

Univerzitet u Beogradu
Prirodno-matematički fakultet

Do MA

Momir V. Čelić

MNOGOSTEPENE ITERACIONE PROCEDURE
ZA RJEŠAVANJE SINGULARNIH
KONAČNO-ELEMENTNIH JEDNAČINA

doktorska disertacija

OSNOVNA ORGANIZACIJA UDRUŽENOG RADA
ZA MATEMATIKU, MEHANIČKU I ASTRONOMIJU
BIBLIOTEKA

Број: Dokt. 185/1

Датум: 17. 03. 1986.

Banjaluka, 1986.

S A D R Ź A J

0.	PREDGOVOR	IV
1.	ELIPTIČKI GRANIČNI ZADACI I METODI ZA NJIHOVO PRIBLIŽNO RJEŠAVANJE	1
1.1.	Osnovni pojmovi i oznake	1
1.2.	Metod konačnih razlika	6
1.3.	Metod konačnih elemenata	8
2.	ITERACIONI METODI ZA RJEŠAVANJE VELIKIH SISTEMA LINEARNIH JEDNAČINA	11
2.1.	Karakteristike sistema jednačina dobijenih primjenom metoda konačnih razlika i konačnih elemenata	11
2.2.	Pregled iteracionih metoda za rješavanje sistema sa regularnom matricom	12
2.3.	Iteracioni metodi za rješavanje sistema sa singularnom matricom	21
3.	MNOGOSTEFENI ITERACIONI METODI	31
3.1.	Uvodne napomene	31
3.2.	Oznake, ideja i komponente mnogostepenih metoda	32
3.3.	Konvergenција i efikasnost mnogostepene iteracione procedure	45
4.	PRIMJENA METODA REGULARIZACIJE ZA RJEŠAVANJE SINGULARNIH KONAČNO-ELEMENTNIH JEDNAČINA	50
4.1.	O nekorektno postavljenim zadacima i metodu regularizacije	50

4.2. Rješavanje singularnih i loše uslovljenih sistema linearnih algebarskih jednačina	56
4.3. Formiranje regularizacionog algoritma za rješavanje konačno-elementnih aproksimacija singularnih eliptičkih graničnih zadataka	58
5. ANALIZA MNOGOSTEPENE ITERACIONE PROCEDURE ZA RJEŠAVANJE SINGULARNIH ZADATAKA	69
5.1. Uvodne napomene	69
5.2. Postavka zadatka, oznake i pretpostavke	70
5.3. Definicija k-stepene iteracione procedure	77
5.4. Dokaz konvergencije	79
5.5. Računski posao i efikasnost	87
LITERATURA	89

ОСНОВНА ОРГАНИЗАЦИЈА УДРУЖЕНОГ РАДА
ЗА МАТЕМАТИКУ, МЕХАНИКУ И АСТРОНОМИЈУ
БИБЛИОТЕКА

Б р о ј: _____

Д а т у м: _____

o. PREDGOVOR

Mnoge fizičke pojave se mogu opisati parcijalnim diferencijalnim jednačinama, pa metodi za njihovo približno rješavanje dobijaju sve značajnije mjesto u savremenoj numeričkoj analizi. U metode koji se najčešće koriste spadaju metod konačnih razlika (metod mreža) i metod konačnih elemenata. Ono što je zajedničko za oba ta metoda je to da njihovo korištenje dovodi do sistema algebarskih jednačina. Odredjene specifičnosti tih sistema zahtijevaju korištenje specijalnih metoda za njihovo rješavanje koji će uzimati u obzir karakteristike tih sistema i omogućiti nalaženje rješenja sa upotrebom manjeg broja računskih operacija u poredjenju sa opštim metodima linearne algebre. Sa pojavom računara važno mjesto medju tim specijalnim metodima dobili su iteracioni metodi (vidjeti [25, 26, 45]). U početku su se oni koristili samo za rješavanje sistema sa regularnom matricom, da bi kasnije bili korišteni za rješavanje singularnih, kako saglasnih tako i nesaglasnih, a takodje i sistema sa pravougaonom matricom (vidjeti [16]). Različiti iteracioni metodi se medjusobno upoređuju po efikasnosti, tj. po faktorima konvergencije i po broju računskih operacija potrebnih da se norma greške (ili ostatka) smanji odredjen broj (ϵ^{-1}) puta. Pokazalo se da je najefikasniji metod za rješavanje mrežnih jednačina metod promjenljivih pravaca za koji je broj potrebnih računskih operacija jednak $O(n^2 \ln n \ln \epsilon^{-1})$, tj. $O(N^2 \ln N \ln \epsilon^{-1})$ gdje je N^2 - broj nepoznatih. Ali i za taj metod je potreban velik broj računskih

operacija kada je h malo. U želji da se dobije ekonomičan metod za koji ce potreban broj operacija biti proporcionalan sa brojem nepoznatih, Fedorenko je 1961. godine (vidjeti [58]) predložio jedan iteracioni metod sa koji je taj broj jednak

$O(N^2 h \epsilon^{-1})$. Tek ce 80-ih godina Fedorenkova ideja biti značajnije iskorištena što ce dovesti do pojave mnogostepenih ili višemrežnih metoda (multilevel, multigrid methods), čija se ideja sastoji u korištenju većeg broja mreža i kombinovanju klasičnih iteracionih metoda sa dobrim glačajućim svojstvima i krupno-mrežne korekcije. Pokazalo se da ti metodi imaju faktor konvergencije koji ne zavisi od diskretizacionog parametra h zbog čega su oni vrlo brzi, a da je potreban broj računskih operacija proporcionalan broju nepoznatih, što ove metode čini asimptotski optimalnim iteracionim metodima. I pored ovih prednosti u odnosu na klasične iteracione metode, ovaj metod se jos uvijek ne koristi dovoljno u praksi na što utiče možda i složenost strukture jednog iteracionog kruga u poredjenju sa krugom klasičnih metoda. Osim toga, ovaj metod je još uvijek nepoznat u širim krugovima, pogotovu onim inženjerskim. Za sada ne postoji ni jedna knjiga koja bi na pristupačan način govorila o ovom metodu. Ostaje jedina mogućnost za upoznavanjem korištenjem nekih osnovnih radova (vidjeti [6, 7, 10, 24]).

U Kelnu je 1981.g. održana konferencija posvećena ovom metodu poslije koje je objavljen zbornik radova što je ostavilo velik trag i uticalo na sve veće interesovanje za ovaj metod.

U prvoj polovini 1985.g. je u Americi održana druga konferencija što je znak da je ovaj metod i šire prihvacen.

Za sada se mnogostepeni metodi uspješno primjenjuju za rješavanje eliptičkih jednačina, za probleme sopstvenih vrijednosti i zadatke bifurkacije, paraboličke i integralne jednačine.

Ovaj rad se bavi analizom mnogostepenih metoda za rješavanje singularnih samokonjugovanih eliptičkih zadataka, pri čemu se ti zadaci tretiraju kao nekorektni, te se prethodno vrši njihova regularizacija u smislu Tihonova.

Rad je podijeljen na 5 glava.

U prvoj glavi su definisani eliptički granični zadaci (klasična i varijaciona formulacija), uvedene oznake i funkcionalni prostori koji će se kasnije koristiti. Navedeni su i glavni metodi za približno rješavanje eliptičkih graničnih zadataka - metod konačnih razlika i metod konačnih elemenata.

Druga glava je posvećena klasičnim iteracionim metodima. Ukazano je na osnovni nedostatak ovih metoda pri korištenju za rješavanje mrežnih i konačno-elementnih jednačina - sporu konvergenciju i velik računski posao zbog zavisnosti faktora konvergencije od diskretizacionog parametra h .

U trećoj glavi je opisan mnogostepeni iteracioni metod - asimptotski optimalan iteracioni metod. Za njegovu primjenu se stalno pretpostavljalo da posmatrani zadatak ima jedinstveno rješenje, tj. da nula nije tačka spektra operatora zadatka.

Glavni rezultati su u glavama 4 i 5.

Vršeci konačno-elementnu aproksimaciju samokonjugovanih, nenegativno odredjenih eliptičkih graničnih zadataka, dobijamo singularne sisteme algebarskih jednačina. Pošto su to nekorektno postavljani zadaci, za njihovo rješavanje se koristi Tihonovljevi

metod regularizacije. Za rješavanje dobijenog sistema je formiran jedan regularizacioni algoritam koji liči na uprošćenu regularizaciju. Tako se dobija regularizovani sistem za čije se rješavanje koristi mnogostepena iteraciona procedura. Dokazana je konvergencija te procedure nezavisno od diskretizacionog parametra h , a takodje je dokazano da je ukupan računski posao proporcionalan broju nepoznatih. Pošto se zbog korištenja mnogostepenog metoda vrši formiranje i rješavanje sistema na raznim stepenima koji imaju oblik regularizovanog sistema, predložena su dva načina za izbor parametara regularizacije na svakom stepenu.

Banjaluka, novembra 1985.

M. V. Čelić

1. ELIPTIČKI GRANIČNI ZADACI I METODI ZA NJIHOVO PRIBLIŽNO RJEŠAVANJE

1.1. Osnovni pojmovi i oznake

Neka je $\Omega \in \mathbb{R}^n$ proizvoljna ograničena, jednostruko povezana oblast, $\partial\Omega$ njena granica i $\bar{\Omega} = \Omega \cup \partial\Omega$.

Neka funkcija $u = u(x_1, x_2, \dots, x_n)$ unutar oblasti Ω zadovoljava jednačinu

$$(1) \quad Lu \equiv - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} + cu = f,$$

a na granici $\partial\Omega$ jedan od graničnih uslova

$$(2) \quad u|_{\partial\Omega} = 0$$

$$(3) \quad \left(\frac{\partial u}{\partial n} + \sigma u \right) |_{\partial\Omega} \equiv \left[\sum_{i,j=1}^n a_{ij} \frac{\partial u}{\partial x_j} \cos(\nu, x_i) + \sigma u \right] |_{\partial\Omega} = 0.$$

Sa ν smo označili spoljnu normalu na granicu $\partial\Omega$.

Zadatak (1)-(2) nazivamo prvim graničnim ili Dirichlet-ovim,

a zadatak (1)-(3) trecim graničnim zadatkom ako je $\sigma \neq 0$, i drugim graničnim ili Neumann-ovim zadatkom ako je $\sigma = 0$.

Koeficijenti a_{ij} , b_i , c i σ su funkcije od $x = (x_1, x_2, \dots, x_n)$.

Smatracemo da je ispunjen uslov eliptičnosti, tj. da postoji konstanta $C_0 > 0$ takva da za sve realne brojeve $\xi_1, \xi_2, \dots, \xi_n$, koji nisu svi jednaki nuli, u svakoj tački $x \in \bar{\Omega}$, važi nejednakost

$$(4) \quad 0 < C_0 \sum_{i=1}^n \xi_i^2 \leq \sum_{i,j=1}^n a_{ij} \xi_i \xi_j.$$

U slučaju kada su koeficijenti jednačine a_{ij} , b_i , c i desna strana f dovoljno glatke funkcije, funkcija $u \in C^2(\Omega) \cap C(\bar{\Omega})$ koja u oblasti Ω zadovoljava jednačinu (1), a na granici $\partial\Omega$

neki od graničnih uslova (2)-(3), se naziva klasičnim rješenjem odgovarajućeg graničnog zadatka (vidjeti [18,32,33,37]).

$C^k(\bar{\Omega})$ je standardna oznaka za prostor svih k puta neprekidno diferencijabilnih na $\bar{\Omega}$ funkcija. Norma u $C^k(\bar{\Omega})$ se definiše jednakošću

$$\|u\|_{C^k(\bar{\Omega})} = \max_{x \in \bar{\Omega}} \sum_{|\alpha| \leq k} |D^\alpha u(x)|,$$

gdje je $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, α_i su nenegativni cijeli brojevi, $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$, $D^0 u(x) = u(x)$, $D^\alpha u(x)$ označava parcijalni izvod reda $|\alpha|$, tj.

$$D^\alpha u(x) = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

Pri tome je $C^0(\bar{\Omega}) = C(\bar{\Omega})$.

Sa $L_2(\Omega)$ ćemo označavati prostor svih mjerljivih na Ω funkcija koje su na Ω integrabilne zajedno sa svojim kvadratom. U tom prostoru se norma definiše sa

$$\|u\| = \left(\int_{\Omega} |u|^2 dx \right)^{1/2}.$$

Često ćemo stavljati $L_2(\Omega) = H^0(\Omega)$, a uvedenu normu $\|\cdot\|$ ćemo ponekad označavati i sa $\|\cdot\|_0$.

Prostor $L_2(\Omega)$ je Hilbert-ov za skalarni proizvod

$$(u, v) = \int_{\Omega} u(x)v(x) dx.$$

Za cijeli broj $m \geq 1$ ćemo sa $H^m(\Omega)$ označavati prostore Soboljeva reda m . To su prostori koji se sastoje od svih funkcija iz $L_2(\Omega)$ koje imaju sve uopštene izvode do reda m zaključno, koji su takodje integrabilni na Ω zajedno sa svojim kvadratima (vidjeti [15,33]).

Ako sa $\mathcal{D}(\Omega)$ označimo prostor beskonačno puta diferencijabilnih na Ω funkcija sa kompaktnim nosačem u Ω , kažemo da je funkcija $D^*u \in L_2(\Omega)$ uopšteni izvod funkcije u reda $|*|$ ako je

$$\int_{\Omega} D^*u \varphi dx = (-1)^{|*|} \int_{\Omega} u D^*\varphi dx \quad \text{za sve } \varphi \in \mathcal{D}(\Omega).$$

(vidjeti [31]).

Norma u $H^m(\Omega)$ se definiše jednakošću

$$\|u\|_m = \left(\sum_{|*| \leq m} \int_{\Omega} |D^*u|^2 dx \right)^{1/2}.$$

Prostori $H^m(\Omega)$ su Hilbert-ovi prostori za skalarnu proizvodnju

$$(u, v)_m = \sum_{|*| \leq m} (D^*u, D^*v).$$

Koristimo i prostore

$$H_0^m(\Omega) = \overline{\mathcal{D}(\Omega)},$$

gdje se zatvaranje uzima u smislu norme $\|\cdot\|_m$.

Pridružimo jednačini (1) integralni identitet

$$(5) \quad \sum_{i,j=1}^n \int_{\Omega} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx + \sum_{i=1}^n \int_{\Omega} b_i \frac{\partial u}{\partial x_i} v dx + \int_{\Omega} c u v dx = \int_{\Omega} f v dx,$$

gdje je v proizvoljna glatka funkcija jednaka nuli na granici $\partial\Omega$ oblasti Ω . Ako je $u(x)$ npr. dva puta neprekidno diferencijabilna funkcija u Ω , i ako su svi koeficijenti jednačine (1) glatke funkcije, lako se pokazuje da su jednakosti (1) i (5) za tu funkciju ekvivalentne. Ali, (5) ima smisla i za funkciju u koja ima izvode samo prvog reda

a_{ij} , b_i , c i f koje nisu diferencijabilne funkcije.

Izraz $\int_{\Omega} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx$ ima smisla za sve funkcije u i v iz prostora $H^1(\Omega)$, pa (5) važi ne samo za sve funkcije $v \in \mathcal{D}(\Omega)$ već i za sve funkcije $v \in H_0^1(\Omega)$.

U vezi sa ovim se uvodi pojam uopštenog rjesenja (v. [32, 33, 37]).

Funkcija $u \in H^1(\Omega)$ se naziva uopštenim rjesenjem graničnog zadatka (1)-(2) za $f \in L_2(\Omega)$, ako ona zadovoljava (5) za sve funkcije $v \in H_0^1(\Omega)$ i granični uslov (2), pri čemu se jednakost u tom graničnom uslovu podrazumijeva u smislu traga (v. [57]).

Funkcija $u \in H^1(\Omega)$ se naziva uopštenim rjesenjem graničnog zadatka (1)-(3) za $f \in L_2(\Omega)$, ako za nju važi

$$(5') \int_{\Omega} \left[\sum_{i,j=1}^m a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + \sum_{i=1}^m b_i \frac{\partial u}{\partial x_i} v + c u v \right] dx + \int_{\partial\Omega} \sigma u v ds = \int_{\Omega} f v dx, \quad v \in H^1(\Omega)$$

Za eliptičke granične zadatke se osim formulacije (1)-(2)

ili (1)-(3) koristi i varijaciona formulacija (v. [8, 18, 22]):

naci funkciju $u \in V$ tako da je

$$(6) \quad a(u, v) = (f, v) \quad \text{za sve } v \in V,$$

gdje je $V = H_0^1(\Omega)$ u slučaju prvog i $V = H^1(\Omega)$ u slučaju drugog i trećeg graničnog zadatka, a sa $a(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$ je označena bilinearna forma

$$(7) \quad a(u, v) = \int_{\Omega} \left[\sum_{i,j=1}^m a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + \sum_{i=1}^m b_i \frac{\partial u}{\partial x_i} v + c u v \right] dx$$

koja odgovara operatoru L iz (1) za prvi i drugi granični zadatak, dok je u slučaju trećeg graničnog zadatka

$$(7') \quad a(u, v) = \int_{\Omega} \left[\sum_{i,j=1}^m a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + \sum_{i=1}^m b_i \frac{\partial u}{\partial x_i} v + c u v \right] dx + \int_{\partial\Omega} \sigma u v ds$$

Govoricemo da je bilinearna forma $a(\cdot, \cdot)$ simetrična ako je

$$a(u, v) = a(v, u) \quad \text{za sve } u, v \in V.$$

Bilinearna forma $a(\cdot, \cdot)$ je pozitivna ako je

$$a(u, u) > 0 \quad \text{za sve } u \in V \text{ osim za } u=0,$$

nenegativna ili nenegativno odredjena ako je

$$a(u, u) \geq 0 \quad \text{za sve } u \in V,$$

i pozitivno odredjena ili V -eliptična ako je

U vezi sa ovim se uvodi pojam uopštenog rješenja (vidjeti [32,33,37]).

Funkcija $u \in H^1(\Omega)$ se naziva uopštenim rješenjem graničnog zadatka (1)-(2) za $f \in L_2(\Omega)$, ako ona zadovoljava (5) za sve funkcije $v \in H_0^1(\Omega)$ i granični uslov (2), pri čemu se jednakost u tom graničnom uslovu podrazumijeva u smislu traga (vidjeti [37]).

Funkcija $u \in H^1(\Omega)$ se naziva uopštenim rješenjem graničnog zadatka (1)-(3) za $f \in L_2(\Omega)$, ako za nju važi (5) za sve funkcije $v \in H^1(\Omega)$.

Za eliptičke granične zadatke se osim formulacije oblika (1)-(2) ili (1)-(3) koristi i varijaciona (slaba) formulacija (vidjeti [8,18,22]):

naći funkciju $u \in V$ tako da je

$$(6) \quad a(u, v) = (f, v) \quad \text{za sve } v \in V,$$

gdje je $V = H_0^1(\Omega)$ u slučaju prvog i $V = H^1(\Omega)$ u slučaju drugog i trećeg graničnog zadatka, a sa $a(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$ je označena bilinearna forma

$$(7) \quad a(u, v) = \int_{\Omega} \left[\sum_{i,j=1}^n a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} v + cuv \right] dx$$

koja odgovara diferencijalnom operatoru L iz (1).

Govoricemo da je bilinearna forma $a(\cdot, \cdot)$ simetrična ako je

$$a(u, v) = a(v, u) \quad \text{za sve } u, v \in V.$$

Bilinearna forma $a(\cdot, \cdot)$ je pozitivna ako je

$$a(u, u) > 0 \quad \text{za sve } u \in V \text{ osim za } u=0,$$

nenegativna ili nenegativno odredjena ako je

$$a(u, u) \geq 0 \quad \text{za sve } u \in V,$$

i pozitivno odredjena ili V -eliptična ako je

$$\alpha(u, u) \geq \alpha \|u\|_V^2 \quad \text{za sve } u \in V,$$

gdje je $\alpha > 0$ i $\|\cdot\|_V$ - norma u prostoru V .

Za operator L kažemo da je samokonjugovan ako je

$$(Lu, v) = (u, Lv)$$

za sve funkcije u, v iz oblasti definisanosti operatora L .

Ako je L samokonjugovan, biće odgovarajuća bilinearna forma $\alpha(\cdot, \cdot)$ simetrična.

Operator L je pozitivan ako je $(Lu, u) > 0$ za sve funkcije $u \neq 0$ iz oblasti definisanosti operatora L , nenegativan ili nenegativno odredjen ako je $(Lu, u) \geq 0$ za sve u i pozitivno odredjen ako je $(Lu, u) \geq \alpha \|u\|^2$, $\alpha > 0$.

Pošto ćemo se mi baviti eliptičkim zadacima sa nenegativno odredjenim operatorima, primjetimo da će njima odgovarati nenegativno odredjena bilinearna forma.

Skup

$$\ker L = \{u : Lu = 0\}$$

ćemo zvati jezgrom operatora L .

Mi ćemo se u ovom radu baviti samokonjugovanim eliptičkim graničnim zadacima kojima odgovara varijaciona (slaba) formulacija oblika (6) sa simetričnom bilinearnom formom $\alpha(\cdot, \cdot)$.

Pokazuje se da je u tom slučaju zadatak (6) ekvivalentan ekstremalnom zadatku:

naci $u \in V$ tako da je

$$J(u) = \inf_{v \in V} J(v),$$

gdje je $J: V \rightarrow \mathbb{R}$ linearni funkcional definisan sa

$$J(v) = \alpha(v, v) - 2(f, v).$$

Ako je bilinearna forma $\alpha(\cdot, \cdot)$ neprekidna i V -eliptična, tj. ako postoji $\alpha > 0$ tako da je

$$\alpha(v, v) \geq \alpha \|v\|_V^2 \quad \text{za sve } v \in V,$$

a linearna forma $f(v) = (f, v)$ neprekidna, tada zadatak (6) ima jedinstveno rješenje, što je sadržaj Lax-Milgram-ovog tvrdjenja (vidjeti [8, 22]).

V -eliptičnost bilinearne forme povlači njenu pozitivnu odredjenost. Mi ćemo razmatrati samokonjugovane, nenegativno odredjene granične zadatke kojima će odgovarati slaba formulacija oblika (6) sa simetričnom i nenegativno odredjenom bilinearom formom. Takvi zadaci ne zadovoljavaju uslove Lax-Milgram-ovog tvrdjenja, pa oni ne moraju za sve funkcije f imati rješenje, a ako ga budu imali, ono neće biti jedinstveno. Ti zadaci će imati rješenje ako je funkcija f ortogonalna na jezgro odgovarajućeg operatora L , pri čemu se ortogonalnost podrazumijeva u smislu L_2 -skalarnog proizvoda (vidjeti [1, 2]). To ćemo uvijek pretpostavljati. Ako je $\ker L \neq \{0\}$, zadatak je singularan i on će imati beskonačno mnogo rješenja. Iz tog skupa se na jedinstven način izdvaja rješenje koje je ortogonalno na skup $\ker L$. To će biti rješenje sa minimalnom L_2 -normom. Bavićemo se odredjivanjem aproksimacije baš tog rješenja.

1.2. Metod konačnih razlika

Jedan od vrlo rasprostranjenih metoda za približno rješavanje graničnih zadataka za parcijalne diferencijalne jedna-

čine, pa samim tim i za eliptičke granične zadatke, je metod konačnih razlika ili metod mreža (vidjeti [9, 35, 47]).

Osnovna ideja ovog metoda se sastoji u prelasku od diferencijalne jednačine na njen diferencijski analog - diferencijsku šemu. Da bi se mogla napisati ta diferencijska sema potrebno je:

1. zamijeniti oblast neprekidne promjene argimenta oblašcu njegove diskretne promjene,
2. zamijeniti diferencijalni operator nekim diferencijskim operatorom, a takodje formulisati diferencijski analog za granične uslove.

U našem slučaju ovo znači da se oblast Ω u kojoj je zadana jednačina (1) zamjenjuje oblašcu Ω_h od konačno mnogo tačaka koja se naziva mrežom. Za ostvarenje drugog zahtjeva treba sve parcijalne izvode zamijeniti količnicima razlika. Po ostvarenju ove procedure dolazimo do diskretnog eliptičkog graničnog zadatka kojeg ćemo zapisati u obliku operatorske jednačine

$$L_h u_h = f_h .$$

Taj zadatak u stvari predstavlja sistem linearnih algebarskih jednačina, pa se zadatak približnog rješavanja polaznog eliptičkog graničnog zadatka svodi na odredjivanje rjesenja dobijenog sistema jednačina. Poslije ostvarenog prelaska na diferencijski zadatak treba provjeriti da li i u kom smislu diferencijski zadatak aproksimira diferencijalni, te da li i u kom smislu rjesenje prvog konvergira ka rjesenju drugog zadatka. Mi se u ovom radu nećemo interesovati za te probleme, već ćemo pažnju posvetiti problemima vezanim za efek-

tivno rješavanje sistema dobijenih diskretizacijom neprekidnih graničnih zadataka. Često se za algebarske jednačine dobijene primjenom metoda mreža koristi i naziv mrežne jednačine (vidjeti [45]).

1.3. Metod konačnih elemenata

Uz metod konačnih razlika, najčešće korišćen metod za numeričko rješavanje graničnih zadataka za parcijalne diferencijalne jednačine je metod konačnih elemenata. Specijalno, za rješavanje eliptičkih graničnih zadataka taj metod pokazuje i neke prednosti u odnosu na metod konačnih razlika.

Metod konačnih elemenata je u suštini varijacioni metod (vidjeti [22, 44, 36]) i za njegovu primjenu se koristi varijaciona (slaba) formulacija graničnog zadatka. Njega karakterišu sljedeća svojstva:

1. oblast u kojoj se rješava zadatak se dijeli na podoblasti ili konačne elemente,

2. nepoznata funkcija se aproksimira funkcijama specijalnog oblika na svakom konačnom elementu, pa samim tim i na cijeloj oblasti,

3. ta aproksimacija se stavlja u jednačinu koja predstavlja varijacionu (slabu) formulaciju polaznog zadatka, što dovodi do sistema jednačina u kojima se kao nepoznate veličine javljaju parametri aproksimacije čijim određivanjem dobijamo približno rjesenje polaznog zadatka.

Najčešće se u dvodimenzionom slučaju vrši trijangulacija

oblasti Ω , tj. ta oblast se dijeli na trouglove, dok se nepoznata funkcija na tim trouglovima aproksimira polinomima. Posmatrano na cijeloj oblasti, nepoznata funkcija se aproksimira funkcijama koje su dio po dio polinomi. Mi samo za aproksimaciju koristiti polinome prvog stepena, tj. linearne konačne elemente. Pretpostavlja se da u prostoru funkcija koje su dio po dio polinomi određenog stepena, postoji baza od funkcija koje imaju "male" nosače.

Koristeći metod konačnih elemenata, zadatak oblika (6) koji se rješava u prostoru V se zamjenjuje zadatkom:

naci $u_h \in V_h$ tako da je

$$(8) \quad a(u_h, v_h) = (f, v_h) \quad \text{za sve } v_h \in V_h,$$

gdje je V_h konačnodimenzioni potprostor prostora V koji se sastoji od funkcija koje su dio po dio polinomi određenog stepena. Za taj prostor se koristi i naziv prostor konačnih elemenata.

Raznim trijangulacijama T_h oblasti Ω će odgovarati razni prostori konačnih elemenata V_h , a svakom tom prostoru diskretni zadatak oblika (8).

Ako je data baza $\{\psi_1^h, \psi_2^h, \dots, \psi_N^h\}$ prostora V_h , gdje je

$N = \dim V_h$, rješenje zadatka (8) tražimo u obliku

$$(9) \quad u_h = \sum_{i=1}^N u_i \psi_i^h$$

gdje su u_i nepoznati koeficijenti. Da bismo odredili te koeficijente, uvrstavamo to rješenje u (8) odakle, stavljajući

$v_h = \psi_j^h$, $j = 1, 2, \dots, N$, dobijamo sistem jednačina

$$(10) \quad \sum_{i=1}^N a(\psi_i^h, \psi_j^h) u_i = (f, \psi_j^h), \quad j = 1, 2, \dots, N.$$

Rješavanjem ovog sistema određujemo nepoznate koeficijente što nam omogućava da korištenjem (9) dobijemo rješenje zadatka (8), koje predstavlja aproksimaciju rješenja zadatka (6). Budući da je broj N obično velik, zbog količine računskog posla potrebnog za rješavanje sistema (10), poželjno je da matrica tog sistema bude što rjedja, tj. da bude što manji broj elemenata te matrice koji su različiti od nule. To se postiže pogodnim izborom baze prostora V_k . Iz formule (7) za bilinearnu formu $a(\cdot, \cdot)$ je jasno da će koeficijent $a(\varphi_i^k, \varphi_j^k)$ biti jednak nuli kada je mjera presjeka nosača funkcija φ_i^k i φ_j^k jednaka nuli. Zbog toga se za bazu prostora V_k biraju funkcije sa malim nosačima, tzv. finitne funkcije. Za jednačine (10) dobijene primjenom metoda konačnih elemenata se koristi i naziv konačno-elementne jednačine. U ovom radu ćemo se baviti problemom rješavanja konačno-elementnih jednačina, a rezultate vezane za druge aspekte teorije konačnih elemenata ćemo direktno koristiti imajući u vidu mnogobrojne knjige posvećene metodu konačnih elemenata (vidjeti [1, 2, 8, 23, 36, 38, 43, 44]).

ОСНОВНА ОРГАНИЗАЦИЈА НАУЧНОГ РАДА
ЗА МАТЕМАТИКУ, МЕХАНИКУ И АЕРОКОСМИЈУ
В И Б Л И О Т Е К А

Б р о ј: _____

Д а т у м: _____

2. ITERACIONI METODI ZA RJEŠAVANJE VELIKIH SISTEMA LINEARNIH JEDNAČINA

Diskretizacijom eliptičkih graničnih zadataka primjenom metoda konačnih razlika ili metoda konačnih elemenata se dobijaju veliki sistemi linearnih algebarskih jednačina koji se najčešće rješavaju primjenom iteracionih metoda. U ovoj glavi ćemo navesti neke od najčešće korištenih iteracionih metoda, kako za sisteme sa regularnom, tako i za sisteme sa singularnom matricom, kao i neke rezultate vezane za brzinu konvergencije ovih metoda.

2.1. Karakteristike sistema jednačina dobijenih primjenom metoda konačnih razlika i konačnih elemenata

Za razliku od opštih sistema linearnih algebarskih jednačina, sistemi dobijeni primjenom metoda konačnih razlika i konačnih elemenata imaju određene specifičnosti:

- a) velik broj nepoznatih, b) broj nenultih elemenata matrice je mali u poredjenju sa ukupnim brojem elemenata,
- c) specifičnost mjesta na kojim se nalaze nenulti elementi,
- d) velika razbacanost spektra.

U slučaju sistema sa simetričnom matricom karakteristika d) znači da sopstvene vrijednosti matrice pripadaju intervalu koji se neograničeno povećava kada parametar diskretizacije $h \rightarrow 0$.

Ove specifičnosti eliptičkih mrežnih i konačno-elementnih jednačina nameću potrebu za primjenom ekonomičnih algoritama za njihovo numeričko rješavanje. Direktni metodi se primjenjuju, ali za vrlo usku klasu zadataka. Za rješavanje ovakvih zadataka prednost imaju iteracioni metodi.

2.2. Pregled iteracionih metoda za rješavanje sistema sa regularnom matricom

Neka je primjenom nekog od metoda za približno rješavanje eliptičkih graničnih zadataka dobijen sistem jednačina

$$(1) \quad Ay = b$$

sa regularnom matricom A ($\det A \neq 0$) koji treba riješiti.

Za matricu A ćemo još pretpostaviti da je simetrična i pozitivna (kao operator).

U svakom iteracionom metodu se polazi od neke početne aproksimacije y_0 rješenja jednačine (1) i sukcesivno se određuju aproksimacije $y_1, y_2, \dots, y_k, y_{k+1}, \dots$ gdje je k - broj iteracije. Aproksimacija y_{k+1} se izražava preko prethodnih aproksimacija rekurentnom formulom

$$y_{k+1} = F_k(y_0, y_1, \dots, y_k),$$

gdje je F_k neka funkcija koja zavisi od A , b i k .

Iteracioni metod je reda m ako svaka sljedeća aproksimacija zavisi samo od m prethodnih, tj.

$$y_{k+1} = F_k(y_{k-m+1}, y_{k-m+2}, \dots, y_k).$$

U praksi se najčešće koriste iteracioni metodi sa $m=1$ ili $m=2$. Od izbora funkcije F_k zavisi struktura iteracionih sema.

Ako je ta funkcija linearna, za odgovarajući iteracioni metod se kaže da je linearan. Metodi kod kojih funkcija ne zavisi od k se nazivaju stacionarnim iteracionim metodima. Mi ćemo razmatrati linearne iteracione metode prvog reda koji se mogu napisati u kanonskom obliku

$$(2) \quad B \frac{y_{k+1} - y_k}{\tau_{k+1}} + A y_k = b, \quad \kappa = 0, 1, \dots \quad \text{za sve } y_0,$$

gdje je B - regularna matrica, κ - broj iteracije, $\tau_1, \tau_2, \dots, \tau_{k+1}, \dots$ - iteracioni parametri, $\tau_k > 0$.

Ako je $B = E$ - jedinična matrica, imaćemo

$$(3) \quad \frac{y_{k+1} - y_k}{\tau_{k+1}} + A y_k = b, \quad \kappa = 0, 1, \dots \quad \text{za sve } y_0.$$

U ovom slučaju se y_{k+1} određuje po eksplicitnoj formuli

$$y_{k+1} = y_k - \tau_{k+1} (A y_k - b),$$

pa govorimo o eksplicitnom iteracionom metodu.

U opštem slučaju, za $B \neq E$, iteracioni metod je implicitan: za određivanje y_{k+1} treba riješiti jednačinu

$$(4) \quad B y_{k+1} = B y_k - \tau_{k+1} (A y_k - b) = \phi_k, \quad \kappa = 0, 1, \dots$$

pri čemu je prirodan zahtjev da računski posao potreban za rješavanje sistema $B y_{k+1} = \phi_k$ bude manji od računskog posla potrebnog za rješavanje sistema (1).

Ako je poznato y_0 , sukcesivno određujemo y_1, y_2, \dots . Iteracioni metod ima smisla ako on konvergira, tj. ako

$$\|y_k - y\| \rightarrow 0 \quad \text{kada } k \rightarrow \infty,$$

gdje je $\|u\| = (\sum u_i^2)^{1/2}$ vektorska norma koju ćemo i ubuduće stalno koristiti, a y tačno rješenje sistema (1).

Obično se zadaje neka greška $\epsilon > 0$ (relativna greška $\|y_k - y\| / \|y_0 - y\|$) sa kojom je potrebno naći približno rješenje sistema (1). Pri

tome se računanje završava kada je ispunjen uslov

$$\|y_n - y\| \leq \varepsilon \|y_0 - y\|.$$

Ovaj uslov je nepogodan za praktičnu provjeru jer je y nepoznat vektor, i on može biti zamijenjen uslovom

$$(5) \quad \|y_n - y\|_D \leq \varepsilon \|y_0 - y\|_D,$$

gdje je D neka simetrična i pozitivna matrica, $\|u\|_D = \sqrt{(Du, u)}$,

a $(u, v) = \sum_i u_i v_i$ skalarni proizvod dva vektora. Stavljajući

npr. $D = A^2$, dobijamo

$$\|Ay_n - b\| \leq \varepsilon \|Ay_0 - b\|.$$

Ovaj uslov je pogodan za praktičnu primjenu.

Greškom k -te iteracije ćemo zvati razliku

$$z_k = y_k - y.$$

Greški odgovara jednačina

$$B \frac{z_{k+1} - z_k}{\tau_{k+1}} + A z_k = 0, \quad n = 0, 1, \dots, \quad z_0 = y_0 - y,$$

odakle dobijamo

$$z_{k+1} = (E - \tau_{k+1} B^{-1} A) z_k$$

odnosno

$$z_{k+1} = S_{k+1} z_k, \quad S_{k+1} = E - \tau_{k+1} B^{-1} A.$$

Isključujući z_k, z_{k-1}, \dots, z_1 , imamo za $k = n-1$

$$z_n = T_n z_0, \quad T_n = S_n S_{n-1} \dots S_2 S_1.$$

Odavde dobijamo

$$\|z_n\|_D = \|T_n z_0\|_D \leq \|T_n\|_D \cdot \|z_0\|_D \quad \text{ili} \quad \|z_n\|_D \leq q_n \|z_0\|_D, \quad q_n = \|T_n\|_D.$$

Iteracioni proces ćemo završiti kada bude $q_n \leq \varepsilon$.

Ako je $n = n(\varepsilon)$ najmanji broj za koji je ispunjena nejednakost

$q_n \leq \varepsilon$, odnosno nejednakost (5), onda je ukupan broj računskih operacija potrebnih za nalaženje približnog rjesenja sistema (1) sa zadanom tačnošću ε jednak

$$Q(\varepsilon) = \sum_{k=1}^{n(\varepsilon)} Q_k,$$

gdje je Q_k - broj računskih operacija potrebnih za računanje jedne iteracije, tj. za rješavanje sistema (4). Pretpostavljajući da je $Q_k = q_0$ za sve k , dobijamo

$$Q(\varepsilon) = n(\varepsilon)q_0.$$

Ovaj broj se nastoji minimizirati, a to se postiže odgovarajućim izborom matrice B i parametara τ_{k+1} u (2).

Uzimajući $B = E$ i $\tau_{k+1} = \tau$, dobijamo najjednostavniji iteracioni metod za rješavanje (1), tzv. metod proste iteracije:

$$(6) \quad \frac{y_{k+1} - y_k}{\tau} + Ay_k = b, \quad k = 0, 1, \dots,$$

ili

$$(6) \quad y_{k+1} = Ty_k + \tau b, \quad \text{gdje je } T = E - \tau A.$$

Postoje i druge varijante metoda proste iteracije.

Uzimajući $B = D = \text{diag} A$, $\tau_{k+1} = 1$ za sve k , iz (2) dobijamo

$$(7) \quad D(y_{k+1} - y_k) + Ay_k = b, \quad k = 0, 1, \dots$$

ili

$$(7') \quad y_{k+1} = Ty_k + D^{-1}b, \quad \text{gdje je } T = E - D^{-1}A.$$

Ovo je Jacobi-jev iteracioni metod (vidjeti [25, 26, 45]).

Za $B = D = \text{diag} A$, $\tau_{k+1} = \omega$ za sve k , dobijamo Jacobi-jev

ω -metod

$$(8) \quad D \frac{y_{k+1} - y_k}{\omega} + Ay_k = b, \quad k = 0, 1, \dots$$

ili

$$(8') \quad y_{k+1} = Ty_k + \omega D^{-1}b, \quad \text{gdje je } T = E - \omega D^{-1}A.$$

U praksi se često koristi i Gauss-Seidel-ov iteracioni metod, i to u jednoj od dvije forme

$$(9) \quad (D + A^+)(y_{k+1} - y_k) + Ay_k = b, \quad k = 0, 1, \dots,$$

$$(10) \quad (D + A^-)(y_{k+1} - y_k) + Ay_k = b, \quad k = 0, 1, \dots,$$

koje se dobijaju iz (2) za $\tau_{k+1}=1$ i $B=D+A^+$, odnosno $B=D+A^-$, gdje je A^+ gornja trougaona, a A^- donja trougaona matrica sa nulama na glavnoj dijagonali, a D je dijagonalna matrica tako da je

$$A = A^- + D + A^+.$$

Iteracione šeme (9) i (10) mogu biti napisane i u obliku

$$(9') \quad y_{k+1} = T y_k + (D+A^+)^{-1} b, \quad \text{gdje je } T = E - (D+A^+)^{-1} A,$$

$$(10') \quad y_{k+1} = T y_k + (D+A^-)^{-1} b, \quad \text{gdje je } T = E - (D+A^-)^{-1} A.$$

Stavljajući u (2) $B = D + \omega A^-$, $\tau_{k+1} = \omega$ za sve k , dobijamo metod relaksacije

$$(11) \quad (D + \omega A^-) \frac{y_{k+1} - y_k}{\omega} + A y_k = b, \quad k=0,1,\dots,$$

ili

$$(11') \quad y_{k+1} = T y_k + \omega (D + \omega A^-)^{-1} b, \quad \text{gdje je } T = E - \omega (D + \omega A^-)^{-1} A.$$

Za $\omega=1$ ovo je Gauss-Seidelov metod (10), dok u slučaju $\omega > 1$ koristimo naziv metod gornje relaksacije.

Svi navedeni primjeri iteracionih metoda predstavljaju primjere tzv. stacionarnih metoda kojima odgovara kanonski zapis

$$(12) \quad B \frac{y_{k+1} - y_k}{\tau} + A y_k = b, \quad k=0,1,\dots$$

sa regularnom matricom B .

U vezi sa konvergencijom tih metoda važi

T e o r e m a 1. (vidjeti [45,46])

Neka je A simetrična i pozitivna matrica reda N i neka je ispunjen uslov

$$(13) \quad (Bx, x) > \frac{\tau}{2} (Ax, x) \quad \text{ili} \quad B > \frac{\tau}{2} A \quad \text{za sve } x \in \mathbb{R}^N.$$

Tada iteracioni metod (12) konvergira, tj.

$$\|z_k\|_A = \|y_k - y\|_A \rightarrow 0 \quad \text{kada } k \rightarrow \infty \quad \square$$

Uslov (13) je dovoljan uslov konvergencije i on je pogodan za praktično utvrđivanje kada neki metod konvergira.

Zapisujući iteracione metode u obliku

$$(14) \quad y_{k+1} = T y_k + \ell,$$

za neku matricu T i neki vektor ℓ , moguće je formulirati neophodan i dovoljan uslov konvergencije tog iteracionog metoda.

T e o r e m a 2. (vidjeti [25, 26])

Iteracioni metod (14) konvergira ako i samo ako je

$$(15) \quad \rho(T) < 1,$$

gdje je $\rho(T) = \max\{|\lambda| : \lambda \text{ sopstvena vrijednost } T\}$ -spektralni radijus matrice T . \square

S obzirom na (14), za grešku će važiti formula

$$z_{k+1} = T z_k = T^{k+1} z_0, \quad z_0 = y_0 - y,$$

pa će metod (14) konvergirati kada je $\|T\| < 1$, gdje je

$\|T\| = \sqrt{\rho(T^T T)}$ -spektralna norma matrice T . Uslov

$$\|T\| < 1$$

je dovoljan, ali ne i neophodan za konvergenciju iteracionog metoda (14).

Primjenom ovih teorema je moguće dokazati konvergenciju razmatranih iteracionih metoda.

Šod metoda proste iteracije (6) je $B = E$, pa uzimajući u

obzir da je $E \gg A / \|A\|$, imamo

$$B - \frac{\tau}{2} A = E - \frac{\tau}{2} A \gg \frac{1}{\|A\|} A - \frac{\tau}{2} A = \left(\frac{1}{\|A\|} - \frac{\tau}{2} \right) A > 0$$

za $\frac{1}{\|A\|} - \frac{\tau}{2} > 0$. Kako je za simetričnu pozitivnu matricu A ,

$\|A\| = \lambda_N = \max\{\lambda : \lambda \text{ sopstvena vrijednost } A\}$, metod proste

iteracije konvergira za sve vrijednosti τ za koje je

$$\tau < \frac{2}{\lambda_N}.$$

Što se tiče Jacobi-jevih metoda (7) i (8), pokazuje se da ako konvergira metod (7), tada Jacobi-jev ω -metod konvergira za $0 < \omega \leq 1$ (vidjeti [26]).

U slučaju Gauss-Seidel-ovog metoda (9) je $B = D + A^+$, $\tau = 1$, pa je

$$\begin{aligned} B - \frac{\tau}{2} A &= D + A^+ - \frac{1}{2} A = D + A^+ - \frac{1}{2} (A^- + D + A^+) = \\ &= \frac{1}{2} D + \frac{1}{2} (A^+ - A^-), \end{aligned}$$

odakle dobijamo

$$\begin{aligned} ((B - \frac{1}{2} A)x, x) &= \frac{1}{2} (Dx, x) + \frac{1}{2} ((A^+ - A^-)x, x) = \frac{1}{2} (Dx, x) + \\ &+ \frac{1}{2} (A^+x, x) - \frac{1}{2} (A^-x, x) = \frac{1}{2} (Dx, x) \end{aligned}$$

jer je zbog simetričnosti matrice A $(A^+x, x) = (A^-x, x)$.

Zbog pretpostavke da je matrica A pozitivna, bice i matrica D pozitivna, što se vidi iz uslova

$$(A\zeta, \zeta) = (D\zeta, \zeta) = a_{ii} (\zeta^i)^2 > 0, \quad \text{tj. } a_{ii} > 0 \quad i=1, \dots, N,$$

za $\zeta = (0, \dots, \zeta^i, \dots, 0)$.

Na osnovu ovoga zaključujemo da je za metod (9) ispunjen uslov konvergencije (13). I metod (10) tada konvergira, što se dokazuje analogno.

Konačno, za metod relaksacije je $B = D + \omega A^-$, $\tau = \omega$, pa je

$$\begin{aligned} B - \frac{\tau}{2} A &= D + \omega A^- - \frac{\omega}{2} (A^- + D + A^+) = \\ &= (1 - \frac{\omega}{2}) D + \frac{\omega}{2} (A^- - A^+), \end{aligned}$$

odakle dobijamo da je

$$((B - \frac{\tau}{2} A)x, x) = (1 - \frac{\omega}{2}) (Dx, x) > 0 \quad \text{za } 0 < \omega < 2.$$

Na taj način, u slučaju simetrične i pozitivne matrice A , metod relaksacije konvergira za sve ω za koje je $0 < \omega < 2$.

Sama činjenica da neki iteracioni metod konvergira nije dovoljna da bi se sudilo o njegovoj pogodnosti za praktičnu primjenu. Cilj je da se dobiju ekonomični iteracioni metodi. Računski posao $Q(\epsilon)$ se može minimizirati ako se biraju iteracioni metodi koji brzo konvergiraju, tj. koji poslije malog broja iteracionih koraka $n(\epsilon)$ dostižu zadanu tačnost ϵ . Na brzinu konvergencije metoda proste iteracije (6) se može uticati pogodnim izborom parametra τ . Pokazuje se da je u tom smislu optimalna vrijednost parametra τ data sa

$$\tau = \tau_0 = \frac{2}{\gamma_1 + \gamma_2},$$

gdje su γ_1 i γ_2 ocjena odozdo najmanje i ocjena odozgo najveće sopstvene vrijednosti matrice A respektivno (vidjeti [45, 46, 47]). Za tako izabrani parametar τ se dobija ocjena

$$\|y_{k+1} - y\| \leq q_n \|y_k - y\|,$$

gdje je $q_n = \frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1}$ - faktor konvergencije.

Kako je

$$\|y_n - y\| \leq \left(\frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1}\right)^n \|y_0 - y\|,$$

nejednakost

$$\|y_n - y\| \leq \epsilon \|y_0 - y\|$$

će biti ispunjena za sve

$$n \geq n(\epsilon) = \left\lceil \frac{\ln \epsilon^{-1}}{\ln q_n^{-1}} \right\rceil + 1.$$

Za sisteme oblika (1) dobijene diskretizacijom eliptičkih graničnih zadataka sa parametrom diskretizacije h je najčešće $\gamma_1 = C_1$, $\gamma_2 = C_2 h^{-2}$, pa je u tim slučajevima $n(\epsilon) = O(h^{-2} \ln \epsilon^{-1})$, pa kako je za navedene metode $q_n = O(h^{-2})$, ukupan broj račun-

skih operacija potrebnih za određivanje aproksimacije sistema (1) sa tačnošću ϵ će biti $Q(\epsilon) = O(h^{-4} \ln \epsilon^{-1})$. Faktor konvergencije q metoda proste iteracije zavisi od h na sljedeći način

$$q = 1 - O(h^2),$$

pa to znači da ćemo za male vrijednosti parametra h imati vrlo sporu konvergenciju. Pokazuje se da Gauss-Seidel-ov metod nešto brže konvergira (vidjeti [26, 45]), ali da je i ta konvergencija spora i da faktor konvergencije zavisi od h na sličan način kao i kod metoda proste iteracije. Kod oba ova metoda je $n(\epsilon)$ reda N^2 . Taj broj se bitno smanjuje ako se koristi metod relaksacije sa optimalnim relaksacionim parametrom ω (vidjeti [26]). Pokazuje se da je u tom slučaju pri rješavanju mrežnih i konačno-elementnih jednačina

$n(\epsilon) = O(h^{-1} \ln \epsilon^{-1})$, $Q(\epsilon) = O(h^{-3} \ln \epsilon^{-1})$, dok faktor konvergencije q na sljedeći način zavisi od diskretizacionog parametra h

$$q = 1 - O(h).$$

Istu brzinu konvergencije kao i metod relaksacije sa optimalnim relaksacionim parametrom ω ima i Richardson-ov metod

$$\frac{y_{k+1} - y_k}{\tau_{k+1}} + A y_k = b, \quad k = 0, 1, \dots$$

sa tzv. čebiševljevim izborom parametara τ_{k+1} (vidjeti [45, 47]). Na kraju ovog pregleda iteracionih metoda za rješavanje sistema sa regularnom matricom dobijenih diskretizacijom eliptičkih graničnih zadataka, pomenimo i metod promjenljivih pravaca Peaceman-Rachford-a (vidjeti [20]), za koji je $n(\epsilon) = O(h \ln \epsilon^{-1})$, a $Q(\epsilon) = O(h^{-3} \ln \epsilon^{-1})$. Što predstavlja poboljšanje čak i

u odnosu na Richardson-ov i metod relaksacije.

Zajednički nedostatak svih navedenih metoda je zavisnost broja $n(\epsilon)$ od parametra h , tj. zavisnost faktora konvergencije q_h od broja h , gdje je q_h faktor smanjenja greške na jednom iteracionom koraku

$$\|z_k\|_D \leq q_h \|z_{k-1}\|_D,$$

gdje je D neka simetrična i pozitivna matrica.

Postavlja se pitanje odredjivanja iteracionog metoda kod koga faktor konvergencije neće zavisiti od diskretizacionog parametra h .

2.3. Iteracioni metodi za rješavanje sistema sa singularnom matricom

Iteracioni metodi su se u početku primjenjivali samo za rješavanje regularnih sistema jednačina. Jedan od prvih koji je počeo koristiti iteracione metode za rješavanje singularnih linearnih sistema bio je H.B.Keller (vidjeti [13]). On je prvi i formulisao neophodan i dovoljan uslov konvergencije iteracionih metoda za singularne sisteme, koji je kasnije preformulisan u terminima spektralnog radijusa.

T e o r e m a 3. (vidjeti [16, 21])

Neophodan i dovoljan uslov za konvergenciju iteracionog procesa

$$y_{k+1} = T y_k + l$$

za rješavanje singularnih linearnih sistema je

$$\rho(T) < 1,$$

pri čemu, ako je moduo neke sopstvene vrijednosti λ jednak 1, tada je $\lambda = 1$ i to je prost pol rezolventnog operatora

$$R(\lambda, T) = (\lambda E - T)^{-1}. \quad \square$$

posmatraćemo sistem linearnih algebarskih jednačina

$$(16) \quad Ay = b$$

sa simetričnom, nenegativno odredjenom matricom A reda N , za koju je $\det A = 0$. U početku ćemo pretpostavljati da je taj sistem saglasan, tj. da ima rješenja. Uslov singularnosti matrice A znači da sistem $Ay = 0$ ima i neko nenulto rješenje. Sa $\ker A$ ćemo označavati jezgro matrice A , tj.

$$\ker A = \{y \in \mathbb{R}^N : Ay = 0\},$$

a sa $\text{im} A$ skup vektora oblika $u = Ay$, tj.

$$\text{im} A = \{u \in \mathbb{R}^N : u = Ay \text{ za neko } y \in \mathbb{R}^N\}.$$

S obzirom na ove oznake i simetričnost matrice A , važi sljedeće razlaganje

$$(17) \quad \mathbb{R}^N = \ker A \oplus \text{im} A.$$

Ovo znači da se svaki vektor $y \in \mathbb{R}^N$ može predstaviti u obliku $y = \tilde{y} + \bar{y}$, gdje je $\tilde{y} \in \ker A$, $\bar{y} \in \text{im} A$ i $(\tilde{y}, \bar{y}) = 0$. Uslov saglasnosti sistema (16) znači da $b \in \text{im} A$, tj. da je vektor b ortogonalan na potprostor $\ker A$. Rješenje sistema (16) sa minimalnom normom je jedinstveno i njega ćemo zvati normalnim rješenjem. To će biti rješenje $y^* = \tilde{y}^* + \bar{y}^*$ čija je komponenta \tilde{y}^* iz $\ker A$ jednaka nuli, tj. $\tilde{y}^* = 0$. Za rješavanje sistema (16) možemo koristiti iteracionu proceduru

$$(18) \quad B \frac{y_{k+1} - y_k}{\tau} + Ay_k = b, \quad k = 0, 1, \dots, \text{ za sve } y_0,$$

gdje je B - simetrična i pozitivno odredjena matrica, τ - iteracioni parametar.

T e o r e m a 4. (vidjeti [39])

Iteracioni proces (18) sa parametrom

$$(19) \quad \tau = \frac{2}{\gamma_1 + \gamma_2},$$

gdje je $\gamma_1 > 0$ - ocjena odozdo najmanje nenulte, a $\gamma_2 > \gamma_1$ - ocjena odozgo najveće sopstvene vrijednosti matrice $C = B^{-1/2} A B^{-1/2}$, konvergira brzinom geometrijske progresije sa količnikom

$$(20) \quad \rho_k = \frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1}$$

ka jednom od rješenja sistema (16) odredjenom početnom iteracijom y_0 .

D o k a z.

Iteracioni proces (18) se može predstaviti u obliku

$$(21) \quad \frac{x_{k+1} - x_k}{\tau} + C x_k = g, \quad x_0 = B^{1/2} y_0,$$

gdje je $x = B^{1/2} y$, $g = B^{-1/2} b$, $C = B^{-1/2} A B^{-1/2}$. Ovo je moguće zbog pretpostavke da je B pozitivno odredjena, pa postoji $B^{-1/2}$ i $B^{1/2}$, i one su takodje pozitivno odredjene matrice. Šemu (21) možemo shvatiti kao iteracionu proceduru za rješavanje sistema

$$(22) \quad Cx = g.$$

Označavajući sa x^* normalno rješenjetog sistema, a sa x_k κ -tu iteraciju odredjenu po formuli (21), za grešku $z_k = x_k - x^*$ važi

$$\frac{z_{k+1} - z_k}{\tau} + C z_k = 0, \quad \kappa = 0, 1, \dots, \quad z_0 = x_0 - x^*$$

ili

$$(23) \quad z_{k+1} = (E - \tau C) z_k, \quad \kappa = 0, 1, \dots, \quad z_0 = x_0 - x^*.$$

Kako je matrica C simetrična i nenegativno odredjena, ona

ima sistem ortogonalnih sopstvenih vektora $\{\psi_j\}$, $j=1,2,\dots,N$, koji odgovaraju sopstvenim vrijednostima

$$0 = \mu_1 = \dots = \mu_r \leq \gamma_1 \leq \mu_{r+1} \leq \mu_{r+2} \leq \dots \leq \mu_N \leq \gamma_2.$$

Jasno je da je $\ker C = \text{span}\{\psi_1, \dots, \psi_r\}$, a $\text{im} C = \text{span}\{\psi_{r+1}, \dots, \psi_N\}$, gdje $\text{span}\{\mu_1, \dots, \mu_k\}$ označava prostor generisan vektorima $\mu_1, \mu_2, \dots, \mu_k$.

Greški \tilde{z}_k odgovara razvoj po sopstvenim vektorima

$$\tilde{z}_k = \sum_{j=1}^N C_j^{(k)} \psi_j, \quad k=0,1,\dots$$

Stavljajući taj razvoj u (23), dobijamo

$$\sum_{j=1}^N C_j^{(k+1)} \psi_j = \sum_{j=1}^N (1 - \tau \mu_j) C_j^{(k)} \psi_j,$$

odakle, zbog ortogonalnosti sopstvenih vektora ψ_j dobijamo

$$C_j^{(k+1)} = (1 - \tau \mu_j) C_j^{(k)} = \dots = (1 - \tau \mu_j)^{k+1} C_j^{(0)}, \quad j = r+1, \dots, N$$

i

$$C_j^{(k+1)} = C_j^{(k)} = \dots = C_j^{(0)}, \quad j = 1, 2, \dots, r.$$

Zaključujemo da komponenta $\tilde{z}_k = \sum_{j=1}^r C_j^{(k)} \psi_j$ greške \tilde{z}_k iz $\ker C$ ostaje nepromijenjena u toku iteracionog procesa i da je ona jednaka projekciji početne greške \tilde{z}_0 na $\ker C$. Kako rješenja sistema (22) mogu imati proizvoljne projekcije na $\ker C$, za konvergenciju iteracionog procesa (21) je dovoljno dokazati da projekcija greške \tilde{z}_k na $\text{im} C$ teži nuli kada $k \rightarrow \infty$, tj. da je

$$\lim_{k \rightarrow \infty} |C_j^{(k)}| = 0 \quad \text{za } j = r+1, \dots, N.$$

Da bi ovaj uslov bio ispunjen potrebno je da bude

$$|1 - \tau \mu_j| < 1 \quad \text{za } j = r+1, \dots, N$$

odnosno

$$0 < \tau < \frac{2}{\mu_N}.$$

Parametar τ izabran po formuli (19) očigledno zadovoljava ovaj uslov.

Koristeći razlaganje (17) i predstavljajući \bar{x}_k u obliku

$$\bar{x}_k = \tilde{x}_k + \bar{x}_k,$$

iz (23) dobijamo

$$\tilde{x}_{k+1} = \tilde{x}_k, \quad \bar{x}_{k+1} = \left(E - \frac{2}{\gamma_1 + \gamma_2} C \right) \bar{x}_k,$$

zbog čega je

$$\|\bar{x}_{k+1}\| \leq \left\| E - \frac{2}{\gamma_1 + \gamma_2} C \right\| \|\bar{x}_k\|,$$

gdje je

$$\left\| E - \frac{2}{\gamma_1 + \gamma_2} C \right\| = \rho \left(E - \frac{2}{\gamma_1 + \gamma_2} C \right).$$

Kako je

$$\begin{aligned} \rho \left(E - \frac{2}{\gamma_1 + \gamma_2} C \right) &= \max_{r+1 \leq j \leq n} \left| 1 - \frac{2}{\gamma_1 + \gamma_2} \mu_j \right| \leq \max_{\gamma_1 \leq \mu_j \leq \gamma_2} \left| 1 - \frac{2}{\gamma_1 + \gamma_2} \mu_j \right| = \\ &= \frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1} = q_n. \end{aligned}$$

Na taj način se projekcija greške na $\text{im} C$ smanjuje brzinom geometrijske progresije sa količnikom q_n , dok projekcija greške na $\ker C$ ostaje nepromijenjena i određuje se polaznom iteracijom. Naime, ako je $x_0 = \tilde{x}_0 + \bar{x}_0$, $\tilde{x}_0 \in \ker C$, $\bar{x}_0 \in \text{im} C$, iteracioni proces (21) konvergira ka rješenju $x = \tilde{x}_0 + x^*$ sistema (22). Jasno je da u slučaju nulte početne iteracije x_0 , iteracioni proces konvergira normalnom rješenju.

Lako se pokazuje da je izbor $\tau = \frac{2}{\gamma_1 + \gamma_2}$ optimalan jer to daje najbržu konvergenciju. Na taj način su dobijeni rezultati analogni rezultatima dobijenim za metod proste iteracije za sisteme sa regularnom matricom.

I u ovom slučaju se pri diskretizaciji singularnih eliptičkih graničnih zadataka za koje je $\gamma_1 = C_1, \gamma_2 = C_2 h^{-2}$, dobija

$n(\varepsilon) = O(h^{-2} \ln \varepsilon^{-1})$ i $Q(\varepsilon) = O(h^{-4} \ln \varepsilon^{-1})$, što pokazuje da iteracioni metodi za rješavanje sistema sa singularnim matricama imaju isti nedostatak kao i iteracioni metodi za rješavanje regularnih zadataka, a to je zavisnost faktora konvergencije q od diskretizacionog parametra h .

U praksi se diskretizacijom singularnih graničnih zadataka zbog grešaka računanja najčešće dobijaju nesaglasni sistemi, a često je i nemoguće utvrditi da li je sistem saglasan ili ne. Razmotrićemo primjenu iteracionih metoda za rješavanje nesaglasnih sistema oblika (16). Koristeći (17), dobijamo razlaganje vektora b u obliku $b = \tilde{b} + \bar{b}$ gdje $\tilde{b} \in \ker A$, $\bar{b} \in \text{im} A$. Sistem će biti nesaglasan ako je $\tilde{b} \neq 0$. Za nesaglasne sisteme uvodimo pojam uopštenog rješenja. Uopštenim rješenjem (kvazirješenjem, pseudorješenjem) sistema (16) se naziva vektor $u \in \mathbb{R}^n$ za koji je

$$\|Au - b\| = \min_{y \in \mathbb{R}^n} \|Ay - b\|,$$

a odredjivanje uopštenog rješenja u stvari predstavlja metod najmanjih kvadrata (vidjeti [14]).

Pokazuje se da je skup uopštenih rješenja uvijek neprazan i da uopšteno rješenje sistema (16) predstavlja rješenje uvijek saglasnog sistema $Au = \bar{b}$. Zbog toga uopšteno rješenje nije jedinstveno i ono se odredjuje do na vektor iz $\ker A$. Uopšteno rješenje sa minimalnom normom se naziva normalnim rješenjem. Ono je jedinstveno i njegova projekcija na $\ker A$ je 0 , tj. normalno rješenje pripada $\text{im} A$. Uopšteno rješenje nesaglasnog sistema (16) se može definisati i kao rješenje uvijek saglasnog sistema

$$A^2 u = A b.$$

Formiranje ovog saglasnog sistema se u praksi ne preporučuje jer to povećava računski posao. Za dobijanje uopštenog rješenja nesaglasnog sistema (16) se može iskoristiti iteracioni proces (18) sa matricom B koja je još komutativna sa A , tj. $AB = BA$. Iteracioni proces cemo završiti kada bude ispunjen uslov

$$\| \tau_{k+1} - \tau_k \| \leq \varepsilon$$

gdje je $\tau_k = b - Ay_k$ - defekt, a $\varepsilon > 0$ zadani broj.

T e o r e m a 5.

Iteracioni proces (18) za rješavanje nesaglasnog sistema (16) sa parametrom τ odredjenim po formuli (19) daje niz vektora y_k , $k = 1, 2, \dots$, za koje važi relacija

$$(24) \quad \lim_{k \rightarrow \infty} \| \tau_{k+1} - \tau_k \| = 0.$$

D o k a z.

Ponovo cemo iteracionu šemu (18) napisati u obliku

$$\frac{x_{k+1} - x_k}{\tau} + C x_k = g, \quad x_0 = B^{1/2} y_0,$$

gdje je $x = B^{1/2} y$, $g = B^{-1/2} b = \tilde{g} + \bar{g}$, $\tilde{g} \in \ker C$, $\bar{g} \in \text{im} C$. Zbog uslova $AB = BA$, bice $\ker C = \ker A$, $\text{im} C = \text{im} A$. Označavajući sa x^* normalno rješenje nesaglasnog sistema $Cx = g$, a sa $z_k = x_k - x^*$, dobijamo

$$(25) \quad z_{k+1} = (E - \tau C) z_k + \tau \tilde{g}, \quad z_0 = x_0 - x^*.$$

Koristeći razvoje za z_k i \tilde{g} po sopstvenim vektorima $\{\psi_j\}$ matrice C

$$z_k = \sum_{j=1}^n C_j^{(k)} \psi_j, \quad \tilde{g} = \sum_{j=1}^r \beta_j \psi_j,$$

dobijamo iz (25)

$$(26) \quad \begin{aligned} C_j^{(k+1)} &= (1 - \tau \mu_j) C_j^{(k)} = \dots = (1 - \tau \mu_j)^{k+1} C_j^{(0)}, \quad j = r+1, \dots, N \\ C_j^{(k+1)} &= C_j^{(k)} + \tau \beta_j = \dots = C_j^{(0)} + (k+1) \tau \beta_j, \quad j = 1, 2, \dots, r \end{aligned}$$

Ranije je bilo dokazano da se projekcija greške na $\text{im } C$ smanjuje brzinom geometrijske progresije čiji je količnik $q_N = \frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1}$.

Iz (26) slijedi da je

$$\tilde{x}_k = \tilde{x}_0 + k \tau \tilde{q},$$

odakle zaključujemo da će razmatrani iteracioni proces divergirati zbog stalnog nagomilavanja greške u $\ker C$, tj.

imaćemo $\|\tilde{x}_k\| \rightarrow \infty$, $k \rightarrow \infty$. Ali mi želimo da dokažemo relaciju

(24). U tom cilju ćemo ocijeniti izraz $\|B^{-\frac{1}{2}} r_{k+1} - B^{-\frac{1}{2}} r_k\|$:

$$\begin{aligned} \|B^{-\frac{1}{2}} r_{k+1} - B^{-\frac{1}{2}} r_k\| &= \|B^{-\frac{1}{2}} (b - Ay_{k+1}) - B^{-\frac{1}{2}} (b - Ay_k)\| = \\ &= \|B^{-\frac{1}{2}} A B^{-\frac{1}{2}} x_{k+1} - B^{-\frac{1}{2}} A B^{-\frac{1}{2}} x_k\| = \|C x_{k+1} - C x_k\| = \\ &= \|C(\tilde{x}_{k+1} + \bar{x}_{k+1}) - C(\tilde{x}_k + \bar{x}_k)\| = \|C\tilde{x}_{k+1} - \bar{q} - C\tilde{x}_k + \bar{q}\| = \\ &= \|C(\tilde{x}_{k+1} - \tilde{x}_k)\| = \|C(\tilde{x}_{k+1} - \tilde{x}_k)\|. \end{aligned}$$

Kako je

$$\begin{aligned} \|C(\tilde{x}_{k+1} - \tilde{x}_k)\| &\leq \gamma_2 \|\tilde{x}_{k+1} - \tilde{x}_k\| \leq \gamma_2 (\|\tilde{x}_{k+1}\| + \|\tilde{x}_k\|) \leq \\ &\leq \gamma_2 (q_N^{k+1} \|\tilde{x}_0\| + q_N^k \|\tilde{x}_0\|) = \gamma_2 q_N^k (1 + q_N) \|\tilde{x}_0\|, \end{aligned}$$

imaćemo ocjenu

$$\|B^{-\frac{1}{2}} r_{k+1} - B^{-\frac{1}{2}} r_k\| \leq \gamma_2 q_N^k (1 + q_N) \|\tilde{x}_0\|.$$

Na osnovu ove ocjene i nejednakosti

$$\nu \|r_{k+1} - r_k\| \leq \|B^{-\frac{1}{2}} r_{k+1} - B^{-\frac{1}{2}} r_k\|,$$

gdje je ν najmanja sopstvena vrijednost pozitivne matrice $B^{-\frac{1}{2}}$,

dobijamo (24), što je i trebalo dokazati. \square

Neka je $\{y_k\}$ niz određen po iteracionoj semi (18). Definišimo novi niz $\{u_k\}$ po formuli

$$(27) \quad u_k = y_k - \kappa (y_{k+1} - y_k).$$

T e o r e m a 6.

Niz vektora u_k određenih po formuli (27) konvergira ka uopštenom rješenju u sistema (16) određenom početnom iteracijom y_0 .

D o k a z.

Neka je $u = \tilde{y}_0 + \mu^*$ uopšteno rješenje sistema (16) čija je projekcija na $\ker A$ jednaka \tilde{y}_0 , dok smo sa $\mu^* \in \text{im } A$ označili normalno rješenje sistema (16). Tada je $B^{\frac{1}{2}}\mu = \tilde{x}_0 + x^*$, pa je

$$\begin{aligned} u_k - u &= y_k - \kappa (y_{k+1} - y_k) - u = B^{-\frac{1}{2}} [x_k - \kappa (x_{k+1} - x_k)] - \mu = \\ &= B^{-\frac{1}{2}} [x_k - \kappa (x_{k+1} - x_k) - B^{\frac{1}{2}}\mu] = B^{-\frac{1}{2}} [\tilde{x}_k + \bar{x}_k - \kappa (\tilde{x}_{k+1} + \bar{x}_{k+1} - \\ &\quad - \tilde{x}_k - \bar{x}_k) - \tilde{x}_0 - x^*] = B^{-\frac{1}{2}} [\tilde{x}_0 + \kappa \tilde{v}\tilde{g} + \bar{x}_k - \kappa (\tilde{x}_0 + \kappa \tilde{v}\tilde{g} + \tilde{v}\tilde{g} + \bar{x}_{k+1} - \\ &\quad - \tilde{x}_0 - \kappa \tilde{v}\tilde{g} - \bar{x}_k) - \tilde{x}_0 - x^*] = B^{-\frac{1}{2}} [\bar{x}_k - x^* - \kappa (\bar{x}_{k+1} - \bar{x}_k)] = \\ &= B^{-\frac{1}{2}} [\bar{z}_k - \kappa (\bar{z}_{k+1} - \bar{z}_k)], \end{aligned}$$

odakle se dobija ocjena

$$\begin{aligned} \|u_k - u\| &= \|B^{-\frac{1}{2}} [\bar{z}_k - \kappa (\bar{z}_{k+1} - \bar{z}_k)]\| \leq \|B^{-\frac{1}{2}}\| \cdot [\|\bar{z}_k\| + \kappa (\|\bar{z}_{k+1}\| + \|\bar{z}_k\|)] \leq \\ &\leq \|B^{-\frac{1}{2}}\| \cdot [1 + \kappa (1+q)] \|\bar{z}_k\| \leq \|B^{-\frac{1}{2}}\| [1 + \kappa (1+q)] q^k \|\bar{z}_0\| \end{aligned}$$

tj.

$$\|u_k - u\| \leq \|B^{-\frac{1}{2}}\| \cdot [1 + \kappa (1+q)] q^k \|\bar{z}_0\|.$$

Odavde, uzimajući u obzir da je $q < 1$ slijedi

$$\lim_{k \rightarrow \infty} \|u_k - u\| = 0. \quad \square$$

Formiranje niza vektora u_k po formuli (27) se vrši sa ciljem

uništavanja komponente vektora y_k iz $\ker A$ koja se stalno povećava. Pri tome se to uvrštavanje ne mora vršiti na svakom koraku, već se poslije nekog broja k_0 iteracija može određivati vektor u_k . Jasno je da faktor konvergencije ovih iteracionih metoda za rješavanje mrežnih i konačno-elementnih nesaglasnih jednačina takodje zavisi od parametra h . Dokazacemo da je moguće otkloniti taj nedostatak.

ОСРОВОК ПРОИЗВОДЊА УСТРОЈЕНОГ РАДА
ЗА МАТЕМАТИКУ, МЕХАНИКУ И АСТРОНОМИЈУ
С. С. М. Т. С. А.

Б р о ј: _____

Д а т у м: _____

3. MNOGOSTEPENI ITERACIONI METODI

3.1. Uvodne napomene

U prethodnoj glavi smo vidjeli da je glavni nedostatak navedenih iteracionih metoda pri rješavanju diskretnih eliptičkih graničnih zadataka zavisnost faktora konvergencije od diskretizacionog parametra h . Fedorenko je 1961.g. predložio jedan iteracioni metod za rješavanje diskretnih eliptičkih zadataka koji je on nazvao relaksacionim, koji se bazira na sasvim novoj ideji - korištenju krupnijih mreža (vidjeti [58]). Koristeći tu ideju pri rješavanju Dirichlet-ovog zadatka za Poisson-ovu jednačinu u kvadratnoj oblasti, Fedorenko je dobio ocjenu brzine konvergencije svog relaksacionog metoda koja ne zavisi od koraka najsitnije mreže (vidjeti [59]). On je takodje dokazao da je za smanjenje norme početne greške ϵ^{-1} puta potrebno $C \cdot N^2$ aritmetičkih operacija, gdje je N^2 - broj čvorova najsitnije mreže, što je predstavljalo poboljšanje i u odnosu na, do tada najefikasniji metod za rješavanje mrežnih aproksimacija za Poisson-ovu jednačinu - metod promjenljivih pravaca Peaceman-Rachford-a, za koji je bilo potrebno $C N^2 \ln N$ aritmetičkih operacija za postizanje istog efekta. Kasnije je Bahvalov (vidjeti [28]) ispitivao konvergenciju Fedorenkovog metoda i dobio iste rezultate za diferencijski analog prvog graničnog zadatka u pravougaoniku za opštu eliptičku parcijalnu jednačinu sa glatkim koeficijentima

$$\alpha_{11} \frac{\partial^2 u}{\partial x^2} + 2\alpha_{12} \frac{\partial^2 u}{\partial x \partial y} + \alpha_{22} \frac{\partial^2 u}{\partial y^2} + \alpha_1 \frac{\partial u}{\partial x} + \alpha_2 \frac{\partial u}{\partial y} + \alpha u = f$$

Na kraju je Astrahancev (vidjeti [27]) dobio analogan rezultat za diferencijsku aproksimaciju trećeg graničnog zadatka za samokonjugovanu eliptičku jednačinu u proizvoljnoj dvodimenzionoj oblasti sa glatkom granicom.

Iako Fedorenkova ideja nije ostala nezapažena, čini se da ona ipak u praksi nije dovoljno korištena. Možda je na to uticala složenija organizacija računanja nego što je to slučaj za neki od klasičnih iteracionih metoda. Višemrežna Fedorenkova ideja je aktuelizirana ponovo 80-ih godina kada je Brandt otkrio Fedorenkove radove [58, 59]. Koristeći tu ideju, on je došao do mnogostepenih metoda za rješavanje diskretnih eliptičkih graničnih zadataka (vidjeti [6]).

3.2. Oznake, ideja i komponente mnogostepenih metoda

Mnogostepeni iteracioni metodi su se prvo počeli primjenjivati za rješavanje mrežnih jednačina, da bi se tek kasnije počeli koristiti za konačno-elementne jednačine. Zbog jednostavnosti, objasnićemo osnovne ideje mnogostepenih metoda na primjeru mrežnih jednačina, iako ćemo glavne rezultate ovog rada dobiti za konačno-elementne jednačine.

Neka eliptičkom graničnom zadatku

$$\begin{aligned} Lu &= f & u & \in \Omega \\ lu &= 0 & \text{na } \partial\Omega \end{aligned}$$

odgovara diskretni zadatak

$$(1) \quad L_h u_h = f_h \quad \text{na } \Omega_h,$$

gdje je $\Omega_h \subset \bar{\Omega}$ mreža. Pretpostavimo da je u operatorsku jednačinu (1) uključen i diskretni analog graničnog uslova.

L_h je linearni diferencijalski operator definisan na prostoru mrežnih funkcija $\mathcal{G}(\Omega_h)$ odredjenih na mreži Ω_h , a u_h i f_h su mrežne funkcije iz $\mathcal{G}(\Omega_h)$. U opštem slučaju je

$$L_h W_h(x) = \sum_{k \in V} s_k W_h(x + kh),$$

gdje je $k = (k_1, k_2) \in \mathbb{Z}^2$, $V \subset \mathbb{Z}^2$, $x \in \Omega_h$, \mathbb{Z} je skup cijelih brojeva, a s_k su neki koeficijenti.

Ovo je moguće zapisati i u obliku

$$L_h W_h(x) = [s_k]_h W_h(x) = \begin{bmatrix} \vdots & & \\ s_{-1,1} & s_{0,1} & s_{1,1} \\ \dots & s_{-1,0} & s_{0,0} & s_{1,0} & \dots \\ s_{-1,-1} & s_{0,-1} & s_{1,-1} \\ \vdots & & \end{bmatrix}_h W_h(x)$$

Najčešće se koriste diferencijalski operatori kojima odgovaraju 5-tačkaste i 9-tačkaste šeme

$$\begin{bmatrix} s_{0,1} \\ s_{-1,0} & s_{0,0} & s_{1,0} \\ s_{0,-1} \end{bmatrix}_h, \quad \begin{bmatrix} s_{-1,1} & s_{0,1} & s_{1,1} \\ s_{-1,0} & s_{0,0} & s_{1,0} \\ s_{-1,-1} & s_{0,-1} & s_{1,-1} \end{bmatrix}_h.$$

Identifikovaćemo operator L_h sa $[s_k]_h$, tj. stavljamo

$$L_h \hat{=} [s_k]_h.$$

Označavajući sa $\mathcal{G}(\Omega_{2h})$ prostor mrežnih funkcija odredjenih na mreži Ω_{2h} sa dva puta većim korakom, definisamo operatore koji uspostavljaju vezu između prostora $\mathcal{G}(\Omega_h)$ i $\mathcal{G}(\Omega_{2h})$. Operatorom restrikcije (suženja) ćemo nazivati operator

$$I_h^{2h} : \mathcal{G}(\Omega_h) \rightarrow \mathcal{G}(\Omega_{2h})$$

koji se definiše na sljedeći način:

$$(I_{2h}^{2h} w_h)(x) = \sum_{\kappa \in V} \hat{t}_\kappa w_h(x + \kappa h), \quad x \in \Omega_{2h}, \quad V \subset \mathbb{Z}^2,$$

tj.

$$I_{2h}^{2h} \hat{=} [\hat{t}_\kappa]_{2h}^{2h} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \hat{t}_{-1,1} & \hat{t}_{0,1} & \hat{t}_{1,1} \\ \vdots & \vdots & \vdots \\ \dots & \hat{t}_{-1,0} & \hat{t}_{0,0} & \hat{t}_{1,0} & \dots \\ \vdots & \vdots & \vdots \\ \hat{t}_{-1,-1} & \hat{t}_{0,-1} & \hat{t}_{1,-1} \\ \vdots & \vdots & \vdots \end{bmatrix}_{2h}^{2h}$$

Najčešće se za operator restrikcije uzima

$$I_{2h}^{2h} \hat{=} \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}_{2h}^{2h}.$$

Operator interpolacije (produženja)

$$I_{2h}^h : \mathcal{G}(\Omega_{2h}) \rightarrow \mathcal{G}(\Omega_h)$$

se definiše na nešto složeniji način. On proizvoljnoj funkciji $w_{2h} \in \mathcal{G}(\Omega_{2h})$ pridružuje funkciju $w_h \in \mathcal{G}(\Omega_h)$ koju odredjimo na sljedeci način:

$$w_h(x) = \sum_{y \in \Omega_{2h}} w_{h,y} \hat{t}_\kappa, \quad \text{gdje je } w_{h,y} = \begin{cases} w_{2h}(y) \hat{t}_\kappa, & x = y + \kappa h, \kappa \in V \\ 0, & x = y + \kappa h, \kappa \notin V \end{cases}$$

gdje je $V \subset \mathbb{Z}^2$, a \hat{t}_κ su neki koeficijenti. Operator I_{2h}^h ćemo identifikovati sa šemom $] \hat{t}_\kappa [_{2h}^h$, tj.

$$I_{2h}^h \hat{=}] \hat{t}_\kappa [_{2h}^h = \begin{bmatrix} \vdots & \vdots & \vdots \\ \hat{t}_{-1,1} & \hat{t}_{0,1} & \hat{t}_{1,1} \\ \vdots & \vdots & \vdots \\ \dots & \hat{t}_{-1,0} & \hat{t}_{0,0} & \hat{t}_{1,0} & \dots \\ \vdots & \vdots & \vdots \\ \hat{t}_{-1,-1} & \hat{t}_{0,-1} & \hat{t}_{1,-1} \\ \vdots & \vdots & \vdots \end{bmatrix}_{2h}^h$$

U praksi se najčešće koristi bilinearna interpolacija

$$I_{2h}^h \hat{=} \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}_{2h}^h.$$

Za operator $L_h : \mathcal{G}(\Omega_h) \rightarrow \mathcal{G}(\Omega_h)$ se pretpostavlja da ima

inverz L_a^{-1} . Uvodimo sljedeće oznake:

u_a - tačno rješenje jednačine (1)

u_a^j - neka aproksimacija tačnog rješenja

$v_a^j = u_a - u_a^j$ - greška (korekcija) aproksimacije u_a^j

$d_a^j = f_a - L_a u_a^j$ - defekt (ostatak) aproksimacije u_a^j

S obzirom na uvedene oznake je

$$(2) \quad L_a v_a^j = d_a^j.$$

Ova jednačina se zove defektnom jednačinom. Ona je ekvivalentna jednačini (1) jer je svejedno koju od tih jednačina ćemo tačno riješiti. Ako riješimo tačno defektnu jednačinu (2), tada je tačno rješenje jednačine (1)

$$u_a = u_a^j + v_a^j$$

Iako se čini da se zamjenom jednačine (1) defektnom jednačinom (2) ništa ne postiže, pokazace se da nije tako.

Defektna jednačina (2) i njena aproksimacija će odigrati suštinsku ulogu u formiranju mnogostepenih iteracionih metoda. Aproksimaciju defektne jednačine (2) je moguće izvršiti zamjenom operatora L_a "prostijim" operatorom \hat{L}_a čiji se inverz \hat{L}_a^{-1} lakše određuje. Zbog te zamjene se ne dobija tačna korekcija v_a^j već \hat{v}_a^j , tj.

$$(3) \quad \hat{L}_a \hat{v}_a^j = d_a^j.$$

Rješenje jednačine (3) će predstavljati aproksimaciju tačne korekcije v_a^j , pa će popravljena vrijednost $u_a^j + \hat{v}_a^j$ biti samo aproksimacija tačnog rješenja u_a koju možemo uzeti za sljedeću iteraciju, tj.

$$u_a^{j+1} = u_a^j + \hat{v}_a^j.$$

Na ovaj način je određena jedna iteraciona procedura čiji

je iteracioni operator

$$E_h - B_h L_h : G(\Omega_h) \rightarrow G(\Omega_h),$$

gdje je E_h - jedinični operator, a $B_h = \hat{L}_h^{-1} \sim \bar{L}_h^{-1}$.

Za grešku i defekt važe jednakosti

$$\begin{aligned} v_h^{j+1} &= (E_h - B_h L_h) v_h^j \\ d_h^{j+1} &= (E_h - L_h B_h) d_h^j, \quad j=0,1,\dots \end{aligned}$$

Ranije smo vidjeli da konvergencija iteracionih metoda zavisi od spektralnog radijusa iteracionog operatora. Norme $\|E_h - B_h L_h\|$ i $\|E_h - L_h B_h\|$ (ne samo spektralne) daju stvarne faktore redukcije greške i defekta na jednom iteracionom koraku.

Višemrežna ideja se sastoji u tome da se za aproksimaciju \hat{L}_h koristi aproksimacija operatora L_h operatorom $L_H : G(\Omega_H) \rightarrow G(\Omega_H)$, gdje je Ω_H neka krupnija mreža. Ovo znači da se defektna jednačina (2) zamjenjuje jednačinom

$$(4) \quad L_H \hat{v}_H^j = d_H^j.$$

Kako je $\dim G(\Omega_H) \ll \dim G(\Omega_h)$, bice mnogo lakše odrediti inverz L_H^{-1} čija se egzistencija pretpostavlja.

Takodje se pretpostavlja i egzistencija transfer-operatora I_h^H i I_H^h .

Opisacemo jedan iteracioni korak metoda koji se bazira na aproksimaciji defektne jednačine (2) jednačinom (4):

1. izračunati defekt $d_h^j = f_h - L_h u_h^j$
2. suziti defekt $d_H^j = I_h^H d_h^j$
3. tačno riješiti jednačinu $L_H \hat{v}_H^j = d_H^j$
4. interpolirati izračunatu korekciju $\hat{v}_h^j = I_H^h \hat{v}_H^j$
5. odrediti novu aproksimaciju $u_h^{j+1} = u_h^j + \hat{v}_h^j$.

Ovaj iteracioni proces ima iteracioni operator

$$K_{\varepsilon}^H = E_{\varepsilon} - I_H^{\varepsilon} L_H^{\wedge} I_{\varepsilon}^H L_{\varepsilon},$$

tj.

$$(5) \quad u_{\varepsilon}^{j+1} = (E_{\varepsilon} - I_H^{\varepsilon} L_H^{\wedge} I_{\varepsilon}^H L_{\varepsilon}) u_{\varepsilon}^j + b.$$

Zvacemo ga procesom krupno-mrežne korekcije.

Operator $I_H^{\varepsilon} L_H^{\wedge} I_{\varepsilon}^H$ predstavlja aproksimaciju operatora L_{ε}^{\wedge} .

L e m a 1.

Proces krupno-mrežne korekcije (5) koristen kao iteracioni proces nije konvergentan.

D o k a z.

Ovo je posljedica činjenice da, za neko $w_{\varepsilon} \in G(\Omega_{\varepsilon})$ koje nije mrežna nula-funkcija, a za koju je $L_{\varepsilon} w_{\varepsilon} \in \ker(I_{\varepsilon}^H)$, važi

$$(E_{\varepsilon} - I_H^{\varepsilon} L_H^{\wedge} I_{\varepsilon}^H L_{\varepsilon}) w_{\varepsilon} = E_{\varepsilon} w_{\varepsilon} = w_{\varepsilon}$$

pa je

$$\rho(E_{\varepsilon} - I_H^{\varepsilon} L_H^{\wedge} I_{\varepsilon}^H L_{\varepsilon}) \gg 1.$$

Time je narušen neophodan i dovoljan uslov konvergencije iteracionog metoda (5). \square

Ovo u stvari znači da krupno-mrežna defektna jednačina (4) u opštem slučaju nije razumna aproksimacija defektne jednačine (2).

Ranije smo rekli da Jacobi-jev, Gauss-Seidel-ov i drugi klasični iteracioni metodi nemaju dobra svojstva konvergencije jer njihov faktor konvergencije zavisi od parametra h .

Ovo se može prevazici kombinovanjem tih metoda sa procesom krupno-mrežne korekcije. Ta kombinacija daje dvostepene metode koji čine osnovu mnogostepenih metoda. Spektralni radijusi Jacobi-jevog i Gauss-Seidel-ovog metoda za Poisson-ovu jednačinu su veličine reda $1 - O(h^2)$, dok metod relaksacije sa optimalnim relaksacionim parametrom ima spektralni radijus

reda $1 - O(h)$, što je nešto bolje, ali ipak loše kada $h \rightarrow 0$. Ovo znači da će redukcija greške biti spora. Medjutim, pokazuje se da ta redukcija nije ravnomjerna i da je na nekim harmonikama ona brza. Redukcija greške primjenom klasičnih iteracionih metoda se može analizirati razvijanjem greške u diskretni Fourier-ov red. Pri tome se grubo pravi razlika između glatkih (niskofrekventnih) i neglatkih (visokofrekventnih) komponenata greške. Poznato je da su glatke komponente greške odgovorne za sporu asimptotsku konvergenciju jer se njihova amplituda sporo smanjuje. Iako relaksacioni metodi sporo konvergiraju zbog spore redukcije glatkih komponenata greške, ovi su metodi vrlo efikasni u gušenju visokofrekventnih komponenata greške, tj. u smanjenju njihovih amplituda. Zbog ovoga će poslije nekoliko relaksacionih koraka u greški dominantnu ulogu imati niskofrekventne komponente, tj. njen glatki dio, pa kažemo da relaksacioni metodi imaju dobra glačajuća svojstva. Ona se mjere glačajucim faktorom, tj. brojem koji predstavlja najveći (najlošiji) faktor kojim se redukuju visoke frekvencije greške.

Predstavljajući početnu grešku v_k^j u obliku

$$v_k^j = \sum_{i < K} a_i^j \varphi_k^i + \sum_{i > K} a_i^j \varphi_k^i = (v_k^j)_L + (v_k^j)_H$$

gdje je $(v_k^j)_L$ - niskofrekventna, a $(v_k^j)_H$ - visokofrekventna komponenta greške. Funkcije φ_k^i su sopstvene funkcije relaksacionog operatora S_k , a K je neki broj o kojem ćemo kasnije nešto više reći. Primjenjujući relaksacionu proceduru sa iteracionim operatorom S_k , poslije ν iteracionih koraka ćemo dobiti grešku

$$v_k^{j+\nu} = S_k^\nu v_k^j$$

kojoj odgovara razvoj

$$v_k^{j+\nu} = \sum_{i < K} d_i^j (\mu_i)^\nu \psi_k^i + \sum_{i \geq K} d_i^j (\mu_i)^\nu \psi_k^i = (v_k^{j+\nu})_L + (v_k^{j+\nu})_H.$$

Funkcije ψ_k^i za $i < K$ su niske frekvencije, dok su ψ_k^i za $i \geq K$ visoke frekvencije. Za niske frekvencije je karakteristično da su one "vidljive" i na krupnijoj mreži Ω_H , dok visoke frekvencije nisu. Među sopstvenim vrijednostima μ_i , sa $i < K$, su i one sopstvene vrijednosti koje su vrlo bliske jedinici po modulu, dok je $|\mu_i| \leq q < 1$ za $i \geq K$. Zbog ovoga će poslije ν relaksacionih koraka biti smanjene amplitude visokih frekvencija bar sa faktorom q^ν . Glačajući faktor će biti jednak $q = \max_{i \geq K} |\mu_i|$. Redukcija one niske frekvencije koja odgovara sopstvenoj vrijednosti koja je veličina reda $1 - O(h^2)$, ako se koristi Jacobi-jev ili Gauss-Seidel-ov metod, ili reda $1 - O(h)$ u slučaju metoda gornje relaksacije, je neznatna. Na taj način je niskofrekventni dio $(v_k^{j+\nu})_L$ dominantan u odnosu na $(v_k^{j+\nu})_H$.

Restrikcijom greške $v_k^{j+\nu}$ na mrežu Ω_H se dobija

$$v_H^{j+\nu} = \sum_{i < K} \beta_i^{j+\nu} \psi_k^i$$

gdje se funkcije ψ_k^i sada posmatraju samo u tačkama mreže Ω_H . Pri tome nije velika razlika između koeficijenata $\beta_i^{j+\nu}$ i $d_i^j (\mu_i)^\nu$. Pošto je greška $v_k^{j+\nu}$ glatka, ona će biti dobro aproksimirana na krupnijoj mreži Ω_H . Zbog ovoga se prije primjene krupno-mrežne korekcije, tj. prije prelaska na jednačinu (4), vrši glačanje korištenjem nekog relaksacionog metoda. Svaki korak dvostepenog iteracionog metoda se sastoji od glačanja i krupno-mrežne korekcije. Struktura jednog iteracionog koraka, tj.

odredjivanje vrijednosti u_k^{j+1} ako je poznato u_k^j , izgleda ovako:

1. Prvi dio glačanja

-izračunati \bar{u}_k^j primjenom ν_1 koraka nekog relaksacionog metoda na u_k^j :

$$\bar{u}_k^j = \text{Relax}^{\nu_1}(u_k^j, L_k, f_k)$$

2. Krupno-mrežna korekcija

-izračunati defekt $\bar{d}_k^j = f_k - L_k \bar{u}_k^j$

-suziti defekt $\bar{d}_H^j = I_k^H \bar{d}_k^j$

-riješiti tačno jednačinu $L_H \hat{v}_H^j = \bar{d}_H^j$

-interpolirati korekciju $\hat{v}_k^j = I_H^k \hat{v}_H^j$

-izračunati popravljenu aproksimaciju $\bar{u}_k^j + \hat{v}_k^j$

3. Drugi dio glačanja

-izračunati u_k^{j+1} primjenom ν_2 koraka datog relaksacionog metoda na $\bar{u}_k^j + \hat{v}_k^j$:

$$u_k^{j+1} = \text{Relax}^{\nu_2}(\bar{u}_k^j + \hat{v}_k^j, L_k, f_k).$$

Ovaj složeni iteracioni proces se može zapisati u obliku

$$(6) \quad u_k^{j+1} = M_k^H u_k^j + b$$

gdje je iteracioni operator

$$(7) \quad M_k^H = S_k^{\nu_2} K_k^H S_k^{\nu_1}$$

izražen pomocu relaksacionog operatora S_k i operatora krupno-mrežne korekcije $K_k^H = E_k - I_H^k L_H^{-1} I_k^H L_k$.

Sljedeće komponente čine dvostepene metode:

a) relaksaciona procedura koju karakteriše operator S_k

b) brojevi relaksacionih koraka ν_1, ν_2

c) krupna mreža Ω_H

d) sitno-krupni transfer-operator I_k^H (restrikcija)

e) krupno-sitni transfer-operator I_H^h (interpolacija)

f) krupno-mrežni operator L_H .

Pokazuje se da izbor ovih komponenti bitno utiče na efikasnost dvostepenog metoda, ali se čini da nema opšteg pravila za njihov izbor sa ciljem postizanja optimalnih rezultata (vidjeti [24]). Što se tiče izbora krupnije mreže Ω_H , najčešće se koristi tzv. standardno ukрупnjavanje, tj. polazeci od mreže Ω_h sa korakom $h = (h_1, h_2)$, prelazi se na mrežu Ω_H sa korakom $H = 2h = (2h_1, 2h_2)$. O drugim načinima za izbor Ω_H vidjeti [24]. Što se tiče krupno-mrežnog diferencijskog operatora L_H , on se obično konstruiše na Ω_H na analogan način na koji je konstruisan i operator L_h na Ω_h , ali to nije obavezno. U principu L_H može biti bilo koji diferencijski operator definisan na $G(\Omega_H)$ koji u nekom smislu aproksimira L_h . U slučaju konačno-elementnih jednačina se često koristi tzv. Galerkinov pristup po kome se transfer-operatori koriste za odredjivanje krupno-mrežnog operatora L_H , tj. uzima se

$$L_H = I_h^H L_h I_H^h.$$

Za glačanje se umjesto primjene relaksacione procedure sa operatorom S_h ν_1 puta prije i ν_2 puta poslije krupno-mrežne korekcije, mogu koristiti različite procedure na svakom koraku. Primjena relaksacionih metoda za glačanje nije obavezna. Bilo koja iteraciona procedura koja dobro redukuje visoke frekvencije, a čija primjena ne zahtijeva mnogo računskih operacija, se u principu može koristiti za glačanje u okviru mnogostepenih metoda.

za analizu svojstava konvergencije dvostepenih iteracionih metoda treba odrediti spektralni radijus $\rho(M_n^H)$ ili normu $\|M_n^H\|$. Računski eksperimenti vezani za primjenu dvostepenih metoda za rješavanje Dirichlet-ovog zadatka za Poisson-ova jednačinu pokazuju da se pogodnim izborom komponenti ovog metoda dobija vrlo brza iteraciona procedura. U tom slučaju je čak efektivno odredjen spektralni radijus $\rho(M_n^H)$ (vidjeti [24]). Eksperimenti pokazuju, a to se moglo i očekivati, da nema redukcije greške ako je $\nu_n = 0$, što znači da je obavezno prije primjene krupno-mrežne korekcije izvršiti glačanje. Dok relaksacioni metodi redukuju visoke frekvencije, krupno-mrežna korekcija redukuje niske na koje relaksacija nije bitno uticala. Rezultat primjene prvo relaksacije, a zatim krupno-mrežne korekcije je vrlo brza redukcija greške. Naknadno glačanje se može i izbjeći jer visoke frekvencije koje se stvaraju poslije primjene krupno-mrežne korekcije neće imati velike amplitude pa se i ne moraju smanjivati. Dvostepeni metodi se rijetko koriste u praksi. Oni služe samo kao osnova mnogostepenih metoda. Najčešće ni mreža Ω_n nije dovoljno krupna da bi tačno rješavanje krupno-mrežne defektne jednačine $L_n \hat{v}_n^j = \bar{d}_n^j$ bilo ekonomično. Zbog toga se ta jednačina rješava približno. Za približno rješavanje te jednačine prirodno je iskoristiti dvostepeni metod čime se uvodi i treća mreža sa krupnijim korakom nego što je korak mreže Ω_n . Ako je faktor konvergencije tog dvostepenog metoda dovoljno mali, dovoljno je izvršiti samo nekoliko, recimo γ iteracionih koraka, da bi se nasla dobra aproksimacija rješenja \hat{v}_n^j . Ova ideja može biti primje-

njena rekurzivno što dovodi do korištenja sve krupnijih i krupnijih mreža. Na najkrupnijoj mreži može biti iskorišten neki direktni metod ili čak neki klasični iteracioni metod ako on na najkrupnijoj mreži ima dovoljno dobra svojstva konvergencije.

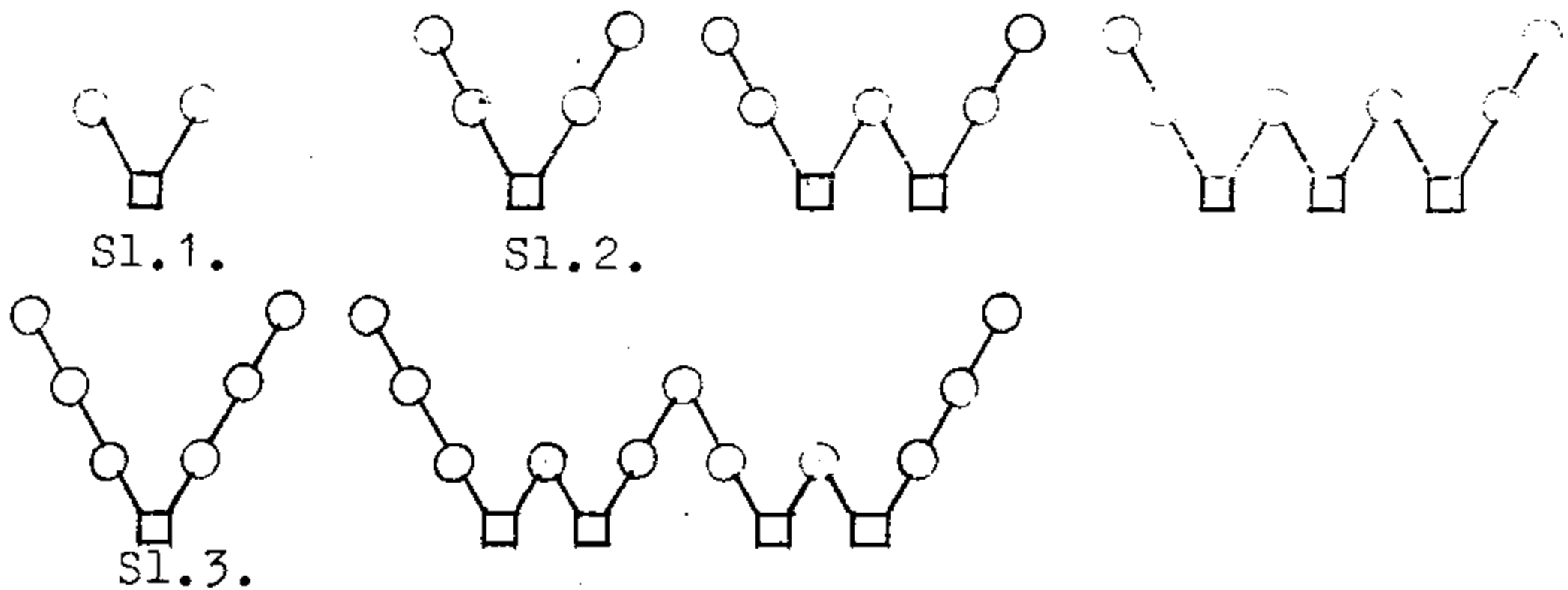
Struktura jednog koraka (kruga) mnogostepenog iteracionog metoda je moguće predstaviti grafički, ako se uvedu sljedeće grafičke oznake:

○- oznaka za glačanje

□- oznaka za tačno rješavanje defektne jednačine

\- oznaka za sitno-krupni transfer

/- oznaka za krupno-sitni transfer



Na Sl. 1.1. je prikazana struktura jednog kruga dvostepenog metoda, Sl. 1.2. prikazuje strukturu trostepenih metoda sa $\gamma=1$, $\gamma=2$ i $\gamma=3$ iteracija redom na nižem stepenu, dok je na Sl. 1.3. prikazan krug četverostepenog metoda sa $\gamma=1$ i $\gamma=2$.

Ako je $\gamma=1$, iteracioni krug nazivamo V-krugom, a ako je $\gamma=2$ onda govorimo o W-krugu. Jasno je da se sa povećanjem broja iteracija γ na nižim stepenima struktura iteracionog kruga uslošnjava što dovodi do povećanja računskog posla. Zbog toga su V i W-krugovi najinteresantniji za praksu.

Opisacemo jedan korak κ -stepenog iteracionog metoda.

Neka je dat niz mreža $\Omega_{h_\ell} = \Omega_\ell$ sa koracima h_ℓ , $\ell = 1, 2, \dots, \kappa$, tako da je Ω_1 najkrupnija mreža, dok je mreža Ω_i krupnija od Ω_{i+1} za $i = 1, 2, \dots, \kappa - 1$. Neka su na prostorima $G(\Omega_\ell)$ mrežnih funkcija definisani sljedeći operatori

$$L_\ell: G(\Omega_\ell) \rightarrow G(\Omega_\ell), \quad S_\ell: G(\Omega_\ell) \rightarrow G(\Omega_\ell), \quad \ell = 1, 2, \dots, \kappa$$

$$I_\ell^{\ell-1}: G(\Omega_\ell) \rightarrow G(\Omega_{\ell-1}), \quad I_{\ell-1}^\ell: G(\Omega_{\ell-1}) \rightarrow G(\Omega_\ell), \quad \ell = 2, 3, \dots, \kappa$$

Na svakoj mreži ćemo posmatrati jednačinu oblika

$$L_\ell u_\ell = f_\ell, \quad \ell = 1, 2, \dots, \kappa.$$

S_ℓ je linearni operator koji odgovara relaksacionom metodu,

$I_\ell^{\ell-1}$ je operator restrikcije, a $I_{\ell-1}^\ell$ - operator interpolacije.

Naš je cilj da riješimo zadatak $L_\kappa u_\kappa = f_\kappa$ na najsitnijoj mreži Ω_κ , dok nam zadaci na krupnijim mrežama služe samo kao pomoćni.

Pretpostavimo da su brojevi ν_1 , ν_2 i γ fiksirani i da ne zavise od κ i ℓ . Ako je data neka aproksimacija u_κ^j tačnog rješenja u_κ , sljedeća aproksimacija u_κ^{j+1} se računa ovako:

a) ako je $\kappa = 2$ to je dvostepeni metod koji je već opisan

b) ako je $\kappa > 2$

1. Prvi dio glačanja

- izračunati \bar{u}_κ^j primjenom $\nu_1 \gg 0$ glačajucih koraka na u_κ^j :

$$\bar{u}_\kappa^j = \text{Relax}^{\nu_1}(u_\kappa^j, L_\kappa, f_\kappa)$$

2. Krupno-mražna korekcija

- izračunati defekt $\bar{d}_\kappa^j = f_\kappa - L_\kappa \bar{u}_\kappa^j$

- suziti defekt $\bar{d}_{\kappa-1}^j = I_\kappa^{\kappa-1} \bar{d}_\kappa^j$

- izračunati jedno približno rješenje $\tilde{v}_{\kappa-1}^j$ defektne jednačine

$$(*) \quad L_{\kappa-1} \hat{v}_{\kappa-1}^j = \bar{d}_{\kappa-1}^j$$

na Ω_{k-1} primjenom γ_{k-1} iteracija $k-1$ -stepenog metoda (koristeći mreže $\Omega_{k-2}, \dots, \Omega_1$ i odgovarajuće mrežne operatore) za rješavanje jednačine (*) sa nultom početnom iteracijom,

- interpolirati korekciju $\tilde{v}_k^j = I_{k-1}^k \tilde{v}_{k-1}^j$
- izračunati popravljenu aproksimaciju na Ω_k $\bar{u}_k^j + \tilde{v}_k^j$

3. Drugi dio glačanja

- izračunati u_k^{j+1} primjenom ν_2 glačajucih koraka na $\bar{u}_k^j + \tilde{v}_k^j$:

$$u_k^{j+1} = \text{Relax}^{\nu_2}(\bar{u}_k^j + \tilde{v}_k^j, L_k, f_k).$$

3.3. Konvergenција i efikasnost mnogostepene iteracione procedure

L e m a 2. (vidjeti [24])

Iteracioni operator M_k opisanog k -stepenog metoda je dat sljedećim rekurentnim formulama

$$M_2 = S_2^{\nu_2} (E_2 - I_1^2 L_1^{-1} I_2^2 L_2) S_2^{\nu_2}$$

$$M_\ell = S_\ell^{\nu_2} [E_\ell - I_{\ell-1}^\ell (E_{\ell-1} - M_{\ell-1}^{\nu_2}) L_{\ell-1}^{-1} I_\ell^{\ell-1} L_\ell] S_\ell^{\nu_2}, \ell = 3, \dots, k. \quad \square$$

Razlika između iteracionog operatora dvostepenog $(h_\ell, h_{\ell-1})$ metoda kojim se rješava jednačina oblika $L_\ell u_\ell = f_\ell$, tj. operatora

$$M_\ell^{\ell-1} = S_\ell^{\nu_2} (E_\ell - I_{\ell-1}^\ell L_{\ell-1}^{-1} I_\ell^{\ell-1} L_\ell) S_\ell^{\nu_2}$$

i operatora M_ℓ ℓ -stepenog metoda je u tome što je operator

$$L_{\ell-1}^{-1} \text{ iz } M_\ell^{\ell-1} \text{ zamijenjen sa operatorom } (E_{\ell-1} - M_{\ell-1}^{\nu_2}) L_{\ell-1}^{-1}.$$

To odražava činjenicu da je krupno-mrežna jednačina $L_{\ell-1} \hat{v}_{\ell-1}^j = \hat{d}_{\ell-1}^j$ riješena približno sa γ iteracija $\ell-1$ -stepenog metoda sa

nultom početnom iteracijom.

Iteracioni operator M_ℓ ($\ell = 3, 4, \dots, \kappa$) se može napisati i u obliku

$$M_\ell = M_\ell^{\ell-1} + A_{\ell-1}^\ell M_{\ell-1}^\ell A_\ell^{\ell-1},$$

gdje je

$$A_{\ell-1}^\ell = S_\ell^{\nu_2} I_{\ell-1}^\ell : G(\Omega_{\ell-1}) \rightarrow G(\Omega_\ell),$$

$$A_\ell^{\ell-1} = L_{\ell-1}^{\nu_1} I_\ell^{\ell-1} L_\ell S_\ell^{\nu_1} : G(\Omega_\ell) \rightarrow G(\Omega_{\ell-1}).$$

Koristeći ovaj zapis iteracionog operatora M_ℓ i pretpostavljajući da su poznate ocjene za norme $\|M_\ell^{\ell-1}\|$, $\|A_{\ell-1}^\ell\|$ i $\|A_\ell^{\ell-1}\|$, $\ell \leq \kappa$, može se izvesti ocjena za $\|M_\ell\|$.

L e m a 3. (vidjeti [24])

Neka sljedeće ocjene važe uniformno u odnosu na $\ell \leq \kappa$:

$$\|M_\ell^{\ell-1}\| \leq \sigma, \quad \|A_{\ell-1}^\ell\| \cdot \|A_\ell^{\ell-1}\| \leq C.$$

Tada važi ocjena

$$\|M_\ell\| \leq \eta_\ell,$$

gdje je η_ℓ definisano rekurentno sa

$$\eta_2 = \sigma, \quad \eta_\ell = \sigma + C \eta_{\ell-1}, \quad \ell = 3, \dots, \kappa. \quad \square$$

Korištenjem ove ocjene, može se dokazati konvergencija mnogo-stepenih metoda nezavisno od $\ell = \ell_\ell$.

Broj iteracija γ na nižem stepenu može biti biran i tako da zavisi od stepena, tj. da bude $\gamma = \gamma_\ell$. Najčešće se razmatraju dva načina izbora:

$$\text{a) } \gamma_\ell = 2, \ell = 1, 2, \dots, \kappa-1 \quad \text{b) } \gamma_\ell = \begin{cases} 1, & \text{ako je } \ell \text{ parno} \\ 2, & \text{ako je } \ell \text{ neparno} \end{cases}$$

T e o r e m a 1. (vidjeti [24])

Ako se parametri γ_ℓ biraju na način a) i ako je $4C\sigma \leq 1$, tada

važi uniformna ocjena

$$\|M_k\| \leq \eta = \frac{1 - \sqrt{1 - 4C\sigma}}{2C} \leq 2\sigma,$$

dok pri izboru b) i uslovu $4C^2(1+C)\sigma \leq 1$, važi

$$\|M_k\| \leq \begin{cases} \frac{1 - \sqrt{1 - 4C^2(1+C)\sigma}}{2C^2} \leq 2\sigma(1+C), & k \text{ parno} \\ \frac{1 - 2C^2\sigma - \sqrt{1 - 4C^2(1+C)\sigma}}{2C^3} \leq \frac{\sigma(1+2C)}{C}, & k \text{ neparno } \square \end{cases}$$

Pošto $\|M_k\|$ predstavlja faktor redukcije greške na jednom koraku, zaključujemo da faktor konvergencije ne zavisi od koraka mreže $h = h_k$.

Ako je $\gamma_l = 2$ i ako je σ dovoljno malo, bice $\eta \approx \sigma$. Ovo znači da, ako dvostepeni metod dovoljno brzo konvergira, tada odgovarajući mnogostepeni metod ima slična svojstva konvergencije.

Ako je γ_l izabrano na način b), tada je ocjena za $\|M_k\|$ nešto lošija, tj. mnogostepeni metod sporije konvergira, ali je zato manji računski posao zbog jednostavnije strukture iteracionih krugova.

Isključuje se upotreba V-krugova ($\gamma_l = 1$), jer to dovodi do iteracionog procesa čiji faktor konvergencije zavisi od $h = h_k$ (vidjeti [24]).

Ovdje je samo kratko prikazan jedan od pristupa za dokazivanje nezavisnosti faktora konvergencije od koraka mreže.

Postoje i drugi načini za taj dokaz (vidjeti [3, 10, 12, 18]).

U većini tih dokaza se zahtijeva dovoljan broj glačajućih koraka, a isključuje se upotreba V-krugova.

U radu [5] je prvi put dat dokaz konvergencije mnogostepenog metoda za rješavanje pozitivno određenih konačno-elementnih jednačina koja ne zavisi od parametra h , a koji dozvoljava i korištenje proizvoljnog broja glačajućih koraka i upotrebu V-krugova. To svakako ima značaj jer utiče na smanjenje računskog posla.

Činjenica da neki metod ima faktor konvergencije koji ne zavisi od h ništa ne govori o efikasnosti metoda dok se ne uzme u obzir i računski posao. Pokazuje se da je broj aritmetičkih operacija potrebnih za jedan mnogostepeni krug proporcionalan broju čvorova najsitnije mreže.

Tako je u slučaju standardnog ukрупnjavanja za koje je $N_\ell \approx 4N_{\ell-1}$, $\ell = 2, 3, \dots, k$, gdje je N_ℓ - broj čvorova mreže Ω_ℓ , pri $\gamma_\ell = \gamma$ za sve ℓ , ukupan računski posao Q_k jednog k -stepenog kruga dat sa (vidjeti [24])

$$Q_k = \begin{cases} \frac{4}{3} C N_k & , \gamma = 1 \\ 2 C N_k & , \gamma = 2 \\ 4 C N_k & , \gamma = 3 \\ O(N_k \ln N_k) & , \gamma = 4 \end{cases}$$

Ako je $\gamma \leq 3$, naš k -stepeni iteracioni metod ce biti asimptotski optimalan jer tada konvergencija ne zavisi od h , dok je ukupan računski posao proporcionalan sa brojem čvorova. Ako se koriste neka druga ukрупnjavanja za koja je $\gamma_\ell = \gamma$ i $N_\ell \approx \tau N_{\ell-1}$, $\tau > 1$, tada je ukupan računski posao dat sa

$$Q_k = \begin{cases} \frac{\tau}{\tau - \gamma} C N_k & , \gamma < \tau \\ O(N_k \ln N_k) & , \gamma = \tau \end{cases}$$

U ovom slučaju će κ -stepeni metod biti asimptotski optimalan za $\gamma < \tau$.

ОСНОВНИ ОРГАНИЗАЦИЈА УЗДРУЖЕНОГ РАДА
ЗА МАТЕМАТИКУ, ФИЗИКУ И АСТРОНОМИЈУ
В И Ђ Л С Т Ђ А А

Б р о ј: _____

Д а т у м: _____

4. PRIMJENA METODA REGULARIZACIJE ZA RJEŠAVANJE SINGULARNIH KONAČNO-ELEMENTNIH JEDNAČINA

U ovoj glavi ćemo opisati metod regularizacije koji se koristi za rješavanje nekorektno postavljenih zadataka, a onda ćemo dokazati njegovu primjenljivost za rješavanje velikih sistema linearnih algebarskih jednačina koji se dobijaju konačno-elementnom aproksimacijom samokonjugovanih nenegativno odredjenih eliptičkih graničnih zadataka u dvodimenzionalnoj oblasti Ω . Bice predložena dva načina za izbor regularizacionih parametara.

4.1. O nekorektno postavljenim zadacima i metodu regularizacije

Medju matematičkim zadacima posebno mjesto zauzimaju zadaci čija su rješenja nestabilna u odnosu na male promjene polaznih podataka. Za njih je karakteristično to da proizvoljno male promjene polaznih podataka mogu dovesti do velikih promjena rješenja. Ti zadaci su loše postavljeni i oni pripadaju klasi tzv. nekorektno postavljenih zadataka. Za zadatak

$$(1) \quad A\mathfrak{z} = u, \quad \mathfrak{z} \in Z, \quad u \in U,$$

gdje su Z i U neki metrički prostori, a $A: Z \rightarrow U$ operator, kažemo da je korektan ili korektno postavljen ako su ispunjeni uslovi: a) jednačina (1) ima rjesenje za svako $u \in U$, b) rjesenje \mathfrak{z} se određuje jednoznačno zadavanjem u , c) rjesenje \mathfrak{z} neprekidno zavisi od u u metrikama Z i U , tj. ono

je stabilno u odnosu na poremećaje desne strane. Ako bar jedan od ovih uslova nije ispunjen, kažemo da je zadatak (1) nekorektan ili nekorektno postavljen.

Pokazuje se da je klasa nekorektno postavljenih zadataka veoma široka i da su ti zadaci značajni, iako se dugo smatralo da ti zadaci nemaju praktični značaj. Interes za tu klasu zadataka je naglo porastao poslije pojave Tihonovljevihi radova [48, 49] u kojima je uveden metod regularizacije za rješavanje nekorektno postavljenih zadataka.

Pretpostavimo da za neko $\bar{u} \in U$ zadatak (1) ima rješenje \bar{x} . Često je umjesto tačne desne strane \bar{u} za koju zadatak (1) ima rješenje, poznata neka njena aproksimacija u_δ i broj $\delta > 0$ tako da je $g_u(u_\delta, \bar{u}) \leq \delta$, gdje je sa $g_u(\cdot, \cdot)$ označena metrika u prostoru U . Na taj način mi umjesto tačnih polaznih podataka (A, \bar{u}) imamo približne polazne podatke (A, u_δ) i ocjenu njihove greške δ . Zadatak se sastoji u tome da se na osnovu poznatih polaznih podataka (A, u_δ, δ) nađe element x_δ koji će biti aproksimacija elementa \bar{x} sa svojstvom stabilnosti u odnosu na male promjene u_δ . Za aproksimaciju x_δ se ne može uzeti tačno rješenje jednačine (1) sa desnom stranom u_δ jer ono u opstem slučaju ne postoji za svako $u \in U$, a ako postoji, nije stabilno u odnosu na male promjene desne strane pa se može bitno razlikovati od rješenja \bar{x} .

Za operator $R(u, \delta)$ iz U u Z ćemo reci da je regularizacioni za jednačinu (1) (u odnosu na element \bar{u}), ako on zadovoljava svojstva:

1. postoji takav broj $\delta_1 > 0$, da je operator $R(u, \delta)$ odredjen

za svako δ , $0 \leq \delta \leq \delta_1$, i svako $u_\delta \in U$ tako da je

$$\varrho_U(u_\delta, \bar{u}) \leq \delta,$$

2. za svako $\varepsilon > 0$, postoji $\delta_\varepsilon = \delta_\varepsilon(\varepsilon, u_\delta) \leq \delta_1$, tako da iz nejednakosti

$$\varrho_U(u_\delta, \bar{u}) \leq \delta \leq \delta_\varepsilon$$

slijedi nejednakost

$$\varrho_Z(z_\delta, \bar{z}) \leq \varepsilon,$$

gdje je

$$z_\delta = R(u_\delta, \delta).$$

Na taj način se z_δ određuje pomoću operatora koji zavisi od parametra δ - mjere greške polaznih podataka, i za to z_δ važi

$$\varrho_Z(z_\delta, \bar{z}) \rightarrow 0 \quad \text{kada } \delta \rightarrow 0.$$

Često se koristi i druga definicija regularizacionog operatora. Kaže se da je operator $R(u, \alpha)$ iz U u Z , koji zavisi od parametra α , regularizacioni ako za njega važi:

1. postoje takvi brojevi $\delta_1 > 0$ i $\alpha_1 > 0$, da je operator $R(u, \alpha)$ određen za svako $\alpha \in (0, \alpha_1)$ i svako $u \in U$ za koje je

$$\varrho_U(u, \bar{u}) \leq \delta_1,$$

2. postoji takav funkcional $\alpha = \alpha(u, \delta)$ određen na skupu

$U_{\delta_1} = \{u : \varrho_U(u, \bar{u}) \leq \delta_1\}$ elemenata $u \in U$, tako da za svako $\varepsilon > 0$, postoji takav broj $\delta(\varepsilon) \leq \delta_1$ da, ako $\tilde{u} \in U$ i $\varrho_U(\tilde{u}, \bar{u}) \leq \delta \leq \delta(\varepsilon)$, važi

$$\varrho_Z(\bar{z}, z_\alpha) \leq \varepsilon,$$

gdje je

$$z_\alpha = R(\tilde{u}, \alpha(\tilde{u}, \delta)).$$

Ova definicija za $\alpha = \delta$ predstavlja u stvari prvu definiciju.

Ni u jednoj od ovih definicija se ne pretpostavlja jednoznačnost operatora regularizacije.

Ako je $g_u(\bar{u}, u_\delta) \leq \delta$, tada se za približno rješenje zadatka (1) sa približnom desnom stranom u_δ , uzima element $z_\alpha = R(u_\delta, \alpha)$ dobijen pomoću regularizacionog operatora $R(u, \alpha)$, gdje je $\alpha = \alpha(u_\delta) = \alpha_1(\delta)$ saglasno sa greškom δ polaznog podatka u_δ (vidjeti [48,49]). To rješenje z_α se naziva regularizovanim rješenjem jednačine (1). Brojni parametar α se naziva parametrom regularizacije. Svaki izbor parametra regularizacije α saglasnog sa greškom δ polaznih podataka u_δ određuje regularizacioni operator, tj. regularizacioni algoritam za određivanje približnih rješenja jednačine (1) stabilnih u odnosu na male promjene desne strane. Ako je poznato da je $g_u(u_\delta, \bar{u}) \leq \delta$, saglasno definiciji regularizacionog operatora, moguće je izabrati parametar regularizacije $\alpha = \alpha(u_\delta)$ tako da, kada greška $\delta \rightarrow 0$, regularizovano rješenje $z_\alpha = R(u_\delta, \alpha(u_\delta))$ teži (u metriki prostora Z) traženom tačnom rjesenju \bar{z} , tj. $g_Z(z_\alpha(u_\delta), \bar{z}) \rightarrow 0$. Zbog ovoga i ima smisla za približno rješenje zadatka (1) uzeti regularizovano rješenje.

Prema tome, problem određivanja rješenja zadatka koje će biti stabilno u odnosu na male promjene desne strane se svodi na određivanje regularizacionog operatora, pri čemu je potrebno izabrati parametar regularizacije na osnovu informacije o greski sa kojom se zadaje desna strana. Opisani metod za određivanje približnog rjesenja se naziva metodom regularizacije (vidjeti [54,55]).

Postoji više načina za određivanje regularizacionog operatora. Jedan od često korištenih načina je varijacioni princip izbora (vidjeti [54]) po kome se, kada je $g_u(u_\delta, \bar{u}) \leq \delta$, pribli-

žno rješenje traži u klasi Q_δ elemenata iz Z za koje je $g_u(Az, u_\delta) \leq \delta$. Klasa Q_δ predstavlja skup mogućih rješenja. Ali, za približno rješenje zadatka (1) nije moguće uzeti proizvoljno rješenje iz Q_δ jer je ta klasa široka i takvo "približno rješenje" ne bi bilo u opštem slučaju stabilno u odnosu na male promjene desne strane. Zbog ovoga je neophodan princip izbora mogućih rješenja koji će obezbijediti da se iz klase Q_δ izabere približno rješenje koje će biti stabilno u odnosu na male promjene desne strane. Za taj princip izbora moguće je uzeti varijacioni princip. U tom slučaju se izbor vrši uz pomoć specijalnih funkcionala $\Omega[z]$. Kažemo da je ne-negativni funkcional $\Omega[z]$, definisan na svuda gustom u Z skupu $Z_\lambda \subseteq Z$, stabilizirajući funkcional ako:

- a) element \bar{z} pripada njegovoj oblasti definisanosti
- b) za svaki broj $d > 0$, skup $Z_{\lambda, d} = \{z \in Z_\lambda : \Omega[z] \leq d\}$ je kompaktna u Z .

Neka je $\Omega[z]$ - stabilizirajući funkcional na $Z_\lambda \subseteq Z$. Prema varijacionom principu izbora, razmatraju se samo oni elementi skupa Q_δ na kojima je definisan funkcional $\Omega[z]$, tj. elementi z skupa $Z_{\lambda, \delta} = Z_\lambda \cap Q_\delta$, i medju tim elementima se za z_δ uzima element koji minimizira funkcional $\Omega[z]$.

Egzistencija takvog elementa se dokazuje u [54]. Tako izabrani element se može shvatiti kao rezultat primjene nekog operatora \tilde{R} koji zavisi od δ , na desnu stranu $u = u_\delta$ jednačine (1), tj.

$$z_\delta = \tilde{R}(u_\delta, \delta).$$

Pokazuje se da ovako definisani varijacioni princip izbora

predstavlja regularizacioni algoritam za rješavanje zadatka (1), tj. da je operator $\tilde{R}(u, \delta)$ regularizacioni za (1) (vidjeti [54]). To znači da se element $x_\delta = \tilde{R}(u_\delta, \delta)$ može uzeti za aproksimaciju rješenja x .

Da bi se otklonili neki nedostaci varijacionog principa izbora, za formiranje regularizacionih operatora se koristi i Lagrange-ov metod koji se sastoji u određivanju elementa x_α na kom funkcional

$$(2) \quad P_\alpha[x, u_\delta] = g_\alpha^2(Ax, u_\delta) + \alpha \Omega[x]$$

dostiže svoj infimum, pri čemu se parametar α određuje iz uslova

$$(3) \quad g_\alpha(Ax_\alpha, u_\delta) = \delta.$$

Broj $g_\alpha(Ax, u_\delta)$ se često naziva defektom, pa se izbor parametra α iz uslova (3) naziva izborom na osnovu defekta (vidjeti [54, 41]). Element x_α određen na ovaj način se može shvatiti kao rezultat primjene nekog operatora R_α koji zavisi od α , na desnu stranu $u = u_\delta$ jednačine (1), tj.

$$x_\alpha = R_\alpha(u_\delta, \alpha),$$

pri čemu se $\alpha = \alpha(\delta)$ određuje na osnovu defekta, tj. iz (3).

Pokazuje se da je taj operator regularizacioni. Parametar $\alpha = \alpha(\delta)$ koji se određuje iz (3) nije jednoznačan (vidjeti [54]), pa zbog toga nije jedinstven ni izbor regularizacionog operatora. Pitanje izbora parametra regularizacije je jedno od osnovnih pitanja metoda regularizacije i ono se razmatralo u mnogim radovima (vidjeti npr. [29, 30]). U slučaju kada je nepoznata greška polaznih podataka δ , nije moguće odrediti parametar regularizacije na osnovu defekta. U tom slučaju se

parametar λ određuje po kriterijumu kvazioptimalnosti ili kriterijumu količnika. Ovi načini izbora parametra regularizacije su predloženi u [50], dok je njihovu opravdanost pri rješavanju singularnih i loše uslovljenih sistema linearnih algebarskih jednačina dokazao Leonov u radu [34].

4.2. Rješavanje singularnih i loše uslovljenih sistema linearnih algebarskih jednačina

Klasi nekorektno postavljenih zadataka pripadaju i zadaci rješavanja sistema linearnih algebarskih jednačina sa singularnom ili loše uslovljenom matricom. Pojam loše uslovljenosti sistema nije sasvim precizno definisan (vidjeti [56]), ali je poznato da primjer takvih sistema predstavljaju sistemi kod kojih male promjene desnih strana mogu dovesti do velikih promjena rješenja. Sistem je singularan ako je matrica tog sistema singularna, tj. ako je njena determinanta jednaka nuli. U tom slučaju matrica tog sistema ima bar jednu sopstvenu vrijednost jednaku nuli. Matrica loše uslovljenih sistema ima neke sopstvene vrijednosti bliske nuli pa se često kaže da su loše uslovljeni sistemi "približno singularni". Ako se računanje vrši sa konačnom tačnošću, u najvećem broju slučajeva nije ni moguće utvrditi da li je neki sistem singularan ili samo loše uslovljen. Zbog toga je i nemoguće u okvirima zadate tačnosti praviti razliku između singularnih i loše uslovljenih sistema, pa se za njihovo rješavanje koriste isti metodi. U radovima [51, 52] su predloženi metodi za

rješavanje tih zadataka čijim korištenjem se dobijaju rješenja stabilna u odnosu na male promjene polaznih podataka. Radi se u stvari o primjeni metoda regularizacije za određivanje normalnog rješenja sistema

$$(4) \quad A\bar{z} = \bar{u}$$

gdje je $A = \{a_{ij}\}$, $\bar{z} = \{z_j\}$, $\bar{u} = \{u_i\}$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, a m i n ne moraju biti jednaki.

Uvedeci funkcional

$$(5) \quad P_\alpha[z, u_\delta] = \|Az - u_\delta\|^2 + \alpha \Omega[z], \quad \Omega[z] = \|z\|^2,$$

gdje je $\|\cdot\|$ - ranije uvedena norma vektora, pokazuje se da se za svako $\alpha > 0$ taj funkcional minimizira na jedinstvenom elementu z^* koji se može odrediti iz sistema

$$(6) \quad (A^T A + \alpha E)z = A^T u_\delta,$$

gdje je A^T - transponovana matrica matrice A .

Važi sljedeća

T e o r e m a 1. (vidjeti[54])

Neka je $\bar{z}^{(0)}$ normalno rješenje sistema $A\bar{z} = \bar{u}$ i neka je umjesto tačne desne strane \bar{u} zadat vektor u_δ takav da je $\|u_\delta - \bar{u}\| \leq \delta$.

Neka su dalje $\beta_1(\delta)$ i $\beta_2(\delta)$ proizvoljne, neprekidne na $[0, \delta_2]$ i pozitivne na $(0, \delta_2]$ funkcije koje monotonno teže nuli kada $\delta \rightarrow 0$ i takve da je

$$\frac{\delta}{\beta_1(\delta)} \leq \beta_2(\delta), \quad \beta_2(0) = 0.$$

Tada za proizvoljnu pozitivnu na $(0, \delta_2]$ funkciju $\alpha = \alpha(\delta)$ koja zadovoljava uslove

$$\frac{\delta}{\beta_1(\delta)} \leq \alpha(\delta) \leq \beta_2(\delta)$$

vektori $z^{(\delta)}$ koji minimiziraju funkcionale $P_{\alpha(\delta)}[z, u_\delta]$ konvergiraju ka normalnom rješenju $\bar{z}^{(0)}$ sistema (4) kada $\delta \rightarrow 0$, tj.

$$\lim_{\delta \rightarrow 0} \|z^{(\delta)} - \bar{z}^{(0)}\| = 0. \quad \square$$

Na osnovu ove teoreme zaključujemo da će minimizacija funkcionala (5), odnosno rješavanje sistema (6), određivati jedan regularizacioni algoritam za određivanje normalnog rješenja sistema (4) ako je parametar regularizacije $\alpha(\delta)$ izabran tako da

$$\frac{\delta}{\alpha(\delta)} \rightarrow 0 \quad \text{kada} \quad \delta \rightarrow 0.$$

Vektore $\bar{x}^{(\delta)}$ možemo shvatiti kao rezultate primjene nekih operatora $R(u, \alpha(\delta))$, koji su na osnovu prethodne teoreme regularizacioni u smislu Tihonovljeve definicije. Na sličan način se približno određuje normalno rješenje ako je i matrica sistema zadana približno (vidjeti [51, 52, 54]).

4.3. Formiranje regularizacionog algoritma za rješavanje konačno-elementnih aproksimacija singularnih eliptičkih graničnih zadataka

Medju zadacima matematičke fizike javljaju se i oni čije se rješenje ne određuje na jedinstven način. Standardan primjer je Neumann-ov zadatak za Poisson-ovu jednačinu (vidjeti [53]). Pri diskretizaciji ovog zadatka, npr. primjenom metoda konačnih elemenata, dobija se singularan sistem algebarskih jednačina sa simetričnom i nenegativno određenom matricom. Iako polazni zadatak ima beskonačno rješenja, zbog greške računanja desne strane odgovarajućeg sistema jednačina, najčešće se dobija nesaglasan sistem. Ali, mi i iz takvog sistema želimo da izvučemo informaciju o rješenju polaznog zadatka. Zbog toga se posmatraju uopštena rješenja dobijenog sistema

jednačina. Pošto polazni zadatak ima beskonačno mnogo rješenja, postavlja se pitanje od kojeg od tih rješenja ćemo dobiti aproksimaciju primjenom nekog približnog metoda. Za zadatak sa nejedinstvenim rješenjem je za praksu često interesantnije određivanje izvoda ili neke njihove kombinacije koja se određuje na jedinstven način. Zbog toga je dovoljno odrediti bilo koje od tih rješenja. Tako je u slučaju Neumann-ovog zadatka za Poisson-ovu jednačinu prvi izvod rješenja zadatka funkcija koja se određuje na jedinstven način, pa je za njegovo približno određivanje dovoljno odrediti bilo koje približno rješenje zadatka, a zatim njega približno diferencirati. Mi ćemo se baviti određivanjem normalnog rješenja.

Razmotricemo jedan regularizacioni algoritam za rješavanje sistema jednačina dobijenih konačno-elementnom diskretizacijom samokonjugovanih nenegativno određenih eliptičkih graničnih zadataka sa nejedinstvenim rješenjem ne vezujući se isključivo za Neumann-ov zadatak za Poisson-ovu jednačinu. Tom eliptičkom graničnom zadatku odgovara varijaciona formulacija:

naci $u \in H^1(\Omega)$ tako da je

$$(7) \quad a(u, v) = (f, v) \quad \text{za sve } v \in H^1(\Omega),$$

gdje je $a(\cdot, \cdot)$ simetrična, nenegativna bilinearna forma, a $\Omega \subset \mathbb{R}^2$ poligonalna oblast.

Neka je diskretizacijom takvog zadatka dobijen sistem algebarskih jednačina reda N

$$(8) \quad Ax = b$$

koji je, s obzirom na prirodu neprekidnog zadatka singularan

ili loše uslovljen. Matrica A je simetrična i nenegativno određena pri čemu su neke njene sopstvene vrijednosti ili jednake nuli ili vrlo bliske nuli. U svakom slučaju, zadatak (8) je nekorektno postavljen, pa je za određivanje njegovog približnog rješenja stabilnog u odnosu na male promjene desne strane potrebno primjeniti prethodno opisani Tihonovljev metod regularizacije. Fadejeva je uradu [57] predložila jedan metod za povećanje stabilnosti rješenja loše uslovljenih sistema koji se sastoji u pomaku spektra matrice, tj. u prelasku sa sistema (8) na sistem

$$(9) \quad (A + \alpha E)x = b,$$

i pokazala da, ako se parametar α izabere tako da bude mnogo manji od najmanje pozitivne sopstvene vrijednosti u slučaju singularne matrice, odnosno mnogo manji od najmanje sopstvene vrijednosti koja nije u grupi sopstvenih vrijednosti koje su bliske nuli u slučaju loše uslovljene matrice, onda rješenje sistema (9) aproksimira normalno rješenje sistema (8) i to rješenje je stabilno u odnosu na male poremećaje početnih podataka. Rješenje sistema (9), označimo ga sa x_α , može biti shvaćeno i kao rezultat primjene nekog operatora R koji zavisi od α i koji je regularizacioni u smislu Tihonova, pa se metod Fadejeve može shvatiti kao jedna varijanta Tihonovljevog metoda regularizacije u kojoj se regularizacioni operator određuje primjenom pomaka spektra, tj. u tom slučaju je $R(\alpha) = (A + \alpha E)^{-1}$.

Da je za određivanje regularizacionog operatora za zadatak (8) koristen varijacioni princip izbora, umjesto sistema (9)

označene prva i druga varijacija funkcionala P_α .

Ako je $\delta^2 P_\alpha[u, f_{\sigma_j}] > 0$ za sve $u \in \mathcal{M}_j$, zaključujemo da je funkcional (15) strogo konveksan, pa on svoj minimum dostiže na jedinstvenom elementu kojeg ćemo označavati sa u^* . Ostaje da se dokaže nejednakost (16).

Pošto je u^* element na kom se dostiže minimum funkcionala $P_\alpha[u, f_{\sigma_j}]$, mora biti

$$\delta P_\alpha[u^*, f_{\sigma_j}] = 0,$$

tj.

$$\alpha(u^*, v) + \lambda(u^*, v) = (f_{\sigma_j}, v) \quad \text{za sve } v \in \mathcal{M}_j.$$

Stavljajući $v = u^*$, dobijamo

$$\alpha(u^*, u^*) + \lambda(u^*, u^*) = (f_{\sigma_j}, u^*)$$

odakle, zbog nenegativnosti forme $\alpha(\cdot, \cdot)$, a poslije primjene Cauchy-Schwarz-ove nejednakosti, slijedi nejednakost (16). \square Rješenje navedenog ekstremalnog zadatka se određuje iz uslova $\delta P_\alpha[u, f_{\sigma_j}] = 0$, tj. u^* je rješenje zadatka

$$(17) \quad \alpha(u, v) + \lambda(u, v) = (f_{\sigma_j}, v) \quad \text{za sve } v \in \mathcal{M}_j.$$

Ovo je u stvari varijacioni zadatak kome odgovara sistem (13). U slučaju da je desna strana sistema (12) određena tačno, imali bismo umjesto (17) varijacioni zadatak tog oblika, samo što bi umjesto funkcije f_{σ_j} stajala funkcija f . Ako sa \bar{u} označimo normalno rješenje zadatka (14) sa $f_{\sigma_j} = f$, pokazuje se da se u^* može uzeti za aproksimaciju tog rješenja stabilnu u odnosu na male promjene funkcije f_{σ_j} , tj. važi

T e o r e m a 2 .

Neka je \bar{u}^* element na kom se dostiže minimum funkcionala

$$P_\alpha(u, f) = \alpha(u, u) - 2(f, u) + \lambda(u, u).$$

ada

$$\|\bar{u}^* - \bar{u}\| \rightarrow 0 \quad \text{kada } \alpha \rightarrow 0.$$

D o k a z .

ako je

$$P_\alpha[\bar{u}^*, f] = \min_{u \in M_j} P_\alpha[u, f],$$

će

$$P_\alpha[\bar{u}^*, f] \leq P_\alpha[\bar{u}, f],$$

j.

$$\alpha(\bar{u}^*, \bar{u}^*) - 2(f, \bar{u}^*) + \alpha(\bar{u}^*, \bar{u}^*) \leq \alpha(\bar{u}, \bar{u}) - 2(f, \bar{u}) + \alpha(\bar{u}, \bar{u}),$$

dnosno

$$\alpha(\bar{u}^*, \bar{u}^*) - 2(f, \bar{u}^*) + \alpha(\bar{u}^*, \bar{u}^*) \leq -\alpha(\bar{u}, \bar{u}) + \alpha(\bar{u}, \bar{u}).$$

odavde, uzimajući u obzir da je

$$(\bar{u}^*, f) = \alpha(\bar{u}, \bar{u}^*),$$

dobijamo

$$\alpha(\bar{u}^* - \bar{u}, \bar{u}^* - \bar{u}) + \alpha\|\bar{u}^*\|^2 \leq \alpha\|\bar{u}\|^2,$$

odakle, zbog nenegativnosti bilinearne forme $\alpha(\cdot, \cdot)$, konačno slijedi

$$(18) \quad \|\bar{u}^*\| \leq \|\bar{u}\|.$$

Ova nejednakost označava da je skup $\{\bar{u}^*\}$ ograničen, pa će to biti kompaktan skup. Zbog toga iz svakog niza (\bar{u}^{α_n}) , gdje $\alpha_n \rightarrow 0$ kada $n \rightarrow \infty$, možemo izdvojiti konvergentan podniz, označimo ga sa $(\bar{u}^{\alpha_{n_k}})$. Neka je

$$\lim_{k \rightarrow \infty} \bar{u}^{\alpha_{n_k}} = \bar{u}^{(0)}$$

ako je svaki član tog niza rješenje odgovarajućeg ekstremalnog zadatka, prema ranijem je

$$\alpha(\bar{u}^{\alpha_{n_k}}, v) + \alpha_{n_k}(\bar{u}^{\alpha_{n_k}}, v) = (f, v) \quad \text{za sve } v \in Y_{ij}$$

odakle, puštajući da $k \rightarrow \infty$, dobijamo

$$a(\bar{u}^{(n)}, v) = (f, v) \quad \text{za sve } v \in \mathcal{M}_j,$$

a je $\bar{u}^{(n)}$ jedno od rješenja diskretnog singularnog zadatka (14) sa $f_{\delta_j} = f$, tj. konačno-elementne aproksimacije zadatka (7) koja odgovara prostoru \mathcal{M}_j . Na osnovu (18) je takodje

$$\|\bar{u}^{(n)}\| \leq \|\bar{u}\| \quad \text{za sve } n,$$

dakle se dobija

$$\|\bar{u}^{(n)}\| \leq \|\bar{u}\|.$$

Iako norma rješenja $\bar{u}^{(n)}$ zadatka (14) sa $f_{\delta_j} = f$ ne može biti manja od norme normalnog rješenja tog zadatka, mora biti

$$\|\bar{u}^{(n)}\| = \|\bar{u}\|,$$

odotle, zbog jedinstvi normalnog rješenja slijedi da je

$$\bar{u}^{(n)} = \bar{u}.$$

Ako bi niz $(\bar{u}^{(n)})$ imao još jedan konvergentan podniz koji bi konvergirao nekom elementu $\bar{u}^{(n)}$, na analogan bi se način moglo pokazati da je $\bar{u}^{(n)} = \bar{u}$. Ovo u stvari znači da taj niz ima samo jednu tačku nagomilavanja, tj. da je on konvergentan i da je

$$\lim_{n \rightarrow \infty} \bar{u}^{(n)} = \bar{u}.$$

Zbog proizvoljnosti brojnog nula-niza (n_k) , zaključujemo da je

$$\lim_{k \rightarrow \infty} \bar{u}^{(n_k)} = \bar{u}. \quad \square$$

Razmotrimo sada slučaj kada je umjesto funkcije f poznata njeena aproksimacija takva da je

$$\|f_{\delta} - f\| \leq \delta.$$

T e o r e m a 3 .

Ako je $u^{d(\delta)}$ element na kom se dostiže minimum funkcionala (15). Tada za $d = d(\delta)$, izabrano tako da

$$(19) \quad d(\delta) \rightarrow 0 \quad \text{i} \quad \frac{\delta}{d(\delta)} \rightarrow 0 \quad \text{kada } \delta \rightarrow 0,$$

važi

$$\|u^{\alpha(\delta)} - \bar{u}\| \rightarrow 0 \quad \text{kada } \delta \rightarrow 0.$$

D o k a z .

Kako je

$$P_{\alpha}[\bar{u}^{\alpha(\delta)}, f] = \min_{u \in M_j} P_{\alpha}[u, f]$$

$$P_{\alpha}[\bar{u}^{\alpha(\delta)}, f_{\delta}] = \min_{u \in M_j} P_{\alpha}[u, f_{\delta}],$$

biće

$$\alpha(\bar{u}^{\alpha(\delta)}, v) + d(\bar{u}^{\alpha(\delta)}, v) = (f, v)$$

i

$$\alpha(u^{\alpha(\delta)}, v) + d(u^{\alpha(\delta)}, v) = (f_{\delta}, v),$$

odakle slijedi da je

$$\alpha(\bar{u}^{\alpha(\delta)} - u^{\alpha(\delta)}, v) + d(\bar{u}^{\alpha(\delta)} - u^{\alpha(\delta)}, v) = (f - f_{\delta}, v),$$

tj. funkcional $P_{\alpha}[u, f - f_{\delta}]$ dostiže svoj minimum na elementu

$$u = \bar{u}^{\alpha(\delta)} - u^{\alpha(\delta)}.$$

Kako je na osnovu Leme 1

$$\|\bar{u}^{\alpha(\delta)} - u^{\alpha(\delta)}\| \leq \frac{1}{\alpha(\delta)} \|f - f_{\delta}\|$$

biće

$$\begin{aligned} \|\bar{u}^{\alpha(\delta)} - \bar{u}\| &\leq \|u^{\alpha(\delta)} - \bar{u}^{\alpha(\delta)}\| + \|\bar{u}^{\alpha(\delta)} - \bar{u}\| \leq \frac{1}{\alpha(\delta)} \|f - f_{\delta}\| + \|\bar{u}^{\alpha(\delta)} - \bar{u}\| \leq \\ &\leq \frac{\delta}{\alpha(\delta)} + \|\bar{u}^{\alpha(\delta)} - \bar{u}\|. \end{aligned}$$

Pošto po pretpostavci $\frac{\delta}{\alpha(\delta)} \rightarrow 0$ kada $\delta \rightarrow 0$, a $\|\bar{u}^{\alpha(\delta)} - \bar{u}\| \rightarrow 0$ kada $\delta \rightarrow 0$

na osnovu pretpostavke i Teoreme 2, dobijamo da

$$\|u^{\alpha(\delta)} - \bar{u}\| \rightarrow 0 \quad \text{kada } \delta \rightarrow 0$$

što je i trebalo dokazati. \square

Na osnovu dokazane teoreme zaključujemo da minimizacija fun-

kcionala $P_{\alpha}[u, f_{\delta}]$ definisanog sa (15), odnosno rješavanje

zadatka (17) sa parametrom $\alpha = \alpha(\delta)$ izabranim tako da važi (19),

generiše algoritam koji je regularizacioni u smislu Tihonova.

Prema tome, rješenje sistema (13) se može uzeti za aproksimaciju normalnog rješenja sistema (10), pa se rješavanje tog sistema sa singularnom ili loše uslovljenom matricom može zamijeniti rješavanjem sistema (13). Tako dobijene sisteme oblika (13) mi ćemo rješavati primjenom mnogostepenih iteracionih procedura. Tačnije, mi ćemo regularizovati sistem koji odgovara konačno-elementnoj aproksimaciji u prostoru \mathcal{M}_κ , tj. na stepenu $j = \kappa$, pri čemu će nam se kao pomoćni javljati sistemi oblika (13) na stepenima $j = \kappa - 1, \dots, 1$. I ti sistemi će se rješavati na isti način, tj. primjenom j -stepenih procedura za $j > 1$, dok će se jedino sistem na stepenu $j = 1$ rješavati nekim direktnim metodom. Da bi se formirali sistemi oblika (13) na svim stepenima $j = \kappa, \kappa - 1, \dots, 1$, treba prethodno riješiti problem izbora parametara regularizacije. Pri izboru se mora voditi računa o saglasnosti tog parametra sa greškom δ , tj. o uslovu (19). Kod nas se situacija komplikuje posto mi moramo riješiti κ sistema oblika (13) sa različitim vrijednostima greške desne strane $\delta = \delta_j$. Postavlja se pitanje da li mi za svaki sistem moramo odrediti parametar α , ili će jedna vrijednost α odgovarati svim sistemima.

Greska desne strane tih sistema, koja nastaje zbog primjene formula za numeričku integraciju, može biti izražena u funkciji parametra h_j koji karakteriše prostor \mathcal{M}_j . Neka je ta greska na stepenu j jednaka

$$\delta_j = O(h_j^\beta), \quad \beta > 0.$$

U slučaju kada budemo koristili isti parametar α na svim stepenima, parametre regularizacije ćemo birati nasljedeci način:

$$(20) \quad d_j = \frac{1}{2} \frac{p^{p/2}}{v_k} \quad \text{za } j = k, k-1, \dots, 1.$$

Kada budemo koristili različite vrijednosti parametra d , stavljaćemo

$$(21) \quad d_k = \frac{1}{2} \frac{p^{p/2}}{v_k},$$

dok ćemo ostale parametre odredjivati po formuli

$$(22) \quad d_{j-1} = 2d_j, \quad j = k, k-1, \dots, 2.$$

Ovakvi izbori parametara regularizacije su u saglasnosti sa (19). Ovo bi takodje odgovaralo rezultatima vezanim za izbor parametra regularizacije po kriterijumima kvazioptimalnosti i količnika za metod uprošćene regularizacije (vidjeti [34]). Izbor (20) će imati smisla ukoliko broj k nije suviše velik što bi garantovalo ispunjenje uslova (19) na svim stepenima. U protivnom je d potrebno birati na osnovu (21) i (22).

ОСНОВНА ОРГАНИЗАЦИЈА УДРУЖЕНОГ РАДА
ЗА МАТЕМАТИКУ, МЕХАНИКУ И АСТРОНОМИЈУ
БИБЛИОТЕКА

Број: _____

Датум: _____

5. ANALIZA MNOGOSTEPENE ITERACIONE PROCEDURE ZA RJEŠAVANJE SINGULARNIH ZADATAKA

U glavi 4 je za rješavanje sistema jednačina dobijenih konačno-elementnom diskretizacijom samokonjugovanih nenegativno odredjenih eliptičkih graničnih zadataka predložen jedan regularizacioni algoritam. U ovoj glavi ćemo analizirati k -stepeni iteracioni metod za rješavanje sistema jednačina koji se javlja u vezi sa tim regularizacionim algoritmom i dokazati konvergenciju tog metoda nezavisno od h .

1. Uvodne napomene

Kao što smo vidjeli u glavi 3, mnogostepeni metodi se sa uspjehom primjenjuju za rješavanje velikih sistema linearnih jednačina koji se dobijaju diskretizacijom eliptičkih graničnih zadataka metodom konačnih razlika ili konačnih elemenata. Još je Bahvalov u svom radu [28] dokazao da se ovaj metod može primjeniti za proizvoljan eliptički operator drugog reda sa neprekidnim koeficijentima uz jedno, kako ga on naziva, prirodno ograničenje, koje se sastoji u tome da operator nije singularan, tj. da tačka nula nije tačka spektra operatora. U svim kasnijim dokazima konvergencije mnogostepenih iteracija (vidjeti [3, 4, 10, 12, 18, 19, 27]) se podrazumijeva ovo ograničenje. Mnogostepeni metod se dakle primjenjivao samo za granične zadatke sa jedinstvenim rjesenjem. Razlog za to je što se pri primjeni ovog metoda za rješavanje odgovarajućeg sistema jednačina dolazi do sistema na naj-

rupnijoj mreži (najnižem stepenu) kojeg treba riješiti tačno, pa se uvijek pretpostavljala egzistencija inverzne matrice tog sistema. Tek je u radovima [3 , 10] nagoviještena mogućnost primjene ovog metoda i za singularne granične zadatke, ali ne za sve, već samo za one iz te klase za koje je poznato jezgro operatora. Naime, u oba rada je razmatran Neumann-ov zadatak za Poisson-ovu jednačinu pri čemu se koristi jezgro odgovarajućeg operatora koje se sastoji od svih konstantnih funkcija. Znanje tog jezgra je omogućilo uvođenje faktor-prostora u kome posmatrani zadatak više nije singularan. Međutim, ovako se moglo postupiti samo u ovom specijalnom slučaju.

Dokazacemo da se mnogostepena iteraciona procedura može primijeniti za rješavanje sistema dobijenih konačnoelementnom diskretizacijom eliptičkih graničnih zadataka uz pretpostavku da je odgovarajući eliptički operator drugog reda samokonjugovan i nenegativno odredjen i da jezgro tog operatora sadrži i neku nenultu funkciju. Nikakvu drugu informaciju o jezgru operatora nećemo koristiti.

5.2. Postavka zadatka, oznake i pretpostavke

Kao prototip samokonjugovanog nenegativno odredjenog eliptičkog graničnog zadatka koristićemo Neumann-ov zadatak za Poisson-ovu jednačinu uz napomenu da ni u jednom trenutku nećemo koristiti nikakvu informaciju o jezgru operatora tog zadatka.

akle, razmatracemo granični zadatak

$$1) \quad \begin{aligned} -\operatorname{div}(a \nabla u) &= f && \text{u } \Omega \\ \frac{\partial u}{\partial n} &= 0 && \text{na } \partial\Omega, \end{aligned}$$

gdje je Ω - poligonalni domen u \mathbb{R}^2 . Za funkciju $a(x)$ se pretpostavlja da $a \in C^1(\bar{\Omega})$ i da postoje pozitivne konstante \underline{a} i \bar{a} takve da je

$$\underline{a} \leq a(x) \leq \bar{a} \quad \text{za } x \in \bar{\Omega}.$$

Zadanku (1) odgovara varijaciona (slaba) formulacija (vidjeti [8]):

traži se $u \in H^1(\Omega)$ tako da je

$$2) \quad a(u, v) = (f, v) \quad \text{za sve } v \in H^1(\Omega),$$

gdje je

$$3) \quad a(u, v) = \int_{\Omega} a \nabla u \nabla v \, dx, \quad (f, v) = \int_{\Omega} f v \, dx.$$

Često je da je bilinearna forma $a(\cdot, \cdot)$ simetrična i nenegativno određena, tj. $a(u, u) \geq 0$ za sve $u \in H^1(\Omega)$, pri čemu postoji funkcija $v \in H^1(\Omega)$ takva da je $v \neq 0$ i $a(v, v) = 0$.

U glavi 1 smo definisali prostore Soboljeva $H^m(\Omega)$ za nenegativne cijele brojeve m . Za brojeve m koji su pozitivni nisu cijeli, prostori $H^m(\Omega)$ se definišu interpolacijom, dok se prostori $H^m(\Omega)$ za negativne brojeve m definišu kao dualni prostori prostora $H_0^{-m}(\Omega)$ (vidjeti [15]).

Za diskretizaciju zadatka (1), odnosno (2), koristićemo metod konačnih elemenata.

Ako je \mathcal{T}_h fiksirana trijangulacija oblasti Ω . Za proizvoljan trougao $T \in \mathcal{T}_h$ uvodimo oznake

$$h_T = \operatorname{diam} T$$

$$h_T \cdot d_T \quad - \text{dijametar kruga upisanog u } T.$$

definiramo

$$4) \quad h_1 = \max_{T \in \mathcal{T}_1} h_T, \quad \delta_0 = \min_{T \in \mathcal{T}_1} \alpha_T, \quad \delta_1 = \min_{T \in \mathcal{T}_1} \frac{h_T}{h_1}.$$

što znači da ćemo pretpostavljati da je trijangulacija regularna i uniformna i da su δ_0 i δ_1 mjere regularnosti i uniformnosti trouglova iz \mathcal{T}_1 (vidjeti [8]).

Što ćemo razmatrati mnogostepenu proceduru, uvodimo nove trijangulacije \mathcal{T}_j oblasti Ω , za $j > 1$. Konstruisaćemo ih induktivno na sljedeći način: svaki trougao $T \in \mathcal{T}_{j-1}$ dijele imo spajanjem sredina stranica na četiri trougla koja će pripadati trijangulaciji \mathcal{T}_j . Tako konstruisane trijangulacije imaju konstante regularnosti i uniformnosti δ_0 i δ_1 kao i početna trijangulacija \mathcal{T}_1 . Za $h_j = \max_{T \in \mathcal{T}_j} h_T$ imamo

$$h_j = h_1 \cdot 2^{1-j}.$$

U svakoj trijangulaciji \mathcal{T}_j oblasti Ω odgovara N_j -dimenzioni prostor konačnih elemenata sastavljen od neprekidnih i po dio linearnih funkcija. Te prostore ćemo označavati sa \mathcal{M}_R ili, češće, sa \mathcal{M}_j .

Obzirom na način na koji su konstruisane trijangulacije \mathcal{T}_j , imamo

$$\mathcal{M}_j \subset H^1(\Omega) \text{ i } \mathcal{M}_j \subset \mathcal{M}_{j+1} \text{ za sve } j \gg 1.$$

U svakom prostoru \mathcal{M}_j odgovara sljedeća konačno-elementna aprosimacija zadatka (2):

Naći $u_j \in \mathcal{M}_j$ tako da je

$$5) \quad a(u_j, v) = (f, v) \quad \text{za sve } v \in \mathcal{M}_j.$$

Čak će cilj biti da zadatak (5) riješimo u prostoru \mathcal{M}_R , gdje je $R > 1$, dok će zadaci tog oblika u prostorima \mathcal{M}_j za $j < R$, biti uključeni samo kao pomoćni.

rešavanje zadatka (5) je ekvivalentno rešavanju sistema linearnih algebarskih jednačina

$$A^{(j)} x^{(j)} = b^{(j)}$$

gde je $A^{(j)}$ matrica čiji su elementi

$$A_{i,k}^{(j)} = \alpha(\varphi_i^{(j)}, \varphi_k^{(j)}),$$

$b^{(j)}$ je vektor sa komponentama $b_i^{(j)} = (f, \varphi_i^{(j)})$ i $\{\varphi_i^{(j)}, i = 1, 2, \dots, N_j\}$ baza prostora M_j .

Matrica $A^{(j)}$ je simetrična, a lako se pokazuje da je i negativno određena, tj. da je za svaki vektor $x \in \mathbb{R}^{N_j}$

$$x^T A^{(j)} x > 0.$$

U ovom primjeru je matrica $A^{(j)}$ singularna, ali se u opštem slučaju diskretizacijom singularnih graničnih zadataka mogu pojaviti pored singularnih i loše uslovljeni sistemi. Mi smo već u glavi 4 rekli da nećemo praviti razliku između tih sistema i da ćemo za njihovo rešavanje koristiti metod regularizacije. Kako se komponente vektora $b^{(j)}$ često određuju primjenom formula za numeričku integraciju, može se desiti da singularni sistemi (6) budu nesaglasni, pa ćemo određivati opštena rešenja tog sistema.

U skladu sa što smo rekli u glavi 4, za rešavanje sistema (6) ćemo koristiti regularizacioni algoritam koji se sastoji u zamjeni sistema (6) sa sistemom

$$(A^{(j)} + \alpha_j M^{(j)}) x^{(j)} = b^{(j)},$$

gde je $\alpha_j > 0$ - parametar regularizacije na stepenu j , a $M^{(j)}$ je matrica čiji su elementi

$$M_{i,k}^{(j)} = (\varphi_i^{(j)}, \varphi_k^{(j)}).$$

Uz proizvoljnom vektoru $x = (x_1, x_2, \dots, x_{N_j})^T$ iz \mathbb{R}^{N_j} pri-

uzimo funkciju $u = \sum_{i=1}^{N_j} x_i \varphi_i^{(j)}$ iz \mathcal{M}_j , tada je

$$x^T M^{(j)} x = (u, u) > 0,$$

za sve $x \neq 0$, što znači da je matrica $M^{(j)}$ pozitivno određena. Kako je matrica $A^{(j)}$ nenegativno određena, bice matrica $A^{(j)} + \alpha_j M^{(j)}$ sistema (7) pozitivno određena, pa će taj zadatak imati jedinstveno rješenje za svaki vektor $b^{(j)}$. Sistemu (7) odgovara slaba formulacija:

naći $u_j^* \in \mathcal{M}_j$ tako da je

$$3) \quad a(u_j^*, v) + \alpha_j(u_j^*, v) = (f, v) \quad \text{za sve } v \in \mathcal{M}_j.$$

Zadatak (8) je konačno-dimenzioni zadatak i on predstavlja aproksimaciju zadatka:

naći $u^* \in H^1(\Omega)$ tako da je

$$3) \quad a(u^*, v) + \alpha_j(u^*, v) = (f, v) \quad \text{za sve } v \in H^1(\Omega).$$

Skalarni proizvod (\cdot, \cdot) na $H^1(\Omega)$ se identifikuje sa dualnom normom na $H^1(\Omega) \times H^1(\Omega)$, pa je desna strana jednačine (9) definirana za sve $f \in H^{-1}(\Omega)$. Za rješavanje sistema linearnih algebarskih jednačina (7) dobijenog konačno-elementnom aproksimacijom zadatka (9) ćemo primjeniti mnogostepenu iteracionu proceduru.

Pozitivno određenoj bilinearnoj formi $a(\cdot, \cdot) + \alpha_j(\cdot, \cdot)$ odgovara energetska norma

$$\|u\|_{\alpha_j}^2 = a(u, u) + \alpha_j(u, u).$$

Za zadatak (9) moramo uvesti sljedeću pretpostavku regularnosti:

Postoji konstanta θ , $0 < \theta \leq 1$, takva da za sve $f \in H^{\theta-1}(\Omega)$

zadatak (9) ima jedinstveno rješenje $u^* \in H^{1+\theta}(\Omega)$ takvo da je

$$10) \quad \|u^*\|_{1+\theta} \leq C \|f\|_{\theta-1}.$$

ajuci u vidu način na koji su uvedeni prostori konačnih elemenata \mathcal{M}_j zaključujemo da su oni affine familije u smislu Arlet-a (vidjeti [8]), te za njih važe sljedeća svojstva:

Ako $u \in H^s(\Omega)$, $1 \leq s \leq 1 + \theta$, gdje je $0 < \theta \leq 1$, tada postoji $u_j \in \mathcal{M}_j$ tako da je

$$1) \quad \|u - u_j\|_0 + h_j \|u - u_j\|_1 \leq C h_j^s \|u\|_s,$$

za svako $u_j \in \mathcal{M}_j$ važi

$$2) \quad \|u_j\|_{s_2} \leq C h_j^{s_1 - s_2} \|u_j\|_{s_1}, \quad \text{za sve } 0 \leq s_1 \leq s_2 \leq 1,$$

vidjeti [1, 2, 23]).

Svojstvo a) je standardno aproksimaciono svojstvo prostora konačnih elemenata \mathcal{M}_j , a b) predstavlja tzv. inverznu postavku. Primjenjujući mnogostepenu iteracionu proceduru, ćemo u prilici da rješavamo zadatke u nešto opštijoj formi nego što je zadatak (8):

Naći $x \in \mathcal{M}_j$ tako da je

$$3) \quad a(x, v) + d_j(x, v) = G(v) \quad \text{za sve } v \in \mathcal{M}_j,$$

gdje je G linearni funkcional.

U odnosu na svaki prostor \mathcal{M}_j definišemo sopstvene funkcije $\psi_i^{(j)}$ odgovarajuće sopstvene vrijednosti $\lambda_i^{(j)}$, $1 \leq i \leq N_j$, za koje je

$$4) \quad a(\psi_i^{(j)}, v) + d_j(\psi_i^{(j)}, v) = \lambda_i^{(j)} (\psi_i^{(j)}, v) \quad \text{za sve } v \in \mathcal{M}_j.$$

Ulog pozitivne odredjenosti bilinearne forme $a(\cdot, \cdot) + d_j(\cdot, \cdot)$

či sve sopstvene vrijednosti $\lambda_i^{(j)}$ pozitivne. Ne umanjujući

opštost, možemo pretpostaviti da je

$$5) \quad \lambda_1^{(j)} \leq \lambda_2^{(j)} \leq \dots \leq \lambda_{N_j}^{(j)}$$

$$(\psi_i^{(j)}, \psi_k^{(j)}) = \delta_{ik}, \quad a(\psi_i^{(j)}, \psi_k^{(j)}) + d_j(\psi_i^{(j)}, \psi_k^{(j)}) = \lambda_i^{(j)} \delta_{ik},$$

gdje je δ_{ik} - Kronecker-ov simbol.

Na sličan način definišemo sopstvene funkcije i sopstve-

vrijednosti u odnosu na formu $\alpha(\cdot, \cdot)$, lako se pokazuje i toj bilinearnoj formi odgovaraju sopstvene funkcije $\psi_i^{(j)}$, da su joj sopstvene vrijednosti

$$6) \quad \mu_i = \lambda_i^{(j)} - \alpha_j, \quad i = 1, 2, \dots, N_j.$$

Prema tome, sopstvene vrijednosti $\lambda_i^{(j)}$ se mogu napisati u obliku

$$\lambda_i^{(j)} = \mu_i + \alpha_j, \quad i = 1, 2, \dots, N_j.$$

U skladu s nejednakost

$$\|u\|_{\alpha_j}^2 \leq C \|u\|_1^2,$$

gdje je $C = \max\{\bar{\alpha}, \alpha_j\}$, i (12) za $s_1 = 0$ i $s_2 = 1$ dobijamo ocjenu

za $\lambda_{N_j}^{(j)}$:

$$7) \quad \lambda_{N_j}^{(j)} = \lambda_{N_j}^{(j)} (\psi_{N_j}^{(j)}, \psi_{N_j}^{(j)}) = \alpha(\psi_{N_j}^{(j)}, \psi_{N_j}^{(j)}) + \alpha_j (\psi_{N_j}^{(j)}, \psi_{N_j}^{(j)}) = \|\psi_{N_j}^{(j)}\|_{\alpha_j}^2 \leq C \|\psi_{N_j}^{(j)}\|_1^2 \leq C R_j^{-2} \|\psi_{N_j}^{(j)}\|_0^2 = C R_j^{-2}.$$

U svaki prostor \mathcal{M}_j uvodimo skalu normi $\|\cdot\|_{\alpha_j, s}$, $-2 \leq s \leq 2$, koje ćemo definirati na sljedeći način:

za proizvoljnom elementu $\chi \in \mathcal{M}_j$, koji u odnosu na bazu od sopstvenih funkcija ψ_i ima razvoj $\chi = \sum_{i=1}^{N_j} c_i \psi_i$, pridružujemo normu

$$8) \quad \|\chi\|_{\alpha_j, s} = \sum_{i=1}^{N_j} c_i^2 \lambda_i^s, \quad -2 \leq s \leq 2.$$

U radi jednostavnosti smo izostavili gornji indeks od $\lambda_i^{(j)}$ i $\psi_i^{(j)}$.

U skladu sa je da je

$$\|\chi\|_{\alpha_j, 0} = \|\chi\|_0, \quad \|\chi\|_{\alpha_j, 1} = \|\chi\|_{\alpha_j}.$$

U nastavku ćemo nam zatrebati još neki rezultati koje ćemo na ovom mjestu navesti.

Pretpostavimo da je sa $b_j(\cdot, \cdot)$ označena simetrična, pozitivno određena bilinearna forma koja je komparabilna sa L_2 -skalarnim proizvodom u smislu da postoji konstanta β nezavisna od R_j , takva da je

$$19) \quad 0 < \beta^{-1} \leq \frac{b_j(v, v)}{(v, v)} \leq \beta \quad \text{za sve } v \in \mathcal{M}_j, v \neq 0.$$

vezi sa tom bilinearnom formom definisemo sopstvene funkcije

$\tilde{\psi}_i^{(j)} = \tilde{\psi}_i$ i sopstvene vrijednosti $\tilde{\lambda}_i^{(j)} = \tilde{\lambda}_i, 1 \leq i \leq N_j$, sa

$$20) \quad a(\tilde{\psi}_i, v) + d_j(\tilde{\psi}_i, v) = \tilde{\lambda}_i b_j(\tilde{\psi}_i, v) \quad \text{za sve } v \in \mathcal{M}_j.$$

ve sopstvene vrijednosti $\tilde{\lambda}_i$ ce takodje biti pozitivne,

kao i ranije možemo, ne umanjujući opštost, pretpostaviti

da je

$$21) \quad \begin{aligned} 0 < \tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots \leq \tilde{\lambda}_{N_j} \\ b_j(\tilde{\psi}_i, \tilde{\psi}_k) = \delta_{ik}, \quad a(\tilde{\psi}_i, \tilde{\psi}_k) + d_j(\tilde{\psi}_i, \tilde{\psi}_k) = \tilde{\lambda}_i \delta_{ik}. \end{aligned}$$

na osnovu (17) i (19) izvodimo ocjenu za $\tilde{\lambda}_{N_j}$:

$$22) \quad \tilde{\lambda}_{N_j} = \max_{0 \neq u \in \mathcal{M}_j} \frac{a(u, u) + d_j(u, u)}{b_j(u, u)} \leq \beta \max_{0 \neq u \in \mathcal{M}_j} \frac{a(u, u) + d_j(u, u)}{(u, u)} = \beta \lambda_{N_j} \leq C h_j^{-2}.$$

za proizvoljnom elementu $x \in \mathcal{M}_j$ odgovara sljedeći razvoj

u sopstvenim funkcijama $\tilde{\psi}_i$:

$$x = \sum_{i=1}^{N_j} \tilde{c}_i \tilde{\psi}_i.$$

definiseмо skalu normi

$$23) \quad \|x\|_{\mathcal{M}_j, s, b}^2 = \sum_{i=1}^{N_j} \tilde{c}_i^2 \tilde{\lambda}_i^{-s}, \quad -2 \leq s \leq 2.$$

u dodatku rada [3] je za takve norme dokazana

L e m a 1.

Postoji konstanta C takva da je, za sve $0 \leq s \leq 1$ i sve $x \in \mathcal{M}_j$

$$24) \quad C^{-1} \|x\|_s \leq \|x\|_{\mathcal{M}_j, s, b} \leq C \|x\|_s. \quad \square$$

3. Definicija k -stepene iteracione procedure

opisacemo algoritam za izvršavanje k -stepene procedure za rješavanje zadatka oblika (13) za $j=k$.

1) Ako je $k=1$, treba zadatak (13) riješiti nekim direktnim metodom, tj. tačno.

Ako je $k > 1$, tada se jedan iteracioni korak k -stepene procedure sastoji od dvije etape. U prvoj etapi se uzima početna vrijednost $z_0 \in M_k$ (to je u stvari prethodna iteracija) i određuje se vrijednost $z_m \in M_k$ (m će biti određeno kasnije) sukcesivno na osnovu formule

$$5) \quad (z_k - z_{k-1}, v) = (\mu_{N_k} + \alpha_k)^{-1} \{ G(v) - \alpha(z_{k-1}, v) - d_k(z_{k-1}, v) \}$$

sve $v \in M_k$.

ka je $\bar{q} \in M_{k-1}$ rješenje zadatka

$$6) \quad \alpha(\bar{q}, v) + d_{k-1}(\bar{q}, v) = G(v) - \alpha(z_m, v) - d_{k-1}(z_m, v) \equiv \bar{G}(v),$$

sve $v \in M_{k-1}$.

načimo sa q , aproksimaciju tačnog rješenja \bar{q} zadatka (26) redjenu primjenom p iteracija $k-1$ -stepene procedure, pri čemu se startovalo od nulte početne iteracije. Broj iteracija p će biti kasnije određen.

redivši na taj način q , stavljamo

$$7) \quad z_{m+1} = z_m + q.$$

Prema tome, jedan iteracioni korak k -stepene procedure se sastoji u prelazu od vrijednosti $z_0 \in M_k$ na vrijednost $z_{m+1} \in M_k$, pri čemu se prelaz vrši u dvije etape na opisani način.

Načinom (25) je definisana procedura glačanja koja ima za cilj redukciju visokofrekventnih komponenti greske. Kao što je poznato, ta procedura ne utiče bitno na niskofrekventne komponente greške, ali se pokazuje da se greška u kojoj dominantnu ulogu imaju glatke, niskofrekventne komponente, može dobro aproksimirati elementom iz M_{k-1} . Rješenje zadatka (6) i predstavlja taj element iz M_{k-1} sa kojim se aproksimira greška. Pošto je dimenzija prostora M_{k-1} skoro četiri puta manja od dimenzije prostora M_k , sa mnogo manje

čunskog posla je moguće odrediti aproksimaciju greške u prostoru M_{k-1} nego u prostoru M_k . Tako određena aproksimacija greške može poslužiti kao popravka prethodne iteracije iz M_k , poslije čega dobijamo sljedeću iteraciju. Jednačina (26) je defektna jednačina. Procedura glačanja definirana sa (25) zahtijeva da se na svakom koraku rješavaju sustemi linearnih jednačina sa Gramm-ovom matricom $M = ((\varphi_i^{(k)}, \varphi_j^{(k)}))$. Ovdje ćemo vidjeti da se taj nedostatak može izbjeći definisanjem glačanja na nešto drugačiji način. Ta izmjena će olakšati posao, a neće uticati na osnovni rezultat o konvergenciji koji ćemo dobiti za k -stepenu proceduru sa glačanjem (25). Korištenjem glačanja u obliku (25) se pojednostavljuje početna analiza. Također i korištenje $\mu_{N_k} + \alpha_k$ u (25) ima za cilj pojednostavljenje dobijanja osnovnih rezultata. Umjesto $(\mu_{N_k} + \alpha_k)^{-1}$ može da stoji neka konstanta $\beta \leq (\mu_{N_k} + \alpha_k)^{-1}$.

4. Dokaz konvergencije

Ukazacemo da je k -stepena iteraciona procedura definirana sa (25)-(27) konvergentna.

T e o r e m a 1.

Ako važi (10) i neka je $p > 1$ cio broj. Tada postoji konstanta $\gamma < 1$ i cio broj $m > 1$, oboje nezavisni od k , tako da, ako je

$$(28) \quad \| \bar{q}_k - q_k \|_{\Delta_{k-1}} \leq \gamma^p \| \bar{q}_k \|_{\Delta_{k-1}},$$

tada je

$$(29) \quad \| z - z_{m+1} \|_{\Delta_k} \leq \gamma \| z - z_0 \|_{\Delta_k}.$$

D o k a z .

ka je

$$\varepsilon_\ell = \bar{z}_\ell - z, \quad 0 \leq \ell \leq m+1$$

ška. Tada je na osnovu (25)

$$b) \quad (\varepsilon_\ell, v) = (\varepsilon_{\ell-1}, v) - (\mu_{N_k} + d_k)^{-1} [\alpha(\varepsilon_{\ell-1}, v) + d_k(\varepsilon_{\ell-1}, v)], \quad 1 \leq \ell \leq m.$$

c) pretpostavimo da greški ε_0 odgovara razvoj

$$\varepsilon_0 = \sum_{i=1}^{N_k} c_i \psi_i,$$

da se korištenjem (30) dokazuje da greški ε_ℓ odgovara razvoj

$$1) \quad \varepsilon_\ell = \sum_{i=1}^{N_k} c_i \left(1 - \frac{\mu_i + d_k}{\mu_{N_k} + d_k} \right)^\ell \psi_i, \quad 0 \leq \ell \leq m.$$

avde se odmah dobija ocjena

$$2) \quad \|\varepsilon_\ell\|_{d_k} \leq \|\varepsilon_0\|_{d_k}.$$

tektnu jednačinu (26) možemo napisati u obliku

$$3) \quad \alpha(\bar{q}_m - \varepsilon_m, v) + d_{k-1}(\bar{q}_m - \varepsilon_m, v) = 0 \quad \text{za sve } v \in \mathcal{M}_{k-1}$$

c) uvedemo skalarni proizvod

$$\langle u, v \rangle_{d_{k-1}} = \alpha(u, v) + d_{k-1}(u, v),$$

da jednakost (33) možemo shvatiti i u sljedećem smislu:

ε_m je ortogonalno na prostor \mathcal{M}_{k-1} u smislu skalarnog proizvoda $\langle \cdot, \cdot \rangle_{d_{k-1}}$ pa je \bar{q}_m ortogonalna projekcija greške ε_m na \mathcal{M}_{k-1} . To znači da je \bar{q}_m element iz \mathcal{M}_{k-1} koji naj-
lje aproksimira ε_m u smislu energetske norme $\|\cdot\|_{d_{k-1}}$.

avljajući u (33) $v = \bar{q}_m$, dobijamo ocjenu

$$1) \quad \|\bar{q}_m\|_{d_{k-1}} \leq \|\varepsilon_m\|_{d_{k-1}}.$$

risteci nejednakosti (28), (34) i (32), dobijamo

$$2) \quad \|\bar{q}_m - q\|_{d_{k-1}} \leq \gamma^p \|\varepsilon_0\|_{d_{k-1}}.$$

redjivanje greške ε_m zamjenjujemo odredjivanjem njene apro-
imacije \bar{q}_m iz \mathcal{M}_{k-1} .

i se dobio sistem

$$(A^2 + \alpha E)x = Ab.$$

omak spektra je kasnije i uvršten u teoriju metoda regularizacije pod nazivom uprošćena regularizacija (vidjeti [40, 42, 41]). Ranije je rečeno da izbor regularizacionog operatora, odnosno regularizacionog algoritma, nije jednoznačan. Dokazamo da ako umjesto jedinične matrice E u (9) stoji pozitivno određena matrica M , koju ćemo kasnije odrediti, da se na taj način biti takodje formiran jedan algoritam za određivanje aproksimacije normalnog rješenja sistema (8) koji će biti regularizacioni u smislu Tihonova. Pretpostavimo da je u vezi sa prostorom linearnih konačnih elemenata M_j

obijen sistem jednačina oblika (8) kojeg ćemo označiti sa (10)

$$A^{(j)} x^{(j)} = b^{(j)}.$$

Učemo praviti razliku između singularnih i loše uslovljenih sistema jer ćemo za rješavanje i jednih i drugih koristiti isti metod. Takvi sistemi imaju jednu zajedničku osobinu - osjetljivi su na male promjene desne strane. U slučaju singularnih sistema te promjene mogu dovesti do nesaglasnih sistema, dok kod loše uslovljenih sistema utiču na veće promjene rješenja.

Pretpostavimo da je umjesto tačne desne strane $\bar{b}^{(j)}$ poznata neka njena aproksimacija $b_{\delta_j}^{(j)}$ i broj $\delta_j > 0$ tako da je

$$(11) \quad \|\bar{b}^{(j)} - b_{\delta_j}^{(j)}\| = \delta_j.$$

Zadatak se sastoji u tome da se na osnovu poznatih podataka $(A^{(j)}, b_{\delta_j}^{(j)}, \delta_j)$ odredi aproksimacija $x_{\delta_j}^{(j)}$ vektora $\bar{x}^{(j)}$, gdje je $\bar{x}^{(j)}$ normalno rješenje sistema (10) sa tačnom desnom stranom $\bar{b}^{(j)}$, tako da ta aproksimacija bude stabilna u odnosu na male pro-

njene desne strane.

Umjesto sistema

$$(12) \quad A^{(j)} x^{(j)} = b_{\delta_j},$$

razmatraćemo sistem

$$(13) \quad (A^{(j)} + \alpha M^{(j)}) x^{(j)} = b_{\delta_j}, \quad \alpha > 0,$$

gdje je $M^{(j)}$ - Gramm-ova matrica (matrica mase) sa elementima

$$M_{ik}^{(j)} = (\varphi_i^{(j)}, \varphi_k^{(j)}),$$

gdje su $\varphi_i^{(j)}$, $i=1,2,\dots,N_j$, bazne funkcije prostora \mathcal{M}_j i $N_j = \dim \mathcal{M}_j$

sistem (12) odgovara varijacionom zadatku:

naći $u_j \in \mathcal{M}_j$ tako da je

$$(14) \quad \alpha(u_j, v) = \langle f_{\delta_j}, v \rangle \quad \text{za sve } v \in \mathcal{M}_j,$$

gdje je f_{δ_j} funkcija takva da je $\|f_{\delta_j} - f\| \leq \delta_j$.

Uvodimo sljedeći linearni funkcional određen na \mathcal{M}_j :

$$(15) \quad P_{\alpha}[u, f_{\delta_j}] = \alpha(u, u) - 2\langle f_{\delta_j}, u \rangle + \alpha(u, u).$$

U vezi sa ovim funkcionalom posmatrajmo ekstremalni zadatak:

naći element u^+ iz prostora \mathcal{M}_j na kome funkcional (15) do-

stiže svoj minimum.

L e m a 1 .

za svako $\alpha > 0$ i svaku funkciju $f_{\delta_j} \in L_2(\Omega)$, definisani ekstre-

malni zadatak ima rješenje, to rješenje je jedinstveno i za

njega važi ocjena

$$(16) \quad \|u^+\| \leq \frac{1}{\alpha} \|f_{\delta_j}\|.$$

D o k a z .

Varijacijom funkcional (15) (vidjeti [27]), dobijamo

$$\delta P_{\alpha}[u, f_{\delta_j}] = 2\alpha(u, v) - 2\langle f_{\delta_j}, v \rangle + 2\alpha(u, v)$$

$$\delta^2 P_{\alpha}[u, f_{\delta_j}] = 2\alpha(v, v) + 2\alpha(v, v),$$

gdje je v proizvoljni element iz \mathcal{M}_j , a sa δP_{α} , $\delta^2 P_{\alpha}$ su

neka su $u = \sum_{i=1}^{N_k} c_i \psi_i$ i $v = \sum_{i=1}^{N_k} \tilde{c}_i \psi_i$ dva proizvoljna elementa iz \mathcal{M}_k .
 tada je

$$\begin{aligned} \langle u, v \rangle_{d_k} &= \alpha(u, v) + d_k(u, v) = \alpha\left(\sum_i c_i \psi_i, \sum_j \tilde{c}_j \psi_j\right) + d_k\left(\sum_i c_i \psi_i, \sum_j \tilde{c}_j \psi_j\right) = \\ &= \sum_{i,j} c_i \tilde{c}_j [\alpha(\psi_i, \psi_j) + d_k(\psi_i, \psi_j)] = \sum_{i,j} c_i \tilde{c}_j \lambda_i \delta_{ij} = \sum_i c_i \tilde{c}_i \lambda_i = \\ &= \sum_i c_i \lambda_i^{\frac{1-\theta}{2}} \cdot \tilde{c}_i \lambda_i^{\frac{1+\theta}{2}} \leq \left(\sum_i c_i^2 \lambda_i^{1-\theta}\right)^{1/2} \cdot \left(\sum_i \tilde{c}_i^2 \lambda_i^{1+\theta}\right)^{1/2} = \\ &= \|u\|_{d_k, 1-\theta} \cdot \|v\|_{d_k, 1+\theta}, \quad 0 \leq \theta \leq 1. \end{aligned}$$

primjenom ove nejednakosti izvodimo ocjenu

$$\begin{aligned} (6) \quad \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}}^2 &= \alpha(\bar{q}_m - \varepsilon_m, \bar{q}_m - \varepsilon_m) + d_{k-1}(\bar{q}_m - \varepsilon_m, \bar{q}_m - \varepsilon_m) = \\ &= -\alpha(\bar{q}_m - \varepsilon_m, \varepsilon_m) - d_{k-1}(\bar{q}_m - \varepsilon_m, \varepsilon_m) = \\ &= -\langle \bar{q}_m - \varepsilon_m, \varepsilon_m \rangle_{d_{k-1}} \leq \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}, 1-\theta} \cdot \| \varepsilon_m \|_{d_{k-1}, 1+\theta} \end{aligned}$$

gdje ćemo ocijeniti faktor $\| \varepsilon_m \|_{d_{k-1}, 1+\theta}$:

$$\begin{aligned} \| \varepsilon_m \|_{d_{k-1}, 1+\theta}^2 &= \sum_{i=1}^{N_k} c_i^2 (\mu_i + d_{k-1})^{1+\theta} \cdot \left(1 - \frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}}\right)^{2m} = \\ &= \sum_{i=1}^{N_k} c_i^2 (\mu_{N_k} + d_{k-1})^\theta \cdot \left(\frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}}\right)^\theta (\mu_i + d_{k-1}) \left(1 - \frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}}\right)^{2m} \leq \\ &\leq (\mu_{N_k} + d_{k-1})^\theta \cdot \max_i \left[\left(\frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}}\right)^\theta \cdot \left(1 - \frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}}\right)^{2m} \right] \cdot \sum_{i=1}^{N_k} c_i^2 (\mu_i + d_{k-1}) \leq \\ &\leq (\mu_{N_k} + d_{k-1})^\theta \cdot \max_{x \in [0,1]} x^\theta (1-x)^{2m} \cdot \| \varepsilon_0 \|_{d_{k-1}}^2. \end{aligned}$$

ko funkcija $x^\theta (1-x)^{2m}$ dostiže svoj maksimum u tački $x = \frac{\theta}{2m+\theta}$,

osnovu prethodne ocjene i (17) konačno dobijamo

$$(7) \quad \| \varepsilon_m \|_{d_{k-1}, 1+\theta} \leq C \nu_k^{-\theta} m^{-\theta/2} \| \varepsilon_0 \|_{d_{k-1}}.$$

taje da se ocijeni faktor $\| \bar{q}_m - \varepsilon_m \|_{d_{k-1}, 1-\theta}$. Za to će nam biti potreban standardni argument dualnosti i Lema 1.

neka $g \in H^{\theta-1}(\Omega)$ i neka $\eta \in H^{1+\theta}(\Omega)$ zadovoljava

$$\alpha(\eta, v) + d(\eta, v) = (g, v) \quad \text{za sve } v \in H^1(\Omega).$$

avljajući $v = \bar{q}_m - \varepsilon_m$, imamo za proizvoljno $\chi \in \mathcal{M}_{k-1}$

$$\begin{aligned} (g, \bar{q}_m - \varepsilon_m) &= \alpha(\eta, \bar{q}_m - \varepsilon_m) + d_{k-1}(\eta, \bar{q}_m - \varepsilon_m) = \\ &= \alpha(\eta - \chi, \bar{q}_m - \varepsilon_m) + d_{k-1}(\eta - \chi, \bar{q}_m - \varepsilon_m) \leq \\ &\leq \| \eta - \chi \|_{d_{k-1}} \cdot \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}}. \end{aligned}$$

og proizvoljnosti $\chi \in \mathcal{M}_{k-1}$ je

$$3) \quad (\varrho, \bar{q}_m - \varepsilon_m) \leq \| \bar{q}_m - \varepsilon_m \| \cdot \inf_{\chi \in \mathcal{M}_{k-1}} \| \eta - \chi \|_{d_{k-1}} \leq C \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}} \cdot \inf_{\chi \in \mathcal{M}_{k-1}} \| \eta - \chi \|_{d_{k-1}} \leq C \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}} \cdot \rho_k^\theta \| \eta \|_{1+\theta} \leq C \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}} \cdot \rho_k^\theta \| \eta \|_{\theta-1},$$

je smo za dobijanje konačne ocjene koristili Lema 1, apro-
imaciono svojstvo (11) sa $s = 1 + \theta$ i pretpostavku regula-
osti (10) za dualni zadatak.

osnovu Leme 1 i (38) je

$$9) \quad \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}, 1-\theta} \leq C \| \bar{q}_m - \varepsilon_m \|_{1-\theta} \leq C \rho_k^\theta \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}}.$$

racajuci se na (36) i koristeći (37) i (39), imamo

$$o) \quad \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}} \leq C m^{-\theta/2} \| \varepsilon_0 \|_{d_{k-1}}.$$

načno je iz (27), (28), (39), (34) i (32)

$$1) \quad \| \varepsilon_{m+1} \|_{d_{k-1}} = \| \varepsilon_m - \varrho \|_{d_{k-1}} \leq \| \bar{q}_m - \varrho \|_{d_{k-1}} + \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}} \leq \gamma^P \| \bar{q}_m \|_{d_{k-1}} + C m^{-\theta/2} \| \varepsilon_0 \|_{d_{k-1}} \leq \gamma^P \| \varepsilon_0 \|_{d_{k-1}} + C m^{-\theta/2} \| \varepsilon_0 \|_{d_{k-1}} = (\gamma^P + C m^{-\theta/2}) \| \varepsilon_0 \|_{d_{k-1}}.$$

ma je potrebna ocjena ovog oblika ali u normi $\| \cdot \|_{d_k}$.

ovom mjestu cemo razmotriti dva različita načina izbora
parametra regularizacije:

stepen	1. način	2. način
k	d	d
$k-1$	d	$2 \cdot d$
\vdots	\vdots	\vdots
2	d	$\frac{k-2}{2} \cdot d$
1	d	$\frac{k-1}{2} \cdot d$

o se koristi 1. način za izbor parametra regularizacije,

osnovu (41) je odmah

$$2) \quad \| \varepsilon_m \|_{d_k} \leq (\gamma^P + C m^{-\theta/2}) \| \varepsilon_0 \|_{d_k}.$$

racajuci γ tako da je $\gamma^P < \frac{1}{2} \gamma$ i m dovoljno veliko da bude

$m^{-\theta/2} < \frac{1}{2} \gamma$, dobijamo iz (42) ocjenu

$$\| \varepsilon_{m+1} \|_{d_k} \leq \gamma \| \varepsilon_0 \|_{d_k},$$

..

$$\| \bar{z} - \bar{z}_{m+1} \|_{d_k} \leq \gamma \| \bar{z} - \bar{z}_0 \|_{d_k},$$

to je i trebalo dokazati.

Uzmotrimo slučaj kada je $d_{k-1} = 2 d_k$.

Ako je uz tu pretpostavku $\| \varepsilon_m \|_{d_k} \leq \| \varepsilon_m \|_{d_{k-1}}$, imaćemo iz (41)

$$\begin{aligned} \| \varepsilon_{m+1} \|_{d_k}^2 &\leq (\gamma^P + C m^{-\theta/2})^2 \| \varepsilon_0 \|_{d_{k-1}}^2 = (\gamma^P + C m^{-\theta/2})^2 [\| \varepsilon_0 \|_{d_k}^2 + d_k(\varepsilon_0, \varepsilon_0)] \leq \\ &\leq (\gamma^P + C m^{-\theta/2})^2 [\| \varepsilon_0 \|_{d_k}^2 + \alpha(\varepsilon_0, \varepsilon_0) + d_k(\varepsilon_0, \varepsilon_0)] = \\ &= (\gamma^P + C m^{-\theta/2})^2 \cdot 2 \| \varepsilon_0 \|_{d_k}^2, \end{aligned}$$

nosno

$$3) \quad \| \varepsilon_{m+1} \|_{d_k} \leq (\sqrt{2} \gamma^P + \sqrt{2} C m^{-\theta/2}) \| \varepsilon_0 \|_{d_k}.$$

Uzajmici u ovom slučaju γ tako da je $\sqrt{2} \gamma^P < \frac{1}{2} \gamma$ i m dovoljno veliko da bude $\sqrt{2} C m^{-\theta/2} < \frac{1}{2} \gamma$, dobijamo traženu nejednakost (29).

Iskreno je da se γ i m mogu izabrati nezavisno od k_j . \square

Ukazaćemo konvergenciju k -stepenog metoda i u L_2 -normi.

U to nam je potrebna jača pretpostavka za zadatak (9). Na-

ime, moramo pretpostaviti H^2 -regularnost, tj. da (10) važi za $\theta=1$

$$\| u^{*j} \|_2 \leq C \| f \|_0.$$

Ovaj zahtjev H^2 -regularnosti nije suviše ograničavajući.

Upravo će biti ispunjen ako je Ω konveksna oblast i ako je fun-

kcija $\alpha(x)$ dovoljno glatka (vidjeti [8]).

T e o r e m a 2.

Ako su ispunjene pretpostavke Teoreme 1 za $\theta=1$. Postoji

konstanta $0 \leq \gamma < 1$ i neki cijeli broj $m \gg 1$, oboje nezavisni

od k , tako da, ako je

$$4) \quad \| \bar{q}_k - q_k \|_0 \leq \gamma^P \| \bar{q}_0 \|_0,$$

da je

$$5) \quad \|z - z_{m+1}\|_0 \leq \gamma \|z - z_0\|_0.$$

D o k a z.

isti način se, kao i u Teoremi 1, korištenjem razvoja greška ε_0 i ε_m po sopstvenim funkcijama, dobija ocjena

$$6) \quad \|\varepsilon_m\|_0 \leq \|\varepsilon_0\|_0$$

risteci se činjenicom da je $\|\cdot\|_{d_{k,0}} = \|\cdot\|_0$ i stavljajući u (38)

$\bar{q}_m - \varepsilon_m$ i $\theta = 1$, dobijamo

$$7) \quad \|\bar{q}_m - \varepsilon_m\|_0 \leq C \rho_k \|\bar{q}_m - \varepsilon_m\|_{d_{k-1}}.$$

bismo dobili nejednakost analognu nejednakosti (35), izvedemo prvo neke pomocne ocjene.

ko je

$$\begin{aligned} \|\bar{q}_m - \varepsilon_m\|_{d_{k-1}}^2 &= \alpha (\bar{q}_m - \varepsilon_m, \bar{q}_m - \varepsilon_m) + d_{k-1} (\bar{q}_m - \varepsilon_m, \bar{q}_m - \varepsilon_m) = \\ &= -\alpha (\bar{q}_m - \varepsilon_m, \varepsilon_m) - d_{k-1} (\bar{q}_m - \varepsilon_m, \varepsilon_m) = -\langle \bar{q}_m - \varepsilon_m, \varepsilon_m \rangle_{d_{k-1}} \leq \\ &\leq \|\bar{q}_m - \varepsilon_m\|_{d_{k-1}} \cdot \|\varepsilon_m\|_{d_{k-1}}, \end{aligned}$$

ce

$$8) \quad \|\bar{q}_m - \varepsilon_m\|_{d_{k-1}} \leq \|\varepsilon_m\|_{d_{k-1}}.$$

risteci razvoj (31) za ε_m , izvodimo ocjenu

$$9) \quad \|\varepsilon_m\|_{d_{k-1}} \leq (\mu_{H_k} + d_{k-1})^{1/2} \|\varepsilon_m\|_0,$$

koje, zbog (46), dobijamo

$$\|\varepsilon_m\|_{d_{k-1}} \leq (\mu_{H_k} + d_{k-1})^{1/2} \|\varepsilon_0\|_0.$$

osnovu (44), (47), (48), (49) i (17) se dobija

$$\begin{aligned} 10) \quad \|\bar{q}_m - q_0\|_0 &\leq \gamma^P \|\bar{q}_m\|_0 \leq \gamma^P (\|\bar{q}_m - \varepsilon_m\|_0 + \|\varepsilon_m\|_0) \leq \gamma^P (C \rho_k \|\bar{q}_m - \varepsilon_m\|_{d_{k-1}} + \|\varepsilon_0\|_0) \leq \\ &\leq \gamma^P [C \rho_k (\mu_{H_k} + d_{k-1})^{1/2} \|\varepsilon_0\|_0 + \|\varepsilon_0\|_0] \leq C \gamma^P \|\varepsilon_0\|_0, \end{aligned}$$

o predstavlja analog od (35).

risteci (47) i (36) za $\theta = 1$, dobijamo

$$\|\bar{q}_m - \varepsilon_m\|_0^2 \leq C \rho_k^2 \|\bar{q}_m - \varepsilon_m\|_{d_{k-1}}^2 = C \rho_k^2 [\alpha (\bar{q}_m - \varepsilon_m, \bar{q}_m - \varepsilon_m) + d_{k-1} (\bar{q}_m - \varepsilon_m, \bar{q}_m - \varepsilon_m)] =$$

$$\begin{aligned}
 &= -C h_k^2 \langle \bar{q}_m - \varepsilon_m, \varepsilon_m \rangle_{d_{k-1}} \leq C h_k^2 \| \bar{q}_m - \varepsilon_m \|_{d_{k-1}, 0} \cdot \| \varepsilon_m \|_{d_{k-1}, 2} = \\
 &= C h_k^2 \| \bar{q}_m - \varepsilon_m \|_0 \cdot \| \varepsilon_m \|_{d_{k-1}, 2}
 \end{aligned}$$

.

$$1) \quad \| \bar{q}_m - \varepsilon_m \|_0 \leq C h_k^2 \| \varepsilon_m \|_{d_{k-1}, 2}.$$

ijenice $\| \varepsilon_m \|_{d_{k-1}, 2}$ koristeći razvoj za ε_m po sopstvenim funkcijama ψ_i :

$$\begin{aligned}
 \| \varepsilon_m \|_{d_{k-1}, 2}^2 &= \sum_{i=1}^{N_k} C_i^2 (\mu_i + d_{k-1})^2 \left(1 - \frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}} \right)^{2m} = \\
 &= \sum_{i=1}^{N_k} C_i^2 \left(\frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}} \right)^2 \cdot (\mu_{N_k} + d_{k-1})^2 \left(1 - \frac{\mu_i + d_{k-1}}{\mu_{N_k} + d_{k-1}} \right)^{2m} \leq \\
 &\leq (\mu_{N_k} + d_{k-1})^2 \cdot \max_{x \in (0, 1)} x^2 (1-x)^{2m} \cdot \| \varepsilon_0 \|_0^2.
 \end{aligned}$$

avde, na sličan način na koji je dobijena ocjena (37), dobijamo

$$2) \quad \| \varepsilon_m \|_{d_{k-1}, 2} \leq C h_k^{-2} m^{-1} \| \varepsilon_0 \|_0.$$

načno je, na osnovu (51), (50) i (52)

$$\begin{aligned}
 \| \varepsilon_{m+1} \|_0 &\leq \| \varepsilon_m - \bar{q}_m \|_0 + \| \bar{q}_m - q_m \|_0 \leq C h_k^2 \| \varepsilon_m \|_{d_{k-1}, 2} + C \gamma^p \| \varepsilon_0 \|_0 \leq \\
 &\leq C (m^{-1} + \gamma^p) \| \varepsilon_0 \|_0.
 \end{aligned}$$

rajuci γ tako da bude $C \gamma^p < \frac{1}{2} \gamma$ i m dovoljno veliko da bude $C m^{-1} < \frac{1}{2} \gamma$, dobijamo (45), čime je Teorema 2 u potpunosti dokazana. \square

nije je rečeno da procedura glačanja definisana sa (25) ima jedan nedostatak jer je na svakom koraku glačanja potrebno rješavati sistem sa matricom $M^{(k)} = ((\varphi_i^{(k)}, \varphi_j^{(k)}))$. Da bi se procedura glačanja pojednostavila, zamijenimo (25) sa

$$3) \quad \mathcal{L}_k(z_l - z_{l-1}, v) = \tilde{\lambda}_{N_k}^{-1} \{ G(v) - \alpha(z_{l-1}, v) - d_k(z_{l-1}, v) \} \quad \text{za sve } v \in \mathcal{M}_k.$$

koristeći nejednakost

$$\| \bar{q}_m - \varepsilon_m \|_{d_k}^2 \leq \| \bar{q}_m - \varepsilon_m \|_{d_{k+1}, 1-\theta, b} \cdot \| \varepsilon_m \|_{d_{k+1}, 1+\theta, b},$$

na analogan način se dokazuje

T e o r e m a 3.

Neka važe uslovi Teoreme 1 i neka je $p > 1$ neki cijeli broj.

la postoji konstanta $0 < \gamma < 1$ i cijeli broj $m \gg 1$ tako da, ako je ispunjen uslov (28) za iteracionu proceduru baziranu na (53), (26) i (27), tada važi uslov (29). \square

Prema tome, zamjena procedure glačanja ne utiče na konvergenciju k -stepene iteracione procedure.

Ako za bazu prostora konačnih elemenata \mathcal{M}_k izaberemo funkcije φ_i koje su jednake 1 u jednom čvoru, a 0 u svim ostalim čvorovima, glačajuća procedura (53) će biti, sa stanovišta računskog posla, mnogo bolja od glačanja definisanog sa (5) ako $b_k(\cdot, \cdot)$ izaberemo tako da dobijemo sistem jednačina sa dijagonalnom matricom $M^{(k)}$, što bi značilo da za glačanje koristimo Jacobi-jevu iteracionu šemu.

Ako su $u = \sum_{i=1}^{N_k} u_i \varphi_i$ i $v = \sum_{i=1}^{N_k} v_i \varphi_i$ dva proizvoljna elementa iz \mathcal{M}_k , definisanjem forme $b_k(\cdot, \cdot)$ na sljedeći način

$$b_k(u, v) = \sum_{i=1}^{N_k} u_i v_i (\varphi_i, \varphi_i),$$

jesto sistema sa matricom $M^{(k)}$, dobićemo sistem sa odgovarajućom dijagonalnom matricom.

Ako definisana forma $b_k(\cdot, \cdot)$ će zadovoljavati uslov (19)

ako je baza prostora \mathcal{M}_k izabrana tako da je norma $\|u\|_{L_2}^2 = \sum_{i=1}^{N_k} k^2 u_i^2$ ekvivalentna sa normom $\|u\|_0$, tj.

$$C^{-1} \|u\|_{L_2} \leq \|u\|_0 \leq C \|u\|_{L_2},$$

je se pod $\|u\|_{L_2}$ podrazumijeva norma vektora $u = (u_1, u_2, \dots, u_{N_k})^T$,

pod $\|u\|_0$ - L_2 -norma funkcije $u = \sum_{i=1}^{N_k} u_i \varphi_i$.

Da taj uslov ispunjava baza $\{\varphi_i\}$ za koju je $\varphi_i = 1$ u jednom čvoru i $\varphi_i = 0$ u ostalim čvorovima (vidjeti [8]).

5. Računski posao i efikasnost

opreme dokazane u prethodnom paragrafu pokazuju da k -stepena iteraciona procedura za rješavanje sistema linearnih gebarskih jednačina oblika (7) na stepenu $j=k$, ima faktor konvergencije γ koji ne zavisi od parametra diskretizacije Δ , zbog čega je taj iteracioni metod mnogo bolji od klasičnih iteracionih metoda koji nemaju to svojstvo. Ali, sama ta činjenica ne govori da smo time dobili optimalan metod. Potrebno je imati u vidu i računski posao koji je potreban za realizaciju tog metoda.

Bismo odredili ukupnu količinu potrebnog računskog posla, odimo sljedeće oznake

- Q_{j-1}^j - računski posao dvostepenog metoda izuzimajući posao potreban za rješavanje defektne jednačine, $j=2, \dots, k$,
- Q_1 - računski posao potreban za tačno rješavanje defektne jednačine na stepenu $j=1$,
- Q_j - računski posao j -stepenog metoda na jednom koraku, $j=2, \dots, k$.

Na osnovu ovih oznaka, naš je cilj da odredimo Q_k . Jasno je da je

$$Q_2 = Q_1^2 + Q_1, \quad Q_j = Q_{j-1}^j + p Q_{j-1}, \quad j=3, \dots, k.$$

Prema tome, računski posao potreban za jedan korak k -stepene iteracione procedure će biti

$$Q_k = \sum_{j=2}^k p^{k-j} Q_{j-1}^j + p^{k-2} Q_1.$$

Pretpostavimo da je broj računskih operacija potrebnih za izvođenje (27) ograničen konstantom $C_1 m N_j = C N_j$, tj.

$$Q_{j-1}^j \sim CN_j.$$

la je

$$Q_k \sim \sum_{j=2}^k p^{k-j} CN_j + p^{k-2} Q_1.$$

obzirom da je

$$N_j \sim \frac{N_k}{4^{k-j}},$$

de

$$Q_k \sim CN_k \sum_{j=2}^k \left(\frac{p}{4}\right)^{k-j} + p^{k-2} Q_1.$$

avde se dobija da je ukupan računski posao na jednom koraku dat sa

$$Q_k = \begin{cases} \frac{4}{3} CN_k & , p=1 \\ 2 CN_k & , p=2 \\ 4 CN_k & , p=3 \\ O(N_k \log N_k) & , p=4 . \end{cases}$$

sno je da će broj računskih operacija biti proporcionalan brojem nepoznatih ako se defektna jednačina rješava $(p-1)$ -stepenim metodom sa $p \leq 3$ iteracija.

da je taj broj iteracija $p > 4$, nećemo imati ekonomičan metod. Zbog toga su za nas interesantni slučajevi $p=2$ i $p=3$.

tim slučajevima će ranije definisana k -stepena iteraciona procedura imati sljedeća dva svojstva:

njen faktor konvergencije ne zavisi od parametra h_k

ukupan broj računskih operacija na jednom iteracionom koraku je proporcionalan broju nepoznatih.

Dakle, ovaj je uveden metod asimptotski optimalan (kada $k \rightarrow \infty$).

LITERATURA

- Aziz A.K., Babuška I., Mathematical foundations of the finite element method, New-York - London, Academic Press, 1972.
- Babuška I., The finite element method, Lecture on Boston international conference, 1970.
- Bank R.E., Dupont T., An optimal order process for solving elliptic finite element equations, *Math. Comp.*, 36, 1981, 35-51.
- Bank R. E., Analysis of a multilevel inverse iteration procedure for eigenvalue problems, *SIAM J. Numer. Anal.* 19, № 5, 1982.
- Braess D., Hackbusch W., A new convergence proof for the multigrid method including the V-cycle, *SIAM J. Numer. Anal.*, № 5, 1983, 967-975.
- Brandt A., Multi-level adaptive solutions to boundary-value problems, *Math. Comp.*, 31, 1977, 333-390.
- Brandt A., Guide to multigrid development, u [11]
- Ciarlet Ph., The finite element method for elliptic problems, North-Holland Publishing Company, Amsterdam - New-York-Oxford, 1978.
- Forsythe G.E., Wasow W.R., Finite difference methods for partial differential equations, Wiley, New-York, 1960.
- Hackbusch W., On the convergence of multi-grid iterations, *Beiträge zur Numer. Math.*, 9, 1981, 213-239.
- Hackbusch W., Trottenberg U., (editors), Multigrid methods, Proceedings of the conference held at Köln-Porz 1981., *Lecture Notes in mathematics*, 960, Springer-Verlag, Berlin, 1982.

- . Hackbusch W., Multi-grid convergence theory, u [11]
- . Keller B.H., On the solution of singular and semidefinite linear systems by iteration, SIAM J. Numer. Anal., 2, 1965, 1-290.
- . Lawson C.L., Hanson H.J., Solving least squares problems, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1974.
- . Lions J.L., Magenes E., Non-homogeneous boundary value problems and applications, vol I, Springer Verlag, Berlin - Heidelberg - New-York, 1972.
- . Marek I., Iterative methods of solving linear systems with a rectangular matrix, Report 8132, Katholieke Universiteit, Nijmegen, The Netherlands, 1981.
- . Molchanov I.N., Nikolenko L.D., On an approach to integrating boundary problems with a non-unique solution, Inform. Process. Lett., 1, № 4, 1972, 168-172.
- . Nicolaidis R.A., On the ℓ^2 -convergence of an algorithm for solving finite element equations, Math. Comp., 31, 1977, 2-906.
- . Nicolaidis R.A., On multigrid convergence in the indefinite case, Math. Comp., 32, 1978, 1082-1086.
- . Peaceman D.W., Rachford H.H., The numerical solution of parabolic and elliptic differential equations, J. Soc. Indust. Math., 3, № 1, 1955, 28-41.
- . Plemmons R.J., Regular splittings and the discrete Neumann problem, Numer. Math., 25, 1976, 153-161.
- . Rektorys K., Variational methods in mathematics, science

- d engineering, D. Reidel Publishing Company, Dordrecht-Holland, Dordrecht-USA, London-England, 1980.
- 1. Strang G., Fix G., An analysis of the finite element method, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.
 - 2. Stüben K., Trottenberg U., Multigrid methods: fundamental algorithms, model problem analysis and applications, u [11]
 - 3. Varga R.S., Matrix iterative analysis, Prentice Hall, Englewood Cliffs, New Jersey, 1962.
 - 4. Young D.M., Iterative solution of large linear systems, Academic Press, New-York - London, 1971.
 - 5. Астраханцев Г.П., Об одном итерационном методе решения точных эллиптических задач, ЖВММФ, II, №2, 1971, 439-448.
 - 6. Бахвалов Н.С., О сходимости одного релаксационного метода при естественных ограничениях на эллиптический оператор, ЖВММФ, 6, №5, 1966, 861-883.
 - 7. Винокуров В.А. Два замечания о выборе параметра регуляризации, ЖВММФ, 12, №2, 1972, 481-483.
 - 8. Гордонова В.И., Морозов В.А., Численные алгоритмы выбора параметра в методе регуляризации, ЖВММФ, 13, №3, 1973, 539-545.
 - 9. Колмогоров А.Н., Фомин С.В., Элементы теории функций функционального анализа, Наука, Москва, 1981.
 - 10. Ладыженская О.А., Краевые задачи математической физики, Наука, Москва, 1973.
 - 11. Ладыженская О.А., Уральцева Н.Н., Линейные и квазилинейные уравнения эллиптического типа, Наука, Москва, 1973.

1. Леонов А.С., К обоснованию выбора параметра регуляризации по критериям квазиоптимальности и отношений, ЖВМиМФ, 18, 1978, 1363-1376.
2. Марчук Г.И., Методы вычислительной математики, Наука, Москва, 1980.
3. Марчук Г.И., Агошков В.И., Введение в проекционно-сеточные методы, Наука, Москва, 1981.
4. Михайлов В.П., Дифференциальные уравнения в частных производных, Наука, Москва, 1983.
5. Митчелл Э., Уэйт Р., Метод конечных элементов для уравнений с частными производными, Мир, Москва, 1981.
6. Молчанов И.Н., Численные методы решения некоторых задач теории упругости, Наукова думка, Киев, 1979.
7. Молчанов И.Н., Николенко Л.Д., Об одном прямом методе решения задач Неймана для уравнения Пуассона, ЖВМиМФ, 13, №6, 1973, 1607-1612.
8. Морозов В.А., О принципе невязки при решении несовместных уравнений методом регуляризации А.Н.Тихонова, ЖВМиМФ, 13, №5, 1973, 1099-1111.
9. Музилёв Н.В., Об алгоритме упрощенной регуляризации, ЖВМиМФ, 15, №3, 1975, 772-775.
10. Норри Д., де Фриз Ж., Введение в метод конечных элементов, Мир, Москва, 1981.
11. Оганесян Л.А., Руховец Л.А., Вариационно-разностные методы решения эллиптических уравнений, АН АССР, Ереван, 1979.
12. Самарский А.А., Николаев Е.С., Методы решения сеточных уравнений, Наука, Москва, 1978.

1. Самарский А.А., Введение в численные методы, Наука, Москва, 1982.
2. Самарский А.А., Теория разностных схем, Наука, Москва, 1983.
3. Тихонов А.Н., О решении некорректно поставленных задач методом регуляризации, ДАН СССР, 1963, 151, №3, 501-504.
4. Тихонов А.Н., О регуляризации некорректно поставленных задач, ДАН СССР, 1963, 153, №1, 49-52.
5. Тихонов А.Н., Гласко В.В., О приближенном решении интегральных уравнений Фредгольма первого рода, ЖВМиМФ, 4, №3, 1964, 44-57.
6. Тихонов А.Н., О некорректных задачах линейной алгебры и устойчивом методе их решения, ДАН СССР, 1965, 163, №6, 591-594.
7. Тихонов А.Н., Об устойчивости алгоритмов для решения вырожденных систем линейных алгебраических уравнений, ЖВМиМФ, 5, №4, 1965, 718-722.
8. Тихонов А.Н., Самарский А.А., Уравнения математической физики, Наука, Москва, 1977.
9. Тихонов А.Н., Арсенин В.Я., Методы решения некорректных задач, Наука, Москва, 1979.
10. Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г., Регуляризирующие алгоритмы и априорная информация, Наука, Москва, 1983.
11. Фаддеев Д.К., Фаддеева В.Н., Вычислительные методы линейной алгебры, Физматгиз, Москва, 1963.
12. Фаддеева В.Н., Сдвиг для систем с плохо-обусловленными матрицами, ЖВМиМФ, 5, №5, 1965, 907-911.

3. Федоренко Р.П., Релаксационный метод решения разностными эллиптическими уравнениями, ЖВММФ, 1, №5, 1961, 922-927.
9. Федоренко Р.П., О скорости сходимости одного итерационного процесса, ЖВММФ, 4, №3, 1964, 559-564.

ОБЩЕУЧЕБНИКОВСКАЯ НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАДА
ЗА МАТЕМАТИКУ, МЕХАНИКУ И АСТРОНОМИЈУ
БЕОГРАД

Број: _____

Датум: _____