dokler ni dosezena predp.tocnost (se srednja tocka ne sprem. vec)
    k = k + 1;
    xap = 0.5*(f2*(x1^2-x3^2)+f1*(x3^2-x2^2)+f3*(x2^2-x1^2))/(f2*(x1-x3)+f1*(x3-x2)+f3*(x2-x1)); % 1. nacin-analiticno

    if xap < x2 % ce je minimum trenutne parabole manjsi od srednje tocke
        x3 = x2; % naj desna tocka postane srednja tocka
        f3 = funkc_a_p(x3,falfas); % in funkcija v njej
    else % sicer je minimum trenutne parabole vecji od srednje tocke
        x1 = x2; % naj leva tocka postane srednja tocka
        f1 = funkc_a_p(x1,falfas); % in funkcija v njej
    end
    xkold = x2; % stara srednja tocka
    x2 = xap; % nova srednja tocka naj bo minimum trenutne parabole
    f2 = funkc_a_p(xap,falfas); % in funkcija v njej
    xknew = x2; % nova srednja tocka
end

alfa_opt = xknew

```

```
return
```

```
%-----  
%          Generacija nase funkcije  
%-----  
function falfa = funkc_a_p(x,falfas)  
  
syms alfas  
  
falfa = subs(falfas,alfas,x);  
  
return
```

### 9.1.8 Algoritem Newtonove metode za več spremenljivk (Matlab)

```
%-----  
% ols_newt_multi  
%  
% dim = dimenzija problema, sf je številka testne funkcije  
%  
% klic funkcije: [X,Y,GRAD,HESSE] = ols_newt_multi(3,2,[1 -1 1]); ali [X,Y,GRAD,HESSE] = ols_newt_multi(2,1,[-1.5 1]);  
%-----  
function [X,Y,GRAD,HESSE] = ols_newt_multi(dim,sf,x0)  
  
close all  
clc  
disp('Newton_multi')  
  
MaxIter = 200  
TolX = 1e-6  
TolFun = 1e-6  
  
x = x0  
fx0 = F(x0,sf,1)  
fx = fx0  
  
X = x; Y = fx; GRAD = []; HESSE = [];  
  
if det(F(x,sf,3))==0  
    disp('Pozor:Hessejeva matrika je singularna!!!')  
    return  
end  
  
for k = 1:MaxIter  
    g = F(x,sf,2)  
    H = F(x,sf,3)  
    GRAD = [GRAD;g];  
    HESSE = [HESSE;H];  
    x = x - (inv(H)*g)'  
    fx = F(x,sf,1)  
    X = [X;x];Y = [Y;fx];  
    if ( norm(x-x0)<TolX) & (abs(fx-fx0)<TolFun) )  
        break  
    end  
    x0 = x  
    fx0 = fx  
end  
  
g = F(x,sf,2);  
GRAD = [GRAD;g];  
HESSE = [HESSE;H];  
  
disp('koncni rezultat (Newton multi):')
```

```

xfinal = x
fxfinal = fx

if dim == 2 % pri 2D problemu lahko se sliko za konvergiranje k resitvi narisemo
    ols_spaceplotter_steep([X Y]);
    figure(1)
    title('Konvergenca k resitvi pri Newton multi metodi')
    figure(2)
    title('Konvergenca k resitvi pri Newton multi metodi')
end

return

```

```

%-----
%                PODPROGRAMI
%-----

%-----
% ols_spaceplotter_steep.m: Prikaz konvergence (steepest) pri 2D problemu
%-----

```

```
function ols_spaceplotter_steep(pts)
```

```

% Inputs:
% Pts - 2D data array of x,y,z points
%
% Outputs:
% x1 - vector of x1 axis values
% x2 - vector of x2 axis values
% f - 2D array of function values

```

```

% Close all open figures
close all

```

```

%-----
% generacija prve slike
%-----

```

```
% Define Plot Limits
```

```

d = 0.2;
x1_min = -3;
x1_max = 5;
x2_min = -3;
x2_max = 5;

```

```
% Define contour lines
```

```
c = [-2 -1 -0.2 -0.1 -0.05 0 1 1.8 2 3 4 6 8 10];
```

```
% Define x1 and x2
```

```

x1 = x1_min:d:x1_max;
x2 = x2_min:d:x2_max;

```

```
warning off
```

```
% Compute function values
```

```

for i = 1:length(x1)
    for j = 1:length(x2)
        f(j,i) = F2d([x1(i),x2(j)]);
    end
end

```

```
warning on
```

```
% Display Plots
```

```
% Contour Plot
```

```

figure
hold on

```

```

[C,h] = contour(x1,x2,f,c,'-k');
clabel(C,h);

% Plot function only if no point data supplied
if nargin == 1
    plot3(pts(1,:),pts(2,:),pts(3,:),r',LineWidth',2)
    plot3(pts(1,:),pts(2,:),pts(3,:),-ro',LineWidth',2)
end

show_opt(pts)

hold off
xlabel('x1');
ylabel('x2');
title('Konvergenca k minimumu pri Newton multi metodi (vred.funk.na contour krogih)')

%-----
% generacija druge slike
%-----

d = 0.05;
x1_min = -0.5;
x1_max = 2;
x2_min = -0.5;
x2_max = 2;

% Define x1 and x2
x1 = x1_min:d:x1_max;
x2 = x2_min:d:x2_max;

warning off

% Compute function values
for i = 1:length(x1)
    for j = 1:length(x2)
        f(j,i) = F2d([x1(i),x2(j)]);
    end
end

warning on

figure
mesh(x1,x2,f)
hold on
[C,h] = contour(x1,x2,f,c,'-k');
clabel(C,h);

% Plot function only if no point data supplied
if nargin == 1
    plot(pts(1,:),pts(2,:),r',LineWidth',2)
    plot(pts(1,:),pts(2,:),-ro',LineWidth',2)
end

hold off
xlabel('x1');
ylabel('x2');
zlabel('f = 4*x1^2 - sin(x1+x2)+ 2*x2^2')
title('Konvergenca k minimumu pri steepest metodi')

return

%-----
% show_opt(pts)
%-----
function show_opt(pts)

d = axis;

x1opt=pts(1,end);

```

```

x2opt=pts(2,end);
X2OPT=linspace(d(3),x2opt,10);
X2X = x1opt*ones(length(X2OPT),1);
plot(X2X,X2OPT,'-', 'LineWidth',1.5)
X1OPT=linspace(d(1),x1opt,10);
X1X = x2opt*ones(length(X1OPT),1);
plot(X1OPT,X1X,'-', 'LineWidth',1.5)
%plot(x1opt,x2opt,'p', 'LineWidth',4)
text(x1opt,d(3),num2str(x1opt))
text(d(1),x2opt,num2str(x2opt))
return

```

```

%-----
% Nasa funkcija (pri 2D prikazu ols_space_plotter_steep)
% (mora biti seveda ista kot v naslednjem podprogramu)
%-----
function f = F2d(x)

```

```

x1 = x(1);
x2 = x(2);
f = 4*x1^2 - sin(x1+x2)+ 2*x2^2;

return

```

```

%-----
% F - nabor nasih testnih funkcij
%-----
function objekt1 = F(x,sf,objekt)

```

```

if sf == 1
    x1 = x(1);
    x2 = x(2);
    f = 4*x1^2 - sin(x1+x2)+ 2*x2^2;
    df = [8*x1-cos(x1+x2) 4*x2-cos(x1+x2)];
    d2f = [8+sin(x1+x2) sin(x1+x2); sin(x1+x2) 4+sin(x1+x2)];
elseif sf == 2
    x1 = x(1);
    x2 = x(2);
    x3 = x(3);
    f = x1^4+x1^3-x1+x2^4-x2^2+x2+x3^2-x3+x1*x2*x3;
    df = [4*x1^3+3*x1^2-1+x2*x3 4*x2^3-2*x2+x1*x3+1 2*x3-1+x1*x2]; % TU SEM 1 PRISTEL V 2. CLENU!!!!!!
    d2f = [12*x1^2+6*x1 x3 x2; x3 12*x2^2-2 x1; x2 x1 2];
end

if objekt == 1
    objekt1 = f;
elseif objekt == 2
    objekt1 = df;
else
    objekt1 = d2f;
end

return

```

### ***9.1.9 Algoritem DFP QuasiNewtonove metode za več spremenljivk s pomožno analitično enodimenzionalno optimizacijo v Symbolic toolboxu (Matlab)***

```

% dfp_m.m

close all

```

```

clear

TOL = 1E-3;
N = 100;
dim = 2;

syms x1s x2s

fs = input('vnesi funkcijo v pravilni sintaksi (npr. x1s^2-2*x1s*x2s+2*x2s^2-x1s-2*x2s:');

x0 = input('x0= (npr. [0 0])')

dfx1s = diff(fs,x1s)
dfx2s = diff(fs,x2s)

dfs = [dfx1s dfx2s]

g0 = subs(dfs,{x1s,x2s},{x0(1),x0(2)})
g0old = g0;

G0 = eye(dim);

for i = 1:N
    clear x1 alfa
    syms x1 alfa
    x1 = x0' - alfa*G0*g0';
    falfa = subs(fs,{x1s,x2s},{x1(1),x1(2)})
    dfalfa = diff(falfa,alfa)
    alfa_opt = solve(dfalfa)
    x1 = subs(x1,alfa,alfa_opt)
    g0 = subs(dfs,{x1s,x2s},{x1(1),x1(2)})
    dg0 = g0 - g0old;
    dx0 = x1 - x0';
    M = dx0*dx0'/(dx0*dg0')
    N = (G0*dg0*dg0'*G0)/(dg0*dg0')
    G0 = G0 + M - N
    if norm(double(dx0)) < TOL
        break
    else
        x0 = x1';
        g0old = g0;
    end
end

disp('optimum je:')
x1

```

## LITERATURA

- [1] Ahuja, R. K.: Network flows: Theory, algorithms, and Applications, Prentice Hall, 1993.
- [2] Chong, E. K. P.: An Introduction To Optimization, Wiley-Interscience, 2 Sub Edition, 2001.
- [3] Čižman A.: Operacijske raziskave, Teorija in uporaba v organizaciji, Moderna organizacija Kranj, 2004.
- [4] Dragan, D.: Optimizacija logističnih sistemov, prosojnice, računalniški programi in ostala gradiva z vaj, Fakulteta za logistiko, Univerza v Mariboru, Celje, 2008.
- [5] Dragan, D.: Optimizacija logističnih sistemov, Zbirka rešenih nalog, Fakulteta za logistiko, Univerza v Mariboru, Celje, 2008.
- [6] Dragan D.: Uporaba Hooke-Jeevesove metode nelinearnega programiranja pri načrtovanju optimalnih helikopterskih poti, Osnutek članka za zbornik, Fakulteta za logistiko, Celje, 2008.
- [7] Dragan, D.: Upravljanje logističnih sistemov, učbenik, Fakulteta za logistiko, Univerza v Mariboru, Celje, 2009.
- [8] Fletcher, R.: Practical Methods Of Optimization, John Wiley & Sons, 2 Sub Edition, 2000.
- [9] Fogiel M., The Operations Research Problem Solver, Research And Education Association, New Jersey, 1989.
- [10] Hillier, F.S.: Introduction to Operations Research, McGraw-Hill, 2001.
- [11] Kunič J.: Metode optimiranja, Moderna organizacija Kranj, 1986.
- [12] Kunič J.: Operacijske raziskave, Moderna organizacija Kranj, 1988.
- [13] Matko D., Bratkovič F.: Računalniško inženirstvo v vodenju sistemov, Fakulteta za elektrotehniko in računalništvo, Univerza v Ljubljani, Ljubljana, 1993.
- [14] McKeown, J.: An introduction to unconstrained optimisation, Bristol, New York, Cambridge: ESM, cop. 1990.
- [15] Onwubiko, C. O.: Introduction to Engineering design optimization, Upper Saddle River: Prentice-Hall, cop. 2000.
- [16] Paoluzzi A.: Geometric programming for Computer-Aided Design, Wiley, 2003.



- [17] Petrič J., Zlobec S.: Nelinearno programiranje, Naučna knjiga, Beograd, 1983.
- [18] Petrič J., Kojič Z., Šarenac L.: Operaciona istraživanja, Zbirka rešenih zadataka, knjiga 1, Naučna knjiga, Beograd, 1996.
- [19] Pike R.: Optimization for Engineering Systems, Van Nostrand Reinhold Company, 1986.
- [20] Povh J., Pustavrh S.: Zbirka rešenih nalog iz operacijskih raziskav, Novo Mesto, 2001.
- [21] Rao, Engineering Optimization: Theory and Practice, Wiley-Interscience, 3 Sub Edition, 1996.
- [22] Rardin, R.: Optimization In Operations Research, Prentice Hall, 1st Edition, 1997.
- [23] Salkin, M. S. : Integer Programming, Addison-Wesley, 1974.
- [24] Taha H.: Operations Research, An Introduction, Sixth Edition, Prentice-Hall, 1997.
- [25] Usenik J.: Kvantitativne metode v logistiki, Valvazorjev raziskovalni center Krško, 2006.
- [26] Vadnal A.: Linearno programiranje, Ekonomska Fakulteta, Ljubljana, 1961.
- [27] Vadnal A.: Diskretno dinamično programiranje, Državna založba Slovenije, Ljubljana, 1976.
- [28] Vanderplaats, G. N.: Numerical Optimization Techniques For Engineering Design, Vanderplaats Research and Development, Inc, 3rd Ed., 4th Printing Edition, 2001.
- [29] Vrančić, D., Dragan, D.,: Optimizacija logističnih sistemov, prosojnica s predavanj, Fakulteta za logistiko, Univerza v Mariboru, Celje, 2008.
- [30] Winston, W., L. : Operations Research, Applications and Algorithms, Duxbury press, International Thomson Publishing, 1994, ISBN 0-534-20971-8.
- [31] Zadnik Stirn L.: Metode operacijskih raziskav za poslovno odločanje, Novo Mesto, 2001.
- [32] Shewchuk J. R.: An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, Carnegie Mellon University, 1994.