Numerička analiza

25. predavanje

Autor: Nela Bosner

Predavač: Nela Bosner

nela@math.hr

web.math.hr/~nela/nad.html

PMF – Matematički odjel, Zagreb

Sadržaj predavanja

- Numeričko rješavanje parcijalnih diferencijalnih jednadžbi:
  - Numeričko rješavanje paraboličke jednadžbe — difuzijske jednadžbe .
    - Eksplicitna metoda konačnih razlika.
    - Potpuna implicitna metoda konačnih razlika.
    - Crank–Nicolsonova metoda
  - Numeričko rješavanje eliptičke jednadžbe — Poissonove jednadžbe.

Numeričko rješavanje difuzijske jednadžbe

Difuzijska jednadžba

Difuzijska jedadžba opisuje razne procese u praksi:

- u fizici: distribuciju topline po vremenu u nekom objektu
- u financijama: ponašanje vrijednosti opcija (financijski instrument koji dozvoljava "klađenje" da li će vrijednost te imovine rasti ili padati)

Opći oblik difuzijske (paraboličke parcijalne diferencijalne) jednadžbe je

$$\frac{\partial u(x,t)}{\partial t} - \nabla \cdot (a(x)\ \nabla u(x,t)) = f(x,t), \quad t \in [0,T],\ x \in \Omega$$

pri čemu je potrebno još zadati

- inicijalni uvjet za $t = 0$, i
- rubne uvjete za $x \in \partial\Omega$.

Difuzijska jednadžba (nastavak)

Mi ćemo zbog jednostavnosti promatrati jednostavniji oblik jednadžbe

$$\frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t), \qquad x \in [a,b],\ t \in [0,T],$$

s Dirichletovim rubnim uvjetima.

Standardna metoda za dobivanje aproksimacija rješenja parcijalne diferencijalne jednadžbe, kao što je difuzijska, je diskretizacija pomoću konačnih diferencija.

- U toj metodi iz danog područja $[a,b] \times [0,T]$ izabran je skup točaka koji čini mrežu.
- U svakoj točki mreže derivacije u diferencijalnoj jednadžbi zamijenjuju se sa kvocijentima koji se približavaju derivaciji kada mreža postaje sve finija.

Konačne razlike

Parcijalnu derivaciju $\partial u/\partial t$ možemo definirati kao

$$\frac{\partial u}{\partial t}(x,t) = \lim_{\delta t \to 0} \frac{u(x,t+\delta t) - u(x,t)}{\delta t}.$$

Umjesto računanja limesa kada $\delta t \to 0$, uzet ćemo $\delta t > 0$ koji je vrlo mali, i budući da vrijedi

$$\frac{\partial u}{\partial t}(x,t) = \frac{u(x,t+\delta t) - y(x,t)}{\delta t} + \mathcal{O}(\delta t)$$

definirat ćemo aproksimaciju

$$\frac{\partial u}{\partial t}(x,t) \approx \frac{u(x,t+\delta t) - y(x,t)}{\delta t}.$$

Ovime smo dobili konačnu razliku od $\partial y/\partial t$, a konačnu razliku ovoga oblika posebno ćemo još zvati i konačna razlika unaprijed.

Konačne razlike (nastavak)

Alternativno možemo definirati

$$\frac{\partial u}{\partial t}(x,t) = \lim_{\delta t \to 0} \frac{u(x,t) - u(x,t-\delta t)}{\delta t},$$

tako da je na sličan način aproksimacija dana sa

$$\frac{\partial u}{\partial t}(x,t) \approx \frac{u(x,t) - u(x,t-\delta t)}{\delta t}.$$

Konačnu razliku ovoga oblika zovemo konačna razlika unazad.

Također možemo primijetiti da je

$$\frac{\partial u}{\partial t}(x,t) = \lim_{\delta t \to 0} \frac{u(x,t+\delta t) - u(x,t-\delta t)}{2\delta t}, \text{ i}$$

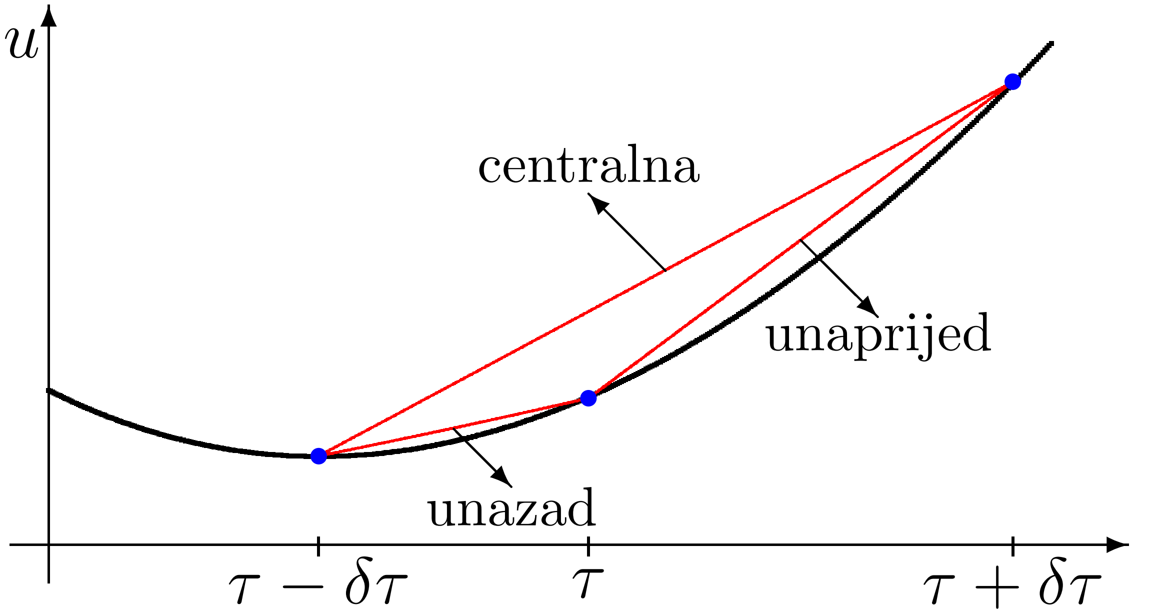
$$\frac{\partial u}{\partial t}(x,t) = \frac{u(x,t+\delta t) - u(x,t-\delta t)}{2\delta t} + \mathcal{O}\left((\delta t)^2\right),$$

Konačne razlike (nastavak)

pa možemo definirati centralnu konačnu razliku

$$\frac{\partial u}{\partial t}(x,t) \approx \frac{u(x,t+\delta t) - u(x,t-\delta t)}{2\delta t}.$$

Vidimo da je centralna konačna razlika točnija.



Konačna razlika unazad, unaprijed, i centralna.

Konačne razlike u difuzijskoj jednadžbi

- Kada se primijenjuju na difuzijsku jednadžbu, konačne razlike unaprijed i unazad koje aproksimiraju $\partial u/\partial t$ vode do eksplicitne odnosno implicitne metode konačnih razlika.
- Centralna konačna razlika gornjeg oblika po varijabli t se ne koriste u praksi jer daje nestabilne metode.
- Centralna konačna razlika oblika

$$\frac{\partial u}{\partial t}(x,t) \approx \frac{u(x,t+\delta t/2) - u(x,t-\delta t/2)}{\delta t}$$

pojavljuje se u Crank–Nicolsonovoj shemi za konačne razlike.

Konačne razlike u difuzijskoj jednadžbi (nast.)

Parcijalne derivacije po varijabli x možemo definirati na analogan način:

- konačna razlika unaprijed

$$\frac{\partial u}{\partial x}(x,t) \approx \frac{u(x+\delta x,t) - u(x,t)}{\delta x}$$

- konačna razlika unazad

$$\frac{\partial u}{\partial x}(x,t) \approx \frac{u(x,t) - u(x-\delta x,t)}{\delta x}$$

- centralna konačna razlika

$$\frac{\partial u}{\partial x}(x,t) \approx \frac{u(x+\delta x,t) - u(x-\delta x,t)}{2\delta x}$$

Konačne razlike u difuzijskoj jednadžbi (nast.)

Za drugu parcijalnu derivaciju $\partial^2 u/\partial x^2$ možemo definirati simetričnu centralnu konačnu razliku kao

- konačnu razliku unaprijed od konačnih razlika unazad
- konačnu razliku unazad od konačnih razlika unaprijed

U oba slučaja dobivamo

$$\frac{\partial^2 u}{\partial x^2}(x,t) = \frac{\frac{u(x+\delta x,t)-u(x,t)}{\delta x} - \frac{u(x,t)-u(x-\delta x,t)}{\delta x}}{\delta x} + \mathcal{O}\left((\delta x)^2\right)$$
$$\approx \frac{u(x+\delta x,t) - 2u(x,t) + u(x-\delta x,t)}{(\delta x)^2}.$$

Mreža za difuzijsku jednadžbu

Kako bismo mogli primijeniti metodu konačnih razlika na difuzijsku jednadžbu moramo podijeliti

- x os na ekvidistantne čvorove sa razmakom od δx
- t os na ekvidistantne čvorove sa razmakom od δt

pri čemu uzimamo

$$\delta x = \frac{b-a}{n}, \qquad \delta t = \frac{T}{m}.$$

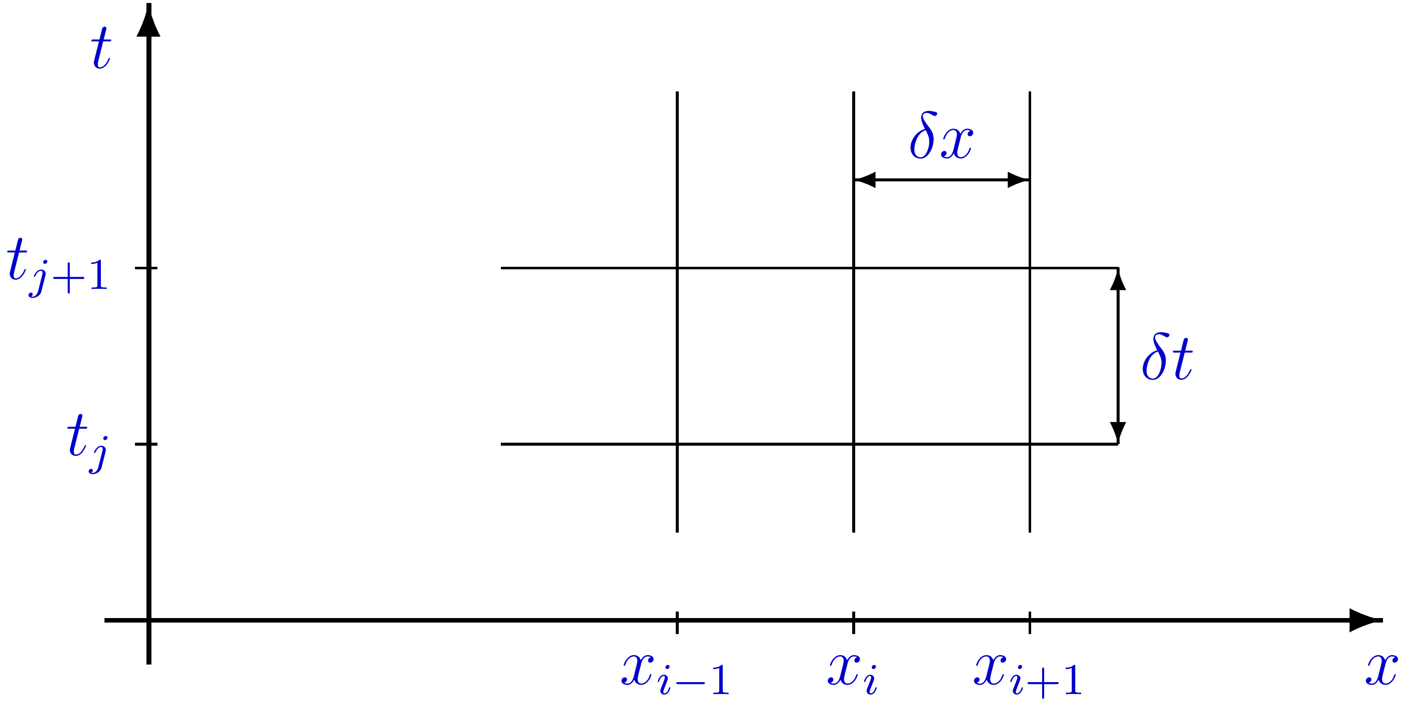
Ovime na (x,t) ravnini, unutar domene $[a,b] \times [0,T]$, definiramo mrežu, pri čemu čvorovi mreže imaju oblik

$$(x_i, t_j) = (a + i\delta x, j\delta t), \qquad i = 0, \ldots n, \; j = 0, \ldots, m.$$

U tom slučaju računat ćemo aproksimativno rješenje samo u čvorovima mreže, i pišemo

$$u_{i,j} \approx u(a + i\delta x, j\delta t).$$

Mreža za difuzijsku jednadžbu (nastavak)



Oznake na mreži.

Eksplicitna metoda konačnih razlika

Razmatramo difuzijsku jednadžbu

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},$$

sa rubnim i inicijalnim uvjetima

$$\begin{aligned} u(a,t) &= y_a(t), \\ u(b,t) &= u_b(t), \\ u(x,0) &= u_0(x). \end{aligned}$$

Želimo naći aproksimaciju rješenja u čvorovima mreže koristeći

- konačnu razliku unaprijed za $\partial u / \partial t$,
- simetričnu centralnu konačnu razliku za $\partial^2 u / \partial x^2$.

Eksplicitna metoda konačnih razlika (nastavak)

Pri tome se difuzijska jednadžba transformira u

$$\frac{u_{i,j+1} - u_{i,j}}{\delta t} + \mathcal{O}(\delta t) = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\delta x)^2} + \mathcal{O}\left((\delta x)^2\right).$$

Zanemarujući izraze $\mathcal{O}(\delta t)$ i $\mathcal{O}\left((\delta x)^2\right)$ dobivamo diferencijsku jednadžbu

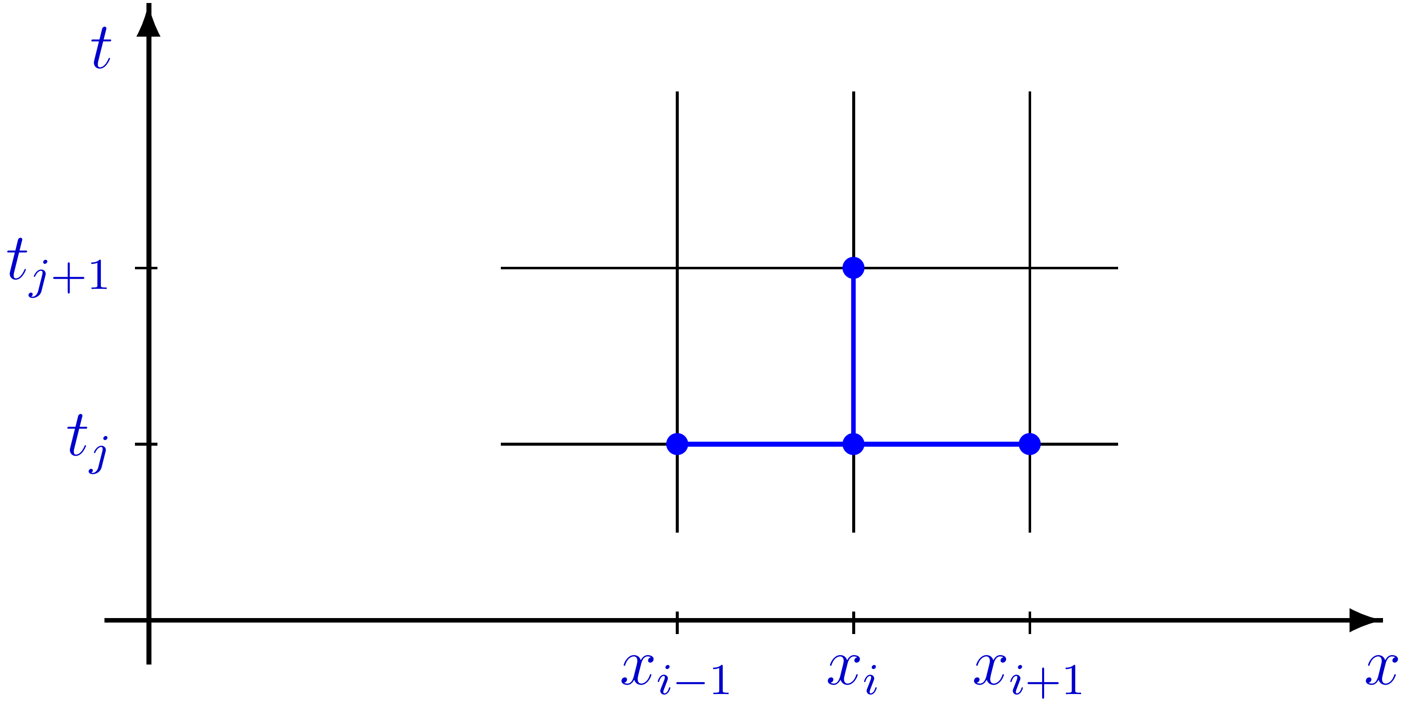
$$u_{i,j+1} = \lambda u_{i-1,j} + (1 - 2\lambda) u_{i,j} + \lambda u_{i+1,j},$$

gdje je

$$\lambda = \frac{\delta t}{(\delta x)^2}, \qquad \text{Courantov broj.}$$

Ako u vremenskom koraku j znamo $u_{i,j}$ za sve vrijednosti od i, tada $u_{i,j+1}$ možemo izračunati eksplicitno.

Eksplicitna metoda konačnih razlika (nastavak)



$u_{i,j+1}$ ovisi samo o $u_{i-1,j}$, $u_{i,j}$ i $u_{i+1,j}$.

Eksplicitna metoda konačnih razlika (nastavak)

Sada možemo riješiti diferencijsku jednadžbu za

$$0 < i < n, \qquad 0 < j \leq m.$$

Rubne uvjete koristimo za određivanje $u_{0,j}$ i $u_{n,j}$:

$$\begin{aligned} u_{0,j} &= u_a(j\delta t), & 0 < j \leq m, \\ u_{n,j} &= u_b(j\delta t), & 0 < j \leq m. \end{aligned}$$

Za pokretanje ove iterativne metode koristimo inicijalni uvjet

$$u_{i,0} = u_0(a + i\delta x), \qquad 0 \leq i \leq n.$$

Iterativna metoda završava za $j = m$ i rješenjem

$$u_{i,m}, \quad 0 \leq i \leq n,$$

što predstavlja aproksimaciju rješenja za $u(a + i\delta x, T)$, $0 \leq i \leq n$.

Algoritam eksplicitne metode konačnih razlika

$\lambda = \frac{\delta t}{(\delta x)^2}$;

for $i = 0 : n$

  $u_{stari}(i) = u_0(a + i \cdot \delta x)$;

end

for $j = 1 : m$

  $u_{novi}(0) = u_a(j \cdot \delta t)$;

  $u_{novi}(n) = u_b(j \cdot \delta t)$;

  for $i = 1 : (n-1)$

    $u_{novi}(i) = \lambda \cdot u_{stari}(i-1) + (1 - 2 \cdot \lambda) \cdot u_{stari}(i) +$
    $+ \lambda \cdot u_{stari}(i+1)$;

  end

  $u_{stari} = u_{novi}$;

end

$u = u_{stari}$;

Stabilnost eksplicitne metode konačnih razlika

- Stabilnost numeričke metode je u bliskoj vezi sa numeričkom greškom.
- Metoda konačnih razlika je stabilna ako greška učinjena u jednom koraku metode ne utječe na povečanje greške u koracima koji slijede.
- Kod neutralno stabilne metode greška ostaje konstantna u svim koracima.
- Ako greške opadaju i po mogućnosti se prigušuju, kažemo da je numerička metoda stabilna.
- Ako, s druge strane, greška raste sa povečanjem broja koraka, aproksimativno rješenje divergira, i kažemo da je numerička metoda nestabilna.

Stabilnost eksplicitne metode konačnih razlika

Za daljnju analizu korisno je sve vrijednosti $u_{i,j}$ za fiksni vremenski korak j organizirati u vektor

$$u^{(j)} = [\ u_{1,j} \quad \cdots \quad u_{n-1,j}\ ]^T.$$

Za matrični oblik diferencijske jednadžbe definiramo $(n-1) \times (n-1)$ matricu A

$$A = \begin{bmatrix} 1-2\lambda & \lambda & 0 \cdots & 0 & \\ \lambda & 1-2\lambda & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \lambda \\ 0 & \cdots & 0 & \lambda & 1-2\lambda \end{bmatrix}.$$

Stabilnost eksplicitne metode konačnih razlika

Tada eksplicitnu metodu možemo zapisati kao

$$u^{(j+1)} = Au^{(j)} + u_r^{(j)},$$

gdje je

$$u_r^{(j)} = \lambda[\ u_{0,j} \quad 0 \quad \cdots \quad 0 \quad u_{n,j}\ ]^T.$$

Pretpostavimo sada da našu numeričku metodu izvršavamo na računalu u aritmetici konačne preciznosti. U tom slučaju ćemo u svakom koraku j umjesto $u^{(j+1)}$ izračunati aproksimativnu vrijednost $\tilde{u}^{(j+1)}$ koja sadrži i greške zaokruživanja. Neka je

$$\tilde{u}^{(j+1)} = A\tilde{u}^{(j)} + \tilde{u}_r^{(j)} + f^{(j+1)},$$

gdje $f^{(j+1)}$ sadrži greške zaokruživanja koje su se dogodile kod računanja $A\tilde{u}^{(j)} + \tilde{u}_r^{(j)}$.

Stabilnost eksplicitne metode konačnih razlika

Definirajmo ukupnu grešku

$$e^{(j)} = \tilde{u}^{(j)} - u^{(j)}, \quad j \geq 0.$$

Možemo zaključiti da vrijedi sljedeće:

$$e^{(j+1)} = Ae^{(j)} + (\tilde{u}_r^{(j)} - u_r^{(j)}) + f^{(j+1)}.$$

Usredotočimo se sada na utjecaj grešaka zaokruživanja $e^{(0)}$ kod računanja inicijalnog uvjeta (za $j = 0$). Zbog jednostavnosti, pretpostavimo da su svi daljnji koraci ($j > 0$) izračunati egzaktno, tj. da je

$$\tilde{u}_r^{(j)} = u_r^{(j)}, \quad f^{(j)} = 0, \quad j > 0.$$

Tada imamo

$$e^{(j+1)} = Ae^{(j)} = A^{j+1}e^{(0)}.$$

Stabilnost exsplicitne metode konačnih razlika

Da bi metoda bila stabilna, greška $e^{(0)}$ mora biti prigušena, a za to treba biti

$$\lim_{j\to\infty} A^j e^{(0)} = 0,$$

što nam je poznata situacija iz iterativnih metoda za sustave linearnih jednadžbi. Od tamo znamo da će metoda biti stabilna ako i samo ako je

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

gdje je $\rho(A)$ spektralni radijus matrice A. Dakle, da bi metoda bila stabilna zahtijevamo da za sve svojstvene vrijednosti $\mu_1(A), \ldots, \mu_{n-1}(A)$ od A vrijedi

$$|\mu_k(A)| < 1, \quad k = 1, \ldots, n-1.$$

Stabilnost eksplicitne metode konačnih razlika

Slijedeći korak je računanje svojstvenih vrijednosti od A. U tu svrhu, matricu pišemo kao

$$A = I - \lambda \cdot \underbrace{\begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix}}_{=G}.$$

Preostaje nam sada samo naći svojstvene vrijednosti $\mu_k(G)$ od matrice G, jer su tada

$$\mu_k(A) = 1 - \lambda \cdot \mu_k(G), \quad k = 1, \ldots, n-1.$$

Svojstvene vrijednosti TST matrice

G je matrica sa posebnom strukturom i njen oblik je poznat pod imenom TST, gdje “TST” dolazi od Toeplitzova (konstantna duž svih dijagonala), simetrična i tridijagonalna.

Teorem. Neka je G $m \times m$ TST matrica sa dijagonalnim elementima α i vandijagonalnim elementima β. Tada su svojstvene vrijednosti od G dane sa

$$\mu_k(G) = \alpha + 2\beta \cos\left(\frac{k\pi}{m+1}\right), \quad k = 1, \ldots, m,$$

a odgovarajući ortonormirani svojstveni vektori dani su, po komponentama, sa

$$q_\ell^{(k)} = \sqrt{\frac{2}{m+1}} \sin\left(\frac{\ell k\pi}{m+1}\right), \quad \ell, k = 1, \ldots, m.$$

Svojstvene vrijednosti TST matrice (nastavak)

Dokaz. Pretpostavimo da je $\beta \neq 0$, jer bi inače G bio multipl od identitete, čime je tvrdnja teorema trivijalna.

Pretpostavimo da je μ svojstvena vrijednost od G sa odgovarajućim svojstvenim vektorom q. Ako definiramo $q_0 = q_{m+1} = 0$, jednakost $Aq = \mu q$ možemo napisati u obliku

$$\beta q_{\ell-1} + (\alpha - \mu) q_\ell + \beta q_{\ell+1} = 0, \quad \ell = 1, \ldots, m.$$

To je linearna diferencijska jednadžba, za koju ponovo razmatramo karakteristični polinom

$$\chi(z) = \beta + (\alpha - \mu) z + \beta z^2.$$

Ako korijene ovog polinoma označimo sa z_1 i z_2, tada općenito rješenje diferencijske jednadžbe ima oblik

$$q_\ell = c_1 z_1^\ell + c_2 z_2^\ell, \quad \text{za konstante } c_1, c_2,$$

gdje su konstante određene iz rubnih uvijeta $q_0 = q_{m+1} = 0$.

Svojstvene vrijednosti TST matrice (nastavak)

Korijeni od $\chi(z)$ su

$$z_{1,2} = \frac{\mu - \alpha \pm \sqrt{(\mu - \alpha)^2 - 4\beta^2}}{2\beta}, \qquad (*)$$

a iz uvjeta $q_0 = 0$ slijedi da je $c_1 + c_2 = 0$. S druge strane, iz uvjeta $q_{m+1} = 0$ slijedi:

$$c_1 z_1^{m+1} + c_2 z_2^{m+1} = 0,$$

odnosno, zbog prethodne napomene je $c_2 = -c_1$, a kako nas interesira netrivijalno rješenje slijedi $c_1 \neq 0$, imamo

$$z_1^{m+1} = z_2^{m+1}.$$

Postoji $m + 1$ rješenja ove jednadžbe, i to

$$z_2 = z_1 \exp\left(\frac{2\pi k\iota}{m+1}\right), \quad k = 0, 1, \ldots, m, \ \ \iota = \sqrt{-1}.$$

Svojstvene vrijednosti TST matrice (nastavak)

Međutim, slučaj za $k = 0$ se može odbaciti, jer je tada $z_2 = z_1$, a odatle je

$$q_\ell = c_1 z_1^\ell + c_2 z_2^\ell = c_1 z_1^\ell - c_1 z_1^\ell = 0.$$

Množeći prethodnu jednakost koja povezuje z_1 i z_2 sa $\exp(-\pi k\iota/(m+1))$ i uvrštavajući vrijednosti za z_1 i z_2 iz jednadžbe za korijene $(*)$ dobivamo

$$\left(\mu - \alpha + \sqrt{(\mu-\alpha)^2 - 4\beta^2}\right) \exp\left(\frac{-\pi k\iota}{m+1}\right) =$$

$$= \left(\mu - \alpha - \sqrt{(\mu-\alpha)^2 - 4\beta^2}\right) \exp\left(\frac{\pi k\iota}{m+1}\right).$$

Nakon potrebnih skračivanja, i dijeljenja jednakosti sa 2 imamo

Svojstvene vrijednosti TST matrice (nastavak)

$$\sqrt{(\mu-\alpha)^2-4\beta^2}\cos\left(\frac{k\pi}{m+1}\right)=(\mu-\alpha)\iota\sin\left(\frac{k\pi}{m+1}\right),$$

a nakon kvadriranja obaju strana jednakosti i rješavanja kvadratne jednadžbe po μ

$$\mu^2-2\alpha\mu+\alpha^2-4\beta^2\cos^2\left(\frac{\pi k}{m+1}\right)=0,$$

dobivamo rješenja

$$\mu_{1,2}=\alpha\pm 2\beta\cos\left(\frac{\pi k}{m+1}\right),\qquad k=1,\ldots,m.$$

Ako uzmemo rješenje sa znakom plus dobivamo tražene svojstvene vrijednosti, dok rješenje sa znakom minus samo ponavlja te iste vrijednosti, pa se stoga može odbaciti.

Svojstvene vrijednosti TST matrice (nastavak)

Uvrštavajući dobiveni izraz za svojstvene vrijednosti u jednadžbu za korijene $(*)$ dobivamo

$$z_{1,2} = \cos\left(\frac{k\pi}{m+1}\right) \pm \iota \sin\left(\frac{k\pi}{m+1}\right),$$

pa je zbog toga

$$q_\ell^{(k)} = c_1(z_1^\ell - z_2^\ell) = 2c_1 \iota \sin\left(\frac{\pi k \ell}{m+1}\right), \quad k, \ell = 1, \ldots, m.$$

Ako uzmemo da je $c_1 = -(\iota/2)\sqrt{2/(m+1)}$, onda se lako provjeri da svaki od vektora $q^{(k)}$ ima normu jedan, jer je

Svojstvene vrijednosti TST matrice (nastavak)

$$\begin{aligned}
\sum_{\ell=1}^{m} \left(q_\ell^{(k)}\right)^2 &= \frac{2}{m+1} \sum_{\ell=1}^{m} \sin^2\left(\frac{\ell k \pi}{m+1}\right) \\
&= \frac{2}{m+1} \cdot \frac{1}{2} \sum_{\ell=1}^{m} \left[1 - \cos\left(\frac{2\ell k \pi}{m+1}\right)\right] \\
&= \frac{2}{m+1} \left[\frac{m}{2} - \frac{1}{2} \sum_{\ell=1}^{m} \cos\left(\frac{2\ell k \pi}{m+1}\right)\right]
\end{aligned}$$

odakle se, prelaskom sa kosinusa na realni dio kompleksnog broja, dobiva da je suma kosinusa u zadnjoj jednakosti jednaka -1, čime smo dobili traženi rezultat.

Svojstveni vektori su ortogonalni, budući da je matrica simetrična.

Stabilnost eksplicitne metode konačnih razlika

Dakle, svojstvene vrijednosti naše matrice G dane su sa

$$\begin{aligned}\mu_k(G) =& 2 - 2\cos\left(\frac{k\pi}{n}\right) \\ =& 4\sin^2\left(\frac{k\pi}{2n}\right), \quad k = 1, \ldots, n-1.\end{aligned}$$

Dalje vrijedi

$$\mu_k(A) = 1 - 4\lambda\sin^2\left(\frac{k\pi}{2n}\right), \; k = 1, \ldots, n-1.$$

Da bi uvjet stabilnosti $|\mu_k(A)| < 1$ bio zadovoljen, mora vrijediti

$$\left|1 - 4\lambda\sin^2\left(\frac{k\pi}{2n}\right)\right| < 1, \; k = 1, \ldots, n-1.$$

Stabilnost eksplicitne metode konačnih razlika

Budući da je $\lambda > 0$ po definiciji slijedi da je uvijek $1 - 4\lambda \sin^2\left(\frac{k\pi}{2n}\right) < 1$. S druge strane, uvjet

$$-1 < 1 - 4\lambda \sin^2\left(\frac{k\pi}{2n}\right)$$

se može pojednostavniti na

$$\lambda \sin^2\left(\frac{k\pi}{2n}\right) < \frac{1}{2}.$$

Dakle, gornji uvjet stabilnosti ekvivalentan je dvijema jednadžbama

$$\lambda > 0$$

$$\lambda \sin^2\left(\frac{k\pi}{2n}\right) < \frac{1}{2}.$$

Stabilnost exsplicitne metode konačnih razlika

Najveći izraz sa sinusom je

$$\sin\left(\frac{(n-1)\pi}{2n}\right) < 1,$$

a ako povećavamo dimenziju matrice A $d = n - 1$ tada vrijedi

$$\lim_{d\to\infty} \sin\left(\frac{d\pi}{2(d+1)}\right) = 1.$$

Ovime smo dobili konačan uvjet.

Teorem. Za

$$0 < \lambda \leq \frac{1}{2}$$

eksplicitna metoda konačnih razlika za difuzijsku jednadžbu je stabilna. ∎

Stabilnost eksplicitne metode konačnih razlika

- Ovaj kriterij stabilnosti daje uvjet na veličinu koraka:

$$0 < \delta t \leq \frac{(\delta x)^2}{2}.$$

- Kao posljedica uvjeta stabilnosti, parametri m i n ne mogu biti izabrani nezavisno jedan od drugoga.
- Ako moramo izračunati rješenje sa velikom točnošću, tada δx mora biti malen, što daje kvadratnu ogradu za δt koji mora biti još manji.
- Zato nam je praktičnije naći numeričku metodu koja je bezuvjetno stabilna.

Konvergencija eksplicitne metode kon. razlika

Da bi numerička metoda bila uopće korisna u primjenama, mora biti konvergentna:

- aproksimacije moraju težiti točnom rješenju kada δx i δt teže ka nuli

Prvo nas zanima lokalna pogreška diskretizacije, to je ostatak koji dobijemo kada u relaciju koja definira metodu uvrstimo točno rješenje:

$$\epsilon_{i,j+1} = u(x_i, t_{j+1}) - \lambda u(x_{i-1}, t_j) - (1 - 2\lambda) u(x_i, t_j) - \lambda u(x_{i+1}, t_j).$$

Izraze u lokalnoj pogrešci diskretizacije razvijamo u Taylorov red i dobivamo

Konvergencija eksplicitne metode kon. razlika

$$
\begin{aligned}
&\epsilon_{i,j+1} = \\
&= u(x_i,t_j) + \delta t \frac{\partial u}{\partial t}(x_i,t_j) + \frac{1}{2}(\delta t)^2 \frac{\partial^2 u}{\partial t^2}(x_i,t_j) + \mathcal{O}((\delta t)^3) - \\
&\quad - \lambda \left[u(x_i,t_j) - \delta x \frac{\partial u}{\partial x}(x_i,t_j) + \frac{1}{2}(\delta x)^2 \frac{\partial^2 u}{\partial x^2}(x_i,t_j) - \frac{1}{6}(\delta x)^3 \frac{\partial^3 u}{\partial x^3}(x_i,t_j) + \right. \\
&\quad \left. + \frac{1}{24}(\delta x)^4 \frac{\partial^4 u}{\partial x^4}(x_i,t_j) + \mathcal{O}((\delta x)^5) \right] - (1-2\lambda) u(x_i,t_j) - \\
&\quad - \lambda \left[u(x_i,t_j) + \delta x \frac{\partial u}{\partial x}(x_i,t_j) + \frac{1}{2}(\delta x)^2 \frac{\partial^2 u}{\partial x^2}(x_i,t_j) + \frac{1}{6}(\delta x)^3 \frac{\partial^3 u}{\partial x^3}(x_i,t_j) + \right. \\
&\quad \left. + \frac{1}{24}(\delta x)^4 \frac{\partial^4 u}{\partial x^4}(x_i,t_j) + \mathcal{O}((\delta x)^5) \right]
\end{aligned}
$$

Konvergencija eksplicitne metode kon. razlika

Ako uvrstimo da je $\delta t = \lambda(\delta x)^2$, sređivanjem prethodnog izraza dobivamo

$$\begin{aligned}
\epsilon_{i,j+1} &= \delta t \left(\frac{\partial u}{\partial t}(x_i, t_j) - \frac{\partial^2 u}{\partial x^2}(x_i, t_j) \right) + \\
&\quad + \frac{(\delta t)^2}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t_j) - \lambda \frac{(\delta x)^4}{12} \frac{\partial^4 u}{\partial x^4}(x_i, t_j) + \mathcal{O}((\delta t)^3) + \mathcal{O}((\delta x)^5) \\
&= \frac{(\delta t)^2}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t_j) - \lambda \frac{(\delta x)^4}{12} \frac{\partial^4 u}{\partial x^4}(x_i, t_j) + \mathcal{O}((\delta t)^3) + \mathcal{O}((\delta x)^5)
\end{aligned}$$

jer je u rješenje difuzijske jednadžbe. Dakle, vrijedi

$$\begin{aligned}
u(x_i, t_{j+1}) &= \lambda u(x_{i-1}, t_j) + (1 - 2\lambda) u(x_i, t_j) + \lambda u(x_{i+1}, t_j) + \\
&\quad + \mathcal{O}((\delta t)^2) + \mathcal{O}((\delta x)^4) (\approx \mathcal{O}((\delta x)^4)).
\end{aligned}$$

Konvergencija exsplicitne metode kon. razlika

Drugim riječima, ako bismo u vremenskom koraku j metodi dali egzaktne vrijednosti rješenja u svim čvorovima, pogreška u $(j+1)$-vom koraku bi bila reda veličine $\mathcal{O}(\delta t(\delta x)^2)$.

Ako definiramo

$$u_{egz}^{(j)} = [\ u(x_1, t_j) \quad \cdots \quad u(x_{n-1}, t_j)\ ]^T$$

$$\epsilon^{(j+1)} = [\ \epsilon_{1,j+1} \quad \cdots \quad \epsilon_{n-1,j+1}\ ]^T$$

tada izraz sa lokalnom pogreškom diskretizacije možemo napisati i u matričnom obliku kao

$$u_{egz}^{(j+1)} = Au_{egz}^{(j)} + u_r^{(j)} + \epsilon^{(j+1)}, \qquad j = 0, \ldots m-1.$$

Pri tome je za svaki indeks j

$$\|\epsilon^{(j)}\|_\infty \leq K\delta t(\delta x)^2.$$

Konvergencija eksplicitne metode kon. razlika

Teorem. Promatramo eksplicitnu metodu konačnih razlika za difuzijsku jednadžbu u kojoj je Courantov broj λ konstantan kada $\delta t \to 0$ i $\delta x \to 0$, i još je $\lambda \leq 1/2$. Neka je rješenje aproksimirano na vremenskom intervalu $[0, T]$, s korakom $\delta t = \lambda(\delta x)^2$, te neka je $m = \lfloor T/\delta t \rfloor$. Tada metoda konvergira, tj. za $e^{(j)} = u^{(j)} - u_{egz}^{(j)}$ vrijedi

$$\lim_{\delta x \to 0} \max_{j=0,\ldots,m} \|e^{(j)}\|_\infty = 0.$$

Dokaz. Prvo primijetimo da je inicijalna vrijednost rješenja zadana, pa je (u egzaktnoj aritmetici) $e^{(0)} = 0$. Drugo, za $\lambda \leq 1/2$ je

$$\|A\|_\infty = \lambda + 1 - 2\lambda + \lambda = 1.$$

Konvergencija exsplicitne metode kon. razlika

Oduzimanjem matričnih oblika iteracije metode i izraza za lokalnu pogrešku diskretizacije dobivamo

$$e^{(j+1)} = Ae^{(j)} - \epsilon^{(j+1)},$$

pa onda induktivno možemo zaključiti da je za svaki j

$$e^{(j)} = A^j e^{(0)} - \sum_{i=0}^{j-1} A^i \epsilon^{(j-i)} = -\sum_{i=0}^{j-1} A^i \epsilon^{(j-i)}.$$

Sada uzimanjem norme dobivamo

$$\begin{aligned} \|e^{(j)}\|_\infty &\leq \sum_{i=0}^{j-1} \|A\|_\infty^i \|\epsilon^{(j-i)}\|_\infty \leq jK\delta t(\delta x)^2 \\ &\leq mK\delta t(\delta x)^2 \leq TK(\delta x)^2, \end{aligned}$$

Konvergencija eksplicitne metode kon. razlika

jer je $m\delta t \leq T$. ■

Iz dokaza prethodnog teorema jasno je vidljivo da

- Konzistentnost: lokalna pogreška diskretizacije zadovoljava $\|e^{(j)}\|/\delta t \to 0$ kada $\delta x \to 0$.
- Stabilnost: Courantov broj zadovoljava $\lambda \leq 1/2$

osiguravaju konvergenciju metode.

Potpuna implicitna metoda konačnih razlika

- Implicitne metode se koriste kako bi se izbjegla ograničenja vezana uz stabilnost eksplicitne metode.
- Ove metode nam omogućuju da koristimo mreže u x koordinati sa velikim brojem čvorova, bez da moramo uzeti jako mali δt.
- Jedna od implicitnih metoda je i potpuna implicitna metoda konačnih razlika, koja računa aproksimaciju rješenja difuzijske jednadžbe u čvorovima mreže koristeći
  - konačnu razliku unazad za $\partial u/\partial t$,
  - simetričnu centralnu konačnu razliku za $\partial^2 u/\partial x^2$.

Potpuna implicitna metoda konačnih razlika (n.)

Pri tome se difuzijska jednadžba transformira u

$$\frac{u_{i,j+1} - u_{i,j}}{\delta t} + \mathcal{O}(\delta t) = \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{(\delta x)^2} + \mathcal{O}\left((\delta x)^2\right).$$

Zanemarujući izraze $\mathcal{O}(\delta t)$ i $\mathcal{O}\left((\delta x)^2\right)$ dobivamo diferencijsku jednadžbu

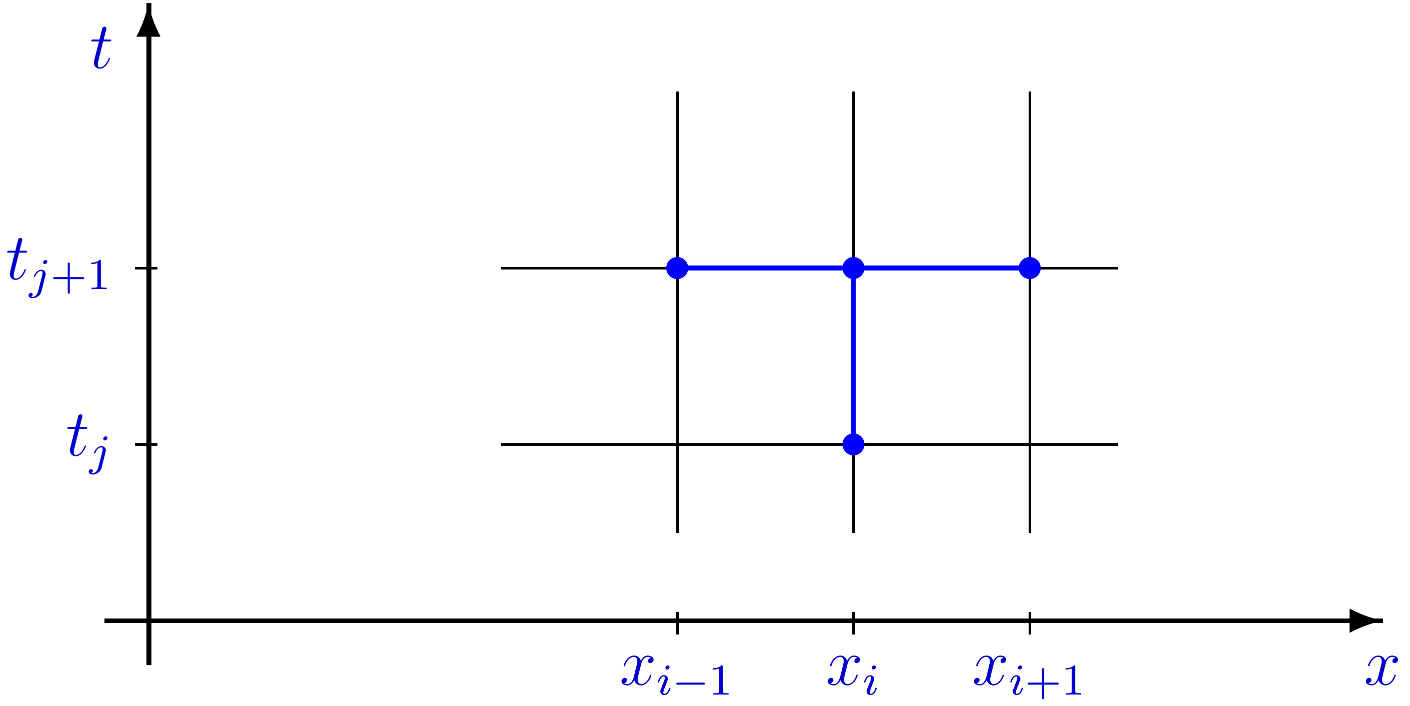
$$-\lambda u_{i-1,j+1} + (1 + 2\lambda)u_{i,j+1} - \lambda u_{i+1,j+1} = u_{i,j},$$

gdje je opet

$$\lambda = \frac{\delta\tau}{(\delta x)^2}.$$

U potpunoj implicitnoj metodi $u_{i-1,j+1}$, $u_{i,j+1}$ i $u_{i+1,j+1}$ implicitno ovise o $u_{i,j}$.

Potpuna implicitna metoda konačnih razlika (n.)



$u_{i-1,j+1}$, $u_{i,j+1}$ i $u_{i+1,j+1}$ ovise o $u_{i,j}$.

Potpuna implicitna metoda konačnih razlika (n.)

- nove vrijednosti se ne mogu razdvojiti i eksplicitno izračunati iz starih vrijednosti.
- Radi se o simultanom rješavanju jednadžbi, odnosno o rješavanju sustava linearnih jednadžbi.

Sada možemo riješiti diferencijsku jednadžbu za

$$0 < i < n, \qquad 0 < j \leq m.$$

Rubne uvjete koristimo za određivanje $u_{0,j}$ i $u_{n,j}$:

$$\begin{aligned} u_{0,j} &= u_a(j\delta t), & 0 < j \leq m, \\ u_{n,j} &= u_b(j\delta t), & 0 < j \leq m. \end{aligned}$$

Za pokretanje ove iterativne metode koristimo inicijalni uvjet

$$u_{i,0} = u_0(a + i\delta x), \qquad 0 \leq i \leq n.$$

Potpuna implicitna metoda konačnih razlika (n.)

Iterativna metoda završava za $j = m$ i rješenjem

$$u_{i,m}, \quad 0 \leq i \leq n,$$

što predstavlja aproksimaciju rješenja za $u(a + i\delta x, T)$, $0 \leq i \leq n$.

S obzirom da moramo rješavati sustave, sada nam i u fazi računanja treba matrični oblik diferencijske jednadžbe.

Potpuna implicitna metoda konačnih razlika (n.)

Definiramo $(n-1)\times(n-1)$ matricu A

$$A = \begin{bmatrix} 1+2\lambda & -\lambda & 0\cdots & 0 & \\ -\lambda & 1+2\lambda & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\lambda \\ 0 & \cdots & 0 & -\lambda & 1+2\lambda \end{bmatrix},$$

i vektor desne strane sustava

$$b = u^{(j)} + u_r^{(j+1)},$$

gdje su

$$\begin{aligned} u^{(j)} &= \begin{bmatrix} u_{1,j} & \cdots & u_{n-1,j} \end{bmatrix}^T, \\ u_r^{(j+1)} &= \lambda \begin{bmatrix} u_{0,j+1} & 0 & \cdots & 0 & u_{n,j+1} \end{bmatrix}^T. \end{aligned}$$

Potpuna implicitna metoda konačnih razlika (n.)

Sada potpunu implicitnu metodu možemo napisati u matričnom obliku kao

$$Au^{(j+1)} = u^{(j)} + u_r^{(j+1)} = b^{(j)}.$$

Vektor $u_r^{(j+1)}$ se pojavljuje zbog rubnih uvjeta, npr. iz prve jednadžbe slijedi

$$(1 + 2\lambda)u_{1,j+1} - \lambda u_{2,j+1} = u_{1,j} + \lambda u_{0,j+1}.$$

Pokazat ćemo da je matrica A regularna pa se korak implicitne metode može napisati eksplicitno kao

$$u^{(j+1)} = A^{-1} \left(u^{(j)} + u_r^{(j+1)} \right).$$

Algoritam potpune implicitne metode kon. razl.

definiraj rješavač s matricom A; $\lambda = \frac{\delta t}{(\delta x)^2}$;

for $i = 0 : n$

$\quad u(i) = u_0(a + i \cdot \delta x)$;

end

for $j = 1 : m$

$\quad$ **for** $i = 1 : n - 1$

$\quad\quad b(i) = u(i)$;

$\quad$ **end**

$\quad u(0) = u_a(j \cdot \delta t)$;

$\quad u(n) = u_b(j \cdot \delta t)$;

$\quad b(1) = b(1) + \lambda \cdot u(0)$;

$\quad b(n-1) = b(n-1) + \lambda \cdot u(n)$;

$\quad$ riješi sustav $Au(1 : n-1) = b$;

end

Stabilnost potpune implicitne metode kon. razl.

- Kao i kod eksplicitne metode pretpostavimo da našu numeričku metodu izvršavamo na računalu u aritmetici konačne preciznosti.
- U tom slučaju ćemo u svakom koraku j umjesto $u^{(j+1)}$ izračunati aproksimativnu vrijednost $\tilde{u}^{(j+1)}$ koja sadrži i greške zaokruživanja.
- Neka je

  $$\tilde{u}^{(j+1)} = A^{-1} \left(\tilde{u}^{(j)} + \tilde{u}_r^{(j+1)}\right) + f^{(j+1)},$$

  gdje $f^{(j+1)}$ sadrži greške zaokruživanja koje su se dogodile kod računanja $A^{-1}\left(\tilde{u}^{(j)} + \tilde{u}_r^{(j+1)}\right)$.

Definirajmo ukupnu grešku

$$e^{(j)} = \tilde{u}^{(j)} - u^{(j)}, \quad j \geq 0.$$

Stabilnost potpune implicitne metode kon. razl.

Možemo zaključiti da vrijedi sljedeće:

$$e^{(j+1)} = A^{-1}e^{(j)} + A^{-1}(\tilde{u}_r^{(j+1)} - u_r^{(j+1)}) + f^{(j+1)}.$$

Usredotočimo se opet na utjecaj grešaka zaokruživanja $e^{(0)}$ kod računanja inicijalnog uvjeta (za $j = 0$). Zbog jednostavnosti, pretpostavimo da su svi daljnji koraci ($j > 0$) izračunati egzaktno, tj. da je

$$\tilde{u}_r^{(j)} = u_r^{(j)}, \quad f^{(j)} = 0, \quad j > 0.$$

Tada imamo

$$e^{(j+1)} = A^{-1}e^{(j)} = A^{-(j+1)}e^{(0)}.$$

Da bi metoda bila stabilna, greška $e^{(0)}$ mora biti prigušena, a za to treba biti

$$\lim_{j\to\infty} A^{-j}e^{(0)} = 0.$$

Stabilnost potpune implicitne metode kon. razl.

Znamo da će metoda biti stabilna ako i samo ako je

$$\rho(A^{-1}) < 1,$$

Dakle, da bi metoda bila stabilna zahtijevamo da za sve svojstvene vrijednosti $\mu_1(A^{-1}), \ldots, \mu_{n-1}(A^{-1})$ od A^{-1} vrijedi

$$|\mu_k(A^{-1})| < 1, \quad k = 1, \ldots, n-1.$$

Dalje vrijedi da je $\mu_k(A^{-1}) = (\mu_k(A))^{-1}$, a svojstvene vrijednosti matrice A dobit ćemo iz rastava

Stabilnost potpune implicitne metode kon. razl.

$$A = I + \lambda \cdot \underbrace{\begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix}}_{=G}.$$

Budući da znamo svojstvene vrijednosti $\mu_k(G)$ od matrice G, tada je

$$\mu_k(A^{-1}) = \frac{1}{1 + \lambda \cdot \mu_k(G)}, \quad k = 1, \ldots, n-1.$$

Kako je

Stabilnost potpune implicitne metode kon. razl.

$$\mu_k(G) = 4\sin^2\left(\frac{k\pi}{2n}\right), \ \ k = 1, \ldots, n-1,$$

imamo

$$\mu_k(A) = 1 + 4\lambda\sin^2\left(\frac{k\pi}{2n}\right), \ \ k = 1, \ldots, n-1.$$

Jer je $\lambda > 0$, a $\sin\left(\frac{k\pi}{2n}\right) \neq 0$ za $k = 1, \ldots, n-1$, onda je

$$\mu_k(A) > 1$$

$$0 < \mu_k(A^{-1}) < 1, \quad k = 1, \ldots, n-1.$$

Dakle, za bilo koji $\lambda > 0$ je $0 < \mu_k(A^{-1}) < 1$, što znači da je potpuna implicitna metoda konačnih razlika bezuvjetno stabilna.

Stabilnost potpune implicitne metode kon. razl.

S druge strane, vidimo da su sve svojstvene vrijednosti matrice A pozitivne, što znači da je matrica pozitivno definitna.

- Zbog toga za rješavanje sustava $Au^{(j+1)} = b^{(j)}$ možemo koristiti metode
  - faktorizaciju Choleskog
  - Gauss–Seidelovu i SOR metodu
  - metodu konjugiranih gradijenata

  koje su specijalno prilagođene za tridijagonalnu matricu.
- Kod efikasno implementirane eksplicitne i implicitne metode broj operacija je istog reda veličine, pa rješavanje sustava kod implicitne metode ne predstavlja preveliki dodatni trošak u odnosu na eksplicitnu.

Crank–Nicolsonova metoda

- Crank–Nicolsonova metoda je također implicitna metoda koja nema problema sa stabilnošću, ali ima grešku diskretizacije derivacije $\partial u/\partial t$ reda veličine $\mathcal{O}\left((\delta t)^2\right)$.
- Crank–Nicolsonova metoda računa aproksimaciju rješenja difuzijske jednadžbe u čvorovima mreže tako da uzima srednju vrijednost diferencijskih jednadžbi eksplicitne i potpuno implicitne metode.

Dakle, ako koristimo konaču razliku unaprijed za $\partial u/\partial t$ dobivamo eksplicitnu metodu

$$\frac{u_{i,j+1} - u_{i,j}}{\delta t} + \mathcal{O}(\delta t) = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\delta x)^2} + \mathcal{O}\left((\delta x)^2\right),$$

Crank–Nicolsonova metoda (nastavak)

a ako koristimo konaču razliku unazad za $\partial u/\partial t$ dobivamo potpunu implicitnu metodu

$$\frac{u_{i,j+1} - u_{i,j}}{\delta t} + \mathcal{O}(\delta t) = \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{(\delta x)^2} + \mathcal{O}\left((\delta x)^2\right).$$

Srednja vrijednost tih dviju jednadžbi je

$$\frac{u_{i,j+1} - u_{i,j}}{\delta t} + \mathcal{O}(\delta t) =$$

$$= \frac{1}{2}\left(\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\delta x)^2} + \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{(\delta x)^2}\right) +$$

$$+\mathcal{O}\left((\delta x)^2\right).$$

Crank–Nicolsonova metoda (nastavak)

Ovom metodom zapravo aproksimiramo vrijednost difuzijske jednadžbe u točki $(x_i, t_{j+\frac{1}{2}})$, koja se nalazi na pola puta između (x_i, t_j) i (x_i, t_{j+1}):

$$\frac{\partial u}{\partial t}(x_i, t_{j+\frac{1}{2}}) = \frac{\partial^2}{\partial x^2}(x_i, t_{j+\frac{1}{2}}).$$

U ovom slučaju, prvu derivaciju po varijabli t aproksimiramo centralnom konačnom razlikom

$$\frac{\partial u}{\partial t}(x_i, t_{j+\frac{1}{2}}) \approx \frac{u_{i,j+1} - u_{i,j}}{\delta t},$$

a drugu derivaciju po varijabli x sa

$$\frac{\partial^2}{\partial x^2}(x_i, t_{j+\frac{1}{2}}) \approx \frac{1}{2}\left(\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\delta x)^2} + \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{(\delta x)^2}\right).$$

Crank–Nicolsonova metoda (nastavak)

Provjerimo točno koliku smo grešku napravili u varijabli t. Za egzaktne vrijednosti $u_{i,j} = u(x_i, t_j)$ imamo

$$
\begin{aligned}
u_{i,j+1} &= u_{i,j+\frac{1}{2}} + \frac{\delta t}{2} \frac{\partial u}{\partial t}(x_i, t_{j+\frac{1}{2}}) + \frac{(\delta t)^2}{8} \frac{\partial^2 u}{\partial t^2}(x_i, t_{j+\frac{1}{2}}) + \mathcal{O}\left((\delta t)^3\right) \\
u_{i,j} &= u_{i,j+\frac{1}{2}} - \frac{\delta t}{2} \frac{\partial u}{\partial t}(x_i, t_{j+\frac{1}{2}}) + \frac{(\delta t)^2}{8} \frac{\partial^2 u}{\partial t^2}(x_i, t_{j+\frac{1}{2}}) + \mathcal{O}\left((\delta t)^3\right).
\end{aligned}
$$

To znači da za centralnu konačnu razliku vrijedi

$$
\frac{u_{i,j+1} - u_{i,j}}{\delta t} = \frac{\partial u}{\partial t}(x_i, t_{j+\frac{1}{2}}) + \mathcal{O}\left((\delta t)^2\right).
$$

S druge strane, moramo još provjeriti odnos desne strane u Crank–Nicolsonovoj iteraciji sa $\frac{\partial^2}{\partial x^2}(x_i, t_{j+\frac{1}{2}})$.

Crank–Nicolsonova metoda (nastavak)

Uzimanjem srednje vrijednosti prethodno izraženih konačnih razlika unaprijed i unazad, dobivamo jednakost

$$\begin{aligned}
&\frac{1}{2}\left(\frac{u_{i+1,j}-2u_{i,j}+u_{i-1,j}}{(\delta x)^2}+\frac{u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}}{(\delta x)^2}\right)= \\
&\quad=\frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}(x_i,t_j)+\frac{\partial^2 u}{\partial x^2}(x_i,t_{j+1})\right)+\mathcal{O}\left((\delta x)^2\right) \\
&\quad=\frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}(x_i,t_{j+\frac{1}{2}})-\frac{\delta x}{2}\frac{\partial^3 u}{\partial x^3}(x_i,t_{j+\frac{1}{2}})+\right. \\
&\qquad\left.+\frac{\partial^2 u}{\partial x^2}(x_i,t_{j+\frac{1}{2}})+\frac{\delta x}{2}\frac{\partial^3 u}{\partial x^3}(x_i,t_{j+\frac{1}{2}})\right)+\mathcal{O}\left((\delta x)^2\right) \\
&\quad=\frac{\partial^2 u}{\partial x^2}(x_i,t_{j+\frac{1}{2}})+\mathcal{O}\left((\delta x)^2\right).
\end{aligned}$$

Crank–Nicolsonova metoda (nastavak)

Znači, aproksimirajući difuzijsku jednadžbu u točki $(x_i, t_{j+\frac{1}{2}})$ iteracijama Crank–Nicolsonove metode napravili smo grešku reda veličine $\mathcal{O}\left((\delta x)^2\right) + \mathcal{O}\left((\delta t)^2\right)$.

Srednja vrijednost tih dviju jednadžbi točije sada glasi

$$\frac{u_{i,j+1} - u_{i,j}}{\delta t} + \mathcal{O}\left((\delta t)^2\right) =$$

$$= \frac{1}{2}\left(\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\delta x)^2} + \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{(\delta x)^2}\right) +$$

$$+\mathcal{O}\left((\delta x)^2\right).$$

Crank–Nicolsonova metoda (nastavak)

Zanemarujući izraze $\mathcal{O}\left((\delta t)^2\right)$ i $\mathcal{O}\left((\delta x)^2\right)$ dobivamo diferencijsku jednadžbu

$$-\frac{\lambda}{2}u_{i-1,j+1} + (1+\lambda)u_{i,j+1} - \frac{\lambda}{2}u_{i+1,j+1} =$$

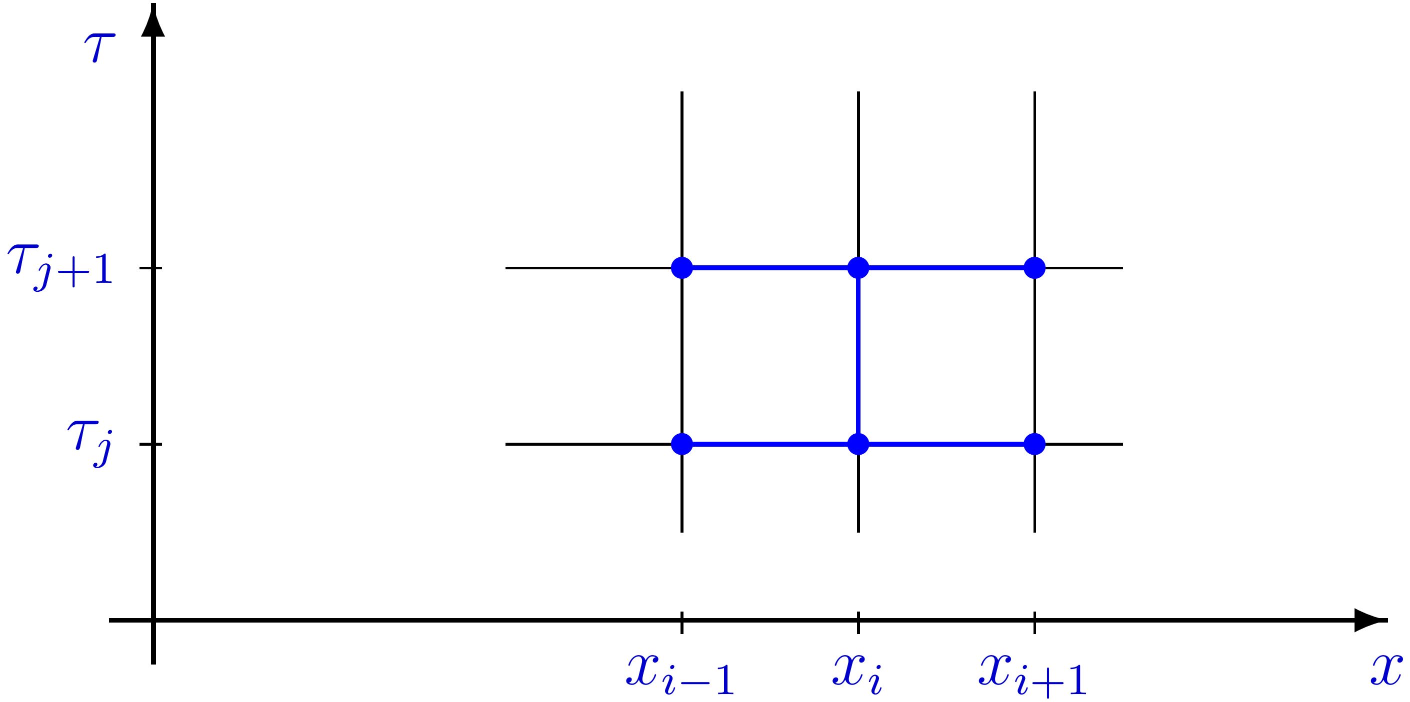
$$= \frac{\lambda}{2}u_{i-1,j} + (1-\lambda)u_{i,j} + \frac{\lambda}{2}u_{i+1,j},$$

gdje je opet

$$\lambda = \frac{\delta t}{(\delta x)^2}.$$

U Crank–Nicolsonovoj metodi $u_{i-1,j+1}$, $u_{i,j+1}$ i $u_{i+1,j+1}$ implicitno ovise o $u_{i-1,j}$, $u_{i,j}$ i $u_{i+1,j}$.

Crank–Nicolsonova metoda (nastavak)



$u_{i-1,j+1}$, $u_{i,j+1}$ i $u_{i+1,j+1}$ ovise o $u_{i-1,j}$, $u_{i,j}$ i $u_{i+1,j}$

Crank–Nicolsonova metoda (nastavak)

Sada možemo riješiti diferencijsku jednadžbu za

$$0 < i < n, \qquad 0 < j \le m.$$

Rubne uvjete koristimo za određivanje $u_{0,j}$ i $u_{n,j}$:

$$\begin{aligned} u_{0,j} &= u_a(j\delta t), & 0 < j \le m, \\ u_{n,j} &= u_b(j\delta t), & 0 < j \le m. \end{aligned}$$

Za pokretanje ove iterativne metode koristimo inicijalni uvjet

$$u_{i,0} = u_0(a + i\delta x), \qquad 0 \le i \le n.$$

Iterativna metoda završava za $j = m$ i rješenjem

$$u_{i,m}, \quad 0 \le i \le n,$$

što predstavlja aproksimaciju rješenja za $u(a + i\delta x, T)$, $0 \le i \le n$.

Crank–Nicolsonova metoda (nastavak)

Također ćemo i ovdje morati rješavati sustave, pa nam opet treba matrični oblik diferencijske jednadžbe. Definiramo $(n-1) \times (n-1)$ matricu A

$$A = \begin{bmatrix} 1+\lambda & -\frac{\lambda}{2} & 0 \cdots & 0 & \\ -\frac{\lambda}{2} & 1+\lambda & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\frac{\lambda}{2} \\ 0 & \cdots & 0 & -\frac{\lambda}{2} & 1+\lambda \end{bmatrix},$$

Crank–Nicolsonova metoda (nastavak)

i $(n-1) \times (n-1)$ matricu B

$$B = \begin{bmatrix} 1-\lambda & \frac{\lambda}{2} & 0 \cdots & 0 & \\ \frac{\lambda}{2} & 1-\lambda & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \frac{\lambda}{2} \\ 0 & \cdots & 0 & \frac{\lambda}{2} & 1-\lambda \end{bmatrix}.$$

Isto tako nam još trebaju vektori

$$u^{(j)} = [\ u_{1,j} \quad \cdots \quad u_{n-1,j}\ ]^T,$$

$$u_r^{(j)} = \lambda[\ u_{0,j} \quad 0 \quad \cdots \quad 0 \quad u_{n,j}\ ]^T.$$

Crank–Nicolsonova metoda (nastavak)

Tada Crank–Nicolsonovu metodu možemo napisati u matričnom obliku kao

$$Au^{(j+1)} = Bu^{(j)} + \frac{1}{2}u_r^{(j)} + \frac{1}{2}u_r^{(j+1)} = b^{(j)}.$$

Pokazat ćemo da je matrica A regularna pa se korak Crank–Nicolsonove metode može napisati eksplicitno kao

$$u^{(j+1)} = A^{-1}\left(Bu^{(j)} + \frac{1}{2}u_r^{(j)} + \frac{1}{2}u_r^{(j+1)}\right).$$

Algoritam Crank–Nicolsonove metode

definiraj rješavač s matricom A; $\lambda = \frac{\delta t}{(\delta x)^2}$;

for $i = 0 : n$

$\quad u(i) = u_0(a + i \cdot \delta x)$;

end

for $j = 1 : m$

$\quad$ **for** $i = 1 : n - 1$

$\qquad b(i) = \frac{\lambda}{2} u(i-1) + (1-\lambda) u(i) + \frac{\lambda}{2} u(i+1)$;

$\quad$ **end**

$\quad u(0) = y_a(j \cdot \delta t)$;

$\quad u(n) = u_b(j \cdot \delta t)$;

$\quad b(1) = b(1) + \frac{\lambda}{2} \cdot u(0)$;

$\quad b(n-1) = b(n-1) + \frac{\lambda}{2} \cdot u(n)$;

$\quad$ riješi sustav $Au(1 : n-1) = b$;

end

Stabilnost Crank–Nicolsonove metode

- Kao i do sada pretpostavimo da našu numeričku metodu izvršavamo na računalu u aritmetici konačne preciznosti.
- U tom slučaju ćemo u svakom koraku j umjesto $u^{(j+1)}$ izračunati aproksimativnu vrijednost $\tilde{u}^{(j+1)}$ koja sadrži i greške zaokruživanja.
- Neka je

$$\tilde{u}^{(j+1)} = A^{-1}\left(B\tilde{u}^{(j)} + \frac{1}{2}\tilde{u}_r^{(j)} + \frac{1}{2}\tilde{u}_r^{(j+1)}\right) + f^{(j+1)},$$

gdje $f^{(j+1)}$ sadrži greške zaokruživanja koje su se dogodile kod računanja $A^{-1}\left(B\tilde{u}^{(j)} + \frac{1}{2}\tilde{u}_r^{(j)} + \frac{1}{2}\tilde{u}_r^{(j+1)}\right)$.

Definirajmo ukupnu grešku

$$e^{(j)} = \tilde{u}^{(j)} - u^{(j)}, \quad j \geq 0.$$

Stabilnost Crank–Nicolsonove metode (nast.)

Možemo zaključiti da vrijedi sljedeće:

$$e^{(j+1)} = A^{-1}Be^{(j)} + \frac{1}{2}A^{-1}(\tilde{u}_r^{(j)} - u_r^{(j)} + \tilde{u}_r^{(j+1)} - u_r^{(j+1)}) + f^{(j+1)}.$$

Usredotočimo se opet na utjecaj grešaka zaokruživanja $e^{(0)}$ kod računanja inicijalnog uvjeta (za $j = 0$). Zbog jednostavnosti, pretpostavimo da su svi daljnji koraci ($j > 0$) izračunati egzaktno, tj. da je

$$\tilde{u}_r^{(j)} = u_r^{(j)}, \quad f^{(j)} = 0, \quad j > 0.$$

Tada imamo

$$e^{(j+1)} = A^{-1}Be^{(j)} = (A^{-1}B)^{j+1}e^{(0)}.$$

Da bi metoda bila stabilna, greška $e^{(0)}$ mora biti prigušena, a za to treba biti

$$\lim_{j\to\infty} (A^{-1}B)^j e^{(0)} = 0.$$

Stabilnost Crank–Nicolsonove metode (nast.)

Znamo da će metoda biti stabilna ako i samo ako je

$$\rho(A^{-1}B) < 1,$$

Dakle, da bi metoda bila stabilna zahtijevamo da za sve svojstvene vrijednosti $\mu_1(A^{-1}B), \ldots, \mu_{n-1}(A^{-1}B)$ od $A^{-1}B$ vrijedi

$$|\mu_k(A^{-1}B)| < 1, \quad k = 1, \ldots, n-1.$$

Tražene svojstvene vrijednosti dobit ćemo iz rastava matrica A i B

Stabilnost Crank–Nicolsonove metode (nast.)

$$A = I + \frac{\lambda}{2} \cdot \underbrace{\begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix}}_{=G}.$$

$$B = I - \frac{\lambda}{2} \cdot \underbrace{\begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix}}_{=G}.$$

Stabilnost Crank–Nicolsonove metode (nast.)

Sada jednakost $Ae^{(j+1)} = Be^{(j)}$ možemo napisati kao

$$\begin{aligned}\left(I + \frac{\lambda}{2}G\right)e^{(j+1)} &= \left(I - \frac{\lambda}{2}G\right)e^{(j)}\\ (2I + \lambda G)\,e^{(j+1)} &= (2I - \lambda G)\,e^{(j)}\\ &=(4I - (2I + \lambda G))e^{(j)}.\end{aligned}$$

Ako definiramo $C = 2I + \lambda G$, tada je

$$Ce^{(j+1)} = (4I - C)e^{(j)},$$

Odnosno

$$e^{(j+1)} = (4C^{-1} - I)e^{(j)}.$$

Stabilnost Crank–Nicolsonove metode (nast.)

Budući da znamo svojstvene vrijednosti $\mu_k(G)$ od matrice G, tada je

$$\mu_k(C) = 2 + \lambda\mu_k(G), \quad k = 1, \ldots, n-1,$$

i

$$\mu_k(4C^{-1} - I) = \frac{4}{2 + \lambda\mu_k(G)} - 1, \quad k = 1, \ldots, n-1.$$

Kako je

$$\mu_k(G) = 4\sin^2\left(\frac{k\pi}{2n}\right), \; k = 1, \ldots, n-1,$$

Stabilnost Crank–Nicolsonove metode (nast.)

zbog toga što je $\lambda > 0$, i $\sin^2\left(\frac{k\pi}{2n}\right) > 0$ za $k = 1, \ldots, n-1$, vrijedi da je

$$\begin{aligned} \mu_k(G) &> 0, \\ \mu_k(C) = 2 + \lambda\mu_k(G) &> 2, \\ 0 < \frac{4}{\mu_k(C)} = \frac{4}{2 + \lambda\mu_k(G)} &< 2, \\ -1 < \mu_k(4C^{-1} - I) &< 1, \quad k = 1, \ldots, n-1. \end{aligned}$$

Dakle, za bilo koji $\lambda > 0$ je $|\mu_k(4C^{-1} - I)| < 1$, što znači da je Crank–Nicolsonova metoda bezuvjetno stabilna.

Stabilnost Crank–Nicolsonove metode (nast.)

- Kao i kod potpuno implicitne metode, može se vidjeti da su sve svojstvene vrijednosti matrice A pozitivne, što znači da je matrica pozitivno definitna.
- Zbog toga za rješavanje sustava $Au^{(j+1)} = b^{(j)}$ možemo koristiti metode
  - faktorizaciju Choleskog
  - Gauss–Seidelovu i SOR metodu
  - metodu konjugiranih gradijenata

  koje su specijalno prilagođene za tridijagonalnu matricu.

Numeričko rješavanje Poissonove jednadžbe

Poissonova jednadžba

Poissonova jednadžba opisuje razne procese u fizici:

- prijenos topline,
- procese u elektrostatici,
- i gravitacijskom polju.

Ona je zapravo stacionarni oblik (koji ne ovisi o vremenu) difuzijske jednadžbe uz uvjet da je funkcija a konstantna, čiji je općeniti oblik

$$-\Delta u(x,t) = f(x,t), \qquad x \in \Omega$$

pri čemu je potrebno još zadati

- rubne uvjete za $x \in \partial\Omega$.

1D Poissonova jednadžba

Promotrimo prvo najjednostavniji slučaj Poissonove jednadžbe u jednoj dimenziji — na nekom segmentu, recimo $[0,1]$. Jednadžba glasi

$$-\frac{d^2u(x)}{dx^2} = f(x), \qquad x \in [0,1],$$

pri čemu je funkcija f zadana, a u nepoznata funkcija.

- Da bi Poissonova jednadžba bila dobro zadana, moramo još zadati rubne uvjete na rubu tog segmenta.
- U ovom slučaju, uzmimo najjednostavnije rubne uvjete, tj. zahtijevajmo da na rubu segmenta za funkciju u vrijedi

$$u(0) = u(1) = 0.$$

Diskretizacija 1D Poissonova jednadžba

Da bismo numerički riješili 1D Poissonovu jednadžbu, zajedno s rubnim uvjetima, moramo je diskretizirati, tj. odabrati niz točaka x_i u kojima želimo naći približno rješenje.

- Neka su točke x_i ekvidistantne, tj. neka je

$$x_i = ih, \quad h = \frac{1}{N+1}, \quad i = 0, \ldots, N+1.$$

  Sada imamo $N+1$ podsegmenata, tako da dobivena matrica (slično kao kod difuzijske jednadžbe) bude dimenzija $N \times N$.

- Također, neka je približno rješenje jednadžbe u točkama x_i označeno s $u_i \approx u(x_i)$, i neka je funkcija s desne strane u tim točkama $f_i = f(x_i)$.

Diskretizacija 1D Poissonova jednadžba (nast.)

Drugu derivaciju aproksimirat ćemo simetričom konačnom razlikom:

$$\frac{d^2u}{dx^2}(x_i) \approx \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2}.$$

Sada uvrstimo aproksimaciju za drugu derivaciju u diferencijalnu jednadžbu, za sve točke x_i. Dobivamo

$$-u_{i-1} + 2u_i - u_{i+1} = h^2 f(x_i), \quad i = 1, \ldots, N.$$

U matričnom obliku, ova jednadžba glasi:

$$G_N u = h^2 f,$$

pri čemu je G_N matrica koja se pojavljivala i kod difuzijske jednadžbe (samo što je sada dimenzije N):

Diskretizacija 1D Poissonova jednadžba (nast.)

$$G_N = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix},$$

$u = [u_1, u_2, \ldots u_N]^T$ je nepoznati vektor rješenja, a
$f = [f_1, f_2, \ldots f_N]^T$ vektor desne strane sustava.

Iz teorema o svojstvenim vrijednostima TLS matrica znamo da su svojstvene vrijednosti matrice G_N jednake

$$\lambda_j = 2\left(1 - \cos\left(\frac{\pi j}{N+1}\right)\right) = 4\sin^2\left(\frac{\pi j}{2(N+1)}\right), \quad j = 1, \ldots, N.$$

Uvjetovanost matrice G_N

Najveća svojstvena vrijednost prethodne matrice približno je jednaka 4, dok je najmanja svojstvena vrijednost λ_1 približno jednaka

$$\lambda_1 \approx 2\left(1 - \left(1 - \frac{\pi^2}{2(N+1)^2}\right)\right) = \left(\frac{\pi}{N+1}\right)^2.$$

U prethodnoj formuli koristili smo aproksimaciju funkcije kosinus s njezina prva dva člana Taylorovog reda.

Sada je odmah jasno da je matrica G_N pozitivno definitna i da je njezina uvjetovanost približno jednaka

$$\kappa(G_N) = \frac{\lambda_N}{\lambda_1} \approx \frac{4(N+1)^2}{\pi^2}.$$

Uvjetovanost matrice G_N (nastavak)

To znači da uvjetovanost brzo raste s porastom broja podintervala (podsjetite se veze uvjetovanosti i rješenja linearnih sustava).

2D Poissonova jednadžba

Sada promatramo slučaj Poissonove jednadžbe (eliptičku parcijalnu diferencijalnu) u dvije dimenzije — na nekom kvadratu, recimo $[0,1] \times [0,1]$. Jednadžba glasi

$$-\frac{\partial^2 u(x,y)}{\partial x^2} - \frac{\partial^2 u(x,y)}{\partial y^2} = f(x,y), \qquad (x,y) \in [0,1] \times [0,1],$$

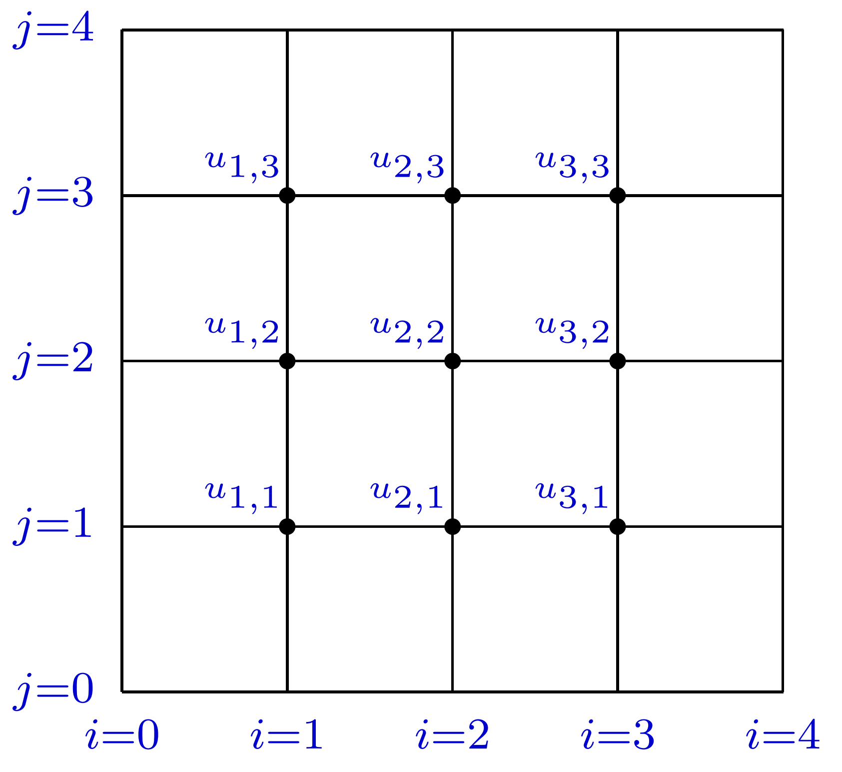
uz rubni uvjet $u = 0$, tj. funkcija u je jednaka 0 na rubu kvadrata.

Kvadrat podijelimo u mrežu čvorova, a da nam bude jednostavnije, pretpostavimo da je i ta mreža kvadratna, tj. korak u x i y smjeru je jednak

$$h = \frac{1}{N+1}.$$

Diskretizacija 2D Poissonove jednadžbe

Uz tako definirane korake, unutarnji čvorovi mreže su točke (x_i, y_j), gdje je $x_i = ih$, $y_j = jh$, za $i, j = 1, \ldots, N$. Dakle, imamo $n := N^2$ unutarnjih čvorova mreže. Takva mreža za $N = 3$ izgleda ovako:



Diskretizacija 2D Poissonove jednadžbe (nast.)

Vrijednost aproksimacije rješenja u čvoru (x_i, y_j) označavamo s $u_{i,j} \approx u(ih, jh)$, a funkcijsku vrijednost s $f_{i,j} = f(ih, jh)$.

Druge parcijalne derivacije aproksimirat ćemo simetričim konačnim razlikama:

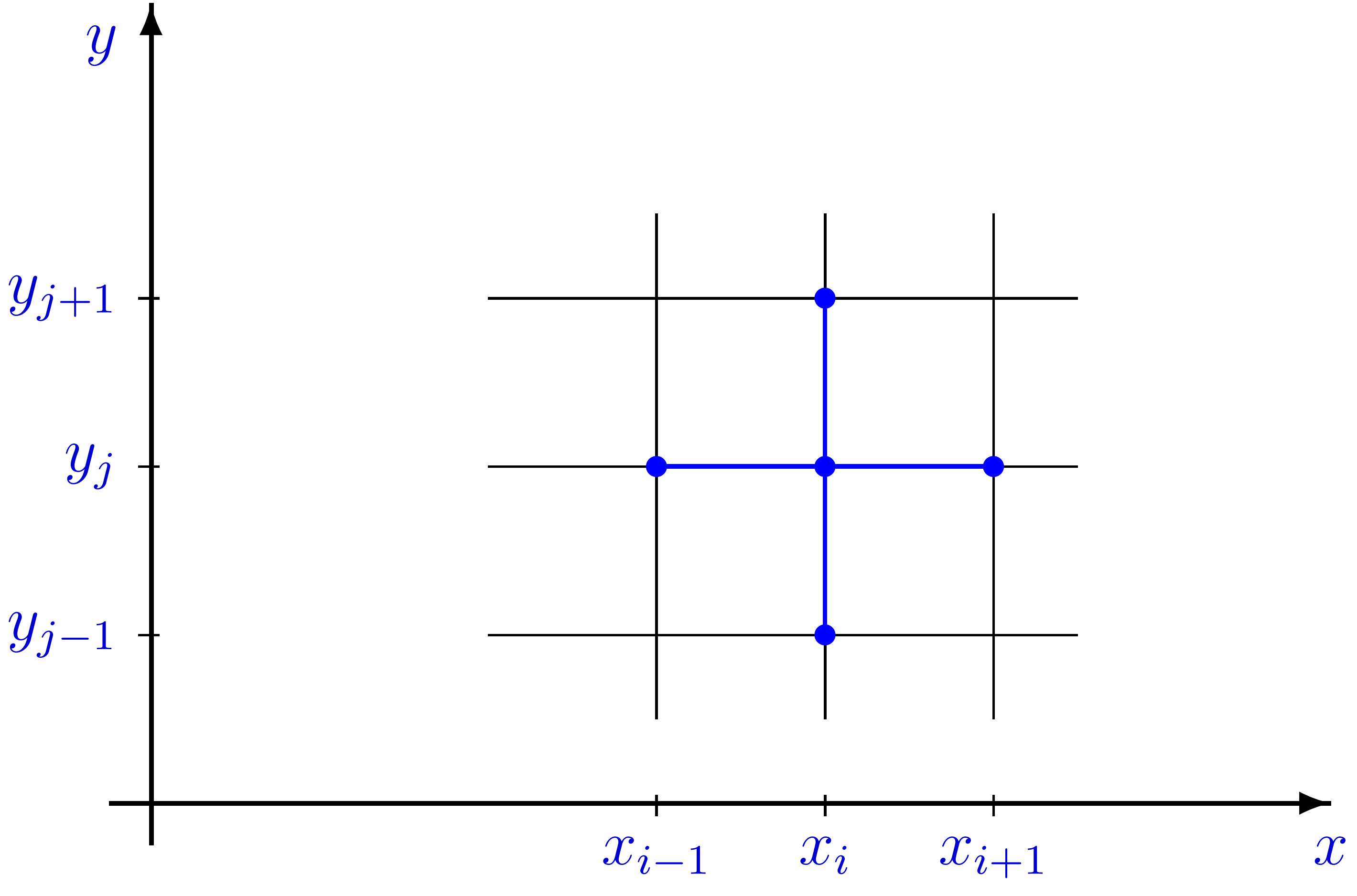
$$\frac{\partial^2 u}{\partial x^2}(x_i, y_j) \approx \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2},$$

$$\frac{\partial^2 u}{\partial y^2}(x_i, y_j) \approx \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2}.$$

Uvrstimo li te aproksimacije derivacija u diferencijalnu jednadžbu, dobivamo

$$4u_{i,j} - u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1} = h^2 f_{i,j}, \quad i, j = 1, \ldots, N.$$

Diskretizacija 2D Poissonove jednadžbe (nast.)



$u_{i,j}$ ovisi o $u_{i-1,j}$, $u_{i+1,j}$, $u_{i,j-1}$ i $u_{i,j+1}$.

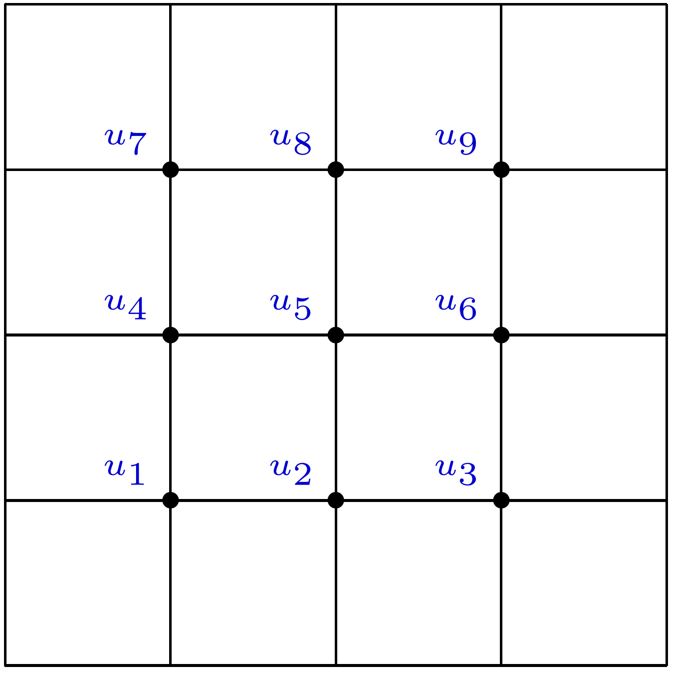
Numeriranje čvorova u kvadratnoj mreži

Pitanje je kako treba napisati ove jednadžbe, tako da se dobije linearni sustav s nekom strukturom. Postoje dva načina da bi se to napravilo.

- Jedan je sekvencijalno numeriranje $u_{i,j}$ po recima ili stupcima (slijeva nadesno, ili zdesna nalijevo, odozgo nadolje ili odozdo nagore),
- a drugi tzv. crveno–crni poredak čvorova, kao kod šahovnice.

Sekvencijalno numeriranje čvorova

Ako $u_{i,j}$ sekvencijalno numeriramo po recima odozdo nagore, na primjer za $N = 3$, dobivamo ovakav poredak čvorova:



Dakle, lako zamjenjujemo $u_{i,j}$ s u_k, gdje $k = (j-1)N + i$.

Aproksimativni lin. sustav za 2D Poisson. jedn.

Ako se na isti način transformiraju i $f_{i,j}$ u f_k, onda dobivamo linearni sustav

$$G_{N\times N} u = h^2 f,$$

gdje je $u = [u_1, u_2, \ldots, u_{N\times N}]^T$, $f = [f_1, f_2, \ldots, f_{N\times N}]^T$, a matrica $G_{N\times N}$ ima N blok-redaka i blok-stupaca, svaki dimenzije N. $G_{N\times N}$ je $N^2 \times N^2$ matrica oblika

$$G_{N\times N} = \begin{bmatrix} G_N + 2I_N & -I_N & & \\ -I_N & \ddots & \ddots & \\ & \ddots & \ddots & -I_N \\ & & -I_N & G_N + 2I_N \end{bmatrix},$$

pri čemu je I_N jedinična matrica reda N, a opet je

Aproksimativni lin. sustav za 2D Poisson. jedn.

$$G_N = \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}.$$

G_N matrica koja nastaje diskretizacijom odgovarajuće 1D Poissonove jednadžbe.

Na primjer, za $N = 3$, matrica linearnog sustava je

Aproksimativni lin. sustav za 2D Poisson. jedn.

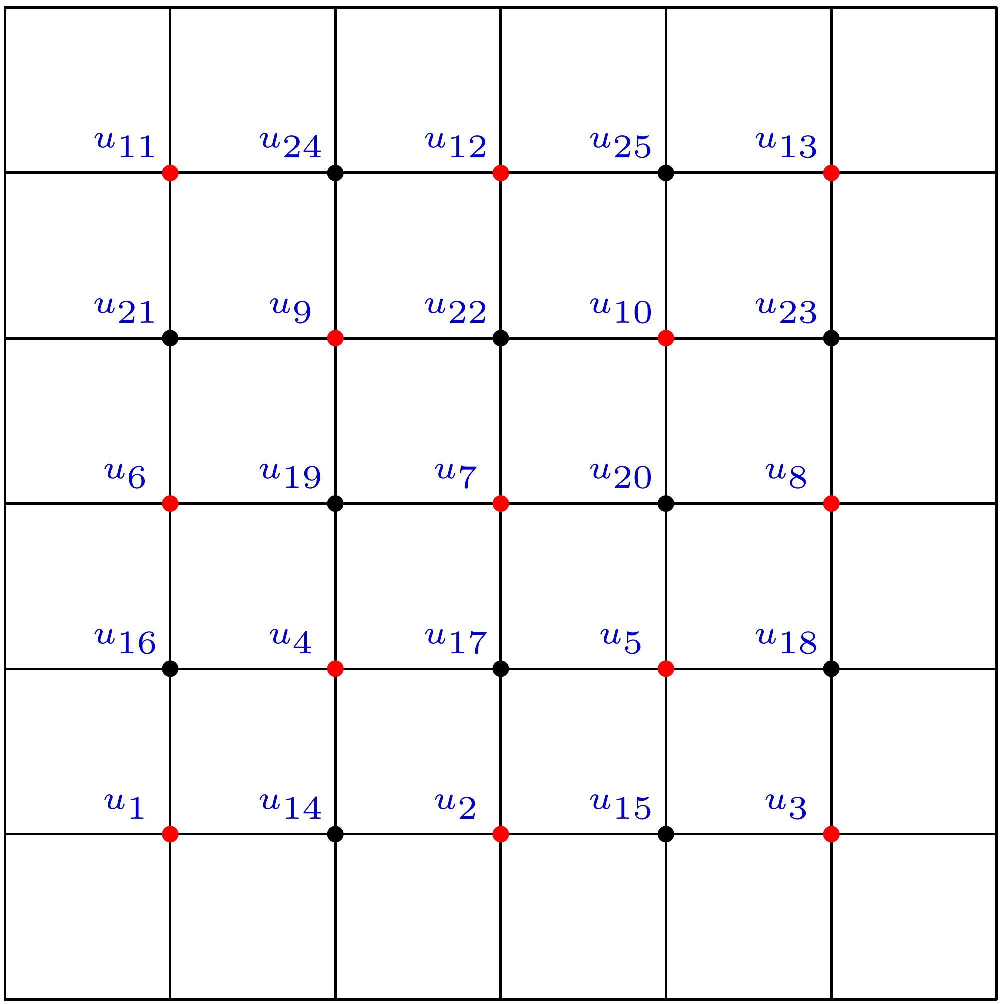
$$G_{3\times 3} = \left[\begin{array}{ccc|ccc|ccc} 4 & -1 & & -1 & & & & & \\ -1 & 4 & -1 & & -1 & & & & \\ & -1 & 4 & & & -1 & & & \\ \hline -1 & & & 4 & -1 & & -1 & & \\ & -1 & & -1 & 4 & -1 & & -1 & \\ & & -1 & & -1 & 4 & & & -1 \\ \hline & & & -1 & & & 4 & -1 & \\ & & & & -1 & & -1 & 4 & -1 \\ & & & & & -1 & & -1 & 4 \end{array}\right]$$

Crveno-crno numeriranje čvorova

Ako čvorove $u_{i,j}$ poredamo u tzv. crveno–crni poredak, dobit ćemo konzistentno poredanu matricu (definicija kasnije).

Crveno–crni poredak dobivamo tako da ih obojamo poput šahovske ploče: svaki crveni čvor (osim rubnog) je okružen s četiri crna susjeda i obratno.
Na primjer, za $N = 5$ takvo crveno–crno bojanje čvorova izgleda ovako:

Crveno-crno numeriranje čvorova (nastavak)



Crveno-crno numeriranje čvorova (nastavak)

Ako zatim sve čvorove koji su crveno obojani popišemo prije crnih (dodijelimo im indekse prije crnih), ili obratno, dobit ćemo blok matricu oblika

$$PG_{N\times N}P^T = \begin{bmatrix} D_1 & G_{12} \\ G_{21} & D_2 \end{bmatrix}.$$

Lako je vidjeti da su dijagonalni blokovi baš dijagonalne matrice, jer ne postoji veza između dva crvena ili dva crna čvora (osim čvora sa samim sobom).

Konkretno, crveno–crni poredak za matricu $G_{3\times 3}$ daje

Crveno-crno numeriranje čvorova (nastavak)

$$PG_{3\times 3}P^T = \left[\begin{array}{ccccc|cccc}
4 & & & & & -1 & -1 & & \\
& 4 & & & & -1 & & -1 & \\
& & 4 & & & -1 & -1 & -1 & -1 \\
& & & 4 & & & -1 & & -1 \\
& & & & 4 & & & -1 & -1 \\
\hline
-1 & -1 & -1 & & & 4 & & & \\
-1 & & -1 & -1 & & & 4 & & \\
& -1 & -1 & & -1 & & & 4 & \\
& & -1 & -1 & -1 & & & & 4
\end{array}\right].$$

Svojstvene vrijednosti matrice $G_{N\times N}$

Teorem. Svojstvene vrijednosti matrice $G_{N\times N}$ su

$$\lambda_{j,k} = 4\left(\sin^2\left(\frac{j\pi}{2(N+1)}\right) + \sin^2\left(\frac{k\pi}{2(N+1)}\right)\right), \quad j,k = 1,\ldots,N,$$

a odgovarajući svojstveni vektori su

$$u_{m,\ell}^{(j,k)} = \frac{2}{N+1}\sin\left(\frac{mj\pi}{N+1}\right)\sin\left(\frac{\ell k\pi}{N+1}\right), \quad j,k,m,\ell = 1,\ldots,N,$$

gdje $u_{m,\ell}^{(j,k)}$ označava komponentu koja odgovara točki mreže (m,ℓ) svojstvenog vektora, pridruženog svojstvenoj vrijednosti $\lambda_{j,k}$.

Svojstvene vrijednosti matrice $G_{N\times N}$ (nastavak)

Dokaz. Neka je λ neka svojstvena vrijednost od $G_{N\times N}$ sa odgovarajućim svojstvenim vektorom u, koji se može particionirati u oblik

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}, \quad u_\ell = \begin{bmatrix} u_{1,\ell} \\ \vdots \\ u_{N,\ell} \end{bmatrix}, \quad \ell = 1, \ldots, N.$$

Jednakost $G_{N\times N} u = \lambda u$ može se tada napisati u obliku

$$T u_{\ell-1} + (S - \lambda I_N) u_\ell + T u_{l+1} = 0, \quad \ell = 1, \ldots, N, \qquad (*)$$

gdje smo postavili da je $T = -I_N$, $S = G_N + 2I_N$, i $u_0 = u_{N+1} = 0$.

Svojstvene vrijednosti matrice $G_{N\times N}$ (nastavak)

Prema teoremu o svojstvenim vrijednostim i svojstvenim vektorima TST matrica, možemo napisati da je $S = Q\Lambda_S Q^T$ i $T = Q\Lambda_T Q^T$, gdje su Λ_S i Λ_T dijagonalne matrice, sa dijagonalnim elementima jednakim

$$\lambda_{S,j} = 4 - 2\cos\left(\frac{j\pi}{N+1}\right), \qquad \lambda_{T,j} = -1.$$

m-ti element j-tog stupca od Q jednak je

$$q_m^{(j)} = \sqrt{\frac{2}{N+1}}\sin\left(\frac{mj\pi}{N+1}\right), \quad m, j = 1, \ldots, N.$$

Pomnožimo blok-rekurziju $(*)$ sa Q^T slijeva kako bismo dobili

$$\Lambda_T v_{\ell-1} + (\Lambda_S - \lambda I)v_\ell + \Lambda_T v_{\ell+1} = 0, \quad v_\ell = Q^T u_\ell, \quad \ell = 1, \ldots, N.$$

Svojstvene vrijednosti matrice $G_{N\times N}$ (nastavak)

Budući da su ovdje sve matrice dijagonalne, jednakosti duž vertikalnih linija u mreži imaju oblik

$$\lambda_{T,j} v_{j,\ell+1} + \lambda_{S,j} v_{j,\ell} + \lambda_{T,j} v_{j,\ell-1} = \lambda v_{j,\ell}, \quad j = 1, \ldots, N.$$

Ako je, za fiksnu vrijednost od j, vektor $[v_{j,1} \; \ldots \; v_{j,N}]^T$ svojstveni vektor TST matrice

$$\begin{bmatrix} \lambda_{S,j} & \lambda_{T,j} & & \\ \lambda_{T,j} & \ddots & \ddots & \\ & \ddots & \ddots & \lambda_{T,j} \\ & & \lambda_{T,j} & \lambda_{S,j} \end{bmatrix},$$

sa odgovarajućom svojstvenom vrijednošću λ, i ako uzmemo da su sve ostale komponenete vektora v jednake 0, tada će prethodna jednakost biti zadovoljena.

Svojstvene vrijednosti matrice $G_{N\times N}$ (nastavak)

Ponovo prema teoremu o TST matricama, svojstvene vrijednosti ove matrice jednake su

$$\begin{aligned}\lambda_{j,k} &= \lambda_{S,j} + 2\lambda_{T,j}\cos\left(\frac{k\pi}{N+1}\right)\\ &= 4 - 2\cos\left(\frac{j\pi}{N+1}\right) - 2\cos\left(\frac{k\pi}{N+1}\right)\\ &= 4\sin^2\left(\frac{j\pi}{2(N+1)}\right) + 4\sin^2\left(\frac{k\pi}{2(N+1)}\right),\end{aligned}$$

za $k = 1, \ldots, N$. Pripadajući svojstveni vektori su

$$v_{j,\ell}^{(j,k)} = \sqrt{\frac{2}{N+1}}\sin\left(\frac{\ell k\pi}{N+1}\right).$$

Svojstvene vrijednosti matrice $G_{N\times N}$ (nastavak)

Budući da je

- ℓ-ti blok od $u^{(j,k)}$ jednak Q puta ℓ-ti blok od v,
- i budući da je samo j-ta komponenta ℓ-tog bloka od v različita od nule,

imamo

$$u_{m,\ell} = q_m^{(j)} v_{j,\ell}^{(j,k)} = \frac{2}{N+1} \sin\left(\frac{mj\pi}{N+1}\right) \sin\left(\frac{\ell k\pi}{N+1}\right).$$

Dobivanjem svojstvenih vrijednosti $\lambda_{j,k}$ i odgovarajućih svojstvenih vektora $u^{(j,k)}$ za svako $j = 1, \ldots, N$, sada imamo sve N^2 svojstvene parove od $G_{N\times N}$. ■

Uvjetovanost matrice $G_{N\times N}$

Korolar. Najmanja i najveća svojstvena vrijednost od $G_{N\times N}$ ponašaju se kao

$$\lambda_{min} \approx 2\pi^2 h^2 + \mathcal{O}(h^4) \qquad \text{i} \qquad \lambda_{max} \approx 8 + \mathcal{O}(h^2)$$

kada $h \to 0$, tako da se uvjetovanost matrice $G_{N\times N}$ ponaša kao

$$\kappa(G_{N\times N}) \approx \frac{4}{\pi^2} h^{-2} + \mathcal{O}(1) = \mathcal{O}(N^2).$$

Dokaz. Najmanja svojstvena vrijednost od $G_{N\times N}$ je ona sa $j = k = 1$, a najveća je ona sa $j = k = N$:

$$\lambda_{min} = 8\sin^2\left(\frac{\pi h}{2}\right), \qquad \lambda_{max} = 8\sin^2\left(\frac{\pi}{2} - \frac{\pi h}{2}\right).$$

Uvjetovanost matrice $G_{N\times N}$ (nastavak)

Razvojem funkcija $\sin(x)$ i $\sin(\pi/2 - x) = \cos(x)$ u Taylorov red dobivamo traženi rezultat, a dijeljenjem λ_{max} sa λ_{min} dobivamo ocjenu za broj uvjetovanosti. ■

Napomena. Kako bi diskretizacija konačnim razlikama bila što točnija poželjno je uzeti što manji h. S druge strane, uvjetovanost aproksimativnog linearnog sustava dramatičnon raste kao $\mathcal{O}(h^{-2})$, što predstavlja veliki izazov za njegovo numeričko rješavanje.

Rješavanje pomoću standardnih iteracija

Sada ćemo pokazati da matrica $G_{N\times N}$ spada u jednu posebnu klasu matrica, za koje postoje vrlo konkretni rezultati o konvergenciji standardnih iteracija.

U tu svrhu iznijet ćemo niz definicija i teorema bez dokaza, jer su dokazi prilično tehnički.

Definicija. Matrica A ima svojstvo (A) ako postoji matrica permutacije P takva da vrijedi

$$PAP^T = \begin{bmatrix} D_1 & A_{12} \\ A_{21} & D_2 \end{bmatrix},$$

gdje su D_1 i D_2 dijagonalne matrice. ■

Matrice s konzistentnim poretkom

Od sada pa nadalje pretpostavljamo da je matrica particionirana kao $A = L + D + R$, gdje je D regularna dijagonalna matrica, L strogo donjetrokutasta, a R strogo gornjetrokutasta.

Definicija. Neka je $\alpha \neq 0$. Definirajmo familiju matrica

$$T_J(\alpha) = -\alpha D^{-1} L - \frac{1}{\alpha} D^{-1} R.$$

Vidimo da je $T_J(1) = T_J$ matrica iteracije u Jacobijevoj metodi. ■

Propozicija. Za matrice A sa svojstvom (A), svojstvene vrijednosti matrica $T_J(\alpha)$ ne ovise o α, s tim da D, L i R dobivamo iz rastava matrice PAP^T. ■

Matrice s konzistentnim poretkom (nastavak)

Definicija. Neka je A proizvoljna matrica takva da je

$$A = L + D + R$$

s tim da je D regularna i

$$T_J(\alpha) = -\alpha D^{-1} L - \frac{1}{\alpha} D^{-1} R.$$

Ako svojstvene vrijednosti matrice $T_J(\alpha)$ ne ovise o α, onda kažemo da A ima konzistentan poredak (engl. consistent ordering). ■

Korolar. Ako A ima konzistentan poredak, tada je

$$\rho(T_{GS}) = (\rho(T_J))^2,$$

što znači da Gauss–Seidelova metoda konvergira dvostruko brže nego Jacobijeva metoda (ako barem jedna od njih kovergira). ■

Optimalni parametar SOR(ω) metode

Kod SOR(ω) metode, one vrijednosti parametra ω za koje je $\rho(T_{SOR(\omega)})$ globalno najmanji, zovemo optimalnim relaksacijskim parametrima i označavamo s ω_{opt}.

Teorem. Pretpostavimo da matrica A ima konzistentan poredak i da matrica T_J u Jacobijevoj metodi ima samo realne svojstvene vrijednosti. Onda SOR(ω) metoda konvergira za bilo koji početni vektor ako i samo ako je $\mu := \rho(T_J) < 1$ (tj. Jacobijeva metoda konvergira) i vrijedi $0 < \omega < 2$. Dodatno, za $\mu < 1$ onda vrijedi i

$$\omega_{\text{opt}} = \frac{2}{1 + \sqrt{1 - \mu^2}},$$

$$\rho(T_{SOR(\omega_{\text{opt}})}) = \omega_{\text{opt}} - 1 = \frac{\mu^2}{(1 + \sqrt{1 - \mu^2})^2} = \frac{1 - \sqrt{1 - \mu^2}}{1 + \sqrt{1 - \mu^2}},$$

Optimalni parametar SOR(ω) metode (nast.)

a za sve $\omega \in \langle 0, 2\rangle$ vrijedi

$$\rho(T_{SOR(\omega)}) = \begin{cases} 1 - \omega + \frac{1}{2}\omega^2\mu^2 + \omega\mu\sqrt{1 - \omega + \frac{1}{4}\omega^2\mu^2}, & \text{za } 0 < \omega \leq \omega_{\text{opt}}, \\ \omega - 1, & \text{za } \omega_{\text{opt}} \leq \omega \leq 2. \end{cases}$$

■

Teorem. Neka je A simetrična (hermitska) i pozitivno definitna matrica i pretpostavimo da A ima konzistentan poredak. Onda matrica iteracije T_J u Jacobijevoj metodi ima samo realne svojstvene vrijednosti i vrijedi $\rho(T_J) < 1$, tj. Jacobijeva metoda konvergira. ■

Konvergencija stand. iter. za sustav s $G_{N\times N}$

Može se pokazati da matrica $G_{N\times N}$, i svaka matrica dobivena bilo kojom numeracijom čvorova, ima konzistentni poredak.

Dakle, da bismo ispitali konvergenciju iterativnih metoda, dovoljno je naći spektralni radijus matrice T_J. Prvo, nađimo rastav (cijepanje) matrice $G_{N\times N}$

$$G_{N\times N} = 4I_{N\times N} - (4I_{N\times N} - G_{N\times N}),$$

pa je $M = 4I_{N\times N}$, $N = (4I_{N\times N} - T_{N\times N})$,

$$T_J = M^{-1}N = (4I_{N\times N})^{-1}\,(4I_{N\times N} - G_{N\times N}) = I_{N\times N} - \frac{1}{4}\,G_{N\times N}.$$

Drugim riječima, T_J je polinom od $G_{N\times N}$, pa ako je $\lambda_{i,j}$ svojstvena vrijednost od $G_{N\times N}$, onda je $1 - \lambda_{i,j}/4$ svojstvena vrijednost od T_J.

Konvergencija stand. iter. za sustav s $G_{N\times N}$

Znamo da svojstvene vrijednosti matrice $G_{N\times N}$ na alternativni način možemo napisati kao

$$\lambda_{i,j} = 4 - 2\left(\cos\frac{\pi i}{N+1} + \cos\frac{\pi j}{N+1}\right),$$

odakle slijedi da je

$$\rho(T_J) = \max_{i,j}\left|1 - \frac{\lambda_{i,j}}{4}\right| = \left|1 - \frac{\lambda_{1,1}}{4}\right| = \left|1 - \frac{\lambda_{N,N}}{4}\right| = \cos\frac{\pi}{N+1}.$$

Odmah je vidljivo da porastom N argument kosinusa ide prema nuli, pa će $\rho(T_J)$ biti sve bliže 1, a iterativne će metode sve sporije konvergirati.

Konvergencija stand. iter. za sustav s $G_{N\times N}$

Čak štoviše, možemo procijeniti $\rho(T_J)$ za velike N. Dovoljno dobra aproksimacija bit će prva dva člana u Taylorovom redu za funkciju kosinus

$$\rho(T_J) = \cos\frac{\pi}{N+1} \approx 1 - \frac{\pi^2}{2(N+1)^2}.$$

Spektralni radijus za Gauss–Seidelovu metodu lako je dobiti koristeći

$$\rho(T_{GS}) = \big(\rho(T_J)\big)^2 = \cos^2\frac{\pi}{N+1}.$$

Približno, kvadriranjem prva dva člana u Taylorovom redu, vrijedi

$$\rho(T_{GS}) \approx 1 - \frac{\pi^2}{(N+1)^2}.$$

Konvergencija stand. iter. za sustav s $G_{N\times N}$

Konačno, dobivamo

$$\omega_{opt} = \frac{2}{1+\sin\frac{\pi}{N+1}}, \qquad \rho(T_{SOR(\omega_{opt})}) = \frac{\cos^2\frac{\pi}{N+1}}{\left(1+\sin\frac{\pi}{N+1}\right)^2}.$$

Veličinu $\rho(T_{SOR(\omega_{opt})})$ možemo približno ocijeniti za velike N. Vrijedi

$$\frac{\cos^2\frac{\pi}{N+1}}{\left(1+\sin\frac{\pi}{N+1}\right)^2} = \frac{1-\sin\frac{\pi}{N+1}}{1+\sin\frac{\pi}{N+1}} = 1 - \frac{2\sin\frac{\pi}{N+1}}{1+\sin\frac{\pi}{N+1}}$$

$$\approx 1 - 2\sin\frac{\pi}{N+1} \approx 1 - \frac{2\pi}{N+1}.$$

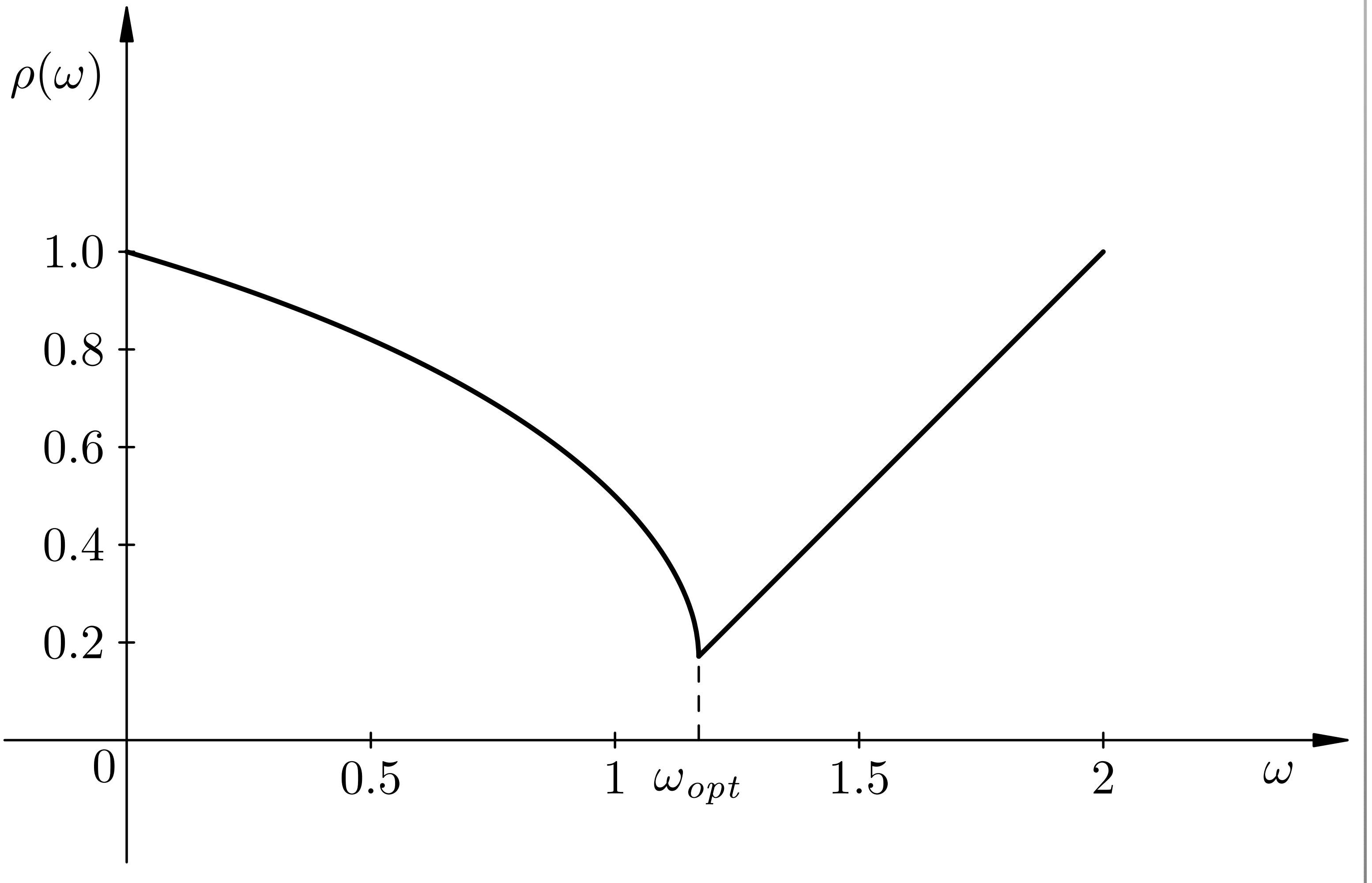
Konvergencija stand. iter. za sustav s $G_{N\times N}$

Primijetite da spektralni radijus i kod Jacobijeve metode i kod Gauss–Seidelove metode ima oblik $1 - O(1/N^2)$, dok kod SOR metode s optimalnim parametrom spektralni radijus ima oblik $1 - O(1/N)$, što pokazuje da bi SOR za optimalni izbor parametra morao biti reda veličine N puta brži i od Jacobijeve i od Gauss–Seidelove metode.

Pogledajmo kako se ponaša $\rho(T_{SOR(\omega)})$ kao funkcija od ω, te ovisno o N, kako se ponaša ω_{opt}. Redom, za $N = 3$, 9, 16, 25, dobivamo sljedeće grafove.

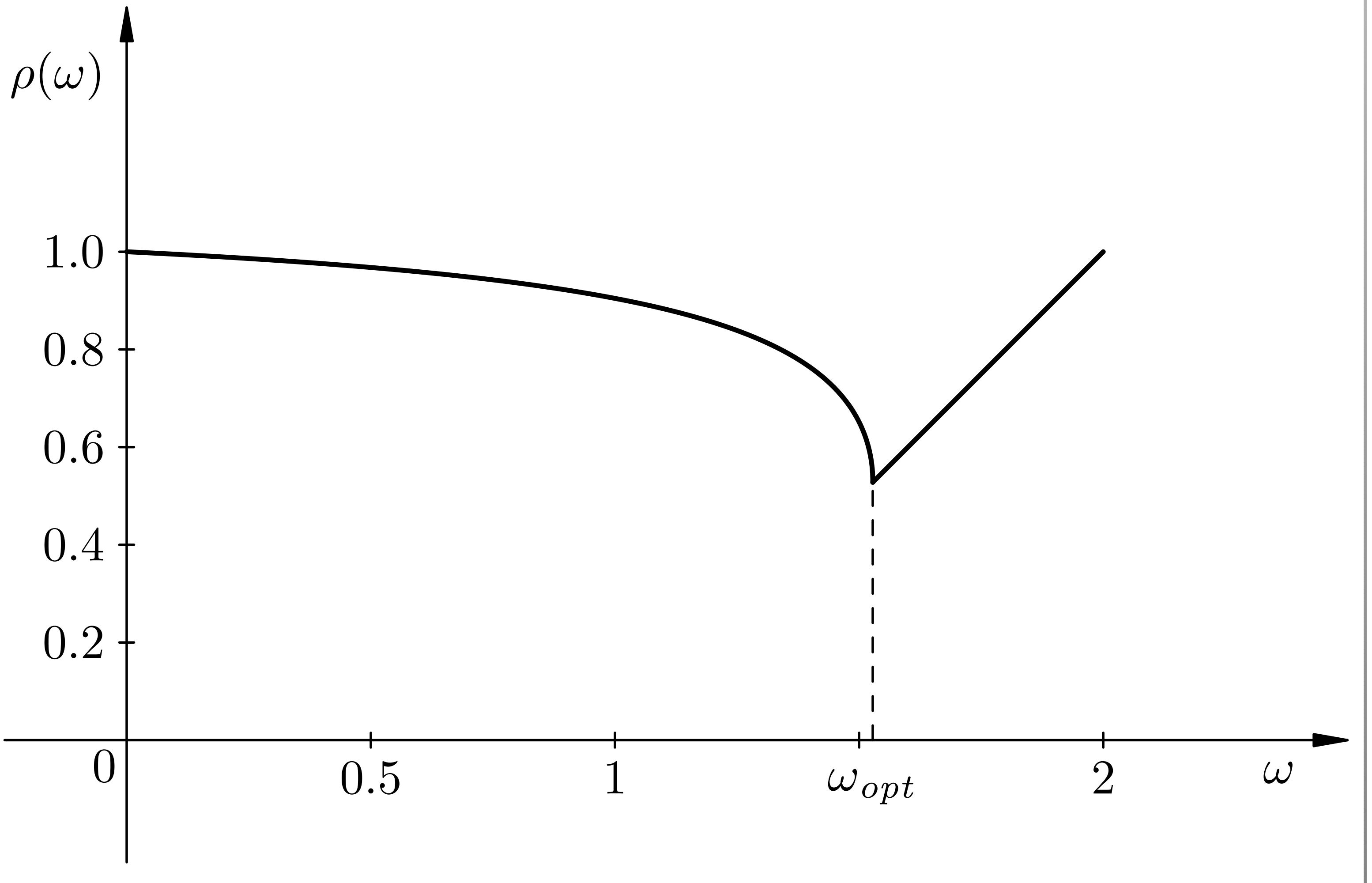
Konvergencija SOR(ω) za sustav s $G_{3\times 3}$

$N = 3$



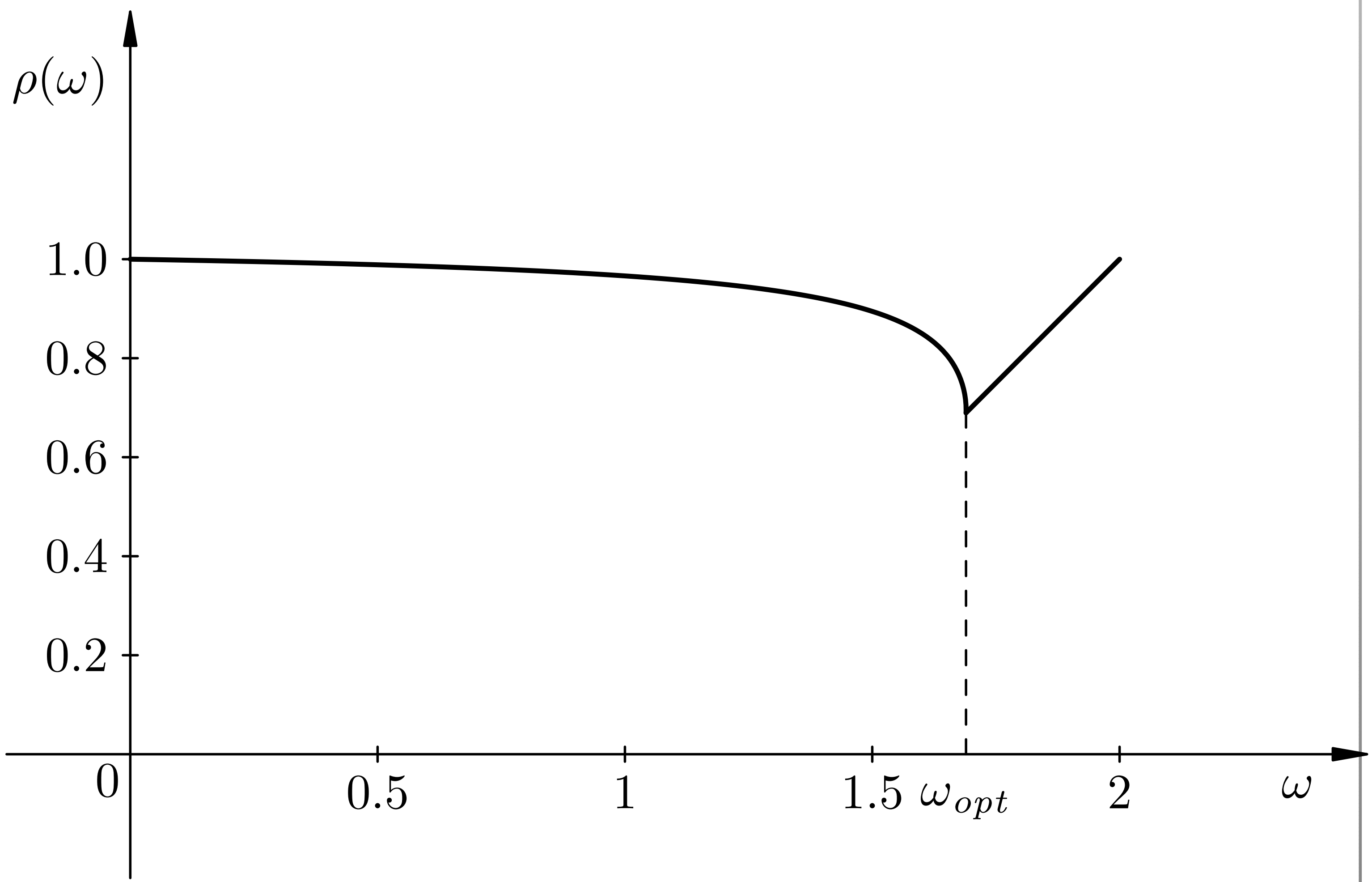
Konvergencija SOR(ω) za sustav s $G_{9\times 9}$

$N = 9$



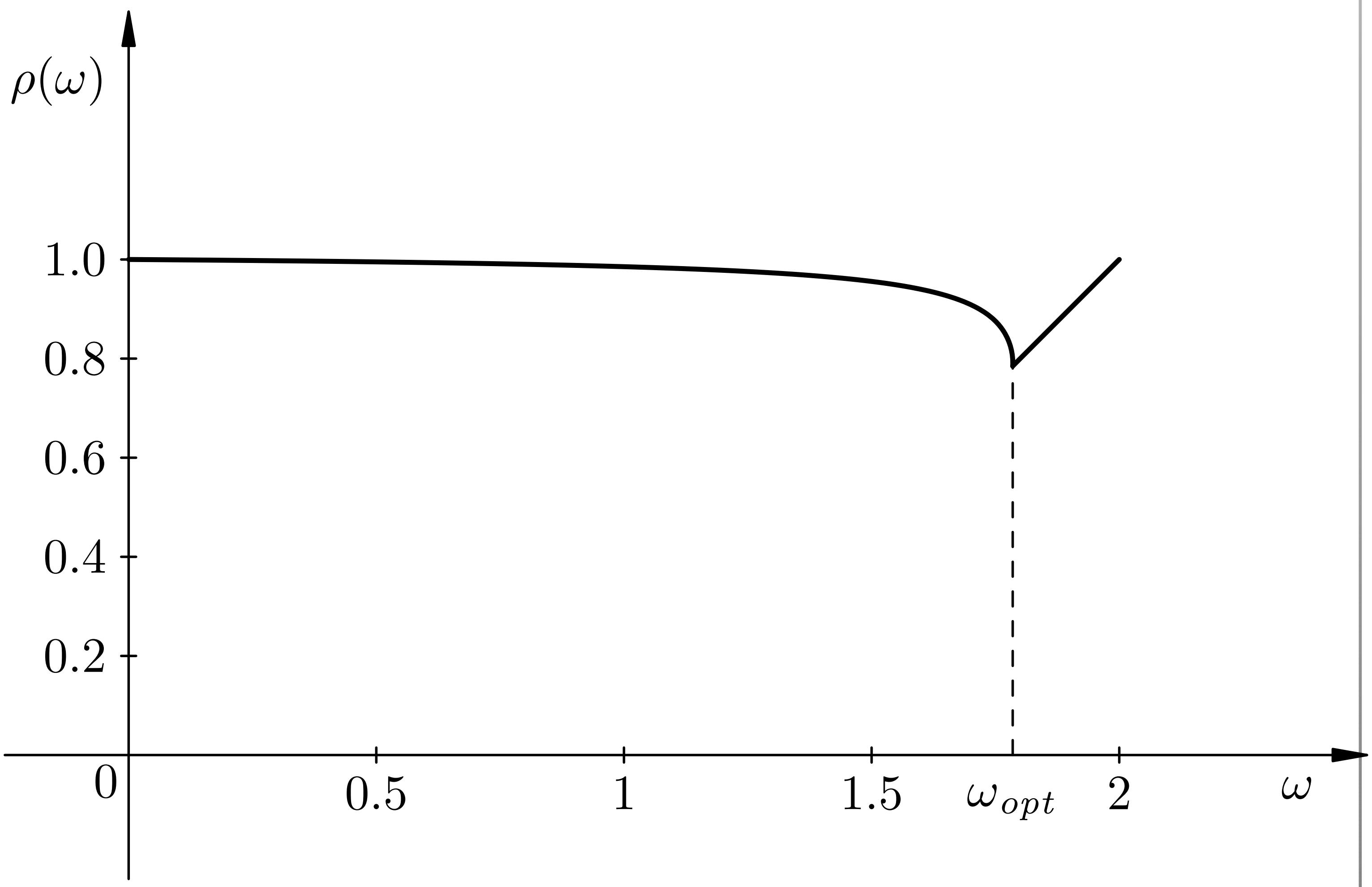
Konvergencija SOR(ω) za sustav s $G_{16\times 16}$

$N = 16$



Konvergencija SOR(ω) za sustav s $G_{25\times 25}$

$N = 25$



Konvergencija SOR(ω) za sustav s $G_{N\times N}$

Kao što smo očekivali, optimalni se parametar pomiče prema 2, a $\rho(T_{SOR(\omega_{opt})})$ postaje sve bliže 1.