UNIVERZITET U BEOGRADU

PRIRODNO-MATEMATIČKI FAKULTET

DOKTORSKA DISERTACIJA

- PRILOG IZUČAVANJU ALGORITAMA OPTIMIZACIJE -

Beograd, juna 1980. NADA DJURANOVIĆ-MILIČIĆ

SADRŽAJ

0. UVOD

Optimizacioni problemi odavno interesuju matematičare, fizičare, inženjere i druge. Mogućnost korišćenja metoda diferencijalnog i varijacionog računara za rešavanje odredjenih vrsta optimizacionih problema u geometriji, mehanici i fizici bila je poznata i primenjivana još od sredine osamnaestog veka.

U poslednjih tridesetak godina došlo je do velikog interesovanja za novu klasu optimizacionih problema, koji se najčešće zovu problemi matematičkog programiranja, a koji se ne mogu rešavati klasičnim metodama. Problem matematičkog programiranja, koji možemo ovako formulisati: naći

$$(0.1) \qquad \min \{ \phi(x) \mid x \in X \},$$

gde je X - skup definisan odredjenim ograničenjima na promenljivu $x \in R^n$, je ponikao u teorijskoj ekonomiji, ali se takodje javlja u obliku važnih praktičnih problema u industriji, vojnim disciplinama, inženjerskim zadacima.

Specijalno, interesovanje za probleme nelinearnog programiranja je posebno poraslo posle pojave rada H.W.Kuhn-a i A.W.Tucker-a /32/ 1951. god. u kojem su izloženi potrebni i dovoljni uslovi za optimalna rešenja problema nelinearnog programiranja.

Danas se u svetu ovom problematikom bavi veliki broj matematičara. Postoje brojni specijalizovani časopisi koji objavljuju najnovije rezultate iz ove oblasti. Pre svega pažnja je posvećena nalaženju što efikasnijih metoda za rešavanje problema (0.1). Pri tom se pod efikasnošću podrazumeva brza (teorijska) konvergencija, ekonomičnost za rad računara i numerička stabilnost metode.

Takozvane gradijentne metode, ili metode prvoga reda su dosta jednostavne, zahtevaju samo nalaženje prvih izvoda, ali imaju sporu - ne veću od linearne, brzinu konvergencije. Metode koje koriste druge izvode su mahom razne modifikacije poznate Newton-ove metode - brzo konvergiraju, ali su zbog potrebe da se nadje hesijan i inverzni hesijan veoma skupe i znatno opterećuju memoriju računara. Zbog toga su napori mnogih matematičara bili usmereni ka nalaženju metoda koje koriste samo prve izvode, a imaju ponašanje metoda (brzinu konvergencije) koje koriste druge izvode (metodâ drugog reda). Takve su metode promenljive metrike, odnosno, metode konjugovanih pravaca. Medjutim, kod metoda promenljive metrike i konjugovanih pravaca, kao uostalom, i kod većine ostalih optimizacionih metoda korak duž datog pravca se odredjuje iz uslova minimizacije funkcije (duž pravca), a to je vrlo skup proces. Zato se počela da posvećuje velika pažnja definisanju novih algoritama za nalaženje koraka koji bi bili pre svega ekonomičniji, a koji istovremeno obezbedjuju dobru brzinu konvergencije (videti na primer, Nazareth /40/, Danilin /11/, /12/).

Metode koje ne zahtevaju nalaženje izvoda su najjednostavnije i za praksu veoma važne jer se mogu primeniti i onda kada otkažu gradijentne metode. U najnovijim radovima Toint-a i Callier-a /59/, /60/ objedinjuju se dosadašnji rezultati važnijih metoda koje ne koriste izvode i istovremeno korak ne mora da bude dobijen iz uslova minimizacije duž pravca.

Ova teza se bavi upravo pomenutim problemima: modifikacijama algoritama za nalaženje koraka, kao i metodama za uslovnu optimizaciju koje ne koriste izvode. Raspored izlaganja je sledeći. U prvom poglavlju se daju najvažnije definicije, stavovi, kao i algoritmi čije se modifikacije izlažu u trećem i četvrtom poglavlju. U drugom poglavlju se izlažu poznate projektivne metode i one optimizacione metode koje ne koriste izvode čije se projektivne varijante izlažu u petom poglavlju.

Pri tom se za dokaze stavova upućuje na odgovarajuće reference. Dokaz teoreme 2.2 je originalan.

U trećem poglavlju se izlažu modifikacije i poopštenja metoda za nalaženje koraka duž datog pravca: Curry-Altman-ove metode /9/, /1/, me-

tode Ostrowskog /42/, Céa-Goldstein-ove metode /27/, /54/. Ove su modifikacije zasnovane na korišćenju svojstava takozvanih "prinudnih funkcija" (definicija 1.9). Kod svake modifikacije definiše se algoritam, dokazuje njegova dobra definisanost i, uz odgovarajuće pretpostavke o nizu vektora pravaca dokazuje se da generisani niz tačaka konvergira ka optimalnom rešenju (definicija 1.1) datog optimizacionog problema, a za modifikovanu metodu Ostrowskog i poopštenu Céa-Goldstein-ovu metodu daje se i ocena brzine konvergencije.

Pri tom se prvo izlažu algoritmi za bezuslovnu optimizaciju, a zatim odgovarajuće uslovne varijante dobijenih modifikacija.

U četvrtom poglavlju izložene su četiri modifikacije metode Danilin-Pšeničnog /13/, koje se sastoje u tome da za dati niz vektora pravaca definisan kao u metodi Danilin-Pšeničnog, definišemo različite korake. Dokazana je konvergencija ka optimalnom rešenju kod svake od modifikacija, i data je ocena brzine konvergencije.

U petom poglavlju izlažu se u 5.1 projektivne varijante bezuslovnih optimizacionih metoda koje ne koriste izvode: metode lokalnih varijacija /4/, Powell-Zangwill-ove metode /46/, /62/ i ciklične koordinatne metode. Ovde se na prirodan način, kako to do sada nije radjeno, vrši adaptacija bezuslovnih optimizacionih metoda na slučaj kada su ograničenja data u obliku linearnih jednakosti, odnosno nejednakosti. U 5.2 je dato poopštenje rezultata Danilin-a /11/, koji se odnose na optimizaciju uz ograničenja u obliku nelinearnih jednakosti.

1. POGLAVLJE

1.1 Potrebni i dovoljni uslovi optimalnosti

Opšti problem nelinearnog programiranja možemo formulisati na sledeći način: naći

$$(1.1) \qquad \min \{ \phi(x) \mid x \in X \},$$

gde je $\quad X = \{ x \in D \mid \; f_i(x) \leq 0, \quad i = 1, \dots, m \},$

pri čemu su funkcije $\phi : D \subset R^n \to R$ i $f_i : D \subset R^n \to R$, $i \in I_o$, $I_o = \{1, \dots, m\}$ diferencijabilne na otvorenom skupu D.

Funkcija ϕ se zove funkcija cilja, funkcije f_i, $i \in I_o$ se zovu funkcije ograničenja ili samo - ograničenja. Skup X je dopustivi skup, a svako $x \in X$ nazivamo dopustivim rešenjem. Za ograničenje f_j, $j \in I_o$, ćemo reći da je aktivno u datoj tački $x \in X$ ako je $f_j(x) = 0$.

Za bilo koje $x \in X$ i $\varepsilon \geq 0$ definišimo skup indeksa $I(x,\varepsilon) \subset I_o$ na sledeći način:

$$I(x,\varepsilon) = \{ j \in I_o \mid -\varepsilon \leq f_j(x) \leq 0 \}.$$

Specijalno, $I(x,0)$ će označavati skup indeksa koji odgovaraju onim ograničenjima koja su aktivna u x.

Definicija 1.1. Optimalnim rešenjem ili globalnim minimumom problema (1.1) nazivamo onu tačku $x^* \in X$ za koju važi

$$(1.2) \qquad \phi(x) \geq \phi(x^*) \qquad \text{za } \forall x \in X .$$

Definicija 1.2. *Lokalnim minimumom problema (1.1) nazivamo tačku $x^* \in X$ ako postoji $\varepsilon > 0$ tako da važi*

$$(1.3) \qquad \phi(x) \geq \phi(x^*) \qquad \text{za } \forall x \in \{x \in X \mid \|x - x^*\| \leq \varepsilon\}.$$

Ako su nejednakosti (1.2), odnosno (1.3) stroge, tačka $x^* \in X$ je strogi globalni minimum, odnosno strogi lokalni minimum.

Sada ćemo izložiti dve leme koje ćemo koristiti u dokazu poznate Kuhn-Tucker-ove teoreme. Prethodno definišemo dopustivi vektor pravca.

Definicija 1.3. *Neka $x \in X$. Za vektor pravca $p \in R^n$ ćemo reći da je dopustiv u tački x ako postoji $\alpha_d > 0$ takvo da $x - \alpha p \in X$ za $\forall \alpha \in (0, \alpha_d]$.*

Lema 1.1. *Neka je x dopustivo rešenje problema (1.1). Neka je $\phi : D \subset R^n \to R$ diferencijabilna u tački x. Potreban uslov da je x optimalno rešenje problema (1.1) jeste da ne postoji dopustivi vektor pravca $p \in R^n$ koji zadovoljava nejednakost*

$$(1.4) \qquad \langle \nabla\phi(x), p \rangle > 0.$$

Dokaz. *Pretpostavimo suprotno, tj. pretpostavimo da postoji $\alpha_d > 0$ takvo da $x - \alpha p \in X$ za $\forall \alpha \in [0, \alpha_d]$ i da vektor p zadovoljava uslov (1.4). Tada, pošto je, na osnovu diferencijabilnosti ϕ*

$$\lim_{\alpha \to 0} \frac{\phi(x) - \phi(x - \alpha p)}{\alpha} = \langle \nabla\phi(x), p \rangle,$$

iz (1.4) sledi da je

$$\lim_{\alpha \to 0} \frac{\phi(x) - \phi(x - \alpha p)}{\alpha} > 0\,.$$

Iz definicije granične vrednosti sledi da mora postojati $\bar{\alpha} > 0$ *takvo da je za svako* $\alpha \neq 0$, $-\bar{\alpha}<\alpha<\bar{\alpha}$, $\dfrac{\phi(x) - \phi(x - \alpha p)}{\alpha} > 0$, *odnosno da je* $\phi(x - \alpha p) < \phi(x)$ *za* $\forall \alpha \in (0,\bar{\alpha})$. *Sledi da ako je* $0 < \alpha < \min\{\alpha_d, \bar{\alpha}\}$ *da je* $\phi(x - \alpha p) < \phi(x)$ *i pri tom je* $x - \alpha p$ *dopustivo rešenje. Prema tome,* x *nije optimalno rešenje problema (1.1), što je suprotno sa pretpostavkom teoreme.*

Lema 1.2. *(Farkas /18/)* *Za date vektore* $a_i \in R^n$, $i = 1,\dots,t$ *i* $g \in R^n$ *ne postoji vektor* $p \in R^n$ *koji zadovoljava relacije*

$$< g,p> > 0 \quad i \quad <a_i,p> \geq 0, \qquad i = 1,\dots,t$$

ako i samo ako se g *može izraziti kao linearna kombinacija oblika*

$$g = \sum_{i=1}^{t} \lambda_i a_i, \quad \lambda_i \leq 0, \quad i = 1,\dots,t.$$

Dokaz. *Pogledati na primer, u Zangwill /63/, str. 290.*

Pre nego što formulišemo teoremu Kuhn-Tucker-a, definišimo pojam regularnosti.

Definicija 1.4. *Kažemo da je u tački* $x \in X$ *zadovoljen uslov regularnosti ako su gradijent vektori aktivnih ograničenja u* x, $\nabla f_i(x)$, $i \in I(x,o)$, *linearno nezavisni, tj. ako važi*

$$\text{rang}\ (\nabla f_i(x),\ \ i \in I(x,o)) = q < n \ \text{kadgod je kard}\ I(x,o) = q > 0. \tag{1.5}$$

Za dopustivi skup X *kažemo da je regularan ako je u svakoj tački* $x \in X$ *zadovoljen uslov (1.5).*

Teorema 1.3. *(Kuhn-Tucker /32/)* *Neka su funkcije* $\phi : D \subset R^n \to R$ *i* $f_i : D \subset R^n \to R$, $i \in I_o$, *diferencijabilne na otvorenom skupu D. Neka je* x^* *- dopustivo rešenje, pri čemu je* $f_i(x^*) = 0$ *za* $i \in I(x^*,o)$. *Tada, ako je u tački* x^* *zadovoljen uslov regularnosti (1.5), potreban uslov da je* x^* *- optimalno rešenje problema (1.1) jeste da važi sledeće*

$$(1.6) \qquad \nabla\phi(x^*) = \sum_{i \in I(x^*,o)} \lambda_i \nabla f_i(x^*), \quad \lambda_i \leq 0, \quad i \in I(x^*,o).$$

Dokaz. *Pretpostavimo da ne važi relacija (1.6). Tada, na osnovu leme 1.2 sledi da postoji vektor* $p \in R^n$ *takav da za* $x = x^*$ *zadovoljava nejednakost (1.4) i uslov*

$$(1.7) \qquad \langle \nabla f_i(x^*), p \rangle \geq 0, \qquad i \in I(x^*,o).$$

Pošto u tački x^* *važi uslov regularnosti, postoji vektor* $q \in R^n$ *takav da važe stroge nejednakosti*

$$\langle \nabla f_i(x^*), q \rangle > 0, \qquad i \in I(x^*,o).$$

Zamenimo p sa $p + \varepsilon q$, *gde je* ε *ma koji pozitivan broj takav da važi* $\langle \nabla\phi(x^*), p + \varepsilon q \rangle > 0$. *Dakle, sve nejednakosti iz (1.7) postaju stroge. Sledi da funkcije ograničenja zadovoljavaju relaciju* $f_i(x^* - \alpha p) < 0$, $i \in I(x^*,o)$ *za sve vrednosti* α *iz nekog odredjenog intervala, tj. za* $\alpha \in (o, \alpha^*)$. *Neka je* $\bar{\alpha} > 0$ *takvo da* $x - \alpha p \in X$ *za* $\forall \alpha \in [o, \bar{\alpha}]$. *Tada je za svako* α, $0 < \alpha < \min\{\alpha^*, \bar{\alpha}\}$ *vektor p dopustiv u tački* x^*. *Sledi da p zadovoljava pretpostavke leme 1.1, a odatle sledi da* x^* *nije optimalno rešenje problema (1.1), što je kontradiktorno sa pretpostavkom teoreme.*

Uobičajeno je da se Kuhn-Tucker-ovi uslovi (1.6) pišu u sledećem obliku

$$(1.8) \qquad \nabla\phi(x^*) = \sum_{i=1}^{m} \lambda_i \nabla f_i(x^*), \quad \lambda_i \leq 0, \quad i = 1,\dots,m$$

$$(1.9) \qquad \lambda_i \nabla f_i(x^*) = 0, \quad i = 1,\ldots,m.$$

Definicija 1.5. *Tačka $x^* \in X$ koja zadovoljava uslove (1.8) i (1.9) naziva se Kuhn-Tucker-ovom tačkom ili stacionarnom tačkom problema (1.1).*

Teorema 1.3 je mogla da se dokaže i pod pretpostavkama slabijim od uslova (1.5). Naime, postoje i drugačije formulisani uslovi regularnosti (constraint qualifications) i manje strogi od uslova (1.5). Pomenimo: Kuhn-Tucker-ov uslov regularnosti /32/, Slater-ov uslov regularnosti /57/, Arrow-Hurwicz-Uzawa uslov regularnosti /2/, Karlin-ov uslov regularnosti /29/ i druge.

Mangasarian O.L. detaljno razmatra razne uslove regularnosti, kao i veze izmedju njih u /36/.

Posledica 1.3 se odnosi na specijalan slučaj problema (1.1) kada su sva ograničenja zadata u obliku jednakosti.

Posledica 1.3. *Razmatramo optimizacioni problem:*

$$(1.10) \qquad \min\{\phi(x) \mid x \in X\}, \quad X = \{x \in D \mid f_i(x) = 0, \quad i = 1,\ldots,m\}.$$

Neka su funkcije $\phi: D \subset R^n \to R$ i $f_i: D \subset R^n \to R$, $i \in I_o = \{1,\ldots,m\}$ diferencijabilne na otvorenom skupu D. Neka je x^* - optimalno rešenje problema (1.10) i neka je u tački x^* zadovoljen uslov regularnosti (1.5). Tada postoje neograničeni po znaku množioci λ_i, $i \in I_o$, takvi da važi

$$\nabla\phi(x^*) = \sum_{i=1}^{m} \lambda_i \nabla f_i(x^*).$$

Dokaz. *Pogledati na pr. u Hadley /28/, str.68.*

Teorema 1.3 daje samo potrebne uslove optimalnosti u x^*, (1.8) i (1.9). Da bi oni bili i dovoljni, potrebno je da učinimo još neke dodatne pretpostavke o funkciji cilja ϕ i funkcijama ograničenjâ f_i, $i \in I_o$.

Zato ćemo prvo dati neke definicije i stavove.

Definicija 1.6. *Za funkciju* $\phi : D \subset R^n \to R$ *kažemo da je pseudokonveksna na otvorenom i konveksnom skupu* D *ako je* ϕ *diferencijabilna i za* $\forall x_1, x_2 \in D$ *važi*

$$\langle \nabla\phi(x_1), (x_2 - x_1) \rangle \geqslant 0 \Rightarrow \phi(x_2) \geqslant \phi(x_1).$$

Ako je $x_1 \neq x_2$, *a nejednakost stroga, kažemo da je* ϕ *strogo pseudokonveksna funkcija.*

Lema 1.4. *Neka je* $\phi : D \subset R^n \to R$ *pseudokonveksna funkcija na otvorenom i konveksnom skupu* D. *Pretpostavimo da je* $\nabla\phi(x^*) = 0$ *za neko* $x^* \in D$. *Tada je* $\nabla\phi(x^*) = 0$ *potreban i dovoljan uslov da je tačka* x^* *globalni minimum funkcije* ϕ. *Ako je* ϕ *strogo pseudokonveksna,* x^* *je jedinstveni globalni minimum.*

Dokaz. *Dokaz sledi neposredno iz definicije 1.6.*

Definicija 1.7. *Za funkciju* $\phi : D \subset R^n \to R$ *kažemo da je kvazikonveksna na konveksnom skupu* D *ako za* $\forall x_1, x_2 \in D$ *i* $\forall \lambda \in [0,1]$ *važi*

$$\phi\left[\lambda x_1 + (1-\lambda)x_2\right] \leq \max\{\phi(x_1),\ \phi(x_2)\}.$$

Ako je $x_1 \neq x_2$, $0 < \lambda < 1$, *a nejednakost stroga, kažemo da je* ϕ *strogo kvazikonveksna funkcija.*

Lema 1.5. *Nivoski skup* $L = \{x \in D \mid \phi(x) \leq \beta\}$ *je konveksan za* $\forall \beta \in R$ *ako i samo ako je* $\phi : D : R^n \to R$ *kvazikonveksna funkcija na konveksnom skupu* D.

Dokaz. *Videti na pr. u Avriel /3/, str. 145.*

Definicija 1.8. *Za funkciju $\phi : D \subset R^n \to R$ kažemo da je konveksna na konveksnom skupu D ako za $\forall x_1, x_2 \in D$ i $\forall \lambda \in [o,1]$ važi*

$$\phi\left[\lambda x_1 + (1-\lambda)x_2\right] \leq \lambda\phi(x_1) + (1-\lambda)\phi(x_2).$$

Ako je $x_1 \neq x_2, o<\lambda<1$, a nejednakost stroga, kažemo da je ϕ strogo konveksna.

Lema 1.6. *Ako je diferencijabilna funkcija $\phi : D \subset R^n \to R$ konveksna na otvorenom i konveksnom skupu D, ona je i pseudokonveksna i strogo kvazikonveksna.*

Dokaz. *Dokaz neposredno sledi iz definicijà 1.6, 1.7 i 1.8.*

Teorema 1.7. *Neka je funkcija $\phi : D \subset R^n \to R$ pseudokonveksna, a funkcije $f_i : D \subset R^n \to R$, $i \in I_o$ diferencijabilne i kvazikonveksne na otvorenom konveksnom skupu D. Neka $x^* \in X$ i neka je u x^* zadovoljen uslov regularnosti (1.5). Tada je potreban i dovoljan uslov da je x^* optimalno rešenje problema (1.1) da je x^* - Kuhn-Tucker-ova tačka problema (1.1).*

Dokaz. *Videti na primer u Zangwill /63/, str. 47.*

1.2. O najvažnijim pristupima rešavanju problema (1.1)

Postoje dva osnovna pristupa rešavanju problema (1.1). Kod prvoga pristupa se o ograničenjima vodi računa na indirektan način i problem uslovne optimizacije (1.1) se svodi na problem (odnosno niz problema) bezuslovne optimizacije. Najvažnije metode sa ovakvim pristupom predstavlju metode kaznenih funkcija koje su najviše izučavali: Fiacco i Mc Cormick /19/, Poljak /45/, Polak /44/, Luenberger /34/, Conn i Pietrzykowski /7/ i drugi.

Drugi pristup karakteriše direktno razmatranje ograničenja. Metode sa ovakvim pristupom možemo nazvati metodama dopustivih pravaca u širem smislu. Ove metode su iteracione i generišu niz tačaka $\{x_k\}$ sledećeg oblika:

$$(1.11) \qquad x_{k+1} = x_k - \alpha_k p_k, \quad x_{k+1} \in X, \quad \phi(x_{k+1}) \leq \phi(x_k), \quad k = 0,1,2,\ldots,$$

gde je p_k - dopustivi pravac u tački x_k (definicija 1.3), a α_k - veličina koraka duž pravca p_k. Pomenimo Zountendijk-ove metode dopustivih pravaca /64/, zatim razne projektivne metode, kao: Rosen-ovu /52/, /53/, Goldfarb-ovu /25/, Ritter-ovu /50/, Best i Ritter-ovu /5/ i druge.

O projektivnim metodama biće posebno reči u drugom poglavlju, a mi ćemo ovde izložiti osnovna stanovišta od kojih se polazilo u definisanju iteracionih metoda oblika (1.11) i analizi njihovih konvergencijskih svojstava.

Monotoni iteracioni algoritmi su vrlo često definisani pomoću skupovnog preslikavanja $A: V \to V$ u kojem je slika $A(x)$ ma koje tačke $x \in V$, podskup skupa V. Korišćenjem svojstava skupovnog preslikavanja došlo se do konvergencijskih rezultata za optimizacione algoritme najopštijeg oblika, kao na primer u Zangwill /63/, Polak /44/, Luenberger /34/ itd.

Na drugoj strani, razvio se pristup monotonim optimizacionim algoritmima koji koristi takozvane "prinudne" ("forcing") funkcije: Ortega i Rheinboldt /41/, Daniel /10/, Rauch /49/, V.Kovačević /30/ i drugi.

Kakve su to "prinudne" funkcije? Razni autori ih različito definišu, ali je u svim tim raznim definicijama zajedničko sledeće: nenegativna "prinudna" funkcija μ *igra ulogu "indikatora optimalnosti" i iteracija* x_k *se može ponoviti samo ako je* $\mu(x_k) = 0$. *Ako je* $\mu(x_k)>0$, *zahtevaćemo da sledeća iteracija* x_{k+1} *daje umanjenje funkcije cilja* ϕ *koje iznosi bar* $\mu(x_k)$, *tj. da važi*

$$\phi(x_k) - \phi(x_{k+1}) \geqslant \mu(x_k).$$

Termin "pristup pomoću prinudnih funkcija" se koristi, pošto s jedne strane μ *"prinudjava" ("forces") funkciju cilja* ϕ *da opada ako je* $\mu(x_k) > 0$, *i, na drugoj strani, konvergencija niza vrednosti funkcije* $\{\phi(x_k)\}$ *će imati za posledicu konvergenciju niza* $\{\mu(x_k)\}$ *ka nuli (niz* $\{\mu(x_k)\}$ *će konvergencijom niza* $\{\phi(x_k)\}$ *biti "prinudjen" da teži ka nuli).*

Medjutim, ovakvim pristupom nisu doskora bili obuhvaćeni optimizacioni algoritmi onako opšteg karaktera kao kod skupovnog pristupa. 1977. R.R.Meyer objavljuje rad /37/ sa pristupom konvergencijskoj analizi optimizacionih algoritama, zasnovanom na prinudnim funkcijama, u kojem izlaže rezultate koji ne samo da su opštiji od rezultata dobijenih skupovnim pristupom, već su, štaviše, često lakši za primenu. Zatim, Meyer-ovi rezultati mogu se jednostavno proširiti i na klasu nekih algoritama koji se ne mogu direktno razmatrati pomoću skupovnog pristupa.

Meyer takodje pokazuje kako se zadatom skupovnom preslikavanju (kojim je definisan dati optimizacioni algoritam) mogu pridružiti prinudne funkcije sa gore navedenim svojstvima, i obrnuto, kako se, kod algoritama definisanim pomoću prinudnih funkcija može definisati odgovarajuće skupovno preslikavanje i na taj način pokazuje ekvivalenciju ovih dvaju pristupa.

Meyer u /37/ na kraju pokazuje da su algoritmi Ortega i Rheinboldta /41/ samo specijalan slučaj pristupa izloženog u ovom radu.

Pošto se u ovoj tezi u poglavlju 3. i 4. razmatraju monotoni optimizacioni algoritmi koji se definišu pomoću prinudnih funkcija, Meyer-ovi rezultati bacaju svakako novo svetlo na značaj ovakvih pristupa.

1.3. Algoritmi za nalaženje koraka

Razmatramo iteracione optimizacione algoritme koji generišu nizove tačaka $\{x_k\}$ *oblika (1.11). Niz* $\{x_k\}$ *je definisan ako je definisan niz vektora pravaca* $\{p_k\}$ *i niz koraka* $\{\alpha_k\}$*, odnosno, ako su nam poznati algoritmi za generisanje niza* $\{p_k\}$*, odnosno* $\{\alpha_k\}$.

Elkin /17/, Topkis i Vienott /61/, Ortega i Rheinboldt /41/, Daniel /10/, Rauch /49/ izučavaju algoritme koji generišu vektore pravaca i algoritme za izračunavanje koraka nezavisno jedne od drugih, što im omogućuje da raznim kombinacijama ovih algoritama dodju do različitih optimizacionih metoda. Ovde ćemo posebno izložiti one algoritme za nalaženje koraka, čijim modifikacijama se bavimo u trećem poglavlju.

Prvo ćemo izložiti algoritme koji se odnose na bezuslovnu optimizaciju

$$(1.12) \qquad \min\{\phi(x)\,|\,x \in D\},$$

gde je $\phi: D \subset R^n \to R$ *neprekidno diferencijabilna funkcija na otvorenom skupu* D.

Pre nego što predjemo na definisanje samih algoritama, daćemo neke definicije i leme koje se koriste u dokazima u poglavlju 3. (i 4).

Definicija 1.9. *Preslikavanje* $\sigma: [0,\infty) \to [0,\infty)$ *jeste prinudna funkcija ako za ma koji niz* $\{t_k\} \subset [0,\infty)$

$$\lim_{k\to\infty} \sigma(t_k) = 0 \quad \text{povlači} \quad \lim_{k\to\infty} t_k = 0$$

i $\sigma(t) > 0$ *za* $t > 0$.

Definicija 1.10. *Pretpostavimo da je funkcija* $\phi: D \subset R^n \to R$ *ravnomerno neprekidna na skupu* $D_o \subset D$. *Tada se funkcija* w *definisana relacijom*

$$(1.13) \qquad w(t) = \sup\{ \|\phi(x) - \phi(y)\| \mid x,y \in D_o, \ \|x-y\| \leq t\}$$

za svako $t \geq 0$ *za koje je* $w(t) < +\infty$ *zove moduo neprekidnosti* ϕ *na* D_o.

Lema 1.8. *Pretpostavimo da je funkcija* $\phi: D \subset R^n \to R$ *ravnomerno neprekidna na konveksnom skupu* $D_o \subset D$. *Tada je moduo neprekidnosti* ϕ *na* D_o *definisan i monotono rastuća ravnomerno neprekidna funkcija na* $[o,+\infty)$ *i* $w(o) = 0$.

Dokaz. *Videti u Ortega i Rheinboldt /41/, str. 64.*

Definicija 1.11. *Pretpostavimo da je funkcija* $\phi: D \subset R^n \to R$ *neprekidno diferencijabilna i pretpostavimo da na nekom skupu* $D_o \subset D$ *važi*

$$\gamma = \sup\{ \|\nabla\phi(x) - \nabla\phi(y)\| \mid x,y \in D_o\} > 0 .$$

Tada preslikavanje $s: [o,\infty) \to [o,\infty)$ *definisano sa*

$$(1.14) \qquad s(t) = \begin{cases} \inf\{ \|x-y\| \mid x,y \in D_o, \ \|\nabla\phi(x) - \nabla\phi(y)\| \geq t\}, & t \in [o,\gamma), \\ \lim\limits_{s \to \gamma_-} s(t), & t \in [\gamma,+\infty), \end{cases}$$

je takozvani obratni moduo neprekidnosti od $\nabla\phi$ *na* D_o *(videti /41/, str.482).*

Vidimo da je s *uvek definisana i monotono rastuća funkcija i da je* $s(o) = 0$.

Lema 1.9. *Pretpostavimo da* $\phi: D \subset R^n \to R$ *ima ravnomerno neprekidne izvode na* $D_o \subset D$ *i da je* γ *iz definicije 1.11 pozitivno. Tada je* $s(t) > 0$ *za svako* $t > 0$. *Dakle,* s *je prinudna funkcija.*

Dokaz. *Dokaz sledi neposredno iz definicije 1.11.*

Lema 1.10. *Neka je funkcija* $\phi: D \subset R^n \to R$ *dva puta neprekidno diferencijabilna na otvorenom skupu* D *i neka postoji takvo* m, $o < m < \infty$ *da za* $\forall x \in D$ *i* $\forall y \in R^n$ *važi*

$$\langle y, H(x)y \rangle \geqslant m \| y \|^2 ,$$

gde je $H(x)$ *– hesijan funkcije* ϕ *u tački* x. *Tada je funkcija* ϕ *strogo konveksna i za ma koju tačku* $x_o \in D$ *skup*

$$L = \{ x \in D \mid \phi(x) \leq \phi(x_o) \}$$

je ograničen (i konveksan).

Dokaz. *Videti u Polak /44/, str. 340.*

Definicija 1.12. *Kažemo da je skup* $S \subset R^n$ *povezan ako se ne može predstaviti kao unija dvaju disjunktnih nepraznih otvorenih skupova. Maksimalni povezani podskup, tj. povezani podskup koji nije sadržan ni u kojem većem povezanom podskupu, zove se komponenta skupa.*

Definicija 1.13. *Kažemo da je skup* $S \subset R^n$ *lučno povezan ako za* $\forall x, y \in S$ *postoji neprekidno preslikavanje* $p: [o,1] \to S$ *tako da je* $p(o) = x$ *i* $p(1) = y$.

Očigledno je svaki lučno povezan skup i povezan i svaki otvoren povezan skup je lučno povezan.

Lema 1.11. *Neka je $\phi: D \subset R^n \to R$ neprekidna funkcija na otvorenom skupu D i neka je diferencijabilna na kompaktnom skupu L^o, gde smo sa L^o označili lučno-povezanu komponentu nivoskog skupa $L = \{ x \in D \mid \phi(x) \leq \phi(x_o) \}$ za neko $x_o \in D$. Tada, za svako $x \in L^o$ i $p \in R^n$ za koje važi $\langle \nabla\phi(x), p \rangle > 0$, postoji $\alpha^* > 0$ takvo da je $\phi(x) = \phi(x - \alpha^* p)$ i $[x, x - \alpha^* p] \subset L^o$. Specijalno, ako je $\eta > 0$ ma koji broj sa svojstvom da je*

$$(1.15) \qquad \phi(x - \alpha p) < \phi(x) \text{ za } \forall\, x - \alpha p \in (x, x - \eta p] \cap L^o,$$

tada $[x, x - \eta p] \subset L^o$.

Dokaz. *Videti u /41/, str. 480.*

Lema 1.12. *Neka je $\phi: D \subset R^n \to R$ neprekidno diferencijabilna funcija na kompaktnom skupu $D_o \subset D$ i pretpostavimo da je skup stacionarnih tačaka od ϕ u D_o konačan. Neka je $\{x_k\} \subset D_o$ ma koji niz takav da je $\lim\limits_{k \to \infty} (x_k - x_{k+1}) = 0$ i $\lim\limits_{k \to \infty} \nabla\phi(x_k) = 0$. Tada je $\lim\limits_{k \to \infty} x_k = x^*$ i $\nabla\phi(x^*) = 0$.*

Dokaz. *Videti u /41/, str. 476.*

Lema 1.13. *Pretpostavimo da je funkcija $\phi: D \subset R^n \to R$ neprekidno diferencijabilna na $D_o \subset D$ i da je moduo neprekidnosti*

$$w(t) \equiv \sup \{ \| \nabla\phi(x) - \nabla\phi(y) \| \mid \| x - y \| \leq t,\ x, y \in D_o \}$$

od $\nabla\phi$ na D_o dobro definisan i neprekidan na $[0, \infty)$. Tada za $\forall\, x \in D_o$ i $p \neq 0$ važi

$$(1.16)\qquad \phi(x-\alpha p) \leq \phi(x) - \alpha\langle\nabla\phi(x),p\rangle + \alpha\|p\|\int_0^1 \omega(t\alpha\|p\|)dt,$$

za $\forall\alpha$ takvo da $[x,x-\alpha p]\subset D_0$. Specijalno, ako $\nabla\phi$ zadovoljava Lipschitz-ov uslov

$$\|\nabla\phi(x) - \nabla\phi(y)\| \leq \gamma\|x-y\| \quad \text{za } \forall x,y\subset D_0,\quad \gamma > 0.$$

važi

$$(1.17)\qquad \phi(x-\alpha p)\leq\phi(x) - \alpha\langle\nabla\phi(x),p\rangle + \frac{\gamma}{2}\alpha^2\|p\|^2.$$

Dokaz. Videti u Ortega, Rheinboldt /41/, str. 254.

Sada možemo izložiti i sáme algoritme za nalaženje koraka koji se odnose na problem (1.12).

Razmatramo iterativne algoritme za nalaženje optimalnog rešenja problema (1.12) koji generišu nizove tačaka $\{x_k\}$ oblika

$$x_{k+1} = x_k - \alpha_k p_k,\qquad k = 0,1,2\ldots,$$

pri čemu pretpostavimo da vektor pravca p_k zadovoljava sledeće:

$$p_k \neq 0 \quad i \quad \langle\nabla\phi(x_k),\ p_k\rangle \geqslant 0.$$

Korak $\alpha_k \geqslant 0$ izračunavamo pomoću jednog od sledećih algoritamâ:

Curry-Atlman-ov algoritam /41/, str. 251.

$$\alpha_k = 0 \quad \text{ako je } \langle\nabla\phi(x_k),p_k\rangle = 0;$$

inace

$$(1.18)\qquad \alpha_k = \min\{\alpha\geqslant 0 \mid \langle\nabla\phi(x_k-\alpha p_k),p_k\rangle = \mu\langle\nabla\phi(x_k),p_k\rangle\},$$

gde je $\mu\in[0,1)$ fiksiran broj.

Céa-Goldstein-ov algoritam /54/, str. 93.

$\alpha_k = 0$ *ako je* $<\nabla\phi(x_k),p_k> = 0$; *inače*

$\alpha_k > 0$ *je takav broj da važe sledeće relacije*

$$(1.19) \qquad \phi(x_k) - \phi(x_k-\alpha_k p_k) \geq \mu\alpha_k <\nabla\phi(x_k),p_k>$$

$$(1.20) \qquad \phi(x_k) - \phi(x_k-2\alpha_k p_k) < 2\mu\alpha_k <\nabla\phi(x_k),p_k>,$$

gde je $0<\mu<1$ *fiksiran broj.*

Algoritam Ostrowskog /41/, str. 481.

Pretpostavljamo da funkcija $\phi: D\subset R^n \to R$ *zadovoljava sledeći uslov na otvorenom skupu D:*

$$\|\nabla\phi(x) - \nabla\phi(y)\| \leq \gamma \|x-y\| \quad za \ \forall x,y\in D, \quad \gamma>0$$

Korak α_k *se definiše na sledeći način:*

$$(1.21) \qquad \alpha_k = \begin{cases} 0 & , \text{ako je } <\nabla\phi(x_k),p_k> = 0; \\ \dfrac{1}{\gamma\|p_k\|} \dfrac{<\nabla\phi(x_k),p_k>}{\|p_k\|} & , \text{inače.} \end{cases}$$

Za slučaj uslovne optimizacije vršene su razne adaptacije postojećih algoritama za nalaženje koraka u problemima bezuslovne optimizacije, kao npr. u: Daniel /10/, Ritter /51/.

Ovde ćemo podrobnije izložiti pristup S.W. Rauch-a na osnovu njegovog rada /49/, jer ćemo ga koristiti u poglavlju 3. u definisanju uslovne varijante dobijenih modifikacija bezuslovnih algoritama za generisanje koraka α_k.

S.W.Rauch razmatra sledeći problem nelinearnog programiranja

(1.22) $\min\{\phi(x) \mid x \in X\}, \quad X = \{x \in D \mid f_j(x) \leq 0, \quad j \in I_o = \{1,\dots,m\}\},$

a funkcije $\phi, f_j : D \subset R^n \to R$, $j \in I_o$, su neprekidno diferencijabilne na otvorenom konveksnom skupu D.

Rauch u /49/ proširuje Elkin-ovu /17/ teoriju konvergencije za bezuslovnu optimizaciju na problem oblika (1.22).

Ako je x_o proizvoljna tačka iz X, L^o se definiše kao povezana konponenta nivoskog skupa $L = \{x \in D \mid \phi(x) \leq \phi(x_o)\}$ koja sadrži x_o. Rauch dakle pretpostavlja da su sledeće pretpostavke uvek ispunjene:

1. $\operatorname{int} X = \{x \in D \mid f_j(x) < 0 \text{ za } \forall j \in I_o\} \neq \emptyset$;

2. svaka tačka skupa $\{x \in X \mid f_j(x) = 0$ za neko $j \in I_o\}$ jeste tačka nagomilavanja od $\operatorname{int}(X)$;

3. $L^o \cap X$ je zatvoreno i $\operatorname{int}(L^o \cap X) \neq \emptyset$;

4. ϕ je ograničena odozdo na $L^o \cap X$.

Rauch zatim daje definiciju prinudne funkcije koja predstavlja poopštenje pojma prinudne funkcije koju je definisao Elkin u /17/.

Definicija 1.14. Funkcija $\Phi: [0,\infty) \times \dots \times [0,\infty) \subset R^m \to R^+$ jeste prinudna funkcija od m promenljivih ako za svakih m nizova $\{t_i^k\} \subset [0,\infty)$, $i = 1,\dots,m$ važi

(1.23) $$\lim_{k\to\infty} \Phi(t_1^k,\dots,t_m^k) = 0 \Rightarrow \lim_{k\to\infty} t_i^k = 0$$

za bar jedno i, $1 \leq i \leq m$.

Za ma koji skup indeksa $I \subset I_o$ označavamo sa I^N, odnosno I^L potskupove indeksa $j \in I$ za koje su f_j nelinearne, odnosno linearne funkcije.

Rauch razmatra odredjene minimalne uslove koje mora da zadovoljava vektor pravca $p_k \in R^n$. Jasno, p_k mora biti dopustiv pravac u x_k (definicija 1.3).

Neka su $\mu_j: [o,\infty) \to [o,\infty)$, $j \in I_o^N$ date prinudne funkcije, gde je $\mu_j(o) = 0$ i, za svako $x \in X$, skup indeksa $I \subset I_o$ i $\nu \in [o,\infty)$, razmotrimo skup

$$K(x,I,\nu) = \{p \in R^n \mid \|p\| = 1,\ \langle\nabla f_j(x),p\rangle \geq \mu_j(\nu),\ j \in I^N,\ \langle\nabla f_j(x),p\rangle \geq o, j \in I^L\}$$

Iz teoreme o srednjoj vrednosti sledi da je nuždan uslov da je p dopustiv u tački x da $p \in K(x,I(x,o),o)$, ali u opštem slučaju ovaj uslov nije dovoljan. Da bismo do te ocene mogli da dodjemo, pretpostavićemo da je skup $L^O \cap X$ kompaktan i da je u datoj tački $x_k \in L^O \cap X$ izabran pravac takav da važi

$$(1.24) \qquad p_k \in K(x_k,\ I(x_k,\varepsilon_k),\ \nu_k), \quad \langle\nabla\phi(x_k),p_k\rangle \geq 0$$

za neko $\nu_k > 0$ i $\varepsilon_k > 0$.

Osnovu za Rauch-ove uslovne verzije postojećih bezuslovnih algoritama za nalaženje koraka α_k čine sledeće dve leme.

Lema 1.14. *Pretpostavimo da je skup $L^O \cap X$ kompaktan i da vektor $p_k \in R^n$ zadovoljava uslov (1.24) u tački $x_k \in L^O \cap X$.*
Ako je $\tilde{\alpha} = \sup\{\alpha \geq o \mid x_k - tp_k \in L^O$ za $\forall t \in [o,\alpha)\}$ i

$$(1.25) \qquad \alpha^* = \sup\{\alpha \geq 0 \mid x_k - tp_k \in X \quad \text{za } \forall t \in [o,\alpha)\},$$

tada je

$$(1.26) \qquad \alpha^* \geq \min\{\tilde{\alpha},\ \hat{F}(\nu_k,\varepsilon_k)\},$$

gde je $\hat{F}: [o,\infty) \times [o,\infty) \subset R^2 \to [o,\infty)$ prinudna funkcija od dve promenljive koja zavisi samo od $L^O \cap X$.

Dokaz. *Videti u Rauch /49/, str. 211.*

Lema 1.15. *(Poopštenje Elkin-ove leme iz /17/)* *Pretpostavimo da je skup* $L^o \cap X$ *kompaktan i da vektor* $p_k \in R^n$ *zadovoljava uslov (1.24) i da je* $<\nabla\phi(x_k), p_k> > 0$. *Tada* $L^o \cap X$ *sadrži otvoren interval* $(x_k, x_k - \alpha_o p_k)$ *za neko* $\alpha_o > 0$. *Dalje, ako je* $\hat{\alpha} > 0$ *ma koji broj takav da je*

$$(1.27) \qquad \phi(x_k - \alpha p_k) < \phi(x_k) \quad za \ \forall\, x_k - \alpha p_k \in (x_k, x_k - \hat{\alpha} p_k] \cap L^o \cap X,$$

tada $[x_k, x_k - \hat{\alpha} p_k] \cap X \subset L^o \cap X$.

Dokaz. *Videti u Rauch /49/, str. 212.*

Posle izlaganja uslovnih verzija poznatih alogaritama za nalaženje koraka α_k, *Rauch izlaže nekoliko algoritama za nalaženje vektora pravca* p_k *takvog da zadovoljava uslov (1.24). Ovi algoritmi, kombinovani sa algoritmima za nalaženje koraka, generišu nizove tačaka* $\{x_k\}$ *koji konvergiraju ka Kuhn-Tucker-ovoj tački problema (1.22) (definicija 1.5).*

Primedba. *Mangasarian O.L. je u /35/ pokazao da su Kuhn-Tucker-ovi uslovi /18/ i /19/ za tačku* $x^* \in X$ *(za problem 1.22) ekvivalentni sledećim uslovima*

$$(1.28) \qquad P_q(x^*)\nabla\phi(x^*) = 0, \qquad q = kard\ I\ (x^*, o) \ i$$

$$(1.29) \qquad \lambda_q(x^*) \leq 0 \qquad , \qquad q = kard\ I\ (x^*, o),$$

gde je $P_i(x^*) = I - N_i^T(x^*)\,[N_i(x^*)\ N_i(x^*)^T]^{-1}\ N_i(x^*)$, $i = 1, \ldots, q$

ortogonalna projekcija iz R^n *u ortogonalni komplement prostora koji obrazuju vektori* $n_i(x^*) = \nabla f_i(x^*)$, $i \in I(x^*, o)$, $N_i^T(x^*) = (n_1(x^*), n_2(x^*), \ldots, n_i(x^*))$, *a* $\lambda_q(x^*) = [N_q(x^*)\ N_q(x^*)^T]^{-1}\ N_q(x^*)\ \nabla\phi(x^*)$.

Ako je skup X regularan (definicija 1.4), rang $N_i = i < n$.

Izloženi algoritmi za nalaženje pravca uključuju Rosen-ov /52/, /53/ i Zountendijk-ov /64/ alogaritam. ε - procedura koja je uključena u ove algoritme odgovara Zountendijk-ovim zahtevima protiv kruženja. (Kruženje je pojava da, kad se ne vodi računa o ograničenjima koja nisu aktivna, ali su "skoro" aktivna u posmatranoj tački, dodje do konvergencije niza $\{x_k\}$ ka nestacionarnoj tački).

Alogaritam za ε

Korak 1. Ako je $\| P_q(x) \ \nabla\phi(x) \| = 0$, $q = \mathrm{kard}\, I(x,o)$, staviti $\varepsilon = 0$;

Korak 2. Inače, za fiksiranu monotono rastuću prinudnu funkciju $\hat{\mu}: [o,\infty) \to [o,\infty)$ takvu da je $\hat{\mu}(o) = 0$ i $\hat{\mu}(t) > 0$ kada je $t > 0$ i fiksirano $\xi > 1$, neka je ε prvi broj u nizu $\{\xi^{-j}\}_{j=0}^{\infty}$ takav da je $P_q(x)$, $q = \mathrm{kard}\, I(x,\varepsilon)$ dobro definisano i $\hat{\mu}(\| P_q(x) \ \nabla\phi(x) \|) \geqslant \varepsilon$.

Iz pretpostavljene regularnosti skupa X i osobina projekcije sledi da je ε dobro definisano.

Lema 1.16. Neka je skup X regularan, neka je x_o ma koja tačka iz skupa X takva da je skup $L^o \cap X$ kompaktan i neka je skup

$$\Omega = \{x \in L^o \cap X \mid \| P_q(x) \ \nabla\phi(x) \| = 0, \quad q = \mathrm{kard}\, I(x,o)\}$$

konačan. Neka je niz $\{x_k\} \subset L^o \cap X$ ma koji niz koji zadovoljava uslov:

$$\lim_{k\to\infty} \| P_{q_k}(x_k) \ \nabla\phi(x_k) \| = 0, \quad q_k = \mathrm{kard} I(x_k, \varepsilon_k) \tag{1.30}$$

za neki dati pozitivan niz $\{\varepsilon_k\}$ takav da je $\lim\limits_{k\to\infty} \varepsilon_k = 0$. Ako se Ω sastoji samo od jedne tačke, ili ako Ω sadrži konačno mnogo tačaka i

$\lim\limits_{k\to\infty} \| x_k - k_{k+1} \| = 0$, *tada je*

$$\lim_{k\to\infty} x_k = x^*, \quad \| P_q(x^*)\nabla\phi(x^*)\| = 0, \quad q = \text{kard } I(x^*,0).$$

Zatim, svi granični skupovi I^* *niza* $\{I(x_k,\varepsilon_k)\}$ *su sadržani u* $I(x^*,0)$.

Dokaz. *Videti u Rauch /49/, str. 224.*

1.4. Metoda Danilin-Pšeničnog /48/, str. 80.

U četvrtom poglavlju ćemo izložiti nekoliko modifikacija metode Danilin-Pšeničnog /13/, /48/. Zato ćemo ovde ukratko izložiti metodu Danilin-Pšeničnog, kao i najvažnije stavove koje su u vezi sa njom dali autori.

Razmatramo sledeći problem optimizacije: naći

$$\min \{\phi(x) \,|\, x \in R^n\},$$

gde pretpostavljamo da je funkcija $\phi: R^n \to R$ *dva puta neprekidno diferencijabilna.*

Neka je zadat beskonačan niz tačaka $\{x_k\} \subset R^n$. *Tom nizu pridružujemo niz* $\{y_k\} \subset R^n$ *koji definišemo na sledeći način:*

$$(1.31) \qquad y_k = x_k + r_k,$$

gde su vektori r_k *takvi da su zadovoljeni sledeći uslovi:*

1. *ako je* Δ_k *- determinanta čiji su stupci vektori* $\dfrac{r_k}{\|r_k\|}, \dots, \dfrac{r_{k-n+1}}{\|r_{k-n+1}\|}$, *tada je za* $\forall k \geq n-1$ $\Delta_k \geq \varepsilon$, *gde je* ε *- proizvoljno mali pozitivan broj.*

2. $\| r_k \| \to 0$ *kada* $k \to \infty$.

Za $k \geq n-1$ definišimo matricu A_k pomoću sledećeg sistema jednačina

(1.32) $\quad A_k r_{k-i} = e_{k-i}, \qquad i = 0,1,\dots,n-1$

gde je $e_{k-i} = \nabla\phi(y_{k-i}) - \nabla\phi(x_{k-i})$, a r_k i y_k - elementi niza (1.31).

Metoda Danilin-Pšeničnog se definiše sledećom iteracijom:

(1.33) $\quad x_{k+1} = x_k - \alpha_k p_k, \qquad k = 0,1,2\dots,$

gde je vektor pravca $p_k = A_k^{-1}\,\nabla\phi(x_k)$ za $k \geq n-1$, a matrica A_k je zadata sistemom (1.32). (Za $k = 0,1,\dots,n-2$ za p_k se može uzeti npr. sam gradijent $\nabla\phi(x_k)$). Korak α_k se izračunava pomoću sledećeg algoritma:

Korak 1. Uzimamo $\alpha = 1$ i izračunavamo tačku $x = x_k - \alpha p_k$.

Korak 2. Izračunavamo $\phi(x_k - \alpha p_k)$.

Korak 3. Proveravamo da li važi nejednakost

(1.34) $\quad \phi(x_k) - \phi(x) \geq \varepsilon\alpha \langle\nabla\phi(x_k),p_k\rangle, \quad 0<\varepsilon<\frac{1}{2}$.

Korak 4. Ako je nejednakost 1.34 zadovoljena, tada vrednost $\alpha = 1$ uzimamo za α_k; inače, vršimo umanjivanje α sve do tada dok uslov (1.34) nije zadovoljen.

Napomenimo ovde da su u radovima Danilina /12/ i Bulanij-a i Danilina /6/ izloženi samo specijalni slučajevi gornje metode dobijeni odredjenim izborom vektora r_k.

Ovde ćemo navesti i stavove iz /48/ koji se odnose na metodu.

Lema 1.17. Neka je $\{x_k\}$ ograničeni niz, $\|x_{k+1}-x_k\| \to 0$, $k\to\infty$ i za svako $k \geq n-1$ matrica A_k je definisana sistemom jednačina (1.32). Tada je

(1.35) $\quad \lim\limits_{k\to\infty} \|A_k - H(x_k)\| = 0,$

gde je $H(x)$ - hesijan funkcije ϕ u tački x.

Dokaz. *Videti u /48/, str. 80.*

Lema 1.18. *Neka je $\phi: R^n \to R$ dva puta neprekidno diferencijabilna funkcija takva da postoje m i M, $0 < m < M < \infty$ da je zadovoljen uslov:*

$$(1.36) \qquad m\|y\|^2 \leq \langle y, H(x)y\rangle \leq M\|y\|^2 \qquad \text{za } \forall x, y \in R^n$$

i neka je niz $\{x_k\}$ takav da je $\phi(x_{k+1}) \leq \phi(x_k)$ i $\langle\nabla\phi(x_k), x_{k+1} - x_k\rangle \to 0$, $k \to \infty$. Tada $\|x_{k+1} - x_k\| \to 0$, $k \to \infty$.

Dokaz. *Videti u /48/, str. 82.*

1.5. O brzini konvergencije

Kod razmatranja iteracionog alogaritma za nalaženje tačaka bilo lokalnog bilo globalnog minimuma važnu ulogu igraju globalna odnosno, lokalna konvergencija. Globalna konvergencija razmatra problem da li će generisani niz konvergirati ka rešenju ako se za početnu tačku uzme proizvoljna tačka. Lokalna konvergencija karakteriše krajnju brzinu konvergencije (tj. ponašanje proizvoljno dalekog dela niza) u opštem slučaju, odredjuje relativne prednosti jednog algoritma u odnosu na drugi.

Definicije Q i R faktora konvergencije, koje ćemo ovde dati, su motivisane činjenicom da se u razmatranju odredjenog iteracionog procesa, kod kojeg je x^* jedna od tačaka nagomilavanja, prirodno nameću ocene oblika

$$(1.37) \qquad \|x_{k+1} - x^*\| \leq \beta\|x_k - x^*\|^p, \quad \forall k \geq k_0.$$

Definisaćemo ove indikatora asimptotske brzine konvergencije za proizvoljan konvergentan niz.

Definicija 1.15. *Neka je* $\{x_k\} \subset R^n$ *proizvoljan konvergentan niz sa graničnom vrednošću* x^*. *Tada su veličine*

$$(1.38)\qquad Q_p\{x_k\} = \begin{cases} 0 & \text{, ako je } x_k = x^* \text{ za sve osim konačno mnogo } k; \\ \limsup\limits_{k\to\infty} \dfrac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} & \text{, ako je } x_k \neq x^* \text{ za sve osim konačno mnogo } k; \\ \infty & \text{, inače.} \end{cases}$$

Definisane za $\forall p \in [1,+\infty)$, *količnički faktori konvergencije ili Q - faktori u odnosu na normu* $\|\ \|$ *na* R^n.

Ako je $Q_p = Q_p\{x_k\} < +\infty$ *za neko* $p \in [1,\infty)$, *tada za* $\forall \varepsilon > 0\ \exists k_o$ *takvo da (1.37) važi sa* $\beta = Q_p + \varepsilon$.

Definicija 1.16. *Neka je* $\{x_k\} \subset R^n$ *proizvoljan konvergentan niz sa graničnom vrednošću* x^*. *Tada je Q - red (ili često, samo - red) konvergencije niza* $\{x_k\}$ *definisan na sledeći način:*

$$(1.39)\qquad Q_Q\{x_k\} = \begin{cases} +\infty & \text{ako je } Q_p\{x_k\} = 0\ ,\ \forall p \in [1,\infty) \\ \inf\{p \in [1,\infty) \mid Q_p\{x_k\} = \infty\} & \text{inače.} \end{cases}$$

Od posebnog je interesa slučaj kada je Q - red konvergencije jednak 1. Kažemo da je konvergencija niza $\{x_k\}$ *Q - linearna (ili sámo - linearna), ako je* $0 < Q_1 < 1$. *Ako je* $Q_1 = 0$, *kažemo da niz* $\{x_k\}$ *ima Q - superlinearnu (ili sámo - superlinearnu) konvergenciju u* x^*.

Imajući u vidu da ako je $Q_1 < +\infty$ *da za* $\forall \varepsilon > 0\ \exists k_o(\varepsilon)$ *tako da ocena (1.37) važi sa* $\beta = Q_1 + \varepsilon$, *vidimo da je definicija linearne konvergencije niza ekvivalentna sa sledećim uslovom:*

$$(1.40)\qquad \|x_{k+1} - x^*\| \leq \beta \|x_k - x^*\| \quad \forall k \geq k_o, \quad 0 < \beta < 1;$$

odnosno, za $k_o = 0$, *sa*

(1.41) $\|x_{k+1}-x^*\| \leq \beta \|x_k - x^*\| \leq \ldots\ldots \leq \beta^{k+1} \|x_o - x^*\|, \quad 0 < \beta < 1.$

Analogno, definicija superlinearne konvergencije je ekvivalentna sa sledećom ocenom

(1.42) $\|x_{k+1} - x^*\| \leq \beta_k \|x_k - x^*\| \quad$ *gde* $\beta_k \to 0, \quad k \to \infty$

Primetimo da su definicije linearne, odnoso superlinearne konvergencije date relacijama (1.40), (1.41) i (1.42) najviše uobičajene.

Uzimanjem pogodnog korena umesto količnika uzastopnih grešaka, dobijamo korenske (ili srednje) faktore konvergencije.

Definicija 1.17. *Neka je* $\{x_k\} \subset R^n$ *ma koji niz koji konvergira ka* x^*. *Tada su brojevi*

$$R_p\{x_k\} = \begin{cases} \limsup\limits_{k\to\infty} \|x_k - x^*\|^{\frac{1}{k}}, & \text{ako je } p = 1; \\ \limsup\limits_{k\to\infty} \|x_k - x^*\|^{\frac{1}{p^k}}, & \text{ako je } p > 1 \end{cases}$$

faktori korenske (srednje) konvergencije ili, kraće, R - faktori.

Definicija 1.18. *Neka je* $\{x_k\} \subset R^n$ *ma koji niz koji konvergira ka* x^*. *Tada se veličina*

$$O_R\{x_k\} = \begin{cases} \infty & \text{ako je } R_p\{x_k\} = o, \ \forall p \in [1, \infty) \\ \inf\{p \in [1,\infty) \mid R_p\{x_k\} = 1\}, & \text{inače} \end{cases}$$

naziva R - red konvergencije niza $\{x_k\}$.

Za razliku od Q - faktora niza, R - faktori su uvek nezavisni od norme na R^n, *kao što tvrdi sledeća lema:*

Lema 1.19. *Neka je* $\{x_k\} \subset R^n$ *proizvoljan niz takav da konvergira ka* x^*. *Tada je* $R_p\ \{x_k\}$ *nezavisan od norme na* R^n *za* $\forall p \in [1,\infty)$.

Dokaz. *Videti u /41/, str. 288.*

Primetimo da, kada $\{x_k\}$ *konvergira ka* x^*, *da tada uvek postoji* $k_o \geq 0$ *takvo da je*

$$0 \leq \|x_k - x^*\| < 1 \quad , \qquad k \geq k_o,$$

pa je odatle, $0 \leq R_p\ \{x_k\} \leq 1$ *za* $\forall p \geq 1$.

Ako je $0 < R_1\ \{x_k\} < 1$, *kažemo da je konvergencija niza* $\{x_k\}$ *u* x^* *R - linearna, dok ako je* $R_1 = 0$, *kažemo da je konvergencija R - superlinearna.*

Kada $\{x_k\} \subset R^n$ *konvergira ka* x^* *i* $0 < Q_p\{x_k\} < +\infty$ *i* $0 < R_p\{x_k\} < 1$ *za neko* $p > 1$, *tada uvek možemo dobiti jednu od relacija:* $R_p\{x_k\} < Q_p\{x_k\}$ *ili* $R_p\{x_k\} > Q_p\{x_k\}$ *izborom pogodne norme. Medjutim, ovo se ne može postići za* $p = 1$, *kao što tvrdi sledeća lema:*

Lema 1.20. *Neka niz* $\{x_k\} \subset R^n$ *konvergira ka* x^*. *Tada je*

$$R_1\ \{x_k\} \leq Q_1\ \{x_k\}$$

u svakoj normi.

Dokaz. *Videti u /41/, str. 296.*

Na kraju navedimo jednu teoremu o oceni brzine konvergencije koju ćemo koristiti u poglavlju 3.

Teorema 1.21. *(Vidi /41/, str.477)* *Neka je* $\phi : D \subset R^n \to R$ *neprekidno diferencijabilna funkcija na otvorenom skupu* $D_o \subset D$ *i pretpostavimo da* $\{x_k\} \subset D_o$ *konvergira ka* $x^* \in D_o$. *Pretpostavimo da je* $\nabla\phi(x^*) = 0$, *da je* ϕ *dva puta neprekidno diferencijabilna u nekoj okolini od* x^*, *da je hesijan* $H(x^*)$ *nesingularan i da postoji* $\eta > 0$ *i* k_o *takvo da važi*

$$\phi(x_k) - \phi(x_{k+1}) \geq \eta \|\nabla\phi(x_k)\|^2, \quad \forall k \geq k_o.$$

Tada je $R_1\{x_k\} < 1$.

Dokaz. *Videti /41/, str. 477.*

2. POGLAVLJE

2.1. O projektivnim metodama

Projektivne metode predstavljaju posebnu grupu metoda dopustivih pravaca koje su prvenstveno definisane za rešavanje problema (1.1) kada su funkcije f_i, $i \in I_o$ *linearne, tj. problema oblika*

$$(2.1) \qquad \min\{\phi(x)\,|\,x \in X\}, \quad X = \{x \in R^n \,|\, \langle a_i, x\rangle \leq b_i,\ i \in I_o\}$$

$$a_i \in R^n, \quad b_i \in R, \quad i \in I_o \quad i \quad I_o = \{1,\dots,m\}.$$

Pri tom se pretpostavlja da je $int(X) \neq \emptyset$, *tj. da za* $\forall i \in I_o$ *postoji* $x \in X$ *takvo da je* $\langle a_i, x\rangle < b_i$.

Iz date dopustive tačke x_k *dopustivi pravac* p_k *se odredjuje u linearnoj mnogostrukosti definisanoj presekom aktivnih ograničenja u tački* x_k. *U slučaju nelinearnih ograničenja pravac* p_k *se projektuje u presek tangentnih hiperravni ograničenja aktivnih u* x_k, *pa se onda posebnim postupkom dobijena tačka "vraća" u dopustivi skup.*

Klasična projektivna metoda jeste Rosen-ova /52/, /53/ koja je ustvari projektivna varijanta Cauchy-jeve metode najbržeg spuštanja /8/.

Opisaćemo kako izgleda k-ti ciklus Rosen-ovog algoritma.

Na početku k-tog ciklusa je dato sledeće:
$x_k \in X$, $\nabla\phi(x_k)$, $I(x_k,0) = \{i \in I_o \,|\, \langle a_i, x_k\rangle = b_i\}$, $A_k^T = (a_i)$, $i \in I(x_k,0)$,
$P_k = I - A_k^T (A_k A_k^T)^{-1} A_k$.

Korak 1. *Neka je* $\lambda_k = (A_k A_k^T)^{-1} A_k \nabla\phi(x_k) = (\lambda_{ik})$, $i \in I(x_k, 0)$ *i neka je* $e_k \in I(x_k,0)$ *takvo da važi*

$$\lambda_{e_k k} = max \{\lambda_{ik}, i \in I(x_k,0)\}.$$

Stavimo

$$p_k = \begin{cases} P_k\nabla\phi(x_k), & ako\ je\ \|P_k\nabla\phi\ (x_k)\| \neq 0; \\ \bar{P}_k\nabla\phi(x_k), & ako\ je\ \|P_k\nabla\phi(x_k)\| = 0 \quad i\ \lambda_{e_k k} > 0, \end{cases}$$

gde je $\bar{P}_k = I - \bar{A}_k^T(\bar{A}_k\bar{A}_k^T)^{-1}\bar{A}_k = (a_i),\quad i \in \bar{I}(x_k,0),\ \bar{I}(x_k,0) = I(x_k,0)\backslash\{e_k\}.$

Korak 2. *Neka je*

$$\alpha_k^* = min\left\{\frac{\langle a_i,x_k\rangle - b_i}{\langle a_i,p_k\rangle}\ \middle|\ \langle a_i,p_k\rangle > 0,\ i \in I_o\right\},$$

$$\hat{\alpha}_k = sup\ \{\alpha > 0\ |\ \phi\ (x_k - \mu p_k) > \phi(x_k - \alpha p_k)\quad za\ \forall\mu < \alpha\}\ ,$$

$$\alpha_k = min\ \{\hat{\alpha}_k, \alpha_k^*\}.$$

Korak 3. *Staviti* $x_{k+1} = x_k - \alpha_k p_k$. *Izračunati* $\nabla\phi(x_{k+1})$.
Staviti $I(x_{k+1},0) = \{i \in I_o\ |\langle a_i,x_{k+1}\rangle = b_i\},\ A_{k+1}^T = (a_i),\ i \in I(x_{k+1},0)$
$P_{k+1} = I - A_{k+1}^T\ (A_{k+1}A_{k+1}^T)^{-1}\ A_{k+1}$.

Zameniti k sa k+1 i ići na korak 1.

Kriterijum zaustavljanja. *Stati ako je* $P_k\nabla\phi(x_k) = 0$ *i* $\lambda_{e_k k} \leq 0$.

Rosen-ova teorema, koju ćemo sad izložiti predstavlja samo specijalan slučaj Kuhn-Tucker-ove teoreme (teorema 1.3) koji se odnosi na problem (2.1).

Teorema 2.1. *(Rosen /52/)* *Neka je $\phi : R^n \to R$ neprekidno diferencijabilna strogo konveksna funkcija. Neka x^* pripada preseku q hiperravni. Pretpostavimo da je u x^* zadovoljen uslov regularnosti (1.5). Tada je x^* optimalno rešenje problema (2.1) ako i samo ako važi:*

$$(2.2) \qquad P_q(x^*)\nabla\phi(x^*) = 0 \quad i$$

$$q = \text{kard } I(x^*,0)$$

$$(2.3) \qquad \lambda_q(x^*) = 0, \ \lambda_q(x^*) = (A_q A_q^T)^{-1} A_q \nabla\phi(x^*), \ A_q^T = (a_i), \ i \in I(x^*,0)$$

Dokaz. *Videti u /52/.*

Medjutim, niz tačaka $\{x_k\}$ generisan gornjim originalnim Rosen-ovim algoritmom ne mora da konvergira ka optimalnom rešenju problema (2.1) zbog fenomena takozvanog kruženja (zastrevanija). Kao što smo već pomenuli, kruženje je pojava da ako razmatramo samo ograničenja koja su aktivna u posmatranoj tački x_k, a ne vodimo računa o onim "skoro" aktivnim, može doći do toga da korak α_k teži nuli i onda kada smo daleko od rešenja. Polak /44/ je dao ε - varijantu Rosen-ove metode i dao korektan dokaz konvergencije ka optimalnom rešenju problema (2.1).

Shanno D.F. u /55/ predlaže ubrzanje Rosen-ove metode za specijalan slučaj problema (2.1).

Goldfarb /25/ je u svojoj doktorskoj disertaciji dao projektivnu varijantu metoda promenljive metrike Davidon-Fletcher-Powell-a /15/, /20/, ali nije dao dokaz konvergencije dobijenog niza tačaka ka optimalnom rešenju problema (2.1) za funkciju cilja opšteg oblika, već samo za kvadratnu funkciju. Iako je ova metoda u praksi pokazala dobre rezultate, do dokaza konvergencije za opšti oblik funkcije cilja nije se ni do danas došlo.

Ritter /52/, Best i Ritter /5/ uvodjenjem odgovarajućih procedura protiv kruženja obezbedjuju konvergenciju ka optimalnom rešenju problema (2.1) nizova tačaka $\{x_k\}$ generisanih njihovim projektivnim metodama.

Pomenimo ovde i projektivnu metodu Danilin-a /11/ iako je ona više od teorijskog značaja. Danilin razmatra sledeći problem

$$(2.4)\quad \min\{\phi(x)\mid x\in X\}\ ,\quad X=\{x\in R^n\mid f(x)=0\}\ ,\quad f=(f_1,\dots,f_m),$$

gde su funkcije ϕ, $f_i: R^n\to R$ *neprekidno diferencijabilne, a* X *- glatka* $(n-m)$ *dimenziona mnogostrukost, tj u ma kojoj tački* $x\in X$ *rang matrice* ∇f *je jednak* m. *Ako pretpostavimo, odredjenosti radi, da važi za* $\forall\, x\in X$

$$(2.5)\quad \det\left(\frac{\partial f_i(x)}{\partial x_j}\right)\neq 0,\ i,j=1,\dots,m,\ \text{tada}$$

na osnovu teoreme o implicitnim funkcijama u okolini tačke x *sistem* $f(x)=0$ *definiše funkciju* $y=y(t)$, $y\in R^m$, $y=(x_1,\dots,x_m)$, $t\in R^{n-m}$, $t=(x_{m+1},\dots,x_n)$, $x=(y,t)$. *(vidi /11/, str.88)*

Problem (2.7) se na taj način svodi na problem bezuslovne minimizacije funkcije $F(t)=\phi(t,y(t))$. *Danilin u /11/ izlaže iteracione metode oblika*

$$t_{k+1}=t_k-\alpha_k p_k,\quad y_{k+1}=y(t_{k+1}),\quad x_{k+1}=(y_{k+1},\ t_{k+1})$$

i pri tom izlaže nekoliko izbora za vektor p_k, *a korak* α_k *definiše tako da zadovoljava uslov*

$$F(t_k)-F(t)\geq \varepsilon\alpha<\nabla F(t_k),p_k>,\qquad 0<\varepsilon<\frac{1}{2}\,.$$

U poglavlju 5. ćemo dati poopštenje Danilin-ovih rezultata.

Gill i Murray u /22/ daju originalan pristup projektivnim metodama u kojem zadatak uslovne optimizacije (2.1) takodje svode na zadatak bezuslovne minimizacije u prostoru R^q, *gde je* q *- broj ograničenja aktivnih u tački* x_k.

Izložićemo Gill-Murray-ev algoritam u kratkim crtama.

Napišimo problem (2.1) korišćenjem matrice ograničenja, t. u obliku:

$$(2.6)\quad \min\{\phi(x)\mid A^Tx\leq b\},\quad \text{gde je } A - n{\times}m \text{ matrica}$$

Neka je x_k ma koja dopustiva tačka sa odgovarajućom $n \times q$ matricom aktivnih ograničenja A_k i desnom stranom - vektorom b_k. Da bi tačka $x_{k+1} = x_k - \alpha_k p_x$ zadovoljavala ograničenja koja su aktivna u x_k, p_k mora pripadati prostoru ortogonalnom na prostor razapet normalama (gradijentima) aktivnih ograničenja. Analogno Newton-ovoj metodi, gde je pravac definisan bezuslovnim minimumom kvadratne funkcije, Gill i Murray definišu pravac p_k kao minimum kvadratne aproksimacije funkcije ϕ uz uslov da p_k pripada pomenutom prostoru.

Naime, ako imamo kvadratnu aproksimaciju funkcije ϕ u okolini tačke x_k

$$\phi(x_k+p) \approx Q(p) = \phi(x_k) + \langle\nabla\phi(x_k),p\rangle + \frac{1}{2}\langle H(x_k)p,p\rangle ,$$

vektor pravca p_k je rešenje sledećeg optimizacionog problema:

$$(2.7) \qquad \min \{Q(p) \mid A_k^T p = 0\} .$$

Gill i Murray svode problem (2.6) na bezuslovnu optimizaciju na sledeći način.

Označimo sa M linearnu mnogostrukost definisanu presekom q aktivnih ograničenja u tački x_k:

$$M = \{x \in R^n \mid A_k^T x = b_k\} .$$

Vektorski prostor $M_o = \{ y \in R^n \mid A_k^T y = 0\}$ je paralelan sa M i, ako su vrste od A_k^T linearno nezavisne, $\dim M_o = n-q$. Prema tome, svaki vektor iz M_o se može izraziti kao linearna kombinacija od $n-q$ bazisnih vektora $\{z_j\}_{j=1}^{n-q}$. Ovo povlači da ako je Z - matrica, čije su kolone z_j, tada za svako $y \in M_o$ postoji $n-q$ skalara v_j takvih da je

$$y = Zv, \qquad v = (v_1,\dots,v_{n-q})$$

Ako izrazimo p u ovom obliku i zamenimo u (2.7), dobićemo (odbacivanjem konstantnog člana):

(2.8) $\min \{\frac{1}{2} < v, Z^T H(x_k) Zv > + < \nabla \phi(x_k), Zv > | v \in R^{n-q} \}$,

što je sada bezuslovni problem sa n-q promenljivih. Ako je matrica $Z^T H_k Z$ pozitivno definitna, rešenje v^* problema (2.8) će biti rešenje sledeće matrične jednačine:

(2.9) $Z^T H(x_k) Z v^* = -Z^T \nabla\phi(x_k)$.

Rešenje p^* problema (2.7) će onda glasiti $p^* = Zv^*$.

Kada smo vektor pravca p_k našli na ovaj način, nalazimo korak α_k na isti način kao u Rosen-ovoj metodi, i nalazimo tačku $x_{k+1} = x_k - \alpha_k p_k$ i proces iteriramo dok ne dodjemo do rešenja problema (2.6).

Karakteristika ovakvog prilaza jeste da što je veći broj ograničenja aktivnih u x_k, manja je dimenzija sistema jednačina (2.9) koji treba rešiti da bismo odredili p_k.

Gill i Murray su razvili nekoliko efikasnih postupaka za nalaženje matrice Z- tj. bazisnih vektora n-q dimenzionog prostora, kao i za efikasno dobijanje nove matrice iz prethodne za slučaj kada se odgovarajuće aktivno ograničenje uklanja, odnosno, kada novo neaktivno ograničenje postaje aktivno u tački x_{k+1}. O ovome se detaljnije može naći u radovima /23/, /24/, /26/.

Navešćemo neke od ovih postupaka.

1. Formirajmo matricu T na sledeći način

$$T = \begin{bmatrix} A_k^T \\ V_k \end{bmatrix},$$

tj. na q linearno nezavisnih vrsta matrice A_k^T dodato je n-q proizvoljnih linearno nezavisnih vrsta koje čine matricu V_k.

Nije teško videti da matricu Z čine upravo n-q poslednjih kolona od T^{-1}, tj.

$$(T^T)^{-1} = \begin{bmatrix} S \\ Z^T \end{bmatrix}, \text{ na primer.}$$

Medjutim, ovaj metod ima dva osnovna nedostatka: prvi - kada zbog nezgodnog izbora Z može da se desi da projekcija gradijenta $\nabla\phi(x_k)_p = Z^T\nabla\phi(x_k)$ bude velika čak i kad je x_k blizu stacionarne tačke, drugi - nepogodan izbor Z može dovesti do vrlo slabo uslovljene matrice $H(x_k)_p = Z^T H(x_k) Z$ - projektovanog hesijana, čak i kad je polazni problem dobro uslovljen.

Sledeći postupak za nalaženje Z obezbedjuje da projektovani hesijan $H(x_k)_p$ ima minimalni kondition number - odnosno veličinu uslovljenosti.

Pri tom veličina uslovljenosti $k(H(x_k)_p)$, koja se definiše kao

$$k(H(x_k)_p = \| H(x_k)_p \| \, \| H(x_k)_p^+ \|$$

gde je $H(x_k)_p^+ = (H(x_k)_p^T H(x_k)_p)^{-1} H(x_k)_p$ - pseudoinverz od $H(x_k)_p$, zavisi od $k(H(x_k))$ i od $[k(Z)]^2$. (/22/, str. 60)

2. Pretpostavimo da je V_k proizvoljno i da je poznata LQ faktorizacija matrice T, tj.

$$T = LQ,$$

gde je L - donja trougaona matrica, a Q - ortogonalna matrica. Q ćemo razdeliti na sledeći način

$$Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix},$$

Q_1 je $q \times n$ matrica, a Q_2 je $(n-q) \times n$ matrica. Nije teško pokazati da je $Z = Q_2^T$ (Vidi /22/, str. 61).

2.2. O optimizacionim metodama koje ne koriste izvode

Sve metode koje smo dosad razmatrali u oba poglavlja koriste prve, a neke i druge parcijalne izvode funkcije cilja i pomoću njih odredjuju egzistenciju i položaj traženog minimuma.

Ipak u problemima optimizacije u praksi često je izračunavanje samih vrednosti funkcije cilja odnosno ograničenjâ veoma dugo i komplikovano, pa je nalaženje eksplicitnih izraza za potrebne izvode još teže ili čak i nemoguće. U ovakvim slučajevima pribegava se numeričkom diferenciranju, tj. aproksimacijama izvoda pomoću konačnih razlika. Medjutim, greške (aproksimacije, zaokruženja) koje povlače ovakvi pristupi, mogu da dovedu do sasvim spore konvergencije algoritma, ili čak da dovedu dotle da algoritam uopšte ne konvergira iako je teorijski konvergencija obezbedjena. Stewart /54/ je na ovakav način modifikovao Davidon-Fletcher-Powell-ovu metodu, Gill i Murray njihovu gore pomenutu quazi-Newton-ovu metodu, itd.

Zbog navedenih teškoća koje prate numeričko diferenciranje za praksu su veoma značajne optimizacione metode koje zahtevaju samo izračunavanje vrednosti funkcije cilja. Ovakve metode, zbog svoje prirode, ne zahtevaju jake pretpostavke o funkciji cilja, (obično samo neprekidnost) pa zato imaju vrlo široko polje primene: npr. kada su izvodi prekidni, kada vrednosti funkcije predstavljaju rezultat eksperimenta itd.

Medjutim, dok gradijentne metode imaju dokaze optimalnosti, većina metoda koje koriste izvode ih nema i zato one najčešće daju ne optimalno, već neko "dobro" rešenje, koje za praktične svrhe aproksimira optimalno rešenje. Takodje, ovde su i problemi dokaza konvergencije prisutni za razliku od metoda sa dobrom teorijskom podlogom. Efikasnost ovih metoda je u praksi mnogo slabija od gradijentnih metoda pošto koriste vrlo malo podataka o posmatranom optimalnom problemu.

Metode koje ne koriste izvode su ipak, uprkos svim navedenim nedostacima vrlo važne i interesantne za izučavanje jer se u praksi mogu primeniti onda kada sve druge metode otkažu.

Sada ćemo izložiti nekoliko poznatijih metoda za bezuslovnu optimizaciju koje ne koriste izvode.

Ciklična koordinatna metoda

Razmatramo sledeći problem: naći

(*) $\min\{\phi(x) \mid x \in R^n\}$.

Pretpostavićemo da postoji takva tačka $\bar{x} \in R^n$ *da je nivoski skup* $\{x \in R^n \mid \phi(x) \leq \phi(\bar{x})\}$ *ograničen.*

Algoritam

Korak 1. *Naći takvu tačku* $x_0^1 \in R^n$ *da je skup*

$L(x_0^1) = \{x \in R^n \mid \phi(x) \leq \phi(x_0^1)\}$ *ograničen. Staviti* $k = 1$.

Korak 2. *Za* $i = 1,\ldots,n$ *izračunati*

$x_i^k = x_{i-1}^k + \alpha_i^k e_i$, *gde je*

$\phi(x_{i-1}^k + \alpha_i^k e_i) = \min\limits_{\alpha} \phi(x_{i-1}^k + \alpha e_i)$, *a* $e_i, i=1,\ldots,n$ *su jedinični koordinatni vektori.*

Korak 3. *Staviti* $x_0^{k+1} = x_n^k$.

Naći $\alpha_s^k = \max\{|\alpha_i^k|, i = 1,\ldots,n\}$.

Ako je $\alpha_s^k = 0$, *ići na korak 4; inače zameniti* k *sa* $k+1$ *i ići na korak 2.*

Korak 4. *Staviti* $x_{opt} = x^{o}_{k+1}$; *STOP.*

Teorema 22. *Neka je* $\phi: R^n \to R$ *konveksna diferencijabilna funkcija. Tada je niz* $\{x^k_i\}$, $k=1,2,\ldots, i=0,1,\ldots,n$, *generisan gornjim algoritmom ili konačan i njegova poslednja tačka je optimalno rešenje problema (2.13), ili beskonačan i svaka njegova tačka nagomilavanja je optimalno rešenje problema (2.13).*

Dokaz. *Iz konveksnosti i neprekidnosti funkcije* ϕ *na* R^n *sledi zatvorenost skupa* $L(x^1_o)$. *Odavde i iz ograničenosti skupa* $L(x^1_o)$ *(korak 1) sledi kompaktnost* $L(x^1_o)$.

Po konstrukciji je $\phi(x^k_i) \le \phi(x^k_{i-1}) \le \phi(x^1_o)$ *za* k *i* $i=1,2,\ldots,n$. *Odatle sledi da* $\{x^k_i\} \subset L(x^1_o)$.

Iz kompaktnosti skupa $L(x^1_o)$ *sledi da postoji* $K \subset N$ *tako da važi*

$$(2.11) \quad x^k_i \to x_i, \quad k \in K, \; k \to \infty, \quad x_i \in L(x^1_o), \quad i = 1,\ldots,n.$$

Niz $\{\phi(x^k_i)\}$ *je monotono opadajući i ograničen odozdo, dakle, konvergentan. Prema tome važi*

$$(2.12) \quad \phi(x^k_i) \to \bar{\phi}, \quad k \to \infty, \quad i = 1,\ldots,n.$$

Dakle, iz (2.11), (2.12) i neprekidnosti ϕ *sledi*

$$(2.13) \quad \lim_{k\to\infty} \phi(x^k_i) = \lim_{\substack{k\in K\\ k\to\infty}} \phi(x^k_i) = \phi(x_i) = \bar{\phi}, \quad i = 1,\ldots,n$$

Odavde sledi da je $\phi(x_{i+1}) = \phi(x_i)$, $i = 1,\ldots,n$.

Dokažimo da je bilo koja od tačaka nagomilavanja x_i *optimalno rešenje. Pokažimo da to važi npr. za* $i = 1$ *i označimo* x_1 *sa* x^*. *Iz (2.11) sledi*

$$x^k_1 \to x_1 = x^*, \quad k \in K, \; k \to \infty$$

Iz $x_1^k = x_1^k + \alpha_2^k e_2$ sledi, za $k \in K$, $k \to \infty$

$$x_2 = x^* + \alpha_2 e_2,$$

gde je $\alpha_2 = \lim\limits_{\substack{k\to\infty \\ k\in K}} \alpha_2^k$ i $\phi(x^* + \alpha_2 e_2) \leq \phi(x^* + \alpha e_2)$ za $\forall \alpha \in R$.

Ako pretpostavimo da x^* nije optimalno rešenje problema (2.10), biće zbog diferencijabilnosti, $\phi(x_i) < \phi(x^*)$ za neko $i \in \{2,\dots,n\}$, a to je kontradiktorno sa (2.13). Dakle, zbog (2.13) će važiti

$$\min_{\alpha} \phi(x^* + \alpha e_i) = \phi(x^*), \quad i = 1,\dots,n.$$

U slučaju kada je niz $\{x_i^k\}$ konačan, dokaz je trivijalan i sledi neposredno iz algoritma.

Metoda lokalnih varijacija /44/, str. 63.

Algoritam. Uvedimo oznake $d_1 = e_1$, $d_2 = -e_1, \dots, d_{2n-1} = e_n$, $d_{2n} = -e_n$, gde su e_i, $i = 1,\dots,n$ – jedinični koordinatni vektori.

Korak 0. Izabrati takvu tačku $x_o \in R^n$ da je skup $L(x_o) = \{x \in R^n \mid \phi(x) \leq \phi(x_o)\}$ ograničen. Izabrati $\varepsilon_o > 0$.

Korak 1. Staviti $k = 0$; staviti $x = x_o$ i izračunati $\phi(x)$.

Korak 2. Staviti $\varepsilon = \varepsilon_k$.

Korak 3. Staviti $j = 1$.

Korak 4. Izračunati $\phi(x + \varepsilon d_j)$.

Korak 5. *Ako je* $\phi(x+\varepsilon d_j) < \phi(x)$, *zameniti* x *sa* $x+\varepsilon d_j$ *i ići na korak 3; inače ići na korak 6.*

Korak 6. *Ako je* $j < 2n$, *zameniti* j *sa* $j+1$ *i ići na korak 4; inače ići na korak 7.*

Korak 7. *Staviti* $x_{k+1} = x$; *staviti* $\varepsilon_{k+1} = \frac{\varepsilon}{2}$; *zameniti* k *sa* $k+1$ *i ići na korak 2.*

Teorema 2.3. *Neka je* $\phi\colon R^n \to R$ *neprekidno diferencijabilna funkcija. Tada je svaka tačka nagomilavanja niza* $\{x_k\}$ *generisanog gornjim algoritmom stacionarna tačka problema (2.10).*

Dokaz. *Videti u /44/, str. 64.*

Powell-Zangwill-ova metoda /62/, str. 294.

Označimo sa $s_1^o, \ldots, s_n^o$ *koordinatne pravce u* R^n. *Pretpostavićemo da su pravci normalizovani, tj. da važi* $\|s_i^o\| = 1$, $i = 1,\ldots,n$.

Algoritam. *Skalar* $0 < \varepsilon$ 1 *je zadat.*

Korak 0. *Naći tačku* $x_o \in R^n$ *takvu da je skup*

$L = \{x \in R^n \mid \phi(x) \leq \phi(x_o)$ *ograničen. Staviti* $\delta^o = 1$. *Staviti* $k = 0$.

Korak 1. *Za* $i = 1,\dots,n$ *izračunati* $\alpha_i^k \in R$ *tako da minimizira* $\phi(x_{i-1}^k + \alpha_i s_i^k)$ *i staviti* $x_i^k = x_{i-1}^k + \alpha_i^k s_i^k$.

Korak 2. *Staviti* $\lambda^k = \|x_n^k - x_o^k\|$ *i* $s_{n+1}^k = \dfrac{x_n^k - x_o^k}{\lambda^k}$. *Ako je* $\lambda^k \leq \varepsilon$, *STOP; inače izračunati* α_{n+1}^k *tako da minimizira* $\phi(x_n^k + \alpha_{n+1} s_{n+1}^k)$ *i staviti* $x_o^{k+1} = x_{n+1}^k = x_n^k + \alpha_{n+1}^k s_n^k$.

Korak 3. *Staviti* $\alpha_p^k = \max\{\alpha_i^k \mid i = 1,\dots,n\}$.

a) Ako je $\dfrac{\alpha_p^k \delta^k}{\lambda^k} \geq \varepsilon$, *staviti* $s_i^{k+1} = s_i^k$ *za* $i \neq p$, $s_p^{k+1} = s_{n+1}^k$ *i staviti* $\delta^{k+1} = \dfrac{\alpha_p^k \delta^k}{\lambda^k}$.

b) Ako je $\dfrac{\alpha_p^k \delta^k}{\lambda^k} < \varepsilon$, *staviti* $s_i^{k+1} = s_i^k$, $i = 1,\dots,n$ *i staviti* $\delta^{k+1} = \delta^k$.

Zameniti k *sa* $k+1$ *i ići na korak 2.*

Lema 2.4. $\delta^k = \det[s_1^k,\dots,s_n^k]$, *gde je* $[s_1^k,\dots,s_n^k]$ *matrica čije su kolone vektori* $s_1^k,\dots,s_n^k$.

Dokaz. *Videti u /62/, str. 294.*

Posledica 2.4. $\mathrm{Det}[s_1^k,\dots,s_n^k] \geq \varepsilon$ *za svako* k.

Dokaz. *Korak 3. obezbedjuje da važi* $\delta^k \geq \varepsilon$ *za svako* k.

Teorema 2.5. *Neka je $\phi : R^n \to R$ strogo konveksna neprekidno diferencijabilna funkcija. Tada je 1. svaka tačka nagomilavanja x^* niza $\{x_i^k\}$ generisanog Powell-Zangwill-ovim algoritmom takodje tačka nagomilavanja niza $\{ x_j^k \}$;*

2. x^* je jedinstveno optimalno rešenje problema (2.13).

Dokaz. *Videti u Zangwill /62/, str. 295.*

Većina od dosad poznatih metoda za uslovnu optimizaciju, koje ne koriste izvode, u praksi je dala slabe rezultate. Polazilo se od toga da je dovoljno da se u poznate metode za bezuslovnu optimizaciju uključe jednostavne procedure koje bi vodile računa o ograničenjima. Na primer, u tački u kojoj neko ograničenje nije zadovoljeno možemo staviti da je funkcija cilja jednaka nekoj velikoj konstantnoj vrednosti i na taj način izvršiti odbacivanje nedopustivih tačaka. Rezultati Friedman-a i Pinder-a /21/ pokazuju da ovakav pristup primenjen u praksi ima ozbiljne nedostatke.

Zatim, iteracije su se definisale i ovako: u dopustivoj oblasti se krećemo kao u bezuslovnom slučaju, a kad naidjemo na ograničenje definišemo kretanje tako da ono prati granicu dopustive oblasti. Ovakav pristup imaju metode Singer-a /56/ i Mugele-a /39/. Ipak, iako je pokazano da su ovakve metode efektivne kad se primene na odredjene specifične probleme, u opštem slučaju nisu pogodne.

Najefektivnojom se u praksi pokazala metoda Davies-a i Swann-a /16/, gde se vektor pravca definiše tako da pripada preseku aktivnih (linearnih) ograničenja. Ovde se pošlo od sledeće metode za bezuslovnu optimizaciju.

Na k-toj interaciji data je tačka x_o^k i ortogonalni pravci p_i^k, $i = 1,\ldots,n$. Nalazimo sukcesivno tačke minimuma x_i^k, $i = 1,\ldots,n$ funkcije cilja ϕ duž pravaca p_i^k, $i = 1,\ldots,n$. Formirajmo vektore d_i, $i = 1,\ldots,n$ na sledeći način: $d_i = x_n^k - x_{i-1}^k$. Vektore pravaca za sledeću iteraciju definišimo na sledeći način:

$$p_1^{k+1} = d_1, \quad p_i^{k+1} = d_i - N_{i-1}\alpha_i, \quad i = 2,3,\dots,n,$$

gde je $N_{i-1} = \left[p_1^{k+1}, p_2^{k+1}, \dots, p_{i-1}^{k+1} \right]$, *a* $\alpha_i = (N_{i-1}^T N_{i-1})^{-1} N_{i-1}^T d_i$, *čime je obezbedjena ortogonalnost vektora* p_i^{k+1}, $i = 1,\dots,n$. *Dakle,*

$$(2.14) \quad p_i^{k+1} = \left[I - N_{i-1}(N_{i-1}^T N_{i-1})^{-1} N_{i-1}^T \right] d_i = P_{i-1} d_i,$$

gde je P_{i-1} *projekciona matrica koja projektuje* d_i *u potprostor ortogonalan na već izabrane pravce.*

Ako su zadata ograničenja u obliku jednakosti

$$(2.15) \quad A^T x = b, \quad A - n \times m \ \text{matrica}, \ m < n,$$

problem je u stvari od n-m dimenzija (pod pretpostavkom da je rang A = m) i samo toliki broj vektora pravaca će nam i biti potreban. Za vektore pravaca $p_1^{k+1}, \dots, p_m^{k+1}$ *uzimamo vektore normala ograničenja* a_i, $i = 1,\dots,m$, *pa je* $N_m = [a_1,\dots,a_m]$, *a ostali vektori se mogu naći korišćenjem relacija (2.14). Ako tačka* x^{k+1} *zadovoljava uslov (2.15) i, ako se krećemo duž vektora* $p_{m+1}^{k+1}, \dots, p_n^{k+1}$, *tada će sve tačke* x_i^{k+1}, $i = 1,\dots,n$, *biti dopustive.*

Proširenje na slučaj linearnih ograničenja u obliku nejednakosti je prirodno, jer svako ograničenje koje postane aktivno u toku jedne iteracije tretiramo kao ograničenje u obliku jednakosti, a to smo upravo opisali. Vektori pravaca u opštem slučaju više nisu ortogonalni, ali, ako su ograničenja linearno nezavisna i vektori pravaca biće linerano nezavisni i razapinjaće ceo parametarski prostor.

3. POGLAVLJE

Dat je problem bezuslovne optimizacije

$$(3.1) \qquad \min \{\phi(x) \mid x \in D\},$$

gde je $\phi: D \subset R^n \to R$ *neprekidno diferencijabilna funkcija na otvorenom skupu* D.

Razmatramo iterativne algoritme za nalaženje optimalnog rešenja problema (3.1) koji generišu nizove tačaka $\{x_k\}$ *oblika:*

$$(3.2) \qquad x_{k+1} = x_k - \alpha_k p_k, \quad k = 0,1,2,\ldots,$$

pri čemu pretpostavljamo da vektor pravca p_k *zadovoljava sledeće:*

$$(3.3) \qquad p_k \neq 0, \quad \langle \nabla\phi(x_k), p_k \rangle \geq 0.$$

Korak $\alpha_k \geq 0$ *izračunavamo pomoću nekog od algoritama koje ćemo sada izložiti.*

3.1. Modifikovani Curry-Altman-ov algoritam /41/, str. 251.

Za niz $\{x_k\}$ *koji zadovoljava (3.2) i (3.3) definišemo* α_k *na sledeći način:*

$\alpha_k = 0$, *ako je* $\langle \nabla\phi(x_k), p_k \rangle = 0$; *inače*

$\alpha_k \in I_k$, *gde* I_k *predstavlja prvi interval pozitivnih rešenja nejednačine*

(3.4) $\quad \langle \nabla\phi(x_k - \alpha p_k), p_k \rangle \quad \sigma\,(\langle \nabla\phi(x_k), p_k \rangle),$

pri čemu je $\sigma: [o,+\infty) \to [o,+\infty)$ *prinudna funkcija (definicija 1.9) takva da zadovoljava uslov* $\sigma(t) \leq \delta t$, $0 < \delta < 1$ *za svako* $t \geq 0$.

Sledeće leme pokazuju da je ovaj algoritam dobro definisan, tj. da pod odredjenim uslovima interval I_k *postoji.*

Lema 3.1. *Neka je* $\phi: D \subset R^n \to R$ *neprekidno diferencijabilna funkcija na* D *i pretpostavimo da* $x_k, \alpha_k > 0$ *i* p_k *zadovoljavaju uslov* $\langle \nabla\phi(x_k), p_k \rangle > 0$, $[x_k, x_k - \alpha_k p_k] \subset D$ *i*

(3.5) $\quad \langle \nabla\phi(x_k - \alpha_k p_k), p_k \rangle \leq \sigma\,(\langle \nabla\phi(x_k), p_k \rangle),$

gde je $\sigma: [o,+\infty) \to [o,+\infty)$ *prinudna funkcija takva da je* $\sigma(t) \leq \delta t$, $0 < \delta < 1$ *za svako* $t \geq 0$. *Tada postoji* $\bar{\alpha}$, $0 < \bar{\alpha} < \alpha_k$ *koje zadovoljava nejednačinu (3.4).*

Dokaz. *Definišimo funkciju* $F: [0,1] \to R$ *na sledeći način:*

$$F(\lambda) = \frac{\langle \nabla\phi(x_k - \lambda\alpha_k p_k), p_k \rangle}{\langle \nabla\phi(x_k), p_k \rangle}\,.$$

Vidimo da je $F(o) = 1$, *i, zbog (3.5)*

$$F(1) = \frac{\langle \nabla\phi(x_k - \alpha_k p_k), p_k \rangle}{\langle \nabla\phi(x_k), p_k \rangle} \leq \frac{\sigma\,(\langle \nabla\phi(x_k), p_k \rangle)}{\langle \nabla\phi(x_k), p_k \rangle} \leq \delta < 1.$$

Iz pretpostavki leme sledi da je F *neprekidna na* $[0,1]$. *Prema tome,* $F(\lambda)$ *poprima sve vrednosti izmedju* $\dfrac{\sigma(\langle \nabla\phi(x_k), p_k \rangle)}{\langle \nabla\phi(x_k), p_k \rangle}$ *i 1. Dakle, postoji takvo* $\bar{\lambda} \in (0,1)$ *da važi*

$$\frac{\sigma(\langle \nabla\phi(x_k), p_k \rangle)}{\langle \nabla\phi(x_k), p_k \rangle} < F(\bar{\lambda}) < 1,$$

odnosno, za $\bar{\alpha} = \bar{\lambda}\alpha_k$ *biće zadovoljena nejednačina (3.4).*

Lema 3.2. *Neka je* $\phi: D \subset R^n \to R$ *neprekidno diferencijabilna funkcija na otvorenom skupu D i neka je lučno povezana komponenta* L^o *nivoskog skupa* $L = \{x \ D \mid \phi(x) \leq \phi(x_o)\}$ *kompaktna za neko* $x_o \in D$*. Neka niz* $\{x_k\}$ *zadovoljava relacije (3.2), (3.3) i (3.4). Tada važi sledeće:*

1. svaki korak α_k *koji je zadat modifikovanim Curry-Altman-ovim algoritmom je oblika*

$$\alpha_k = q_k \bar{\alpha}_k, \qquad 0 \leq q_k < 1,$$

gde je $\bar{\alpha}_k$ *- najmanje pozitivno rešenje poopštene Curry-Altman-ove jednačine*

$$(3.6) \qquad <\nabla\phi(x_k - \alpha p_k), p_k> = \sigma(<\nabla\phi(x_k), p_k>);$$

2. $x_k \in L^o$;

3. $\phi(x_k) \geq \phi\left[\lambda x_k + (1-\lambda)x_{k+1}\right] \geq \phi(x_{k+1}), \quad \forall \lambda \in [0,1]$.

Dokaz *ćemo izvesti metodom matematičke indukcije po k. Pretpostavimo da* $x_k \in L^o$ *i da je* $<\nabla\phi(x_k), p_k> > 0$*. Tada je* $\sigma(<\nabla\phi(x_k), p_k>) > 0$ *na osnovu definicije 1.9. Iz leme 1.11. sledi da postoji takvo* $\beta_k > 0$ *da* $[x_k, x_k - \beta_k p_k] \subset L^o$ *i* $\phi(x_k) = \phi(x_k - \beta_k p_k)$.

Prema tome, na osnovu Lagrange-ove teoreme, postoji $\bar{\beta} \in (0, \beta_k)$ *tako da važi*

$$0 = \phi(x_k) - \phi(x_k - \beta_k p_k) = \beta_k <\nabla\phi(x_k - \bar{\beta} p_k), p_k>, \text{ tj.}$$

$$<\nabla\phi(x_k - \bar{\beta} p_k), p_k> = 0.$$

Označimo sa $\nu(\alpha)$ funkciju $<\nabla\phi)x_k-\alpha p_k),p_k>$. Pri tom je $\nu(o) = <\nabla\phi(x_k),p_k> > 0$ i $\nu(\bar{\beta}) = 0$. Funkcija $\nu(\alpha)$, kao neprekidna poprima sve vrednosti izmedju 0 i $<\nabla\phi(x_k),p_k>$. Sledi da postoji $\bar{\alpha}_k \in (0,\bar{\beta})$ takvo da važi

$$\nu(\bar{\alpha}_k) = \sigma[\nu(o)] \leq \delta\nu(0), \qquad 0<\delta<1,$$

odnosno, da jednačina (3.6) mora imati bar jedno rešenje na intervalu $(0,\bar{\beta})$. Pošto je leva strana jednačine (3.6) neprekidna funkcija, postojaće i najmanje rešenje ove jednačine.

Označimo sa $F(\alpha)$ sledeću funkciju:

$$F(\alpha) = <\nabla\phi(x_k-\alpha p_k),p_k> - \sigma(<\nabla\phi(x_k),p_k>).$$

Vidimo da je

$$F(0) = <\nabla\phi(x_k),p_k> - \sigma(<\nabla\phi(x_k),p_k>) \geqslant <\nabla\phi(x_k),p_k> - \delta<\nabla(x_k),p_k> =$$

$$=(1-\delta) <\nabla\phi(x_k),p_k> \quad 0.$$

Prema tome, $F(\alpha) > 0$ za svako $\alpha \in [0,\bar{\alpha}_k)$, tj.

$$(3.7) \quad <\nabla\phi(x_k-\alpha p_k),p_k> > \sigma(<\nabla\phi(x_k),p_k>) \quad za \ \forall\alpha \in [0,\bar{\alpha}_k).$$

Dakle, bilo koje $\alpha \in [0,\bar{\alpha}_k)$ predstavlja korak definisan modifikovanim Curry-Altman-ovim algoritmom. Odavde sledi da je $\alpha_k = q_k\,\bar{\alpha}_k$, $0 \leq q_k < 1$, čime smo dokazali tvrdjenje 1. Odavde dalje sledi da je $x_{k+1} = x_k - \alpha_k p_k$ dobro definisano i da, pošto $[x_k,x_k-\beta_k p_k] \subset L^O$, a $\bar{\alpha}_k \in (o,\beta_k)$ i $[x_k,x_{k+1}] \subset L^O$. Dakle, $x_{k+1} \in L^O$, čime je dokazano tvrdjenje 2.

Dalje, kako izvod funkcije $\Phi(\alpha) = \phi\,(x_k-\alpha p_k)$ glasi $\Phi'(\alpha) = -<\nabla\phi(x_k - \alpha p_k),p_k>$, iz (3.6) i (3.7) sledi da je $\Phi'(\alpha) < 0$ za $\alpha \in [0,\bar{\alpha}_k]$, tj. da funkcija $\Phi(\alpha)$ monotono opada na intervalu $[0,\bar{\alpha}_k]$. Odavde sledi da važi

$$(3.8) \qquad \phi(x_k) \geqslant \phi[\lambda x_k + (1-\lambda)x_{k+1}] \geqslant \phi(x_{k+1}) \quad za \ \lambda \in [0,1],$$

čime smo dokazali i tvrdjenje 3.

Teorema 3.3. *Neka su zadovoljene pretpostavke iz leme 3.2. Tada, ako je*

(3.9) $\alpha_k \geq q\bar{\alpha}_k, \quad 0 < q < 1$ *i*

(3.10) $\| p_k \| \leq M$ *za svako* k

važi sledeće:

1. $\phi(x_k) - \phi(x_{k+1}) \geq \mu(<\nabla\phi(x_k), p_k>)$,

   gde je $\mu: [0,+\infty) \to [0,+\infty)$ *prinudna funkcija;*

2. $\lim_{k\to\infty} <\nabla\phi(x_k), p_k> = 0.$

Dokaz. *Na osnovu Lagrange-ove teoreme iz (3.7) i (3.9) će slediti*

(3.10') $\phi(x_k) - \phi(x_{k+1}) = \alpha_k <\nabla\phi(x_k - \hat{\alpha}p_k), p_k>$ *za neko* $\hat{\alpha} \in (0,\alpha_k)$

$$> \alpha_k \ \sigma(<\nabla\phi(x_k), p_k>),$$

(3.11) $\phi(x_k) - \phi(x_{k+1}) > q\,\bar{\alpha}_k \ \sigma(<\nabla\phi(x_k), p_k>).$

Ako je $\nabla\phi(x) \equiv 0$ *za* $x \in L^0$, *tvrdjenja teoreme su trivijalna. Zato pretpostavimo da je* $\nabla\phi(x) \not\equiv 0$ *za* $x \in L^0$. *Tada iz pretpostavki teoreme, na osnovu leme 1.9 sledi da je obratni modul neprekidnosti od* $\nabla\phi$ *na* L^0 *prinudna funkcija.*

Dalje, kako je $\bar{\alpha}_k$ *rešenje jednačine (3.6), važiće sledeća jednakost:*

$$<\nabla\phi(x_k),p_k> - \sigma(<\nabla\phi(x_k),p_k>) = <\nabla\phi(x_k),p_k> - <\nabla\phi(x_k-\bar{\alpha}_k p_k),p_k> .$$

Odavde, pošto je $\sigma(t) \leq \delta t$, $0<\delta<1$ *dalje sledi*

$$(3.12) \qquad <\nabla\phi(x_k),p_k> - \delta<\nabla\phi(x_k),p_k> \leq <\nabla\phi(x_k),p_k> - <\nabla\phi(x_k-\bar{\alpha}_k p_k),p_k>,$$

odnosno, korišćenjem Cauchy-Swartz-ove nejednakosti i uslova (3.10) iz (3.12) sledi:

$$(3.13) \qquad (1-\delta)<\nabla\phi(x_k),p_k> \leq \|\nabla\phi(x_k)-\nabla\phi(x_k-\bar{\alpha}_k p_k)\|\,\|p_k\| \leq M\|\nabla\phi(x_k)-\nabla\phi(x_k-\bar{\alpha}_k p_k)\| .$$

(Iz (3.13), na osnovu definicije 1.11. sledi

$$\bar{\alpha}_k\|p_k\| \geq s\left[\frac{1-\delta}{M}<\nabla\phi(x_k),p_k>\right], \text{ odnosno,}$$

$$(3.14) \qquad \bar{\alpha}_k \geq \frac{1}{\|p_k\|}\, s\left[\frac{1-\delta}{M}<\nabla\phi(x_k),p_k>\right] \geq \frac{1}{M}\, s\left[\frac{1-\delta}{M}<\nabla\phi(x_k),p_k>\right];$$

gde je sa s označen obratni moduo neprekidnosti od $\nabla\phi$.

Iz relacija (3.11) i (3.14) dobijamo sledeće:

$$\phi(x_k) - \phi(x_{k+1}) > \frac{q}{M}\, s\left[\frac{1-\delta}{M}<\nabla\phi(x_k),p_k>\right]\sigma(<\nabla\phi(x_k),p_k>), \text{ tj.}$$

$$(3.15) \qquad \phi(x_k) - \phi(x_{k+1}) > \mu(<\nabla\phi(x_k),p_k>)$$

gde je $\mu: [0,+\infty) \to [0,+\infty)$ *prinudna funkcija oblika*

$$\mu(t) = \frac{q}{M} s\left(\frac{1-\delta}{M} t\right) \sigma(t).$$

Dalje, kako je zbog pretpostavki teoreme funkcija ϕ *ograničena odozdo na* L^O*, a iz tvrdjenja 3 leme 3.2. sledi da je* $\phi(x_k) \geq \phi(x_{k+1})$*, važiće sledeće:*

$$(3.16) \qquad \lim_{k\to\infty}\left[\phi(x_k) - \phi(x_{k+1})\right] = 0.$$

Iz (3.15) sledi, na osnovu definicije 1.9. prinudne funkcije da je $\lim_{k\to\infty} \langle\nabla\phi(x_k),p_k\rangle = 0.$

Posledica 3.3. *Neka niz* $\{x_k\}$ *zadovoljava relacije (3.2), (3.3), (3.4) i uslov (3.9), s tim što je u relaciji (3.4) funkcija* $\sigma(t)$ *oblika* $\sigma(t) = \delta t$, $0<\delta<1$. *Neka su zadovoljene pretpostavke leme 3.2. Tada niz* $\{x_k\}$ *ima sledeće osobine:*

1. $x_k \in L^0$;
2. $\phi(x_k) - \phi(x_{k+1}) \geq \mu(\langle\nabla\phi(x_k), \frac{p_k}{\|p_k\|}\rangle)$,

   gde je $\mu: [0,+\infty) \to [0,+\infty)$ *prinudna funkcija;*
3. $\lim_{k\to\infty} \frac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|} = 0.$

(Vidi /41/,str.483)

Dokaz. *Iz leme 3.2 sledi da* $x_k \in L^0$. *Iz relacije (3.11) sledi, za* $\sigma(t) = \delta t$:

$$\phi(x_k)-\phi(x_{k+1}) \geq q\delta\bar{\alpha}_k \langle\nabla\phi(x_k),p_k\rangle. \tag{3.17}$$

Deobom prve nejednakosti u (3.13) sa $\|p_k\|$ *dobijamo:*

$$(1-\delta)\frac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|} \leq \|\nabla\phi(x_k)-\nabla\phi(x_k-\bar{\alpha}_k p_k)\|. \tag{3.18}$$

Iz (3.18) na osnovu definicije 1.11. funkcije s sledi

$$\bar{\alpha}_k \geq \frac{1}{\|p_k\|}\, s\left[(1-\delta)\langle\nabla\phi(x_k), \frac{p_k}{\|p_k\|}\rangle\right]. \tag{3.19}$$

Iz (3.17) i (3.19) dobijamo sledeću relaciju:

$$(3.20)\quad \phi(x_k)-\phi(x_{k+1}) > q\delta\,\frac{\langle\nabla\phi(s_k),p_k\rangle}{\|p_k\|}\; s\left[(1-\delta)\langle\nabla\phi(x_k),\frac{p_k}{\|p_k\|}\rangle\right] =$$

$$= \mu\left(\frac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|}\right),$$

gde je $\mu: [0,+\infty) \to [0,+\infty)$ prinudna funkcija oblika $\mu(t) = q\delta t s\,[(1-\delta)t]$.

Iz ograničenosti ϕ na L^o i monotonosti $\{\phi(x_k)\}$ sledi da $\phi(x_k)-\phi(x_{k+1})\to 0$, $k\to\infty$.

Na osnovu toga iz (3.20) sledi da $\dfrac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|} \to 0, k\to\infty$.

Teorema 3.4. *Neka je $\phi: D\subset R^n\to R$ dva puta neprekidno diferencijabilna funkcija na otvorenom skupu D koja zadovoljava uslov:*

$$(3.21)\quad m_o\|y\|^2 \le \langle H(x)y,y\rangle \quad za\ \forall x\in D,\ \forall y\in R^n,\quad m_o > 0,$$

gde je $H(x)$ hesijan funkcije ϕ u tački x. Neka niz $\{x_k\}$ zadovoljava relacije (3.2), (3.3) i (3.4) i neka je funkcija $\sigma(t)$ iz relacije (3.4) oblika $\sigma(t) = \delta t$, $0<\delta<1$. Neka je zadovoljen uslov (3.9). Tada je

$$\lim_{k\to\infty}\|x_{k+1}-x_k\| = 0.$$

(Videti /41/, str. 485.)

Dokaz. Iz (3.21) sledi, na osnovu leme 1.10, da je nivoski skup $L = \{x\in D \mid \phi(x)\le\phi(x_o)\}$ konveksan (dakle i lučno povezan) i ograničen za neko $x_o\in D$ i da, prema tome, postoji M_o, $m_o< M_o<\infty$ tako da važi

$$(3.22)\quad m_o\|y\|^2 \le \langle H(x)y,y\rangle \le M_o\|y\|^2 \quad za\ \forall x\in L,\ \forall y\in R^n.$$

Ako sa $\bar{\alpha}_k$ označimo najmanje pozitivno rešenje jednačine (3.6) (kada je $\sigma(t) = \delta t$, $0<\delta<1$), tada, na osnovu Taylor-ove formule i korišćenjem relacije (3.22) dobijamo:

$$\langle \nabla\phi(x_k), p_k\rangle - \delta\langle \nabla\phi(x_k), p_k\rangle = \langle \nabla\phi(x_k), p_k\rangle - \langle \nabla\phi(x_k - \bar{\alpha}_k p_k), p_k\rangle =$$

$$= \bar{\alpha}_k \int_0^1 \langle H(x_k - \Theta\bar{\alpha}_k p_k) p_k, p_k\rangle \, d\Theta \leq \bar{\alpha}_k M_o \| p_k \|^2 .$$

Odavde dalje sledi

$$(3.23) \quad (1-\delta) \left\langle \nabla\phi(x_k), \frac{p_k}{\| p_k \|} \right\rangle \leq \bar{\alpha}_k M_o \| p_k \| .$$

Analogno, opet korišćenjem (3.6) i (3.22) zaključujemo da važi

$$(3.24) \quad (1-\delta) \left\langle \nabla\phi(x_k), \frac{p_k}{\| p_k \|} \right\rangle \geq \bar{\alpha}_k m_o \| p_k \| .$$

Iz (3.23) i (3.24) sada sledi

$$\bar{\alpha}_k M_o \| p_k \| \geq (1-\delta) \left\langle \nabla\phi(x_k), \frac{p_k}{\| p_k \|} \right\rangle \geq \bar{\alpha}_k m_o \| p_k \| .$$

Odavde, pošto je $\alpha_k < \bar{\alpha}_k$, *dalje sledi*

$$(3.24') \quad (1-\delta) \left\langle \nabla\phi(x_k), \frac{p_k}{\| p_k \|} \right\rangle > \alpha_k m_o \| p_k \| = m_o \| x_{k+1} - x_k \| .$$

Pošto $\left\langle \nabla\phi(x_k), \frac{p_k}{\| p_k \|} \right\rangle \to 0$, $k \to \infty$ *na osnovu tvrdjenja 2 posledice 3.3., a* $1-\delta > 0$, *iz (3.24') sledi tvrdjenje teoreme.*

3.2. Poopšteni algoritam Ostrowskog

Neka je $\phi : D \subset R^n \to R$ *neprekidno diferencijabilna funkcija na otvorenom skupu D i neka važi sledeći uslov*

$$(3.25) \quad \| \nabla\phi(x) - \nabla\phi(y) \| \leq \gamma \| x - y \| \qquad \text{za } \forall x, y \in L^o, \quad \gamma > 0,$$

gde je sa L^o *označen lučno povezan deo nivoskog skupa* $L = \{x \in D \mid \phi(x) \leq \phi(x_o)\}$ *za neko* $x_o \in D$.

Razmatramo niz $\{x_k\}$ *sa svojstvima (3.2) i (3.3), pri čemu je* α_k *zadato na sledeći način*

$$(3.26)\quad \alpha_k = \begin{cases} 0, & \text{ako je } \langle\nabla\phi(x_k),p_k\rangle = 0; \\ \dfrac{1}{\gamma\|p_k\|}\, d\left(\dfrac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|}\right), & \text{ako je } \langle\nabla\phi(x_k),p_k\rangle > 0, \end{cases}$$

gde je $d: [0,\infty) \to [0,\infty)$ *prinudna funkcija takva da je* $d(t) \leq \delta t$, $0<\delta<2$ *za* $\forall t \geq 0$.

Teorema 3.5. *Neka funkcija* ϕ *zadovoljava gornje pretpostavke i neka je skup* L^0 *kompaktan. Tada za niz* $\{x_k\}$ *koji zadovoljava relacije (3.2), (3.3) i (3.26) važe sledeća tvrdjenja:*

1. $x_k \in L^0$;
2. $\phi(x_k) \geq \phi(x_{k+1})$;
3. $\lim\limits_{k\to\infty} \dfrac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|} = 0$;
4. $\lim\limits_{k\to\infty} (x_k - x_{k+1}) = 0$.

Dokaz. *Dokaz ćemo sprovesti primenom metode matematičke indukcije u odnosu na k. Pretpostavimo da* $x_k \in L^0$ *i da je* $\langle\nabla\phi(x_k),p_k\rangle > 0$. *Tada iz leme 1.13. sledi da za svako* $\alpha > 0$ *takvo da* $[x_k, x_k - \alpha p_k] \subset D$ *važi sledeće:*

$$(3.27)\quad \phi(x_k) - \phi(x_k - \alpha p_k) \geq \alpha\langle\nabla\phi(x_k),p_k\rangle - \frac{1}{2}\gamma\alpha^2\|p_k\|^2.$$

Specijalno za $\forall\, x_k - \alpha p_k \in (x_k, x_k - \alpha_k p_k] \cap L^0$ *važiće*

$$\phi(x_k)-\phi(x_k-\alpha p_k) \geq \alpha(\langle\nabla\phi(x_k),p_k\rangle - \frac{1}{2}\gamma\alpha\|p_k\|^2) \geq$$

$$\alpha\left[\frac{\|p_k\|}{\delta}\, d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right) - \frac{1}{2}\|p_k\|\, d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right)\right] = \qquad (jer\ je\ t \geqslant \frac{d(t)}{\delta})$$

$$= \alpha\,\|p_k\|\; d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right)\left(\frac{1}{\delta} - \frac{1}{2}\right),\ tj.$$

$$(3.27') \qquad \phi(x_k)-\phi(x_k - \alpha p_k) \geqslant \alpha\,\|p_k\|\; d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right)\frac{2-\delta}{2\delta} > 0 \quad jer\ je.$$

konstanta $\frac{2-\delta}{2\delta} > 0$ *zbog* $0<\delta<2$.

Iz (3.27') sledi, na osnovu leme 1.11. da $[x_k,x_{k+1}] \subset L^0$, *a odatle da* $x_{k+1} \in L^0$.

Dalje, zamenom izraza za α_k *iz (3.26) u (3.27) dobijamo*

$$\phi(x_k)-\phi(x_k - \alpha_k p_k) \geqslant \frac{1}{\gamma\|p_k\|}\, d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right) <\nabla\phi(x_k),p_k> -$$

$$- \frac{1}{2\gamma}\left[d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right)\right]^2 \geqslant \frac{2-\delta}{2\gamma\delta}\left[d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right)\right]^2,\quad tj.$$

$$(3.28) \qquad \phi(x_k) - \phi(x_{k+1}) \geqslant \frac{2-\delta}{2\gamma\delta}\,\mu\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right),$$

gde je $\mu(t) = [d(t)]^2$ *prinudna funkcija. Iz (3.28) sledi da je* $\phi(x_k) > \phi(x_{k+1})$. *Dalje, iz monotonosti i ograničenosti* ϕ *odozdo na* L^0 *sledi da* $\phi(x_k) - \phi(x_{k+1}) \to 0$, $k\to\infty$. *Iz (3.28) sada na osnovu definicije prinudne funkcije, sledi da* $\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|} \to 0$, $k\to\infty$.

Kako je

$$(3.29) \qquad \|x_{k+1} - x_k\| = \alpha_k\|p_k\| = \frac{1}{\gamma} d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right),$$

iz (3.29) i činjenice da $d\left(\frac{<\nabla\phi(x_k),p_k>}{\|p_k\|}\right) \to 0$, $k\to\infty$ *sledi tvrdjenje 4.*

3.3. Poopštenje Céa-Goldstein-ovog algoritma

Razmatramo niz tačaka $\{x_k\}$ *sa svojstvima (3.2) i (3.3) pri čemu se korak* α_k *definiše na sledeći način:*

$$\alpha_k = 0 \quad \text{ako je } \langle\nabla\phi(x_k), p_k\rangle = 0;$$

inače se za α_k *uzima rešenje sledećih nejednačina:*

$$\phi(x_k) - \phi(x_k - \alpha\, p_k) \geq \alpha\, \sigma\, (\langle\nabla\phi(x_k), p_k\rangle) \tag{3.30}$$

$$\phi(x_k) - \phi(x_k - 2\alpha\, p_k) < 2\alpha\, \sigma(\langle\nabla\phi(x_k), p_k\rangle), \tag{3.31}$$

gde je $\sigma\colon [0,+\infty) \to [0,+\infty)$ *prinudna funkcija takva da je* $\sigma(t) \leq \delta t$, $0<\delta<1$.

U sledećoj lemi ćemo pokazati da je ovaj algoritam dobro definisan, tj. da pod odredjenim uslovima postoji $\alpha_k > 0$ *koje zadovoljava (3.30) i (3.31).*

Lema 3.6. *Pretpostavimo da je* $\phi\colon R^n \to R$ *neprekidno diferencijabilna funkcija ograničena odozdo i pretpostavimo da* x_k, $\alpha_k > 0$ *i* p_k *zadovoljavaju uslov* $\langle\nabla\phi(x_k), p_k\rangle > 0$ *i*

$$\phi(x_k) - \phi(x_k - \alpha_k p_k) < \alpha_k\, \sigma\,(\langle\nabla\phi(x_k), p_k\rangle), \tag{3.32}$$

gde je $\sigma\colon [0,+\infty) \to [0,+\infty)$ *prinudna funkcija takva da je* $\sigma(t) \leq \delta t$, $0<\delta<1$. *Tada postoji* $0<\bar{\alpha}_k<\alpha_k$ *takvo da su zadovoljene nejednačine (3.30) i (3.31).*

Dokaz. *Definišimo funkciju* $F\colon [0,1] \to R$ *na sledeći način:*

$$F(\lambda) = \begin{cases} \dfrac{\phi(x_k) - \phi(x_k - \lambda\alpha_k p_k)}{\lambda\alpha_k <\nabla\phi(x_k),p_k>} & , \quad \lambda \in (0,1] \; ; \\ 1 & , \quad \lambda = 0. \end{cases}$$

Iz L'Hospital-ovog pravila sledi da kada $\lambda \to 0$, $F(\lambda)\to 1$, *pa je* F *zbog pretpostavki leme neprekidna na* $[0,1]$.

Prema tome, pošto je $F(0) = 1$ *i*

$$F(1) = \frac{\phi(x_k) - \phi(x_k - \alpha_k p_k)}{\alpha_k <\nabla\phi(x_k),p_k>} < \frac{\sigma(<\nabla\phi(x_k),p_k>)}{<\nabla\phi(x_k),p_k>} \leq \delta < 1,$$

a F, *zbog neprekidnosti poprima sve vrednosti izmedju* $\dfrac{\sigma(<\nabla\phi(x_k),p_k>)}{<\nabla\phi(x_k),p_k>}$ *i* 1, *postojaće takvo* $\bar{\lambda}\in (0,1)$ *da važi* $\dfrac{\sigma(<\nabla\phi(x_k),p_k>)}{<\nabla\phi(x_k),p_k>} \leq F(\bar{\lambda}) < 1$,

odnosno, za $\bar{\alpha}_k = \bar{\lambda}\alpha_k$ *biće zadovoljena relacija (3.30).*

Sada pretpostavimo da neko $\alpha > 0$ *zadovoljava nejednačinu (3.30), ali da ne zadovoljava (3.31), tj. da imamo ovakvu situaciju:*

$$\phi(x_k) - \phi(x_k - \alpha p_k) \geq \alpha\, \sigma(<\nabla\phi(x_k),p_k>)$$

$$\phi(x_k) - \phi(x_k - 2\alpha p_k) \geq 2\alpha\, \sigma(<\nabla\phi(x_k),p_k>),$$

odnosno, ako ne postoji i *takvo da* $2^i\alpha$ *zadovoljava (3.31), imaćemo*

$$(3.33) \quad \phi(x_k) - \phi(x_k - 2^i\alpha p_k) \geq 2^i\alpha\sigma(<\nabla\phi(x_k),p_k>), \quad i = 0,1,2,\ldots$$

Kada $i\to\infty$, *desna strana nejednakosti (3.33) teži* ∞, *odnosno,* $\phi(x_k - 2^i\alpha p_k) - \phi(x_k)\to -\infty$, $i\to\infty$, *a to je suprotno sa pretpostavkom da je* ϕ *ograničena odozdo.*

Prema tome, za neko i *će važiti:*

$$\phi(x_k) - \phi(x_k - 2^{i-1}\alpha p_k) \geq 2^{i-1}\alpha\, \sigma(<\nabla\phi(x_k),p_k>)$$

$$\phi(x_k) - \phi(x_k - 2^i\alpha p_k) < 2^i\alpha\sigma(<\nabla\phi(x_k),p_k>).$$

Teorema 3.7. *Neka je $\phi: R^n \to R$ neprekidno diferencijabilna funkcija i neka je $L = \{x \in R^n | \phi(x) \leq \phi(x_0)\}$ konveksan i ograničen skup. Neka je niz $\{x_k\}$ definisan sa (3.2), (3.3), (3.30) i (3.31). Tada, ako je $\|p_k\| \leq M$ za $\forall k$, važi sledeće:*

1. $x_k \in L$;
2. $\phi(x_k) \geq \phi(x_{k+1})$;
3. $\phi(x_k) - \phi(x_{k+1}) \geq \mu(\langle\nabla\phi(x_k), p_k\rangle)$, gde je μ prinudna funkcija;
4. $\langle\nabla\phi(x_k), p_k\rangle \to 0, \quad k \to \infty$.

Dokaz. *Dokaz ćemo izvesti primenom metode matematičke indukcije u odnosu na k. Pretpostavimo da x_k pripada L i da je $\langle\nabla\phi(x_k), p_k\rangle > 0$. Iz leme 3.6. sledi da je $x_{k+1} = x_k - \alpha_k p_k$ dobro definisano. Pošto iz (3.30) sledi da je $\phi(x_k) > \phi(x_{k+1})$, sledi da x_{k+1} pripada L.*

Iz pretpostavki teoreme sledi da je moduo neprekidnosti ω od $\nabla\phi$ na L, na osnovu leme 1.8 ravnomerno neprekidna monotono rastuća funkcija na $[0, +\infty)$, $\omega(0) = 0$, dakle, jedna prinudna funkcija.

Ako pretpostavimo da $\nabla\phi$ nije konstanta na L, funkcija $\eta(t) = \int_0^1 \omega(ts)ds$ je dobro definisana i strogo rastuća. Prema tome, η je prinudna funkcija.

Iz leme 1.13. sada sledi

$$(3.34) \quad \phi(x_k) - \phi(x_k - 2\alpha_k p_k) \geq 2\alpha_k \langle\nabla\phi(x_k), p_k\rangle - 2\alpha_k \|p_k\| \eta(2\alpha_k \|p_k\|).$$

Iz (3.34) korišćenjem (3.31) sledi

$$2\alpha_k \sigma(\langle\nabla\phi(x_k), p_k\rangle) \geq 2\alpha_k \langle\nabla\phi(x_k), p_k\rangle - 2\alpha_k \|p_k\| \eta(2\alpha_k \|p_k\|).$$

Dalje, kako je $\sigma(t) \leq \delta t$, $0<\delta<1$ za svako $t \geq 0$, odavde sledi

$$(3.35) \quad \delta\langle\nabla\phi(x_k), p_k\rangle \geq \langle\nabla\phi(x_k), p_k\rangle - \|p_k\| \eta(2\alpha_k \|p_k\|).$$

Iz (3.35) sledi, zbog $\|p_k\| \leq M$ za svako k

$$(3.36)\qquad \eta(2\alpha_k \|p_k\|) \geq \frac{1-\delta}{\|p_k\|} <\nabla\phi(x_k), p_k> \geq \frac{1-\delta}{M} <\nabla\phi(x_k), p_k>.$$

Iz (3.36) sledi da će za svaku funkciju $d: [0,+\infty) \to [0,+\infty)$ *koja je strogo rastuća i* $d(t) \geq \eta(t)$ *za* $\forall t \geq 0$ *(dakle, prinudna funkcija), važiti:*

$$2\alpha_k \|p_k\| \geq d^{-1}\left(\frac{1-\delta}{M} <\nabla\phi(x_k), p_k> \right),$$

odnosno

$$(3.37)\qquad \alpha_k \geq \frac{1}{2\|p_k\|} d^{-1}\left(\frac{1-\delta}{M} <\nabla\phi(x_k), p_k> \right) \geq \frac{1}{2M} d^{-1}\left(\frac{1-\delta}{M} <\nabla\phi(x_k), p_k> \right).$$

Iz (3.30) sada na osnovu (3.37) sledi

$$\phi(x_k) - \phi(x_k - \alpha_k p_k) \geq \frac{1}{2M} d^{-1}\left(\frac{1-\delta}{M} <\nabla\phi(x_k), p_k> \right) \sigma(<\nabla\phi(x_k), p_k>), \text{ odnosno,}$$

$$(3.38)\qquad \phi(x_k) - \phi(x_k - \alpha_k p_k) \geq \mu(<\nabla\phi(x_k), p_k>),$$

gde je μ *prinudna funkcija oblika* $\mu(t) = \frac{1}{2M} d^{-1}\left(\frac{1-\delta}{M} t \right) \sigma(t)$.

Kako je ϕ *ograničena odozdo na* L *i* $\phi(x_k) \geq \phi(x_{k+1})$, *sledi da* $\phi(x_k) - \phi(x_{k+1}) \to 0$, $k \to \infty$.

Iz (3.38) sledi da $\mu(<\nabla\phi(x_k), p_k>) \to 0$, $k \to \infty$, *a odavde sledi tvrdjenje 4. teoreme.*

3.4. Teoreme konvergencije

Ovde ćemo izložiti pod kojim će uslovima niz $\{x_k\}$ *koji zadovoljava relacije (3.2), (3.3), a korak* α_k *je definisan jednim od izložena tri algoritma, konvergirati ka optimalnom rešenju problema (3.1).*

Teorema 3.8. *Neka su zadovoljene pretpostavke teoreme 3.3. Tada, ako je niz pravaca* $\{p_k\}$ *takav da* $\|\nabla\phi(x_k)\| \to 0$ *kada* $\langle\nabla\phi(x_k),p_k\rangle \to 0$, $k\to\infty$ *važi sledeće: svaka tačka nagomilavanja niza* $\{x_k\}$ *je stacionarna tačka problema (3.1).*

Dokaz. *Na osnovu pretpostavki teoreme i iz teoreme 3.3 sledi da* $\|\nabla\phi(x_k)\| \to 0$, $k\to\infty$. *Dalje, iz leme 3.2 sledi da je niz* $\{x_k\}$ *sadržan u kompaktnom skupu* L^0. *Prema tome, postoji podniz* $\{x_k\}\subset L^0$, $k\in K\subset\{0,1,2,\ldots\}$ *takav da* $x_k\to x^*$, $k\in K$, $k\to\infty$, $x^*\in L^0$. *Iz neprekidnosti* $\nabla\phi$ *sledi da je* $\nabla\phi(x^*) = 0$.

Posledica 3.8. *Neka su zadovoljene pretpostavke teoreme 3.7. i neka je niz* $\{p_k\}$ *takav da* $\|\nabla\phi(x_k)\| \to 0$ *kada* $\langle\nabla\phi(x_k),p_k\rangle \to 0$, $k\to\infty$. *Tada važi tvrdjenje teoreme 3.8.*

Dokaz. *Dokaz sledi iz teoremâ 3.7 i 3.8.*

Teorema 3.9. *Neka važe pretpostavke teoreme 3.8 i neka uz to niz tačaka* $\{x_k\}$ *zadovoljava uslov*

$$(3.39)\quad \rho(\|\nabla\phi(x_k)\|) \geq \mu(\|x_k - x_{k+1}\|),$$

gde je $\mu : [0,\infty)\to[0,\infty)$ *prinudna funkcija, a* $\rho : [0,\infty)\to[0,\infty)$ *je takva funkcija da* $\rho(t)\to 0$ *kada* $t\to 0$ *i* $\rho(t) = 0$ *ako i samo ako je* $t = 0$. *Tada, ako* $\nabla\phi$ *ima konačno mnogo nula na* L^0, *važi sledeće*

$$x_k\to x^*,\quad k\to\infty\quad,\qquad \nabla\phi(x^*) = 0,$$

tj. niz $\{x_k\}$ *konvergira ka jednoj od stacionarnih tačaka* ϕ.

Dokaz. *Iz teoreme 3.8. sledi da* $\|\nabla\phi(x_k)\| \to 0$, $k\to\infty$. *Iz svojstava funkcije* ρ *sledi da*

$$(3.40)\quad \rho(\|\nabla\phi(x_k)\|)\to 0,\quad k\to\infty.$$

Iz (3.39) i (3.40) sledi da

(3.41) $\mu\,(\|x_k - x_{k+1}\|) \to 0,\ k \to \infty.$

Iz (3.41), na osnovu definicije pri nudne funkcije, sledi

(3.42) $\|x_k - x_{k+1}\| \to 0,\ k \to \infty.$

Iz (3.42) i teoreme 3.8 sledi, na osnovu leme 1.12, da $x_k \to x^*$, $k \to \infty$ *i* $\nabla\phi\,(x^*) = 0$.

Posledica 3.9. *Neka su zadovoljene pretpostavke posledice 3.8, uslov (3.39) i neka skup* $\{x \in L \mid \nabla\phi(x)=0\}$ *ima konačno mnogo tačaka. Tada važi tvrdjenje teoreme 3.9.*

Dokaz. *Dokaz je analogan dokazu teoreme 3.9.*

Teorema 3.10. *Neka su zadovoljene pretpostavke teoreme 3.5. Tada, ako je niz* $\{p_k\}$ *takav da* $\|\nabla\phi(x_k)\| \to 0$ *kada* $\frac{\langle\nabla\phi(x_k), p_k\rangle}{\|p_k\|} \to 0, k \to \infty$ *svaka tačka nagomilavanja niza* $\{x_k\}$ *je stacionarna tačka.*

Dokaz. *Dokaz sledi iz teoreme 3.5 analogno dokazu teoreme 3.8.*

Posledica 3.10. *Neka su zadovoljene pretpostavke teoreme 3.10. Tada, ako* $\nabla\phi$ *ima konačan broj nula na* L^0, *niz* $\{x_k\}$ *konvergira ka jednoj od stacionarnih tačaka problema (3.1).*

Dokaz. *Iz teoreme 3.5 sledi da* $\|x_k - x_{k+1}\| \to 0,\ k \to \infty$. *Odavde, i iz pretpostavki posledice, na osnovu leme 1.12 sledi tvrdjenje.*

Teorema 3.11. *Neka su zadovoljene pretpostavke teoreme 3.5, s tim što prinudna funkcija d zadovoljava uslov*

$$\varepsilon t \le d(t) \le \delta t,\quad 0<\varepsilon<\delta< 2 \qquad \text{za } \forall t \ge 0.$$

Neka je niz vektora pravaca $\{p_k\}$ *takav da važi*

$$\langle\nabla\phi(x_k), p_k\rangle \ge \beta\,\|\nabla\phi(x_k)\|\,\|p_k\|\ , \qquad 0<\beta\le 1.$$

Tada, ako $\nabla\phi$ *ima konačan broj nula u* L^0, *sledi, na osnovu posledice 3.10, da niz* $\{x_k\}$ *konvergira ka jednoj od stacionarnih tačaka problema (3.1).*

Ako je uz to ϕ *dva puta neprekidno diferencijabilna u okolini tačke* x^* *i ako je hesijan* $H\,(x^*)$ *nesingularan, brzina konvergencije niza* $\{x_k\}$ *ka* x^* *je u najmanju ruku R-linearna.*

Dokaz. *Na osnovu relacije (3.28) u dokazu teoreme 3.5 imamo da važi*

$$\phi(x_k) - \phi(x_{k+1}) \geqslant \frac{2-\delta}{2\gamma\delta}\left[d\left(\frac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|}\right)\right]^2 \geqslant \frac{\varepsilon^2\beta^2(2-\delta)}{2\gamma\delta}\|\nabla\phi(x_k)\|^2, \quad tj.$$

$$\phi(x_k)-\phi(x_{k+1}) \geqslant \eta\ \|\nabla\phi(x_k)\|^2, \text{ gde je } \eta = \frac{\varepsilon^2\beta^2(2-\delta)}{2\gamma\delta} > 0.$$

Odavde, na osnovu teoreme 1.21 sledi da je brzina konvergencije niza $\{x_k\}$ *ka* x^* *u najmanju ruku R - linearna.*

Teorema 3.12. *Neka važe pretpostavke teoreme 3.8, odnosno posledice 3.8, odnosno teoreme 3.10. Tada, ako je funkcija* ϕ *pseudokonveksna, svaka tačka nagomilavanja niza* $\{x_k\}$ *jeste optimalno rešenje problema (3.1).*

Ako je funkcija ϕ *strogo pseudokonveksna, niz* $\{x_k\}$ *konvergira ka tački* x^**, gde je* x^* *- jedinstveno optimalno rešenje problema (3.1).*

Dokaz. *Dokaz sledi iz teoreme 3.8, odnosno posledice 3.8, odnosno teoreme 3.10 i definicije 1.6 i leme 1.4.*

Teorema 3.13. *Neka je* $\phi: R^n \to R$ *dva puta neprekidno diferencijabilna funkcija takva da postoji* $m > 0$ *tako da važi:*

$$(3.43) \qquad m\|y\|^2 \leq \langle y, H(x)y\rangle$$

za svako $x \in R^n$ *i svako* $y \in R^n$. *Dalje, neka je niz vektora* $\{p_k\}$ *takav da je* $\|p_k\| \leq M$ *za svako k i*

$$(3.44) \qquad \langle\nabla\phi(x_k), p_k\rangle \geqslant \beta\|\nabla\phi(x_k)\|\,\|p_k\|, \qquad 0<\beta\leq 1$$

i neka funkcija σ *iz uslova (3.30) i (3.31) zadovoljava uslov* $\delta_1 t \leq \sigma(t) \leq \delta_2 t$, $0<\delta_1<\delta_2<1$.

Tada je brzina konvergencije niza $\{x_k\}$ *generisanog poopštenim Céa-Goldstein-ovim algoritmom u najmanju ruku R-linearna.*

Dokaz. *Iz uslova (3.43) sledi da je funkcija* ϕ *, na osnovu leme 1.10 strogo konveksna, a nivoski skup* $L = \{x \in R^n \mid \phi(x) \leq \phi(x_o)\}$ *konveksan i kompaktan. Dalje, pošto je* ϕ *dva puta neprekidno diferencijabilna iz (3.43) takodje sledi da postoji* $K > 0, K \geqslant m$, *takvo da važi*

(3.45) $$m\|y\|^2 \le \langle y, H(x)y\rangle \le K\|y\|^2$$

za $\forall x \in L$, $\forall y \in R^n$.

Iz teoreme 3.12 sledi da $x_k \to x^*$, $k\to\infty$, gde je x^* – jedinstveno optimalno rešenje problema (3.1).

Da bi smo skratili pisanje, označimo sa F sledeću funkciju:

$F(x,p,\alpha)=\phi(x)-\phi(x-\alpha p)$.

Iz Taylor-ove teoreme sledi:

$$F(x_k,p_k,\alpha)=\alpha\langle\nabla\phi(x_k),p_k\rangle-\alpha^2\int_0^1(1-t)\langle p_k,H(x_k-t\alpha p_k)p_k\rangle dt.$$

Odavde, na osnovu (3.45) sledi

(3.46) $$F(x_k,p_k,\alpha)\geq \alpha\langle\nabla\phi(x_k),p_k\rangle-\frac{\alpha^2}{2}K\|p_k\|^2.$$

Uslov (3.31) možemo napisati u ovom obliku:

$F(x_k,p_k,2\alpha)<2\alpha\sigma(\langle\nabla\phi(x_k),p_k\rangle)$,

odnosno, uvodjenjem funkcije $\bar{F}$,

$\bar{F}(x_k,p_k,2\alpha)=F(x_k,p_k,2\alpha)-2\alpha\sigma(\langle\nabla\phi(x_k),p_k\rangle)<0.$

Iz (3.44) i (3.46) sledi

$$\bar{F}(x_k,p_k,2\alpha)\geq 2\alpha\langle\nabla\phi(x_k),p_k\rangle-2\alpha^2K\|p_k\|^2-2\alpha\,\sigma(\langle\nabla\phi(x_k),p_k\rangle)\geq$$

$$\geq 2\alpha\langle\nabla\phi(x_k),p_k\rangle-2\alpha\delta_2\langle\nabla\phi(x_k),p_k\rangle-2\alpha^2K\|p_k\|^2\geq$$

$$\geq 2\alpha\beta(1-\delta_2)\|\nabla\phi(x_k)\|\,\|p_k\|-2\alpha^2K\|p_k\|^2.$$

Dalje, pošto α_k mora zadovoljavati uslov $\bar{F}(x_k,p_k,2\alpha_k)<0$, iz poslednje nejednakosti sledi da mora važiti

$$\alpha_k>\frac{\beta(1-\delta_2)\|\nabla\phi(x_k)\|}{K\|p_k\|}.$$

Iz ove relacije i iz (3.30) sledi

$$\phi(x_k)-\phi(x_k-\alpha p_k)\geq\alpha_k\sigma(\langle\nabla\phi(x_k),p_k\rangle)>$$

$$> \frac{\beta^2 \delta_1(1-\delta_2)}{K} \, \|\nabla\phi(x_k)\|^2 = \eta \|\nabla\phi(x_k)\|^2, \text{ gde je } \quad \eta = \frac{\beta^2 \ \delta_1(1-\delta_2)}{K} > 0.$$

Iz poslednje nejednakosti, na osnovu teoreme 1.21 sledi da je brzina konvergencije niza $\{x_k\}$ *ka* x^* *u najmanju ruku R-linearna.*

3.5. Uslovne verzije algoritama za nalaženje koraka

Razmatramo problem: naći

$$(3.47) \qquad \min\{\ \phi(x) |\ x\in X\}, \qquad \text{gde je } X=\{x\in D\,|\,f_j(x)\leq 0,\ j\in I_0=\{1,\dots,m\}\},$$

a funkcije $\phi, f_i : D\subset R^n \to R, i\in I_0$ *su neprekidno diferencijabilne na otvorenom konveksnom skupu* D *u pri tom dopustivi skup* X*, odnosno funkcije ograničenjâ* f_i, $i\in I_0$, *zadovoljavaju Rauch-ove pretpostavke 1,2 i 3 iz odeljka 1.3.*

Označimo sa L^0 *lučno povezanu komponentu nivoskog skupa* $L=\{x\in D\,|\,\phi(x)\leq \phi(x_0)\}$ *za neko* $x_0\in D$.

Ovde ćemo izložiti uslovne verzije algoritama za nalaženje koraka iz odeljka 3.1.

Teorema 3.14. *(Uslovna verzija modifikovanog Curry-Altman-ovog algoritma). Neka je* $\phi : D\subset R^n \to R$ *neprekidno diferencijabilna funkcija na otvorenom konveksnom skupu* D*. Neka je* $L^0\cap X$ *kompaktan skup. Dalje pretpostavimo da* $x_k\in L^0\cap X$*, da vektor pravca* $p_k\in R^n$ *zadovoljava uslov (1.24) i da je* α_k *definisano modifikovanim Curry-Altman-ovim algoritmom. Neka je* $\alpha_k{}^*$ *dato sa (1.25). Definišimo* α_k *na sledeći način:*

$$(3.48) \qquad \alpha_k = \min\{\ \alpha_k',\ \alpha_k''\ \}.$$

Tada, ako je $\alpha_k' \geq q\ \bar{\alpha}_k$, $0\leq q<1$, *gde je* $\bar{\alpha}_k$ *– najmanje pozitivno rešenje poopštene Curry-Altman-ove jednačine (3.6), važi sledeće:*

1. $x_{k+1} = x_k - \alpha_k p_k \in L^0\cap X;$

2. $\phi(x_k) \geq \phi[\lambda x_k + (1-\lambda)x_k] \geq \phi(x_{k+1}), \qquad \lambda\in[0,1];$

3. $\phi(x_k)-\phi(x_{k+1})\geq F(<\nabla\phi(x_k),p_k>,\nu_k,\varepsilon_k)$, *gde je F prinudna funkcija (definicija 1.9) od tri promenljive koja zavisi samo od* $L^0\cap X$.

Dokaz. *Pre svega, iz pretpostavki teoreme sledi da su pretpostavke Rauch-a o problemu (3.47) zadovoljene. Dokaz ćemo izvesti metodom metematičke indukcije u odnosu na k. Pretpostavimo da* $x_k\in L^0\cap X$ *i da je* $<\nabla\phi(x_k), p_k>>0$. *Tada iz leme 1.15 sledi da postoji takvo* $\alpha_d>0$ *da* $(x_k,x_k-\alpha_d p_k)\subset L^0\cap X$. *Iz leme 3.2 sledi da je* $\alpha_k'>0$. *Ako* $x_k-\alpha p_k\in(x_k,x_k-\alpha_k' p_k]\cap L^0\cap X$, *na osnovu Lagrange-ove teoreme imamo:*

$$\phi(x_k)-\phi(x_k-\alpha p_k)=\alpha <\nabla\phi(x_k-\hat{\alpha}p_k),p_k> \text{ za neko } \hat{\alpha}\in(0,\alpha). \tag{3.49}$$

Iz (3.49) i relacije (3.10) u dokazu teoreme 3.3 sledi:

$$\phi(x_k)-\phi(x_k-\alpha p_k)>\alpha\sigma(<\nabla\phi(x_k),p_k>)>0.$$

Prema tome, relacija (1.27) važi sa $\hat{\alpha}=\alpha_k'$. *Pošto odavde sledi da su pretpostavke leme 1.15 zadovoljene, sledi da* $[x_k,x_k-\alpha_k' p_k]\cap X\subset L^0\cap X$.

Dalje, kako je $\alpha_k\leq\alpha_k^*$, *sledi da je* $[x_k,x_k-\alpha_k p_k]\cap X=[x_k,x_k-\alpha_k p_k]\subset L^0\cap X$, *a odavde, da* $x_{k+1}=x_k-\alpha_k p_k\in L^0\cap X$. *Dalje, pošto je* $\alpha_k\leq\alpha_k'$, *iz relacije (3.8) u dokazu leme 3.2 sledi*

$$\phi(x_k)\geq\phi\ \lambda[x_k+(1-\lambda)x_{k+1}]\geq\phi(x_{k+1}) \qquad \text{za } \forall\lambda\in[0,1].$$

Na osnovu Lagrange-ove teoreme, relacijâ (3.7) i (3.10) i činjenice da je $\alpha_k\leq\alpha_k'$, *imamo:*

$$\phi(x_k)-\phi(x_{k+1})=\alpha_k<\nabla\phi(x_k-\alpha_s p_k),p_k>>\alpha_k\sigma(<\nabla\phi(x_k),p_k>),\quad \alpha_s\in(0,\alpha_k). \tag{3.50}$$

Zatim, korišćenjem relacijâ (3.14), (3.48) i (1.26) dobijamo:

$$\alpha_k=\min\{\alpha_k',\alpha_k^*\}\geq\min\{\alpha_k',\ \min\{\alpha_k',\hat{F}(\nu_k,\varepsilon_k)\}\}=\min\{\alpha_k',\hat{F}(\nu_k,\varepsilon_k)\}\geq \tag{3.51}$$

$$\geq\min\{qs\,[(1-\delta)<\nabla\phi(x_k),p_k>,\hat{F}(\nu_k,\varepsilon_k)\}.$$

Iz (3.50) i (3.51) sledi

$$\phi(x_k)-\phi(x_{k+1}) > \min\{qs[(1-\delta)\ \langle\nabla\phi(x_k),p_k\rangle],\hat{F}(\nu_k,\varepsilon_k)\}\sigma(\ \langle\nabla\phi(x_k),p_k\rangle)=$$

$$=F(\ \langle\nabla\phi(x_k),p_k\rangle,\ \nu_k,\varepsilon_k),$$

gde je F prinudna funkcija oblika $F(t_1,t_2,t_3)=\min\{qs\ [(1-\delta)t_1],\hat{F}(t_2,t_3)\}\sigma(t_1)$.

Teorema 3.15. *(Uslovna verzija poopštenog algoritma Ostrowskog)* *Pretpostavimo da je skup* $L^0\cap X$ *kompaktan i da funkcija* ϕ *zadovoljava uslov:*

$$(3.52)\qquad \|\nabla\phi(x)-\nabla\phi(y)\|\leq \gamma\|x-y\| \qquad \text{za } \forall x,y\in L^0\cap X.$$

Dalje, prepostavimo da $x_k\in L^0\cap X$, *da vektor pravca* $p_k\in R^n$ *zadovoljava uslov (1.24) i da je* $\alpha_k^{?}$ *definisano pooptenim algoritmom Ostrowskog. Neka je* $\alpha_k{}^*$ *dato sa (1.25). Definišimo* α_k *na sledeći način:*

$$(3.53)\qquad \alpha_k=\min\{\ \alpha_k^{?},\alpha_k{}^*\}=\min\{\frac{d(\ \langle\nabla\phi(x_k),p_k\rangle)}{\gamma},\alpha_k{}^*\}.$$

Tada važi sledeće:

1. $x_{k+1}=x_k-\alpha_k p_k\in L^0\cap X;$

2. $\phi(x_k)-\phi(x_{k+1})\geq F(\ \langle\nabla\phi(x_k),p_k\rangle,\nu_k,\varepsilon_k)$, *gde je F prinudna funkcija od tri promenljive koja zavisi samo od* $L^0\cap X$.

Dokaz. *Iz pretpostavki teoreme sledi da su Rauch-ove pretpostavke o problemu (3.47) ispunjene. Dokaz ćemo izvesti metodom matematičke indukcije u odnosu na k. Pretpostavimo da* $x_k\in L^0\cap X$ *i da je* $\langle\nabla\phi(x_k),p_k\rangle>o$. *Tada iz leme 1.15 sledi da postoji* $\alpha_d>o$ *takvo da* $(x_k,x_k-\alpha_d p_k)\subset L^0\cap X$. *Iz (3.52) sledi, na osnovu leme 1.13, da za* $[x_k,x_k-\alpha p_k]\subset L^0\cap X$ *važi:*

$$\phi(x_k)-\phi(x_k-\alpha p_k)\geq\alpha\ \langle\nabla\phi(x_k),p_k\rangle-\frac{1}{2}\gamma\alpha^2.$$

Ako sada uzmemo $x_k-\alpha p_k\in(x_k,x_k-\alpha_k^{?}p_k]\cap L^0\cap X$, *važiće*

$$\phi(x_k)-\phi(x_k-\alpha p_k)\geq\alpha\left[\ \langle\nabla\phi(x_k),p_k\rangle-\frac{1}{2}\gamma\alpha\right]\geq$$

$$\geq \alpha\left[\frac{1}{\delta}\, d(\langle\nabla\phi(x_k), p_k\rangle) - \frac{1}{2}\gamma\alpha\right] \geq \text{ (pošto je } t \geq \frac{d(t)}{\delta}, 0<\delta<2 \text{ za } \forall t\geq 0)$$

$$\geq \alpha\left[\frac{\gamma\alpha'_k}{\delta} - \frac{1}{2}\gamma\alpha'_k\right] = \alpha\gamma\alpha'_k\,\frac{2-\delta}{2\delta} > 0 \quad \text{(pošto je } \alpha'_k = \frac{d(\langle\nabla\phi(x_k), p_k\rangle)}{\gamma} \text{ i}$$

$\alpha \leq \alpha'_k$, $0 < \delta < 2$).

Prema tome, relacija (1.27) važi za $\hat{\alpha} = \alpha'_k$. Pošto odavde sledi da su pretpostavke leme 1.15 zadovoljene, sledi da $[x_k, x_k - \alpha'_k p_k] \cap X \subset L^0 \cap X$. Dalje, kako je $\alpha_k \leq \alpha_k^*$ sledi da je $[x_k, x_k - \alpha_k p_k] \cap X = [x_k, x_k \alpha_k p_k] \subset L^0 \cap X$ i dakle,

$x_{k+1} = x_k - \alpha_k p_k \in L^0 \cap X$.

Dakle i za $[x_k, x_{k+1}]$ važi na osnovu leme 1.13:

$$\phi(x_k) - \phi(x_k - \alpha_k p_k) \geq \alpha_k \langle\nabla\phi(x_k), p_k\rangle - \frac{1}{2}\gamma\alpha_k^2 \geq \text{(na osnovu (3.53))} \geq \gamma\alpha_k^2\,\frac{2-\delta}{2\delta} =$$

$$= \frac{\gamma(2-\delta)}{2\delta}\left[\min\left\{\frac{d(\langle\nabla\phi(x_k), p_k\rangle)}{\gamma}, \alpha_k^*\right\}\right]^2 \geq \quad \text{(na osnovu (1.26))}$$

$$\geq \frac{\gamma(2-\delta)}{2\delta}\left[\min\left\{\frac{d(\langle\nabla\phi(x_k), p_k\rangle)}{\gamma}, \min\left\{\frac{d(\langle\nabla\phi(x_k), p_k\rangle)}{\gamma}, \hat{F}(\nu_k, \varepsilon_k)\right\}\right\}\right]^2 =$$

$$= \frac{\gamma(2-\delta)}{2\delta}\left[\min\left\{\frac{d(\langle\nabla\phi(x_k), p_k\rangle)}{\gamma}, \hat{F}(\nu_k, \varepsilon_k)\right\}\right]^2 = \frac{\gamma(2-\delta)}{2\delta} F(\langle\nabla\phi(x_k), p_k\rangle, \nu_k, \varepsilon_k),$$

gde je F prinudna funkcija oblika $F(t_1, t_2, t_3) = \left[\min\left\{\frac{d(t_1)}{\gamma}, \hat{F}(t_2, t_3)\right\}\right]^2$.

Teorema 3.16. *(Uslovna verzija poopštenog Céa-Goldstein-ovog algoritma - uporedi /49/, str. 216.)*

Neka je skup $L \cap X$ kompaktan, neka $x_k \in L \cap X$, vektor $p_k \in R^n$ zadovoljava uslov (1.15) i neka je $\alpha'_k \geq 0$ ma koji broj definisan poopštenim Céa-Goldstein-ovim algoritmom. Neka je α_k^* dato sa (1.25) i neka je $\alpha_k = \min\{\alpha'_k, \alpha_k^*\}$. Dalje, neka je:

$$(3.54)\qquad \lambda_k = \begin{cases} \lambda_k^*, \text{ ako je } \alpha_k=\alpha_k^* \text{ i } \phi(x_k)-\phi(x_k-\alpha_k^* p_k)< \alpha_k^* \sigma(\langle\nabla\phi(x_k),p_k\rangle); \\ 1, \text{ inače}, \end{cases}$$

gde je $\lambda_k^* \in (0,1)$ izabrano tako da (3.30) i (3.31) važi sa $\alpha=\lambda_k^* \alpha_k^*$. Tada:

1. $x_{k+1}=x_k-\lambda_k \alpha_k p_k \in L\cap X$;

2. $\phi(x_k)-\phi(x_{k+1})\geq F(\langle\nabla\phi(x_k),p_k\rangle, \nu_k, \sigma_k)$, gde je F prinudna funkcija od tri promenljive koja zavisi samo od $L\cap X$.

Dokaz. *Rauch-ove pretpostavke o problemu (3.47) su očigledno zadovoljene, a dokaz ćemo kao i dosad, izvesti metodom matematičke indukcije u odnosu na k. Ako pretpostavimo da $x_k\in L\cap X$ i da je $\langle\nabla\phi(x_k),p_k\rangle>0$, iz leme 3.6 sledi da postoji $\alpha_k' >0$ koje zadovoljava (3.30) i (3.31). Takodje iz leme 3.6 sledi da je λ_k dobro definisano. Dalje, iz (3.54) sledi da je $\lambda_k\alpha_k\leq \alpha_k^*$ i*

$$\phi(x_k)-\phi(x_{k+1})=\phi(x_k)-\phi(x_k- \lambda_k \alpha_k p_k) \geqslant \lambda_k \alpha_k\sigma(\langle\nabla\phi(x_k),p_k\rangle)> 0.$$

Prema tome, $x_{k+1}\in L\cap X$. Iz (1.26) sledi da je $\alpha_k^* \geq \min\{ \alpha_k', \hat{F}(\nu_k, \varepsilon_k)\}$. Dalje, iz konveksnosti skupa D sledi da $[x_k,x_{k+1}]\subset D$ i da važi ocena (3.37), tj.:

$$(3.55)\qquad \min\{ \alpha_k', \lambda_k^*\alpha_k^*\}\geq \frac{1}{2}\, d^{-1} \left[(1-\delta) \langle\nabla\phi(x_k),p_k\rangle\right],$$

gde je d^{-1} prinudna funkcija. Sada, iz (3.30), (3.54), (3.55) i (1.26) sledi

$$\phi(x_k)-\phi(x_k- \lambda_k\alpha_k p_k)\geq \lambda_k \alpha_k\sigma(\langle\nabla\phi(x_k),p_k\rangle)=\min\{\alpha_k', \lambda_k^*\alpha_k^*, \alpha_k^*\}\sigma(\langle\nabla\phi(x_k),p_k\rangle)\geq$$

$$\geq \min\{ \frac{1}{2}\, d^{-1}\left[(1-\delta) \langle\nabla\phi(x_k),p_k\rangle\right],\hat{F}(\nu_k, \varepsilon_k)\}\sigma(\langle\nabla\phi(x_k),p_k\rangle)=$$

$$= F(\langle\nabla\phi(x_k),p_k\rangle, \nu_k, \varepsilon_k),$$

gde je F prinudna funkcija oblika $F(t_1,t_2,t_3)=\sigma(t_1)\min\{ \frac{1}{2}\, d^{-1}[(1-\delta)t_1],\hat{F}(t_2,t_3)\}$.

U algoritmima opisanim teoremama 3.14, 3.15 i 3.16 korak α_k, pored toga što obezbedjuje dopustivost tačaka (uz pretpostavku da pravac $p_k\in R^n$ zadovoljava uslov (1.24), zadovoljava i takozvani princip "dovoljnog opadanja":

$$(3.56)\qquad \phi(x_k)-\phi(x_{k+1})\geq F(\langle\nabla\phi(x_k),p_k\rangle, \nu_k, \sigma_k),$$

gde je F prinudna funkcija od tri promenljive.

Sada ćemo izložiti teoremu u kojoj dokazujemo da pod odredjenim uslovima niz $\{x_k\}$konvergira ka stacionarnoj tački problema (3.47).

Teorema 3.17. *Razmotrimo iteraciju* $x_{k+1}=x_k-\alpha_k p_k$, $k=0,1,\ldots$, *gde je* ε_k *izabrano pomoću algoritma za* ε *(str.22,I poglavlje),* $p_k\in R^n$ *je ma koji pravac takav da zajedno sa* α_k *dobijenim pomoću bilo kojeg od algoritama opisanih teoremama (3.14) i (3.15) imamo da* $x_{k+1}\in L^0\cap X$ *i relacija (3.56) važi sa*

$$\nu_k = \| P_{q_k}(x_k)\ \nabla\phi(x_k)\| \ , q_k=\mathrm{kard} I(x_k,\varepsilon_k).$$

Tada, ako je niz $\{p_k\}$ *takav da* $\| P_{q_k}(x_k)<\nabla\phi(x_k)\|\to 0$ *ako* $<\nabla\phi(x_k),p_k>\to 0$ *kada* $k\to\infty$, *i skup*

$$\{x\in L^0\cap X|\ \| P_q(x)\ \nabla\phi(x)\| =0,\quad q=\mathrm{kard}\ I(x,0)\}$$

sadrži samo jednu tačku x^*, *sledi da je* $\lim_{k\to\infty} x_k = x^*$.

Dokaz. *Iz pretpostavki teoreme sledi da važi*

$$\phi(x_k)-\phi(x_{k+1})\geq F(\ <\nabla\phi(x_k),p_k>,\ \| P_{q_k}(x_k)\ \nabla\phi(x_k)\| \ ,\varepsilon_k),\quad q_k=\mathrm{kard}\ I(x_k,\varepsilon_k).$$

Iz leme 1.16 i definicije 1.14 prinudne funkcije sledi da $<\nabla\phi(x_k),p_k>\to 0$ *ili* $\| P_{q_k}(x_k)\ \nabla\phi(x_k)\|\to 0$ *ili* $\varepsilon_k\to 0$ *kada* $k\to\infty$. *Pošto iz algoritma za* ε *sledi da* $\varepsilon_k\to 0$ *ako i samo ako* $\| P_{q_k}(x_k)\ \nabla\phi(x_k)\|\to 0$, $k\to\infty$ *a* $<\nabla\phi(x_k),p_k>\to 0$ *povlači, prema pretpostavci, da* $\| P_{q_k}(x_k)\ \nabla\phi(x_k)\|\to 0, k\to\infty$, *u svakom slučaju će slediti da* $\| P_{q_k}(x_k)\nabla\phi(x_k)\|\to 0$, $k\to\infty$, *a iz leme 1.16 na osnovu ovoga sledi da* $x_k\to x^*$, $k\to\infty$.

Posledica 3.17. *Ako učinimo analogne pretpostavke o skupu* $L\cap X$, *dobićemo analogno tvrdjenje i za algoritam opisan teoremom 3.16.*

4. POGLAVLJE

Dat je problem bezuslovne optimizacije : naći

$$(4.1) \qquad \min\{\phi(x) \mid x \in R^n\},$$

gde je $\phi: R^n \to R$ *dva puta neprekidno diferencijabilna funkcija.*

U ovom poglavlju ćemo izložiti neke modifikacije metode Danilin-Pšeničnog izložene u I poglavlju. Modifikacije se sastoje u tome što ćemo za vektore pravca uzeti vektore definisane metodom Danilin-Pšeničnog, a korak α_k *ćemo definisati novim algoritmima.*

Drugim rečima, iterativni algoritmi za nalaženje optimalnog rešenja problema (4.1), koje ćemo ovde izložiti, generišu nizove tačaka $\{x_k\}$ *oblika:*

$$(4.2) \qquad x_{k+1} = x_k - \alpha_k p_k, \qquad k=0,1,2\ldots \quad \text{gde je}$$

$$(4.3) \qquad p_k = A_k^{-1}\,\nabla\phi(x_k) \quad \text{za } \forall k \geq n-1$$

(za k=0, 1, ...,n-2 može se uzeti npr. $p_k = \nabla\phi(x_k)$*), pri čemu je matrica* A_k *definisana sistemom (1.32) i pretpostavljamo da zadovoljava uslov*

$$(4.4) \qquad \langle A_k^{-1}\nabla\phi(x_k), \nabla\phi(x_k)\rangle > 0 \qquad \text{za } \forall\, k \geq n-1$$

Korak α_k *će biti definisan jednim od algoritama koji slede.*

Primedba. *Ako je* $\nabla\phi(x_k)$ *različito od nule, iz (4.4) sledi da je* p_k *različito od nule.*

4.1. I modifikacija metode Danilin-Pšeničkog

Neka niz $\{x_k\}$ *zadovoljava relacije (4.2), (4.3) i (4.4). Korak* α_k *definišimo na sledeći način:*

$\alpha_k \in I_k$*, gde* I_k *predstavlja prvi interval pozitivnih rešenja nejednačine*

$$(4.5) \qquad \langle\nabla\phi(x_k - \alpha p_k), p_k\rangle > \varepsilon\langle\nabla\phi(x_k), p_k\rangle, \qquad 0 < \delta < 1.$$

Teorema 4.1. *Neka je* $\phi: R^n \to R$ *dva puta neprekidno diferencijabilna funkcija koja zadovoljava uslov*

(4.6) $$<H(x)y,y> \geq m \,\| y \|^2 \quad za \; \forall \, x,y \in R^n, \; m>0,$$

(gde je sa $H(x)$ označen hesijan funkcije ϕ u tački x). Neka je niz $\{x_k\}$ zadat relacijama (4.2), (4.3), pri čemu je ispunjen uslov (4.4). Neka je korak α_k definisan upravo izloženim algoritmom. Tada, ako je $\alpha_k \geq q\bar{\alpha}_k$, $0 \leq q < 1$, gde je $\bar{\alpha}_k$ – najmanje pozitivno rešenje Curry-Altman-ove jednačine

(4.7) $$<\nabla\phi(x_k-\alpha p_k),p_k> = \delta<\nabla\phi(x_k),p_k> \quad , \; 0<\delta<1,$$

važi sledeće:

1. $x_k \in L = \{x \in R^n \mid \phi(x) \leq \phi(x_o)\}$;

2. $\phi(x_{k+1}) < \phi(x_k)$;

3. $\| x_k - x^* \| \to 0$, $k \to \infty$, gde je x^* – jedinstveno optimalno rešenje problema (4.1).

Dokaz. Pretpostavimo da je $\nabla\phi(x_k) \neq 0$. Pošto, zbog (4.4) važi da je

$$<\nabla\phi(x_k),p_k> = <\nabla\phi(x_k),A_k^{-1}\nabla\phi(x_k)> > 0 \quad za \; \forall \, k \geq n-1,$$

sledi da su zadovoljene pretpostavke posledice 3.3 i teoreme 3.4, odakle imamo da važi sledeće: $x_k \in L$, gde je L – konveksan i kompaktan skup, $\phi(x_k) > \phi(x_{k+1})$,

(4.8) $$m \,\| y \|^2 \leq <H(x)y,y> \leq M \,\| y \|^2 \quad za \; \forall \, x \in L, \; \forall \, y \in R^n,$$

(4.9) $$\lim_{k\to\infty} \frac{<\nabla\phi(x_k),p_k>}{\| p_k \|} = 0,$$

(4.10) $$\lim_{k\to\infty} \| x_k - x_{k+1} \| = 0.$$

Iz (4.9) i činjenice da $x_k \in L$ sledi da su zadovoljene pretpostavke leme 1.17, a odatle imamo da važi

(4.11) $$\lim_{k\to\infty} \| A_k - H(x_k) \| = 0.$$

Iz (4.11), na osnovu relacije (4.8), sledi da se za proizvoljne M_1 i m_1 takve da je $M_1 \geq M$ i $0 < m_1 \leq m$ može naći broj $k_o > 0$ takav da za $\forall \, k \geq k_o$ i za $\forall \, y \in R^n$ važi

(4.12) $\quad m_1 \| y \|^2 \leq \langle A_k y, y \rangle \leq M_1 \| y \|^2$.

Iz (4.12) sledi da za $\forall k > k_0$ važi

(4.13) $\quad \langle \nabla\phi(x_k), p_k \rangle = \langle A_k p_k, p_k \rangle \geq m_1 \| p_k \|^2$, odnosno,

(4.14) $$\frac{\langle \nabla\phi(x_k), p_k \rangle}{\| p_k \|} \geq m_1 \| p_k \| .$$

Iz (4.9) i (4.14) sledi da

(4.15) $\quad \| p_k \| \to 0,\ k \to \infty.$

Na osnovu relacije (4.12) imamo da važi, počev od nekog k

(4.16) $\quad \| \nabla\phi(x_k) \| = \| A_k p_k \| \leq M_1 \| p_k \|$.

Iz (4.16) sledi da $\| \nabla\phi(x_k) \| \to 0$, $k \to \infty$. Odavde, zatim, na osnovu stroge konveksnosti ϕ (relacija (4.8)) i iz činjenice da niz $\{x_k\}$ pripada kompaktnom skupu L sledi da niz $\{x_k\}$ konvergira ka jedinstvenom optimalnom rešenju x^* problema (4.1).

Teorema 4.2. *Neka su zadovoljene pretpostavke teoreme 4.1. Tada je brzina konvergencije niza tačaka $\{x_k\}$ ka x^* u najmanju ruku R-linearna.*

Dokaz. Iz relacije (3.17) sledi da važi

(4.17) $\quad \phi(x_k) - \phi(x_{k+1}) \geq q\, \delta \bar\alpha_k \langle \nabla\phi(x_k), p_k \rangle$.

Dalje, pošto je $\bar\alpha_k$ rešenje jednačine (4.7), sledi

$(1-\delta)\ \langle \nabla\phi(x_k), p_k \rangle = \langle \nabla\phi(x_k), p_k \rangle - \langle \nabla\phi(x_k - \bar\alpha_k p_k), p_k \rangle$.

Primenom Lagrange-ove teoreme na $\nabla\phi$ na intervalu $[x_k, x_k - \bar\alpha_k p_k] \subset L$ i korišćenjem relacije (4.8) dobijamo:

(4.18) $$(1-\delta)\ \langle \nabla\phi(x_k), p_k \rangle = \int_0^1 \langle H(x_k - \theta\bar\alpha_k p_k) p_k, p_k \rangle\, d\theta \leq M\, \bar\alpha_k \| p_k \|^2 .$$

Iz (4.18) sledi

(4.19) $$\bar\alpha_k \geq \frac{1-\delta}{M}\ \frac{\langle \nabla\phi(x_k), p_k \rangle}{\| p_k \|^2} .$$

Kombinacijom (4.17) i (4.19) dobijamo

(4.20) $\quad \phi(x_k)-\phi(x_{k+1}) \geq \frac{q\delta\ (1-\delta)}{M} \quad \frac{<\nabla\phi(x_k),p_k>^2}{\|p_k\|^2}$.

Dalje, iz relacije (4.12) sledi da postoji $k_o > 0$ *takvo da važi za* $\forall k \geq k_o$:

(4.21) $\quad <\nabla\phi(x_k),p_k> = <A_k^{-1}\nabla\phi(x_k),\ \nabla\phi(x_k)> \geq \frac{1}{M_1}\ \|\nabla\phi(x_k)\|^2$ *i*

(4.22) $\quad \|p_k\| = \|A_k^{-1}\nabla\phi(x_k)\| \leq \frac{1}{m_1}\ \|\nabla\phi(x_k)\|$.

Iz (4.20), (4.21) i (4.22) dobijamo

$$\phi(x_k)-\phi(x_{k+1}) \geq \frac{q\delta\ (1-\delta)m_1^2}{M\ M_1^2}\ \|\nabla\phi(x_k)\|^2 \quad za\ \forall k \geq k_o,\ tj.$$

(4.23) $\quad \phi(x_k)-\phi(x_{k+1}) \geq \eta\,\|\nabla\phi(x_k)\|^2$, *gde je* $\eta = \frac{q\delta\ (1-\delta)m_1^2}{M\ M_1^2} > 0.$

Iz (4.23) sledi da su zadovoljene pretpostavke teoreme 1.21, a odatle sledi tvrdjenje teoreme.

4.2. II Modifikacija metode Danilin-Pšeničnog

Neka niz $\{x_k\}$ *zadovoljava relacije (4.2), (4.3) i (4.4), a korak* α_k *odredimo iz sledećeg uslova*

(4.24) $\quad \phi(x_k)-\phi(x_k-\alpha p_k) \geq \alpha\ d(\ <\nabla\phi(x_k),p_k>\ ),$

gde je $d: [0,\infty) \to [0,\infty)$ *funkcija takva da važi* $\delta_1 t \leq d(t) \leq \delta_2 t$ *za* $\forall t \geq 0$, $0 < \delta_1 < \delta_2 < 1$. *(Očigledno je da je d jedna prinudna funkcija).*

U sledećoj lemi ćemo pokazati da je korak α_k *dobro definisan.*

Lema 4.3. *Neka je* $\phi: R^n \to R$ *neprekidno diferencijabilna funkcija i pretpostavimo da* x_k, $\alpha_k > 0$ *i* $p_k \in R^n$ *zadovoljavaju uslov* $<\nabla\phi(x_k),p_k> > 0$ *i*

$$\phi(x_k)-\phi(x_k-\alpha_k p_k) < \alpha_k d(\ <\nabla\phi(x_k),p_k>\ ),$$

gde je funkcija $d: [0,\infty) \to [0,\infty)$ *takva da važi* $\delta_1 t \leq d(t) \leq \delta_2 t$ *za* $\forall t \geq 0$, $0 < \delta_1 < \delta_2 < 1$. *Tada postoji* $0 < \bar{\alpha}_k < \alpha_k$ *takvo da važi (4.24).*

Dokaz: *Definišimo funkciju* $F: [0,1]\to R$ *na sledeći način:*

$$F(t)=\begin{cases}\dfrac{\phi(x_k)-\phi(x_k-t\alpha_k p_k)}{t\alpha_k <\nabla\phi(x_k),p_k>}, & t\in(0,1] \ ;\\ 1, & t=0.\end{cases}$$

Iz L'Hospital-ovog pravila sledi da $F(t)\to 0$, $t\to 0$, *a odatle sledi da je* F *neprekidna na* $[0,1]$ *i pri tom važi:* $F(0)=1$ *i*

$$F(1)=\frac{\phi(x_k)-\phi(x_k-\alpha_k p_k)}{\alpha_k <\nabla\phi(x_k),p_k>} < \frac{d(<\nabla\phi(x_k),p_k>)}{<\nabla\phi(x_k),p_k>}\leq\delta_2<1.$$

Iz neprekidnosti F *sledi da* F *poprima sve vrednosti izmedju* $\dfrac{d(<\nabla\phi(x_k),p_k>)}{<\nabla\phi(x_k),p_k>}$ *i 1 za* $t\in[0,1]$ *.Odatle sledi da* $\exists\,\bar t\in(0,1)$ *takvo da važi* $F(1)\leq F(\bar t)\leq F(0)$, *odnosno, da postoji* $\bar\alpha_k=\bar t\ \alpha_k$ *takvo da važi*

$$\phi(x_k)-\phi(x_k-\bar\alpha_k p_k)\geq\bar\alpha_k\ d(<\nabla\phi(x_k),p_k>.$$

U sledećoj teoremi pokazaćemo da niz $\{x_k\}$*pod odredjenim uslovima konvergira ka optimalnom rešenju problema (4.1).*

Teorema 4.4. *Neka su zadovoljene pretpostavke teoreme 4.1, s tim što je korak* α_k *definisan tako da zadovoljava uslov (4.24). Tada važe tvrdjenja 1,2 i 3 teoreme 4.1 i tvrdjenje teoreme 4.2.*

Dokaz. *Iz relacije (4.24) sledi da je* $\phi(x_{k+1})<\phi(x_k)$. *Odavde, ako pretpostavimo da* $x_k\in L$ *sledi da* $x_{k+1}\in L=\{x\in R^n \mid \phi(x)\leq\phi(x_0)\}$. *Kako je* ϕ *monotona i ograničena odozdo na konvesnom i kompaktnom skupu* L, *sledi da*

(4.25) $\qquad \phi(x_k)-\phi(x_{k+1})\to 0,\ k\to\infty.$

Iz (4.24) i (4.25) sledi da

(4.26) $\qquad \alpha_k d(<\nabla\phi(x_k),p_k>)\to 0,\ k\to\infty.$

Pošto je $d(t)\geq\delta_1 t$, *iz (4.26) dalje sledi:*

$$(4.27)\qquad \delta_1\, \alpha_k \langle \nabla\phi(x_k), p_k\rangle \le \alpha_k\, d(\langle \nabla\phi(x_k), p_k\rangle) \to 0, k\to\infty \text{ i}$$

$$(4.28)\qquad \delta_1\, \alpha_k \langle \nabla\phi(x_k), p_k\rangle = \delta_1 \langle \nabla\phi(x_k), x_k - x_{k+1}\rangle \to 0, \quad k\to\infty.$$

Iz (4.28) i monotonosti niza $\{\phi(x_k)\}$ sledi da niz $\{x_k\}$ zadovoljava pretpostavke leme 1.18, a odatle sledi

$$(4.29)\qquad \|x_k - x_{k+1}\| \to 0, \quad k\to\infty.$$

Iz (4.29) iz činjenice da $\{x_k\}$ pripada kompaktnom skupu L sledi da su zadovoljene pretpostavke leme 1.17, a odatle dalje sledi da važi

$$(4.30)\qquad \|A_k - H(x_k)\| \to 0, \quad k\to\infty.$$

Iz (4.30) dalje sledi, kao i u dokazu teoreme 4.1. da važi relacija:

$$(4.31)\qquad m_1 \|y\|^2 \le \langle A_k y, y\rangle \le M_1 \|y\|^2 \qquad \text{za } \forall k \ge k_0.$$

Da bismo dokazali konvergenciju niza $\{x_k\}$ ka optimalnom rešenju problema (4.1), pokažimo pre sve da je niz $\{\alpha_k\}$ ograničen odozdo.

Na osnovu Taylor-ove formule imamo

$$(4.32)\qquad \phi(x_k) - \phi(x_k - \alpha p_k) = \alpha \langle \nabla\phi(x_k), p_k\rangle - \frac{\alpha^2}{2} \langle H(\xi_k) p_k, p_k\rangle =$$

$$(\xi_k = x_k - \theta\alpha p_k,\ \theta \in (0,1))$$

$$= \alpha\, d(\langle \nabla\phi(x_k), p_k\rangle) \left[\frac{\langle \nabla\phi(x_k), p_k\rangle}{d(\langle \nabla\phi(x_k), p_k\rangle)} - \frac{\alpha}{2} \frac{\langle H(\xi_k) p_k, p_k\rangle}{d(\langle \nabla\phi(x_k), p_k\rangle)} \right].$$

Iz (4.32) sledi da će (4.24) biti zadovoljeno ako važi

$$(4.33)\qquad \frac{\langle \nabla\phi(x_k), p_k\rangle}{d(\langle \nabla\phi(x_k), p_k\rangle)} - \frac{\alpha}{2} \frac{\langle H(\xi_k) p_k, p_k\rangle}{d(\langle \nabla\phi(x_k), p_k\rangle)} \ge 1.$$

Dalje, imajući u vidu da je $\delta_1 t \le d(t) \le \delta_2 t$, kao i relacije (4.8) i (4.12), vidimo da će relacija (4.33) biti zadovoljena ako važi

$$\frac{\langle \nabla\phi(x_k), p_k\rangle}{d(\langle \nabla\phi(x_k), p_k\rangle)} - \frac{\alpha}{2} \frac{\langle H(\xi_k) p_k, p_k\rangle}{d(\langle \nabla\phi(x_k), p_k\rangle)} \ge \frac{\langle \nabla\phi(x_k), p_k\rangle}{\delta_2 \langle \nabla\phi(x_k), p_k\rangle} - \frac{\alpha}{2} \frac{M \|p_k\|^2}{\delta_1 \langle \nabla\phi(x_k), p_k\rangle} \ge$$

$$\geq \frac{1}{\delta_2} - \frac{\alpha M}{2\delta_1 m_1} \geq 1 . \tag{4.34}$$

Iz (4.34) sledi da je (4.33), a time i (4.24) zadovoljeno ako je

$$0 < \alpha \leq \frac{2\delta_1 m_1(1-\delta_2)}{M\delta_2} . \tag{4.35}$$

Iz (4.35) sledi da postoji konstanta $\alpha^* \in (0, \frac{2\delta_1 m_1(1-\delta_2)}{M\delta_2})$ takva, da će počev od nekog k biti zadovoljena nejednačina (4.34) ukoliko je

$$\frac{2\delta_1 m_1(1-\delta_2)}{M\delta_2} \geq \alpha_k \geq \alpha^*.$$

Prema tome, postoji $\alpha^* > 0$ takvo da je $\alpha_k \geq \alpha^*$ za dovoljno veliko k. Na osnovu ovoga iz relacije (4.29) sledi

$$\|p_k\| = \frac{1}{\alpha_k} \|x_k - x_{k+1}\| \leq \frac{1}{\alpha^*} \|x_k - x_{k+1}\| \to 0, \quad k\to\infty . \tag{4.36}$$

Dalje, iz (4.36) sledi, na osnovu relacije (4.12)

$$\|\nabla\phi(x_k)\| = \|A_k p_k\| \leq M_1 \|p_k\| \to 0, \quad k\to\infty . \tag{4.37}$$

Iz (4.37) zaključujemo kao i u dokazu teoreme 4.1. da niz $\{x_k\}$ konvergira ka jedinstvenom optimalnom rešenju problema (4.1).

Ostalo je još da ocenimo brzinu konvergencije.

Iz relacije (4.24) sledi na osnovu definicije funkcije α i relacije (4.12):

$$\phi(x_k) - \phi(x_{k+1}) \geq \alpha_k\, \alpha(<\nabla\phi(x_k), p_k>) \geq \alpha_k \delta_1 <\nabla\phi(x_k), p_k> =$$

$$= \alpha_k \delta_1 <\nabla\phi(x_k), A_k^{-1}\nabla\phi(x_k)> \geq \frac{\alpha^*\delta_1}{M_1} \|\nabla\phi(x_k)\|^2 , \text{ tj.}$$

$$\phi(x_k) - \phi(x_{k+1}) \geq \eta \|\nabla\phi(x_k)\|^2 \text{ za } \forall k \geq k_0, \quad \eta = \frac{\alpha^*\delta_1}{M_1} > 0. \tag{4.38}$$

Iz (4.38) sledi da su zadovoljene predpostavke teoreme 1.21., a odatle sledi da je brzina konvergencije u najmanju ruku R-linearna.

4.3. III modifikacija metode Danilin-Pšeničnog

Pretpostavimo da je funkcija $\phi: R^n \to R$ dva puta neprekidno diferencijabilna i da zadovoljava uslov (4.6). Na osnovu leme 1.10. sledi da je skup L kompaktan i da važi relacija (4.8). Iz (4.8), na osnovu Lagrange-ove teoreme sledi

$$\|\nabla\phi(x)-\nabla\phi(y)\| \leq M\|x-y\| \quad za \quad \forall\, x,y \in L,$$

tj. da $\nabla\phi$ *zadovoljava Lipsitz-ov uslov na* L *sa Lipsitz-ovom konstantom* M.

Prema tome, možemo koristiti poopšteni algoritam Ostrowskog za definisanje koraka α_k.

Neka je niz $\{x_k\}$ *definisan relacijama (4.2), (4.3), neka je zadovoljen uslov (4.4), a korak* α_k *definišemo na sledeći način*

$$\alpha_k = \frac{1}{M\|p_k\|}\, d\left(\frac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|}\right), \tag{4.39}$$

gde je $d: [0,\infty)\to[0,\infty)$ *prinudna funkcija takva da je* $\delta_1 t \leq d(t) \leq \delta_2 t$, $0<\delta_1<\delta_2$, $0<\delta_2<2$ *za* $\forall t \geqslant 0$.

Teorema 4.5. *Neka su zadovoljene pretpostavke teoreme 4.1., s tim što je korak* α_k *definisan relacijom (4.39). Tada važe tvrdjenja 1., 2. i 3. teoreme 4.1. i tvrdjenje teoreme 4.2.*

Dokaz. *Pretpostavimo da je* $\nabla\phi(x_k)\neq 0$. *Pošto su zadovoljene pretpostavke teoreme 3.5, sledi da važe tvrdjenja teoreme 3.5:* $x_k \in L$, $\phi(x_k)\geqslant\phi(x_{k+1})$,

$$\|x_k-x_{k+1}\|\to 0,\ k\to\infty,\quad \frac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|}\to 0,\ k\to\infty.$$

Iz ovih relacija, analogno dokazu teoreme 4.1 sledi da $\{x_k\}$ *teži ka jedinstvenom optimalnom rešenju problema (4.1).*

Dalje, iz relacije (3.28) sledi da važi

$$\phi(x_k)-\phi(x_{k+1}) \geqslant \frac{2-\delta_2}{2M\delta_2}\left[d\left(\frac{\langle\nabla\phi(x_k),p_k\rangle}{\|p_k\|}\right)\right]^2. \tag{4.40}$$

Iz (4.40), pošto je $d(t)\geqslant\delta_1 t$, *sledi*

$$\phi(x_k)-\phi(x_{k+1}) \geqslant \frac{2-\delta_2}{2M\delta_2}\,\delta_1^2\,\frac{\langle\nabla\phi(x_k),p_k\rangle^2}{\|p_k\|^2}. \tag{4.41}$$

Iz (4.41) analogno dokazu teoreme 4.2 sledi da važi ocena

$$\phi(x_k)-\phi(x_{k+1}) \geqslant \eta\,\|\nabla\phi(x_k)\|^2,\quad \eta>0,\ za\ \forall k\geqslant k_0. \tag{4.42}$$

Iz (4.42) sledi, na osnovu teoreme 1.21, da je konvergencija niza $\{x_k\}$ *R-linearna.*

4.4. IV modifikacija metode Danilin-Pšeničnog

Neka niz $\{x_k\}$ *zadovoljava relacije (4.2), (4.3) i (4.4), a korak* α_k *nadjimo pomoću sledećeg algoritma.*

Algoritam

Korak 1. *Uzmimo* $\alpha = \bar{\alpha}_k = \min\{\dfrac{d(<\nabla\phi(x_k),p_k>)}{\|p_k\|^3}, 1\}$, *pri čemu je*

$d: [0,\infty) \to [0,\infty)$ *funkcija takva da je* $\delta_1 t \leq d(t) \leq \delta_2 t$ *za* $\forall t \geq 0$, $0<\delta_1<\delta_2$

(d je očigledno prinudna funkcija).

Korak 2. *Ispitujemo da li je zadovoljena nejednakost*

$$\phi(x_k) - \phi(x_k-\alpha p_k) \geq \varepsilon\alpha^2 d(<\nabla\phi(x_k),p_k>), \quad 0<\varepsilon<1/2. \tag{4.43}$$

Korak 3. *Ukoliko je relacija (4.43) zadovoljena,* $\alpha_k = \bar{\alpha}_k$; *inače, umanjujemo* $\bar{\alpha}_k$ *sve dotle dok relacija (4.43) nije zadovoljena.*

U sledećoj teoremi pokazaćemo da niz $\{x_k\}$ *pod odredjenim uslovima konvergira ka optimalnom rešenju problema (4.1).*

Teorema 4.6. *Neka su zadovoljene pretpostavke teoreme 4.1, s tim što je korak* α_k *definisan gornjim algoritmom. Tada važe tvrdjenja 1., 2. i 3. teoreme 4.1.*

Ako je uz to $\delta_1 > \dfrac{\delta_2}{2} + \varepsilon\delta_2^2$, *niz* $\{x_k\}$ *konvergira super-linearno ka* x^*.

Dokaz. *Pošto ćemo u dokazu koristiti lemu 1.17, pre svega pokažimo da pod uslovima teoreme niz* $\{x_k\}$ *zadovoljava pretpostavke leme 1.17.*

Dokaz izvodimo metodom matematičke indukcije. Pretpostavimo da $x_k \in L$ *i da je* $\nabla\phi(x_k) \neq 0$.

Iz uslova (4.4) sledi da važi

$$<\nabla\phi(x_k),p_k> = <\nabla\phi(x_k),A_k^{-1}\nabla\phi(x_k)> > 0 \quad \text{za } \forall k \geq n-1. \tag{4.44}$$

Dalje, na osnovu Taylor-ove teoreme imamo za $\alpha > 0$:

$$\phi(x_k) - \phi(x_k-\alpha p_k) = \alpha<\nabla\phi(x_k),p_k> - \frac{\alpha^2}{2}<H(\xi_k)p_k,p_k> =$$

$$(\xi_k = x_k - \theta_k\alpha p_k, \quad \theta_k \in (0,1))$$

$$= \alpha d(<\nabla\phi(x_k),p_k>)\left[\frac{<\nabla\phi(x_k),p_k>}{d(<\nabla\phi(x_k),p_k>)} - \frac{\alpha}{2}\,\frac{<H(\xi_k)p_k,p_k>}{d(<\nabla\phi(x_k),p_k>)}\right].$$

Iz (4.44) sledi da je $d(<\nabla\phi(x_k),p_k>) > 0$. *Prema tome, nejednakost (4.43) će biti zadovoljena ako je*

$$\frac{\langle\nabla\phi(x_k),p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} - \frac{\alpha}{2}\frac{\langle H(\xi_k)p_k,p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} \geq \alpha\varepsilon \qquad \text{, odnosno,}$$

(4.45) $$\frac{1}{\alpha}\frac{\langle\nabla\phi(x_k),p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} - \frac{1}{2}\frac{\langle H(\xi_k)p_k,p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} \geq \varepsilon \qquad ,0<\varepsilon<\frac{1}{2}\,.$$

Kako je $\frac{\langle\nabla\phi(x_k),p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} > 0$ *i, zbog (4.6),* $\frac{\langle H(\xi_k)p_k,p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} > 0$, *za neko* $\alpha=\alpha_k$ *će biti zadovoljena nejednačina (4.45), a time i relacija (4.43).*
Time smo pokazali da je korak α_k *dobro definisan.*

Iz (4.43) sada sledi da je $\phi(x_{k+1}) < \phi(x_k)$, *a odavde da* $x_{k+1}\in L$. *Kako je* ϕ *ograničena odozdo na konveksnom i kompaktnom skupu L, važiće* $\phi(x_k)-\phi(x_{k+1})\to 0$, $k\to\infty$. *Na osnovu ovoga, iz relacije (4.43) sledi*

(4.46) $$\alpha_k^2\, d(\langle\nabla\phi(x_k),p_k\rangle)\to 0,\ k\to\infty\,.$$

Dalje, kako je $\alpha_k \leq \bar{\alpha}_k \leq \frac{d(\langle\nabla\phi(x_k),p_k\rangle)}{||p_k||^3}$, *biće* $d(\langle\nabla\phi(x_k),p_k\rangle) \geq \alpha_k ||p_k||^3$.

Odavde, na osnovu (4.46) sledi

(4.47) $$\alpha_k^3\ ||p_k||^3 \leq \alpha_k^2 d(\langle\nabla\phi(x_k),p_k\rangle)\to 0, \qquad k\to\infty\,.$$

Iz (4.47) sledi $||x_{k+1}-x_k||=\alpha_k||p_k||\to 0$, $k\to\infty$. *Prema tome, niz* $\{x_k\}$ *zadovoljava predpostavke lene 1.17, odakle sledi da važi relacija (4.30). Odatle dalje sledi da važi relacija (4.31), kao i:*

(4.48) $$m_1||p_k||^2 \leq \langle A_k p_k,p_k\rangle = \langle\nabla\phi(x_k),p_k\rangle \leq M_1||p_k||^2, \quad \forall k\geq k_0.$$

Iz relacije (4.48) sledi, da počev od ne kog k važi:

(4.49) $$\frac{d(\langle\nabla\phi(x_k),p_k\rangle)}{||p_k||^3} \geq \frac{\delta_1\langle\nabla\phi(x_k),p_k\rangle}{||p_k||^3} \geq \frac{\delta_1 m_1}{||p_k||}.$$

Iz (4.31) i ograničenosti $\nabla\phi$ *na L sledi:*

$$||p_k|| = ||A_k^{-1}\ \nabla\phi(x_k)|| \leq \frac{1}{m_1}K \qquad \text{za}\ \forall k\geq n-1.$$

Iz (4.48) i (4.49) dobijamo:

$$(4.50)\qquad \frac{d(\langle\nabla\phi(x_k),p_k\rangle)}{||p_k||^3} \geq \frac{\delta_1 m_1}{m_2 K} = \bar{\alpha} > 0.$$

Pošto je $\bar{\alpha}_k = \min\{ \dfrac{d(\langle\nabla\phi(x_k),p_k\rangle)}{||p_k||^3}, 1\}$ iz (4.50) sledi da će za dovoljno veliko k važiti $\bar{\alpha}_k \geq \alpha > 0$.

Dalje, imajući u vidu da je $\delta_1 t \leq d(t) \leq \delta_2 t$, kao i relaciju (4.8), vidimo da će relacija (4.45) biti zadovoljena ako je zadovoljena sledeća relacija:

$$\frac{1}{\alpha}\,\frac{\langle\nabla\phi(x_k),p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} - \frac{1}{2}\,\frac{\langle H(\xi_k)p_k,p_k\rangle}{d(\langle\nabla\phi(x_k),p_k\rangle)} \geq \frac{1}{\alpha}\,\frac{\langle\nabla\phi(x_k),p_k\rangle}{\delta_2\langle\nabla\phi(x_k),p_k\rangle} - \frac{1}{2}\,\frac{M||p_k||^2}{\delta_1\langle\nabla\phi(x_k),p_k\rangle} \geq$$

$$(4.51)\qquad \geq \frac{1}{\alpha\delta_2} - \frac{M}{2\,\delta_1}\,\frac{||p_k||^2}{m_1||p_k||^2} = \frac{1}{\alpha\delta_2} - \frac{M}{2m_1\delta_1} \geq \varepsilon.$$

Iz (4.51) sledi

$$(4.52)\qquad 0<\alpha \leq \frac{2\delta_1 m_1}{\delta_2(2\delta_1 m_1\varepsilon+M)}.$$

Iz (4.52) sledi da postoji konstanta $0<\alpha^* < \dfrac{2\delta_1 m_1}{\delta_2(2\delta_1 m_1\varepsilon+M)}$ takva da će, počev od nekog k biti zadovoljena nejednačina (4.51) ukoliko je

$\dfrac{2\delta_1 m_1}{\delta_2\,(2\delta_1 m_1\varepsilon+M)}\ \alpha_k \geq \alpha^*$. Prema tome, postoji $\alpha^* > 0$ takvo da je $\alpha_k \geq \alpha^*$ za dovoljno veliko k.

Iz činjenice da je $\bar{\alpha}_k \geq \alpha$ i $\alpha_k \geq \alpha^*$ za dovoljno veliko k sledi da postoji konstanta $C > 0$ takva da počev od nekog k važi $\alpha_k \geq C > 0$. Odavde sada sledi:

$$(4.53)\qquad ||p_k|| = \frac{1}{\alpha_k}||x_{k+1} - x_k|| \leq \frac{||x_{k+1}-x_k||}{C} \to 0, k\to\infty .$$

Iz relacijâ (4.30) i (4.53) dalje sledi

(4.54) $||\nabla\phi(x_k)||=||A_k p_k||\leq M_1||p_k||\to 0,\ k\to\infty$.

Iz stroge konveksnosti ϕ, *iz činjenice da* $\{x_k\}\subset L$ *i iz relacije* (4.54) *sledi da niz* $\{x_k\}$ *konvergira ka jedinstvenom optimalnom rešenju* (4.1). *Pre nego što predjeno na ocenu brzine konvergencije, pokažimo da je, počev od nekog* k, $\alpha_k = 1$.

Iz relacija (4.50) *i* (4.53) *sledi* $\dfrac{d(<\nabla\phi(x_k),p_k>)}{||p_k||^3}\geq\dfrac{\delta_1 m_1}{||p_k||}\to\infty$, $k\to\infty$.

Odavde i iz definicije $\bar{\alpha}_k$ *sledi, da je počev od negok* k, $\bar{\alpha}_k = 1$. *Takodje vidimo da je za dovoljno veliko* k *relaci ja* (4.51), *a time i* (4.43) *i* (4.45) *zadovoljena ako je* $\alpha_k=1$. *Naime, za* $\alpha=\alpha_k=1$, *ocenimo izraz iz relacije* (4.51):

$\dfrac{1}{\delta_2}-\dfrac{M}{2m_1\delta_1}>\dfrac{2m_1\delta_1-M\delta_2}{2m_1\delta_2^2}>\varepsilon$, *a odavde sledi* $\delta_1>\dfrac{\delta_2}{2}+\varepsilon\,\delta_2^2$, *što je po pretpostavci zadovoljeno. Pošto je* $\alpha_k=1$ *za dovoljno veliko* k *važiće sledeća relacija:*

(4.55) $||x_{k+1}-x_k||^2=<x_{k+1}-x^*,\ x_{k+1}-x^*>=<x_k-x^*-A_k^{-1}\nabla\phi(x_k),x_{k+1}-x^*>$.

Na osnovu Lagrange-ove teoreme važi sledeće:

(456) $<A_k^{-1}\nabla\phi(x_k),x_{k+1}-x^*>=<A_k^{-1}(\nabla\phi(x_k)-\nabla\phi(x^*)),\ x_{k+1}-x^*>=$

$=<A_k^{-1}H(\eta_k)(x_k-x^*),x_{k+1}-x^*>,\qquad \eta_k=x_k+\theta(x_k-x^*),\quad \theta\in(0,1)$.

Iz (4.55) *i* (4.56) *sada sledi*

$$||x_{k+1}-x^*||^2=<(I-A_k^{-1}H(\eta_k))(x_k-x^*),x_{k+1}-x^*>=$$
$$=<(A_k^{-1}(A_k-H(\eta_k))(x_k-x^*),x_{k+1}-x^*>\leq$$
$$\leq||A_k^{-1}||\,||A_k-H(\eta_k)||\,||x_k-x^*||\ \ ||x_{k+1}-x^*||\leq$$
$$\leq\frac{1}{m_1}\ ||A_k-H(\eta_k)||\ ||x_k-x^*||\ ||x_{k+1}-x^*||.$$

Odavde dalje sledi:

$||x_{k+1}-x^*||\leq\gamma_k||x_k-x^*||$, *gde je* $\gamma_k=\dfrac{1}{m_1}\,||A_k-H(\eta_k)||$.

Na osnovu leme 1.21 i ravnomerne neprekidnosti $H(x)$ *na* L *imamo:*

$$||A_k - H(\eta_k)|| \leq ||A_k - H(x_k)|| + ||H(x_k) - H(\eta_k)|| \to 0, k \to \infty ,$$

odakle sledi da $\gamma_k = \frac{1}{m_1} \; ||A_{k_*} - H(\eta_k)|| \to 0,\ k \to \infty$ *, odakle sledi da niz* $\{x_k\}$ *konvergira superlinearno ka* x^*.

Primedba. *Pošto algoritmi izloženi u radovima Danilin-a /12/ i Bulanij-Danilin-a /6/ predstavljaju samo specijalne slučajeve algoritma izloženog u /13/, izložene četiri modifikacije važe i za ove specijalne slučajeve.*

5. POGLAVLJE

5.1. Projektivne metode

Razmatramo sledeći problem: naći

(5.1). $\min\ \{\phi(x) \mid x \in X\}$, gde je $X=\{\ x \in R^n \mid A^T x=b\}$,

a A - $n \times m$ matrica, $m \leq n$.

Pretpostavićemo da je rang$A$ $=m$

Označimo sa Z- $n \times (n-m)$ matricu koja zadovoljava uslov

(5.2) $A^T Z=0$.

(O načinima izračunavanja matrice Z pogledati u /22/, str. :+, /23/, str. 52}, /26/, str. :: /. Pretpostavićemo da kolone matrice Z, vektori z_i, $i=1,\ldots,$ $n-m$ zadovoljavaju uslov

(5.3) $\|z_i\| = 1, \qquad i=1,\ldots, n-m.$

Iz izlaganja u odeljku 2.1 sledi da vektor z_i, $i=1,\ldots, n-m$ predstavlja i-ti bazisni vektor $(n-m)$ - dimenzionog prostora paralelnog linearnoj mnogostrukosti X definisanoj presekom od m linearno nezavisnih hiperravni. Dakle je rang$Z$ $=n-m$.

Projektivna Powell-Zangwill-ova metoda

Staviti $s_i^o = z_i$, gde je z_i, $i=1,\ldots, n-m$ i-ta kolona matrice Z, definisane uslovom (5.2), takvi da važi (5.3).

Algoritam *Skalar $0 < \varepsilon \leq 1$ je zadat.*

Korak 0. *Naći tačku $x_o^o \in X$ takvu da je nivoski skup $L=\{x \in X \mid \phi(x) \leq \phi(x_o^o)\}$ ograničen. Staviti $\delta^o=1$. Staviti $k=0$.*

Korak 1. *Za $i=1, \ldots, n-m$ izračunati $\alpha_i^k \in R$ tako da minimizira*

$$\phi(x_{i-1} + \alpha s_i^k) \text{ i staviti } x_i^k = x_{i-1} + \alpha_i^k s_i^k$$

Korak 2. *Izračunati* $\lambda^k = \| x_{n-m}^k - x_0^k \|$. *Ako je* $\lambda^k \leq \varepsilon$, *STOP; inače izračunati* $s_{n-m+1}^k = \dfrac{x_{n-m}^k - x_0^k}{\lambda^k}$ *i* α_{n-m+1}^k *tako da minimizira*

$\phi(x_{n-m}^k + \alpha s_{n-m+1}^k)$ *i staviti* $x_0^{k+1} = x_{n-m+1}^k = x_{n-m}^k + \alpha_{n-m+1}^k s_{n-m+1}^k$.

Korak 3. *Naći* $\alpha_p^k = \max\{\alpha_i^k,\ i=1,\ldots,n-m\}$.

a) Ako je $\dfrac{\alpha_p^k \delta^k}{\lambda^k} \geq \varepsilon$, *staviti* $s_p^{k+1} = s_{n-m+1}^k$ *i* $s_i^{k+1} = s_i^k$ *za* $i \neq p$, $i = 1,\ldots,n-m$. *Staviti* $\delta^{k+1} = \dfrac{\alpha_p^k \delta^k}{\lambda^k}$.

b) Ako je $\dfrac{\alpha^k \delta^k}{\lambda^k} < \varepsilon$, *staviti* $s_i^{k+1} = s_i^k$, $i=1,\ldots,n-m$ *i staviti* $\delta^{k+1} = \delta^k$. *Zameniti* k *sa* $k+1$ *i ići na korak 2.*

Kao što se vidi, algoritam projektivne Powell-Zangwill-ove metode je u potpunosti analogan originalnom Powell-Zangwill-ovom algoritmu (odeljak 2.1), što je prirodno jer smo uvodjenjem matrice (odnosno prostora) Z zadatak (5.1) sveli na zadatak bezuslovne optimizacije u $n-m$ *- dimenzionom prostoru.*

Teorema 5.1. *Neka je* $\phi: R^n \to R$ *neprekidno diferencijabilna strogo konveksna funkcija i neka je* rang$A=m$. *Tada niz* $\{x_i^k\}$, $i=0, 1,\ldots, n-m$, $k=0, 1,\ldots$ *generisan projektivnim Powell-Zangwill-ovim algoritmom konvergira ka optimalnom rešenju problema (5.1).*

Dokaz. *Egzistencija i jedinstvenost optimalnog rešenja* x^* *problema (5.1) sledi iz kompaktnosti i konveksnosti skupa L i stroge konveksnosti funkcije* ϕ. *Naime, svaka tačka niza* $\{x_i^k\}$ *pripada skupu L jer je po konstrukciji*

$\phi(x_i^k) \leq \phi(x_{i-1}^k)$, $i=0, 1,\ldots, n-m$ *i* $\phi(x_{n-m+1}^k) \leq \phi(x_{n-m}^k)$ *i* $x_i^k \in X$ *za* $i=0,1,\ldots,$ $n-m$, $k=0,1,2,\ldots$

Dokaz da niz $\{x_i^k\}$ *teži ka takvoj tački* $x^* \in X$ *u kojoj važi*

(5.4) $$\langle \nabla\phi(x^*), s_i \rangle = 0, \quad i=1,\dots, n-m,$$

gde $x_i^k \to x_i$, $s_i^k \to s_i$, $\|s_i\| = 1$, $k \in K \subset \{0,1,2,\dots\}$, $k\to\infty$ *i vektori* s_i, $i=1,\dots,$ $n-m$ *su linearno nezavisni, se izvodi na potpuno isti način kao u bezuslovnom slučaju (videti npr u /62/ str.294.).*

Iz (5.4) sada sledi da je gradijent $\nabla\phi(x^*)$ *ortogonalan na* $n-m$ *bazisnih vektora* $n-m$ *- dimenzionog prostora paralelnog mnogostrukosti* X*, dakle,* $\nabla\phi(x^*)$ *je ortogonalan na* X*. Dakle, u tački* x^* *je ispunjen potreban i (zbog konveksnosti) dovoljan uslov za minimum* ϕ *na* X *(teorema 2.1). Kako je zbog stroge konvesnosti tačka* x^* *jedinstvena, sledi da ceo niz* $\{x_i^k\}$ *konvergira ka* x^*.

Projektivna metoda lokalnih varijacija

Uvedimo oznake: $d_1=z_1$, $d_2=-z_1,\dots,d_{2(n-m)-1}=z_{n-m}$, $d_{2(n-m)}=-z_{n-m}$.

Algoritam

Korak 0. *Naći takvu tačku* $x_0 \in X$ *da je skup* $L=\{x\in X \mid \phi(x)\le \phi(x_0)\}$ *ograničen. Izabrati* $\varepsilon_0 > 0$.

Korak 1. *Staviti* $k=0$, $x=x_0$ *i izračunati* $\phi(x)$.

Korak 2. *Staviti* $\varepsilon = \varepsilon_k$.

Korak 3. *Staviti* $j=1$.

Korak 4. *Izračunati* $\phi(x+\varepsilon d_j)$

Korak 5. *Ako je* $\phi(x+\varepsilon d_j) < \phi(x)$; *zameniti* x *sa* $x+\varepsilon d_j$ *i ići na korak 3; inače ići na korak 6.*

Korak 6. *Ako je* $j < 2(n-m)$, *zameniti* j *sa* $j+1$ *i ići na korak 4; inače ići na korak 7.*

Korak 7. *Staviti* $x_{k+1}=x$, *staviti* $\varepsilon_{k+1}=\frac{\varepsilon}{2}$. *Ako je* $\varepsilon_{k+1}=0$, *ići na korak 8; Inače zameniti* k *sa* $k+1$ *i ići na korak 2.*

Korak 8. *Staviti* $x_{opt}=x_{k+1}$; *STOP.*

I ovde je analogija sa bezuslovnim algoritmom prirodna, kao i u prethodnom slučaju.

Teorema 5.2. *Neka je* $\phi: R^n \to R$ *neprekidno diferencijabilna funkcija i neka je rang*$A=m$. *Tada svaka tačka nagomilavanja* x^* *niza* $\{x_k\}$ *generisanog projektivnim algoritmom lokalnih varijacija zadovoljava uslov* $P(x^*)\nabla\phi(x^*)=0$, *gde je* P *- operator ortogonalnog projektovanja iz* R^n *u ortogonalni komplement prostora koji obrazuju kolone* $a_i, i=1,\dots,m$ *matrice* A, *tj. u tački* x^* *je zadovoljen potreban uslov za minimum funkcije* ϕ *na skupu* X.

Dokaz. *Iz neprekidnosti* ϕ *i ograničenosti* L *(korak 0) sledi kompaktnost skupa* L. *Dakle, počevši proces od* $x=x_k$ *i* $\varepsilon=\varepsilon_k$ *algoritam može da konstruiše samo konačan broj tačaka u ciklusu izmedju koraka 3. i 6. Dokaz da važi*

$$\langle\nabla\phi(x^*), z_j\rangle = 0, \quad \text{za } j=1,2,\dots,n-m, \tag{5.5}$$

gde $x_k \to x^*, k \in K \subset \{0,1,2,\dots\}$, $k\to\infty$, *izvodi se analogno dokazu za bezuslovni slučaj (videti npr. Polak /41/, str. 63).*

Iz relacije (5.5) vidimo da je gradijent $\nabla\phi(x^*)$ *normalan na* $n-m$ *bazisnih vektora koji razapinju linearnu mnogostrukost* X. *Odatle sledi da je ortogonalna projekcija gradijenta* $\nabla\phi(x^*)$ *na* X *nula, a to je i trebalo dokazati.*

Pošto je u slučaju kada su linearna ograničenja data u obliku nejednakosti optimizacioni algoritam znatno složeniji, izložićemo projektivnu varijantu samo ciklične koordinatne metode. Na analogan način bi se izložene projektivne varijante Powell-Zangwill-ove metode i metode lokalnih varijacija formulisane za slučaj ograničenja u obliku jednakosti, adaptirale za slučaj ograničenja u obliku linearnih nejednakosti.

Projektivna ciklično-koordinatna metoda

Razmatramo sledeći problem: naći

$$\min\ \{\phi(x) \mid x\in X\}\ , \quad X=\{\, x\in R^n \mid A^T x \le b\}, \tag{5.6}$$

gde je $A=(a_j)$, $j\in\{1,\dots,m\}$ *-* $n\times m$ *matrica ograničenja. Neka je*

$I(x_i^k)=\{\, j\in\{1,\dots,m\} \mid \langle a_j, x_i^k\rangle = b_j\} = \{\, i_1,\dots,i_q\}\ ,\ i_1< i_2< \dots <i_q.$

Pretpostavićemo:

(5.7) da je $X \neq \emptyset$ i da postoji takva tačka $\bar{x} \in X$ da je skup $L=\{ x \in X \mid \phi(x) \leq \phi(\bar{x})\}$ ograničen;

(5.8) da su vektori a_j, $j \in I(x_i^k)$ linearno nezavisni za $\forall k \circ i\ \forall i$;

(5.9) da je funkcija $\phi: R^n \to R$ konveksna i neprekidno diferencijabilna.

Algoritam

Korak 0 Naći takvu tačku $x_o^1 \in X$ da je skup $L=\{ x \in X \mid \phi(x) \leq \phi(x_o^1)\}$ ograničen. Stavimo $I=I(x_o^1)$. Neka je kard $I=q$ i neka su $e_i, i=1,\dots,n-q$ jedinični koordinatni vektori iz prostora R^{n-q}. Ako je $q=0$, staviti $Z_I=E$, gde je E - jedinična $n \times n$ matrica, inače naći matricu Z_I iz uslova $A_I^T Z_I=0$, pri čemu je $A_I=(a_i)$, $i \in I$. Staviti $k=1$.

Korak 1. Staviti $i=1$.

Korak 2. Izračunati α_i^k iz uslova

$\phi(x_{i-1}^k + \alpha_i^k Z_I e_i)=\min\{ \phi(x_{i-1}^k + \alpha Z_I e_i) \mid x_{i-1}^k + \alpha Z_I e_i \in X\}$. Staviti $x_i^k=x_{i-1}^k + \alpha_i^k Z_I e_i$. Ako je $I(x_i^k) \supset I(x_{i-1}^k)$, ići na korak 3; inače ići na korak 4.

Korak 3. Staviti $x_o^{k+1} = x_i^k$, zameniti k sa $k+1$, q sa kard $I(x_i^k)$, I sa $I(x_i^k)$, izračunati Z_I i ići na korak 1.

Korak 4. Ako je $i=n-q$, ići na korak 5; inače zameniti i sa $i+1$ i ići na korak 2.

Korak 5. Naći $\alpha_s^k = \max\{ |\alpha_i^k|\ ,\ i=1,2,\dots, n-q\}$. Ako je $\alpha_s^k=0$, ići na korak 6; Inače staviti $x_o^{k+1}=x_{n-q}^k$, zameniti k sa $k+1$ i ići na korak 1.

Korak 6. Ako je $q \neq 0$, ići na korak 7; inače staviti $x=x_o^{k+1}$ i ići na krak 14.

Korak 7. Staviti $j=1$.

Korak 8. *Zameniti* I *sa* $I \setminus \{i_j\}$. *Staviti* $x_0^{k+1}=x_{n-q}^k$, *zameniti* k *sa* $k+1$, q *sa* $q-1$. *Izračunati* Z_I. *Staviti* $i=1$.

Korak 9. *Izračunati* α_i^k *iz uslova*

$$\phi(x_{i-1}^k + \alpha_i^k Z_I e_i) = \min \{\phi(x_{i-1}^k + \alpha Z_I e_i) \mid x_{i-1}^k + \alpha Z_I e_i \in X\}.$$

Staviti $x_i^k = x_{i-1}^k + \alpha_i^k Z_I e_i$. *Ako je* $I(x_i^k) \supset I(x_{i-1}^k)$, *ići na korak 11; inače ići na korak 10.*

Korak 10. *Ako je* $i=n-q$, *ići na korak 12; inače zameniti* i *sa* $i+1$ *i ići na korak 9.*

Korak 11. *Ako je* $\phi(x_i^k) \geq \phi(x_0^k)$, *ići na korak 13; inače ići na korak 3.*

Korak 12. *Ako je* $\phi(x_{n-q}^k) \gtrsim \phi(x_0^k)$, *ići na korak 13; inače staviti* $x_0^{k+1}=x_{n-q}^k$, *zameniti* k *sa* $k+1$ *i ići na korak 1.*

Korak 13. *Zameniti* I *sa* $I \cup \{i_j\}$, q *sa* $q+1$, j *sa* $j+1$. *Staviti* $x=x_0^k$. *Ako je* $j > q$, *ići na korak 14; inače ići na korak 7.*

Korak 14. $x_{opt} = x$; *STOP.*

Teorema 5.3. *Pretpostavimo da su zadovoljeni uslovi (5.7), (5.8 i (5.9). Tada izloženi algoritam svodi rešavanje problema (5.6) na rešavanje konačnog broja podproblema minimizacije funkcije cilja* ϕ *na linearnim mnogostrukostima. Pri tom je u svakoj tački nagomilavanja niza* $\{x_i^k\}$ *generisanog algoritmom projekcija gradijenta nula.*

Dokaz. *Vidimo da na k-toj iteraciji mogu nastupiti dva slučaja:*

Slučaj 1: $I(x_i^k)=I(x_{i-1}^k)$ *za* $i=1,\ldots,n-q$, *gde je kard* $I(x_i^k)=q$; *tj. posle minimizacje funkcije cilja* ϕ *duž svih bazisnih pravaca ostajemo u mnogostrukosti koja je definisana skupom indeksa* $I(x_0^k)=I=\{i_1,\ldots,i_q\}$;

Slučaj 2: $I(x_i^k) \supset I(x_{i-1}^k)$, *kada za neko* $i \in \{1,\ldots,n-q\}$ *dolazimo na granicu linearne mnogostrukosti koja je definisana skupom indeksa* $I(x_{i-1}^k)$.

Razmotrimo slučaj 1.

Kada se realizuje slučaj 1, ispitujemo na koraku 5 algoritma da li smo našli tačku minimuma na posmatranoj mnogostrukosti, tj. da li je

$\max\limits_{i=1,\dots,n-q} |\alpha_i^k| = 0$. Ako je $\max\limits_{i=1,\dots,n-q} |\alpha_i^k| \neq 0$, stavljamo $x_0^{k+1} = x_{n-q}^k$ i proces

traženja minimuma na ovoj mnogostrukosti iteriramo sve dok, eventualno na nekoj (pod) iteraciji $k+t$, $t \geq 0$, ne dobijemo tačke x_i^{k+t}, $i \in \{1,\dots, n-q\}$ za koje je ispunjen uslov $\max\limits_{i=1,\dots,n-q} |\alpha_i^{k+t}| = 0$. Tada tačka x_{n-q}^{k+t} predstavlja optimalno rešenje podproblema minimizacije funkcije ϕ na posmatranoj mnogostrukosti. Zatim, na koraku 7 odbacujemo ograničenje koje odgovara prvom indeksu i_1 da vidimo da li funkcija ϕ opada na mnogostrukosti definisanoj skupom indeksa $I \setminus \{i_1\}$, a za početnu tačku x_0^{k+1} uzimamo upravo dobijenu tačku x_{n-q}^{k+t}.

Ako funkcija ϕ na novodobijenom preseku ne opada, indeks i_1 vraćamo polaznom skupu indeksa I (korak 13) i sada odstranjujemo iz skupa aktivnih ograničenja ono ograničenje koje odgovara indeksu i_2 $(i_1 < i_2)$ i ponavaljamo prethodno ispitivanje. Ovim postupkom ćemo ili doći do zaključka da odbacivanjem bilo kojeg od q ograničenjâ koja odgovaraju skupu indeksa I ne dolazi do opadanja funkcije (slučaj kada je $j > q$ na koraku 13) tj. da je tačka x_0^k - tačka optimuma (korak 14), ili da odbacivanjem nekog ograničenja koje odgovara indeksu i_j, $j \in \{1,\dots,q\}$, dolazi do opadanja funkcije. Ukoliko se realizuje poslednja situacja, pristupamo traženju minimuma funkcije ϕ na skupu ograničenja koji odgovara skupu indeksa $\{i \mid i \in I \setminus \{i_j\}\}$, tj. vraćamo se koraku 1, odnosno koraku 3.

Pokažimo da se slučaj 1 može realizovati samo konačan broj puta. Videli smo da je u ovom slučaju, neposredno pre nego što se neko ograničenje odbaci (korak 8) nadje-na najmanja vrednost funkcije ϕ na mnogostrukosti definisanoj presekom aktivnog skupa ograničenja i zatim sledeća iteracija umanjuje vrednost funkcije ϕ. Zato se ne možemo vratiti skupu ograničenja koji je bio aktivan neposredno pre odbacivanja jednog od ograničenja. Pošto postoji samo konačan broj kombinacija skupova aktivnih ograničenja (2^q-broj podskupova skupa od q elemenata), slučaj 1 se može realizovati samo konačan broj puta.

Razmotrimo sdda slučaj 2. Tada, bar za jedno $j \notin I$ postaje aktivno novo ograničenje, tj. u tački $x_0^{k+1} = x_i^k$ skup indeksa koji odgovara skupu aktivnih ograničenja zadovoljava relaciju $I(x_0^{k+1}) \supset I(x_0^k)$. Tačka x_i^k se za neko $i \in \{1,\dots, n-q\}$ nalazi na granici mnogostrukosti definisane skupom indeksa $I = \{i_1,\dots,i_q\}$.

Na taj je način dimenzija mnogostrukosti u novoj tački x_o^{k+1} *manja od dimenzije mnogostrukosti u tački* x_o^k *jer je dodato najmanje jedno novo ograničenje.*

Kako su vektori $a_i, i \in I$, *po pretpostavci, linearno nezavisni, jasno je da se ovo proširivanje skupa aktivnih ograničenja mora prekinuti posle konačnog broja koraka koji nije veći od* n, *gde je* n *- dimenzija prostora* R^n $(X \subset R^n)$.

Prema tome, uzastopna realizacija slučaja 2 može dovesti do mnogostrukosti nulte dimenzije koja se sastoji samo iz jedne tačke koja je ujedno i globalni minimum funkcije cilja ϕ *na toj mnogostrukosti, a to je upravo slučaj 1, ili ćemo za neko* $k \geq k_o$ *ostati u istoj mnogostrukosti, a to je opet slučaj 1.*

Dakle, posle konačnog broja ponavljanja slučaja 2 mora nastupiti slučaj 1, a slučaj 1 se, kako smo gore pokazali, može takodje javiti samo konačan broj puta, samo, naravno, sám broj iteracija može biti beskonačan.

Medjutim, tačka nagomilavanja niza $\{x_i^k\}$ *generisanog gornjim algoritmom ne mora biti optimalno rešenje problema (5.6). Može se naime desiti da, ostajući u jednoj mnogostrukosti niz konvergira ka tački u kojoj je projekcija gradijenta nula.*

Primedba. *Zahtev da važi uslov (5.8) o linearnoj nezavisnosti gradijenata aktivnih ograničenja nije bitan za definisanje algoritma. Ovaj uslov se uvodi da bi se eliminisali tzv. degenerativni slučajevi koji bi zahtevali specijalna dodatna razmatranja. Ako se javi linearna zavisnost (degeneracija) kada nova ograničenja postanu aktivna, nije jednostavno odrediti linearno nezavisna ograničenja koja definišu granicu dopustive oblasti. Ako algoritam vodi u takvu tačku* x, *odgovarajućom malom izmenom vektora* b *(videti Dantzig /14/ možemo postići izdvajanje ograničenja koja definišu degenerativnu ivicu ili vrh.*

Srećom, ova pojava se vrlo retko javlja u praksi zato što greška zaokruživanja koja se javlja kod izračunavanja matrice ograničenja prirodno uključuje pomenute izmene.

5.2. Poopštenje Danilin-ovih rezultata

Danilin u /11/ razmatra sledeći problem: naći

(5.10) $\min \{\phi(x) \mid x \in X\}$, $X=\{ x \in R^n \mid f(x)=0\}$, $f=(f_1,\ldots,f_m)$, *gde su funkcije* $\phi, f_i : R^n \to R$ *neprekidno difere-ncijabilne.*

Pretpostavićemo da je u svakoj tački $x \in X$ *zadovoljen uslov*

(5.11) $\det\left(\frac{\partial f_i}{\partial x_j}\right) \neq 0, \; i,j=1,\ldots,m.$

Na osnovu teoreme o implicitnim funkcijama u okolini tačke x *sistem* $f(x)=0$ *definiše funkciju* $y=y(t)$, $y \in R^m$, $t \in R^{n-m}$, $y=(x_1,\ldots,x_m)$, $t=(x_{m+1},\ldots,x_n)$, $x=(y,t)$.

Definišimo funkciju $F: R^{n-m} \to R$ *na sledeći način:* $F(t)=\phi(y(t),t)$.

Razmatramo optimizacioni algoritam za rešavanje problema (5.10) koji generiše niz tačaka $\{x_k\}$ *sledećeg oblika:*

(5.12) $x_{k+1}=(y_{k+1},t_{k+1})$, $t_{k+1}=t_k-\alpha_k p_k$, $y_{k+1}=y(t_{k+1})$, $\|p_k\|=1$, *gde pravac* p_k *i korak* α_k *zadovolja vaju sledeće relacije*

(5.13) $\langle \nabla F(t_k), p_k \rangle \geq \rho \, \|\nabla F(t_k)\| \, \|p_k\|$, $0<\rho\leq 1$,

(5.14) $F(t_k)-F(t_k-\alpha_k p_k) \geq \alpha_k d(\langle \nabla F(t_k), p_k \rangle)$, *gde je* $d: [0,\infty) \to [0,\infty)$ *funkcija takva da važi* $\delta_1 t \leq d(t) \leq \delta_2 t$ *za* $\forall t \geq 0$, $0<\delta_1<\delta_2<1$ *(d je očigledno jedna prinudna funkcija).*

Teorema 5.4. *(Videti /11/, str. 88) Neka su funkcije* ϕ, $f_i: R^n \to R$, $i=1,\ldots,m$ *dva puta neprekidno diferencijabilne, neka funkcije* f_i *i* $i=1,\ldots,m$ *zadovoljavaju uslov (5.11) i neka je skup* $L=\{x \quad X \mid \phi(x) \quad \phi(x_0)\}$ *ograničen za p izvoljno izabranu tačku* $x_0 \quad X$. *Tada niz (5.12) ima sledeća svojstva:*

1. $\{x_k\} \subset L$;
2. $\phi(x_{k+1}) \leq \phi(x_k)$;
3. $\|\nabla F(t_k)\| \to 0$ *za* $k \to \infty$.

Dokaz. *Mogućnost konstrukcije niza (5.12) sledi iz uslova (5.11): za dovoljno male vrednosti parametra* α *tačka* t_{k+1} *se nalazi u okolini tačke* t_k *gde je definisana funkcija* $y(t_k)$. *U toj okolini je funkcija* $F(t)=\phi(t,y(t))$ *dva puta neprekidno diferencijabilna prema pretpostavkama teoreme.*

Prema tome, na osnovu Taylor-ove teoreme, važiće:

$$F(t_k)-F(t_{k+1})=\alpha_k <\nabla F(t_k),p_k>-\frac{\alpha_k^2}{2}<\nabla^2F(\xi_k)p_k,p_k>,\ \xi_k=t_k-\theta_k\alpha_k p_k,\ \theta_k\in(0,1),$$

gde je ∇^2F- *hesijan funkcije* F. *Dalje imamo:*

$$(5.15)\quad F(t_k)-F(t_{k+1})=\alpha_k d(<\nabla F(t_k),p_k>)\left[\frac{<\nabla F(t_k),p_k>}{d(<\nabla F(t_k),p_k>)}-\frac{\alpha_k}{2}\frac{<\nabla^2F(\xi_k)p_k,p_k>}{d(<\nabla F(t_k),p_k>)}\right].$$

Iz (5.15) sledi da će relacija (5.14) biti zadovoljena ako je $\alpha_k>0$ *dovoljno malo. Prema tome,* $\phi(x_k)\geq\phi(x_{k+1})$ *i* $\{x_k\}\subset L$. *Dokažimo da* $||\nabla F(t_k)||\to 0$, $k\to\infty$. *Iz neprekidne diferencijabilnosti funkcija* f_i, $i=1,\dots,m$ *na zatvorenom i ograničenom skupu* L, $\det(f_y(x))$ *je ravnomerno neprekidna funkcija na* L. *Zato, ako je apsolutna vrednost determinante u proizvoljnoj tački* $\hat{x}\in L$ *jednaka* μ, *tj.* $|\det(f_y(x))|=\mu>0$, *postoji konstanta* $\nu>0$ *takva da je u svakoj tački skupa* L *koja pripada sferi* $S_{\hat{x}}$ *radijusa* ν *(sa centrom u tački* $\hat{x}$*), apsolutna vrednost determinante* $|\det(f_y(x))|\geq\frac{\mu}{2}>0$.

Osim toga, zbog ograničenosti skupa L *i neprekidnosti prvih i drugih parcijalnih izvoda funkcija* f_i, *za svaku tačku* $\hat{x}\in L$, *ovi izvodi će biti ograničeni nekom konstantom* M *u sferi* $S_{\hat{x}}$ *radijusa* ν. *Odavde sledi, prema teoremi o implicitnim funkcijama, da u nekom paralelepipedu* $[\hat{x}_i\pm\varepsilon\hat{x}_i]$, $i=1,\dots,n$ *koji pripada sferi* $S_{\hat{x}}(\nu)$, *sistem* $f(x)=0$ *definiše bar jednu dvaput neprekidno diferencijabilnu vektor funkciju* $y(t)$, *pri čemu u tom paralelepipedu važi:*

$$(5.16)\qquad ||\nabla y(t)||\leq N_1,\quad ||\nabla^2y(t)||\leq N_2.$$

Zbog ograničenosti izvodâ $\nabla y(t)$, $\nabla^2y(t)$, *a takodje zbog ograničenosti prvi h i drugih ivzoda funkcije* ϕ *na skupu* L, *izvodi funkcije* F *u paralelepipedu* $[x_i\pm\varepsilon x_i]$ *takodje su ograničeni:* $||\nabla F(t)||\leq K_1$, $||\nabla^2F(t)||\leq K_2$.

Prema tome, postoji konstanta $\varepsilon\geq 0$ *takva da, ako je* $\alpha_k\leq\varepsilon$, *tačka* x_{k+1} *će biti u sferi* S_{x_k} *radijusa* ν. *Neka je*

$$(5.17)\qquad ||x_{k+1}-x_k||^2=||t_{k+1}-t_k||^2+||y_{k+1}-y_k||^2=\nu^2.$$

Kako odavde sledi da $x_{k+1}\in S_{x_k}$, *sledi da* x_{k+1} *pripada i paralelepipedu* $[x_i^k\pm\varepsilon x_i^k]$, $i=1,\dots,n$ *u kojem za izvode funkcije* $y(t)$ *važe ocene (5.16).*

Prema tome, $||y_{k+1}-y_k|| \leq N_3||t_{k+1}-t_k||$, *pa iz (5.17) sledi*

$(N_3^2+1)||t_{k+1}-t_k||^2 = \alpha_k^2 N_4^2 ||p_k||^2 \geq \nu^2$, *odakle,* $\alpha_k \geq \frac{\nu}{N_4||p_k||^2} = \frac{\nu}{N_4}$.

Odavde vidimo da jednakost $||x_{k+1}-x_k|| = \nu$ *važi za vrednosti parametra* $\alpha_k \geq \frac{\nu}{N_4}$, *tj. za* ε *se može uzeti bilo koja konstanta koja nije veća od* $\frac{\nu}{N_4}$

Sada, korišćenjem relacije (5.15), imajući u vidu osobine funkcije d kao i činjenicu da je $\nabla^2 F$ *ograničeno u paralelepipedu* $[x_i^k \pm \varepsilon x_i^k]$, *zaključujemo da će relacija (5.14) biti zadovoljena ako važi:*

(5.18)

$$\frac{\langle\nabla F(t_k),p_k\rangle}{d(\langle\nabla F(t_k),p_k\rangle)} - \frac{\alpha_k}{2}\frac{\langle\nabla^2 F(\xi_k)p_k,p_k\rangle}{d(\langle\nabla F(t_k),p_k\rangle)} \geq \frac{1}{\delta_2} - \frac{\alpha_k}{2\,\delta_1}\frac{K_2}{\rho||\nabla F(t_k)||} \geq 1.$$

Iz (5.18) sledi $0 < \alpha_k \leq \frac{2(1-\delta_2)\,\delta_1\rho||\nabla F(t_k)||}{M\,\delta_2} \leq \frac{2(1-\delta_2)\delta_1\rho K_1}{M\,\delta_2}$.

Prema tome, uslov (5.14) će biti zadovoljen ako je

$$\alpha_k = \min\left\{\varepsilon\,,\ \frac{2\,\delta_1(1-\delta_2)\rho K_1}{M\,\delta_2}\right\}.$$

Iz neprekidnosti ϕ *i ograničenosti skupa L sledi da je* ϕ *ograničena odozdo, iz (5.14) sledi da je* $\phi(x_{k+1}) \leq \phi(x_k)$. *Prema tome,*

(5.19) $\phi(x_{k+1}) - \phi(x_k) \to 0, \quad k\to\infty$.

Iz (5.19) i (5.14) dalje sledi $d(\langle\nabla F(t_k),p_k\rangle) \to 0,\ k\to\infty$, *a odavde, na osnovu svojstava funkcije d:*

(5.20) $\langle\nabla F(t_k),p_k\rangle \to 0, k\to\infty$.

Iz (5.20) i (5.13) dobijamo

$$||\nabla F(t_k)|| \to 0,\ k\to\infty\ ,$$

što je i trebalo dokazati.

DODATAK

Razmatramo sledeći problem: naći

$$\min \{ 10\xi_1^2 + 20\xi_2^2 - 70\xi_1 - 160\xi_2 + 442{,}5 \mid x = (\xi_1, \xi_2) \quad R^2 \}$$

sa tačnošću od $\varepsilon = 10^{-2}$. *Koristiti iteracioni algoritam oblika*

$$(1) \quad \begin{aligned} x_{k+1} &= x_k - \alpha_k p_k, \quad k = 0,1,2\ldots, \alpha_k > 0, \\ p_k &= g_k, \text{ gde je } g_k = \nabla\phi(x_k), \end{aligned}$$

pri čemu α_k *računati pomoću algoritama za nalaženje koraka izloženih u odeljku 1.3, kao i pomoću njihovih modifikacija izloženih u poglavlju 3.*

Pošto je ovde funkcija cilja ϕ *oblika*

$$\phi(x) = \langle x, Qx \rangle - \langle a, x \rangle + 442{,}5,$$

gde je

$$Q = 10\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad a = 10\begin{bmatrix} 7 \\ 10 \end{bmatrix},$$

strogo konveksna neprekidno diferencijabilna funkcija, potreban i dovoljan uslov za minimum glasi:

$$\nabla\phi(x) = 0,$$

odnosno

$$\nabla\phi(x) = 2Qx - a = 10\begin{bmatrix} 2\xi_1 - 7 \\ 4\xi_2 - 16 \end{bmatrix} = 0\,.$$

Sledi da funkcija ϕ *postiže minimalnu vrednost 0 u jedinstvenoj tački minimuma* $x^* = \begin{bmatrix} 3{,}5 \\ 4 \end{bmatrix}$.

Modifikovani Curry-Altman-ov algoritam

Za niz $\{x_k\}$ *koji zadovoljava (1),* α_k *je definisano na sledeći način:*

$\alpha_k = 0$, *ako je* $\langle \nabla\phi(x_k), p_k \rangle = 0$, *inače*

$\alpha_k \in I_k$, *gde* I_k *predstavlja prvi interval pozitivnih rešenja nejednačine*

$$(2)\quad \langle \nabla\phi(x_k - \alpha p_k),\ p_k\rangle > \sigma(\langle \nabla\phi(x_k), p_k\rangle),$$

pri čemu je $\sigma: [0,\infty)\to[0,\infty)$ *prinudna funkcija takva da zadovoljava uslov* $\sigma(t) \leq \delta t,\ 0<\delta<1$ *za* $t \geq 0$.

Kako je $p_k = g_k$, *uslov (2) će glasiti*

$$(3)\quad \langle \nabla\phi(x_k - \alpha g_k),\ g_k\rangle > \sigma(\| g_k\|^2).$$

Za prinudnu funkciju σ *možemo uzeti, na primer, funkciju* $\sigma(t) = \ln(1 + t),\ t \geq 0$, *pa će uslov (3) glasiti:*

$$(4)\quad \langle \nabla\phi(x_k - \alpha g_k),\ g_k\rangle > \ln\,(1 + \| g_k\|^2).$$

Uzmimo $x_0 = \begin{bmatrix}0\\0\end{bmatrix}$. *Tada je* $g_0 = \begin{bmatrix}-70\\-160\end{bmatrix}$, $\| g_0\|^2 = 30500$, $\nabla\phi(x_0 - \alpha g_0) = 10\begin{bmatrix}140\alpha - 7\\640\alpha - 16\end{bmatrix}$.

Uslov (4) će u tački x_0 *glasiti*

$$10\begin{bmatrix}140\alpha - 7 & 640\alpha - 16\end{bmatrix}\cdot\begin{bmatrix}-70\\-160\end{bmatrix} > \ln 30501.$$

Rešavanjem ove nejednačine dobijamo dopustivi interval za α_0: $0 < \alpha < 0{,}0271743$.

Uzmimo $\alpha_0 = 0{,}027$.

Sledeće približno rešenje će, prema (1), glasiti

$$x_1 = x_0 - \alpha_0 g_0 = \begin{bmatrix}1{,}89\\4{,}32\end{bmatrix},\qquad g_1 = \begin{bmatrix}-32{,}2\\12{,}8\end{bmatrix}.$$

Formirajmo nejednačinu (4) u tački x_1:

$$\begin{bmatrix}-32{,}2 + 644\alpha & 12{,}8 - 512\alpha\end{bmatrix}\begin{bmatrix}-32{,}2\\12{,}8\end{bmatrix} > \ln 1201{,}68.$$

Odavde sledi da je $0 < \alpha < 0{,}0437365$. *Neka je* $\alpha_1 = 0{,}043$. *Sledi*

$$x_2 = x_1 - \alpha_1 g_1 = \begin{bmatrix}3{,}2746\\3{,}7696\end{bmatrix},\quad g_2 = \begin{bmatrix}-4{,}508\\-9{,}216\end{bmatrix}.$$

Dopustivi interval za α_2 *dobijamo rešavanjem nejednačine*

$$\begin{bmatrix}-4{,}508 + 90{,}16\alpha & -9{,}216 + 368{,}64\alpha\end{bmatrix}\cdot\begin{bmatrix}-4{,}508\\-9{,}216\end{bmatrix} > \ln 106{,}25672.$$

Dobijamo $0<\alpha<0{,}0264446$; *neka je* $\alpha_2 = 0{,}026$.

Sledeće približno rešenje će, na osnovu (1), glasiti:

$$x_3 = x_2 - \alpha_2 g_2 = \begin{bmatrix}3{,}391808\\4{,}009216\end{bmatrix},\ \text{a}\ g_3 = \begin{bmatrix}-2{,}16384\\0{,}36864\end{bmatrix}.$$

Uslov (4) će u tački x_3 glasiti

$$\begin{bmatrix}-2,16384 + 43,2768\alpha & 0,36864 - 14,7456\alpha\end{bmatrix} \begin{bmatrix}-2,16384 \\ 0,36864\end{bmatrix} > \ln 5,818098,$$

odakle sledi $0<\alpha<0,0308551$. *Uzmimo* $\alpha_3 = 0,03$. *Dobijamo*

$$x_4 = x_3 - \alpha_3 g_3 = \begin{bmatrix}3,4567232 \\ 3,9981568\end{bmatrix}, \quad g_4 = \begin{bmatrix}-0,865536 \\ -0,073728\end{bmatrix}.$$

Analogno, rešavanjem nejednačine

$$\begin{bmatrix}-0,865536 + 17,31072\alpha & -0,073728 + 2,94912\alpha\end{bmatrix} \begin{bmatrix}-0,865536 \\ -0,073728\end{bmatrix} > \ln 1,7545888,$$

dobijamo $0<\alpha<0,0126544$. *Uzmimo* $\alpha_4 = 0,012$, *sledi*

$$x_5 = x_4 - \alpha_4 g_4 = \begin{bmatrix}3,4671096 \\ 3,9990415\end{bmatrix}, \quad g_5 = \begin{bmatrix}-0,657808 \\ -0,03834\end{bmatrix}.$$

Kako je $\phi(x_4) = 0,018796$, $\phi(x_5) = 0,010836$, $|\phi(x_5) - \phi(x_4)| = 0,0079598 < 0,01$, *sledi da* $\phi(x_5)$ *predstavlja približnu vrednost minimuma sa datom tačnošću* $\varepsilon = 0,01$.

Curry-Altman-ov algoritam

Za niz $\{x_k\}$ *koji je definisan sa (1),* α_k *definišemo na sledeći način*

$\alpha_k = 0$, *ako je* $\langle\nabla\phi(x_k), p_k\rangle = 0$, *inače*

α_k *predstavlja najmanje pozitivno rešenje jednačine*

$\langle\nabla\phi(x_k - \alpha p_k), p_k\rangle = \mu\langle\nabla\phi(x_k), p_k\rangle$ *za fiksirano* $0 \doteq \mu < 1$, *odnosno, za* $p_k = g_k$

$$(5) \quad \langle\nabla\phi(x_k - \alpha g_k), g_k\rangle = \mu\|g_k\|^2, \quad 0 \doteq \mu < 1.$$

Uzmimo opet za početnu tačku $x_0 = \begin{bmatrix}0 \\ 0\end{bmatrix}$. *Imamo da je*

$$g_0 = \begin{bmatrix}-70 \\ -160\end{bmatrix}, \quad \|g_0\|^2 = 30500, \quad \nabla\phi(x_0 - \alpha g_0) = 10\begin{bmatrix}140\alpha-7 \\ 640\alpha-16\end{bmatrix}.$$

Uslov (5) u tački x_0 će glasiti

$$\begin{bmatrix}1400\alpha - 70 & 6400\alpha - 160\end{bmatrix} \begin{bmatrix}-70 \\ -160\end{bmatrix} = 30500\mu.$$

Rešavanjem ove jednačine, za razne vrednosti $\mu \in [0,1)$, *dobijamo*

za $\mu = 0,2$: $\quad \alpha_0 = 0,022$,

za $\mu = 0,3$: $\quad \alpha_0 = 0,019$,

$$\text{za } \mu = 0{,}4: \qquad \alpha_0 = 0{,}016,$$
$$\text{za } \mu = 0{,}8: \qquad \alpha_0 = 0{,}05.$$

Kako smo za $\mu = 0{,}2$ *dobili najveću vrednost koraka* α_0, *možemo očekivati najbržu konvergenciju upravo za ovu vrednost* μ. *Zato ćemo ostala približna rešenja nalaziti iz uslova (5) sa* $\mu = 0{,}2$.

Prema tome, za $\mu = 0{,}2$, $\alpha_0 = 0{,}022$, *na osnovu (1) izračunaćemo*

$$x_1 = x_0 - \alpha_0 g_0 = \begin{bmatrix} 1{,}54 \\ 3{,}52 \end{bmatrix}, \quad g_1 = \begin{bmatrix} -39{,}2 \\ -19{,}2 \end{bmatrix}.$$

Uslov (5) u tački x_1 *će glasiti*

$$[-39{,}2 + 784\alpha \quad -19{,}2 + 768\alpha] \cdot \begin{bmatrix} -39{,}2 \\ -19{,}2 \end{bmatrix} = 0{,}2 \cdot 1905{,}28.$$

Odavde dobijamo $\alpha_1 = 0{,}034$ *i sledeća iteracija glasi*

$$x_2 = x_1 - \alpha_1 g_1 = \begin{bmatrix} 2{,}9828 \\ 4{,}1728 \end{bmatrix}.$$

Nalazimo $g_2 = \begin{bmatrix} -12{,}544 \\ 6{,}912 \end{bmatrix}$ *i formiramo uslov (5) za tačku* x_2

$$[-12{,}544 + 250{,}88\alpha \quad 6{,}912 - 276{,}48\alpha] \begin{bmatrix} -12{,}544 \\ 6{,}912 \end{bmatrix} = 0{,}2 \cdot 205{,}12765,$$

odakle sledi da je $\alpha_2 = 0{,}032$ *i*

$$x_3 = x_2 - \alpha_2 g_2 = \begin{bmatrix} 3{,}274208 \\ 3{,}951616 \end{bmatrix}, \quad g_3 = \begin{bmatrix} -4{,}51584 \\ -1{,}93536 \end{bmatrix}.$$

Analogno izračunavamo $\alpha_3 = 0{,}035$, *pa je*

$$x_4 = x_3 - \alpha_3 g_3 = \begin{bmatrix} 3{,}4322624 \\ 4{,}0193536 \end{bmatrix}, \quad g_4 = \begin{bmatrix} -1{,}354752 \\ 0{,}774144 \end{bmatrix}.$$

Zatim izračunavamo $\alpha_4 = 0{,}032$ *i dobijamo*

$$x_5 = x_4 - \alpha_4 g_4 = \begin{bmatrix} 3{,}4756144 \\ 3{,}9945809 \end{bmatrix}.$$

Dalje dobijamo $\phi(x_4) = 0{,}053375$, $\phi(x_5) = 0{,}0065339$. *Medjutim, kako je* $|\phi(x_5) - \phi(x_4)| = 0{,}0468411 > 0{,}01$, *približna vrednost* x_5 *ne predstavlja rešenje sa traženom tačnošću* $\varepsilon = 0{,}01$. *Zato moramo izračunati sledeću približnu vrednost. Dobijamo* $\alpha_5 = 0{,}034$, *pa je*

$$x_6 = x_5 - \alpha_5 g_5 = \begin{bmatrix} 3{,}4921966 \\ 4{,}0019508 \end{bmatrix},$$

i

$$\phi(x_6) = 0{,}000685042.$$

Pošto je $|\phi(x_6) - \phi(x_5)| = 0{,}0058489 < 0{,}01$, $\phi(x_6)$ *predstavlja približnu vrednost minimuma sa traženom tačnošću.*

Ako sada uporedimo rezultate dobijene primenom Curry-Altman-ovog i modifikovanog Curry-Altman-ovog algoritma, vidimo da modifikovani algoritam generiše niz $\{x_k\}$ *koji brže konvergira ka rešenju nego niz generisan originalnim Curry-Altman-ovim algoritmom.*

Poopšteni Céa-Goldstein-ov algoritam

α_k *iz relacije (1) definišemo na sledeći način:*

$$\alpha_k = 0 \quad \text{ako je} \ \langle \nabla\phi(x_k), p_k \rangle = 0;$$

inače se za α_k *uzima ono* α *koje zadovoljava uslove*

$$\phi(x_k) - \phi(x_k - \alpha p_k) \geqslant \alpha\sigma(\langle \nabla\phi(x_k), p_k \rangle),$$

$$\phi(x_k) - \phi(x_k - 2\alpha p_k) < 2\alpha\sigma(\langle \nabla\phi(x_k), p_k \rangle),$$

gde je $\sigma: [0,\infty) \to [0,\infty)$ *prinudna funkcija takva da je* $\sigma(t) \leq \delta t$, $0<\sigma<1$, *za* $\forall t \geqslant 0$.

Ako opet uzmemo $\sigma(t) = ln(1+t)$, $p_k = g_k$, *gornji uslovi dobijaju sledeći oblik:*

$$(6) \quad \begin{aligned} &\phi(x_k) - \phi(x_k - \alpha g_k) \geqslant \alpha ln(1 + \|g_k\|^2), \\ &\phi(x_k) - \phi(x_k - 2\alpha g_k) < 2\alpha ln(1 + \|g_k\|^2). \end{aligned}$$

Uzmimo opet za početnu tačku $x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ *i formirajmo uslove (6) u tački* x_0. *Dobićemo sledeći sistem nejednačina:*

$$-561000\alpha^2 + 30500\alpha \geqslant 10{,}3255\alpha$$

$$-4 \cdot 561000\alpha^2 + 2 \cdot 30500\alpha < 2 \cdot 10{,}3255\alpha.$$

Rešavanjem ovog sistema dobijamo dopustivi interval vrednosti za α_0: $0{,}0271743 < \alpha \doteq 0{,}0543487$. *Uzmimo, na primer,* $\alpha_0 = 0{,}03$. *Sledi da je*

$$x_1 = x_0 - \alpha_0 g_0 = \begin{bmatrix} 2{,}1 \\ 4{,}8 \end{bmatrix}, \quad g_1 = \begin{bmatrix} -28 \\ 32 \end{bmatrix}.$$

Sistem nejednačina (6) u tački x_1 *će glasiti:*

$$-28320\alpha^2 + 1808\alpha \geqslant 7{,}50053\alpha$$

$$-4 \cdot 28320\alpha^2 + 2 \cdot 1808\alpha < 2 \cdot 7{,}50053\alpha.$$

Rešenje ovog sistema je interval $0,0317884 < \alpha \leq 0,0635769$. *Neka je* $\alpha_1 = 0,04$. *Dobijamo*

$$x_2 = x_1 - \alpha_1 g_1 = \begin{bmatrix} 3,22 \\ 3,52 \end{bmatrix}, \quad g_2 = \begin{bmatrix} -5,6 \\ -19,2 \end{bmatrix}.$$

Dalje, rešavanjem sistema nejednačina

$$-\ 7686,4\alpha^2 + 400\alpha \geq 5,99396\alpha$$

$$-\ 4\cdot 7686,4\alpha^2 + 2\cdot 400\alpha < 2\cdot 5,99396\alpha$$

dobijamo $0,02563 < \alpha \leq 0,05126001$. *Uzmimo, na primer,* $\alpha_2 = 0,03$, *dobijamo*

$$x_3 = x_2 - \alpha_2 g_2 = \begin{bmatrix} 3,388 \\ 4,096 \end{bmatrix}, \quad g_3 = \begin{bmatrix} -2,24 \\ 3,84 \end{bmatrix}.$$

Na analogan način izračunavamo dopustivi interval za α_3: $0,02424 < \alpha \leq 0,04848$. *Ako uzmemo* $\alpha_3 = 0,03$, *dobijamo*

$$x_4 = x_3 - \alpha_3 g_3 = \begin{bmatrix} 3,4552 \\ 3,9808 \end{bmatrix}, \quad g_4 = \begin{bmatrix} -0,896 \\ -0,768 \end{bmatrix}.$$

Dopustivi interval za α_4 *je* $0,01312 < \alpha \leq 0,02624$, *uzimamo* $\alpha_4 = 0,026$ *i dobijamo*

$$x_5 = x_4 - \alpha_4 g_4 = \begin{bmatrix} 3,478496 \\ 4,000768 \end{bmatrix}, \quad g_5 = \begin{bmatrix} -0,43008 \\ 0,02072 \end{bmatrix}.$$

Za α_5 *uzimamo neku vrednost iz intervala* $0,00412 < \alpha \leq 0,00824$, *koja predstavlja rešenje sistema nejednačina (6) obrazovanog u tački* x_5. *Tako za* $\alpha_5 = 0,008$ *dobijamo*

$$x_6 = x_5 - \alpha_5 g_5 = \begin{bmatrix} 3,4819366 \\ 4,0005222 \end{bmatrix}.$$

Kako je dalje $\phi(x_5) = 0,00463601$, $\phi(x_6) = 0,00326831$ *i*

$$|\phi(x_6) - \phi(x_5)| = 0,0013677 < 0,01,$$

$\phi(x_6)$ *predstavlja približnu vrednost minimuma sa traženom tačnošću* $\varepsilon = 0,01$.

Céa-Goldstein-ov algoritam

α_k *iz (1) definišemo na sledeći način:*

$$\alpha_k = 0, \text{ ako je } \langle \nabla\phi(x_k), p_k \rangle = 0,$$

inače se za α_k *uzima ono* α *koje zadovoljava uslove*

$$\phi(x_k) - \phi(x_k - \alpha p_k) \geqslant \alpha\mu < \nabla\phi(x_k), p_k >,$$

$$\phi(x_k) - \phi(x_k - 2\alpha p_k) < 2\alpha\mu < \nabla\phi(x_k), p_k >, \quad 0 < \mu < 1,$$

odnosno za $p_k = g_k$

$$(7) \quad \begin{aligned} &\phi(x_k) - \phi(x_k - \alpha g_k) \geqslant \alpha\mu \| g_k \|^2 , \\ &\phi(x_k) - \phi(x_k - 2\alpha g_k) < 2\sigma\mu \| g_k \|^2 \end{aligned}$$

Neka je $\mu = 0,4$. Uzmimo $x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Rešavanjem sistema (7) dobijamo da je dopustivi interval za α_0: $0,01631 < \alpha \leq 0,03262$. Neka je $\alpha_0 = 0,02$, dobijamo

$$x_1 = x_0 - \alpha_0 g_0 = \begin{bmatrix} 1,4 \\ 3,2 \end{bmatrix}, \quad g_1 = \begin{bmatrix} -42 \\ -32 \end{bmatrix}.$$

Rešavanjem sistema (7) u tački x_1 dobijamo $0,021941 < \alpha \leq 0,0438824$. Uzmimo $\alpha_1 = 0,03$, dobijamo

$$x_2 = x_1 - \alpha_1 g_1 = \begin{bmatrix} 2,66 \\ 4,16 \end{bmatrix}, \quad g_2 = \begin{bmatrix} -16,8 \\ 6,4 \end{bmatrix}.$$

Na analogan način dobijamo sledeće približne vrednosti:

$$x_3 = \begin{bmatrix} 3,164 \\ 3,968 \end{bmatrix}, \qquad g_3 = \begin{bmatrix} -6,72 \\ -1,28 \end{bmatrix};$$

$$x_4 = \begin{bmatrix} 3,3656 \\ 4,0064 \end{bmatrix} \qquad g_4 = \begin{bmatrix} -2,688 \\ 0,256 \end{bmatrix}$$

$$x_5 = \begin{bmatrix} 3,451616 \\ 3,998208 \end{bmatrix} \qquad g_5 = \begin{bmatrix} -0,96768 \\ -0,07168 \end{bmatrix}, \quad \phi(x_5) = 0,0234743;$$

$$x_6 = \begin{bmatrix} 3,4825817 \\ 4,0005017 \end{bmatrix} \qquad g_6 = \begin{bmatrix} -0,348366 \\ 0,020068 \end{bmatrix}, \phi(x_6) = 0,003039.$$

Pošto je

$$|\phi(x_6) - \phi(x_5)| = 0,0204353 > 0,01,$$

moramo tražiti sledeće približno rešenje.

Dobijamo

$$x_7 = \begin{bmatrix} 3,4965163 \\ 3,9996989 \end{bmatrix}, \quad \phi(x_7) = 0,000123174.$$

Kako je

$$|\phi(x_7) - \phi(x_6)| = 0,002916 < 0,01,$$

sledi da $\phi(x_7)$ predstavlja vrednost minimuma sa traženom tačnošću $\varepsilon = 0,01$.

Uporedjivanjem rezultata dobijenih primenom Céa-Goldstein-ovog i poopštenog Céa-Goldstein-ovog algoritma, vidimo da modifikovani algoritam generiše niz $\{x_k\}$ *koji brže konvergira ka rešenju.*

Primena projektivne ciklično-koordinatne metode

Razmotrićemo sledeći problem (Hadley, /28/, str. 10): naći $\min_{x\in R^2} \phi(x)$, *gde je*

$$\phi(x) = 10(x_1 - 3{,}5)^2 + 20(x_2 - 4)^2,$$

uz uslov da tačka $x = (x_1,x_2)$ *zadovoljava sledeće nejednačine:*

$$\begin{aligned} x_1 + x_2 &\leq 6, \\ x_1 - x_2 &\leq 1, \\ 2x_1 + x_2 &\geq 6, \\ \tfrac{1}{2}x_1 - x_2 &\geq -4, \\ x_1 \geq 1,\ x_2 &\geq 0. \end{aligned}$$

Optimalno rešenje ovog problema je, kako je pokazano u /28/, str 10, $x^* = (2{,}5;\ 3{,}5)$, $\phi(x^*) = 15$.

Navedeni problem rešićemo primenom algoritma projektivne ciklično-koordinatne metode.

Uvedimo sledeće oznake:

$$\begin{aligned}
f_1 &\equiv x_1 + x_2 - 6 \leq 0, & a_1^T &= [1\ \ 1]; \\
f_2 &\equiv x_1 - x_2 - 1 \leq 0, & a_2^T &= [1\ \ -1]; \\
f_3 &\equiv -2x_1 - x_2 + 6 \leq 0, & a_3^T &= [-2\ \ -1]; \\
f_4 &\equiv -\tfrac{1}{2}x_1 + x_2 - 4 \leq 0, & a_4^T &= [-\tfrac{1}{2}\ \ 1]; \\
f_5 &\equiv -x_1 + 1 \leq 0, & a_5^T &= [-1\ \ 0]; \\
f_6 &\equiv -x_2 \leq 0, & a_6^T &= [0\ \ -1].
\end{aligned}$$

Uzmimo za početnu tačku x^0 *vrh konveksnog poliedra definisanog gornjim nejednačinama, koji se nalazi u preseku ograničenja* f_2 *i* f_3:

$$x^0 = (A^T)^{-1} b_0, \quad A^T = \begin{bmatrix} a_2^T \\ a_3^T \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & -1 \end{bmatrix}, \quad b_0 = \begin{bmatrix} 1 \\ -6 \end{bmatrix},$$

$$x^0 = (\tfrac{7}{3},\ \tfrac{4}{3}), \qquad \phi(x^0) = 155{,}8333.$$

Ispitajmo da li kretanjem iz x^0 u dopustivom pravcu možemo dobiti manju vrednost funkcije. Odstranimo jedno od aktivnih organičenja u x^0, na primer, f_3. Pogledajmo da li kretanjem duž ograničenja f_2 iz tačke x^0 dolazi do opadanja funkcije ϕ. Kako je

$$f_2 = x_1 - x_2 - 1 = 0 \text{ za } x_2 = x_1 - 1, \text{ tj. } x_1 = t, \; x_2 = t - 1, \; t \in R,$$

dopustive tačke će biti oblika $x = (t, t - 1)$. Pri tom je očigledno za $x = x^0$, $t_0 = 7/3$. Dalje je

$$F(t) = \phi(t, t - 1) = 10(t - 3{,}5)^2 + 20(t - 1 - 4)^2 = 30t^2 - 270t + 622{,}5.$$

Nadjimo minimum funkcije F duž datog pravca:

$$F'(t) = 60t - 270 = 0 \rightarrow t^* = 4{,}5 \rightarrow x_1 = 4{,}5, \; x_2 = 3{,}5.$$

Proverimo da li je $(4{,}5;\ 3{,}5)$ dopustiva tačka. Kako je $4{,}5 + 3{,}5 - 6 = 2 > 0$, vidimo da nije zadovoljeno prvo ograničenje. Kako je F strogo konveksna, funkcija će opadati za $\forall t \in [7/3;\ 4{,}5]$. Prema tome, F će u tački preseka ograničenja f_1 i f_2 imati manju vrednost od one u tački x^0. Nadjimo tačku preseka f_2 i f_1 :

$$x_1 = t, \; x_2 = t - 1 \Rightarrow t + t - 1 - 6 = 0 \Rightarrow t = 3{,}5$$

Dakle, presek f_1 i f_2 je $x^{(1)} = (3{,}5;\ 2{,}5)$. Napustimo sada ograničenje f_2 i idimo po f_1 iz tačke $x^{(1)}$:

$$f_1 \equiv x_1 + x_2 - 6 = 0 \Rightarrow x_2 = 6 - x_1 \Rightarrow x_1 = t, \; x_2 = 6 - t \Rightarrow x = (t, 6-t);$$

$$F(t) = \phi(t,\ 6-t) = 10(t - 3{,}5)^2 + 20(2 - t)^2 = 30t^2 - 150t + 202{,}5;$$

$$F'(t) = 60t - 150 = 0 \Rightarrow t = 2{,}5.$$

Odavde sledi da funkcija F postiže minimum duž pravca f_1 u tački $x^{(2)} = (2{,}5;\ 3{,}5)$.

Tačka $(2{,}5;\ 3{,}5)$ je očigledno dopustivo rešenje. Ispitajmo da li je optimalno. Uklonimo (prema koraku 8. algoritma) jedino aktivno ograničenje f_1 u tački $x^{(2)}$ i pogledajmo da li funkcija ϕ opada u pravcu koordinatne ose x_1, odnosno ose x_2. Pogledajmo pre svega ponašanje funkcije u pravcu x_1-ose, odnosno u tačkama oblika $x = (2{,}5 + t;\ 3{,}5)$:

$$F(t) = \phi(2{,}5 + t;\ 3{,}5) = 10(t-1)^2 + 20 \cdot (0{,}5)^2 = 10t^2 - 20t + 15.$$

Kako je $F'(t) = 0$ za $t = 1$, funkcija F će opadati za $\forall t \in (-\infty, 1]$ i rasti za $\forall t \in [1, +\infty)$. Pogledajmo sada za koje t su dopustive tačke $(2{,}5 + t; 3{,}5)$:

$f_1 \equiv 2,5 + t + 3,5 - 6 \leq 0 \Rightarrow t \leq 0;$

$f_1 \equiv 2,5 + t - 3,5 - 1 \leq 0 \Rightarrow t \leq 2;$

$f_3 \equiv -2x_1 - x_2 + 6 = -2t - 2,5 \leq 0 \Rightarrow t \geq -1,25;$

$f_4 \equiv -\frac{1}{2}x_1 + x_2 - 4 = -\frac{1}{2}t - 1,7 \leq 0 \Rightarrow t \geq -3,4;$

$f_5 \equiv -x_1 + 1 = -1,5 - t \leq 0 \Rightarrow t \geq -1,5;$

$f_6 \equiv x_2 \geq 0$ *ispunjeno jer je* $x_2 = 3,5$.

Prema tome, dopustivi interval za t *je* $-1,25 \leq t \leq 0$. *Sledi, na osnovu prethodnoga, da će za dopustive* $t \in [-1,25;\ 0]$ *vrednosti funkcije biti veće nego u tački* $x^{(2)} = (2,5;\ 3,5)$ *što znači, da u pravcu* x_1*-ose nema opadanja funkcije.*

Na analogan način ispitujemo ponašanje funkcije ϕ *u pravcu* x_2*-ose, tj. u tačkama oblika* $x = (2,5;\ 3,5 + t)$:

$F(t) = \phi(2,5;\ 3,5 + t) = 20t^2 - 20t + 15,$

$F'(t) = 0 \Rightarrow t = 0,5.$

Dakle $F(t)$ *postiže minimum u pravcu* x_2*-ose za* $t = 0,5$.

Ispitajmo dopustivost tačaka oblika $x = (2,5;\ 3,5 + t)$:

$f_1 \equiv t \leq 0 \Rightarrow t \leq 0;$

$f_2 \equiv -2 - t \leq 0 \Rightarrow t \geq -2;$

$f_3 \equiv -2,5 - t \leq 0 \Rightarrow t \geq -2,5;$

$f_4 \equiv t - 1,7 \leq 0 \Rightarrow t \leq 1,7;$

$f_5 \equiv -2,5 + 1 \leq 0$

$f_6 \equiv -3,5 - t \leq 0 \Rightarrow t \geq -3,5.$

Vidimo da je dopustivi interval za t $-2 \leq t \leq 0$. *Medjutim, pošto funkcija* f *opada za* $t \in (-\infty;\ 0,5)$, *vidimo da ni u pravcu* x_2*-ose funkcija* ϕ *ne opada za dopustive* t.

Prema tome, tačka $x^* = 2,5;\ 3,5)$ *je optimalno rešenje i* $\phi(x^*) = 15$.

Primena projektivne Powell-Zangwill-ove metode

Razmatramo sledeći problem: naći minimum funkcije

$$(*)\qquad \phi(x) = 2x_1^2 + 3x_2^2 + x_3^2 - 6x_1x_2 + 4x_2x_3 - 10x_1 + 7x_3$$

uz uslov da navedene promenljive zadovoljavaju jednačinu:

$$-x_1 + 3x_2 - 2x_3 = 4.$$

Gradijent i Hesijan funkcije ϕ su oblika:

$$\nabla\phi(x) = \begin{bmatrix} 4x_1 - 6x_2 - 10 \\ 6x_2 - 6x_1 + 4x_3 \\ 2x_3 + 4x_2 + 7 \end{bmatrix}, \qquad H(x) = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 6 & 2 \\ 0 & 2 & 2 \end{bmatrix}$$

Iz pozitivne definitnosti Hesijana $H(x)$ sledi stroga konveksnost funkcije ϕ, odakle sledi da izloženi problem ima jedinstveno optimalno rešenje x^*.

Lagrange-ova funkcija za izloženi problem će glasiti:

$$L(x,\lambda) = 2x_1^2 + 3x_2^2 + x_3^2 - 6x_1x_2 + 4x_2x_3 - 10x_1 + 7x_3 + \lambda(-x_1 + 3x_2 - 2x_3 - 4).$$

Odavde sledi da potreban uslov za minimum glasi:

$$\frac{\partial L}{\partial x_1} = 4x_1 - 6x_2 - 10 - \lambda = 0,$$

$$\frac{\partial L}{\partial x_2} = 6x_2 - 6x_1 + 4x_3 + 3\lambda = 0,$$

$$\frac{\partial L}{\partial x_3} = 2x_3 + 4x_2 + 7 - 2\lambda = 0,$$

$$\frac{\partial L}{\partial x_4} = -x_1 + 3x_2 - 2x_3 - 4 = 0, \qquad -x_1 + 3x_2 - 2x_3 = 4.$$

Rešavanjem ovog sistema nalazimo optimalno rešenje i traženu vrednost: $\lambda^* = 28$, $x^* = (26{,}545454;\ 11{,}363636;\ 1{,}772728)$, $\phi(x^*) = -\ 182{,}\ 52309$.

Potražimo sada optimalno rešenje problema (*) primenom projektivnog Powell-Zangwill-ovog algoritma.

Ovde je broj nepoznatih, odnosno dimenzija prostora: $n = 3$, broj ograničenja: $m = 1$, a $n - m = 2$ je dimenzija prostora kojeg razapinju $n - m = 2$ linearno nezavisnih rešenja homogene jednačine $-x_1+3x_2-2x_3=0$. Bazisna rešenja ove jednačine, odnosno bazisni vektori $n - m$-dimenzionog prostora će glasiti:

$$z_1 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \qquad z_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

Kako je $\| z_1 \| = \sqrt{10}$, $\| z_2 \| = \sqrt{5}$, jedinični bazisni vektori, koji su nam potrebni u algoritmu:

$$z_1 = \begin{bmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \\ 0 \end{bmatrix} \qquad z_2 = \begin{bmatrix} -\frac{2}{\sqrt{10}} \\ 0 \\ \frac{1}{\sqrt{5}} \end{bmatrix}, \text{ tj.} \qquad \begin{aligned} z_1 &= (0{,}94868;\ 0{,}31623;\ 0) \\ z_2 &= (-0{,}89443;\ 0;\ 0{,}44721). \end{aligned}$$

Za početnu tačku x_0^0 uzmimo bilo koju dopustivu tačku. Iz stroge konveksnosti funkcije ϕ slediće da je uslov o ograničenosti nivoskog skupa ispunjen. Neka je, na primer, $x_0^0 = (-4,0,0)$. Stavimo $k = 0$, $\delta^0 = 1$, $s_1^0 = z_1$, $s_2^0 = z_2$, $\epsilon = 10^{-4}$.

Sada, prema koraku 1 algoritma nalazimo sledeće tačke

$$x_1^0 = x_0^0 + \alpha_1^0 s_1^0,$$

gde α_1^0 odredjujemo iz uslova

$$\phi(x_0^0+\alpha_1^0 s_1^0) = \min_\alpha \phi(x_0^0+\alpha s_1^0) = \min_\alpha \phi(-4+0{,}94868\alpha;\ 0{,}31623\alpha;0) =$$

$$= \min_\alpha(0{,}29999\alpha^2-17{,}07616\alpha+const.),$$

sledi

$$\alpha_1^0 = 28{,}46122, \quad x_1^0 = (23{,}00058;\ 9{,}00029;\ 0).$$

Dalje je

$$x_2^0 = x_1^0 + \alpha_2^0 s_2^0,$$

gde α_2^0 odredjujemo iz uslova

$$\phi(x_1^0+\alpha_2^0 s_2^0) = \min_\alpha \phi(x_1^0+\alpha s_2^0) = \min_\alpha \phi(23{,}00058-0{,}89443\alpha;\ 9{,}00029;\ 0{,}44721\alpha) =$$

$$= \min_\alpha(1{,}80001\alpha^2-5{,}81493\alpha+const.)$$

Dobijamo

$$\alpha_2^0 = 1{,}61525, \quad x_2^0 = (21{,}55585;\ 9{,}00029;\ 0{,}72236).$$

Prema koraku 2 algoritma nalazimo

$$\lambda^0 = \|x_2^0 - x_0^0\| = 27{,}10403,$$

$$s_3^0 = \frac{x_2^0 - x_0^0}{\lambda^0} = (0,94288;\ 0,33206;\ 0,02663),$$

$$x_3^0 = x_2^0 + \alpha_3^0 s_3^0,$$

pri čemu se α_3^0 *odredjuje iz uslova*

$$\min_{\alpha} = \phi(x_2^0 + \alpha s_3^0) = \phi(x_1^0 + \alpha_3^0 s_3^0).$$

Dobija se $\alpha_3^0 = 3,60166$, *pa je*

$$x_0^1 = x_3^0 = (24,95178;\ 10,19626;\ 0,81834).$$

Prema koraku 3 algoritma nalazimo

$$\alpha_p^0 = max\{\ 28,46122;\ 1,61525\} = 28,46122 = \alpha_1^0.$$

Pošto je

$$\frac{\alpha_p^0 \cdot \delta^0}{\lambda^0} = \frac{28,461215 \cdot 1}{27,10403} = 1,05007 > \varepsilon, \quad \delta^1 = 1,05007,$$

sledi da ćemo na sledećoj iteraciji (k = 1) za bazisne vektore uzeti

$$s_1^1 = s_3^0 = (0,94288;\ 0,33206;\ 0,02665),$$

$$s_2^1 = s_2^0 = (-0,89443;\ 0;\ 0,44721).$$

Dakle, za iteraciju k = 1 imamo spremne: početni vektor x_0^1 *i bazisne vektore* s_1^1 *i* s_2^1 *i prethodno opisani postupak iteriramo, ali sada sa vektorima* x_0^1, s_1^1 *i* s_2^1. *Dobijamo sledeće rezultate:*

$$x_1^1 = x_1^0 + \alpha_1^1 s_1^1 = (24,95184;\ 10,19628;\ 0,81834),\ \alpha_1^1 = 0,00006;$$

$$x_2^1 = x_1^1 + \alpha_2^1 s_2^1 = (24,08093;\ 10,19628;\ 1,25379),\ \alpha_2^1 = 0,97370;$$

$$\lambda^1 = \| x_2^1 - x_0^1 \| = 0,97365;$$

$$s_3^1 = (-0,89442;\ 0,00002;\ 0,44723);$$

$$x_0^2 = x_3^1 = x_2^1 + \alpha_3^1 s_3^1 = (24,08086;\ 10,19628;\ 1,25383),\ \alpha_3^1 = 0,00008;$$

$$\alpha_p^1 = max\{\ 0,00006;\ 0,97370\} = 0,97370 = \alpha_2^1.$$

Pošto je

$$\frac{\alpha_p^1 \cdot \delta^1}{\lambda^1} = \frac{0,97370 \cdot 1,05007}{0,97365} = 1,05012 > \varepsilon, \quad \delta^2 = 1,05012,$$

Za sledeću iteraciju (k =2) bazisni vektori će biti

$$s_1^2 = s_1^1 = (0,94288;\ 0,33206;\ 0,02665),$$

$$s_2^2 = s_3^1 = (-0,89442;\ 0,00002;\ 0,44723).$$

Dalje, prema koraku 1 algoritma dobijamo

$$x_1^2 = x_0^2 + \alpha_1^2 s_1^2 = (25,75869;\ 10,78717;\ 1,30125),\quad \alpha_1^2 = 1,77947,$$

$$x_2^2 = (25,32834;\ 10,78718;\ 1,51643),\quad \alpha_2^2 = 0,48115.$$

Dalje, prema koraku 2, sledi

$$\lambda^2 = \| x_2^2 - x_0^2 \| = 1,40511,$$

$$s_3^2 = (0,88782;\ 0,42054;\ 0,18689),$$

$$x_0^3 = x_3^2 = x_2^2 + \alpha_3^2 s_3^2 = (26,54727;\ 11,36454;\ 1,77302).$$

Prema koraku 3 algoritma nalazimo

$$\alpha_p^2 = max\{ 1,77947;\ 0,48115\} = 1,77947 = \alpha_1^2.$$

Pošto je

$$\frac{\alpha_p^2 \delta^2}{\lambda^2} = 1,32990 > \varepsilon,$$

sledi da će na sledećoj iteraciji (k = 3) bazisni vektori biti

$$s_1^3 = s_3^2 = (0,88782;\ 0,42054;\ 0,18689),$$

$$s_2^3 = s_2^2 = (-0,89442;\ 0,00002;\ 0,44723).$$

Dalje, na osnovu koraka 1, dobijamo

$$x_1^3 = x_0^3 + \alpha_1^3 s_1^3 = (26,54725;\ 11,36455;\ 1,77302),\quad \alpha_1^3 = -0,00002,$$

$$x_2^3 = (26,54721;\ 11,36455;\ 1,77304).$$

Prema koraku 2 je

$$\lambda^3 = \| x_2^3 - x_0^3 \| = 0,00006.$$

Kako je $\varepsilon = 0,0001$, sledi da je $\lambda^3 < \varepsilon$, što znači da je x_2^3 optimalno rešenje sa traženom tačnošću $\varepsilon = 10^{-4}$.

Dakle, primenom projektivne Powell-Zangwill-ove metode dobili smo

$$x^* = (26{,}54721;\ 11{,}36455;\ 1{,}77304),$$

$$\phi(x^*) = -182{,}5328.$$

REZULTATI PROGRAMA NA RAČUNARU

Ovde ćemo izložiti rezultate dobijene na računaru IBM 360/44. Programi na FORTRAN jeziku za optimizacione algoritme dati su u Prilogu.

Prvo ćemo izložiti dobijene rezultate kod minimizacije Fletcher--Powell-ove funkcije /47/

$$\phi(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4.$$

Potreban uslov za minimum funkcije ϕ glasi

$$\frac{\partial\phi}{\delta x_1} = 0, \quad tj. \quad 2(x_1 + 10x_2) + 40(x_1 - x_4)^3 = 0,$$

$$\frac{\partial\phi}{\delta x_2} = 0, \quad tj. \quad 20(x_1 + 10x_2) + 4(x_2 - 2x_3)^3 = 0,$$

$$\frac{\partial\phi}{\delta x_3} = 0, \quad tj. \quad 10(x_3 - x_4) - 8(x_2 - 2x_3)^3 = 0,$$

$$\frac{\partial\phi}{\delta x_3} = 0, \quad tj. \quad -10(x_3 - x_4) - 40(x_1 - x_4)^3 = 0.$$

Rešavanjem ovog sistema dobijamo $x_1 = x_4$, $x_2 = 2x_3$, $x_1 = -10x_2$, $x_4 = -20x_3$, odakle sledi da je optimalno rešenje $x^* = (0,0,0,0)$ i $\min_{x \in R^4} \phi(x) = \phi(x^*) = 0$.

Za početnu tačku uzimamo $x^0 = (-3,-1,0,1)$.

Rezultati primene poopštenog Céa-Goldstein-ovog algoritma:

broj iteracija	55,
broj računanja vrednosti funkcije	534,
broj računanja gradijenta	55.

Rezultati primene Céa-Goldstein-ovog algoritma:

μ	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
broj iteracija	77	83	75	63	67	64	56	81
br. rač. vr. funkcije	746	835	825	768	787	861	605	742
br. rač. gradijenta	77	83	75	63	67	64	56	81

Rezultati dobijeni primenom Nelder-Mead-ovog algoritma (Petrić J., /43/, str. 115.)

broj iteracija	*63,*
broj računanja vrednosti funkcije	*258.*

Uporedjivanjem rezultata dobijenih primenom tri izložena optimizaciona algoritma, vidimo da najbržu konvergenciju ima poopšteni Céa-Goldstein-ov algoritam.

Sada dajemo dobijene rezultate kod minimizacije Rosenbrock-ove funkcije (Rosenbrock H.H., Computer J., 3, 174, 1960)

$$\phi(x) = 100(x_2-x_1^2)^2 + (1-x_1)^2.$$

Rešavanjem sistema

$$\frac{\delta\phi}{\delta x_1} = -400x_1(x_2-x_1^2) - 2(1-x_1) = 0,$$

$$\frac{\delta\phi}{\delta x_2} = 200(x_2-x_1^2) = 0,$$

dobijamo optimalno rešenje $x^* = (1,1)$, $\phi(x^*) = 0$.

Za početnu tačku uzimamo $x^0 = (-1,2;\ 1)$.

Rezultati primene poopštenog Céa-Goldstein-ovog algoritma:

broj iteracija	*41,*
broj računanja vrednosti funkcije	*286,*
broj računanja gradijenta	*41.*

Rezultati primene Nelder-Mead-ovog algoritma:

broj iteracija	*57,*
broj računanja vrednosti funkcije	*231.*

Rezultati primene Céa-Goldstein-ovog algoritma:

μ	*0,4*	*0,7*	*0,9*
broj iteracija	*51*	*68*	*ne postiže se*
br. rač. vr. funkcija	*591*	*784*	*tražena tačnost ni*
br. rač. gradijenta	*51*	*68*	*u 100 iteracija*

Iz izloženog se vidi da poopšteni Céa-Goldstein-ov algoritam obezbedjuje najbržu konvergenciju.

PRILOG

Optimizacija Fletcher-Powell-ove funkcije

Fortran program za Céa-Goldstein-ov algoritam

```
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NG,NF,NIMX
      DIMENSION X(20)
      EXTERNAL F
      NG=0
      NF=0
C
      READ(5,10)N,(X(I),I=1,N)
   10 FORMAT(I2,10F7.0)
   15 READ(5,20,END=60)XMI,NIMX
   20 FORMAT(F10.0,I3)
      WRITE(6,30)N,XMI,(I,X(I),I=1,N)
   30 FORMAT(' 1BROJ PROMENLJIVIH =',I2,' ,MI=',F5.3/
     * ' POCETNE VREDNOSTI'/1X,6('X(',I1,')=',E14.7,3X))
C
      CALL MINI(N,X,F,K,XMI)
C
      WRITE(6,40) K,(I,X(I),I=1,N)
   40 FORMAT(///' REZULTAT DOBIJEN U',I4,' ITERACIJA'/
     * ' VREDNOSTI ARGUMENATA'/
     * 1X,6('X(',I1,')=',E14.7,3X))
      FF=F(N,X)
      WRITE(6,50) FF
   50 FORMAT(/' VREDNOST FUNKCIJE =',E14.7)
      WRITE(6,55)NF,NG
   55 FORMAT(' BROJ RACUNANJA FUNKCIJE=',I5/' BROJ RACUNANJA GRADIJENTA=
     *',I3)
C
      GO TO 15
   60 STOP
      END
```

```
      FUNCTION F(N,X)
C     FLETCHER - POWELL-OVA FUNKCIJA
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NG,NF,NIMX
      NF=NF+1
      DIMENSION X(N)
C
      F=(X(1)+10.*X(2))**2+5.*(X(3)-X(4))**2+ (X(2)-2.*X(3))**4+
     *10.*(X(1)-X(4))**4
C
      RETURN
      END
```

```
      SUBROUTINE DER(N,X,G)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NG,NF,NIMX
      DIMENSION X(N),G(N)
      NG=NG+1
C
      G(1)=2.*(X(1)+10.*X(2))+40.*(X(1)-X(4))**3
      G(2)=20.*(X(1)+10.*X(2))+4.*(X(2)-2.*X(3))**3
      G(3)=10.*(X(3)-X(4))-8.*(X(2)-2.*X(3))**3
      G(4)=-10.*(X(3)-X(4))-40.*(X(1)-X(4))**3
C
      RETURN
      END


      SUBROUTINE SEARCH(N,X,F,G,XMI,ALFA)
C
C   CEA- GOLDSTEIN-OV ALGORITAM
C
      IMPLICIT REAL*8(A-H,O Z)
      DIMENSION X(N),G(N),Y(20)
      LOGICAL POC/.TRUE./
      FX=F(N,X)
C
C     PROVERA POGODNOSTI PRETHODNOG ALFA
C
      S=0.0
      DO 10 I=1,N
   10 S=S+G(I)*G(I)
      IF(S)20,20,30
   20 ALFA=0.0
      RETURN
   30 IF(POC) GO TO 60
      DO 40 I=1,N
   40 Y(I)=X(I)-ALFA*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.ALFA*XMI*S) GO TO 6
      DO 50 I=1,N
   50 Y(I)=X(I)-2.0*ALFA*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.2.0*ALFA*S*XMI) RETURN
C
C  TRAZENJE ALFA KOJE ZADOVOLJAVA PRVU NEJEDNACINU
C
   60 ALFA=0.5
      POC=.FALSE.
   65 DO 70 I=1,N
   70 Y(I)=X(I)-ALFA*G(I)
      FY=F(N,Y)
      IF(FX-FY.GE.ALFA*XMI*S) GO TO 80
      ALFA=ALFA/2.0
      GO TO 65
C
C    TRAZENJE ALFA KOJE ZADOVOLJAVA I DRUGU NEJEDNACINU
C
   80 DO 90 I=1,N
   90 Y(I)=X(I)-2.0*ALFA*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.2.0*XMI*S*ALFA) RETURN
C
      A=ALFA
      B=2.0*ALFA
   95 C=0.5*(A+B)
      DO 100 I=1,N
```

```
  100 Y(I)=X(I)-C*G(I)
      FY=F(N,Y)
      IF(FX-FY.GE.C*XMI*S) GO TO 110
      B=C
      GO TO 95
  110 DO 120 I=1,N
  120 Y(I)=X(I)-2.0*C*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.2.0*C*XMI*S) GO TO 130
      A=C
      GO TO 95
  130 ALFA=C
      RETURN
      END


      SUBROUTINE CONVRG(N,X,Y,F,G,IPAS)
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION X(N),Y(N),G(N)
C
      FTOL=0.001
C
C     ISPITATI VREDNOSTI FUNKCIJE
      FX=F(N,X)
      FY=F(N,Y)
      IF(DABS(FX-FY).GT.FTOL) GO TO 60
C
C     KRITERIJUM KONVERGENCIJE JE ZADOVOLJEN
      IPAS=1
      RETURN
C     KONVERGENCIJA NIJE POSTIGNUTA
   60 IPAS=2
      RETURN
      END


      SUBROUTINE MINI(N,X,F,K,XMI)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NC,NF,NIMX
      EXTERNAL F
      DIMENSION X(N),G(20),Y(20)
C
      K=0
   10 K=K+1
C
      CALL DER(N,X,G)
C
      CALL SEARCH(N,X,F,G,XMI,ALFA)
C
      DO 20 I=1,N
   20 Y(I)=X(I)-ALFA*G(I)
C
      CALL CONVRG(N,X,Y,F,G,IPAS)
C
      DO 30 I=1,N
   30 X(I)=Y(I)
C
      IF(IPAS.EQ.1) RETURN
      IF(K.LT.NIMX) GO TO 10
      WRITE(6,40)NIMX
   40 FORMAT(//' NE POSTIZE SE TACNOST U',I3,' ITERACIJA')
      RETURN
C
      END
```

```
CJ FRCMENLJIVIH = 4
ETNE VREDNOSTI
)=-0.3CCCCCCC C1   X(2)=-0.1CCCCCOD C1   X(3)= C.C  X(4)= 0.1000000C 01

.200
.TAT CCBIJEN V  77 ITERACIJA
.CSTI ARGUMENATA
=-0.4001..1.C 00 X(2)= 0.3555436D-C1 X(3)=-0.1783477C CC X(4)=-C.2184851C (

CST FUNKCIJE = C.4459296C-C1
RACUNANJA FUNKCIJE=  746
RACUNANJA GRACIJENTA= 77

.3CC
LTAT CCBIJEN V  83 ITERACIJA
NCSTI ARGUMENATA
=-0.400839.C CC X(2)= .381.3.9D-C1 X(3)=-C.1787436C CC X(4)=-0.2190917C (

NCST FUNKCIJE = C.4393304C-C1
 RACUNANJA FUNKCIJE=  835
 RACUNANJA GRACIJENTA= 83

.400
.LTAT CCBIJEN V  75 ITERACIJA
INCSTI ARGUMENATA
=-0.39893C8C CC X(2)= C.38C2774D-C1 X(3)=-C.178C5C9C CC X(4)=-C.2179099C

INCST FUNKCIJE = .4316..D-C1
I RACUNANJA FUNKCIJE=  825
I RACUNANJA GRACIJENTA= 75

C.5CC
.LTAT CCBIJEN V  63 ITERACIJA
INCSTI ARGUMENATA
)=-C.4C583C8C CC X(2)= C.398937SD-C1 X(3)=-C.18C63CCC CC X(4)=-C.2222953C

INCST FUNKCIJE = C.4597112C-C1
J RACUNANJA FUNKCIJE= .768
J RACUNANJA GRACIJENTA= 63

C.6CC
.LTAT CCBIJEN V  67 ITERACIJA
ENCSTI ARGUMENATA
)=-C.4C4878SC CC X(2)= 0.3974615D-C1 X(3)=-C.18.2838C CC X(4)=-0.2216959C (

ENCST FUNKCIJE = C.4557C32C-C1
J RACUNANJA FUNKCIJE=  787
J RACUNANJA GRACIJENTA= 67

C.7CC
ZULTAT CCBIJEN V  64 ITERACIJA
EDNCSTI ARGUMENATA
1)=-0.420058CC CC X(2)= 0.4036596D-C1 X(3)=-C.1856646C CC X(4)=-C.2311946C

EDNCST FUNKCIJE = 0.5208471C-C1
IJ RACUNANJA FUNKCIJE=  761
IJ RACUNANJA GRACIJENTA= 64
```

```
,600
```



```
.TAT CCBIJEN V  56 ITERACIJA
\CSTI ARGUMENATA
=-0.455537CC CC X(2)= C.4595I91D-C1 X(3)=-C.1555€C5C CC X(4)=-C.2568503C C

,CST FUNKCIJE = C.7266127D-C1
 RACUNANJA FUNKCIJE=  6C5
 RACUNANJA GRADIJENTA= 56


.SCC
LTAT CCBIJEN V  81 ITERACIJA
NCSTI ARGUMENATA
=-0.467893CC CC X(2)= C.4793834D-C1 X(3)=-C.2 67749C   X(4)=-.275 858C

NCST FUNKCIJE = C.89517C2C-C1
 RACUNANJA FUNKCIJE=  742
 RACUNANJA GRADIJENTA= 81
CP       C
```

Fortran program za poopšteni Céa-Goldstein-ov algoritam

```
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NG,NF,NIMX
      DIMENSION X(20)
      EXTERNAL F
      NG=0
      NF=0

      READ(5,10)N,(X(I),I=1,N)
   10 FORMAT(I2,10F7.0)
   15 READ(5,20,END=60)XMI,NIMX
   20 FORMAT(F10.0,I3)
      WRITE(6,30)N,(I,X(I),I=1,N)
   30 FORMAT('IBROJ PROMENLJIVIH =',I2/
     * ' POCETNE VREDNOSTI'/1X,6('X(',I1,')=',E14.7,3X))

      CALL MINI(N,X,F,K)

      WRITE(6,40) K,(I,X(I),I=1,N)
   40 FORMAT(///' REZULTAT DOBIJEN V',I4,' ITERACIJA'/
     * ' VREDNOSTI ARGUMENATA'/
     * 1X,6('X(',I1,')=',E1 .7,3X))
      FF=F(N,X)
      WRITE(6,50) FF
   50 FORMAT(/' VREDNOST FUNKCIJE =',E14.7)
      WRITE(6,55)NF,NG
   55 FORMAT(' BROJ RACUNANJA FUNKCIJE=',I5/' BROJ RACUNANJA GRADIJENTA=
     *',I3)

      GO TO 15
   60 STOP
      END
```

```
      FUNCTION F(N,X)
C     FLETCHER - POWWELE-CVA FUNKCIJA
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NG,NF,NIMX
      NF=NF+1
      DIMENSION X(N)
C
      F=(X(1)+10.*X(2))**2+5.*(X(3)-X(4))**2+ (X(2)-2.*X(3))**4+
     *10.*(X(1)-X(4))**4
C
      RETURN
      END


      SUBROUTINE DER(N,X,G)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NG,NF,NIMX
      DIMENSION X(N),G(N)
      NG=NG+1
C
      G(1)=2.*(X(1)+10.*X(2))+40.*(X(1)-X(4))**3
      G(2)=20.*(X(1)+10.*X(2))+4.*(X(2)-2.*X(3))**3
      G(3)=10.*(X(3)-X(4))-8.*(X(2)-2.*X(3))**3
      G(4)=-10.*(X(3)-X(4))-40.*(X( )-X(4))**3
C
      RETURN
      END

      SUBROUTINE SEARCH(N,X,F,G,ALFA)
C
C    POOPSTENI SEA-GOLDSTEIN-OV ALGORITAM
C
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION X(N),G(N),Y 20)
      LOGICAL POC/.TRUE./
      FX=F(N,X)
C
C     PROVERA POGODNOSTI PRETHODNOG ALFA
C
      S=0.0
      DO 10 I=1,N
   10 S=S+G(I)*G(I)
      XL=DLOG(1.+S)
      IF(S)20,20,30
   20 ALFA=0.0
      RETURN
   30 IF(POC) GO TO 60
      DO 40 I=1,N
   40 Y(I)=X(I)-ALFA*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.ALFA*XL) GO TO 60
      DO 50 I=1,N
   50 Y(I)=X(I)-2.0*ALFA*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.2.0*ALFA*XL)RETURN
C
C    TRAZENJE ALFA KOJE ZADOVOLJAVA PRVU NEJEDNACINU
C
   60 ALFA=0.5
      POC=.FALSE.
   65 DO 70 I=1,N
   70 Y(I)=X(I)-ALFA*G(I)
      FY=F(N,Y)
```

```
      IF(FX-FY.GE.ALFA*XL) GO TO 80
      ALFA=ALFA/2.0
      GO TO 65
C
C     TRAZENJE ALFA KOJE ZADOVOLJAVA I DRUGU NEJEDNACINU
C
   80 DO 90 I=1,N
   90 Y(I)=X(I)-2.0*ALFA*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.2.0*ALFA*XL) RETURN
C
      A=ALFA
      B=2.0*ALFA
   95 C=0.5*(A+B)
      DO 100 I=1,N
  100 Y(I)=X(I)-C*G(I)
      FY=F(N,Y)
      IF(FX-FY.GE.C*XL) GO TO 110
      B=C
      GO TO 95
  110 DO 120 I=1,N
  120 Y(I)=X(I)-2.0*C*G(I)
      FY=F(N,Y)
      IF(FX-FY.LT.2.0*C*XL) GO TO 130
      A=C
      GO TO 95
  130 ALFA=C
      RETURN
      END

      SUBROUTINE CONVRG(N,X,Y,F,G,IPAS)
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION X(N),Y(N),G(N)
C
      FTOL=0.001
C
C     ISPITATI VREDNOSTI FUNKCIJE
      FX=F(N,X)
      FY=F(N,Y)
      IF(DABS(FX-FY).GT.FTOL) GO TO 60
C
C     KRITERIJUM KONVERGENCIJE JE ZADOVOLJEN
      IPAS=1
      RETURN
C     KONVERGENCIJA NIJE POSTIGNUTA
   60 IPAS=2
      RETURN
      END

      SUBROUTINE MINI(N,X,F,K)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON NG,NF,NIMX
      EXTERNAL F
      DIMENSION X(N),G(20),Y(20)
C
      K=0
   10 K=K+1
C
      CALL DER(N,X,G)
C
      CALL SEARCH(N,X,F,G,ALFA)
C
      DO 20 I=1,N
   20 Y(I)=X(I)-ALFA*G(I)
```

```
CALL CONVRG(N,X,Y,F,G,IPAS)
DO 30 I=1,N
X(I)=Y(I)
IF(IPAS.EQ.1) RETURN
IF(K.LT.NIMX) GO TO 10
WRITE(6,40)NIMX
0 FORMAT(//' NE POSTIZE SE TACNOST U',I3,' ITERACIJA')
RETURN
```

```
PROMENLJIVIH = 4
NE VREDNOSTI
-0.3CC... D  1    X(2)=-... 1    X(3)= ...    X(4)= ...D 01

TAT DOBIJEN U  55 ITERACIJA
OSTI ARGUMENATA
-0.4561981D 00    X(2)= 0.4512251D-01X(3)=-0.1983104D 00X(4)=-0.2544402D (

OST FUNKCIJE = 0.7.426..D-..
RACUNANJA FUNKCIJE=  534
RACUNANJA GRADIJENTA= 55
P        0
```

Fortran program za Nelder-Mead-ov algoritam

```
IMPLICIT REAL*8(A-H,O-Z)

MAIN LINE PROGRAM FOR NELDER AND MEAD MINIMIZATION ROUTINE.

NX JE BROJ NEZAVISNIH PROMENLJIVIH
STEP JE POCETNI KORAK
X(I) JE POCETNI VEKTOR
DATA KARTICE SU SLEDECEG OBLIKA
BR. KARTICE       PARAMETRI        FORMAT          KOLONE
  1                NX               I5              1 DO 5
  1                STEP             F10.5           6 DO 15
  2                X(I)             F10.5           1 DO ..
DIMENSION XI(50,50),X 50),SUM(50)
COMMON/C/ X,XI,STEP,SUM,NX,IN,KI,NF
NF=0
1 FORMAT(I5,F10.5)
00 READ1,NX,STEP
IF(NX) 998,999,998
98 READ 2,(X(I),I=1,NX)
2 FORMAT(10F10.5)
ALFA=1.0
BETA=0.5
GAMA=2.0
DIFER=0.
XNX=NX
IN=1
CALL SUMR
PRINT 102,SUM(1),(X(I),I=1,NX)
PRINT 1002,STEP
PRINT 103
```

```
C3 FCRMAT(4X,18H VREDNCST FLNKCIJE,15X,3HX1=,15X,3HX2=,15X,3HX3=,15X,
  1'X4=',
  1   16X,16HPROMENA FLNKCIJE)
 2 FCRMAT(1H1,12X,25HPCCETNA VREDNCST FLNKCIJE,F10.5/' KCLCNA X JE'
  1,/5X,1C(E11.4,2X)I
 2 FCRMAT(12X,'STEP=',F6.2)
   K1=NX+1
   K2=NX+2
   K3=NX+3
   K4=NX+4
   CALL START
 5 DC 3 I=1,K1
   DC 4 J=1,NX
 4 X(J)=X1(I,J)
   IN=I
   CALL SUMR
 3 CONTINUE
   NACI NAJVECU VREDNCST    SUM(I) U SIMPLEKSU
 6 SUMH=SUM(1)
   INDEX=1
   DC 7 I=2,K1
   IF(SUM(I).LE.SUMHI GC TC 7
   SUMH=SUM(I)
   INDEX=I
 7 CCNTINUE
   NACI NAJMANJU VREDNCST  SUM(II U SIMPLEKSU
   SUML=SUM(1)
   KCUNT=1
   DC 8 I=2,K1
   IF(SUML.LE.SUM(I)I GC TC 8
   SUML=SUM(1)
   KCUNT=I
 8 CCNTINUE
   NACI CENTRCID TACAKA KCD KCJIH JE I RAZLICITC CD INDEKSA
   DC 9 J=1,NX
   SUM2=0.
   DC 1C I=1,K1
   SUM2=SUM2+X1(I,J)
   X1(K2,J)=1./XNX*(SUM2-X1(INDEX,J))
   NACI REFLEKCIJU GCRNJE TACKE KRCZ CENTRCID
   X1(K3,J)=(1.+ALFAI*X1(K2,J)-ALFA*X1(INDEX,J)
   X(J)=X1(K3,J)
   IN=K3
   CALL SUMR
   IF(SUM(K3).LT.SUML) GC TC 11
   NACI DRUGU NAJVECU VREDNCST U SIMPLEKSU
   IF(INDEX.EQ.1) GO TC 39
   SUMS=SUM(1)
   GC TC 39
   SUMS=SUM(2)
   DC 12 I=1,K1
   IF((INDEX-I).EQ.OI GC TC 12
   IF(SUM(I).LE.SUMSI GC TC 12
   SUMS=SUM(I)
   CCNTINUE
   IF(SUM(K3).GT.SUMS) GC TC 13
   GC TC 14
```

```
  FORMIRATI EKSRANZIJU NOVCE MINIMUMA AKO JE REFLEKCIJA DALA JEDAN MINIM
.1 DO 15 J=1,NX
  X1(K4,J)=(1-GAMA)*X1(K2,J)+GAMA*X1(K3,J)
.5 X(J)=X1(K4,J)
  IN=K4
  CALL SUMR
  IF(SUM(K4).LT.SUMR) GO TO 16
  GO TO 14
3 IF(SUM(K3).GT.SUMR) GO TO 17
  DO 18 J=1,NX
8 X1(INDEX,J)=X1(K3,J)

.7 DO 19 J=1,NX
  X1(K4,J)=BETA*X1(INDEX,J)+(1.-BETA)*X1(K2,J)
.9 X(J)=X1(K4,J)
  IN=K4
  CALL SUMR
  IF(SUMR.GT.SUM(K4)) GO TO 16
  REDUKOVATI SIMPLEKS POLOVLJENJEM AKO JE REFLEKCIJA DALA VECU VREDNOST (
  MAXIMUMA
  DO 20 J=1,NX
  DO 20 I=1,K1
. X1(I,J)=0.5*(X1(I,J)+X1(KOUNT,J))
  DO 29 I=1,K1
  DO 30 J=1,NX
C X(J)=X1(I,J)
  IN=I
  CALL SUMR
9 CONTINUE
  GO TO 26
6 DO 21 J=1,NX
  X1(INDEX,J)=X1(K4,J)
1 X(J)=X1(INDEX,J)
  IN=INDEX
  CALL SUMR
  GO TO 26
4 DO 22 J=1,NX
  X1(INDEX,J)=X1(K3,J)
2 X(J)=X1(INDEX,J)
  IN=INDEX
  CALL SUMR
5 DO 23 J=1,NX
3 X(J)=X1(K2,J)
  IN=K2
  CALL SUMR
  DIFER=0.
  DO 24 I=1,K1
4 DIFER=DIFER+(SUM(I)-SUM(K2))**2
  DIFER=1./XNX*DSQRT(DIFER)
  PRINT 101,SUMR,(X1(KOUNT,J),J=1,NX ),DIFER
1 FORMAT(2(2X,E16.6),3(7X,E16.6),12X,E16.6)
  IF(DIFER.GE.0.0001) GO TO  28
  WRITE(6,55)NF
5 FORMAT(' BROJ RACUNANJA FUNKCIJE=',I5)
  GO TO 100
3 CONTINUE
  END
```

```
      SUBROUTINE START
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION A(50,50),X1(50,50),X(50),SUM(50)
      COMMON/C/ X,X1,STEP,SUM,NX,IN,K1,NF
      VN=NX
      STEP1=STEP/(VN* SQRT(2.))*(DSQRT(VN+1.)+VN-1.)
      STEP2=STEP/(VN* SQRT(2.))*(DSQRT(VN+1.)-1.)
      DO 1 J=1,NX
    1 A(1,J)=0.
      DO 2 I=2,K1
      DO 2 J=1,NX
      A(I,J)=STEP2
      L=I-1
      A(I,L)=STEP1
    2 CONTINUE
      DO 3 I=1,K1
      DO 3 J=1,NX
    3 X1(I,J)=X(J)+A(I,J)
      RETURN
      END


      SUBROUTINE SUMR
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON/C/ X,X1,STEP,SUM,NX,IN,K1,NF
      NF=NF+1
      DIMENSION X1(50,50),X(50),SUM(50)
      FLETCHER - POWELL-OVA FUNKCIJA
      SUM(IN)=(X(1)+10.*X(2))**2+5.*(X(3)-X(4))**2+ (X(2)-2.*X(3))**4+
     *10.*(X(1)-X(4))**4
      RETURN
      END

        POCETNA VREDNOST FUNKCIJE=215.00000
 X JE
0.30000 01  -0.20000 01   0.0            0.10000 01
      STEP=  0.40
VREDNOST FUNKCIJE      X1=          X2=             X3=             X4=


.195609D-01     -0.291891D 00   0.289548D-01   -0.149172D 00  -0.142930D 0
RACUNANJA FUNKCIJE=  258

AT DOBIJEN V  63 ITERACIJA
```

Literatura

1. *Altman M., Generalized gradient methods of minimizing a functional, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. 14, pp. 313-318, 1966.*
2. *Arrow, K. J., Z. Hurwicz, H. Uzawa, Constraint qualifications in maximization problems, Nav. Res. Log. Quart. 8, pp. 175-191, 1961.*
3. *Avriel M., Nonlinear programming, analysis and methods, Prentice-Hall, Inc., Englewoord Cliffs, New Jersey, 1976.*
4. *Baničuk, N. V., Petrov, V. M., Černousko F. L., Čislennoe rešenie variacionnih i kraevih zadač metodom lokalnih variacij, ŽVM i MF, 6, 6, str. 947-961, 1966.*
5. *Best M. J., Ritter K., An accelerated conjugate direction method to solve linearly constrained minimization problems, Journal of Computer and System Science, vol. 11, pp.295-322, 1975.*
6. *Bulanij A. P., Danilin, J. M., Kvazinjutonovskije algoritmi minimizacii, osnovannije na postroenii sistem sopražonnih vektorov, ŽVM i MF, tom 18, No. 4, str. 877-885, 1978.*
7. *Conn A. R. Petrzykowski T, A penalty function method converging directly to a constrained optimum, SIAM J.N.A., vol. 14, No. 2, pp 348-375, 1977.*
8. *Cauchy A. L., Méthode générale pour la résolution des systemes d'equations simultaneés, Compt. Rend., 25, pp 536-538, 1847.*
9. *Curry H., The method of steepest descent for nonlinear minimization problems, Quart. Appl. Math., 2, pp 258-261, 1944.*
10. *Daniel J. W., The approwimate minimization of functionals, Prentice Hall, Inc., Englewoord Cliffs, N. Y., 1971.*
11. *Danilin J. M., O minimizacii funkcij v zadačah s ograničenijami tipa ravenstv, Kibernetika, No. 2, str. 88-95, 1971.*
12. *Danilin J. M., Metodi sopražennih napravlenij ne trebujuščije rešenija odnomernih zadač minimizacii, DAN SSSR, Tom 218, No. 3, str. 513-516, 1974.*
13. *Danilin J. M., Pšeničnij B. N., O metodah minimizacii s uskorennoj shodimostju, ŽVM i MF, 10, No. 6, str. 1341-1354, 1970.*
14. *Dantzig G. S., Linear programming and extensions, Princeton University Press, Princeton, N. Y., 1963.*

15. Davidon, W. C., Variable metric method for minimization, A.E.C. Research and Development Report, ANL-5990, 1959.

16. Davies D.,Swann W. H., Review of constrained optimization. In "Optimization" (R. Fletcher, ed.) Academic Press, London, 1969.

17. Elkin R., Convergence theorems for Gauss-Siedel and other minimization algorithms, Doctoral Thesis, University of Maryland, College Park, 1968.

18. Farkas J., Über die Theorie der einfachen Ungleichungen, Journal für die Reine und Angewandte Mathematik, 124, pp 1-24, 1902.

19. Fiacco A. V., McCormick G. P., Nonlinear programming: sequential unconstrained minimization techniques, Wiley, New York, 1968.

20. Fletcher R., Powell M. J. D., A rapidly convergent descent method for minimization, Computer Y. 6, pp 163-168, 1963.

21. Friedman, P., Pinder K. L., Optimization of a simulation model of a chemical plant, Ind. Eng. Chem. Process Des Develop. 11, pp 512-520, 1972.

22. Gill, P. E., Murray W., Numerical methods for constrained optimization, Academic Press, London, 1974.

23. Gill P. E., Golub G. H., Murray W., Saunders M. A., Methods for modifying matrix factorizations, Mathematics of computation, vol. 28, No. 126, pp. 505-535, 1974.

24. Gill P. E., Murray W., Numerically stable methods for quadratic programming, Mathematical Programming 14, pp 349-372, 1978.

25. Goldfarb D., Extension of Davidon's variable metric method to maximization under linear inequality and equality constraits, SIAM J. Appl. Math., vol 17, No. 4, 1969.

26. Goldfarb D., Matrix factorizations in optimization of nonlinear functions subject to linear constraints, Mathematical Programming 10, pp 1-31, 1975.

27. Goldstein A. A., Constructive real analysis, New York, 1967.

28. Hadley G., Nonlinear and dynamic programming, Addison-Wesley, Reading, Mass, 1964.

29. Karlin S., Mathematical methods and theory in games, programming and economics, vols I, II, Addison Wesley Publishing Company, Inc., Reading, Mass., 1959.

30. Kovačević V., A convergence theory for linearly constrained nonlinear programming problems, doktorska teza, 1977.

31. Kuester J. Z., Mize J. H., Optimization techniques with Fortran, McGraw-Hill Inc., 1973.
32. Kuhn H. W., Tucker A. W., Nonlinear programming in J. Neuman (ed.), "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability", pp 481-492, University of California Press, Berkeley, California, 1951.
33. Levitin E. S., Poljak B. T., Metodi minimizacii pri naličii ograničenij, ŽVM i MF, tom 6, str. 787-822, 1966.
34. Luenberger D. G., Introduction to linear and nonlinear programming, Addison Wesley Publishing Company, 1973.
35. Mangasarian O., Equivalence in nonlinear programming, Naval. Res. Logist. Quart. 10, pp. 299-306, 1963.
36. Mangasarian O., Nonlinear programming, McGraw-Hill, New York, 1969.
37. Meyer R. R., A comparison of the forcing function and point-to-set mapping approaches to convergence analysis, SIAM J. Control and Optimization, vol. 15, No. 4, pp 699-715, 1977.
38. Meyer R. R., A convergence theory for a class of antijamming strategies, JOTA, vol. 21, No. 3, 1977.
39. Mugele R. A., A program for optimal control of nonlinear process, IBB Systems Journal, 1, pp 2-17, 1962.
40. Nazareth L., A conjugate direction algorithm without line searches, JOTA, vol. 23, No. 3, pp 373-38., 1977.
41. Ortega J., Rheinboldt W., Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.
42. Ostrowski A., Solution of equations and systems of equations, 2nd ed., Academic Press, New York, 1966.
43. Petrić J., Nelinearno programiranje, Privredno-finansijski vodič, Beograd, 1979.
44. Polak E., Computational methods in optimization, a unified approach, Academic Press, New York, 1971.
45. Poljak B. T. O skorosti shodimosti metoda štrafnih funkcij, ŽVM i MF, tom 11, No. 1, str. 3-11, 1971.
46. Powell M. J. D., An efficient method for finding the minimum of a function of several variables without calculating derivations, Computer Journal, vol. 7, pp 155-162, 1964.

47. Powell M. J. D., Computer J., 7, 155, 1964.

48. Pšeničnij B. N., Danilin J. M., Čislennie metodi v ekstremaljnih zadačah, Nauka, Moskva, 1975.

49. Rauch S. W., Convergence theory for a class of nonlinear programming problems, SIAM J. Numer. Anal.,vol. 10, No. 1, 1973.

50. Ritter K., A superlinearly convergent method for minimization problems with linear inequality constraints, Mathematical Programming 4, pp 44-71, 1973.

51. Riter K., A method of conjugate directions for linearly constrained non-linear programming problems, SIAM J. Numer. Anal., vol. 12, No. 3, pp 273-303, 1975.

52. Rosen J. B., The gradient projection method for nonlinear programming, Part I. Linear Constraints, J. Soc. Indust. Appl. Math. 1, 1960.

53. Rosen J. B., The gradient projection method for nonlinear programming. Part II. Nonlinear constraints, J. Soc. Indust. Appl. Math. 4, 1961.

54. Sea J., Optimizacija, teorija i algoritmi, Mir, Moskva, 1973.

55. Shanno D. F., An accelerated gradient projection method for linearly constrained nonlinear estimation, SIAM j. Appl. Math., vol. 18, No.2, pp 322-334, 1970.

56. Singer E., Simulation and optimization of oil refinery design, Chem. Eng. Progr. Symp. Ser. 37, 58, 1962.

57. Slater M., Lagrange multipliers revisited: a contribution to nonlinear programming, in "Cowles Commission Discussion Paper", Mathematics 403, 1950.

58. Stewart G. W., A modification of Davidon's minimization method to accept difference approximations to derivatives, JACM 14, 1967.

59. Toint P. L., Callier F. M., On the uniform nonsingularity of matrices of search directions and the rate of convergence in minimization algorithms, JOTA, vol. 23, N . 4, pp 511-529, 1977.

60. Toint P. L., Callier F. M., On the accelerating property of an algorithm for function minimization without calculating derivatives, JOTA, vol. 23, No. 4, pp 531-547, 1977.

61. Topkis D. M., Veinott A. F., On the convergence of some feasible directions algorithms for nonlinear programming, J. SIAM Control 2, pp 268-279, 1967.

62. Zangwill W. I., *Minimizing a function without calculating derivatives*, Computer Journal, vol. 10, pp 293-296, 1967.

63. Zangwill W. I., *Nonlinear programming - a unified approach*, Prentice--Hall, Inc., Engelwood Cliffs, W. J., 1969.

64. Zountendijk G., *Methods of feasible directions*, Elsevier Publishing Company, 1960.

REGISTAR