# Numeričko rešavanje nelinearnih jednačina i sistema jednačina

## Uvodni pojmovi

**Definicija.** Neka je $X$ neprazan skup. Funkcija $d : X \times X \to \mathbb{R}$ koja zadovoljava uslove:

$$\begin{aligned} &1^\circ \quad d(x,y) \geq 0, \qquad d(x,y) = 0 \ \Leftrightarrow \ x = y, \\ &2^\circ \quad d(x,y) = d(y,x), \\ &3^\circ \quad d(x,y) \leq d(x,z) + d(z,y), \end{aligned}$$

zove se *metrika* ili *rastojanje*, a $(X,d)$, ili samo $X$, je *metrički prostor*.

**Primeri.** $1^\circ$ U skupu $\mathbb{R}$ metrika se definiše na sledeći način:

$$d(x,y) = |x-y|.$$

$2^\circ$ U skupu $\mathbb{R}^n = \{\mathbf{x} = (x_1, x_2, \ldots, x_n) \,|\, x_1, x_2, \ldots, x_n \in \mathbb{R}\}$ metrika može da se definiše na jedan on sledećih načina:

$$d(\mathbf{x},\mathbf{y}) = \left( \sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} \qquad (p \geq 1).$$

Najčešće se koriste sledeće metrike:

$$\begin{aligned} p=1: \quad & d(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i| \ , \\ p=2: \quad & d(\mathbf{x},\mathbf{y}) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} \ , \\ p\to\infty: \quad & d(\mathbf{x},\mathbf{y}) = \max_{1\le i\le n} |x_i - y_i| \ . \end{aligned}$$

Neka je $(X,d)$ metrički prostor.

**Definicija.** Funkcija $a:\mathbb{N}\to X,\ \ a(n)=a_n$ zove se *niz* u prostoru $X$ i označava se sa $\{a_n\}_{n\in\mathbb{N}}$.

**Definicija.** Niz $\{a_n\}_{n\in\mathbb{N}}$ je *konvergentan* ako postoji tačka $a\in X$ takva da

$$(\forall\varepsilon>0)(\exists n_0\in\mathbb{N})(\forall n\ge n_0)\ d(a_n,a)<\varepsilon\,.$$

Tačka $a$ je *granična vrednost* ili *granica* niza $\{a_n\}_{n\in\mathbb{N}}$, što se označava sa

$$\lim_{n\to\infty} a_n = a \qquad \text{ili} \qquad a_n\to a \quad (n\to\infty)\,.$$

Ako niz nije konvergentan, on je *divergentan*.

**Definicija.** Niz $\{a_n\}_{n\in\mathbb{N}}$ je *Košijev* ako važi

$$(\forall\varepsilon>0)(\exists n_0\in\mathbb{N})(\forall m,n\ge n_0)\ d(a_m,a_n)<\varepsilon\,,$$

ili,

$$(\forall\varepsilon>0)(\exists n_0\in\mathbb{N})(\forall n\ge n_0)(\forall p\in\mathbb{N})\ d(a_{n+p},a_n)<\varepsilon\,.$$

Niz $\{a_n\}_{n\in\mathbb{N}}$ u prostoru $\mathbb{R}$ ili $D\subset\mathbb{R}$ zove se *realni niz.* Kada se radi o realnim nizovima u prostoru sa uobičajenom metrikom $d(x,y)=|x-y|$, $x,y\in\mathbb{R}$, prethodne definicije dobijaju sledeći oblik.

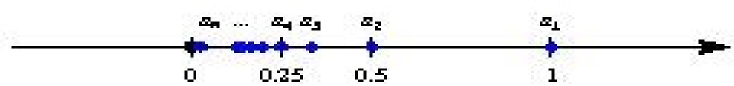
Realni niz $\{a_n\}_{n\in\mathbb{N}}$ je *konvergentan* u $D\subseteq\mathbb{R}$ ako postoji tačka $a\in D$ takva da

$$(\forall\varepsilon>0)(\exists n_0\in\mathbb{N})(\forall n\ge n_0)\ |a_n-a|<\varepsilon\,.$$

Realni niz $\{a_n\}_{n\in\mathbb{N}}$ je *Košijev* ako važi

$$(\forall\varepsilon>0)(\exists n_0\in\mathbb{N})(\forall n\ge n_0)(\forall p\in\mathbb{N})\ |a_{n+p}-a_n|<\varepsilon\,.$$

**Primer 1.** Niz $\{a_n\}_{n\in\mathbb{N}}$, gde je $a_n = \dfrac{1}{n}$, $n \in \mathbb{N}$, je konvergentan u $\mathbb{R}$ i $\lim\limits_{n\to\infty} a_n = 0$. Zaista, za proizvoljno $\varepsilon > 0$ postoji $n_0 = \left[\dfrac{1}{\varepsilon}\right] + 1$ tako da za svako $n \geq n_0$ važi $|a_n| = \dfrac{1}{n} < \varepsilon$. Ipak, iako svi članovi niza pripadaju skupu $\mathbb{R}^+ = (0, +\infty)$, on nije konvergentan u $\mathbb{R}^+$, jer $\lim\limits_{n\to\infty} a_n = 0 \notin \mathbb{R}^+$.




**Primer 2.** Poznato je da je $\sqrt{2} = 1.41421\ldots$ iracionalan broj koji u decimalnom zapisu ima beskonačno mnogo decimala. Neka je $b_n$, $n = 0, 1, 2, \ldots$, racionalan broj koji se dobija odbacivanjem svih decimala broja $\sqrt{2}$ posle $n$-te, tj. $b_0 = 1$, $b_1 = 1.4$, $b_2 = 1.41$, $b_3 = 1.414$, $b_4 = 1.4142$, itd. Dobijeni niz $\{b_n\}_{n\in\mathbb{N}}$ je Košijev. Zaista, za proizvoljno (malo) $\varepsilon > 0$ postoji $n_0 = [-\log_{10}\varepsilon] + 1$ takvo da za svako $n > n_0$ i svako $p \in \mathbb{N}$ važi

$$|b_{n+p} - b_n| \leq 9 \cdot 10^{-(n+1)} < 10^{-n} < \varepsilon.$$

Jasno je da je $\lim\limits_{n\to\infty} b_n = \sqrt{2} \notin \mathbb{Q}$, pa je niz konvergentan u $\mathbb{R}$, ali ne i u $\mathbb{Q}$.

**Teorema.** *U skupu $\mathbb{R}$ niz $\{a_n\}_{n\in\mathbb{N}}$ je konvergentan ako i samo ako je Košijev.*

**Definicija.** Metrički prostor u kome svaki Košijev niz konvergira zove se *kompletan metrički prostor.*

Skup $\mathbb{R}$ i svaki zatvoreni interval $[\alpha, \beta] \subset \mathbb{R}$ su kompletni metrički prostori.

**Definicija.** Neka je realna funkcija $f(x)$ definisana u nekoj okolini tačke $x_0 \in \mathbb{R}$, osim, možda, u tački $x_0$. Funkcija ima *graničnu vrednost* $A \in \mathbb{R}$ u tački $x_0$, tj.

$$\lim_{x\to x_0} f(x) = A,$$

ako važi

$$(\forall\varepsilon > 0)(\exists\delta > 0) \quad 0 < |x - x_0| < \delta \;\Rightarrow\; |f(x) - A| < \varepsilon\,.$$

**Definicija.** Funkcija $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, je *neprekidna u tački* $x_0 \in D$ ako je

$$\lim_{x\to x_0} f(x) = f(x_0).$$

Funkcija je *neprekidna na skupu* $D$ ako je neprekidna u svakoj tački $x \in D$. Skup svih funkcija neprekidnih na skupu $D$ označava se sa $C(D)$.

**Definicija.** Funkcija $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, koja je neprekidna u tački $x_0 \in D$, ima *prvi izvod u tački* $x_0$ ako postoji konačna granična vrednost

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Ako funkcija ima konačan prvi izvod u $x_0$, ona je *diferencijabilna u tački* $x_0$. Funkcija je *diferencijabilna na skupu* $D$ ako je diferencijabilna u svakoj tački $x \in D$. Ako je funkcija diferencijabilna na skupu $D$ i njen izvod je neprekidna funkcija na $D$, ona je *neprekidno-diferencijabilna na* $D$. Skup svih funkcija neprekidno-diferencijabilnih na skupu $D$ označava se sa $C^1(D)$.

**Definicija.** *Izvod reda* $n \in \mathbb{N}$ realne funkcije $f(x)$ je

$$f^{(n)}(x) = \big(f^{(n-1)}(x)\big)', \ n = 1, 2, \ldots, \quad f^{(0)}(x) = f(x).$$

Ako funkcija ima izvod reda $n$ na skupu $D$, ona je $n$ *puta diferencijabilna na* $D$. Ako je još $f^{(n)}(x)$ neprekidna funkcija na $D$, ona je $n$ *puta neprekidno-diferencijabilna na* $D$. Skup svih funkcija $n$ puta neprekidno-diferencijabilnih na skupu $D$ označava se sa $C^n(D)$.

**Teorema.** *Neka je funkcija* $f(x)$ $n$ *puta diferencijabilna u nekoj okolini tačke* $x_0 \in \mathbb{R}$. *Tada važi*

$$f(x) = T_n(x) + o\big((x - x_0)^n\big), \qquad x \to x_0,$$

*gde je*

$$T_n(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

*Tejlorov polinom stepena* $n$ *funkcije* $f(x)$ *u okolini tačke* $x_0$.

**Definicija.** Neka je $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, definisana u nekoj okolini tačke $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. *Parcijalni izvod funkcije* $f$ *po promenljivoj* $x_j$ *u tački* $\mathbf{x}$ je

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f(x_1, x_2, \ldots, x_j, \ldots, x_n)}{h}.$$

*Parcijalni izvodi drugog reda funkcije $f$* su

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right).$$

**Teorema.** *Neka funkcija $f(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, ima neprekidne parcijalne izvode u nekoj okolini tačke $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_n)$. Tada važi*

$$f(\mathbf{x}) = T_1(\mathbf{x}) + R_1(\mathbf{x}),$$

*gde je*

$$T_1(\mathbf{x}) = f(\mathbf{a}) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{a})\,(x_j - a_j)$$

*Tejlorov polinom prvog stepena funkcije $f(\mathbf{x})$ u okolini tačke $\mathbf{a}$, a $R_1(\mathbf{x}) = o\big(d(\mathbf{x}, \mathbf{a})\big)$, $\mathbf{x} \to \mathbf{a}$.*

## Opšta teorija iterativnih procesa

Neka je data funkcija $f : [\alpha, \beta] \to \mathbb{R}$ i neka jednačina

$$f(x) = 0 \tag{1}$$

ima rešenje $x^* \in [\alpha, \beta]$. Jedan od načina za približno rešavanje jednačine, tj. određivanje približnog rešenja je formiranje niza $\{x_k\}_{k\in\mathbb{N}}$ u $[\alpha, \beta]$ takvog da je

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

Jednačina (1) može da se na beskonačno mnogo načina predstavi u obliku

$$x = \Phi(x) \qquad (\Phi : [\alpha, \beta] \to \mathbb{R}) \tag{2}$$

(na primer: $x = x + \lambda f(x)$ za proizvoljno $\lambda \in \mathbb{R}$ ili $x = x + \big(f(x)\big)^{2\mu}$ za $\mu \in \mathbb{R}$). Ova jednačina, kao ekvivalentna polaznoj, ima isto rešenje $x^*$, a njen oblik sugeriše formiranje niza koji konvergira ka $x^*$:

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, 2, \ldots, \tag{3}$$

pri čemu je $x_0 \in [\alpha, \beta]$ početna vrednost. Niz $\{x_k\}_{k\in\mathbb{N}}$ se zove *iterativni niz*, funkcija $\Phi : [\alpha, \beta] \to \mathbb{R}$ je *iterativna funkcija*, a formulom (3) definisan je *iterativni proces* koji se naziva *metod proste iteracije* ili *opšti iterativni metod* za rešavanje jednačine (2), odnosno (1).

Iterativni metod ima smisla samo ako iterativni niz konvergira ka tačnom rešenju, a kao približno rešenje jednačine može da se uzme onaj član niza $x_k$ za koji važi:

$$|x_k - x_{k-1}| < \varepsilon$$

gde je $\varepsilon$ unapred zadata tačnost.

U cilju određivanja uslova pod kojima iterativni proces zaista konvergira ka tačnom rešenju jednačine, najpre će biti formulisano i dokazano tvrđenje poznato pod nazivom *Banahov stav o fiksnoj tački*.

**Definicija.** Neka je $X$ metrički prostor. Funkcija $F : X \to X$ je kontrakcija ako postoji $q \in [0,1)$ tako da

$$(\forall x, y \in X)\ d\big(f(x), f(y)\big) \le q\, d(x,y).$$

**Definicija.** Tačka $x^* \in X$ je *fiksna* ili *nepokretna tačka* funkcije $F : X \to X$ ako važi

$$F(x^*) = x^*.$$

**Teorema.** *Neka je $X$ kompletan metrički prostor i $F : X \to X$ kontrakcija. Tada postoji jedinstvena fiksna tačka funkcije $F$ i ona je granična vrednost niza $\{x_k\}_{k\in\mathbb{N}}$ definisanog sa $x_{k+1} = F(x_k)$, $k = 0, 1, 2, \ldots$, za proizvoljno $x_0 \in X$.*

*Dokaz.* Bez umanjenja opštosti, dokaz će biti izveden za slučaj realne funkcije realne promenljive, tj. kada je $X = \mathbb{R}$ (ili $X = [\alpha, \beta] \subset \mathbb{R}$) sa uobičajenom metrikom $d(x,y) = |x - y|$, $x, y \in \mathbb{R}$.

Neka je $x_0 \in \mathbb{R}$ proizvoljno i $x_{k+1} = F(x_k)$, $k = 0, 1, 2, \ldots$. Dokazaćemo najpre da je ovako formiran niz $\{x_k\}_{k\in\mathbb{N}}$ Košijev. Kako je $F$ kontrakcija, postoji $q \in [0,1)$ tako da je

$$\begin{aligned}|x_{k+1} - x_k| &= |F(x_k) - F(x_{k-1})| \le q|x_k - x_{k-1}| \\ &\le q^2|x_{k-1} - x_{k-2}| \le \cdots \le q^k|x_1 - x_0|, \quad k \in \mathbb{N}_0.\end{aligned}$$

Zato za proizvoljne $k, p \in \mathbb{N}$ važi

$$\begin{aligned}|x_{k+p} - x_k| &= |x_{k+p} - x_{k+p-1} + x_{k+p-1} - x_{k+p-2} + \cdots + x_{k+1} - x_k| \\ &\le |x_{k+p} - x_{k+p-1}| + |x_{k+p-1} - x_{k+p-2}| + \cdots + |x_{k+1} - x_k| \\ &\le q^{k+p-1}|x_1 - x_0| + q^{k+p-2}|x_1 - x_0| + \cdots + q^k|x_1 - x_0| \\ &= \big(q^{p-1} + q^{p-2} + \cdots + 1\big)q^k|x_1 - x_0| \\ &= \frac{1 - q^p}{1 - q}\, q^k|x_1 - x_0| \le \frac{|x_1 - x_0|}{1 - q}\, q^k.\end{aligned}$$

Neka je $\varepsilon > 0$ proizvoljno. Tada, za svako $k \geq k_0$, $k_0 = \left[ \dfrac{\log \frac{(1-q)\varepsilon}{|x_1 - x_0|}}{\log q} \right] + 1$, važi

$$|x_{k+p} - x_k| \leq \frac{|x_1 - x_0|}{1-q}\, q^k < \varepsilon,$$

što znači da je niz $\{x_k\}_{k\in\mathbb{N}}$ Košijev. Metrički prostor $\mathbb{R}$ je kompletan, pa je niz $\{x_k\}_{k\in\mathbb{N}}$ konvergentan, tj. postoji $x^* = \lim\limits_{k\to\infty} x_k$.

Dokažimo da je $x^*$ fiksna tačka funkcije $F$. Za svako $k \in \mathbb{N}$ važi

$$\begin{aligned}
|x^* - F(x^*)| &= |x^* - x_k + x_k - F(x_k) + F(x_k) - F(x^*)| \\
&\leq |x^* - x_k| + |x_k - F(x_k)| + |F(x_k) - F(x^*)| \\
&\leq |x^* - x_k| + |x_k - x_{k+1}| + q|x_k - x^*| \\
&= (1+q)|x_k - x^*| + |x_k - x_{k+1}| \\
&\leq (1+q)|x_k - x^*| + q^k|x_1 - x_0|.
\end{aligned}$$

Zbog konvergencije nizova $\{x_k\}_{k\in\mathbb{N}}$ i $\{q^k\}_{k\in\mathbb{N}}$, desna strana nejednakosti teži 0 kad $k \to \infty$, pa je $|x^* - F(x^*)| = 0$, tj. $x^* = F(x^*)$.

Pokažimo još da je fiksna tačka $x^*$ je jedinstvena. Pretpostavimo da postoji još jedna $x^{**}$ takva da je $x^{**} = F(x^{**})$. Tada važi nejednakost

$$|x^* - x^{**}| = |F(x^*) - F(x^{**})| \leq q|x^* - x^{**}|,$$

koja je zbog činjenice da je $0 \leq q < 1$, moguća samo ako je $x^* = x^{**}$. □

Postavimo sada uslove za konvergenciju iterativnog procesa (3), tj. metoda proste iteracije.

**Teorema.** *Neka je $\Phi(x)$ neprekidna funkcija koja zadovoljava sledeće uslove:*
1° $\Phi : [\alpha, \beta] \to [\alpha, \beta]$;
2° $\Phi(x)$ *ima izvod na* $[\alpha, \beta]$ *takav da za svako* $x \in [\alpha, \beta]$ *važi*

$$|\Phi'(x)| \leq q < 1.$$

*Tada jednačina (2) odnosno (1) ima jedinstveno rešenje* $x^* \in [\alpha, \beta]$ *koje je granična vrednost niza* $\{x_k\}_{k\in\mathbb{N}}$ *definisanog sa* $x_{k+1} = \Phi(x_k)$, $k = 0, 1, 2, \ldots$, *za proizvoljno* $x_0 \in [\alpha, \beta]$.

*Dokaz.* Pod navedenim uslovima funkcija $\Phi(x)$ je kontrakcija. Zaista, za proizvoljne $x, y \in [\alpha, \beta]$ na osnovu Lagranžove teoreme o srednjoj vrednosti postoji $\xi$ između njih tako važi

$$|\Phi(x) - \Phi(y)| = |\Phi'(\xi)|\,|x - y| \leq q|x - y|, \quad 0 \leq q < 1.$$

Segment $[\alpha, \beta]$ je kompletan metrički prostor, pa prema Banahovom stavu postoji jedinstvena tačka $x^* \in [\alpha, \beta]$ za koju važi $x^* = \Phi(x^*)$, tj. koja je rešenje jednačine (2), a koja se dobija kao granična vrednost niza $x_{k+1} = \Phi(x_k)$, $k = 0, 1, 2, \ldots$, za proizvoljno $x_0 \in [\alpha, \beta]$. $\square$

Za približno rešenje dobijeno metodom proste iteracije može se proceniti greška na sledeći način.

**Posledica.** *Neka funkcija $\Phi(x)$ zadovoljava uslove prethodne teoreme. Tada važi:*

$$|x_k - x^*| \leq \frac{q^k}{1-q}|x_1 - x_0|.$$

*Dokaz.* Za proizvoljan član $x_k$ niza $\{x_n\}_{n\in\mathbb{N}}$ važi procena

$$|x_k - x^*| = |\Phi(x_{k-1}) - \Phi(x^*)| \leq q|x_{k-1} - x^*| \leq q^2|x_{k-2} - x^*| \leq \cdots \leq q^k|x_0 - x^*|.$$

Kako je

$$|x_0 - x^*| \leq |x_0 - x_1| + |x_1 - x^*| \leq |x_1 - x_0| + q|x_0 - x^*|$$

i $0 \leq q < 1$, važi

$$|x_0 - x^*| \leq \frac{1}{1-q}|x_1 - x_0|,$$

a stoga i

$$|x_k - x^*| \leq \frac{q^k}{1-q}|x_1 - x_0|. \quad \square$$

Za primenu iterativnih procesa ključna je njegova konvergencija, ali su važni i brzina konvergencije i ponašanje odstupanja približnog od tačnog rešenja posle velikog broja iteracija.

Neka je (3) konvergentan iterativni proces.

**Definicija.** *Red konvergencije* iterativnog procesa je $r$ ako je

$$|x_{k+1} - x^*| = \mathcal{O}\big(|x_k - x^*|^r\big), \qquad k \to \infty,$$

tj. ako postoji konstanta $A > 0$ takva da je $|x_{k+1} - x^*| \leq A|x_k - x^*|^r$ za dovoljno veliko $k$. Veličina

$$\lim_{k\to\infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^r} = C_r$$

zove se *asimptotska konstanta greške* ili *faktor konvergencije* iterativnog procesa.
Za određivanje reda konvergencije u praksi koristi se sledeće tvrđenje.

**Teorema.** *Neka je $\Phi : [\alpha, \beta] \to [\alpha, \beta]$ $r$ puta diferencijabilna. Ako važi*

$$\begin{aligned} \Phi(x^*) &= x^*, \\ \Phi'(x^*) &= \Phi''(x^*) = \cdots = \Phi^{(r-1)}(x^*) = 0, \\ \Phi^{(r)}(x^*) &\neq 0, \end{aligned}$$

*tada iterativni proces (3) ima red konvergencije $r$. Ako je $\Phi \in C^r[\alpha, \beta]$, tada je asimptotska konstanta greške*

$$C_r = \frac{|\Phi^{(r)}(x^*)|}{r!}.$$

*Dokaz.* Neka iterativni proces (3) konvergira ka tački $x^* \in [\alpha, \beta]$. Razvojem funkcije $\Phi$ po Tejlorovoj formuli u okolini tačke $x^*$ dobija se

$$\begin{aligned} x_{k+1} &= \Phi(x_k) \\ &= \Phi(x^*) + \frac{\Phi'(x^*)}{1!}(x_k - x^*) + \frac{\Phi''(x^*)}{2!}(x_k - x^*)^2 \\ &\quad + \cdots + \frac{\Phi^{(r-1)}(x^*)}{(r-1)!}(x_k - x^*)^{r-1} + \frac{\Phi^{(r)}(\xi_k)}{r!}(x_k - x^*)^r, \end{aligned}$$

gde je $\xi_k$ neka tačka između $x_k$ i $x^*$. Prema uslovima teoreme

$$x_{k+1} - x^* = \frac{\Phi^{(r)}(\xi_k)}{r!}(x_k - x^*)^r,$$

tj.

$$|x_{k+1} - x^*| = \frac{|\Phi^{(r)}(\xi_k)|}{r!}|x_k - x^*|^r = A|x_k - x^*|^r,$$

gde je

$$A = \frac{|\Phi^{(r)}(\xi_k)|}{r!}.$$

Konstanta $A$ je pozitivna zbog uslova $\Phi^{(r)}(x^*) \neq 0$. Prema definiciji, red konvergencije iterativnog procesa je $r$. Osim toga, ako je $\Phi^{(r)}(x)$ neprekidna funkcija, važi i

$$\lim_{k\to\infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^r} = \lim_{k\to\infty} \frac{|\Phi^{(r)}(\xi_k)|}{r!} = \frac{|\Phi^{(r)}(x^*)|}{r!},$$

pa je asimptotska konstanta greške metoda jednaka

$$C_r = \frac{|\Phi^{(r)}(x^*)|}{r!}. \quad \square$$

Ako je iterativni proces konvergentan, brzina konvergencije može značajno da se ubrza povećanjem reda konvergencije. Sledeće dve teoreme daju načine za povećanje reda konvergencije.

**Teorema.** *Neka iterativni proces $x_{k+1} = \Phi(x_k)$, $k = 0, 1, 2, \ldots$, konvergira ka $x^*$ sa redom konvergencije $r$, neka je funkcija $\Phi$ diferencijalna $r + 1$ puta u okolini tačke $x^*$ i neka je $\Phi'(x^*) \neq r$. Tada je iterativni proces*

$$x_{k+1} = x_k - \frac{x_k - \Phi(x_k)}{1 - \frac{1}{r}\Phi'(x_k)}, \qquad k = 0, 1, 2, \ldots$$

*konvergentan sa redom konvergencije najmanje $r + 1$.*

**Teorema.** *Neka iterativni proces $x_{k+1} = \Phi(x_k)$, $k = 0, 1, 2, \ldots$, konvergira ka $x^*$ sa redom konvergencije $r \geq 2$ i neka je funkcija $\Phi$ diferencijalna $r + 1$ puta u okolini tačke $x^*$. Tada je iterativni proces*

$$x_{k+1} = \Phi(x_k) - \frac{1}{r}\Phi'(x_k)\big(x_k - \Phi(x_k)\big), \qquad k = 0, 1, 2, \ldots$$

*konvergentan sa redom konvergencije najmanje $r + 1$.*

Osim iterativnih procesa oblika (3), za rešavanje jednačina koriste se i procesi oblika

$$x_{k+m} = \Phi(x_{k+m-1}, x_{k+m-2}, \ldots, x_k), \qquad k = 0, 1, 2, \ldots,$$

pri čemu je neophodno $m$ početnih vrednosti $x_0, x_1, \ldots, x_{m-1}$. Red konvergencije ovakvih metoda, za razliku od prethodnih, ne mora da bude ceo broj.

# Metodi za rešavanje nelinearnih jednačina

## Njutnov metod

Neka je data jednačina

$$f(x) = 0,$$

pri čemu funkcija $f$ zadovoljava sledeće uslove: $f \in C^1[\alpha, \beta]$ i $f'(x) \neq 0$ za svako $x \in [\alpha, \beta]$.

Pretpostavimo da jednačina ima prosto izolovano rešenje $x^* \in [\alpha, \beta]$. Neka je $x_0 \in [\alpha, \beta]$ proizvoljna tačka. Prema Tejlorovoj formuli u okolini $x_0$ važi

$$\begin{aligned} f(x^*) &= f(x_0) + f'(x_0)(x^* - x_0) + o(x^* - x_0), \qquad x^* \to x_0 \\ 0 &\approx f(x_0) + f'(x_0)(x^* - x_0), \\ x^* &\approx x_0 - \frac{f(x_0)}{f'(x_0)} \qquad (f'(x_0) \neq 0 \text{ jer je } x_0 \text{ prosta nula funkcije } f(x)). \end{aligned}$$

Poslednja formula sugeriše konstrukciju iterativnog procesa

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, 2, \ldots,$$

sa početnom vrednošću $x_0 \in [\alpha, \beta]$. Ovaj metod rešavanja jednačina zove se *Njutnov metod*, *Njutn-Rafsonov metod* ili *metod tangente*. O njegovoj konvergenciji govore sledeće dve teoreme.

**Teorema.** *Neka jednačina*

$$f(x) = 0$$

*ima rešenje $x^* \in [\alpha, \beta]$, pri čemu funkcija $f$ zadovoljava sledeće uslove: $f \in C^2[\alpha, \beta]$ i $f'(x) \neq 0$ za svako $x \in [\alpha, \beta]$. Tada postoji segment $U(x^*) = [x^* - \delta, x^* + \delta]$ takav da iterativni proces*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, 2, \ldots,$$

*konvergira za svako $x_0 \in U(x^*)$.*

*Dokaz.* Za iterativnu funkciju Njutnovog metoda

$$\Phi(x) = x - \frac{f(x)}{f'(x)}$$

važi

$$\Phi'(x) = 1 - \frac{f'^2(x) - f(x) f''(x)}{f'^2(x)} = \frac{f(x) f''(x)}{f'^2(x)},$$

tj.,

$$\Phi(x^*) = x^*, \qquad \Phi'(x^*) = 0.$$

Zbog neprekidnosti funkcija $f(x)$, $f'(x)$ i $f''(x)$ i uslova $f'(x) \neq 0$ na $[\alpha, \beta]$, i funkcija $\Phi'(x)$ je neprekidna na $[\alpha, \beta]$, pa postoji segment $[x^* - \delta, x^* + \delta] = U(x^*)$ takav da je $|\Phi'(x)| \leq q < 1$ za svako $x \in U(x^*)$. Osim toga, važi i

$$\begin{aligned} x \in U(x^*) \quad &\Rightarrow \quad |x - x^*| \leq \delta \\ &\Rightarrow \quad |\Phi(x) - x^*| = |\Phi(x) - \Phi(x^*)| = |\Phi'(\xi)||x - x^*| \\ &\qquad \leq q|x - x^*| < |x - x^*| < \delta \\ &\Rightarrow \quad \phi(x) \in U(x^*), \end{aligned}$$

što znači da $\Phi : U(x^*) \to U(x^*)$. Prema tome, ispunjeni su uslovi za konvergenciju Njutnovog iterativnog procesa na $U(x^*)$. $\square$

**Teorema.** *Neka su zadovoljeni uslovi prethodne teoreme. Tada Njutnov iterativni proces ima red konvergencije $r = 2$ i asimptotsku konstantu greške*

$$C_2 = \frac{|f''(x^*)|}{2|f'(x^*)|}.$$

*Dokaz.* Prema definicionoj formuli iterativnog niza važi

$$\begin{gathered} x_{k+1} - x^* = x_k - x^* - \frac{f(x_k) - f(x^*)}{f'(x_k)}, \\ f(x_k) - f(x^*) = f'(x_k)(x_k - x^*) - f'(x_k)(x_{k+1} - x^*). \end{gathered}$$

S druge strane, razvoj funkcije $f$ po Tejlorovoj formuli oko tačke $x_k$ daje

$$f(x^*) - f(x_k) = f'(x_k)(x^* - x_k) + \frac{f''(\xi_k)}{2!}(x^* - x_k)^2, \qquad \xi_k \text{ između } x^* \text{ i } x_k.$$

Sabiranjem ovih jednakosti dobija se

$$\begin{gathered} 2f'(x_k)(x_{k+1} - x^*) = f''(\xi_k)(x^* - x_k)^2, \\ \frac{x_{k+1} - x^*}{(x^* - x_k)^2} = \frac{f''(\xi_k)}{2f'(x_k)}. \end{gathered}$$

Zbog neprekidnosti funkcija $f'(x)$ i $f''(x)$ važi

$$\lim_{k \to \infty} \frac{x_{k+1} - x^*}{(x^* - x_k)^2} = \frac{f''(x^*)}{2f'(x^*)},$$

što dokazuje tvrđenje teoreme. $\square$

Ako je rešenje $x^* \in [\alpha, \beta]$ višestruko, Njutnov metod konvergira, ali sa redom konvergencije $r = 1$. Postoje, međutim, modifikacije ovog metoda koje zadržavaju kvadratnu konvergenciju.

Ako je višestrukost rešenja poznata, tj. ako se zna da je $x^*$ nula funkcije $f(x)$ reda $m$, tada se ona prosta nula funkcije $F(x) = \sqrt[m]{f(x)}$. Primenom Njutnovog metoda na rešavanje jednačine $F(x) = 0$ dobija se iterativni proces

$$x_{k+1} = x_k - m\frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \ldots, \qquad x_0 \in U(x^*).$$

Ako je višestrukost rešenja poznata, tada se Njutnov metod primenjuje na određivanje proste nule funkcije $F(x) = \dfrac{f(x)}{f'(x)}$. Na taj način dobija se iterativni niz

$$x_{k+1} = x_k - \frac{f(x_k)f'(x_k)}{f'^2(x_k) - f(x_k)f''(x_k)}, \quad k = 0, 1, 2, \ldots, \qquad x_0 \in U(x^*).$$

## Metod sečice

Primena Njutnovog metoda za rešavanje jednačine

$$f(x) = 0$$

zahteva diferencijabilnost funkcije $f(x)$. U praksi se često dešava da $f'(x)$ ne postoji ili da je izračunavanje njegove vrednosti u svakoj iteraciji previše računski zahtevno. U tom slučaju može da se izvrši aproksimacija

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

pa iterativni niz dobija oblik

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}f(x_k), \qquad k = 0, 1, 2, \ldots,$$

pri čemu su potrebne dve početne vrednosti $x_0, x_1 \in [\alpha, \beta]$. Rešavanje jednačine primenom ovog iterativnog niza zove se *metod sečice*.

Pod izvesnim uslovima koje zadovoljava funkcija $f(x)$, konvergencija metoda sečice je zagarantovana.

**Teorema.** *Neka jednačina*

$$f(x) = 0$$

*ima rešenje* $x^* \in [\alpha, \beta]$, *pri čemu funkcija* $f$ *zadovoljava sledeće uslove:* $f \in C^2[\alpha, \beta]$ *i* $f'(x) \neq 0$ *za svako* $x \in [\alpha, \beta]$. *Tada postoji segment* $U(x^*) = [x^* - \delta, x^* + \delta]$ *takav da iterativni proces*

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k), \qquad k = 0, 1, 2, \ldots,$$

*konvergira za svako* $x_0, x_1 \in U(x^*)$. *Pri tome, red konvergencije i asimptotska konstanta greške su*

$$r = \frac{1 + \sqrt{5}}{2} \approx 1.62, \qquad C_r = \Big| \frac{f''(x^*)}{2f'(x^*)} \Big|^{1/r} = \Big| \frac{f''(x^*)}{2f'(x^*)} \Big|^{\frac{\sqrt{5}-1}{2}}.$$

## Metod regula falsi

*Metod regula falsi* je modifikacija metoda sečice takva da se jedna od početnih vrednosti zadržava u svim iteracijama:

$$x_{k+1} = x_k - \frac{x_k - x_0}{f(x_k) - f(x_0)} f(x_k), \qquad k = 0, 1, 2, \ldots,$$

pri čemu se početne vrednosti $x_0, x_1 \in U(x^*)$ moraju uzeti tako da budu sa različitih strana rešenja jednačine $x^*$, tj. mora da važi $f(x_0) \cdot f(x_1) < 0$.

Ako je $f \in C^1[\alpha, \beta]$, iterativna funkcija ovog metoda

$$\Phi(x) = x - \frac{x - x_0}{f(x) - f(x_0)} f(x)$$

zadovoljava uslove $\Phi(x^*) = x^*$, $\Phi'(x^*) \neq 0$, pa ako metod konvergira, red konvergencije je $r = 1$.

## Metod Stefensena

Ako se u Njutnovom iterativnom nizu vrednost prvog izvoda funkcije u svakoj iteraciji zameni na sledeći način:

$$f'(x_k) \approx \frac{f\big(x_k + f(x_k)\big) - f(x_k)}{f(x_k)},$$

dobija se

$$x_{k+1} = x_k - \frac{\big(f(x_k)\big)^2}{f\big(x_k + f(x_k)\big) - f(x_k)}, \qquad k = 0, 1, 2, \dots,$$
$$x_0 \in [\alpha, \beta].$$

Metod rešavanja jednačina primenom ovog iterativnog procesa zove se *metod Stefensena.*

Ako važi $f \in C^2[\alpha, \beta]$, $f'(x) \neq 0$ na $[\alpha, \beta]$, metod konvergira sa redom konvergencije i asimptotskom konstantom greške

$$r = 2, \qquad C_r = \Big| \frac{f''(x^*)}{2f'(x^*)}\big(f'(x^*) + 1\big) \Big| .$$

## Metod polovljenja intervala

Ovaj metod rešavanja jednačina nije iterativni proces, iako se odvija u broju koraka koji zavisi od zahtevane tačnosti. Jednostavan je za primenu, ali je relativno spor, pa se koristi uglavnom za grubu lokalizaciju rešenja.

Neka jednačina

$$f(x) = 0$$

ima rešenje $x^* \in [\alpha, \beta]$. Tada je sigurno $f(\alpha) \cdot f(\beta) < 0$. *Metod polovljenja intervala* zasniva se na konstrukciji niza intervala $\{[x_k, y_k]\}$ takvih da je

$$y_{k+1} - x_{k+1} = \frac{y_k - x_k}{2}, \quad k = 0, 1, 2, \dots, \qquad f(x_k) \cdot f(y_k) < 0,$$

počevši od $[x_0, y_0] = [\alpha, \beta]$. Na taj način je obezbeđeno da $x^* \in [x_k, y_k]$ za svako $k \in \mathbb{N}$, tj. da je

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

Proces se prekida kad je dužina intervala $[x_k, y_k]$ manja od nekog unapred zadatog broja $\varepsilon > 0$, a za približno rešenje se uzima sredina tog intervala:

$$x^* \approx z_k = \frac{x_k + y_k}{2}.$$

Greška metoda, tj. odstupanje približnog od tačnog rešenja je

$$|z_k - x^*| \leq \frac{1}{2^{k+1}}(\beta - \alpha).$$

## Metodi za simultano određivanje nula polinoma

Zbog specifičnosti i posebnog značaja algebarskih jednačina oblika

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \qquad (a_0, a_1, \ldots, a_{n-1} \in \mathbb{C})$$

postoje mnogi metodi namenjeni njihovom pribliňom rešavanju. Prethodno opisanim metodima mogu se naći rešenja sukcesivno, jedno po jedno na razdvojenim intervalima. Međutim, posebnu klasu čine metodi kojima se odjednom (simultano) određuju sva rešenja jednačine.

Neka polinom

$$P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

ima $n$ prostih (različitih) nula $r_1, r_2, \ldots, r_n$. Njegovim predstavljanjem u faktorisanom obliku

$$P(x) = (x - r_1)(x - r_2)\cdots(x - r_n) = \prod_{j=1}^{n}(x - r_j) = (x - r_i)\prod_{\substack{j=1\\ j\neq i}}^{n}(x - r_j)$$

za svaku nulu $r_i$, $\ i = 1, 2, \ldots, n$, dobija se

$$x - r_i = \frac{P(x)}{\prod\limits_{\substack{j=1\\ j\neq i}}^{n}(x - r_j)}, \qquad \text{tj.} \qquad r_i = x - \frac{P(x)}{\prod\limits_{\substack{j=1\\ j\neq i}}^{n}(x - r_j)}.$$

Ovo sugeriše konstrukciju iterativnog metoda koji je poznat kao *Vajerštrasov metod za simultano određivanje nula polinoma.* Ako su $x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}$ neke aproksimacije nula polinoma $r_1, r_2, \ldots, r_n$ redom, tada se, uzimajući $x = x_i^{(k)}$ i $r_j \approx x_j^{(k)}$, dobija poboljšana aproksimacija nule $r_i$:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{P(x_i^{(k)})}{\prod\limits_{\substack{j=1\\ j\neq i}}^{n}\left(x_i^{(k)} - x_j^{(k)}\right)}, \qquad i = 1, 2, \ldots, n.$$

Na ovaj način se u svakom koraku $k = 0, 1, 2, \ldots$ određuju aproksimacije svih nula polinoma. Drugim rečima, dobija se niz $\left\{\left(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\right)\right\}$ koji koji konvergira ka $(r_1, r_2, \ldots, r_n)$ ako su mu početne vrednosti $\left(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}\right)$ dovoljno blizu. Konvergencija metoda je kvadratna $(r = 2)$.

Postoji važna modifikacija Vajerštrasovog metoda dobijena Gaus-Zajdelovim pristupom. Taj pristup podrazumeva da se u $k$-toj iteraciji koriste ne samo vrednosti iz prethodne, već sve izračunate vrednosti iz tekuće iteracije. Posupak se može predstaviti sledećim formulama:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{P(x_i^{(k)})}{\prod\limits_{j=1}^{i-1}\big(x_i^{(k)} - x_j^{(k)}\big)\prod\limits_{j=i+1}^{n}\big(x_i^{(k)} - x_j^{(k)}\big)}, \quad i = 1, 2, \ldots, n,$$
$$k = 0, 1, 2, \ldots$$

sa početnim vrednostima $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$.

## Metodi za rešavanje sistema nelinearnih jednačina

Neka je dat sistem nelinearnih jednačina

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ f_2(x_1, x_2, \ldots, x_n) &= 0, \\ &\vdots \\ f_n(x_1, x_2, \ldots, x_n) &= 0, \end{aligned} \tag{4}$$

gde su funkcije $f_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$, definisane u nekoj oblasti $D \subset \mathbb{R}^n$. Uvođenjem oznaka

$$\mathbf{x} = [x_1 \;\; x_2 \;\; \cdots \;\; x_n]^T, \qquad \mathbf{f} = [f_1 \;\; f_2 \;\; \cdots \;\; f_n]^T, \qquad \mathbf{0} = [0 \;\; 0 \;\; \cdots \;\; 0]^T,$$

sistem se može zapisati kao vektorska jednačina

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}. \tag{5}$$

Neka je $\mathbf{x}^* = [x_1^* \;\; x_2^* \;\; \cdots \;\; x_n^*]^T \in D$ tačno rešenje jednačine (5) odnosno sistema (4), a $\mathbf{x}^{(k)} = [x_1^{(k)} \;\; x_2^{(k)} \;\; \cdots \;\; x_n^{(k)}]^T \in D$ neka njegova aproksimacija, tj. približno rešenje. Ako funkcije $f_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$, imaju neprekidne parcijalne izvode na $D$, Tejlorovim razvojem u okolini tačke $\mathbf{x}^{(k)}$ dobija se

$$\begin{aligned} f_i(x_1^*, \ldots, x_n^*) = f_i&(x_1^{(k)}, \ldots, x_n^{(k)}) \\ &+ \frac{\partial f_i}{\partial x_1}(x_1^{(k)}, \ldots, x_n^{(k)})(x_1^* - x_1^{(k)}) + \cdots + \frac{\partial f_i}{\partial x_n}(x_1^{(k)}, \ldots, x_n^{(k)})(x_n^* - x_n^{(k)}) \\ &+ r_i^{(k)}, \qquad i = 1, 2, \ldots, n, \end{aligned}$$

gde je $r_i^{(k)}$, $i = 1, 2, \ldots, n$, ostatak u Tejlorovoj formuli. Kako je

$$f_i(x_1^*, \ldots, x_n^*) = 0, \qquad i = 1, 2, \ldots, n,$$

prethodni sistem jednakosti postaje

$$\begin{aligned} 0 &= f_i(x_1^{(k)}, \ldots, x_n^{(k)}) \\ &+ \frac{\partial f_i}{\partial x_1}(x_1^{(k)}, \ldots, x_n^{(k)})(x_1^* - x_1^{(k)}) + \cdots + \frac{\partial f_i}{\partial x_n}(x_1^{(k)}, \ldots, x_n^{(k)})(x_n^* - x_n^{(k)}) \\ &+ r_i^{(k)}, \qquad i = 1, 2, \ldots, n, \end{aligned}$$

Ako se uvedu oznake $\mathbf{r}^{(k)} = [r_1^{(k)} \ r_2^{(k)} \ \cdots \ r_n^{(k)}]^T$ i

$$W = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}$$

sistem se može predstaviti u vektorskom obliku

$$\mathbf{0} = \mathbf{f}(\mathbf{x}^{(k)}) + W(\mathbf{x}^{(k)})(\mathbf{x}^* - \mathbf{x}^{(k)}) + \mathbf{r}^{(k)}.$$

Zanemarivanjem ostatka $\mathbf{r}^{(k)}$ i pod pretpostavkom da je $W(\mathbf{x}^{(k)})$ regularna matrica, važi

$$\begin{aligned} \mathbf{0} &\approx \mathbf{f}(\mathbf{x}^{(k)}) + W(\mathbf{x}^{(k)})(\mathbf{x}^* - \mathbf{x}^{(k)}), \\ -\mathbf{f}(\mathbf{x}^{(k)}) &\approx W(\mathbf{x}^{(k)})(\mathbf{x}^* - \mathbf{x}^{(k)}), \\ \mathbf{x}^* - \mathbf{x}^{(k)} &\approx -W(\mathbf{x}^{(k)})^{-1}\mathbf{f}(\mathbf{x}^{(k)}), \\ \mathbf{x}^* &\approx \mathbf{x}^{(k)} - W(\mathbf{x}^{(k)})^{-1}\mathbf{f}(\mathbf{x}^{(k)}). \end{aligned}$$

Ovo sugeriše konstrukciju iterativnog procesa

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - W(\mathbf{x}^{(k)})^{-1}\mathbf{f}(\mathbf{x}^{(k)}), \quad k = 0, 1, 2, \ldots, \qquad \mathbf{x}^{(0)} \in D,$$

koji je poznat kao *metod Njutn-Kantoroviča*. On se može primeniti za rešavanje sistema nelinearnih jednačina ako iterativni niz $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ konvergira ka tačnom rešenju $\mathbf{x}^*$, tj.

$$\lim_{k\to\infty} d(\mathbf{x}^{(k)}, \mathbf{x}^*) = 0,$$

pri čemu je $d(\mathbf{x}, \mathbf{y})$ izabrana metrika u $\mathbb{R}^n$. Proces se zaustavlja u onoj iteraciji $k$ kada je $d\big(\mathbf{x}^{(k)}, \mathbf{x}^{(k-1)}\big) < \varepsilon$, gde je $\varepsilon > 0$ neki unapred zadati broj, a za približno rešenje jednačine (5) odnosno sistema (4) se uzima $\mathbf{x}^* \approx \mathbf{x}^{(k)}$.

Matrica $W(\mathbf{x}) = \left[\dfrac{\partial f_i}{\partial x_j}(\mathbf{x})\right]_{n\times n}$ zove se *Jakobijeva matrica* sistema funkcija $f_i(\mathbf{x})$, $i = 1, 2, \ldots, n$.

Ako je $\mathbf{x}^{(0)}$ dovoljno blizu tačnog rešenja $\mathbf{x}^*$, Jakobijeva matrica $W(\mathbf{x})$ regularna za svako $\mathbf{x} \in D$ i ako su drugi parcijalni izvodi $\dfrac{\partial^2 f_i}{\partial x_k \partial x_j}$, $i, j, k = 1, 2, \ldots, n$, ograničeni na $D$, tada iterativni niz $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ konvergira ka $\mathbf{x}^*$ sa redom konvergencije $r \geq 2$.

Da bi se izbeglo zahtevno određivanje inverzne matrice $W(\mathbf{x}^{(k)})^{-1}$ u svakoj iteraciji $k \in \mathbb{N}$, u praksi se često koristi *modifikovani metod Njutn-Kantorovića*

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - W(\mathbf{x}^{(0)})^{-1}\mathbf{f}(\mathbf{x}^{(k)}), \quad k = 0, 1, 2, \ldots, \qquad \mathbf{x}^{(0)} \in D.$$

Ovde se u svakoj iteraciji koristi inverzna matrica $W(\mathbf{x}^{(0)})^{-1}$ i ona može biti određena nekim od numeričkih metoda. Nedostatak ovog metoda je njegova linearna konvergencija, tj. potrebno je više iteracija za dostizanje zadovoljavajuće tačnosti nego kod metoda Njutn-Kantorovića.

# Numerički metodi u linearnoj algebri

## Uvodni pojmovi

Neka je dat sistem linearnih jednačina

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n. \end{aligned} \tag{6}$$

Uvodeći oznake

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

on može biti zapisan u vektorsko-matričnom obliku

$$A\mathbf{x} = \mathbf{b}. \tag{7}$$

Sistem (6) odnosno (7) ima jedinstveno rešenje $\mathbf{x}^* = [x_1\ x_2\ \ldots\ x_n]^T$ ako i samo ako je matrica $A$ regularna (ima inverznu matricu), tj. ako je $\det A \neq 0$.

U nastavku će biti navedeni neki opšti pojmovi i tvrđenja iz teorije matrica koji su potrebni za dalje izlaganje. Iako većina navedenih tvrđenja važi u istom ili nešto izmenjenom obliku i za kompleksne matrice, ovde će biti reči samo o realnim matricama, tj. o matricama iz skupa

$$\mathcal{M}_{n\times n} = \left\{ [a_{ij}]_{n\times n} \;\middle|\; a_{ij} \in \mathbb{R},\ i,j = 1,2,\ldots,n \right\}.$$

**Definicija.** Skalar $\lambda \in \mathbb{R}$ i vektor $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ su *sopstvena vrednost* i *sopstveni vektor* matrice $A \in \mathcal{M}_{n\times n}$ ako važi

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Skup $Sp(A)$ svih sopstvenih vrednosti matrice $A$ zove se *spektar*, a

$$\rho(A) = \max\big\{|\lambda| \;\big|\; \lambda \in Sp(A)\big\}$$

je *spektralni radijus* matrice $A$.

**Definicija.** *Karakteristični polinom* matrice $A \in \mathcal{M}_{n\times n}$ je polinom stepena $n$ definisan sa

$$P(\lambda) = \det(A - \lambda I).$$

**Teorema.** *Sopstvene vrednosti matrice $A$ su nule njenog karakterističnog polinoma.*

Matrica $A \in \mathcal{M}_{n\times n}$ može imati najviše $n$ različitih sopstvenih vrednosti.

**Teorema.** *Ako je $\lambda$ sopstvena vrednost matrice $A$, tada je $\lambda^k$ $(k \in \mathbb{N})$ sopstvena vrednost matrice $A^k$. Ako je $A$ regularna, tada je $1/\lambda$ sopstvena vrednost matrice $A^{-1}$.*

**Definicija.** Matrica $A \in \mathcal{M}_{n\times n}$ je *pozitivno-definitna* (*pozitivno-semidefinitna*) ako za svaki vektor $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ važi

$$\mathbf{x}^T A\mathbf{x} > 0 \qquad (\mathbf{x}^T A\mathbf{x} \geq 0).$$

**Teorema.** *Matrica $A \in \mathcal{M}_{n\times n}$ je pozitivno-definitna (pozitivno-semidefinitna) ako i samo ako su sve njene sopstvene vrednosti pozitivne (nenegativne).*

Primetimo da je za proizvoljnu matricu $A \in \mathcal{M}_{n\times n}$ matrica $A^T A$ pozitivno-semidefinitna, pa su sve njene sopstvene vrednosti nenegativne.

**Definicija.** Neka je $X$ vektorski prostor nad poljem skalara $\mathbb{R}$. Funkcija $\|\cdot\| : X \to \mathbb{R}$ takva da za svako $\mathbf{x}, \mathbf{y} \in X$ i svako $\lambda \in \mathbb{R}$ važi

$$\begin{aligned}
&1^\circ \quad \|\mathbf{x}\| \geq 0, \qquad \|\mathbf{x}\| = 0 \;\Leftrightarrow\; \mathbf{x} = \mathbf{0},\\
&2^\circ \quad \|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|,\\
&3^\circ \quad \|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,
\end{aligned}$$

zove se *norma*, a $(X, \|\cdot\|)$, ili samo $X$, je *normirani prostor.*

U vektorskom prostoru $\mathbb{R}^n$ od posebnog interesa su norme oblika

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \qquad (p \geq 1).$$

Najčešće korišćene norme su:

$$p = 1: \quad \|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| \ ,$$

$$p = 2: \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \ , \qquad \text{(Euklidska norma)}$$

$$p \to \infty: \quad \|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i| \ .$$

Može se pokazati da je $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$ za svaki vektor $\mathbf{x} \in \mathbb{R}^n$.

U vektorskom prostoru $\mathcal{M}_{n\times n}$ takođe mogu da se definišu različite norme. Pri izboru konkretne matrične norme posebno je važan njen odnos prema izabranoj vektorskoj normi.

**Definicija.** Norma matrice $\|A\|$, $A \in \mathcal{M}_{n\times n}$, je *saglasna sa normom vektora* $\|\mathbf{x}\|$, $\mathbf{x} \in \mathbb{R}^n$, ako za proizvoljne $A, B \in \mathcal{M}_{n\times n}$ i $\mathbf{x} \in \mathbb{R}^n$ važi:

$$\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|,$$
$$\|AB\| \leq \|A\|\|B|.$$

Ako je norma $\|A\|$ saglasna sa normom $\|\mathbf{x}\|$ i ako za svaku matricu $A \in \mathcal{M}_{n\times n}$ postoji vektor $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$, takav da je

$$\|A\mathbf{x}\| = \|A\|\|\mathbf{x}\|,$$

tada je norma matrice $\|A\|$ *potčinjena normi vektora* (*indukovana normom vektora*) $\|\mathbf{x}\|$.

Najčešće korišćene norme u prostoru $\mathcal{M}_{n\times n}$ su:

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| \ , \qquad \text{potčinjena normi} \qquad \|\mathbf{x}\|_1,$$

$$\|A\|_2 = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} \ , \qquad \text{saglasna sa normom} \qquad \|\mathbf{x}\|_2,$$

$$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| \ , \qquad \text{potčinjena normi} \qquad \|\mathbf{x}\|_\infty.$$

Norma $\|A\|_2$ (Frobeniusova ili Šmitova norma) je saglasna sa Euklidskom normom vektora $\|\mathbf{x}\|_2$, ali joj nije potčinjena. Matrična norma potčinjena Euklidskoj normi vektora je *spektralna norma* definisana sa

$$\|A\|_{sp} = \sigma(A) = \sqrt{\max|\lambda(A^T A)|},$$

gde su sa $\lambda(A^T A)$ označene sopstvene vrednosti matrice $A^T A$.

**Napomena 1.** Za kompleksne matrice $A = [a_{ij}]_{n\times n}$, $a_{ij} \in \mathbb{C}$, $i, j = 1, 2, \ldots, n$, spektralna norma se definiše kao

$$\|A\|_{sp} = \sigma(A) = \sqrt{\max|\lambda(A^* A)|},$$

gde je $A^* = \overline{A}^T$.

**Napomena 2.** Ako je matrice $A$ realna i simetrična, tj. $A^T = A$, tada je $A^T A = A^2$, pa za sopstvene vrednosti važi $\lambda(A^T A) = \lambda(A)^2$. Zato je spektralna norma takve matrice jednaka njenom spektralnom radijusu:

$$\|A\|_{sp} = \sigma(A) = \max|\lambda(A)| = \rho(A).$$

Isto važi i za hermitske kompleksne matrice, tj. one za koje je $A^* = A$.

**Teorema.** *Neka je $A \in \mathcal{M}_{n\times n}$ proizvoljna matrica i $\rho(A)$ njen spektralni radujus. Ako je $\|A\|$ proizvoljna norma matrice $A$, tada važi:*

$$\rho(A) \le \|A\|.$$

*Ako je $A$ simetrična matrica ($A^T = A$), tada je njena spektralna norma jednaka spektralnom radijusu, tj.*

$$\sigma(A) = \rho(A)$$

*i ona ima najmanju vrednost od svih normi u $\mathcal{M}_{n\times n}$ koje su saglasne sa nekom normom vektora u $\mathbb{R}^n$.*

# Direktni metodi za rešavanje sistema jednačina

Direktni metodi za rešavanje sistema linearnih jednačina su oni kod kojih se posle konačnog broja koraka teorijski dobija tačno rešenje. Ipak, kako se prilikom procesa izračunavanja javlja greška zaokrugljivanja međurezultata, dobijeni konačni rezultat je ograničene tačnosti. Štaviše, grešku koja se pojavljuje teško je kontrolisati i povećava se sa porastom broja računskih operacija. Zato je od posebne važnosti za ove metode analiza grešaka, stabilnost sistema u odnosu na male promene koeficijenata, kao i što manji broj računskih operacija.

Kao predstavnici direktnih metoda ovde će biti predstavljeni Gausov, Gaus-Žordanov metod i neki faktorizacioni metodi.

## Gausov metod

Neka sistem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1,\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2,\\ &\vdots\\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n, \end{aligned}$$

odnosno

$$A\mathbf{x} = \mathbf{b},$$

ima jedinstveno rešenje $\mathbf{x}^* = [x_1\ \ x_2\ \ \ldots\ \ x_n]^T$. Gausov metod se bazira na svođenje kvadratnog sistema primenom elementarnih transformacija na trougaoni oblika

$$\begin{aligned} r_{11}x_1 + r_{12}x_2 + \cdots + r_{1n}x_n &= c_1,\\ r_{22}x_2 + \cdots + r_{2n}x_n &= c_2,\\ &\vdots\\ r_{nn}x_n &= c_n, \end{aligned} \tag{8}$$

tj.,

$$R\mathbf{x} = \mathbf{c}, \qquad \text{gde je} \qquad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n}\\ & r_{22} & & r_{2n}\\ & & \ddots & \vdots\\ & & & r_{nn} \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_1\\ c_2\\ \vdots\\ c_n \end{bmatrix}. \tag{9}$$

Kako polazni sistem ima jedinstveno rešenje $\mathbf{x}^*$, ono je i jedinstveno rešenje njemu ekvivalentnog trougaonog sistema (8), što znači da je $\det R \neq 0$. Matrica $R$ je gornjetrougaona (čiji su svi elementi ispod glavne dijagonale jednaki nuli), pa je

$$\det R = r_{11}r_{22}\cdots r_{nn} \neq 0, \qquad \text{tj.,} \qquad r_{ii} \neq 0, \quad i = 1, 2, \ldots, n.$$

Prema tome, tačno rešenje se dobija sukcesivnim rešavanjem $n$ jednačina sistema (8) sa po jednom nepoznatom.

Formiranje sistema (8) izvodi se u $n-1$ koraka. Označimo $a_{ij} = a_{ij}^{(1)}$, $b_i = b_i^{(1)}$, $i, j = 1, 2, \ldots, n$.

Neka je $a_{11}^{(1)} \neq 0$. Množenjem prve jednačine sa $m_{i1} = \dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}}$, $i = 2, 3, \ldots, n$, i oduzimanjem od $i$-te jednačine redom dobija se ekvivalentan sistem kod koga nepoznata $x_1$ figuriše samo u prvoj jednačini:

$$
\begin{aligned}
a_{11}^{(1)} x_1 + a_{12}^{(1)} x_2 + \cdots + a_{1n}^{(1)} x_n &= b_1^{(1)}, \\
a_{22}^{(2)} x_2 + \cdots + a_{2n}^{(2)} x_n &= b_2^{(2)}, \\
&\vdots \\
a_{n2}^{(2)} x_2 + \cdots + a_{nn}^{(2)} x_n &= b_n^{(2)},
\end{aligned}
$$

gde je

$$
a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1} a_{1j}^{(1)}, \quad b_i^{(2)} = b_i^{(1)} - m_{i1} b_1^{(1)}, \qquad i, j = 2, 3, \ldots, n.
$$

Ako je $a_{22}^{(2)} \neq 0$, ponavljanjem postupka može da se eliminiše nepoznata $x_2$ iz svih jednačina osim prve i druge, pa treća, itd. Posmatrajmo sistem pre $k$-tog koraka:

$$
\begin{aligned}
a_{11}^{(1)} x_1 + a_{12}^{(1)} x_2 + \cdots + a_{1k}^{(1)} x_k + a_{1,k+1}^{(1)} x_{k+1} + \cdots + a_{1n}^{(1)} x_n &= b_1^{(1)}, \\
a_{22}^{(2)} x_2 + \cdots + a_{2k}^{(2)} x_k + a_{2,k+1}^{(2)} x_{k+1} + \cdots + a_{2n}^{(2)} x_n &= b_2^{(2)}, \\
&\vdots \\
a_{kk}^{(k)} x_k + a_{k,k+1}^{(k)} x_{k+1} + \cdots + a_{kn}^{(k)} x_n &= b_k^{(k)}, \\
a_{k+1,k}^{(k)} x_k + a_{k+1,k+1}^{(k)} x_{k+1} + \cdots + a_{k+1,n}^{(k)} x_n &= b_{k+1}^{(k)}, \\
&\vdots \\
a_{nk}^{(k)} x_k + a_{n,k+1}^{(k)} x_{k+1} + \cdots + a_{nn}^{(k)} x_n &= b_n^{(k)}.
\end{aligned}
$$

Ako je $a_{kk}^{(k)} \neq 0$, množenjem $k$-te jednačine eliminacionim faktorom

$$
m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \qquad i = k+1, k+2, \ldots, n, \tag{10}
$$

i oduzimanjem od $i$-te jednačine redom dobija se

$$
\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + \quad a_{1k}^{(1)}x_k + \quad a_{1,k+1}^{(1)}x_{k+1} + \cdots + \quad a_{1n}^{(1)}x_n &= b_1^{(1)},\\
a_{22}^{(2)}x_2 + \cdots + \quad a_{2k}^{(2)}x_k + \quad a_{2,k+1}^{(2)}x_{k+1} + \cdots + \quad a_{2n}^{(2)}x_n &= b_2^{(2)},\\
&\vdots\\
a_{kk}^{(k)}x_k + \quad a_{k,k+1}^{(k)}x_{k+1} + \cdots + \quad a_{kn}^{(k)}x_n &= b_k^{(k)},\\
a_{k+1,k+1}^{(k+1)}x_{k+1} + \cdots + a_{k+1,n}^{(k+1)}x_n &= b_{k+1}^{(k+1)},\\
&\vdots\\
a_{n,k+1}^{(k)}x_{k+1} + \cdots + \quad a_{nn}^{(k+1)}x_n &= b_n^{(k+1)},
\end{aligned}
$$

gde je

$$
a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)}, \quad b_i^{(k+1)} = b_i^{(k)} - m_{ik}b_k^{(k)}, \qquad i,j = k+1,\ldots,n. \tag{11}
$$

Pod uslovom $a_{kk}^{(k)} \neq 0$, $k = 1,2,\ldots,n$, posle $n-1$ koraka dobija se trougaoni sistem

$$
\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + \quad a_{1n}^{(1)}x_n &= b_1^{(1)},\\
a_{22}^{(2)}x_2 + \cdots + a_{2n}^{(2)}x_n &= b_2^{(2)},\\
&\vdots\\
a_{nn}^{(n)}x_n &= b_n^{(n)},
\end{aligned}
$$

a sukcesivnim rešavanjem jednačina od $n$-te ka prvoj dobija se

$$
x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}}, \qquad x_k = \frac{1}{a_{kk}^{(k)}}\Big(b_k^{(k)} - \sum_{i=k+1}^{n} a_{ki}^{(k)}x_i\Big), \quad k = n-1, n-2, \ldots, 2, 1.
$$

Formulama (10)-(11) predstavljen je osnovni Gausov metod za rešavanje sistema linearnih jednačina. Uslov za realizaciju metoda je $a_{kk}^{(k)} \neq 0$, $k = 1,2,\ldots,n$. Ako je matrica sistema $A$ regularna, ovaj uslov se uvek može obezbediti odgovarajućom permutacijom jednačina sistema. Elementi $a_{kk}^{(k)}$ zovu se *glavni elementi* ili *pivoti*.

## Gaus-Žordanov metod

Za razliku od Gausovog metoda koji se zasniva na transformaciji sistema $A\mathbf{x} = \mathbf{b}$ u trougaoni $R\mathbf{x} = \mathbf{c}$, cilj Gaus-Žordanovog metoda je dobijanje ekvivalentnog sistema oblika $D\mathbf{x} = \mathbf{c}$, gde je $D$ dijagonalna matrica, i njegovo jednostavno rešavanje. Postupak eliminacije promenljivih iz jednačina sistema je sličan Gausovom postupku, ali se u svakom koraku vrši eliminacija odgovarajuće promenljive iz svih jednačina sistema, osim iz jedne. U $k$-tom koraku sistem ima oblik

$$\begin{array}{l}
a_{11}^{(1)}x_1 + a_{1k}^{(k)}x_k + a_{1,k+1}^{(k)}x_{k+1} + \cdots + a_{1n}^{(k)}x_n = b_1^{(k)},\\
\qquad a_{22}^{(2)}x_2 + a_{2k}^{(k)}x_k + a_{2,k+1}^{(k)}x_{k+1} + \cdots + a_{2n}^{(k)}x_n = b_2^{(k)},\\
\qquad\qquad\vdots\\
\qquad\qquad a_{kk}^{(k)}x_k + a_{k,k+1}^{(k)}x_{k+1} + \cdots + a_{kn}^{(k)}x_n = b_k^{(k)},\\
\qquad\qquad a_{k+1,k}^{(k)}x_k + a_{k+1,k+1}^{(k)}x_{k+1} + \cdots + a_{k+1,n}^{(k)}x_n = b_{k+1}^{(k)},\\
\qquad\qquad\vdots\\
\qquad\qquad a_{nk}^{(k)}x_k + a_{n,k+1}^{(k)}x_{k+1} + \cdots + a_{nn}^{(k)}x_n = b_n^{(k)}.
\end{array}$$

Ako je $a_{kk}^{(k)} \neq 0$, množenjem $k$-te jednačine eliminacionim faktorom

$$m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \qquad i = 1, 2, \ldots, n,\ i \neq k,$$

i oduzimanjem od $i$-te jednačine redom dobija se

$$\begin{array}{l}
a_{11}^{(1)}x_1 + a_{1,k+1}^{(k+1)}x_{k+1} + \cdots + a_{1n}^{(k+1)}x_n = b_1^{(k+1)},\\
\qquad a_{22}^{(2)}x_2 + a_{2,k+1}^{(k+1)}x_{k+1} + \cdots + a_{2n}^{(k+1)}x_n = b_2^{(k+1)},\\
\qquad\qquad\vdots\\
\qquad\qquad a_{kk}^{(k)}x_k + a_{k,k+1}^{(k+1)}x_{k+1} + \cdots + a_{kn}^{(k+1)}x_n = b_k^{(k+1)},\\
\qquad\qquad a_{k+1,k+1}^{(k+1)}x_{k+1} + \cdots + a_{k+1,n}^{(k+1)}x_n = b_{k+1}^{(k+1)},\\
\qquad\qquad\vdots\\
\qquad\qquad a_{n,k+1}^{(k)}x_{k+1} + \cdots + a_{nn}^{(k+1)}x_n = b_n^{(k+1)},
\end{array}$$

gde je

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)}, \quad b_i^{(k+1)} = b_i^{(k)} - m_{ik}b_k^{(k)}, \quad i,j = 1,2,\ldots,n, \quad i,j \neq k.$$

U $n$-tom koraku dobija se dijagonalni sistem

$$a_{kk}^{(k)}x_k = b_k^{(n)}, \qquad k = 1, 2, \ldots, n,$$

čijim se rešavanjem

$$x_k = \frac{b_k^{(n)}}{a_{kk}^{(k)}}, \qquad k = 1, 2, \ldots, n,$$

dobija rešenje polaznog sistema.

I kod Gaus-Žordanovog metoda uslov za realizaciju je $a_{kk}^{(k)}$, $k = 1, 2, \ldots, n$, što se uvek može obezbediti permutacijama jednačina ili promenljivih ako je $\det A \neq 0$.

Iako je dobijeni dijagonalni sistem jednostavniji za rešavanje od trougaonog, ipak je za primenu Gaus-Žordanovog metoda potreban veći broj operacija nego za primenu Gausovog.

## Faktorizacioni metodi. LR faktorizacija

Faktorizacioni metodi se zasnivaju na predstavljanju matrice sistema u obliku proizvoda nekih matrica pogodnog oblika, a zatim na sukcesivnom rešavanju novih sistema jednačina. Posebno su pogodni kada je matrica sistema retka (ima mnogo elemenata koji su jednaki nuli).

Ovde će biti predstavljen *metod Čoleskog* koji je baziran na tzv. *LR-faktorizaciji* (ili *LU-faktorizaciji*) matrice sistema, tj. predstavljanju matrice sistema u obliku proizvoda jedne donje-trougaone (levo-trougaone) i jedne gornje-trougaone (desno-trougaone) matrice:

$$A = LR, \qquad L = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}, \quad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}.$$

Na taj način, sistem jednačina postaje

$$LR\mathbf{x} = \mathbf{b},$$

pa se može rešavati u dva koraka, rešavanjem dva trougaona sistema

$$L\mathbf{y} = \mathbf{b}, \qquad R\mathbf{x} = \mathbf{y},$$

najpre

$$y_1 = \frac{b_1}{l_{11}}, \qquad y_i = \frac{1}{l_{ii}}\Big(b_i - \sum_{k=1}^{i-1} l_{ik} y_k\Big), \quad i = 2, 3, \ldots, n,$$

a zatim

$$x_n = \frac{y_n}{r_{nn}}, \qquad x_i = \frac{1}{r_{ii}}\Big(y_i - \sum_{i=i+1}^{n} r_{ik}x_k\Big), \quad i = n-1, n-2, \ldots, 2, 1.$$

**Teorema.** *Matrica $A \in \mathcal{M}_{n\times n}$ ima LR-faktorizaciju ako i samo ako su sve podmatrice $A_k$, $k = 1, 2, \ldots, n$, matrice $A$, sastavljene od njenih prvih $k$ vrsta i kolona, regularne.*

Primetimo da su u trougaonim matricama $L$ i $R$ poznati elementi iznad glavne dijagonale ($l_{ij} = 0$ za $j > i$) i ispod glavne dijagonale ($r_{ij} = 0$ za $i > j$) redom. Nepoznatih elemenata $l_{ij}$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, i$ i $r_{ij}$, $i = 1, 2, \ldots, n$, $j = i, i+1, \ldots, n$ ima po $n+(n-1)+\cdots+2+1 = n(n+1)/2$ u svakoj od matrica $L$ i $R$, dakle $n^2+n$ ukupno. Oni se mogu odrediti izjednačavanjem elemenata proizvoda $LR$ sa elementima matrice $A$, kojih ima $n^2$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} l_{11}r_{11} & l_{11}r_{12} & \cdots & l_{11}r_{1n} \\ l_{21}r_{11} & l_{21}r_{12}+l_{22}r_{22} & & l_{21}r_{1n}+l_{22}r_{2n} \\ \vdots & & & \\ l_{n1}r_{11} & l_{n1}r_{12}+l_{n2}r_{22} & & l_{n1}r_{1n}+l_{n2}r_{2n}+\cdots+l_{nn}r_{nn} \end{bmatrix}.$$

Prema tome, kada $LR$-faktorizacija matrice postoji, ona nije jedinstvena. Ako je, međutim, unapred zadato $n$ elemenata, na primer $r_{ii}(\neq 0)$, $i = 1, 2, \ldots, n$, ostali elementi se mogu odrediti sukcesivno. Iz prve kolone dobija se

$$l_{i1} = \frac{a_{i1}}{r_{11}}, \quad i = 1, 2, 3, \ldots, n,$$

a iz prve vrste

$$r_{1j} = \frac{a_{1j}}{l_{11}}, \quad j = 2, 3, \ldots, n.$$

Zatim, korišćenjem druge kolone imamo

$$l_{i2} = \frac{1}{r_{22}}\big(a_{i2} - l_{i1}r_{12}\big), \quad i = 2, 3, \ldots, n,$$

a korišćenjem druge vrste i

$$r_{2j} = \frac{1}{l_{22}}\big(a_{1j} - l_{21}r_{1j}\big), \quad j = 3, \ldots, n.$$

Nastavljajući postupak, dobijaju se najpre elementi $j$-te kolone matrice $L$

$$l_{ij} = \frac{1}{r_{ii}}\big(a_{ij} - \sum_{k=1}^{i-1} l_{ik} r_{kj}\big), \quad i = 2, \dots, n,$$

a zatim i elementi $i$-te vrste matrice $R$

$$r_{ij} = \frac{1}{l_{ii}}\big(a_{ij} - \sum_{k=1}^{i-1} l_{ik} r_{kj}\big), \quad j = 3, \dots, n.$$

Objedinjujući prethodno rečeno dobijamo formule za sukcesivno određivanje elemenata u $LR$-faktorizaciji matrice $A$:

$$\begin{aligned}
&l_{11} = \frac{a_{11}}{r_{11}}, \quad l_{i1} = \frac{a_{i1}}{r_{11}}, \quad r_{1i} = \frac{a_{1i}}{l_{11}},\\
&l_{ii} = \frac{1}{r_{ii}}\big(a_{ii} - \sum_{k=1}^{i-1} l_{ik} r_{ki}\big),\\
&l_{ji} = \frac{1}{r_{ii}}\big(a_{ji} - \sum_{k=1}^{i-1} l_{jk} r_{ki}\big), \quad r_{ij} = \frac{1}{l_{ii}}\big(a_{ij} - \sum_{k=1}^{i-1} l_{ik} r_{kj}\big),\\
&\quad i = 2, 3 \dots, n, \quad j = i+1, i+2, \dots, n.
\end{aligned} \tag{12}$$

U praksi se najčešće uzima $r_{ii} = 1$, $i = 1, 2, \dots, n$, ili $l_{ii} = 1$, $i = 1, 2, \dots, n$.

Za $LR$-faktorizaciju matrice može da se koristi i Gausov algoritam. Setimo se da se Gausovim algoritmom matrica transformiše u gornje-trougaonu matricu, tj.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \qquad \longrightarrow \qquad R = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n)} \end{bmatrix},$$

gde je

$$\begin{aligned}
a_{1j}^{(1)} &= a_{1j}, \qquad j = 1, 2, \dots, n,\\
m_{ik} &= \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \qquad k = 1, 2, \dots, n, \quad i = k+1, \dots, n,\\
a_{ij}^{(k+1)} &= a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}, \qquad k = 1, 2, \dots, n, \quad i, j = k+1, \dots, n.
\end{aligned}$$

Eliminacioni faktori $m_{ik}$, $i = 2, 3, \ldots, n$, $k = 1, 2, \ldots, i-1$, su u tom slučaju nedijagonalni elementi matrice $L$, tj.

$$L = \begin{bmatrix} 1 & & & \\ m_{21} & 1 & & \\ \vdots & & \ddots & \\ m_{n1} & m_{n2} & \cdots & 1 \end{bmatrix}.$$

Zaista, posmatrajući elemente na mestu $(i, j)$ u proizvodu

$$LR = \begin{bmatrix} 1 & & & \\ m_{21} & 1 & & \\ \vdots & & \ddots & \\ m_{n1} & m_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n)} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ m_{21}a_{11}^{(1)} & m_{21}a_{12}^{(1)} + a_{22}^{(2)} & & m_{21}a_{1n}^{(1)} + a_{2n}^{(2)} \\ \vdots & & & \\ m_{n1}a_{11}^{(1)} & m_{n1}a_{12}^{(1)} + m_{n2}a_{22}^{(2)} & & m_{n1}a_{1n}^{(1)} + m_{n2}a_{2n}^{(2)} + \cdots + a_{nn}^{(n)} \end{bmatrix}$$

i imajući u vidu da je

$$m_{ik}a_{kj}^{(k)} = a_{ij}^{(k)} - a_{ij}^{(k+1)}, \quad k = 1, 2, \ldots, n, \quad i, j = k+1, \ldots, n,$$

dobijamo sledeće:

za $i \le j$ :

$$\begin{aligned} (LR)_{ij} &= \sum_{k=1}^{n} l_{ik} r_{kj} = \sum_{k=1}^{i-1} m_{ik} a_{kj}^{(k)} + a_{ij}^{(i)} = \sum_{k=1}^{i-1} \big(a_{ij}^{(k)} - a_{ij}^{(k+1)}\big) + a_{ij}^{(i)} \\ &= a_{ij}^{(1)} - a_{ij}^{(2)} + a_{ij}^{(2)} - a_{ij}^{(3)} + \cdots + a_{ij}^{(i-1)} - a_{ij}^{(i)} + a_{ij}^{(i)} \\ &= a_{ij}^{(1)} = a_{ij}, \end{aligned}$$

za $i > j$ :

$$\begin{aligned} (LR)_{ij} &= \sum_{k=1}^{n} l_{ik} r_{kj} = \sum_{k=1}^{j} m_{ik} a_{kj}^{(k)} = \sum_{k=1}^{j-1} \big(a_{ij}^{(k)} - a_{ij}^{(k+1)}\big) + \frac{a_{ij}^{(j)}}{a_{jj}^{(j)}} a_{jj}^{(j)} \\ &= a_{ij}^{(1)} - a_{ij}^{(2)} + a_{ij}^{(2)} - a_{ij}^{(3)} + \cdots + a_{ij}^{(j-1)} - a_{ij}^{(j)} + a_{ij}^{(j)} \\ &= a_{ij}^{(1)} = a_{ij}. \end{aligned}$$

Predstavljeni faktorizacioni metod može da se primeni na rešavanje sistema jednačina ako je matrica sistema $A$ takva da su joj svi glavni minori različiti od nule, ali se primena može proširiti i na sve sisteme čija je matrica regularna. Kada se za $LR$-faktorizaciju primenjuje Gausov algoritam sa izborom glavnog elementa, onda se dobija faktorizacija matrice $A' = LR$, gde je $A'$ matrica koja je nastala permutovanjem vrsta matrice $A$. Način permutovanja vrsta treba zapamtiti u indeksnom nizu $I = (p_1, p_2, \ldots, p_{n-1})$, pri čemu je $p_k$ redni broj vrste iz koje se uzima glavni element u $k$-tom eliminacionom koraku. Zatim se permutovanjem elemenata vektora slobodnih članova $\mathbf{b}$ na isti način dobija vektor $\mathbf{b}'$ i tek se onda rešavaju dobijeni trougaoni sistemi jednačina. Ako sa $P$ označimo permutacionu matricu koja je nastala permutacijom vrsta jedinične matrice na način opisan indeksnim nizom $I$, tada je $A' = PA$, $\mathbf{b}' = P\mathbf{b}$, a navedeni postupak može da se opiše na sledeći način:

$$\begin{gathered} A\mathbf{x} = \mathbf{b}, \\ PA\mathbf{x} = P\mathbf{b}, \\ LR\mathbf{x} = \mathbf{b}', \\ R\mathbf{x} = \mathbf{y}, \quad L\mathbf{y} = \mathbf{b}'. \end{gathered}$$

Metod Čoleskog posebno je pogodan za rešavanje sistema čija je matrica simetrična i pozitivno-definitna. Tada je $L = R^T$, tj. koristi se faktorizacija $A = R^T R$, a formule (12) dobijaju formu

$$\begin{aligned} & r_{11} = \sqrt{a_{11}}, \quad r_{1i} = \frac{a_{1i}}{r_{11}}, \\ & r_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2}, \quad r_{ij} = \frac{1}{r_{ii}}\big(a_{ij} - \sum_{k=1}^{i-1} r_{ki} r_{kj}\big), \\ & \quad i = 2, 3 \ldots, n, \quad j = i+1, i+2, \ldots, n. \end{aligned}$$

Ova varijanta metoda Čoleskog poznata je i pod imenom *metod kvadratnog korena.*

## Iterativni metodi

Iterativni metodi za rešavanje sistema linearnih jednačina

$$A\mathbf{x} = \mathbf{b}$$

zasnivaju se na formiranju niza vektora $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ koji konvergira ka tačnom rešenju $\mathbf{x}^*$.

**Definicija.** Niz vektora $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$, $\mathbf{x}^{(k)} = [x_1^{(k)} \;\; x_2^{(k)} \;\; \dots \;\; x_n^{(k)}]^T \in \mathbb{R}^n$, konvergira ka vektoru $\mathbf{x}^* = [x_1^* \;\; x_2^* \;\; \dots \;\; x_n^*]^T \in \mathbb{R}^n$, tj.

$$\lim_{n\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*,$$

ako je $\displaystyle\lim_{n\to\infty} x_j^{(k)} = x_j^*, \quad j = 1, 2, \dots, n.$

Niz matrica $\{A^{(k)}\}_{k\in\mathbb{N}}$, $A^{(k)} = \left[a_{ij}^{(k)}\right]_{n\times n} \in \mathcal{M}_{n\times n}$, konvergira ka matrici $A^* = \left[a_{ij}^*\right]_{n\times n} \in \mathcal{M}_{n\times n}$, tj.

$$\lim_{n\to\infty} A^{(k)} = A^*,$$

ako je $\displaystyle\lim_{n\to\infty} a_{ij}^{(k)} = a_{ij}^*, \quad i, j = 1, 2, \dots, n.$

Konvergencija nizova vektora i matrica može se definisati i korišćenjem norme.

**Definicija.** Niz vektora $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ konvergira ka vektoru $\mathbf{x}^* = [x_1^* \; x_2^* \; \dots \; x_n^*]^T \in \mathbb{R}^n$ ako je

$$\lim_{n\to\infty} \|\mathbf{x}^{(k)} - \mathbf{x}^*\| = 0,$$

gde je $\|\cdot\|$ proizvoljno izabrana norma u $\mathbb{R}^n$.

Niz matrica $\{A^{(k)}\}_{k\in\mathbb{N}}$ konvergira ka matrici $A^* = \left[a_{ij}^*\right]_{n\times n} \in \mathcal{M}_{n\times n}$ ako je

$$\lim_{n\to\infty} \|A^{(k)} - A^*\| = 0,$$

gde je $\|\cdot\|$ proizvoljno izabrana norma u $\mathcal{M}_{n\times n}$.

Može se pokazati da su navedene definicije ekvivalentne.

**Teorema.** *Neka je $A \in \mathcal{M}_{n\times n}$. Ako je $\|A\| < 1$, gde je $\|\cdot\|$ proizvoljno izabrana norma u $\mathcal{M}_{n\times n}$, tada je*

$$\lim_{k\to\infty} A^k = O,$$

*gde je $O = [\,0\,]_{n\times n}$.*

**Teorema.** *Neka je $A \in \mathcal{M}_{n\times n}$. Tada važi:*

$$\lim_{k\to\infty} A^k = O \qquad \textit{ako i samo ako} \qquad \rho(A) < 1,$$

*gde je* $O = [\,0\,]_{n\times n}$ *i* $\rho(A)$ *spektralni radijus matrice* $A$.

**Teorema.** *Neka je* $A \in \mathcal{M}_{n\times n}$. *Matrični red* $\sum_{k=0}^{\infty} A^k$ *konvergira ako i samo ako je* $\lim_{k\to\infty} A^k = O$. *Tada važi*

$$\sum_{k=0}^{\infty} A^k = I + A + A^2 + \cdots = (I-A)^{-1}. \tag{13}$$

Iterativni niz vektora $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ koji konvergira ka tačnom rešenju $\mathbf{x}^*$ sistema linearnih jednačina $A\mathbf{x} = \mathbf{b}$ može se formirati rekurzivno

$$\mathbf{x}^{(k)} = F_k(\mathbf{x}^{(k-1)}), \quad k = 1, 2, \ldots,$$

gde je $\mathbf{x}^{(0)}$ početna vrednost. Proces se zaustavlja kad je $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| < \varepsilon$, gde je $\varepsilon > 0$ unapred zadata tačnost i $\|\cdot\|$ izabrana norma u $\mathbb{R}^n$, a približno rešenje sistema jednačina je $\mathbf{x}^{(k)}$.

Ako iterativna funkcija $F_k$ ne zavisi od $k$ metod je starcionaran.

## Metod proste iteracije

Najjednostavniji stacionarni iterativni metod za rešavanje sistema linearnih jednačina je *metod proste iteracije* ili *opšti iterativni metod*. Ako se polazni sistem predstavi u obliku

$$\mathbf{x} = B\mathbf{x} + \beta, \tag{14}$$

tada se niz formira rekurzivno kao

$$\mathbf{x}^{(k)} = B\mathbf{x}^{(k-1)} + \beta, \quad k = 1, 2, \ldots,$$

gde je $\mathbf{x}^{(0)} \in \mathbb{R}^n$ početna vrednost. U praksi se najčešće uzima $\mathbf{x}^{(0)} = \beta$. U skalarnom obliku, metod proste iteracije predstavlja se na sledeći način:

$$\begin{aligned}
x_1^{(k)} &= b_{11}x_1^{(k-1)} + b_{12}x_2^{(k-1)} + \cdots + b_{1n}x_n^{(k-1)} + \beta_1,\\
x_2^{(k)} &= b_{21}x_1^{(k-1)} + b_{22}x_2^{(k-1)} + \cdots + b_{2n}x_n^{(k-1)} + \beta_2,\\
&\ \vdots\\
x_n^{(k)} &= b_{n1}x_1^{(k-1)} + b_{n2}x_2^{(k-1)} + \cdots + b_{nn}x_n^{(k-1)} + \beta_n.
\end{aligned}$$

Polazni sistem jednačina $A\mathbf{x} = \mathbf{b}$ može na beskonačno mnogo načina da se predstavi u obliku (14) pogodnom za formiranje iterativnog niza. Pitanje konvergencije metoda, tj. iterativnog niza $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ je ključno za njegovu primenu. Da li metod konvergira zavisi od iterativne matrice $B$.

**Teorema.** *Metod proste iteracije konvergira ka tačnom rešenju sistema (14) ako je $\|B\| < 1$, gde je $\|\cdot\|$ proizvoljno izabrana norma u $\mathcal{M}_{n\times n}$ saglasna sa odgovarajućom normom vektora u $\mathbb{R}^n$.*

*Dokaz.* Za proizvoljno izabrani početni vektor $\mathbf{x}^{(0)} \in \mathbb{R}^n$ važi

$$\begin{aligned}
\mathbf{x}^{(1)} &= B\mathbf{x}^{(0)} + \beta,\\
\mathbf{x}^{(2)} &= B\mathbf{x}^{(1)} + \beta = B\big(B\mathbf{x}^{(0)} + \beta\big) + \beta = B^2\mathbf{x}^{(0)} + (I + B)\beta,\\
&\ \vdots\\
\mathbf{x}^{(k)} &= B^k\mathbf{x}^{(0)} + (I + B + \cdots + B^{k-1})\beta, \quad k = 3, 4, \ldots.
\end{aligned}$$

S obzirom na uslov $\|B\| < 1$, na osnovu (13) važi

$$\lim_{k\to\infty} B^k = O \qquad \text{i} \qquad \sum_{k=0}^{\infty} B^k = I + B + B^2 + \cdots = (I - B)^{-1},$$

pa se u graničnom procesu dobija

$$\begin{aligned}
\mathbf{x}^* = \lim_{k\to\infty} \mathbf{x}^{(k)} &= \lim_{k\to\infty} \Big(B^k\mathbf{x}^{(0)} + \big(I + B + \cdots + B^{k-1}\big)\beta\Big)\\
&= \lim_{k\to\infty} B^k\mathbf{x}^{(0)} + \big(I + B + B^2 + \cdots\big)\beta = (I - B)^{-1}\beta.
\end{aligned}$$

Dobijeni vektor $\mathbf{x}^*$ je tačno rešenje sistema, jer važi

$$\mathbf{x}^* = (I - B)^{-1}\beta \quad \Rightarrow \quad (I - B)\mathbf{x}^* = \beta \quad \Rightarrow \quad \mathbf{x}^* = B\mathbf{x}^* + \beta. \quad \square$$

Procena greške metoda i potrebnog broja iteracija za postizanje unapred zadate tačnosti data je sledećom teoremom.

**Teorema.** *Ako je $\|B\| < 1$, gde je $\|\cdot\|$ proizvoljno izabrana norma u $\mathcal{M}_{n\times n}$ saglasna sa odgovarajućom normom vektora u $\mathbb{R}$, za proizvoljno izabrani početni vektor $\mathbf{x}^{(0)} \in \mathbb{R}^n$ i proizvoljno $k \in \mathbb{N}$ važi:*

$$\begin{aligned}
\|\mathbf{x}^{(k)} - \mathbf{x}^*\| &\le \|B\|\,\|\mathbf{x}^{(k-1)} - \mathbf{x}^*\|,\\
\|\mathbf{x}^{(k)} - \mathbf{x}^*\| &\le \|B\|^k\,\|\mathbf{x}^{(0)} - \mathbf{x}^*\|,\\
\|\mathbf{x}^{(k)} - \mathbf{x}^*\| &\le \|B\|^k\,\|\mathbf{x}^{(0)}\| + \frac{\|B\|^k\,\|\beta\|}{1 - \|B\|}.
\end{aligned}$$

*Specijalno, ako je* $\mathbf{x}^{(0)} = \beta$*, važi*

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \le \frac{\|B\|^{k+1}\,\|\beta\|}{1 - \|B\|}.$$

*Dokaz.* Prva nejednakost dobija se iz saglasnosti normi matrica i vektora:

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| = \|B\mathbf{x}^{(k-1)} + \beta - B\mathbf{x}^* - \beta\| \le \|B\|\,\|\mathbf{x}^{(k-1)} - \mathbf{x}^*\|,$$

a druga sukcesivnom primenom prve nejednakosti. Za dokaz treće nejednakosti koriste se izrazi za $\mathbf{x}^{(k)}$ i $\mathbf{x}^*$ iz dokaza prethodne teoreme:

$$\begin{aligned}
\|\mathbf{x}^{(k)} - \mathbf{x}^*\| &= \|B^k\,\mathbf{x}^{(0)} + (I + B + \cdots + B^{k-1})\beta - (I - B)^{-1}\beta\| \\
&\le \|B\|^k\,\|\mathbf{x}^{(0)}\| + \|(I + B + \cdots + B^{k-1}) - (I + B + \cdots)\|\|\beta\| \\
&\le \|B\|^k\,\|\mathbf{x}^{(0)}\| + \|B\|^k\,\big(\|I\| + \|B\| + \|B\|^2 + \cdots\big)\,\|\beta\| \\
&\le \|B\|^k\,\|\mathbf{x}^{(0)}\| + \frac{\|B\|^k\,\|\beta\|}{1 - \|B\|}.
\end{aligned}$$

Specijalno, ako je $\mathbf{x}^{(0)} = \beta$, dobija se

$$\begin{aligned}
\|\mathbf{x}^{(k)} - \mathbf{x}^*\| &= \|B^k\,\beta + (I + B + \cdots + B^{k-1})\beta - (I - B)^{-1}\beta\| \\
&= \|(I + B + \cdots + B^k) - (I + B + \cdots)\|\|\beta\| \\
&\le \|B\|^{k+1}\|\beta\|\,\big(\|I\| + \|B\| + \|B\|^2 + \cdots\big) \\
&= \frac{\|B\|^{k+1}\,\|\beta\|}{1 - \|B\|}. \quad \square
\end{aligned}$$

Pod uslovima prethodne teoreme može se izvesti i vrlo korisna i relacija

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \le \frac{\|B\|}{1 - \|B\|}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|,$$

koja omogućava procenu greške na osnovu razlike između dveju uzastopnih iteracija.

Iz prethodne teoreme može se uočiti da je konvergencija metoda proste iteracije linearna, a dovoljan uslov za konvergenciju je da proizvoljna norma iterativne matrice bude manja od 1. Imajući u vidu da spektralni radijus matrice nije veći od ni od koje njene norme, može da se formuliše potreban i dovoljan uslov za konvergenciju.

**Teorema.** *Metod proste iteracije konvergira ka tačnom rešenju sistema (14) ako i samo ako su sve sopstvene vrednosti matrice* $B$ *po modulu manje od* 1.

Drugim rečima, da bi metod proste iteracije

$$\mathbf{x}^{(k)} = B\mathbf{x}^{(k-1)} + \beta, \quad k = 1, 2, \ldots,$$

konvergirao, potrebno je i dovoljno da za sve nule $\lambda_i(B)$, $i = 1, 2, \ldots, n$, karakterističnog polinoma matrice $B$

$$P(\lambda) = \det(B - \lambda I) = \begin{vmatrix} b_{11} - \lambda & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} - \lambda & & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & & b_{nn} - \lambda \end{vmatrix}$$

važi $|\lambda_i(B)| < 1$.

Pošto određivanje sopstvenih vrednosti matrice može da bude složen problem, u praksi se za proveru konvergencije metoda češće primenjuje uslov koji koristi neku od normi matrica.

## Jakobijev metod

*Jakobijev metod* za rešavanje sistema jednačina

$$A\mathbf{x} = \mathbf{b}$$

je specijalni slučaj metoda proste iteracije u kome se iterativna matrica dobija izdvajanjem dijagonalnih elemenata matrice $A$. Neka je

$$D = \operatorname{diag} A = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & 0 \\ 0 & & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

regularna matrica, tj $a_{kk} \neq 0$, $k = 1, 2, \ldots, n$. Tada je $A = D - (D - A)$, pa sistem može da se predstavi u obliku

$$\begin{aligned} A\mathbf{x} &= \mathbf{b}, \\ D\mathbf{x} - (D - A)\mathbf{x} &= \mathbf{b}, \\ D\mathbf{x} &= (D - A)\mathbf{x} + \mathbf{b}, \end{aligned}$$

tj.,

$$\mathbf{x} = D^{-1}(D-A)\mathbf{x} + D^{-1}\mathbf{b}.$$

Odgovarajući iterativni metod za rešavanje sistema jednačina je poznat pod imenom *Jakobijev metod*:

$$\mathbf{x}^{(k)} = D^{-1}(D-A)\mathbf{x}^{(k-1)} + D^{-1}\mathbf{b}, \qquad k = 1, 2, \ldots,$$

pri čemu je $\mathbf{x}^{(0)}$ proizvoljan početni vektor, najčešće $\mathbf{x}^{(0)} = D^{-1}\mathbf{b}$. Kako je

$$D^{-1} = \begin{bmatrix} \dfrac{1}{a_{11}} & & & \\ & \dfrac{1}{a_{22}} & & 0 \\ 0 & & \ddots & \\ & & & \dfrac{1}{a_{nn}} \end{bmatrix}, \; D - A = \begin{bmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & 0 & & -a_{2n} \\ \vdots & & & \\ -a_{n1} & -a_{n2} & & 0 \end{bmatrix},$$

skalarni oblik Jakobijevog metoda je

$$\begin{aligned} x_1^{(k)} &= \phantom{-\frac{a_{21}}{a_{11}}x_1^{(k-1)}} - \frac{a_{12}}{a_{11}}x_2^{(k-1)} - \cdots - \frac{a_{1n}}{a_{11}}x_n^{(k-1)} + \frac{b_1}{a_{11}}, \\ x_2^{(k)} &= -\frac{a_{21}}{a_{11}}x_1^{(k-1)} \phantom{- \frac{a_{12}}{a_{11}}x_2^{(k-1)}} - \cdots - \frac{a_{2n}}{a_{22}}x_n^{(k-1)} + \frac{b_2}{a_{22}}, \\ &\vdots \\ x_n^{(k)} &= -\frac{a_{n1}}{a_{nn}}x_1^{(k-1)} - \frac{a_{n2}}{a_{nn}}x_2^{(k-1)} - \cdots \phantom{- \frac{a_{1n}}{a_{11}}x_n^{(k-1)}} + \frac{b_n}{a_{nn}}. \end{aligned}$$

Kao i svaki metod proste iteracije, Jakobijev metod konvergira ako je bar neka norma iterativne matrice manja od 1. Potreban i dovoljan uslov konvergencije je da sve sopstvene vrednosti iterativne matrice budu po modulu manje od 1. Kako je iterativna matrica Jakobijevog metoda $B = D^{-1}(D-A)$, a njen karakteristični polinom je

$$\begin{aligned} P(\lambda) &= \det(B - \lambda I) = \det\left(D^{-1}(D-A) - \lambda I\right) \\ &= \det\left(D^{-1}\left(D - A - \lambda D\right)\right) = (-1)^n \det\left(D^{-1}\right) \det\left(\lambda D + (A - D)\right), \end{aligned}$$

sopstvene vrednosti, kao nule karakterističnog polinoma, rešenja su jednačine

$$\det\left(\lambda D + (A-D)\right) = 0.$$

Imajući u vidu da je

$$\lambda D + (A - D) = \begin{bmatrix} \lambda a_{11} & & & \\ & \lambda a_{22} & & 0 \\ 0 & & \ddots & \\ & & & \lambda a_{nn} \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & & 0 \end{bmatrix},$$

može da se formuliše potreban i dovoljan uslov konvergencije Jakobijevog metoda.

**Teorema.** *Jakobijev metod konvergira ka tačnom rešenju sistema (14) ako i samo ako za sve nule $\lambda_i$, $i = 1, 2, \ldots, n$, jednačine*

$$\begin{vmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \lambda a_{22} & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & & \lambda a_{nn} \end{vmatrix} = 0$$

*važi $|\lambda_i| < 1$.*

## Gaus-Zajdelov metod i varijanta Nekrasova

U primeni metoda proste iteracije

$$\begin{aligned} x_1^{(k)} &= b_{11}x_1^{(k-1)} + b_{12}x_2^{(k-1)} + \cdots + b_{1n}x_n^{(k-1)} + \beta_1, \\ x_2^{(k)} &= b_{21}x_1^{(k-1)} + b_{22}x_2^{(k-1)} + \cdots + b_{2n}x_n^{(k-1)} + \beta_2, \\ &\vdots \\ x_n^{(k)} &= b_{n1}x_1^{(k-1)} + b_{n2}x_2^{(k-1)} + \cdots + b_{nn}x_n^{(k-1)} + \beta_n \end{aligned}$$

u $k$-toj iteraciji se sukcesivno izračunavaju koordinate vektora $\mathbf{x}^{(k)}$. Ako se, međutim, pri određivanju koordinate $x_j^{(k)}$ u $k$-toj iteraciji koristi sve što je do tada izračunato, pa i sve već određene vrednosti $x_1^{(k)}, x_2^{(k)}, \ldots, x_{j-1}^{(k)}$, dobija se novi iterativni metod poznat pod imenom *Gaus-Zajdelovim metod*:

$$\begin{aligned} x_1^{(k)} &= b_{11}x_1^{(k-1)} + b_{12}x_2^{(k-1)} + \cdots + b_{1n}x_n^{(k-1)} + \beta_1, \\ x_2^{(k)} &= b_{21}x_1^{(k)} + b_{22}x_2^{(k-1)} + \cdots + b_{2n}x_n^{(k-1)} + \beta_2, \\ &\vdots \\ x_n^{(k)} &= b_{n1}x_1^{(k)} + b_{n2}x_2^{(k)} + \cdots + b_{nn}x_n^{(k-1)} + \beta_n. \end{aligned}$$

Za dobijanje iterativne matrice uočimo matrični zapis gornjih jednakosti:

$$\mathbf{x}^{(k)} = B_1\mathbf{x}^{(k)} + B_2\mathbf{x}^{(k-1)} + \beta,$$

pri čemu je

$$B_1 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ b_{21} & 0 & & 0 \\ \vdots & & & \\ b_{n1} & b_{n2} & & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & & b_{2n} \\ \vdots & & & \\ 0 & 0 & & b_{nn} \end{bmatrix}.$$

Zato je

$$\begin{aligned} (I - B_1)\mathbf{x}^{(k)} &= B_2\mathbf{x}^{(k-1)} + \beta, \\ \mathbf{x}^{(k)} &= (I - B_1)^{-1}B_2\mathbf{x}^{(k-1)} + (I - B_1)^{-1}\beta, \end{aligned}$$

pa je iterativna matrica $B' = (I - B_1)^{-1}B_2$, a njen karakteristični polinom

$$\begin{aligned} P(\lambda) &= \det(B' - \lambda I) = \det\big((I - B_1)^{-1}B_2 - \lambda I\big) \\ &= \det\big((I - B_1)^{-1}\left(B_2 - \lambda(I - B_1)\right)\big) \\ &= \det\big((I - B_1)^{-1}\big)\det\left(\lambda B_1 + B_2 - \lambda I\right). \end{aligned}$$

Prema tome, uslovi konvergencije Gaus-Zajdelovog metoda dati su sledećom teoremom.

**Teorema.** *Gaus-Zajdelov metod konvergira ka tačnom rešenju sistema (14) ako i samo ako za sve nule $\lambda_i$, $i = 1, 2, \ldots, n$, jednačine*

$$\begin{vmatrix} b_{11} - \lambda & b_{12} & \cdots & b_{1n} \\ \lambda b_{21} & b_{22} - \lambda & & b_{2n} \\ \vdots & & & \\ \lambda b_{n1} & \lambda b_{n2} & & b_{nn} - \lambda \end{vmatrix} = 0$$

*važi $|\lambda_i| < 1$.*

Istim pristupom na primenu Jakobijevog metoda dobija se *Gaus-Zajdelov metod, varijanta Nekrasova*:

$$\begin{aligned} x_1^{(k)} &= \phantom{-\frac{a_{21}}{a_{11}}x_1^{(k)}} - \frac{a_{12}}{a_{11}}x_2^{(k-1)} - \cdots - \frac{a_{1n}}{a_{11}}x_n^{(k-1)} + \frac{b_1}{a_{11}}, \\ x_2^{(k)} &= -\frac{a_{21}}{a_{11}}x_1^{(k)} \qquad\qquad - \cdots - \frac{a_{2n}}{a_{22}}x_n^{(k-1)} + \frac{b_2}{a_{22}}, \\ &\vdots \\ x_n^{(k)} &= -\frac{a_{n1}}{a_{nn}}x_1^{(k)} - \frac{a_{n2}}{a_{nn}}x_2^{(k)} - \cdots \qquad\qquad + \frac{b_n}{a_{nn}}. \end{aligned}$$

Predstavljanjem gornjih jednakosti u matričnom obliku, na sličan način kao u prethodnim slučajevima može da se odredi potreban i dovoljan uslov za konvergenciju metoda.

**Teorema.** *Gaus-Zajdelov metod, varijanta Nekrasova, konvergira ka tačnom rešenju sistema (14) ako i samo ako za sve nule* $\lambda_i$, $i = 1, 2, \ldots, n$, *jednačine*

$$\begin{vmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & & a_{2n} \\ \vdots & & & \\ \lambda a_{n1} & \lambda a_{n2} & & \lambda a_{nn} \end{vmatrix} = 0$$

*važi* $|\lambda_i| < 1$.

## Stabilnost sistema linearnih jednačina

Nezavisno od metoda koji se primenjuje za rešavanje sistema linearnih jednačina, tačnost dobijenog rešenja umnogome zavisi od osetljivosti samog sistema na male promene koeficijenata.

**Primer.** Sistem linearnih jednačina

$$\begin{aligned} x + 2y &= 4, \\ 2x + 4.01y &= 8.01 \end{aligned}$$

ima jedinstveno rešenje $(x, y) = (2, 1)$. Ako se u drugoj jednačini u slobodnom članu greškom izostavi 0 i umesto polaznog rešava sistem

$$\begin{aligned} x + 2y &= 4, \\ 2x + 4.01y &= 8.1, \end{aligned}$$

rešenje postaje $(x, y) = (-16, 10)$. U ovom slučaju mala promena koeficijenata $(< 0.1)$ drastično utiče na rešenje sistema.

Sistemi čija su rešenja nestabilna u odnosu na male promene koeficijenata zovu se *slabo uslovljeni sistemi* (*ill–conditioned systems*). Koliko je sistem stabilan pokazuje *faktor uslovljenosti sistema* (*condition number*):

$$k(A) = \|A\|\|A^{-1}\|,$$

gde je $A$ matrica sistema. Faktor uslovljenosti zavisi od izbora norme, ali je uvek $k(A) \geq 1$. Zaista, imajući u vidu osobine matričnih normi, važi:

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| = k(A).$$

Ukoliko je $k(A)$ mnogo veći od 1, sistem je slabo uslovljen.

U navedenom primeru matrica sistema je regularna, ali je njena determinanta jednaka 0.01, dakle vrlo bliska 0. To je prvi znak slabe uslovljenosti. Ona se može potvrditi i izračunavanjem faktora uslovljenosti. Uzimajući za normu matrice $\|A\|_\infty = \max_i \sum_j |a_{ij}|$, važi:

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 \\ 2 & 4.01 \end{bmatrix}, \qquad && \|A\|_\infty = \max\{1+2, 2+4.01\} = 6.01, \\ A^{-1} &= \begin{bmatrix} 401 & -200 \\ -200 & 100 \end{bmatrix}, && \|A\|_\infty = \max\{401+200, 200+100\} = 601, \end{aligned}$$

pa je faktor uslovljenosti $k(A) = 3612 \gg 1$.

## Inverzija matrica, direktni i iterativni metodi

Ako je $A$ regularna matrica, njena inverzna matrica $A^{-1}$ je ona za koju važi

$$AA^{-1} = A^{-1}A = I.$$

U tom smislu, inverzna matrica $A^{-1}$ matrice $A$ može da se shvati kao rešenje matrične jednačine

$$AX = I.$$

Ako se njene kolone i kolone jedinične matrice označe sa $\mathbf{x}_j$ i $\mathbf{e}_j$, $j = 1, 2, \ldots, n$, redom, tada matrična jednačina dobija oblik

$$\begin{aligned} A[\mathbf{x}_1 \;\; \mathbf{x}_2 \;\; \cdots \;\; \mathbf{x}_n] &= [\mathbf{e}_1 \;\; \mathbf{e}_2 \;\; \cdots \;\; \mathbf{e}_n], \\ [A\mathbf{x}_1 \;\; A\mathbf{x}_2 \;\; \cdots \;\; A\mathbf{x}_n] &= [\mathbf{e}_1 \;\; \mathbf{e}_2 \;\; \cdots \;\; \mathbf{e}_n], \end{aligned}$$

što znači da je svaki od vektora-kolona $\mathbf{x}_j$ matrice $A^{-1}$ rešenje sistema jednačina

$$A\mathbf{x} = \mathbf{e}_j, \qquad j = 1, 2, \ldots, n.$$

Stoga, problem određivanja inverzne matrice svodi se na rešavanje $n$ sistema linearnih jednačina koji imaju zajedničku matricu, a razlikuju se samo po vektorima slobodnih članova.

Direktni metodi za rešavanje sistema linearnih jednačina pogodni su za primenu u određivanju inverzne matrice, jer se transformacije u cilju eliminacije promenljivih izvršavaju odjednom. Tako, primenom Gausovog algoritma vektori-kolone inverzne matrice dobijaju se rešavanjem $n$ trougaonih sistema linearnih jednačina oblika

$$R\mathbf{x} = \mathbf{e}'_j, \quad j = 1, 2, \ldots, n,$$

gde je $R$ gornjetrougaona matrica dobijena transformacijama matrice $A$, a $\mathbf{e}'_j$ vektori dobijeni odgovarajućim transformacijama vektora $\mathbf{e}_j$, $j = 1, 2, \ldots, n$.

Pri određivanju inverzne matrice Gaus-Žordanovim metodom izbegava se formalno rešavanje sistema jednačina. Zaista, odgovarajućim trasformacijama matrica $A$ dobija oblik dijagonalne matrice, koja se jednostavno svodi na jediničnu. Tada se iz svakog od $n$ dobijenih sistema jednačina direktno čita po jedna kolona matrice $A^{-1}$. Ovaj algoritam šematski može da se predstavi na sledeći način:

$$
\begin{aligned}
[\,A \mid I\,] &= \left[\begin{array}{cccc|cccc}
a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\
a_{21} & a_{22} & & a_{2n} & 0 & 1 & & 0 \\
\vdots & & & & & & & \\
a_{n1} & a_{n2} & & a_{nn} & 0 & 0 & & 1
\end{array}\right] \\
&\longrightarrow \left[\begin{array}{cccc|cccc}
a_{11}^{(1)} & 0 & \cdots & 0 & e'_{11} & e'_{12} & \cdots & e'_{1n} \\
0 & a_{22}^{(2)} & & 0 & e'_{21} & e'_{22} & & e'_{2n} \\
\vdots & & & & & & & \\
0 & 0 & & a_{nn}^{(n)} & e'_{n1} & e'_{n2} & & e'_{nn}
\end{array}\right] \\
&\longrightarrow \left[\begin{array}{cccc|cccc}
1 & 0 & \cdots & 0 & c_{11} & c_{12} & \cdots & c_{1n} \\
0 & 1 & & 0 & c_{21} & c_{22} & & c_{2n} \\
\vdots & & & & & & & \\
0 & 0 & & 1 & c_{n1} & c_{n2} & & c_{nn}
\end{array}\right] = [\,I \mid A^{-1}\,]\,.
\end{aligned}
$$

Za inverziju matrica većih dimenzija posebno su pogodni iterativni metodi. Ovde će biti predstavljena klasa metoda sa proizvoljnim redom konvergencije, sa posebnim naglaskom na metod sa kvadratnom konvergencijom.

Neka je $X_0$ neka matrica takva da je $\|F_0\| < 1$, gde je $F_0 = I - AX_0$. Tada je

$$
\begin{aligned}
AX_0 &= I - F_0, \\
X_0 &= A^{-1}(I - F_0), \\
A^{-1} &= X_0(I - F_0)^{-1}.
\end{aligned}
$$

Na osnovu relacije (13) važi

$$(I - F_0)^{-1} = I + F_0 + F_0^2 + \cdots + F_0^{r-1} + F_0^r + \cdots\,,$$

pa je

$$A^{-1} \approx X_1 = X_0(I + F_0 + F_0^2 + F_0^3 + \cdots + F_0^{r-1}).$$

Ovo razmatranje sugeriše formiranje cele klase iterativnih procesa sa redom konvergencije $r$, $r \in \mathbb{N}$:

$$X_{k+1} = X_k(I+F_k+F_k^2+F_k^3+\cdots+F_k^{r-1}), \quad F_k = I-AX_k, \qquad k = 0,1,2,\ldots,$$

gde je $X_0$ matrica takva da je $\|I - AX_0\| < 1$.

Specijalno, za $r = 2$, dobija se *Šulcov metod*

$$X_{k+1} = X_k(I + F_k) = X_k(I + I - AX_k) = X_k(2I - AX_k).$$

**Teorema.** *Neka je $X_0$ matrica takva da je $\|I - AX_0\| \le q < 1$. Tada niz matrica $\{X_k\}_{k\in\mathbb{N}}$ definisan sa*

$$X_{k+1} = X_k(2I - AX_k), \qquad k = 0,1,2,\ldots,$$

*konvergira i važi*

$$\lim_{k\to\infty} X_k = A^{-1}, \qquad \|X_k - A^{-1}\| \le \|I - AX_0\|^{2^k} \frac{\|X_0\|}{1 - \|I - AX_0\|}.$$

Kad konvergira, Šulcov metod konvergira sa redom konvergencije 2.

# Aproksimacija funkcija

## Uvodni pojmovi

Opšti problem aproksimacije funkcija može da se formuliše na sledeći način: za zadatu funkciju $f(x)$ definisanu na diskretnom ili kontinualnom skupu tačaka $X \in \mathbb{R}$ treba odrediti funkciju $\Phi(x) = \Phi(x; a_0, a_1, \ldots, a_n)$, tj. parametre $a_0, a_1, \ldots, a_n \in \mathbb{R}$, tako da $\Phi(x)$ može da zameni funkciju $f(x)$ po izvesnim kriterijumima. Ovde će biti rešavan problem u kome je $\Phi(x)$ linearna po parametrima, tj.

$$\Phi(x) = a_0\Phi_0(x) + a_1\Phi_1(x) + \cdots + a_n\Phi_n(x),$$

gde je $\{\Phi_0(x), \Phi_1(x), \ldots, \Phi_n(x)\}$ unapred zadati sistem funkcija. Funkcija $\Phi(x)$ se zove *aproksimaciona funkcija.*

Na osnovu izbora sistema funkcija $\{\Phi_0(x), \Phi_1(x), \ldots, \Phi_n(x)\}$ razlikuju se, na primer, polinomske, trigonometrijske, eksponencijalne ili druge aproksimacije. Najširu praktičnu primenu imaju polinomske aproksimacije

$$\Phi(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n.$$

Razlog tome su, između ostalog, dobra aproksimativna svojstva, o čemu govori u literaturi poznata Vajerštrasova teorema.

**Teorema.** *Ako je $f(x)$ funkcija neprekidna na $[a, b]$, tada za svako $\varepsilon > 0$ postoji $n \in \mathbb{N}$ i polinom $P_n(x)$ stepena $n$ tako da za svako $x \in [a, b]$ važi $|f(x) - P_n(x)| < \varepsilon$.*

U zavisnosti od kriterijuma za izbor parametara u aproksimacionoj funkciji razlikuju se vrste aproksimacije.

Ako se parametri $a_0, a_1, \ldots, a_n \in \mathbb{R}$ određuju iz uslova da su vrednosti aproksimacione funkcije $\Phi(x)$ jednake vrednostima funkcije $f(x)$ u zadatim tačkama $x_0, x_1, \ldots, x_n \in X$, u pitanju je *interpolacija.*

Ako je kriterijum za izbor parametara minimizacija greške (odstupanja) pri zameni funkcije $f(x)$ aproksimacionom funkcijom $\Phi(x)$ na skupu $X$, tada se radi o *najboljoj aproksimaciji*. Veličina greške aproksimacije $\delta(x) = f(x) - \Phi(x)$ meri se nekom normom u prostoru funkcija, a u zavisnosti od izbora norme dodatno se razlikuju vrste najboljih aproksimacija.

## Interpolacija funkcija

Neka je funkcija $f(x)$ definisana na skupu $X \subseteq [a, b]$ i neka su zadate tačke $x_0, x_1, \ldots, x_n \in X$ u kojima funkcija ima vrednosti $f(x_i) = f_i$, $i = 0, 1, \ldots, n$. Postupak zamene funkcije $f(x)$ funkcijom

$$\Phi(x) = a_0\Phi_0(x) + a_1\Phi_1(x) + \cdots + a_n\Phi_n(x),$$

tako da su ispunjeni uslovi

$$\Phi(x_i) = f_i, \qquad i = 0, 1, \ldots, n.$$

zove se *interpolacija*. Funkcija $\Phi(x)$ je *interpolaciona funkcija* funkcije $f(x)$, a tačke $x_i$, $i = 0, 1, \ldots, n$ su *čvorovi interpolacije*. Koeficijenti $a_0, a_1, \ldots, a_n \in \mathbb{R}$ mogu da se dobiju rešavanjem sistema linearnih jednačina

$$\Phi(x_i) = a_0\Phi_0(x_i) + a_1\Phi_1(x_i) + \cdots + a_n\Phi_n(x_i) = f_i, \qquad i = 0, 1, \ldots, n, \tag{15}$$

ili, u matričnom obliku,

$$\begin{bmatrix} \Phi_0(x_0) & \Phi_1(x_0) & \cdots & \Phi_n(x_0) \\ \Phi_0(x_1) & \Phi_1(x_1) & & \Phi_n(x_1) \\ \vdots & & & \\ \Phi_0(x_n) & \Phi_1(x_n) & & \Phi_n(x_n) \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

Postavljeni interpolacioni problem ima jedinstveno rešenje ako i samo ako je matrica sistema regularna, tj. $\det\,[\Phi_j(x_i)]_{i,j=0}^n \neq 0$. Da bi ovaj uslov bio ispunjen za promenljivi izbor čvorova, sistem funkcija $\{\Phi_0(x), \Phi_1(x), \ldots, \Phi_n(x)\}$ mora da bude takav da ne postoji linearna kombinacija

$$a_0\Phi_0(x) + a_1\Phi_1(x) + \cdots + a_n\Phi_n(x)$$

koja ima $n + 1$ različitih nula na $X$. Za takav sisteme funkcija kaže se da je *Čebiševljev sistem* (ima *Haarovo svojstvo*) i samo Čebiševljevi sistemi mogu da se koriste za interpolaciju.

**Teorema.** *Neka su funkcije* $\Phi_k(x)$ $n+1$ *puta diferencijabilne na* $[a,b]$. *Sistem funkcija* $\{\Phi_0(x), \Phi_1(x), \ldots, \Phi_n(x)\}$ *je Čebiševljev ako i samo ako za svako* $x \in [a,b]$ *važi*

$$\begin{vmatrix} \Phi_0(x) & \Phi_1(x) & \cdots & \Phi_k(x) \\ \Phi_0'(x) & \Phi_1'(x) & & \Phi_k'(x) \\ \vdots & & & \\ \Phi_0^{(k)}(x) & \Phi_1^{(k)}(x) & & \Phi_k^{(k)}(x) \end{vmatrix} \neq 0, \qquad k = 0, 1, \ldots n.$$

Sistem funkcija sa navedenim svojstvom koji se najčešće primenjuje u praksi je sistem algebarskih polinoma $\{1, x, x^2, \ldots, x^n\}$ i na taj način se rešava polinomska interpolacija. Za zadati iterpolacioni zahtev, tj. zadati skup različitih čvorova $x_i \in [a,b]$ i zadate vrednosti funkcije u njima $f_i = f(x_i)$, $i = 0, 1, \ldots, n$, interpolacioni polinom

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

postoji i jedinstven je. Zaista, ako se sistem jednačina (15) predstavi u obliku

$$\begin{aligned} a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n &= f_0, \\ a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n &= f_1, \\ &\vdots \\ a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_n x_n^n &= f_n, \end{aligned}$$

njegova determinanta je Vandermondova determinanta

$$\begin{vmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & & x_1^n \\ \vdots & & & & \\ 1 & x_n & x_n^2 & & x_n^n \end{vmatrix} = \prod_{i>j} (x_i - x_j) \neq 0,$$

pa postoji jedinstveno rešenje $(a_0, a_1, a_2, \ldots, a_n)$. Koeficijenti interpolaconog polinoma teorijski mogu da se dobiju rešavanjem ovog sistema, ali se u praksi, zbog njegove slabe uslovljenosti, postupa drugačije. U zavisnosti od načina formiranja razlikuju se vrste interpolacionih polinoma.

## Lagranžov interpolacioni polinom

Neka je funkcija $f(x)$ definisana na skupu $X \subseteq [a,b]$ i neka su zadate tačke $x_0, x_1, \ldots, x_n \in X$ u kojima funkcija ima vrednosti $f(x_i) = f_i$, $i = 0, 1, \ldots, n$.

Polinom

$$P_n(x) = \sum_{k=0}^{n} f_k L_k(x),$$

gde je, za $k = 0, 1, \ldots, n$, polinom $L_k(x)$ definisan sa

$$L_k(x) = \frac{(x - x_0)(x - x_1)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_n)}{(x_k - x_0)(x_k - x_1)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)},$$

zove se *Lagranžov interpolacioni polinom.*

### Njutnov interpolacioni polinom sa podeljenim razlikama

### Njutnovi interpolacioni polinomi sa konačnim razlikama

### Hermitov interpolacioni polinom

## Najbolje aproksimacije

Neka je $\{\Phi_k(x) \mid k = 0, 1, \ldots, n\}$ skup polinoma koji čine bazu u prostoru $\mathcal{P}_n$ polinoma stepena ne većeg od $n$ (što je najjednostavnije obezbeđeno ako je stepen polinoma $\Phi_k(x)$ jednak $k$). Problem najboljih polinomskih aproksimacija funkcija $f(x)$ definisanih na skupu $X \subseteq [a, b]$ i predstavlja određivanje parametara $a_k$, $k = 0, 1, \ldots, n$, u funkciji oblika

$$\Phi(x) = a_0\Phi_0(x) + a_1\Phi_1(x) + \cdots + a_n\Phi_n(x)$$

tako da je $\|\delta_n\| = \|f - \Phi\|$, gde je $\|\cdot\|$ izabrana norma u prostoru funkcija, minimalno. Za tako izabrane parametre $a_k = a_k^*$, $k = 0, 1, \ldots, n$, funkcija

$$\Phi^*(x) = a_0^*\Phi_0(x) + a_1^*\Phi_1(x) + \cdots + a_n^*\Phi_n(x)$$

zove se *najbolja aproksimacija* funkcije $f(x)$ po normi $\|\cdot\|$ u skupu polinoma stepena ne većeg od $n$, a $\|\delta_n^*\| = \|f - \Phi^*\|$ je *veličina najbolje aproksimacije.*

Prema izboru norme razlikuju se vrste najbolje aproksimacije funkcije.

Neka je $F$ vektorski prostor funkcija $\varphi$ koje su definisane na $X \subseteq [a,b]$. Ako one još imaju i osobinu da su $\varphi^p(x)$ apsolutno integrabilne na $(a,b)$ $(p \geq 1)$, norma može da se definiše na sledeći način:

$$\|\varphi\|_p = \left(\int_a^b |\varphi(x)|^p \, dx\right)^{1/p}.$$

Najčešće korišćene norme su:

$$p = 2: \qquad \|\varphi\|_2 = \sqrt{\int_a^b \big(\varphi(x)\big)^2 \, dx},$$

$$p \to \infty: \qquad \|\varphi\|_\infty = \max_{a \leq x \leq b} |\varphi(x)|,$$

a odgovarajuće najbolje aproksimacije su srednje–kvadratna i Čebiševljeva mini–maks aproksimacija redom. Norma na $F$ može da se definiše i kao

$$\|\varphi\|_p = \left(\sum_{j=0}^{m} |\varphi(x_j)|^p\right)^{1/p},$$

pri čemu su $x_j \in X$, $j = 0, 1, \ldots, m$ $(m > n)$. Specijalni i najčešće korišćeni slučajevi su

$$p = 2: \qquad \|\varphi\|_2 = \sqrt{\sum_{j=0}^{m} \big(\varphi(x_j)\big)^2 \, dx},$$

$$p \to \infty: \qquad \|\varphi\|_\infty = \max_{0 \leq j \leq m} |\varphi(x_j)|,$$

a odgovarajuće aproksimacije su diskretna srednje–kvadratna i diskretna mini–maks aproksimacija.

## Srednje–kvadratna aproksimacija

Već je rečeno da se u prostoru funkcija $\varphi(x)$ takvih da su $\varphi^2(x)$ integrabilne na $(a,b)$ (tj. integrali $\int_a^b \big(\varphi(x)\big)^2 \, dx$ konačni) može uvesti norma

$$\|\varphi\| = \sqrt{\int_a^b \big(\varphi(x)\big)^2 \, dx}$$

koja je indukovana skalarnim proizvodom

$$(\varphi, \psi) = \int_a^b \varphi(x)\psi(x)\, dx.$$

Ovaj koncept se može uopštiti uvođenjem *težinske funkcije* $p : (a, b) \to \mathbb{R}_0^+$, skalarnog proizvoda

$$(\varphi, \psi) = \int_a^b p(x)\varphi(x)\psi(x)\, dx$$

i odgovarajuće norme

$$\|\varphi\| = \sqrt{(\varphi, \varphi)} = \sqrt{\int_a^b p(x)\big(\varphi(x)\big)^2\, dx}.$$

Aproksimaciona funkcija oblika

$$\Phi(x) = a_0\Phi_0(x) + a_1\Phi_1(x) + \cdots + a_n\Phi_n(x)$$

takva da je $\|\delta_n\| = \|f - \Phi\|$ minimalno zove se *najbolja srednje–kvadratna aproksimacija* funkcije $f(x)$ sa težinom $p(x)$ na intervalu $(a, b)$ u skupu polinoma stepena ne većeg od $n$.

Norma odstupanja aproksimacione funkcije $\|\delta_n\|$ je minimalno ako se za parametre $a_k$, $k = 0, 1, \ldots, n$ uzmu one vrednosti za koje funkcija $\|\delta_n\|^2 = \|f - a_0\Phi_0 - a_1\Phi_1 - \cdots - a_n\Phi_n\|^2$ dostiže minimum. Posmatrajmo zato funkciju

$$\begin{aligned} J(a_0, a_1, \ldots, a_n) = \|\delta_n\|^2 = (\delta_n, \delta_n) = \int_a^b p(x)(\delta_n(x))^2\, dx \\ = \int_a^b p(x)\big(f(x) - a_0\Phi_0(x) - a_1\Phi_1(x) - \cdots - a_n\Phi_n(x)\big)^2\, dx \end{aligned}$$

i odredimo njenu stacionarnu tačku iz uslova

$$\frac{\partial}{\partial a_k} J(a_0, a_1, \ldots, a_n) = 0, \qquad k = 0, 1, \ldots, n.$$

Kako je

$$
\begin{aligned}
&\frac{\partial}{\partial a_k} J(a_0, a_1, \ldots, a_n) \\
&\quad = \frac{\partial}{\partial a_k} \left( \int_a^b p(x)\big(f(x) - a_0\Phi_0(x) - a_1\Phi_1(x) - \cdots - a_n\Phi_n(x)\big)^2 \, dx \right) \\
&\quad = \int_a^b \frac{\partial}{\partial a_k} \left( p(x)\big(f(x) - a_0\Phi_0(x) - a_1\Phi_1(x) - \cdots - a_n\Phi_n(x)\big)^2 \right) dx \\
&\quad = 2\int_a^b p(x)\big(f(x) - a_0\Phi_0(x) - a_1\Phi_1(x) - \cdots - a_n\Phi_n(x)\big)\big(-\Phi_k(x)\big) \, dx \\
&\quad = -2 \left( \int_a^b p(x) f(x) \Phi_k(x) \, dx - \sum_{j=0}^{n} a_j \int_a^b p(x)\Phi_j(x)\Phi_k(x) \, dx \right),
\end{aligned}
$$

parametri $a_k$, $k = 0, 1, \ldots, n$ mogu da se dobiju rešavanjem sistema linearnih jednačina

$$
\sum_{j=0}^{n} a_j \int_a^b p(x)\Phi_j(x)\Phi_k(x) \, dx = \int_a^b p(x)\Phi_k(x) f(x) \, dx, \qquad k = 0, 1, \ldots, n,
$$

tj.

$$
\sum_{j=0}^{n} a_j (\Phi_j, \Phi_k) = (f, \Phi_k), \qquad k = 0, 1, \ldots, n,
$$

ili, u razvijenom obliku,

$$
\begin{aligned}
(\Phi_0, \Phi_0)a_0 + (\Phi_1, \Phi_0)a_1 + \cdots + (\Phi_n, \Phi_0)a_1 &= (f, \Phi_0), \\
(\Phi_0, \Phi_1)a_0 + (\Phi_1, \Phi_1)a_1 + \cdots + (\Phi_n, \Phi_1)a_1 &= (f, \Phi_1), \\
&\vdots \\
(\Phi_0, \Phi_n)a_0 + (\Phi_1, \Phi_n)a_1 + \cdots + (\Phi_n, \Phi_n)a_1 &= (f, \Phi_n).
\end{aligned}
$$

Ako je $\{\Phi(x) \mid k = 0, 1, \ldots, n\}$ skup polinoma koji su ortogonalni u odnosu na skalarni proizvod

$$
(\varphi, \psi) = \int_a^b p(x)\varphi(x)\psi(x) \, dx,
$$

tj. važi

$$
(\Phi_i, \Phi_k) = \int_a^b p(x)\Phi_i(x)\Phi_k(x) \, dx = \begin{cases} 0, & i \neq k, \\ \|\Phi_k\|^2, & i = k, \end{cases}
$$

sistem je dijagonalan i rešenje je

$$a_k^* = \frac{(f, \Phi_k)}{(\Phi_k, \Phi_k)}, \qquad k = 0, 1, \ldots, n.$$

U tom slučaju je

$$\begin{aligned}
&\frac{\partial^2}{\partial a_i \partial a_k} J(a_0, a_1, \ldots, a_n) \\
&\qquad = -2\frac{\partial}{\partial a_i}\left(\int_a^b p(x)\Phi_k(x)f(x)\,dx - \sum_{j=0}^{n} a_j \int_a^b p(x)\Phi_k(x)\Phi_j(x)\,dx\right) \\
&\qquad = 2\int_a^b p(x)\Phi_k(x)\Phi_i(x)\,dx = 2(\Phi_i, \Phi_k),
\end{aligned}$$

pa je ispunjen dovoljan uslov da u stacionarnoj tački $(a_0^*, a_1^*, \ldots, a_n^*)$ funkcija $J(a_0, a_1, \ldots, a_n)$ dostiže minimum:

$$d^2 J(a_0^*, a_1^*, \ldots, a_n^*) = \sum_{i,k=0}^{n} \frac{\partial^2 J}{\partial a_i \partial a_k}\, da_i\, da_k = 2\sum_{k=0}^{n} \|\Phi_k\|^2\, da_k^2 > 0.$$

Za ovako izabrane parametre u aproksimacionoj funkciji važi

$$\begin{aligned}
\|\delta_n^*\|^2 &= (f - \Phi^*, f - \Phi^*) = (f, f) - 2(f, \Phi^*) + (\Phi^*, \Phi^*) \\
&= (f, f) - 2\left(f, \sum_{k=0}^{n} a_k^* \Phi_k\right) + \left(\sum_{j=0}^{n} a_j^* \Phi_j, \sum_{k=0}^{n} a_k^* \Phi_k\right) \\
&= (f, f) - 2\sum_{k=0}^{n} a_k^* (f, \Phi_k) + \sum_{j=0}^{n}\sum_{k=0}^{n} a_j^* a_k^* (\Phi_j, \Phi_k).
\end{aligned}$$

Koristeći ortogonalnost funkcija $\Phi_k(x)$ i jednakosti kojima su određeni parametri $a_k^*$, za veličinu najbolje aproksimacije se dobija

$$\begin{aligned}
\|\delta_n^*\|^2 &= (f, f) - 2\sum_{k=0}^{n} \frac{(f, \Phi_k)}{(\Phi_k, \Phi_k)}(f, \Phi_k) + \sum_{k=0}^{n} \frac{(f, \Phi_k)^2}{(\Phi_k, \Phi_k)^2}(\Phi_k, \Phi_k) \\
&= (f, f) - \sum_{k=0}^{n} \frac{(f, \Phi_k)^2}{(\Phi_k, \Phi_k)}.
\end{aligned}$$

Na osnovu izloženog, može se formalizovati postupak određivanja najbolje srednje–kvadratne aproksimacije u skupu polinoma stepena ne većeg od $n$.

1° Formirati niz polinoma $\{Q_0(x), Q_1(x), \ldots, Q_n(x)\}$, ortogonalnih u odnosu na težinu $p(x)$ na intervalu $(a, b)$, primenom Gram–Šmitovog postupka ili na neki drugi način.

2° Odrediti parametre u aproksimacionoj funkciji

$$\Phi(x) = a_0 Q_0(x) + a_1 Q_1(x) + \cdots + a_n Q_n(x)$$

prema formulama

$$a_k = \frac{(f, Q_k)}{(Q_k, Q_k)}, \qquad k = 0, 1, \ldots, n.$$

3° Proceniti veličinu najbolje aproksimacije prema formuli

$$\|\delta_n^*\| = \sqrt{(f, f) - \sum_{k=0}^{n} \frac{(f, \Phi_k)^2}{(\Phi_k, \Phi_k)}}.$$

U praksi, prvi korak se može preskočiti. U literaturi i u raznim programskim paketima se mogu naći polinomi ortogonalni na različitim intervalima i sa različitim težinskim funkcijama. U sledećoj tabeli su dati polinomi prvog, drugog i trećeg stepena iz klasa najpoznatijih predstavnika tzv. klasičnih ortogonalnih polinoma.

| Naziv klase | polinomi do trećeg stepena | interval ortogonalnosti | težinska funkcija |
|---|---|---|---|
| Ležandrovi | $P_1(x) = x$<br>$P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$<br>$P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$ | $(-1, 1)$ | $p(x) = 1$ |
| Čebiševljevi prve vrste | $T_1(x) = x$<br>$T_2(x) = 2x^2 - 1$<br>$T_3(x) = 4x^3 - 3x$ | $(-1, 1)$ | $p(x) = \frac{1}{\sqrt{1 - x^2}}$ |
| Lagerovi | $L_1(x) = -x + 1$<br>$L_2(x) = \frac{1}{2}x^2 - 2x + 1$<br>$L_3(x) = -\frac{1}{6}x^3 + \frac{3}{2}x^2 - 3x + 1$ | $(0, +\infty)$ | $p(x) = e^{-x}$ |
| Hermitovi | $H_1(x) = 2x$<br>$H_2(x) = 4x^2 - 2$<br>$H_3(x) = 8x^3 - 12x$ | $(-\infty, +\infty)$ | $p(x) = e^{-x^2}$ |

### Diskretna srednje–kvadratna aproksimacija

Ako je funkcija $f(x)$ zadata svojim vrednostima na diskretnom skupu tačaka $f(x_j) = f_j$, $j = 0, 1, 2, \ldots, m$ $(m > n)$, a norma odstupanja aproksimacione funkcije $\delta_n(x) = f(x) - \Phi(x)$ se definiše kao

$$\|\delta_n\| = \sqrt{\sum_{j=0}^{m} \big(\delta_n(x_j)\big)^2},$$

javlja se problem *diskretne srednje–kvadratne aproksimacija* ili *metoda najmanjih kvadrata*[1]. Napomenimo da se on u slučaju $m \le n$ ovaj problem svodi na interpolaciju.

Da bismo odredili parametre $a_k^*$, $k = 0, 1, \ldots, n$, u aproksimacionoj funkciji tako da je norma odstupanja bila minimalna, odredimo najpre stacionarne tačke funkcije

$$J(a_0, a_1, \ldots, a_n) = \|\delta_n\|^2 = \sum_{j=0}^{m} \big(f_j - a_0\Phi_0(x_j) - a_1\Phi_1(x_j) - \cdots - a_n\Phi_n(x_j)\big)^2.$$

Za svako $i = 0, 1, \ldots, n$ važi

$$\begin{aligned}\frac{\partial}{\partial a_i} J(a_0, a_1, \ldots, a_n) &= \sum_{j=0}^{m} \frac{\partial}{\partial a_i} \big(f_j - a_0\Phi_0(x_j) - a_1\Phi_1(x_j) - \cdots - a_n\Phi_n(x_j)\big)^2 \\ &= -2\sum_{j=0}^{m} \Phi_i(x_j)\big(f_j - a_0\Phi_0(x_j) - a_1\Phi_1(x_j) - \cdots - a_n\Phi_n(x_j)\big) \\ &= -2\sum_{j=0}^{m} \Phi_i(x_j)\big(f_j - a_0\Phi_0(x_j) - a_1\Phi_1(x_j) - \cdots - a_n\Phi_n(x_j)\big),\end{aligned}$$

pa se stacionarna tačka dobija kao rešenje sistema jednačina

$$\frac{\partial}{\partial a_i} J(a_0, a_1, \ldots, a_n) = 0, \qquad i = 0, 1, \ldots, n,$$

tj.,

$$\sum_{j=0}^{m} \Phi_i(x_j)\big(f_j - a_0\Phi_0(x_j) - a_1\Phi_1(x_j) - \cdots - a_n\Phi_n(x_j)\big) = 0, \qquad i = 0, 1, \ldots, n,$$

[1] least square approximation

ili, u razvijenom obliku,

$$
\begin{aligned}
a_0 \sum_{j=0}^{m} \Phi_0(x_j)\Phi_0(x_j) + \cdots + a_n \sum_{j=0}^{m} \Phi_0(x_j)\Phi_n(x_j) &= \sum_{j=0}^{m} \Phi_i(x_j) f_j, \\
&\vdots \\
a_0 \sum_{j=0}^{m} \Phi_n(x_j)\Phi_0(x_j) + \cdots + a_n \sum_{j=0}^{m} \Phi_n(x_j)\Phi_n(x_j) &= \sum_{j=0}^{m} \Phi_n(x_j) f_j,
\end{aligned}
$$

Ako uvedemo oznake

$$
\mathbf{f} = [f_0 \;\; f_1 \;\; \cdots \;\; f_m]^T, \quad \mathbf{a} = [a_0 \;\; a_1 \;\; \cdots \;\; a_n]^T,
$$

$$
X = \begin{bmatrix} \Phi_0(x_0) & \Phi_1(x_0) & \cdots & \Phi_n(x_0) \\ \Phi_0(x_1) & \Phi_1(x_1) & \cdots & \Phi_n(x_1) \\ \vdots & & & \\ \Phi_0(x_m) & \Phi_1(x_m) & \cdots & \Phi_n(x_m) \end{bmatrix} = [\Phi_j(x_i)]_{(m+1)\times(n+1)},
$$

može se uočiti da je koeficijent prethodnog sistema jednačina na mestu $(i,k)$, $i,k = 0,1,\ldots,n$, jednak skalarnom proizvodu $i$-te i $k$-te vektor-kolone matrice $X$, ili proizvodu (u matričnom smislu) $i$-te vrste matrice matrice $X^T$ i $k$-te kolone matrice $X$:

$$
\sum_{j=0}^{m} \Phi_i(x_j)\Phi_k(x_j) = [\Phi_i(x_0) \;\; \Phi_i(x_1) \;\; \cdots \;\; \Phi_i(x_m)] \begin{bmatrix} \Phi_k(x_0) \\ \Phi_k(x_1) \\ \vdots \\ \Phi_k(x_m) \end{bmatrix},
$$

a slobodni član u $i$-toj jednačini je proizvod $i$-te vrste matrice matrice $X^T$ i vektora $\mathbf{f}$:

$$
\sum_{j=0}^{m} \Phi_i(x_j) f_j = [\Phi_i(x_0) \;\; \Phi_i(x_1) \;\; \cdots \;\; \Phi_i(x_m)] \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_m \end{bmatrix}.
$$

Tako, sistem može da se predstavi u matričnom obliku

$$
X^T X \mathbf{a} = X^T \mathbf{f},
$$

a njegovo rešenje je

$$
\mathbf{a}^* = (X^T X)^{-1} X^T \mathbf{f}.
$$

U praksi, određivanje parametara u aproksimacionoj funkciji obavlja se na potpuno prirodan način. Ako pođemo od zahteva da su sve vrednosti aproksimacione funkcije jednake zadatim vrednostima, tj. $\Phi(x_j) = f_j$, $j = 0, 1, \ldots, m$, dobija se preodređen pravougaoni sistem jednačina

$$a_0\Phi_0(x_j) + a_1\Phi_1(x_j) + \cdots + a_n\Phi_n(x_j) = f_j, \quad j = 0, 1, \ldots, m,$$

ili, u matričnom obliku,

$$X\mathbf{a} = \mathbf{f},$$

koji u opštem slučaju ne mora da ima rešenje. Množenjem sa $X^T$ dobija se kvadratni sistem

$$X^TX\mathbf{a} = X^T\mathbf{f},$$

isti onaj koji je dobijen iz uslova najbolje aproksimacije.

Postupak dobijanja najbolje diskretne srednje–kvadratne aproksimacije na osnovu podataka $f(x_j) = f_j$, $j = 0, 1, \ldots, m$, ili, drugim rečima, određivanje funkcije $\Phi(x)$ koja najbolje opisuje skup podataka $(x_j, f_j)$, $j = 0, 1, \ldots, m$, u literaturi i praksi zove se još i *fitovanje podataka*.

Naravno, izbor oblika aproksimacione funkcije je od ključnog značaja. Opisani postupak moguće je sprovesti samo ako je aproksimaciona funkcija linearna po parametrima:

$$\Phi(x) = a_0\Phi_0(x) + a_1\Phi_1(x) + \cdots + a_n\Phi_n(x).$$

Ako je aproksimaciona funkcija polinom prvog stepena $\Phi(x) = a_0 + a_1x$, rešava se vrlo rasprostranjen problem poznat pod imenom *linearna regresija*. Uz određene modifikacije postupak linearne regresije može da se primeni i u slučaju nekih nelinearnih aproksimacionih funkcija.

$1^\circ$ Ako je aproksimaciona funkcija oblika $\Phi(x) = a_0e^{a_1x}$, tada se umesto nje može da se posmatrati funkcija

$$\Psi(x) = \log(\Phi(x)) = \log a_0 + a_1x = b_0 + b_1x, \qquad b_0 = \log a_0,\ b_1 = a_1,$$

a za odgovarajuće parametre uzeti $a_0^* = e^{b_0^*}$, $a_1^* = b_1^*$, gde su $b_0^*, b_1^*$ vrednosti parametara određene postupkom linearne regresije funkcijom $\Psi(x)$ nad podacima $(x_j, \log f_j)$, $j = 0, 1, \ldots, m$.

$2^\circ$ Ako je aproksimaciona funkcija oblika $\Phi(x) = a_0x^{a_1}$, tada se linearizacija može izvršiti na sledeći način:

$$\begin{aligned}\log(\Phi(x)) &= \log(a_0x^{a_1}) = \log\left(a_0e^{\log x^{a_1}}\right) = \log\left(a_0e^{a_1\log x}\right)\\ &= \log a_0 + a_1\log x,\\ \Psi(X) &= b_0 + b_1X, \qquad b_0 = \log a_0,\ b_1 = a_1,\ X = \log x.\end{aligned}$$

Tada se za odgovarajuće parametre uzimaju $a_0^* = e^{b_0^*}$, $a_1^* = b_1^*$, gde su $b_0^*, b_1^*$ vrednosti parametara određene postupkom linearne regresije funkcijom $\Psi(X) = b_0 + b_1 X$ nad podacima $(\log x_j, \log f_j)$, $j = 0, 1, \ldots, m$.

3° Ako je aproksimaciona funkcija oblika $\Phi(x) = \dfrac{1}{a_0 + a_1 x}$, tada se rešava linearna regresija funkcijom

$$\Psi(x) = \frac{1}{\Phi(x)} = a_0 + a_1 x$$

nad podacima $(x_j, 1/f_j)$, $j = 0, 1, \ldots, m$.