

UNIVERZITET CRNE GORE
MAŠINSKI FAKULTET PODGORICA

OPTIMIZACIONE METODE U KONSTRUKCIJAMA

U Podgorici, 2006.godine

Prof. dr Olivera Jovanović

[] SADRZAJ

1. PARAMETARSKA IDENTIFIKACIJA SISTEMA

- 1.1. Postupak identifikacije sistema
- 1.2. Kriterijum odstupanja
- 1.3. Algoritam za optimizaciju sistema

2. OPTIMIZACIJA

- 2.1. Uvod u optimizaciju
- 2.2. Primjer optimizacije - optimizacija cijene padobrana
- 2.3. Matematička osnova problema

3. JEDNODIMENZIONALNA OPTIMIZACIJA

- 3.1. Metode upoređivanja vrijednosti funkcije
- 3.2. Metod zlatnog reza
- 3.3. Njutnova metoda
- 3.4. Metoda sječice
- 3.5. Metode aproksimacije polinomom

4. VISEDIMENZIONALNA NEOGRANICENA OPTIMIZACIJA BEZ IZRACUNAVANJA DERIVACIJA

- 4.1. Metoda traženja po koordinatnim osama
- 4.2. Hooke – Jeeves-ov metod
- 4.3. Powell-ov metod

5. VISEDIMENZIONALNA NEOGRANICENA OPTIMIZACIJA ZA DERIVABILNE FUNKCIJE

- 5.1. Gradijentne metode
 - 5.1.1. Gradijentne metode prvog reda
 - 5.1.1.1 Cauchy-eva metoda najstrmijeg opadanja
 - 5.1.1.2 Fletcher-Reeves-ov metod konjugovanih gradijenata
 - 5.1.2. Gradijentne metode drugog reda
 - 5.1.2.1 Njutnova metoda
 - 5.1.2.2 Gauss- Njutnova metoda
- 5.2. Metode varijabilne metrike

6. Literatura

1. PARAMETARSKA IDENTIFIKACIJA SISTEMA

1.1. POSTUPAK IDENTIFIKACIJE SISTEMA

Pod parametarskom identifikacijom sistema podrazumijeva se proces definisanja oblika matematičkog modela razmatranog fizičkog sistema kada su za isti poznate funkcije ulaza i odgovora. Veoma jasnu interpretaciju parametarske identifikacije dao je Bekey (1970). Ona se sastoji iz tri dijela i to:

1. Definisanje oblika modela koji obuhvata izbor diferencijalnih jednačina i selekciju nepoznatih parametara.
2. Izbor kriterijuma odstupanja (formiranje funkcije greske) kojim se obezbedjuje "tačnost podešavanja" odgovora modela i odgovora fizičkog sistema, ako su oba pobuđena istom funkcijom ulaza.
3. Definisanje algoritma za sistematsko podešavanje nepoznatih parametara, tako da razlika u odgovorima između modela i fizičkog sistema, bude minimalna.

Prvi i najvažniji korak kod procesa identifikacije sistema predstavlja definisanje oblika matematičkog modela razmatranog fizičkog sistema. Za neadekvatno odabrani matematički model nemoguće je zadovoljiti uslov o kriterijumu odstupanja, bez obzira koliko se dobar algoritam razvije za sistematsko podešavanje parametara. Dobrim izborom matematičkog modela povećava se mogućnost za kvalitativnije upoređenje poznatih fizičkih veličina.

1.2. KRITERIJUM ODSUPANJA

Ocjena valjanosti modela izražava se kroz kriterijum ekvivalencije, što se podrazumijeva neki kriterijum greške. U ovom slučaju greška predstavlja razliku između matematičkog modela i fizičkog sistema ako su oba pobuđeni istom funkcijom ulaza.

Funkcija greške se može formirati na više načina i to kao: .

$$J(\beta, T) = \int_0^T |x(\beta, T) - y(t)|^T |x(\beta, T) - y(t)| dt, \quad (1.1)$$

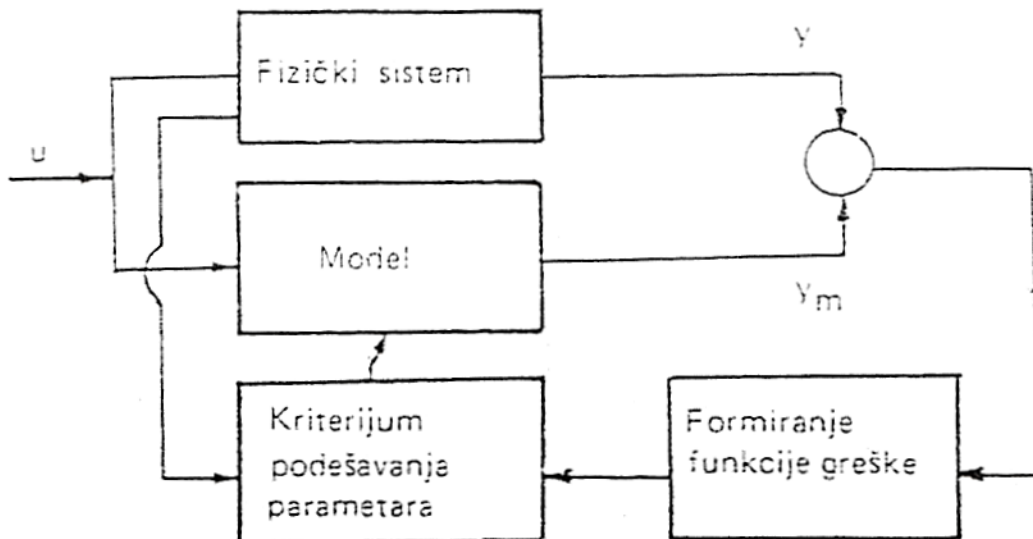
a) Suma kvadrata pojedinih grešaka, gdje je β vektor nepoznatih parametara za identifikaciju.

b) Težinska suma kvadrata pojedinih funkcija:

$$J(\beta, T) = \sum_{k=1}^n t_k J_k(\beta, T)$$

Izborom težinskih faktora može se uzeti u obzir relativna važnost pojedinih komponenti.

c) Kriterijum maksimalne apsolutne vrijednosti koji se manje koristi prethodna dva $J(\beta) = \max J_k(\beta)$.
 Izbor kriterijuma odstupanja zauzima značajno mjesto u postupku identifikacije, posebno kada je riječ o ocjenjivanju parametara. Imajući u vidu da na objekat i model djeluje jedan te isti ulazni signal (sl. 1.1) i da se upoređuju vrijednosti izlaza objekta i izlaza modela, potrebno je tako izabrati kriterijum koji će omogućiti da se izvrši neophodna korekcija modela. Činjenica da nepoznati parametri nijesu dostupni mjerenju zahtijeva da kriterijum odstupanja bude funkcional izlaznih signala ili greška ocjene parametara, imajući u vidu da je izlazni signal funkcionalno povezan sa numeričkim vrijednostima parametara.



Sl.1.1. Blok dijagram identifikacije sistema (po Bekey)

Rješavanje postavljenih zadataka identifikacije svodi se na iznalaženje minimuma funkcije greške. Međutim, treba postaviti još neka dopunska pitanja koja mogu unijeti više svjetla u ovu složenu problematiku. Tako, na primjer, kako izabrati funkciju greške, da li ta funkcija posjeduje minimum, postoji li jedinstveno rješenje, da li je jedinstvenost rješenja povezana sa klasom, ulaznih signala itd.

U opštem slučaju grešku možemo definisati analitički kao $J=J(e)$. Kriterijum odstupanja razmatraćemo sa više gledišta uslovljenih načinom na koji je ona definisana:

Ako razmatramo slučaj na slici 1.1 tada govorimo o izlaznoj grešci: $e=y-y_M=y-M(u)$, gdje je $M(u)$ izlazni signal modela pobuđenog ulaznim signalom u .

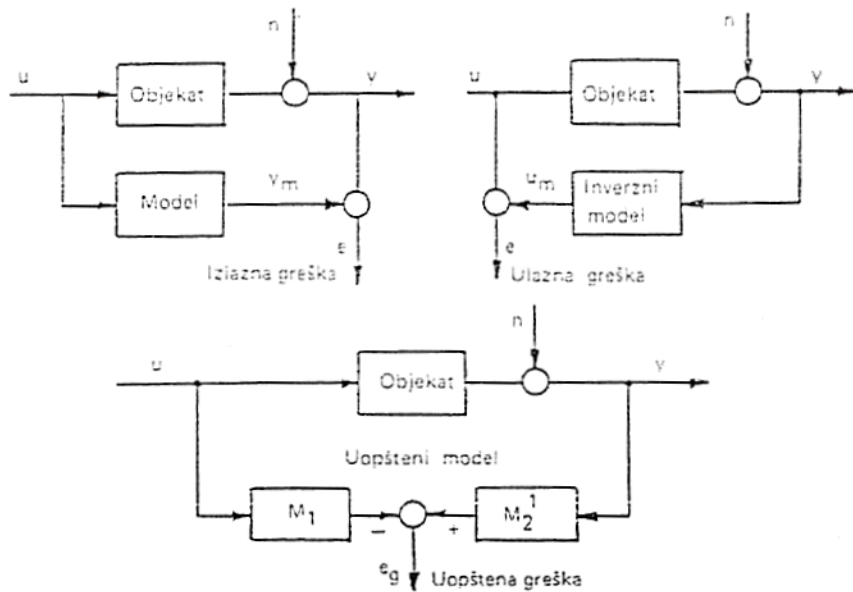
Ako je $e=u-u_M = u-M^{-1}(y)$, onda je to greška ulaza. $M^{-1}(y)$ je inverzni model tj. model koji za dati izlazni signal može reprodukovati jedinstveni (jedan i samo jedan) ulazni signal.

Grešku je moguće definisati i na sljedeći način:

$$e=M_2^{-1}(y)-M_1(u), \text{ gdje } M_2^{-1}(y) \text{ predstavlja inverzni model.}$$

Na ovaj način definisana greška naziva se uopštena ili generisana greška.

Na slici 1.2, data je interpretacija prethodno definisanih grešaka u vidu blok dijagrama.



SI.1.2. Blok dijagram izlazne greške, greške ulaza i uopstene greške po Eykhoff-u

Ukoliko razmatramo postupak ocjenjivanja nepoznatih parametara modela u cilju što boljeg podešavanja nepoznatog modela stvarnim karakteristikama objekta, tada posmatramo blok $J(\beta)$ prikazan na slici 1.1. Blok $J(\beta)$ koji prima i obrađuje informaciju o dinamičkom ponašanju izlaznih veličina objekta i modela može se predstaviti na sljedeći način:

$$J=J(y, y_M) = J(e) = \int_0^T e^q(t, \beta) dt, \quad q = 1, 2, 3 \dots$$

Postavlja se pitanje izbora parne ili neparne podintegralne funkcije greške. Izbor parnog funkcionala dovodi do zavisnosti kriterijuma ekvivalencije i podešljivih parametara modela što sa računskog stanovišta olakšava posao.

Kriterijum odstupanja treba da bude definisan kao razlika između analitičkih i eksperimentalnih veličina koje su međusobno nezavisne.

1.3. ALGORITAM ZA OPTIMIZACIJU SISTEMA

Treći korak u procesu parametarske identifikacije po Bekey predstavlja izbor algoritma za sistematsko podešavanje parametara matematičkog modela sve dok se ne dobije minimum funkcije greške. Ukoliko se radi o funkciji jedne promjenljive $J(\beta, T)$, minimum funkcije i vrijednost nepoznatog parametra traži se iz uslova:

$$\nabla J(\beta, T)_{\beta} = \beta_k = 0 \quad (1.2)$$

i to su problemi koji su predmet istraživanja u jednodimenzionalnoj optimizaciji.

Ako je riječ o funkciji greške kao funkciji više promjenljivih tada primjena izraza (1.2) vodi ka dobijanju sistema simultanih nelinearnih algebarskih jednačina po vektoru nepoznatih parametara β , čije analitičko rješenje je po prirodi veoma složeno. Zato je potrebno razviti i primijeniti takav algoritam za optimizaciju sistema koji će sigurno voditi ka rješenju problema.

Suština razvoja algoritama za optimizaciju sistema jeste minimiziranje funkcije kriterijuma odstupanja koja predstavlja neku površ, odnosno nalaženje vektora nepoznatih parametara koji leži u minimalnoj tački te površine.

Danas su u upotrebi mnoge metode za rješavanje problema minimizacije funkcije u zavisnosti od broja parametara. One su uglavnom iterativne metode i proces nalaženja minimuma počinje odabiranjem početnog vektora β_0 na osnovu koga se definiše novi vektor β_1 . Proces se ponavlja i definiše se novi vektor β_2 , zatim β_0 i tako dalje do traženog vektora β^* . Ukoliko je vektor β^* u oblasti traženog minimuma onda je to i rješenje problema. Međutim, može se desiti da se i ne dostigne numerički minimum što zavisi od izbora početnog vektora β_0 , kao i od razvijenog algoritma minimizacije koji treba da garantuje postizanje totalnog, a ne nekog lokalnog minimuma funkcije greške.

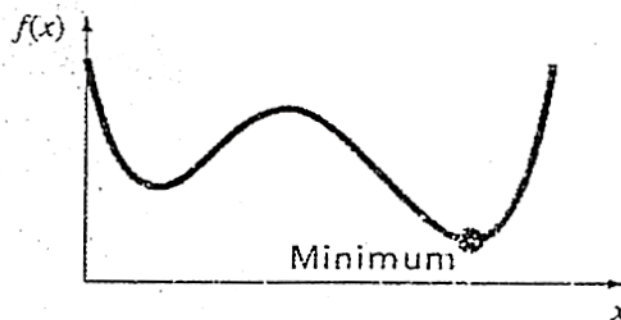
Ukoliko je funkcija kriterijuma odstupanja, odnosno funkcija greške, funkcija više promjenjivih tada se, za definisanje minimuma višedimenzionalne površine koriste metode bezuslovne optimizacije bez izracunavanja derivacija funkcije i gradijentne metode.

2. OPTIMIZACIJA

2.1. UVOD U OPTIMIZACIJU

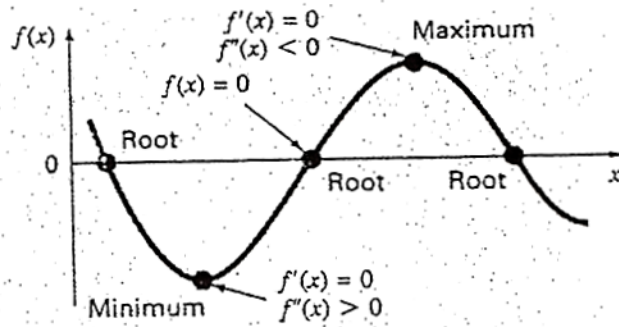
Optimizacija je primijenjena matematička disciplina u kojoj se nalaze ekstremne vrijednosti funkcije cilja. U konstruktivnom masinstvu optimizacijom se projektuju masine i djelovi masina sa najboljim osobinama kao što su najveća pouzdanost sistema, najveća efikasnost masina, najveća prilagodjenost svrsi, najmanja proizvodna cijena itd. Realni uslovi za praktičnu primjenu optimizacionih metoda stvoreni su spojem klasične i numeričke matematike i informacionih tehnologija.

Za optimalno projektovanje konstrukcija u tehnici metode nelinearnog programiranja su osnovne metode jer su funkcije cilja i funkcije ograničenja uglavnom nelinearne prirode. Problemi optimizacije uključuju određivanje jedne ili više vrijednosti nezavisno promjenljive koja se odnosi na „najbolju“ ili optimalnu vrijednost funkcije. Zato je potrebno identifikovati tačke maksimuma i minimuma funkcije Sl 2.1. Takvi problemi se javljaju kao rutinski u inženjerskoj literaturi bilo da se radi o funkcijama jedne ili više promjenljivih.



Sl 2.1 – Minimum funkcije

Optimizacija i problemi rješavanja jednačina su povezani u smislu da oba uključuju pretpostavljanje i traženje tačke na funkciji. Rješavanje jednačine je traženje nula funkcije oblika $f(x) = 0$, a problemi optimizacije su traženje maksimuma i minimuma funkcije i rješavaju se sa jednačinom $f'(x) = 0$ (Sl 2.2).



Sl 2.2 - Funkcija jedne promjenljive na kojoj se vidi razlika između korijena jednadžine i tačaka optimuma

Optimum je tačka gdje je kriva ravna i u matematičkom smislu $f'(x)=0$. Drugi izvod funkcije određuje maksimum sa $f''(x)<0$ i minimum $f''(x)>0$. Sada je razumljivo zašto se za određivanje optimuma razvila posebna strategija i to difrenciranje funkcije i nalaženje nula nove funkcije. Činjenica je da neke optimizacione metode nalaze rješenje putem jednačine $f'(x)=0$. Međutim treba reći da je ovakvo rješavanje komplikovano kada se $f'(x)$ ne može naci analitički, kao kod funkcija koje su predstavljene tačkastim vrijednostima. Tada se koriste konacne razlike umjesto izvoda funkcije. Posmatranje optimizacije kao rješavanje korijena odgovarajuće jednadžine pogodno je u jednostavnijim slučajevima kod jednodimenzionih problema gdje je razumijevanje problema potpuno.

2.2. Primjer optimizacije - optimizacija cijene padobrana

Postavka problema

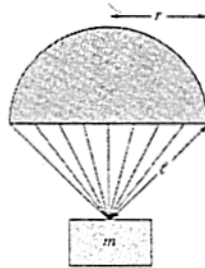
Bazični problem optimizacije i korišćenja numeričkih metoda ilustrovaćemo na primjeru padobrana. Na ovom tipično inženjerskom primjeru, ispitacemo brzinu udara padobrana o zemlju. Neka je zadatak spustiti putem padobrana pakete hrane u određenoj blizini izbjegličkog kampa u ratnim uslovima na najnižoj visini aviona od 500m. Padobran je prikazan na Sl 2.3 i otvara se odmah po napuštanju aviona. Da bi se smanjila šteta na paketima prilikom njihovog bacanja brzina udara treba da je manja od kritične brzine $v_c = 20 \frac{m}{s}$. Površina padobrana je pola sfere $A = 2r^2\pi$. Dužina svakog od 16 užadi koji spajaju padobran sa teretom je $l = r\sqrt{2}$. Vučna sila za padobran je linearna funkcija površine padobrana, odnosno

$$c = k_c A$$

gdje je:

c – vučni koeficijent (kg/s) i

k_c – koeficijent proporcionalnosti efekta površine na vuču (kg/sm²).



Sl 2.3 - Otvoren padobran

Takođe, težina svakog tereta je pojedinačno : $m = \frac{M_t}{n}$, gdje je M_t – ukupni teret ispušten padobranom (kg) i n – broj paketa.

Konačno, cijena svakog padobrana može se predstaviti sljedećim izrazom

$$\text{cijena padobrana} = c_0 + c_1 + c_2 \cdot A^2$$

gdje su c_0 , c_1 , c_2 – koeficijenti cijene. Konstanta c_0 je osnovna cijena padobrana. Nelinearan član $c_2 A^2$, koji se odnosi na površinu, dat je zbog toga što je skuplje konstruisati veliki nego mali padobran. Cilj je određivanje poluprečnika padobrana (r) i broja padobrana (n) uz minimalnu cijenu koštanja jednog padobrana pri čemu treba da je zadovoljen i kriterijum kritične brzine udara o zemlju v_c .

Rješenje :

Cilj je odrediti broj i veličinu padobrana uz minimalnu cijenu koštanja aviotransporta. Problem je sa određivanjem minimuma sa ograničenjem zato što brzina udara pošiljke o zemlju treba da je manja od kritične vrijednosti. Cijena koštanja ukupnog aviotransporta dobija se kada se individualna cijena padobrana pomnoži sa brojem padobrana (n) i tako se dobija funkcija cilja koju treba minimizirati.

$$C = n(c_0 + c_1 + c_2 A^2).$$

Sljedeće je postaviti ograničenja kojih u ovom zadatku ima dva i to:

- $v \leq v_c$,
- $n \geq 1$ (broj pošiljki mora biti ≥ 1).

Ovim je definisan zadatak optimizacije i mora se zaključiti da je to nelinearni ograničen problem. Iako je ovaj problem široko formulisan, on se može postaviti i na sljedeći način :

- Kako odrediti brzinu udara v ?

Brzina tijela koje pada može se izračunati po formuli :

$$v = \frac{g \cdot m}{c} \left(1 - e^{-\frac{c}{m}t} \right) \quad (2.1)$$

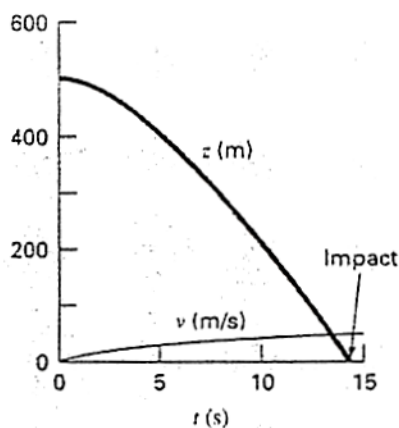
gdje je v – brzina (m/s)

g – ubrzanje zemljine teže (m/s^2)

m – masa tereta (kg) koji se spusta jednim padobranom

t – vrijeme (s)

Mada izraz (1) daje veza između brzine v i vremena t , neophodno je znati vrijeme padanja tijela. Zbog toga nam je potrebna veza između distance (z), udaljenosti tijela koje se spusta padobranom od zemlje, i vremena padanja t . Funkcija $z = f(t)$ je data na Sl. 2.4.



Sl. 2.4 - Kriva promjene rastojanja z od vremena t

Integraljenjem jednačine (2.1) dobija se udaljenost padobrana od zemlje pri padanju u zavisnost od vremena t :

$$z = \int_0^t \frac{g \cdot m}{c} \left(1 - e^{-\frac{c}{m}t} \right) dt$$

i njen konačan oblik :

$$z = z_0 - \frac{g \cdot m}{c} \cdot t + \frac{g \cdot m^2}{c^2} \left(1 - e^{-\frac{c}{m}t} \right)$$

gdje je z_0 – početna visina padanja.

Vrijeme padanja dobićemo iz ove jednačine uzimajući da je $z = 0$:

$$f(t) = 0 = z_0 - \frac{g \cdot m}{c} \cdot t + \frac{g \cdot m^2}{c^2} \left(1 - e^{-\frac{c}{m}t} \right)$$

i zamjenom u izraz za brzinu (2.1) dobijamo brzinu udara o zemlju.

Konačno, postavka ovog problema optimizacije je sljedeća: treba naći minimum cijene

$C = n(c_0 + c_1 + c_2 A^2)$ sa prvim ograničenjem $v \leq v_c$ i drugim ograničenjem $n \geq 1$, gdje je :

$$A = 2\pi r^2$$

$$l = r\sqrt{2}$$

$$c = k_c A$$

$$m = \frac{M_t}{n}$$

uz poznate relacije:

$$z = z_0 - \frac{g \cdot m}{c} \cdot t + \frac{g \cdot m^2}{c^2} \left(1 - e^{-\frac{c}{m}t} \right)$$

$$v = \frac{g \cdot m}{c} \left(1 - e^{-\frac{c}{m}t} \right)$$

Ovaj problem ćemo optimizacionim metodama riješiti kasnije.

Kod optimizacionih problema u širokoj tehničkoj praksi postoje neki fundamentalni elementi koji su prepoznatljivi kod svih i to :

- funkcija cilja tj. funkcija čija se ekstremna vrijednost traži (maksimum i minimum) i ona je obično naš cilj
- promjenljive koje figurišu u funkciji cilja i određuju njenu složenost. Mogu biti realne i cjelobrojne i u našem slučaju kod razmatranog primjera su r , t i n .
- ograničenja koja se javljaju za određene promjenljive i određuju oblast u kojima su rješenja važeća.

Mada funkcija cilja i ograničenja mogu biti predstavljeni preko prostih jednačina one mogu biti osnova za kompleksne zavisnosti i modele što je prikazano u navedenom primjeru. Tako su određivanje vremena padanja tereta i brzina udara zahtijevali komplikovana računanja, dok primjena optimizacionih metoda je mnogo efikasnija i korisnija.

2.3. Matematička osnova problema

Postoje mnogi matematički koncepti i operacije koje čine osnovu optimizacije. To su kvadratne matrice, Jacobijan i Hesse-ova matrica, koveksne i konkavne matrice itd.

Optimizacija ili problemi matematičkog programiranja generalno se mogu postaviti na sljedeći način :

- naći x tako da $f(x)$ bude minimum ili maksimum pri ograničenju

$$\begin{aligned} d_i(x) &\leq a_i & i = 1, \dots, m \\ e_i(x) &= b_i & i = 1, \dots, p \end{aligned} \quad (2.2)$$

gdje je x n -dimenzioni vektor, $f(x)$ - funkcija cilja, $d_i(x)$ – ograničenja u vidu nejednačina i $e_i(x)$ – ograničenja oblika jednačina, a_i i b_i su konstante.

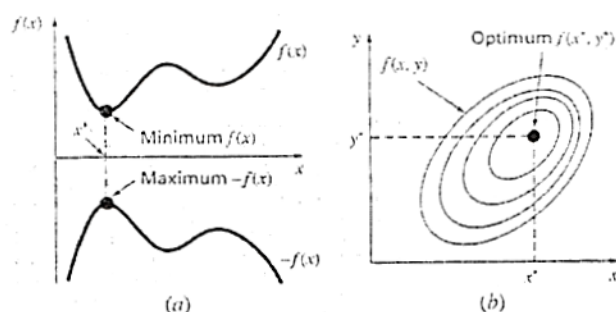
Problemi optimizacije u zavisnosti od oblika funkcije cilja mogu biti :

- ako su $f(x)$ i ograničenja linearni i imamo linearno programiranje
- ako je $f(x)$ kvadratna funkcija i ograničenja linearna, onda je kvadratno programiranje
- ako je $f(x)$ nije linearno ili kvadratno i (ili) ograničenja su nelinearna imamo nelinearno programiranje.

Dalje, ako su uključena ograničenja (2.2) onda je optimizacija sa ograničenjem, a ako ograničenja nijesu uključena onda je optimizacija bez ograničenja. Za ograničene probleme broj stepeni slobode dat je sa $n-p$. Generalno, za dobijanje rješenja mora biti $p \leq n$. Za $p > n$ problem je preograničen.

Drugi način klasifikovanja problema optimizacije je po dimenzionalnosti. Postoje jednodimenzionalni i višedimenzionalni problemi. Kako im ime kaže jednodimenzionalne metode za funkciju cilja imaju funkciju samo jedne promjenljive pa se maksimum i minimum nalaze samo u pravcu jedne promjenljive. Kod višedimenzionih problema funkcija cilja je funkcija dvije ili više promjenljivih i tu vizueliziramo "bregove" i "doline" i tražimo optimalno rješenje.

Međutim, kao i kod običnog pješaćenja, ne ograničavamo se na kretanje strogo u jednom pravcu već ispitujemo topografije i efikasno stižemo do cilja. Na kraju, proces traženja maksimuma i minimuma je identičan, jer se ispituju funkcije $f(x)$ ili $-f(x)$. Ova ekvivalentnost je ilustrovana graficki za jednodimenzionalan problem funkcijom kao na Sl. 2.5.



Sl. 2.5. – Funkcije $f(x)$ i $-f(x)$

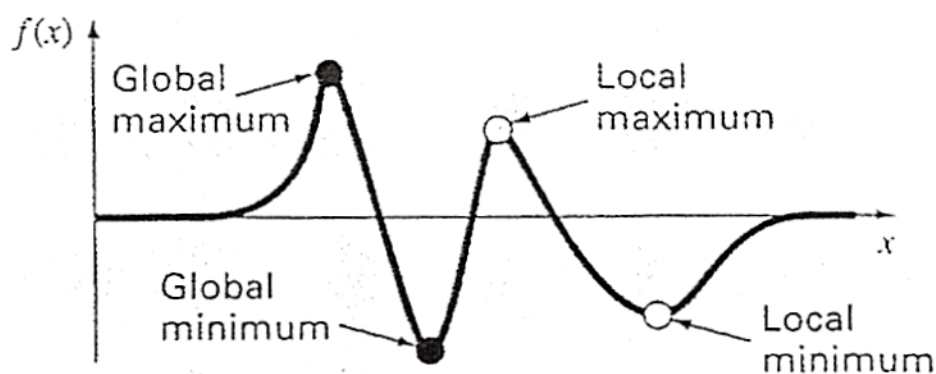
Sljedeće poglavlje, poslije uvodnog dijela je jednodimenziona neograničena optimizacija. Prezentovane su metode za nalaženje maksimuma i minimuma funkcije sa jednom promjenljivom. Pokazane su tri metode :

- metod zlatnog presjeka
- Njutnova metoda
- metod kvadratne interpolacije

Ove metode su potrebne i za višedimenzionu optimizaciju. Dalje, obrađeno je poglavlje u kome se rješavaju dva tipa problema i neograničene multidimenzione optimizacije. Prvi tip problema rješava se bez traženja izvoda funkcije. Poznate metode u ovoj grupi su Powel-ova i Hook-Jeeves-ova metoda. Drugoj grupi pripadaju gradijentne metode. Detaljno je obrađena metoda najbržeg porasta odnosno spuštanja . Metode koje pripadaju ovoj grupi su Njutnova metoda, Gauss-Njutnova, Njutn-Rapsonova itd.

3. JEDNODIMENZIONALNA OPTIMIZACIJA

U ovom dijelu su opisane tehnike za nalaženje maksimuma i minimuma funkcije jedne promjenljive. Postoje tri načina da se ovaj problem ublaži. Na Sl 3.1. je prikazana funkcija jedne promjenljive, kao i lokalni i globalni maksimum i minimum. Kao što se sa slike vidi dvije tačke lijevo odnose se na globalni, a dvije tačke desno na lokalni maksimum i minimum. Razlikovati globalni od lokalnih minimuma može biti težak zadatak u nekim situacijama. Promjena polazne tačke i praćenje da li usvojena metoda dovodi do boljeg rješenja ili vodi u istu tačku. Mada sva tri načina mogu biti korisni, ne postoji siguran put kako odrediti globalni optimum. Međutim, iako treba biti oprezan u odnosu na ovo pitanje, na sreću postoje kod brojnih inženjerskih problema globalni optimum se može odrediti na nedvosmislen način.



Sl 3.1. - Funkcija jedne promjenljive sa optimalnim tačkama

Jednodimenzionalne metode mogu se podijeliti na metode koje upoređuju vrijednosti funkcije i na taj način ocjenjuju interval u kome se nalazi optimalno rješenje i na druge metode koje koriste izvode funkcije i kod kojih se optimalno rješenje traži rješavanjem jednačine $f'(x)=0$.

Da bi odredili efikasnost numeričkih metoda kod nelinearnog programiranja poželjno je za svaku metodu znati stepen konvergencije. Što je stepen konvergencije veći to je metoda efikasnija (brža). Pretpostavimo za funkciju koja zavisi od n promjenljivih neka metoda daje niz vektora x^k koji teže ka optimalnom rješenju x^* ,

$$\|x\| = \sqrt{\sum_{i=1}^u x_i^2}$$

Ovo je Euklidova norma (dužina) vektora x . Ako postoji broj p i neko $\alpha \neq 0$ tako da je

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^p} = \alpha,$$

tada ćemo p zvati stepenom konvergencije niza x^k , dok je $\|x^k - x^*\|$ greška k -te aproksimacije. Ako je $p = 1$ tada je stepen konvergencije linearan, a ako je $p > 1$ stepen konvergencije je superlinearan. Specijalno za $p = 2$ stepen konvergencije je kvadratni. Za neke metode i klase funkcija stepen konvergencije biće uvijek linearan, za druge superlinearan.

3.1. Metode upoređivanja vrijednosti funkcije

Kod ovih metoda upoređenjem vrijednosti funkcije ocjenjuje se interval u kome se nalazi optimalno rješenje. Što se više vrijednosti uporedi informacija o optimalnom rješenju biće bolja i interval će biti manji. Kod ovih metoda dužina intervala zavisi samo od broja upoređivanja, a ne od same funkcije. Ove metode su tipično jednodimenzione i ne mogu se proširiti sa više promjenljivih. Kod ovih metoda važno je samo da funkcija bude unimodalna. Najbolja metoda ovog tipa je Fiboccijeva. Kao pojednostavljenje Fibonccijeve metode dobijamo *Metodu zlatnog presjeka*. Ovo su spore, ali pouzdane metode za lociranje tačke optimalnog rješenja.

Za derivabilne funkcije optimalno rješenje se traži veoma efikasno rješavanjem jednačine $f'(x)=0$. Metoda se odlikuje brзом konvergencijom. Za povoljno "glatke" funkcije (funkcije sa nekoliko neprekidnih izvoda) koristi se *Njutnova metoda*. Pojedostavljena *Njutnova metoda* je metoda sječice. Veoma popularne metode su metode aproksimacije polinomom kao što su *metoda parabole* i *kubna metoda*.

3.2. Metod zlatnog reza

Kada tražimo rješenje nelinearne jednačine, tada tražimo vrijednost promjenljive x tako da je $f(x)=0$. Rješavanjem optimizacionih problema ima za cilj nalaženje promjenljive x u kojoj funkcija $f(x)$ ima maksimum ili minimum.

Metod zlatnog presjeka je jednodimenziona metoda kojom se optimalno rješenje funkcije dobija upoređujući vrijednosti funkcije u pojedinim tačkama. Interesatno svojstvo ove metode je da dužina konačnog intervala u kome se nalazi x_{opt} ne zavisi od izbora

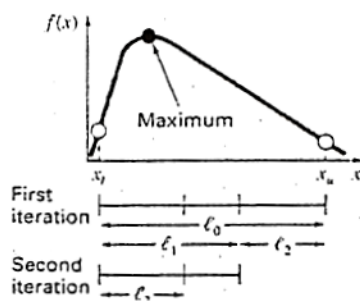
funkcije f . Za bilo koju unimodalnu funkciju f , poslije n upoređivanja vrijednosti funkcije, dobijamo istu ocjenu za dužinu intervala koji sadrži optimalnu tačku (x_{opt}). Dobijanje tačke x_{opt} metodom zlatnog presjeka u biti je slično sa metodama određivanja nule funkcije $f(x)$ koje se baziraju na polovljenju intervala u kome se nalazi rješenje jednačine.

Naime, nađe se donja (x_d) i gornja (x_g) tačka nekog intervala promjenljive x u kome se nalazi rješenje x_0 i nađe polovljenjem intervala srednja tačka

$$x_s = \frac{x_d + x_g}{2}$$

Tada se izračuna vrijednost funkcije u tim tačkama nađe promjena znaka funkcije. Na osnovu toga se zaključuje u kojem se intervalu od dva nova (x_d, x_s) ili (x_s, x_g) nalazi traženo rješenje. Zatim se odbaci jedna od tačaka (x_d ili x_g) i traženje nastavlja u jednom od ovih intervala ponavljajući postupak polovljenja intervala i upoređivanja vrijednosti funkcije sve dok se ne dobije zadovoljavajuće rješenje.

Sada ćemo razviti proceduru za nalaženje optimuma za funkciju jedne promjenljive. Zbog jednostavnosti fokusiraćemo našu pažnju na nalaženje maksimuma. Kao kod polovljenja intervala pri određivanju nule funkcije, startovaćemo sa definisanjem intervala koji sadrži rješenje, odnosno maksimum funkcije. Zadržaćemo oznake intervala x_d i x_g kao i kod polovljenja intervala u prethodnom primjeru. Međutim, razlika sa navedenim primjerom polovljenja intervala je u tome što sada razvijamo novu strategiju nalaženja maksimuma u okviru intervala, kao na Sl 3.2. Koristićemo vrijednosti funkcije u tri tačke i uvođenjem četvte tačke provjeravamo gdje se nalazi maksimum da li u okviru prve tri ili zadnje tri tačke.



Sl 3.2. –Početne iteracije metode zlatnog reza

Posebno važno je izabrati srednju tačku. Ona treba da zadovolji dva uslova :

$$l_0 = l_1 + l_2$$

$$\frac{l_1}{l_0} = \frac{l_2}{l_1}$$

Iz ovih jednačina dobija se

$$\frac{l_1}{l_1 + l_2} = \frac{l_2}{l_1}$$

Ako uvedemo oznaku $R = \frac{l_2}{l_1}$ kao odnos ovih dužina dobićemo :

$$1 + R = \frac{1}{R}$$

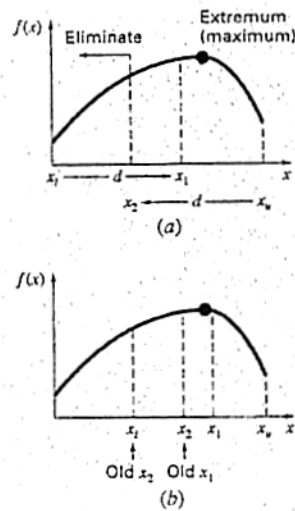
ili

$$R^2 + R - 1 = 0$$

Rjesenje ove jednacine

$$R = \frac{-1 + \sqrt{1 - 4(-1)}}{2} = \frac{-1 + \sqrt{5}}{2} = 0.61803\dots$$

Ova vrijednost je bila poznata prije antičke grčke i naziva se zlatni presjek ili zlatni rez. Ovo je ključni element kod metode zlatnog presjeka na kome se metod konceptualno zasniva i omogućava da se optimalna vrijednost nalazi efikasno. Kao sto je prikazano na Sl 3.3, u metodi se startuje sa dvije početne vrijednosti x_d i x_g koje okružuju lokalnu ekstremnu vrijednost funkcije $f(x)$.



Sl 3.3 Metod zlatnog reza

Zatim, biraju se dvije unutrašnje tacke x_1 i x_2 u skladu sa zlatnim presjekom:

$$d = R(x_g - x_d)$$

$$x_1 = x_d + d$$

$$x_2 = x_g - d$$

Pošto se nadu vrijednosti funkcije u ovim tačkama mogu nastupiti sljedeći slučajevi :

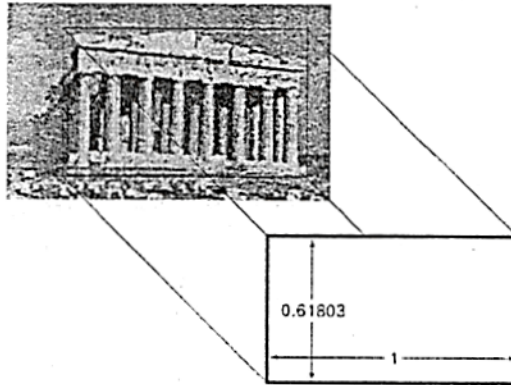
- ako je $f(x_1) > f(x_2)$, tada interval lijevo od x_2 nije biran jer ne sadrži maksimum funkcije i treba ga eliminisati
- ako je $f(x_2) > f(x_1)$, tada desno od tačke x_1 neće biti maksimum funkcije i taj interval treba eliminisati.

Na ovaj način se veoma efikasno dolazi do rješenja i interval koji sadrži maksimum se rapidno smanjuje. Metoda zlatnog presjeka je pouzdana. Njen nedostatak je što je prilično spora i njen stepen konvergencije je u najboljem slučaju linearan.

U mnogim kulturama, određenim brojevima su pripisivana neka svojstva. Tako na primjer na Zapadu su bili bliski sa "srećnim 7" ili "petak 13". Antički Grci su, navedeni broj R, zvali zlatni presjek

$$R = \frac{\sqrt{5}-1}{2} = 0,61803\dots$$

Ovaj broj je bio korišćen za mnoge svrhe uključujući i projektovanje građevina, kao sto je prikazano na Sl 3.4.



Sl 3.4. – Partenon u Atini projektovan u skladu sa „zlatnim“ pravougaonikom

Ovaj broj je važan i kod niza tzv. "Fibonaccijevi brojevi". Kod ovog niza je svaki član zbir prethodna dva, a količnik susjednih brojeva teži zlatnom odnosu 0,618. Članovi niza redjaju se na sljedeći način:

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89,....

U grupu metoda upoređivanja vrijednosti funkcije važnu ulogu ima i Fibonaccijeva metoda koja u poređenju sa metodom zlatnog presjeka je bolja, jer daje manji interval kao rješenje optimizacije. Međutim, nedostatak ove metode sastoji se u tome što treba prvo odrediti broj iteracija n koji garantuje da optimalna tačka x leži unutar željenog intervala, pa tek onda traže tačke unutar intervala. Znači, da za primjenu ove metode treba unaprijed znati broj koraka.

Za razliku od metoda upoređivanja vrijednosti funkcije, postoje metode koje zahtijevaju da funkcija f bude derivabilna. Vrijednost funkcije f i njenih derivacija u nekoj aproksimaciji x^* optimalnog rješenja koriste se za određivanje bolje aproksimacije x_{k+1} . Metode ovog tipa su Njutnova metoda, metoda sječice. To su klasične metode poznate u nauci za određivanje korijena jednačine $f(x) = 0$. U slučaju traženja optimalnih rješenja one se koriste za izračunavanje korijena jednačine $f'(x)=0$ tj. na izračunavanju stacionarnih tačaka.

Metode aproksimacije polinomom čine posebnu grupu metoda za određivanje optimalnog rješenja funkcije f . Polazi se od ideje da se funkcija f aproksimira s nekim polinomom na nekom intervalu koji sadrži optimalno rješenje. Metode toga tipa zavise od stepena polinoma kojim se aproksimira funkcija.

Primjer 3.1:

Koristeći metodu zlatnog reza naći maximum funkcije

$$f(x) = 2 \sin x - \frac{x^2}{10}$$

u intervalu $x_1=0$ i $x_u = 4$.

Rešenje:

Prvo, u metodi zlatnog reza se kreiraju dvije unutrašnje vrijednosti x_1 i x_2 .

$$d = \frac{\sqrt{5}-1}{2}(4-0) = 2.472$$

$$x_1 = 0 + 2.472 = 2.472$$

$$x_2 = 4 - 2.472 = 1.528$$

Vrijednosti funkcije u unutrašnjim tačkama x_1 i x_2 su:

$$f(x_2) = f(1.528) = \frac{1.528^2}{10} - 2 \sin(1.528) = 1.765$$

$$f(x_1) = f(2.472) = 0.63$$

Sobzirom da je $f(x_2) > f(x_1)$ maximum je u intervalu definisanom sa x_1 , x_2 i x_u . Tako, za novi interval brojeva granica je $x_1=0$ a x_1 postaje gornja granica, dakle $x_u=2.472$. Tada ranije x_2 postaje novo x_1 i $x_1=1.528$. Sada nećemo ponovo računati $f(x_1)$ jer je ta vrijednost već nađena u prethodnoj iteraciji $f(1.528)=1.765$.

Ono što ostaje da se izračuna je novoj iteraciji jeste vrijednost d i x_2 .

$$d = \frac{\sqrt{5}-1}{2}(2.472-0) = 1.528$$

$$x_2 = 4 - 1.528 = 0.944$$

Vrijednost funkcije u x_2 je $f(0.944)=1.531$. Ova vrijednost je manja od vrijednosti funkcije u tački x_1 i maximum se nalazi negdje između tačaka x_2 , x_1 i x_u . Ponavljajući proceduru, dobije se tabela sa vrijednostima:

i	x_1	$f(x_1)$	x_2	$f(x_2)$	x_1	$f(x_1)$	x_u	$f(x_u)$	d
1	0	0	1.5279	1.7674	2.4721	0.6300	4.000	-3.1136	2.4721
2	0	0	0.9443	1.5310	1.5279	1.7674	2.4721	0.6300	1.5279
3	0.9443	1.5310	1.5279	1.7674	1.8885	1.5423	2.4721	0.6300	0.9443
4	0.9443	1.5310	1.3050	1.7595	1.5279	1.7674	1.8885	1.5432	0.5836
5	1.3050	1.7595	1.5279	1.7674	1.6656	1.7136	1.8885	1.5432	0.3607
6	1.3050	1.7595	1.4427	1.7755	1.5279	1.7647	1.6656	1.7136	0.2229
7	1.3050	1.7595	1.3910	1.7742	1.4427	1.7755	1.5279	1.7647	0.1378
8	1.3050	1.7742	1.4427	1.7755	1.4752	1.7732	1.5279	1.7647	0.0851

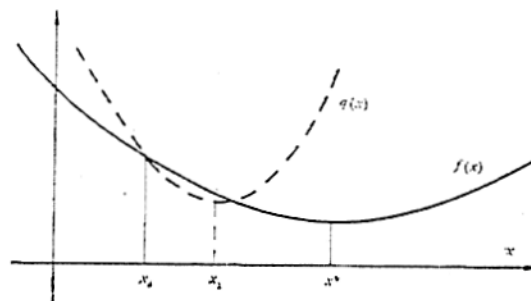
Osjenčene vrijednosti predstavljaju maximume za svaku iteraciju. Poslije osme iteracije maximum se javlja u tački $x = 1.4427$ sa vrijednošću funkcije u njoj 1.7755. Tako rezultat konvergira vrijednosti 1.7757 u tački $x = 1.4276$.

3.3. Njutnova metoda

Ideja Njutnove metode je jednostavna i sastoji se u tome da se vrijednost funkcije f u obliku x_0 može aproksimirati sa prva tri člana Tajlorovog reda kao :

$$f(x) \approx q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

gdje je x_0 početna aproksimacija stacionarne tačke x^* . Funkcija $q(x)$ je parabola, a njena ekstremna tačka x_1 , uzima se za novu aproksimaciju tačke x^* , kao što je prikazano na Sl 3.5.



Sl 3.5. Ideja Njutnove metode

Vrijednost x_1 , s obzirom da je tačka minimuma parabole $q(x)$ nalazi se iz uslova $q'(x) = 0$, pa je

$$q'(x) = f'(x_0) + f''(x_0)(x_1 - x_0) = 0$$

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)}$$

Ponavljajući postupak dobija se rekurzivna formula :

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \quad (*)$$

za računanje stacionarne tačke funkcije f koja se zove " Njutnova metoda ". Njutnova metoda je, za dovoljno " glatke " funkcije, veoma efikasna i stepen konvergencije je kvadratni za dobro pogođene početne aproksimacije x_0 .

Primjer 3.2:

Koristeći Njutnovu metodu naći maximum funkcije

$$f(x) = 2 \sin x - \frac{x^2}{10}$$

sa početnom vrijednošću $x_0 = 2.5$.

Rešenje: Prvi i drugi izvod funkcije mogu se prikazati kao:

$$f'(x) = 2 \cos x - \frac{x}{5}$$

$$f''(x) = -2 \sin x - \frac{x}{10}$$

može biti zamijenjen u Eq i daje:

$$x_{i+1} = x_i - \frac{2 \cos x_i - x_i / 5}{-2 \sin x_i - 1/5}$$

Zamjena početnih vrijednosti priraštaja je:

$$x_1 = 2.5 - \frac{2 \cos 2.5 - 2.5/5}{-2 \sin 2.5 - 1/5} = 0.99508$$

a vrijednost funkcije u njoj je 1.57859. Druga iteracija daje:

$$x_1 = 0.995 - \frac{2 \cos 0.995 - 0.995/5}{-2 \sin 0.995 - 1/5} = 1.46901$$

a vrijednost funkcije u njoj je 1.77385.

Proces se ponavlja, a rezultati su prikazani u tabeli:

i	x	f(x)	f'(x)	f''(x)
0	2.5	0.57194	-2.10229	-1.39694
1	0.99508	1.57859	0.88985	-1.87761
2	1.46901	1.77385	-0.09058	-2.18965
3	1.42764	1.77573	-0.00020	-2.17954
4	1.42755	1.77573	0.00000	-2.17952

Tako poslije četiri iteracije rezultat konvergira brzo na tačnu vrijednost.

3. 4. Metoda sječice

Kao što se vidi u relaciji (*) Njutnova metoda se zasniva na nalaženju prvog i drugog izvoda u svakoj aproksimaciji x_k . Ukoliko se drugi izvod funkcije f zamijeni sa konačnim razlikama

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

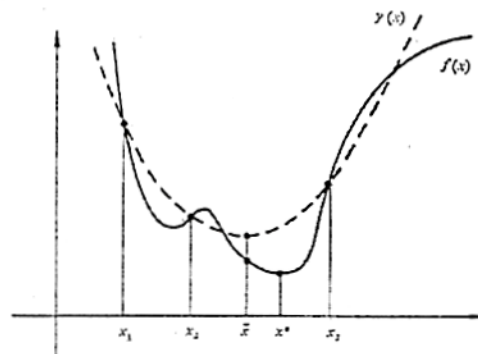
tada Njutnova metoda dobija oblik

$$x_{k+1} = x_k - \frac{f'(x_k)(x_k - x_{k-1})}{f'(x_k) - f'(x_{k-1})}$$

i računanje stacionarne tačke funkcije f na ovaj način naziva se metoda sječice. Tačke x_0 i x_1 su proizvoljne. Prednost metode sječice u odnosu na Njutnovu metodu je u tome što se ne izračunava drugi izvod funkcije. Međutim, metoda sječice je sporija od Njutnove metode, a brža od metode upoređivanja vrijednosti funkcija. Koeficijent konvergencije metode sječice je superlinearan.

3.5 Metode aproksimacije polinomom

Ideja ove metode bazira se na činjenici da se funkcija f , čija se optimalna tačka x^* traži, dobro aproksimira polinomom na intervalu koji sadrži tačku x^* . Kao što se kroz dvije tačke može provući samo jedna prava, tako se i kroz tri tačke može provući samo jedna parabola. Za neke uočene tri tačke x_0, x_1, x_2 kao na Sl.3.6., u intervalu koji sadrži minimum funkcije treba naći vrijednosti funkcija u njima $f(x_0), f(x_1)$ i $f(x_2)$.

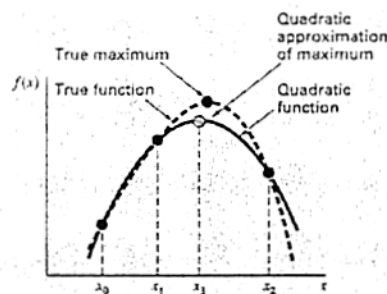


Sl. 3.6. Kvadratna interpolacija za traženje minimuma funkcije

Tada se kroz te tačke provlači parabola $y = a + bx + cx^2$ čija se optimalna tačka nalazi iz uslova $y'(x) = b + 2cx = 0$, odnosno rješavanjem sistema jednačina dobija se vrijednost optimalne tačke x_3 parabole $y(x)$ koja je aproksimacija optimalne tačke x^* :

$$x_3 = \frac{f(x_0)(x_1^2 - x_2^2) + f(x_1)(x_2^2 - x_0^2) + f(x_2)(x_0^2 - x_1^2)}{2f(x_0)(x_1 - x_2) + 2f(x_1)(x_2 - x_0) + 2f(x_2)(x_0 - x_1)} \quad (**)$$

Sada postoje četiri tačke, a za novu kvadratnu aproksimaciju dovoljne su tri. Dalje se, jednostavnom strategijom sličnom metodi zlatnog presjeka, zaključuje da je jedna od tačaka suvišna i postupak se nastavlja sa tri tačke sve dok se ne dobije zadovoljavajuća tačnost aproksimacije. Za funkciju koja ima maksimum grafički prikaz dat je na Sl.3.7.



Sl.3.7. Kvadratna interpolacija za funkciju koja ima maksimum

Ako se aproksimacija vrši sa polinomom trećeg reda tada imamo *Kubnu metodu*. U metodi parabole, za razliku od Njutnove metode ne koristimo izvode funkcije f , pa se ona može koristiti za iznalaženje optimalnih tačaka nederivabilnih funkcija. Stepennost konvergencije obje metode (Metode parabole i Kubne metode) je superlinearna.

Primjer 3.3:

Koristeći kvadratnu interpolaciju naći maximum funkcije:

$$f(x) = 2 \sin x - \frac{x^2}{10}$$

sa početnim vrijednostima $x_0 = 0$, $x_1 = 1$ i $x_2 = 4$.

Rešenje: Vrijednost funkcije u tri proizvoljne tačke je:

$$\begin{array}{ll} x_0 = 0 & f(x_0) = 0 \\ x_1 = 1 & f(x_1) = 1.5829 \\ x_2 = 4 & f(x_2) = -3.1136 \end{array}$$

i koristeći jednačinu (**) dobija se:

$$x_3 = \frac{0(1^2 - 4^2) + 1.5829(4^2 - 0^2) + (-3.1136)(0^2 - 1^2)}{2(0)(1-4) + 2(1.5829)(4-0) + 2(-3.1136)(0-1)} = 1.5055$$

a vrijednost funkcije u tački x_3 je $f(1.5055) = 1.7691$.

Strategija metode je slična metodi zlatnog reza u dijelu da treba naći novu tačku u procesu približavanja optimalnoj vrijednosti, a odbaciti jednu od nacrtanih tačaka u svakoj iteraciji.

Kako je vrijednost funkcije u novoj tački veća od vrijednosti funkcije u tački x_1 to se za donju granicu bira tačka x_1 , a x_0 se odbacuje. Zato u sljedećoj iteraciji uzimamo vrijednosti:

$$\begin{array}{ll} x_0 = 1 & f(x_0) = 1.5829 \\ x_1 = 1.5055 & f(x_1) = 1.7691 \\ x_2 = 4 & f(x_2) = -3.1136 \end{array}$$

i koristeći jednačinu (**) dobija se x_3 .

$$x_3 = \frac{1.5829(1.5055^2 - 4^2) + 1.7691(4^2 - 1^2) + (-3.1136)(1^2 - 1.5055^2)}{2(1.5829)(1.5055-4) + 2(1.7691)(4-1) + 2(-3.1136)(1-1.5055)} = 1.4903$$

a vrijednost funkcije u njoj $f(1.4903) = 1.7714$.

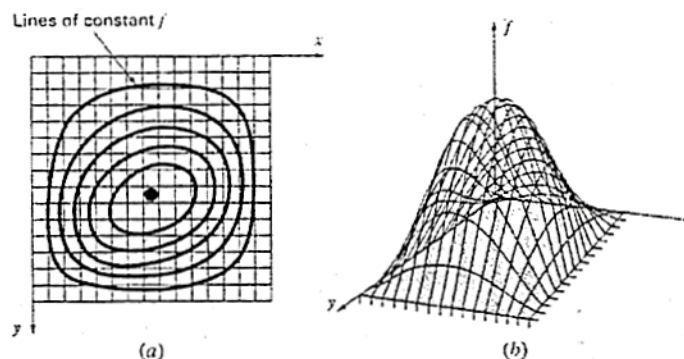
Proces se ponavlja, a rezultati su prikazani u tabeli:

i	x_0	$f(x_0)$	x_1	$f(x_1)$	x_2	$f(x_2)$	x_3	$f(x_3)$
1	0.0000	0.0000	1.0000	1.5829	4.000	3.1136	1.5055	1.7691
2	1.0000	1.5829	1.5055	1.7691	4.000	3.1136	1.4903	1.7714
3	1.0000	1.5829	1.4903	1.7714	1.5055	1.7691	1.4256	1.7757
4	1.0000	1.5829	1.4256	1.7757	1.4903	1.7714	1.4266	1.7757
5	1.4256	1.7757	1.4266	1.7757	1.4903	1.7714	1.4275	1.7757

Tako poslije pet iteracija rezultati konvergiraju vrijednosti funkcije 1.7757 u tački $x = 1.4276$.

4. VISEDIMENZIONALNA NEOGRANICENA OPTIMIZACIJA BEZ IZRACUNAVANJA DERIVACIJA

U ovom dijelu se analiziraju metode za nalaženje optimalnih rješenja funkcija više promjenjivih. Kod jednodimenzionalnih problema imaginacija je kretanje u jednom pravcu, kod dvodimenzionalnih slučajeva to su npr. planine i uvale (Sl 4.1), a kod višedimenzionalnih nije moguća odgovarajuća imaginacija.



Sl 4.1. Funkcija više promjenjivih i njihove linije konstantnih vrijednosti

Ograničićemo razmatranje na slučajeve sa dvije promjenljive zato što se često pokazuje da se ovo razmatranje najbolje prati vizuelno.

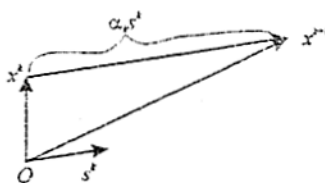
Metode kod višedimenzionalne neograničene optimizacije mogu se klasifikovati na više načina. U ovom slučaju podjelu ćemo izvršiti u zavisnosti od toga da li se u metodama traže izvodi funkcije ili ne.

Prvu klasu čine negradientne ili direktne metode kod kojih ne treba tražiti parcijalne izvode funkcije cilja po parametrima i gradientne metode za čiju promjenu je potrebno da je funkcija derivabilna. Strategija traženja minimuma treba da je takva da se sa što manjim brojem izračunavanja ciljne funkcije odredi položaj tačke u n-dimenzionalnom prostoru za koji ciljna funkcija poprima najmanju vrijednost.

Druga klasa metoda primjenjiva je samo na diferencijabilne funkcije (dok je prva klasa primjenjiva i u nediferencijabilnom slučaju). Metode bezuslovne optimizacije bez izračunavanja izvoda primjenjuju se i u slučajevima kada je funkcija diferencijabilna, ali je izračunavanje izvoda komplikovano, ili nemoguće (ukoliko se radi o rezultatima mjerenja, tj. analitički izraz nije poznat). Metode bezuslovne optimizacije su po pravilu iterativnog tipa. Njima se nalazi niz tačaka $\{x^k\} \in \mathbb{R}^n$ korišćenjem relacije:

$$x^{k+1} = x^k + a_k \cdot s^k, \quad k = 0, 1, 2, \dots$$

Sljedeća vrijednost x^{k+1} u nizu $\{x^k\}$ dobija se sabiranjem prethodne vrijednosti sa proizvodom $a_k s^k$, gdje je $s^k \in \mathbb{R}^n$ vektor koji određuje pravac kretanja iz tačke x^k , a $a_k > 0$ broj koji određuje dužinu koraka u pravcu s^k .



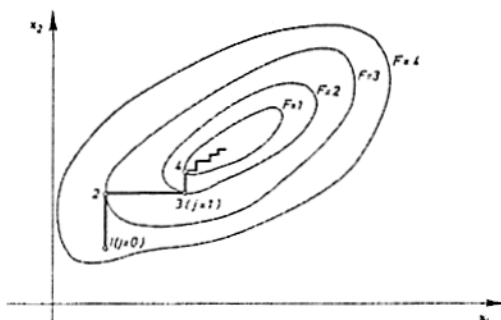
Sl 4.2 Generisanje niza $\{x^k\}$

Potrebno je da niz pravaca kretanja $\{s^k\}$ i niz koraka $\{a^k\}$ obezbijedi da niz vrijednosti funkcije $\{f(x^n)\}$ monotono opada i konvergira vrijednosti u nekom od lokalnih minimuma problema.

Od direktnih metoda ili metoda bezuslovne optimizacije bez izračunavanja derijacija opisaćemo tri metode: traženje minimuma po koordinatnim osama, postupak po Hook-u i Jeeves-u i Powell-ovu metodu.

4.1. Metoda traženja po koordinatnim osama

Minimum po ovoj metodi, traži uzastopno po koordinatnim osama jednim od postupaka za traženje po pravcu. Za dvodimenzionalan problem geometrijska ilustracija je prikazana na Sl 4.3.



Sl 4.3. Traženje minimuma po koordinatnim osama

Jedna iteracija u dijagramu toka kod ove metode sastoji se od traženja po svim koordinatnim smjerovima. Postupak se zaustavlja ako je zadovoljen kriterijum kojim se zahtijeva da apsolutna vrijednost promjene parametara u dvije uzastopne iteracije bude manja od unaprijed zadatog malog broja ε , tako da bude:

$$|x^{k+1} - x^k| \leq \varepsilon, \quad \text{gdje je } k \text{ broj iteracije}$$

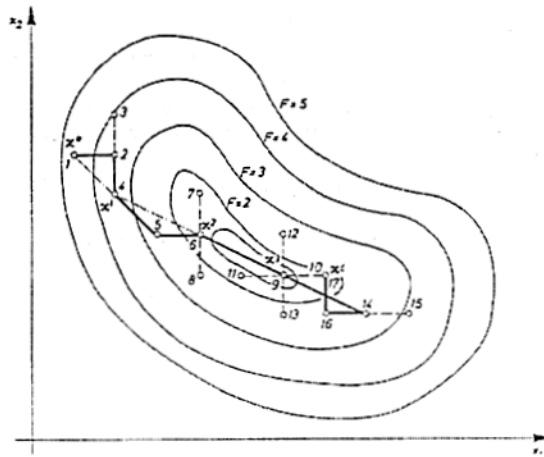
Ovaj postupak je veoma jednostavan ali neefikasan jer ne uzima u obzir nikakvo saznanje, o funkciji cilja iz prethodne iteracije. Osnovni nedostatak metode je da je kretanje prema minimumu vrlo sporo, tj. u malim koracima, ako postoji "uska dolina" u smjeru koji se ne poklapa sa jednim od koordinatnih smjerova što se može vidjeti i na slici 4.3.

4.2. Hooke – Jeeves-ov metod

Ova metoda je znatno poboljšanje prethodnog postupka, naročito u traženju minimuma uzduž uske doline. Osnovna ideja metode jeste naći pravac uzduž takve doline. Polazi se od tačke x_0 i prave se probni koraci dužine a_k u pravcima koordinatnih osa i kreće se tamo gdje ima poboljšanja, a ako to nije moguće korak se polovi.

Svaka iteracija se sastoji od dva dijela: istraživanja i kretanja u odabranom smjeru. Poboljšanjem se smatra tačka u kojoj je vrijednost funkcije manja od vrijednosti funkcije u

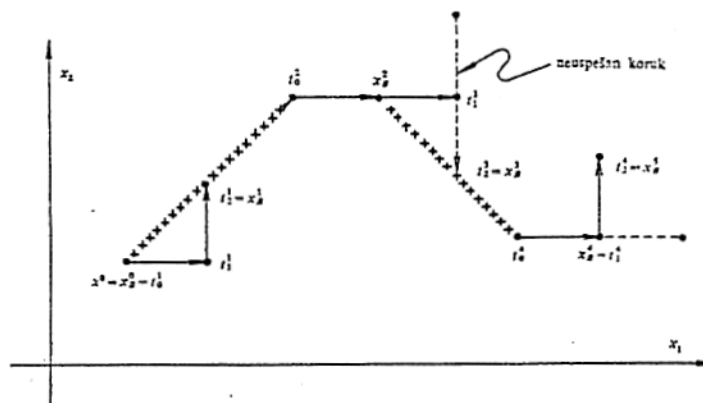
prethodnoj bazičnoj tački. Kada se dobije nova bazična tačka x_1 , tada se kroz te dvije bazične tačke dobija pravac po kojem treba ići za dužinu koja je jednaka udaljenosti tih dviju tačaka, kao na slici 4.4.



Sl 4.4. Postupak po Hook-u i Jeeves-u

Tako dolazimo do nove bazične tačke x_2 . Znači, radi se o iterativnom postupku u kome niz bazičnih tačaka x_0, x_1, x_2, \dots teži prema nekom lokalnom minimumu x^* funkcije f . U svakoj bazičnoj tački pretražuju se koordinatne ose poslije čega nalazimo novu bazičnu tačku x_{k+1} . Tada se pravi skok u pravcu određenom posljednjim dvijema bazičnim tačkama i dobija se nova bazična tačka iz koje se vrši ispitivanje.

Ako nove vrijednosti funkcija ne zadovoljavaju kriterijum o minimumu tada se korak smanjuje. Ako ni tada ne dobijemo funkciju manju od vrijednosti funkcije u bazičnoj tački onda se radi o aproksimaciji lokalnog minimuma.



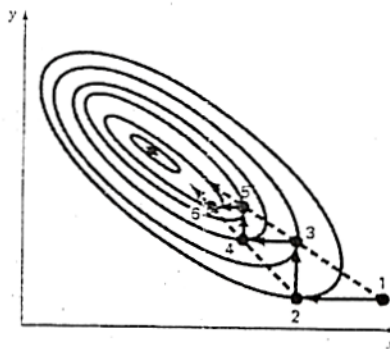
Sl 4.5. Ilustracija Hooke-Jeevesove metode

Na slici Sl 4.5. je data ilustracija Hook-Jeeves-ove metode primijenjena na funkciju dvije promjenjive. Punom linijom je prikazan uspješan korak, a isprekidanom neuspješan korak. Tačkastom linijom predstavljeni su uspješni skokovi u pravcu određenom bazičnim tačkama.

Metoda Hooke-Jeeves-a je jednostavna i pouzdana i lako se programira na računaru. Međutim njen nedostatak je što sporo konvergira optimalnom rješenju.

Može se reći da su navedene metode direktnog traženja minimuma su spore. Zadatak je naći efikasnije prilaze optimalnom rješenju bez traženja parcijalnih izvoda funkcije. Osnovna ideja ovakvih metoda bila bi varirati jednu promjenjivu, dok su druge konstantne. Time bi se problem sveo na traženje minimuma koristeći Metode jednodimenzionalne optimizacije.

Grafički prikaz za funkciju dvije promjenjive dat je na Sl 4.6. Kao što se sa slike vidi polazi se od tačke 1, traži minimum u pravcu promjenjive x, dok je ordinatay konstantna i tako se dolazi do tačke 2. Zatim se uzima x konstantno, a nekom od jednodimenzionalnih metoda nalazi se minimum u pravcu y i dobija se tačka 3. Postupak se nastavlja i redaju se tačke 4, 5, 6 i td cime se tako se ide ka optimalnom rješenju.



Sl 4.6. Pravci koji vode ka optimumu

Međutim, mnogo je efikasnije uočiti pravce određene tačkama 1-3 i 2-4. U njihovom presjeku se nalazi tačka minimuma veoma efikasno. Najbolji algoritam sa ovakvom osnovnom idejom nalaženja minimuma dao je Powell.

Primjer 4.1. Metodom Hookea i Jeevesa odredićemo lokalni minimum funkcije

$$f(x) = f(x_1, x_2, x_3) = x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3$$

Neka je polazna bazična tačka, recimo

$$t_0^1 = x_B^0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

i dužina koraka $\delta = 0,1$. Korak duž pozitivnog dijela prve koordinatne ose vodi do tačke

$$x_B^0 + d_1^0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0,1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,1 \\ 0 \\ 0 \end{bmatrix}.$$

Budući da je $f(x_B^0 + d_1^0) = 0,1^4 + 0,1^3 - 0,1 = -0,0989 < f(x_B^0) = 0$, prvi korak je uspješan i stavljamo

$$t_1^1 = x_B^0 + d_1^0.$$

Sada iz tačke t_1^1 pretražujemo drugu koordinatnu osu: Korak duž pozitivnog dijela druge koordinatne ose vodi do

$$t_1^1 + d_2^0 = \begin{bmatrix} 0,1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0,1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,1 \\ 0,1 \\ 0 \end{bmatrix}.$$

Budući da je $f(t_1^1 + d_2^0) = -0,0089 > f(t_1^1) = -0,0989$, ovaj korak vodi na brdo i zato je neuspješan. Za korak duž negativnog dijela druge koordinatne ose:

$$t_1^1 - d_2^0 = \begin{bmatrix} 0,1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0,1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,1 \\ -0,1 \\ 0 \end{bmatrix}$$

nalazimo da je $f(t_1^1 - d_2^0) = -0,2088 < f(t_1^1)$, ovaj korak vodi u dolinu i zato je uspješan; stavljamo

$$t_2^1 = t_1^1 - d_2^0$$

i iz ove tačke pretražujemo treću koordinatnu osu. Korak u pozitivnom smjeru te ose daje tačku

$$t_2^1 + d_3^0 = \begin{bmatrix} 0 \\ -0,1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0,1 \end{bmatrix} = \begin{bmatrix} 0,1 \\ -0,1 \\ 0,1 \end{bmatrix}$$

Budući da je $f(t_2^1 + d_3^0) = -0,2998 < f(t_2^1)$, ovaj korak vodi u dolinu i zato je uspješan. Stavljamo $t_3^1 = t_2^1 + d_3^0$.

Na isti način iscrpili smo sve koordinatne ose u prvoj iteraciji; zato t_3^1 postaje nova bazična tačka, tj.

$$x_B^1 = t_3^1.$$

Sada dolazi do skoka: Novu bazičnu tačku pokušaćemo pomaknuti u smjeru $x_B^1 - x_B^0$, tj. postaviti u tačku

$$t_0^2 = x_B^1 + (x_B^1 - x_B^0) = \begin{bmatrix} 0,2 \\ -0,2 \\ 0,2 \end{bmatrix}.$$

Budući da je $f(t_0^2) = -0,5968 < -0,2998 = f(x_B^1)$, skok je bio uspješan. (Ovo znači da se dolina nastavlja i opada preko x_B^0 i x_B^1 bar do t_0^2). Sada bismo t_0^2 mogli proglasiti za bazičnu tačku; međutim, uobičajeno je da se bazična tačka određuje poslije prvog uspješnog koraka iz t_0^2 .

Korak iz t_0^2 u smjeru pozitivnog dijela prve ose vodi do

$$t_0^2 + d_1^1 = \begin{bmatrix} 0,2 \\ -0,2 \\ 0,2 \end{bmatrix} + \begin{bmatrix} 0,1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,3 \\ -0,2 \\ 0,2 \end{bmatrix}.$$

Budući da je $f(t_0^2 + d_1^1) = -0,6813 < f(t_0^2)$, zakoračili smo u dolinu i zato je ovaj korak uspješan.

Tačku $t_0^2 + d_1^1$ proglašavamo novom bazičnom tačkom, tj.

$$x_B^2 = t_0^2 + d_1^1$$

i iz x_B^2 nastavljamo sa koracima iz početka: prvo duž prve koordinatne ose, zatim duž druge itd.

Rezultati ovih proračuna dati su u tabeli.

Tabela

k	Komponente bazične tačke x_B^k			Dužina koraka
	x_1	x_2	x_3	δ
0	0	0	0	0,1
1	0,1	-0,1	0,1	0,1
2	0,3	-0,2	0,2	0,1
3	0,4	-0,3	0,3	0,1
4	0,5	-0,5	0,4	0,1
5	0,5	-0,6	0,5	0,1
6	0,5	-0,8	0,6	0,1
7	0,5	-0,9	0,7	0,1
8	0,6	-0,9	0,8	0,1
9	0,55	-0,95	0,75	0,05
10	0,575	-0,950	0,775	0,025
11	0,5750	-0,9376	0,7750	0,0125
12	0,56875	-0,93760	0,76875	0,00625
13	0,571875	-0,940625	0,768750	0,003125
14	0,570312	-0,939062	0,767187	0,0015625
15	0,570703	-0,939062	0,767187	0,00039063
16	0,570703	-0,939257	0,767187	0,00019531
17	0,570703	-0,939257	0,767187	0,00000001

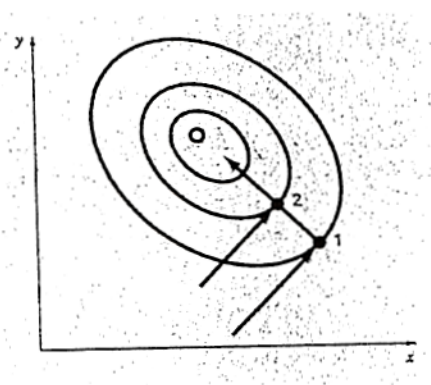
Bazična tačka x_B^{17} daje aproksimaciju lokalnog minimuma

$$x_1^* = 0,570703 \quad x_2^* = -0,939257 \quad x_3^* = 0,767187$$

Eventualno poboljšanje može se postići sa manjom dužinom koraka $\delta < 10^{-8}$.

4.3. Powell-ov metod

Powell-ova metoda sastoji se u efikasnom traženju tačke minimuma preko konjugovanih pravaca. Metoda je bazirana na činjenici da se spajanjem tačaka 1 i 2, (Sl 4.7.), koje su dobijene kao minimumi na pravcima iz različitih startnih pozicija, dobija pravac koji vodi ka minimumu.



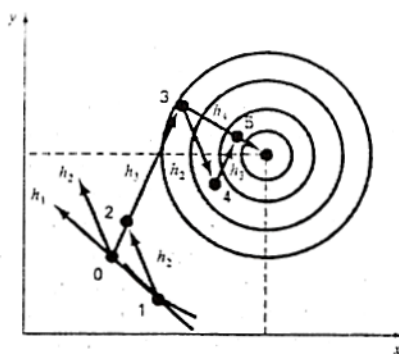
Sl.4.7. Konjugovani pravci

Taj pravac se naziva konjugovani pravac. Ovu metodu je najbolje objasniti preko kvadratne funkcije $f(x,y)$. Svako sljedeće traženje minimuma na pravcu se dobija u konačnom broju koraka bez obzira na startnu tačku. Primjena uprošćene verzije Powell-ove metode za nalaženje minimuma funkcije:

$$f(x,y) = c - (x-a)^2 - (y-b)^2, \text{ gdje su } a, b, c - \text{konstante}$$

prikazana je na Sl 4.8.

Izaberimo početnu tačku O i pravce h_1 i h_2 koji nemoraju biti konjugovani. Iz tačke O idemo pravcem h_1 do minimuma na tom pravcu 1, kao na Sl 4.8. Iz tačke 1 traženjem minimuma na pravcu h_2 dolazimo do tačke 2.

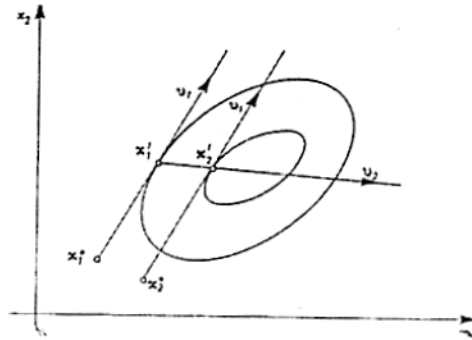


Sl.4.8. Powell-ov metod

Spajanjem tačaka O i 2 dolazimo do pravca h_3 i minimuma na njemu tačke 3. Pravac h_3 je po Powell-u konjugovani pravac, jer se dobija spajanjem tačaka minimuma na dva paralelna pravca. Iz tačke 3 traženjem minimuma na pravcu h_2 dolazi se do tačke 4. Iz tačke 4 se traži minimum na konjugovanom pravcu h_3 i dolazi se do tačke 5. Spajanjem tačaka 3 i 5 dobija se konjugovani pravac h_4 koji nas vodi direktno ka minimumu.

Powell-ova metoda je veoma efikasna i ima kvadratnu konvergenciju. Koristi se za nalaženje optimalnih rješenja kod nederivabilnih funkcija.

Na Sl 4.9. prikazani su konjugovani smjerovi v_1 i v_2 za kvadratnu funkciju dviju promjenljivih. Smjer v_2 je dobijen spajanjem minimuma na dva paralelna pravca smjera v_1 .

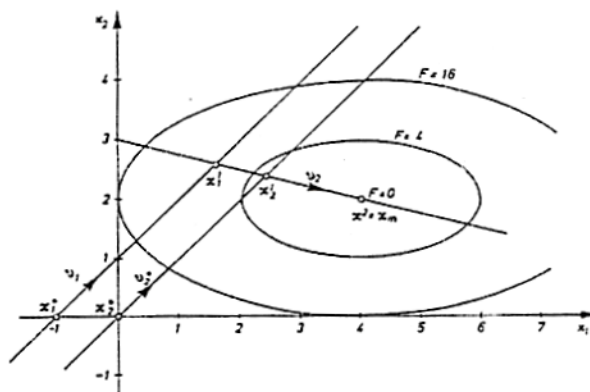


Sl 4.9. Konjugovani smjerovi kvadratne funkcije

Kao primjer za nalaženje minimuma u konjugovanim smjerovima, po metodi Powell-a neka posluži funkcija:

$$f(x_1, x_2) = (x_1 - 4)^2 + (x_2 - 2)^2$$

Krive jednakih vrijednosti funkcije prikazane su na Sl 4.10. Minimum funkcije se nalazi u tacki $x_{\min} = (4, 2)$.



Sl 4.10. Primjer odredjivanja minimuma funkcije $f(x_1, x_2) = (x_1 - 4)^2 + (x_2 - 2)^2$

jedinicni vektori $e^1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e^2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e^3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, tada se poslije trece iteracije dobijaju

sljedeci rezultati:

$$x_3 = \begin{bmatrix} 0.579 \\ -0.932 \\ 0.769 \end{bmatrix}.$$

S obzirom da funkcija f nije kvadratna strogo konveksna funkcija, ne mozemo ocekivati da vrijednosti iz trece iteracije predstavljaju minimum funkcije. U poredjenju sa gradijentnim metodama moze se zakljuciti da su rezultati tacni samo do prvog decimalnog mjesta. Zato se vrijednosti x_3 uzimaju za novu pocetnu aproksimaciju i postupak se ponavlja sa jedinicnim vektorima kao polaznim smjerovima.

Problem koji moze da se javi kod Powell-ove metode, gledano čisto sa numeričkog stanovišta, sastoji se u tome da skup vektora postaje linearno zavisian. Tada se ovim postupkom ne dobijaju optimalna rješenja. Ovaj nedostatak je uočio i sam Powell i sugerisao da se posle doređenog broja iteracija nastavi sa koordinatnim osama kao pravcima konvergencije. Međutim, korekcijom ovog nedostatka gubi se svojstvo kvadratnog završavanja metoda gubi na efikasnosti. Ovim problemom bavio se i Zangwill i dosao do metode koja je superornija od Powell-ove. Medjutim, Powell-ova metoda je i dalje davala bolje rezultate od Zangwill-ove metode. Ono cime se najbolje moglo doskociti pojavi linearno zavisnih smjerova jeste ponovno uvođenje koordinatnih smjerova umjesto izracunatih konjugovanih pravaca.

I pored svojih nedostataka Powell-ova metoda uz inteligentno koriscenje ostaje kao najbolja u grupi metoda bez izracunavanja derivacija. No, uprkos svojim nedostacima, nema čvrste evidencije da postoji metoda bez izracunavanja derivacija, koja je bolja od konjugovane Powell-ove metode.

Istraživanja sa Hook-Jeeves-ovom i Powell-ovom metodom pokazala su da ove metode treba koristiti uglavnom, u slučaju nederivabilnih funkcija ili kada se derivacija ne može lako izračunati. U slučaju derivabilnih funkcija postoje efikasnije i brže metode, koje ce biti obradjene u daljem dijelu teksta .

5. VISEDIMENZIONALNA NEOGRANICENA OPTIMIZACIJA ZA DERIVABILNE FUNKCIJE

5.1. Gradijentne metode

Gradijentne metode se mogu primijeniti samo za derivabilne funkcije. Ovdje ćemo posmatrati metode bezuslovne optimizacije za derivabilne funkcije sa više promjenljivih.

Da bi uopšte govorili o gradijentnim metodama potrebno je prethodno opisati neke matematičke pojmove kao što su derivacija funkcije, gradijent, Jakobijeva i Hesse-ova matrica.

Derivacija funkcije u nelinearnom programiranju je važna iz više razloga. Derivacije opisuju ponašanje funkcija i na taj način klasifikuju funkcije na linearne, konveksne, konkavne itd. One služe za formulisanje uslova optimalnosti, tj. za provjeravanje da li je neko dopustivo rješenje optimalno ili ne. Ako dopustivo rješenje nije optimalno, derivacije se koriste za formulisanje numeričkih metoda koje služe za nalaženje boljih dopustivih rješenja.

Označimo sa $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ vektorsku funkciju definisanu na \mathbb{R}^n (Euklidov prostor n -koordinatni) sa vrijednostima u \mathbb{R}^m , tj

$$f(x) = \begin{bmatrix} f^1(x) \\ f^2(x) \\ \dots \\ f^m(x) \end{bmatrix}$$

Derivacija funkcije f u proizvoljnoj tački X je matrica sa m vrsta i n kolona:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f^1(x)}{\partial x_1} & \frac{\partial f^1(x)}{\partial x_2} & \dots & \frac{\partial f^1(x)}{\partial x_n} \\ \frac{\partial f^2(x)}{\partial x_1} & \frac{\partial f^2(x)}{\partial x_2} & \dots & \frac{\partial f^2(x)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f^m(x)}{\partial x_1} & \frac{\partial f^m(x)}{\partial x_2} & \dots & \frac{\partial f^m(x)}{\partial x_n} \end{bmatrix}$$

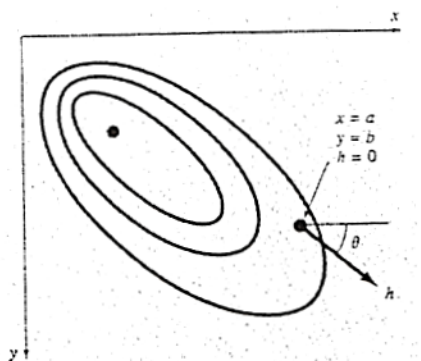
Ovako dobijena matrica naziva se Jakobijeva matrica. Ona predstavlja prvu derivaciju funkcije.

Ako imamo skalarnu funkciju f zavisnu od n promjenjivih, njena derivacija u tački X je prva vrsta Jakobijeve matrice:

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right]$$

Ova matrica vrsta naziva se gradijent.

Neka je data funkcija dvije promjenljive $f(x,y)$ i tacka $A(a,b)$ na njoj Sl 5.1. Ako zelimo da znamo nagib povrsi u toj tacki definisemo pravac h koji sa x -osom gradi ugao θ . Vrijednosti duz nove ose h formiraju novu funkciju $g(h)$. Ako poziciju tacke A uzmemo za $h = 0$, tada se nagib u toj tacki dobija kao prvi izvod funkcije g u toj tacki: $g'(0)$. Ovaj nagib, koji se naziva pravac izvoda, moze se izracunati preko parcijalnih izvoda duz x i y ose kao



Sl 5.1. Gradijent u tacki (a,b) je u pravcu ose h i gradi ugao Θ sa x -osom

$$g'(0) = \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta$$

gdje su parcijalni izvodi izracunati za tacku $x = a, y = b$.

S obzirom na postavljeni cilj, namece se pitanje koji je pravac najveceg porasta vrijednosti funkcije. Odgovor je gradijent, koji se definise kao:

$$\nabla f = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j}$$

On predstavlja pravac izvoda funkcije $f(x,y)$ u tacki (a,b) . Funkcija najbrze raste u pravcu gradijenta i najbrze opada u pravcu antigradijenta.

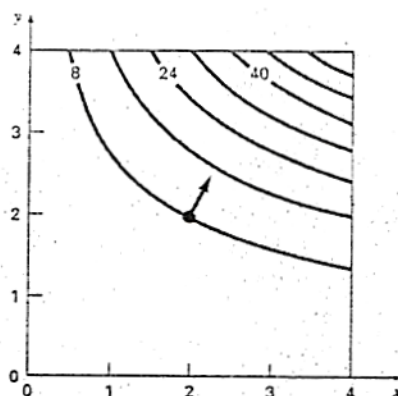
Primjer 5.1: Naci vrijednost gradijenta i nacrtati ga za funkciju

$$f(x, y) = xy^2 \quad \text{u tacki } (2,2).$$

Resenje: Vrijednost funkcije u tacki $(2, 2)$ je $f(x, y) = xy^2 = 2 (2)^2 = 8$

Parcijalni izvodi funkcije f su: $\frac{\partial f}{\partial x} = y^2 = 2^2 = 4$, $\frac{\partial f}{\partial y} = 2xy = 8$ pa je gradijent $\nabla f = 4\vec{i} + 8\vec{j}$.

Ugao koji pravac gradijenta zaklapa sa x-osom je $\text{tg } \vartheta = \frac{8}{4}$,
 $\vartheta = \text{tg}^{-1} 2 = 1.107 [\text{rad}] = 63.4^\circ$
 Intenzitet vektora gradijenta $|\nabla f| = 8.944$ graficki je prikazan na 5.2.



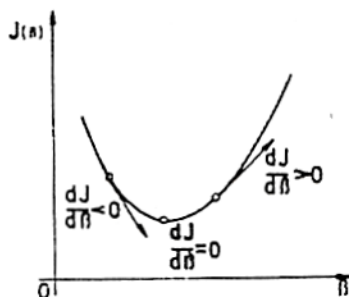
Sl 5.2. Gradijent u tacki $(2,2)$

Druga derivacija funkcije f u tački x ima oblik:

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

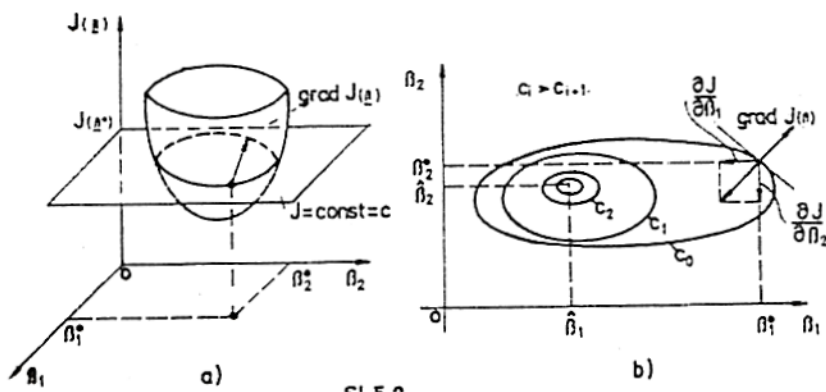
Druga derivacija skalarne funkcije f u tački x je simetrična matrica n -tog reda poznata kao Hesse-ova matrica.

Za funkciju jedne promjenljive prvi izvod je nagib tangente u nekoj tački i ukoliko je jednak nuli dobijamo stacionarnu tačku Sl 5.3.



Sl 5.3. Minimum funkcije jedne promjenljive

Za konveksnu funkciju dvije promjenljive pravac gradijenta je prikazan na Sl 5.4.



Sl 5.4. Funkcija dvije promjenljive

Da li je ta tačka max ili min funkcije zavisi od znaka drugog izvoda funkcije u toj tački. Za $y''(x) > 0$ imamo minimum i za $y''(x) < 0$ imamo maksimum funkcije.

Za funkciju dvije promjenljive Hesse-ova matrica je:

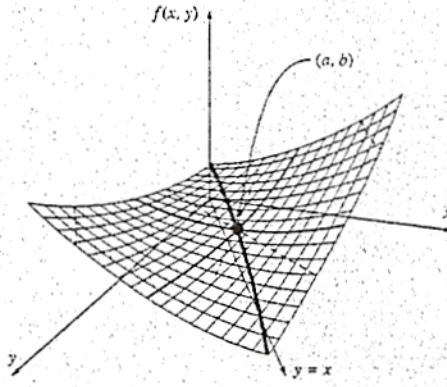
$$|H| = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2$$

Mogu nastupiti sljedeći slučajevi:

Ako je $|H| > 0$ i $\frac{\partial^2 f}{\partial x^2} > 0$ tada $f(x,y)$ ima lokalni minimum.

Ako je $|H| > 0$ i $\frac{\partial^2 f}{\partial x^2} < 0$ tada $f(x,y)$ ima lokalni maksimum.

Ako je $|H| < 0$ tada $f(x,y)$ ima sedlastu tačku, znači nema ni maksimuma ni minimuma kao što se vidi na Sl.5.5 za tačku (a,b) .



Sl 5.5. Sedlasta tacka

Ako tražimo optimalno rješenje (maksimum ili minimum) za funkciju više promjenljivih potrebno je da Hesse-ova matrica bude pozitivno definitna. Za konveksne funkcije svaki lokalni minimum je globalni minimum. Za konkavne funkcije svaki lokalni maksimum je globalni maksimum.

Bekey (1970) u grupu gradijentnih metoda uvrstio je metod najbržeg opadanja, Newton-Raphson, Newton i Gauss-Newton-ov metod. Pomenute metode koriste informaciju o intenzitetu nagiba lokalne konfiguracije površine, a njihova razlika sastoji se u broju članova Tajlorovog reda koji se uzima u analizu, što direktno utiče na brzinu konvergencije.

Osnovna iterativna jednačina gradijentnih metoda data je izrazom:

$$\beta = \beta_{i-1} + \alpha d_{i-1}$$

pri čemu je d_{i-1} pravac vektora, a α skalarna veličina. Kako gradijentne metode baziraju na razvoju Taylorovog reda i uključivanju različitog broja njegovih članova to je jasno da osnovna razlika ovih metoda leži u postupku za definisanje vektora d_{i-1} koji se može napisati kao

$$d_{i-1} = K \nabla J(\beta, T)$$

odnosno, razlika metoda je u definisanju matrice K što je prikazano u sljedećoj tabeli.

M e t o d	K
Najstrmijeg spusta	$K = k I$
Newton-Raphson	$K_{ij} = \frac{J(\beta, T)}{ \nabla J(\beta, T) ^2}$
Newton	$K = H^{-1} = \left \frac{\partial^2 J(\beta, T)}{\partial \beta_j \partial \beta_k} \right ^{-1}$
Gauss-Newton	$K = S^{-1} = \left 2 \int_0^1 \nabla J(\beta, T) \nabla J^T(\beta, T) dt \right ^{-1}$

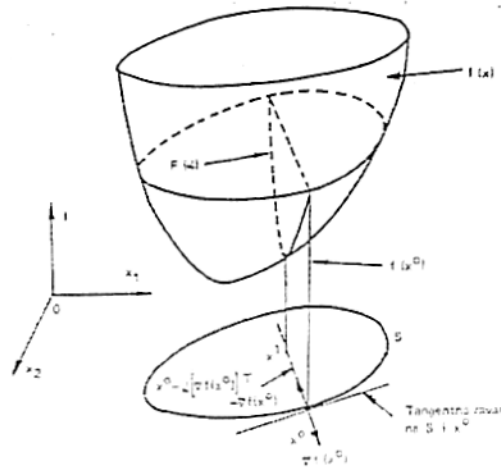
Gradijentne iterativne metode grubo se mogu podijeliti na dvije klase.

5.1. 1. Gradijentne metode prvog reda

U prvu klasu spadaju metode koje koriste samo prvu derivaciju funkcije f i zovu se Metode prvog reda.

5.1.1.1 Cauchy-eva metoda najstrmijeg opadanja

Najstarija Metoda prvog reda je Cauchyeva metoda najstrmijeg spusta koja potiče još od 1847. godine. Ideja metode prikazana je na sl.5.6. Kao sve numeričke metode nelinearne optimizacije i Cauchy-eva metoda je iterativna.



Sl.5.6 Ideja Cauchy-eva metode

Prvo se specificira neka proizvoljna aproksimacija β_0 . Zatim se u tački β_0 određuje gradijent $\nabla f(\beta_0)$. Gradient pokazuje smjer najbržeg porasta funkcije f u okolini β_0 . Kako se traži minimum (a ne maksimum) funkcije f , važniji je smjer najbržeg opadanja funkcije f , a on je određen negativnim gradijentom $-\nabla f(\beta_0)$. Funkcija f sada se pretražuje u tom smjeru, odnosno rješava se problem jednodimenzionalne optimizacije

$$\text{Min } f(\beta_0 - \alpha | \nabla f(\beta_0) |^T)$$

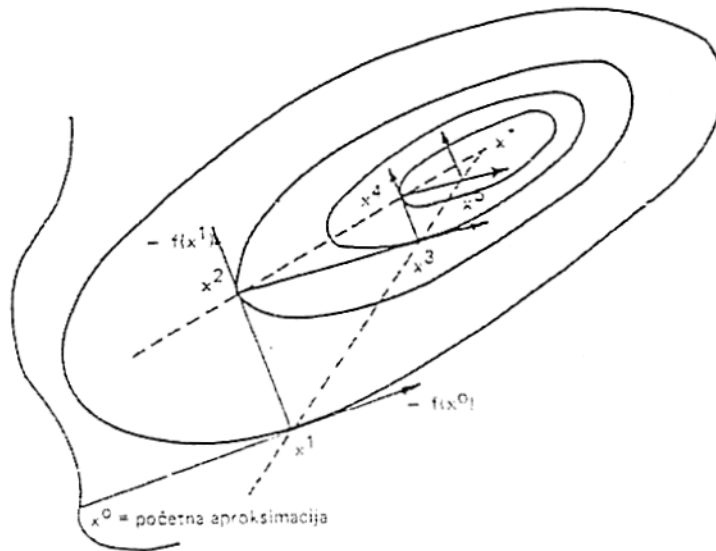
gdje se potroši dosta vremena za određivanje dužine koraka. Sada je $f(\beta_0 - \alpha | \nabla f(\beta_0) |^T)$ funkcija samo jedne promjenljive α koju možemo napisati kao

$$F(\alpha) = f(\beta_0 - \alpha | \nabla f(\beta_0) |^T)$$

Geometrijska interpretacija funkcije $F(\alpha)$ je očigledna, i u slučaju dvije promjenljive prikazana je na Sl. 5.7.

Vazno svojstvo Metode najstrmijeg opadanja je da su gradijenti, izračunati u dvije proizvoljne uzastopne aproksimacije, međusobno ortogonalni, tj.

$$\nabla f(x^{k+1})[\nabla f(x^k)]^T = 0, k = 0,1,2$$



Sl. 5.7 Konvergencija Cauchy-e

Cauchy-eva metoda ima linearni stepen konvergencije i odlikuje se dobrim napredovanjem prema tački minimuma β^* iz "dalekih" početnih aproksimacija, ali zato sporom konvergencijom u blizini optimalne tačke. Tipična konvergencija Cauchy-eve metode za funkciju f sa dvije promjenljive i elipsastim presjecima prikazana je na slici 5.7.

Primjer 5.2. Primijenimo Cauchyevu metodu na funkciju

$$f(x) = x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3$$

Iteraciju ćemo započeti iz ishodišta i sa računom ćemo prestati kad bude ispunjeno pravilo zaustavljanja sa $K = 0,1$, tj.

$$\sigma_k \leq 0,1 \cdot \sigma_0,$$

U normiranju vektora u^k koristićemo Čebiševljevu dužinu. (Podsjetimo se da je Čebiševljeva dužina, ili norma, vektora $z = (z_i)$, broj $\|z\| = \max \{|z_1|, \dots, |z_n|\}$). Budući da je $x^0 = (0, 0, 0)^T$ i $\nabla f(x^0) = [-1, 1, -1]$, nalazimo da je $\|[\nabla f(x^0)]^T\| = 1$ i zato je

$$u^0 = \frac{[\nabla f(x^0)]^T}{\|[\nabla f(x^0)]^T\|} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

Funkciju f sada ćemo pretražiti u smjeru

$$x^0 - \sigma u^0 = \begin{bmatrix} \sigma \\ -\sigma \\ \sigma \end{bmatrix}, \quad \sigma > 0.$$

Budući da je

$$F(\sigma) = f(x^0 - \sigma u^0) = f(\sigma, -\sigma, \sigma) = 2\sigma^4 - 3\sigma$$

Minimum ove funkcije (izračunat Metodom parabole) je $\sigma_0 = 0,71308$.

0,7211

Zato je prva izračunata aproksimacija

$$x^1 = x^0 - \sigma_0 u^0 = \begin{bmatrix} 0,71308 \\ -0,71308 \\ 0,71308 \end{bmatrix}$$

Dok je vrijednost funkcije u ishodištu $f(x^0) = 0$, u prvoj izračunatoj aproksimaciji vrijednost je znatno manja $f(x^1) = -1,62213$; očigledno je da se radi o velikom skoku.

Računanje se sada nastavlja iz x^1 . Da bi se zadovoljilo pravilo zaustavljanja bilo je potrebno izračunati trinaest iteracija. Rezultati tih proračuna prikazani su u tabeli.

tabela				
k	x_1^k	x_2^k	x_3^k	$f(x^k)$
1	0,71308	-0,71308	0,71308	-1,62213
2	0,52536	-0,90297	0,72361	-1,90097
3	0,56270	-0,93966	0,72774	-1,91020
4	0,56916	-0,93058	0,7499	-1,91122
5	0,56759	-0,93946	0,75407	-1,91158
6	0,57072	-0,93658	0,76153	-1,91170
7	0,56969	-0,93925	0,76310	-1,91175
8	0,57080	-0,93848	0,76578	-1,91176
9	0,57044	-0,93954	0,76634	-1,91177
10	0,57083	-0,93917	0,76731	-1,91177
11	0,57070	-0,93955	0,75751	-1,91177

12	0,57085	-0,93942	0,76786	-1,91177
13	0,57080	-0,93955	0,76794	-1,91177

Čebiševljeva dužina gradijenta u tački x^{13} manja je od 0,001. Ovo znači da je x^{13} blizu stacionarne tačke (u kojoj gradijent mora biti jednak nula-vektoru i zato je njegova dužina jednaka nuli). Budući da je funkcija strogo konveksna u okolini tačke x^{13} zaključujemo da x^{13} aproksimira (izolovani) lokalni minimum.

5.1.1.2 Fletcher-Reeves-ov metod konjugovanih gradijenata

U glavi 4.3 je prikazano kako je Powell-ov metod konjugovanih pravaca povećao efikasnost traženja ekstremnih vrijednosti funkcije u odnosu na metod traženja po koordinatnim osama. Tako se, koristeći konjugovane gradijente može poboljšati linearni stepen konvergencije Cauchy-eve metode najbrzeg opadanja. Optimizacione metode koristeći konjugovane gradijente za definisanje pravaca istraživanja dobijaju kvadratni stepen konvergencije. One nalaze optimalnu vrijednost za kvadratnu funkciju za određen broj koraka bez obzira na početnu vrijednost promjenljive.

Powell-ov metod ima kvadratni stepen konvergencije i ne zahtijeva informaciju o vrijednosti gradijenta.

Primjer 5.3 Koristeći Fletcher-Reeves-ov metod odrediti minimum funkcije

$$F(x) = 10x_1^2 + 3x_2^2 - 10x_1x_2 + 2x_1$$

polazeći od početne vrijednosti $x_0 = [0,0]^T$.

Gradijent funkcije F je

$$\nabla F(x) = \begin{bmatrix} 20x_1 - 10x_2 + 2 \\ -10x_1 + 6x_2 \end{bmatrix}$$

Prva iteracija:

Pravac negativnog gradijenta u tački x_0 je

$$g_0 = -\nabla F(x_0) = \begin{bmatrix} -2 \\ 0 \end{bmatrix},$$

$$v_0 = g_0 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \quad x_0 + sv_0 = \begin{bmatrix} -2s \\ 0 \end{bmatrix}.$$

Treba naci minimum funkcije jedne promjenljive

$$f(s) = F(x_0 + sv_0) = 10(2s)^2 + 3(0)^2 - 10(-2s)(0) + 2(-2s) = 40s^2 - 4s$$

$$f'(s) = 80s - 4 = 0, \quad s = 0.05$$

Tacka minimuma na pravcu $x_0 + sv_0$ je

$$x_1 = x_0 + sv_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + 0.05 \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.1 \\ 0 \end{bmatrix}$$

Druga iteracija:

$$g_1 = -\nabla F(x_1) = \begin{bmatrix} -20(-0.1) + 10(0) - 2 \\ 10(-0.1) - 6(0) \end{bmatrix} = \begin{bmatrix} 0 \\ -1.0 \end{bmatrix}$$

$$\gamma = \frac{g_1 g_1}{g_0 g_0} = \frac{1.0}{4} = 0.25$$

$$v_1 = g_1 + \gamma v_0 = \begin{bmatrix} 0 \\ -1.0 \end{bmatrix} + 0.25 \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.5 \\ -1.0 \end{bmatrix}$$

$$x_1 + sv_1 = \begin{bmatrix} -0.1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -0.5 \\ -1.0 \end{bmatrix} = \begin{bmatrix} -0.1 - 0.5s \\ -s \end{bmatrix}$$

$$f(s) = F(x_1 + sv_1) = 10(-0.1 - 0.5s)^2 + 3(-s)^2 - 10(-0.1 - 0.5s)(s) + 2(-0.1 - 0.5s) = 0.5s^2 - s - 0.1$$

$$f'(s) = s - 1 = 0 \quad s = 1.0$$

Minimum funkcije F(x) je :

$$x_2 = x_1 + sv_1 = \begin{bmatrix} -0.1 \\ 0 \end{bmatrix} + 1.0 \begin{bmatrix} -0.5 \\ -1.0 \end{bmatrix} = \begin{bmatrix} -0.6 \\ -1.0 \end{bmatrix}$$

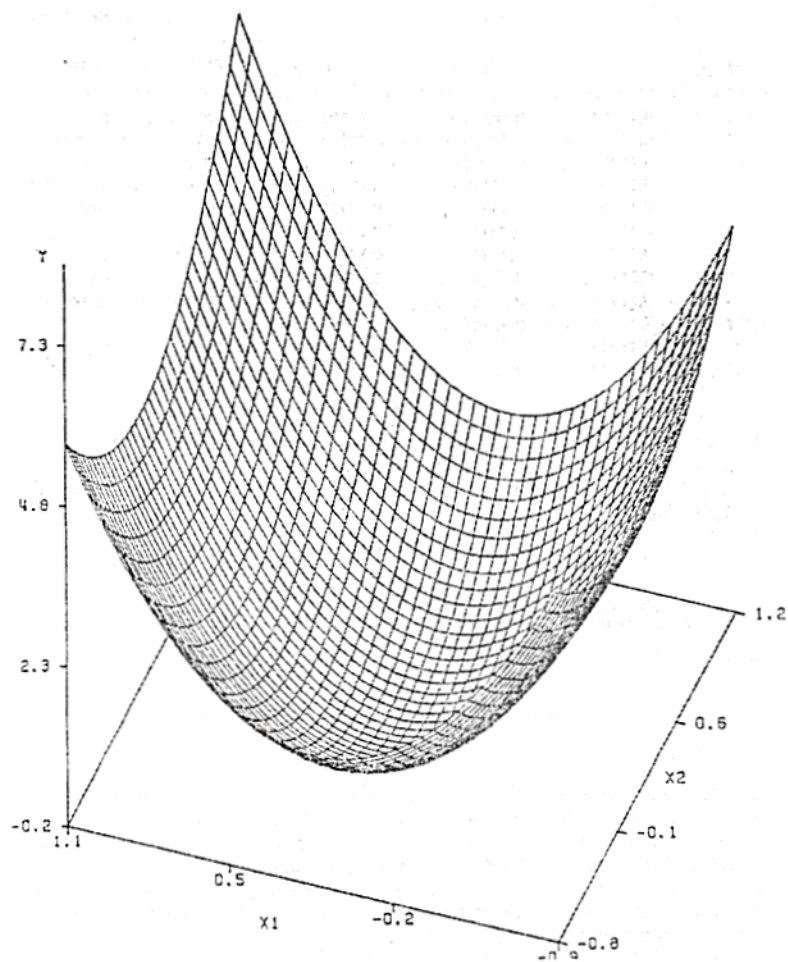
5.1.2. Gradijentne metode drugog reda

U ovu klasu spadaju metode koje koriste pored prve i drugu derivaciju ili neku njenu aproksimaciju i zovu se Metode drugog reda. Jedna od tih metoda je i Newton-ova metoda.

5.1.2.1. Njutnova metoda

Njutnova metoda za izracunavanje ekstremnih vrijednosti funkcije vise promjenljivih je uopstenje jednodimenzionalne Njutnove metode koja je izlozena u poglavlju 3.3. Odlikuje se superlinearnom konvergencijom, dakle, kad konvergira, čini to brže od Cauchyve metode. Njutnova metoda je brza ali "hirovita". Brzina Njutnove metode zavisi od funkcije f , načina na koji se vektori normalizuju i strategije koja se koristi u primjeni metode.

Kod metoda drugog reda osnovne poteškoće koje se javljaju vezane su za Hesseovu matricu. Minimum funkcije traži se samo za konveksne površi kao na Sl.5.8., što se matematički definiše Hesse-ovom matricom.



Sl .5.8. Strogo konveksna funkcija

Za dvaput neprekidno derivabilnu funkciju f kažemo da je konveksna ako je Hesseova matrica $\nabla^2 f(\beta)$ pozitivno definitna za svako β .

Njutnove metode pripadaju gradijentnim metodama drugog reda. Po istraživanjima koja je vršio naučnik McNiven, one bezuslovno konvergiraju pod pretpostavkom da je veličina koraka reducirana u svakoj sljedećoj iteraciji.

Njutnove metode bazira na razvoju funkcije kriterijuma odstupanja u Taylorov red u okolini tacke β_{i-1} uzimajući prva tri člana tog reda što u matričkom obliku izgleda:

$$J(\beta_i, T) = J(\beta_{i-1}, T) + \nabla J^T(\beta_{i-1}, T)(\beta_i - \beta_{i-1}) + \frac{1}{2}(\beta_i - \beta_{i-1})^T \nabla^2 J(\beta_{i-1}, T)(\beta_i - \beta_{i-1})$$

gdje je $\nabla J(\beta_i, T)$ vektor gradijenta, a $\nabla^2 J(\beta_i, T)$ Hesse-ova matrica. Minimum funkcije $J(\beta, T)$ postiže se ako su drugi i treći član jednačine istovremeno jednaki nuli, odnosno:

$$\nabla J(\beta_{i-1}, T) + \nabla^2 J(\beta_{i-1}, T)(\beta_i - \beta_{i-1}) = 0$$

što sređivanjem po vektoru β_i daje:

$$\beta_i = \beta_{i-1} - \nabla^2 J^{-1}(\beta_{i-1}, T) \cdot \nabla J(\beta_{i-1}, T) \quad (4.1)$$

Jednačina (4.1) poznata je kao jednačina Njutnovog metoda i u njoj inverzna Hesse-ova matrica predstavlja i pravac i intenzitet vektora spuštanja.

Primjer 5.4. Treba odrediti tačku lokalnog minimuma za funkciju

$$f(x) = x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3$$

pomoću Newtonove metode.

Ovdje je gradijent:

$$\nabla f(x) = (4x_1^3 + 3x_1^2 + x_2 x_3 - 1, 4x_2^3 - 2x_2 + x_1 x_3 + 1, 2x_3 + x_1 x_2 - 1)$$

i Hesseova matrica:

$$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 + 6x_1 & x_3 & x_2 \\ x_3 & 12x_2^2 - 2 & x_1 \\ x_2 & x_1 & 2x_3 \end{bmatrix}$$

Budući da je Hesseova matrica singularna u ishodištu, ne možemo uzeti ishodište za početnu aproksimaciju! (Zbog upoređivanja, primijetimo da Cauchy-eva metoda ne koristi drugu derivaciju i zato smo je mogli započeti iz ishodišta u primjeru 4.2). Sada ćemo iteraciju početi npr. iz tačke $x^0 = (1, -1, 1)^T$ i prekinuti je kada bude najveća komponenta gradijenta $[\nabla f(x^k)]^T$ po apsolutnoj vrijednosti manja od $\varepsilon = 10^{-5}$. U sljedećoj tabeli izložena je vrlo brza konvergencija. Rezultati su zaokruženi na pet decimalnih mjesta.

tabela				
k	x_1^k	x_2^k	x_3^k	$f(x^k)$
0	1	-1	1	-1
1	0,71037	-0,95427	0,83232	-1,83669
2	0,59255	-0,94165	0,77825	-1,91014
3	0,57150	-0,93962	0,76847	-1,91177
4	0,57086	-0,93956	0,76818	-1,91177
5	0,57086	-0,93956	0,76818	-1,91177

Željeno optimalno rješenje je:

$$x^* = x^4 = \begin{bmatrix} 0,57086 \\ -0,93956 \\ 0,76818 \end{bmatrix}.$$

Ovaj rezultat podudara se sa rezultatom, koji smo dobili pomoću Cauchy-eva metode, na tri decimalna mjesta.

Napomena: Jedan od načina da se smanji hirovitost Njutnove metode je da se u svakoj iteraciji funkcija f pretraži iz tačke x^k u smjeru

$$g(x^k) = [\nabla^2 f(x^k)]^{-1} [\nabla f(x^k)]^T$$

Ovako modifikovana Newtonova metoda naziva se Gauss-Njutnov metod i računa aproksimacije po formuli

$$x^{k+1} = x^k - \sigma_k g(x^k), \quad k = 0, 1, 2, \dots \quad (5.11)$$

gdje je σ_k optimalna dužina koraka, tj. rješenje jednodimenzionalnog problema

$$\text{Min } f(x^k - \sigma g(x^k)).$$

5.1.2.2. Gauss-Njutnova metoda

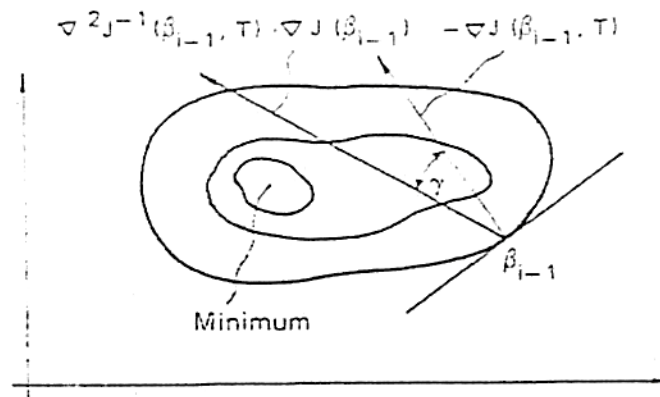
Da bi se smanjila "hirovitost" metode i obezbijedio uslov u kojem će se greška smanjivati u svakoj iteraciji, u jednacini (4.1) uvodi se pozitivni skalar α , tako da se vrši podešavanje veličine koraka u svakoj iteraciji, pa takva jednačina ima oblik:

$$\beta_i = \beta_{i-1} - \alpha \cdot \nabla^2 J^{-1}(\beta_{i-1}, T) \cdot \nabla J(\beta_{i-1}, T) \quad (4.2)$$

Tako se dobija jednačina modifikovane Njutnove metode, u literaturi poznata kao Gauss-Njutnova metoda.

Kod metode najbržeg spusta, ili Cauchy-ove metode koristi se pravac negativnog gradijenta dok kod Njutnovih metoda Hesse-ovom matricom vrši se zaokretanje pravca bliže minimumu za ugao γ prikazan na slici 5.9.

$$\gamma = \frac{\nabla J(\beta_{i-1}, T) \cdot \nabla^2 J^{-1}(\beta_{i-1}, T) \nabla J(\beta_{i-1}, T)}{\|\nabla J(\beta_{i-1}, T)\| \cdot \|\nabla^2 J^{-1}(\beta_{i-1}, T) \cdot \nabla J(\beta_{i-1}, T)\|}$$



SI. 5.9. Pravac istraživanja minimuma po Gauss-Njutnovoju metodi

Rešavajući zadatak optimizacije problem je sveden na istraživanje minimuma na pravcu zarotiranom od gradijenta za ugao γ . Koristeci neku od ranije navedenih metoda jednodimenzionalne optimizacije određuje se dužina koraka α .

Koju metodu ćemo primijeniti za određivanje dužine koraka α zavisi od konkretnog slučaja. Postoje metode upoređivanja vrijednosti funkcije za koje je karakteristično to da su tipično jednodimenzionalne metode što znači da se ne mogu lako uprostiti na probleme sa n promjenljivih. Za dovoljno "glatke" funkcije primjenjuje se Njutnova metoda. Ova metoda ima brzu konvergenciju (kad konvergira!). Pojednostavljena Njutnova metoda daje Metodu sječice.

Za primjenu veoma popularne metode za nalaženje minimuma su takozvane Metode aproksimacije polinoma. Takve metode su Metoda parabole ili Metoda kvadratne interpolacije, zatim Metoda kvadratne ekstrapolacije i smatraju se brzim i pouzdanim metodama kada funkcija nije suviše komplikovana. Kada se vrši izbor za primjenu optimizacionih metoda bitan faktor je i efikasnost numeričke metode za izračunavanje optimalnog rješenja pa u vezi s tim poželjno je za svaku metodu znati tzv. stepen konvergencije. Naravno, što je stepen konvergencije veći, metoda je efikasnija.

Neka je vektor β_{i-1} dobijen kao jedan od iterativnih vektora u toku procesa optimizacije funkcije više promjenljivih. Promjena vektora nepoznatih parametara β mijenja se po jednačini (4.2), gdje je očigledno da je novi vektor β_i funkcija nepoznate vrijednosti α koja se određuje nalaženjem minimuma funkcije:

$$J(\alpha, T) = J[\beta_{i-1} - \alpha \nabla^{-2} J(\beta_{i-1}, T) \cdot \nabla J(\beta_{i-1}, T)]$$

po pravcu koji je određen negativnim gradijentom funkcije kriterijuma odstupanja zaokrenutim za vrijednost inverzne Hesse-ove matrice. Tako se stiže do novog vektora β_i koji se nalazi na istoj površini. Nagib funkcije kriterijuma odstupanja nalazi se iz jednačine:

$$\frac{\partial}{\partial \alpha} J(\alpha, T) = -\nabla J^T(\beta_{i-1}, T) \cdot \nabla^{-2} J(\beta_{i-1}, T) \cdot \nabla J(\beta_{i-1}, T)$$

Zadatak se svodi na traženje tačke na pravcu d tako da skalarni proizvod gradijenta u toj tački i pravca d daje nulu. Grafička interpretacija ovakvog zahtjeva bila bi naći tačku na pravcu d u kojoj je gradijent normalan na pravac d .

Nalaženje tačke minimuma na pravcu teče na sljedeći način. Prvo se za odabranu početnu vrijednost vektora nepoznatih parametara odredi greška i gradijent u toj tački, a zatim izračuna inverzna Hesse-ova matrica drugog reda, što pomnoženo sa gradijentom daje pravac d u toj tački. Tada se izračuna skalarni proizvod vektora gradijenta i pravca d i memoriše ta vrijednost za dalje upoređivanje.

Na istom pravcu d odredi se još jedna tačka biranjem skalara α i nađe gradijent u njoj. Tada se pomnože vektor gradijenta u novoj tački sa pravcem d . Ukoliko je taj proizvod u početnoj tački istraživanja bio negativan a u drugoj pozitivan tada sigurno postoji tačka između te dvije u kojoj je navedeni skalarni proizvod jednak nuli i ona predstavlja minimum. U suprotnom, ukoliko je proizvod gradijenta u početnoj tački ispitivanja i odabranog pravca negativan, a u drugoj pozitivan radi se o maksimumu. Program je koncipiran tako da se druga tačka na pravcu d može birati dok se ne dobije prava vrijednost pogodna za dalja istraživanja, a zatim vrši traženje tačke minimuma upoređivanjem vrijednosti proizvoda vektora gradijenta i pravca d , tako što se proces ponavlja sve dotle dok se ne dobije vrijednost nagiba u novoj tački manja od unaprijed izabrane tolerancije zaustavljanja računskog programa.

Najzad, treba voditi računa i o tome da nalaženje vrijednosti vektora nepoznatih parametara u toku procesa integracije ne isključuje fizičke osobine što znači da tačka treba da bude u granicama mogućeg prostora.

5.2. Metode varijabilne metrike

Najbolje osobine Cauchy-eve i Njutnove metode ujedinjene su u Metode varijabilne metrike. One imaju linearan stepen konvergencije i odlikuju se kvadratnim završavanjem. Od mnogih metoda ovog tipa mogu se napomenuti Metoda Davidona, Fletchera i Powella.

Primjer 5.5 Zbog upoređivanja sa Njutnovom metodom na nekvadratnoj funkciji, primijenimo DFP metodu na funkciju

$$f(x) = x_1^4 + x_1^3 - x_1 + x_2^4 - x_2^2 + x_2 + x_3^2 - x_3 + x_1 x_2 x_3$$

Polazeći opet od početne aproksimacije

$$x_0 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

i sa izborom $H_0 = I$, dobijamo sljedeće rezultate:

tabela				
k	x_1^k	x_2^k	x_3^k	$f(x^k)$
0	1	-1	1	-1
1	0,64701	-1	1	-1,85478
2	0,59389	-0,91655	0,91655	-1,88632
3	0,58529	-0,93804	0,76954	-1,91099
4	0,57132	-0,94059	0,76944	-1,91177

Poslije upoređivanja sa rezultatima iz primjera 5.3, zaključujemo da su za gornju funkciju Njutnova i DFP metoda podjednako efikasne.

6. LITERATURA

1. AVRIEL, M.
[1996] *Nonlinear Programming: Analysis and Methods*, Prentice-Hall, Inc., Englewood Cliffs, N. J.
2. JENSEN, P.A. I BARD, J.F
[2003] *Operations Research, Models and Methods*, John Wiley&Sons, Inc.
3. JANG , J.S., SUN, C.T.
[1997] *Neuro-Fuzzy and Soft Computing*, Prentice Hall
4. BEER, F., JOHNSTON, R.
[1998] *Vector mechanics for Engineers*, McGraw-Hill
5. FOGEL, D., ROBINSON, C.,
[2003] *Computational Intelligence*, John Wiley&Sons
6. MERIAM, J.L., KRAIGE, L.G.
[2003] *Dynamics*, John Wiley&Sons
7. MERIAM, J.L., KRAIGE, L.G.
[2003] *Statics*, John Wiley&Sons
8. MEDSKER, L., LIEBOWITZ, J.
[1994] *Design and Development of Expert Systems and Neural Networks*, Macmilan College
9. FOWLER, B.,
[1996] *Dynamics*, Addison-Wesley
10. KARTALOPOULOS, S.V.
[1996] *Understanding Neural Networks and Fuzzy Logic*, IEEE Press