UNIVERZITET U NOVOM SADU

PRIRODNO - MATEMATIČKI FAKULTET

DEPARTMAN ZA MATEMATIKU I INFORMATIKU

Rešavanje problema minimizacije sa ograničenjima primenom kaznenih funkcija

\- master rad -

Mentor: Student:

Prof. dr Sanja Rapajić Angela Štajer

Novi Sad, 2014

Sadržaj

1 Uvod

Operaciona istraživanja su oblast koju čine raznovrsni naučni pristupi donošenju odluka, koji pomoću kvantitativnih metoda i algoritama nastoje da omoguće najbolje rešenje velikih i složenih problema koji se javljaju u svim oblastima života. Neke metode operacionih istraživanja pojavile su se u prvoj polovini 20. veka, ali najviše su se razvile tokom i nakon Drugog svetskog rata. Glavni cilj pri upotrebi ovih metoda bio je optimiziranje organizacije velikih ratnih operacija i način upotrebe ratnih sredstava, pa otuda i potiče naziv ove oblasti.

Operaciona istraživanja imaju za zadatak proučavanje velikih i složenih sistema (tehničkih, ekonomskih, poslovnih, itd.). Istraživanje na ovim sistemima uglavnom predstavlja izuzetno težak zadatak, koji je po pravilu veoma skup, a ponekad i nemoguć, naročito u fazi njihovog planiranja, projektovanja i uvođenja u rad. Zbog toga je veoma bitna celokupna analiza sistema, gde veoma važnu ulogu igraju mnogobrojne matematičke discipline kao što su: matematička analiza, statistička teorija, teorija igara, matematička logika, teorija verovatnoće, linearno, nelinearno i dinamičko programiranje, heurističko-matematičko programiranje, teorija masovnog opsluživanja i dr.

Naučni pristup donošenju odluka obično uključuje upotrebu jednog ili više matematičkih modela, na osnovu kojeg će biti moguće predviđanje i upoređivanje posledica različitih varijanti u procesu odlučivanja. Matematički model je matematičko predstavljanje stvarne situacije, koji služi za izvođenje potrebnih analiza na osnovu kojih se može doći do traženih odgovora u vezi sa postavljenim problemom. Na osnovu eksperimentisanja na matematičkom modelu treba zaključiti koje je rešenje najbolje i njega treba predložiti donosiocima odluke da ga implementiraju u praksi. Da bi se za neko rešenje moglo reći da je optimalno, tj. najbolje, treba imati meru kojom se određuje njegov kvalitet i koja omogućava njegovo poređenje sa drugim mogućim rešenjima.

Mera kvaliteta u matematičkom modelu se dobija pomoću funkcije, kojom se svakom rešenju pridružuje odgovarajuća vrednost. Ta funkcija pokazuje efikasnost izvršenja zadatka radi postizanja cilja i naziva se ***kriterijumska funkcija*** ili ***funkcija cilja***.

Da bi optimizacija bila što uspešnija, potrebno je naći rešenje koje daje ekstremnu vrednost funkcije cilja. Promenljive koje treba odrediti obično su uslovljene međusobnim relacijama i ograničenjima, koji čine ***skup ograničenja***. Svako rešenje koje zadovoljava postojeća ograničenja naziva se ***dopustivim***. Jedan optimizacioni zadatak može imati više dopustivih rešenja, koja formiraju skup dopustivih rešenja ili *dopustivi skup*. Koeficijenti u matematičkom modelu nazivaju se i *parametrima modela* ili *parametrima sistema*.

Kako je glavni zadatak optimizacije nalaženje rešenja koje daje ekstremnu vrednost funkcije cilja, ***optimalno rešenje*** optimizacionog zadatka je ono dopustivo rešenje u kome funkcija cilja dostiže svoju optimalnu vrednost (minimalnu, ako je u pitanju problem minimizacije ili maksimalnu, ako je u pitanju problem maksimizacije). Vrednost funkcije cilja koja odgovara optimalnom rešenju naziva se ***optimalna vrednost*** ili ***optimum***.

Koliko se uspešno može odrediti optimalno rešenje nekog optimizacionog zadatka zavisi, pre svega, od osobina funkcije cilja, tipa ograničenja i primenjene metode. Zbog toga je osnovni zadatak u teoriji optimizacije razvoj metoda za rešavanje zadataka globalne optimizacije.

Ovaj rad se bavi rešavanjem nelinearnih problema optimizacije, što znači da je reč o nelinearnom programiranju. Najpre će biti uvedene neke teorijske osnove nelinearnog programiranja i potrebni i dovoljni uslovi za određivanje lokalnog ekstrema za probleme optimizacije sa i bez ograničenja.

Problemi sa ograničenjima se mogu rešavati na dva suštinski različita načina. Prvi je određivanje optimuma uz zadovoljavanje ograničenja, a drugi je transformacija problema na način da se zadovoljavanje ograničenja "ugradi" u funkciju cilja, i tada se transformirani problem rešava kao zadatak bez ograničenja. Metode koje pripadaju drugoj grupi pristupa problema, i koje će biti predstavljene u ovom radu su, metode unutrašnjih i spoljašnjih kaznenih funkcija.

Kada se dati problem sa ograničenjima pretvori u problem bez ograničenja, potrebno je rešiti takav optimizacioni zadatak. Za rešavanje takvih problema najčešće se koriste gradijentna metoda, Njutnova metoda i njene modifikacije, metoda konjugovanih gradijenata, koje su objašnjene u radu teorijski i prikazane kroz primere. Za rešavanje nelinearnih problema vrlo često se koriste kombinovane metode, pa je zato u ovom radu posebna pažnja posvećena metodama koje predstavljaju kombinaciju kaznenih funkcija sa metodom konjugovanih gradijenata i Njutnovom metodom, ili nekim modifikacijama Njutnove metode. Prikazani su algoritmi i teorijski rezultati konvergencije datih postupaka. Predstavljena teorija objašnjena je kroz primere. Neki od primera u radu rešeni su pomoću računara, primenom programskih paketa *Matlab*® i *Mathematica*®.

2 Nelinearna optimizacija

Nelinearna optimizacija (nelinearno programiranje - NP) podrazumeva određivanje ekstremne (minimalne ili maksimalne) vrednosti funkcije cilja nad skupom ograničenja, pri čemu je funkcija cilja nelinearna i/ili skup ograničenja definisan nelinearnim algebarskim jednačinama ili nejednačinama. Za rešavanje problema NP ne postoji univerzalna metoda, već je za svaki konkretan slučaj potrebno naći novu metodu, ili prilagoditi neku od postojećih. U velikom broju slučajeva nije lako pronaći prikladnu metodu za određivanje optimalnog rešenja formulisanog zadatka, što znači da još uvek postoji veliki broj teško rešivih zadataka NP.

Tokom godina razvijene su mnogobrojne metode optimizacije pomoću kojih se mogu rešavati razni zadaci nelinearnog programiranja. Sve te metode su specijalizovane za različite tipove zadataka koji se razlikuju po obliku matematičkog modela, tj. po obliku i dimenzijama funkcije cilja i skupa ograničenja.

Kako bi se procenila primenljivost pojedinih algoritama za rešavanje problema NP, potrebno je odrediti broj računskih operacija koje treba obaviti u procesu rešavanja. Neki algoritmi u određenim zadacima nisu uvek primenljivi, čak i kada se ima u vidu primena savremenih računara. Ta primena zavisi od više faktora, naročito od karaktera i dimenzije matematičkog modela.

2.1 Osnovni pojmovi

Dat je opšti oblik problema nelinearne optimizacije

$$\min(\max) f(x_1, x_2, \dots, x_n) \tag{1}$$
$$g_i(x_1, x_2, \dots, x_n) \le 0,\ i = 1, \dots, m,$$

gde je $f: R^n \to R$ funkcija cilja, a $g_i: R^n \to R$, $i = 1, \dots, m$ su funkcije ograničenja.

R^n označava euklidski prostor dimenzije n; $x \in R^n$, tj. $x = [x_1, x_2, \dots, x_n]^T$.

Problem (1) se može zapisati u ekvivalentnom obliku

$$\min(\max) f(x)$$
$$g_i(x) \le 0, i = 1, \dots, m$$

ili

$$\min_{x \in X}(\max_{x \in X}) f(x)$$

gde je $X = \{x \in R^n,\ g_i(x) \le 0,\ i = 1, \dots, m\}$ dopustiv skup.

Ovakav oblik nelinearnog programiranja se naziva problem ***optimizacije sa ograničenjima***.

Ukoliko je skup ograničenja u problemu $X = R^n$, tj. ako ne postoje ograničenja onda je reč o problemu ***optimizacije bez ograničenja*** tj.

$$\min_{x \in R^n} (\max_{x \in R^n}) f(x).$$

Problem minimizacije se uvek može svesti na problem maksimizacije i obrnuto. Dakle,

$$\max_{x \in X} f(x) = \min_{x \in X} (-f(x)).$$

Dopustivi skup se bez umanjenja opštosti uvek može definisati preko ograničenja tipa nejednakosti jer se ograničenje oblika jednakosti $h(x) = 0$ može zameniti sa dve nejednakosti $h(x) \leq 0,\ h(x) \geq 0$ tj. sa $h(x) \leq 0,\ -h(x) \leq 0$.

Specijalan slučaj optimizacije sa ograničenjima je problem oblika

$$\min(\max) f(x)$$
$$g_i(x) = 0, \qquad i = 1, \ldots, m,$$

gde je skup ograničenja dat jednakostima.

Ovakav problem se naziva ***klasični problem optimizacije sa ograničenjima*** i rešava se direktno, bez svođenja na ograničenja tipa nejednakosti.

Definicija 1. Svako $x \in X$ gde je X dopustiv skup se naziva *dopustivo rešenje*. Ograničenje $g_i(x) \leq 0$ je *aktivno* u dopustivom rešenju $\bar{x}$ ako je $g_i(\bar{x}) = 0$, a *neaktivno* ako je $g_i(\bar{x}) < 0$.

Definicija 2. Dopustivo rešenje u kome funkcija cilja dostiže svoju optimalnu vrednost naziva se ***optimalno rešenje*** optimizacionog zadatka i označava se sa x^*.

Definicija 3. Vrednost funkcije cilja koja odgovara optimalnom rešenju naziva se ***optimalna vrednost*** ili ***optimum*** i označava se sa $f^* = f(x^*)$.

2.2 Pojam lokalnog i globalnog ekstrema

Neka je dat problem

$$\min_{x \in X} f(x) \tag{2}$$

gde je X dopustiv skup.

Definicija 4. Tačka $x^* \in X$ je *globalni minimum* funkcije f na skupu X, ili *globalni minimum* problema (2), ako je

$$f(x^*) \leq f(x) \text{ za sve } x \in X,$$

tj. ako u njoj funkcija f dostiže svoju minimalnu vrednost na X.

Definicija 5. Tačka $x^* \in X$ je *strogi globalni minimum* funkcije f na skupu X, ili *strogi globalni minimum* problema (2), ako je

$$f(x^*) < f(x) \text{ za sve } x \in X, x \neq x^*.$$

U slučaju kada je funkcija f neprekidna a skup X zatvoren i ograničen, f dostiže svoju minimalnu vrednost, tj. postoji (bar jedan) globalni minimum $x^* \in X$.

Definicija 6. Tačka $x^* \in X$ je *lokalni minimum* funkcije f na skupu X, ili *lokalni minimum* problema (2), ako postoji $\delta > 0$ tako da je

$$f(x^*) \leq f(x) \text{ za sve } x \in X \text{ za koje je } \|x - x^*\| < \delta.$$

Definicija 7. Tačka $x^* \in X$ je *strogi lokalni minimum* funkcije f na skupu X, ili *strogi lokalni minimum* problema (2), ako postoji $\delta > 0$ tako da važi uslov

$$f(x^*) < f(x) \text{ za sve } x \in X \text{ za koje je } \|x - x^*\| < \delta, x \neq x^*.$$

Analogno se mogu uvesti pojmovi globalnog i lokalnog maksimuma. Globalni, odnosno lokalni minimumi i maksimumi nazivaju se globalni, odnosno lokalni ekstremi.

Strogi lokalni minimumi se obično mogu odrediti uz pomoć prvih i drugih izvoda u tački x^*. Na ovoj činjenici se baziraju mnogi numerički postupci, koji spadaju u metode nelinearnog programiranja. U jednostavnim slučajevima, kada je skup X zatvoren i ograničen, globalni minimum se nalazi pretraživanjem skupa svih lokalnih minimuma, posle čega se izabere onaj sa najmanjom vrednošću funkcije cilja ali je ovo moguće samo u jednostavnim problemima.

2.3 Optimizacija bez ograničenja - Potrebni i dovoljni uslovi optimalnosti

Optimalno rešenje zadatka optimizacije se može odrediti posmatranjem vrednosti funkcije cilja u svim dopustivim rešenjima razmatranog problema. Optimalno rešenje je ono rešenje u kome funkcija cilja dostiže maksimalnu ili minimalnu vrednost. Ispitivanje svih dopustivih rešenja i upoređivanje vrednosti funkcije cilja u svim dopustivim rešenjima je moguće uraditi samo u najtrivijalnijim problemima. Zbog toga se razvijaju postupci, pomoću kojih se najpre generišu kandidati za optimalno rešenje, a onda se proverava da li je neki od njih optimalno rešenje. Posebna pažnja se posvećuje proveravanju uslova koje treba da zadovolji optimalno rešenje.

Pri nalaženju optimalnog rešenja datog zadatka može se desiti da se minimalna (maksimalna) vrednost funkcije cilja dostiže u različitim tačkama, što znači da rešenje problema nije jedinstveno. U opštem slučaju korisno je naći uslove koje neko rešenje mora da zadovolji da bi bilo optimalno, dok druga dopustiva rešenja mogu ali ne moraju da ih zadovolje. Takvi uslovi se nazivaju ***potrebni uslovi*** za optimalnost, odnosno ***potrebni uslovi optimalnosti***.

Neka je dat problem

$$\min_{x \epsilon R^n} f(x),$$

gde je funkcija cilja $f: R^n \to R$ definisana i diferencijabilna na čitavom prostoru R^n.

Definicija 8. Neka je $x = [x_1, \dots, x_n]^T$ i neka je $f(x) \epsilon C^1(R^n)$. Tada je

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_2} \end{bmatrix} \qquad (3)$$

gradijent funkcije $f(x)$.

Potrebni uslovi da tačka x^* bude lokalni minimum funkcije f dati su sledećom teoremom.

Teorema 1. [7] Neka je funkcija f diferencijabilna u tački x^*. Ako je x^* lokalni minimum funkcije f tada je $\nabla f(x^*) = 0$.

Definicija 9. Tačke koje zadovoljavaju uslov $\nabla f(x^*) = 0$ nazivaju se ***stacionarne tačke***.

Stacionarne tačke se određuju iz jednačine $\nabla f(x) = 0$ tj. rešavanjem sistema od n jednačina sa n nepoznatih, oblika

$$\frac{\partial f}{\partial x_1} = 0, \dots, \frac{\partial f}{\partial x_n} = 0.$$

Uslov koji garantuje da je posmatrano rešenje optimalno naziva se ***dovoljan uslov*** optimalnosti. Da je neko rešenje optimalno, može se tvrditi samo ukoliko ono zadovoljava uslov optimalnosti koji je istovremeno i potreban i dovoljan, i obrnuto, samo optimalna rešenja zadovoljavaju potrebne i dovoljne uslove za optimalnost.

Definicija 10. Simetrična matrica $Q = \left[q_{ij}\right]_{n\times n}$ je ***pozitivno definitna*** ako je

$$x^T Qx > 0 \text{ za sve } x \in R^n,\ x \neq 0,$$

a ***pozitivno semidefinitna*** ako je

$$x^T Qx \geq 0 \text{ za sve } x \in R^n.$$

Definicija 11. Funkcija

$$q(x) = x^T Qx = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x_i x_j$$

naziva se ***kvadratna forma***. Kvadratna forma je pozitivno definitna (pozitivno semidefinitna) ukoliko je matrica Q pozitivno definitna (pozitivno semidefinitna).

Na osnovu narednih lema jednostavno se proverava da li je matrica pozitivno definitna ili pozitivno semidefinitna.

Lema 1. [7] **Silvesterov kriterijum.** Simetrična matrica $Q = \left[q_{ij}\right]_{n\times n}$ je pozitivno definitna ako i samo ako su svi glavni minori matrice Q pozitivni, tj.

$$D_1 = q_{11} > 0,\ D_2 = \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} > 0, \dots,\ D_n = \begin{vmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \cdots & q_{nn} \end{vmatrix} > 0.$$

Lema 2. [7] Simetrična matrica $Q = \left[q_{ij}\right]_{n\times n}$ je pozitivno semidefinitna ako i samo ako su svi minori simetrični u odnosu na glavnu dijagonalu nenegativni.

Definicija 12. Neka je $x = [x_1, \dots, x_n]^T$ i neka je $f(x)\epsilon C^2(R^n)$. Tada je

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \tag{4}$$

Hesijan funkcije $f(x)$.

Teoreme koje slede daju dovoljne uslove optimalnosti drugog reda za stroge lokalne ekstreme.

Teorema 2. [7] Neka je funkcija f dva puta neprekidno diferencijabilna u nekoj okolini tačke x^*. Ako je $\nabla f(x^*) = 0$ i ako je matrica $\nabla^2 f(x^*)$ pozitivno definitna, tada je x^* strogi lokalni minimum funkcije f.

Teorema 3. [7] Neka je $\nabla f(x^*) = 0$ i neka su $D_1, \dots, D_n$ glavni minori matrice $\nabla^2 f(x^*)$

- ako je $D_j > 0, j = 1, \dots, n$ tada je x^* strogi lokalni minimum funkcije f
- ako je $(-1)^j D_j > 0,\ j = 1, \dots, n$ tada je x^* strogi lokalni maksimum funkcije f.

Ukoliko ne postoji pravilnost znaka glavnih minora treba uzeti u obzir potrebne uslove optimalnosti drugog reda koji su dati u narednoj teoremi.

Teorema 4. [7] Neka je funkcija f dva puta neprekidno diferencijabilna u nekoj okolini tačke x^*. Ako je x^* lokalni minimum funkcije f tada je $\nabla f(x^*) = 0$ i matrica $\nabla^2 f(x^*)$ je pozitivno semidefinitna.

2.4 Optimizacija sa ograničenjima - Potrebni i dovoljni uslovi optimalnosti

Ograničenja optimizacionog problema definišu dopustiv skup problema, pa zato, ukoliko ona postoje, moraju biti uzeta u obzir pri izboru strategije za rešavanje takvog problema. Pre nego što se izabere postupak kojim će se rešavati dati problem, potrebno je proveriti model problema, jer loše postavljena ograničenja mogu dopustiv skup preterano smanjiti ili čak problem učiniti nerešivim.

Kada su u pitanju problemi sa ograničenjima, postoje dva suštinski različita pristupa njihovom rešavanju. Prvi je određivanje optimuma uz zadovoljavanje ograničenja, a drugi je transformacija problema na način da se zadovoljavanje ograničenja "ugradi" u funkciju cilja, i tada se transformirani problem rešava kao zadatak bez ograničenja.

Posmatra se klasičan problem minimizacije sa ograničenjima, kod koga su sva ograničenja zadata jednačinama

$$\min f(x) \tag{5}$$
$$h_i(x) = 0, \quad i = 1, \dots, m,$$

gde su funkcije $f: R^n \to R$ i $h_i: R^n \to R,\ i = 1, \dots, m$ definisane i diferencijabilne na čitavom prostoru R^n.

Ovaj problem se može rešiti metodom eliminacije promenljivih ili metodom Lagranžovih množitelja.

Neka je sa $J(x)$ označen ***Jakobijan*** preslikavanja $h(x)$ tj.

$$J(x) = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix}.$$

2.4.1 Metoda eliminacije promenljivih

Osnova ove metode se sastoji u svođenju problema (5) na problem optimizacije bez ograničenja. Ideja je da se promenljive $x_1, \dots, x_m$ izraze iz sistema jednačina

$$\begin{gathered} h_1(x_1, \dots, x_n) = 0 \\ \vdots \\ h_m(x_1, \dots, x_n) = 0 \end{gathered} \tag{6}$$

kao

$$\begin{gathered} x_1 = \varphi_1(x_{m+1}, \dots, x_n) \\ \vdots \\ x_m = \varphi_m(x_{m+1}, \dots, x_n) \end{gathered}$$

i uvrste u funkciju cilja. Na taj način se eliminiše m promenljivih i problem (5) se svodi na problem minimizacije bez ograničenja:

$$\min f(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) = \min F(x_{m+1}, \dots, x_n). \tag{7}$$

Ako tačka $(x_{m+1}^*, \dots, x_n^*)$ predstavlja rešenje problema (7) sledi da je tačka $(\varphi_1(x_{m+1}^*, \dots, x_n^*), \dots, \varphi_m(x_{m+1}^*, \dots, x_n^*), x_{m+1}^*, \dots, x_n^*)$ rešenje problema (5).

Sistem (6) je moguće rešiti analitičkim putem samo u jednostavnim slučajevima, na primer kada su funkcije $h_1, \dots, h_m$ linearne. Kada su funkcije nelinearne, da bi se mogla efektivno izvršiti eliminacija, Jakobijan preslikavanja $h(x)$ mora biti punog ranga, tj. $rangJ(x) = m$ na R^n. Dakle, eliminacija promenljivih nije opravdana i jednostavna u nelinearnom slučaju.

2.4.2 Metoda Lagranžovih množitelja

Kod ove metode se koriste potrebni i dovoljni uslovi optimalnosti izraženi preko Lagranžove funkcije.

Definicija 13. *Lagranžova funkcija* pridružena problemu (5) je definisana na sledeći način:

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x),$$

gde su $\lambda_i,\ i = 1,2,\dots,m$ *Lagranžovi množitelji, a* $\lambda = [\lambda_1,\dots,\lambda_m]^T$ *vektor Lagranžovih množitelja.*

Slede potrebni uslovi da tačka x^* bude lokalni minimum problema (5).

Teorema 5. [7] Neka su funkcije $f, h_1, \dots, h_m$ neprekidno diferencijabilne u nekoj okolini tačke x^* i neka je $J(x)$ Jakobijan funkcije $h(x)$. Ako je x^* lokalni minimum problema (5) i ako je rang $J(x^*) = m$, tada postoji λ^* tako da je $\nabla L(x^*,\lambda^*) = 0$.

Ako je rang $J(x) = m$ u svakom dopustivom rešenju x, na osnovu prethodne teoreme sledi da su stacionarne tačke Lagranžove funkcije, potencijalni lokalni minimumi problema (5).

Rešavanje jednačine $\nabla L(x,\lambda) = 0$ je ekvivalentno sa $\nabla_x L(x,\lambda) = 0$ i $\nabla_\lambda L(x,\lambda) = 0$, tj. sa

$$\begin{aligned} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) &= 0 \\ h_i(x) = 0, \qquad i &= 1,2,\dots,m. \end{aligned}$$

Dovoljni uslovi za strogi lokalni minimum problema (5) formulisani su u sledećoj teoremi.

Teorema 6. [7] Neka su funkcije $f, h_1, \dots, h_m$ dvaput neprekidno diferencijabilne u nekoj okolini tačke x^*. Ako je $\nabla L(x^*,\lambda^*) = 0$ i ako je matrica $\nabla^2_{xx} L(x^*,\lambda^*)$ pozitivno definitna na $T(x^*)$, tada je x^* strogi lokalni minimum problema (5), gde je

$$\nabla^2_{xx} L(x,\lambda) = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{bmatrix}$$

Hesijan Lagranžove funkcije, a $T(x^*) = \{x \in R^n | J(x^*)x = 0\}$ - tangentni prostor, tj. $T(x^*)$ predstavlja skup rešenja homogenog sistema od m linearnih jednačina sa n nepoznatih, pri čemu je $J(x^*)$ Jakobijan funkcije h u x^*.

Uslov pozitivne definitnosti se može proveriti pomoću uopštenog Silvesterovog kriterijuma.
Neka je sa $H(x^*, \lambda^*)$ označena matrica dimenzije $(m+n) \times (m+n)$ koja se sastoji od četiri bloka

$$H(x^*, \lambda^*) = \begin{bmatrix} 0 & J(x^*) \\ J^T(x^*) & \nabla_{xx}^2 L(x^*, \lambda^*) \end{bmatrix},$$

gde je 0 matrica dimenzije $m \times m$ čiji su svi elementi jednaki nuli, a $J(x^*)$ je Jakobijan preslikavanja h u x^*.

Teorema 7. [7] **Uopšteni Silvesterov kriterijum.** Neka je rang $J(x^*) = m$ i neka su $D_1, \dots, D_{m+n}$ glavni minori matrice $H(x^*, \lambda^*)$. Ako je

$$(-1)^m D_{2m+1} > 0, \dots, (-1)^m D_{m+n} > 0,$$

tada je matrica $\nabla_{xx}^2 L(x^*, \lambda^*)$ pozitivno definitna na $T(x^*)$.

Slede dovoljni uslovi za strogi lokalni minimum i maksimum.

Teorema 8. [7] Neka je $\nabla L(x^*, \lambda^*) = 0$, rang $J(x^*) = m$ i neka su $D_1, \dots, D_{n+m}$ glavni minori matrice $H(x^*, \lambda^*)$

- ako je $(-1)^m D_{2m+j} > 0,\ j = 1, \dots, n-m$, tada je x^* strogi lokalni minimum problema (5).
- ako je $(-1)^{m+j} D_{2m+j} > 0,\ j = 1, \dots, n-m$, tada je x^* strogi lokalni maksimum problema (5).

3 Metode za rešavanje problema optimizacije bez ograničenja

Kao što je pomenuto, problem optimizacije bez ograničenja je problem oblika

$$\min_{x \in R^n} f(x)$$

gde je funkcija cilja $f: R^n \to R$ definisana na prostoru R^n.

Postupci za rešavanje problema optimizacije bez ograničenja dele se na negradijentne i gradijentne metode. Negradijentne metode koriste samo vrednost funkcije, a ne i njene izvode. One su jednostavne, ali sporije konvergiraju i mogu se primeniti za rešavanje problema optimizacije funkcija koje nisu diferencijabilne. Gradijentne metode koriste izvode funkcije i dele se na metode prvog i drugog reda.

Metode prvog reda koriste samo prvi izvod funkcije cilja. Najpoznatija i najstarija metoda prvog reda je Košijeva metoda najbržeg pada, koja potiče iz 1847. godine. Ova metoda je pouzdana, jer se dolazi do optimuma iz bilo koje početne aproksimacije, čime se postiže globalna konvergencija ali je relativno spora, tj. linearno konvergentna.

Metode drugog reda koriste prvi i drugi izvod funkcije cilja. Ove metode brže konvergiraju od metoda prvog reda, obično su superlinearno ili kvadratno konvergentne, ali su skuplje i nekad manje pouzdane.

Definicija 14. Neka niz $\{x^k\}$ konvergira ka x^* kad $k \to \infty$, tj. $\lim_{k\to\infty} x^k = x^*$. Ako važi

$i)$ $\|x^{k+1} - x^*\| \le c\|x^k - x^*\|, 0 < c < 1,$ tada $\{x^k\}$ konvergira linearno ka x^* sa koeficijentom brzine c.

$ii)$ $\|x^{k+1} - x^*\| \le c_k\|x^k - x^*\|, \lim_{k\to\infty} c_k = 0,$ tada $\{x^k\}$ konvergira superlinearno ka x^*.

$iii)$ $\|x^{k+1} - x^*\| \le c\|x^k - x^*\|^p,\ c > 0$ tada $\{x^k\}$ konvergira ka x^* sa redom p.

3.1 Metoda najbržeg pada

Metoda najbržeg pada, poznata kao Košijeva (Cauchy) metoda bazirana je na činjenici da funkcija najbrže opada u pravcu negativnog gradijenta. Zbog toga je neophodna pretpostavka za primenu ove metode da je funkcija f bar jednom neprekidno diferencijabilna. Košijeva metoda je iterativna metoda. Najpre se zada početna iteracija x^o.

Pošto funkcija najbrže opada u pravcu negativnog gradijenta, naredna iteracija x^1 se traži upravo u pravcu $-\nabla f(x^o)$. Dužina koraka α_0 se nalazi rešavanjem problema jednodimenzionalne optimizacije

$$min\, f(x^0 - \alpha \cdot \nabla f(x^o)). \qquad (8)$$

Funkcija $f(x^0 - \alpha \cdot \nabla f(x^o))$ je funkcija jedne promenljive α, pa se može označiti sa $f_0(\alpha) = f(x^0 - \alpha \cdot \nabla f(x^o))$. Odavde sledi, da se problem optimizacije (8) može zapisati u sledećem obliku

$$\min_{\alpha > 0} f_0(\alpha).$$

Za rešavanje ovog problema se koristi neka od metoda optimizacije funkcija sa jednom promenljivom.

Kada se reši problem (8) dobije se neko lokalno optimalno rešenje $\alpha = \alpha_0$, koje predstavlja dužinu početnog koraka u smeru $-\nabla f(x^o)$. Tada se naredna iteracija x^1 dobija na sledeći način

$$x^1 = x^0 - \alpha_0 \cdot \nabla f(x^o),$$

nakon čega se postupak ponavlja, tj. iterativno pravilo glasi

$$x^{k+1} = x^k - \alpha_k \cdot \nabla f(x^k),$$

gde je α_k rešenje jednodimenzionalnog problema

$$\min_{\alpha > 0} f_k(\alpha), \qquad (9)$$

pri čemu je $f_k(\alpha) = f(x^k - \alpha \cdot \nabla f(x^k))$.

Potrebno je napomenuti nekoliko činjenica.

1. Smer gradijenta se ne menja ako se gradijent normalizuje, tj.ako se umesto vektora $\nabla f(x^k)$ koristi vektor

$$u^k = \frac{(\nabla f(x^k))}{\|(\nabla f(x^k))\|}, \quad k = 0,1,2,\ldots$$

Vektor u^k ima jediničnu dužinu, tj.$\|u^k\| = 1,\ k = 0,1,2,\ldots$, što doprinosi stabilnosti metode.

2. Kada se rešava jednodimenzionalni problem (9), može se desiti da u nekoj k −toj iteraciji funkcija $f_k(\alpha)$ ima nekoliko lokalnih minimuma. U tom slučaju poželjno je unapred odrediti koji od njih će se sačuvati.

3. Metoda najbržeg pada u početku vrlo dobro napreduje prema optimalnom rešenju, ali se brzo uspori. Ovo usporavanje se može otkriti pomoću nekog kriterijuma, kao npr.

$$\|\alpha_k \cdot u^k\| \leq \|K \cdot \alpha_0 \cdot u^0\|, \tag{10}$$

gde su vektori u^k normalizovani, $\alpha_k > 0$ i broj K je konstanta $(0 < K < 1)$ koja se zadaje unapred. Zbog ovih osobina, kriterijum (10) je analogan sledećem kriterijumu

$$\alpha_k \leq K \cdot \alpha_0, \tag{11}$$

koji se može koristiti upravo kao izlazni kriterijum. Najčešće korišćeni kriterijum zaustavljanja je

$$\|\nabla f(x^k)\| < \varepsilon,$$

gde je $\varepsilon > 0$ unapred zadat broj.

4. Važna osobina metoda najbržeg pada jeste da su vektori gradijenata, koji su izračunati u dve proizvoljne uzastopne aproksimacije, međusobno ortogonalni, tj.

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

Algoritam 1. Metoda najbržeg pada (Košijeva metoda)

Korak 1. Izabrati početnu tačku $x^0 \in R^n$, staviti $k = 0$. Zadati $\varepsilon > 0$ za izlazni kriterijum.

Korak 2. Ako je $\|\nabla f(x^k)\| < \varepsilon$ tada je $x^k = x^*$ rešenje. Inače, preći na korak 3.

Korak 3. Izračunati korak α_k. Formira se jednodimenzionalni problem

$$f_k(\alpha) = f(x^k - \alpha \cdot \nabla f(x^k)),$$

pa se α_k dobija kao rešenje problema

$$\min_{\alpha > 0} f_k(\alpha).$$

Korak 4. Sledeću iteraciju x^{k+1} odrediti kao

$$x^{k+1} = x^k - \alpha_k \cdot \nabla f(x^k)$$

zameniti k sa $k + 1$ i preći na Korak 2.

Naredne teoreme odnose se na konvergenciju metode najbržeg pada.

Ova metoda globalno konvergira ka stacionarnoj tački funkcije $f(x)$ pod blagim uslovima.

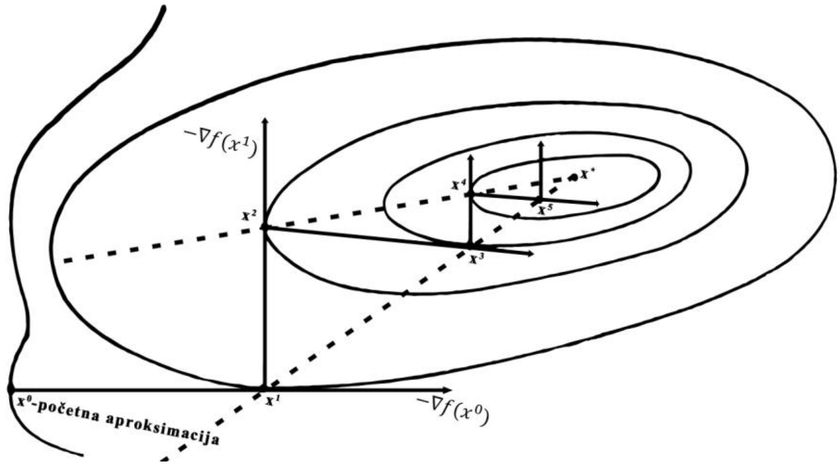
Teorema 9. [8] Neka je $f(x) \in C^1(R^n)$, neka je skup $X_0 = \{x \in R^n | f(x) \leq f(x^0)\}$ ograničen i neka je niz $\{x^k\}$ generisan metodom najbržeg pada. Tada važi:

1. niz $\{f(x^k)\}$ je monotono opadajući niz, ako je $\nabla f(x^k) \neq 0$,

2. ili je $\nabla f(x^k) = 0$ u svakoj iteraciji ili je svaka tačka nagomilavanja x^* niza $\{x^k\}$ stacionarna tačka funkcije f, tj. niz generisan metodom najbržeg pada konvergira ka stacionarnoj tački i

$$\lim_{k \to \infty} \nabla f(x^k) = \nabla f(x^*) = 0.$$

Konvergencija metode najbržeg pada za funkciju sa dve promenljive i elipsastim presecima (posle prve aproksimacije) data je na slici 1.



Slika 1. Konvergencija metode najbržeg pada

Teorema 10. [2] Neka je $f \epsilon C^2(R^n)$ i neka je x^* lokalni minimum od f. Osim toga neka je $\bar{a} > 0$ najmanji, a $\bar{A} > 0$ najveći karakteristični koren matrice Hesijana $\nabla^2 f(x^*)$. Ako niz $\{x^k\}$ generisan metodom najbržeg pada konvergira ka x^* tada niz $\{f(x^k)\}$ konvergira linearno ka $f(x^*)$ sa koeficijentom brzine ne većim od $(\frac{\bar{A}-\bar{a}}{\bar{A}+\bar{a}})^2$.

Prethodna teorema ukazuje da odnos $r = \frac{\bar{A}}{\bar{a}}$ direktno utiče na brzinu metode najbržeg pada. Ako je r veliko matrica je loše uslovljena i konvergencija se jako usporava.

Metoda najbržeg pada je pouzdana i laka za implementaciju, ali relativno sporo konvergira. Mana ove metode je što se u njoj najviše vremena obično potroši na jednodimenzionalnu optimizaciju kojom se računa optimalna dužina koraka α_k.

Primer 1. Rešiti problem optimizacije

$$\min f(x) = 6{x_1}^2 + {x_2}^2 + x_1 \cdot x_2 - 2x_1 - 3x_2$$

metodom najbržeg pada iz početne tačke $x^0 = [1,1]^T$.

Rešenje:

Neka je kriterijum za zaustavljanje $\|\nabla f(x^k)\| \leq 0.00001$.

Gradijent funkcije f je

$$\nabla f(x) = \begin{bmatrix} 12x_1 + x_2 - 2 \\ 2x_2 + x_1 - 3 \end{bmatrix},$$

a vrednost gradijenta u početnoj tački $x^0 = [1,1]^T$ je

$$\nabla f(x^0) = \begin{bmatrix} 11 \\ 0 \end{bmatrix}.$$

Kako je $\|\nabla f(x^0)\| \neq 0$, sledi da $x^0 = [1,1]^T$ nije rešenje problema. Da bi se izračunala sledeća iteracija, potrebno je odrediti korak α_0. Formira se jednodimenzionalni problem

$$\min f_0(\alpha),$$

gde je

$$f_0(\alpha) = f(x^0 - \alpha \cdot \nabla f(x^0)) = f([1,1]^T - \alpha \cdot [11,0]^T)$$

$$f_0(\alpha) = f\left(\begin{bmatrix} 1 - 11\alpha \\ 1 \end{bmatrix}\right) =$$

$$6 \cdot (1 - 11\alpha)^2 + 1 + (1 - 11\alpha) \cdot 1 - 2 \cdot (1 - 11\alpha) - 3 \cdot 1.$$

Korak α_0 se dobija kao rešenje ovog jednodimenzionalnog problema i sledi $\alpha_0 = \frac{1}{12}$.

Sada se računa sledeća iteracija

$$x^1 = x^0 - \alpha_0 \cdot (\nabla f(x^o))$$

pa je

$$x^1 = \begin{bmatrix} 1/12 \\ 1 \end{bmatrix}.$$

Dalje je

$$\nabla f(x^1) = \begin{bmatrix} 0 \\ -11/12 \end{bmatrix},$$

a pošto je $\|\nabla f(x^1)\| \neq 0$, x^1 nije rešenje i postupak se nastavlja. Na analogan način dobijaju se ostali članovi iterativnog niza

$$x^2 = \begin{bmatrix} 1/_{12} \\ 35/_{24} \end{bmatrix}$$

$$x^3 = \begin{bmatrix} 13/_{288} \\ 35/_{24} \end{bmatrix}$$

$$x^4 = \begin{bmatrix} 13/_{288} \\ 851/_{576} \end{bmatrix}$$

$$x^5 = \begin{bmatrix} 301/_{6912} \\ 851/_{576} \end{bmatrix}$$

$$x^6 = \begin{bmatrix} 0.0435403 \\ 1.47769 \end{bmatrix}.$$

Primer 2. Naći minimum funkcije

$$f(x) = {x_1}^2 - x_1 x_2 - 4x_1 + {x_2}^2 - x_2$$

metodom najbržeg pada iz početne tačke $x^0 = [0,0]^T$.

Rešenje:

Zadatak je rešenjen u programskom paketu Matlab. Naredni M-file-ovi koriste se za određivanje minimuma ove funkcije.

Prvi M-file se odnosi na računanje gradijenta funkcije f:

```
function g= grad(f,x,h)
%racuna gradijent funkcije f(x)
if nargin<3, h=.00001; end
h2= 2*h; N= length(x);
I= eye(N);
for n=1:N
g(n)= (feval(f,x+I(n,:)*h)-feval(f,x-I(n,:)*h))/h2;
end
```

Drugi M-file se odnosi na rešvanje datog problema pomoću metode najbržeg pada.

```
function [xo,fo] = opt_steep(f,x0,TolX,TolFun,alpha0,MaxIter)
%za minimizaciju funkcije koristi se metoda najbrzeg pada
if nargin < 6, MaxIter = 100; end %maksimalan broj iteracija
```

```
if nargin < 5, alpha0 = 10; end %pocetna duzina koraka
if nargin < 4, TolFun = 1e-8; end %|f(x)| < TolFun
if nargin < 3, TolX = 1e-6; end %|x(k)- x(k - 1)|<TolX
x = x0; fx0 = feval(f,x0); fx = fx0;
alpha = alpha0; kmax1 = 25;
warning = 0;
for k = 1: MaxIter
g = grad(f,x); g = g/norm(g);
alpha = alpha*2; %probni korak u smeru negativnog gradijenta
fx1 = feval(f,x - alpha*2*g);
for k1 = 1:kmax1 %nalazi se optimalna duzina koraka
fx2 = fx1; fx1 = feval(f,x-alpha*g);
if fx0 > fx1+TolFun && fx1 < fx2 - TolFun %fx0 > fx1 < fx2
den = 4*fx1 - 2*fx0 - 2*fx2; num = den - fx0 + fx2;
alpha = alpha*num/den;
x = x - alpha*g
fx = feval(f,x)
break;
else alpha = alpha/2;
end
end
if k1 <= kmax1, warning = warning + 1;
%nije se nasla optimalna duzina koraka
else warning = 0;
end
if warning >= norm((x - x0),2) < TolX&&abs((fx - fx0) < TolFun);
end
x0 = x; fx0 = fx;
end
xo = x;
fo = fx;
if k1 == MaxIter, fprintf('Just best in %d iterations',MaxIter)
end
```

Kada se program izvrši u Matlab prozoru, dobija se:

```
>> f = inline('x(1)*(x(1) - 4 - x(2)) + x(2)*(x(2)- 1)','x');
>> x0 = [0 0], TolX = 1e-6; TolFun = 1e-8; alpha0 = 1; MaxIter = 100;
>> [xo,fo] = opt_steep(f,x0,TolX,TolFun,alpha0,MaxIter)

xo =

    3.0000    1.9999


fo =

   -7.0000
```

Dakle, minimum funkcije f je u tački $x^0 = [3,2]^T$ i vrednost funkcije u tački minimuma je $f^* = -7$. U sledećoj tabeli su navedene sve iteracije dobijene metodom najbržeg pada.

k	$\boldsymbol{x_1^k}$	$\boldsymbol{x_2^k}$	$\boldsymbol{f(x^k)}$
0	0	0	0
1	2.6154	0.6538	-5.5577
2	2.3819	1.5879	-6.7028
3	2.9208	1.7226	-6.9388
4	2.8726	1.9151	-6.9874
5	2.9837	1.9429	-6.9974
6	2.9738	1.9825	-6.9995
7	2.9966	1.9882	-6.9999
8	2.9946	1.9964	-7.0000
9	2.9993	1.9976	-7.0000
10	2.9989	1.9993	-7.0000
11	2.9999	1.9995	-7.0000
12	2.9998	1.9998	-7.0000
13	3.0000	1.9999	-7.0000

Tabela 1. Vrednosti iteracija dobijenih metodom najbržeg pada

Kao što je rečeno, najviše vremena kod ove metode se potroši na nalaženje optimalne dužine koraka α_k. Košijeva metoda se može pojednostaviti ukoliko se dužina koraka fiksira, tj. ako je α_k isto za svako $k = 0,1,2,\dots$ i na taj način se dobija modifikovana Košijeva metoda koja je jednostavnija za primenu, gde se aprokimacije računaju po sledećem pravilu

$$x^{k+1} = x^k - \alpha \cdot \nabla f(x^k), \qquad k = 0,1,2,\dots$$

Da bi modifikovana Košijeva metoda globalo konvergirala, potrebno je da funkcija f bude dva puta neprekidno diferencijabilna i strogo konveksna. Korak α treba da bude dovoljno mali, naprimer

$$0 < \alpha < \frac{2}{\bar{A}},$$

gde je $\bar{A} > 0$ najveći karakteristični koren Hesijana funkcije f u rešenju x^*. Modifikovana Košijeva metoda konvergira sporije od originalne Košijeve metode.

3.2 Njutnova metoda

Njutnova metoda je jedna od najpoznatijih metoda za minimizaciju funkcija više promenljivih bez ograničenja. Pored prvog, koristi se i drugi izvod funkcije cilja (Hesijan), čime se postiže brža konvergencija prema rešenju, zbog čega je potrebno da funkcija cilja ima neprekidne parcijalne izvode drugog reda.

Ideja metode je da se funkcija cilja aproksimira sa prva dva člana svog razvoja u Tejlorov red, pa pošto kvadratni deo razlaganja bolje aproksimira funkciju nego linearni, postupak koji se bazira na razvoju drugog reda konvergira brže nego postupak koji koristi razvoj prvog reda.

Posmatra se problem

$$\min_{x \in R^n} f(x), \tag{12}$$

gde je $f(x) \in C^2(R^n)$. Neka je $x^0 \in R^n$ početna iteracija. Ako je poznata $k-$ta aproksimacija x^k, tada se priraštaj funkcije $f(x)$ u tački x^k može predstaviti na sledeći način

$$f(x) = f(x^k) + \langle \nabla f(x^k), x - x^k \rangle + \frac{1}{2}\langle \nabla^2 f(x^k)(x - x^k), x - x^k \rangle + o\left(\|x - x^k\|^2\right).$$

Koristeći kvadratni deo ovog priraštaja

$$f_k(x) = f(x^k) + \langle \nabla f(x^k), x - x^k \rangle + \frac{1}{2}\langle \nabla^2 f(x^k)(x - x^k), x - x^k \rangle$$

naredna iteracija x^{k+1} se određuje kao rešenje problema

$$\min_{x \in R^n} f_k(x),$$

tj. rešavanjem jednačine $\nabla f_k(x) = 0$.

$$\nabla f_k(x) = \nabla f(x^k) + \nabla^2 f(x^k)(x - x^k) = 0$$
$$\nabla^2 f(x^k)(x - x^k) = -\nabla f(x^k).$$

Pod pretpostavkom da je Hesijan $\nabla^2 f(x^k)$ regularna matrica, množenjem sa leve strane sa $(\nabla^2 f(x^k))^{-1}$ dobija se

$$x - x^k = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k).$$

Odavde sledi opšte iterativno pravilo Njutnovog postupka

$$x^{k+1} = x^k - (\nabla^2 f(x^k))^{-1} \nabla f(x^k), \quad k = 0,1,2, \dots \tag{13}$$

Kada su u pitanju problemi velikih dimenzija, tj. kad je n veliko, izračunavanje inverznog Hesijana $(\nabla^2 f(x^k))^{-1}$ je skupo, pa se tada naredna iteracija x^{k+1} dobija na sledeći način

$$\left(\nabla^2 f(x^k)\right)s^k = -\nabla f(x^k)$$
$$x^{k+1} = x^k + s^k, \qquad k = 0,1,2,\ldots \tag{14}$$

Dakle, u svakoj iteraciji rešava se sistem linearnih jednačina i s^k se dobija kao rešenje tog sistema, i na taj način se izbegava izračunavanje inverznog Hesijana.

Sledi algoritam za Njutnovu metodu.

Algoritam 2. Njutnova metoda

Korak 1. Izabrati početnu tačku $x^0 \in R^n$, staviti $k = 0$. Zadati $\varepsilon > 0$ za izlazni kriterijum.

Korak 2. Ako je $\|\nabla f(x^k)\| < \varepsilon$ tada je $x^k = x^*$ rešenje. Inače, preći na korak 3.

Korak 3. Izračunati sledeću iteraciju: $x^{k+1} = x^k - (\nabla^2 f(x^k))^{-1}\nabla f(x^k)$. Staviti $k = k + 1$, i preći na korak 2.

Naredne teoreme odnose se na konvergenciju Njutnovog postupka.

Teorema 11. [7] Neka je tačka x^0 takva da je skup $X_0 = \{x \epsilon R^n | f(x) \leq f(x^0)\}$ ograničen. Pretpostavimo da je funkcija *f* dva puta neprekidno diferencijabilna na X_0 i da je matrica $\nabla^2 f(x)$ pozitivno definitna za $x \epsilon X_0$. Neka je $\{x^k\}$ niz generisan Njutnovim postupkom. Tada $\{x^k\}$ konvergira ka jedinstvenom rešenju x^* problema (12).

Ako važi dodatna pretpostavka da je f tri puta neprekidno diferencijabilna, može se pokazati da je brzina konvergencije kvadratna.

Teorema 12. [2] Neka je $f \epsilon C^3(R^n)$, x^* je lokalni minimumum funkcije f i neka je Hesijan $\nabla^2 f(x^*)$ pozitivno definitna matrica. Tada postoji okolina tačke x^* takva da za x^0 iz te okoline, niz generisan Njutnovim postupkom kvadratno konvergira ka x^*, tj. važi ocena

$$\|x^{k+1} - x^*\| \leq c\|x^k - x^*\|^2.$$

Dokaz:

Neka postoje $\rho > 0, \beta_1 > 0, \beta_2 > 0$ tako da za svako x za koje je

$$\|x - x^*\| < \rho$$

važi

$$\|\nabla^2 f(x)^{-1}\| < \beta_1$$

i

$$\|\nabla f(x^*)^T - \nabla f(x)^T - \nabla^2 f(x)(x^* - x)\| < \beta_2 \|x - x^*\|^2.$$

Pretpostavimo da je x^k odabrano tako da važi

$$\beta_1 \beta_2 \|x^k - x^*\| < 1$$

i

$$\|x^k - x^*\| < \rho.$$

Tada je

$$\begin{aligned}
\|x^{k+1} - x^*\| &= \|x^k - x^* - \nabla^2 f(x^k)^{-1} \nabla f(x^k)^T\| \\
&= \|\nabla^2 f(x^k)^{-1}[\nabla f(x^*)^T - \nabla f(x^k)^T - \nabla^2 f(x^k)(x^* - x^k)]\| \\
&\leq \|\nabla^2 f(x^k)^{-1}\| \beta_2 \|x^k - x^*\|^2 \\
&\leq \beta_1 \beta_2 \|x^k - x^*\|^2 \\
&< \|x^k - x^*\|.
\end{aligned}$$

Poslednja nejednakost pokazuje da je nova tačka bliža rešenju x^* nego prethodna tačka, a pretposlednja nejednakost pokazuje da je red konvergencije dva, tj. da je Njutnov metod kvadratno kovergentan. Dakle, Njutnova metoda je lokalno kvadratno konvergentna.

U slučaju kada funkcija cilja ima oblik kvadratne forme

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c,$$

gde je A pozitivno definitna matrica, $A \in R^{n,n}, b \in R^n, c \in R$ do rešenja se dolazi već u prvoj iteraciji za svaku početnu aproksimaciju.

Dokaz za ovo tvrđenje je vrlo jednostavan. Naime, kako je

$$\nabla f(x) = Ax + b \quad \text{i} \quad \nabla^2 f(x) = A,$$

kada se ovo zameni u (13) prva iteracija u Njutnovoj metodi je

$$x^1 = x^0 - A^{-1}(Ax^0 + b) = -A^{-1}b.$$

Kako je

$$\nabla f(x^1) = Ax^1 + b = -AA^{-1}b + b = 0.$$

Sledi da je rešenje $x^* = x^1 = -A^{-1}b$.

Primer 3. Njutnovom metodom odrediti minimum funkcije

$$f(x) = {x_1}^4 + {x_2}^2 - 2x_1x_2 + 1$$

iz početne tačke $x^0 = [1,1]^T$, gde je $x = [x_1, x_2]^T \epsilon R^2$, $\varepsilon = 10^{-3}$.

Rešenje:

Gradijent funkcije $f(x)$ je oblika

$$\nabla f(x) = \begin{bmatrix} 4{x_1}^3 - 2x_2 \\ 2x_2 - 2x_1 \end{bmatrix},$$

a Hesijan

$$\nabla^2 f(x) = \begin{bmatrix} 12{x_1}^2 & -2 \\ -2 & 2 \end{bmatrix}.$$

Pošto je $\|\nabla f(x^0)\| \geq 10^{-3}$, dakle $x^0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ nije rešenje.

Prva iteracija je

$$x^1 = x^0 - (\nabla^2 f(x^0))^{-1} \nabla f(x^0),$$

gde je

$$\nabla^2 f(x^0) = \begin{bmatrix} 12 & -2 \\ -2 & 2 \end{bmatrix} \quad \text{i} \quad (\nabla^2 f(x^0))^{-1} = \frac{1}{20}\begin{bmatrix} 2 & 2 \\ 2 & 12 \end{bmatrix}.$$

Odavde sledi

$$x^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{1}{20}\begin{bmatrix} 2 & 2 \\ 2 & 12 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4/5 \\ 4/5 \end{bmatrix}.$$

Kako je $\nabla f(x^1) = \begin{bmatrix} 56/125 \\ 0 \end{bmatrix}$, sledi $\|\nabla f(x^1)\| \geq 10^{-3}$. Dakle, $x^1 = \begin{bmatrix} 4/5 \\ 4/5 \end{bmatrix}$ nije rešenje.

Druga iteracija je

$$x^2 = x^1 - (\nabla^2 f(x^1))^{-1} \nabla f(x^1),$$

pri čemu je

$$\nabla^2 f(x^1) = \begin{bmatrix} 192/25 & -2 \\ -2 & 2 \end{bmatrix} \quad \text{i} \quad (\nabla^2 f(x^1))^{-1} = \frac{25}{284}\begin{bmatrix} 2 & 2 \\ 2 & 192/25 \end{bmatrix}.$$

Odavde sledi

$$x^2 = \begin{bmatrix} 4/5 \\ 4/5 \end{bmatrix} - \frac{25}{284}\begin{bmatrix} 2 & 2 \\ 2 & 192/25 \end{bmatrix}\begin{bmatrix} 56/125 \\ 0 \end{bmatrix} = \begin{bmatrix} 256/355 \\ 256/355 \end{bmatrix} = \begin{bmatrix} 0.721126 \\ 0.721126 \end{bmatrix}.$$

Sada je $\|\nabla f(x^2)\| = 0.057755581 \geq 10^{-3}$, pa odavde sledi da x^2 nije rešenje problema.

Treća iteracija je

$$x^3 = x^2 - (\nabla^2 f(x^2))^{-1}\nabla f(x^2),$$

$$x^3 = \begin{bmatrix} 0.707505277 \\ 0.707505277 \end{bmatrix}.$$

Kako je $\|\nabla f(x^3)\| = 0.001595331 < 10^{-3}$, sledi da je

$$x^3 = \begin{bmatrix} 0.707505277 \\ 0.707505277 \end{bmatrix}$$

rešenje problema.

Primer 4. Naći minimum funkcije

$$f(x) = {x_1}^4 + {x_1}^3 - x_1 + {x_2}^4 - {x_2}^2 + x_2 + {x_3}^2 - x_3 + x_1x_2x_3$$

Njutnovom metodom iz početne tačke $x^0 = [1, -1, 1]^T$ gde je $x = [x_1, x_2, x_3]^T \in R^3$, $\varepsilon = 10^{-5}$.

Rešenje:

Zadatak je rešen u programskom paketu Matlab. Formirana su tri M-file-a; prvi za funkciju f, drugi za gradijent funkcije f i treći za Hesijan funkcije f.

M-file za funkciju f:

```
function res =f2(x)
res=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);
```

M-file za gradijent funkcije f:

```
function res=gradf2(x)
res=[4*(x(1)^3)+3*(x(1)^2)+x(2)*x(3)-1;4*(x(2)^3)-
2*x(2)+x(1)*x(3)+1;2*x(3)+x(1)*x(2)-1];
```

M-file za Hesijan funkcije f:

```
function res=hesianf2(x)
res=[12*(x(1)^2)+6*x(1)  x(3)  x(2);x(3)  12*(x(2)^2)-2  x(1);x(2)  x(1)  2];
```

Zatim se pravi M-file, kojim se određuje minimum funkcije f Njutnovom metodom.

```
function res=njutn2(x0,eps)
k=1;
xk=x0;
xk1=xk-hesianf2(xk)^(-1)*gradf2(xk);
vrednostf=f2(xk1);
while norm(gradf2(xk),'inf')>eps
    xk=xk1;
    xk1=xk-hesianf2(xk)^(-1)*gradf2(xk);
    vrednostf=f2(xk1);
k=k+1;
end
res=xk1;
k
vrednostf
```

Potom se program izvršava u Matlab prozoru:

```
>> njutn2([1;-1;1],0.0001)
k =
    5

ans =
   0.5709
  -0.9396
   0.7682
```

Odavde sledi da je tačka minimuma $x^* = [0.5709, -0.9396, 0.7682]^T$ i ona je nađena u petoj iteraciji. Vrednost funkcije cilja u optimalnoj tački je $f^* = -1.9118$. U sledećoj tabeli su navedene sve iteracije dobijene Njutnovom metodom.

k	x_1^k	x_2^k	x_3^k	$f(x^k)$
0	1	-1	1	-1
1	0.7104	-0.9543	0.8323	-1.8367
2	0.5926	-0.9417	0.7782	-1.9101
3	0.5715	-0.9396	0.7685	-1.9118
4	0.5709	-0.9396	0.7682	-1.9118
5	0.5709	-0.9396	0.7682	-1.9118

Tabela 2. Vrednosti iteracija dobijenih Njutnovim postupkom

Nedostatak Njutnove metode u slučaju velikog broja promenljivih i komplikovanijih oblika funkcije f, je da zahteva mnogo računskih operacija te je skupa. Naime, u svakoj iteraciji potrebno je izračunati n^2 elemenata matrice Hesijana i rešiti sistem linearnih jednačina, što je skupo kad je reč o problemima velikih dimenzija. Osim toga, Njutnova metoda je lokalno konvergentna, pa zahteva modifikacije u slučaju kada je početna iteracija daleko od rešenja. Jedna od modifikacija je oblika

$$x^{k+1} = x^k - \alpha_k \cdot (\nabla^2 f(x^k))^{-1} \nabla f(x^k), \ \ k = 0,1,2,\ldots$$

gde se korak α_k određuje minimizacijom $f(x^{k+1})$. U blizini rešenja očekuje se $\alpha_k \approx 1$. Ostale modifikacije Njutnovog postupka mogu se podeliti u dve grupe i to su kvazi-Njutnovi postupci i netačni Njutnovi postupci.

3.3 Kvazi-Njutnovi postupci

Za rešavanje problema optimizacije bez ograničenja često se koristi metod koji spada u grupu kvazi-Njutnovih postupaka i zove se modifikovani Njutnov metod. Iterativno pravilo ovog postupka je

$$x^{k+1} = x^k - \alpha_k S_k \nabla f(x^k), \;\; k = 0,1,2,\ldots$$

gde je S_k simetrična matrica dimenzije $n \times n$, a α_k se određuje kao minimum od $f(x^{k+1})$. U slučaju da je matrica S_k baš inverzni Hesijan i $\alpha_k = 1$, u pitanju je Njutnov metod, a ukoliko je $S_k = E$ dobija se metod najbržeg pada. Kao i ranije, potrebno je pretpostaviti da je matrica S_k pozitivno definitna i ona treba da bude dovoljno dobra aproksimacija inverznog Hesijana.

Za dokaz konvergencije ovog postupka, radi jednostavnosti, posmatran je specijalan slučaj kada je funkcija cilja kvadratna

$$f(x) = \frac{1}{2} x^T Q x - b^T x$$

gde je Q simetrična, pozitivno definitna matrica i $b \epsilon R^n$. Tada je

$$x^{k+1} = x^k - \alpha_k S_k \nabla f(x^k), \;\; k = 0,1,2,\ldots \tag{15a}$$

gde je

$$\nabla f(x^k) = Q x^k - b. \tag{15b}$$

Pošto je $f(x)$ kvadratna funkcija, α_k se može eksplicitno izraziti kao

$$\alpha_k = \frac{\nabla f(x^k)^T S_k \nabla f(x^k)}{\nabla f(x^k)^T S_k Q S_k \nabla f(x^k)}. \tag{15c}$$

Sledi teorema o konvergenciji u slučaju kada je f kvadratna funkcija.

Teorema 13. [2] Neka je x^* jedinstveno rešenje problema (12) i neka je niz $\{x^k\}$ generisan modifikovanim Njutnovim postupkom. Definiše se funkcija $E(x) = \frac{1}{2}(x - x^*)^T Q (x - x^*)$. Tada za svako $k = 0, 1, 2, \ldots$ važi

$$E(x^{k+1}) \leq \left(\frac{B_k - b_k}{B_k + b_k}\right)^2 E(x^k), \tag{16}$$

gde su b_k i B_k redom najmanji i najveći karakteristični koreni matrice $S_k Q$.

Dokaz:

Direktnom zamenom sledi

$$\frac{E(x^k) - E(x^{k+1})}{E(x^k)} = \frac{(\nabla f(x^k)^T S_k \nabla f(x^k))^2}{\left(\nabla f(x^k)^T S_k Q S_k \nabla f(x^k)\right)(\nabla f(x^k)^T Q^{-1} \nabla f(x^k))}.$$

Ako se stavi da je

$$T_k = S_k^{\frac{1}{2}} Q S_k^{\frac{1}{2}} \qquad \text{i} \qquad p_k = S_k^{\frac{1}{2}} \nabla f(x^k)$$

tada je

$$\frac{E(x^k) - E(x^{k+1})}{E(x^k)} = \frac{(p_k{}^T p_k)^2}{(p_k{}^T T_k p_k)(p_k{}^T T_k{}^{-1} p_k)}.$$

Iz Kantorovičeve nejednakosti sledi

$$E(x^{k+1}) \leq \left(\frac{B_k - b_k}{B_k + b_k}\right)^2 E(x^k)$$

gde su b_k i B_k redom najmanji i najveći karakteristični koreni matrice T_k. Kako je

$$S_k^{\frac{1}{2}} T_k S_k{}^{-\frac{1}{2}} = S_k Q$$

sledi da je matrica $S_k Q$ slična matrici T_k te ima iste karakteristične korene.

Iz ove teoreme se vidi, da ako je matrica S_k dovoljno dobra aproksimacija za Q^{-1}, onda će se najmanji i najveći karakteristični koreni malo razlikovati pa će konvergencija biti brza. U slučaju kada funkcija cilja nije kvadratna, priča je analogna. Dakle matrica S_k treba da bude dovoljno dobra aproksimacija za inverzni Hesijan $\nabla^2 f(x^k)^{-1}$, da bi postupak brže konvergirao.

Kantorovičeva nejednakost

Neka je Q pozitivno definitna, simetrična matrica dimenzije $n \times n$. Tada za bilo koji vektor x važi

$$\frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} \geq \frac{4\bar{a}\bar{A}}{(\bar{a} + \bar{A})^2}$$

gde su $\bar{a}$ najmanji i $\bar{A}$ najveći karakteristični koreni matrice Q.

U grupu kvazi-Njutnovih postupaka spada i metod koji kombinuje najbolja svojstva metode najbržeg pada i Njutnove metode. Ideja je da se u početku primenjuje metoda najbržeg pada koja je globalnog karaktera, a kad se dođe blizu rešenja da se primeni Njutnova metoda koja je brza. Neophodna pretpostavka za primenu ove metode jeste da je $f \epsilon C^1(R^n)$.

Na početku se počinje Košijevom metodom najbržeg pada, što daje velike početne skokove

$$x^{k+1} = x^k - \alpha_k \cdot \nabla f(x^k),$$

a završava se Njutnovom metodom sa $\alpha_k \approx 1$, što će garantovati kvadratno završavanje

$$x^{k+1} = x^k - \alpha_k \cdot \left(\nabla^2 f\left(x^k\right)\right)^{-1} \nabla f\left(x^k\right).$$

Ovde se umesto Hesijana koristi neka njegova aproksimacija H_k, tj.

$$x^{k+1} = x^k - \alpha_k \cdot H_k \cdot \nabla f(x^k), \qquad k = 0,1,2,\dots$$

Na početku metode se stavlja da je $H_0 = E$, tj. H_0 je jedinična matrica. Tada je

$$x^1 = x^0 - \alpha_0 \cdot \nabla f(x^0),$$

što se poklapa sa Košijevom metodom.

Kao i ranije, α_k, $k = 0,1,2,\dots$ se nalazi rešavanjem problema jednodimenzionalne optimizacije funkcije f iz tačke x^k u smeru

$$d^k = -H_k \cdot \nabla f(x^k), \;\; k = 0,1,2,\dots$$

Kako bi metoda imala kvadratno završavanje, potrebno je da u svakom koraku matrica H_k, $k = 0,1,2,\dots$ bude „upotpunjena“ tako da dobro aproksimira inverzni Hesijan

$$(\nabla^2 f(x^k))^{-1}$$

kada je x^k blizu tačke optimuma.

U slučaju kada funkcija cilja ima oblik kvadratne forme

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c,$$

gde je A pozitivno definitna matrica, posle n iteracija matrica H_n treba da bude jednaka inverznom Hesijanu funkcije f, tj. $H_n = A^{-1}$.

Postoje različiti načini „upotpunjavanja“ matrice H_k, a ovde je naveden jedan od njih.

Matrica H_k se upotpunjuje pomoću dve matrice A_k i B_k tako da važi

$$H_{k+1} = H_k + A_k + B_k, \qquad k = 0,1,2,...$$

gde matrice A_k i B_k imaju rang jednak jedinici, što znači da u opštem slučaju matrica $A_k + B_k$ ima rang jednak broju dva. Ova metoda se naziva Davidon-Fletcher-Powell (DFP) metoda.

Algoritam 3. DFP metoda

Korak 1. Izabrati početnu aproksimaciju i početnu pozitivno definitnu matricu H_0 (često se uzima $H_0 = E$). Izračunati gradijent funkcije f u tački x^0, tj. $g^0 = \nabla f(x^0)$. Odrediti kriterijum za zaustavljanje, npr. sa nekim malim $K > 1$. Staviti $k = 0$.

Korak 2. Izračunati smer pretraživanja funkcije f iz tačke x^k

$$d^k = -H_k g^k.$$

Korak 3. Naći $\alpha_k > 0$ kao rešanje problema jednodimenzionalne optimizacije u smeru d^k

$$\min f(x^k + \alpha d^k).$$

Korak 4. Izračunati

$$z^k = \alpha_k d^k$$

i novu aproksimaciju

$$x^{k+1} = x^k + z^k.$$

Ako je $\alpha_k \leq K \cdot \alpha_0$ algoritam se zaustavlja i optimalno rešenje je $x^* = x^{k+1}$. Ako je $\alpha_k > K \cdot \alpha_0$ nastaviti sa algoritmom.

Korak 5. Matrica H_k se upotpunjava na sledeći način:

$$g^{k+1} = \nabla f(x^{k+1})$$

$$w^k = g^{k+1} - g^k$$

$$A_k = \frac{1}{\langle z^k, w^k \rangle}(z^k)(z^k)^T$$

$$B_k = -\frac{1}{\langle w^k, H_k w^k \rangle} H_k w^k (w^k)^T H_k$$

$$H_{k+1} = H_k + A_k + B_k.$$

Staviti $k = k + 1$ i vratiti se na Korak 2.

Potrebno je napomenuti da se vektor d^k u koraku 2. može normalizovati, tj. za d^k se može uzeti

$$d^k = -\frac{H_k g^k}{\|H_k g^k\|}.$$

Ova normalizacija doprinosi stabilnosti metode. Takođe, za gradijente izračunate u dve uzastopne iteracije važi relacija „ortogonalnosti“

$$\langle g^{k+1}, H_k g^k \rangle = 0, \qquad k = 0,1,2,\ldots$$

Lema 3. [8] U DFP metodi svaka matrica H_k, izračunata u Algoritmu 4 je pozitivno definitna.

Dokaz:

Dokaz se izvodi indukcijom. Po pretpostavci, početno H_0 je pozitivno definitno. Pretpostavlja se da je i H_k za neko $k > 0$ pozitivno definitno. Treba pokazati da iz ovoga sledi da je i H_{k+1} pozitivno definitna matrica.

Kako su H_k, A_k, B_k simetrične matrice, sledi da je i njihov zbir H_{k+1} simetrična matrica. Treba još pokazati da je

$$\langle \xi, H_{k+1}\, \xi \rangle > 0, \qquad \text{za svako } \xi \neq 0.$$

Iz definicije d^k i g^k iz Algoritma 4 sledi

$$\nabla f(x^k) d^k = -\nabla f(x^k) H_k g^k = -\langle g^k, H_k g^k \rangle < 0,$$

što znači da je d^k smer u kome funkcija f opada iz tačke x^k i zbog toga je $\alpha_k > 0$. Kako je matrica H_k po pretpostavci pozitivno definitna, sledi

$$H_k = S_k \cdot S_k, \tag{17}$$

gde je S_k neka pozitivno definitna matrica. Neka za neki proizvoljan vektor $\xi \neq 0$, važi

$$p = S_k \cdot \xi \quad \text{ i } \quad q = S_k \cdot w^k.$$

Sada je

$$\langle \xi, H_{k+1}\xi \rangle = \langle \xi, H_k \xi \rangle + \frac{\langle z^k, \xi \rangle^2}{\langle z^k, w^k \rangle} - \frac{\langle \xi, H_k w^k \rangle^2}{\langle w^k, H_k w^k \rangle} = \langle p, p \rangle + \frac{\langle z^k, \xi \rangle^2}{\langle z^k, w^k \rangle} - \frac{\langle p, q \rangle^2}{\langle q, q \rangle^2},$$

pa na osnovu (17) i uvedenih oznaka, sledi

$$\langle \xi, H_{k+1}\xi \rangle = \frac{\langle p, p \rangle \langle q, q \rangle - \langle p, q \rangle^2}{\langle q, q \rangle^2} + \frac{\langle z^k, \xi \rangle^2}{\langle z^k, w^k \rangle} \geq \frac{\langle z^k, \xi \rangle^2}{\langle z^k, w^k \rangle}$$

jer je $\langle p, q \rangle^2 \leq \langle p, p \rangle \langle q, q \rangle$ prema tzv. nejednačini Cauchy-Schwartz-Bunjakovskog.

Sada je

$$\begin{aligned}\langle z^k, w^k\rangle &= \langle z^k, g^{k+1} - g^k\rangle \\ &= -\langle z^k, g^k\rangle \quad \text{zbog svojstva ortogonalnosti} \\ &= \alpha_k\langle g^k, H_k g^k\rangle, \text{ prema definiciji } z^k.\end{aligned}$$

Kako je $\alpha_k > 0$ i isto tako (zbog pozitivne definitnosti matrice H_k) sledi $\langle g^k, H_k g^k\rangle > 0$ a odavde je $\langle z^k, w^k\rangle > 0$.

Iz ovoga sledi da je $\langle \xi, H_{k+1}\xi\rangle > 0$ za svako $\xi \neq 0$, što je i trebalo dokazati.

Potrebno je napomenuti, da se u dokazu pretpostavilo da je $g^k \neq 0$, tj. da x^k nije stacionarna tačka. U slučaju da ovo ne važi, algoritam se završava u k-toj iteraciji i gornji rezultat važi samo za $H_i, i = 0, 1, \dots, k$.

Lema 4. [6] Neka je funkcija cilja kvardatna forma

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$$

sa pozitivno definitnom matricom A, $A \in R^{n,n}, b \in R^n, c \in R$. Tada važi

$$\langle d^k, Ad^l\rangle = 0 \text{ za } k \neq l \quad \text{ i } \quad H_k\mathrm{A}d^l = d^l \text{ za } k > l.$$

Teorema 14. [6] Kada se DFP algoritam primeni na funkciju cilja koja je kvadratna forma

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c,$$

gde je A pozitivno definitna simetrična matrica, minimum se dobija u najviše n iteracija. Ako se rešenje dobija u n-tom koraku onda je $H_n = A^{-1}$.

<u>Dokaz:</u>

Neka su u DFP algoritmu dobijeni vektori $d^k, k = 0, 1, \dots, n-1$ (ako rešenje nije dobijeno ranije). Oni su na osnovu Leme 5 linearno nezavisni i mogu generisati ceo prostor R^n. Koristeći Lemu 4 važi $H_n\mathrm{A}d^k = d^k$ za $k = 0, 1, \dots, n-1$. Pošto vektori d^k čine bazu prostora, sledi da je $H_n\mathrm{A}x = x$, za svako $x \epsilon R^n$, pa sledi $H_n = A^{-1}$. Dalje je prema Lemi 4

$$\langle d^k, Ad^n\rangle = 0, k = 0,1, \dots, n-1,$$

$$\langle d^n, AH_n\nabla f(x^n)\rangle = 0,$$

$$\langle d^k, \nabla f(x^n)\rangle = 0.$$

Pošto je $\nabla f(x^n)$ normalno na n linearno nezavisnih vektora $d^k, k = 0, 1, \dots, n-1$ sledi $\nabla f(x^n) = 0$, tj. $x^n = x^*$.

Primer 5. Odrediti minimum funkcije

$$f(x_1, x_2) = {x_1}^2 - 2x_1x_2 + 2{x_2}^2 - x_1 - 2x_2$$

DFP metodom iz početne tačke $x^0 = [1,1.5]^T$.

Rešenje:

Kako je data funkcija kvadratna i strogo konveksna, sledi da će se rešenje naći u najviše dva koraka. Zbog toga se ne određuje pravilo zaustavljanja. Za početnu pozitivno definitnu matricu H_0 se bira da je jedinična matrica.

Neka je gradijent funkcije

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 - 2x_2 - 1 \\ -2x_1 + 4x_2 - 2 \end{bmatrix}$$

tada je gradijent u početnoj tački

$$g^0 = \nabla f(x^0) = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

Dalje, je

$$d^0 = -H_0 g^0 = -\nabla f(x^0) = \begin{bmatrix} 2 \\ -2 \end{bmatrix}.$$

Sada se traži optimalna dužina koraka α_0. Ono se nalazi kao rešenje sledećeg problema minimizacije

$$f_0(\alpha) = f(x^0 + \alpha d^0) = f\left(\begin{bmatrix} 1 + 2\alpha \\ 1.5 - 2\alpha \end{bmatrix}\right) = -1.5 - 8\alpha + 20\alpha^2.$$

Odavde sledi

$${f_0}'(\alpha) = 0 \text{ kada je } \alpha = 0.2,$$

pa je $\alpha_0 = 0.2$ i

$$z^0 = 0.2d^0 = \begin{bmatrix} 0.4 \\ -0.4 \end{bmatrix}.$$

Sada je prva iteracija

$$x^1 = x^0 + z^0 = \begin{bmatrix} 1.4 \\ 1.1 \end{bmatrix}.$$

Sledi upotpunjavanje matrice H_1.

$$g^1 = \nabla f(x^1) = \begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix},$$

$$w^0 = g^1 - g^0 = \nabla f(x^1) - \nabla f(x^0) = \begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix} - \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1.6 \\ -2.4 \end{bmatrix},$$

$$A_0 = \frac{1}{\langle z^0, w^0 \rangle}(z^0)(z^0)^T = \frac{1}{1.6}\begin{bmatrix} 0.16 & -0.16 \\ -0.16 & 0.16 \end{bmatrix}$$

$$B_0 = \frac{1}{\langle w^0, H_0 w^0 \rangle} H_0 w^0 (w^0)^T H_0 = \frac{1}{8.32}\begin{bmatrix} 2.56 & -3.84 \\ -3.84 & 5.76 \end{bmatrix}$$

$$H_1 = H_0 + A_0 + B_0 = \begin{bmatrix} 0.79231 & 0.36154 \\ 0.36154 & 0.40769 \end{bmatrix}.$$

Računa se druga iteracija.

$$d^1 = -H_1 g^1 = -\begin{bmatrix} 0.79231 & 0.36154 \\ 0.36154 & 0.40769 \end{bmatrix}\begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix} = \begin{bmatrix} 0.46154 \\ 0.30769 \end{bmatrix}.$$

Optimalna dužina koraka α_1 se nalazi kao rešenje sledećeg problema minimizacije

$$f_1(\alpha) = f(x^1 + \alpha d^1) = f\left(\begin{bmatrix} 1.4 + 0.46154\alpha \\ 1.1 + 0.30769\alpha \end{bmatrix}\right).$$

Odavde sledi

$$f_1{}'(\alpha) = 0 \text{ kada je } \alpha = 1.3,$$

pa je $\alpha_1 = 1.3$ i

$$z^1 = 1.3 d^1 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}.$$

Sada je druga iteracija

$$x^2 = x^1 + z^1 = \begin{bmatrix} 2 \\ 1.5 \end{bmatrix}$$

što je i traženi minimum.

3.4 Metoda konjugovanih gradijenata

Metoda konjugovanih gradijenata je nastala sa idejom da se dobije postupak koji je brži od metode najbržeg pada, a s druge strane da se izbegne izračunavanje Hesijana i rešavanje sistema jednačina, tj. da se dobije postupak koji je jeftiniji od Njutnovog postupka.

Postupak konjugovanih gradijenata je modifikacija postupka najbržeg pada, u smislu da se kod ove metode drugačije bira pravac pretraživanja.

Kod ovog postupka uzastopni pravci nisu međusobno ortogonalni, kao što je to slučaj kod metode najbržeg pada, već su konjugovani u odnosu na simetričnu pozitivno definitnu matricu.

Definicija 15. Neka je A pozitivno definitna matrica reda n. Pravci $p_0, p_1, \dots, p_m,\ m \leq n$, su konjugovani u odnosu na matricu A, tj. A-konjugovani, ako važi

$$\langle p_i, Ap_j \rangle = 0, i \neq j.$$

Lema 5. [6] Pravci konjugovani u odnosu na pozitivno definitnu matricu su linearno nezavisni.

Definicija 16. Neka je Y^m linearni vektorski potprostor dimenzije *m* u prostoru R^n. Skup $a + Y^m$, gde je $a \in R^n$, naziva se *linearna mnogostrukost* dimenzije *m*. Mnogostrukost dimenzije $n - 1$ se naziva *hiperravan*, a linearna mnogostrukost dimenzije 1 je *prava*.

Lema 6. [6] Skup

$$G = \{x \in R^n,\ \langle c, x \rangle = \gamma\},\ c \in R^n,\ c \neq 0,\ \gamma \in R$$

je hiperravan prostora R^n. Vektor c je vektor normale hiperravni G.

Sledi algoritam za metodu konjugovanih gradijenata.

Algoritam 4. Metoda konjugovanih gradijenata

Korak 1. Izabrati početnu tačku $x^0 \in R^n$. Zadati $\varepsilon > 0$ za izlazni kriterijum.

Korak 2. Ako je $\|\nabla f(x^0)\| < \varepsilon$ tada je $x^0 = x^*$ rešenje. Inače, preći na korak 3.

Korak 3. Staviti $p_0 = \nabla f(x^0)$.

Korak 4. Naći α_k kao rešenje problema $\min_{\alpha \geq 0} f(x^k - \alpha \cdot p_k)$.

Korak 5. Staviti $x^{k+1} = x^k - \alpha_k \cdot p_k$.

Korak 6. Izračunati $\nabla f(x^{k+1})$. Ako je $\|\nabla f(x^{k+1})\| < \varepsilon$ tada je $x^{k+1} = x^*$ rešenje, inače izračunati

$$p_{k+1} = \nabla f(x^{k+1}) - \beta_k \cdot p_k$$

gde se β_k se računa na jedan od sledeća tri načina

$$i)\quad \beta_k = \frac{\langle \nabla^2 f(x^{k+1})p_k, \nabla f(x^{k+1})\rangle}{\langle p_k, \nabla^2 f(x^{k+1})p_k\rangle}$$

$$ii)\quad \beta_k = -\frac{\|\nabla f(x^{k+1})\|^2}{\|\nabla f(x^k)\|^2}$$

$$iii)\quad \beta_k = -\frac{\langle \nabla f(x^{k+1}) - \nabla f(x^k), \nabla f(x^{k+1})\rangle}{\|\nabla f(x^k)\|^2}$$

Korak 7. Zameniti k sa $k+1$ i ići na korak 4.

Za primenu ove metode potrebno je da je $f \epsilon C^1(R^n)$, a čak i $f \epsilon C^2(R^n)$ ako β_k određujemo sa i).

Specijalan slučaj metode konjugovanih gradijenata je slučaj kada funkcija cilja ima oblik kvadratne forme

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c,$$

gde je A simetrična, pozitivno definitna matrica reda n, $b, x \in R^n, c \in R$. Tada se do rešenja dolazi slično kao u Algoritmu 2, sa tim da se α_k i β_k računaju na sledeći način

$$\beta_k = \frac{\langle Ap_k, \nabla f(x^{k+1})\rangle}{\langle p_k, Ap_k\rangle}$$

$$\alpha_k = \frac{\|\nabla f(x^k)\|^2}{\langle Ap_k, p_k\rangle}.$$

Potrebno je napomenuti da ako funkcija cilja ima oblik kvadratne forme, tada metoda konjugovanih gradijenata daje rešenje u najviše *n* koraka, gde je *n* dimenzija problema.

Teorema 15. [6] Neka je funkcija cilja oblika kvadratne forme

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$$

gde je $A \in R^{n,n}, b \in R^n, c \in R$ i neka su pravci p_k određeni sa

$$p_k = \nabla f(x^k) - \beta_{k-1}p_{k-1}$$

gde je

$$\beta_{k-1} = \frac{\langle \nabla f(x^k), Ap_{k-1} \rangle}{\langle Ap_{k-1}, p_{k-1} \rangle}$$

i $p_0 = \nabla f(x^0), p_{k-1} \neq 0$. Tada važi:

1) Pravci p_k su konjugovani u odnosu na matricu A i vektori $Ap_1, Ap_2, \dots, Ap_k$ su linearno nezavisni.

2) x^* pripada linearnoj mnogostrukosti dimenzije $n-k$, koja je data izrazom

$$G_{n-k} = \{x \epsilon R^n; \langle Ap_{i-1}, x - x^i \rangle = 0, i = 1, \dots, k\}.$$

3) Metoda konjugovanih gradijenata daje minimum kvadratne forme u najviše n koraka.

<u>Dokaz:</u>

1) Teorema se dokazuje indukcijom. Prvo treba pokazati da tvrđenje važi za $k = 1$, tj. da važi

$$\langle Ap_0, p_1 \rangle = \langle Ap_1, p_0 \rangle = 0.$$

Zbog simetričnosti matrice je

$$\langle Ap_0, p_1 \rangle = \langle Ap_1, p_0 \rangle,$$

a iz definicije koraka β_0 sledi

$$\begin{aligned}\langle Ap_0, p_1 \rangle &= \langle Ap_0, \nabla f(x^1) - \beta_0 p_0 \rangle \\ &= \langle Ap_0, \nabla f(x^1) - \frac{\langle Ap_0, \nabla f(x^1) \rangle}{\langle Ap_0, p_0 \rangle} p_0 \rangle \\ &= \frac{\langle Ap_0, \nabla f(x^1) \langle Ap_0, p_0 \rangle - \langle Ap_0, \nabla f(x^1) \rangle p_0 \rangle}{\langle Ap_0, p_0 \rangle} \\ &= \frac{\langle Ap_0, \nabla f(x^1) (\langle Ap_0, p_0 \rangle - \langle Ap_0, p_0 \rangle) \rangle}{\langle Ap_0, p_0 \rangle} = 0.\end{aligned}$$

Sad još treba dokazati da su Ap_0 i Ap_1 linearno nezavisni vektori. Pretpostavlja se suprotno, tj. da postoje konstante a_1 i a_2 takve da je

$$a_1 Ap_0 + a_2 Ap_1 = 0,$$

pri čemu je bar jedna od konstanti a_1 ili a_2 različita od nule. Neka je $a_1 \neq 0$. Množenjem ove jednakosti skalarno sa p_1 dobija se da je $a_2 = 0$, a množenjem sa p_0 dobija se da je $a_1 = 0$, što je u kontradikciji sa pretpostavkom. Dakle, vektori su linearno nezavisni.

Dokaz za k je analogan prethodnom.

2) Kako je x^* minimum, sledi da je $\nabla f(x^*) = Ax^* - b = 0$, tj. $x^* = A^{-1}b$. Sada je

$$\begin{aligned}\langle Ap_{k-1}, x^* - x^k\rangle &= \langle Ap_{k-1}, A^{-1}b - x^k\rangle = \langle A^{-1}Ap_{k-1}, b - Ax^k\rangle = -\langle p_{k-1}, \nabla f(x^k)\rangle \\ &= 0, k = 1, \dots, n.\end{aligned}$$

Dakle, $x^* \epsilon\, G_{n-k}$.

3) Kada je $\nabla f(x^{k-1}) = 0$ sledi da je $x^* = x^{k-1}$. Kada je pak, $\nabla f(x^{k-1}) \neq 0$ sprovodi se postupak konjugovanih gradijenata. Kako je $x^* \epsilon\, G_0 \subset G_1 \subset \dots \subset G_{n-1} \subset R^n$ sledi da je G_0 dimenzije nula, dakle $G_0 = \{x^*\}$. Ukoliko pre n-tog koraka ne postoji slučaj kada je $\nabla f(x^{k-1}) = 0$, sledi da je $\nabla f(x^n) = 0$, tj. $x^n = x^*$.

Ukoliko pak funkcija cilja nema oblik kvadratne forme, tada ne postoji garancija u koliko koraka se dolazi do rešenja. Može se desiti da dođe do nagomilavanja grešaka, pa se zato nakon određenog broja koraka radi restart tj. uzima da je $\beta_k = 0$, tj. za pravac pretraživanja se bira pravac negativnog gradijenta i ta modifikacija se naziva metoda konjugovanih gradijenata sa restartom. Ova metoda je uopštenje metode najbržeg pada. Njena brzina konvergencije je data istom formulom kao kod metode najbržeg pada, s tim što su veliki karakteristični koreni eliminisani iz razmatranja, o čemu govori naredna teorema.

Teorema 16. [2] Neka je $f \in C^2(R^n)$ i neka je x^* lokalni minimum od $f(x)$. Neka je $\nabla^2 f(x^*)$ simetrična i pozitivno definitna matrica čiji je najmanji karakteristični koren $\bar{a}$, a najveći $\bar{A}$ i koja ima $(n-m)$ karakterističnih korena u intervalu $[\bar{a}, b], \bar{a} > 0$, a preostalih m karakterističnih korena većih od b. Ako niz $\{x^k\}$ generisan metodom konjugovanih gradijenata sa restartom svakih $(m+1)$ koraka (tj. x^{k+1} se dobija iz x^k na osnovu $(m+1)$ konjugovano gradijentnih koraka) konvergira ka x^*, tada $\{f(x^k)\}$ konvergira ka $f(x^*)$ sa koeficijentom brzine $\left(\frac{b-\bar{a}}{b+\bar{a}}\right)^2$.

Primer 6. Metodom konjugovanih gradijenata odrediti minimum funkcije

$$f(x) = \frac{3}{2}{x_1}^2 + \frac{5}{2}{x_2}^2 - x_1.x_2 - 5x_1$$

gde je $x = [x_1, x_2]^T \epsilon\, R^2$, a početna tačka je $x^0 = [0,0]^T$.

<u>Rešenje:</u>

Funkcija cilja se može zapisati kao

$$f(x) = \frac{1}{2}(3{x_1}^2 + 5{x_2}^2 - 2x_1 \cdot x_2) - 5x_1$$

i ona ima oblik kvadratne forme, gde je

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 5 \end{bmatrix}, \qquad b = \begin{bmatrix} -5 \\ 0 \end{bmatrix}.$$

S obzirom na to da je funkcija cilja kvadratna, do rešenja se dolazi u najviše dva koraka.

Gradijent funkcije cilja je oblika

$$\nabla f(x) = Ax + b$$

tj.

$$\nabla f(x) = \begin{bmatrix} 3x_1 - x_2 - 5 \\ -x_1 + 5x_2 \end{bmatrix}.$$

Pošto je $\|\nabla f(x^0)\| = \sqrt{25} \neq 0$, dakle $x^0 = [0,0]^T$ nije rešenje.

Traži se prva aproksimacija

$$x^1 = x^0 - \alpha_0 \cdot p_0$$

gde je

$$p_0 = \nabla f(x^0) = \begin{bmatrix} -5 \\ 0 \end{bmatrix}$$

$$\alpha_0 = \frac{\|\nabla f(x^0)\|^2}{\langle Ap_0, p_0 \rangle} = \frac{25}{\langle \begin{bmatrix} 3 & -1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} -5 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \end{bmatrix} \rangle} = \frac{25}{75} = \frac{1}{3}.$$

Sada je $x^1 = \begin{bmatrix} -5 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} -5 \\ 0 \end{bmatrix} = \begin{bmatrix} 5/3 \\ 0 \end{bmatrix}$.

Kako je $\nabla f(x^1) = \begin{bmatrix} 0 \\ -5/3 \end{bmatrix}$ i $\|\nabla f(x^1)\| \neq 0$, sledi da x^1 nije rešenje.

Traži se druga aproksimacija

$$x^2 = x^1 - \alpha_1 \cdot p_{1,}$$

gde je

$$p_1 = \nabla f(x^1) - \beta_1 \cdot p_0$$

$$\beta_1 = \frac{\langle Ap_0, \nabla f(x^1) \rangle}{\langle Ap_0, p_0 \rangle} = \frac{\langle \begin{bmatrix} -15 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ -5/3 \end{bmatrix} \rangle}{75} = -\frac{1}{9}$$

$$\alpha_1 = \frac{\|\nabla f(x^1)\|^2}{\langle Ap_1, p_1 \rangle} = \frac{25/9}{\langle \begin{bmatrix} 3 & -1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} -5/9 \\ -5/3 \end{bmatrix}, \begin{bmatrix} -5/9 \\ -5/3 \end{bmatrix} \rangle} = \frac{3}{14}.$$

Sada je

$$p_1 = \begin{bmatrix} 0 \\ -\frac{5}{3} \end{bmatrix} + \frac{1}{9}\begin{bmatrix} -5 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{5}{9} \\ -\frac{5}{3} \end{bmatrix},$$

pa je $x^2 = \begin{bmatrix} \frac{5}{3} \\ 0 \end{bmatrix} - \frac{3}{14}\begin{bmatrix} -\frac{5}{9} \\ -\frac{5}{3} \end{bmatrix} = \begin{bmatrix} \frac{25}{14} \\ \frac{5}{14} \end{bmatrix}$.

Pošto je $\nabla f(x^2) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, tj $\|\nabla f(x^2)\| = 0$, sledi da je $x^2 = \begin{bmatrix} \frac{25}{14} \\ \frac{5}{14} \end{bmatrix}$ rešenje problema.

Primer 7. Odrediti minimum funkcije

$$f(x) = {x_1}^2 - x_1 x_2 - 4x_1 + {x_2}^2 - x_2$$

metodom konjugovanih gradijenata, sa početnom iteracijom $x^0 = [0,0]^T$.

<u>Rešenje:</u>

Zadatak je rešenjen u programskom paketu Matlab. Naredni M-file-ovi koji se koriste pri određivanju minimuma ove funkcije.

Prvi M-file se odnosi na računanje gradijenta funkcije f:

```
function g= grad(f,x,h)
%racuna gradijent funkcije f(x)
if nargin<3, h=.00001; end
h2= 2*h; N= length(x);
I= eye(N);
for n=1:N
g(n)= (feval(f,x+I(n,:)*h)-feval(f,x-I(n,:)*h))/h2;
end
```

Sledeći M-file se odnosi na određivanje minimuma.

```
function [xo,fo] = opt_conjg(f,x0,TolX,TolFun,alpha0,MaxIter,KC)
%KC = 2: Koristi metod konjugovanih gradijenata
if nargin < 7, KC = 0; end
if nargin < 6, MaxIter = 100; end
if nargin < 5, alpha0 = 10; end
if nargin < 4, TolFun = 1e-8; end
if nargin < 3, TolX = 1e-6; end
N = length(x0); nmax1 = 20; warning = 0; h = 1e-4;%dimenzije promenljive
x = x0; fx = feval(f,x0); fx0 = fx;
for k = 1: MaxIter
xk0 = x; fk0 = fx; alpha = alpha0;
g = grad(f,x,h); s = -g;
for n = 1:N
alpha = alpha0;
fx1 = feval(f,x + alpha*2*s);
for n1 = 1:nmax1 %Nalazi optimalnu duzinu koraka
fx2 = fx1; fx1 = feval(f,x+alpha*s);
```

```
if fx0 > fx1 + TolFun && fx1 < fx2 - TolFun %fx0 > fx1 < fx2
den = 4*fx1 - 2*fx0 - 2*fx2; num = den-fx0 + fx2;
alpha = alpha*num/den;
x = x+alpha*s; fx = feval(f,x);
break;
elseif n1 == nmax1/2
alpha = -alpha0; fx1 = feval(f,x + alpha*2*s);
else
alpha = alpha/2;
end
end
x0 = x
fx0 = fx
if n < N
g1 = grad(f,x,h);
if KC <= 1, s = - g1 +(g1 - g)*g1'/(g*g'+ 1e-5)*s;
else s = -g1 + g1*g1'/(g*g'+ 1e-5)*s;
end
g = g1;
end
if n1 >= nmax1, warning = warning+1;
%nije nadjena optimalna duzina koraka
else warning = 0;
end
end
if warning >= (norm(x - xk0,2)<TolX&&abs(fx - fk0)< TolFun),break;
end
end
xo = x; fo = fx;
if k == MaxIter, fprintf('Just best in %d iterations',MaxIter),
end
```

Sada se u Matlab prozoru izvršava program:

```
>> f = inline('x(1).^2 - 4*x(1) - x(1).*x(2) + x(2).^2 - x(2)','x');
>> x0 =[0 0]; TolX = 1e-4; TolFun = 1e-4; alpha0 = 10; MaxIter = 100;
>> [xo,fo] = opt_conjg(f,x0,TolX,TolFun,alpha0,MaxIter,2)

x0 =

    2.6154    0.6538


fx0 =

   -5.5577


x0 =

    3.0000    2.0000


fx0 =

   -7.0000
```

xo =

3.0000 2.0000
fo =

-7.0000

Dakle, minimum funkcije je $[3,2]^T$. U sledećoj tabeli su navedene sve iteracije dobijene metodom konjugovanih gradijenata.

k	x_1^k	x_2^k	$f(x^k)$
0	0	0	0
1	2.6154	0.6538	-5.5577
2	3.0000	2.0000	-7.0000

Tabela 3. Vrednosti iteracija dobijenih metodom konjugovanih gradijenata

Dakle, metodom konjugovanih gradijenata rešenje je dobijeno u drugoj iteraciji, a isti ovaj zadatak je rešen prethodno metodom najbržeg pada gde je rešenje dobijeno u 13. iteraciji.

Metode konjugovanih gradijenata se odlikuju kvadratnim završavanjem, međutim numerički rezultati pokazuju da su one osetljive na izbor dužine koraka, pa mogu biti neefikasne u nekim slučajevima. One daju dobre rezultate pri rešavanju problema bez ograničenja velikih dimenzija, jer ne koriste loše uslovljene matrice za izračunavanje pravaca pretraživanja.

4 Optimizacija sa ograničenjima - Metode kaznenih funkcija

Posmatra se problem minimizacije sa ograničenjima oblika

$$\min_{x \in X} f(x), \tag{18}$$

gde je X dopustiv skup.

Među najjednostavnije i najpopularnije metode za rešavanje problema ovog tipa spadaju metode kaznenih funkcija. One svode problem (18) na niz problema optimizacije bez ograničenja.

Prvu kaznenu funkciju sa ograničenjima tipa jednakosti uveo je Kourant (*Courant*) 1943. godine, a 1969. godine Pietrgikovski (*Pietrgykowski*) je analizirao ovaj pristup za rešavanje nelinearnih problema. Značajan napredak u rešavanju praktičnih problema korišćenjem metoda kaznenih funkcija dali su Fiako i Mekormik, u svom radu pod nazivom tehnike sekvencijalne neograničene minimizacije (*SUMT*), a 1968. godine oni su predstavili efikasne algoritme za optimizaciju korišćenjem metoda kaznenih funkcija. Ovom tematikom bavio se i Flečer (*Fletcher*) koji je uveo neke nove koncepte kaznenih funkcija.

Metode kaznenih funkcija su procedure za rešavanje problema optimizacije sa ograničenjima pomoću niza problema optimizacije bez ograničenja. Kod ovih postupaka aproksimacija se postiže dodavanjem kaznenog člana funkciji cilja u delu domena koji ne zadovoljava data ograničenja (nedopustivo područje). Ovako se pogoršava vrednost tako modifikovane funkcije cilja, i tačke u nedopustivom području će postati "loše" (nekonkurentne), što proces traženja prilikom optimizacije „tera" iz nedopustivog područja. Kako su ograničenja „ugrađena" u modifikovanu funkciju cilja, problem postaje problem bez ograničenja, pa se za njegovo rešavanje mogu primeniti postupci koji se koriste u tu svrhu.

Modifikovana funkcija cilja se u dopustivom području mora poklapati sa zadatom funkcijom cilja. Zbog toga je važno da parametar koji određuje težinu kazne ima zanemarljiv uticaj na iznos funkcije cilja na dopustivom skupu.

Međutim, u delu domena koji ne zadovoljava ograničenja, funkcija cilja treba da prima velike vrednosti. To se postiže dodavanjem kaznenog člana, koji je definisan na takav način. Kako se funkcija cilja modifikuje, moguće je primeniti postupke za optimizaciju problema bez ograničenja, a da ograničenja ipak budu zadovoljena.

Nameće se važno pitanje, koliko dobro problem bez ograničenja vrši aproksimaciju originalnog problema sa ograničenjima. Ovo je od suštinskog značaja za ispitivanje toga, da li rešenje problema bez ograničenja konvergira ka rešenju originalnog problema.

Drugo pitanje, koje se postavlja kod ovakvih problema jeste, kako da se reši dati problem bez ograničenja, kada njegova funkcija cilja sadrži kazneni član, jer je obično struktura pomoćnog problema bez ograničenja mnogo složenija, što može usporiti konvergenciju algoritama koji se koriste za rešavanje takvih problema.

Ukoliko funkcilja cilja i ograničenja nisu diferencijabilni u domenu, tada se za rešavanje niza problema optimizacije bez ograničenja koriste negradijentne metode, kao što su napr. metoda Hooke-Jeeves, Nelder-Mead simpleks postupak, Powellov postupak i dr.

Ukoliko su pak funkcija cilja i ograničenja diferencijabilne, tada se mogu primeniti gradijentne metode. Neke od najčešće korišćenih metoda u tu svrhu su postupak konjugovanih gradijenata, Njutnov postupak, kvazi-Njutnovi postupci i konjugovani gradijenti sa restartom.

Metode kaznenih funkcija su jednostavne za primenu, a u ovom radu su predstavljene metode spoljašnjih i unutrašnjih (barijernih) kaznenih funkcija, kao i njihove kombinacije sa metodom konjugovanih gradijenata i Njutnovom metodom.

4.1 Metoda spoljašnjih kaznenih funkcija

Kao što je rečeno, osnovna ideja metoda kaznenih funkcija jeste svođenje problema (18) na niz pomoćnih problema.

Problem (18) se može približno rešiti rešavanjem niza problema minimizacije bez ograničenja oblika

$$\min_{x \in X_0} F_k(x), \quad k = 1,2,\ldots \tag{19}$$

gde je $X \subset X_0 \subseteq R^n$.

Funkcije $F_k(x)$ su pomoćne funkcije, koje se određuju tako da se sa porastom broja k sve manje razlikuju od funkcije cilja $f(x)$ na skupu X, a van dopustivog skupa X trpe beskonačno veliku kaznu, što se postiže uvođenjem kaznenog dodatka.

Skup X_0 je najčešće ceo prostor R^n ili neki njegov podskup. U ulozi skupa X_0 obično se uzima skup koji je zadat na jednostavan način, tako da minimizacija pomoćnih funkcija $F_k(x)$ na skupu X_0 ne predstavlja veliki problem.

Definicija 17. Kaznene funkcije skupa X na skupu X_0 $(X \subset X_0)$ su funkcije $\{q_k(x)\}$, $k = 1,2,\ldots$ koje su definisane i nenegativne na skupu X_0 sa osobinom

$$\lim_{k\to\infty} q_k(x) = \begin{cases} 0, & x \in X \\ +\infty, & x \in X_0 \setminus X\,. \end{cases}$$

Proširena funkcija $F_k(x)$ u (19) ima oblik

$$F_k(x) = f(x) + q_k(x), \quad x \in X_0, \qquad k = 1,2,\ldots$$

pa se rešava niz problema

$$\min F_k(x) = \min(f(x) + q_k(x)), \;\; x \in X_0, \;\; k = 1,2,\ldots \tag{20}$$

Neka je $F_{k_*} = inf_{x \in X_0} F_k(x) > -\infty$, $k = 1,2,\ldots$. Rešenje problema (19) za fiksirano k se obeležava sa x^k, odnosno

$$F_k(x^k) = F_{k_*}.$$

Za široku klasu problema važi $\lim_{k\to\infty} F_{k_*} = f^*$ ili $\lim_{k\to\infty} x^k = x^*$, gde je $f^* = \min f(x),\; x \in X$, a x^* je optimalno rešenje problema (18).

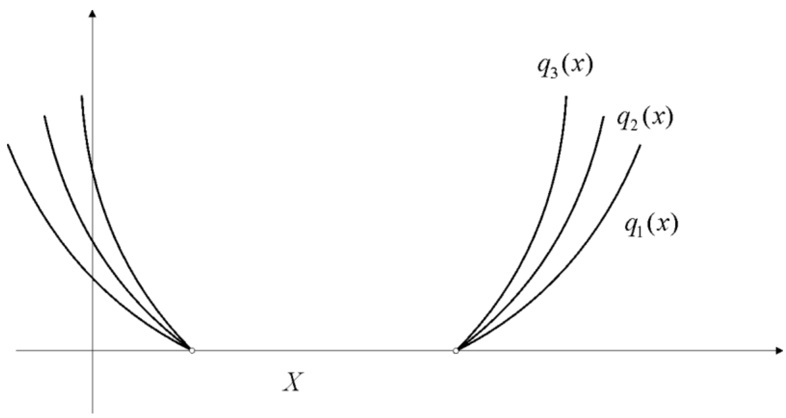
Ukoliko se neki od problema (20) ne mogu tačno rešiti, rešavaju se približno nekom metodom tako da važi

$$F_k(x^k) \le F_{k_*} + \varepsilon_k, \qquad \varepsilon_k > 0, \; \lim_{k\to\infty} \varepsilon_k = 0, \qquad k = 1,2,\ldots$$

U dopustivom području problema, vrednosti ukupne modifikovane funkcije cilja su manje, jer tu nema kaznenog dodatka, zbog čega se ne stvara „zid“ kako tok traženja ne bi izašao iz dopustivog područja, već kazneni dodatak “gura” tok traženja optimuma u dopustivo područje. Kod spoljašnjih kaznenih funkcija početna tačka i druge tačke za vreme procesa traženja optimuma ne moraju biti u dopustivom području.

Oblik kaznenih funkcija osigurava porast modifikovane funkcije cilja van dopustivog skupa, dok u dopustivom skupu nema kazne, tj. kazneni dodatak je nula.

Na slici 2. se vidi geometrijska interpretacija jednog niza spoljašnjih kaznenih funkcija $\{q_k(x)\}$



Slika 2. Niz spoljašnjih kaznenih funkcija

Neka je dopustiv skup X definisan na sledeći način

$$X = \{x \in R^n,\ \ g_i(x) \leq 0, i = 1, \dots, m,\ \ g_i(x) = 0, i = m+1, \dots, s\}. \qquad (21)$$

Jedna od najčešće korišćenih spoljašnjih kaznenih funkcija za skup X je

$$q_k(x) = t_k q(x)$$

gde je

$$q(x) = \sum_{i=1}^{m} max\{0, g_i(x)\}^p + \sum_{i=m+1}^{s} \left|g_i(x)\right|^p,\ \ p \geq 1,\ \ x \in X_0 \qquad (22)$$

ili

$$q(x) = \sum_{i=1}^{s} (g_i(x) + \left|g_i(x)\right|)^2/2\,,$$

gde su $t_k > 0, k = 1,2, \dots$ sa osobinom

$$\lim_{k\to\infty} t_k = +\infty,$$

i zovu se *kaznene konstante (kazneni parametri).*

Ukoliko su funkcije $g_i(x) \in C^r(X_0)$ za $i = 1, \dots, s$, tada je za bilo koje $p > r$, $q(x) \in C^r(X_0)$.

U praksi se obično za p uzima vrednost 2, jer to obezbeđuje da je $q(x)$ klase C^1, tj. da ima neprekidan prvi izvod.

Ako se kaznena funkcija $q_k(x)$ bira na pomenut način, tada ona nasleđuje dobra svojstva funkcija $g_i(x)$, tj. ako su funkcije $g_1, \dots, g_s$ diferencijabilne i/ili konveksne, to će važiti i za $q_k(x)$.

Metoda spoljašnjih kaznenih funkcija data je sledećim algoritmom.

Algoritam 5. Metoda spoljašnjih kaznenih funkcija

Korak 1. Neka je $\{t_k\}$, $k = 1, 2 \dots$ niz koji teži beskonačnosti, tako da je za svako k, $t_k \geq 0$, $t_{k+1} > t_k$.

Korak 2. Za svako $k = 1, 2 \dots$ definiše se funkcija

$$F_k(x) = f(x) + q_k(x)$$

i rešava se problem

$$\min F_k(x), x \in X_0$$

odakle se dobija rešenje x^k.

U koraku 2. se pretpostavlja da za svako k problem minimizacije ima rešenje.

Za konvergenciju navedenog postupka bitan je odnos funkcije cilja i ograničenja u datom problemu.

Definicija 18. Ograničenja $g_i(x)$, $i = 1, \dots, s$ su ***saglasna*** sa funkcijom cilja $f(x)$ na skupu X_0 ako za bilo koji niz $\{x^k\} \in X_0$ sa osobinom:

$$\begin{aligned} &\lim_{k\to\infty} sup \;\; g_i(x^k) \leq 0, \; i = 1, \dots, m \\ &\lim_{k\to\infty} \;\; g_i(x^k) = 0, \; i = m+1, \dots, s \end{aligned} \tag{23}$$

sledi

$$\lim_{k\to\infty} inf\, f(x^k) \geq f^* = \inf f(x), x \in X .$$

Teorema 17. [6] Neka je $F_k(x) = f(x) + t_k q(x)$, $k = 1, \dots, s$ gde je $q(x)$ definisana sa (22). Neka su ograničenja u (21) saglasna sa funkcijom cilja $f(x)$ na X_0. Tada važi

$$\lim_{k\to\infty} F_{k_*} = f^*.$$

Teorema 18. [6] Neka je $F_k(x) = f(x) + t_k q(x),\ k = 1, \dots, s$ gde je $q(x)$ definisana sa (22). Neka je

$$f^{**} = \inf_{x \in X_0} f(x) > -\infty.$$

Tada je

$$\lim_{k\to\infty} F_{k_*} = f^*$$

ako i samo ako su ograničenja saglasna sa funkcijom cilja na X_0.

Definicija 19. Tačka $(x^*, \lambda^*) \in X_0 \times \Lambda_0$ je ***sedlasta tačka*** funkcije Lagranža $L(x, \lambda)$ pridružene problemu (18) gde je skup X definisan sa (21), ako važi:

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*) \qquad \forall x \in X_0, \forall \lambda \in \Lambda_0$$

gde je $\Lambda_0 = \{\lambda_i \geq 0,\ i = 1, \dots, m\}$.

Lema 7. [6] Ako funkcija Lagranža pridružena problemu (18) sa dopustivim skupom X (21) ima sedlastu tačku, onda su ograničenja $g_i(x),\ i = 1, \dots, s$ saglasna sa funkcijom cilja $f(x)$ na X_0.

Dokaz:

Neka je $(x^*, \lambda^*) \in X_0 \times \Lambda_0$ sedlasta tačka funkcije Lagranža $L(x, \lambda)$,

$$L(x, \lambda) = f(x) + \textstyle\sum_{i=1}^{s} \lambda_i g_i(x),$$

$x \in X_0,\ \lambda \in \Lambda_0$, $\Lambda_0 = \{\lambda \in R^s ,\ \lambda_i \geq 0,\ i = 1, \dots, m\}$. Neka niz $\{x^k\} \in X_0$ ima osobinu (23).

Tada je

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) = f^* \leq L(x, \lambda^*)\, , \qquad x \in X_0, \qquad \lambda \in \Lambda_0\, .$$

Kako je $\{x^k\} \in X_0$, sledi

$$f^* \leq L(x^k, \lambda^*) = f(x^k) + \sum_{i=1}^{s} \lambda_i^* g_i(x^k).$$

Zbog (23) je

$$\lim_{k\to\infty} \inf \sum_{i=1}^{s} \lambda_i^* g_i(x^k) \leq 0$$

pa sledi

$$f^* \leq \lim_{k\to\infty} \inf f(x^k),$$

a to znači da su ograničenja saglasna sa funkcijom cilja.

Teorema 19. [6] Ako funkcija Lagranža problema (18) sa dopustivim skupom X definisanim sa (21) ima sedlastu tačku i ako je funkcija cilja $f(x)$ ograničena sa donje strane na X_0, tada za niz $\{x^k\}$ sa osobinom (23) važi

$$\lim_{k\to\infty} f(x^k) = \lim_{k\to\infty} F_k(x^k) = \lim_{k\to\infty} F_{k_*} = f^*.$$

Dokaz:

Dokaz sledi direktno primenom Teoreme 18 i Leme 7.

U sledećoj lemi navedene su nejednakosti koje proizilaze direktno iz definicije x^k i nejednakosti $t_{k+1} > t_k$.

Lema 8. [2]

$$i)\ \ F_k(x^k) \le F_{k+1}(x^{k+1})$$

$$ii)\ \ q(x^k) \ge q(x^{k+1})$$

$$iii)\ \ f(x^k) \le f(x^{k+1})$$

Dokaz:

i) Iz definicije $F_k(x)$ sledi

$$F_{k+1}(x^{k+1}) = f(x^{k+1}) + t_{k+1}q(x^{k+1}) \ge f(x^{k+1}) + t_k q(x^{k+1}) \ge f(x^k) + t_k q(x^k) = F_k(x^k)$$

ii) Važi da je

$$f(x^k) + t_k q(x^k) \le f(x^{k+1}) + t_k q(x^{k+1})$$

$$f(x^{k+1}) + t_{k+1} q(x^{k+1}) \le f(x^k) + t_{k+1} q(x^k).$$

Sabiranjem prethodna dva izraza dobija se sledeća nejednakost

$$(t_{k+1} - t_k) q(x^{k+1}) \le (t_{k+1} - t_k) q(x^k)$$

odakle sledi *ii).*

iii) Takođe važi

$$f(x^{k+1}) + t_k q(x^{k+1}) \ge f(x^k) + t_k q(x^k),$$

koristeći nejednakost pod *ii)* dobija se njednakost *iii).*

Lema 9. [2] Neka je x^* rešenje problema (19). Tada za svako k važi

$$f(x^*) \geq F_k(x^k) \geq f(x^k).$$

Dokaz:

$$f(x^*) = f(x^*) + t_k q(x^*) \geq f(x^k) + t_k q(x^k) \geq f(x^k).$$

Teorema 20. [2] Neka je $\{x^k\}$ niz generisan Algoritmom 5. Tada, svaka granična tačka niza predstavlja rešenje problema (19).

Dokaz:

Neka da je niz $\{x^k\}$, $k \in K$ konvergentan podniz od $\{x^k\}$ čija je granica $\bar{x}$. Kako je funkcija cilja f neprekidna, sledi

$$\lim_{k \in K} f(x^k) = f(\bar{x}).$$

Neka je f^* vrednost funkcije cilja u optimalnoj tački problema (19). Tada, prema Lemi 8 i 9, niz vrednosti $F_k(x^k)$ je neopadajući i ograničen sa f^*. Dakle,

$$\lim_{k \in K} F_k(x^k) = F^* \leq f^*.$$

Kada se prethodna dva izraza oduzmu, dobija se

$$\lim_{k \in K} t_k q(x^k) = F^* - f(\bar{x}).$$

Pošto $q(x^k) \geq 0$ i $t_k \to \infty$, sledi

$$\lim_{k \in K} q(x^k) = 0.$$

Kada se iskoristi da je funkcija q neprekidna, sledi da je $q(\bar{x}) = 0$, što znači da je granična tačka $\bar{x}$ dopustiva za (19). Na osnovu Leme 9 je $f(x^k) \leq f^*$, pa je

$$f(\bar{x}) = lim_{k \in K} f(x^k) \leq f^*.$$

Metoda spoljašnjih kaznenih funkcija je ilustrovana kroz nekoliko primera.

Primer 8. Dati problem optimizacije sa ograničenjima

$$\min(3x^2 - 8x)$$
$$x - 1 \leq 0.$$

rešiti pomoću metode spoljašnjih kaznenih funkcija.

Rešenje:

Zadatak se rešava tako, što se problem svodi na rešavanje niza zadataka oblika

$$\min_{x \in R} F_k(x),$$

a pri tom se koriste kaznene funkcije oblika (22). Tada je funkcija $F_k(x)$ definisana na sledeći način

$$F_k(x) = 3x^2 - 8x + t_k \max\{0, x-1\}^2, x \epsilon X_0 = R$$

i važi da su t_k pozitivni brojevi takvi da $t_k \to \infty,\ k \to \infty$.

Može se uočiti da je funkcija $F_k(x)$ konveksna i neprekidno diferencijabilna za svako k, pa se minimum može naći diferenciranjem. Posmatraju se dva slučaja:

1. $\max\{0, x-1\} = 0$.

Tada je $F_k(x) = 3x^2 - 8x$ i $F_k{}'(x) = 6x - 8 = 0$, odakle je $x = \frac{4}{3}$, što se odbacuje jer ne zadovoljava uslov $x - 1 \le 0$.

2. $\max\{0, x-2\} = x - 2$.

Tada je modifikovana funkcija cilja oblika

$$F_k(x) = 3x^2 - 8x \quad + t_k(x-2)^2$$

Dalje je

$$F_k{}'(x) = 6x - 8 + 2t_k(x-2) = 0$$

pa je rešenje prethodne jednačine

$$x_k = \frac{2t_k + 8}{2t_k + 6}.$$

Kako ovo rešenje zadovoljava uslov $\max\{0, x-2\} = x - 2$, usvaja se.
S obzirom da je $\lim_{k\to\infty} t_k = \infty$ sledi

$$\lim_{k\to\infty} x_k = \lim_{t_k\to\infty} \frac{2 + \frac{8}{t_k}}{2 + \frac{6}{t_k}} = 1$$

tj. $\lim_{k\to\infty} x_k = 1$.

Pošto je ovo problem konveksnog programiranja i $f(x) \in C^1$, da bi se pokazalo da je x^* zaista rešenje, potrebno je koristiti sledeću teoremu.

Teorema 21. [6] Neka je X konveksan skup, $f(x)$ konveksna na X i $f(x) \in C^1(X)$ i neka je X_* skup tačaka minimuma funkcije $f(x)$ na skupu X, tj.

$$X_* = \{x^* \in X | f(x^*) = f^* = \min f(x), x \in X\}.$$

Tada važi

$$x^* \in X_* \text{ ako i samo ako je } \langle \nabla f(x^*), x - x^* \rangle \geq 0, \forall x \in X.$$

Dakle, da bi se pokazalo da je $x^* = 1$ zaista rešenje Primera 4, treba pokazati da važi

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0$$

tj.

$$\langle \nabla f(x^*), x - x^* \rangle = \langle -2, x - 1 \rangle \geq 0$$

Ovo je tačno, što znači da $x^* = 1$ jeste rešenje polaznog problema i $f^* = f(1) = -5$ je optimalna vrednost funkcije cilja.

Primer 9. Neka je dat sledeći problem minimizacije

$$\min(x_1^2 + x_1 x_2 + x_2^2)$$
$$x_1 + x_2 - 2 = 0.$$

Rešiti ga metodom spoljašnjih kaznenih funkcija.

Rešenje:

Umesto ovog problema, rešava se niz zadataka oblika

$$\min_{(x_1,x_2) \in R^2} F_k(x_1, x_2),$$

gde je

$$F_k(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2 + t_k(x_1 + x_2 - 2)^2$$

modifikovana funkcija cilja. Pošto je $F_k(x)$ neprekidno diferencijabilna i konveksna funkcija, rešava se sistem jednačina

$$\frac{\partial F_k}{\partial x_1} = 2x_1 + x_2 + 2t_k(x_1 + x_2 - 2) = 0$$
$$\frac{\partial F_k}{\partial x_2} = x_1 + 2x_2 + 2t_k(x_1 + x_2 - 2) = 0.$$

Odatle se dobije

$$x_{1k} = x_{2k} = \frac{4t_k}{3 + 4t_k},$$

pa je

$$x_1^* = x_2^* = \lim_{t_k \to \infty} \frac{4t_k}{4t_k + 3} = 1.$$

Dakle, niz $[x_1^k, x_2^k]^T$ konvergira i $\lim_{k\to\infty}[x_1^k, x_2^k] = [1,1]^T$.

Početni problem je problem konveksnog programiranja i $f(x) \in C^1$. Zbog toga je potrebno pokazati da $x^* = [x_1^*, x_2^*]^T$ zaista rešenje. To se pokazuje pomoću Teoreme 21, tj. treba pokazati da je

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0.$$

Kako je

$$\begin{aligned}\langle \nabla f(x^*), x - x^* \rangle &= \left\langle \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} \right\rangle = 3(x_1 - 1) + 3(x_2 - 1) \\ &= 3x_1 + 3x_2 - 6 = 3(x_1 + x_2 - 2) = 0,\end{aligned}$$

zbog ograničenja $x_1 + x_2 - 2 = 0$, sledi da je

$$x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

zaista rešenje polaznog problema, a optimalna vrednost funkcije cilja je

$$f^* = 1^2 + 1 + 1^2 = 3.$$

Napomena:

Tačno rešenje se u prethodnim primerima tražilo diferenciranjem. Ovo je bilo moguće jer su funkcije $F_k(x)$ diferencijabilne i jednostavne. Kada je u pitanju problem optimizacije većih dimenzija, koriste se neke od postojećih metoda za optimizaciju bez ograničenja. Nedostatak tih metoda jeste da one daju samo približno rešenje tih problema.

Kako su kaznene funkcije kvadratne, one osiguravaju neprekidnost nagiba modifikovane funkcije cilja, mada drugi izvodi nisu neprekidni na rubu dopustivog područja. Zbog toga se više preporučuju kvazi-Njutnovi postupci za rešavanje problema bez ograničenja, jer se kod aproksimacije Hesijana koristi samo prvi izvod funkcije cilja.

Primer 10. Naći minimum funkcije

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

metodom spoljašnjih kaznenih funkcija uz ograničenje

$$x_2 \leq 3 - x_1$$

polazeći iz početne tačke $[5,5]^T$.

Rešenje:

Problem je rešen u programskom paketu *Matlab*®, korišćenjem kaznene funkcije oblika

$$q(x) = \sum_{i=1}^{s} (g_i(x) + |g_i(x)|)^2/2.$$

Modifikovana funkcilja cilja je

$$F_k(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 3)^2 + 4 + t_k\left(\frac{\left((x_2 + x_1 - 3) + |x_2 + x_1 - 3|\right)^2}{2}\right).$$

U programskom paketu *Matlab*® napravi se M-file za rešavanje datog problema.

Izgled M-file-a je sledeći:

```
function f=FunkcijaKaznena(x)
fun=(x(1)-2)^2+(x(2)-3)^2+4;
kaz=10;
ogr1=(x(1)+x(2)-3);
%spoljasnja kaznena funkcija
pen=(ogr1+abs(ogr1))^2/2;
f=fun+kaz*pen;
```

Sada se izvršava program:

```
>> FunkcijaCilja=@FunkcijaKaznena;
BrojVarijabli=2;
x0=[5 5];%spoljasnja pocetna tacka
options=optimset('LargeScale','off');
[x,fval,exitflag,output]=fminunc(FunkciljaCilja,x0,options)
```

Za razne vrednosti kaznenog parametra dobijaju se sledeći rezultati, prikazani u Tabeli 4.:

tačka	kaznena konstanta	optimum $[x_1, x_2]$	funkcija cilja	iznos ograničenja
0		[5 , 5]	17	početna tačka
1	0.0001	[2 , 3]	4	2
2	0.1	[1.71 , 2.71]	4.57	1.42
3	1	[1.2 , 2.2]	5.6	0.4
4	10	[1.02 , 2.02]	5.95	0.04
5	1000	[1.0002 , 2.0002]	5.99	0.0004

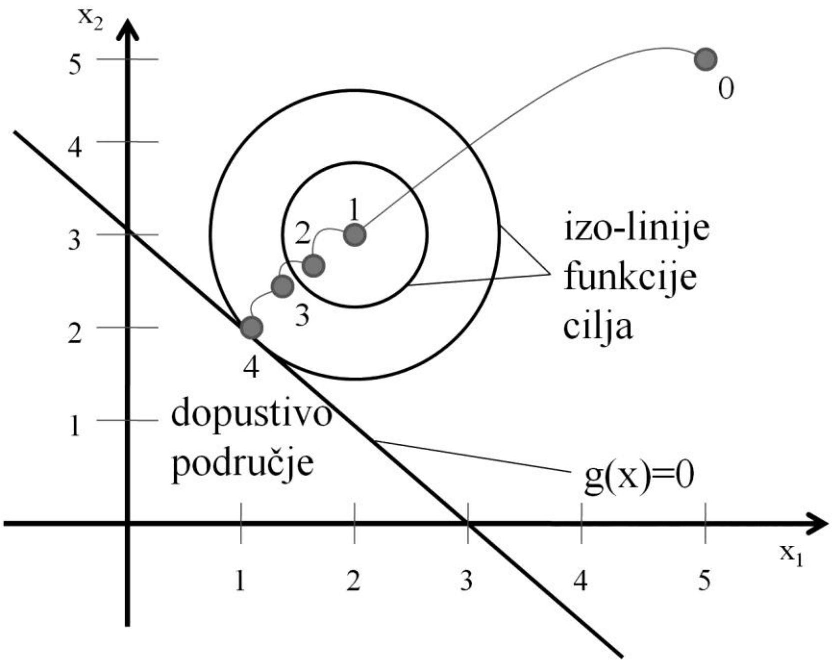
Tabela 4. Rezultati dobijeni metodom spoljašnjih kaznenih funkcija

Iz tabele se vidi da je optimalno rešenje

$$x^* = \lim_{k\to\infty} x_k = [1,2]^T$$

i optimalna vrednost funkcije cilja je $f^* = 6$.

Geometrijska interpretacija toka traženja optimuma data je na slici 3.



Slika 3. Tok traženja optimuma

Primer 11. Rešiti problem

$$\min(x_1^2 x_2^2)$$

$$x_1 + x_2 = 0$$
$$1 - x_1 - x_2^2 = 0$$

metodom spoljašnjih kaznenih funkcija, gde je $x = [x_1, x_2]^T \in R^2$.

<u>Rešenje:</u>

Modifikovana funkcija cilja je oblika

$$F_k(x_1, x_2) = x_1^2 x_2^2 + t_k(x_1 + x_2)^2 + t_k(1 - x_1 - x_2^2)^2.$$

Problemi

$$\min F_k(x_1, x_2), \qquad x \in R^2,\ k = 1,2, \dots\ .$$

se rešavaju približno koristeći funkciju **FindMinimum**[·] u programskom paketu *Mathematica*®.

Uzimajući pozitivan rastući niz parametara kazne, dobija se

za $t_k = 1$
FindMinimum[(x^2)*(y^2)+(x+y)^2+(1-x-y^2)^2, {x,0},{y,0}]
{0.113106,{x→0.541556,y→-0.550815}}

za $t_k = 10$
FindMinimum[(x^2)*(y^2)+10*(x+y)^2+10*(1-x-y^2)^2, {x,0},{y,0}]
{0.141549,{x→0.607655,y→-0.609932}}

za $t_k = 100$
FindMinimum[(x^2)*(y^2)+100*(x+y)^2+100*(1-x-y^2)^2, {x,0},{y,0}]
{0.145448,{x→0.616957,y→-0.617204}}

za $t_k = 1\,000$
FindMinimum[(x^2)*(y^2)+1000*(x+y)^2+1000*(1-x-y^2)^2, {x,0},{y,0}]
{0.145853,{x→0.617926,y→-0.617951}}

za $t_k = 10\,000$
FindMinimum[(x^2)*(y^2)+10000*(x+y)^2+10000*(1-x-y^2)^2, {x,0},{y,0}]
{0.145894,{x→0.618023,y→-0.618026}}

za $t_k = 100\,000$
FindMinimum[(x^2)*(y^2)+100000*(x+y)^2+100000*(1-x-y^2)^2, {x,0},{y,0}]
{0.145898,{x→0.618033,y→-0.618033}}

za $t_k = 1\,000\,000$
FindMinimum[(x^2)*(y^2)+1000000*(x+y)^2+1000000*(1-x-y^2)^2, {x,0},{y,0}]
{0.145898,{x→0.618034,y→-0.618034}}

za $t_k = 10\,000\,000$
FindMinimum[(x^2)*(y^2)+10000000*(x+y)^2+10000000*(1-x-y^2)^2, {x,0},{y,0}]
{0.145898,{x→0.618034,y→-0.618034}}

Sledi da je optimalno rešenje dobijeno kada je $t_k = 10\,000\,000$

$$x_1^* \approx 0.618033$$
$$x_2^* \approx -0.618033$$

i optimalna vrednost funkcije cilja je

$$f^* \approx 0.145898$$

pri tome su ograničenja zadovoljena sa greškom reda 10^{-6}.

4.2 Metoda unutrašnjih kaznenih funkcija (barijernih funkcija)

Metoda unutrašnjih kaznenih funkcija, nasuprot postupku spoljašnih kaznenih funkcija, zahteva dopustivu, "unutrašnju" početnu tačku. Kada se tokom postupka traženja optimuma trenutna tačka približi ograničenju, iznos modifikovane funkcije cilja $B_k(x)$ mora značajno da poraste zbog kaznenog dodatka, da bi se postupak traženja vratio u dopustivo područje. Zbog toga, odabrana kaznena funkcija za ograničenja nejednakosti u dopustivom području mora imati zanemarljiv ("mali") doprinos iznosu funkcije cilja, a progresivno veliki iznos u blizini rubova dopustive oblasti. Unutrašnja kaznena funkcija na ovaj način formira takozvanu barijeru na rubovima dopustivog područja, koja sprečava da postupak traženja izađe iz dopustivog skupa. Dakle, kazneni dodatak predstavlja barijeru protiv izlaska iz dopustive oblasti. Ova metoda generiše niz dopustivih tačaka, čija granica predstavlja optimalno rešenje originalnog problema.

Neka je dat problem

$$\min_{x \in X \subseteq R^n} f(x) \tag{24}$$

gde je X dopustiv skup.

Pri rešavanju problema (24) konstruiše se konvergentan niz $\{x^k\}$, $k = 1,2, \ldots$ čiji elementi aproksimiraju rešenje, a leže van nekog zabranjenog skupa $V \subset X$. Pri tome se može desiti da je $\lim_{k\to\infty} x^k = x^* \in V$.

Ukoliko je granica skupa X zadata komplikovanim izrazima potrebna je izolacija nekog podskupa skupa X, jer bi u suprotnom računanje tačaka x^k zahtevalo dosta vremena, a moglo bi usporiti ili dovesti u pitanje konvergenciju. Ideja je da se optimalnoj tački x^* približava kroz skup $X \setminus V$, a kada se dođe u blizinu tačke x^*, ukoliko ona pripada skupu V omogućava se „preskakanje" barijere.

Definicija 20. Neka je $V \subset X$ i neka je funkcija $B(x)$ definisana na skupu $X \setminus V \neq \emptyset$ sa nenegativnim i konačnim vrednostima. Funkcija $B(x)$ se naziva ***barijera*** ili ***barijerna funkcija*** skupa V, ukoliko za svaki niz $\{v^k\} \in X \setminus V$ koji konvergira ka nekoj tački $v \in V$, važi

$$\lim_{k\to\infty} B(v^k) = +\infty.$$

Umesto problema (24) rešava se niz problema oblika

$$\min_{x \in X \setminus V} B_k(x), \quad k = 1,2, \ldots \tag{25}$$

gde je

$$B_k(x) = f(x) + a_k B(x)$$

i a_k je niz pozitivnih brojeva sa osobinom $\lim_{k\to\infty} a_k = 0$.

Problem (25) je takođe problem sa ograničenjima, ali se može rešiti nekom metodom za optimizaciju bez ograničenja, polazeći od tačke koja je unutrašnja. Pošto funkcija cilja teži ka beskonačnosti u blizini granice dopustivog skupa, minimum funkcije $B_k(x)$ će pripadati unutrašnjosti skupa X. Zato je sa računarske tačke gledišta problem (25) ustvari problem bez ograničenja.

Neka je $\min_{x\in X\setminus V} B_k(x) = B_{k_*}$. Kako je $B_k(x) \geq f(x) \geq f^*$za $x \in X \setminus V$, sledi da je $B_{k_*} \geq f^*$. Ako je $f^* > -\infty$, problemi (25) imaju rešenje. To rešenje može biti određeno primenom neke metode koja daje približno rešenje, pa se niz x^k može definisati sa

$$x^k \in X \setminus V, \quad B_k(x^k) \leq B_{k_*} + \varepsilon_k, \quad k = 1,2,\ldots$$

gde je $\varepsilon_k \geq 0$ i $\lim_{k\to\infty} \varepsilon_k = 0$ (ε_k je greška metode).

Zahtev da skup dopustivih rešenja sadrži unutrašnje tačke je od velike važnosti za konvergenciju metode. Očigledno je da ovo isključuje mnoge probleme koji sadrže ograničenja oblika jednakosti.

Definiše se dopustiv skup X na sledeći način

$$X = \{x \in R^n,\ g_i(x) \leq 0, i = 1, \ldots, m\}.$$

Tada se unutrašnjost skupa X može označiti sa $intX$, tj.

$$intX = \{x \in R^n,\ g_i(x) < 0, i = 1, \ldots, m\},$$

a sa ∂X granica skupa X ($\partial X = X \setminus intX$).

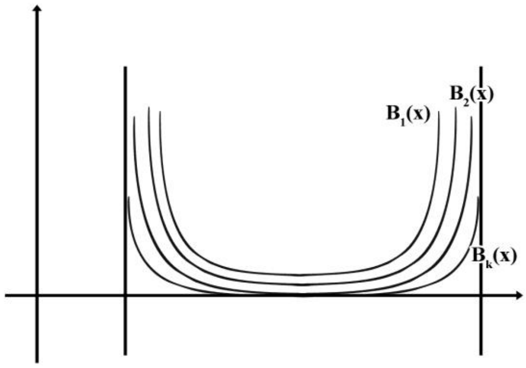
Barijerne funkcije mogu biti definisane na neki od sledećih načina

$$B(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)} \tag{26}$$

$$B(x) = \sum_{i=1}^{m} \max\left\{-\ln\left(-\ g_i(x)\right), 0\right\}^p,\ \ p \geq 1. \tag{27}$$

Obično se za p uzima vrednost 1.

Geometrijska interpretacija niza unutrašnjih kaznenih funkcija $\{B_k\}$ data je na slici 4.



Slika 4. Niz unutrašnjih kaznenih funkcija

Sledi algoritam za rešavanje problema optimizacije sa ograničenjim pomoću metode unutrašnjih kaznenih funkcija.

Algoritam 6. Metoda barijernih funkcija

Korak 1. Neka je $\{a_k\}$, $k = 1, 2$ ... niz koji teži ka nuli, tako da je za svako k, $a_k \geq 0$, $a_{k+1} < a_k$.

Korak 2. Za svako $k = 1, 2$ ... definiše se funkcija

$$B_k(x) = f(x) + a_k B(x)$$

i rešava problem

$$\min B_k(x), x \in intX$$

odakle se dobija rešenje x^k.

Sledi teorema o konvergenciji.

Teorema 22. [6] Neka je $V \subset X$, $X \setminus V \neq \emptyset$, $i\ f^* = f^{**}$, gde je

$$f^* = \inf_{x \in X} f(x), \quad f^{**} = \inf_{X \setminus V} f(x) > -\infty .$$

Neka je niz $\{x^k\}$ određen tako da je $B_k(x^k) = B_{k_*}$. Tada važi

$$lim_{k\to\infty} B_{k_*} = f^*, \quad lim_{k\to\infty} a_k B(x^k) = 0.$$

Ako je X zatvoren i ograničen skup, a funkcija $f(x)$ poluneprekidna funkcija od dole na X, tada niz $\{x^k\}$ konvergira ka $x^* \in X_*$, gde je X_* skup tačaka minimuma funkcije $f(x)$ na skupu X.

Metoda unutrašnjih kaznenih funkcija zahteva da početna tačka x^0 pripada unutrašnjosti skupa X, što znači da unutrašnjost dopustivog skupa X mora biti neprazna. Određivanje početne tačke otežava primenu ove metode. Međutim, za razliku od metode spoljašnjih kaznenih funkcija, metoda unutrašnjih kaznenih funkcija generiše uvek niz dopustivih tačaka.

Slede primeri za ilustraciju metode unutrašnjih kaznenih funkcija.

Primer 12. Metodom barijernih funkcija, rešiti problem optimizacije sa ograničenjima

$$\min(x_1 + x_2)$$
$$x_2 - x_1^2 \geq 0$$
$$x_1 \geq 0$$

Rešenje:

U rešavanju problema, koristiće se kaznena funkcija oblika (27). Dakle, rešava se problem

$$\min_{x \in R^2} B_k(x)$$

gde je

$$B_k(x) = x_1 + x_2 - a_k(\ln(x_2 - x_1^2) + \ln x_1)$$

i a_k je niz pozitivnih brojeva

$$a_k = \frac{1}{2^k}, \quad k = 1,2, \dots \ .$$

za koje važi $\lim_{k\to\infty} a_k = 0$. Unutrašnjost skupa X je

$$intX = \{[x_1, x_2]^T \in R^2 | x_2 - x_1^2 > 0, x_1 > 0\}$$

Kako je funkcija cilja diferencijabilna, rešenje se dobija rešavanjem sistema jednačina

$$\frac{\partial B_k}{\partial x_1} = 1 - a_k \frac{x_2 - 3x_1^2}{x_1(x_2 - x_1^2)} = 0$$
$$\frac{\partial B_k}{\partial x_2} = 1 - a_k \frac{1}{x_2 - x_1^2} = 0.$$

Za svako $x_1 \neq 0$ i $x_2 - x_1^2 \neq 0$ ovaj sistem se svodi na sistem

$$x_1(x_2 - x_1^2) - a_k(x_2 - 3x_1^2) = 0$$
$$a_k = x_2 - x_1^2.$$

Kada se zameni izraz $x_2 - x_1^2$ sa a_k u prvu jednačinu sistema, dobija se

$$2x_1^2 a_k + x_1 a_k - {a_k}^2 = 0.$$

Iz uslova $x_1 \geq 0$ sledi

$$x_1 = \frac{-1 + \sqrt{1 + 8a_k}}{4}.$$

Za različite k rešenje sistema je niz parova $[x_{1k}, x_{2k}]^T$, gde je

$$x_{1k} = \frac{-1 + \sqrt{1 + 8a_k}}{4}, \qquad x_{2k} = x_{1k}^2 + a_k.$$

Pošto je $\lim_{k\to\infty}[x_{1k}, x_{2k}]^T = [0,0]^T$, sledi da je optimalno rešenje problema $x^* = [0,0]^T$, a vrednost funkcije cilja $f^* = 0$. Kako bi se proverilo da je ovo zaista rešenje, potrebno je proveriti uslove Teoreme 21. Zaista,

$$\langle \nabla f(x^*), x - x^* \rangle = x_1 + x_2 \geq x_1 + x_1^2 \geq 0$$

za $x \in X = \{x \in R^2 | x_2 - x_1^2 \geq 0, x_1 \geq 0\}$, pa sledi da je tačka $x^* = [0,0]^T$ rešenje problema.

Primer 13. Metodom barijernih funkcija rešiti problem

$$\min(x_1 + \frac{(4 + x_2)^3}{3})$$
$$x_1 \geq 1$$
$$x_2 \geq 4.$$

Rešenje:

Ograničenja se mogu zapisati u obliku $g_i \leq 0$, tj.

$$1 - x_1 \leq 0$$
$$4 - x_2 \leq 0.$$

Funkcija cilja se modifikuje uvođenjem kaznenih funkcija oblika (26)

$$B_k(x) = x_1 + \frac{(4 + x_2)^3}{3} - a_k\left(\frac{1}{1 - x_1} + \frac{1}{4 - x_2}\right).$$

pa se rešava niz problema minimizacije bez ograničenja oblika

$$\min_{x \in R^2} B_k(x).$$

Problem se svodi na rešavanje sistema jednačina

$$\frac{\partial B_k}{\partial x_1} = 1 - \frac{a_k}{(1-x_1)^2} = 0$$
$$\frac{\partial B_k}{\partial x_2} = (4+x_2)^2 - \frac{a_k}{(4-x_2)^2} = 0,$$

čije rešenje je

$$x_{1k} = 1 - \sqrt{a_k}$$
$$x_{2k} = \left(16 - \sqrt{a_k}\right)^{1/2}.$$

Kako je $\lim\limits_{k\to\infty} a_k = 0$, dobija se rešenje početnog problema

$$x_1^* = 1$$
$$x_2^* = 4.$$

Kako je

$$\langle \nabla f(x^*), x - x^* \rangle = \langle \begin{bmatrix} 1 \\ 64 \end{bmatrix}, \begin{bmatrix} x_1 - 1 \\ x_2 - 4 \end{bmatrix} \rangle = (x_1 - 1) + 64(x_2 - 4) \geq 0$$

sledi da je $[x_1^*, x_2^*]^T = [1,4]^T$ rešenje problema, a vrednost funkcije cilja u rešenju je

$$f^* = 171.6.$$

Kada se među ograničenjima problema nalaze i ograničenja oblika jednakosti, skup dopustivih rešenja nema „unutrašnjost“, pa se metoda unutrašnjih kaznenih funkcija ne može primeniti za rešavanje takvog problema. Dakle, posmatra se problem oblika

$$\min f(x)$$
$$g_i(x) \leq 0, i = 1, \dots, p$$
$$h_j(x) = 0, j = 1, \dots, q.$$

U ovom slučaju može se koristiti metoda mešovitih kaznenih funkcija data u [8].

Ukoliko ograničenja tipa nejednakosti imaju unutrašnju tačku, tada se ona tretiraju kao da se radi o metodama unutrašnjih kaznenih funkcija, dok se ograničenja tipa jednakosti posmatraju kao u metodama spoljašnjih kaznenih funkcija. Dakle, u ovim slučajevima, proširena funkcija cilja problema sa ograničenjima ima sledeći oblik

$$\phi_k(x) = f(x) + t_k q(x) + a_k \sum_{j=1}^{q} \left| h_j(x) \right|^p,$$

gde je $q(x)$ jedna od već pomenutih kaznenih funkcija, a $p \geq 1$.

5 Kombinacija metoda kaznenih funkcija i metoda optimizacije bez ograničenja

Već su pomenuta dva suštinski važna pitanja kod kaznenih funkcija i to su: koliko dobro pomoćni problem bez ograničenja aproksimira polazni problem sa ograničenjima, i kako rešiti pomoćni problem bez ograničenja čija funkcija cilja sadrži kazneni član.

Sa povećavanjem kaznenog parametra t_k (kod spoljašnjih kaznenih funkcija), i smanjivanjem barijernog koeficijenta a_k (kod barijernih funkcija) kad $k \to \infty$ poboljšava se aproksimacija polaznog problema, ali struktura pomoćnog problema postaje mnogo složenija i komplikovanija što može usporiti konvergenciju algoritma koji se koristi za rešavanje pomoćnog problema bez ograničenja. Upravo iz tog razloga direktna primena Košijeve metode najbržeg pada nije preporučljiva, ali se zato Njutnov postupak i njegove modifikacije uspešno mogu primeniti za rešavanje pomoćnog problema optimizacije bez ograničenja kao i metoda konjugovanih gradijenata sa restartom koja je veoma pogodna u tu svrhu.

U narednim poglavljima predstavljene su metode za rešavanje pomoćnog problema bez ograničenja koji sadrži kazneni član.

5.1 Kombinacija metode Lagranžovih množitelja i kaznenih funkcija

Neka je dat polazni problem sa ograničenjima oblika

$$\begin{gathered} \min f(x) \\ g_i(x) \leq 0, i = 1,2, \dots, m. \end{gathered} \tag{28}$$

Ako dopustiv skup sadrži ograničenja tipa jednakosti, ona se uvek mogu predstaviti preko nejednakosti. Najopštija klasa spoljašnjih kaznenih funkcija za problem (28) je

$$q(x) = \gamma(g^+(x)) \tag{29}$$

gde je γ neprekidna funkcija $\gamma: R^m \to R$, a $g_i^+(x) = \max\{0, g_i(x)\}, i = 1, 2, \dots, m$ su komponente od $g^+(x)$.

Na primer, za

$$q(x) = \frac{1}{2}\sum_{i=1}^{m} g_i^+(x)^2 = \frac{1}{2}|g^+(x)|^2$$

važi

$$\gamma(y) = \frac{1}{2}|y|^2.$$

Rečeno je da se problem (28) aproksimira nizom problema oblika

$$\min_{x \epsilon X_0 = R^n} F_k(x)\,, k = 1, 2, \ldots \tag{30}$$

gde je $F_k(x) = f(x) + t_k q(x)$, a $q(x)$ je dato sa (29).

Većina algoritama za rešavanje ovog pomoćnog problema bez ograničenja (30) zahteva da funkcija cilja bude klase $C^1(R^n)$, pa je zato potrebno da $q(x)\,\epsilon\, C^1(R^n)$. Pretpostavlja se da su $f, g\,\epsilon\, C^1(R^n)$.

Neka je

$$\nabla g_i^+(x) = \begin{cases} \nabla g_i(x), & za\ g_i(x) \geq 0 \\ 0, & za\ g_i(x) < 0 \end{cases}$$

a $\nabla g^+(x)$ je matrica $m \times n$ sa vrstama $\nabla g_i^+(x), i = 1,2, \ldots, m$. Nažalost, ∇g^+ obično ima prekide u tačkama gde je $g_i^+(x) = 0$, za neko $i = 1,2, \ldots, m$, tj. u aktivnim ograničenjima. Da bi $q(x)\,\epsilon\, C^1(R^n)$, potrebno je da $\gamma\,\epsilon\, C^1(R^n)$. Pretpostavlja se da ako

$$y = [y_1, y_2, \ldots, y_n]^T, \qquad \nabla y = [\nabla \gamma_1, \nabla \gamma_2, \ldots, \nabla \gamma_n]^T$$

tada

$$y_i = 0 \quad \Longrightarrow \quad \nabla \gamma_i = 0. \tag{31}$$

Uz ovu pretpostavku važi da $\gamma\,\epsilon\, C^1(R^n)$ i izvod od $\gamma(g^+(x))$ po x je neprekidan oblika

$$\nabla \gamma\big(g^+(x)\big) \nabla g(x).$$

Na ovaj način je obezbeđeno $F_k(x) \epsilon\, C^1(R^n)$.

Jedan od načina za rešavanje pomoćnog problema bez ograničenja (30) je metoda Lagranžovih množitelja. Pošto su zadovoljeni uslovi za njenu primenu, za fiksno k problem (30) ima rešenje x^k i važi

$$\nabla f(x^k) + t_k \nabla \gamma\Big(g^+(x^k)\Big) \nabla g(x^k) = 0,$$

što se može zapisati kao

$$\nabla f(x^k) + (\lambda^k)^T \nabla g(x^k) = 0,$$

gde je

$$(\lambda^k)^T = t_k \nabla \gamma\Big(g^+(x^k)\Big) \tag{32}$$

vektor Lagranžovih množitelja.

Dakle, za svako k rešenje problema (30) je vektor x^k, kome se pridružuje vektor Lagranžovih množitelja $(\lambda^k)^T$ koji je takođe određen nakon rešavanja problema (30).

Ako je x^* regularno rešenje polaznog problema (28), tada postoji jedinstven vektor Lagranžovih množitelja λ^* koji odgovara rešenju x^*.

Definicija 21. Tačka $\bar{x}$ koja zadovoljava ograničenja problema (28) je regularna tačka ako su gradijenti aktivnih ograničenja u toj tački linearno nezavisni.

Naredna teorema tvrdi da λ^k konvergira ka λ^*.

Teorema 23. [2] Neka je $\{x^k\}$ niz generisan metodom spoljašnjih kaznenih funkcija za rešavanje problema (28), sa kaznenom funkcijom oblika (29), gde je $\gamma \epsilon C^1(R^n)$ za koje važi (31). Neka su $(\lambda^k)^T = t_k \nabla\gamma\left(g^+(x^k)\right)$ odgovarajući vektori Lagranžovih množitelja za $x^k, k = 1, 2, \dots$. Tada, ako $\{x^k\}$ konvergira ka regularnom rešenju x^* polaznog problema (28), onda $\{\lambda^k\}$ konvergira ka vektoru λ^* Lagranžovih množitelja koji odgovara rešenju x^*.

Važno je napomenuti da u opštem slučaju ako $\{x^k\}$ konvergira ka x^*, tada pošto $\{\lambda^k\}$ konvergira ka λ^*, niz $\{x^k\}$ konvergira ka x^* preko spoljašnjosti dopustivog skupa.

5.2 Struktura Hesijana

Pomoćni problem (30) može se rešiti i drugim metodama za optimizaciju bez ograničenja. Određivanje Hesijana funkcije $F_k(x)$ i poznavanje njegove strukture je važno za primenu Njutnovog postupka i određivanje brzine konvergencije postupka najbržeg pada, pa sledi analiza strukture Hesijana koja je predstavljena u [2].

Pretpostavlja se i dalje da je $q(x) \epsilon C^1(R^n)$, ali ne i $C^2(R^n)$. Štaviše, drugi izvod najpoznatije kaznene funkcije

$$q(x) = \frac{1}{2}|g^+(x)|^2$$

ima prekide kada je $g_i(x) = 0$ za neko i, što znači da Hesijan ima prekide na rubu dopustive oblasti, gde se najčešće nalazi rešenje polaznog problema. Međutim, pošto niz generisan metodom spoljašnjih kaznenih funkcija konvergira ka rešenju preko spoljašnjosti dopustivog skupa, to znači da je u tačkama niza x^k Hesijan najčešće dobro definisan, pa se može primeniti standardna analiza.

Kako je $F_k(x) = f(x) + t_k q(x)$, gde je $q(x)$ dato sa (29), sledi da je Hesijan funkcije $F_k(x)$ u oznaci Q_k oblika

$$Q_k(x) = \nabla^2 f(x) + t_k \nabla\gamma\big(g^+(x)\big)G(x) + t_k \nabla g^+(x)^{\mathrm{T}}\Gamma\big(g^+(x)\big)\nabla g^+(x)$$

gde su G i Γ Hesijani od g i γ respektivno. Za fiksirano t_k, koristeći (32) i Hesijan Lagranžove funkcije

$$L_k(x^k) = \nabla^2 f(x^k) + (\lambda^k)^T G(x^k)$$

dobija se da je Hesijan funkcije $F_k(x)$ oblika

$$Q_k(x^k) = L_k(x^k) + t_k \nabla g^+(x^k)^{\mathrm{T}} \Gamma\Big(g^+(x^k)\Big) \nabla g^+(x^k).$$

Prvi sabirak $L_k(x^k)$ konvergira ka Hesijanu Lagranžove funkcije pridružene polaznom problemu (28), kad $x^k \to x^*$, i ta granična vrednost ne zavisi od kaznenog koeficijenta t_k. Drugi sabirak Hesijana $Q_k(x^k)$ je matrica ranga koji je jednak broju aktivnih ograničenja u rešenju x^* i koja teži ka beskonačnosti kad $t_k \to \infty$.

Teorema 24. [2] Neka je x^* regularno rešenje polaznog problema (28). Ako postoji r aktivnih nedegenerisanih ograničenja u x^*, tada Hesijan $Q_k(x^k)$ ima r karakterističnih korena koji teže ka beskonačnosti kad $t_k \to \infty$. Ostalih $(n-r)$ karakterističnih korena, iako zavise od kaznenog koeficijenta t_k, teže ka konačnim vrednostima i to baš ka karakterističnim korenima matrice L_M koja predstavlja restrikciju Hesijana Lagranžove funkcije $L(x^*)$ na tangentni podprostor M aktivnih ograničenja.

Dokaz ovoga zahteva složenu analizu koja je ovde izostavljena, a detaljno je data u [2].

Ovakva specifična struktura čini Hesijan loše uslovljenom matricom što predstavlja problem za primenu određenih postupaka, a pre svega postupka najbržeg pada. Naime, brzina konvegencije metode najbržeg pada zavisi od uslovnog broja Hesijana u rešenju x^*. Ako je uslovni broj Hesijana veliki, tj. ako je Hesijan loše uslovljena matrica, to dovodi do velikog usporavanja metode najbržeg pada. Kako je to ovde slučaj, znači da metoda najbržeg pada nije pogodna za rešavanje pomoćnog problema optimizacije bez ograničenja (30).

Analogna diskusija važi ako se polazni problem (28) rešava metodom barijernih funkcija. I tada je Hesijan loše uslovljen i ima neprikladnu strukturu.

5.3 Kombinacija Njutnove metode i kaznenih funkcija

Jedan od načina za efikasno rešavanje pomoćnog problema bez ograničenja (30) je primena Njutnovog postupka ili neke njegove modifikacije. Ovaj postupak je kvadratno konvergentan, ali nepogodna struktura Hesijana ne utiče na njegovu brzinu. Međutim, s obzirom da je Hesijan loše uslovljen, posebna pažnja mora se posvetiti načinu određivanja inverznog Hesijana.

Posmatra se polazni problem oblika

$$\begin{gathered} \min f(x) \\ h_i(x) = 0, \; i = 1, 2, \dots, m \;, m < n \end{gathered} \tag{33}$$

i umesto njega rešava se problem bez ograničenja

$$\min_{x \epsilon X_0 = R^n} F_k(x), \qquad k = 1, 2, \dots$$

gde je $F_k(x) = f(x) + t_k q(x)$, a $q(x) = \frac{1}{2}|h(x)|^2$.

Za rešavanje ovog problema bez ograničenja pogodan je modifikovan Njutnov postupak predstavljen u [2], oblika

$$x^{k+1} = x^k - \alpha_k[\mathrm{I} + t_k \nabla h(x^k)^T \nabla h(x^k)]^{-1} \nabla \mathrm{F_k}(x^k)^{\mathrm{T}},$$

gde je α_k izabrano minimizacijom $F_k(x^{k+1})$. Matrica $\mathrm{I} + t_k \nabla h(x^k)^T \nabla h(x^k)$ je pozitivno definitna i može se efikasno invertovati, iako je loše uslovljena.

Na osnovu teoreme o konvergenciji modifikovanog Njutnovog postupka, brzina konvergencije ovog postupka je određena karakterističnim korenima matrice

$$[\mathrm{I} + t_k \nabla h(x^k)^T \nabla h(x^k)]^{-1} Q_k(x^k), \tag{34}$$

gde je $Q_k(x^k)$ Hesijan funkcije $F_k(x)$ u x^k.

Kad $t_k \to \infty$, matrica (34) ima m karakterističnih korena koji teže ka istoj vrednosti B, a preostalih $(n-m)$ karakterističnih korena teže ka karakterističnim korenima matrice L_M u rešenju x^* problema (33). Neka su $\bar{a}$ i $\bar{A}$ najmanji i najveći karakteristični koreni matrice L_M. Ako interval $[\bar{a}\,, \bar{A}]$ sadrži vrednost B, tada će ovaj postupak kad $t_k \to \infty$ konvergirati sa koeficijentom brzine $\left(\frac{\bar{A}-\bar{a}}{\bar{A}+\bar{a}}\right)^2$. U suprotnom će konvergencija biti spora.

Ukoliko su dostupne informacije drugog reda, Njutnov metod je efikasan za primenu kod kaznenih funkcija jer je brz. Međutim, ako ove informacije nisu dostupne, onda se koriste postupci prvog reda.

5.4 Kombinacija metode konjugovanih gradijenata i kaznenih funkcija

Metoda konjugovanih gradijenta sa restartom je posebno pogodna za rešavanje pomoćnog problema bez ograničenja (30), pogotovo kada polazni problem ima malo aktivnih ograničenja. Ako postoji m aktivnih ograničenja, tada će brzina konvergencije postupka sa $(m+1)$ konjugovano gradijentnih koraka biti nezavisna od kaznenog koeficijenta t_k.

Posmatra se problem koji ima samo ograničenja tipa jednakosti

$$\begin{gathered} \min f(\boldsymbol{x}) \\ h_i(x) \;=\; 0, i = 1,2,\ldots, m\,, m < n, \end{gathered} \tag{35}$$

tj. njemu pridružen

$$\min_{x \epsilon X_0 = R^n} F_k(x),\; k = 1, 2, \ldots \tag{36}$$

gde je $F_k(x) = f(x) + t_k q(x)$, a $q(x) = \frac{1}{2}|h(x)|^2$.

Pošto problem (35) ima m aktivnih ograničenja, Hesijan funkcije $F_k(x)$ ima m karakterističnih korena reda veličine t_k, a preostalih $(n-m)$ karakterističnih korena su blizu karakterističnih korena matrice L_M koja odgovara polaznom problemu (35). Dakle, Hesijan funkcije $F_k(x)$ ima m velikih karakterističnih korena. Međutim, ako se postupak konjugovanih gradijenata primeni sa $(m+1)$ koraka, eliminisaće se negativan efekat tih m velikih karakterističnih korena i koeficijent brzine konvergencije neće zavisiti od njih.

Teorema 25. [2] Neka je x^{k+1} dobijeno primenom $(m+1)$ konjugovano gradijentnih koraka iz x^k, i neka je $x^k \to \bar{x}$ gde je $\bar{x}$ rešenje problema (36). Tada niz $\{f(x^k)\}$ konvergira linearno ka $f(\bar{x})$ sa koeficijentom brzine približno jednakom $\left(\frac{\bar{A}-\bar{a}}{\bar{A}+\bar{a}}\right)^2$, gde su $\bar{a}$ i $\bar{A}$ najmanji i najveći karakteristični koren matrice $L_M(\bar{x}\,)$.

Ovaj postupak je posebno efikasan za relativno male vrednosti m, mnogo je efikasniji od metode najbržeg pada, a neznatno skuplji.

U radu [12] dat je postupak koji predstavlja kombinaciju metoda kaznenih funkcija i konjugovanih gradijenata, a koristi se za rešavanje problema optimizacije portfolija.

Neka je dato n investicija sa stopama prinosa $R_j,\ j=1,\dots,n$, i standardnim odstupanjima $\sigma_j, j=1,\dots,n$. Očekivane vrednosti za stope prinosa označene su sa $\mu_j = E(R_j), j=1,\dots,n$. Neka je $x_j, j=1,\dots,n$ udeo j-te investicije u portfoliju. Prinos portfolija je

$$\sum_{j=1}^{n} R_j x_j,$$

a očekivani prinos portfolija je

$$\sum_{j=1}^{n} x_j \mu_j.$$

Neka je $V \in R^{n,n}$ kovarijansna matrica sa elementima v_{ij} koji predstavljaju kovarijansu i-te i j-te investicije, onda je varijansa celog portfolija

$$\sum_{i,j=1}^{n} x_i x_j v_{ij}.$$

Neka je vektor težinskih koeficijenata $x=[x_1,x_2,\dots,x_n]^T$, vektor prinosa $R=[R_1,R_2,\dots,R_n]^T$, vektor očekivanih prinosa $\mu=[\mu_1,\mu_2,\dots,\mu_n]^T$. Tada je $R^T x$ prinos portfolija, $\mu^T x$ očekivani prinos portfolija, a $x^T V x$ varijansa portfolija. Rizik portfolija se meri varijansom.

Pošto se razmatra skup investicija sa idejom da se plasira sav raspoloživi kapital onda treba da važi

$$\sum_{j=1}^{n} x_j = 1.$$

Problem optimizacije portfolija može se formulisati sledećim osnovnim modelom

$$\max\ \mu^T x$$
$$\min x^T V x \tag{37}$$

uz ograničenja

$$e^T x = 1$$
$$0 \leq x \leq 1$$

gde je $e = [1,1,\dots,1]^T$.

Ovaj osnovni model može se transformisati na standardni problem kvadratnog programiranja

$$\min f(x) = -(1-\lambda)\mu^T x + \lambda x^T V x$$

uz ograničenja (38)

$$e^T x = 1$$
$$a \leq x \leq b$$

gde su $a, b \in R^n$ dati vektori koji predstavljaju donju i gornju granicu vektora odluke, a λ je koeficijent averzije prema riziku. Ako je $\lambda = 0$ znači da se maksimizira prinos ne uzimajući u obzir rizik. S druge strane, ako je $\lambda = 1$, tada se minimizira rizik ne uzimajući u obzir prinos.

Za fiksirano $\lambda \in (0,1)$, problem (38) je problem kvadratnog programiranja. Pošto je kovarijansna matrica pozitivno semidefinitna, problem (38) je problem konveksnog kvadratnog programiranja. Postoje mnogobrojni postupci koji se koriste za rešavanje ovakvih problema i oni su efikasni u slučaju kad je kovarijansna matrica V retka matrica, ali većina njih nije pogodna za primenu u slučaju kada V nije retka. Postupak konjugovanih gradijenata ne zahteva izračunavanje Hesijana već samo gradijenata funkcije cilja, a pošto se koristi za rešavanje problema optimizacije bez ograničenja, prvi korak jeste da se problem (38) transformiše u takav oblik.

Neka je

$$c = [c_1, c_2, \dots, c_n]^T, \quad c = -(1-\lambda)\mu, \qquad Q = \left[q_{ij}\right]_{n\times n} = 2\lambda V.$$

Tada je problem (38) ekvivalentan sledećem problemu

$$\min f(x) = c^T x + \frac{1}{2} x^T Q x$$

uz ograničenja (39)

$$e^T x = 1$$
$$a \leq x \leq b.$$

Pošto je matrica kovarijanse V simetrična i pozitivno semidefinitna, to takođe važi i za matricu Q, pa znači da je $f(x)$ konveksna funkcija.

Za ograničenje tipa jednakosti $e^T x = 1$ i ograničenja tipa nejednakosti $a \leq x \leq b$ definiše se kaznena funkcija $q_k : R^{n+1} \to R$

$$q_k(x) = \frac{t_k}{2}[(e^T x - 1)^2 + ||\min\{x - a, 0\}||^2 + ||\min\{b - x, 0\}||^2]$$

gde je $t_k > 0$ kazneni parametar, a $||\cdot||$ označava vektorsku normu 2. Ako je x dopustivo rešenje problema (39), onda je

$$q_k(x) = 0.$$

Problem (39) sa ograničenjima se svodi na niz problema bez ograničenja

$$F_k(x) = c^T x + \frac{1}{2}x^T Q x + \frac{t_k}{2}[(e^T x - 1)^2 + \|\min\{x - a, 0\}\|^2 + \|\min\{b - x, 0\}\|^2].$$

Funkcija $F_k(x)$ je deo po deo kvadratna funkcija, po delovima neprekidno diferencijabilna, i ako je Q pozitivno semidefinitna, onda je F_k po delovima konveksna kvadratna funkcija.

$F_k(x)$ ima sledeću, kompaktniju formu:

$$F_k(x) = \bar{c}_0 + \bar{c}(x; t_k)^T x + \tfrac{1}{2}x^T \bar{Q}(x; t_k)x \qquad (40)$$

gde je $\bar{c}_0$ konstanta.

Teorema 26. [12] Neka je dat niz $\{t_k\}$, takav da $t_k \to +\infty$ kad k $\to +\infty$. Neka je x^k tačan globalni minimum za $F_k(x)$. Tada je svaka tačka nagomilavanja x^* niza $\{x^k\}$ rešenje problema (38).

<u>*Dokaz*:</u>

Neka je $\bar{x}$ globalno rešenje problema (38). Onda, za svaku dopustivu tačku x, važi

$$f(\bar{x}) \leq f(x).$$

Kako je x^k tačan globalni minimum funkcije $F_k(x)$ za fiksno t_k, sledi da

$$F_k(x^k) \leq F_k(\bar{x}),$$

pa prema tome važi

$$f(x^k) + \frac{t_k}{2}[(e^T x^k - 1)^2 + ||\min\{x^k - a, 0\}||^2 + ||\min\{b - x^k, 0\}||^2] \leq$$
$$f(\bar{x}) + \tfrac{t_k}{2}[(e^T \bar{x} - 1)^2 + ||\min\{\bar{x} - a, 0\}||^2 + ||\min\{b - \bar{x}, 0\}||^2] = f(\bar{x}),$$

gde poslednja jednakost sledi iz dopustivosti $\bar{x}$. Odavde sledi

$$(e^T x^k - 1)^2 + ||\min\ \{x^k - a, 0\}||^2 + ||\min\ \{b - x^k, 0\}||^2 \leq \frac{2}{t_k}[f(\bar{x}) - f(x^k)].$$

Neka je x^* tačka nagomilavanja niza $\{x^k\}$. Bez smanjenja opštosti, pretpostavlja se da

$$\lim_{k \to +\infty} x^k = x^*.$$

Kada $k \to \infty$ iz prethodne nejednakosti sledi

$$\begin{aligned} 0 &\leq (e^T x^* - 1)^2 + ||\min\ \{x^* - a, 0\}||^2 + ||\min\ \{b - x^*, 0\}||^2 \\ &\leq \lim_{k \to +\infty} \frac{2}{t_k}[f(\bar{x}) - f(x^k)] = 0, \end{aligned}$$

gde poslednja jednakost sledi iz $t_k \to +\infty$. Odavde je

$$e^T x^* = 1, \quad a \leq x^* \leq b$$

pa sledi da je x^* dopustiva tačka.

U nastavku se dokazuje da je x^* globalni minimum problema (38).

Kako je

$$\begin{aligned} f(x^*) &\leq f(x^*) + \lim_{k \to +\infty} \frac{t_k}{2}[(e^T x^k - 1)^2 + ||\min\ \{x^k - a, 0\}||^2 + ||\min\ \{b - x^k, 0\}||^2] \\ &\leq f(\bar{x}) \end{aligned}$$

sledi da je x^* globalni minimum funkcije f.

Teorema 27. [12] Neka je x^* rešenje problema (38). Onda je x^* globalni minimum za $F_k(x)$ za svako t_k.

Na osnovu prethodne dve teoreme zasnovan je algoritam za rešavanje problema (38).

Kao što je već rečeno, metod konjugovanih gradijenata je prikladan za rešavanje problema optimizacije bez ograničenja jer ne zahteva puno memorije i nije skup.

Pošto je broj investicija obično veliki, reč je o problemu velikih dimenzija i kako matrica $\bar{Q}(x; t_k)$ najčešće nije retka matrica, metod konjugovanih gradijenata je pogodan za rešavanje problema (40). Međutim, funkcija $F_k(x)$ nije klasična kvadratna funkcija, pa se standardne procedure za minimiziranje kvadratne funkcije ne mogu direktno primeniti. U datom radu formulisan je novi algoritam za rešavanje problema (40). Funkcija F_k se modifikuje na sledeći način.

Što se tiče koeficijenata kvadratnih članova u

$$\frac{t_k}{2}[||\min\{x-a,0\}||^2 + ||\min\{b-x,0\}||^2]$$

matrica $Q = [q_{ij}]$ se modifikuje prema sledećem pravilu:

$$q_{ij} = \begin{cases} q_{ij}, & \text{ako } i \neq j, \\ q_{ii}, & \text{ako } i = j, a_i \leq x_i \leq b_i, \\ (q_{ii} + t_k), & \text{ako } i = j, \ a_i > x_i \text{ ili } x_i > b_i. \end{cases} \tag{41}$$

Što se tiče koeficijenata linearnih članova u

$$\frac{t_k}{2}[||\min\{x-a,0\}||^2 + ||\min\{b-x,0\}||^2]$$

$c = [c_1, c_2, \dots, c_n]^T$ se modifikuje prema sledećem pravilu:

$$c_i = \begin{cases} c_i, & \text{ako } a_i \leq x_i \leq b_i \\ (c_i - t_k a_i), & \text{ako } x_i < a_i \\ (c_i - t_k b_i), & \text{ako } x_i > b_i. \end{cases} \tag{42}$$

Dalje, definiše se

$$\bar{Q} \triangleq Q + t_k e e^T, \quad \bar{c} \triangleq c - t_k e.$$

Metod konjugovanih gradijenata biće upotrebljen za minimizaciju kvadratne funkcije:

$$\min F_k(x) = c_0 + \bar{c}^T x + \frac{1}{2} x^T \bar{Q} x, \tag{43}$$

gde je t_k dati kazneni parametar. Lako je uočiti da

$$\nabla F_k(x) = \bar{Q}x + \bar{c}.$$

Iako postoje razne varijante metode konjugovanih gradijenata za rešavanje (43), sve one su bazirane na sledeća dva koraka.

1. Odrediti pravac pretraživanja iz trenutne iteracije x^l

$$d^l = \begin{cases} -(\bar{Q}x^0 + \bar{c}) & za\ l = 0, \\ -(\bar{Q}x^l + \bar{c}) + \beta_l d^{l-1} & za\ l \geq 1, \end{cases} \tag{44}$$

gde je β_l izabrano tako da je d^l konjugovani pravac od d^{l-1} u odnosu na matricu $\bar{Q}$.

2. Duž pravca d^l izabrati veličinu koraka α_l tako da u novoj iteraciji

$$x^{l+1} = x^l + \alpha_l d^l$$

apsolutna vrednost funkcije F_k dovoljno opada.

Sledeća lema predstavlja način za određivanje pravca pretraživanja.

Lema 10. [12] Ako je

$$\beta_l = \frac{(\bar{Q}x_l+\bar{c})^T\bar{Q}d^{l-1}}{\left(d^{l-1}\right)^T\bar{Q}d^{l-1}} \tag{45}$$

$$d^l = -(\bar{Q}x^l + \bar{c}) + \beta_l d^{(l-1)},$$

onda je d^l konjugovani pravac od d^{l-1} u odnosu na $\bar{Q}$ i tad se metod naziva HS metod.

U slučaju da je veličina koraka α_l oblika

$$\alpha_l = -\frac{\left(\bar{Q}x^l+\bar{c}\right)^T d^l}{\left(d^l\right)^T\bar{Q}d^l} \tag{46}$$

važi sledeća teorema o globalnoj konvergenciji.

Teorema 28. [12] Neka x^0 je proizvoljna početna iteracija i $\{x^l\}$ niz generisan koracima 1. i 2., tada je

$$\lim_{l\to+\infty} F_l(x^l) = -\infty$$

ili

$$\liminf \nabla F_l(x^l) = 0.$$

Ako je x^* tačka nagomilavanja niza $x^l,\ l = 1, 2, ...$, onda je x^* globalni minimum F_l.

Napomena:

Ako se β_l računa na sledeći način

$$\beta_l = \frac{\|\bar{Q}x^l + \bar{c}\|^2}{\|\bar{Q}x^l + \bar{c}\|^2}$$

tada takođe važi tvrđenje teoreme i tad je u pitanju FR metod.

Algoritam 7. Kazneni algoritam baziran na metodi konjugovanih gradijenata

Korak 1. Date su konstante $t_0 > 1, \lambda \in [0,1], \delta > 0, \varepsilon > 0$ i ρ. Uneti vektor očekivanog prinosa μ i izračunati Q i c. Izabrati početnu iteraciju x^0. Postaviti $k \coloneqq 0,\ l \coloneqq 0,$ i $x^l \coloneqq x^k$.

Korak 2. Ako

$$\|\bar{Q}x^l + \bar{c}\| \leq \varepsilon$$

tada

$$x^k \coloneqq x^l$$

i preći na korak 5; u suprotnom, preći na korak 3.

Korak 3. Odrediti pravac pretraživanja d^l na osnovu (44) i (45).

Korak 4. Izračunati α_l koristeći (46) i odrediti

$$x^l := x^l + \alpha_l d^l$$

Vratiti se na korak 2.

Korak 5. Proveriti dopustivost za x^k kod problema (39). Ako je

$$q_k(x^k) \leq \delta$$

algoritam se završava; u suprotnom, preći na korak 6.

Korak 6. Postaviti $l := 0, x^l := x^k, t_k := \rho t_k$. Na novoj iteraciji x^k, modifikovati matricu Q i vektor c na osnovu (41) i (42). Postaviti $k := k + 1$, i vratiti se na korak 2.

Napomena:

1. U algoritmu, indeks k označava broj ažurirajućih kaznenih parametara, a l označava broj iteracija kod metoda konjugovanih gradijenata za potproblem bez ograničenja (43).

2. Za neko fiksno t_k, lako je uvideti da uslov

$$q_k(x^k) = \frac{t_k}{2}[(e^T x^k - 1)^2 + ||\min\{x^k - a, 0\}||^2 + ||\min\{b - x^k, 0\}||^2] \leq \delta$$

implicira da je x^k dopustivo rešenje. Iz Teoreme 26 sledi da je x^k globalni minimum polaznog problema (38) ako je x^k globalni minimum problema (43).

Numerički eksperimenti

U radu [12] Zhong Wan, ShaoJun Zhang, and YaLin Wang su testirali efikasnost predloženog algoritma na realnim podacima sa kineske berze iz 2007. Računske procedure su implementirane na MATLAB-u 6.5.

Početno rešenje je izabrano tako da važi

$$e^T x^0 = 1,$$

granični vektor a je nula vektor, a b je vektor čije su komponente 1. Početni kazneni parametar je $t_0 = 10$, koeficijent averzije prema riziku $\lambda = 0{,}5$, $\varepsilon = 10^{-7}$ i $\delta = 10^{-4}$.

Algoritam 7 je primenjen na deset realnih problema čije su dimenzije u opsegu od 10 do 100. Očekivane stope prinosa za svaku akciju uzete su iz mesečnih podataka sa berze u Kini 2007. godine.

U Tabeli 5 predstavljeni su numerički rezultati, pri čemu je n dimenzija problema. U trećoj i četvrtoj koloni dato je vreme procesiranja kada je β_l dobijeno metodama HS i FR, k ukazuje na broj kaznenih parametara, t_k je vrednost kaznenog parametra, a $q_k(x^*)$ označava vrednost kaznenog člana.

Problem	n	CPU od HS	CPU od FR	k	t_k	$q_k(x^*)$
1	10	$1''$	$1''$	3	10^4	$2.5652e-005$
2	20	$1''$	$8''$	3	10^4	$1.3314e-005$
3	30	$1''$	$5''$	3	10^4	$5.9817e-005$
4	40	$4''$	$5'8''$	4	10^5	$1.1302e-005$
5	50	$2''$	$11''$	4	10^5	$2.4787e-005$
6	60	$2''$	$46''$	3	10^4	$8.5817e-005$
7	70	$4''$	$12''$	4	10^5	$1.7759e-005$
8	80	$4''$	$1'43''$	4	10^5	$2.8436e-005$
9	90	$2''$	$1'26''$	4	10^5	$9.6836e-005$
10	100	$3''$	$6'48''$	4	10^5	$4.8977e-005$

Tabela 5. Numeričke performanse Algoritma 7

U Tabeli 6 data je lista dobijenog optimalnog rešenja za svaki problem.

	Optimalno rešenje x^*
Problem 1	$x^*(3) = 0.6691;\ x^*(6) = 0.3311;$ ostale komponente x^* su nule
Problem 2	$x^*(18) = 1.0000;$ ostale komponente x^* su nule
Problem 3	$x^*(4) = 0.6985;\ x^*(28) = 0.3021;$ ostale komponente x^* su nule
Problem 4	$x^*(38) = 1.0000;$ ostale komponente x^* su nule
Problem 5	$x^*(38) = 0.8859;\ x^*(49) = 0.1142;$ ostale komponente x^* su nule
Problem 6	$x^*(14) = 1.0000$ ostale komponente x^* su nule
Problem 7	$x^*(51) = 0.8486;\ x^*(61) = 0.1516;$ ostale komponente x^* su nule
Problem 8	$x^*(21) = 0.1175;\ x^*(78) = 0.8827;$ ostale komponente x^* su nule
Problem 9	$x^*(9) = 0.2215;\ x^*(25) = 0.6690;$ ostale komponente x^* su nule
Problem 10	$x^*(67) = 0.3086;\ x^*(79) = 0.5138;\ x^*(97) = 0.1797;$ ostale komponente x^* su nule

Tabela 6. Optimalna rešenja datih deset problema

Numerički rezultati pokazuju efikasnost datog algoritma pri rešavanju realnog problema.

6 Zaključak

U svakodnevnom životu za projektovanje, izgradnju i održavanje nekog inženjerskog sistema, inženjeri moraju da donesu mnoge tehnološke i upravljačke odluke u nekoliko faza. Njihov cilj je povećati dobit i smanjiti troškove, pri čemu se mora voditi računa o raspoloživim resursima. Pošto se maksimiziranje dobiti i minimiziranje troškova može predstaviti kao funkcija više promenljivih, optimizacija se može definisati kao proces pronalaženja uslova koji daju maksimalnu ili minimalnu vrednost funkcije cilja, najčešće nad nekim skupom ograničenja.

Problem optimizacije može biti problem sa ili bez ograničenja. Prisustvo ograničenja u nelinearnom programiranju otežava optimizaciju, u odnosu na programiranje bez ograničenja.

Metode spoljašnjih i unutrašnjih kaznenih funkcija spadaju u jednostavne, često korišćene metode za optimizaciju funkcija sa ograničenjima. Metode spoljašnjih kaznenih funkcija aproksimiraju problem sa ograničenjima nizom problema bez ograničenja sa nametnutom velikom kaznom za tačke daleko od dopustivog skupa. Sa povećavanjem kaznenog parametra ka beskonačnosti, aproksimacija biva tačnija i niz rešenja pomoćnih problema bez ograničenja konvergira ka rešenju polaznog problema sa ograničenjima preko spoljašnjosti dopustivog skupa.

Metode unutrašnjih kaznenih funkcija aproksimiraju problem sa ograničenjima nizom problema (koji su sa kompjuterske tačke gledišta) bez ograničenja sa nametnutom velikom kaznom u blizini granice dopustivog skupa. Za razliku od spoljašnjih kaznenih funkcija, ove metode su primenljive samo kod problema u kojima dopustiv skup ima nepraznu unutrašnjost. Sa smanjivanjem barijernog koeficijenta ka nuli, povećava se preciznost aproksimacije i niz rešenja pomoćnih problema bez ograničenja konvergira ka rešenju polaznog problema preko unutrašnjosti dopustivog skupa.

Osnovna karakteristika ovih metoda je loše uslovljen Hesijan funkcije cilja pomoćnog problema bez ograničenja. Ova matrica je zbir Hesijana Lagranžove funkcije pridružene polaznom problemu i matrice koja teži ka beskonačnosti, čiji je rang jednak broju aktivnih ograničenja. Zbog ovakve nezgodne strukture Hesijana, primena metode najbržeg pada nije preporučljiva jer je jako spora. Međutim, ova specifična struktura Hesijana ne utiče na brzinu modifikovanog Njutnovog postupka, pa je on efikasan za rešavanje ovakvih problema.Metoda konjugovanih gradijenata sa restartom je takođe pogodna za rešavanje pomoćnog problema bez ograničenja, posebno u slučaju kad polazni problem ima mali broj aktivnih ograničenja.

Literatura

[1] Damir Vučina: Metode inženjerske numeričke optimizacije, FESB, Split, 2005.

[2] David G. Luenberger, Yinju Ye: Linear and Nonlinear Programming, Springer, third edition, 2008.

[3] Dr. Jovan J.Petrić: Operaciona istraživanja, knjiga 1. Savremena administracija, Beograd, 1979.

[4] Glevitzky Béla: Operációkutatás I, Debrecen, 2003.

[5] Gradimir V. Milovanović, Predrag S. Stanimirović, Simbolička implementacija nelinearne optimizacije, Elektronski fakultet u Nišu, Edicija monografije, Niš, 2002, X+236 (ISBN 86-80135-67-4).

[6] Katarina Surla, Zagorka Lozanov-Crvenković: Operaciona istraživanja, Prirodno-matematički fakultet, Novi Sad, 2002.

[7] S.Krčevinac, M.Čangalović, V.Kovačević-Vujčić, M.Martić, M.Vujošević: Operaciona istraživanja, Fakultet organizacionih nauka, Beograd, 2006.

[8] S.Zlobec, J.Petrić: Nelinearno programiranje, Naučna knjiga, Beograd, 1989.

[9] Sanjay Sharma: Applied nonlinear programming, New Age International (P) Ltd., 2006.

[10] V.Vujčić, M.Ašić, N.Miličić: Matematičko programiranje, Matematički institut, Beograd, 1980.

[11] Won Young Yang, Wenwu Cao, Tae-Sang Chung, John Morris, APPLIED NUMERICAL METHODS USING MATLAB, Published by John Wiley & Sons, Inc., Hoboken, New Jersey., 2005.

[12] Zhong Wan, Shao Jun Zhang, YaLin Wang: Penalty Algorithm Based on Conjugate Gradient Method for Solving Portfolio Management Problem, Hindawi Publishing Corporation, Journal of Inequalities and Applications, Volume 2009, Article ID 970723, 16 pages doi: 10.1155/2009/970723

[13] Ф.П. Васильев: Численные методы решения экстремальных задач, Москва << Наука>> Физико-математической литературы, 1980.

Biografija

Angela Štajer (rođ. Gal) je rođena 2. novembra 1987. godine u Novom Sadu. Osnovnu školu „Jožef Atila” i Elektrotehničku školu ”Mihajlo Pupin”, smer elektrotehničar telekomunikacija je završila u Novom Sadu. 2006. godine je upisala osnovne akademske studije na Prirodno-matematičkom fakultetu u Novom Sadu, smer matematika finansija. Studije je završila u avgustu 2010. godine. Iste godine je upisala master studije primenjene matematike, modul matematika finansija, na istom fakultetu. Zaključno sa junskim ispitnim rokom 2014. godine, položila je sve predviđene ispite, kao i grupu pedagoško-psihološko-metodičkih predmeta. Od oktobra 2010. godine radi kao profesor matematike u Srednjoj školi „Lukijan Mušicki“ u Temerinu.

Novi Sad, oktobar 2014. Angela Štajer

UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI FAKULTET
KLJUČNA DOKUMENTACIJSKA INFORMACIJA

Redni broj: RBR	
Identifikacioni broj: IBR	
Tip dokumentacije: TD	Monografska dokumentacija
Tip zapisa: TZ	Tekstualni štampani materijal
Vrsta rada: VR	Master rad
Autor: AU	Angela Štajer
Mentor: MN	Prof. dr Sanja Rapajić
Naslov rada: NR	Rešavanje problema minimizacije sa ograničenjima primenom kaznenih funkcija
Jezik publikacije: JP	Srpski (latinica)
Jezik izvoda: JI	s/e
Zemlja publikovanja: ZP	Republika Srbija
Uže geografsko područje: UGP	Vojvodina
Godina: GO	2014
Izdavač: IZ	Autorski reprint
Mesto i adresa: MA	Trg Dositeja Obradovića 4, 21000 Novi Sad

Fizički opis: FO	6 poglavlja/ 78 strana/ 4 slike/ 6 tabela
Naučna oblast: NO	Matematika
Naučna disciplina: ND	Operaciona istraživanja
Ključne reči: PO UDK	Nelinearno programiranje, spoljašnje kaznene funkcije, unutrašnje kaznene funkcije
Čuva se: ČU	U biblioteci Departmana za matematiku i informatiku Prirodno-matematičkog fakulteta u Novom Sadu
Izvod: IZ	U radu su predstavljene metode spoljašnjih i unutrašnjih (barijernih) kaznenih funkcija, koje se koriste za rešavanje problema minimizacije sa ograničenjima. Kako se problem sa ograničenjima pomoću ovih metoda svodi na problem bez ograničenja, predstavljene su i mnoge metode za rešavanje ovakvih problema.
Datum prihvatanja teme od strane NN veća: DP	10.05.2012.
Datum odbrane: DO	
Članovi komisije: KO	
Predsednik:	dr Nataša Krejić, redovni profesor Prirodno- matematičkog fakulteta u Novom Sadu
Član:	dr Zorana Lužanin, redovni profesor Prirodno-matematičkog fakulteta u Novom Sadu
Član:	dr Sanja Rapajić, vanredni profesor Prirodno-matematičkog fakulteta u Novom Sadu

UNIVERSITY OF NOVI SAD

FACULTY OF SCIENCE

KEY WORDS DOCUMENTATION

Accession number: ANO	
Identification number: INO	
Document type: DT	Monograph type
Type of record: TR	Printed text
Contents code: CC	master's thesis
Author: AU	Angela Štajer
Mentor: MN	Prof. Dr Sanja Rapajić
Title: XI	Solving the constrained minimization problems by penalty functions
Language of text: LT	Serbian (latin)
Language of abstract: LA	s/e
Country of publication: CP	Republic of Serbia
Locality publication: UGP	Vojvodina
Publication year: PU	2014
Publisher: PU	Author's reprint
Publ.place: PP	Trg Dositeja Obradovića 4, 21000 Novi Sad

Physical description: PD	6 chapters/ 78 pages/ 4 pictures/ 6 tables
Scientific field: SF	Mathematics
Scientific discipline: SD	Operations Research
Key words: UC	Nonlinear programming, penalty functions, barrier functions
Holding data: HD	Department of Mathematics and Informatics' Library, Faculty of Sciences, University of Novi Sad
Abstract: AB	Penalty and barrier methods are presented in this paper. These methods are used for constrained minimization problems. The main idea of these methods is to transform the constrained minimization problem to the problem of unconstrained minimization, so many methods for unconstrained minimization are also presented.
Accepted by the Scientific Board on: DP	10. 05. 2012.
Defended:	
Thesis defend board: President:	Dr Nataša Krejić, Full Professor, Faculty of Sciences, University of Novi Sad
Member:	Dr Zorana Lužanin, Full Professor, Faculty of Sciences, University of Novi Sad
Member:	Dr Sanja Rapajić, Associate Professor, Faculty of Sciences, University of Novi Sad