NUMERIČKA ANALIZA

1. INTERPOLACIJA

U numeričkim metodama razmatraju se matematičke formule i računski procesi koji mogu da budu efektivno sprovedeni. To znači da mogu da budu sprovedeni za konačno mnogo izvedenih računskih radnji i da mogu da budu dovedeni do poznatih konkretnih brojnih vrijednosti, najčešće do približnih vrijednosti.

U glavi 1. govorimo o zadatku interpolacije. Dato je nekoliko tačaka u ravni, a treba odrediti kruvu liniju da prolazi kroz sve te tačke.

1.1. LAGRANŽOV INTERPOLACIONI POLINOM

Razmotrimo funkciju $f\colon R \to R$ ili $f\colon [a,b] \to R$, Neka je $n \geq 0$. Razmotrimo $n+1$ međusobno različitih tačaka $x_0, \ldots, x_n$ na realnoj osi. Neka su date vrijednosti $f(x_i)$ za $0 \leq i \leq n$. Želimo da odredimo funkciju $L_n\colon R \to R$ ili $L_n\colon [a,b] \to R$ koja zadovoljava uslov $L_n(x_i) = f(x_i)$ za $0 \leq i \leq n$. Za funkciju L_n se kaže da je interpolaciona funkcija. Kažemo da želimo da riješimo zadatak o interpolaciji za podatke $(x_i, f(x_i))$ sa $0 \leq i \leq n$. Treba definisati klasu funkcija $\{L_n\}$ u kojoj se traži funkcija L_n. Uzmimo da je to klasa svih mogućih polinoma $R \to R$ čiji je stepen $\leq n$. Dakle, polazimo od predstavljanja $L_n(x) = a_0 + a_1 x + \ldots + a_n x^n$, gdje su $a_j \in R$ $(0 \leq j \leq n)$ zasad neodređene veličine. Iz uslova interpolacije $L_n(x_i) = f(x_i)$, $0 \leq i \leq n$ imamo $a_0 + a_1 x_i + \ldots + a_n x_i^n = f(x_i)$, $0 \leq i \leq n$ ili

$$\begin{bmatrix} 1 & x_0 & \cdots & x_0^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ \vdots \\ f(x_n) \end{bmatrix}.$$

Imamo linearni sistem od $n+1$ jednačina sa $n+1$ nepoznatih. Njegova determinanta D je poznata Vandermondova determinanta i znamo da je $D = \prod_{0 \leq i < j \leq n}(x_j - x_i)$. Po uslovu je $x_i \neq x_j$ za $i \neq j$. Tako da je $D \neq 0$. Zato sistem ima jedinstveno rješenje. Dakle, postavljeni zadatak o interpolaciji ima rješenje u razmatranoj klasi $\{L_n\}$, i to jedinstveno. Na redu je konstrukcija interpolacionog polinoma $L_n = L_n(x)$, odnosno dobijanje njegovog eksplicitnog izraza. Mogao bi da bude riješen linearni sistem, a mi ćemo izvršiti konstrukciju drugim putem.

Razmotrimo proizvod $\omega_{n+1}(x) = \prod_{i=0}^{n}(x - x_i)$. Funkcija ω_{n+1} je polinom stepena tačno $n+1$ i zadovoljava $\omega_{n+1}(x_i) = 0$ za $0 \leq i \leq n$. U izrazu za $\omega_{n+1}(x)$ izostavimo jedan faktor, recimo prvi faktor $x - x_0$. Vidimo da je funkcija $y(x) = \prod_{i=1}^{n}(x - x_i)$ polinom stepena tačno n i da zadovoljava $y(x_i) = 0$ za $1 \leq i \leq n$. A $y(x_0) = \prod_{i=1}^{n}(x_0 - x_i) \neq 0$. Prema tome, lako formiramo funkcije $\ell_i = \ell_i(x)$ da zadovoljavaju: $\ell_i(x_j) = 0$ za $j \neq i$ i $\ell_i(x_i) = 1$. Upravo

$$\ell_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} \quad \text{ili se piše samo} \quad \ell_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}.$$

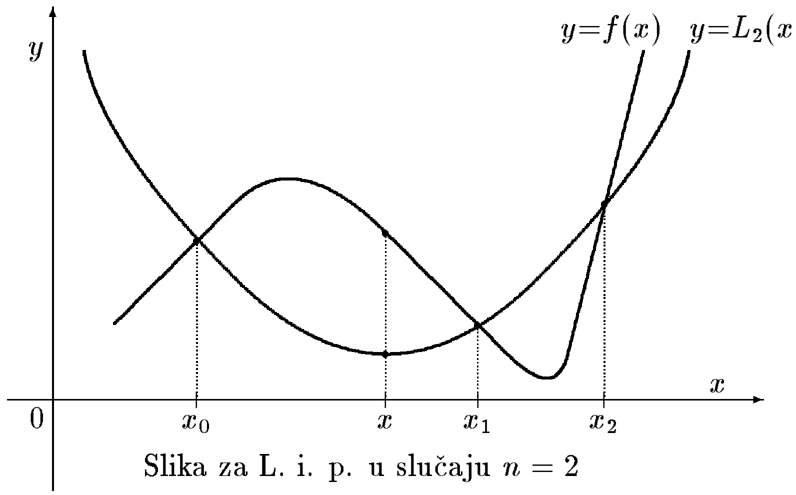
Na kraju, funkcija $L_n(x) = \ell_0(x) f(x_0) + \ldots + \ell_n(x) f(x_n)$ predstavlja rješenje postavljenog zadatka, za L_n se kaže da predstavlja Lagranžov interpolacioni polinom, pišemo

$$L_n(x) = \sum_{i=0}^{n} f(x_i) \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}.$$

U numeričkoj praksi, veličina $f(x)$ nije poznata, a veličina $L_n(x)$ je poznata, odnosno može da bude izračunata, tako da je $L_n(x)$ približna zamjena za $f(x)$. Ovdje je x jedna tačka na realnoj osi koja nije čvor, tj. koja ne pripada mreži čvorova, tj. koja se ne poklapa ni sa kojim x_i, $0 \leq i \leq n$. Ponekad se upotrebljava sljedeća terminologija. Ako tačka x za koju se interpolacija vrši pripada intervalu $[a, b]$ obrazovanom od strane čvorova onda se kaže da se vrši interpolacija u užem smislu. A u suprotnom slučaju se kaže da se vrši ekstrapolacija. Ovdje je $a = \min_{i=0,\dots,n} x_i$ i $b = \max_{i=0,\dots,n} x_i$.

Tabela Ulazni podaci za zadatak o interpolaciji

i	x_i	$f(x_i)$
0	x_0	$f(x_0)$
1	x_1	$f(x_1)$
$\cdots$	$\cdots$	$\cdots$
n	x_n	$f(x_n)$



Slika za L. i. p. u slučaju $n = 2$

Donekle sličan zadatku o interpolaciji je ZADATAK O NAJBOLJOJ APROKSIMACIJI, engl. best fit. Skup ulaznih podataka $(x_i, f(x_i))$, $1 \leq i \leq n$ je istog oblika kao maločas. Kod aproksimacije se ne traži više da pojedinačno odstupanje $r_i = f(x_i) - L_n(x_i)$ bude $= 0$, već se sada traži da $r_i = f(x_i) - g(x_i)$ bude što je moguće manje. Ovdje je g funkcija koja predstavlja rješenje zadatka o aproksimaciji. Traži se ustvari da vrijednost izraza $r^2 = \sum_{i=1}^n |r_i|^2$ bude što je moguće manja, pa se zato kaže i – **metoda najmanjih kvadrata**. Iz koje klase funkcija $\{g\}$ treba izabrati g? U najprostijem slučaju g je linearna funkcija. Ili je g polinom stepena $\leq k$, gdje je $k \leq n - 2$. Ili je g oblika $g(x) = a + \frac{b}{x}$ ili $g(x) = ae^{bx}$ ili itd.

Najbolja aproksimacija sa $g(x) = ax + b$ za $(x_i, f(x_i)) = (x_i, f_i)$, gdje je $1 \leq i \leq n$ (linearna regresija, linearni trend):

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i f_i \\ \sum_{i=1}^n f_i \end{bmatrix}.$$

Navedena formula (za određivanje a i b) lako se izvodi iz uslova za stacionarnu tačku funkcije $F = F(a, b)$, odnosno iz uslova za njenu tačku minimuma. Znamo da uslov za stacionarnu tačku glasi: $\frac{\partial F(a,b)}{\partial a} = 0$, $\frac{\partial F(a,b)}{\partial b} = 0$. Uvedena je oznaka $F(a, b) = \sum_{i=1}^n (ax_i + b - f_i)^2$. Vidimo da je $r^2 = F(a, b)$.

Uzmimo da treba naći zavisnost oblika $y = ae^{bx}$, na osnovu podataka $\{(x_i, y_i),\ 1 \leq i \leq n\}$. Svođenje na prethodni slučaj pomoću: $y = ae^{bx} \Rightarrow \ln y = \ln a + bx$. Treba primijeniti model prave linije na podatke $\{(x_i, \ln y_i),\ 1 \leq i \leq n\}$. Kao rezultat, imaćemo b i $\ln a$.

Primjer 1: Naći najbolju aproksimaciju oblika $y = ax + b$ po metodi najmanjih kvadrata za podatke

x	1	2	3	4
y	1	4	6	8

. Izrada: $\sum x_i^2 = 30 \quad \sum x_i = 10 \quad \sum x_i y_i = 59 \quad \sum y_i = 19$

$$\begin{cases} 30a + 10b = 59 \\ 10a + 4b = 19 \end{cases} \qquad a = 2{,}3 \quad b = -1.$$ Odgovor: $y = 2{,}3x - 1$.

Primjer 2: Po metodi najmanjih kvadrata, naći aproksimaciju oblika $y = ae^{bx}$ (svođenjem na linearni slučaj) za podatke

x	1	2	3	4
y	1	4	10	20

. Izrada: $\ln y = Y$

x	1	2	3	4
Y	0	1,39	2,30	2,99

$Y = Ax + B$ $\begin{cases} 30A + 10B = 21{,}64 \\ 10A + 4B = 6{,}68 \end{cases}$ $A = 0{,}99$ $B = -0{,}80$ $e^{-0{,}80} = 0{,}45$. Odgovor: $y = 0{,}45e^{0{,}99x}$.

1.2. OCJENA GREŠKE ZA LAGRANŽOV INTERPOLACIONI POLINOM

Ova sekcija predstavlja nastavak prethodne, tako da se sve oznake preuzimaju. Neka $x \in [a,b]$ ili $x \in R$, s tim da x nije čvor. Razmotrimo numerički odgovor $f(x) \approx L_n(x)$ na pitanje: čemu je jednako $f(x)$. Greškom ili greškom numeričkog odgovora nazivamo razliku $r(x) = f(x) - L_n(x)$. U ovoj sekciji treba da dobijemo izraz za $r(x)$, odnosno treba da ocijenimo sa gornje strane $|r(x)|$. Pretpostavlja se da $f \in C^{n+1}(R)$ ili $f \in C^{n+1}[a,b]$. Uvedimo funkciju φ relacijom

$$\varphi(t) = f(t) - L_n(t) - K\omega_{n+1}(t), \tag{1}$$

gdje je K zasad neodređena konstanta. Izaberimo K tako da bude $\varphi(x) = 0$. Dakle, neka je

$$K = \frac{f(x) - L_n(x)}{\omega_{n+1}(x)}. \tag{2}$$

Sada je vrijednost K određena. Pogledajmo rješenja jednačine $\varphi(t) = 0$. Vidimo da φ ima bar $n+2$ nule. Upravo, njene nule su sigurno $t = x$, $t = x_9$, …, $t = x_n$. Naime, $\varphi(x_i) = f(x_i) - L_n(x_i) - K\omega_{n+1}(x_i) = f(x_i) - f(x_i) - K \cdot 0 = 0$. Tačke $x, x_0, \ldots, x_n$ obrazuju na realnoj osi interval $[a,b]$. Dakle, neka bude $a = \min(x, x_0, \ldots, x_n)$ i $b = \max(x, x_0, \ldots, x_n)$. Nabrojane tačke obrazuju na realnoj osi $n+1$ malih intervala. (Rolova teorema: ako je f neprekidna na $[a,b]$, diferencijabilna u (a,b) i $f(a) = f(b)$ onda postoji broj $a < \xi < b$ takav da je $f'(\xi) = 0$.) Rolova teorema govori da između dvije nule funkcije postoji bar jedna nula njenog izvoda. Tako da funkcija φ' ima bar $n+1$ nulu unutar $[a,b]$. Slično, φ'' ima bar n nula. Itd. Na kraju, $\varphi^{(n+1)}$ ima bar jednu nulu: postoji $\xi \in (a,b)$ takav da je $\varphi^{(n+1)}(\xi) = 0$. S druge strane, jasno je da je $L_n^{(n+1)}(t) \equiv 0$ i da je $\omega_{n+1}^{(n+1)}(t) \equiv (n+1)!$ Ako se (1) diferencira $n+1$ puta po t onda imamo $\varphi^{(n+1)}(t) = f^{(n+1)}(t) - K(n+1)!$ Uvrstimo $t = \xi$: $\varphi^{(n+1)}(\xi) = f^{(n+1)}(\xi) - K(n+1)!$ ili $0 = f^{(n+1)}(\xi) - K(n+1)!$ ili

$$K = \frac{f^{(n+1)}(\xi)}{(n+1)!}. \tag{3}$$

Upoređivanjem (2) i (3) dobijamo rješenje, tj. traženi izraz za grešku:

$$r(x) = f(x) - L_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)\omega_{n+1}(x), \tag{4}$$

$\xi = \xi(x) \in (a,b)$. Često se (4) piše u obliku:

$$|f(x) - L_n(x)| \leq \frac{1}{(n+1)!} M_{n+1}|\omega_{n+1}(x)|, \text{ gdje je } M_{n+1} = \max_{t\in[a,b]} |f^{(n+1)}(t)|.$$

U formuli (4): greška $= \ldots$, a zatim u susjednoj formuli: $|\text{greška}| \leq \ldots$.

1.3. PODIJELJENE RAZLIKE I NJIHOVA SVOJSTVA

U ovoj sekciji uvodi se pojam podijeljene razlike i utvrđuju se neka njihova jednostavna svojstva. Kaže se konačna razlika ili razlika ili engl. difference za izraz $f(x_1) - f(x_0)$, a kaže se podijeljena razlika za izraz $\dfrac{f(x_1) - f(x_0)}{x_1 - x_0}$, a za taj izraz koriste se razne oznake, kao $[x_0, x_1[$ ili (f, x_0, x_1) ili $f[x_0, x_1]$. Dakle, neka su o funkciji $f\colon R \to R$ dati na raspolaganje oni isti podaci kao u prethodnim sekcijama: $(x_i, f(x_i))$, $0 \le i \le n$. Podijeljena razlika prvog reda funkcije f označava se sa $f[x_i, x_j]$ i jednaka je

$$f[x_i, x_j] = \frac{f(x_j) - f(x_i)}{x_j - x_i}. \tag{1}$$

Slično, drugog reda

$$f[x_i, x_j, x_k] = \frac{f[x_j, x_k] - f[x_i, x_j]}{x_k - x_i}.$$

Uopšte, razlika reda k definiše se pomoću razlika reda $k-1$, i to

$$f[x_0, \ldots, x_k] = \frac{f[x_1, \ldots, x_k] - f[x_0, \ldots, x_{k-1}]}{x_k - x_0}.$$

Podijeljena razlika nultog reda označava se sa $f[x_i]$ i jednaka je $f(x_i)$.

Sljedeća lema utvrđuje jedno svojstvo podijeljenih razlika. Lako se vidi da je podijeljena razlika jednaka jednoj linearnoj kombinaciji vrijednosti funkcije u obuhvaćenim čvorovima. Lema daje izraz za koeficijente linearne kombinacije.

Lema. Za svako $k \ge 1$ važi jednakost

$$f[x_1, \ldots, x_k] = \sum_{j=1}^{k} f(x_j) \frac{1}{\prod_{i \ne j} (x_j - x_i)} \tag{2}$$

$(\prod_{i=1,\ldots,j-1,j+1,\ldots,k})$.

Lemu ćemo dokazati primjenom principa matematičke indukcije po k. Za $k = 1$ razmatrana jednakost svodi se na $f[x_1] = f(x_1)$ i očito je istinita. Za $k = 2$ ona se svodi na definicionu relaciju (1). Uzmimo da je jednakost (2) tačna za $k = \ell$ i dokažimo da je ona tada tačna i za $k = \ell + 1$. Prva naredna transformacija oslanja se na definiciju, a druga na indukcijsku pretpostavku:

$$f[x_1, \ldots, x_{\ell+1}] = \frac{1}{x_{\ell+1} - x_1} \Big(f[x_2, \ldots, x_{\ell+1}] - f[x_1, \ldots, x_\ell[\Big) =$$

$$\frac{1}{x_{\ell+1} - x_1} \left(\sum_{j=2}^{\ell+1} f(x_j) \frac{1}{\prod_{\substack{2 \le i \le \ell+1 \\ i \ne j}} (x_j - x_i)} - \sum_{j=1}^{\ell} f(x_j) \frac{1}{\prod_{\substack{1 \le i \le \ell \\ i \ne j}} (x_j - x_i)} \right). \tag{3}$$

Čemu je jednak koeficijent c_j uz $f(x_j)$ u posljednjem izrazu? Vidimo da je $1 \le j \le \ell + 1$. Uzmimo prvo da je $1 < j < \ell + 1$. Tada je

$$c_j = \frac{1}{x_{\ell+1} - x_1} \left(\frac{1}{\prod_{\substack{2 \le i \le \ell+1 \\ i \ne j}} (x_j - x_i)} - \frac{1}{\prod_{\substack{1 \le i \le \ell \\ i \ne j}} (x_j - x_i)} \right) =$$

$$\frac{1}{x_{\ell+1}-x_1}\cdot\frac{1}{\prod_{\substack{1\le i\le \ell+1\\ i\ne j}}(x_j-x_i)}\Big((x_j-x_1)-(x_j-x_{\ell+1})\Big)=\frac{1}{\prod_{\substack{1\le i\le \ell+1\\ i\ne j}}(x_j-x_i)}.$$

Vidimo da c_j ima predviđeni oblik. Uzmimo sada da je $j=1$. Veličina $f(x_1)$ pojavljuje se samo u drugom sabirku desne strane formule (3) i očito je

$$c_1=\frac{1}{x_{\ell+1}-x_1}\left(\prod_{\substack{1\le i\le \ell\\ i\ne j}}(x_j-x_i)\right)^{-1}=\left(\prod_{\substack{1\le i\le \ell+1\\ i\ne j}}(x_j-x_i)\right)^{-1},$$

ima predviđeni oblik. Slično kada je $j=\ell+1$, tj. za $c_{\ell+1}$. Lema je dokazana.

Iz leme neposredno proizilaze sljedeća dva svojstva podijeljenih razlika.

1. Podijeljena razlika je linearni operator od funkcije f, tj. važi $(\alpha_1 f_1+\alpha_2 f_2)[x_1,\ldots,x_k]=\alpha_1 f_1[x_1,\ldots,x_k]+\alpha_2 f_2[x_1,\ldots,x_k]$.

2. Podijeljena razlika $f[x_1,\ldots,x_k]$ je simetrična funkcija svojih argumenata $x_1,\ldots,x_k$, tj. ne mijenja se pri bilo kakvoj njihovoj permutaciji. Recimo, $f[x_1,x_2]=f[x_2,x_1]$ i $f[x_1,x_2,x_3]=f[x_3,x_1,x_2]$ i slično.

Na kraju, uobičajeno je da se podijeljene razlike prikazuju u obliku tabele. U nastavku je data tabela u slučaju kada su poznate vrijednosti funkcije f u tačkama $x_0,\ldots,x_n$:

$f(x_0)$	$f[x_0,x_1]$	$f[x_0,x_1,x_2]$	$\ldots$	$f[x_0,x_1,\ldots,x_n]$
$f(x_1)$	$f[x_1,x_2]$	$\ldots$		
$f(x_2)$	$\ldots$			
$\ldots$				
	$f[x_{n-1},x_n]$			
$f(x_n)$				

1.4. NJUTNOVA INTERPOLACIONA FORMULA SA PODIJELJENIM RAZLIKAMA

Neka su opet poznate vrijednosti $f(x_0),\ldots,f(x_n)$ funkcije $f\colon R\to R$ u $n+1$ međusobno različitih tačaka $x_0,\ldots,x_n$. U ovoj sekciji će: (a) L. i. p. biti prikazan u drugom obliku. Naime, u 1.1. je bila dokazana jedinstvenost interpolacionog polinoma, tako da Nj. i. p. koji slijedi predstavlja samo drugi oblik jednog te istog polinoma (drugi prikaz, drugi zapis). Još će: (b) biti izvedena formula o vezi između podijeljene razlike $(n+1)$–vog reda neke funkcije i njenog $(n+1)$–vog izvoda.

Kao prvo, dokažimo jedan identitet. Vidi se da je

$$f(x)-L_n(x)=f(x)-\sum_{i=0}^{n}f(x_i)\prod\nolimits_{j=0\ \ j\ne i}^{n}\frac{x-x_j}{x_i-x_j}=$$

$$\prod_{i=0}^{n}(x-x_i)\cdot\left(\frac{f(x)}{\prod_{i=0}^{n}(x-x_i)}+\sum_{i=0}^{n}\frac{f(x_i)}{(x_i-x)\prod_{j=0\ \ j\ne i}^{n}(x_i-x_j)}\right)=\omega_{n+1}(x)f[x,x_0,\ldots,x_n],$$

uz primjenu leme iz prethodne sekcije. Ranija oznaka $\omega_{n+1}(x)=\prod_{i=0}^{n}(x-x_i)$.

$$\text{Identitet}\quad f(x)-L_n(x)=\omega_{n+1}(x)f[x,x_0,\ldots,x_n].\tag{1}$$

S druge strane, prepišimo formulu (4) iz sekcije 1.2.

$$r(x) = f(x) - L_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)\omega_{n+1}(x), \quad \xi = \xi(x) \in (a,b),$$

izraz za grešku L. i. p. Uporedimo formulu (1) i prepisanu formulu. Tako dobijamo

$$f[x, x_0, \ldots, x_n] = \frac{1}{(n+1)!} f^{(n+1)}(\xi) \tag{2}$$

za neko $\xi \in (a,b)$, gdje je $a = \min(x, x_0. \ldots, x_n)$, $b = \max(x, x_0, \ldots, x_n)$, $f \in C^{n+1}[a,b]$, $n \geq 0$.

Na lijevoj strani formule (2) napisana je podijeljena razlika $(n+1)$–vog reda funkcije f.

Uradili smo (b) i na redu je (a).

Prelazimo na izvođenje Nj. i. f. Uvedimo oznaku $\omega_{k+1}(x) = \prod_{i=0}^{k}(x - x_i)$. Neka $L_k = L_k(x)$ označava L. i. p. funkcije f po mreži čvorova $\{x_0, \ldots, x_k\}$. Tako da je naravno $L_k(x_i) = f(x_i)$ za $i = 0, \ldots, k$. Funkcija $L_m(x) - L_{m-1}(x)$ predstavlja očito polinom stepena manjeg ili jednakog m, a za $x = x_0$, ..., $x = x_{m-1}$ očito je

$$L_m(x) - L_{m-1}(x) = f(x) - f(x) = 0.$$

Primjera radi, kvadratni polinom $p = p(x)$ čije su nule $x = 3$ i $x = 4$ prikazuje se kao $p(x) = A_2(x-3)(x-4)$. Tako da se polinom $L_m(x) - L_{m-1}(x)$ prikazuje kao

$$L_m(x) - L_{m-1}(x) = A_m \omega_m(x),$$

a treba odrediti A_m. Supstitucija $x = x_m$:

$$L_m(x_m) - L_{m-1}(x_m) = A_m \omega_m(x_m),$$

$$f(x_m) - L_{m-1}(x_m) = A_m \omega_m(x_m). \tag{3}$$

S druge strane, upotrebimo (1) kada je $n = m - 1$:

$$f(x) - L_{m-1}(x) = \omega_m(x) f[x, x_0, \ldots, x_{m-1}],$$

$$\text{supstitucija } x = x_m: \quad f(x_m) - L_{m-1}(x_m) = \omega_m(x_m) f[x_m, x_0, \ldots, x_{m-1}],$$

$$f(x_m) - L_{m-1}(x_m) = \omega_m(x_m) f[x_0, \ldots, x_m] \tag{4}$$

jer je podijeljena razlika simetrična funkcija svojih argumenata.

Uporedimo (3) i (4). Tako $A_m = f[x_0, \ldots, x_m]$ i zato

$$L_m(x) - L_{m-1}(x) = f[x_0, \ldots, x_m]\omega_m(x).$$

Nastavljamo

$$L_n(x) = L_0(x) + (L_1(x) - L_0(x)) + (L_2(x) - L_1(x)) + \ldots + (L_n(x) - L_{n-1}(x)),$$

$$L_n(x) = L_0(x) + f[x_0, x_1]\omega_1(x) + f[x_0, x_1, x_2]\omega_2(x) + \ldots + f[x_0, x_1, \ldots, x_n]\omega_n(x),$$

$$L_n(x) = f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \ldots + f[x_0, x_1, \ldots, x_n](x - x_0)\ldots(x - x_{n-1}). \tag{5}$$

Interpolacioni polinom zapisan u obliku (5) naziva se Njutnovim interpolacionim polinomom sa podijeljenim razlikama.

Na kraju, želimo da zapišemo zajedno, tj. u jednoj formuli izraz za interpolacioni polinom i izraz za njegovu grešku, želimo da napišemo formulu oblika "tačna vrijednost = približna vrijednost + greška", tj. formulu $f(x) = L_n(x) + r(x)$. Na osnovu (1) i (5) imamo:

$$f(x) = f(x_0) + f[x_0, x_1]\omega_1(x) + f[x_0, x_1, x_2]\omega_2(x) + \ldots +$$

$$f[x_0, \ldots, x_n]\omega_n(x) + f[x, x_0, \ldots, x_n]\omega_{n+1}(x).$$

Najzad, zapazimo jedno dobro svojstvo koje posjeduje Njutnova interpolaciona formula sa podijeljenim razlikama, neka vrsta "aditivnosti". Zamislimo ovakvu situaciju. Mi procjenjujemo $f(x)$ pomoću $L_4(x)$, tj. koristimo čvorove $x_0, \ldots, x_4$. U okviru toga, ocijenimo odgovarajuću grešku $f(x) - L_4(x)$, pri čemu se ispostavi da je greška isuviše velika, ispostavi se da nismo zadovoljni dobijenom preciznošću. Tada, radi dobijanja bolje aproksimacije, uključimo u račun još jedan čvor x_5. Zapaziti da prilikom računanja zbira $L_5(x)$ ne moramo da ponovo računamo sve njegove sabirke, već da je dovoljno da se dosad izračunatoj približnoj vrijednosti $L_4(x)$ doda još jedan sabirak; uporedi izraze za $L_4(x)$ i $L_5(x)$.

Naknadno dodato: generalizacija pojma podijeljene razlike: $f[x, x] = \lim_{\varepsilon\to 0} f[x, x + \varepsilon] = f'(x)$, $f[x, x, x] = \lim_{\varepsilon\to 0} f[x, x + \varepsilon, x + 2\varepsilon] = \frac{1}{2!}f''(x)$ i slično, $f[x_1, x_1, x_2] = \lim_{\varepsilon\to 0} f[x_1, x_1 + \varepsilon, x_2]$ i slično.

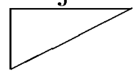
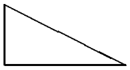
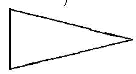
1.5. KONAČNE RAZLIKE

Konačne razlike čine osnovni aparat u numeričkim metodama, a definišu se samo u slučaju ravnomjerne (ekvidistantne) mreže čvorova. U ovoj sekciji daju se definicije i svojstva. Ekvidistantna mreža čvorova definiše se pomoću početnog čvora $x_0 \in R$ i svog koraka $h > 0$. Čvorovi su $x_i = x_0 + ih$, obično za i važi $i = 0, 1, \ldots, n$, gdje je $n \geq 1$. Sljedeća tablica predstavlja primjer. Mreža je definisana sa $x_0 = 0 \quad h = 0{,}1 \quad n = 5$. Tablica se odnosi na funkciju $y(x) = e^{2x}$. Prikazane su njene vrijednosti u čvorovima y_i i odgovarajuće konačne razlike prvog reda Δy_i i drugog reda $\Delta^2 y_i$:

i	x_i	y_i	Δy_i	$\Delta^2 y_i$
0	0	1	0,22140	0,04902
1	0,1	1,22140	0,27042	0,05988
2	0,2	1,49182	0,33030	0,07312
3	0,3	1,82212	0,40342	0,08932
4	0,4	2,22554	0,49274	
5	0,5	2.71828		

Razmotrimo funkciju $f\colon R \to R$ ili $f\colon [a, b] \to R$. Označimo sa f_i njene vrijednosti u čvorovima mreže, tj. $f_i = f(x_i)$ za $i = 0, 1, \ldots, n$. Za jednu te istu konačnu razliku koriste se tri različita naziva. Izraz $f_{i+1} - f_i$ naziva se konačnom razlikom prvog reda. Za taj izraz koriste se sljedeće tri različite oznake: $\Delta f_i = f_{i+1} - f_i$ konačna razlika unaprijed, $\nabla f_{i+1} = f_{i+1} - f_i$ konačna razlika unazad i $\delta f_{i+1/2} = f_{i+1} - f_i$ ili svejedno $f^1_{i+1/2} = f_{i+1} - f_i$ centralna konačna razlika. Razlike višeg reda definišu se na osnovu razlika nižeg reda: $\Delta^m f_i = \Delta(\Delta^{m-1} f_i) = \Delta^{m-1} f_{i+1} - \Delta^{m-1} f_i$, $\nabla^m f_i = \nabla(\nabla^{m-1} f_i) = \nabla^{m-1} f_i - \nabla^{m-1} f_{i-1}$ i $\delta^m f_i = \delta(\delta^{m-1} f_i) = \delta^{m-1} f_{i+1/2} - \delta^{m-1} f_{i-1/2}$ ili svejedno $f^m_i = f^{m-1}_{i+1/2} - f^{m-1}_{i-1/2}$. Vidimo da se za centralne razlike koriste dvije različite oznake

$\delta^m f_i$ i f_i^m. U slučaju centralne razlike $\delta^m f_i = f_i^m$, ako je m paran onda je i cio, a ako je m neparan onda je i polucio. Napomenimo da se rijetko koriste ∇ i δ.

Tabela razlika unaprijed Δ^m ima oblik  , tabela razlika unazad ∇^m ima oblik  , a tabela centralnih razlika δ^m ima oblik  . U nastavku je prikazana tablica centralnih razlika:

x	f	$\delta f{=}f^1$	$\delta^2 f{=}f^2$	$\delta^3 f{=}f^3$	
x_0	f_0				$\vdots$
		$f^1_{1/2}$			
x_1	f_1		f^2_1		$\vdots$
		$f^1_{3/2}$		$f^3_{3/2}$	
x_2	f_2		f^2_2		$\vdots$
		$f^1_{5/2}$		$f^3_{5/2}$	
x_3	f_3		f^2_3		
		$f^1_{7/2}$		$\vdots$	
x_4	f_4		$\vdots$		
		$\vdots$			
$\vdots$	$\vdots$				

Na primjer: $\Delta f_i = f_{i+1} - f_i$, $\Delta^2 f_i = f_{i+2} - 2f_{i+1} + f_i$, $\Delta^3 f_i = f_{i+3} - 3f_{i+2} + 3f_{i+1} - f_i$, $\Delta^4 f_i = f_{i+4} - 4f_{i+3} + 6f_{i+2} - 4f_{i+1} + f_i$, itd. Vidimo da je konačna razlika jedna linearna kombinacija vrijednosti funkcije u čvorovima koji su obuhvaćeni. Sljedeća lema utvrđuje čemu su jednaki koeficijenti linearne kombinacije.

Lema 1. Za svako $m \geq 1$ važi jednakost

$$\Delta^m f_i = \sum_{j=0}^{m} (-1)^j \binom{m}{j} f_{i+m-j}.$$

Lemu ćemo dokazati indukcijom. Za $m = 1$ jednakost se svodi na $\Delta f_i = f_{i+1} - f_i$ i očito je tačna. Uzmimo da je jednakost tačna za $m = l$ i dokažimo da je tada jednakost tačna i za $m = l + 1$:

$$\Delta^{l+1} f_i = \Delta^l f_{i+1} - \Delta^l f_i = \sum_{j=0}^{l} (-1)^j \binom{l}{j} f_{i+1+l-j} - \sum_{j=0}^{l} (-1)^j \binom{l}{j} f_{i+l-j} =$$

$$f_{i+1+l} + \sum_{j=1}^{l} (-1)^j \binom{l}{j} f_{i+1+l-j} - \sum_{j=0}^{l-1} (-1)^j \binom{l}{j} f_{i+l-j} - (-1)^l f_i =$$

$$f_{i+1+l} + \sum_{j=1}^{l} (-1)^j \binom{l}{j} f_{i+1+l-j} - \sum_{j=1}^{l} (-1)^{j-1} \binom{l}{j-1} f_{i+l-j+1} + (-1)^{l+1} f_i =$$

$$\text{pomoću iz kombinatorike} \quad \binom{l}{j} + \binom{l}{j-1} = \binom{l+1}{j} \quad \text{imamo}$$

$$f_{i+1+l} + \sum_{j=1}^{l} (-1)^j \binom{l+1}{j} f_{i+1+l-j} + (-1)^{l+1} f_i =$$

$$\sum_{j=0}^{l+1}(-1)^j\binom{l+1}{j}f_{i+l+1-j},$$

što je i trebalo. Lema je dokazana.

Lema pokazuje da je operator uzimanja konačne razlike od neke funkcije linearan po toj funkciji, tj. da važi jednakost $\Delta^m(\alpha f+\beta g)_i=\alpha\Delta^m f_i+\beta\Delta^m g_i$.

Podijeljena razlika $f[x_i,\ldots,x_{i+m}]$ reda m obrazuje se po vrijednostima funkcije u $m+1$ čvorova $x_i,\ldots,x_{i+m}$ koji su raspoređeni bilo ravnomjerno bilo neravnomjerno. Konačna razlika $\Delta^m f_i$ reda m obrazuje se po tim istim čvorovima jedino kada oni čine ravnomjernu mrežu.

Lema 2. Važi jednakost

$$\Delta^m f_i=h^m m!f[x_i,\ldots,x_{i+m}].$$

Dokazuje se indukcijom. Za $m=1$ imamo $\Delta f_i=hf[x_i,x_{i+1}]$ ili $\Delta f_i=h\frac{f_{i+1}-f_i}{x_{i+1}-x_i}$ ili $\Delta f_i=f_{i+1}-f_i$. Indukcijski korak:

$$\Delta^{l+1}f_i=\Delta^l f_{i+1}-\Delta^l f_i=$$

$$\left(h^l l!f[x_{i+1},\ldots,x_{i+l+1}]-h^l l!f[x_i,\ldots,x_{i+l}]\right)\frac{h(l+1)}{x_{i+l+1}-x_i}=$$

$$h^{l+1}(l+1)!f[x_i,\ldots,x_{i+l+1}],\quad \text{dokazano.}$$

Bila je napisana formula (2) u sekciji 1.4. $f[x_i,\ldots,f_{i+m}]=\frac{1}{m!}f^{(m)}(\xi)$.

Napisana formula (1.4.2) izražava vezu između podijeljene razlike određenog reda i izvoda funkcije tog istog reda. Jednostavnim njenim kombinovanjem sa prethodnom lemom, dobija se veza konačnih razlika i izvoda, kako slijedi.

Lema 3. Ako $f\in C^m[x_i,x_{i+m}]$ onda važi

$$\Delta^m f_i=h^m f^{(m)}(\xi)\quad \text{za neko}\ \ x_i<\xi<x_{i+m}.$$

Naravno da se ista jednakost može zapisati i u drugim oznakama: $\Delta^m f_i=\nabla^m f_{i+m}=\delta^m f_{i+m/2}=f^m_{i+m/2}=h^m f^{(m)}(\xi)$ za neko $\xi\in(x_i,x_{i+m})=(x_0+ih,x_0+(i+m)h)$.

1.6. NJUTNOVE INTERPOLACIONE FORMULE SA KONAČNIM RAZLIKAMA

U ovoj sekciji izvode se dvije najpoznatije interpolacione formule. I one predstavljaju samo drugi zapis L. i. p. Račun je jednostavan jer je sve pripremljeno u prethodnim sekcijama.

Razmotrimo ravnomjernu mrežu čvorova čiji je osnovni čvor $x=x_0\in R$ i čiji je korak $h>0$. Sami čvorovi su $x_i=x_0+ih$ za razne cijele i. Pored nezavisno promjenljive x koristi se i druga nezavisno promjenljiva t, kako je to uobičajeno kada se radi o ravnomjernoj mreži. Dvije promjenljive povezane su sljedećom relacijom o smjeni promjenljive: $x=x_0+ht$ ili svejedno $t=\frac{x-x_0}{h}$.

Razmotrimo funkciju f i označimo sa f_i njene vrijednosti u čvorovima x_i.

Neka je $L_n=L_n(x)$ L. i. p. za funkciju f za mrežu čvorova $x_0,x_1,\ldots,x_n$. Prepišimo formulu (1.4.5) iz sekcije 1.4.

$$L_n(x)=f(x_0)+f[x_0,x_1](x-x_0)+\ldots+f[x_0,\ldots,x_n](x-x_0)\ldots(x-x_{n-1}),$$

sa podijeljenih razlika prelazimo na konačne razlike po lemi iz prethodne sekcije

$$L_n(x) = f_0 + \frac{1}{h}\Delta f_0(x - x_0) + \ldots + \frac{1}{n!h^n}\Delta^n f_0(x - x_0)\ldots(x - x_{n-1}),$$

sa x prelazimo na t

$x - x_0 = ht,\ x - x_1 = h(t-1),\ x - x_2 = h(t-2),\ \ldots,$

$$L_n(x_0 + ht) = f_0 + \Delta f_0 t + \frac{1}{2!}\Delta^2 f_0 t(t-1) + \ldots + \frac{1}{n!}\Delta^n f_0 t(t-1)\ldots(t-n+1), \qquad (I)$$

ovo je I Nj. i. f. ili Nj. i. f. za interpolaciju unaprijed.

Moglo bi se naravno umjesto $\Delta f_0, \Delta^2 f_0, \ldots, \Delta^n f_0$ pisati i $f^1_{1/2}, f^2_1, \ldots, f^n_{n/2}$.

Što se tiče izraza za grešku formule (I), kako se radi samo o drugom prikazu jednog te istog polinoma to možemo naravno da iskoristimo raniju formulu za grešku (1.2.4), samo što ćemo toj formuli dati nešto drukčiji oblik, kako slijedi:

$$f(x) - L_n(x) = f(x_0 + ht) - L_n(x_0 + ht) = \frac{1}{(n+1)!} f^{(n+1)}(\xi) \prod_{i=0}^{n}(x - x_i),$$

$$f(x) - L_n(x) = f(x_0 + ht) - L_n(x_0 + ht) = \frac{1}{(n+1)!} f^{(n+1)}(\xi) h^{n+1} t(t-1)\ldots(t-n),$$

gdje je ξ neki broj takav da je $\min(x_0, x) < \xi < \max(x, x_n)$.

Razmotrimo sada $n+1$ čvorova $x_{-n} < \ldots < x_{-1} < x_0$. Neka je sada $L_n = L_n(x)$ L. i. p. za f po toj mreži čvorova. I dalje je $x = x_0 + ht$. Sličnim izvođenjem dobija se druga Nj. i. f. ili Nj. i. f. za interpolaciju unazad:

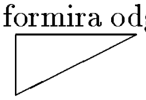
$$L_n(x) = f(x_0) + f[x_0, x_{-1}](x - x_0) + \ldots + f[x_0, \ldots, x_{-n}](x - x_0)\ldots(x - x_{-(n-1)}),$$

$$L_n(x_0 + ht) = f_0 + \nabla f_0 t + \frac{1}{2!}\nabla^2 f_0 t(t+1) + \ldots + \frac{1}{n!}\nabla^n f_0 t(t+1)\ldots(t+n-1), \qquad (II)$$

a za grešku formule (II) važi sljedeći izraz:

$$r(x) = f(x) - L_n(x) = f(x_0 + ht) - L_n(x_0 + ht) = \frac{1}{(n+1)!} f^{(n+1)}(\xi) h^{n+1} t(t+1)\ldots(t+n),$$

$\min(x_{-n}, x) < \xi < \max(x, x_0)$. Ili $x_{-n} < \xi < x_0$, kada se radi o interpolaciji u užem smislu.

Pogledajmo još jednom prvu Nj. i. f. Uobičajeno je da se ta formula zapisuje preko konačnih razlika unaprijed. Tako da se prethodno formira odgovarajuća tablica konačnih razlika unaprijed, ta tablica ima gornje trougaoni oblik  . Vidi se da sve potrebne brojne vrijednosti čitamo samo iz gornjeg reda tablice.

Ako koristimo mrežu $\{x_0, x_1, x_2\}$ onda je

$$L_2(x) = f_0 + \frac{x - x_0}{h}\Delta f_0 + \frac{(x - x_0)(x - x_1)}{2h^2}\Delta^2 f_0 \quad \text{ili}$$

$$L_2(x_0 + ht) = f_0 + t\Delta f_0 + \frac{1}{2}t(t-1)\Delta^2 f_0.$$

Dodajmo još jedan čvor, da bismo sa većom preciznošću saznali $f(x_0 + ht)$. Ako koristimo mrežu $\{x_0, x_1, x_2, x_3\}$ onda je

$$L_3(x_0 + ht) = f_0 + t\Delta f_0 + \frac{1}{2}t(t-1)\Delta^2 f_0 + \frac{1}{6}t(t-1)(t-2)\Delta^3 f_0.$$

Vidimo da se doda jedan sabirak, kada se sa L_2 prelazi na L_3.

Zapišimo L_3 zajedno sa izrazom za grešku:

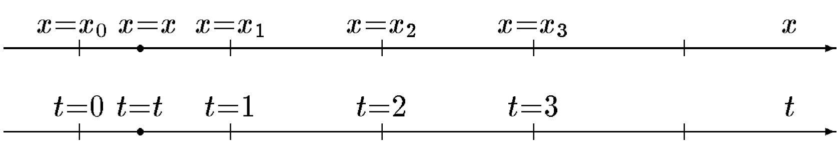
$$f(x_0 + ht) = L_3(x_0 + ht) + \frac{1}{24}t(t-1)(t-2)(t-3)h^4 f^{IV}(\xi) \quad \text{ili}$$

$$f(x) = f_0 + \frac{x - x_0}{h}\Delta f_0 + \frac{(x-x_0)(x-x_1)}{2!h^2}\Delta^2 f_0 + \frac{(x-x_0)(x-x_1)(x-x_2)}{3!h^3}\Delta^3 f_0 +$$
$$\frac{1}{4!}(x-x_0)(x-x_1)(x-x_2)(x-x_3)f^{IV}(\xi).$$

Kaže se da je I Nj. i. f. slična Tejlorovoj formuli.

Pogledajmo mali primjer. Neka je $h = 0{,}1 \quad x_0 = 0 \quad x_1 = 0{,}1 \quad x_2 = 0{,}2 \quad x_3 = 0{,}3$. Poznata je veza $x = x_0 + ht$. Recimo, ako je $x = 0{,}04$ tada je $t = 0{,}4$.

Sljedeća slika prikazuje vezu dvije promjenljive x i t:



Pogledajmo mali primjer za II Nj. i. f. Neka bude $h = 0{,}1 \quad x_{-3} = 9{,}7 \quad x_{-2} = 9{,}8 \quad x_{-1} = 9{,}9 \quad x_0 = 10$. Stalno je $x = x_0 + ht$. Recimo, ako je $x = 9{,}87$ tada je $t = -1{,}3$.

Iskustvo pokazuje da je najbolje da se interpolacioni polinom obrazuje pomoću vrijednosti funkcije u malom broju tačaka, obično manje ili jednako od pet ili šest tačaka. Pogledajmo tri moguća slučaja. Neka imamo veliku tablicu vrijednosti funkcije f po ravnomjernoj mreži, recimo od $x = 0$ do $x = 10$ sa korakom $h = 0{,}1$. Na osnovu datih informacija treba, za dato x koje nije čvor, primjenom aparata interpolacije, naći dobru približnu vrijednost za broj $f(x)$. Tačna vrijednost $f(x)$ očito je van našeg domašaja. Prvi slučaj: tačka x je blizu početka mreže. Tada treba primijeniti I Nj. i. f. Drugi slučaj: tačka x je blizu kraja mreže, blizu krajnjeg desnog čvora. Tada je pogodno da se primijeni II Nj. i. f. I još treći mogući slučaj: za dati broj x može da se kaže da se nalazi u sredini tablice. Ako se opredijelimo da koristimo šest čvorova (oni su uzastopni) onda uzimamo tri lijevo od x i tri desno od x, zato što se tako minimizuje greška jer se minimizuje proizvod $\omega_{n+1}(x)$, a i intuitivno je jasno da je tako povoljno. A za sprovođenje samog računanja treba upotrebiti bilo koji od raznih, a međusobno ekvivalentnih, zapisa (prikaza) jednog te istog interpolacionog polinoma; recimo L. i. p. ili Nj. i. f. sa podijeljenim razlikama ili I Nj. i. f. (ako I Nj. i. f. onda sa x_0 označimo krajnji lijevi od šest čvorova).

1.7. INTERPOLACIJA SA VIŠESTRUKIM ČVOROVIMA

U slučaju L. i. p. pretpostavljalo se da su u čvorovima date samo vrijednosti funkcije. Razmotrimo sada jedan opštiji zadatak interpolacije. V. sliku. Neka mrežu čine čvorovi $x_1, \ldots, x_n \in [a, b]$ među kojima nema poklapanja. Ovdje je $a = \min(x, x_1, \ldots, x_n)$ i $b = \max(x, x_1, \ldots, x_n)$. Neka su u čvorovima date vrijednosti funkcije $f(x_i)$ i njenih izvoda $f^{(j)}(x_i)$ do reda $m_i - 1$ uključeno, $j = 1, \ldots, m_i - 1$ za $i = 1, \ldots, n$. Prema tome, u svakoj tački x_i poznate su brojne vrijednosti $f(x_i), f'(x_i), \ldots, f^{(m_i-1)}(x_i)$, $i = 1, \ldots, n$. Tako da ulaznih veličina ukupno ima $m_1 + \ldots + m_n = s + 1$. Želimo da odredimo interpolacionu funkciju koja odgovara ovim ulaznim podacima. Opredijelimo se da interpolacionu funkciju tražimo u klasi svih polinoma stepena manjeg od $s + 1$. Za H_s se kaže da je Hermitov interpolacioni polinom ili da je interpolacioni polinom sa višestrukim čvorovima. Za $m_i \geq 1$ se kaže da predstavlja višestrukost čvora x_i, $i = 1, \ldots, n$. U tački $x = 10$, funkcija $y = (x - 10)^2$ ima dvostruku nulu (nulu reda 2).

Mi treba da razmotrimo sljedeća četiri pitanja: (1) postojanje i. p. (2) jedinstvenost i. p. (3) konstrukcija i. p. tj. dobijanje eksplicitnog izraza za i. p. i (4) ocjena greške: kolika se greška čini kada se kaže da je $f(x)$ približno jednako $H_s(x)$ za određeno $x \in [a, b]$.

Dakle, $H_s = H_s(x)$ treba da bude polinom stepena $\leq s$ koji zadovoljava uslove: $H_s^{(j)}(x_i) = f^{(j)}(x_i)$, $j = 0, \ldots, m_i - 1$, $i = 1, \ldots, n$. Ima $s + 1$ uslova. Opšti izraz za polinom stepena $\leq s$ glasi: $H_s(x) = a_s x^s + \ldots + a_1 x + a_0$. Ima $s + 1$ koeficijenata, tj. $s + 1$ stepeni slobode.

Oko jedinstvenosti. Iz algebre je poznato sljedeće tvrđenje: jedini polinom $p = p(x)$ stepena $\leq s$ koji ima $s + 1$ nula, pri čemu se pojedina nula broji svojom višestrukošću, jeste $p(x) \equiv 0$. Dopustimo da postoje dva polinoma $H_s^{(1)}(x)$ i $H_s^{(2)}(x)$ stepena $\leq s$ da oba zadovoljavaju sve uslove (svih $s + 1$ ograničenja). Neka bude $H_s^{(3)}(x) = H_s^{(2)}(x) - H_s^{(1)}(x)$. Imamo da je $H_s^{(3)(j)}(x) = 0$, $j = 0, \ldots, m_i - 1$, $i = 1, \ldots, n$. $H_s^{(3)}$ je polinom stepena $\leq s$ i ima barem $s + 1$ nula (brojano sa višestrukostima). Zato je $H_s^{(3)}(x) \equiv 0$, tj. $H_s^{(1)}(x) \equiv H_s^{(2)}(x)$. Ovim je jedinstvenost dokazana.

Oko postojanja. Dovedimo u vezu $s + 1$ ograničenja i $s + 1$ stepeni slobode. Dakle, treba da budu zadovoljene sljedeće jednakosti (ima ih $s + 1$):

$$H_s^{(j)}(x_i) = f^{(j)}(x_i), \quad j = 0, \ldots, m_i - 1, \quad i = 1, \ldots, n. \tag{1}$$

Mi smo upravo napisali sistem od $s + 1$ linearnih jednačina sa $s + 1$ nepoznatih $a_s, \ldots, a_0$. Razmotrimo i odgovarajući homogeni sistem:

$$H_s^{(j)}(x_i) = 0, \quad j = 0, \ldots, m_i - 1, \quad i = 1, \ldots, n. \tag{2}$$

Ima li homogeni sistem drugog rješenja osim trivijalnog rješenja $a_s = \ldots = a_0 = 0$? Kao što je maločas navedeno prilikom analize jedinstvenosti, iz (2) slijedi da je $H_s(x) \equiv 0$; slijedi da je $a_s = \ldots = a_0 = 0$. Homogeni sistem (2) ima samo trivijalno rješenje. Zato i sistem (1) ima jedinstveno rješenje, ma kakve da su brojne vrijednosti $\{f^{(j)}(x_i)\}$ koje čine njegovu desnu stranu (njegove slobodne članove). Ovim je postojanje dokazano; v. i šablon kasnije.

Sada o konstrukciji. Brojne vrijednosti $\{f^{(j)}(x_i)\}$ ne učestvuju u matrici sistema (1) već jedino učestvuju na desnoj strani sistema. Zato će pojedina nepoznata a_k biti linearna kombinacija tih brojnih vrijednosti. Broj a_k je koeficijent polinoma. Kako je $H_s(x) = a_s x^s + \ldots + a_1 x + a_0$ to polinom $H_s(x)$ može da bude prikazan u sljedećem obliku (kada se linearne kombinacije uvrste u $H_s(x) = \ldots$):

$$H_s(x) = \sum_{i=1}^{n} \sum_{j=0}^{m_i-1} c_{ij}(x) f^{(j)}(x_i),$$

gdje su $c_{ij}(x)$ polinomi stepena $\leq s$. Nećemo izvoditi eksplicitne izraze za $c_{ij}(x)$, zato što su ti izrazi glomazni, već ćemo samo navesti jedan primjer kasnije.

Još o ocjeni greške. Neka $x \in [a,b]$ i neka x nije čvor. Treba ocijeniti razliku $f(x) - H_s(x)$. Neka $f \in C^{s+1}[a,b]$. Uvedimo proizvod $\omega_{s+1}(t) = \prod_{i=1}^{n}(t-x_i)^{m_i}$. Tako da je ω_{s+1} polinom stepena tačno $s+1$ čiji je najstariji koeficijent $=1$, pa će biti $\omega_{s+1}^{(s+1)}(t) = (s+1)!$ Uvedimo u razmatranje i funkciju φ relacijom

$$\varphi(t) = f(t) - H_s(t) - K\omega_{s+1}(t), \tag{3}$$

gdje je K zasad neodređena brojna vrijednost. Vrijednost K biramo tako da u tački interpolacije $t = x$ bude zadovoljen uslov $\varphi(x) = 0$. Dakle, mi stavljamo $K = \dfrac{f(x) - H_s(x)}{\omega_{s+1}(x)}$. Rekli smo da je tačka nterpolacije $t = x$ fiksirana. Koliko nula ima funkcija $\varphi = \varphi(t)$? Tačka $t = x_i$ je nula reda m_i za svako $i = 1, \ldots, n$. Plus $t = x$. Ukupno ima nula, uzimajući u obzir njihove višestrukosti, najmanje $m_1 + \ldots + m_n + 1 = s + 2$. Odavde, po Rolovoj teoremi, prvi izvod $\varphi' = \varphi'(t)$ ima bar $s+1$ nula na intervalu $[a,b]$, uračunavajući višestrukosti. Slično, φ'' ima bar s nula. Itd. Funkcija $\varphi^{(s+1)}$ ima bar jednu nulu. Označimo tu nulu sa ξ: $\varphi^{(s+1)}(\xi) = 0$, $\xi \in (a,b)$. Diferencirajmo relaciju (3) $s+1$ puta. Tako $\varphi^{(s+1)}(t) = f^{(s+1)}(t) - K(s+1)!$ Uvrstimo ovdje $t = \xi$. Tako $0 = f^{(s+1)}(\xi) - K(s+1)!$ Odranije imamo da je $K = \dfrac{f(x) - H_s(x)}{\omega_{s+1}(x)}$. Tako

$$f(x) - H_s(x) = \frac{1}{(s+1)!} f^{(s+1)}(\xi)\omega_{s+1}(x).$$

Ovdje je $\xi = \xi(x)$; ξ zavisi od x. Dobili smo izraz za grešku.

Razmotrimo jedan primjer. Neka imamo $n = 3$ čvora $x = -1$, $x = 0$ i $x = 1$. Neka su čvorovi $x = -1$ i $x = 1$ jednostruki (prosti), a $x = 0$ neka je dvostruk, tako da je $s + 1 = 4$. Dakle, date su četiri brojne vrijednosti, označavaju se kao $f(-1)$, $f(0)$, $f'(0)$, $f(1)$ ili kao f_{-1}, f_0, f_0', f_1. Treba odrediti polinom trećeg stepena $H_3 = H_3(x)$ da zadovoljava:

$$H_3(-1) = f_{-1}, \quad H_3(0) = f_0, \quad H_3'(0) = f_0', \quad H_3(1) = f_1.$$

U okviru primjera, sastavićemo eksplicitni izraz za H_3 i napisaćemo izraz za grešku. V. sliku za H_3. Za eksplicitni izraz, treba da se pođe od predstavljanja $H_3(x) = a_3x^3 + a_2x^2 + a_1x + a_0$. Dobiće se:

$$H_3(x) = -\frac{1}{2}x^2(x-1)f_{-1} - (x+1)(x-1)f_0 + \frac{1}{2}(x+1)x^2 f_1 - (x+1)x(x-1)f_0'.$$

Izraz za grešku:

$$f(x) - H_3(x) = \frac{1}{4!} f^{IV}(\xi)(x+1)x^2(x-1), \quad \xi \in (-1,1) \quad (\text{ako } x \in [-1,1]).$$

<u>Šablon:</u> $\qquad H_s(x) = a_s x^s + \ldots + a_1 x + a_0,$

$$H_s'(x) = s a_s x^{s-1} + \ldots + a_1, \quad \text{itd.}$$

$$H_s(x_1) = f(x_1) \Rightarrow a_s x_1^s + \ldots + a_1 x_1 + a_0 = f(x_1),$$

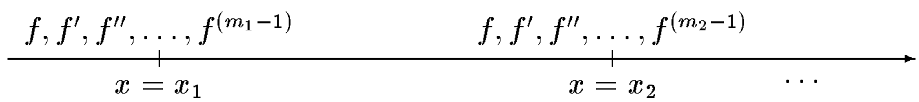
$$H_s'(x_1) = f'(x_1) \Rightarrow s a_s x_1^{s-1} + \ldots + a_1 = f'(x_1), \quad \text{itd.}$$

$$H_s(x_2) = f(x_2) \Rightarrow a_s x_2^s + \ldots + a_1 x_2 + a_0 = f(x_2), \quad \text{itd.} \quad \text{itd.}$$

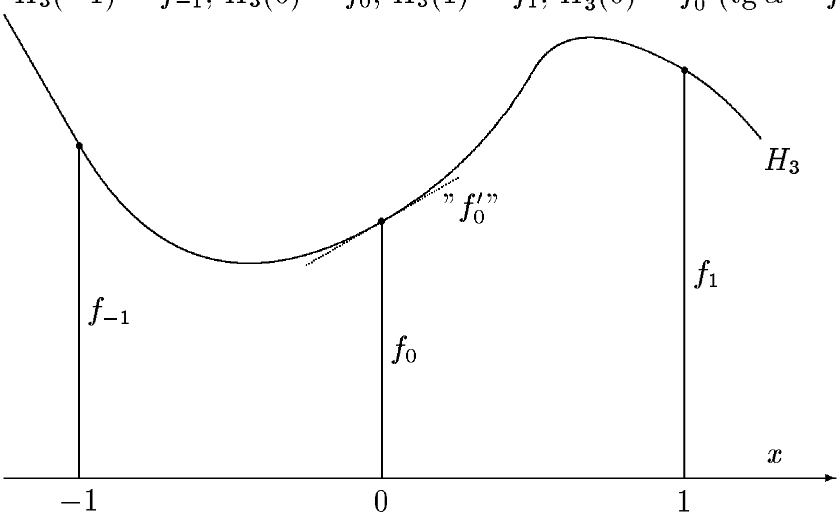
$$\begin{bmatrix} x_1^s & \cdots & x_1 & 1 \\ s x_1^{s-1} & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ x_2^s & \cdots & x_2 & 1 \\ \vdots & \ddots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} a_s \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f'(x_1) \\ \vdots \\ f(x_2) \\ \vdots \end{bmatrix}.$$

Ako su na desnoj strani sve nule onda postoji samo trivijalno rješenje $\Rightarrow \det M \neq 0$.

Slika za H. i. p.



Slika za primjer sa Hermitovim i. polinomom H_3: $H_3(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0$, $H_3(-1) = f_{-1}$, $H_3(0) = f_0$, $H_3(1) = f_1$, $H_3'(0) = f_0'$ ($\text{tg}\,\alpha = f_0'$, koef. pravca tangente):



1.8. INTERPOLACIJA POMOĆU SPLAJNA

Razmotrimo opet zadatak o interpolaciji funkcije. Neka je n prirodan broj i neka je $[a, b]$ zatvoreni interval na realnoj osi. Neka je po intervalu $[a, b]$ postavljeno $n + 1$ čvorova $a = x_0 < x_1 < \ldots < x_n = b$ i neka je u svakom čvoru poznata brojna vrijednost $f(x_i) = f_i \in R$. Na osnovu tih podataka o funkciji f, treba naći približnu vrijednost za $f(x)$, gdje je x data tačka iz intervala $[a, b]$. Iskustvo pokazuje da interpolacija pomoću polinoma visokog stepena obično ne daje zadovoljavajuće aproksimacije. Zato razmotrimo drugu mogućnost za rješavanje postavljenog zadatka. Neka interpolaciona funkcija $s = s(x)$ na svakom malom intervalu $[x_0, x_1], \ldots, [x_{n-1}, x_n]$ bude polinom čiji stepen nije visok. Za $s = s(x)$ se kaže da je splajn, a kaže se da se vrši interpolacija pomoću splajna ili pomoću dio–po–dio polinoma; engl. spline – krivuljar. Ako je s polinom trećeg stepena na $[x_{i-1}, x_i]$ onda se kaže da je s kubni splajn (ovo ćemo raditi). Kada dobijemo splajn, onda ćemo mi, kada je data tačka x, izračunati brojnu vrijednost $s(x)$. Broj $s(x)$ i predstavlja naš odgovor, tj. kazaćemo da je $f(x) \approx s(x)$. Treba da se dobije i ocjena za grešku $r(x) = f(x) - s(x)$. Radi jednostavnosti računanja, razmotrićemo samo slučaj ekvidistantno raspoređenih čvorova $x_i = a + ih$ za $0 \le i \le n$, gdje je $h = \dfrac{b-a}{n}$.

Definicija. Za podatke n, a, b i $\{f_i\}_{i=0}^{n}$, kubnim splajnom naziva se funkcija $s\colon [a, b] \to R$ koja zadovoljava sljedeća tri uslova: (a) na svakom malom intervalu $[x_{i-1}, x_i]$, $s(x)$ je polinom trećeg stepena, (b) $s \in C^2[a, b]$ i (c) $s(x_i) = f_i$ za $0 \le i \le n$ (uslov interpolacije).

Jasno,

$$s(x) = s_i(x) \qquad \text{za} \qquad x_{i-1} \le x \le x_i,$$

gdje je $s_i(x)$ polinom trećeg stepena. Napišimo predstavljanje polinoma trećeg stepena u nešto prilagođenom obliku:

$$s_i(x) = a_i + b_i(x - x_i) + \frac{c_i}{2}(x - x_i)^2 + \frac{d_i}{6}(x - x_i)^3, \qquad 1 \le i \le n.$$

Imamo $4n$ slobodnih veličina $\{a_i, b_i, c_i, d_i\}_{i=1}^{n}$. Očito je

$$s_i'(x) = b_i + c_i(x - x_i) + \frac{d_i}{2}(x - x_i)^2, \qquad s_i''(x) = c_i + d_i(x - x_i).$$

Neprekidnost splajna i njegovog prvog i drugog izvoda je pod znakom pitanja samo u tačkama dodira $x_1, \ldots, x_{n-1}$ dva susjedna mala intervala. Uslov (b) možemo da rastavimo na uslove: (b_1) $s(x_i - 0) = s(x_i + 0)$, tj. $s_i(x_i) = s_{i+1}(x_i)$, (b_2) $s'(x_i - 0) = s'(x_i + 0)$, tj. $s_i'(x_i) = s_{i+1}'(x_i)$ i (b_3) $s''(x_i - 0) = s''(x_i + 0)$, tj. $s_i''(x_i) = s_{i+1}''(x_i)$, $1 \le i \le n-1$. Tako da (b) daje $3(n-1)$ uslova, a uslov interpolacije (c) daje još $n+1$ uslova; ukupno $4n-2$ uslova. Broj stepeni slobode je $4n - (4n-2) = 2$.

Iskoristimo prvo (b_1) i (c) u obliku $s_i(x_{i-1}) = f_{i-1}$ i $s_i(x_i) = f_i$ za $1 \le i \le n$:

$$s_i(x_{i-1}) = f_{i-1} \qquad a_i + b_i(x_{i-1} - x_i) + \frac{c_i}{2}(x_{i-1} - x_i)^2 + \frac{d_i}{6}(x_{i-1} - x_i)^3 = f_{i-1}$$

$$a_i - b_i h + \frac{c_i}{2}h^2 - \frac{d_i}{6}h^3 = f_{i-1} \qquad i = 1, \ldots, n$$

$$s_i(x_i) = f_i \qquad a_i = f_i \qquad i = 1, \ldots, n \qquad \text{(svi } a_i \text{ su određeni i izlaze iz računa)}$$

$$b_i h - \frac{c_i}{2}h^2 + \frac{d_i}{6}h^3 = f_i - f_{i-1} \qquad i = 1, \ldots, n \tag{1}$$

(b_2): $$s_i'(x_i) = s_{i+1}'(x_i) \qquad b_i + c_i(x_i - x_i) + \frac{d_i}{2}(x_i - x_i)^2 = b_{i+1} + c_{i+1}(x_i - x_{i+1}) + \frac{d_{i+1}}{2}(x_i - x_{i+1})^2$$

$$b_i = b_{i+1} - c_{i+1}h + \frac{d_i}{2}h^2 \qquad i = 1, \ldots, n-1 \tag{2}$$

(b$_3$): $\quad s''_i(x_i) = s''_{i+1}(x_i) \qquad c_i + d_i(x_i - x_i) = c_{i+1} + d_{i+1}(x_i - x_{i+1})$

$$c_i = c_{i+1} - d_{i+1}h \qquad i = 1, \ldots, n-1. \tag{3}$$

Na račun dva stepena slobode koji su ostali, treba dodati neka dva nova uslova. Postoji nekoliko običnih načina da se to izvede, a mi ćemo primijeniti jedan od načina. Možemo smatrati da funkcija f zadovoljava $f''(a) = 0$ i $f''(b) = 0$ (ponekad se kaže da su ovo dva granična uslova). Tada možemo tražiti da bude $s''(a) = 0$ i $s''(b) = 0$. Dakle, dopunimo definiciju splajna sljedećim uslovom: (d) $s''(a) = 0$ i $s''(b) = 0$. Izrazimo sada ova dva nova uslova preko naših oznaka. $s''(a) = 0$ znači $s''_1(x_0) = 0$, tj. $c_1 + d_1(x_0 - x_1) = 0$ ili

$$c_1 - d_1 h = 0. \tag{4}$$

$s''(b) = 0$ znači $s''_n(x_n) = 0$, tj. $c_n + d_n(x_n - x_n) = 0$ ili

$$c_n = 0. \tag{5}$$

Ima $3n$ jednačina (1)–(5) i $3n$ nepoznatih $\{b_i, c_i, d_i\}_{i=1}^n$. Zapišimo (3) i (4) zajedno kao

$$c_i = c_{i+1} - d_{i+1}h \qquad i = 0, \ldots, n-1, \tag{6}$$

stavljajući da je

$$c_0 = 0, \tag{7}$$

gdje je uvedena pomoćna promjenljiva c_0.

Razmotrimo sada (1)–(2), (5)–(7). Želimo da dobijemo sistem u kome se pojavljuju samo $\{c_i\}_{i=0}^n$:

prepišimo (1) $\qquad b_i h - \dfrac{c_i}{2}h^2 + \dfrac{d_i}{6}h^3 = f_i - f_{i-1} \qquad i = 1, \ldots, n$

pomjerimo indeks za jedan naviše $\qquad b_{i+1}h - \dfrac{c_{i+1}}{2}h^2 + \dfrac{d_{i+1}}{6}h^3 = f_{i+1} - f_i \qquad i = 0, \ldots, n-1$

oduzmimo dvije relacije:

$$(b_{i+1} - b_i)h - \frac{c_{i+1} - c_i}{2}h^2 + \frac{d_{i+1} - d_i}{6}h^3 = f_{i+1} - 2f_i + f_{i-1} \qquad i = 1, \ldots, n-1 \qquad \Big/ : h$$

Mi uvrstimo $b_{i+1} - b_i = \ldots$ u relaciju (2) čime

$$c_{i+1}h - \frac{d_{i+1}}{2}h^2 = \frac{c_{i+1} - c_i}{2}h - \frac{d_{i+1} - d_i}{6}h^2 + \frac{f_{i+1} - 2f_i + f_{i-1}}{h} \qquad i = 1, \ldots, n-1 \tag{$*_1$}$$

(b_j su eliminisani).

S druge strane, (6) govori da je $\qquad d_{i+1}h = c_{i+1} - c_i \qquad i = 0, \ldots, n-1 \qquad (*_2)$

isto $\qquad d_i h = c_i - c_{i-1} \qquad i = 1, \ldots, n \qquad (*_3)$

Supstitucijom $(*_2)$ i $(*_3)$ u relaciju $(*_1)$:

$$c_{i+1}h - \frac{h}{2}(c_{i+1} - c_i) = \frac{c_{i+1} - c_i}{2}h - \frac{c_{i+1} - c_i - c_i + c_{i-1}}{6}h + \frac{f_{i+1} - 2f_i + f_{i-1}}{h} \qquad \Big/ \cdot \frac{6}{h}$$

(i d_j su eliminisani)

$$6c_{i+1}-3(c_{i+1}-c_i)=3(c_{i+1}-c_i)-(c_{i+1}-2c_i+c_{i-1})+\frac{6}{h^2}(f_{i+1}-2f_i+f_{i-1}) \qquad i=1,\dots,n-1$$

$$\begin{cases} c_{i-1}+4c_i+c_{i+1}=\dfrac{6}{h^2}(f_{i+1}-2f_i+f_{i-1}), \qquad i=1,\dots,n-1, \\ \qquad\qquad c_0=0, \quad c_n=0. \end{cases} \tag{8}$$

Sistem linearnih jednačina (8) ima $n+1$ jednačina i ima $n+1$ nepoznatih $\{c_i\}_{i=0}^{n}$. Taj sistem ima jedinstveno rješenje, kako će kasnije biti pokazano (uzimamo obavezu).

Kada su $\{c_i\}_{i=0}^{n}$ određeni onda se na osnovu (6) neposredno izračunaju svi $\{d_i\}_{i=1}^{n}$. Zatim se na osnovu (1) dobiju svi $\{b_i\}_{i=1}^{n}$. Sada su svi polinomi $\{s_i(x)\}_{i=1}^{n}$ određeni. Mi smo riješili zadatak o interpolaciji pomoću kubnog splajna koji se opisuje uslovima (a)–(d). Bliže rečeno, dokazano je postojanje i jedinstvenost rješenja $s=s(x)$ i dat je postupak za njegovu konstrukciju: formirati sistem (8), riješiti taj sistem, itd.

Numerička metoda je konstruisana.

Samo napominjemo da se ponekad umjesto para uslova $f''(a)=0$, $f''(b)=0$ (koji vode do tzv. prirodnog kubnog splajna) uzima par uslova $f'(a)=0$, $f'(b)=0$ ili $f(a)=f(b)$, $f'(a)=f'(b)$ ili nešto slično.

Kao primjer, prikazaćemo sistem (8) u slučaju $n=6$. Može se smatrati da su nepoznate $\{c_i\}_{i=0}^{n}$ ili se može smatrati da su nepoznate $\{c_i\}_{i=1}^{n-1}$. Tako da je matrica sistema linearnih jednačina dimenzije 7×7 ili 5×5:

$$\begin{array}{lcl}
c_0=0 & & \\
c_0+4c_1+c_2=\dots & & 4c_1+c_2=\dots \\
c_1+4c_2+c_3=\dots & & c_1+4c_2+c_3=\dots \\
c_2+4c_3+c_4=\dots & \text{ili} & c_2+4c_3+c_4=\dots \\
c_3+4c_4+c_5=\dots & & c_3+4c_4+c_5=\dots \\
c_4+4c_5+c_6=\dots & & c_4+4c_5=\dots \\
c_6=0 & &
\end{array}$$

$$M=\begin{bmatrix} 1&0&0&0&0&0&0\\ 1&4&1&0&0&0&0\\ 0&1&4&1&0&0&0\\ 0&0&1&4&1&0&0\\ 0&0&0&1&4&1&0\\ 0&0&0&0&1&4&1\\ 0&0&0&0&0&0&1 \end{bmatrix} \qquad \text{ili} \qquad M=\begin{bmatrix} 4&1&0&0&0\\ 1&4&1&0&0\\ 0&1&4&1&0\\ 0&0&1&4&1\\ 0&0&0&4&1 \end{bmatrix}.$$

Sada ćemo razdužiti obavezu (o jedinstvenosti rješenja).

Definicija. Za kvadratnu matricu $A=[a_{ij}]\in R^{n\times n}$ kaže se da je *dijagonalno dominantna* (po vrstama) ako za svako $i=1,\dots,n$ važi nejednakost $|a_{ii}|>\sum_{j=1\quad j\neq i}^{n}\quad|a_{ij}|$.

Suma po $j=1,\dots,i-1,i+1,\dots,n$.

Vidimo da je matrica M sistema (8) dijagonalno dominantna jer je $4 > 1 + 1$.

Teorema. Ako je A dijagonalno dominantna onda je A regularna ($\det A \neq 0$).

Dokaz. Razmotrimo homogeni sistem linearnih jednačina $A\mathbf{x} = 0$, gdje je $\mathbf{x} = (x_1, \ldots, x_n)$, kao vektor–kolona, tj. razmotrimo

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0 \qquad \text{za} \qquad i = 1, \ldots, n.$$

Dopustimo da razmatrani sistem ima netrivijalno rješenje $(x_1, x_2, \ldots, x_n) \neq (0, 0, \ldots, 0)$. Među brojevima $|x_1|, |x_2|, \ldots, |x_n|$ uočimo najveći, odnosno ako ima više jednakih najvećih uočimo bilo koji od njih. Neka uočenom broju odgovara indeks ℓ, odnosno neka je $|x_\ell|$ najveći; $|x_i| \leq |x_\ell|$ za $i = 1, \ldots, n$. Imamo da je $|x_\ell| > 0$, tj. da je $x_\ell \neq 0$. Dopušteno rješenje naravno zadovoljava svih n jednačina sistema, pa posebno zadovoljava i ℓ–tu jednačnu (jednačinu $i = \ell$), tako da je

$$a_{\ell 1}x_1 + \ldots + a_{\ell n}x_n = 0$$

$$a_{\ell\ell}x_\ell = -a_{\ell 1}x_1 - \ldots - a_{\ell,\ell-1}x_{\ell-1} - a_{\ell.\ell+1}x_{\ell+1} - \ldots - a_{\ell n}x_n \qquad |a+b| \leq |a| + |b|$$

$$|a_{\ell\ell}| \cdot |x_\ell| \leq |a_{\ell 1}| \cdot |x_1| + \ldots + |a_{\ell,\ell-1}| \cdot |x_{\ell-1}| + |a_{\ell.\ell+1}| \cdot |x_{\ell+1}| + \ldots + |a_{\ell n}| \cdot |x_n|$$

$$|a_{\ell\ell}| \cdot |x_\ell| \leq |a_{\ell 1}| \cdot |x_\ell| + \ldots + |a_{\ell,\ell-1}| \cdot |x_\ell| + |a_{\ell.\ell+1}| \cdot |x_\ell| + \ldots + |a_{\ell n}| \cdot |x_\ell| \qquad / : |x_\ell|$$

$$|a_{\ell\ell}| \leq |a_{\ell 1}| + \ldots + |a_{\ell,\ell-1}| + |a_{\ell.\ell+1}| + \ldots + |a_{\ell n}|.$$

A uslov dijagonalne dominantnosti za ℓ–tu vrstu (za vrstu $i = \ell$) govori upravo suprotno da je $|a_{\ell\ell}| > |a_{\ell 1}| + \ldots + |a_{\ell,\ell-1}| + |a_{\ell.\ell+1}| + \ldots + |a_{\ell n}|$. Dobili smo kontradikciju. Ne može postojati netrivijalno (nenulto) rješenje. Homogeni sistem ima dakle samo nulto rješenje. Pokazali smo da je $\det A \neq 0$. Teorema je dokazana.

Zapaziti da je matrica sistema (8) trodijagonalna, što predstavlja jedno značajno dobro svojstvo razmatrane numeričke metode.

Definicija. Za kvadratnu matricu $A = [a_{ij}] \in R^{n\times n}$ kaže se da je *trodijagonalna* ako je ispunjen sljedeći uslov: $a_{ij} \neq 0 \Rightarrow j = i - 1$ ili $j = i$ ili $j = i + 1$.

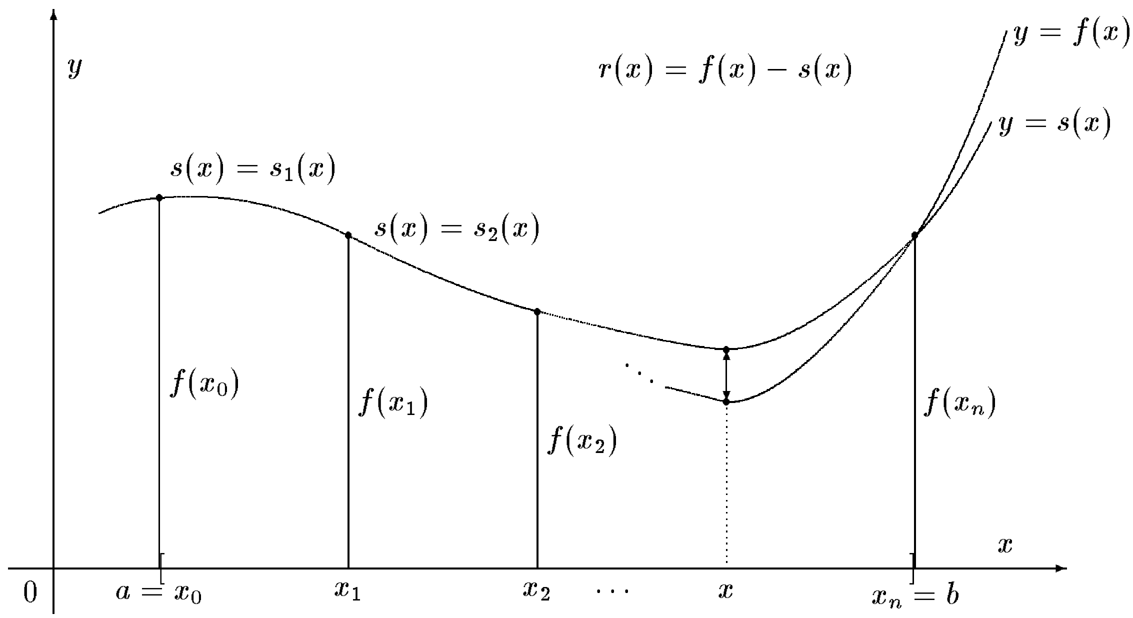
Ako je matrica sistema linearnih jednačina trodijagonalna onda to omogućava da se sistem znatno lakše (brže) riješi postupkom uzastopne eliminacije nepoznatih, zbog prisustva velikog broja nula. Poznato je sljedeće o vremenskom trošku rješavanja sistema linearnih jednačina oblika $n \times n$ (trošak se mjeri potrebnim brojem izvršenih aritmetičkih operacija): za opšti ili puni sistem trošak iznosi $t_n = O(n^3)$, a za trodijagonalni iznosi svega $t_n = O(n)$.

Sljedeću teoremu o ocjeni greške navodimo bez dokaza.

Teorema. Neka $f \in C^4[a, b]$. Neka je $M_4 = \max\limits_{a\leq x\leq b} |f^{IV}(x)|$. Tada važe nejednakosti:

$$\max_{a\leq x\leq b} |f(x) - s(x)| \leq M_4 h^4, \quad \max_{a\leq x\leq b} |f'(x) - s'(x)| \leq M_4 h^3 \quad \text{i} \quad \max_{a\leq x\leq b} |f''(x) - s''(x)| \leq M_4 h^2.$$

Vidimo da za svako fiksirano $x \in [a, b]$ važi relacija $s(x) \to f(x)$ kad $h \to 0$, odnosno kad $n \to \infty$. Zato se kaže da razmatrana numerička metoda konvergira; kad korak mreže $h \to 0$ onda greška metode $r(x) = f(x) - s(x) \to 0$. Kaže se da metoda ima četvrti red ili stepen konvergencije, budući da je $r(x) = O(h^4)$ kad $h \to 0$. Značajno je i što $s'(x) \to f'(x)$, a isto tako i što $s''(x) \to f''(x)$. Drugim riječima, izvod splajna može dobro da posluži za aproksimaciju izvoda funkcije, a drugi izvod splajna za aproksimaciju drugog izvoda funkcije.



1.9. NUMERIČKO DIFERENCIRANJE

Na osnovu date tablice vrijednosti funkcije treba procijeniti vrijednost izvoda funkcije u nekoj tački. U ovoj sekciji biće izvedene formule za numeričko diferenciranje, biće dobijen odgovarajući izraz za grešku i biće navedeni primjeri formula za numeričko diferenciranje. Neka je $n \geq 1$ i razmotrimo na realnoj osi $n+1$ međusobno različitih tačaka $x_0, \ldots, x_n$. Neka su poznate vrijednosti funkcije f u čvorovima, tj. neka su date brojne vrijednosti $f(x_i) = f_i \in R$. Neka $x \in R$ i neka je $k \geq 1$. Treba procijeniti $f'(x)$ ili uopšte $f^{(k)}(x)$. Stavimo $a = \min(x, x_0, \ldots, x_n)$ i $b = \max(x, x_0, \ldots, x_n)$. Pretpostavlja se da $f \in C^{n+k+1}[a,b]$.

Neka je $L_n = L_n(x)$ L. i. p. za f po mreži $\{x_i\}_{i=0}^n$. Mi ćemo uzeti da je $f'(x) \approx L_n'(x)$ i uopšte da je $f^{(k)}(x) \approx L_n^{(k)}(x)$. Nema teškoća da se izvodi polinoma $L_n(x)$ izračunaju tačno. Kaže se da se razmatraju formule za numeričko diferenciranje koje su interpolacionog tipa. Kasnije će biti dati primjeri konkretnih realizacija takvih formula, u zavisnosti od mreže, od x i od k (u zavisnosti od podataka).

Prelazimo na ocjenu greške formule $f^{(k)}(x) \approx L_n^{(k)}(x)$, tj. na dobijanje izraza za grešku $r(x) = f^{(k)}(x) - L_n^{(k)}(x)$. Odranije iz sekcije 1.4. znamo formule (1) i (2), kako su tamo bile numerisane:

$$f(x) - L_n(x) = f[x, x_0, \ldots, x_n]\omega_{n+1}(x) \quad \text{i} \quad f[x, x_0, \ldots, x_n] = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \ \ \xi \in (a,b);$$

oznaka $\omega_{n+1}(x) = \prod_{i=0}^n (x - x_i)$. Primijenimo na $f(x) - L_n(x) = f[x, x_0, \ldots, x_n]\omega_{n+1}(x)$ Lajbnicovu formulu za k–ti izvod proizvoda $(u(x)v(x))^{(k)} = \sum_{j=0}^k \binom{k}{j} u^{(j)}(x)v^{(k-j)}(x)$:

$$r(x) = f^{(k)}(x) - L_n^{(k)}(x) = \sum_{j=0}^{k} \binom{k}{j} (f[x, x_0, \ldots, x_n])^{(j)} \omega_{n+1}^{(k-j)}(x).$$

Razmotrimo izraz

$$A = f[x, x+\varepsilon, \ldots, x+j\varepsilon, x_0, \ldots, x_n].$$

S jedne strane, A predstavlja podijeljenu razliku reda j funkcije $f[t, x_0, \ldots, x_n]$ po mreži čvorova $t = x$, $t = x+\varepsilon$, ..., $t = x+j\varepsilon$. Formula oblika formule (2) iz sekcije 1.4. govori o vezi podijeljenih razlika i izvoda funkcije i daje nam $A = \dfrac{1}{j!}(f[\xi(\varepsilon), x_0, \ldots, x_n])^{(j)}$, gdje je $x < \xi(\varepsilon) < x + j\varepsilon$. Kad $\varepsilon \to 0$ onda je očito $\lim_{\varepsilon \to 0} \xi(\varepsilon) = x$, tako da postoji i $\lim_{\varepsilon \to 0} A$;

$$\lim_{\varepsilon \to 0} A = \frac{1}{j!}(f[x, x_0, \ldots, x_n])^{(j)}.$$

S druge strane, A predstavlja podijeljenu razliku reda $n+j+1$ funkcije $f = f(t)$ po naznačenih $n+j+2$ čvorova $t = x$, $t = x+\varepsilon$, ..., $t = x_n$. Opet upotrebimo poznatu formulu o vezi podijeljenih razlika i izvoda funkcije: $A = \dfrac{1}{(n+j+1)!} f^{(n+j+1)}(\xi_j(\varepsilon))$, gdje je $a < \xi_j(\varepsilon) < b$. Uvedimo oznake $m_1 = \min_{t \in [a,b]} f^{(n+j+1)}(t)$ i $m_2 = \max_{t \in [a,b]} f^{(n+j+1)}(t)$. Tako da je $m_1 \leq f^{(n+j+1)}(\xi_j(\varepsilon)) \leq m_2$. Maločas smo pokazali da postoji $\lim_{\varepsilon \to 0} A$, tako da znamo da postoji i $\lim_{\varepsilon \to 0} f^{(n+j+1)}(\xi_j(\varepsilon))$. Vidimo da je $m_1 \leq \lim_{\varepsilon \to 0} f^{(n+j+1)}(\xi_j(\varepsilon)) \leq m_2$. Funkcija $f^{(n+j+1)}$ je neprekidna na intervalu $[a,b]$. Poznata je teorema da neprekidna na zatvorenom intervalu funkcija uzima sve svoje međuvrijednosti $m \in [m_1, m_2]$. Zato postoji tačka $\xi_j \in [a,b]$ takva da je

$$\lim_{\varepsilon \to 0} A = \frac{1}{(n+j+1)!} f^{(n+j+1)}(\xi_j).$$

Uporediti dva prikazivanja za $\lim_{\varepsilon \to 0} A$. Prema tome

$$r(x) = f^{(k)}(x) - L_n^{(k)}(x) = \sum_{j=0}^{k} \binom{k}{j} \frac{j!}{(n+j+1)!} f^{(n+j+1)}(\xi_j) \omega_{n+1}^{(k-j)}(x), \tag{1}$$

$\xi_j \in [a,b]$ za $j = 0, \ldots, k$. Može se pokazati da je $a < \xi_j < b$.

Formula (1) važi za svako $x \in [a,b]$, tj. ona važi i kada se tačka x poklapa sa nekim čvorom (kada je $x = x_i$ za neko i), s tim da tada prethodno izvođenje treba da bude malo prilagođeno. Samo se napominje da važi $\lim_{\varepsilon \to 0} f[x, x+\varepsilon, \ldots, x+k\varepsilon] = \frac{1}{k!} f^{(k)}(x)$, ako $k \geq 1$, f dovoljno glatka, što može da posluži kao definicija $f[\underbrace{x, \ldots, x}_{k+1}]$.

$$f^{(k)}(x) - L_n^{(k)}(x) = \sum_{j=0}^{k} \frac{k!}{(k-j)!(n+j+1)!} f^{(n+j+1)}(\xi_j) \omega_{n+1}^{(k-j)}(x), \text{ jer } \binom{k}{j} = \frac{k!}{j!(k-j)!}$$

$$\left| f^{(k)}(x) - L_n^{(k)}(x) \right| \leq \sum_{j=0}^{k} \frac{k!}{(k-j)!(n+j+1)!} M_{n+j+1} \left| \omega_{n+1}^{(k-j)}(x) \right|, \quad M_{n+j+1} = \max_{t \in [a,b]} \left| f^{(n+j+1)}(t) \right|$$

Prelazimo na primjere. Biće navedena tri primjera, sva tri se odnose na slučaj ekvidistantne mreže čvorova.

1. Formula za prvi izvod u čvoru, jednostrana formula. Definišimo podatke koji odgovaraju ovom specijalnom slučaju. Imamo $n+1$ čvor $x_i = x_0 + ih$, gdje je $i = 0, \ldots, n$, a treba da se procijeni $f'(x_0)$. Treba napisati izraz za odgovarajući L. i. p. Zatim treba diferencirati taj izraz i onda naravno treba uvrstiti $x = x_0$. Onda još samo treba konkretizovati formulu za grešku (1).

Tokom rada, prikazaćemo L. i. p. $L_n = L_n(x)$ u obliku I Nj. i. f, uz upotrebu obične smjene $x = x_0 + ht$:

$$L_n(x) = L_n(x_0 + ht) =$$

$$f_0 + \Delta f_0 t + \frac{1}{2!} \Delta^2 f_0 t(t-1) + \ldots + \frac{1}{n!} \Delta^n f_0 t(t-1) \ldots (t-n+1);$$

$$\frac{dy}{dt} = h \frac{dy}{dx}, \qquad L_n'(x) = L_n'(x_0 + ht) =$$

$$\frac{1}{h} \left(\Delta f_0 \frac{d}{dt} t + \frac{1}{2!} \Delta^2 f_0 \frac{d}{dt} t(t-1) + \ldots + \frac{1}{n!} \Delta^n f_0 \frac{d}{dt} t(t-1) \ldots (t-n+1) \right);$$

$$t = 0, \qquad L_n'(x_0) = \frac{1}{h} \left(\Delta f_0 - \frac{1}{2} \Delta^2 f_0 + \ldots + \frac{(-1)^{n+1}}{n} \Delta^n f_0 \right). \tag{2}$$

Za $k = 1$ (1) glasi

$$f'(x) - L'_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_0)\omega'_{n+1}(x) + \frac{1}{(n+2)!} f^{(n+2)}(\xi_1)\omega_{n+1}(x), \quad \xi_0, \xi_1 \in (a,b).$$

Da se ocijeni greška za formulu (2), u formuli (1) se stavi $k = 1$ i $x = x_0$. Znamo da je $\omega_{n+1}(x) = \prod_{i=0}^{n}(x - x_i)$. Vidi se da je $\omega_{n+1}(x_0) = 0$. Izračunati $\omega'_{n+1}(x)$ i $\omega'_{n+1}(x_0)$. Kada se sprovede jednostavni račun onda se dobije

$$f'(x_0) - L'_n(x_0) = \frac{(-1)^n}{n+1} f^{(n+1)}(\xi) h^n \quad \text{za neko } \xi \in (x_0, x_n). \tag{3}$$

Tako da (2) predstavlja formulu za numeričko diferenciranje, u smislu $f'(x_0) \approx L'_n(x_0)$, a (3) izražava njenu grešku. Navedimo neke konkretne slučajeve formule (2):

$$n = 1: \qquad f'(x_0) \approx \frac{1}{h}\Delta f_0 = \frac{f(x_1) - f(x_0)}{h} \tag{4}$$

$$\text{ili} \quad f'(x_i) = \frac{1}{h}\Big(f(x_{i+1}) - f(x_i)\Big) - \frac{h}{2} f''(\xi), \quad x_i < \xi < x_{i+1}, \quad \text{ako } f \in C^2[x_i, x_{i+1}],$$

$$n = 2: \qquad f'(x_0) \approx \frac{1}{h}\left(\Delta f_0 - \frac{1}{2}\Delta^2 f_0\right) = \frac{-f(x_2) + 4f(x_1) - 3f(x_0)}{2h}.$$

2. Formula za prvi izvod u tački koja se nalazi na sredini između dva čvora, simetrična formula. Razmotrimo mrežu čvorova $x_i = x_0 + ih$, gdje je $-(\ell - 1) \le i \le \ell$ i razmotrimo tačku $\overline{x} = x_{1/2} = x_0 + \frac{h}{2}$. Recimo. kada je $\ell = 2$ onda mrežu čine čvorovi $\{x_{-1}, x_0, x_1, x_2\}$ i tada je $\omega_4(x) = (x - x_{-1}) \cdot \ldots \cdot (x - x_2)$. Izostavlja se odgovarajući račun. Važi:

$$f'(\overline{x}) = f'\left(x_0 + \frac{h}{2}\right) = \frac{1}{h}\sum_{j=0}^{\ell-1} \frac{(-1)^j}{(2j+1)!}\left(\frac{1}{2}\cdot\frac{3}{2}\cdot\ldots\cdot\left(j - \frac{1}{2}\right)\right)^2 \delta^{2j+1} f_{1/2} +$$

$$\frac{(-1)^\ell}{(2\ell+1)!}\left(\frac{1}{2}\cdot\frac{3}{2}\cdot\ldots\cdot\left(\ell - \frac{1}{2}\right)\right)^2 f^{(2\ell+1)}(\xi) h^{2\ell}$$

u smislu "tačno = približno + greška".

Za grešku: $\omega_{2\ell}(x) = \prod(x - x_i)$ (u proizvodu: $i = -(\ell - 1)$ do $i = \ell$), $\omega'_{2\ell}(\overline{x}) = 0$. Neki slučajevi:

$$\ell = 1: \qquad f'\left(x_0 + \frac{h}{2}\right) \approx \frac{1}{h}\delta f_{1/2} = \frac{1}{h}\Big(f(x_1) - f(x_0)\Big) \tag{5}$$

$$\text{ili} \quad \left(\text{umjesto } \frac{h}{2} \text{ pisati } h\right) \quad f'(x_i) = \frac{1}{2h}\Big(f(x_{i+1}) - f(x_{i-1})\Big) - \frac{h^2}{6} f'''(\xi),$$

$$\ell = 2: \qquad f'\left(x_0 + \frac{h}{2}\right) \approx \frac{1}{h}\left(\delta f_{1/2} - \frac{1}{24}\delta^3 f_{1/2}\right) =$$

$$\frac{1}{24h}\Big(-f(x_2) + 27f(x_1) - 27f(x_0) + f(x_{-1})\Big).$$

3. Drugi izvod u čvoru, simetrična formula. $f''(x_0)$ aproksimira se preko vrijednosti funkcije u čvorovima $x_{-\ell}$, $x_{-(\ell-1)}$, …, x_ℓ (ima ih $2\ell+1$), gdje je $x_i = x_0 + ih$ za $i = -\ell, -(\ell-1), \dots, \ell$:

$$f''(x_0) = \frac{1}{h^2}\sum_{j=1}^{\ell} \frac{2(-1)^{j-1}}{(2j)!}((j-1)!)^2 \delta^{2j} f_0 + \frac{2(-1)^\ell}{(2\ell+2)!}(\ell!)^2 f^{(2\ell+2)}(\xi) h^{2\ell}$$

(tačno = približno + greška). Za grešku:

$$\omega_{2\ell+1}(x) = \prod_{i=-\ell}^{\ell}(x - x_i), \quad \omega_{2\ell+1}(x_0) = 0,$$

$$\omega'_{2\ell+1}(x_0) = (-1)^\ell (\ell!)^2 h^{2\ell}, \quad \omega''_{2\ell+1}(x_0) = 0.$$

Ako je $\ell = 1$ onda $\quad f''(x_0) \approx \frac{1}{h^2}\delta^2 f_0 = \frac{1}{h^2}\Big(f(x_1) - 2f(x_0) + f(x_{-1})\Big) \qquad (6)$

ili $\quad f''(x_i) = \frac{1}{h^2}\Big(f(x_{i+1}) - 2f(x_i) + f(x_{i-1})\Big) - \frac{h^2}{12} f^{IV}(\xi).$

Ako je $\ell = 2$ onda $\quad f''(x_0) \approx \frac{1}{h^2}\left(\delta^2 f_0 - \frac{1}{12}\delta^4 f_0\right) =$

$$\frac{1}{12h^2}\Big(-f(x_2) + 16f(x_1) - 30f(x_0) + 16f(x_{-1}) - f(x_{-2})\Big).$$

U zaključku, ako se po ekvidistantnoj mreži koja se sastoji od $n+1$ čvorova procjenjuje k–ti izvod funkcije u nekoj tački x onda za grešku važi relacija $r(x) = O(h^{n-k+1})$ za male vrijednosti koraka mreže h. Međutim, ako je tačka x postavljena simetrično u odnosu na čvorove onda se okolnosti poboljšavaju (greška se smanjuje); u tom slučaju važi $r(x) = O(h^{n-k+2})$, red aproksimacije se poveća za jedan. Takvo svojstvo imaju formule iz drugog i trećeg primjera, za $f'(\overline{x})$ i za $f''(x_0)$.

U zaključku, za formule (4)–(6) kaže se da predstavljaju osnovne formule za numeričko diferenciranje. Postoji lak i neposredan način da se one dokažu. Izvršiti razvoj funkcije f po Tejlorovoj formuli do odgovarajućeg izvoda u okolini tačke x. U okviru toga, i odgovarajući izrazi za grešku mogu da budu provjereni (izvedeni).

1.10. NESTABILNOST NUMERIČKOG DIFERENCIRANJA. TRI VRSTE GREŠKE U NUMERIČKIM METODAMA

Mi ćemo sada razmotriti pojam nestabilnosti formula za numeričko diferenciranje (nestabilnosti u odnosu na približnost ulaznih veličina) na jednom jednostavnom primjeru formule za numeričko diferenciranje (na jednostavnom modelu). Tako da izlaganje neće biti tehnički opterećeno, a sve karakteristike razmatrane pojave (nestabilnosti) lijepo se vide.

Neka imamo dva čvora x_0 i $x_1 = x_0 + h$ i dvije odgovarajuće vrijednosti funkcije $f(x_0) = f_0$ i $f(x_1) = f_1$ i neka treba da se procijeni $f'(x_0)$. Kako je rađeno:

$$f'(x_0) \approx \frac{1}{h}\Big(f_1 - f_0\Big) \qquad (1)$$

$$f'(x_0) = \frac{1}{h}\Big(f_1 - f_0\Big) + r_1 \qquad r_1 = -\frac{1}{2}f''(\xi)h \qquad x_0 < \xi < x_1$$

$$|r_1| \leq \frac{1}{2}M_2 h \qquad M_2 = \max_{t\in[a,b]} |f''(t)| \qquad x_0, x_1 \in [a,b] \qquad f \in C^2[a,b].$$

Za r_1 se kaže da predstavlja grešku ili grešku metode.

Neka su f_0 i f_1 ulazne veličine. Uzmimo sada da umjesto sa tačnim vrijednostima ulaznih veličina f_0 i f_1 raspolažemo samo sa odgovarajućim približnim vrijednostima f_0^* i f_1^*. Uvedimo oznake za dvije odgovarajuće greške. Neka bude $\varepsilon_0 = f_0 - f_0^*$ i $\varepsilon_1 = f_1 - f_1^*$. Neka je poznata granica greške ulaznih veličina u oznaci E i to $|\varepsilon_0| = |f_0 - f_0^*| \leq E$ i $|\varepsilon_1| = |f_1 - f_1^*| \leq E$. Za E se kaže da predstavlja mjeru greške ulaznih veličina (ulaznih podataka).

U datim okolnostima, mi možemo da efektivno izračunamo jedino broj $t^{**} = \dfrac{1}{h}\Big(f_1^* - f_0^*\Big)$, gdje u računu ušestvuju približne vrijednosti f_0^* i f_1^*. Tačan broj $t = f'(x_0)$ je nedostižan. Nedostižan je i približni broj $t^* = \dfrac{1}{h}\Big(f_1 - f_0\Big)$. Njegovu ulogu preuzima približni broj t^{**} koji predstavlja numerički odgovor (predstavlja rezultat), budući da numerički odgovor glasi $t \approx t^{**}$.

Neka bude $r_2 = t^* - t^{**}$. Za r_2 se kaže da predstavlja grešku izazvanu približnošću ulaznih veličina. Ponekad se za r_2 kaže da predstavlja neotklonjivu grešku, imajući u vidu da nismo u stanju da otklonimo grešku ulaznih veličina (nismo u stanju saznati njihove tačne vrijednosti), tako da će ta greška "proći" kroz računski proces i (znači) odraziće se na krajnji numerički odgovor.

Na redu je procjena veličine r_2:

$$r_2 = t^* - t^{**} = \frac{1}{h}\Big(f_1 - f_0\Big) - \frac{1}{h}\Big(f_1^* - f_0^*\Big) = \frac{1}{h}\Big(f_1 - f_1^*\Big) - \frac{1}{h}\Big(f_0 - f_0^*\Big) = \frac{1}{h}\varepsilon_1 - \frac{1}{h}\varepsilon_0,$$

$$|r_2| \leq \frac{1}{h}|\varepsilon_1| + \frac{1}{h}|\varepsilon_0| \leq \frac{1}{h}E + \frac{1}{h}E = \frac{2E}{h}.$$

Ako je $E = 0$ onda je $r_2 = 0$. Što je E veće to je i r_2 veće.

Na redu je procjena greške numeričkog odgovora, procjena veličine $r = t - t^{**}$:

$$r = t - t^{**}, \qquad r = t - t^* + t^* - t^{**}, \qquad r = r_1 + r_2,$$

$$|r| \leq |r_1| + |r_2|, \qquad |r| \leq \frac{1}{2}M_2 h + \frac{2E}{h}.$$

Na r, r_1 i r_2 gledamo kao na funkcije od h.

Nije ispunjen uslov da $r_2 = r_2(h) \to 0$ kad $h \to 0$. Zato se kaže da je formula (1) nestabilna u odnosu na grešku ulaznih veličina.

Nacrtati grafik funkcije $g = g(x) = \dfrac{1}{2}M_2 x + \dfrac{2E}{x}$ za $x > 0$. Funkcija dostiže minimum za $x = 2\sqrt{E/M_2}$, a sama vrijednost minimuma iznosi $g(2\sqrt{E/M_2}) = 2\sqrt{M_2 E}$. Ako se kao korak h uzme $h = h_0 = 2\sqrt{E/M_2}$ onda se to može smatrati najpovoljnijim izborom koraka. Budući da je $|r(x)| \leq g(x)$ to je $|r(h_0)| \leq 2\sqrt{M_2 E}$.

Pogledajmo odnos između greške ulaznih veličina E i ukupne greške $r(h_0)$. Ako je $E = 10^{-2n}$ (ulazne veličine znamo sa $2n$ tačnih decimala) onda je greška jednaka 10^{-n}, grubo govoreći; $\sqrt{10^{-2n}} = 10^{-n}$. Dakle, od polaznog broja tačnih decimala, u rezultatu je ostalo (sačuvalo se) samo pola decimala tačnih, a pola se izgubilo. Drugim riječima, greška rezultata je znatno veća od greške ulaznih veličina; nepovoljna okolnost.

Prelazimo na opšti slučaj formule za numeričko diferenciranje. Greška formule za numeričko diferenciranje obično ima oblik

$$r = r_1 + r_2 \quad \text{sa} \quad r_1 \sim C_1 h^{n-k+1} \quad \text{i} \quad r_2 \sim C_2 h^{-k},$$

gdje se računa približna vrijednost k–tog izvoda u nekoj tački po ekvidistantnoj mreži čvorova; h je korak mreže, a čvorova ima ukupno $n+1$ na broju. Uporedi sa primjerima iz prethodne sekcije. Tako da negativno svojstvo nestabilnosti imamo kod svake (kod praktično svake) formule za numeričko diferenciranje.

Pogledajmo jedan mali primjer. Neka bude $y = y(x) = e^x$ i $x_0 = 0$, $x_1 = 0{,}1$, $x_2 = 0{,}2$, $y_0 = y(x_0)$, $y_1 = y(x_1)$, $y_2 = y(x_2)$. Tada je $y_0 = 1$, $y_1 = 1{,}10517$, $y_2 = 1{,}22140$; prikazati u obliku tabele. Neka na osnovu nabrojanih podataka treba da bude procijenjen $y''(x_1)$. Primjenom formule iz prethodne sekcije dobijamo sljedeću procjenu:

$$\frac{y_2 - 2y_1 + y_0}{h^2} = \frac{1{,}22140 - 2 \cdot 1{,}10517 + 1}{0{,}1^2} = 1{,}106.$$

Sada imamo jednostavne opservacije. Ulazni podaci y_0, y_1 i y_2 imaju po šest značajnih cifara, a rezultat 1,106 ima svega četiri značajne cifre; došlo je do gubitka dvije značajne cifre. Ulazni podaci imaju grešku reda 10^{-5}, a rezultat ima grešku reda 10^{-3}. Dakle, rezultat ima znatno veću i apsolutnu i relativnu grešku od ulaznih podataka.

Prelazimo na dio: tri vrste greške u numeričkim metodama. Za bilo koju numeričku metodu, greška ili ukupna greška r računa se kao $r = r_1 + r_2 + r_3$, gdje se za treću komponentu r_3 kaže da predstavlja grešku računanja ili grešku operacija. Iz samog naziva je jasno kako se formira r_3. Naime, tokom računanja po datom obrascu redom se izvode naznačene aritmetičke operacije. Rezultat pojedine aritmetičke operacije je približan broj (ponekad se kaže zaokružen broj), makar da su argumenti operacije tačni brojevi. Tako da ukupna računanja po datom obrascu unose dodatnu grešku r_3.

Znamo da računar izvodi aritmetičke operacije (i druge operacije) sa realnim brojevima samo približno tačno.

Od metode do metode, treba voditi računa o sve tri komponente greške: greška metode + greška izazvana približnošću ulaznih veličina + greška računanja.

U zaključku, ako je metoda nestabilna u odnosu na grešku ulaznih veličina onda se to po pravilu posebno naglasi. Ako već ne možemo da izbjegnemo upotrebu takve metode onda treba dobro pratiti (kontrolisati) ponašanje greške r_2. S druge strane, za stabilnu metodu lako može biti da je r_2 zanemarljivo mala u odnosu na grešku metode r_1.

U zaključku, često je realno smatrati da je greška računanja r_3 zanemarljivo mala u odnosu na grešku metode r_1.

Na primjer, kod L. i. p. smo radili kao da je $r_2 = 0$ i $r_3 = 0$, a umjesto "greška metode" govorili smo jednostavno "greška".

1.11. POJAM PRIBLIŽNOG BROJA

Neka je a tačna vrijednost neke veličine i neka je a^* poznata približna vrijednost te veličine. Greškom ili apsolutnom greškom približnog broja a^* naziva se veličina $A(a^*) = a - a^*$. Relativnom greškom približnog broja a^* naziva se veličina $R(a^*) = \frac{A(a^*)}{|a|} = \frac{a-a^*}{|a|}$. Ponekad se u literaturi ove dvije veličine uzimaju po modulu, pa bi se stavilo $A(a^*) = |a - a^*|$ i $R(a^*) = \frac{|a-a^*|}{|a|}$.

Kada piše recimo $a = 7{,}2 \pm 0{,}1$ onda to ima smisao $a^* = 7{,}2$ i $|A(a^*)| \le 0{,}1$, drukčije rečeno $7{,}1 \le a \le 7{,}3$. Relativna greška $R(a^*)$ se često izražava u procentima. Znamo da realni brojevi zapisani u memoriji računara na 32 bita imaju granicu relativne greške 10^{-7}.

U stvarnosti, mi ne znamo $A(a^*)$ niti $R(a^*)$, već znamo samo neke njihove ocjene, a bolje je da su te dvije ocjene što bliže veličinama $A(a^*)$ i $R(a^*)$. Tako se definišu i granica apsolutne greške $\Delta(a^*)$ i granica relativne greške $\delta(a^*)$ približnog broja a^* kao ma koji brojevi koji ispunjavaju $|a - a^*| \le \Delta(a^*)$ odnosno $\frac{|a-a^*|}{|a|} \le \delta(a^*)$.

Značajnom cifrom broja naziva se svaka cifra njegovog zapisa počev od prve nenulte cifre slijeva. Recimo, broj $a^* = 0{,}0\underline{3045}$ ima četiri značajne cifre, one su podvučene, a broj $a^* = 0{,}0\underline{3045000}$ ima sedam značajnih cifara. Za značajnu cifru se kaže da je sigurna (u užem smislu) ako apsolutna greška tog broja ne prevazilazi polovinu vrijednost pozicije koja odgovara toj cifri (polovinu težine dekadnog mjesta). Na primjer $a^* = 0{,}0\underline{3045}$ $\Delta(a^*) = 3 \cdot 10^{-6}$ broj a^* ima četiri sigurne cifre, one su podvučene. Na primjer $a^* = 0{,}0\underline{30450}00$ $\Delta(a^*) = 2 \cdot 10^{-7}$ broj a^* ima pet sigurnih cifara. Značajna cifra naziva se sigurnom u širem smislu ako apsolutna greška ne prevazilazi vrijednost odgovarajuće pozicije. Da pojasnimo: u primjeru $a \approx 0{,}03045$, vrijednost pozicije cifre 3 je 10^{-2}, vrijednost pozicije cifre 4 je 10^{-4}. Da pojasnimo, n–ta decimala nekog približnog broja a^* je sigurna ako $|A(a^*)| \le 0{,}5 \cdot 10^{-n}$, ona je sigurna u širem smislu ako $|A(a^*)| \le 10^{-n}$, samo još da je značajna.

Lako se vidi da je relativna greška približnog broja blisko povezana sa brojem njegovih sigurnih cifara: ako broj ima k sigurnih cifara onda njegova relativna greška iznosi 10^{-k}, grubo govoreći. Zaista, $a^* = 0{,}12345$ ima 5 sigurnih, relativna $\frac{1}{2} \cdot 10^{-4}$, $b^* = 0{,}92345$ ima 5 sigurnih, relativna $\frac{1}{2} \cdot 10^{-5}$. Uputstvo: posmatrajte veličinu $k' = -\log_{10} |R(a^*)|$.

Ako je približan broj prosto napisan sa recimo 3 decimale onda se podrazumijeva da je njegova granica greške 10^{-3} ili $\frac{1}{2} \cdot 10^{-3}$ (konvencija).

1.12. GREŠKA FUNKCIJE

Ako je $x = 4 \pm 0{,}1$ kolika se greška čini kada se kaže da je $\sqrt{x} \approx 2$? Ako je $x_1 = x_1^* \pm \Delta(x_1^*)$ i $x_2 = x_2^* \pm \Delta(x_2^*)$ kolika se greška čini kada se uzme da približan broj $f(x_1^*, x_2^*)$ zamjenjuje tačnu vrijednost $f(x_1, x_2)$, recimo da je $f(x_1, x_2) = x_1^2 x_2$. U nastavku će biti izvedena formula koja daje izraz za odgovarajuću grešku. Kaže se da je to formula za grešku funkcije, da je to formula za grešku u slučaju kada su argumenti funkcije približni brojevi.

Pogledajmo prvo slučaj funkcije od jedne promjenljive $f\colon [a, b] \to R$. Neka je $x \in [a, b]$ tačna vrijednost argumenta, neka je $x^* \in [a, b]$ raspoloživa približna vrijednost i neka je $\Delta(x^*)$ granica greške, tj. neka je $|x - x^*| \le \Delta(x^*)$. Tada je $f(x)$ tačna vrijednost funkcije, a $f(x^*)$ je približna vrijednost. Neka je

$$A = f(x) - f(x^*).$$

Za A se kaže da je greška funkcije. Naš zadatak je da ocijenimo A. Po Lagranžovoj teoremi o srednjoj vrijednosti (po formuli o konačnim priraštajima) imamo

$$A = f'(\xi)(x - x^*), \quad \xi = x^* + \theta(x - x^*), \quad 0 < \theta < 1,$$

$$|A| = |f'(\xi)| \cdot |x - x^*|,$$

$$|A| \leq M_1|x - x^*|, \quad M_1 = \sup_{t\in G} |f'(t)|, \quad G = [x^* - \Delta(x^*), x^* + \Delta(x^*)].$$

Smatramo da je $G \subset [a,b]$, tako da se može svejedno uzeti da je $M_1 = \sup_{t\in[a,b]} |f'(t)|$:

$$|A| \leq M_1 \Delta(x^*). \tag{1}$$

Vidimo da formula (1) rješava postavljeni zadatak. Pretpostavili smo da $f \in C^1[a,b]$.

Dalje, zapažamo da ξ samo neznatno odstupa od x^*. Razmotrimo veličinu

$$L = |f'(x^*)|\Delta(x^*).$$

Možemo pisati $f'(\xi) \approx f'(x^*)$. Lako se vidi da važi nejednakost

$$|A| \leq L + o(\Delta(x^*)) \quad \text{kad} \quad \Delta(x^*) \to 0.$$

Tako da L može da posluži kao zadovoljavajuća zamjena za grešku $|A|$, kada je greška argumenta mala. L se lako računa i zato se, u praktičnom radu, veličina L često uzima kao zamjena za grešku. Za L se kaže da predstavlja linearnu ocjenu za grešku funkcije.

Primjer $f(x) = \sqrt{x}$ $x^* = 4$ $\Delta(x^*) = 0{,}1$. Tada je $|A| \leq 0{,}0252$ i $L = 0.025$. Zapažamo da L malo podbacuje.

Pogledajmo slučaj funkcije od više promjenljivih. Vidjećemo da važe slične okolnosti. Neka su x_i^* približne vrijednosti za x_i sa granicama greške $\Delta(x_i^*)$, tako da je $|x_i - x_i^*| \leq \Delta(x_i^*)$ za $i = 1,\ldots,n$. Uvedimo oznaku $G = [x_1^* - \Delta(x_1^*), x_1^* + \Delta(x_1^*)] \times \ldots \times [x_n^* - \Delta(x_n^*), x_n^* + \Delta(x_n^*)]$. Neka je funkcija f definisana na skupu G. Pretpostavimo da $f \in C^1(G)$. Naš zadatak je da ocijenimo grešku $A = f(x_1,\ldots,x_n) - f(x_1^*,\ldots,x_n^*)$. U tom cilju, uvedimo parametar θ za zatvoreni interval od tačke $(x_1^*,\ldots,x_n^*)$ do tačke $(x_1,\ldots,x_n)$ tako da $\theta = 0$ i $\theta = 1$ odgovaraju početku odnosno kraju intervala. Posmatrajmo funkciju f samo na tom intervalu. Primijenimo Lagranžovu teoremu o srednjoj vrijednosti. Kada se izvod u smjeru izrazi preko parcijalnih izvoda onda imamo

$$A = \sum_{i=1}^{n} \frac{\partial f(\xi_1,\ldots,\xi_n)}{\partial t_i}(x_i - x_i^*), \quad \xi_i = x_i^* + \theta(x_i - x_i^*), \quad i = 1,\ldots,n, \quad 0 < \theta < 1,$$

$$|A| \leq \sum_{i=1}^{n} \left| \frac{\partial f(\xi_1,\ldots,\xi_n)}{\partial t_i} \right| \cdot |x_i - x_i^*|,$$

$$|A| \leq \sum_{i=1}^{n} B_i \Delta(x_i^*), \quad B_i = \sup_{(t_1,\ldots,t_n)\in G} \left| \frac{\partial f(t_1,\ldots,t_n)}{\partial t_i} \right|. \tag{2}$$

Vidimo da formula (2) rješava postavljeni zadatak.

Dalje, kao praktična ocjena za grešku funkcije uzima se veličina

$$L = \sum_{i=1}^{n} \left| \frac{\partial f(x_1^*,\ldots,x_n^*)}{\partial t_i} \right| \Delta(x_i^*),$$

ovo je tzv. linearna ocjena za grešku funkcije.

Razmotrimo dva važna specijalna slučaja.

1. $f(t_1,\ldots,t_n) = t_1 + \ldots + t_n$. Greška zbira jednaka je zbiru grešaka sabiraka.

Ako se sabira 10 približnih brojeva čije su granice greške po $\frac{1}{2} \cdot 10^{-4}$ onda će zbir imati granicu greške $\frac{1}{2} \cdot 10^{-3}$.

2. $f(t_1, \ldots, t_n) = t_1 \cdot \ldots \cdot t_n$. Relativna greška proizvoda približno je jednaka zbiru relativnih grešaka činilaca. Drukčije rečeno, linearna ocjena za relativnu grešku proizvoda jednaka je zbiru relativnih greški činilaca. Za dokaz, dovoljno je izračunati

$$L = \frac{\partial f(x_1, \ldots, x_n)}{\partial x_i} \Delta(x_i), \quad \text{odnosno} \quad \frac{L}{f} = \frac{L}{x_1 \cdot \ldots \cdot x_n}.$$

Dokaz na drugi način (u slučaju $n = 3$):

$$f = f(x_1, x_2, x_3) = x_1 x_2 x_3 \qquad f^* = f(x_1^*, x_2^*, x_3^*) = x_1^* x_2^* x_3^* \qquad A = f - f^*,$$

$$A = x_1 x_2 x_3 \mp x_1^* x_2 x_3 \mp x_1^* x_2^* x_3 - x_1^* x_2^* x_3^* =$$

$$(x_1 - x_1^*) x_2 x_3 + (x_2 - x_2^*) x_1^* x_3 + (x_3 - x_3^*) x_1^* x_2^*,$$

$$\frac{A}{f} = \frac{x_1 - x_1^*}{x_1} + \frac{x_2 - x_2^*}{x_2} \cdot \frac{x_1^*}{x_1} + \frac{x_3 - x_3^*}{x_3} \cdot \frac{x_1^* x_2^*}{x_1 x_2},$$

$$\frac{A}{f} \approx \frac{x_1 - x_1^*}{x_1} + \frac{x_2 - x_2^*}{x_2} + \frac{x_3 - x_3^*}{x_3}, \quad \text{tj.} \quad R(f^*) \approx R(x_1^*) + R(x_2^*) + R(x_3^*).$$

Ako x_1^* ima 6 sigurnih cifara x_2^* 8 x_3^* 10 onda će proizvod $f^* = x_1^* x_2^* x_3^*$ imati 6 sigurnih cifara.

A razlike $=$ A umanjenika $+$ A umanjioca.

R količnika $\approx$ R djeljenika $+$ R djelioca.

Obrnuti problem greške. Razmotrimo funkciju $f = f(x_1, \ldots, x_n)$ i neka je unaprijed data najveća dozvoljena granica greške funkcije Δf, npr. $\Delta f = 0{,}5 \cdot 10^{-2}$. Treba odrediti granice grešaka argumenata $\Delta x_1, \ldots, \Delta x_n$ tako da greška funkcije bude manja ili jednaka od date gornje granice. Kako da se riješi? Postavljeni zadatak se lako rješava ako se upotrebi linearna ocjena za grešku funkcije L, iako nije savršena. Dakle, treba da bude ispunjeno: $(\forall i)$ $|x_i - x_i^*| \le \Delta x_i \Rightarrow |f(x_1, \ldots, x_n) - f(x_1^*, \ldots, x_n^*)| \le \Delta f$. Prema tome, dovoljno je da bude $L \le \Delta f$ ili svejedno

$$\sum_{i=1}^{n} \left| \frac{\partial f(x_1^*, \ldots, x_n^*)}{\partial x_i} \right| \Delta x_i \le \Delta f \quad \text{ili} \quad \sum_{i=1}^{n} \left| \frac{\partial f(x_1, \ldots, x_n)}{\partial x_i} \right| \Delta x_i \le \Delta f.$$

Drukčije zapisano:

$$\sum_{i=1}^{n} B_i \Delta x_i \le \Delta f, \quad \text{gdje je uvedena oznaka} \quad B_i = \sup \left| \frac{\partial f}{\partial x_i} \right| \quad \text{ili} \quad B_i \ge \sup \left| \frac{\partial f}{\partial x_i} \right|.$$

Kako izabrati $\Delta x_1 \ge 0, \ldots, \Delta x_n \ge 0$? Postoje razni pristupi, a popularna su sljedeća tri pristupa. Princip jednakih apsolutnih grešaka: $\Delta x_1 = \ldots = \Delta x_n$. Princip jednakih relativnih grešaka: $\Delta x_1/|x_1| = \ldots = \Delta x_n/|x_n|$. Princip jednakih doprinosa: $B_1 \Delta x_1 = \ldots = B_n \Delta x_n$, tako da je u razmatranom slučaju očito $(\forall i)$ $B_i \Delta x_i = \Delta f/n$ ili $B_i \Delta x_i \le \Delta f/n$. Odgovor treba dati u obliku $\Delta x_1 = \ldots, \ldots, \Delta x_n = \ldots$. Pretpostavlja se da unaprijed raspolažemo sa grubim aproksimacijama za argumente $x_1, \ldots, x_n$, odnosno sa njihovim intervalima. Npr. $7{,}4 < x_1 < 7{,}5$ i slično ostali x_i. U sva tri pristupa iskoristi se naravno nejednakost $B_1 \Delta x_1 + \ldots + B_n \Delta x_n \le \Delta f$ ili $B_1 \Delta x_1 + \ldots + B_n \Delta x_n = \Delta f$.

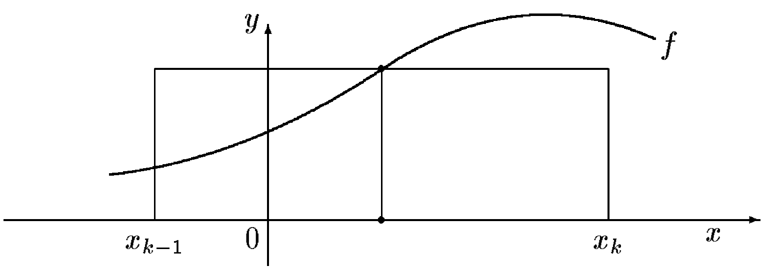
Praktično, uzima se $B_i = |\frac{\partial f(x_1, \ldots, x_n)}{\partial x_i}|$ ($x_1, \ldots, x_n$ iz grubih aproksimacija).

2. NUMERIČKA INTEGRACIJA

2.1. TRI FORMULE

Neka je $[a,b]$ interval na realnoj osi i neka je f funkcija $f\colon [a,b]\to R$. Razmotrimo integral $I=I(f)=\int_a^b f(x)dx$. Kao približna vrijednost služi zbir $S=S(f)=\sum_{k=1}^n c_k f(x_k)$. Za $c_k\in R$ kaže se da je koeficijent kvadraturne sume, za $x_k\in R$ kaže se da je čvor kvadraturne sume ($x_i\neq x_j$ za $i\neq j$); $n\geq 1$. Za zbir $\sum_{k=1}^n c_k f(x_k)$ ili svejedno $\sum_{k=1}^n c_k f_k$ kaže se da predstavlja kvadraturnu sumu, a za formulu oblika $I(f)\approx S(f)$ kaže se da predstavlja kvadraturnu formulu ili formulu za numeričku integraciju. Razmatra se i greška $R=R(f)=I(f)-S(f)$. Za kvadraturnu formulu se kaže da je zatvorenog tipa ako tačke a i b pripadaju mreži čvorova, a otvorenog tipa ako ne pripadaju.

Razmotrićemo tri formule koje imaju jasno geometrijsko tumačenje. Prelazimo na slučaj ekvidistantne mreže čiji je korak $h>0$ i malo prilagođavamo oznake. Neka je $x_0=a$, $x_n=b$, $nh=b-a$ i $x_k=x_0+kh$ za $k=0,1,\ldots,n$. Pisaćemo i $x_{k+1/2}=x_0+(k+\frac{1}{2})h$.



(a) Na malom intervalu $[x_{k-1},x_k]$, površinu ispod grafika funkcije $y=f(x)$ zamijenimo površinom pravougaonika čija je osnovica $[x_{k-1},x_k]$ i čija je visina $f(x_{k-1/2})=f_{k-1/2}$:

$$I_k=\int_{x_{k-1}}^{x_k} f(x)dx,\qquad S_k=\int_{x_{k-1}}^{x_k} f(x_{k-1/2})dx=hf(x_{k-1/2}),\qquad I_k\approx S_k,\tag{1}$$

$$R_k=I_k-S_k=\int_{x_{k-1}}^{x_k}(f(x)-f(x_{k-1/2}))dx,$$

$f\in C^2[a,b]$, Tejlorova formula:

$$f(x)=f(x_{k-1/2})+f'(x_{k-1/2})(x-x_{k-1/2})+\frac{1}{2!}f''(\xi_k(x))(x-x_{k-1/2})^2,\quad \xi_k(x)\in(x_{k-1},x_k),$$

$$R_k=f'(x_{k-1/2})\underbrace{\int_{x_{k-1}}^{x_k}(x-x_{k-1/2})dx}_{=0}+\frac{1}{2!}\int_{x_{k-1}}^{x_k} f''(\xi_k(x))\underbrace{(x-x_{k-1/2})^2}_{\geq 0}dx.$$

Teorema o srednjoj vrijednosti za integrale: f neprekidna, g integrabilna, g ne mijenja znak $\Rightarrow$ $(\exists\xi)$ $a<\xi<b$, $\int_a^b f(x)g(x)dx=f(\xi)\int_a^b g(x)dx$.

$$R_k=\frac{1}{2!}f''(\xi_k)\underbrace{\int_{x_{k-1}}^{x_k}(x-x_{k-1/2})^2dx}_{=h^3/12},\quad \xi_k\in(x_{k-1},x_k),\quad R_k=\frac{1}{24}h^3f''(\xi_k),\tag{2}$$

$$I=\int_a^b f(x)dx=\sum_{k=1}^n\int_{x_{k-1}}^{x_k}f(x)dx=\sum_{k=1}^n I_k,\quad S=\sum_{k=1}^n S_k=h\sum_{k=1}^n f(x_{k-1/2}),\quad I\approx S,\tag{3}$$

$$R = I - S = \sum_{k=1}^{n} R_k = \frac{1}{24} h^3 \sum_{k=1}^{n} f''(\xi_k) = \frac{1}{24} h^3 n \left(\frac{1}{n} \sum_{k=1}^{n} f''(\xi_k) \right).$$

$$\text{(Počinje posebno)} \quad m_1 = \min_{x\in[a,b]} f''(x), \quad m_2 = \max_{x\in[a,b]} f''(x), \quad m_1 \le f''(\xi_k) \le m_2,$$

$$nm_1 \le \sum_{k=1}^{n} f''(\xi_k) \le nm_2, \quad m_1 \le \frac{1}{n} \sum_{k=1}^{n} f''(\xi_k) \le m_2.$$

Funkcija neprekidna na zatvorenom intervalu dostiže sve svoje međuvrijednosti (teorema o međuvrijednostima za neprekidnu funkciju).

$$\exists \xi \in (a,b) \quad \frac{1}{n} \sum_{k=1}^{n} f''(\xi_k) = f''(\xi) \quad \text{(završeno posebno.)}$$

$$R = \frac{1}{24} h^3 n f''(\xi), \quad R = \frac{1}{24}(b-a) h^2 f''(\xi), \tag{4}$$

$$|R| \le \frac{1}{24}(b-a) h^2 M_2, \quad M_2 = \sup_{x\in[a,b]} |f''(x)|.$$

Za (1) se kaže da je osnovna formula pravougaonika, za (2) se kaže da izražava ocjenu greške. Za (3) se kaže da predstavlja (sastavljenu) formulu pravougaonika, a za (4) se kaže da izražava njenu grešku. Iz (4): $R = O(h^2)$ kad $h \to 0$, pa se kaže da je stepen ili red tačnosti ili preciznosti formule pravougaonika jednak $N = 2$. Ako se umjesto $h = 0{,}1$ stavi $h = 0{,}05$ onda se h^2 svede na četvrtinu, a $f''(\xi)$ se promijeni na nepoznat način; po (4) se kaže da se polovljenjem koraka greška svede (grubo govoreći) na četvrtinu prethodne greške. F $\square$: $S = h\sum_{i=1}^{n} f_{i-1/2}$.

(b) Kao približna vrijednost za I_k neka sada služi površina trapeza čija su tjemena $(x_{k-1}, 0)$, $(x_k, 0)$, (x_{k-1}, f_{k-1}) i (x_k, f_k). Stranici trapeza čija su tjemena (x_{k-1}, f_{k-1}) i (x_k, f_k) odgovara jednačina prave linije $y = L_1(x)$, gdje je $L_1(x)$ Lagranžov i. p. za f po mreži čvorova $\{x_{k-1}, x_k\}$. Tako da je očito $S_k = \int_{x_{k-1}}^{x_k} L_1(x)dx$. Poznata je formula za grešku interpolacije:

$$f(x) - L_1(x) = \frac{1}{2!}\omega_2(x) f''(\xi(x)), \quad \xi(x) \in (x_{k-1}, x_k), \quad x \in [x_{k-1}, x_k],$$

$$\omega_2(x) = (x - x_{k-1})(x - x_k), \quad f \in C^2[x_{k-1}, x_k] \quad \text{ili} \quad f \in C^2[a,b].$$

$$\text{Redom:} \quad I_k = \int_{x_{k-1}}^{x_k} f(x)dx, \quad S_k = \frac{h}{2}\Big(f(x_{k-1}) + f(x_k)\Big) \quad \text{(površina trapeza)},$$

$$S_k = \frac{h}{2}\Big(f_{k-1} + f_k\Big), \quad I_k \approx S_k, \tag{5}$$

$$R_k = I_k - S_k = \int_{x_{k-1}}^{x_k} f(x)dx - \int_{x_{k-1}}^{x_k} L_1(x)dx = \frac{1}{2!}\int_{x_{k-1}}^{x_k} f''(\xi(x)) \underbrace{(x - x_{k-1})(x - x_k)}_{\le 0} dx =$$

$$\frac{1}{2!} f''(\xi_k) \int_{x_{k-1}}^{x_k} (x - x_{k-1})(x - x_k)dx, \quad \xi_k \in (x_{k-1}, x_k), \quad R_k = -\frac{1}{12} h^3 f''(\xi_k), \tag{6}$$

$$I = \int_a^b f(x)dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(x)dx = \sum_{k=1}^{n} I_k,$$

$$S = \sum_{k=1}^{n} S_k = h\left(\frac{1}{2}f_0 + \sum_{k=1}^{n-1} f_k + \frac{1}{2}f_n\right), \quad I \approx S, \tag{7}$$

$$R = I - S = \sum_{k=1}^{n} R_k = -\frac{1}{12}h^3 \sum_{k=1}^{n} f''(\xi_k) = -\frac{1}{12}h^3 n f''(\xi), \quad a < \xi < b,$$

$$R = -\frac{1}{12}(b-a)h^2 f''(\xi). \tag{8}$$

Trapezna formula (7) ima grešku reda veličine h^2, odnosno ima red tačnosti $N = 2$, v. (8). Ako se korak prepolovi onda čvorovi iz prvobitne mreže ostaju i u novoj mreži (vrijednosti funkcije se za njih ne računaju ponovo). Rezime o trapeznoj: $I \approx S = h(\frac{1}{2}f_0 + \sum_{i=1}^{n-1} f_i + \frac{1}{2}f_n)$; $R(\frac{h}{2}) \approx \frac{1}{4}R(h)$ (grubo govoreći).

Iz formule (8): $|R| \le \frac{1}{12}(b-a)M_2h^2$, gdje je $M_2 = \max_{a\le x\le b}|f''(x)|$. Ponekad nemamo izraz za $f''(x)$, tako da ne raspolažemo sa M_2. Tada se, radi praktične ocjene greške, može uzeti da veličina $M_2^* = h^{-2}\max_{0\le i\le n-2}|\Delta^2 f_i|$ posluži kao približna zamjena za M_2.

(c) U slučaju Simpsonove formule, za obrazovanje približne vrijednosti S_k za mali interval $[x_{k-1}, x_k]$ iskoriste se tri podatka, iskoriste se vrijednosti funkcije u tačkama x_{k-1}, $x_{k-1/2}$ i x_k. Neka $L_2 = L_2(x)$ bude L. i. p. za funkciju f po mreži čvorova $\{x_{k-1}, x_{k-1/2}, x_k\}$. Površina ispod grafika $y = f(x)$ na intervalu $x_{k-1} \le x \le x_k$ približno je jednaka površini ispod parabole na tom istom intervalu. Sastavite izraz $L_2(x) = \ldots$ (parabola). Izračunajte integral od polinoma $\int_{x_{k-1}}^{x_k} L_2(x)dx$. Dobija se $\int_{x_{k-1}}^{x_k} L_2(x)dx = \frac{1}{6}(f_{k-1}+4f_{k-1/2}+f_k)$. U cilju ocjene greške, uvedimo u razmatranje i i. p. sa višestrukim čvorovima (Hermitov i. p.) $H_3 = H_3(x)$. Neka je $H_3 = H_3(x)$ H. i. p. za funkciju f definisan uslovima:

$$H_3(x_{k-1}) = f(x_{k-1}), \;\; H_3(x_{k-1/2}) = f(x_{k-1/2}), \;\; H_3(x_k) = f(x_k) \;\text{ i }\; H_3'(x_{k-1/2}) = f'(x_{k-1/2}).$$

Sastavite izraz $H_3(x) = \ldots$. Izračunajte $\int_{x_{k-1}}^{x_k} H_3(x)dx$. Dobiće se isti rezultat kao maločas. Dobiće se da je $\int_{x_{k-1}}^{x_k} H_3(x)dx = \frac{1}{6}(f_{k-1} + 4f_{k-1/2} + f_k)$.

Račun se za Simpsonovu formulu odvija po sličnom šablonu kao maločas u slučajevima (a) i (b): mali interval, greška, veliki interval, greška. Formula (9) će biti osnovna Simpsonova formula, a (11) će biti (sastavljena) Simpsonova formula. Pretpostavlja se da $f \in C^4[a,b]$:

$$I_k = \int_{x_{k-1}}^{x_k} f(x)dx, \quad S_k = \frac{h}{6}\Big(f_{k-1} + 4f_{k-1/2} + f_k\Big), \quad I_k \approx S_k, \tag{9}$$

$$R_k = I_k - S_k = \int_{x_{k-1}}^{x_k} f(x)dx - \int_{x_{k-1}}^{x_k} L_2(x)dx =$$

$$\int_{x_{k-1}}^{x_k} f(x)dx - \int_{x_{k-1}}^{x_k} H_3(x)dx = \int_{x_{k-1}}^{x_k} (f(x) - H_3(x))dx.$$

Poznati izraz za grešku za H. i. p. $\quad f(x) - H_3(x) = \dfrac{1}{4!}f^{IV}(\xi(x))\omega_4(x),$

$$x_{k-1} < \xi(x) < x_k, \quad x_{k-1} \le x \le x_k, \quad \omega_4(x) = (x-x_{k-1})(x-x_{k-1/2})^2(x-x_k), \quad f \in C^4[x_{k-1}, x_k].$$

$$R_k = \frac{1}{4!}\int_{x_{k-1}}^{x_k} f^{IV}(\xi(x))\underbrace{\omega_4(x)}_{\le 0}dx = \frac{1}{4!}f^{IV}(\xi_k)\underbrace{\int_{x_{k-1}}^{x_k}\omega_4(x)dx}_{\text{izračunajte}}, \quad \xi_k \in (x_{k-1}, x_k),$$

$$R_k = -\frac{1}{2880}h^5 f^{IV}(\xi_k), \tag{10}$$

$$I = \int_a^b f(x)dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(x)dx = \sum_{k=1}^{n} I_k,$$

$$S = \sum_{k=1}^{n} S_k = \frac{h}{6}\Big(f_0 + f_n + 4\sum_{k=1}^{n} f_{k-1/2} + 2\sum_{k=1}^{n-1} f_k\Big), \quad I \approx S, \tag{11}$$

$$R = I - S = \sum_{k=1}^{n} R_k = -\frac{1}{2880}h^5 \sum_{k=1}^{n} f^{IV}(\xi_k) = -\frac{1}{2880}h^4(b-a)\left(\frac{1}{n}\sum_{k=1}^{n} f^{IV}(\xi_k)\right),$$

$$R = -\frac{1}{2880}(b-a)h^4 f^{IV}(\xi) \qquad \text{za neko } \xi \in (a,b). \tag{12}$$

Vidimo da je $R = O(h^4)$ pa se kaže da je red tačnosti $N = 4$. Ako se korak prepolovi (radi dobijanja manje greške) onda se stare vrijednosti funkcije iskoriste. Polovljenje koraka svodi grešku na otprilike $\frac{1}{16}$ njene ranije vrijednosti.

Kvadraturne formule (b) i (c) se koriste više od (a) jer su pogodne kod polovljenja koraka.

Kvadratura (c) ima upadljivo veći (bolji) red tačnosti od (a) i (b).

Mala dopuna oko (c). Obično se izbjegava da u indeksu pišu polovine, obično se sa h označi rastojanje između dvije susjedne tačke. Drugim riječima, izvršićemo malu izmjenu u oznakama. Kada se izvrši onda se dobije: Simpsonova formula i njen izraz za grešku glase:

$$I = \int_a^b f(x)dx, \quad S = \frac{h}{3}\Big(f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \ldots + 4f_{n-1} + f_n\Big),$$

$$h = \frac{b-a}{n}, \quad x_i = a + ih,\ f_i = f(x_i),\ 0 \le i \le n, \quad \text{broj } n \text{ je obavezno paran}, \quad I \approx S,$$

$$R = I - S = -\frac{1}{180}(b-a)h^4 f^{IV}(\xi) \qquad \text{za neko } \xi \in (a,b).$$

Osnovna Simpsonova: $\int_{x_0}^{x_2} f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + f_2)$ sa greškom $-\frac{1}{90}h^5 f^{IV}(\xi)$. Primjer Simpsonove ($n = 10$): $\int_0^1 f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \ldots + 4f_9 + f_{10})$.

2.2. RUNGEOVO PRAVILO ZA PRAKTIČNU OCJENU GREŠKE

Prvo ćemo izvesti drugi izraz za grešku trapezne formule. Vidjećemo da je taj izraz bolji od izraza iz prethodnog naslova. Iskoristićemo taj izraz da izvedemo tzv. Rungeovo pravilo ili Rungeovu formulu za procjenu greške trapezne formule. Za procjenu greške na Rungeov način ili po Rungeovom principu karakteristično je da se izvrše dva proračuna, tj. da se dobiju dvije približne vrijednosti I_1 i I_2 za jedan te isti tačan broj I. I_1 je dobijeno sa korakom h, a I_2 je dobijeno sa korakom $\frac{h}{2}$. I_2 je dobijeno po mreži koja ima više čvorova, tako da I_2 ima manju grešku od I_1. Kako je približan broj I_2 bolji od I_1, to se I_2 usvaja kao numerički odgovor. Gruba približna vrijednost I_1 ima pomoćnu ulogu, ona služi da se procijeni greška numeričkog odgovora, tj. da se procijeni $R_2 = I - I_2$.

Rungeov princip za dobijanje procjene greške može da bude primijenjen na bilo koju numeričku metodu čiji izraz za grešku ima oblik $R(h) \sim Ch^s$.

Uvedimo potrebne oznake. Razmotrimo funkciju f definisanu na intervalu $[a,b]$ i pretpostavimo da $f \in C^4[a,b]$. Neka je $I = \int_a^b f(x)dx$. Izaberimo $n \geq 1$ i stavimo $h = \frac{b-a}{n}$. Neka je $I_1 = I(h) = h(\frac{1}{2}f(a) + \sum_{i=1}^{n-1} f(a+ih) + \frac{1}{2}f(b))$, trapezna formula sa korakom h i neka je $I_2 = I(\frac{h}{2}) = \frac{h}{2}(\frac{1}{2}f(a) + \sum_{i=1}^{2n-1} f(a+i\frac{h}{2}) + \frac{1}{2}f(b))$, trapezna formula sa korakom $\frac{h}{2}$. Zapaziti da ranije izračunati broj I_1 može da posluži prilikom računanja I_2 jer čvorovi grube mreže pripadaju i detaljnijoj mreži; $I_2 = \frac{1}{2}I_1 + \frac{h}{2}\sum_{i=1}^{n} f(a+(2i-1)\frac{h}{2})$. Uvedimo i oznake za dvije greške: $R_1 = R(h) = I - I_1 = I - I(h)$, $R_2 = R(\frac{h}{2}) = I - I_2 = I - I(\frac{h}{2})$. Odgovor glasi $I \approx I_2$, samo treba R_2 da se ocijeni ($|R_2| \leq \ldots$ ili $R_2 \approx \ldots$).

Razvijamo funkciju f po Maklorenovoj formuli do četvrtog izvoda:

$$f(x) = f(0) + xf'(0) + \frac{1}{2!}x^2 f''(0) + \frac{1}{3!}x^3 f'''(0) + \frac{1}{4!}x^4 f^{IV}(\xi(x)).$$

Za grešku trapezne formule po malom intervalu imamo:

$$r = \int_{-h/2}^{h/2} f(x)dx - \frac{h}{2}\left(f\left(-\frac{h}{2}\right) + f\left(\frac{h}{2}\right)\right) =$$

$$\int_{-h/2}^{h/2} f(0)dx + \frac{1}{2!}f''(0)\int_{-h/2}^{h/2} x^2dx + \frac{1}{4!}\underbrace{\int_{-h/2}^{h/2} f^{IV}(\xi(x))x^4dx}_{\text{t. o srednjoj v.}} -$$

$$\frac{h}{2}\left(2f(0) + 2\frac{1}{2!}\left(\frac{h}{2}\right)^2 f''(0) + \frac{1}{4!}\left(\frac{h}{2}\right)^4 f^{IV}(\xi_1) + \frac{1}{4!}\left(\frac{h}{2}\right)^4 f^{IV}(\xi_2)\right),$$

$$\text{zbog neparnosti } \int_{-h/2}^{h/2} xdx = 0, \quad \int_{-h/2}^{h/2} x^3dx = 0, \quad \text{gdje } \xi_1 = \xi\left(-\frac{h}{2}\right), \quad \xi_2 = \xi\left(\frac{h}{2}\right),$$

$$r = hf(0) + f''(0)\frac{1}{3}\left(\frac{h}{2}\right)^3 + \frac{1}{4!}f^{IV}(\bar{\xi})\int_{-h/2}^{h/2} x^4dx -$$

$$hf(0) - \left(\frac{h}{2}\right)^3 f''(0) - \frac{1}{4!}\left(\frac{h}{2}\right)^5 f^{IV}(\xi_1) - \frac{1}{4!}\left(\frac{h}{2}\right)^5 f^{IV}(\xi_2) =$$

$$-\frac{1}{12}f''(0)h^3 + \rho, \quad |\rho| \leq cM_4h^5, \quad M_4 = \max_{x\in[a,b]}|f^{IV}(x)|.$$

Maločas smo izvršili translaciju po x–osi, samo radi lakšeg pisanja. Neka sada ulogu intervala $[-\frac{h}{2}, \frac{h}{2}]$ preuzme interval $[x_{i-1}, x_i] = [a+(i-1)h, a+ih]$:

$$r_i = \int_{x_{i-1}}^{x_i} f(x)dx - \frac{h}{2}\Big(f(x_{i-1}) + f(x_i)\Big) = -\frac{1}{12}h^3 f''(x_{i-1/2}) + \rho_i,$$

$$|\rho_i| \leq cM_4h^5, \quad \rho_i = O(h^5), \quad \sum_{i=1}^{n}\rho_i = O(h^4).$$

Sabiranjem po $i = 1, \ldots, n$:

$$R(h) = I - I(h) = \sum_{i=1}^{n} r_i = -\frac{1}{12}h^2 \underbrace{\sum_{i=1}^{n} hf''(x_{i-1/2})}_{\text{f. } \square \text{ za } f''} + \sum_{i=1}^{n} \rho_i =$$

$$-\frac{1}{12}h^2 \left(\int_a^b f''(x)dx - \frac{1}{24}(b-a)h^2 f^{IV}(\xi) \right) + O(h^4) = -\frac{1}{12}h^2 \int_a^b f''(x)dx + O(h^4) + O(h^4)$$

(v. formulu pravougaonika i izraz za grešku),

$$R(h) = Ch^2 + O(h^4), \quad C = -\frac{1}{12}\int_a^b f''(x)dx, \quad C \text{ ne zavisi od } h.$$

Mi smo dobili:

$$R(h) = R_1 = I - I(h) = Ch^2 + O(h^4), \quad h \to 0, \quad C = -\frac{1}{12}\big(f'(b) - f'(a)\big). \tag{1}$$

Isto tako:

$$R\left(\frac{h}{2}\right) = R_2 = I - I\left(\frac{h}{2}\right) = C\left(\frac{h}{2}\right)^2 + O\left(\left(\frac{h}{2}\right)^4\right) = \frac{1}{4}Ch^2 + O(h^4).$$

Zapaziti da su brojevi I_1 i I_2 efektivno poznati, što se ne može reći za I i C. Mi imamo:

$$R_1 = I - I_1 = Ch^2 + O(h^4), \quad R_2 = I - I_2 = \frac{1}{4}Ch^2 + O(h^4).$$

$$\text{Oduzimanjem:} \quad I_2 - I_1 = \frac{3}{4}Ch^2 + O(h^4).$$

Ako je $C \neq 0$ onda je

$$\lim_{h\to 0}\frac{R_2}{I_2 - I_1} = \lim_{h\to 0}\frac{\frac{1}{4}Ch^2 + O(h^4)}{\frac{3}{4}Ch^2 + O(h^4)} = \frac{1}{3} \text{ ili } R_2 \sim \frac{1}{3}\big(I_2 - I_1\big) \text{ kad } h \to 0 \text{ ili } R_2 \approx \frac{1}{3}\big(I_2 - I_1\big). \tag{2}$$

Ovo je završna–glavna formula. Broj na desnoj strani formule (2) je efektivno poznat. Za Rungeovu ocjenu (2) kaže se da je praktična zato što je ona efektivno ostvarljiva (njena dobra strana) i zato što je ona samo približna (njena loša strana). Mogli bismo da zapišemo Rungeovu formulu u drugim oznakama kao: $R_{2n} \approx \frac{1}{3}(I_{2n} - I_n)$, ako se radi o trapeznoj formuli.

U slučaju $C = 0$ može se pokazati da važi $|R_2| \leq \frac{1}{3}|I_2 - I_1|$ za male vrijednosti h.

Pogledajmo mali primjer. Razmotrimo $\int_0^1 x^4 dx$ i neka bude $n = 1$, tj. $h = 1$. Tada je $I_1 = 0{,}50000$, $I_2 = 0{,}28125$ i $\frac{1}{3}(I_2 - I_1) = -0{,}07292$. A $R_2 = -0{,}08125$.

Slijede razne dopune

1. Ako $I(\frac{h}{2})$ ne zadovoljava unaprijed traženu preciznost ($R(\frac{h}{2})$ prelazi unaprijed datu dozvoljenu grešku ε, tj. $\frac{1}{3}|I_2 - I_1| > \varepsilon$) onda izračunajte sa korakom $\frac{h}{4}$ novu približnu vrijednost $I(\frac{h}{4})$ i procijenite njenu grešku na Rungeov način. Itd.

2. Neka $I(h)$ znači približnu vrijednost dobijenu po Simpsonovoj formuli. Tada važi $R(h) = I - I(h) = Ch^4 + O(h^6)$, $h \to 0$, C ne zavisi od h. Isto tako važi $R(\frac{h}{2}) \approx \frac{1}{15}(I(\frac{h}{2}) - I(h))$ (procjena greške na Rungeov način). Pretpostavlja se da $f \in C^6[a, b]$.

3. Uopšte, ako je $R(h) \sim Ch^s$ onda važi $R(\frac{h}{2}) \approx \frac{1}{2^s-1}(I(\frac{h}{2}) - I(h))$.

2.3. ROMBERGOVA FORMULA

Mali primjer iz prethodnog naslova: $I(\frac{h}{2}) + \frac{1}{3}(I(\frac{h}{2}) - I(h)) = 0{,}20833$.

U prethodnom naslovu smo vidjeli da broj $b = \frac{1}{3}(I(\frac{h}{2}) - I(h))$ može da posluži kao približna vrijednost za $R(\frac{h}{2}) = I - I(\frac{h}{2})$. Ako je već $R(\frac{h}{2}) \approx b$ onda je očito $I \approx I(\frac{h}{2}) + b$. Dakle, neka broj b služi za popravku približnog broja $I(\frac{h}{2})$, neka sada numerički odgovor bude $I(\frac{h}{2}) + b$. Zanimljivo je da se broj $I(\frac{h}{2}) + b$ poklapa sa približnom vrijednošću za I izračunatom po Simpsonovoj formuli sa korakom $\frac{h}{2}$, tj. po mreži čvorova $a, a + \frac{h}{2}, a + h, \ldots, b$; uvjerite se neposrednim računom. Nije bitno što odgovara baš Simpsonovoj formuli već je bitno što odgovara formuli čiji je stepen tačnosti (četvrti) veći od stepena tačnosti polazne trapezne formule (drugi).

Jedno sredstvo za dobijanje preciznije približne vrijednosti određenog integrala $I = \int_a^b f(x)\,dx$ jeste da se umjesto koraka h primijeni korak $\frac{h}{2}$, zatim $\frac{h}{4}$, itd. Drugo sredstvo bilo bi da se umjesto kvadraturne formule drugog stepena tačnosti primijeni formula četvrtog reda tačnosti, zatim šestog reda, itd. Kod Rombergove formule usaglašemo se koriste jedno i drugo sredstvo, s tim da se prelazak na formulu višeg reda tačnosti vrši pomoću popravke ranije izračunate približne vrijednosti.

U prvom dijelu izlaganja biće izveden izraz za grešku trapezne formule $R(h)$. Upravo, biće pokazano da se greška $R(h)$ razlaže po parnim stepenima h. U drugom dijelu izlaganja biće izvedena formula za popravku.

Prvi dio izlaganja. Neka je $nh = b - a$ i $x_i = a + ih$. Imamo:

$$R(h) = I - I(h) = \int_a^b f(x)dx - h\left(\frac{1}{2}f(x_0) + \sum_{i=1}^{n-1} f(x_i) + \frac{1}{2}f(x_n)\right) =$$

$$\sum_{i=1}^{n}\left[\int_{x_{i-1}}^{x_i} f(x)dx - h\left(\frac{1}{2}f(x_{i-1}) + \frac{1}{2}f(x_i)\right)\right],$$

razviti funkciju f po Tejlorovoj formuli oko tačke $x = x_{i-1/2}$ do šestog izvoda,

$$R(h) = \sum_{i=1}^{n}\left[\frac{1}{24}h^3 f''(x_{i-1/2}) + \frac{1}{1920}h^5 f^{IV}(x_{i-1/2}) + \frac{1}{6!}f^{VI}(\overline{\xi_i})\int_{-h/2}^{h/2} x^6 dx - \right.$$

$$h\left(\frac{1}{2!}\left(\frac{h}{2}\right)^2 f''(x_{i-1/2}) + \frac{1}{4!}\left(\frac{h}{2}\right)^4 f^{IV}(x_{i-1/2}) + \right.$$

$$\left.\left.\frac{1}{2}\cdot\frac{1}{6!}\left(\frac{h}{2}\right)^6 f^{VI}(\xi_{1i}) + \frac{1}{2}\cdot\frac{1}{6!}\left(\frac{h}{2}\right)^6 f^{VI}(\xi_{2i})\right)\right] =$$

$$-\frac{1}{12}h^2 \underbrace{\sum_{i=1}^{n} hf''(x_{i-1/2})}_{\text{f. } \square \text{ za } f''} - \frac{1}{480}h^4 \underbrace{\sum_{i=1}^{n} hf^{IV}(x_{i-1/2})}_{\text{f. } \square \text{ za } f^{IV}} + O(h^6),$$

po sličnosti sa prethodnim naslovom pokažjte da se greška sastavljene formule pravougaonika prikazuje u obliku $c_2h^2 + O(h^4)$, gdje c_2 ne zavisi od h, ali naravno zavisi od podintegralne funkcije,

$$R(h) = -\frac{1}{12}h^2\left(\int_a^b f''(x)dx - c_2(f'')h^2 + O(h^4)\right) -$$

$$\frac{1}{480}h^4\left(\int_a^b f^{IV}(x)dx - c_2(f^{IV})h^2 + O(h^4)\right) + O(h^6) = C_2h^2 + C_4h^4 + O(h^6).$$

Ako se nastavi sa primjenom i dogradnjom postupka od maločas onda će se za grešku sastavljene trapezne formule za funkciju f po intervalu $[a, b]$ sa korakom h dobiti sljedeća formula:

$$R(h) = I - I(h) = C_2h^2 + C_4h^4 + \ldots + C_{2m}h^{2m} + O(h^{2m+2}), \quad h \to 0, \tag{1}$$

gdje veličine $C_2, C_4, \ldots, C_{2m}$ ne zavise od h, a pretpostavlja se da $f \in C^{2m+2}[a, b]$.

Drugi dio izlaganja. Popravljanjem, sa koeficijentom popravke $\frac{1}{3}$, približnih vrijednosti čija greška ima oblik $C_2h^2+C_4h^4+C_6h^6+\ldots$ dobijaju se nove–bolje približne vrijednosti čija greška počinje sa h^4 (čija greška je jednaka $C_4'h^4+C_6'h^6+\ldots$). U nastavku je dat račun koji potkrepljuje ovo tvrđenje. Popravljanjem maločas dobijenih popravljenih vrijednosti, ovog puta uzimajući $\frac{1}{15}$ kao koeficijent popravke, dobijaju se još bolje približne vrijednosti, njihova greška počinje sa h^6. U nastavku je dat račun koji potkrepljuje ovo tvrđenje. Sljedeći koeficijent popravke biće $\frac{1}{63}$. Itd.

Zaista, iz pretpostavke da je $I - I(h) = C_2h^2 + C_4h^4 + C_6h^6 + \ldots$, ako se uvede oznaka $J\left(\frac{h}{2}\right) = I\left(\frac{h}{2}\right) + \frac{1}{3}\Big(I\left(\frac{h}{2}\right) - I(h)\Big)$, slijedi da je

$$I - J\left(\frac{h}{2}\right) = -\frac{1}{3}\Big(I - I(h)\Big) + \frac{4}{3}\left(I - I\left(\frac{h}{2}\right)\right) =$$

$$-\frac{1}{3}\Big(C_2h^2 + C_4h^4 + C_6h^6 + \ldots\Big) + \frac{4}{3}\left(C_2\left(\frac{h}{2}\right)^2 + C_4\left(\frac{h}{2}\right)^4 + C_6\left(\frac{h}{2}\right)^6 + \ldots\right) =$$

$$-\frac{1}{4}C_4h^4 - \frac{5}{16}C_6h^6 + \ldots = C_4'h^4 + C_6'h^6 + \ldots.$$

Zaista, iz pretpostavke da je $I - J(h) = ch^4 + O(h^6)$, ako se uvede oznaka $K\left(\frac{h}{2}\right) = J\left(\frac{h}{2}\right) + \frac{1}{15}\Big(J\left(\frac{h}{2}\right) - J(h)\Big)$, slijedi da je $I - K(\frac{h}{2}) = \ldots = O(h^6)$.

Da se dobije jedna nova popravljena približna vrijednost dovoljne su svega tri–četiri aritmetičke operacije. Time je izvođenje Rombergove formule urađeno. Samo se treba uvjeriti da su formule koje slijede saglasne sa dosad rečenim. Rombergova formula glasi:

$$\begin{cases} S_{0j} = h_j\left(\dfrac{1}{2}f(a) + \displaystyle\sum_{k=1}^{M-1} f(a + kh_j) + \dfrac{1}{2}f(b)\right), \quad j \geq 0, \\ S_{ij} = S_{i-1,j} + \dfrac{1}{4^i - 1}\Big(S_{i-1,j} - S_{i-1,j-1}\Big), \quad i > 0, \quad j \geq i, \end{cases} \tag{2}$$

gdje je $h_j = 2^{-j}h_0$, $M = 2^jn$ i $nh_0 = b - a$.

Formula (2) služi za dobijanje približne vrijednosti za $I = \int_a^b f(x)dx$. Njen parametar je početni korak h_0. Dokaz formule (2) sadržan je u dosadašnjim razmatranjima. Zapaziti da za realizaciju formule (2) nije potrebno da znamo čemu su jednake vrijednosti $C_2, C_4, \ldots, C_{2m}$ iz (1).

Vrijednosti S_{0j} iz prvog reda tabele 1 odgovaraju trapeznoj formul: S_{00} sa korakom h_0, S_{01} sa korakom $h_1 = 2^{-1}h_0$, S_{02} sa korakom $h_2 = 2^{-2}h_0$, itd. Vrijednosti S_{1j} iz drugog reda odgovaraju Simpsonovoj formuli. Vrijednosti S_{2j} iz trećeg reda odgovaraju jednoj formuli čiji je stepen tačnosti šest. Itd. Prilikom računanja vrijednosti S_{0j} treba iskoristiti već ranije izračunati broj $S_{0,j-1}$ koji sadrži polovinu sada potrebnih vrijednosti funkcije u čvorovima, zato smo nacrtali horizontalne strelice. Za računanje približne vrijednosti S_{ij} (kada je $i \geq 1$) dovoljne su svega tri–četiri aritmetičke operacije, vertikalne i kose strelice pokazuju na osnovu kojih ranije izračunatih vrijednosti se izračuna S_{ij}. Tabela 2 govori da precizan redosljed računanja jeste upravo 1, 2, 3, ….

Poslije 3, 6, 10, … računanja se prekidaju za trenutak da bi se provjerilo da li je dostignuta željena tačnost ε (unaprijed određena). Ako je ispunjen izlazni kriterijum $|S_{j-1,j} - S_{jj}| < \varepsilon$ onda se S_{jj} usvaja kao numerički odgovor. Program za računar obično se zaustavlja kada u nizu $S_{00}, S_{11}, S_{22}, \ldots$ dođe do poklapanja neka dva njegova susjedna elementa. Kaže se da se računa sa najvećom mogućom tačnošću ili do mašinske tačnosti. Dalje popravke ne bi koristile, tj. računar prosto ne bi mogao da ih registruje. Dvostruka preciznost (10^{-16}).

$$
\begin{tikzcd}
S_{00} \arrow[r] \arrow[dr] & S_{01} \arrow[r] \arrow[d] \arrow[dr] & S_{02} \arrow[r] \arrow[d] \arrow[dr] & S_{03} \arrow[d] \\
& S_{11} \arrow[dr] & S_{12} \arrow[d] \arrow[dr] & S_{13} \arrow[d] \\
& & S_{22} \arrow[dr] & S_{23} \arrow[d] \\
& & & S_{33}
\end{tikzcd}
$$

Tabela 1 $S_{11} = S_{01} + \frac{1}{3}\Big(S_{01} - S_{00}\Big)$

1	2	4	7
	3	5	8
		6	9
			10

Tabela 2 Redosljed

$I(h)$	$I(\frac{h}{2})$	$I(\frac{h}{4})$	$I(\frac{h}{8})$
	$J(\frac{h}{2})$	$J(\frac{h}{4})$	$J(\frac{h}{8})$
		$K(\frac{h}{4})$	$K(\frac{h}{8})$
			$L(\frac{h}{8})$

Tabela 3

2.4. KVADRATURNE FORMULE U SLUČAJU PRISUSTVA TEŽINSKE FUNKCIJE

Neka je $p = p(x)$ fiksirana funkcija, p: $[a,b] \to R$ za koju se pretpostavlja da je integrabilna na intervalu $[a,b]$. Za funkciju p se kaže da je težinska funkcija ako je ispunjen uslov $p(x) \geq 0$ za svako $x \in [a,b]$. Razmotrimo određeni integral $I = I(f) = \int_a^b f(x)p(x)dx$. Neka je po intervalu $[a,b]$ postavljena mreža čvorova $\{x_i\}_{i=0}^n$ i neka su poznate vrijednosti funkcije u čvorovima $f(x_i) = f_i$. Konstruisaćemo kvadraturnu formulu (interpolacionog tipa) oblika $I(f) \approx S(f)$, gdje je $S(f) = \sum_{i=0}^n c_i f(x_i)$. Funkcija f biće zamijenjena svojim Lagranžovim i. p. po razmatranoj mreži čvorova. Odgovarajući izraz za grešku $R = R(f) = I - S = I(f) - S(f)$ takođe se lako dobija, polazeći od poznatog izraza za grešku interpolacije.

Iz naslova 1.1. i 1.2. poznato je sljedeće:

$$L_n(x) = \sum_{i=0}^n \ell_i(x) f_i, \quad \ell_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^n \frac{x - x_j}{x_i - x_j}, \quad f(x) = L_n(x) + \frac{1}{(n+1)!}\omega_{n+1}(x) f^{(n+1)}(\xi(x)),$$

$$a < \xi(x) < b \ \ (\text{jer } a \leq x \leq b), \quad f \in C^{n+1}[a,b], \quad \omega_{n+1}(x) = \prod_{i=0}^n (x - x_i).$$

Izvođenje kvadraturne formule:

$$f(x) = \sum_{i=0}^n \ell_i(x) f_i + \frac{1}{(n+1)!}\omega_{n+1}(x) f^{(n+1)}(\xi(x)) \quad \Big/ \cdot p(x)$$

$$f(x)p(x) = \sum_{i=0}^n \ell_i(x)p(x) f_i + \frac{1}{(n+1)!}\omega_{n+1}(x) f^{(n+1)}(\xi(x))p(x) \quad \Big/ \int_a^b \dots\, dx$$

$$\underbrace{\int_a^b f(x)p(x)dx}_{I(f)} = \sum_{i=0}^n \underbrace{\Big(\int_a^b \ell_i(x)p(x)dx\Big)}_{c_i} f_i + \underbrace{\frac{1}{(n+1)!}\int_a^b \omega_{n+1}(x) f^{(n+1)}(\xi(x))p(x)dx}_{R(f)},$$

$$I(f) = \sum_{i=0}^n c_i f_i + R(f),$$

$$I(f) \approx S(f) = \sum_{i=0}^n c_i f_i. \tag{1}$$

Za grešku formule (1) imamo:

$$|R(f)| \leq \frac{1}{(n+1)!} M_{n+1} \int_a^b |\omega_{n+1}(x)| p(x)dx, \qquad M_{n+1} = \max_{x\in[a,b]} |f^{(n+1)}(x)|. \tag{2}$$

Neka je mreža ravnomjerna i neka tačke $x = a$ i $x = b$ pripadaju mreži. Drugim riječima, neka čvorovi budu $x_i = a + ih$ za $0 \leq i \leq n$, gdje je $nh = b - a$. Tada se za kvadraturnu formulu oblika (1) kaže da je Njutn–Kotesova kvadraturna formula zatvorenog tipa. Primjera radi, navedimo Njutn–Kotesove formule kada je $n = 3$ i kada je $n = 4$, u slučaju da je težinska funkcija $p(x) \equiv 1$:

$$\int_a^b f(x)dx \approx \frac{b-a}{8}\Big(f_0 + 3f_1 + 3f_2 + f_3\Big) \quad \text{(tri–osminska formula)},$$

$$\int_a^b f(x)dx \approx \frac{b-a}{90}\Big(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4\Big).$$

Zadatak. Neka za kvadraturnu formulu (1), pored njene greške metode $R = R_1$ koja je izražena relacijom (2), *treba proučiti i njenu grešku izazvanu približnošću ulaznih podataka* R_2. Ulaznim podacima smatraju se brojevi $\{f_i\}_{i=0}^n$. Uzimamo da ulazni podaci nisu poznati sasvim tačno, tj. uzimamo da raspolažemo samo sa približnim vrijednostima $\{f_i^*\}_{i=0}^n$. Neka je data mjera greške ulaznih podataka δ: $|f_i - f_i^*| \le \delta$ za $0 \le i \le n$. Ulogu numeričkog odgovora $S = \sum_{i=0}^n c_i f_i$ sada naravno preuzima broj $S^* = \sum_{i=0}^n c_i f_i^*$, tako da je $R_2 = S - S^*$. A naravno da je u datim okolnostima ukupna greška $I - S^*$ jednaka $I - S^* = R_1 + R_2$.

Treba riješiti postavljeni zadatak, tj. treba ocijeniti R_2:

$$R_2 = S - S^* = \sum_{i=0}^n c_i f_i - \sum_{i=0}^n c_i f_i^* = \sum_{i=0}^n c_i(f_i - f_i^*),$$

$$|R_2| = \Big|\sum_{i=0}^n c_i(f_i - f_i^*)\Big| \le \sum_{i=0}^n |c_i(f_i - f_i^*)| = \sum_{i=0}^n |c_i|\cdot|f_i - f_i^*| \le \sum_{i=0}^n |c_i|\delta,$$

$$|R_2| \le C\delta, \quad \text{gdje je } C = \sum_{i=0}^n |c_i|.$$

Na redu je analiza posljednje relacije. Što je veličina C manja to je R_2 manje, odnosno to se greška ulaznih podataka slabije odražava–prenosi na rezultat (na numerički odgovor). Drugim riječima, što je veličina C manja to je formula (1) stabilnija u odnosu na grešku svojih ulaznih podataka, odnosno to je situacija povoljnija. Ili iz suprotnog ugla: ako je broj C velik onda je formula (1) numerički nedovoljno stabilna u odnosu na grešku ulaznih podataka. Zadatak je riješen.

Posmatrajmo relaciju (1) kada je $f(x) \equiv 1$. Tada je $L_n(x) \equiv f(x)$ i zato $R(f) = 0$ odnosno $I(f) = S(f)$. $I(f) = \int_a^b p(x)dx$, $S(f) = \sum_{i=0}^n c_i$, tako da je $\sum_{i=0}^n c_i = \int_a^b p(x)dx = \text{const}$.

Ako su svi koeficijenti $\{c_i\}_{i=0}^n$ pozitivni onda je očito $\sum_{i=0}^n c_i = \sum_{i=0}^n |c_i|$. Međutim, ako među koeficijentima $\{c_i\}_{i=0}^n$ ima i pozitivnih i negativnih onda je očito $\sum_{i=0}^n |c_i| > \sum_{i=0}^n c_i$. Za Njutn–Kotesove formule sa velikim n (recimo sa $n \ge 10$) karakteristično je da su neki koeficijenti pozitivni a neki negativni, što umanjuje njihovu numeričku stabilnost u odnosu na grešku ulaznih podataka, pa se zato te formule u praksi rijetko koriste.

2.5. GAUSOVA KVADRATURNA FORMULA

Priprema o ortogonalnim polinomima

U realnom Hilbertovom prostoru H posmatrajmo n linearno nezavisnih elemenata $\varphi_1, \ldots, \varphi_n$. Primjenom Gram–Šmitovog postupka ortogonalizacije, sistem $\{\varphi_1, \ldots, \varphi_n\}$ može da se ortogonalizuje, tj. može da bude dobijen sistem $\{\psi_1, \ldots, \psi_n\}$ za koji važi: (a) ψ_k je linearna kombinacija od $\varphi_1, \ldots, \varphi_k$ i (b) $\psi_i \perp \psi_j$, tj. $\langle \psi_i, \psi_j \rangle = 0$ za $i \ne j$; $\langle\ ,\ \rangle$ je oznaka za skalarni proizvod. Elementi $\psi_1, \ldots, \psi_n$ određeni su jednoznačno do na multiplikativne konstante.

Razmotrimo konkretni Hilbertov prostor $H = L^2_{p(x)}(a,b)$. Neka je funkcija $p = p(x)$ integrabilna po $[a,b]$ i neka je $p(x) > 0$ skoro svuda na $[a,b]$. Funkcija f pripada H ako i samo ako je $\int_a^b |f(x)|^2 p(x)dx$ konačan. Ako se dvije funkcije razlikuju samo na skupu mjere nula onda se one identifikuju. U H se skalarni proizvod definiše kao $\langle f, g\rangle = \int_a^b f(x)g(x)p(x)dx$. Ovo je Lebegov prostor. Na primjer $L^2(a,b)$, kada je $p(x) \equiv 1$.

U našem Hilbertovom prostoru H razmotrimo njegovih n elemenata $\varphi_i(x) = x^{i-1}$ za $i = 1,\dots,n$. Sistem $\{1, x, x^2, \dots, x^{n-1}\}$ je linearno nezavisan. Njegovom ortogonalizacijom nastaje sistem $\{\psi_1,\dots,\psi_n\}$. Kako je φ_i polinom stepena tačno $i-1$ to je i ψ_i polinom stepena tačno $i-1$. Po n može da se napreduje, tako da se može posmatrati sistem $\{\psi_1, \psi_2, \dots\}$. Izvršimo malo prilagođavanje oznaka: umjesto $\psi_1, \psi_2, \dots$ pisaćemo $\psi_0, \psi_1, \dots$. Tako da je odsad ψ_i polinom stepena tačno i. Za sistem $\{\psi_0, \psi_1, \dots\}$ kaže se da je sistem ili da je niz ortogonalnih polinoma. Navedimo dva primjera.

Primjer 1. Neka je $[a,b] = [-1,1]$ i $p(x) \equiv 1$. Sistem ortogonalnih polinoma u prostoru $L^2(-1,1)$ je sistem tzv. Ležandrovih polinoma $\{P_n\}_{n=0}^{\infty}$ za koje važi relacija: $P_n(x) = \dfrac{1}{2^n n!}\dfrac{d^n}{dx^n}(x^2-1)^n$. Recimo, $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \dfrac{3}{2}x^2 - \dfrac{1}{2}$, $P_3(x) = \dfrac{5}{2}x^3 - \dfrac{3}{2}x$.

Prinjer 2. Neka je $[a,b] = [-1,1]$ i $p(x) = \dfrac{1}{\sqrt{1-x^2}}$. U ovom prostoru $L^2_{p(x)}(-1,1)$, ortogonalni polinomi su tzv. Čebiševljevi polinomi $T_n(x)$ $(n = 0,1,\dots)$ za koje važi relacija: $T_n(x) = \cos(n \arccos x)$. Recimo, $T_0(x) = 1$, $T_1(x) = x$, $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$.

Za dokaz prvog primjera, treba se uvjeriti da se ortogonalizacijom sistema $\{1, x, x^2, \dots\}$ dobija sistem $\{P_0, P_1, P_2, \dots\}$. Ili, pošto je već napisana gotova relacija $P_n(x) = \dots$, dovoljno je uvjeriti se: (a) da je P_n polinom stepena tačno n i (b) da je $P_i \perp P_j$, tj. da je $\langle P_i, P_j\rangle = \int_{-1}^{1} P_i(x)P_j(x)dx = 0$ za $i \neq j$. Slično važi naravno i za drugi primjer. Dokaze za jedan i drugi primjer izostavljamo.

U sljedećoj lemi govori se o jednom svojstvu polinoma iz sistema ortogonalnih polinoma. Prije toga, o faktorizaciji polinoma u realnom području. Polinom $P = P(x)$ na jednoznačan način može da se prikaže u obliku

$$P(x) = c_0 \prod_i (x-\alpha_i)^{m_i} \prod_j (x^2 + \beta_j x + \gamma_j)^{n_j},$$

gdje jednačina $x^2 + \beta_j x + \gamma_j = 0$ nema realnog rješenja; za izraz $x^2 + \beta_j x + \gamma_j$ kaže se da je nesvodljiv trinom. Broj $x = \alpha_i$ je nula polinoma všestrukosti m_i. Polinom P je očito stepena $\sum_i m_i + 2\sum_j n_j$. Recimo

$$P(x) = 2(x-1)^4(x-2)(x-6)^2(x^2+1)^2(x^2+4x+5).$$

Dalje, o broju nula polinoma poznato je sljedeće pravilo: polinom stepena n ima najviše n realnih nula (osim $P \equiv 0$).

Lema. Neka je $\{\psi_0, \psi_1, \dots\}$ sistem ortogonalnih polinoma u prostoru $L^2_{p(x)}(a,b)$, gdje je $\psi_n = \psi_n(x)$ polinom stepena tačno n. Polinom ψ_n ima u otvorenom intervalu (a,b) tačno n međusobno različitih nula.

Dokaz leme. Uočimo sve nule polinoma ψ_n koje pripadaju intervalu (a,b) i čija je višestrukost neparna; neka su to $x_1, \dots, x_\ell$. Dopustimo da je $\ell < n$. Neka je

$$f(x) = \prod_{i=1}^{\ell}(x - x_i) \qquad \text{i} \qquad g(x) = \psi_n(x) f(x).$$

Razmotrimo veličinu

$$A = \int_a^b g(x)p(x)dx = \int_a^b \psi_n(x)\prod_{i=1}^{\ell}(x-x_i)p(x)dx = \int_a^b \psi_n(x)f(x)p(x)dx = \langle \psi_n, f\rangle.$$

Pođimo od faktorizacije za ψ_n i pogledajmo faktorizaciju za g.

Nesvodljivi trinomi u faktorizaciji za ψ_n ne mijenjaju znak uopšte na čitavoj realnoj osi. Faktori $(x-\alpha_i)^{m_i}$, gdje je m_i paran, takođe ne mijenjaju znak. Faktori $(x-\alpha_i)^{m_i}$, gdje je m_i neparan a $\alpha_i \notin (a,b)$, ne mijenjaju znak u (a,b).

Ako je eksponent m_i neparan i $\alpha_i \in (a,b)$ onda se u faktorizaciji za g pojavljuje paran eksponent m_i+1 jer je izvršeno množenje sa $x-\alpha_i$, tj. sa $f(x)$. Tako da je $g(x) \geq 0$ za $x \in [a,b]$, ako je $c_0 > 0$, odnosno $g(x) \leq 0$ za $x \in [a,b]$, ako je $c_0 < 0$. Funkcija g je polinom, g ima samo konačno mnogo nula, tako da je $A = \int_a^b g(x)p(x)dx > 0$ ili < 0 (zavisno od c_0). Dakle, $A \neq 0$.

S druge strane, $A = \langle \psi_n, f \rangle$. Pretpostavili smo da je $\ell < n$. Zato je f polinom stepena nižeg od n. Polinom f može se prikazati kao linearna kombinacija $\psi_0, \ldots, \psi_\ell$; $f = \sum_{i=0}^{\ell} c_i \psi_i$. Važi $\psi_i \perp \psi_n$, tj. $\langle \psi_i, \psi_n \rangle = 0$ za $i = 0, \ldots, \ell \Rightarrow$

$$A = \langle \psi_n, f \rangle = \langle \psi_n, c_0\psi_0 + \ldots + c_\ell \psi_\ell \rangle = c_0 \langle \psi_n, \psi_0 \rangle + \ldots + c_\ell \langle \psi_n, \psi_\ell \rangle = 0.$$

Tako $A = 0$. Dobili smp kontradikciju. Ne može biti $\ell < n$. Mora biti $\ell = n$. Lema je dokazana. Sve nule su jednostruke (proste).

Svi elementi niza ortogonalnih polinoma $\psi_0, \psi_1, \ldots$ određeni su jednoznačno do na multiplikativne konstante. Neka su te konstante izabrane tako da najstariji koeficijent svakog polinoma iz niza bude jednak 1, tj. da bude $\psi_k(x) = x^k + \ldots$.

Prelazimo na Gausovu kvadraturnu formulu

Neka je interval $[a,b]$ fiksiran i neka je težinska funkcija $p = p(x)$ fiksirana. Za težinsku funkciju se pretpostavlja da je integrabilna i da je skoro svuda pozitivna.

Neka je $n \geq 1$. Razmotrimo kvadraturnu formulu sa n čvorova $x_1, \ldots, x_n$:

$$I(f) \approx S(f), \qquad I(f) = \int_a^b f(x)p(x)dx, \qquad S(f) = \sum_{i=1}^{n} c_i f(x_i) \tag{1}$$

i uvedimo sljedeću oznaku za njenu grešku: $R(f) = I(f) - S(f)$.

Definišimo algebarski stepen tačnosti a kvadraturne formule (1). Za (1) se kaže da ima algebarski stepen tačnosti a ako je $R(f) = 0$ (greška je jednaka nuli, kvadraturna formula je tačna) kada je f – bilo koji polinom stepena $\leq$ a. A postoji bar jedan polinom f čiji je stepen jednak a $+ 1$ takav da je za njega $R(f) \neq 0$.

Gaus je razmatrao i riješio sljedeći ekstremalni zadatak: konstruisati kvadraturnu formulu oblika (1) sa najvećim mogućim algebarskim stepenom tačnosti. Vidjećemo da je rješenje zadatka a $= 2n-1$. Ovo nije neočekivano. Zaista, s jedne strane, vidimo da u (1) ima $2n$ stepeni slobode $c_1, \ldots, c_n, x_1, \ldots, x_n$ (pogodno ih izaberi). S druge strane, ima $2n$ uslova ako se traži da bude $R(f) = I(f) - S(f) = 0$ kada je $f(x) = x^k$ za $k = 0, \ldots, 2n-1$.

Dokažimo dvije leme.

Lema 1. Pretpostavimo da kvadraturna formula (1) ima svojstvo da je tačna za sve polinome f čiji je stepen $\leq 2n-1$. Neka je $\omega_n(x) = \prod_{i=1}^{n}(x-x_i)$. Neka je $P_{n-1} = P_{n-1}(x)$ proizvoljni polinom stepena $\leq n-1$. Tada je $\int_a^b \omega_n(x)P_{n-1}(x)p(x)dx = 0$.

Zapaziti da lema 1. ima oblik inplikacije. Ona polazi od toga da postoji kvadraturna formula oblika (1) sa svojstvom "tačna je za sve polinome f čiji je stepen $\leq 2n-1$". Ona ne tvrdi da takva formula postoji. Ona ne dokazuje da takva formula postoji.

Dokaz leme 1. Stavimo $f(x) = \omega_n(x)P_{n-1}(x)$. Proizvod ω_n je polinom stepena tačno n, P_{n-1} je polinom stepena $\leq n-1$, f je polinom stepena $\leq 2n-1$, zato je $R(f) = I(f) - S(f) = 0$, po pretpostavci. Imamo redom:

$$0 = R(f) = I(f) - S(f) = \int_a^b f(x)p(x)dx - \sum_{i=1}^n c_i f(x_i) =$$

$$\int_a^b \omega_n(x)P_{n-1}(x)p(x)dx - \sum_{i=1}^n c_i \left(\omega_n(x)P_{n-1}(x)\right)\Big|_{x=x_i} =$$

$$\int_a^b \omega_n(x)P_{n-1}(x)p(x)dx - \sum_{i=1}^n c_i \cdot 0 \cdot P_{n-1}(x_i) = \int_a^b \omega_n(x)P_{n-1}(x)p(x)dx - 0.$$

Lema 1. je dokazana.

Umjesto $\int_a^b \omega_n(x)P_{n-1}(x)p(x)dx = 0$ možemo pisati $\langle \omega_n, P_{n-1}\rangle = 0$, tj. $\omega_n \perp P_{n-1}$. Polinom ω_n je ortogonalan na sve polinome čiji je stepen $\leq n-1$. Dakle, mora biti $\omega_n = \psi_n$. Ovo određuje koje tačke su čvorovi kvadraturne formule. To su nule polinoma n–tog stepena ψ_n iz niza ortogonalnih polinoma $\{\psi_0, \psi_1, \ldots\}$.

Lema 2. Neka su $x_1, \ldots, x_n$ nule polinoma ψ_n. Neka je kvadraturna formula oblika (1) tačna za sve polinome f čiji je stepen $\leq n-1$. Tada je (1) tačna za sve polinome čiji je stepen $\leq 2n-1$.

Zapaziti da i lema 2. samo hipotetički govori o formuli oblika (1) koja bi sada imala svojstvo "tačna je (greška je nula) za sve polinome čiji je stepen $\leq n-1$".

Dokaz leme 2. Uzmimo ma koji polinom P_{2n-1} čiji je stepen $\leq 2n-1$. Podijelimo ga sa polinomom ψ_n. Ostatak dijeljenja je naravno nižeg stepena od ψ_n. Označimo količnik sa g_{n-1} a ostatak sa r_{n-1}: $P_{2n-1}(x) = g_{n-1}(x)\psi_n(x) + r_{n-1}(x)$, gdje su i g_{n-1} i r_{n-1} izvjesni polinomi stepena $\leq n-1$. Obavezni smo da pokažemo da je $R(P_{2n-1}) = 0$. Imamo redom:

$$R(P_{2n-1}) = R(g_{n-1}\psi_n + r_{n-1}) = \qquad \text{(integral je aditivan, formula (1) je aditivna)}$$

$$R(g_{n-1}\psi_n) + R(r_{n-1}) = \qquad \text{(pretpostavka leme)}$$

$$R(g_{n-1}\psi_n) + 0 = I(g_{n-1}\psi_n) - S(g_{n-1}\psi_n) =$$

$$\int_a^b g_{n-1}(x)\psi_n(x)p(x)dx - \sum_{i=1}^n c_i \left(g_{n-1}(x)\psi_n(x)\right)\Big|_{x=x_i} =$$

$$\langle g_{n-1}, \psi_n\rangle - \sum_{i=1}^n c_i g_{n-1}(x_i) \cdot 0 = 0,$$

jer $\psi_n \perp \psi_0, \ldots, \psi_{n-1} \Rightarrow \psi_n \perp$ linearna kombinacija od $\psi_0, \ldots, \psi_{n-1} \Rightarrow \psi_n \perp g_{n-1}$.

Lema 2. je dokazana.

Uz pomoć dvije leme, postavljeni zadatak sveo se na sljedeće. Čvorovi kvadraturne formule su definisani sa $\psi_n(x) = 0$. Još nisu definisani koeficijenti $c_1, \ldots, c_n$. Traži se da formula bude tačna za proizvoljnu funkciju f koja je polinom čiji je stepen $\leq n-1$, pa će onda automatski biti tačna i za sve polinome do stepena $2n-1$ uključeno, prema lemi 2. Lako se konstruiše formula koja je tačna (čija je greška $R(f)$ jednaka nuli) za polinome stepena $\leq n-1$. To je urađeno u prethodnom naslovu 2.4. gdje piše

$$\int_a^b f(x)p(x)dx \approx \int_a^b L_{n-1}(x)p(x)dx = \int_a^b \left(\sum_{i=1}^n \ell_i(x) f_i p(x)\right) dx = \sum_{i=1}^n c_i f_i,$$

L_{n-1} je L. i. p.

Drugim riječima, formula (1) je interpolacionog tipa.

U ovom trenutku je izvedena Gausova kvadraturna formula. Čvorovi su nule polinoma ψ_n, a koeficijenti se izračunaju po šablonu za kvadraturnu formulu u slučaju prisustva težinske funkcije.

Slijede dvije napomene.

Napomena 1. Ne postoji kvadraturna formula sa n čvorova oblika (1) čiji bi algebarski stepen tačnosti a bio jednak $2n$. Zaista, zamislimo da takva formula postoji; čvorovi su označeni kao $x_1, \dots, x_n$; neka je $\omega_n(x) = \prod_{i=1}^n (x - x_i)$. Posmatrajmo formulu kada je podintegralna funkcija $f(x) = \omega_n^2(x) = \prod_{i=1}^n (x - x_i)^2 \geq 0$. Zapazimo unaprijed da su sve vrijednosti f u čvorovima jednake nuli. Sada, s jedne strane, $I(f) = \int_a^b f(x)p(x)dx > 0$, a s druge strane $S(f) = \sum_{i=1}^n c_i f(x_i) = \sum_{i=1}^n c_i \cdot 0 = 0$. Tako da je $R(f) = I(f) - S(f) \neq 0$. Vidimo da je f polinom stepena tačno $2n$, tako da je dokaz završen. Tek sada je ekstremalni zadatak u potpunosti riješen.

Napomena 2. Svi koeficijenti Gausove kvadraturne formule su pozitivni, tj. za $i = 1, \dots, n$ važi $c_i > 0$. Zaista, neka podintegralna funkcija bude $f(x) = \prod_{j=1\ j\neq i}^n (x - x_j)^2$; f je polinom stepena tačno $2n - 2$; $f(x) \geq 0$. Tada je

$$I(f) = \int_a^b f(x)p(x)dx > 0 \quad \text{i} \quad S(f) = \sum_{j=1}^n c_j f(x_j) = c_i f(x_i) \quad \text{i} \quad R(f) = 0, \quad \text{tj.} \quad I(f) = S(f).$$

Slijedi $c_i f(x_i) > 0$. Slijedi $c_i > 0$ jer je $f(x_i) > 0$. Time je dokaz završen.

Ako se uvrsti $f(x) = 1$ onda $\int_a^b p(x)dx = \sum_{i=1}^n c_i = C$. Veličina C ne zavisi od n.

Svojstvo $c_i > 0$ za $i = 1, \dots, n$ daje numeričku stabilnost Gausove kvadraturne formule u odnosu na grešku ulaznih podataka $f(x_1), \dots, f(x_n)$; v. analizu u 2.4.

Ostaje da se ocijeni greška kvadraturne formule (1)

Neka $f \in C^{2n}[a,b]$. Pridružimo funkciji f njen L. i. p. $L_{n-1} = L_{n-1}(x)$ po tačkama $x_1, \dots, x_n$, stepena $\leq n - 1$. Pridružimo i Hermitov i. p. $H_{2n-1} = H_{2n-1}(x)$, stepena $\leq 2n - 1$, definisan sa $2n$ uslova:

$$H_{2n-1}(x_i) = f(x_i), \qquad H'_{2n-1}(x_i) = f'(x_i), \qquad i = 1, \dots, n$$

(ima n čvorova, svi čvorovi su dvostruki). Dalje imamo:

$$f(x) - H_{2n-1}(x) = \frac{1}{(2n)!} f^{(2n)}(\xi(x))\psi_n^2(x) \qquad \text{(v. u Interpolacija sa višestrukim čvorovima),}$$

$$\text{ponovimo da je} \qquad \psi_n(x) = \prod_{i=1}^n (x - x_i),$$

$$R(f) = I(f) - S(f) = \qquad \text{(formula (1) izvedena je po šablonu)}$$

$$I(f) - I(L_{n-1}) = \qquad (\text{stepen}(L_{n-1}) \leq \text{a} = 2n - 1)$$

$$I(f) - S(L_{n-1}) = \qquad \left(L_{n-1}(x_i) = H_{2n-1}(x_i) = f(x_i),\ \sum_{i=1}^n c_i L_{n-1}(x_i) = \sum_{i=1}^n c_i H_{2n-1}(x_i)\right)$$

$$I(f) - S(H_{2n-1}) = \qquad (\text{stepen}(H_{2n-1}) \le \text{a})$$

$$I(f) - I(H_{2n-1}) = \int_a^b f(x)p(x)dx - \int_a^b H_{2n-1}(x)p(x)dx =$$

$$\int_a^b \frac{1}{(2n)!} f^{(2n)}(\xi(x))\psi_n^2(x)p(x)dx = \qquad (\psi_n^2(x)p(x) \ge 0)$$

$$\frac{1}{(2n)!} f^{(2n)}(\xi) \int_a^b \psi_n^2(x)p(x)dx, \qquad a < \xi < b$$

(posljednji integral je konstanta za konkretnu kvadraturnu formulu). Imamo konačno:

$$R(f) = \frac{1}{(2n)!} f^{(2n)}(\xi) \int_a^b \psi_n^2(x)p(x)dx.$$

Na kraju, navedimo dva specijalna slučaja Gausove kvadraturne formule

Specijalni slučajevi odgovaraju redom prvom i drugom primjeru iz pripreme. Izvođenje (račun) za dva specijalna slučaja se izostavlja. Dokažite sami za vježbu. Sastavite i izraze za grešku.

Primjer 1. Neka je $[a, b] = [-1, 1]$ i $p(x) \equiv 1$. Tada

$$n = 1: \quad \int_{-1}^1 f(x)dx \approx 2f(0), \qquad n = 2: \quad \int_{-1}^1 f(x)dx \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right),$$

$$n = 3: \quad \int_{-1}^1 f(x)dx \approx \frac{5}{9} f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9} f(0) + \frac{5}{9} f\left(\sqrt{\frac{3}{5}}\right).$$

Primjer 2. Neka je $[a, b] = [-1, 1]$ i $p(x) = \dfrac{1}{\sqrt{1-x^2}}$. Tada

$$\int_{-1}^1 \frac{f(x)dx}{\sqrt{1-x^2}} \approx \frac{\pi}{n} \sum_{i=1}^n f_i, \quad \text{gdje je } \ x_i = \cos\frac{(2i-1)\pi}{2n}, \quad f_i = f(x_i) \quad (\text{Hermitova formula}).$$

3. NUMERIČKE METODE ALGEBRE

Numeričkim metodama algebre pripada numeričko rješavanje sljedećih zadataka: naći rješenje sistema linearnih jednačina, izvršiti inverziju matrice, izračunati vrijednost determinante, naći svojstvene / sopstvene vrijednosti i vektore matrice, odrediti nule polinoma.

Razmotrimo zadatak o rješavanju sistema od n linearnih jednačina sa n nepoznatih. Postoje tri vrste numeričkih metoda: (a) tačne ili direktne, (b) iterativne i (c) vjerovatnosne ili stohastičke. Ako je $n < 10^3$ onda treba primijeniti neku metodu oblika (a), ako je $10^3 < n < 10^6$ onda metodu oblika (b), ako je $n > 10^6$ onda metodu oblika (c), po pravilu.

Za numeričku metodu se kaže da je tačna metoda ako je njena greška metode jednaka nuli. Drugim riječima, za numeričku metodu se kaže da je tačna ako ona daje tačan rezultat (greška je jednaka nuli) nakon izvođenja konačno mnogo aritmetičkih i logičkih operacija, pod pretpostavkom da nema greške računanja i greške ulaznih podataka.

3.1. GAUSOVA METODA ELIMINACIJE

Ili Gausova metoda uzastopne eliminacije nepoznatih. To je jedna tačna metoda, služi za rješavanje sistema linearnih jednačina. Neka je $A = [a_{ij}]_{i,j=1}^n$ realna kvadratna matrica oblika $n \times n$. Razmotrimo sistem linearnih jednačina $A\mathbf{x} = \mathbf{b}$ ili

$$\sum_{j=1}^n a_{ij} x_j = b_i, \qquad 1 \le i \le n,$$

$\mathbf{x} = [x_1 \ \ldots \ x_n]^T$, $\mathbf{b} = [b_1 \ \ldots \ b_n]^T$, $n \ge 1$, A – matrica sistema, $\mathbf{b}$ – vektor slobodnih članova, $\mathbf{x}$ – nepoznata.

Formalno gledano, sistem $A\mathbf{x} = \mathbf{b}$ može da bude riješen primjenom Kramerovog pravila (izračuna se $n+1$ determinanta). Međutim, za vremensku složenost ili za potreban broj aritmetičkih operacija t_n tada važi $t_n > n!$ što čini da je taj postupak praktično neupotrebljiv već za recimo $n = 20$. Dodatno, što je broj aritmetičkih operacija veći to je i greška računanja veća. Gausova metoda eliminacije svodi t_n na $t_n = O(n^3)$, a ako se pod t_n podrazumijeva samo broj izvršenih operacija množenja i dijeljenja onda imamo $t_n \sim \frac{1}{3} n^3$.

Izložimo algoritam. Uzmimo da je $a_{11} \neq 0$. Veličina x_1 eliminiše se iz jednačina $i = 2, \ldots, i = n$. Upravo, prva jednačina $i = 1$ podijeli se sa a_{11} pa se onda nova prva jednačina, pomnožena sa $-a_{i1}$, doda i–toj jednačini za $i = 2, \ldots, i = n$. Sistem sada ima oblik:

$$x_1 + \sum_{j=2}^n a_{1j}^{(1)} x_j = b_1^{(1)},$$
$$\sum_{j=2}^n a_{ij}^{(1)} x_j = b_i^{(1)}, \qquad 2 \le i \le n.$$

Uzmimo da je $a_{22}^{(1)} \neq 0$. Veličina x_2 eliminiše se iz jednačina $i = 3, \ldots, i = n$ na sličan način kao maločas. Itd. Time se polazni sistem svede na gornje trougaoni oblik:

$$x_i + \sum_{j=i+1}^n a_{ij}^{(i)} x_j = b_i^{(i)}, \qquad 1 \le i \le n$$

ili

$$\begin{bmatrix} 1 & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & 1 & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{bmatrix}.$$

Ako je $a_{11} = 0$ onda se izvrši pogodna zamjena mjesta jednačina tako da time postane $a_{11} \neq 0$. Slično ako se desi da je $a_{22}^{(1)} = 0$, itd. Zapaziti da postoje dvije mogućnosti: prva: zamjenom mjesta se postigne da su svi eliminatorni elementi $a_{11}, a_{22}^{(1)}, \ldots$ različiti od nule i druga: to ne može da bude postignuto. Lako se vidi da prva mogućnost odgovara slučaju da je $\det A \neq 0$ (matrica A je regularna, sistem $A\mathbf{x} = \mathbf{b}$ ima jedinstveno rješenje), dok druga mogućnost nastaje kada je $\det A = 0$ (matrica A je singularna, nije tačno da sistem $A\mathbf{x} = \mathbf{b}$ ima jedinstveno rješenje). Dakle, tokom realizacije algoritma pokazaće se da li je $\det A \neq 0$ ili je pak suprotno $\det A = 0$.

Prema tome, pretpostavimo da je $\det A \neq 0$.

Gausov algoritam eliminacije sastoji se iz dva dijela: (a) direktni hod i (b) inverzni hod. Direktni hod algoritma obuhvata računanja kojima se polazni sistem $A\mathbf{x} = \mathbf{b}$ svede na gornje trougaoni oblik. U inverznom hodu se pređe put od trougaonog oblika do nalaženja rješenja sistema $\mathbf{x}$.

Ostaje da se izloži (b). Nastali trougaoni sistem se trivijalno rješava. Iz jednačine $i = n$ odredi se x_n, zatim se iz jednačine $i = n - 1$ odredi x_{n-1}, itd. Formula glasi $x_i = b_i^{(i)} - \sum_{j=i+1}^{n} a_{ij}^{(i)} x_j$, $i = n, \ldots, 1$.

3.2. GAUSOVA METODA ELIMINACIJE SA IZBOROM GLAVNOG ELEMENTA

Riječ je opet o jednoj tačnoj metodi za rješavanje sistema linearnih jednačina, a predstavlja jednu modifikaciju ili poboljšanje metode izložene u prethodnom naslovu. U čemu se sastoji modifikacija? Izbor eliminatornog ili glavnog elementa više se ne prepušta slučaju. Glavni element se bira po kriterijumu: neka bude što je moguće veći po apsolutnoj vrijednosti. Dakle, vrši se upravljanje tokom računskih operacija, vrši se vođenje ili pivotiranje. Šta se time dobija? Prethodno, praksa ili iskustvo pokazuje da algoritam iz prethodnog naslova (bez vođenja) ima jednu slabu stranu: uticaj greške računanja često sasvim pokvari numerički rezultat. Takođe se i eventualno prisutna greška ulaznih podataka često veoma negativno odrazi na grešku numeričkog rezultata. Ako se vrši vođenje onda se rezultujuća greška radikalno smanji. Prva aksioma numeričkih metoda – prilikom rješavanja sistema linearnih jednačina direktnom metodom obavezno vršiti izbor glavnog elementa.

U prvom dijelu izlaganja biće izložen sami algoritam. U drugom dijelu izlaganja biće dato obrazloženje za – korisno je vršiti vođenje.

Razmotrimo sistem linearnih jednačina $A\mathbf{x} = \mathbf{b}$. Modifikacija se odnosi na direktni hod algoritma. U prvom koraku se za glavni element izabere najveći po apsolutnoj vrijednosti član čitave matrice A, odnosno najveći po apsolutnoj vrijednosti među brojevima a_{ij}, $1 \leq i \leq n$, $1 \leq j \leq n$. Neka je to a_{kl}. Broj a_{kl} dovodi se na mjesto a_{11}. Jednačine $i = 1$ i $i = k$ zamijene mjesta. Pored toga, neka kolone $j = 1$ i $j = l$ zamijene mjesta, čime promjenljive x_1 i x_l zamijene uloge; voditi računa o ovoj permutaciji. U drugom koraku se za glavni element izabere najveći po apsolutnoj vrijednosti među brojevima $a_{ij}^{(1)}$, $2 \leq i \leq n$, $2 \leq j \leq n$. Slično u trećem koraku, itd. u n–tom koraku. Kada se trougaoni sistem riješi, onda razdužiti permutacije, u obrnutom redosljedu.

Napišimo grubi algoritam. Program ima dvije osnovne promjenljive: matrica A oblika $n \times n$ i niz $\mathbf{b}$ dužine n. Prvo se učitaju ulazni podaci n, A i $\mathbf{b}$.

Za svako i od 1 do n ponovi
{ 1. Među brojevima a_{kl}, $i \le k \le n$, $i \le l \le n$ pronađi najveći po apsolutnoj vrijednosti, neka je to a_{kl}
2. Ako je $a_{kl} = 0$ onda se štampa "det $A = 0$" i program se zaustavlja
3. Neka i–ta i k–ta vrsta matrice A zamijene mjesta i neka b_i i b_k zamijene mjesta
4. Neka i–ta i l–ta kolona matrice A zamijene mjesta i neka se zapamti koja je permutacija kolona izvršena
5. Učini da postane $a_{ii} = 1$, tj. i–tu jednačinu podijeli sa a_{ii}; konkretno:
$a_{ij} \leftarrow 0$ za $1 \le j \le i-1$, $a_{ij} \leftarrow a_{ij}/a_{ii}$ za $i+1 \le j \le n$, $b_i \leftarrow b_i/a_{ii}$, $a_{ii} \leftarrow 1$
6. Za svako m od $i+1$ do n ponovi { Učini da postane $a_{mi} = 0$, tj. i–tu jednačinu pomnoženu sa $-a_{mi}$ dodaj m–toj jednačini; konkretno:
$a_{mj} \leftarrow 0$ za $1 \le j \le i-1$, $a_{mj} \leftarrow a_{mj} - a_{mi}a_{ij}$ za $i+1 \le j \le n$, $b_m \leftarrow b_m - a_{mi}b_i$, $a_{mi} \leftarrow 0$ } }

7. Sada je A gornje trougaona matrica. Za svako i od n naniže do 1 ponovi $x_i \leftarrow b_i - \sum_{j=i+1}^{n} a_{ij}x_j$

8. Razduži permutacije i štampaj rezultat $x_1, \ldots, x_n$.

Zapaziti da pronalaženje glavnog elementa samo malo povećava vremensku složenost t_n.

Može se vršiti samo djelimično vođenje: glavni element traži se samo u koloni, tj. među brojevima $a_{ki}^{(i-1)}$, $i \le k \le n$. Tada nema permutacija.

Za obrazloženje, pogledajmo po kom zakonu se mijenja opšti član matrice a_{mj} u i–tom koraku:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \ldots \\ a_{ii}x_i + a_{i,i+1}x_{i+1} + \ldots + a_{ij}x_j + \ldots + a_{in}x_n = b_i \\ \ldots \\ a_{mi}x_i + a_{m,i+1}x_{i+1} + \ldots + a_{mj}x_j + \ldots + a_{mn}x_n = b_m \\ \ldots \end{cases}$$

$$a_{mj} \leftarrow a_{mj} - \frac{a_{mi}a_{ij}}{a_{ii}}.$$

Sve članove matrice smatramo približnim brojevima zbog greške računanja koja se do datog trenutka akumulirala. Vrše se nova računanja sa članovima matrice. Pogledajmo grešku izraza na desnoj strani u posljednjoj relaciji, u zavisnosti od veličine eliminatornog elementa a_{ii}. Pogledajmo uže grešku izraza $\frac{a_{mi}a_{ij}}{a_{ii}}$. Uvedimo oznake $x = a_{mi}a_{ij}$, $y = a_{ii}$ i $f(x,y) = \frac{x}{y}$. Neka x ima grešku $\Delta(x)$ i neka y ima grešku $\Delta(y)$. Kako se to odražava na grešku broja $f(x,y)$? Naslov Greška funkcije, formula za grešku funkcije od dvije promjenljive:

$$x = x^* + \Delta(x), \qquad y = y^* + \Delta(y), \qquad f(x,y) = f(x^*,y^*) + \Delta(f),$$

$$\frac{\partial f}{\partial x} = \frac{1}{y}, \qquad \frac{\partial f}{\partial y} = -\frac{x}{y^2},$$

$$\Delta(f) \approx \frac{\partial f(x,y)}{\partial x}\Delta(x) + \frac{\partial f(x,y)}{\partial y}\Delta(y) = \frac{1}{y}\Delta(x) - \frac{x}{y^2}\Delta(y).$$

Što je broj $|y| = |a_{ii}|$ veći to je greška $|\Delta(f)|$ manja. Ako se eliminatorni element udvostruči onda se greška približno prepolovi i slično. Pogodno biranje eliminatornog elementa a_{ii} doprinosi iz koraka u korak smanjivanju greške.

3.3. MJERA USLOVLJENOSTI MATRICE

Uvod. Kvadratnoj matrici A biće pridružen broj u oznaci $\mathrm{cond}(A)$ (č. kondicija) ili $M(A)$, a naziva se njenom mjerom uslovljenosti ili kondicijom ili kondicionim brojem. Kod sistema linearnih jednačina $A\mathbf{x} = \mathbf{b}$, greška ulaznih podataka odražava se na grešku numeričkog rezultata u stepenu koji zavisi od mjere uslovljenosti matrice A: što je mjera uslovljenosti veća to se više odražava. Može se pokazati da slične okolnosti važe i kada je riječ o greški računanja.

Napomena. Znamo da se na matricu $A \in R^{n\times n}$ može gledati kao na linearni operator $A\colon R^n \to R^n$, gdje je $\mathbf{y} = A\mathbf{x}$. Tada se $\mathbf{x}$ i $\mathbf{y}$ ponekad zapisuju kao $\mathbf{x} = (x_1, \ldots, x_n)$ odnosno $\mathbf{y} = (y_1, \ldots, y_n)$. To su samo drukčije oznake zato što ostaje naravmo $y_i = \sum_{j=1}^{n} a_{ij}x_j$ $(1 \le i \le n)$.

O normi matrice. U skupu R^n, norma vektora $\mathbf{x}$ u oznaci $\|\mathbf{x}\|$ može da bude uvedena na razne načine; neka $\|\mathbf{x}\|$ označava jednu moguću normu za koju smo se opredijelili. Znamo da važi $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\|+\|\mathbf{y}\|$, nejednakost trougla. Neka je $A \in R^{n\times n}$ kvadratna matrica oblika $n\times n$ ili neka je $A\colon R^n \to R^n$ linearni operator u prostoru R^n. Neka $\|A\|$ označava normu linearnog operatora A koja je saglasna sa uvedenom normom vektora (koja odgovara uvedenoj normi vektora), a definiše se relacijom $\|A\| = \sup_{\mathbf{x}\ne 0} \|A\mathbf{x}\|/\|\mathbf{x}\|$; tzv. indukovana norma. Vidimo da važi nejednakost $\|A\mathbf{x}\| \le \|A\| \cdot \|\mathbf{x}\|$. Označimo jediničnu matricu kao $I \in R^{n\times n}$; važi $\|I\| = 1$. Ako $A \in R^{n\times n}$ i $B \in R^{n\times n}$ onda važi $\|A + B\| \le \|A\| + \|B\|$ i $\|AB\| \le \|A\| \cdot \|B\|$.

Definicija mjere uslovljenosti matrice.

Definicija. Za $A \in R^{n\times n}$, $\det A \ne 0$, stavlja se

$$\mathrm{cond}(A) = \|A\| \cdot \|A^{-1}\|.$$

Ako je $\det A = 0$ onda $\mathrm{cond}(A)$ nije definisano. Neki uzimaju da je tada $\mathrm{cond}(A) = +\infty$.

Koliko najmanje može da bude $\mathrm{cond}(A)$? Iz $AA^{-1} = I$ imamo $\|AA^{-1}\| = \|I\| = 1$ i $\|A\| \cdot \|A^{-1}\| \ge 1$, tj. $\mathrm{cond}(A) \ge 1$.

Ako je broj $\mathrm{cond}(A)$ mali, tj. ako je broj $\mathrm{cond}(A)$ blizu donje granice 1 onda se za matricu A kaže da je dobro uslovljena. Tada je odgovarajući sistem linearnih jednačina $A\mathbf{x} = \mathbf{b}$ manje osjetljiv na grešku ulaznih podataka i na grešku računanja, kao što ćemo vidjeti. Ako je broj $\mathrm{cond}(A)$ veliki, tj. ako je broj $\mathrm{cond}(A)$ blizu gornje granice $+\infty$ onda se za matricu A kaže da je slabo uslovljena. Tada odgovarajući sistem linearnih jednačina $A\mathbf{x} = \mathbf{b}$ ima slaba svojstva numeričke stabilnosti, tj. sistem je nepogodan za numeričko rješavanje, kao što ćemo vidjeti.

Navedimo jedno elementarno svojstvo karakteristike $\mathrm{cond}(A)$. Prethodno, ako je $A\mathbf{x} = \lambda\mathbf{x}$, gdje je $\mathbf{x} \ne 0$ onda se za broj λ kaže da je svojstvena vrijednost matrice A, a za $\mathbf{x}$ se kaže da je odgovarajući svojstveni vektor, kao što je poznato iz linearne algebre. Lako se vidi da je $|\lambda| \le \|A\|$, riječima: norma matrice je veća ili jednaka od apsolutne vrijednosti bilo koje njene svojstvene vrijednosti. Ima li kakve veze između svojstvenih vrijednosti dvije matrice A i A^{-1}?

$$A\mathbf{x} = \lambda\mathbf{x}, \quad A^{-1}A\mathbf{x} = A^{-1}\lambda\mathbf{x}, \quad \mathbf{x} = \lambda A^{-1}\mathbf{x}, \quad A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x},$$

riječima: svojstvene vrijednosti inverzne matrice jednake su recipročnim vrijednostima svojstvenih vrijednosti same matrice, sa jednim te istim odgovarajućim svojstvenim vektorom. Važi: $\det A = 0 \Leftrightarrow \lambda = 0$ je svojstvena vrijednost matrice A. Svojstvo $\mathrm{cond}(A)$:

$$\mathrm{cond}(A) \geq \frac{|\lambda_2|}{|\lambda_1|},$$

gdje su $|\lambda_2|$ i $|\lambda_1|$ redom najveći odnosno najmanji među svim brojevima $|\lambda|$, λ – svojstvena vrijednost matrice A. Zaista,

$$\|A\| \geq |\lambda_2|, \quad \|A^{-1}\| \geq \frac{1}{|\lambda_1|} \quad \Rightarrow \quad \|A\| \cdot \|A^{-1}\| \geq \frac{|\lambda_2|}{|\lambda_1|}.$$

Provjeriti navedeno svojstvo $\mathrm{cond}(A)$ u slučaju $\|\mathbf{x}\| = \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$, A – dijagonalna matrica.

Teorema o uticaju greške ulaznih veličina.

Razmotrimo zadatak o rješavanju sistema linearnih jednačina $A\mathbf{x} = \mathbf{b}$ čiji su ulazni podaci A i $\mathbf{b}$ samo približno poznate veličine. I rješenje $\mathbf{x}$ ćemo saznati samo približno, greška ulaznih podataka očito utiče na grešku rješenja. Uvedimo potrebne oznake. Razmotrimo sistem $A\mathbf{x} = \mathbf{b}$, gdje se pretpostavlja da je $\det A \neq 0$. Neka je A^* raspoloživa približna vrijednost matrice sistema. Uvedimo oznaku δA za odgovarajuće odstupanje, tj. neka je $A = A^* + \delta A$. Neka je $\mathbf{b}^*$ raspoloživa približna vrijednost vektora slobodnih članova. Uvedimo oznaku $\delta\mathbf{b}$ za odgovarajuće odstupanje, tj. neka je $\mathbf{b} = \mathbf{b}^* + \delta\mathbf{b}$. Pretpostavimo da je i $\det A^* \neq 0$. Mi riješimo sistem $A^*\mathbf{x}^* = \mathbf{b}^*$ (sa $\mathbf{x}^*$ smo označili njegovo rješenje) i saopštimo naš numerički odgovor $\mathbf{x} \approx \mathbf{x}^*$. Kolika je greška? Označimo odgovarajuće odstupanje kao $\delta\mathbf{x}$, tj. neka bude $\mathbf{x} = \mathbf{x}^* + \delta\mathbf{x}$; $\delta\mathbf{x} \in R^n$ je apsolutna greška numeričkog odgovora. Dok su $\delta A \in R^{n\times n}$ i $\delta\mathbf{b} \in R^n$ očito apsolutne greške matrice sistema i vektora slobodnih članova, redom. Možemo razmatrati i relativne greške $\|\delta\mathbf{x}\|/\|\mathbf{x}\|$, $\|\delta A\|/\|A\|$ i $\|\delta\mathbf{b}\|/\|\mathbf{b}\|$. Sljedeća teorema izražava relativnu grešku rješenja $\|\delta\mathbf{x}\|/\|\mathbf{x}\|$ preko relativnih grešaka ulaznih podataka i veličine $\mathrm{cond}(A)$. Prije teoreme dolazi jedna lema.

Banahova lema. Neka je C kvadratna matrica koja zadovoljava $\|C\| < 1$. Tada postoji matrica $(I-C)^{-1}$ i važi ocjena

$$\|(I-C)^{-1}\| \leq \frac{1}{1-\|C\|}.$$

Dokaz leme. Za bilo koji vektor $\mathbf{x}$ imamo

$$\|(I-C)\mathbf{x}\| = \|\mathbf{x} - C\mathbf{x}\| \geq \|\mathbf{x}\| - \|C\mathbf{x}\| \geq \|\mathbf{x}\| - \|C\| \cdot \|\mathbf{x}\| = \delta\|\mathbf{x}\|,$$

gdje je $\delta = 1 - \|C\| > 0$. Ako je $(I-C)\mathbf{x} = 0$ onda je $0 \geq \delta\|\mathbf{x}\| \Rightarrow \|\mathbf{x}\| = 0$, $\mathbf{x} = 0$. Vidimo da homogeni sistem $(I-C)\mathbf{x} = 0$ ima samo trivijalno rješenje. Tako da je matrica $I - C$ regularna (invertibilna). Zato u nejednakosti $\|(I-C)\mathbf{x}\| \geq \delta\|\mathbf{x}\|$ možemo uvesti oznake $(I-C)\mathbf{x} = \mathbf{y}$ i $\mathbf{x} = (I-C)^{-1}\mathbf{y}$ pa ta nejednakost dobija oblik

$$\|\mathbf{y}\| \geq \delta\|(I-C)^{-1}\mathbf{y}\| \quad \Rightarrow \quad \|(I-C)^{-1}\mathbf{v}\| \leq \frac{1}{\delta}\|\mathbf{v}\|.$$

Budući da vektor $\mathbf{y}$ prolazi kroz čitav skup R^n to slijedi

$$\|(I-C)^{-1}\| \leq \frac{1}{\delta} = \frac{1}{1-\|C\|}.$$

Lema je dokazana. I – jedinična matrica.

Teorema. Neka matrica A^{-1} postoji i neka je $\|\delta A\|\cdot\|A^{-1}\| < 1$. Tada postoji i matrica $(A-\delta A)^{-1}$ i važi nejednakost

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \frac{1}{1-\mathrm{cond}(A)\frac{\|\delta A\|}{\|A\|}}\mathrm{cond}(A)\left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}+\frac{\|\delta A\|}{\|A\|}\right), \tag{1}$$

gdje je $A\mathbf{x}=\mathbf{b}$ i $(A-\delta A)(\mathbf{x}-\delta\mathbf{x})=\mathbf{b}-\delta\mathbf{b}$.

Dokaz teoreme. $A\mathbf{x}=\mathbf{b} \quad A^*\mathbf{x}^*=\mathbf{b}^* \quad A=A^*+\delta A \quad \mathbf{b}=\mathbf{b}^*+\delta\mathbf{b} \quad \mathbf{x}=\mathbf{x}^*+\delta\mathbf{x}$,

$$(A-\delta A)(\mathbf{x}-\delta\mathbf{x})=\mathbf{b}-\delta\mathbf{b}$$

$$A\mathbf{x}-\delta A\mathbf{x}-A\delta\mathbf{x}+\delta A\delta\mathbf{x}=\mathbf{b}-\delta\mathbf{b} \quad /\cdot(-1)$$

$$(A-\delta A)\delta\mathbf{x}=\delta\mathbf{b}-\delta A\mathbf{x} \quad /\cdot A^{-1}$$

$$(I-A^{-1}\delta A)\delta\mathbf{x}=A^{-1}\delta\mathbf{b}-A^{-1}\delta A\mathbf{x} \qquad (I-A^{-1}\delta A \text{ je invertibilna, lema})$$

$$\delta\mathbf{x}=(I-A^{-1}\delta A)^{-1}(A^{-1}\delta\mathbf{b}-A^{-1}\delta A\mathbf{x})$$

$$\|\delta\mathbf{x}\|\le\|(I-A^{-1}\delta A)^{-1}\|\big(\|A^{-1}\|\cdot\|\delta\mathbf{b}\|+\|A^{-1}\|\cdot\|\delta A\|\cdot\|\mathbf{x}\|\big) \qquad \left(\frac{\|A\mathbf{x}\|}{\|\mathbf{b}\|}=1 \;\Rightarrow\; \frac{\|A\|\cdot\|\mathbf{x}\|}{\|\mathbf{b}\|}\ge 1\right)$$

$$\|\delta\mathbf{x}\|\le\|(I-A^{-1}\delta A)^{-1}\|\left(\|A^{-1}\|\cdot\|\delta\mathbf{b}\|\frac{\|A\|\cdot\|\mathbf{x}\|}{\|\mathbf{b}\|}+\|A^{-1}\|\cdot\|\delta A\|\cdot\|\mathbf{x}\|\frac{\|A\|}{\|A\|}\right) \qquad \text{(lema)}$$

$$\|\delta\mathbf{x}\|\le\frac{1}{1-\|A^{-1}\delta A\|}\left(\mathrm{cond}(A)\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}\|\mathbf{x}\|+\mathrm{cond}(A)\frac{\|\delta A\|}{\|A\|}\|x\|\right) \quad /:\|\mathbf{x}\|$$

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}\le\frac{1}{1-\|A^{-1}\|\cdot\|\delta A\|}\mathrm{cond}(A)\left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}+\frac{\|\delta A\|}{\|A\|}\right).$$

Teorema je dokazana, budući da je

$$\|A^{-1}\|\cdot\|\delta A\|=\|A^{-1}\|\cdot\|\delta A\|\frac{\|A\|}{\|A\|}=\mathrm{cond}(A)\frac{\|\delta A\|}{\|A\|}.$$

Pogledajmo nejednakost (1) kada je $\delta A\approx 0$ i $1-\mathrm{cond}(A)\frac{\|\delta A\|}{\|A\|}\approx 1$. Vidimo da se tada relativna greška rješenja $\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}=\frac{\|\mathbf{x}-\mathbf{x}^*\|}{\|\mathbf{x}\|}$ procjenjuje mjerom uslovljenosti matrice sistema $\mathrm{cond}(A)$ puta zbir relativnih grešaka matrice sistema i vektora slobodnih članova $\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$.

Specijalan slučaj teoreme dobijamo kada stavimo da je $\|\delta A\|=0$, matrica sistema je data tačno. Vidimo da se tada relativna greška desne strane sistema (vektora slobodnih članova) prenosi na relativnu grešku rješenja sistema sa koeficijentom uveličavanja koji je jednak $\mathrm{cond}(A)$. Formulom: ako je $A\mathbf{x}=\mathbf{b}$, $A\mathbf{x}^*=\mathbf{b}^*$, $\mathbf{b}=\mathbf{b}^*+\delta\mathbf{b}$ i $\mathbf{x}=\mathbf{x}^*+\delta\mathbf{x}$ onda je

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}\le\mathrm{cond}(A)\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}. \tag{2}$$

Drugi specijalan slučaj teoreme dobijamo ako stavimo da je $\|\delta\mathbf{b}\| = 0$. Sada je desna strana sistema data tačno, dok je matrica sistema poznata samo približno tačno.

Primjedba. Prilikom rješavanja sistema $A\mathbf{x} = \mathbf{b}$ direktnom metodom, računar ne saopštava tačno rješenje $\mathbf{x}$, već će računar saopštiti približno rješenje $\mathbf{x}^*$, zbog prisustva greške računanja. Kako da se ocijeni greška $\mathbf{x} - \mathbf{x}^*$? Uvrstite $\mathbf{x}^*$ u sistem, tj. izračunajte vektor $\mathbf{b} - A\mathbf{x}^* = \mathbf{b} - \mathbf{b}^* = \delta\mathbf{b}$. Za vektor $\delta\mathbf{b}$ kaže se da predstavlja tzv. ekvivalentnu smetnju. Ono što je računar uradio ekvivalentno je sa time da je tačno (bez greške računanja) riješen sistem čiji je vektor slobodnih članova dat približno. Dovoljno je da se primijeni formula (2).

3.4. ITERATIVNE METODE ZA RJEŠAVANJE SISTEMA LINEARNIH JEDNAČINA

O normi vektora

Neka je $n \geq 1$ i razmotrimo skup R^n. Postoje razne funkcije $\|\ \|\colon R^n \to R$ koje ispunjavaju aksiome norme. Skup R^n zajedno sa jednom takvom normom postaje normirani prostor.

Za dvije norme $\|\ \|_a$ i $\|\ \|_b$ koje su definisane u jednom te istom skupu X kaže se da su ekvivalentne ako postoje konstante $\alpha > 0$ i $\beta > 0$ takve da važi $\alpha\|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq \beta\|\mathbf{x}\|_a$ za svako $\mathbf{x} \in X$. Poznato je sljedeće tvrđenje: bilo koje dvije norme u skupu R^n su ekvivalentne. Zato pitanje konvergencije niza $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ čiji članovi pripadaju R^n ne zavisi od izbora norme i svodi se na pitanje konvergencije po koordinatama (konvergencije n brojnih nizova). Dakle, ako niz konvergira po jednoj normi onda taj niz konvergira i po bilo kojoj drugoj normi, tako da se tada prosto kaže da niz konvergira.

Tri norme u vektorskom prostoru R^n koje se često koriste označavaju se kao $\|\ \|_1$, $\|\ \|_2$ i $\|\ \|_\infty$ a definišu se sa

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \quad \text{i} \quad \|\mathbf{x}\|_\infty = \max_{1\leq i\leq n} |x_i|,$$

gdje je $\mathbf{x} = (x_1, \ldots, x_n) \in R^n$. Znamo da su prva i druga norma specijalni slučajevi norme $\|\mathbf{x}\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}$, gdje je $p \geq 1$. Isto tako znamo da važi $\lim_{p\to+\infty} \|\mathbf{x}\|_p = \|\mathbf{x}\|_\infty$. Slično je $\|\ \|_p$ norma i u skupu C^n, gdje je $1 \leq p \leq +\infty$.

Razmotrimo matricu $A = [a_{ij}]_{i,j=1}^{n} \in R^{n\times n}$. Relacija $\|A\| = \sup_{\mathbf{x}\neq 0} \|A\mathbf{x}\|/\|\mathbf{x}\|$ definiše normu matrice A koja je indukovana normom vektora $\mathbf{x} \in R^n$ u istoj oznaci $\|\mathbf{x}\|$. Za tri uobičajene norme u R^n imamo sljedeće eksplicitne izraze za odgovarajuće norme matrice:

$$\|A\|_1 = \max_{1\leq j\leq n} \left(\sum_{i=1}^{n} |a_{ij}|\right), \quad \|A\|_2 = \sqrt{\max_{1\leq i\leq n} \lambda_i(A^TA)} \quad \text{i} \quad \|A\|_\infty = \max_{1\leq i\leq n} \left(\sum_{j=1}^{n} |a_{ij}|\right);$$

trebalo bi ovo dokazati; A^T je transponovana matrica matrice A, a $\lambda_i(A^TA)$ su svojstvene vrijednosti matrice A^TA. Ako je matrica A simetrična $(A = A^T)$ onda važi $\|A\|_2 = \max_{1\leq i\leq n} |\lambda_i(A)|$, gdje su sa $\lambda_i(A)$ označene svojstvene vrijednosti matrice A.

Brojevi λ^2 su svojstvene vrijednosti matrice A^2.

U slučaju $A \in C^{n\times n}$ ne mijenja se ništa kod $\|A\|_1$ i $\|A\|_\infty$ a mijenja se samo malo kod $\|A\|_2$ kako slijedi: $\|A\|_2 = \sqrt{\max_{1\leq i\leq n} \lambda_i(A^*A)}$, gdje je A^* konjugovana matrica. Ako je A samokonjugovana $(A = A^*)$ onda važi $\|A\|_2 = \max_{1\leq i\leq n} |\lambda_i(A)|$.

Uvod

Biće riječi o metodi proste iteracije za rješavanje sistema linearnih jednačina. Za iterativnu metodu se kaže da je metoda proste iteracije ako se ona zasniva na principu kontrakcije.

Priprema: princip kontrakcije

Kaže se princip kontrakcije ili princip fiksne tačke ili Banahova teorema.

Teorema. Neka je (X, d) kompletan metrički prostor i razmotrimo preslikavanje φ: $X \to X$. Neka φ zadovoljava uslov kontrakcije: postoji broj q $(0 \leq q < 1)$ takav da za sve $x, y \in X$ važi $d(\varphi(x), \varphi(y)) \leq q d(x, y)$. Dalje, neka je izabran $x_0 \in X$ na proizvoljan način i definišimo niz $\{x_n\}_{n=0}^{\infty}$ relacijom: $x_{n+1} = \varphi(x_n)$, $n \geq 0$. Tada: (1) jednačina $\varphi(x) = x$ ima samo jedno rješenje u oznaci X i (2) važi $\lim_{n\to\infty} x_n = X$.

O pojmovima i terminima. Za metrički prostor se kaže da je kompletan ako u njemu svaki Košijev niz konvegira. Za niz $\{x_n\}_{n=1}^{\infty}$ kaže se da je Košijev / fundamentalan ako važi: za svako $\varepsilon > 0$, postoji n_0, takav da za sve $n \geq n_0$ i $p \geq 1$ važi $d(x_{n+p}, x_n) < \varepsilon$. Za razmatrano preslikavanje φ kaže se da predstavlja kontrakciju, a za broj q se kaže da je koeficijent kontrakcije. Za element x koji zadovoljava jednačinu $\varphi(x) = x$ kaže se da predstavlja fiksnu / nepokretnu tačku preslikavanja φ.

Dokaz teoreme. Prvo ćemo pokazati da je niz $\{x_n\}_{n=0}^{\infty}$ Košijev. Imamo redom

$$d(x_{n+2}, x_{n+1}) = d(\varphi(x_{n+1}), \varphi(x_n)) \leq q d(x_{n+1}, x_n),$$

$$d(x_{n+3}, x_{n+2}) = d(\varphi(x_{n+2}), \varphi(x_{n+1})) \leq q d(x_{n+2}, x_{n+1}) \leq q^2 d(x_{n+1}, x_n), \quad \text{itd.}$$

$$d(x_{n+p}, x_n) \leq d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \ldots + d(x_{n+1}, x_n) \leq$$

$$q^{p-1} d(x_{n+1}, x_n) + \ldots + q d(x_{n+1}, x_n) + d(x_{n+1}, x_n) =$$

$$(1 + q + \ldots + q^{p-1}) d(x_{n+1}, x_n) \leq (1 + q + \ldots) d(x_{n+1}, x_n) = \frac{q}{1-q} d(x_{n+1}, x_n) =$$

$$\frac{q}{1-q} d(\varphi(x_n), \varphi(x_{n-1})) \leq \frac{q^2}{1-q} d(x_n, x_{n-1}) = \ldots \leq \frac{q^n}{1-q} d(x_1, x_0).$$

Dobili smo $d(x_{n+p}, x_n) \leq \frac{q^n}{1-q} d(x_1, x_0)$, tako da $d(x_{n+p}, x_n) \to 0$ kad $n \to \infty$, bez obzira na $p \geq 1$. Jeste Košijev. Odatle, budući da je metrički prostor kompletan, niz $\{x_n\}_{n=0}^{\infty}$ je i konvergentan. Označimo sa X njegovu graničnu vrijednost, $X = \lim_{n\to\infty} x_n$.

Preslikavanje φ je neprekidno. Zaista, $d(x, x') < \varepsilon \Rightarrow d(\varphi(x), \varphi(x')) < \varepsilon$. Iz neprekidnosti, $\lim_{n\to\infty} \varphi(x_n) = \varphi(\lim_{n\to\infty} x_n)$.

Primijenimo operaciju $\lim_{n\to\infty}$ na relaciju $x_{n+1} = \varphi(x_n)$. Tako $\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \varphi(x_n)$, $X = \lim_{n\to\infty} \varphi(x_n)$, $X = \varphi(\lim_{n\to\infty} x_n)$, $X = \varphi(X)$, fiksna tačka postoji.

Ne mogu postojati dvije fiksne tačke X_1, X_2 $(X_1 \neq X_2)$. Zaista, tada bi bilo $d(X_1, X_2) = d(\varphi(X_1), \varphi(X_2)) \leq q d(X_1, X_2)$, što je moguće jedino u slučaju $d(X_1, X_2) = 0$, $X_1 = X_2$. Fiksna tačka je jedinstvena. Dokaz je završen.

Naredna teorema predstavlja nastavak prethodne teoreme i služi za ocjenu greške n–te uzastopne aproksimacije x_n.

Teorema. Važe nejednakosti

$$d(X, x_n) \leq \frac{q^n}{1-q} d(x_1, x_0) \qquad (*), \qquad\qquad d(X, x_n) \leq \frac{q}{1-q} d(x_n, x_{n-1}) \qquad (**).$$

Dokaz teoreme. Napišimo nejednakost trougla za tačke x_n, x_{n+1}, X:

$$d(x_n, X) \leq d(x_n, x_{n+1}) + d(x_{n+1}, X) =$$

$$d(\varphi(x_{n-1}), \varphi(x_n)) + d(\varphi(x_n), \varphi(X)) \leq qd(x_{n-1}, x_n) + qd(x_n, X) \quad \Rightarrow$$

$$(1-q)d(x_n, X) \leq qd(x_{n-1}, x_n), \qquad d(x_n, X) \leq \frac{q}{1-q} d(x_{n-1}, x_n),$$

čime smo pokazali (**). Treba (*). Iz uslova kontrakcije:

$$d(X, x_n) \leq \frac{q}{1-q} d(x_n, x_{n-1}) = \frac{q}{1-q} d(\varphi(x_{n-1}), \varphi(x_{n-2})) \leq \frac{q^2}{1-q} d(x_{n-1}, x_{n-2}) = \ldots \leq$$

$$\frac{q^n}{1-q} d(x_1, x_0).$$

Dokaz je završen.

Princip kontrakcije važi i u Banahovom prostoru (u kompletnom normiranom prostoru). Iz norme vektora $\|x\|$ proističe rastojanje po formuli $d(x, y) = \|x - y\|$.

Hilbertov prostor predstavlja specijalan slučaj Banahovog prostora. U Hilbertovom prostoru definisan je skalarni proizvod dva vektora $\langle x, y \rangle$. Iz skalarnog proizvoda proističe norma vektora po formuli $\|x\| = \sqrt{\langle x, x \rangle}$.

Metoda proste iteracije

Neka su dati matrica $A \in R^{n \times n}$ i vektor $\mathbf{b} \in R^n$. Razmotrimo sistem linearnih jednačina

$$A\mathbf{x} = \mathbf{b}. \tag{1}$$

Sistem (1) transformisati u njemu ekvivalentni sistem oblika

$$\mathbf{x} = B\mathbf{x} + \mathbf{c}, \tag{2}$$

$B \in R^{n \times n}$, $\mathbf{c} \in R^n$; sistemi (1) i (2) su ekvivalentni – oni imaju jedna te ista rješenja $\mathbf{x} \in R^n$.

Uzmimo da sistem ili jednačina (1) odnosno (2) ima jedinstveno rješenje $\mathbf{x}$. Za iterativne metode (za metode uzastopnih ili sukcesivnih aproksimacija) karakteristično je da se konstruiše niz vektora (iterativni niz) $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ čiji članovi pripadaju R^n koji bi trebalo da konvergira ka rješenju, tj. trebalo bi da bude $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$ ili svejedno $\lim_{k\to\infty} \|\mathbf{x}^{(k)} - \mathbf{x}\| = 0$.

Kada se kaže da se rješava jednačina $\mathbf{x} = B\mathbf{x} + \mathbf{c}$ i da se primjenjuje metoda proste iteracije tada se ustvari već podrazumijeva da su članovi iterativnog niza definisani relacijom

$$\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c} \qquad \text{za } k = 0, 1, 2, \ldots. \tag{3}$$

Niz će biti definisan kada se još odabere i početna aproksimacija $\mathbf{x}^{(0)} \in R^n$.

Fiksirajmo u R^n jednu normu $\| \ \|$ i sa $\|B\|$ označimo naravno saglasnu normu matrice B.

Teorema o dovoljnim uslovima za konvergenciju metode proste iteracije. Ako je $\|B\| < 1$ onda: (a) sistem (2) ima jedinstveno rješenje $\mathbf{x}$ i (b) iterativni niz (3) konvergira ka $\mathbf{x}$ za bilo koju početnu aproksimaciju $\mathbf{x}^{(0)} \in R^n$.

Dokaz teoreme. Po uslovu $\|B\| < 1$ i po Banahovoj lemi iz prethodnog naslova slijedi da je matrica $I - B$ invertibilna. Znači da sistem $(I - B)\mathbf{x} = \mathbf{c}$ ima jedinstveno rješenje, čime je

(a) dokazano. Uvedimo oznaku $\mathbf{r}^{(k)}$ za grešku k–te aproksimacije, tj. stavimo $\mathbf{r}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$. Imamo redom:

$$\mathbf{x} = B\mathbf{x} + \mathbf{c} \quad \text{i} \quad \mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c} \quad \Rightarrow \quad \mathbf{x} - \mathbf{x}^{(k+1)} = B\mathbf{x} - B\mathbf{x}^{(k)}, \quad \mathbf{r}^{(k+1)} = B\mathbf{r}^{(k)},$$

$$\mathbf{r}^{(1)} = B\mathbf{r}^{(0)}, \qquad \mathbf{r}^{(2)} = B\mathbf{r}^{(1)}, \qquad \ldots \qquad \Rightarrow \qquad \mathbf{r}^{(k)} = B^k\mathbf{r}^{(0)},$$

$$\|\mathbf{r}^{(k)}\| = \|B^k\mathbf{r}^{(0)}\| \leq \|B^k\| \cdot \|\mathbf{r}^{(0)}\| \leq \|B\|^k\|\mathbf{r}^{(0)}\|, \qquad \lim_{k\to\infty} \mathbf{r}^{(k)} = 0.$$

Dakle, dobili smo da je $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$, čime je dokazano (b). Dokaz je završen.

Naredna teorema preuzima oznake i pretpostavke od prethodne teoreme.

Teorema. Stavimo $q = \|B\|$. Važe nejednakosti

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \frac{q^k}{1-q}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|, \tag{4}$$

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \frac{q}{1-q}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|. \tag{5}$$

Dokaz.

$$\text{Mi računamo:} \qquad \mathbf{x} - \mathbf{x}^{(0)} = \mathbf{x} - \mathbf{x}^{(1)} + \mathbf{x}^{(1)} - \mathbf{x}^{(0)},$$

$$\|\mathbf{x} - \mathbf{x}^{(0)}\| \leq \|\mathbf{x} - \mathbf{x}^{(1)}\| + \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| \leq q\|\mathbf{x} - \mathbf{x}^{(0)}\| + \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|,$$

$$(1-q)\|\mathbf{x} - \mathbf{x}^{(0)}\| \leq \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|, \qquad \|\mathbf{x} - \mathbf{x}^{(0)}\| \leq \frac{1}{1-q}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|.$$

a ranije smo pokazali da je $\|\mathbf{r}^{(k)}\| \leq q^k\|\mathbf{r}^{(0)}\|$, tj. $\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq q^k\|\mathbf{x} - \mathbf{x}^{(0)}\|$.

Stavimo $\varphi(\mathbf{x}) = B\mathbf{x} + \mathbf{c}$. Mi računamo:

$$\mathbf{x} - \mathbf{x}^{(k)} = \mathbf{x} - \mathbf{x}^{(k+1)} + \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}, \qquad \text{po aksiomi trougla} \Rightarrow$$

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \|\mathbf{x} - \mathbf{x}^{(k+1)}\| + \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|,$$

imamo da je $\varphi(\mathbf{x}) = \mathbf{x}$, $\varphi(\mathbf{x}^{(k-1)}) = \mathbf{x}^{(k)}$ i $\varphi(\mathbf{x}^{(k)}) = \mathbf{x}^{(k+1)}$,

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq q\|\mathbf{x} - \mathbf{x}^{(k)}\| + q\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|, \qquad (1-q)\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq q\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.$$

Dokaz je završen.

Formule (4) i (5) služe za ocjenu greške k–te aproksimacije $\mathbf{x}^{(k)}$ za bilo koje $k \geq 1$. Iz koraka u korak, greška se množi sa q. Tako da je tempo ili brzima konvergencije prvog reda ili linearna.

(Nismo morali da se zamaramo oko dokaza teorema, budući da se radi o specijalnom slučaju principa kontrakcije kada je $X = R^n$, $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, $\varphi(\mathbf{x}) = B\mathbf{x} + \mathbf{c}$. Iz $\|B\| = q < 1$ slijedi $\|\varphi(\mathbf{y}) - \varphi(\mathbf{x})\| = \|B\mathbf{y} + \mathbf{c} - B\mathbf{x} - \mathbf{c}\| = \|B(\mathbf{y} - \mathbf{x})\| \leq \|B\| \cdot \|\mathbf{y} - \mathbf{x})\| = q\|\mathbf{y} - \mathbf{x}\|$, jeste komtrakcija.)

Analiza teoreme. Može se desiti da je po nekoj normi $\|B\| < 1$, a po nekoj drugoj normi da nije $\|B\| < 1$. Dovoljno je da po jednoj normi bude $\|B\| < 1$ da bi iterativni niz konvergirao ka rješenju, u vezi međusobne ekvivalentnosti svih vektorskih normi nad R^n o čemu je bilo riječi na početku. Nema protivrječnosti u tome što dovoljan uslov $\|B\| < 1$ može da bude ispunjen u jednoj normi, dok nije ispunjen u nekoj drugoj normi.

Kako da se datom sistemu (1) pridruži ekvivalentan sistem oblika (2)? Za ovo postoji više načina, postoji beskonačno mnogo načina. Navedimo jedan način. Neka je $D \in R^{n\times n}$ bilo koja regularna matrica. Jednačina $A\mathbf{x} - \mathbf{b} = 0$ očito je ekvivalentna jednačini $\mathbf{x} = \mathbf{x} + D(A\mathbf{x} - \mathbf{b})$. Odatle:

$$\mathbf{x} = (I + DA)\mathbf{x} - D\mathbf{b} = B\mathbf{x} + \mathbf{c}; \quad B = I + DA, \quad \mathbf{c} = -D\mathbf{b}.$$

Jakobijeva metoda (specijalan slučaj metode proste iteracije)

Primjer: Jakobijeva metoda. Neka polazni sistem (1) zadovoljava $a_{ii} \neq 0$ za $1 \leq i \leq n$. Tada se i–ta jednačina može podijeliti sa a_{ii} i onda iz i–te jednačne izraziti x_i. Vidimo da smo (1) sveli na oblik $\mathbf{x} = B\mathbf{x} + \mathbf{c}$, tj. da smo definisali iterativni niz $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ čiji članovi pripadaju R^n; izabrati $\mathbf{x}^{(0)} \in R^n$. Realizacija:

$$A\mathbf{x} = \mathbf{b} \quad \text{ili} \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \quad \ldots \quad \ldots \end{cases}$$

$$x_i = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}} x_j + \frac{b_i}{a_{ii}} \quad \text{za } 1 \leq i \leq n \quad \text{ili} \quad \mathbf{x} = B\mathbf{x} + \mathbf{c} \quad \text{ili}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & -\dfrac{a_{12}}{a_{11}} & -\dfrac{a_{13}}{a_{11}} & \ldots & -\dfrac{a_{1n}}{a_{11}} \\ -\dfrac{a_{21}}{a_{22}} & 0 & -\dfrac{a_{23}}{a_{22}} & \ldots & -\dfrac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix}}_{=B} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} + \underbrace{\begin{bmatrix} \dfrac{b_1}{a_{11}} \\ \dfrac{b_2}{a_{22}} \\ \vdots \end{bmatrix}}_{=c}$$

$$\Rightarrow \quad \mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c} \quad \text{ili} \quad \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ \vdots \end{bmatrix} = B \cdot \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix} + \mathbf{c} \quad \text{za } k \geq 0. \tag{6}$$

Da li Jakobijev iterativni proces (6) konvergira? Pokušajmo da primijenimo teoremu od maločas. Da li je $\|B\| < 1$? Izaberimo $\|\ \| = \|\ \|_\infty$. Imamo redom:

$$\|B\|_\infty = \max_{1\leq i\leq n} \left\{ \sum_{j=1}^{n} |b_{ij}| \right\} =$$

$$\max\left\{ \left|-\frac{a_{12}}{a_{11}}\right| + \left|-\frac{a_{13}}{a_{11}}\right| + \ldots + \left|-\frac{a_{1n}}{a_{11}}\right|, \left|-\frac{a_{21}}{a_{22}}\right| + \left|-\frac{a_{23}}{a_{22}}\right| + \ldots + \left|-\frac{a_{2n}}{a_{22}}\right|, \ldots \right\} =$$

$$\max\left\{\frac{|a_{12}|+|a_{13}|+\ldots+|a_{1n}|}{|a_{11}|}, \frac{|a_{21}|+|a_{23}|+\ldots+|a_{2n}|}{|a_{22}|}, \ldots\right\}.$$

$\|B\| < 1$? Dovoljno je da matrica A bude dijagonalno dominantna. Mi smo dokazali tvrđenje: ako je matrica A dijagonalno dominantna ($\sum_{j=1\ j\neq i}^{n} |a_{ij}| < |a_{ii}|$ za $1 \leq i \leq n$) onda Jakobijev iterativni proces konvergira. Završen primjer.

Mali konkretni primjer za Jakobijevu metodu kada je $n = 3$: dati su

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ i } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \Rightarrow \quad \mathbf{Ax} = \mathbf{b} \text{ ili } \begin{bmatrix} 10 & 2 & 7 \\ 3 & 10 & 3 \\ -1 & 1 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 12 \\ 14 \\ 16 \end{bmatrix}$$

$$\Rightarrow \quad \text{oblik } \mathbf{x} = \mathbf{Bx} + \mathbf{c} \text{ ili } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -0{,}2 & -0{,}7 \\ -0{,}3 & 0 & -0{,}3 \\ 0{,}2 & -0{,}2 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1{,}2 \\ 1{,}4 \\ 3{,}2 \end{bmatrix} \quad \Rightarrow$$

iteracije $\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}$, gdje je $\mathbf{x}^{(k)} = [x_1^{(k)}\ \ x_2^{(k)}\ \ x_3^{(k)}]^T$, $\mathbf{x}^{(k+1)} = [x_1^{(k+1)}\ \ x_2^{(k+1)}\ \ x_3^{(k+1)}]^T$.

Neophodni i dovoljni uslovi za konvergenciju

U nastavku želimo da nađemo tačne, tj. neophodne i dovoljne uslove da iterativni niz (3) konvergira bez obzira na izbor $\mathbf{x}^{(0)} \in R^n$, tj. za ma kakvu početnu aproksimaciju.

Iz linearne algebre je poznato sljedeće: za svaku matricu $P \in R^{n\times n}$ postoji matrica $Q \in R^{n\times n}$ pomoću koje se P prevodi u njenu Žordanovu kanonsku formu $Q^{-1}PQ$ u smislu da važi jednakost

$$Q^{-1}PQ = \begin{bmatrix} \lambda_1(P) & \alpha_1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2(P) & \alpha_2 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \lambda_{n-1}(P) & \alpha_{n-1} \\ 0 & 0 & 0 & 0 & \ldots & 0 & \lambda_n(P) \end{bmatrix}$$

gdje su $\lambda_i(P) \in C$ svojstvene vrijednosti matrice P, a $\alpha_i \in \{0, 1\}$; za matrice P i $Q^{-1}PQ$ kaže se da su slične.

Lema. Neka za matricu $B \in R^{n\times n}$ važi $\max_{1\leq i\leq n} |\lambda_i(B)| < q$. Tada postoji matrica $D \in R^{n\times n}$ takva da je $\|D^{-1}BD\|_\infty \leq q$.

Dokaz leme. Stavimo $\eta = q - \max_{1\leq i\leq n} |\lambda_i(B)| > 0$. Želimo da odredimo Žordanovu kanonsku formu matrice $\eta^{-1}B$. Dakle, postoji matrica D takva da važi:

$$D^{-1}(\eta^{-1}B)D = \begin{bmatrix} \lambda_1(\eta^{-1}B) & \alpha_1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2(\eta^{-1}B) & \alpha_2 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \lambda_{n-1}(\eta^{-1}B) & \alpha_{n-1} \\ 0 & 0 & 0 & 0 & \ldots & 0 & \lambda_n(\eta^{-1}B) \end{bmatrix} =$$

$$\left[\begin{array}{ccccccc} \eta^{-1}\lambda_1(B) & \alpha_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \eta^{-1}\lambda_2(B) & \alpha_2 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \eta^{-1}\lambda_{n-1}(B) & \alpha_{n-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & \eta^{-1}\lambda_n(B) \end{array}\right] \Big/ \cdot \eta$$

$$D^{-1}BD = \left[\begin{array}{ccccccc} \lambda_1(B) & \eta\alpha_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2(B) & \eta\alpha_2 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda_{n-1}(B) & \eta\alpha_{n-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_n(B) \end{array}\right]$$

Rekli smo $\alpha_i = 0$ ili $\alpha_i = 1$.

Znamo sljedeće: ako je λ svojstvena vrijednost matrice A onda je $c\lambda$ svojstvena vrijednost matrice cA, što je maločas upotrebljeno. Zaista, $A\mathbf{x} = \lambda\mathbf{x} \Rightarrow (cA)\mathbf{x} = (c\lambda)\mathbf{x}$.

Dalje:

$$\|D^{-1}BD\|_\infty = \max\Big\{|\lambda_1(B)| + \eta\alpha_1, |\lambda_2(B)| + \eta\alpha_2, \dots, |\lambda_{n-1}(B)| + \eta\alpha_{n-1}, |\lambda_n(B)|\Big\} \leq$$

$$\max\Big\{|\lambda_1(B)| + \eta, |\lambda_2(B)| + \eta, \dots, |\lambda_{n-1}(B)| + \eta, |\lambda_n(B)|\Big\} \leq \max_{1\leq i\leq n} |\lambda_i(B)| + \eta = q.$$

Lema je dokazana.

Teorema o neophodnim i dovoljnim uslovima za konvergenciju metode proste iteracije. Neka sistem (2) ima jedinstveno rješenje $\mathbf{x}$. Iterativni proces (3) konvergira ka $\mathbf{x}$ (za ma kakvu početnu aproksimaciju $\mathbf{x}^{(0)}$) ako i samo ako su sve svojstvene vrijednosti matrice B po apsolutnoj vrijednosti < 1.

Dokaz teoreme. Uslov je dovoljan. Dato je $|\lambda_i(B)| < 1$. Uzmimo proizvoljno q u granicama $\max_{1\leq i\leq n} |\lambda_i(B)| < q < 1$. Kako su uslovi prethodne leme ispunjeni za matricu B i broj q to postoji matrica D takva da važi $\|D^{-1}BD\|_\infty \leq q$ ili svejedno $\|\Lambda\|_\infty \leq q$, gdje je uvedena oznaka $\Lambda = D^{-1}BD$. Imamo redom:

$$\Lambda = D^{-1}BD \quad \Rightarrow \quad B = D\Lambda D^{-1},$$

$$B^2 = BB = D\Lambda D^{-1}D\Lambda D^{-1} = D\Lambda^2 D^{-1},$$

$$B^3 = D\Lambda^3 D^{-1}, \quad \dots, \quad B^k = D\Lambda^k D^{-1}, \quad \dots,$$

$$\|B^k\|_\infty = \|D\Lambda^k D^{-1}\|_\infty \leq \|D\|_\infty \|\Lambda^k\|_\infty \|D^{-1}\|_\infty \leq \|D\|_\infty \|\Lambda\|_\infty^k \|D^{-1}\|_\infty \leq \|D\|_\infty q^k \|D^{-1}\|_\infty.$$

Odaberimo $\mathbf{x}^{(0)} \in R^n$, čime je niz $\mathbf{x}^{(k)} = B\mathbf{x}^{(k-1)} + \mathbf{c}$ definisan. U prethodnoj teoremi je pokazano da važi $\mathbf{r}^{(k)} = B^k\mathbf{r}^{(0)}$, gdje je uvedena oznaka $\mathbf{r}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$. Slijedi

$$\|\mathbf{r}^{(k)}\|_\infty = \|B^k\mathbf{r}^{(0)}\|_\infty \leq \|B^k\|_\infty \|\mathbf{r}^{(0)}\|_\infty \leq \|D\|_\infty q^k \|D^{-1}\|_\infty \|\mathbf{r}^{(0)}\|_\infty,$$

$$\lim_{k\to\infty} \|\mathbf{r}^{(k)}\|_\infty = 0.$$

Dokazali smo da je $\lim_{k\to\infty} \|\mathbf{x} - \mathbf{x}^{(k)}\|_\infty = 0$, tj. da je $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$ po normi $\| \quad \|_\infty$, $(\forall \mathbf{x}^{(0)})$. Ustvari smo dokazali da je $\lim_{k\to\infty} \|\mathbf{x} - \mathbf{x}^{(k)}\| = 0$, gdje je $\| \quad \|$ bilo koja norma, tj. da je $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$ po bilo kojoj normi, $(\forall \mathbf{x}^{(0)})$, u vezi ekvivalentnosti svih normi u konačno–dimenzionom prostoru.

Uslov je neophodan. Dopustimo da matrica B ima svojstvenu vrijednost $\lambda \in C$ takvu da je $|\lambda| \geq 1$. Označimo sa $\mathbf{v} \in C^n$ odgovarajući svojstveni vektor: $B\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq 0$. Treba pokazati da se može naći bar jedna početna aproksimacija $\mathbf{x}^{(0)}$ takva da odgovarajući iterativni niz $\mathbf{x}^{(k)} = B\mathbf{x}^{(k-1)} + \mathbf{c}$ ne konvergira ka $\mathbf{x}$ kad $k \to \infty$. Odaberimo $\mathbf{x}^{(0)} = \mathbf{x} - \mathbf{v}$. Imamo

$$\mathbf{x}^{(1)} = B\mathbf{x}^{(0)} + \mathbf{c} = B(\mathbf{x} - \mathbf{v}) + \mathbf{c} = B\mathbf{x} - B\mathbf{v} + \mathbf{c} = \mathbf{x} - B\mathbf{v} = \mathbf{x} - \lambda\mathbf{v},$$

$$\mathbf{x}^{(2)} = B\mathbf{x}^{(1)} + \mathbf{c} = B(\mathbf{x} - \lambda\mathbf{v}) + \mathbf{c} = B\mathbf{x} - \lambda B\mathbf{v} + \mathbf{c} = \mathbf{x} - \lambda B\mathbf{v} = \mathbf{x} - \lambda^2\mathbf{v}, \quad \ldots,$$

$$\mathbf{x}^{(k)} = \mathbf{x} - \lambda^k\mathbf{v}, \quad \ldots,$$

$$\mathbf{r}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)} = \mathbf{x} - \mathbf{x} + \lambda^k\mathbf{v} = \lambda^k\mathbf{v},$$

$$\|\mathbf{r}^{(k)}\| = \|\lambda^k\mathbf{v}\| = |\lambda|^k\|\mathbf{v}\| \geq \|\mathbf{v}\| > 0, \qquad \text{nije} \lim_{k\to\infty} \|\mathbf{r}^{(k)}\| = 0,$$

$$\text{nije} \lim_{k\to\infty} \|\mathbf{x} - \mathbf{x}^{(k)}\| = 0 \qquad \Big(\text{nije} \lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}\Big).$$

Teorema je dokazana.

3.5. ZAJDELOVA METODA

Biće izložena Zajdelova metoda (Gaus–Zajdelova metoda).
Razmatra se sistem oblika $A\mathbf{x} = \mathbf{b}$, tj.

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \qquad \text{ili} \qquad \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \cdots \quad \cdots \quad \cdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n \end{cases}$$

U slučaju Jakobijeve metode:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i \right) \quad 1 \leq i \leq n, \qquad k \geq 0. \tag{1}$$

A u slučaju Zajdelove metode:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i \right), \qquad 1 \leq i \leq n, \qquad k \geq 0. \tag{2}$$

Zajdelova metoda je primjer iterativne metode za rješavanje sistema linearnih jednačina $A\mathbf{x} = \mathbf{b}$; $A = [a_{ij}]_{i,j=1}^n \in R^{n\times n}$, $\mathbf{b} = (b_1, \ldots, b_n) \in R^n$, $n \geq 1$. Zajdelova metoda predstavlja malu modifikaciju ili malo poboljšanje Jakobijeve metode. Za jednu i drugu metodu pretpostavlja se da je $a_{ii} \neq 0$ za $1 \leq i \leq n$.

Znamo da "matrica A je dijagonalno dominantna" predstavlja dovoljan uslov za konvergenciju Jakobijeve metode, a vidjećemo da je to i dovoljan uslov za konvergenciju Zajdelove metode. Ako je matrica $A \in R^{n\times n}$ dijagonalno dominantna onda je $\det A \neq 0$; ovo je dokazano ranije, kod kubnog splajna (Interpolacija pomoću splajna).

Neka je matrica A sistema linearnih jednačina $A\mathbf{x} = \mathbf{b}$ koji se rješava dijagonalno dominantna:

$$\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| < |a_{ii}| \qquad \text{za } 1 \leq i \leq n. \tag{3}$$

Tada sistem $A\mathbf{x} = \mathbf{b}$ ima jedinstveno rješenje $\mathbf{x} = (x_1, \ldots, x_n) \in R^n$. Za Jakobijevu metodu važi formula (1). Kod Zajdelove metode računanja se sprovode po formuli (2). Iterativni niz $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$, gdje $\mathbf{x}^{(k)} \in R^n$, biće definisan kada se izabere $\mathbf{x}^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in R^n$. Može se uzeti $\mathbf{x}^{(0)} = (0, \ldots, 0)$.

Dakle, u formuli (2), neka je u toku računanje $(k+1)$–ve iteracije $\mathbf{x}^{(k+1)} = (x_1^{(k+1)}, \ldots, x_n^{(k+1)})$ i neka se konkretno u datom trenutku računa komponenta $x_i^{(k+1)}$. Već su poznate komponente $x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}$ vektora $\mathbf{x}^{(k+1)}$. Upotrebiti ih u datom trenutku, umjesto $x_1^{(k)}, \ldots, x_{i-1}^{(k)}$. Iskustvo pokazuje da Zajdelova metoda daje bolje rezultate od Jakobijeve metode (greška je manja, konvergencija je brža).

Vidimo da se Zajdelova metoda lakše programira od Jakobijeve metode: čim je izračunato $x_i^{(k+1)}$ više nam ne treba $x_i^{(k)}$. Zato se u praksi koristi Zajdelova metoda, a ne Jakobijeva metoda.

Na osnovu (2) je

$$\begin{cases} a_{11}x_1^{(k+1)} + a_{12}x_2^{(k)} + a_{13}x_3^{(k)} + \ldots + a_{1n}x_n^{(k)} = b_1 \\ a_{21}x_1^{(k+1)} + a_{22}x_2^{(k+1)} + a_{23}x_3^{(k)} + \ldots + a_{2n}x_n^{(k)} = b_2 \\ \ldots \quad \ldots \quad \ldots \\ a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} + a_{n3}x_3^{(k+1)} + \ldots + a_{nn}x_n^{(k+1)} = b_n \end{cases}$$

ili $B\mathbf{x}^{(k+1)} + C\mathbf{x}^{(k)} = \mathbf{b}$ za $k \geq 0$, gdje su uvedene oznake $A = B + C$ i

$$B = \begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{bmatrix} \quad \text{i} \quad C = \begin{bmatrix} 0 & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & 0 & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Uvedimo oznaku $\mathbf{r}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$ za grešku k–te aproksimacije $\mathbf{x}^{(k)}$.

Teorema. Pretpostavimo da je

$$\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| \leq q|a_{ii}| \qquad \text{za } 1 \leq i \leq n, \tag{4}$$

gdje je $q < 1$. Tada važi nejednakost

$$\|\mathbf{r}^{(k+1)}\|_\infty \leq q\|\mathbf{r}^{(k)}\|_\infty \qquad \text{za } k \geq 0. \tag{5}$$

Dokaz teoreme. Imamo $A\mathbf{x} = \mathbf{b}$ ili svejedno $B\mathbf{x} + C\mathbf{x} = \mathbf{b}$ i imamo $B\mathbf{x}^{(k+1)} + C\mathbf{x}^{(k)} = \mathbf{b}$. Oduzimanjem: $B\mathbf{x} - B\mathbf{x}^{(k+1)} + C\mathbf{x} - C\mathbf{x}^{(k)} = \mathbf{b} - \mathbf{b}$ ili

$$B\mathbf{r}^{(k+1)} + C\mathbf{r}^{(k)} = 0, \tag{6}$$

čime smo dobili sistem koji se odnosi na grešku $\mathbf{r}^{(k)}$. Posmatrajmo brojeve $|r_1^{(k+1)}|, \ldots, |r_n^{(k+1)}|$ i uočimo najveći među njima. Neka je najveći $|r_l^{(k+1)}|$, tako da je

$$\|\mathbf{r}^{(k+1)}\|_\infty = \|(r_1^{(k+1)}, \ldots, r_n^{(k+1)})\|_\infty = \max_{1\le i\le n} |r_i^{(k+1)}| = |r_l^{(k+1)}|.$$

Napišimo l–tu jednačinu sistema (6):

$$\sum_{j=1}^{l-1} a_{lj} r_j^{(k+1)} + a_{ll} r_l^{(k+1)} + \sum_{j=l+1}^{n} a_{lj} r_j^{(k)} = 0 \qquad \Big/ : a_{ll}$$

$$r_l^{(k+1)} = -\sum_{j=1}^{l-1} \frac{a_{lj}}{a_{ll}} r_j^{(k+1)} - \sum_{j=l+1}^{n} \frac{a_{lj}}{a_{ll}} r_j^{(k)},$$

$$\left|r_l^{(k+1)}\right| \le \sum_{j=1}^{l-1} \left|\frac{a_{lj}}{a_{ll}}\right| \left|r_j^{(k+1)}\right| + \sum_{j=l+1}^{n} \left|\frac{a_{lj}}{a_{ll}}\right| \left|r_j^{(k)}\right|,$$

$$\left|r_l^{(k+1)}\right| \le \sum_{j=1}^{l-1} \left|\frac{a_{lj}}{a_{ll}}\right| \max_{1\le i\le n} |r_i^{(k+1)}| + \sum_{j=l+1}^{n} \left|\frac{a_{lj}}{a_{ll}}\right| \max_{1\le i\le n} |r_i^{(k)}|,$$

$$|r_l^{(k+1)}| \le \alpha \|\mathbf{r}^{(k+1)}\|_\infty + \beta \|\mathbf{r}^{(k)}\|_\infty,$$

gdje su uvedene oznake $\alpha = \sum_{j=1}^{l-1} \left|\frac{a_{lj}}{a_{ll}}\right|$ i $\beta = \sum_{j=l+1}^{n} \left|\frac{a_{lj}}{a_{ll}}\right|$,

$$\|\mathbf{r}^{(k+1)}\|_\infty \le \alpha \|\mathbf{r}^{(k+1)}\|_\infty + \beta \|\mathbf{r}^{(k)}\|_\infty \qquad \Big/ : (1-\alpha)$$

$$\|\mathbf{r}^{(k+1)}\|_\infty \le \frac{\beta}{1-\alpha} \|\mathbf{r}^{(k)}\|_\infty.$$

Po uslovu teoreme je $\alpha + \beta \le q$ (u (4) stavite $i = l$) i $q < 1 \Rightarrow \dfrac{\beta}{1-\alpha} \le q \Rightarrow$

$$\|\mathbf{r}^{(k+1)}\|_\infty \le q \|\mathbf{r}^{(k)}\|_\infty.$$

Teorema je dokazana.

Uzastopnom primjenom nejednakosti (5) nalazimo da je $\|\mathbf{r}^{(k)}\|_\infty \le q^k \|\mathbf{r}^{(0)}\|_\infty$, tako da je $\lim_{k\to\infty} \|\mathbf{r}^{(k)}\|_\infty = 0$. Pokazali smo da je $\lim_{k\to\infty} \mathbf{r}^{(k)} = 0$, odnosno pokazali smo da je $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$. Dakle, ako je ispunjen uslov (4) onda Zajdelov iterativni proces konvergira ka rješenju.

Vidi se da je (3) $\Leftrightarrow$ (4), samo da bi bili $a_{ii} \ne 0$.

3.6. PRIMJER ITERATIVNE METODE (ZA RJEŠAVANJE SISTEMA LINEARNIH JEDNAČINA) VARIJACIONOG TIPA

Oznake za vektore: $\mathbf{x}^{(k)} = \left(x_1^{(k)}, \ldots, x_n^{(k)}\right)$ ili $\mathbf{x}_k = \left(x_1^{(k)}, \ldots, x_n^{(k)}\right)$.

Za numeričku metodu se kaže da je varijaciona ako ona uključuje minimizaciju nekog funkcionala. Za preslikavanje φ se kaže da je funkcional ako su njegove vrijednosti realni brojevi. Postoje razne varijacione metode za rješavanje sistema linearnih jednačina. U ovom naslovu govori se o jednoj takvoj iterativnoj metodi.

U ovom naslovu u prostoru R^n koristi se norma $\|\ \ \|_2$; $\|(x_1, \ldots, x_n)\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$, euklidska norma, biće označavana prosto kao $\|\ \ \|$. Znamo da je ta norma povezana sa skalarnim proizvodom $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$ relacijom $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Iz linearne algebre je poznato sljedeće. Neka je matrica A simetrična ($A = A^T$). Tada A ima n realnih svojstvenih vrijednosti, tj. sve njene svojstvene vrijednosti su realni brojevi; višestrukost svojstvenih vrijednosti je naravno uzeta u obzir. Uredimo ih po veličini: $\lambda_{\min}(A) = \lambda_1 \leq \ldots \leq \lambda_n = \lambda_{\max}(A)$. Neka je dodatno matrica A i pozitivno definitna, u oznaci $A > 0$, što znači da $\mathbf{x} \neq 0 \Rightarrow \langle A\mathbf{x}, \mathbf{x} \rangle > 0$. Tada su sve njene svojstvene vrijednosti pozitivni brojevi $(0 < \lambda_1)$ i njena norma $\|\ \ \|_2$ jednaka je najvećoj njenoj svojstvenoj vrijednosti; $\|A\| = \|A\|_2 = \lambda_n$. Još je i $\det A \neq 0$.

Prelazimo na izlaganje metode minimalne nepovezanosti.

Ako treba riješiti $f(x) = b$ i ako odgovor glasi $x \approx x^*$ onda se kaže da $x - x^*$ predstavlja grešku i još se kaže da $f(x^*) - b$ predstavlja nepovezanost ili ostatak.

Neka $A \in R^{n \times n}$ i $\mathbf{b} \in R^n$ i razmotrimo sistem $A\mathbf{x} = \mathbf{b}$. Sistem $A\mathbf{x} = \mathbf{b}$ ekvivalentan je sistemu $\mathbf{x} = \mathbf{x} + \tau(A\mathbf{x} - \mathbf{b})$ za $\tau \neq 0$. Znači da uzastopne aproksimacije mogu da se računaju po formuli $\mathbf{x}_{k+1} = \mathbf{x}_k + \tau(A\mathbf{x}_k - \mathbf{b})$, $k \geq 0$. Mi možemo da τ podešavamo od koraka do koraka. Mi ćemo uzastopne aproksimacije računati po formuli

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_{k+1}(A\mathbf{x}_k - \mathbf{b}), \qquad k \geq 0. \tag{1}$$

Niz $\{\mathbf{x}_k\}_{k=0}^{\infty}$ je definisan ako se izabere $\mathbf{x}_0 \in R^n$.

(Poznato je tvrđenje: minimum funkcionala $F(\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{b}, \mathbf{x} \rangle$ realizuje se na vektoru $\mathbf{x}$ koji zadovoljava $A\mathbf{x} = \mathbf{b}$. Može se izračunati da gradijent funkcionala iznosi $\nabla F(\mathbf{x}) = 2A\mathbf{x} - 2\mathbf{b}$. Zato je najrentabilnije da se od tačke $\mathbf{x} = \mathbf{x}_k$ krećemo u smjeru $-(A\mathbf{x} - \mathbf{b})$, ako želimo da se smanji ostatak $\mathbf{r}_k = A\mathbf{x}_k - \mathbf{b}$. Zato se razmatrana numerička metoda naziva metodom gradijentnog spusta ili metodom najbržeg spusta.)

Neka $\mathbf{x}$ označava rješenje sistema $A\mathbf{x} = \mathbf{b}$. Možemo sa $\mathbf{d}_k = \mathbf{x} - \mathbf{x}_k$ da označimo grešku k–te aproksimacije $\mathbf{x}_k$. Bolje je što je greška manja. Ako se u izrazu $A\mathbf{x} - \mathbf{b}$ uvrsti $\mathbf{x} = \mathbf{x}_k$, da li će se dobiti nula, šta će se dobiti, bolje je da se dobije što manje. Uvedimo oznaku

$$\mathbf{r}_k = A\mathbf{x}_k - \mathbf{b}, \qquad k \geq 0. \tag{2}$$

Za $\mathbf{r}_k \in R^n$ kaže se da je nepovezanost (lijeve i desne strane sistema). Bolje je što je norma nepovezanosti manja. Niz $\{\mathbf{x}_k\}_{k=0}^{\infty}$, relacija (1), biće definisan ustvari tek kada se odrede i brojevi $\{\tau_{k+1}\}_{k=0}^{\infty}$.

Uvedimo oznaku

$$\varphi(\tau_{k+1}) = \|\mathbf{r}_{k+1}\| = \|A\mathbf{x}_{k+1} - \mathbf{b}\| \geq 0.$$

Želimo da $\varphi(\tau_{k+1})$ ima što je moguće manju vrijednost. To je kriterijum za izbor veličine τ_{k+1}.

Prelazimo na dobijanje eksplicitnog izraza za τ_{k+1}. Za račun je pogodnije da se gleda najmanja moguća vrijednost kvadrata $\varphi^2(\tau_{k+1}) = \|\mathbf{r}_{k+1}\|^2 = \langle \mathbf{r}_{k+1}, \mathbf{r}_{k+1}\rangle$:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_{k+1}(A\mathbf{x}_k - \mathbf{b}) = \mathbf{x}_k + \tau_{k+1}\mathbf{r}_k \qquad \Big/ \cdot A$$

$$A\mathbf{x}_{k+1} = A\mathbf{x}_k + \tau_{k+1}A\mathbf{r}_k \qquad \Big/ - \mathbf{b}$$

$$A\mathbf{x}_{k+1} - \mathbf{b} = A\mathbf{x}_k - \mathbf{b} + \tau_{k+1}A\mathbf{r}_k, \quad \mathbf{r}_{k+1} = \mathbf{r}_k + \tau_{k+1}A\mathbf{r}_k. \tag{$*$}$$

Zatim $\varphi^2(\tau_{k+1}) = \langle \mathbf{r}_{k+1}, \mathbf{r}_{k+1}\rangle = \langle \mathbf{r}_k + \tau_{k+1}A\mathbf{r}_k, \mathbf{r}_k + \tau_{k+1}A\mathbf{r}_k\rangle =$

$$\langle \mathbf{r}_k, \mathbf{r}_k\rangle + \tau_{k+1}\langle \mathbf{r}_k, A\mathbf{r}_k\rangle + \tau_{k+1}\langle A\mathbf{r}_k, \mathbf{r}_k\rangle + \tau_{k+1}^2\langle A\mathbf{r}_k, A\mathbf{r}_k\rangle =$$

$$\langle \mathbf{r}_k, \mathbf{r}_k\rangle + 2\tau_{k+1}\langle \mathbf{r}_k, A\mathbf{r}_k\rangle + \tau_{k+1}^2\langle A\mathbf{r}_k, A\mathbf{r}_k\rangle$$

jer je $A = A^T$. Parabola $y = ax^2 + bx + c$ ima najmanju moguću vrijednost kada je $x = \dfrac{-b}{2a}$, $a > 0$. Prema tome

$$\tau_{k+1} = \frac{-\langle \mathbf{r}_k, A\mathbf{r}_k\rangle}{\langle A\mathbf{r}_k, A\mathbf{r}_k\rangle} = \frac{-\langle \mathbf{r}_k, A\mathbf{r}_k\rangle}{\|A\mathbf{r}_k\|^2}, \qquad k \geq 0. \tag{3}$$

Riješen je zadatak o minimumu funkcionala. Uzastopne aproksimacije su sada definisane. Redosljed računanja je: $\mathbf{x}_0$, $\mathbf{r}_0$, τ_1, $\mathbf{x}_1$, itd.

Teorema o dovoljnim uslovima za konvergenciju metode minimalne nepovezanosti. Neka je $A = A^T$ i $A > 0$. Tada sistem $A\mathbf{x} = \mathbf{b}$ ima jedinstveno rješenje $\mathbf{x}$. Neka je $\mathbf{x}_0 \in R^n$ bilo koji vektor. Tada niz $\{\mathbf{x}_k\}_{k=0}^{\infty}$, relacije (1)–(3), konvergira ka $\mathbf{x}$. Pored toga, važi sljedeća nejednakost:

$$\|A(\mathbf{x}_k - \mathbf{x})\| \leq \rho_0^k \|A(\mathbf{x}_0 - \mathbf{x})\| \qquad \text{za } \ k \geq 0. \tag{4}$$

Svojstvene vrijednosti matrice A su $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. Stavljeno je $\rho_0 = \dfrac{\lambda_n - \lambda_1}{\lambda_1 + \lambda_n} \geq 0$.

Dokaz teoreme. Po konstrukciji je $\varphi(\tau_{k+1}) = \min\limits_{\tau \in R} \varphi(\tau)$. Zato je $\varphi(\tau_{k+1}) \leq \varphi(\tau')$ za bilo koje $\tau' \in R$. Stavimo da je $\tau' = \dfrac{-2}{\lambda_1 + \lambda_n}$. Imamo da je $\varphi(\tau_{k+1}) = \|\mathbf{r}_{k+1}\| = \|(I + \tau_{k+1}A)\mathbf{r}_k\|$, v. (*). Imamo da je $\varphi(\tau') = \|(I + \tau' A)\mathbf{r}_k\|$. Dakle

$$\|(I + \tau_{k+1}A)\mathbf{r}_k\| \leq \|(I + \tau' A)\mathbf{r}_k\|. \tag{$**$}$$

Odredićemo $\|I + \tau' A\|$. Matrica A ima svojstvene vrijednosti $0 < \lambda_1 \leq \ldots \leq \lambda_n$. $\tau' A$ ima svojstvene vrijednosti $\dfrac{-2\lambda_n}{\lambda_1 + \lambda_n} \leq \ldots \leq \dfrac{-2\lambda_1}{\lambda_1 + \lambda_n} < 0$. Zapaziti da je i matrica $\tau' A$ simetrična. Zapaziti da je i matrica $I + \tau' A$ simetrična. Znamo da je tada $\|I + \tau' A\| = \max\limits_{1 \leq i \leq n} |\lambda_i(I + \tau' A)|$. Iz linearne algebre znamo sljedeće: svojstvene vrijednosti matrice $A + cI$ su veće od svojstvenih vrijednosti matrice A za c. Matrica $I + \tau' A$ ima svojstvene vrijednosti po veličini od $1 - \dfrac{2\lambda_n}{\lambda_1 + \lambda_n}$ do $1 - \dfrac{2\lambda_1}{\lambda_1 + \lambda_n}$, tj. od $-\dfrac{\lambda_n - \lambda_1}{\lambda_1 + \lambda_n}$ do $\dfrac{\lambda_n - \lambda_1}{\lambda_1 + \lambda_n}$, od $-\rho_0$ do ρ_0. Dakle, $\|I + \tau' A\| = \rho_0 < 1$.

Nastavljamo od (**):

$\|(I+\tau_{k+1}A)\mathbf{r}_k\| \leq \|(I+\tau' A)\mathbf{r}_k\|,$

$\|\mathbf{r}_{k+1}\| \leq \|(I+\tau' A)\mathbf{r}_k\|,$

$\|\mathbf{r}_{k+1}\| \leq \|I+\tau' A\| \cdot \|r_k\|,$

$\|\mathbf{r}_{k+1}\| \leq \rho_0\|\mathbf{r}_k\|,$

uzastopnom primjenom nalazimo $\quad \|\mathbf{r}_k\| \leq \rho_0^k\|\mathbf{r}_0\|,$

$\|A\mathbf{x}_k - \mathbf{b}\| \leq \rho_0^k\|A\mathbf{x}_0 - \mathbf{b}\|,$

$\|A\mathbf{x}_k - A\mathbf{x}\| \leq \rho_0^k\|A\mathbf{x}_0 - A\mathbf{x}\|,$

$\|A(\mathbf{x}_k - \mathbf{x})\| \leq \rho_0^k\|A(\mathbf{x}_0 - \mathbf{x})\|.$

Dokazali smo (4). Iz (4) slijedi $\lim_{k\to\infty}\|\mathbf{x}_k - \mathbf{x}\| = 0$. Teorema je dokazana.

Znamo da je $\lambda_1\|\mathbf{x}_k - \mathbf{x}\| \leq \|A(\mathbf{x}_k - \mathbf{x})\| \leq \lambda_n\|\mathbf{x}_k - \mathbf{x}\|$.

Formula (4) služi za ocjenu greške k–te sukcesivne aproksimacije $\mathbf{x}_k$.

Primjer sistema $A\mathbf{x} = \mathbf{b}$ u kome je $A^T = A > 0$:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T,$$

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \end{bmatrix} \qquad \begin{pmatrix} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{pmatrix},$$

kvadratna forma: $\langle A\mathbf{x}, \mathbf{x}\rangle = 3x_1^2 + 2x_1x_2 + 2x_2^2$.

3.7. METODA SKALARNOG PROIZVODA

Riješiti tzv. potpuni problem svojstvenih vrijednosti za matricu $A \in R^{n\times n}$ znači odrediti sve njene svojstvene vrijednosti i odrediti sve odgovarajuće svojstvene vektore. Riješiti djelimični problem znači odrediti neke svojstvene vrijednosti ili odrediti jednu svojstvenu vrijednost (dominantnu). Za male vrijednosti n, svojstvene vrijednosti λ mogu da budu određene iz uslova $\det(A - \lambda I) = 0$.

Metoda skalarnog proizvoda (metoda stepena, engl. power method) predstavlja jednu numeričku metodu za rješavanje djelimičnog problema svojstvenih vrijednosti. Razmotrimo jednostavni slučaj kada je matrica A simetrična, mada se ta metoda može primijeniti i na nesimetrične matrice.

U R^n, skalarni proizvod dva vektora: $\langle \mathbf{x}, \mathbf{y}\rangle = x_1y_1 + \ldots + x_ny_n$.

Neka je $A \in R^{n\times n}$ simetrična matrica:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \qquad (a_{ji} = a_{ij}).$$

Označimo njene svojstvene vrijednosti kao $\lambda_i \in R$, $1 \leq i \leq n$, svaka svojstvena vrijednost broji se sa svojom višestrukošću. Neka je numeracija izvršena tako da bude

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|.$$

Označimo odgovarajuće svojstvene vektore kao $\mathbf{e}_i$, $1 \leq i \leq n$; važi $A\mathbf{e}_i = \lambda_i \mathbf{e}_i$. Iz linearne algebre znamo da je $\mathbf{e}_i \perp \mathbf{e}_j$, tj. da je $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ za $i \neq j$; $\langle \mathbf{x}, \mathbf{y} \rangle$ označava skalarni proizvod. Izaberimo svojstvene vektore tako da bude $\langle \mathbf{e}_i, \mathbf{e}_i \rangle = 1$, tj. $\|\mathbf{e}_i\| = 1$ za $1 \leq i \leq n$. Sistem vektora $\{\mathbf{e}_i\}_{i=1}^n$ čini ortonormiranu bazu prostora R^n.

Biće konstruisan niz brojeva $\{\mu_k\}_{k=0}^{\infty}$, $\mu_k \in R$, takav da je $\lim_{k\to\infty} \mu_k = \lambda_1$.

Izaberimo na proizvoljan način ne–nula vektor $\mathbf{x}^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in R^n$. Vektor $\mathbf{x}^{(0)}$ razložimo po bazi $(\mathbf{e}_1, \ldots \mathbf{e}_n)$. Imamo da je $\mathbf{x}^{(0)} = \sum_{i=1}^n c_i \mathbf{e}_i$, gdje $c_i \in R$. Znamo da je $c_i = \langle \mathbf{x}^{(0)}, \mathbf{e}_i \rangle$. Stavimo

$$\mathbf{x}^{(1)} = A\mathbf{x}^{(0)}, \quad \mathbf{x}^{(2)} = A\mathbf{x}^{(1)}, \quad \ldots, \quad \mathbf{x}^{(k)} = A\mathbf{x}^{(k-1)}, \quad \ldots.$$

Imamo da je

$$\mathbf{x}^{(1)} = A\mathbf{x}^{(0)} = A\sum_{i=1}^n c_i \mathbf{e}_i = \sum_{i=1}^n c_i A e_i = \sum_{i=1}^n c_i \lambda_i \mathbf{e}_i.$$

Slično, imamo da je

$$\mathbf{x}^{(2)} = \sum_{i=1}^n c_i \lambda_i^2 \mathbf{e}_i, \quad \ldots, \quad \mathbf{x}^{(k)} = \sum_{i=1}^n c_i \lambda_i^k \mathbf{e}_i, \quad \ldots.$$

Izračunajmo skalarne proizvode $\langle \mathbf{x}^{(k)}, \mathbf{x}^{(k)} \rangle$ i $\langle \mathbf{x}^{(k+1)}, \mathbf{x}^{(k)} \rangle$:

$$\langle \mathbf{x}^{(k)}, \mathbf{x}^{(k)} \rangle = \Big\langle \sum_{i=1}^n c_i \lambda_i^k \mathbf{e}_i, \sum_{j=1}^n c_j \lambda_j^k \mathbf{e}_j \Big\rangle = \sum_{i,j=1}^n c_i c_j \lambda_i^k \lambda_j^k \langle \mathbf{e}_i, \mathbf{e}_j \rangle =$$

$$\sum_{i=1}^n c_i c_i \lambda_i^k \lambda_i^k \langle \mathbf{e}_i, \mathbf{e}_i \rangle = \sum_{i=1}^n c_i^2 \lambda_i^{2k},$$

$$\langle \mathbf{x}^{(k+1)}, \mathbf{x}^{(k)} \rangle = \Big\langle \sum_{i=1}^n c_i \lambda_i^{k+1} \mathbf{e}_i, \sum_{j=1}^n c_j \lambda_j^k \mathbf{e}_j \Big\rangle = \sum_{i,j=1}^n c_i c_j \lambda_i^{k+1} \lambda_j^k \langle \mathbf{e}_i, \mathbf{e}_j \rangle =$$

$$\sum_{i=1}^n c_i c_i \lambda_i^{k+1} \lambda_i^k \langle \mathbf{e}_i, \mathbf{e}_i \rangle = \sum_{i=1}^n c_i^2 \lambda_i^{2k+1}.$$

Stavimo da je

$$\mu_k = \frac{\langle \mathbf{x}^{(k+1)}, \mathbf{x}^{(k)} \rangle}{\langle \mathbf{x}^{(k)}, \mathbf{x}^{(k)} \rangle}, \qquad k \geq 0.$$

Sada je aproksimacioni niz $\{\mu_k\}_{k=0}^{\infty}$ definisan i proces računanja ili algoritam je definisan.

Teorema. Ako je $|\lambda_1| > |\lambda_2|$ i $c_1 = \langle \mathbf{x}^{(0)}, \mathbf{e}_1 \rangle \neq 0$ onda je $\lim_{k\to\infty} \mu_k = \lambda_1$ i

$$|\mu_k - \lambda_1| = O(q^k) \text{ kad } k \to \infty, \quad \text{gdje je } q = \left|\frac{\lambda_2}{\lambda_1}\right|^2 < 1. \tag{1}$$

Dokaz teoreme:

$$\mu_k = \frac{\langle \mathbf{x}^{(k+1)}, \mathbf{x}^{(k)} \rangle}{\langle \mathbf{x}^{(k)}, \mathbf{x}^{(k)} \rangle} = \frac{c_1^2\lambda_1^{2k+1} + c_2^2\lambda_2^{2k+1} + \ldots + c_n^2\lambda_n^{2k+1}}{c_1^2\lambda_1^{2k} + c_2^2\lambda_2^{2k} + \ldots + c_n^2\lambda_n^{2k}}. \tag{2}$$

Vidimo da je $\lim_{k\to\infty} \mu_k = \lambda_1$ jer je $|\lambda_2| < |\lambda_1|, \ldots, |\lambda_n| < |\lambda_1|$, kao $\lim\limits_{n\to\infty} \dfrac{5\cdot 3^{n+1} + 7\cdot 2^{n+1}}{5\cdot 3^n + 7\cdot 2^n} =$ 3. Dalje

$$\mu_k - \lambda_1 = \frac{c_2^2(\lambda_2-\lambda_1)\lambda_2^{2k} + \ldots + c_n^2(\lambda_n-\lambda_1)\lambda_n^{2k}}{c_1^2\lambda_1^{2k} + c_2^2\lambda_2^{2k} + \ldots + c_n^2\lambda_n^{2k}},$$

$$|\mu_k - \lambda_1| \le \frac{1}{c_1^2\lambda_1^{2k}}\left(c_2^2|\lambda_2-\lambda_1|\lambda_2^{2k} + \ldots + c_n^2|\lambda_n-\lambda_1|\lambda_n^{2k}\right) \le$$

$$\frac{1}{c_1^2\lambda_1^{2k}}\left(2c_2^2|\lambda_1|\lambda_2^{2k} + \ldots + 2c_n^2|\lambda_1|\lambda_2^{2k}\right) =$$

$$\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\frac{1}{c_1^2}2|\lambda_1|\left(c_2^2 + \ldots + c_n^2\right) = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right).$$

Teorema je dokazana.

Formula (1) služi za ocjenu greške k–te aproksimacije i govori da je brzina konvergencije (tempo konvergencije) metode skalarnog proizvoda prvog reda ili prvog stepena.

(a) Slučaj $|\lambda_1| = |\lambda_2|$.

Uslov teoreme $|\lambda_1| > |\lambda_2|$ govori da A ima svojstvenu vrijednost koja je dominantna po modulu. Šta će se desiti ako uslov nije ispunjen (nepovoljna okolnost)? Ima više dominantnih po modulu. Sve one su istog znaka ili među njima ima i pozitivnih i negativnih.

Ako je $\lambda_1 = \lambda_2$, $|\lambda_2| > |\lambda_3|$ tada takođe $\mu_k \to \lambda_1$ kad $k \to \infty$ pod uslovom da je $c_1^2 + c_2^2 \neq 0$. Ako je $\lambda_1 = \ldots = \lambda_p$, $|\lambda_p| > |\lambda_{p+1}|$ tada takođe $\mu_k \to \lambda_1$ kad $k \to \infty$ pod uslovom da je $c_1^2 + \ldots + c_p^2 \neq 0$. Izračunajte $\lim_{k\to\infty} \mu_k$ po formuli (2) i uvjerite se.

Ako je $\lambda_1 = \ldots = \lambda_p = -\lambda_{p+1} = \ldots = -\lambda_q$, $|\lambda_q| > |\lambda_{q+1}|$ onda nije istina da niz $\{\mu_k\}_{k=0}^{\infty}$ konvergira ka λ_1, niz $\{\mu_k\}_{k=0}^{\infty}$ ne koristi, taj niz "lažno" konvergira. Uvjerite se posmatranjem dva posebna slučaja kako slijedi. Kada je $\lambda_1 = -\lambda_2$, $|\lambda_2| > |\lambda_3|$ onda je $\lim\limits_{k\to\infty} \mu_k = \dfrac{c_1^2 - c_2^2}{c_1^2 + c_2^2}\lambda_1$ pod uslovom da je $c_1^2 + c_2^2 \neq 0$. Pogledajte i drugi poseban slučaj $\lambda_1 = \lambda_2 = -\lambda_3$, $|\lambda_3| > |\lambda_4|$.

Uzmimo da je nastupila nepovoljna okolnost koja stvara teškoće. Kako da teškoće prevaziđemo? Prvi savjet: primijenite metodu skalarnog proizvoda na matricu $A + cI$ čije su svojstvene vrijednosti $\lambda_i + c$. Drugi savjet: neka niz $\nu_k = \dfrac{\langle \mathbf{x}^{(k+2)}, \mathbf{x}^{(k)} \rangle}{\langle \mathbf{x}^{(k)}, \mathbf{x}^{(k)} \rangle}$ posluži kao aproksimacioni niz. Recimo, ako je $\lambda_1 = -\lambda_2$, $|\lambda_2| > |\lambda_3|$ onda je $\lim\limits_{k\to\infty} \nu_k = \lambda_1^2$ pod uslovom da je $c_1^2 + c_2^2 \neq 0$.

(b) Slučaj $c_1 = \langle \mathbf{x}^{(0)}, \mathbf{e}_1 \rangle = 0$.

Početna aproksimacija $\mathbf{x}^{(0)}$ bira se na slučajan način, ne koristeći bilo kakve informacije o matrici A odnosno o njenim svojstvenim vrijednostima i vektorima. Uostalom, po pravilu, mi i ne raspolažemo sa takvim informacijama. Zato se može desiti da bude $c_1 = 0$ (nepovoljna okolnost). Kada izaberemo $\mathbf{x}^{(0)}$, mi tada ne znamo da li je $c_1 \neq 0$ ili je pak suprotno $c_1 = 0$. Savjet: sprovedite algoritam dva–tri puta sa različitim početnim vektorima $\mathbf{x}^{(0)}$.

Slijede razne dopune

(a) Preveliki brojevi tokom računanja.

Posmatrajmo veličinu

$$\|\mathbf{x}^{(k)}\| = \sqrt{(\mathbf{x}^{(k)}, \mathbf{x}^{(k)})} = \sqrt{\sum_{i=1}^{n} c_i^2 \lambda_i^{2k}} = \sqrt{c_1^2 \lambda_1^{2k} + \ldots + c_n^2 \lambda_n^{2k}}.$$

Ako je $|\lambda_1| > 1$ onda je $\lim_{k\to\infty} \|\mathbf{x}^{(k)}\| = +\infty$, $c_1 \neq 0$. Kod računara će lako doći do prekoračenja. Treba vršiti prilagođavanje veličine $\|\mathbf{x}^{(k)}\|$ s vremena na vrijeme ili vršiti njeno prilagođavanje na svakom koraku. Ulogu vektora $\mathbf{x}^{(k)}$ neka preuzme vektor koji je kolinearan sa $\mathbf{x}^{(k)}$ a čiji je intenzitet $= 1$.

(b) Vektor $\mathbf{e}_1$.

Kada k neograničeno raste onda vektor

$$\mathbf{x}^{(k)} = \sum_{i=1}^{n} c_i \lambda_i^k \mathbf{e}_i = c_1 \lambda_1^k \mathbf{e}_1 + \ldots + c_n \lambda_n^k \mathbf{e}_n$$

postaje skoro kolinearan sa vektorom $\mathbf{e}_1$, pod uslovima $|\lambda_1| > |\lambda_2|$ i $c_1 \neq 0$. Tako da imamo način da približno odredimo svojstveni vektor $\mathbf{e}_1$ koji odgovara dominantnoj po modulu svojstvenoj vrijednosti λ_1.

(c) Vrijednost λ_2.

Ako spektralni podaci simetrične matrice $A = A\mathbf{x}$ glase $\lambda_1, \ldots, \lambda_n$ i $\mathbf{e}_1, \ldots \mathbf{e}_n$ onda spektralni podaci matrice $B = B\mathbf{x} = A\mathbf{x} - \lambda_1 \langle \mathbf{x}, \mathbf{e}_1 \rangle \mathbf{e}_1$ glase $0, \lambda_2, \ldots, \lambda_n$ i $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$. Kada smo odredili, makar i samo približno, λ_1 i $\mathbf{e}_1$ ($\|\mathbf{e}_1\| = 1$), mi izračunamo matricu B. Sada, primjenom metode skalarnog proizvoda na B može da bude približno određena druga svojstvena vrijednost λ_2.

Primjer za metodu skalarnog proizvoda ($n = 2$): $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$, $\mathbf{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{x}^{(1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\mathbf{x}^{(2)} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$, $\mathbf{x}^{(3)} = \begin{bmatrix} 17 \\ 7 \end{bmatrix}$, $\mathbf{x}^{(4)} = \begin{bmatrix} 41 \\ 17 \end{bmatrix}$, $\mu_0 = 2$, $\mu_1 = 2{,}4$, $\mu_2 = 2{,}41379$, $\mu_3 = 2{,}41420$ ($\lambda_1 = \sqrt{2} + 1 = 2{,}41421$).

4. RJEŠAVANJE SISTEMA NELINEARNIH JEDNAČINA

4.1. METODA POLOVLJENJA

Metoda polovljenja intervala (metoda bisekcije) služi za numeričko rješavanje jednačine oblika $f(x) = 0$.

Neka je $[a, b]$ zatvoreni interval na realnoj osi i razmotrimo funkciju $f\colon [a, b] \to R$. Pretpostavimo da je f neprekidna na $[a, b]$, $f \in C[a, b]$. Pretpostavimo da je $f(a)f(b) < 0$, različitog su znaka vrijednosti funkcije u dvije krajnje tačke intervala. Tada, kao što je poznato iz matematičke analize, po Bolcanovoj teoremi, jednačina $f(x) = 0$ ima bar jedno rješenje (funkcija ima bar jednu nulu).

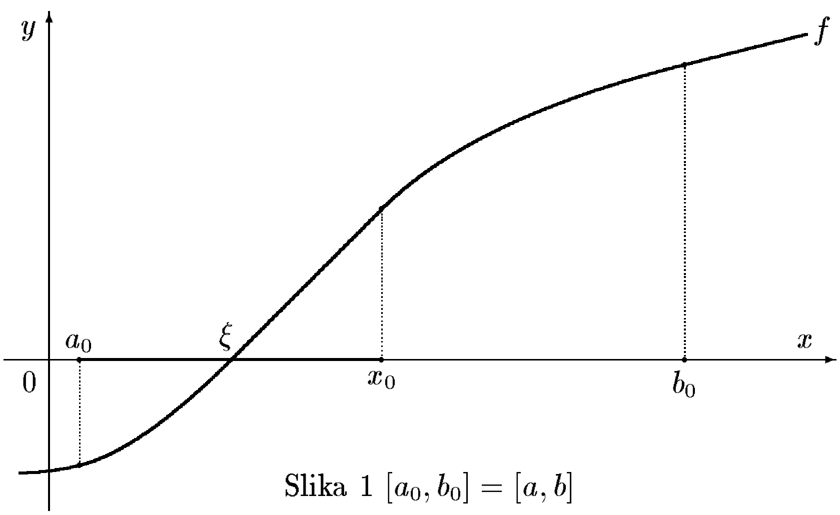
Mi ćemo konstruisati niz $\{x_n\}_{n=0}^{\infty}$ takav da važi $\lim_{n\to\infty} x_n = \xi$, gdje je $f(\xi) = 0$, ξ je rješenje jednačine, $a < \xi < b$.

Prelazimo na izlaganje numeričke metode (na izlaganje algoritma). Stavimo $[a_0, b_0] = [a, b]$ i $x_0 = (a_0 + b_0)/2$, srednja tačka. Izračunajmo $f(x_0)$. Kakvog je znaka $y_0 = f(x_0)$? Ako je $y_0 = 0$ onda smo pronašli rješenje jednačine (x_0 je rješenje). U tom slučaju, treba saopštiti x_0 kao rezultat i završava se rad algoritma. Napominje se da je mogućnost $y_0 = 0$ vrlo slabo vjerovatna, ali ipak moguća.

Prelazimo na slučaj $y_0 \neq 0$. Među brojevima $f(a_0), f(x_0), f(b_0)$ ima i pozitivnih i negativnih. Ili je $f(a_0)f(x_0) < 0$ ili je $f(x_0)f(b_0) < 0$. Ako je $f(a_0)f(x_0) < 0$ onda stavimo $[a_1, b_1] = [a_0, x_0]$, a ako je $f(x_0)f(b_0) < 0$ stavimo $[a_1, b_1] = [x_0, b_0]$. Mi smo definisali novi zatvoreni interval $[a_1, b_1]$ takav da važi $f(a_1)f(b_1) < 0$, različitog su znaka vrijednosti funkcije u dvije krajnje tačke intervala. Stavimo $x_1 = (a_1 + b_1)/2$ (srednja tačka).

Ako je $y_1 = f(x_1) \neq 0$ onda se na sličan način definiše novi interval $[a_2, b_2]$ koji predstavlja lijevu ili desnu polovinu intervala $[a_1, b_1]$ i koji ima svojstvo da obuhvata nulu funkcije (ima svojstvo $f(a_2)f(b_2) < 0$). Stavimo $x_2 = (a_2 + b_2)/2$.

Dalje se nastavlja na isti način. Kada ćemo izaći iz ponavljanja, kako glasi izlazni kriterijum? Obično je unaprijed fiksiran broj koraka n, pa će se iz ponavljanja izaći kada se odredi interval $[a_n, b_n]$. Neka algoritam saopšti $x_n = (a_n + b_n)/2$ kao odgovor i nakon toga neka se zaustavi.



Slika 1 $[a_0, b_0] = [a, b]$

Kolika je greška numeričkog odgovora $\xi \approx x_n$? Uvedimo oznaku za grešku $r_n = \xi - x_n$. Iz dosadašnjeg izlaganja imamo da je $a_n < \xi < b_n$, $b_n - a_n = (b_0 - a_0)/2^n$ i $x_n - a_n = b_n - x_n = (b_0 - a_0)/2^{n+1}$. Prema tome, $|r_n| \leq (b_0 - a_0)/2^{n+1}$ ili svejedno $|r_n| \leq (b - a)/2^{n+1}$.

Drukčiji izlazni kriterijum imamo ako je unaprijed fiksirana granica greške $\varepsilon > 0$, recimo $\varepsilon = \frac{1}{2} \cdot 10^{-5}$. Koliko treba iteracija da se zadovolji $|r_n| < \varepsilon$? Imamo $\frac{b-a}{2^{n+1}} < \varepsilon$ ili $2^{n+1} > \frac{b-a}{\varepsilon}$, pa primijenite $\log_2$.

Analiza numeričke metode.

(a) Nekoliko riječi o slučaju $f(x_i) = 0$. Razmatrani slučaj je privlačan sa stanovišta matematičke analize, prilikom dokaza Bolcanove teoreme. Nije privlačan sa stanovišta numeričkih metoda. Računar je saopštio da je $f(x_i)$ nula, ali mi ne zaborljavamo da je prisutna greška računanja. U takvoj situaciji, potrebne su dodatne mjere, da bi algoritam saopštio pouzdan rezultat.

(b) Iz formule za ocjenu greške vidi se da se iz koraka u korak granica greške množi sa $\frac{1}{2}$. Razmatrana numerička metoda konvergira tempom geometrijske progresije. Kaže se da je njena brzina konvergencije prvog stepena (prvog reda).

4.2. METODA PROSTE ITERACIJE

Uvod

Biće riječi o metodi proste iteracije za rješavanje jedne jednačine sa jednom nepoznatom, kao i sistema od n jednačina sa n nepoznatih. Za iterativnu metodu se kaže da je metoda proste iteracije ako se ona zasniva na principu kontrakcije.

Priprema: princip kontrakcije

Kaže se princip kontrakcije ili princip fiksne tačke ili Banahova teorema.

Teorema. Neka je (X, d) kompletan metrički prostor i razmotrimo preslikavanje φ: $X \to X$. Neka φ zadovoljava uslov kontrakcije: postoji broj q $(0 \le q < 1)$ takav da za sve $x, y \in X$ važi $d(\varphi(x), \varphi(y)) \le qd(x, y)$. Dalje, neka je izabran $x_0 \in X$ na proizvoljan način i definišimo niz $\{x_n\}_{n=0}^{\infty}$ relacijom: $x_{n+1} = \varphi(x_n)$, $n \ge 0$. Tada: (1) jednačina $\varphi(x) = x$ ima samo jedno rješenje u oznaci X i (2) važi $\lim_{n\to\infty} x_n = X$.

O pojmovima i terminima. Za metrički prostor se kaže da je kompletan ako u njemu svaki Košijev niz konvegira. Za niz $\{x_n\}_{n=1}^{\infty}$ kaže se da je Košijev / fundamentalan ako važi: za svako $\varepsilon > 0$, postoji n_0, takav da za sve $n \ge n_0$ i $p \ge 1$ važi $d(x_{n+p}, x_n) < \varepsilon$. Za razmatrano preslikavanje φ kaže se da predstavlja kontrakciju, a za broj q se kaže da je koeficijent kontrakcije. Za element x koji zadovoljava jednačinu $\varphi(x) = x$ kaže se da predstavlja fiksnu / nepokretnu tačku preslikavanja φ.

Dokaz teoreme. Prvo ćemo pokazati da je niz $\{x_n\}_{n=0}^{\infty}$ Košijev. Imamo redom

$$d(x_{n+2}, x_{n+1}) = d(\varphi(x_{n+1}), \varphi(x_n)) \le qd(x_{n+1}, x_n),$$

$$d(x_{n+3}, x_{n+2}) = d(\varphi(x_{n+2}), \varphi(x_{n+1})) \le qd(x_{n+2}, x_{n+1}) \le q^2 d(x_{n+1}, x_n), \quad \text{itd.}$$

$$d(x_{n+p}, x_n) \le d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \ldots + d(x_{n+1}, x_n) \le$$

$$q^{p-1} d(x_{n+1}, x_n) + \ldots + qd(x_{n+1}, x_n) + d(x_{n+1}, x_n) =$$

$$(1 + q + \ldots + q^{p-1}) d(x_{n+1}, x_n) \le (1 + q + \ldots) d(x_{n+1}, x_n) = \frac{q}{1-q} d(x_{n+1}, x_n) =$$

$$\frac{q}{1-q} d(\varphi(x_n), \varphi(x_{n-1})) \le \frac{q^2}{1-q} d(x_n, x_{n-1}) = \ldots \le \frac{q^n}{1-q} d(x_1, x_0).$$

Dobili smo $d(x_{n+p}, x_n) \le \frac{q^n}{1-q} d(x_1, x_0)$, tako da $d(x_{n+p}, x_n) \to 0$ kad $n \to \infty$, bez obzira na $p \ge 1$. Jeste Košijev. Odatle, budući da je metrički prostor kompletan, niz $\{x_n\}_{n=0}^{\infty}$ je i konvergentan. Označimo sa X njegovu graničnu vrijednost, $X = \lim_{n\to\infty} x_n$.

Preslikavanje φ je neprekidno. Zaista, $d(x,x') < \varepsilon \Rightarrow d(\varphi(x),\varphi(x')) < \varepsilon$. Iz neprekidnosti, $\lim_{n\to\infty}\varphi(x_n) = \varphi(\lim_{n\to\infty} x_n)$.

Primijenimo operaciju $\lim_{n\to\infty}$ na relaciju $x_{n+1} = \varphi(x_n)$. Tako $\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \varphi(x_n)$, $X = \lim_{n\to\infty}\varphi(x_n)$, $X = \varphi(\lim_{n\to\infty} x_n)$, $X = \varphi(X)$, fiksna tačka postoji.

Ne mogu postojati dvije fiksne tačke X_1, X_2 $(X_1 \neq X_2)$. Zaista, tada bi bilo $d(X_1,X_2) = d(\varphi(X_1),\varphi(X_2)) \le qd(X_1,X_2)$, što je moguće jedino u slučaju $d(X_1,X_2) = 0$, $X_1 = X_2$. Fiksna tačka je jedinstvena. Dokaz je završen.

Naredna teorema predstavlja nastavak prethodne teoreme i služi za ocjenu greške n–te uzastopne aproksimacije x_n.

Teorema. Važe nejednakosti

$$d(X,x_n) \le \frac{q^n}{1-q} d(x_1,x_0) \qquad (*), \qquad\qquad d(X,x_n) \le \frac{q}{1-q} d(x_n,x_{n-1}) \qquad (**).$$

Dokaz teoreme. Napišimo nejednakost trougla za tačke x_n, x_{n+1}, X:

$$d(x_n,X) \le d(x_n,x_{n+1}) + d(x_{n+1},X) =$$

$$d(\varphi(x_{n-1}),\varphi(x_n)) + d(\varphi(x_n),\varphi(X)) \le qd(x_{n-1},x_n) + qd(x_n,X) \quad \Rightarrow$$

$$(1-q)d(x_n,X) \le qd(x_{n-1},x_n), \qquad d(x_n,X) \le \frac{q}{1-q} d(x_{n-1},x_n),$$

čime smo pokazali $(**)$. Treba $(*)$. Iz uslova kontrakcije:

$$d(X,x_n) \le \frac{q}{1-q} d(x_n,x_{n-1}) = \frac{q}{1-q} d(\varphi(x_{n-1}),\varphi(x_{n-2})) \le \frac{q^2}{1-q} d(x_{n-1},x_{n-2}) = \ldots \le$$

$$\frac{q^n}{1-q} d(x_1,x_0).$$

Dokaz je završen.

Princip kontrakcije važi i u Banahovom prostoru (u kompletnom normiranom prostoru). Iz norme vektora $\|x\|$ proističe rastojanje po formuli $d(x,y) = \|x-y\|$.

Hilbertov prostor predstavlja specijalan slučaj Banahovog prostora. U Hilbertovom prostoru definisan je skalarni proizvod dva vektora $\langle x,y\rangle$. Iz skalarnog proizvoda proističe norma vektora po formuli $\|x\| = \sqrt{\langle x,x\rangle}$.

Slučaj jedne jednačine sa jednom nepoznatom

U pripremnom koraku, data jednačina $f(x) = 0$ transformiše se u ekvivalentnu jednačinu oblika $x = \varphi(x)$ i onda se bavimo njenim rješavanjem.

Teorema. Razmotrimo funkciju $\varphi \in C^1[a,b]$. Neka su ispunjeni uslovi: (1) ako je $a \le x \le b$ onda je $a \le \varphi(x) \le b$ i (2) postoji q takav da je $|\varphi'(x)| \le q < 1$ za $a \le x \le b$. Definišimo niz brojeva $\{x_n\}_{n=0}^{\infty}$ sa $x_{n+1} = \varphi(x_n)$ za $n \ge 0$, gdje $x_0 \in [a,b]$. Tada važi: (1) jednačina $x = \varphi(x)$ ima jedinstveno rješenje na intervalu $[a,b[$ (označimo ga sa ξ) i (2) $\lim_{n\to\infty} x_n = \xi$.

Dokaz teoreme. Po Lagranžovoj teoremi o srednjoj vrijednosti $\varphi(\alpha_2) - \varphi(\alpha_1) = \varphi'(\beta)(\alpha_2 - \alpha_1)$ gdje je $\alpha_1 < \beta < \alpha_2 \Rightarrow |\varphi(\alpha_2) - \varphi(\alpha_1)| = |\varphi'(\beta)| \cdot |\alpha_2 - \alpha_1| \le q|\alpha_2 - \alpha_1|$.

Kako φ: $[a,b] \to [a,b]$ i $x_0 \in [a,b]$ to $x_n \in [a,b]$ za svako n.

Dokažimo da je $\{x_n\}$ Košijev niz. Imamo $|x_{n+1} - x_n| = |\varphi(x_n) - \varphi(x_{n-1})| \le q|x_n - x_{n-1}|$. Slično $|x_{n+2} - x_{n+1}| = |\varphi(x_{n+1}) - \varphi(x_n)| \le q|x_{n+1} - x_n| \le q^2|x_n - x_{n-1}|$. Itd. Isto tako $|x_{n+p} - x_{n+p-1}| \le q^p|x_n - x_{n-1}|$. Na isti način se dokazuje i $|x_n - x_{n-1}| \le q^{n-1}|x_1 - x_0|$. Ukupno

$$|x_{n+p} - x_n| = |x_{n+p} - x_{n+p-1} + \ldots + x_{n+2} - x_{n+1} + x_{n+1} - x_n| \le$$

$$|x_{n+p}-x_{n+p-1}|+\ldots+|x_{n+2}-x_{n+1}|+|x_{n+1}-x_n|\le(q^p+\ldots+q^2+q)|x_n-x_{n-1}|\le$$

$$(q+q^2+\ldots)|x_n-x_{n-1}|=\frac{q}{1-q}|x_n-x_{n-1}|\le\frac{q^n}{1-q}|x_1-x_0|\to 0 \text{ kad } n\to\infty$$

bez obzira na $p\ge 1$.

Na realnoj osi, svaki Košijev niz je konvergentan. Pokazali smo da je niz $\{x_n\}$ konvergentan i odmah uvodimo oznaku $X=\lim_{n\to\infty}x_n$. Iz $a\le x_n\le b$ za svako $n\Rightarrow a\le X\le b$. Iz $x_{n+1}=\varphi(x_n)$ slijedi $X=\lim_{n\to\infty}\varphi(x_n)$ i dalje slijedi (budući da je φ neprekidna funkcija) $X=\varphi(X)$. Znači $X=\xi$. Ne mogu postojati dva rješenja ξ_1 i ξ_2 jer bi tada bilo $\varphi(\xi_1)=\xi_1$, $\varphi(\xi_2)=\xi_2$ i (za neko β) $|\xi_2-\xi_1|=|\varphi(\xi_2)-\varphi(\xi_1)|=|\varphi'(\beta)|\cdot|\xi_2-\xi_1|\le q|\xi_2-\xi_1|$, a znamo da je $q<1$. Dokaz je završen.

Važi nejednakost (za ocjenu greške) $|x_n-\xi|\le\frac{q^n}{1-q}|x_1-x_0|$ za svako n. Isto tako, $|x_n-\xi|\le\frac{q}{1-q}|x_n-x_{n-1}|$ za svako n.

Dokaz druge nejednakosti:

$$|x_n-\xi|\le|x_n-x_{n+1}|+|x_{n+1}-\xi|=|\varphi(x_{n-1})-\varphi(x_n)|+|\varphi(x_n)-\varphi(\xi)|\le$$

$$q|x_{n-1}-x_n|+q|x_n-\xi| \quad\Rightarrow\quad (1-q)|x_n-\xi|\le q|x_{n-1}-x_n| \qquad \Big/ :(1-q)$$

Dokaz prve nejednakosti:

$$|x_n-\xi|\le\frac{q}{1-q}|x_n-x_{n-1}|=\frac{q}{1-q}|\varphi(x_{n-1})-\varphi(x_{n-2})|\le\frac{q^2}{1-q}|x_{n-1}-x_{n-2}|=$$

$$\frac{q^2}{1-q}|\varphi(x_{n-2})-\varphi(x_{n-3})|\le\frac{q^3}{1-q}|x_{n-2}-x_{n-3}|=\ldots\le\frac{q^n}{1-q}|x_1-x_0|.$$

Jasno, po Lagranžovoj teoremi o srednjoj vrijednosti imamo $\varphi(\alpha_2)-\varphi(\alpha_1)=\varphi'(\beta)(\alpha_2-\alpha_1)$, gdje je $\beta=\alpha_1+\theta(\alpha_2-\alpha_1)$ za neko $0<\theta<1\Rightarrow|\varphi(\alpha_2)-\varphi(\alpha_1)|\le q|\alpha_2-\alpha_1|$ za bilo koje $\alpha_1,\alpha_2\in[a,b]$.

(Nismo morali da se zamaramo oko dokaza teoreme, budući da se radi o specijalnom slučaju principa kontrakcije kada je $X=[a,b]$, $d(x,y)=|x-y|$, $q=\max_{x\in[a,b]}|\varphi'(x)|$.)

U zaključku, potrebno je da φ preslikava jedan interval $[a,b]$ u taj isti interval i da važi $|\varphi'(x)|\le q<1$ za $x\in[a,b]$. Tempo konvergencije je linearan (prvog stepena).

Naknadno dodato: za fiksnu tačku ξ preslikavanja φ kaže se da je atraktivna ako je $|\varphi'(\xi)|<1$, da je repulsivna ako je $|\varphi'(\xi)|>1$, da je neutralna ako je $\varphi'(\xi)=\pm1$.

Primjer

Razmotrimo jednačinu $x=\sqrt{1+x}$ na intervalu $1\le x\le 2$. Nacrtati odgovarajuću sliku, nacrtati grafik funkcije $y=\sqrt{1+x}$ na dijelu $1\le x\le 2$, kao i pravu $y=x$. Stavimo $\varphi(x)=\sqrt{1+x}$. Data jednačina ima na tom intervalu jedinstveno rješenje, rješenje može da bude nađeno po metodi proste iteracije (sa proizvoljnom preciznošću) zato što su ispunjeni uslovi teoreme od maločas. Zaista, $1\le x\le 2\Rightarrow\sqrt{2}\le\varphi(x)\le\sqrt{3}$, odnosno funkcija φ prevodi interval $[1,2]$ u samog sebe. Pored toga, $\varphi'(x)=\frac{1}{2\sqrt{1+x}}$ pa je $|\varphi'(x)|\le\frac{1}{2}$ za $x\in[1,2]$. Imamo da je $q=\frac{1}{2}$. Radeći preciznije, možemo pisati da je $q=\frac{1}{2\sqrt{2}}$. Kao x_0 uzme se bilo koja tačka intervala $[1,2]$, recimo stavimo $x_0=1{,}5$. Zatim redom računamo po formuli $x_{n+1}=\sqrt{1+x_n}$. Proces će konvergirati dosta dobrim tempom zato što je koeficijent kontrakcije q manji od $\frac{1}{2}$. Imamo da je $\lim_{n\to\infty}x_n=\xi$, gdje je $\xi=(1+\sqrt{5})/2=1{,}61803$. Uzastopne aproksimacije: $x_1=1{,}58114$, $x_2=1{,}60659$, $x_3=1{,}61449$, $x_4=1{,}61694$, $x_5=1{,}61770$, $x_6=1{,}61793$, ….

Kontrakcija: mali primjer: $\alpha_2 - \alpha_1 = 2 - 1 = 1$, $\varphi(\alpha_2) - \varphi(\alpha_1) = 1{,}73 - 1{,}41 = 0{,}32$, $q = \frac{1}{2\sqrt{2}} = 0{,}35$.

Uopšte, kada treba riješiti jednačinu npr. $x = \sqrt{1+x}$, mi prvo nacrtamo dva grafika $y = x$ i $y = \sqrt{1+x}$ i vidimo gdje se krive sijeku. Tako odredimo otprilike gdje ima rješenja, to je tzv. lokalizacija rješenja, odredimo $[a, b]$. Jednačina $x = \varphi(x) \Rightarrow$ dva grafika $y = x$ i $y = \varphi(x)$.

Kako postići da uslov kontrakcije bude ispunjen?

Budući da se radi o numeričkoj, dosadašnja priča ne vrijedi puno ako se ne da neko uputstvo o pogodnom načinu transformacije iz polaznog oblika $f(x) = 0$ u oblik $x = \varphi(x)$.

Pogledajmo prvo na primjeru. Neka se traži rješenje jednačine $x = \varphi(x)$ na nekom intervalu $[a, b]$. Neka smo za prvi izvod funkcije φ na tom intervalu ustanovili da važi $4 \le \varphi'(x) \le 6$. Ovo vrlo slabo izgleda sa stanovišta primjene metode proste iteracije. Međutim, polazna jednačina je ekvivalentna sa $x = \frac{x+\varphi(x)}{2}$. Izvod funkcije sa desne strane $\frac{x+\varphi(x)}{2}$ kreće se očito između 2,5 i 3,5. Izgleda nam da se nekako može podesiti da se izvod desne strane kreće između -1 i 1.

Uopšte, jednačina $x = \varphi(x)$ očito je ekvivalentna sa jednačinom $x + \lambda x = \varphi(x) + \lambda x$, tj. $x = \frac{\varphi(x)+\lambda x}{1+\lambda}$, $\lambda \neq -1$. Veza izvoda funkcija $\varphi(x)$ i $\frac{\varphi(x)+\lambda x}{1+\lambda}$ je očita. Uspješnost primjene ovog tzv. λ–postupka zavisi od toga – koliko jasnom informacijom o ponašanju izvoda φ' raspolažemo. Tako da će se sukcesivne aproksimacije računati po formuli $x_{n+1} = \frac{\varphi(x_n)+\lambda x_n}{1+\lambda}$. Drukčije se može reći: neka funkcija $\frac{\varphi(x)+\lambda x}{1+\lambda}$ preuzme ulogu funkcije $\varphi(x)$.

Slučaj sistema od dvije jednačine sa dvije nepoznate

U pripremnom koraku, polazni sistem $f(x, y) = 0$, $g(x, y) = 0$ transformiše se u oblik $x = \varphi(x, y)$, $y = \psi(x, y)$.

U pripremnom koraku, odredi se skup $\Omega \subset R^2$ koji sadrži rješenje (ξ, η) razmatranog sistema jednačina. Obično je Ω pravougaonik ili krug. Skup Ω treba da bude zatvoren, da bi metrički prostor bio kompletan. Možemo pisati $\varphi\colon \Omega \to R$, $\psi\colon \Omega \to R$ ili $(\varphi, \psi)\colon \Omega \to R^2$.

U nastavku, mi ćemo uglavnom nastojati da konkretizujemo razne elemente principa kontrakcije. Treba izabrati početnu aproksimaciju $(x_0, y_0) \in \Omega$. Uzastopne aproksimacije računamo po formuli $x_{n+1} = \varphi(x_n, y_n)$, $y_{n+1} = \psi(x_n, y_n)$, $n \ge 0$. Da li važi $\lim_{n\to\infty} x_n = \xi$, $\lim_{n\to\infty} y_n = \eta$?

Uvedimo potrebne oznake. Neka su m_{ij} $(i, j = 1, 2)$ ma koji brojevi koji zadovoljavaju nejednakosti:

$$\sup_{(x,y)\in\Omega} \left|\frac{\partial\varphi(x,y)}{\partial x}\right| \le m_{11}, \qquad \sup_{(x,y)\in\Omega} \left|\frac{\partial\varphi(x,y)}{\partial y}\right| \le m_{12},$$

$$\sup_{(x,y)\in\Omega} \left|\frac{\partial\psi(x,y)}{\partial x}\right| \le m_{21}, \qquad \sup_{(x,y)\in\Omega} \left|\frac{\partial\psi(x,y)}{\partial y}\right| \le m_{22}.$$

Dalje, $M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$. Zatim $q_1 = \|M\|_1 = \max\{m_{11} + m_{21},\, m_{12} + m_{22}\}$. Norma matrice indukovana je normom vektora $\|(x, y)\|_1 = |x| + |y|$.

Teorema 1. Neka su ispunjeni uslovi; (i) skup $\Omega \subset R^2$ je zatvoren i konveksan, (ii) $\varphi, \psi \in C^1(\Omega)$ i $(\varphi, \psi)\colon \Omega \to \Omega$, (iii) $q_1 < 1$. Tada: (1) sistem $x = \varphi(x, y)$, $y = \psi(x, y)$ ima jedinstveno rješenje i (2) važi $\lim_{n\to\infty} x_n = \xi$, $\lim_{n\to\infty} y_n = \eta$ bez obzira na izbor (x_0, y_0).

Dokaz teoreme. Uzmimo $(x_1, y_1), (x_2, y_2) \in \Omega$ i posmatrajmo razliku $\varphi(x_2, y_2) - \varphi(x_1, y_1)$. Po Lagranžovoj teoremi o srednjoj vrijednosti:

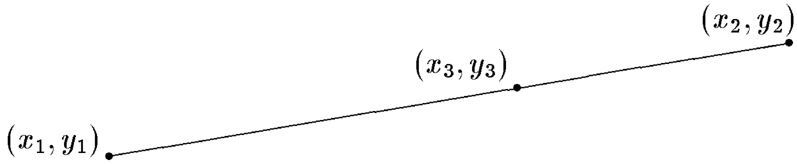
$$\varphi(x_2, y_2) - \varphi(x_1, y_1) = \frac{\partial\varphi(x_3, y_3)}{\partial x}(x_2 - x_1) + \frac{\partial\varphi(x_3, y_3)}{\partial y}(y_2 - y_1),$$

gdje je (x_3, y_3) neka tačka sa duži od (x_1, y_1) do (x_2, y_2), detaljnije se piše $x_3 = x_1 + \theta(x_2 - x_1)$ i $y_3 = y_1 + \theta(y_2 - y_1)$ za neko θ, $0 < \theta < 1$. Konveksnost skupa Ω obezbjeđuje da i $(x_3, y_3) \in \Omega$.

Na isti način

$$\psi(x_2, y_2) - \psi(x_1, y_1) = \frac{\partial\psi(x_4, y_4)}{\partial x}(x_2 - x_1) + \frac{\partial\psi(x_4, y_4)}{\partial y}(y_2 - y_1),$$

gdje je (x_4, y_4) neka (druga) tačka sa iste one duži.



Slika 2

Imamo:

$$|\varphi_2 - \varphi_1| \leq m_{11}|x_2 - x_1| + m_{12}|y_2 - y_1|, \qquad |\psi_2 - \psi_1| \leq m_{21}|x_2 - x_1| + m_{22}|y_2 - y_1|, \qquad (*)$$

gdje su uvedene skraćenice $\varphi_i = \varphi(x_i, y_i)$, $\psi_i = \psi(x_i, y_i)$, $i = 1, 2$.

Uvedimo nove skraćenice $\mathbf{a}_1 = (x_1, y_1)$, $\mathbf{a}_2 = (x_2, y_2)$, $\mathbf{b}_1 = (\varphi_1, \psi_1)$, $\mathbf{b}_2 = (\varphi_2, \psi_2)$. Tako da preslikavanje (φ, ψ) prevodi tačku $\mathbf{a}_1$ u tačku $\mathbf{b}_1$, a prevodi $\mathbf{a}_2$ u $\mathbf{b}_2$.

Saberimo dvije nejednakosti $(*)$:

$$|\varphi_2 - \varphi_1| + |\psi_2 - \psi_1| \leq (m_{11} + m_{21})|x_2 - x_1| + (m_{12} + m_{22})|y_2 - y_1| \leq q_1(|x_2 - x_1| + |y_2 - y_1|),$$

odnosno $\|\mathbf{b}_2 - \mathbf{b}_1\|_1 \leq q_1\|\mathbf{a}_2 - \mathbf{a}_1\|_1$. Time je dokaz teoreme završen, uz pozivanje na princip kontrakcije.

Označimo $q_\infty = \|M\|_\infty = \max\{m_{11} + m_{12}, m_{21} + m_{22}\}$. Norma matrice indukovana je normom vektora $\|(x, y)\|_\infty = \max\{|x|, |y|\}$.

Teorema 2. Neka su ispunjeni uslovi; (i) skup $\Omega \subset R^2$ je zatvoren i konveksan, (ii) $\varphi, \psi \in C^1(\Omega)$ i $(\varphi, \psi)\colon \Omega \to \Omega$, (iii) $q_\infty < 1$. Tada: (1) sistem $x = \varphi(x, y)$, $y = \psi(x, y)$ ima jedinstveno rješenje i (2) važi $\lim_{n\to\infty} x_n = \xi$, $\lim_{n\to\infty} y_n = \eta$ bez obzira na izbor (x_0, y_0).

Zaista, iz $(*)$ imamo, kada se umjesto $|x_2 - x_1|$ odnosno $|y_2 - y_1|$ napiše veći od ta dva broja:

$$|\varphi_2 - \varphi_1| \leq q_\infty \max\{|x_2 - x_1|, |y_2 - y_1|\} \qquad \text{i} \qquad |\psi_2 - \psi_1| \leq q_\infty \max\{|x_2 - x_1|, |y_2 - y_1|\}.$$

Slijedi $\|\mathbf{b}_2 - \mathbf{b}_1\|_\infty \leq q_\infty\|\mathbf{a}_2 - \mathbf{a}_1\|_\infty$ jer $\alpha \leq \gamma$ i $\beta \leq \gamma \Rightarrow \max\{\alpha, \beta\} \leq \gamma$.

Kako postići da uslov kontrakcije bude ispunjen?

Polazni sistem $f(x, y) = 0$, $g(x, y) = 0$ može da se zapiše u obliku:

$$x = x + \alpha f(x, y) + \beta g(x, y), \;\; y = y + \gamma f(x, y) + \delta g(x, y),$$

samo da je $\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} \neq 0$. Konstante α, β, γ i δ biraju se pogodno, tj. biraju se tako da parcijalni izvodi funkcija $\varphi(x, y) = x + \alpha f(x, y) + \beta g(x, y)$ i $\psi(x, y) = y + \gamma f(x, y) + \delta g(x, y)$ budu bliski nuli.

Opšti slučaj sistema od više jednačina sa više nepoznatih

Sistem oblika $x_i = \varphi_i(x_1, \ldots, x_n)$, $1 \le i \le n$ razmatra se po analogiji sa dosadašnjim tekstom.

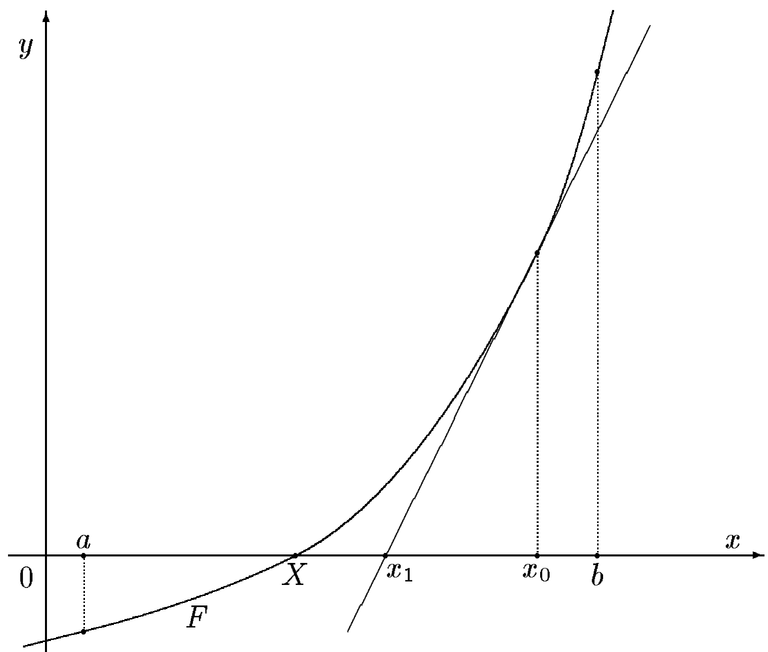
4.3. NJUTNOVA METODA

Uvod

Njutnova metoda (Njutn–Rafsonova metoda) služi za rješavanje jedne jednačine sa jednom nepoznatom oblika $F(x) = 0$, kao i za rješavanje sistema od n jednačina sa n nepoznatih oblika $F_i(x_1, \ldots, x_n) = 0$, $1 \le i \le n$.

Slučaj jedne jednačine sa jednom nepoznatom (metoda tangente)

Razmotrimo neprekidnu funkciju F i razmotrimo jednačinu $F(x) = 0$. Uzmimo da smo, grafičkom metodom ili na neki način, odredili interval koji sadrži rješenje jednačine u oznaci X. Drugim riječima, odredili smo realne brojeve a, b $(a < b)$ takve da je $F(a)F(b) < 0$. Izaberimo početnu aproksimaciju $x_0 \in [a, b]$. Posmatrajmo pravu liniju koja prolazi kroz tačku $(x, y) = (x_0, F(x_0))$. Jednačina prave kroz jednu tačku: $y - F(x_0) = k(x - x_0)$. Želimo da ta prava bude tangenta grafika funkcije $y = F(x)$ u navedenoj tački. Zato $k = F'(x_0)$, jednačina tangente $y - F(x_0) = F'(x_0)(x - x_0)$. Posmatrajmo presječnu tačku tangente i x–ose u oznaci x_1. Izraz za x_1? Jednačina x–ose glasi $y = 0$ pa zato $-F(x_0) = F'(x_0)(x - x_0)$, $x - x_0 = -\frac{F(x_0)}{F'(x_0)}$, $x = x_0 - \frac{F(x_0)}{F'(x_0)}$, $x_1 = x_0 - \frac{F(x_0)}{F'(x_0)}$. Sada imamo novu aproksimaciju x_1. Dalje, na sličan način, postavimo tangentu na grafik funkcije u tački sa grafika čija je apscisa $x = x_1$ i označimo sa x_2 apscisu presječne tačke x–ose i te tangente. Ponavljanjem, dobijamo niz uzastopnih aproksimacija $\{x_n\}_{n=0}^{\infty}$. Poželjno je da bude $\lim_{n\to\infty} x_n = X$.



Slika 3 $F(X) = 0$

Teorema (o dovoljnim uslovima za konvergenciju Njutnove metode). Neka su ispunjeni uslovi: (i) $F \in C^2[a, b]$, (ii) $F(a)F(b) < 0$, (iii) $F'(x)$ je stalnog znaka na $[a, b]$, (iv) F'' je stalnog znaka na $[a, b]$ i (v) tačka $x_0 \in [a, b]$ izabrana je tako da bude $F(x_0)F''(x_0) > 0$. Neka je niz brojeva $\{x_n\}$ definisan sa:

$$x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)} \tag{1}$$

za $n \geq 0$. Tada jednačina $F(x) = 0$ ima jedinstveno rješenje na $[a, b]$ (označimo ga sa X) i važi $\lim_{n\to\infty} x_n = X$.

Dokaz. Iz $F(a)F(b) < 0$ i F' je stalnog znaka slijedi postojanje i jedinstvenost rješenja X. U zavisnosti od toga kakvog su znaka $F'(x)$ i $F''(x)$ moguća su četiri slučaja i to: (1) $F' > 0$, $F'' > 0$, (2) $F' > 0$, $F'' < 0$, (3) $F' < 0$, $F'' > 0$, (4) $F' < 0$, $F'' < 0$. Mi ćemo sprovesti dokaz za prvi slučaj. Za ostala tri slučaja, dokaz je sličan. Dakle, imamo okolnosti kao na slici.

Prvo. Važi $x_n > X$ za svako $n \geq 0$ (zapaziti odmah da je ovaj uslov ekvivalentan sa uslovom $F(x_n) > 0$). Dokazuje se indukcijom. Imamo da je $x_0 > X$ jer je $F''(x) > 0$ i $F(x_0)F''(x_0) > 0$. Ako je $x_n > X$ onda je i $x_{n+1} > X$. Zaista, geometrijski, gledajući od tačke $x = x_n$ unazad, tangenta opada brže od funkcije pa će tangenta presjeći x–osu prije nego funkcija. A analitički se pokaže ako se napiše razvoj po Tejlorovoj formuli čiji ostatak je izražen preko drugog izvoda:

$$F(X) = F(x_n) + F'(x_n)(X - x_n) + \frac{1}{2}F''(\alpha)(X - x_n)^2 \quad \Rightarrow \quad x_{n+1} = X + \frac{F''(\alpha)(X - x_n)^2}{2F'(x_n)} > X,$$

$\alpha = x_n + \theta(X - x_n)$, $0 < \theta < 1$.

Drugo. Niz $\{x_n\}_{n=0}^{\infty}$ je monotono opadajući, tj. važi $x_{n+1} < x_n$ za $n \geq 0$. Zaista, $x_{n+1} - x_n = -F(x_n)/F'(x_n)$, pri čemu je $F(x_n) > 0$ (maločas je pokazano), a takođe $F'(x_n) > 0$ (definicioni uslov prvog slučaja).

Treće. Taj niz je konvergentan budući da je monoton i ograničen (svi njegovi elementi su veći od X). Označimo sa ξ graničnu vrijednost ovog niza $\{x_n\}$.

I četvrto. Na relaciju (1) primijenimo operaciju $\lim_{n\to\infty}$. Niz $\{x_{n+1}\}$ je podniz od $\{x_n\}$ pa i on teži ka ξ. Funkcija F je neprekidna pa važi $\lim_{n\to\infty} F(x_n) = F(\lim_{n\to\infty} x_n)$. Tako da važi sljedeće: $\xi = \xi - F(\xi)/F'(\xi)$. Kako $\xi \in [a, b]$ to je $F'(\xi) \neq 0$. Dakle, $F(\xi) = 0$. Znači da je $\xi = X$. Dokaz je završen.

Teorema (o ocjeni greške). Važi sljedeća nejednakost koja nam služi kao formula za ocjenjivanje greške:

$$|X - x_n| \leq \frac{M_2}{2m_1}(x_n - x_{n-1})^2$$

za svako $n \geq 1$. Ovdje je m_1 ma koji broj za koji važi $m_1 \leq |F'(x)|$ za svako $x \in [a, b]$, a M_2 je bilo koji broj koji zadovoljava uslov $|F''(x)| \leq M_2$ za $x \in [a, b]$.

Ova teorema predstavlja nastavak prethodne teoreme pa od nje preuzima oznake i uslove. Naravno da se može reći da je m_1 infimum od $|F'(x)|$. Broj $m_1 > 0$ sa ovim svojstvom sigurno postoji jer je $F'(x) \neq 0$ i F' neprekidna i $[a, b] \subset R$ zatvoren i ograničen. Slično se može reći da je M_2 supremum modula drugog izvoda.

Dokaz. Razvijmo funkciju $y = F(x)$ po Tejlorovoj formuli u okolini tačke $x = x_{n-1}$ do drugog izvoda:

$$F(x) = F(x_{n-1}) + F'(x_{n-1})(x - x_{n-1}) + \frac{1}{2}F''(\alpha)(x - x_{n-1})^2,$$

α zavisi od x. Iskoristimo ovaj razvoj za $x = x_n$: $F(x_n) = F(x_{n-1}) + F'(x_{n-1})(x_n - x_{n-1}) + \frac{1}{2}F''(\alpha)(x_n - x_{n-1})^2$, α je neka tačka između x_{n-1} i x_n, $|F''(\alpha)| \leq M_2$. Izraz $F(x_{n-1}) + F'(x_{n-1})(x_n - x_{n-1})$ predatavlja linearni dio Tejlorovog razvoja i jednak je u ovoj situaciji nuli po geometrijskoj definiciji prvog izvoda, po definiciji metode tangente. Tako da se ranija formula svodi na $F(x_n) = \frac{1}{2}F''(\alpha)(x_n - x_{n-1})^2$. Slijedi

$$|F(x_n)| \leq \frac{M_2}{2}(x_n - x_{n-1})^2. \tag{2}$$

S druge strane imamo: $F(x_n) = F(x_n) - F(X) = F'(\beta)(x_n - X)$, Lagranžova teorema o srednjoj vrijednosti. Slijedi $|F(x_n)| = |F'(\beta)| \cdot |x_n - X|$, $|F(x_n)| \geq m_1|x_n - X|$ ili svejedno

$$|x_n - X| \leq \frac{|F(x_n)|}{m_1}. \tag{3}$$

Zapaziti usput da se i posljednja relacija (3) može smatrati jednom formulom za ocjenu greške, za ocjenu udaljenosti nekog broja x_n od nule X funkcije F, kao i da izvođenje te relacije nije zavisno od Njutnove metode pa se ta relacija može koristiti za Njutnovu metodu a i za druge metode.

Kombinovanjem (2) i (3) imamo $|x_n - X| \leq \frac{1}{m_1}\frac{M_2}{2}(x_n - x_{n-1})^2$. Dokaz je završen.

Nabrojimo sada četiri glavne karakteristike (1)–(4) Njutnove metode. (1) Red ili brzina konvergencije Njutnove metode je drugi (kvadratni). Znamo da metoda proste iteracije ima samo prvi red konvergencije. Kod metode proste iteracije bilo je $|r_{n+1}| \leq q|r_n|$, gdje je $r_n = \xi - x_n$, a kod Njutnove metode $r_{n+1} \sim cr_n^2$, gdje je $c = -\frac{F''(\xi)}{2F'(\xi)}$. (2) Za uspješnu primjenu Njutnove metode potrebno je da raspolažemo dovoljno dobrom početnom aproksimacijom x_0. (3) Da bi Njutnova metoda konvergirala treba da $|F'|$ bude odvojen od nule. Ili svejedno, treba da $|F'|^{-1}$ bude ograničena veličina. (4) Da bi Njutnova metoda konvergirala, treba da $|F''|$ bude odvojen od $+\infty$ (tj. da bude ograničen).

Kratka naomena o modifikovanoj Njutnovoj metodi

Uzima se $x_{n+1} = x_n - \frac{F(x_n)}{F'(x_0)}$. Stepen konvergencije popušta sa drugog na prvi.

Slučaj sistema od dvije jednačine sa dvije nepoznate

Razmotrimo sistem jednačina $F(x,y) = 0$, $G(x,y) = 0$. Uzmimo da raspolažemo početnom aproksimacijom (x_0, y_0). Kako se, na bazi (x_n, y_n), određuje iduća aproksimacija (x_{n+1}, y_{n+1})? Napišimo dva Tejlorova reda za $F(x,y)$ i $G(x,y)$ u okolini (x_n, y_n) i izjednčimo ih sa nulom, $F(x,y) = 0$, $G(x,y) = 0$:

$$\begin{cases} F(x_n,y_n) + \dfrac{\partial F(x_n,y_n)}{\partial x}(x - x_n) + \dfrac{\partial F(x_n,y_n)}{\partial y}(y - y_n) + \ldots = 0 \\ G(x_n,y_n) + \dfrac{\partial G(x_n,y_n)}{\partial x}(x - x_n) + \dfrac{\partial G(x_n,y_n)}{\partial y}(y - y_n) + \ldots = 0 \end{cases}$$

gdje je $(x,y) = (X,Y)$, korijen. Ako zadržimo samo članove do beskonačno malih prvog reda:

$$\begin{cases} F(x_n,y_n) + \dfrac{\partial F(x_n,y_n)}{\partial x}(x - x_n) + \dfrac{\partial F(x_n,y_n)}{\partial y}(y - y_n) = 0 \\ G(x_n,y_n) + \dfrac{\partial G(x_n,y_n)}{\partial x}(x - x_n) + \dfrac{\partial G(x_n,y_n)}{\partial y}(y - y_n) = 0 \end{cases}$$

gdje je sada $(x,y) = (x_{n+1}, y_{n+1})$. Tako, u matričnom obliku

$$\begin{bmatrix} F(x_n,y_n) \\ G(x_n,y_n) \end{bmatrix} + \begin{bmatrix} \dfrac{\partial F(x_n,y_n)}{\partial x} & \dfrac{\partial F(x_n,y_n)}{\partial y} \\ \dfrac{\partial G(x_n,y_n)}{\partial x} & \dfrac{\partial G(x_n,y_n)}{\partial y} \end{bmatrix} \cdot \begin{bmatrix} x_{n+1} - x_n \\ y_{n+1} - y_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Zatim, po Kramerovom pravilu

$$\Delta = \begin{vmatrix} \dfrac{\partial F(x_n,y_n)}{\partial x} & \dfrac{\partial F(x_n,y_n)}{\partial y} \\ \dfrac{\partial G(x_n,y_n)}{\partial x} & \dfrac{\partial G(x_n,y_n)}{\partial y} \end{vmatrix}, \qquad \Delta_x = \begin{vmatrix} F(x_n,y_n) & \dfrac{\partial F(x_n,y_n)}{\partial y} \\ G(x_n,y_n) & \dfrac{\partial G(x_n,y_n)}{\partial y} \end{vmatrix},$$

$$\Delta_y = \begin{vmatrix} \dfrac{\partial F(x_n,y_n)}{\partial x} & F(x_n,y_n) \\ \dfrac{\partial G(x_n,y_n)}{\partial x} & G(x_n,y_n) \end{vmatrix}, \qquad \begin{cases} x_{n+1} = x_n - \dfrac{\Delta_x}{\Delta} \\ y_{n+1} = y_n - \dfrac{\Delta_y}{\Delta} \end{cases}$$

$(\Delta \neq 0)$.

Opšti slučaj sistema od više jednačina

Neka je $n \geq 1$ i razmotrimo funkcije $F_i \colon \Omega \to R$ $(\Omega \subset R^n)$. Posmatrajmo sistem jednačina $F_i(x_1, \ldots, x_n) = 0$, $1 \leq i \leq n$. Možemo ga zapisati u vektorskom obliku kao $\mathbf{F}(\mathbf{x}) = 0$, $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{F} = (F_1, \ldots, F_n)$. Znači da $\mathbf{F} \colon \Omega \to R^n$. Postupak rješavanja analogan je dosadašnjem tekstu.

U Banahovom prostoru (metoda Njutn–Kantoroviča)

Neka su X i Y dva Banahova prostora, u njima se norma označava redom kao $\| \, \|_X$ i $\| \, \|_Y$. Neka je F operator, $F \colon X \to Y$. Razmotrimo jednačinu $F(x) = 0$. Neka X označava njen korijen.

Uvedimo pojam jakog izvoda ili izvoda po Frešeu preslikavanja F. Za linearni operator $P \colon X \to Y$ (ako takav operator P postoji) kažemo da predstavlja izvod operatora F u tački x ako važi jednakost:

$$\|F(x+\eta) - F(x) - P\eta\|_Y = o(\|\eta\|_X) \qquad \text{kad} \ \ \|\eta\|_X \to 0,$$

odsad ćemo umjesto P pisati $F'(x)$.

Ako polazimo od x_n, onda treba odrediti x_{n+1}. Odgovarajući uslov:

$$F(x_n) + F'(x_n)(x_{n+1} - x_n) = 0, \tag{4}$$

a odatle je

$$x_{n+1} = x_n - (F'(x_n))^{-1} F(x_n), \qquad n = 0, 1, \ldots. \tag{5}$$

Očito se pretpostavlja da $F'(x)$ postoji i da $(F'(x_n))^{-1}$ postoji.

Neka su za neke konstante $a > 0$, $a_1, a_2 \in R$ ispunjena sljedeća dva uslova:

$$\|(F'(x))^{-1}\| \leq a_1 \qquad \text{za} \ \ x \in \Omega_a = \{x \colon \|x - X\|_X < a\} \tag{6}$$

(na lijevoj strani znaka $\leq$ je norma operatora) i

$$\|F(U_1) - F(U_2) - F'(U_2)(U_1 - U_2)\|_Y \leq a_2 \|U_1 - U_2\|_X^2 \qquad \text{za} \ \ U_1, U_2 \in \Omega_a. \tag{7}$$

Uvedimo oznake $c = a_1 a_2$ i $b = \min\left\{a, \dfrac{1}{c}\right\}$.

Teorema. Ako su ispunjeni uslovi (6) i (7) i ako $x_0 \in \Omega_b$ onda Njutnov iterativni proces (5) konvergira i važi sljedeća formula (za ocjenu greške):

$$\|x_n - X\|_X \leq \frac{1}{c}\Big(c\|x_0 - X\|_X\Big)^{2^n}. \tag{8}$$

Dokaz. Vidi sliku 4. Prvo. Svi x_n pripadaju $\Omega_b = \{x \colon \|x - X\|_X < b\}$, što ćemo dokazati indukcijom. Za $n = 0$ ispunjeno je po uslovu teoreme. Treba izvršiti indukcijski korak. Uzmimo da $x_n \in \Omega_b$. Kada u (7) uvrstimo $U_1 = X$, $U_2 = x_n$ tada imamo

$$\|F(X) - F(x_n) - F'(x_n)(X - x_n)\|_Y \leq a_2 \|x_n - X\|_X^2,$$

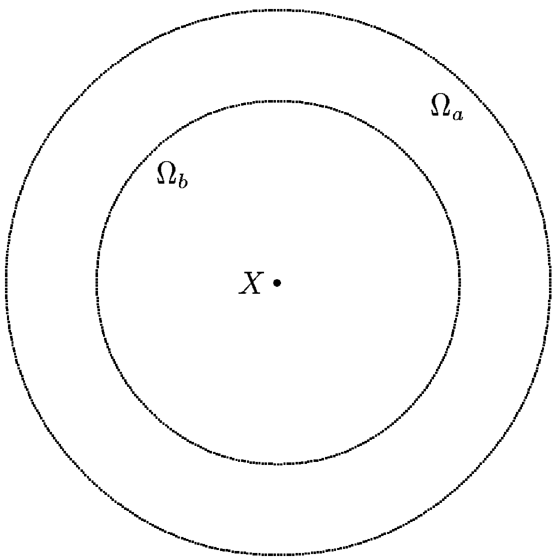
slijedi, budući da je $F(X)=0$, dok za $F(x_n)$ v. formulu (4):

$$\|F'(x_n)(x_{n+1}-X)\|_Y \le a_2\|x_n - X\|_X^2,$$

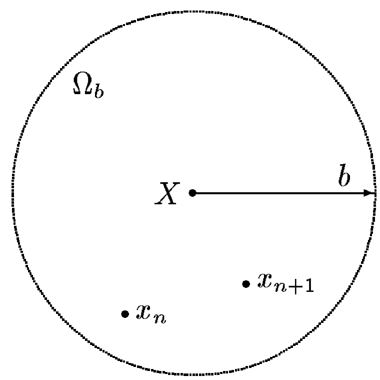
slijedi, pomoću (6):

$$\|x_{n+1}-X\|_X \le c\|x_n - X\|_X^2 \tag{9}$$

i dalje $< cb^2 = (cb)b \le b$. Dobili smo $\|x_{n+1}-X\|_X < b$, odnosno $x_{n+1} \in \Omega_b$. Dokazano je da svi elementi aproksimacionog niza $\{x_n\}$ pripadaju toj otvorenoj lopti. Ova okolnost je prikazana na slici 5. Drugo. Dokažimo formulu (8). Uvedimo oznaku $q_n = c\|x_n - X\|_X$. Iz (9) imamo da je $q_{n+1} \le q_n^2$ za svako n. Recimo, $q_1 \le q_0^2$, $q_2 \le q_1^2$, $q_3 \le q_2^2$ pa $q_3 \le q_0^8$. Dakle, $q_n \le q_0^{2^n}$, indukcijom. Dobili smo da je $c\|x_n - X\|_X \le (c\|x_0 - X\|_X)^{2^n}$, čime je (8) dokazano. I treće. Posmatrajmo nejednakost (8). Budući da je $c\|x_0 - X\|_X < 1$ to desna strana te nejednakosti $\to 0$ kad $n \to \infty$. Zato i njena lijeva strana $\to 0$. Dobili smo da je $\lim_{n\to\infty}\|x_n - X\|_X = 0$, odnosno da je $\lim_{n\to\infty} x_n = X$. Dokaz je završen.



Prostor X, Slika 4,
X – korijen



Slika 5

Slijede dopune.

1. Važno je istaći da se u prethodnoj teoremi pretpostavlja da rješenje X jednačine $F(x)=0$ postoji.

2. Lako se vidi da je rješenje jednačine $F(x)=0$ u lopti Ω_b jedinstveno. Zaista, dopustimo da pored X postoji i neko drugo rješenje Y (važi $F(Y)=0$). Izaberimo $x_0 = Y$. Onda po (5) izlazi da je i $x_1 = Y$, $x_2 = Y$, ..., a u teoremi je dokazano da mora biti $\lim_{n\to\infty} x_n = X$.

5. NUMERIČKE METODE ZA RJEŠAVANJE KOŠIJEVOG ZADATKA ZA OBIČNE DIFERENCIJALNE JEDNAČINE

5.1. UVOD O KOŠIJEVOM ZADATKU I LEMA O DVA RJEŠENJA DIFERENCIJALNE JEDNAČINE

Da se opiše značaj numeričkih metoda za rješavanje običnih diferencijalnih jednačina dovoljne su sljedeće dvije rečenice. Većina procesa koji se proučavaju u fizici i uopšte prirodnim naukama opisuje se parcijalnim diferencijalnim jednačinama, a u jednostavnijim slučajevima – običnim diferencijalnim jednačinama. Veoma su rijetki slučajevi kada može da bude određeno tačno ili egzaktno (ili analitičko) rješenje diferencijalne jednačine.

U ovoj glavi razmatraju se numeričke metode za rješavanje početnog (Košijevog) zadatka za obične diferencijalne jednačine, pa formulišimo zadatak. Dati su diferencijalna jednačina $y' = f(x, y)$ i početni uslov $y(x_0) = y_0$, nepoznata funkcija ili rješenje zadatka $y = y(x)$ razmatra se na intervalu x–ose $[a, b] = [x_0, x_0 + X]$. Prilikom numeričkog rješavanja, pretpostavlja se da rješenje (da tačno rješenje) postoji, da je ono jedinstveno i da je ono jedna dovoljno glatka funkcija. Ako $f \in C^p$ onda $y \in C^{p+1}$.

Kroz svaku tačku oblasti prolazi samo po jedna integralna kriva diferencijalne jednačine.

Prilikom numeričkog rješavanja postavljenog zadatka, mi postavimo po x–osi mrežu čvorova $a = x_0 < x_1 < x_2 < \ldots < x_n = x_0 + X = b$. Označimo sa y_i približnu vrijednost za broj $y(x_i)$, tj. za vrijednost tačnog rješenja $y = y(x)$ u čvoru $x = x_i$. Brojne vrijednosti $\{y_i\}_{i=0}^{n}$ čine naš numerički odgovor, a razmatra se naravno i greška $R_i = y(x_i) - y_i$.

Numeričke metode za Košijev zadatak za o. d. j. dijele se u dvije klase i to: (a) jednokoračne i (b) višekoračne (k–koračne) ili diferencne. U slučaju (a), približna vrijednost y_i određuje se na osnovu približne vrijednosti y_{i-1} koja je u datom trenutku već određena. U slučaju (b), y_i se određuje po nekoliko ranijih približnih vrijednosti, recimo da se određuje po y_{i-4}, y_{i-3}, y_{i-2} i y_{i-1} (tada je $k = 4$). Primjer numeričke metode oblika (a) jeste metoda Runge–Kuta. Primjer numeričke metode oblika (b) jeste Adamsova metoda. Obično je mreža ekvidistantna, tj. $x_i = x_0 + ih$, $nh = b - a = X$.

Od numeričke metode se očekuje da njena greška teži ka nuli kada mreža čvorova na $[a, b]$ postaje sve gušća, odnosno kada broj čvorova neograničeno raste. Takođe se očekuje da greška teži ka nuli određenim tempom, odnosno teži ka nuli što brže kad $n \to \infty$, odnosno kad $h \to 0$. Razmotrimo jednu fiksiranu numeričku metodu. Neka je mreža čvorova ekvidistantna sa korakom $h > 0$. Definicija 1. Neka $x \in [a, b] = [x_0, x_0 + X]$. Neka $z_h(x)$ označava približnu vrijednost za $y(x)$ dobijenu sa korakom h i neka bude $R_h(x) = y(x) - z_h(x)$. Smatramo da je $nh = b - a$ za neki cio broj $n \geq 1$. Za numeričku metodu se kaže da konvergira u tački x ako važi $\lim_{h\to 0} R_h(x) = 0$. Definicija 2. Za numeričku metodu se kaže da konvergira na intervalu $[a, b]$ ako ona konvergira u svakoj tački tog intervala. Definicija 3. Za numeričku metodu se kaže da ima red konvergencije s u tački x ako važi $|R_h(x)| = O(h^s)$ kad $h \to 0$. Definicija 4. Za numeričku metodu se kaže da ima red konvergencije s na intervalu $[a, b]$ ako ona ima red konvergencije s ravnomjerno u svim tačkama tog intervala.

Završavajući uvod, metode koje ćemo raditi uopštavaju se neposredno ili vrlo lako na slučaj implicitno date jednačine, na slučaj sistema jednačina prvog reda, kao i na slučaj jednačine višeg reda.

Na redu je jedna lema iz teorije običnih diferencijalnih jednačina.

Lema. Neka je $f(x, y)$ neprekidna funkcija i neka je neprekidno diferencijabilna po promjenljivoj y. Neka su $Y_1(x)$ i $Y_2(x)$ dva rješenja diferencijalne jednačine $y' = f(x, y)$. Tada važi

$$Y_2(\beta) - Y_1(\beta) = (Y_2(\alpha) - Y_1(\alpha)) \exp\left\{\int_\alpha^\beta f'_y(x, \overline{y}(x))dx\right\}, \tag{$*$}$$

gdje je $\overline{y}(x)$ neki broj između $Y_1(x)$ i $Y_2(x)$.

Dokaz. Imamo $Y_1' = f(x, Y_1)$ i $Y_2' = f(x, Y_2)$, oduzimanjem $(Y_2 - Y_1)' = f(x, Y_2) - f(x, Y_1)$. Na $f(x, Y_2) - f(x, Y_1)$ primijenimo Lagranžovu teoremu (formulu o konačnim priraštajima), pa izlazi

$$f(x, Y_2) - f(x, Y_1) = f'_y(x, \overline{y})(Y_2 - Y_1),$$

gdje je $\overline{y}$ neki broj između $Y_1(x)$ i $Y_2(x)$. Uradimo ovo za razne x, pa tako imamo

$$(Y_2 - Y_1)' = f'_y(x, \overline{y}(x))(Y_2 - Y_1).$$

Posljednje predstavlja linearnu diferencijalnu jednačinu oblika $y' = P(x)y$ po nepoznatoj funkciji $Y_2 - Y_1$. Rješavanjem te jednačine dobija se $(*)$.

Zapaziti da mi ne možemo ništa reći o neprekidnosti ili eventualnoj glatkosti funkcije $\overline{y}(x)$. To nam nije ni potrebno jer je funkcija

$$f'_y(x, \overline{y}(x)) = \frac{f(x, Y_2(x)) - f(x, Y_1(x))}{Y_2(x) - Y_1(x)}$$

neprekidna. Zaista, u brojiocu i imeniocu su neprekidne funkcije, a još je imenilac različit od nule jer se dva rješenja $Y_1(x)$ i $Y_2(x)$ jedne te iste jednačine ne sijeku (jedinstvenost rješenja Košijevog zadatka). Zato su sve gore sprovedene transformacije – ispravne. Lema je dokazana.

Vodeći primjer za ovu lemu: razmotrimo diferencijalnu jednačinu $y' = ay$. Ovdje je očito $f(x, y) = ay$ i $f'_y(x, y) = a$. Opšte rješenje glasi $y(x) = Ce^{ax}$. Razmotrimo dva partikularna rješenja $Y_1(x) = C_1e^{ax}$ i $Y_2(x) = C_2e^{ax}$. Mi upoređujemo rastojanje po visini između ova dva rješenja u dvije tačke na x–osi. Rastojanje u tački $x = \beta$ može da bude znatno veće od rastojanja u ranijoj tački $x = \alpha$:

$$Y_2(\beta) - Y_1(\beta) = (Y_2(\alpha) - Y_1(\alpha))e^{a(\beta-\alpha)}.$$

Mi ćemo ubuduće formulu $(*)$ koristiti obično u nešto obrađenom obliku, kako slijedi. Treba pretpostaviti da funkcija $f = f(x, y)$ zadovoljava Lipšicov uslov po drugoj promjenljivoj y sa konstantom L, tj. da važi $|f(x, y_2) - f(x, y_1)| \leq L|y_2 - y_1|$. Ili treba pretpostaviti da važi $|\frac{\partial f}{\partial y}| \leq L$ (ekvivalentno je). Tada

$$\exp\left\{\int_\alpha^\beta f'_y(x, \overline{y}(x))dx\right\} \leq \exp\left|\int_\alpha^\beta f'_y(x, \overline{y}(x))dx\right| \leq$$

$$\exp\left\{\int_\alpha^\beta |f'_y(x, \overline{y}(x))|dx\right\} \leq \exp\left\{\int_\alpha^\beta L dx\right\} = \exp\{L(\beta - \alpha)\}$$

i

$$|Y_2(\beta) - Y_1(\beta)| \leq |Y_2(\alpha) - Y_1(\alpha)|e^{L(\beta-\alpha)}. \tag{$**$}$$

5.2. OJLEROVA METODA I DRUGI PRIMJERI

Primjer: $y' = x^2 + y^2$, $y(1) = 2$. Imamo da je $f(x_0, y_0) = x_0^2 + y_0^2 = 1 + 4 = 5$. Odavde je $y(x_0 + h) = y(1 + h) \approx y_0 + 5h = 2 + 5h$. Recimo, $y(1{,}01) \approx 2{,}05$.

Postavka zadatka. Razmotrimo Košijev (početni) zadatak $y' = f(x, y)$, $y(x_0) = y_0$ na intervalu $x_0 \le x \le x_0 + X$. Označimo njegovo tačno rješenje kao $y = y(x)$. Treba konstruisati numeričku metodu za dobijanje približnih vrijednosti analitičkog rješenja $y = y(x)$. Postavimo po intervalu $[x_0, x_0 + X]$ ravnomjernu mrežu čvorova čiji je korak h, gdje je $nh = X$. Dakle, stavimo $x_k = x_0 + kh$ za $0 \le k \le n$. Neka y_k označava približnu vrijednost za broj $y(x_k)$ ili svejedno neka y_k bude aproksimacija u čvoru $x = x_k = x_0 + kh$, gdje je $0 \le k \le n$. Veličine $\{y_k\}_{k=0}^n$ biće efektivno određene i one će i predstavljati numerički odgovor. Analiziraćemo i grešku $R_k = y(x_k) - y_k$ za $0 \le k \le n$.

Kako procijeniti $y(x_1)$? Napišimo Tejlorovu formulu za analitičko rješenje $y = y(x)$ vršeći razvoj oko tačke $x = x_0$. Tako

$$y(x_0 + h) = y(x_0) + hy'(x_0) + \frac{1}{2}h^2 y''(\alpha),$$

gdje je $x_0 < \alpha < x_0 + h$. Zanemarimo zadnji sabirak

$$y(x_0 + h) \approx y(x_0) + hy'(x_0), \qquad \text{tj.}$$

$$y(x_1) \approx y(x_0) + hy'(x_0) \qquad \text{ili svejedno} \qquad y_1 = y_0 + hy'(x_0).$$

U vezi date d. j.

$$y_1 = y_0 + hf(x_0, y_0).$$

Kako procijeniti $y(x_2)$? Na sličan način, mi stavljamo

$$y_2 = y_1 + hf(x_1, y_1).$$

Slično i dalje. Dakle, šema za računanje koja i definiše Ojlerovu metodu izražava se sljedećom formulom:

$$y_k = y_{k-1} + hf(x_{k-1}, y_{k-1}) \qquad \text{za} \ \ 1 \le k \le n.$$

Znamo da šema za računanje obuhvata formule pomoću kojih se može sastaviti program za računar (za saznavanje približnih vrijednosti y_k).

Mali primjer za Ojlerovu metodu: $y' = x + y$, $y(0) = 1$. Izaberimo $h = 0{,}1$ i $n = 4$, tako da je $x_1 = 0{,}1$ $x_2 = 0{,}2$ $x_3 = 0{,}3$ $x_4 = 0{,}4$. Tada se dobija numerički rezultat $y_1 = 1{,}1$ $y_2 = 1{,}22$ $y_3 = 1{,}362$ $y_4 = 1{,}5282$. S druge strane, odgovarajuće tačne vrijednosti glase $y(x_1) = 1{,}11034$ $y(x_2) = 1{,}24281$ $y(x_3) = 1{,}39972$ $y(x_4) = 1{,}58365$, budući da je analitičko rješenje $y(x) = 2e^x - x - 1$. Tako da pojedinačne greške $R_k = y(x_k) - y_k$ iznose redom: $R_1 = 0{,}01034$ $R_2 = 0{,}02281$ $R_3 = 0{,}03772$ $R_4 = 0.05545$.

Prelazimo na ocjenu greške. Prvo uvodimo pojam lokalne greške i vršimo njenu ocjenu. Lokalna greška ili greška na koraku definiše se u slučaju Ojlerove metode relacijom:

$$\rho = y(x + h) - [y(x) + hy'(x)], \qquad \text{tj.} \qquad \rho = y(x + h) - [y(x) + hf(x, y(x))].$$

Na primjer, greška metode na prvom koraku jednaka je

$$\rho_1 = y(x_1) - y_1 = y(x_1) - [y_0 + hy'(x_0)] = y(x_1) - [y_0 + hf(x_0, y_0)].$$

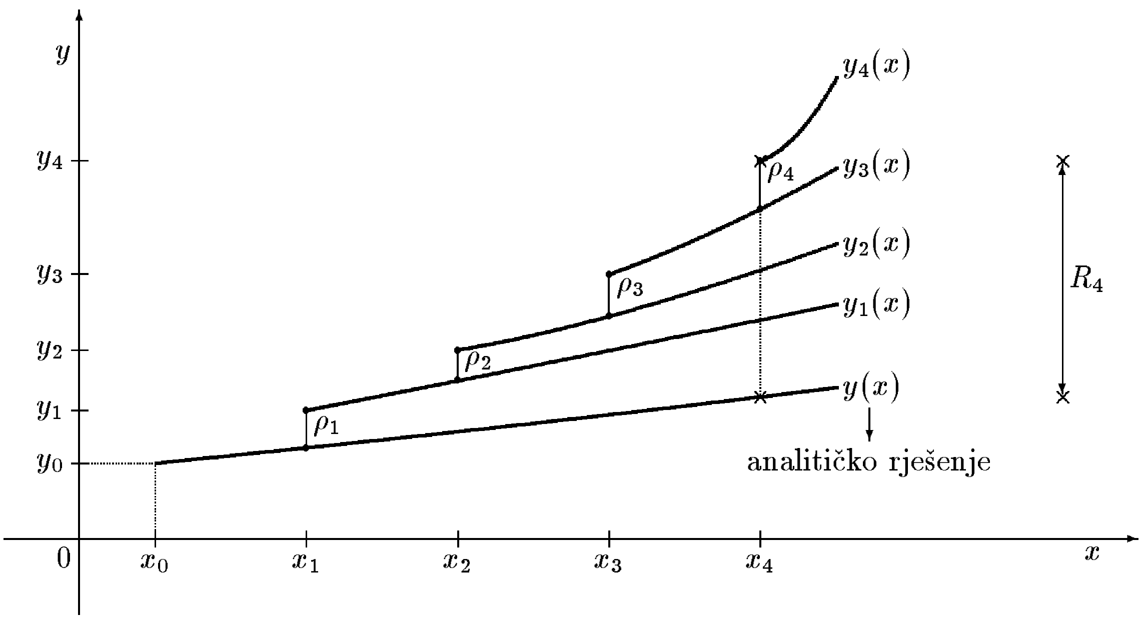
Zanemareni sabirak čini razliku:

$$y(x+h) = y(x) + hy'(x) + \frac{1}{2}h^2y''(\beta), \quad \text{gdje je } x < \beta < x+h \quad \Rightarrow \quad \rho = \frac{1}{2}h^2y''(\beta).$$

Uzmimo da je drugi izvod rješenja K. z. ograničen na intervalu $[x_0, x_0+X]$, tj. da je $|\frac{1}{2}y''(\beta)| \leq C$. Tada je očito $|\rho| \leq Ch^2$. Lokalna greška Ojlerove metode je reda h^2.

Prelazimo na ocjenu greške (globalne greške). Uvedimo potrebne oznake. Kao i dosad, y_k su približne vrijednosti za čvorove $x_k = x_0 + kh$ koje je računar saopštio. Tako da greška u pojedinom čvoru iznosi naravno $R_k = y(x_k) - y_k$ za $0 \leq k \leq n$. Kao i dosad, $y = y(x)$ je analitičko rješenje postavljenog K. z. Ocijenićemo grešku $R_n = y(x_n) - y_n$ u posljednjem čvoru $x_n = x_0 + nh = x_0 + X$.

Uvedimo u razmatranje pomoćne funkcije. Neka funkcija $y_k = y_k(x)$ zadovoljava postavljenu d. j. $y' = f(x,y)$ i neka zadovoljava jednakost $y_k(x_k) = y_k$ kao svoj početni uslov $(0 \leq k \leq n)$. Vidi se da je $y(x) = y_0(x)$. Kao što je maločas rađeno, lokalne greške definisane su relacijom $\rho_k = y_{k-1}(x_k) - y_k = y_{k-1}(x_k) - y_k(x_k)$.



Imamo: $\quad R_n = y(x_n) - y_n(x_n) = y_0(x_n) - y_n(x_n) = \sum_{k=1}^{n}(y_{k-1}(x_n) - y_k(x_n)),$

$$|R_n| \leq \sum_{k=1}^{n} |y_{k-1}(x_n) - y_k(x_n)| \leq \qquad \text{po } (**)$$

$$\sum_{k=1}^{n} |y_{k-1}(x_k) - y_k(x_k)| e^{L(x_n - x_k)} = \sum_{k=1}^{n} |\rho_k| e^{L(x_n - x_k)}.$$

Sada se mijenja definicija konstante C, ona se sada opštije definiše (i zato se ona povećava). Uzmimo da je $|\frac{1}{2}y_k''(x)| \leq C$ za sve $x \in [x_0, x_0+X]$ i $0 \leq k \leq n$. Više od toga (još opštije), stavimo da je C broj takav da važi $|\frac{1}{2}y''(x)| \leq C$ za sve $x \in [x_0, x_0 + X]$, gdje je $y = y(x)$ bilo koje rješenje d. j. $y' = f(x, y)$. Nastavljamo:

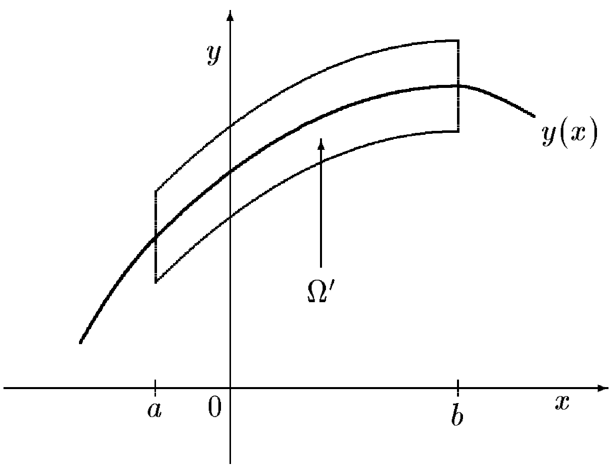
$$|R_n| \leq \sum_{k=1}^{n} Ch^2 e^{L(x_n - x_k)},$$

$$|R_n| \leq \sum_{k=1}^{n} Ch^2 e^{LX} = nCh^2 e^{LX} = CXhe^{LX} = ch$$

($c = \text{const}$). Ili $R_n = O(h)$. Dakle, greška Ojlerove metode je reda h. Zapaziti da R_n može znatno da prevazilazi $|\rho_1| + \ldots + |\rho_n|$. Ponovimo, pod uslovom da je $|\frac{\partial f}{\partial y}| \leq L$ za neko $L > 0$ i da rješenja $y(x)$ d. j. $y' = f(x, y)$ imaju ograničene druge izvode $y''(x)$. Ako $\frac{d}{dx}$ na $y' = f(x, y)$ onda $y'' = f'_x + f'_y y' = f'_x + f'_y f$. Recimo, $y' = x^2 + y^2 \Rightarrow y'' = 2x + 2y(x^2 + y^2)$.

Slika se odnosi na ocjenu greške (globalne greške), $R_4 = y(x_4) - y_4$.

O praktičnoj ocjeni greške Ojlerove metode. Vidjeli smo da se traži ograničenost funkcija y'' i f'_y na skupu $\Omega = [x_0, x_0 + X] \times R$. Realno posmatrano, dovoljna je njihova ograničenost na puno manjem skupu $\Omega' \subset R^2$. Skup Ω' predstavlja usku traku oko analitičkog rješenja $y = y(x)$, $\Omega' = \{(x, y) |\ a \leq x \leq b,\ y(x) - \varepsilon \leq y \leq y(x) + \varepsilon\}$, $[a, b] = [x_0, x_0 + X]$. Obrazloženje: približne vrijednosti y_k samo neznatno odstupaju od $y(x_k)$, $y(x_k) \sim y_k$. Jasno, $(x_k, y_k) \in \Omega'$, $0 \leq k \leq n$. Navedene okolnosti prikazane su na slici:



Kao praktični izraz za grešku služi $R_n = \sum_{k=1}^{n} \rho_k \exp \int_{x_k}^{x_n} f'_y(x, y(x))dx$, gdje treba uzeti $\rho_k = \frac{1}{2}h^2 y''(x_{k-1/2})$, a integral izračunati preko trapezne formule po čvorovima $f'_y(x_i, y_i)$, $k \leq i \leq n$.

Ako primijenimo praktični izraz za grešku na prethodni mali primjer onda dobijamo $R_1 = 0,01050$ $R_2 = 0,02315$ $R_3 = 0,03829$ $R_4 = 0,05630$, što je u dobroj harmoniji sa stvarnim vrijednostima $R_1 = 0{,}01034$ $R_2 = 0{,}02281$ $R_3 = 0{,}03772$ $R_4 = 0.05545$ (gdje piše $y = 2e^x - x - 1$).

Na redu su primjeri jednokoračnih metoda (šema).

Primjer 1: jedna implicitna šema.

Napišimo identitet

$$y(x + h) = y(x) + \int_0^h y'(x + t)dt. \tag{1}$$

Ako za računanje integrala u (1) upotrebimo osnovnu trapeznu formulu $\int_0^h F(t)dt = \frac{h}{2}(F(0) + F(h)) + O(h^3)$ onda

$$y(x + h) = y(x) + \frac{h}{2}\Big(y'(x) + y'(x + h)\Big) + O(h^3) \qquad \text{ili}$$

$$y(x+h) = y(x) + \frac{h}{2}\Big(f(x, y(x)) + f(x+h, y(x+h))\Big) + O(h^3). \tag{2}$$

Tako se dobija numerička formula ili šema za računanje

$$y_{j+1} = y_j + \frac{h}{2}\Big(f(x_j, y_j) + f(x_{j+1}, y_{j+1})\Big) \qquad \text{za } \ j \geq 0. \tag{3}$$

Jednakost (3) ustvari predstavlja jednu jednačinu po nepoznatoj y_{j+1}. Zato se za ovu šemu kaže da je implicitna (da je u neriješenom obliku). Ako se približna vrijednost koja se računa (oblika y_{j+1}) može saznati neposrednim računom onda se za šemu ili metodu kaže da je eksplicitna ili da je u riješenom obliku.

Može se tvrditi da postoji jedinstveno rješenje jednačine (3) jedino ako je h dovoljno malo (jedino za $h \leq h_0$).

Lokalna greška navedene metode iznosi $\rho = O(h^3)$, a njena greška je jednaka $R = O(h^2)$.

Primjer 2: poboljšana ili modifikovana Ojlerova metoda.

Pokušajmo da u prethodnoj metodi nešto promijenimo, da bi se implicitnost otklonila. Ako se na desnoj strani formule (2) broj $y(x+h)$ zamijeni nekim brojem y^* koji od $y(x+h)$ odstupa za $O(h^2)$ onda se ta desna strana izmijeni za $O(h^3)$. Kako je ostatak formule (2) bio $O(h^3)$ i kako se ovom zamjenom uvećava za $O(h^3)$, to će i lokalna greška metode sa y^* biti $O(h^3)$.

Kao y^* može da posluži približna vrijednost za $y(x+h)$ koju daje Ojlerova metoda.

Odavde, šema za računanje modifikovane Ojlerove metode glasi:

$$\begin{cases} y^*_{j+1} = y_j + hf(x_j, y_j), \\ y_{j+1} = y_j + \dfrac{h}{2}\left(f(x_j, y_j) + f(x_{j+1}, y^*_{j+1})\right), \qquad 0 \leq j \leq n-1. \end{cases}$$

Lokalna greška modifikovane Ojlerove metode iznosi $\rho = O(h^3)$, a njena greška je jednaka $R = O(h^2)$.

Primjer 3: Hojnova metoda.

Za računanje integrala u (1) upotrebimo formulu pravougaonika, čime se dobija $y(x+h) \approx y(x) + hf(x + \frac{h}{2}, y(x + \frac{h}{2}))$. Kao približna vrijednost za $y(x + \frac{h}{2})$ može da posluži vrijednost izračunata po Ojlerovoj formuli sa korakom $\frac{h}{2}$. Tako imamo metodu čija je šema za računanje definisana sljedećim parom jednakosti:

$$\begin{cases} y_{j+1/2} = y_j + \dfrac{h}{2}f(x_j, y_j) \\ y_{j+1} = y_j + hf(x_{j+1/2}, y_{j+1/2}) \end{cases}$$

Lako se vidi da lokalna greška Hojnove metode iznosi $\rho = O(h^3)$ i da je njena greška jednaka $R = O(h^2)$.

5.3. OPŠTI SLUČAJ EKSPLICITNE METODE TIPA RUNGE–KUTA

Metoda Runge–Kuta služi za rješavanje Košijevog zadatka $y' = f(x, y)$, $y(x) = y$. Vidimo da je Košijev uslov označen kao $y(x) = y$, umjesto uobičajenog zapisivanja $y(x_0) = y_0$.

Dovoljno je definisati kako se vrši jedan korak, tj. definisati kako se iz tačke x prelazi u tačku $x + h$. Dakle, postavlja se pitanje: kako naći približnu vrijednost u oznaci $z(h)$ za tačnu vrijednost $y(x+h)$. Naravno da se lokalna greška metode ili greška metode na koraku definiše kao $\rho = \rho(h) = y(x+h) - z(h)$.

Šablon oznaka:

po x–osi	tačno	približno	greška
x	$y = y(x)$	$y = y(x)$	0
$x + h$	$y(x + h)$	$z(h)$	$\rho(h)$

.

Evo opšte definicije metode razmatrane klase. U procesu računanja fiksirani su izvjesni brojevi $\alpha_2, \ldots, \alpha_q$, $p_1, \ldots, p_q$, β_{ij} $(1 \le j \le i-1,\ 2 \le i \le q)$. Redom se računaju veličine $k_1(h), \ldots, k_q(h)$, Δy, $z(h)$:

$$\begin{cases} k_1(h) = hf(x,y), \qquad k_2(h) = hf(x+\alpha_2 h, y + \beta_{21}k_1(h)), \qquad \ldots, \\ k_q(h) = hf(x+\alpha_q h, y+\beta_{q1}k_1(h)+\ldots+\beta_{q,q-1}k_{q-1}(h)), \\ \Delta y = \sum_{i=1}^{q} p_i k_i(h), \qquad z(h) = y + \Delta y \end{cases}$$

i zatim se stavlja $y(x+h) \approx z(h)$. Lako se zapaža da sve veličine $k_i(h)$ imaju smisao priraštaja. Zapaža se da mora biti $\sum_{i=1}^{q} p_i = 1$, ako se želi postići makar kakva aproksimacija. Kao što je već rečeno, greška metode na koraku definiše se kao $\rho(h) = y(x+h) - z(h) = y(x+h) - [y(x) + \Delta y]$.

Postavlja se pitanje izbora zasad neodređenih (slobodnih) brojnih vrijednosti α_i, p_i, β_{ij}. Te brojne vrijednosti biće izabrane tako da postane $\rho(0) = \rho'(0) = \ldots = \rho^{(s)}(0) = 0$. Jasno je da je bolje što je broj s veći. To i jeste kriterijum za određivanje parametara α_i, p_i, β_{ij}. Uzima se da broj s ne može da bude povećan, tj. da za neku glatku funkciju $f(x,y) = f_0(x,y)$ važi $\rho^{(s+1)}(0) \ne 0$. Za s se kaže da predstavlja red greške metode.

Možemo napisati Tejlorovu formulu:

$$\rho(h) = \sum_{i=0}^{s} \frac{\rho^{(i)}(0)}{i!}h^i + \frac{\rho^{(s+1)}(\theta h)}{(s+1)!}h^{s+1} \qquad \text{ili} \qquad \rho(h) = \frac{\rho^{(s+1)}(\theta h)}{(s+1)!}h^{s+1} \qquad (0 < \theta < 1).$$

Ispitajmo za prvih nekoliko q.

$\boxed{q = 1}$ $k_1(h) = hf(x,y)$, $\Delta y = p_1 k_1(h)$, $z(h) = y + \Delta y$, $y(x+h) \approx z(h)$,

$$\rho(h) = y(x+h) - y(x) - p_1 hf(x,y)$$

$$\rho'(h) = y'(x+h) - p_1 f(x,y)$$

$$\rho'(0) = y'(x) - p_1 f(x,y) = f(x,y) - p_1 f(x,y) = (1-p_1)f(x,y).$$

Ako želimo da postignemo $\rho'(0) = 0$ onda mora biti $p_1 = 1$. Ovim je jedna konkretna metoda tipa Runge–Kuta definisana. Vidimo da smo ustvari dobili Ojlerovu metodu.

$\boxed{q = 2}$ Opet ćemo izračunati $\rho'(h), \rho''(h), \ldots$ i opet ćemo nastojati da postignemo $\rho'(0) = 0$, $\rho''(0) = 0, \ldots$.

Po opštem šablonu: $k_1(h) = hf(x,y)$, $k_2(h) = hf(x+\alpha_2 h, y+\beta_{21}k_1(h))$, $\Delta y = p_1 k_1(h) + p_2 k_2(h)$, $z(h) = y + \Delta y$, $y(x+h) \approx z(h)$,

$$\rho(h) = y(x+h) - y(x) - p_1 k_1(h) - p_2 k_2(h) = y(x+h) - y(x) - p_1 hf(x,y) - p_2 hf(\overline{x}, \overline{y})$$

$$\text{skraćenice:} \qquad \overline{x} = x + \alpha_2 h, \qquad \overline{y} = y + \beta_{21}k_1(h)$$

$$\rho'(h) = y'(x+h) - p_1 f(x,y) - p_2 f(\overline{x}, \overline{y}) - p_2 h\Big(\alpha_2 f'_x(\overline{x}, \overline{y}) + \beta_{21} f'_y(\overline{x}, \overline{y}) f(x,y)\Big)$$

$$\rho'(0) = (1 - p_1 - p_2) f(x, y)$$

$$\begin{cases} y' = f(x,y) \\ y'' = f'_x + f'_y y' = f'_x + f'_y f \\ y''' = f''_{xx} + 2f''_{xy} f + f''_{yy} f^2 + f'_y y'' = \qquad \text{(izrazite } y'' \text{ po prethodnoj formuli)} \\ y^{IV} = \dots \qquad \text{itd.} \end{cases}$$

$$\rho''(h) = y''(x+h) - 2p_2 \Big[\alpha_2 f'_x(\overline{x}, \overline{y}) + \beta_{21} f'_y(\overline{x}, \overline{y}) f(x,y)\Big] - p_2 h \Big[\alpha_2^2 f''_{xx}(\overline{x}, \overline{y}) +$$

$$2\alpha_2 \beta_{21} f''_{xy}(\overline{x}, \overline{y}) f(x,y) + \beta_{21}^2 f''_{yy}(\overline{x}, \overline{y}) (f(x,y))^2\Big]$$

$$\rho''(0) = (1 - 2p_2\alpha_2) f'_x(x,y) + (1 - 2p_2\beta_{21}) f'_y(x,y) f(x,y)$$

$$\rho'''(h) = \dots, \qquad \rho'''(0) = \dots.$$

Ako želimo da bude $\rho'(0) = 0$ onda mora biti $\underline{1 - p_1 - p_2 = 0}$. Ako želimo da bude $\rho''(0) = 0$ onda mora biti $\underline{1 - 2p_2\alpha_2 = 0}$ i $\underline{1 - 2p_2\beta_{21} = 0}$. Vidimo da imamo tri uslova, a četiri parametra α_2, p_1, p_2, β_{21}. Jedan parametar se bira proizvoljno. Navedimo dvije mogućnosti.

(1) Izaberimo $p_1 = \frac{1}{2}$. Onda izlazi $p_2 = \frac{1}{2}$, $\alpha_2 = 1$, $\beta_{21} = 1$. Vidimo da je ovo modifikovana Ojlerova metoda.

Mi pišemo: $k_1 = hf(x_i, y_i)$, $k_2 = hf(x_i + h, y_i + k_1)$, $y_{i+1} = y_i + \frac{1}{2}(k_1 + k_2)$ (RK2 metoda) ili $k_1^{(i)} = hf(x_i, y_i)$, $k_2^{(i)} = hf(x_i + h, y_i + k_1^{(i)})$, $y_{i+1} = y_i + \frac{1}{2}(k_1^{(i)} + k_2^{(i)})$.

(2) Ako se izabere $p_1 = 0$ onda izlazi $p_2 = 1$, $\alpha_2 = \frac{1}{2}$, $\beta_{21} = \frac{1}{2}$. Vidimo da je ovo posljednja metoda iz prethodne sekcije 5.2. gdje piše Primjer 3. Šema: $k_1 = hf(x_i, y_i)$, $k_2 = hf(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1)$, $y_{i+1} = y_i + k_2$.

Kada kažemo $\rho'(0) = \rho''(0) = 0$ onda mislimo: $\rho'(0) = \rho''(0) = 0$ za svaku d. j. (za svaku funkciju f), a takođe i za bilo koju tačku (x, y) (ima ulogu tačke (x_0, y_0)).

Razmotrimo diferencijalnu jednačinu $y' = y$. Neposredni račun pokazuje da je $\rho'''(0) = y$, nezavisno od izbora parametara α_2, p_1, p_2, β_{21}. To znači da se u slučaju $q = 2$ ne može pronaći formula Runge–Kuta čiji bi stepen tačnosti bio $s = 3$.

$\boxed{q = 3}$ Razmatra se slično dosadašnjem. Može se dokazati da se najviše može postići $s = 3$. Drugim riječima, ne postoji šema za koju bi bilo $s = 4$ (kada je $q = 3$). Evo jedne formule čiji je red tačnosti $s = 3$:

$$\begin{cases} k_1 = hf(x,y), \qquad k_2 = hf\left(x + \dfrac{h}{2}, y + \dfrac{k_1}{2}\right), \qquad k_3 = hf(x + h, y - k_1 + 2k_2), \\ \Delta y = \dfrac{1}{6}\Big(k_1 + 4k_2 + k_3\Big), \qquad z(h) = y + \Delta y, \end{cases}$$

stavlja se $y(x + h) \approx z(h)$.

$\boxed{q = 4}$ Najviše se može postići $s = 4$. Evo jedne šeme čiji je red tačnosti $s = 4$:

$$\begin{cases} k_1 = hf(x,y), \qquad k_2 = hf\left(x + \dfrac{h}{2}, y + \dfrac{k_1}{2}\right), \qquad k_3 = hf\left(x + \dfrac{h}{2}, y + \dfrac{k_2}{2}\right), \\ k_4 = hf(x + h, y + k_3), \qquad \Delta y = \dfrac{1}{6}\Big(k_1 + 2k_2 + 2k_3 + k_4\Big), \qquad z(h) = y + \Delta y. \end{cases}$$

Predložena šema se često koristi u praksi. Ponekad se za nju prosto kaže da je "metoda Runge–Kuta IV reda". Lako se mogu napisati kompletne formule, odnosno formule pomoću kojih se onda sastavi program za računar, kako slijedi. Razmotrimo Košijev zadatak $y' = f(x,y)$, $x_0 \le x \le x_0 + X$, $y(x_0) = y_0$. Neka je $n \ge 1$ i $h = X/n$. Koriste se obične oznake za čvorove $x_i = x_0 + ih$ i približne vrijednosti u čvorovima y_i: RK4:

$$k_1 = hf(x_i, y_i), \qquad k_2 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}\right), \qquad k_3 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}\right),$$
$$k_4 = hf(x_i + h, y_i + k_3), \qquad y_{i+1} = y_i + \frac{1}{6}\Big(k_1 + 2k_2 + 2k_3 + k_4\Big), \qquad 0 \le i \le n-1$$

($k_1, \ldots, k_4$ zavise od i).

Završeno je ispitivanje za razne q. Napominje se da se u praksi koriste samo metode u kojima je $q \le 6$.

Sada ćemo navesti uslove pod kojima je funkcija greške $\rho = \rho(h)$ dovoljno glatka i ograničena na skupu $\Omega = [x_0, x_0 + X] \times R \subset R^2$. Pretpostavimo da $f \in C^{s+2}(\Omega)$ i $f \in B^{s+2}(\Omega)$; c – continuous – neprekidan, b – bounded – ograničen. Zapis $f \in B^{s+2}(\Omega)$ znači da su funkcija f i svi njeni parcijalni izvodi do reda $s+2$ uključeno ograničeni na skupu Ω. Glatkost funkcije f prenosi se na funkcije $k_1(h), \ldots, k_q(h)$ pa se prenosi i na funkciju $\rho = \rho(h)$. Neposredno se vidi da je funkcija $\rho = \rho(h)$ neprekidno diferencijabilna $s+1$ puta. Ograničenost funkcije f i njenih parcijalnih izvoda prenosi se na funkciju $\rho = \rho(h)$. Tako da je $\rho(h) = O(h^{s+1})$, odnosno $|\rho(h)| \le Ch^{s+1}$, gdje je C konstanta. Na primjer: Ojlerova metoda: $\rho(h) = \frac{1}{2!}y''(x + \theta h)h^2$, gdje je $y''(x) = f'_x(x,y) + f'_y(x,y)f(x,y)$, $|y''(x)| \le |f'_x(x,y)| + |f'_y(x,y)| \cdot |f(x,y)|$.

Posljednje, sada ćemo izdvojiti glavni sabirak greške $\rho(h)$. Pretpostavimo da $\rho \in C^{s+2}(\Omega)$ i $\rho \in B^{s+2}(\Omega)$. Tada važi Tejlorov razvoj:

$$\rho(h) = \frac{\rho^{(s+1)}(0)}{(s+1)!}h^{s+1} + \frac{\rho^{(s+2)}(\theta h)}{(s+2)!}h^{s+2},$$

gdje je $0 < \theta < 1$ (različito od ranijeg θ). Slijedi

$$\rho(h) = c(x,y)h^{s+1} + O(h^{s+2}),$$

gdje $c(x,y)$ ne zavisi od h. Glavni sabirak greške je $c(x,y)h^{s+1}$. Zapaža se da $c = c(x,y)$ neprekidno zavisi od (x,y). Primjer: Ojlerova metoda: $\rho(h) = \frac{1}{2!}y''(x)h^2 + \frac{1}{3!}y'''(x + \theta h)h^3$.

5.4. OCJENA GREŠKE ZA METODU RUNGE–KUTA

Ocjena greške metode jedne bilo koje fiksirane metode tipa Runge–Kuta doslovno se poklapa sa ocjenom greške Ojlerove metode sprovedenom u naslovu 5.2. Samo što je kod Ojlerove lokalna greška ρ bila reda h^2, a sada će biti reda h^{s+1}, tako da je tamo greška metode R_n u čvoru $x = x_n = x_0 + nh = x_0 + X$ izašla reda h, a sada će izaći red h^s. Međutim, sada će biti izvršena potpuna analiza greške, odnosno biće uzete u obzir sve tri komponente greške R_n. Drugim riječima, pored greške metode biće uzeta u obzir i greška računanja, kao i greška izazvana približnošću (greškom) ulaznog podatka y_0.

Uvedimo oznake i pretpostavke. Dopustimo da mreža čvorova nije ekvidistantna; neka bude $x_0 < x_1 < \ldots < x_n = x_0 + X$ i $x_{i+1} - x_i = h_i$. Neka se približna vrijednost y_{j+1} računa na osnovu prethodne y_j po formuli oblika $y_{j+1} = \Phi(f, x_j, h_j, y_j)$. Uzmimo da veličinu Φ nismo u stanju da izračunamo potpuno tačno, već da je u stvarnosti računamo sa izvjesnim zaokruživanjem. Zato prestaje da važi $y_{j+1} = \Phi(f, x_j, h_j, y_j)$, već uzimamo da važi $y_{j+1} = \Phi(f, x_j, h_j, y_j) - \delta_{j+1}$. Za δ_{j+1} se kaže da je greška računanja na koraku.

Recimo $y(x_0) = \pi$, $y_0(x_0) = y_0 = 3{,}14$, $R_0 = \pi - 3{,}14$, $|R_0| \le 0{,}002$.

Mi numerički rješavamo Košijev zadatak

$$y' = f(x,y), \qquad x_0 \le x \le x_0 + X, \qquad y(x_0) = y(x_0),$$

čije tačno rješenje je funkcija $y = y(x)$. Može se desiti da je ulazni podatak ili ulazna veličina (to je jedna brojna vrijednost) $y(x_0)$ poznata samo približno. Uzmimo da imamo takve okolnosti i neka je y_0 odgovarajući približni broj. Neka je $R_0 = y(x_0) - y_0 = y(x_0) - y_0(x_0)$. Za R_0 se kaže da predstavlja grešku ulaznih podataka ili ulaznog podatka. I R_0 će se naravno odraziti na rezultujuću ukupnu grešku. Već je upotrebljena oznaka za funkciju $y_0 = y_0(x)$; ona po definiciji zadovoljava $y_0' = f(x, y_0)$, $y_0(x_0) = y_0$; zadovoljava d.j. $y' = f(x,y)$ i zadovoljava približni početni uslov. Oko oznaka zapažamo: u ovoj sekciji y_0 znači približni početni uslov (u ranijim sekcijama značilo je tačni početni uslov), a tačni početni uslov u ovoj sekciji označen je kao $y(x_0)$.

Uvedimo u razmatranje još n rješenja naše d. j. $y' = f(x,y)$. Za $1 \le j \le n$, neka je $y_j = y_j(x)$ funkcija koja zadovoljava razmatranu d.j. i zadovoljava početni uslov u tački $x = x_j$ po približnom broju y_j:

$$y_j' = f(x, y_j), \qquad y_j(x_j) = y_j;$$

$y_j'(x) = f(x, y_j(x))$.

Sa ρ_j ćemo označavati grešku metode na koraku: $\rho_j = y_{j-1}(x_j) - \Phi(f, x_{j-1}, h_{j-1}, y_{j-1})$. Pretpostavićemo da je greška metode na koraku (greška metode na koraku ove fiksirane metode tipa Runge–Kuta čija se ukupna greška razmatra) reda h^{s+1}, odnosno da je $|\rho_j| \le C h_{j-1}^{s+1}$ za svako j.

Za funkciju $\dfrac{\partial f}{\partial y}$ ili svejedno $f_y'(x,y)$ pretpostavićemo da je neprekidna na skupu $\Omega = [x_0, x_0 + X] \times (-\infty, +\infty)$ i da je ograničena na tom skupu, da je za neko $L > 0$ ispunjeno $\left|\dfrac{\partial f}{\partial y}\right| \le L$ za $(x,y) \in \Omega$.

Nas interesuje greška $R_n = y(x_n) - y_n = y(x_n) - y_n(x_n)$ u tački $x_n = x_0 + X$ (u krajnjoj desnoj tački razmatranog intervala).

Teorema. Neka je $|\rho_j| \le C h_{j-1}^{s+1}$ i neka je $\left|\dfrac{\partial f}{\partial y}\right| \le L$ na skupu Ω. Uvedimo oznake $H = \max_{1\le j\le n} h_{j-1}$ i $\delta = \max_{1 \le j \le n} |\delta_j|$. Tada važi nejednakost

$$|R_n| \le e^{LX}\Big(CXH^s + n\delta + |R_0|\Big). \tag{1}$$

Dokaz:

$$R_n = y(x_n) - y_n = y(x_n) - y_n(x_n) = y(x_n) - y_0(x_n) + y_0(x_n) - y_n(x_n) =$$

$$y(x_n) - y_0(x_n) + \sum_{j=1}^{n}\Big(y_{j-1}(x_n) - y_j(x_n)\Big),$$

$$|R_n| \le \underbrace{|y(x_n) - y_0(x_n)|}_{\text{prvi izraz}} + \underbrace{\sum_{j=1}^{n} |y_{j-1}(x_n) - y_j(x_n)|}_{\text{drugi izraz}}.$$

Šablon oznaka: $y_k(x_k) = y_k$ i

po x–osi	približno	tačno	greška
x_0	y_0	$y(x_0)$	$R_0 = y(x_0) - y_0$
x_1	y_1	$y(x_1)$	$R_1 = y(x_1) - y_1$
x_2	y_2	$y(x_2)$	$R_2 = y(x_2) - y_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_n = x_0 + X$	y_n	$y(x_n)$	$R_n = y(x_n) - y_n$
(od a do b)	(računar)	(analitičko)	(razlika)

Za pojedini sabirak u drugom izrazu primijenimo lemu o dva rješenja sa $\alpha = x_j$, $\beta = x_n$, $Y_1(x) = y_{j-1}(x)$, $Y_2(x) = y_j(x)$:

$$y_{j-1}(x_n) - y_j(x_n) = (y_{j-1}(x_j) - y_j(x_j)) \exp\Big\{\int_{x_j}^{x_n} f'_y(x, \overline{y}_j(x))dx\Big\} =$$

$$(y_{j-1}(x_j) - \Phi(f, x_{j-1}, h_{j-1}, y_{j-1}) + \Phi(f, x_{j-1}, h_{j-1}, y_{j-1}) - y_j(x_j)) \exp\Big\{\int_{x_j}^{x_n} f'_y(x, \overline{y}_j(x))dx\Big\} =$$

$$(\rho_j + \delta_j) \exp\Big\{\int_{x_j}^{x_n} f'_y(x, \overline{y}_j(x))dx\Big\},$$

$$|y_{j-1}(x_n) - y_j(x_n)| \le (|\rho_j| + |\delta_j|) \exp\Big\{\int_{x_j}^{x_n} |f'_y(x, \overline{y}_j(x))|dx\Big\} \le$$

$$(Ch_{j-1}^{s+1} + \delta) \exp\Big\{\int_{x_j}^{x_n} L dx\Big\} \le (Ch_{j-1}^{s+1} + \delta) \exp\Big\{LX\Big\}.$$

Sabiranjem:

$$\sum_{j-1}^{n} |y_{j-1}(x_n) - y_j(x_n)| \le \Big(\sum_{j=1}^{n} Ch_{j-1}^{s+1} + n\delta\Big) \exp\Big\{LX\Big\} \le \Big(CXH^s + n\delta\Big) \exp\Big\{LX\Big\}$$

jer je

$$h_0^{s+1} + h_1^{s+1} + \ldots + h_{n-1}^{s+1} = h_0^s h_0 + h_1^s h_1 + \ldots + h_{n-1}^s h_{n-1} \le$$

$$H^s h_0 + H^s h_1 + \ldots + H^s h_{n-1} = H^s X.$$

Slično za prvi izraz, takođe pomoću leme:

$$|y(x_n) - y_0(x_n)| \le |y(x_0) - y_0(x_0)| \exp\Big\{\int_{x_0}^{x_n} |f'_y(x, \overline{y}_0(x))|dx\Big\} \le |R_0| \exp\Big\{LX\Big\}.$$

Na kraju, prvi izraz plus drugi izraz:

$$|R_n| \le \Big(CXH^s + n\delta + |R_0|\Big)\exp\Big\{LX\Big\}.$$

Dokaz je završen.

Pogledajmo završnu formulu (1) koja izražava totalnu ili sveukupnu grešku R_n u krajnjoj desnoj tački intervala od x_0 do $x_0 + X$. Vidimo da greška $R_n \to 0$ ako $H \to 0$ uz istovremeno $n\delta \to 0$ i $R_0 \to 0$; numerička metoda konvergira. Veličina $n\delta$ predstavlja zbirnu grešku računanja, a R_0 je greška ulaznih podataka. Ako se razmatra samo greška metode onda važi nejednakost $|R_n| \le e^{LX}CXH^s$; $|R_n| \le ch^s$ ($c = \text{const}$) ako je mreža ravnomjerna, red konvergencije je h^s ako su lokalne greške reda h^{s+1}.

5.5. ALGORITAM ZASNOVAN NA METODI RUNGE–KUTA

Samo ćemo naznačiti elemente koji čine algoritam. Neka smo se opredijelili za jednu određenu metodu. Označimo sa s red greške metode; obično je $s = 4$. Uzimamo da je prisutna jedino greška metode. Koristićemo formulu za grešku metode:

$$R_h(x_0 + X) = \sum_{j=1}^{n} \rho_j \exp\Big\{\int_{x_j}^{x_n} f'_y(x, \overline{y}_j(x))dx\Big\}, \tag{1}$$

gdje je $X = nh$, v. u 5.4. Koristićemo i formulu u kojoj je izdvojen glavni sabirak lokalne greške:

$$\rho(h) = ch^{s+1} + O(h^{s+2}), \tag{2}$$

gdje je $c = c(x, y)$ neprekidna funkcija, v. na kraju 5.3.

1. Jedan postupak da se ocijeni greška.

Neka su $z_h(x_0 + X)$ i $z_{2h}(x_0 + X)$ dvije približne vrijednosti za $y(x_0 + X)$ dobijene redom sa korakom h odnosno $2h$. Neka su $R_h(x_0 + X)$ i $R_{2h}(x_0 + X)$ odgovarajuće greške:

$$y_(x_0 + X) = z_h(x_0 + X) + R_h(x_0 + X) \qquad \text{i} \qquad y(x_0 + X) = z_{2h}(x_0 + X) + R_{2h}(x_0 + X).$$

Iz (1) i (2) lako se može izvesti sljedeća formula koja izražava vezu između jedne i druge greške:

$$\lim_{h \to 0} \frac{R_{2h}(x_0 + X)}{R_h(x_0 + X)} = 2^s. \tag{3}$$

Zapaziti da je $\lim_{h\to 0} \overline{y}_j(x) = y(x)$, gdje je $y = y(x)$ analitičko rješenje K. z. Isto tako, $\rho(2h) \sim 2^{s+1}\rho(h)$ kad $h \to 0$.

Iz (3) se vrlo lako izvodi sljedeća formula:

$$R_h(x_0 + X) \sim \frac{z_h(x_0 + X) - z_{2h}(x_0 + X)}{2^s - 1} \qquad \text{kad } h \to 0. \tag{4}$$

Ona se izvodi na običan Rungeov način. U numeričkoj praksi, formula (4) koristi se u obliku

$$R_h(x_0 + X) \approx \frac{z_h(x_0 + X) - z_{2h}(x_0 + X)}{2^s - 1},$$

$$\text{RK4:} \quad R_h(x_0 + X) \approx \frac{z_h(x_0 + X) - z_{2h}(x_0 + X)}{15}.$$

Očito je da posljednja napisana formula služi da se dobije procjena za grešku približnog broja $z_h(x_0+X)$. Bitno je što je procjena efektivna (mi smo u mogućnosti da izračunamo izraz na desnoj strani).

2. Drugi postupak da se ocijeni greška.

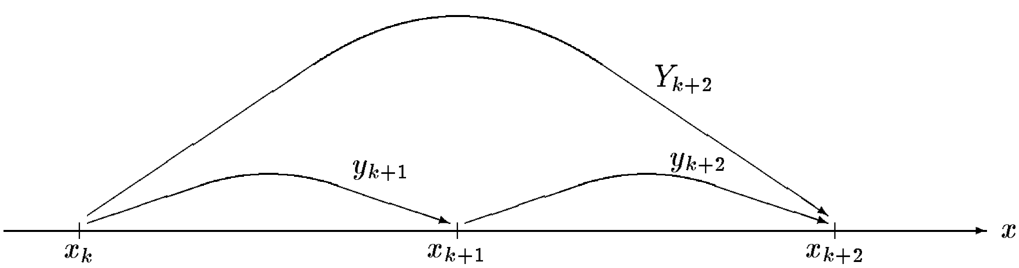
Integral u formuli (1) izračunati po trapeznoj formuli preko $f'_y(x_i,y_i)$, $j \le i \le n$.

Formula (1) postaće praktično upotrebljiva kada se nađu efektivne procjene ρ_j.

Neka je $h>0$ i neka je $X=nh$. Prvo se izvrši glavni proračun. Glavni proračun izvrši se čitav sa korakom h. Dakle, neka y_k označava približnu vrijednost koja se odnosi na čvor $x_k=x_0+kh$ za $1\le k\le n$. Ponovimo: računar je saopštio brojeve $\{y_k\}_{k=1}^n$, oni predstavljaju numerički odgovor i pitamo se kolika je njihova greška. Može se reći da je broj y_{k+1} dobijen na osnovu (x_k,y_k) sa korakom h. Može se reći da je broj y_{k+2} dobijen na osnovu (x_{k+1},y_{k+1}) sa korakom h. V. sliku. U cilju dobijanja efektivne ocjene greške numeričkog odgovora y_n, sada se izvrši i pomoćni proračun, kako slijedi. Na osnovu (x_k,y_k) sa korakom $2h$ izračunati i približnu vrijednost Y_{k+2} koja se (znači) odnosi na čvor x_{k+2}; izračunati za svako parno k. Približna vrijednost Y_{k+2} ima pomoćnu ulogu, ona služi da se procijene lokalne greške ρ_{k+1} i ρ_{k+2} koje se tiču glavnog proračuna, one se odnose na y_{k+1} i y_{k+2}. Lako se vidi da je $\rho_{k+2}\sim\rho_{k+1}$ kad $h\to 0$ (tj. da je $\lim_{h\to 0}\rho_{k+2}/\rho_{k+1}=1$), što se u numeričkoj praksi koristi u obliku $\rho_{k+2}\approx\rho_{k+1}$ za male h. Upotrebom temeljnih formula (1) i (2) dosta lako se može izvesti sljedeća relacija:

$$\rho_{k+1}+\rho_{k+2}\approx\frac{1}{2^s-1}\Big(y_{k+2}-Y_{k+2}\Big), \tag{5}$$

$$\text{RK4:}\quad \rho_{k+1}+\rho_{k+2}\approx\frac{1}{15}\Big(y_{k+2}-Y_{k+2}\Big).$$



3. Podešavanje koraka. Tokom napredovanja po x–osi, treba smanjiti tekući korak h ako lokalna greška postane prevelika.

Ponavljanje o jednokoračnim metodama: RK4: $K_1=hf(x_i,y_i)$, $K_2=hf(x_i+\frac{h}{2},y_i+\frac{K_1}{2})$, $K_3=hf(x_i+\frac{h}{2},y_i+\frac{K_2}{2})$, $K_4=hf(x_i+h,y_i+K_3)$, $y_{i+1}=y_i+\frac{1}{6}(K_1+2K_2+2K_3+K_4)$.

5.6. DIFERENCNE METODE

1. Pojam Adamsove metode i njene lokalne greške

Neka je $p_{k-1}(x)$ L. i. p. funkcije $F(x) = y'(x)$ po mreži čvorova $x_{n-i} = x_n - ih$, gdje je $1 \le i \le k$. Napišimo eksplicitni izraz za $p_{k-1}(x)$ preko $y'(x_{n-k})$, ..., $y'(x_{n-1})$:

$$p_{k-1}(x) = \sum_{i=n-k}^{n-1} \ell_i(x) y'(x_i), \qquad \ell_i(x) = \prod_{\substack{j=n-k \\ j\neq i}}^{n-1} \frac{x - x_j}{x_i - x_j}.$$

Po izrazu za grešku interpolacije:

$$F(x) - p_{k-1}(x) = y'(x) - p_{k-1}(x) = r_{k-1}(x) =$$

$$\frac{1}{k!}\omega_k(x) F^{(k)}(\xi(x)) = \frac{1}{k!}\omega_k(x) y^{(k+1)}(\xi(x)), \qquad \omega_k(x) = \prod_{i=n-k}^{n-1} (x - x_i),$$

ako $y \in C^{k+1}[x_0, x_0 + X]$; $x_{n-k} < \xi(x) < x_n$. Na jednakost $y'(x) = p_{k-1}(x) + r_{k-1}(x)$ primijenite $\int_{x_{n-1}}^{x_n} dx$:

$$y(x_n) - y(x_{n-1}) = \int_{x_{n-1}}^{x_n} p_{k-1}(x)dx + \int_{x_{n-1}}^{x_n} r_{k-1}(x)dx,$$

$$y(x_n) = y(x_{n-1}) + \int_{x_{n-1}}^{x_n} p_{k-1}(x)dx + \rho_n \qquad \text{ili} \qquad y(x_n) \approx y(x_{n-1}) + \int_{x_{n-1}}^{x_n} p_{k-1}(x)dx.$$

Za ρ_n se kaže da je lokalna greška metode ili da je greška metode na koraku. Izračunajte $\int_{x_{n-1}}^{x_n} p_{k-1}(x)dx$. Račun pokazuje da je $\int_{x_{n-1}}^{x_n} p_{k-1}(x)dx$ jedna linearna kombinacija vrijednosti $y'(x_{n-k})$, ..., $y'(x_{n-1})$:

$$\int_{x_{n-1}}^{x_n} p_{k-1}(x)dx = h\Big(b_{-k} y'(x_{n-k}) + \ldots + b_{-1} y'(x_{n-1})\Big),$$

gdje su b_{-k}, ..., b_{-1} konstante. Sprovedite račun u slučajevima $k = 1, \ldots, 4$.

Razmotrimo K. z.

$$y' = f(x, y), \qquad x_0 \le x \le x_0 + X, \qquad y(x_0) = y_0$$

i razmotrimo mogućnost njegovog numeričkog rješavanja po ravnomjernoj mreži čvorova čiji je korak $h = X/N$, tj. po mreži čvorova $x_n = x_0 + nh$, gdje je $0 \le n \le N$. Označimo sa y_n približne vrijednosti u čvorovima za $0 \le n \le N$. Iz prethodnog izvođenja vidimo da smo ustvari konstruisali jednu numeričku metodu za rješavanje postavljenog K. z. Ta metoda izražava se sljedećom šemom za računanje koja ide u računar:

$$y_n = y_{n-1} + h\Big(b_{-k} f(x_{n-k}, y_{n-k}) + \ldots + b_{-1} f(x_{n-1}, y_{n-1})\Big), \qquad k \le n \le N;$$

ovo je eksplicitna (e.) ili ekstrapolaciona Adamsova formula. Takođe se kaže i Adams–Bašfortova metoda. Vidi se da veličine b_{-k}, ..., b_{-1} zavise od k.

U nastavku su date eksplicitne formule u slučajevima $k = 1, \ldots, 4$. Koristi se obična skraćenica $f_i = f(x_i, y_i)$:

$$\boxed{k=1} \quad y_n = y_{n-1} + hf_{n-1}, \qquad \boxed{k=2} \quad y_n = y_{n-1} + \frac{h}{2}\big(3f_{n-1} - f_{n-2}\big),$$

$$\boxed{k=3} \quad y_n = y_{n-1} + \frac{h}{12}\big(23f_{n-1} - 16f_{n-2} + 5f_{n-3}\big),$$

$$\boxed{k=4} \quad y_n = y_{n-1} + \frac{h}{24}\big(55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}\big).$$

Ponekad se umjesto f_i piše y_i'. Vidimo da je $k = 1$ Ojlerova metoda. Za $k = 4$ kaže se da je Adamsova e. metoda IV reda. Provjera ($k = 4$): $b_{-4} = \int_0^1 \frac{(t+2)(t+1)t}{(-1)(-2)(-3)}dt = -\frac{9}{24}$, $b_{-3} = \int_0^1 \frac{(t+3)(t+1)t}{1\cdot(-1)(-2)}dt = \frac{37}{24}$, $b_{-2} = \int_0^1 \frac{(t+3)(t+2)t}{2\cdot 1\cdot(-1)}dt = -\frac{59}{24}$, $b_{-1} = \int_0^1 \frac{(t+3)(t+2)(t+1)}{3\cdot 2\cdot 1}dt = \frac{55}{24}$, $b_{-4}+\ldots+b_{-1} = 1$.

Zapaža se da metoda nije u stanju sama da startuje, odnosno da je za startovanje potrebna i druga, pomoćna metoda. Upravo, približne vrijednosti $y_1, \ldots, y_{k-1}$ izračunaju se po nekoj metodi tipa Runge–Kuta ili po metodi Tejlorovog razvoja $y_{i+1} = y_i + f(x_i, y_i)h + \frac{1}{2}(f_x'(x_i, y_i) + f_y'(x_i, y_i)f(x_i, y_i))h^2$.

Sada ćemo izvršiti ocjenu lokalne greške ρ_n:

$$\omega_k(x) = (x - x_{n-k})\ldots(x - x_{n-1}),$$

$$r_{k-1}(x) = \frac{1}{k!}\omega_k(x)y^{(k+1)}(\xi(x)),$$

$$\min(x_{n-k}, \ldots, x_{n-1}, x) < \xi(x) < \max(x_{n-k}, \ldots, x_{n-1}, x),$$

$$\min(x_{n-k}, x) < \xi(x) < \max(x_{n-1}, x),$$

$$x_{n-k} < \xi(x) < x_n \quad \text{jer} \quad x \in [x_{n-1}, x_n];$$

$$\rho_n = \int_{x_{n-1}}^{x_n} r_{k-1}(x)dx = \frac{1}{k!}\int_{x_{n-1}}^{x_n} \omega_k(x)y^{(k+1)}(\xi(x))dx =$$

t. o srednjoj vrijednosti za integrale, jer $\omega_k(x)$ ne mijenja znak na intervalu $x_{n-1} \le x \le x_n$

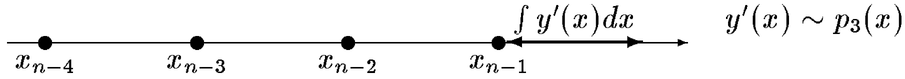
$$= \frac{1}{k!}y^{(k+1)}(\xi_n)\int_{x_{n-1}}^{x_n} \omega_k(x)dx =$$

izračunajte integral, pomoću smjene $x = x_{n-1} + ht$

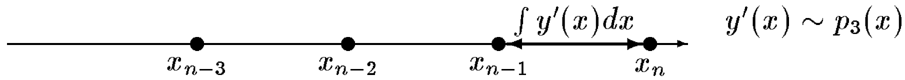
$$= \frac{1}{k!}y^{(k+1)}(\xi_n)C_k h^{k+1}, \quad \xi_n \in (x_{n-k}, x_n), \quad \text{prepišimo:} \quad \rho_n = \frac{1}{k!}C_k y^{(k+1)}(\xi)h^{k+1}.$$

Na primjer, u slučaju $k = 4$ dobija se $C_4 = \int_0^1 (t+3)(t+2)(t+1)t\,dt = \dfrac{251}{30}$, tako da je $\rho_n = \dfrac{251}{720}h^5 y^V(\xi_n)$.

Eksplicitna $(k=4)$



Implicitna $(k=4)$



Prelazimo na implicitnu (i.) ili interpolacionu Adamsovu formulu. Takođe se kaže i Adams–Multonova metoda. Neka sada $p_{k-1}(x)$ bude interpolacioni polinom funkcije $F(x) = y'(x)$ po mreži čvorova $x_{n-i} = x_n - ih$, gdje je $0 \le i \le k-1$. Sada je $\omega_k(x) = (x - x_{n-k+1}) \ldots (x - x_n)$. Izvođenje je slično onome od maločas. Račun pokazuje da je:

$$\int_{x_{n-1}}^{x_n} p_{k-1}(x)dx = h\Big(b_{-k+1}y'(x_{n-k+1}) + \ldots + b_0 y'(x_n)\Big).$$

Prema tome, i. Adamsova formula:

$$y_n = y_{n-1} + h\Big(b_{-k+1}f(x_{n-k+1}, y_{n-k+1}) + \ldots + b_0 f(x_n, y_n)\Big), \qquad k-1 \le n \le N \tag{1}$$

(b_i e. $\neq b_i$ i.).

Ustvari smo napisali jednu jednačinu po nepoznatoj y_n jer y_n figuriše i na lijevoj i na desnoj strani posljednje jednakosti.

U nastavku su date implicitne formule u slučajevima $k = 1, \ldots, 4$:

$$\boxed{k=1} \quad y_n = y_{n-1} + hf_n, \qquad \boxed{k=2} \quad y_n = y_{n-1} + \frac{h}{2}\Big(f_n + f_{n-1}\Big),$$

$$\boxed{k=3} \quad y_n = y_{n-1} + \frac{h}{12}\Big(5f_n + 8f_{n-1} - f_{n-2}\Big),$$

$$\boxed{k=4} \quad y_n = y_{n-1} + \frac{h}{24}\Big(9f_n + 19f_{n-1} - 5f_{n-2} + f_{n-3}\Big).$$

Za posljednju formulu kaže se da je Adamsova i. metoda IV reda.

Naravno da se lokalna greška ρ_n definiše kao $\rho_n = \int_{x_{n-1}}^{x_n} r_{k-1}(x)dx$, gdje je $r_{k-1}(x)$ greška interpolacije. Kada se sprovede račun koji je po svemu sličan onome od maločas onda će se dobiti:

$$\rho_n = \frac{1}{k!} y^{(k+1)}(\xi_n) C_k h^{k+1}, \qquad \xi_n \in (x_{n-k+1}, x_n)$$

(C_k e. $\neq C_k$ i.).

Recimo, $k=4$: $C_4 = \int_0^1 (t+2)(t+1)t(t-1)dt = -\dfrac{19}{30}$, $\rho_n = -\dfrac{19}{720} h^5 y^V(\xi_n)$.

2. Pojam prediktor–korektor metode

Kako da se jednačina (1) riješi brzo i sa dobrom preciznošću? Zapaža se da je pogodno da tu jednačinu rješavamo primjenom metode proste iteracije. Stavimo

$$y_n^{(\ell)} = y_{n-1} + h\Big(b_{-k+1}f(x_{n-k+1}, y_{n-k+1}) + \ldots + b_{-1}f(x_{n-1}, y_{n-1}) + b_0 f(x_n, y_n^{(\ell-1)})\Big)$$

$$\text{ili} \quad y_n^{(\ell)} = \varphi(y_n^{(\ell-1)}) \quad \text{za } \ell \geq 1,$$

gdje je $\varphi(t) = y_{n-1} + h\Big(b_{-k+1}f(x_{n-k+1}, y_{n-k+1}) + \ldots + b_{-1}f(x_{n-1}, y_{n-1}) + b_0 f(x_n, t)\Big)$, tako da je $\varphi'(t) = hb_0 f'_y(x_n, t)$; želimo da bude $\lim_{\ell\to\infty} y_n^{(\ell)} = y_n$, a treba da izaberemo početnu aproksimaciju $y_n^{(0)}$. S jedne strane, za dovoljno male h biće $|\varphi'(t)| \leq q < 1$, odnosno uslov kontrakcije biće ispunjen, ako je $\dfrac{\partial f}{\partial y}$ ograničena funkcija. S druge strane, dovoljno dobru početnu aproksimaciju može da nam pruži neka eksplicitna formula.

Iskustvo pokazuje da je dovoljno da se izvrši samo jedna iteracija. Drugim riječima, ne ide se dalje od $\ell = 1$; stavlja se da je $y_n \approx y_n^{(1)}$. Opisali smo postupak rješavanja jednačine (1).

Pogledajmo još jednom taj postupak. Računaju se dva broja $y_n^{(0)} = y_n^{\text{pred}}$ i $y_n^{(1)} = y_n^{\text{kor}}$, y_n^{kor} se uzima za y_n. Vidimo da se kombinuju jedna eksplicitna šema i jedna implicitna šema. Jedna i druga šema primijene se po jednom. Tako nastaje jedna nova (treća) šema za koju se kaže da je prediktor–korektor tipa (skraćeno p.–k. tipa). Broj y_n^{pred} predskazuje, a $y_n^{\text{kor}} = y_n$ je popravljena vrijednost.

Dvije šeme koje se kombinuju uzimaju se sa jednim te istim redom lokalne greške h^{k+1}, po pravilu. Iskustvo ili praksa pokazuje da upotreba metoda oblika p.–k. daje najbolje aproksimacije. Bašfortov broj služi da Multonova metoda prestane da bude implicitna.

Navedimo jedan primjer šeme oblika p.–k. Adamsova p.–k. šema IV reda data je sa:

$$y_n^* = y_{n-1} + \frac{h}{24}\Big(55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}\Big), \quad f_n^* = f(x_n, y_n^*),$$

$$y_n = y_{n-1} + \frac{h}{24}\Big(9f_n^* + 19f_{n-1} - 5f_{n-2} + f_{n-3}\Big), \quad 4 \leq n \leq N.$$

3. Primjeri diferencnih šema

Budući da je $y' = f(x, y)$ to je

$$\int_{x_n-ph}^{x_n} y'(x)dx = \int_{x_n-ph}^{x_n} f(x, y(x))dx \quad \text{ili} \quad y(x_n) - y(x_n - ph) = \int_{x_n-ph}^{x_n} f(x, y(x))dx.$$

Sada se na posljednji integral primijeni neka kvadraturna formula.

Izaberimo $p = 1$ i neka se primijeni trapezna formula:

$$y(x_n) - y(x_{n-1}) \approx \frac{h}{2}\Big(f(x_{n-1}, y(x_{n-1})) + f(x_n, y(x_n))\Big).$$

Odavde imamo šemu:

$$y_n = y_{n-1} + \frac{h}{2}\Big(f(x_{n-1}, y_{n-1}) + f(x_n, y_n)\Big) \quad \text{ili svejedno} \quad y_n = y_{n-1} + \frac{h}{2}\Big(f_{n-1} + f_n\Big).$$

Koristi se obična oznaka $f_i = f(x_i, y_i)$. Vidimo da je ovo jedna i. šema.

Izaberimo $p = 2$ i neka se primijeni Simpsonova formula:

$$y(x_n) - y(x_{n-2}) \approx \frac{h}{3}\Big(f(x_{n-2}, y(x_{n-2})) + 4f(x_{n-1}, y(x_{n-1})) + f(x_n, y(x_n))\Big).$$

Odavde imamo šemu:

$$y_n = y_{n-2} + \frac{h}{3}\Big(f_{n-2} + 4f_{n-1} + f_n\Big).$$

Ponekad se zapisuje kao $y_n = y_{n-2} + \frac{h}{3}(y'_{n-2} + 4y'_{n-1} + y'_n)$. Vidimo da je ovo jedna i. šema.

Izaberimo $p = 2$ i neka se primijeni formula pravougaonika:

$$y(x) - y(x-2h) \approx 2hf(x-h, y(x-h)) \qquad \text{ili} \qquad y(x_n) - y(x_{n-2}) \approx 2hf(x_{n-1}, y(x_{n-1})).$$

Odavde imamo šemu za računanje:

$$y_n = y_{n-2} + 2hf_{n-1}.$$

Vidimo da je ovo jedna e. šema.

4. Opšti oblik diferencne šeme

Razmotrimo ekvidistantnu mrežu čvorova $x_0, x_1, x_2, \ldots$ čiji je korak $h > 0$; $x_i = x_0 + ih$ za $i \geq 0$. Sljedeća jednakost definiše opšti oblik diferencne (ili višekoračne ili k–koračne) metode:

$$\sum_{i=0}^{k} a_{-i} y_{n-i} - h \sum_{i=0}^{k} b_{-i} f(x_{n-i}, y_{n-i}) = 0. \tag{2}$$

Formula (2) služi za računanje približnog broja y_n, a smatra se da su već poznati približni brojevi y_j, gdje je $j < n$.

Za startovanje šeme treba da su već određeni y_1, …, y_{k-1}. Za njihovo određivanje koristi se neka jednokoračna metoda.

U praksi se upotrebljavaju jedino šeme sa $a_0 \neq 0$. Razmatraćemo samo šeme u kojima je $a_0 \neq 0$.

Diferencne šeme dijele se na eksplicitne i implicitne. Ako je $b_0 = 0$ onda se za šemu kaže da je eksplicitna. Ako je $b_0 \neq 0$ onda se za šemu kaae da je implicitna.

Diferencne šeme dijelimo na one koje su Adamsovog tipa i one koje nisu Adamsovog tipa. Pogledajmo koeficijente a_0, a_{-1}, …, a_{-k}. Ako je $a_{-2} = \ldots = a_{-k} = 0$ onda se za šemu kaže da je Adamsova ili da je Adamsovog tipa; tada je očito $a_0 = 1$ i $a_{-1} = -1$.

Na kraju, lokalna greška ρ_n opšte diferencne šeme izražene formulom (2) definiše se slično kao lokalna greška za ranije navedene posebne slučajeve. Neka je $a_0 = 1$. Formula (2) definiše sljedeću šemu za računanje koja ide u računar:

$$y_n = -\sum_{i=1}^{k} a_{-i} y_{n-i} + h \sum_{i=0}^{k} b_{-i} y(x_{n-i}, y_{n-i}).$$

Ta šema za računanje nastala je od neke formule oblika

$$y(x_n) \approx -\sum_{i=1}^{k} a_{-i} y(x_{n-i}) + h \sum_{i=0}^{k} b_{-i} f(x_{n-i}, y(x_{n-i})).$$

Prema tome (razlika)

$$\rho_n = y(x_n) + \sum_{i=1}^{k} a_{-i} y(x_{n-i}) - h \sum_{i=0}^{k} b_{-i} f(x_{n-i}, y(x_{n-i})).$$

5.7. METODA NEODREĐENIH KOEFICIJENATA

Napišimo opet opšti oblik diferencne šeme:

$$\sum_{i=0}^{k} a_{-i} y_{n-i} - h \sum_{i=0}^{k} b_{-i} f_{n-i} = 0 \qquad \text{ili svejedno} \qquad \sum_{i=0}^{k} a_{-i} y_{n-i} - h \sum_{i=0}^{k} b_{-i} f(x_{n-i}, y_{n-i}) = 0,$$

gdje su a_{-i} i b_{-i} konstante, tj. a_{-i} i b_{-i} ne zavise od h. Posmatrajmo sljedeći izraz:

$$L = \sum_{i=0}^{k} a_{-i} y(x_{n-i}) - h \sum_{i=0}^{k} b_{-i} f(x_{n-i}, y(x_{n-i})) \qquad \text{ili} \qquad L = \sum_{i=0}^{k} a_{-i} y(x_{n-i}) - h \sum_{i=0}^{k} b_{-i} y'(x_{n-i}),$$

gdje je očito uzeto da je funkcija $y = y(x)$ analitičko rješenje d. j. $y' = f(x, y)$. Metoda neodređenih koeficijenata predstavlja jedan mogući način da se izaberu zasad neodređene veličine a_{-i} i b_{-i}, tj. da se pogodno odrede koeficijenti a_{-i} i b_{-i}. Postupa se kako slijedi. Napišimo Tejlorov razvoj funkcije $y = y(x)$ u okolini tačke $x = x_n$:

$$y(x_n + \alpha) = y(x_n) + y'(x_n)\alpha + \frac{1}{2!} y''(x_n)\alpha^2 + \ldots$$

(do nekog stepena α). Diferencirajmo ili svejedno napišimo Tejlorov razvoj funkcije $y' = y'(x)$ u okolini tačke $x = x_n$:

$$y'(x_n + \alpha) = y'(x_n) + y''(x_n)\alpha + \frac{1}{2!} y'''(x_n)\alpha^2 + \ldots.$$

Napišimo jedan i drugi razvoj kada je $\alpha = -kh$, ..., $\alpha = -h$ i $\alpha = 0$ redom. Sve to uvrstimo u L. Na taj način, kada se izvrši sređivanje:

$$L = E_0 y(x_n) + E_1 h y'(x_n) + E_2 h^2 y''(x_n) + \ldots + E_{q-1} h^{q-1} y^{(q-1)}(x_n) + O(h^q).$$

Sada želimo da što više početnih koeficijenata E_p bude jednako nuli. Recimo, želimo da bude $E_0 = E_1 = E_2 = E_3 = 0$. Na osnovu te želje određuju se koeficijenti a_{-i} i b_{-i}.

Pretpostavlja se da je analitičko rješenje $y = y(x)$ dovoljno glatka funkcija.

Nećemo dalje konkretizovati ovaj postupak.

5.8. OCJENA GREŠKE DIFERENCNE METODE

1. Slučaj eksplicitne Adamsove metode

Razmotrimo eksplicitnu Adamsovu šemu:

$$y_n = y_{n-1} + h \sum_{i=1}^{k} b_{-i} f(x_{n-i}, y_{n-i}), \qquad k \le n \le N.$$

Kao i dosad, $y = y(x)$ označava analitičko rješenje razmatranog Košijevog zadatka $y' = f(x, y)$, $y(x_0) = y_0$, a y_n su približne vrijednosti. Isto tako, $x_n = x_0 + nh = x_0 + nX/N$ (mreža čvorova je ravnomjerna).

Brojeve $y_1, \ldots, y_{k-1}$ dala je neka pomoćna metoda. Za grešku tih brojeva pretpostavlja se sljedeće: $|R_j| = |y(x_j) - y_j| \le C_0 h^{k+1}$ za $j = 1, \ldots, k-1$.

Iz 5.6. za lokalnu grešku ρ_n znamo:

$$\rho_n = y(x_n) - y_n^* = C_k h^{k+1} y^{(k+1)}(\xi), \qquad \text{gdje je} \qquad y_n^* = y(x_{n-1}) + h\sum_{i=1}^{k} b_{-i} f(x_{n-i}, y(x_{n-i})).$$

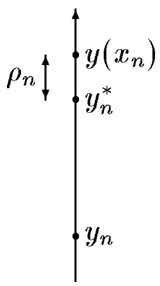
Slijedi

$$|\rho_n| \le C_1 h^{k+1},$$

ako se pretpostavi: $(k+1)$–vi izvod $y^{(k+1)}(x)$ funkcije $y = y(x)$ je ograničen na intervalu $[x_0, x_0+X]$, gdje $y = y(x)$ označava bilo koje rješenje d. j. $y' = f(x,y)$. Ako $f \in C^k(\Omega) = C^k([x_0, x_0+X] \times R)$ onda važi ta ograničenost, ako su još naravno svi parcijalni izvodi funkcije $f = f(x,y)$ do reda k uključeno ograničeni na skupu Ω.

Mi treba da ocijenimo grešku metode $R_n = y(x_n) - y_n$ za razne n.

Mi ćemo u izvođenju izraziti R_n preko $R_{n-k}, \ldots, R_{n-1}$. Mi ćemo napredovati po x–osi od tačke $x = x_0$ prema tački $x = x_N = x_0 + X$.



Slika za 1. Eksplicitna Adamsova

$y(x_n)$ je tačan broj, y_n je približan broj (računar saopštava y_n)

$R_n = y(x_n) - y_n$ je greška, ρ_n je lokalna greška

Vidjećemo da je potrebna još jedna pretpostavka (treća), pa se sada i ona uvodi, kako slijedi. Neka funkcija od dvije promjenljive $f = f(x,y)$ zadovoljava Lipšicov uslov po svojoj drugoj promjenljivoj y sa Lipšicovom konstantom L, tj. neka za ma koje dvije tačke $(x, y_1) \in \Omega$ i $(x, y_2) \in \Omega$ važi nejednakost

$$|f(x,y_1) - f(x,y_2)| \le L|y_1 - y_2|.$$

Za treću pretpostavku potrebno je i dovoljno da funkcija $f'_y(x,y)$ bude ograničena na skupu Ω, tj. da bude $|f'_y(x,y)| \le L$ za svako $(x,y) \in \Omega$.

Imamo redom:

$$R_n = y(x_n) - y_n = y(x_n) - y_n^* + y_n^* - y_n = \rho_n + y_n^* - y_n =$$

$$\rho_n + y(x_{n-1}) - y_{n-1} + h\sum_{i=1}^{k} b_{-i}[f(x_{n-i}, y(x_{n-i})) - f(x_{n-i}, y_{n-i})] =$$

$$\rho_n + R_{n-1} + h\sum_{i=1}^{k} b_{-i} f'_y(x_{n-i}, \widetilde{y}_{n-i})(y(x_{n-i}) - y_{n-i}) =$$

$$\rho_n + R_{n-1} + h\sum_{i=1}^{k} b_{-i} f'_y(x_{n-i}, \widetilde{y}_{n-i}) R_{n-i},$$

$$|R_n| \le |\rho_n| + |R_{n-1}| + h\sum_{i=1}^{k} |b_{-i}||f'_y(x_{n-i}, \tilde{y}_{n-i})||R_{n-i}| \le$$

$$|\rho_n| + |R_{n-1}| + h\sum_{i=1}^{k} |b_{-i}|L|R_{n-i}|.$$

Uvedimo u razmatranje niz brojeva $\{E_n\}_{n=0}^{N}$ sljedećom relacijom:

$$E_n = \max(|R_0|, |R_1|, \dots, |R_n|), \qquad n \ge k.$$

Jasno, $|R_n| \le E_n$. Jasno, uvedeni niz brojeva je neopadajući. Možemo reći da uvedeni brojevi odražavaju "rastući" oblik greške i da majoriraju grešku. Dalje imamo:

$$|R_n| \le |\rho_n| + E_{n-1} + h\sum_{i=1}^{k} |b_{-i}|LE_{n-1},$$

$$|R_n| \le |\rho_n| + E_{n-1} + hBLE_{n-1},$$

$$E_n \le |\rho_n| + E_{n-1}(1 + hBL) \qquad \text{za } \ n = k, k+1, \dots, N. \tag{1}$$

Uvedena je oznaka $B = \sum_{i=1}^{k} |b_{-i}|$.

Odranije raspolažemo sa sljedećim relacijama:

$$E_0 = 0, \qquad E_1 = \dots = E_{k-1} = C_0 h^{k+1}, \qquad |\rho_n| \le C_1 h^{k+1} \qquad (k \le n \le N). \tag{2}$$

Rekurentne formule (1) i (2) primjenjuju se uzastopno u nastavku.

Pisaćemo kao da je $C_1 \ge C_0$. Imamo redom:

$$E_k \le |\rho_k| + (1 + hBL)E_{k-1} \le C_1 h^{k+1} + (1 + hBL)C_0 h^{k+1} \le$$

$$C_1 h^{k+1} + (1 + hBL)C_1 h^{k+1} = C_1 h^{k+1}[1 + (1 + hBL)],$$

$$E_{k+1} \le |\rho_{k+1}| + (1 + hBL)E_k \le C_1 h^{k+1} + (1 + hBL)C_1 h^{k+1}[1 + (1 + hBL)] =$$

$$C_1 h^{k+1}[1 + (1 + hBL) + (1 + hBL)^2],$$

slično $\quad E_{k+2} \le C_1 h^{k+1}[1 + (1 + hBL) + (1 + hBL)^2 + (1 + hBL)^3], \quad$ itd.

$$E_N \le C_1 h^{k+1}[1 + (1 + hBL) + (1 + hBL)^2 + \dots + (1 + hBL)^{N-k+1}],$$

$$E_N \le C_1 h^{k+1}[1 + (1 + hBL) + \dots + (1 + hBL)^{N-1}] = \qquad \text{zbir geometrijske progresije}$$

$$C_1 h^{k+1} \frac{1}{(1 + hBL) - 1}[(1 + hBL)^N - 1] \le C_1 h^{k+1} \frac{1}{hBL}(1 + hBL)^N,$$

$$\text{poznato je} \qquad \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = e, \quad \forall n \ \left(1 + \frac{1}{n}\right)^n < e,$$

$$\text{poznato je} \quad \lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a, \quad \forall n \ \left(1+\frac{a}{n}\right)^n < e^a, \quad a > 0.$$

$$\text{Mi imamo} \quad E_N \le C_1 h^{k+1}\frac{1}{hBL}e^{XBL} = Ch^k$$

$$\text{jer je} \quad (1+hBL)^N = \left(1+\frac{XBL}{N}\right)^N \le e^{XBL}.$$

$$\text{Mi imamo definitivno} \quad |R_N| \le Ch^k.$$

Formulišimo teoremu koju smo upravo dokazali.

Teorema. Neka su ispunjeni sljedeći uslovi. (a) Neka $y = y(x)$ označava bilo koje rješenje d. j. $y' = f(x,y)$. Za svaku funkciju $y = y(x)$ važi $y \in C^{k+1}[x_0, x_0 + X]$. Sve funkcije $y^{(k+1)} = y^{(k+1)}(x)$ su ravnomjerno ograničene na intervalu $[x_0, x_0 + X]$. (b) Funkcija $f'_y(x,y)$ je neprekidna i ograničena na skupu $\Omega = [x_0, x_0+X]\times(-\infty, +\infty)$. (c) Startne približne vrijednosti $y_1, \dots, y_{k-1}$ dobijene su po nekoj šemi tipa Runge–Kuta čiji red lokalne greške iznosi h^{k+1}. Tada važi nejednakost $|R_N| \le Ch^k$, $C = \text{const}$, tj. C ne zavisi od h.

Važe nejednakosti $|R_n| \le Ch^k$, gdje je $0 \le n \le N$.

Kao što je rađeno, skup Ω može se smanjiti do skupa $\Omega' \subset R^2$ koji predstavlja usku traku oko grafika funkcije $y = y(x)$, tj. oko grafika analitičkog rješenja razmatranog Košijevog zadatka. Tačke (x_n, y_n) pripadaju Ω'. Za pojedino $x_0 \le x \le x_0 + X$, skup Ω' obuhvata $y(x) - \varepsilon \le y \le y(x) + \varepsilon$.

2. Slučaj implicitne Adamsove metode

Razmotrimo implicitnu Adamsovu šemu:

$$y_n = y_{n-1} + h\sum_{i=0}^{k-1} b_{-i} f(x_{n-i}, y_{n-i}), \qquad k-1 \le n \le N. \tag{3}$$

Biće opet: znamo da je lokalna greška reda h^{k+1}, a vidjećemo da je greška reda h^k. Izvođenje je slično i iskaz teoreme je sličan pa ćemo iznijeti bez dokaza. Ipak, pojavljuje se jedna razlika, kako slijedi. Kod eksplicitne šeme u iskazu teoreme je bilo: neka je N ma kakav prirodan broj i neka je onda $h = X/N$. Sada u iskazu teoreme imamo: neka je N dovoljno veliki prirodan broj i neka je onda $h = X/N$. Ili: neka je broj $h > 0$ dovoljno mali. Ovim se postiže da jednačina (3) ima jedinstveno rješenje za svako n.

$(\exists h_0 > 0)$ $(\forall h \le h_0)$ (postoje jedinstveni brojevi $\{y_n\}$ & važi $|R_n| \le Ch^k$)

Ima još jedna okolnost karakteristična za implicitnu šemu. Jednačina (3) po nepoznatoj y_n po pravilu bude riješena samo približno tačno. Ako se ona svaki put, tj. za svako n riješi sa greškom do $C_2 h^{k+1}$, gdje je C_2 neka konstanta, onda iskaz teoreme ostaje da važi.

3. Opšti slučaj diferencne metode ili slučaj diferencne metode koja ne mora da bude Adamsovog tipa

Razmotrimo šemu:

$$\sum_{i=0}^{k} a_{-i} y_{n-i} - h\sum_{i=0}^{k} b_{-i} f(x_{n-i}, y_{n-i}) = 0, \qquad a_0 \ne 0, \qquad k \le n \le N. \tag{4}$$

Teorema. (a) Neka lokalna greška ima red veličine h^{s+1}. (b) $|f'_y(x,y)| \leq L$ za $(x,y) \in \Omega$. (c) Red veličine greške brojeva $y_1, \ldots, y_{k-1}$ jednak je h^{s+1}. (d) Neka je ispunjen uslov α, v, niže. Tada je red veličine greške šeme (4) jednak h^s $(\forall h \leq h_0)$.

Ovu teoremu navodimo bez dokaza, a samo ćemo razjasniti šta je to "uslov α".

Razmotrimo jednačinu po nepoznatoj $\alpha \in C$:

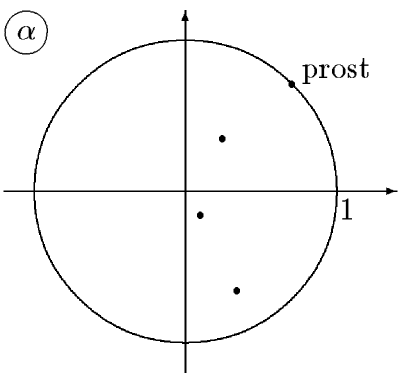
$$a_0\alpha^k + a_{-1}\alpha^{k-1} + \ldots + a_{-k} = 0,$$

ovo je tzv. karakteristična jednačina. Sagledajmo sva njena rješenja (sve njene korijene) u kompleksnoj α–ravni. Uslov α ili uslov korijena glasi: svi korijeni su po modulu ≤ 1 i svi korijeni koji su po modulu $= 1$ su prosti (jednostruki).

Primjer. Šema

$$y_n = -4y_{n-1} + 5y_{n-2} + h(4f_{n-1} + 2f_{n-2})$$

ima lokalnu grešku $O(h^4)$. Njena karakteristična jednačina glasi $\alpha^2 + 4\alpha - 5 = 0$, korijeni su $\alpha_1 = 1$ i $\alpha_2 = -5$, tako da uslov α nije ispunjen. Tu šemu ne treba upotrebljavati.



Slika za 3. Opšta diferencna
Korijeni $\alpha_1, \ldots, \alpha_k$

Napomena za buduće gradivo. Kasnije ćemo u 5.11. raditi Milnovu p.–k. metodu, tako da se izražava preko dvije formule:

$$\text{prva formula} \qquad y_n = y_{n-4} + h\Big(\ldots \quad \ldots \quad \ldots\Big)$$

znači $\alpha^4 - 1 = 0$, slijedi $\alpha_{1,2} = \pm 1$, $\alpha_{3,4} = \pm i$, tako da je uslov α ispunjen,

$$\text{druga formula} \qquad y_n = y_{n-2} + h\Big(\ldots \quad \ldots \quad \ldots\Big)$$

znači $\alpha^2 - 1 = 0$, slijedi $\alpha_{1,2} = \pm 1$, tako da je uslov α ispunjen.

5.9. ADAMSOVA METODA ČETVRTOG REDA

Adamsova p.–k. šema IV reda izražava se pomoću sljedeće dvije formule:

$$\begin{cases} y_n^{\text{pred}} = y_{n-1} + \dfrac{h}{24}\Big(55f(x_{n-1}, y_{n-1}) - 59f(x_{n-2}, y_{n-2}) + 37f(x_{n-3}, y_{n-3}) - 9f(x_{n-4}, y_{n-4})\Big), \\ y_n = y_{n-1} + \dfrac{h}{24}\Big(9f(x_n, y_n^{\text{pred}}) + 19f(x_{n-1}, y_{n-1}) - 5f(x_{n-2}, y_{n-2}) + f(x_{n-3}, y_{n-3})\Big) \end{cases}$$

za $n = 4, 5, \ldots, N$. Njena lokalna greška ρ_n je reda h^5, a njena greška $R_n = y(x_n) - y_n$ je reda h^4. Za lokalnu grešku ρ_n' Adamsove e. šeme IV reda važi:

$$\rho_n' = \frac{1}{4!} C' y^V(\xi') h^5 = \frac{1}{4!}\Big(\frac{251}{30}\Big) y^V(\xi') h^5, \qquad x_{n-4} < \xi' < x_n. \tag{1}$$

A za lokalnu grešku ρ_n razmatrane p.–k. šeme važi:

$$\rho_n = \frac{1}{4!} C y^V(\xi) h^5 + O(h^6) = \frac{1}{4!}\Big(-\frac{19}{30}\Big) y^V(\xi) h^5 + O(h^6), \qquad x_{n-3} < \xi < x_n. \tag{2}$$

Dosad rečeno rađeno je ranije u naslovima 5.6. i 5.8. a sada će biti korišćeno. Zadržavamo ranije oznake. Primjera radi: $h = X/N$ i $y_n = y_n^{\text{kor}}$. U ovom naslovu samo će biti izveden efektivni izraz za ρ_n.

Dva broja y_n^{pred} i y_n izračunati su oslanjanjem na jednu te istu prethodnu približnu vrijednost y_{n-1} (uslov u tački $x = x_{n-1}$). Ta dva broja se razlikuju zato što se razlikuju njihove lokalne greške ρ_n' i ρ_n. Mi pišemo

$$y_n^{\text{pred}} - y_n = (y(x_n) - \rho_n') - (y(x_n) - \rho_n) = \rho_n - \rho_n'. \tag{3}$$

Vidimo da $\xi' \to \xi$ kad $h \to 0$, a mi ćemo to koristiti u obliku $y^V(\xi') \approx y^V(\xi)$.

U nastavku se još samo primijene sredstva aritmetike na formule (1)–(3):

$$y_n^{\text{pred}} - y_n = \rho_n - \rho_n' \approx -\frac{19}{720} y^V(\xi) h^5 - \frac{251}{720} y^V(\xi) h^5 = -\frac{270}{720} y^V(\xi) h^5 \approx \frac{270}{19}\rho_n \approx 14\rho_n$$

(digitron pokazuje $\frac{270}{19} = 14{,}21$). Konačno:

$$\rho_n \approx \frac{1}{14}\Big(y_n^{\text{pred}} - y_n\Big), \qquad \text{u drugim oznakama} \qquad \rho_n \approx \frac{1}{14}\Big(y_n^* - y_n\Big).$$

Praktični izraz za grešku: $R_N \approx \sum_{n=1}^{N} \rho_n \exp\Big\{\int_{x_n}^{x_N} f_y'(x, \overline{y}(x))dx\Big\}$, gdje se ρ_n uzima iz konačne formule, integral izračunati po trapeznoj formuli, staviti $\overline{y}(x_i) \approx y_i$.

Na primjer, d. j. $y' = x^2 + y^2$, znači $f(x,y) = x^2 + y^2 \Rightarrow f_y'(x,y) = 2y$.

5.10. ALGORITAM ZASNOVAN NA DIFERENCNOJ METODI

Razmatra se početni zadatak $y' = f(x,y)$, $x_0 \le x \le x_0 + X$, $y(x_0) = y_0$.

Samo ćemo naznačiti elemente koji čine algoritam. Zadatak je definisan sljedećim parametrima: $f = f(x,y)$, x_0, y_0 i X. Neka smo se opredijelili za jednu određenu diferencnu metodu. Biće izložena dva načina za praktičnu ocjenu greške.

Prvi način. Sprovedu se dva odvojena proračuna. Neka su $z_h(x_0 + X)$ i $z_{2h}(x_0 + X)$ dvije približne vrijednosti za $y(x_0 + X)$ dobijene redom sa korakom h odnosno $2h$. Označimo sa $R_h(x_0 + X)$ i $R_{2h}(x_0 + X)$ odgovarajuće greške:

$$y(x_0 + X) = z_h(x_0 + X) + R_h(x_0 + X) \qquad \text{i} \qquad y(x_0 + X) = z_{2h}(x_0 + X) + R_{2h}(x_0 + X).$$

Uzmimo da je red greške diferencne metode koja se primjenjuje jednak h^4.

Između jedne i druge greške postoji sljedeća veza:

$$\lim_{h\to 0}\frac{R_{2h}(x_0+X)}{R_h(x_0+X)}=16 \qquad \text{ili} \qquad R_{2h}(x_0+X)\approx 16R_h(x_0+X).$$

Na običan Rungeov način, lako se dobija sljedeća procjena greške približnog broja $z_h(x_0+X)$:

$$R_h(x_0+X)\approx\frac{1}{15}\Big(z_h(x_0+X)-z_{2h}(x_0+X)\Big).$$

Drugi način. Izvrši se jedan proračun, a greška se procjenjuje preko lokalnih greški. Koristimo obične oznake $Nh=X$ i $x_n=x_0+nh$. Za grešku važi (za razne n):

$$R_n=\sum_{k=1}^{n}\rho_k\exp\{(x_n-x_k)f'_y(\xi_k,\eta_k)\} \qquad \text{ili} \qquad |R_n|\le\sum_{k=1}^{n}|\rho_k|\exp\{(x_n-x_k)L\}, \qquad L\ge\frac{\partial f}{\partial y}.$$

5.11. MILNOVA METODA

Neka je $Nh=X$ i $x_n=x_0+nh$. Milnova metoda definiše se šemom za računanje:

$$\begin{cases} y_n^{\text{pred}}=y_{n-4}+\dfrac{4h}{3}\Big(2f(x_{n-3},y_{n-3})-f(x_{n-2},y_{n-2})+2f(x_{n-1},y_{n-1})\Big), \\ y_n=y_{n-2}+\dfrac{h}{3}\Big(f(x_{n-2},y_{n-2})+4f(x_{n-1},y_{n-1})+f(x_n,y_n^{\text{pred}})\Big), \qquad n=4,5,\dots,N. \end{cases}$$

Prva formula može da se dobije kada se u identitetu $\int_{x_{n-4}}^{x_n}y'(x)dx=\int_{x_{n-4}}^{x_n}f(x,y(x))dx$ funkcija $F(x)=f(x,y(x))$ zamijeni svojim interpolacionim polinomom po tačkama $x=x_{n-4}$, $x=x_{n-3}$, $x=x_{n-2}$ i $x=x_{n-1}$. Kada se izračuna integral od polinoma onda će se kao jedan sabirak dobiti $0\cdot F(x_{n-4})$. Druga formula može da se dobije kada se napiše $\int_{x_{n-2}}^{x_n}y'(x)dx=\int_{x_{n-2}}^{x_n}f(x,y(x))dx$ i onda se $f(x,y(x))$ zamijeni svojim interpolacionim polinomom koji se odnosi na tačke x_{n-2}, x_{n-1}, x_n i x_{n+1}. Kada integral od polinoma onda će se kao jedan sabirak dobiti $0\cdot f(x_{n+1},y(x_{n+1}))$.

Lokalna greška Milnove metode je reda h^5, a njena greška R_n je reda h^4.

Za lokalnu grešku ρ_n važi praktična procjena:

$$\rho_n\approx\frac{1}{29}\Big(y_n^{\text{pred}}-y_n\Big) \qquad \text{ili u drugim oznakama} \qquad \rho_n\approx\frac{1}{29}\Big(y_n^*-y_n\Big).$$

Ponavljanje o višekoračnim (diferencnim) metodama: Adamsova p.–k. četvrtog: $y_{n+1}^*=y_n+\frac{h}{24}\Big(55f_n-59f_{n-1}+37f_{n-2}-9f_{n-3}\Big)$, $y_{n+1}=y_n+\frac{h}{24}\Big(9f_{n+1}^*+19f_n-5f_{n-1}+f_{n-2}\Big)$, Milnova: $y_{n+1}^*=y_{n-3}+\frac{4h}{3}\Big(2f_{n-2}-f_{n-1}+2f_n\Big)$, $y_{n+1}=y_{n-1}+\frac{h}{3}\Big(f_{n-1}+4f_n+f_{n+1}^*\Big)$.

6. NUMERIČKE METODE ZA RJEŠAVANJE GRANIČNOG ZADATKA ZA OBIČNE DIFERENCIJALNE JEDNAČINE

Razmatraćemo dvije vrste metoda. (a) Diferencne metode, drukčije se kaže – metoda konačnih razlika. (b) Varijacione metode, za neke od njih kaže se – metoda konačnih elemenata. Jednu i drugu vrstu opisaćemo u slučaju diferencijalne jednačine drugog reda.

6.1. METODA KONAČNIH RAZLIKA

Biće konstruisana numerička metoda za dobijanje približnog rješenja graničnog zadatka drugog reda. Biće dokazano da numerička metoda ima tzv. red aproksimacije h^2 i da je ona stabilna u odnosu na svoju desnu stranu i svoja dva granična uslova. Na kraju će biti dokazano da ona konvergira, s tim da je red konvergencije (red tačnosti) takođe h^2.

Prvo se daje postavka analitičkog zadatka koji treba da bude numerički riješen. Neka $p \in C^2[0,X]$, $p(x) \geq 0$ za $x \in [0,X]$, $f \in C^2[0,X]$. Razmotrimo jednačinu

$$-y'' + p(x)y(x) = f(x) \qquad \text{za } \ 0 \leq x \leq X \tag{1}$$

i granične uslove

$$y(0) = a, \qquad y(X) = b. \tag{2}$$

Iz teorije običnih d. j. poznata su sljedeća svojstva graničnog zadatka (1)–(2). Zadatak (1)–(2) ima jedinstveno rješenje $y = y(x)$ i $y \in C^4[0,X]$. Može se posmatrati diferencijalni operator $Ly = -y'' + p(x)y$ čiji je domen skup $\mathcal{D} = \{y\colon\ y \in C^2[0,X],\ y(0) = y(X) = 0\}$. Samo se napominje da je operator L simetričan, tj. važi $\langle Ly_1, y_2\rangle = \langle y_1, Ly_2\rangle$ za sve $y_1, y_2 \in \mathcal{D}$ (za simetričnost nije potrebno $p(x) \geq 0$). Pored toga, operator je i pozirivan, tj. važi $\langle Ly, y\rangle > 0$ za sve $y \in \mathcal{D}$ osim za $y(x) \equiv 0$. U Lebegovom prostoru $L^2(0,X)$, skalarni proizvod dvije funkcije definiše se kao $\langle y_1, y_2\rangle = \int_0^X y_1(x)y_2(x)dx$.

Neka je N prirodan broj i neka je $h = X/N$. Uvedimo ekvidistantnu mrežu čvorova $x_n = nh$ za $0 \leq n \leq N$. Neka je $y = y(x)$ analitičko rješenje razmatranog zadatka (1)–(2), tako da je $y(x_n)$ njegova vrijednost u čvoru $x = x_n$. Neka su y_n odgovarajuće približne vrijednosti, biće dobijene kasnije. Uvedimo i oznaku za grešku (grešku metode): $R_n = y(x_n) - y_n$ za $0 \leq n \leq N$.

Neka je

$$\ell_0(y(x_n)) = -\frac{1}{h^2}\Big(y(x_{n+1}) - 2y(x_n) + y(x_{n-1})\Big)$$

i

$$r_n = -\ell_0(y(x_n)) - y''(x_n) = \frac{1}{h^2}\Big(y(x_{n+1}) - 2y(x_n) + y(x_{n-1})\Big) - y''(x_n).$$

Za r_n se kaže da je greška aproksimacije. Poznata je sljedeća formula:

$$r_n = \frac{1}{12}y^{IV}(\xi_n)h^2, \qquad \text{gdje je } \ x_{n-1} < \xi_n < x_{n+1}.$$

Dokažimo formulu korišćenjem Tejlorovog razvoja

$$y(x_n + \alpha) = y(x_n) + \alpha y'(x_n) + \frac{1}{2!}\alpha^2 y''(x_n) + \frac{1}{3!}\alpha^3 y'''(x_n) + \frac{1}{4!}\alpha^4 y^{IV}(x_n + \theta\alpha)$$

$(0 < \theta < 1)$ kada je $\alpha = h$ i $\alpha = -h$:

$$r_n = \frac{1}{h^2}\Big(y(x_n) + \frac{1}{2!}y''(x_n)h^2 + \frac{1}{4!}y^{IV}(\xi_1)h^4 - 2y(x_n) +$$

$$y(x_n) + \frac{1}{2!}y''(x_n)h^2 + \frac{1}{4!}y^{IV}(\xi_2)h^4\Big) - 2y''(x_n) =$$

$$\frac{1}{4!h^2}\left(y^{IV}(\xi_1) + y^{IV}(\xi_2)\right) = \frac{1}{12}y^{IV}(\xi)h^2, \qquad x_{n-1} < \xi < x_{n+1}.$$

Tako da je greška aproksimacije reda h^2. Ili $|r_n| \le \frac{1}{12}M_4h^2$, gdje je $M_4 = \max_{x\in[0,X]}|y^{IV}(x)|$, gdje je $y = y(x)$ tačno rješenje za (1)–(2).

Neka je $p_n = p(x_n)$ i $f_n = f(x_n)$. Napišimo jednačinu (1) kada je $x = x_n$:

$$-y''(x_n) + p_n y(x_n) = f_n.$$

Uvedimo oznaku

$$\ell(y(x_n)) = -\frac{1}{h^2}\Big(y(x_{n+1}) - 2y(x_n) + y(x_{n-1})\Big) + p_n y(x_n).$$

Imamo

$$\ell(y(x_n)) = f_n - r_n.$$

U numeričkoj metodi, drugi izvod $y''(x)$ u čvoru $x = x_n$ zamjenjuje se razlikom drugog reda koja je sastavljena od efektivno poznatih brojeva $\{y_n\}_{n=0}^N$ (biće dobijeni kasnije), tj. od y_{n-1}, y_n, y_{n+1}, a ne od brojeva $y(x_{n-1})$, $y(x_n)$, $y(x_{n+1})$. Mi ćemo saznati brojeve $\{y_n\}_{n=0}^N$ ako riješimo sljedeći sistem linearnih jednačina:

$$-\frac{1}{h^2}\Big(y_{n+1} - 2y_n + y_{n-1}\Big) + p_n y_n = f_n \qquad \text{ili} \qquad \ell(y_n) = f_n, \qquad 1 \le n \le N-1, \tag{3}$$

$$y_0 = a, \qquad y_N = b. \tag{4}$$

Imamo da je $\ell(y(x_n)) = f_n - r_n$ i $\ell(y_n) = f_n$. Oduzimanjem, uzimajući u obzir linearnost diferencnog izraza ℓ,

$$\ell(y(x_n)) - \ell(y_n) = f_n - r_n - f_n \quad \text{ili} \quad \ell(y(x_n) - y_n) = -r_n \quad \text{ili} \quad \ell(R_n) = -r_n, \quad 1 \le n \le N-1,$$

veza greške metode R_n i greške aproksimacije r_n.

Biće dokazana sljedeća teorema.

Teorema. Sistem linearnih jednačina (3)–(4) ima jedinstveno rješenje $\{y_n\}_{n=0}^N$ i važi sljedeća formula (za ocjenu greške):

$$\max_{0\le n\le N}|R_n| \le \frac{1}{96}X^2M_4h^2.$$

Mi ćemo ustvari dokazati nešto opštiju teoremu koja se odnosi na nešto opštiju situaciju. Prethodna teorema analizira samo grešku metode. Sljedeća teorema uzima u obzir i grešku računanja i grešku izazvanu približnošću ulaznih podataka; ulazni podaci su a i b. Neka veličine $\{y_n\}_{n=0}^N$ više ne zadovoljavaju sistem (3)–(4) nego odsad uzimamo da one zadovoljavaju sljedeći sistem:

$$-\frac{1}{h^2}\Big(y_{n+1} - 2y_n + y_{n-1}\Big) + p_n y_n = f_n + \delta_n \qquad \text{ili} \qquad \ell(y_n) = f_n + \delta_n, \qquad 1 \le n \le N-1, \tag{5}$$

$$y_0 = a - R_0, \qquad y_N = b - R_N. \tag{6}$$

Uslovi (5)–(6) bi se očito sveli na (3)–(4) da je $\delta_n = 0$ i $R_0 = R_N = 0$. Zašto su uvedeni δ_n? Kada se sistem linearnih jednačina riješi onda se njegovo rješenje radi provjere uvrsti u sami taj sistem. Lijeva i desna strana se ne poklope već se razlikuju za δ_n. Razlikuju se zato što se tokom rješavanja sistema akumulirala greška računanja. Još, brojevi δ_n odgovaraju i slučaju kada je desna strana jednačine $f = f(x)$ poznata samo približno, poznata sa nekom greškom. Zašto su potrebni R_0 i R_N? Moguće je da su brojevi a i b koji definišu par graničnih uslova samo približno poznate veličine.

Za mjeru greške računanja uzećemo $\max_{0<n<N} |\delta_n|$. Za mjeru greške ulaznih podataka uzećemo $\max(|R_0|, |R_N|)$. I dalje je $R_n = y(x_n) - y_n$. Sada R_n odražava sve tri komponente greške (metode, računanja i od ulaznih podataka).

Ranija veza $\ell(R_n) = -r_n$ sada u novoj situaciji očito postaje

$$\ell(R_n) = -r_n - \delta_n, \qquad 1 \le n \le N-1.$$

Sada u novoj situaciji važi sljedeća teorema (koja će biti dokazana).

Teorema. Sistem linearnih jednačina (5)–(6) ima jedinstveno rješenje $\{y_n\}_{n=0}^N$ i važi sljedeća formula (za ocjenu greške):

$$\max_{0\le n\le N} |R_n| \le \frac{1}{96} X^2 M_4 h^2 + \frac{1}{8} X^2 \max_{0<n<N} |\delta_n| + \max(|R_0|, |R_N|).$$

Dokaz teoreme. Sistem linearnih jednačina (5) sastoji se od $N-1$ jednačina i ima $N-1$ nepoznatih $\{y_n\}_{n=1}^{N-1}$. Možemo pomnožiti jednačine sa $-h^2$. Označimo sa M matricu sistema, ona je dimenzije $(N-1)\times(N-1)$. Vidimo da je matrica M trodijagonalna. Prvo pitanje: pokazaćemo da je matrica M regularna, tj. da je $\det M \ne 0$; koristićemo sljedeće: $p(x) \ge 0$ za svako $x \in [0, X] \Rightarrow p_n \ge 0$ za svako $n \in \{1, \dots, N-1\}$. Mi pišemo

$$M = \begin{bmatrix} -2-p_1h^2 & 1 & 0 & \cdots & 0 & 0 \\ 1 & -2-p_2h^2 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2-p_{N-1}h^2 \end{bmatrix}.$$

Recimo, u slučaju $p(x) \equiv 0$, $N = 6$ matrica glasi

$$M = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}.$$

Posmatrajmo homogeni sistem $M\mathbf{x} = 0$, stavimo $x_l = \max_{1\le n\le N-1} |x_n|$ i dopustimo da je $|x_l| > 0$. Ne može biti $l = 1$, budući da prva jednačina sistema glasi $-2x_1 + x_2 = 0$ ili $-(2+p_1h^2)x_1 + x_2 = 0$. Na isti način, ne može biti $l = N-1$. Prema tome, l–ta jednačina ima oblik $x_{l-1} - (2+p_lh^2)x_l + x_{l+1} = 0$. Napisana relacija je održiva samo ako je $p_l = 0$ i $x_{l-1} = x_l = x_{l+1}$. Zatim gledamo $(l-1)$–vu jednačinu, itd. (zatim gledamo $(l+1)$–vu jednačinu, itd.). Tako dobijamo $x_1 = \ldots = x_{N-1}$. Dobili smo kontradikciju jer smo dopustili da je $|x_l| > 0$. Ne može biti $|x_l| > 0$, mora biti $|x_l| = 0$. Znači $|x_n| = 0$ za svako $n = 1, \dots, N-1$. Znači da homogeni sistem $M\mathbf{x} = 0$ ima samo trivijalno (nulto) rješenje. Mi smo pokazali da je matrica M regularna.

Zato sistem (5)–(6) ima jedinstveno rješenje. Prvo pitanje je završeno.

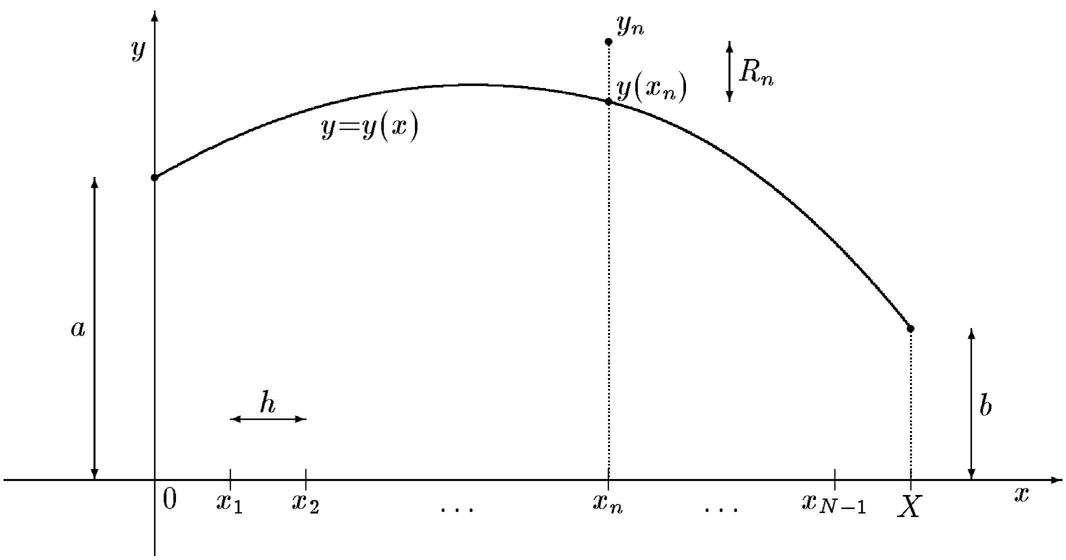
Mi raspolažemo sa

$$\ell(R_n) = -r_n - \delta_n, \qquad 1 \leq n \leq N-1. \tag{7}$$

Bilo bi bolje da raspolažemo sa $R_n = \ldots$. Kao da bi se na relaciju $\ell(R_n) = -r_n - \delta_n$ primijenio operator ℓ^{-1}. O tome je ustvari riječ u nastavku.

Mi raspolažemo i sa

$$|r_n| \leq \frac{1}{12} M_4 h^2, \qquad 1 \leq n \leq N-1. \tag{8}$$



Dokažimo dvije leme (uslov stabilnosti).

Lema 1. Neka je $p_n \geq 0$ za $1 \leq n \leq N-1$. Neka je $\ell(z_n) = -\frac{1}{h^2}\Big(z_{n+1} - 2z_n + z_{n-1}\Big) + p_n z_n$. Neka je $\{z_n\}_{n=0}^{N}$ ma kakav konačan niz brojeva. Ako je $\ell(z_n) \geq 0$ za $1 \leq n \leq N-1$ i $z_0 \geq 0$, $z_N \geq 0$ onda je $z_n \geq 0$ za $1 \leq n \leq N-1$.

Dokaz. Uvedimo oznaku $d = \min\limits_{0 \leq n \leq N} z_n$ i dopustimo da je $d < 0$. Za koje n važi $z_n = d$? Ne može biti $z_0 = d$, niti $z_N = d$. Neka je q najmanji cio broj za koji je $z_q = d$. Imamo $z_{q-1} > d$ i $z_{q+1} \geq d$. Tako da je $-(z_{q+1} - 2z_q + z_{q-1}) < 0$. Još, $p_q \geq 0$ i $z_q = d < 0 \Rightarrow p_q z_q \leq 0$. Sabiranjem

$$\ell(z_q) = -\frac{1}{h^2}\Big(z_{q+1} - 2z_q + z_{q-1}\Big) + p_q z_q < 0.$$

Po uslovu leme je $\ell(z_q) \geq 0$, tako da smo dobili kontradikciju. Dakle, ne može biti $d < 0$, nego mora biti $d \geq 0$. Lema je dokazana.

Lema 2. Neka je $p_n \geq 0$ za $1 \leq n \leq N-1$. Neka je $\ell(z_n) = -\frac{1}{h^2}\Big(z_{n+1} - 2z_n + z_{n-1}\Big) + p_n z_n$. Neka je $\{z_n\}_{n=0}^{N}$ ma kakav konačan niz brojeva. Važi nejednakost

$$\max_{0 \leq n \leq N} |z_n| \leq \max(|z_0|, |z_N|) + \frac{1}{8} X^2 Z, \qquad \text{gdje je } \; Z = \max_{0<n<N} |\ell(z_n)|.$$

Dokaz. Uvedimo u razmatranje niz brojeva

$$\omega_n = |z_0|\frac{X-nh}{X} + |z_N|\frac{nh}{X} + \frac{1}{2}Z(X-nh)nh.$$

Iz eksplicitnog izraza za ω_n je $\omega_n \geq 0$. Neposrednim računom nalazimo da je $-\frac{1}{h^2}\Big(\omega_{n+1} - 2\omega_n + \omega_{n-1}\Big) = Z$, tako da je $\ell(\omega_n) = Z + p_n\omega_n \geq Z$. Dalje, za $1 \leq n \leq N-1$ imamo $\ell(\omega_n \pm z_n) = \ell(\omega_n) \pm \ell(z_n) \geq Z \pm \ell(z_n) \geq 0$. Pored toga, $\omega_0 \pm z_0 = |z_0| \pm z_0 \geq 0$ i $\omega_N \pm z_N = |z_N| \pm z_N \geq 0$. Prema tome, niz brojeva $\{\omega_n \pm z_n\}_{n=0}^N$ zadovoljava sve uslove prethodne leme. Zato imamo $\omega_n \pm z_n \geq 0$ za $0 \leq n \leq N$. Napišimo odvojeno: $\omega_n + z_n \geq 0$ i $\omega_n - z_n \geq 0$. Znači da je $|z_n| \leq \omega_n$. Slijedi da je $|z_n| \leq \max\limits_{0\leq n\leq N} \omega_n$.

Ostaje samo da se izračuna $\max\limits_{0\leq n\leq N} \omega_n$. Imamo:

$$|z_0|\frac{X-nh}{X} + |z_N|\frac{nh}{X} \leq \max(|z_0|,|z_N|)\frac{X-nh}{X} + \max(|z_0|,|z_N|)\frac{nh}{X} = \max(|z_0|,|z_N|),$$

$$\frac{1}{2}Z(X-nh)nh \leq \frac{1}{2}Z\frac{1}{4}X^2 \quad \text{jer je} \quad x(1-x) \leq \frac{1}{4} \quad \text{za } x \in [0,1],$$

$$\text{sabiranjem:} \quad \omega_n \leq \max(|z_0|,|z_N|) + \frac{1}{2}Z\frac{1}{4}X^2.$$

Znači da je $\max\limits_{0\leq n\leq N} \omega_n \leq \max(|z_0|,|z_N|) + \frac{1}{2}Z\frac{1}{4}X^2$. Lema je dokazana jer $(\forall n)\ |z_n| \leq x \Rightarrow \max\limits_{0\leq n\leq N} |z_n| \leq x$.

Lema 2 govori sljedeće. Neka $\{z_n\}_{n=0}^N$ predstavlja rješenje diferencnog graničnog zadatka $\ell(z_n) = f_n$ za $0 < n < N$, $z_0 = a$, $z_N = b$. Tada je ispunjen tzv. uslov stabilnosti, tj. tada važi nejednakost $\|\mathbf{z}\| \leq C_1\|\mathbf{f}\| + C_2\|\mathbf{g}\|$, gdje je $\mathbf{z} = (z_0, z_1, \ldots, z_N)$, $\mathbf{f} = (f_1, \ldots, f_{N-1})$ i $\mathbf{g} = (a, b)$. Imamo u vidu max–normu. Dva diferencna problema: ako su mali $\Delta\mathbf{f}$ i $\Delta\mathbf{g}$ onda je i $\Delta\mathbf{z}$ (oduzimanjem, uzmite u obzir linearnost). Male promjene $\mathbf{f}$ i $\mathbf{g}$ $\Rightarrow$ samo male promjene $\mathbf{z}$.

Slijedi završni dio dokaza teoreme. Primijenimo lemu 2 na niz brojeva $\{R_n\}_{n=0}^N$:

$$\max_{0\leq n\leq N} |R_n| \leq \max(|R_0|,|R_N|) + \frac{1}{8}X^2 \max_{0<n<N} |\ell(R_n)| = \qquad \text{po (7)}$$

$$\max(|R_0|,|R_N|) + \frac{1}{8}X^2 \max_{0<n<N} |-r_n - \delta_n| \leq$$

$$\max(|R_0|,|R_N|) + \frac{1}{8}X^2 \max_{0<n<N} |r_n| + \frac{1}{8}X^2 \max_{0<n<N} |\delta_n| \leq \qquad \text{po (8)}$$

$$\max(|R_0|,|R_N|) + \frac{1}{8}X^2\frac{1}{12}M_4h^2 + \frac{1}{8}X^2 \max_{0<n<N} |\delta_n|.$$

Teorema je dokazana.

U nastavku – dopune.

Dopuna 1. O ocjeni greške.

Posmatrajmo samo grešku metode, tj. neka je $\delta_n = 0$ i $R_0 = R_N = 0$. Red greške je h^2, greška $\approx Ch^2$. Na običan Rungeov način dobija se praktični izraz za grešku. Dakle, neka je x tačka sa intervala $[0, X]$ koja je čvor i po mreži sa korakom h i po mreži sa korakom $2h$. Neka su $z_h(x)$ i $z_{2h}(x)$ dvije odgovarajuće približne vrijednosti. Tada je

$$\text{greška}(z_h(x)) = R_h(x) = y(x) - z_h(x) \approx \frac{1}{3}\Big(z_h(x) - z_{2h}(x)\Big). \tag{9}$$

Drugi pristup ocjeni greške nalazimo ako diferenciramo relaciju $y'' = py - f$: $y''' = p'y + py' - f'$, $y^{IV} = p''y + 2p'y' + py'' - f'' = p''y + 2p'y' + p(py - f) - f''$, preko y_n i y'_n za $1 \leq n \leq N-1$ dolazimo do M_4^*, gdje y' treba procijeniti na bazi $y'(x) \approx \frac{1}{2h}\left(y(x+h) - y(x-h)\right)$, $y'_n \approx \frac{1}{2h}\left(y_{n+1} - y_{n-1}\right)$. Time postaje upotrebljiva formula $\frac{1}{96}X^2M_4h^2$.

Dopuna 2. Trodijagonalni sistem.

Koji postupak može da posluži za rješavanje trodijagonalnog sistema linearnih jednačina (3)? Naravno da je to (pojednostavljeni) Gausov algoritam uzastopne eliminacije nepoznatih. Ako je potrebno napisati program za računar, mogu se koristiti gotove formule (Tomasov algoritam):

$$c_1x_1 + b_1x_2 = d_1, \quad a_ix_{i-1} + c_ix_i + b_ix_{i+1} = d_i, \quad i = 2, \ldots, n-1, \quad a_nx_{n-1} + c_nx_n = d_n \quad (10)$$

$$\Rightarrow \quad \alpha_2 = -\frac{b_1}{c_1}, \quad \beta_2 = \frac{d_1}{c_1}, \quad \alpha_{i+1} = -\frac{b_i}{c_i + a_i\alpha_i}, \quad \beta_{i+1} = \frac{d_i - a_i\beta_i}{c_i + a_i\alpha_i}, \quad i = 1, \ldots, n-1 \quad (a_1 = 0),$$

$$x_n = \frac{d_n - a_n\beta_n}{c_n + a_n\alpha_n}, \quad x_i = \alpha_{i+1}x_{i+1} + \beta_{i+1}, \quad i = n-1, \ldots, 1.$$

Dopuna 3. Ako potencijal nije pozitivan.

Ako se ukine uslov $p(x) \geq 0$ onda u iskazu teoreme treba dodati: za dovoljno male $h > 0$. Dodati i pretpostavku o postojanju jedinstvenog rješenja $y = y(x)$ analitičkog problema (1)–(2).

Dopuna 4. Nešto opštija diferencijalna jednačina.

Ako diferencijalna jednačina glasi $y'' + p(x)y' + q(x)y = f(x)$ onda u numeričkoj metodi imamo ovakav sistem diferencnih jednačina (ovakav diferencni problem):

$$\frac{1}{h^2}\Big(y_{n+1} - 2y_n + y_{n-1}\Big) + p(x_n)\frac{1}{2h}\Big(y_{n+1} - y_{n-1}\Big) + q(x_n)y_n = f(x_n)$$

$(1 \leq n \leq N-1)$. Obrazloženje: poznate su formule: $\lim_{h\to 0} \frac{1}{h^2}\Big(y(x_{n+1}) - 2y(x_n) + y(x_{n-1})\Big) = y''(x)$, $\lim_{h\to 0} \frac{1}{2h}\Big(y(x_{n+1}) - y(x_{n-1})\Big) = y'(x)$.

Dopuna 5. Drugačiji granični uslovi.

Za granične uslove $y(0) = a$, $y(X) = b$ kaže se da predstavljaju granične uslove prve vrste, Dirihleove uslove. Kao što smo vidjeli, u numeričkoj metodi, njima se pridružuju relacije $y_0 = a$, $y_N = b$. Ako u postavci analitičkog problema piše $y'(0) = a$, $y'(X) = b$ (granični uslovu druge vrste, Nojmanovi uslovi) onda u numeričkoj mnetodi treba

$$\frac{1}{h}\Big(y_1 - y_0\Big) = a, \qquad \frac{1}{h}\Big(y_N - y_{N-1}\Big) = b.$$

Obrazloženje: $\lim_{h\to 0} \frac{1}{h}\Big(y(h) - y(0)\Big) = y'(0)$, $\lim_{h\to 0} \frac{1}{h}\Big(y(X) - y(X-h)\Big) = y'(X)$. Pogledajmo i slučaj graničnih uslova treće vrste, Robenovih uslova $\alpha y(0) + \overline{\alpha}y'(0) = a$, $\beta y(X) + \overline{\beta}y'(X) = b$. U numeričkoj metodi, takvim uslovima odgovaraju diferencni uslovi $\alpha y_0 + \overline{\alpha}\frac{1}{h}\Big(y_1 - y_0\Big) = a$, $\beta y_N + \overline{\beta}\frac{1}{h}\Big(y_N - y_{N-1}\Big) = b$ $(\alpha, \overline{\alpha}, a, \beta, \overline{\beta}, b \in R)$. Bilo koji interval, ne mora $[0, X]$.

Ako se granični problem $-y'' + p(x)y = f(x)$, $y(0) = y(1) = 0$ rješava primjenom metode varijacionog tipa (METODE KONAČNIH ELEMENATA) onda se određuje funkcija $y = y(x)$ koja realizuje minimum funkcionala

$$F(y) = \langle Ly, y\rangle - 2\langle f, y\rangle = \int_0^1 \Big((y'(x))^2 + p(x)y^2(x) - 2f(x)y(x)\Big)\,dx,$$

gdje je $Ly = -y'' + p(x)y$ $(p(x) \geq 0)$.

NEKA PREZIMENA U ORIGINALU:

Adams	Adams	Gaus	Gauss	Lebeg	Lebesgue	Rafson	Raphson
Banah	Banach	Hermit	Hermite	Ležandr	Legendre	Romberg	Romberg
Bašfort	Bashforth	Hilbert	Hilbert	Lipšic	Lipschitz	Runge	Runge
Koši	Cauchy	Jakobi	Jacobi	Miln	Milne	Zajdel	Seidel
Kotes	Cotes	Kuta	Kutta	Multon	Moulton	Simpson	Simpson
Ojler	Euler	Lagranž	Lagrange	Njutn	Newton	Tejlor	Taylor

MATLAB je najpoznatiji komercijalni softverski paket za oblast numeričkih metoda.

LITERATURA. Desanka Radunović: Numeričke metode, Akademska misao, Beograd, 2003. Ima na internetu. Slobodno se može preuzeti kao jedan fajl formata pdf, ima 240 strana. Otkucati u Google: desanka radunovic numericke metode. Nalazi se na sajtu elibrary.matf.bg.ac.rs.

NUMERIČKA ANALIZA
SADRŽAJ PREDAVANJA:

1.1 Lagranžov interpolacioni polinom xnumera.tex
1.2 Ocjena greške za Lagranžov interpolacioni polinom xnumera.tex
1.3 Podijeljene razlike i njihova svojstva xnumera.tex
1.4 Njutnova interpolaciona formula sa podijeljenim razlikama xnumera.tex
1.5 Konačne razlike xnumera.tex
1.6 Njutnove interpolacione formule sa konačnim razlikama xnumera.tex
1.7 Interpolacija sa višestrukim čvorovima xnumerb.tex
1.8 Interpolacija pomoću splajna xnumerb.tex
1.9 Numeričko diferenciranje xnumerc.tex
1.10 Nestabilnost numeričkog diferenciranja i tri vrste greške u numeričkim metodama xnumerc.tex
1.11 Pojam približnog broja xnumerc.tex
1.12 Greška funkcije xnumerc.tex
2.1 Tri formule xnumerd.tex
2.2 Rungeovo pravilo za praktičnu ocjenu greške xnumerd.tex
2.3 Rombergova formula xnumerd.tex
2.4 Kvadraturne formule u slučaju prisustva težinske funkcije xnumerd.tex
2.5 Gausova kvadraturna formula xnumerd.tex
3.1 Gausova metoda eliminacije xnumere.tex
3.2 Gausova metoda eliminacije sa izborom glavnog elementa xnumere.tex
3.3 Mjera uslovljenosti matrice xnumere.tex
3.4 Iterativne metode za rješavanje sistema linearnih jednačina xnumere.tex
3.5 Zajdelova metoda xnumere.tex
3.6 Primjer iterativne metode (za rješavanje sistema linearnih jednačina) varijacionog tipa xnumerf.tex
3.7 Metoda skalarnog proizvoda xnumerf.tex
4.1 Metoda polovljenja xnumerg.tex
4.2 Metoda proste iteracije xnumerg.tex
4.3 Njutnova metoda xnumerg.tex
5.1 Uvod o Košijevom zadatku i lema o dva rješenja diferencijalne jednačine xnumerh.tex
5.2 Ojlerova metoda i drugi primjeri xnumerh.tex
5.3 Opšti slučaj eksplicitne metode tipa Runge–Kuta xnumerh.tex
5.4 Ocjena greške za metodu Runge–Kuta xnumerh.tex
5.5 Algoritam zasnovan na metodi Runge–Kuta xnumerh.tex
5.6 Diferencne metode xnumeri.tex
5.7 Metoda neodređenih koeficijenata xnumeri.tex
5.8 Ocjena greške diferencne metode xnumeri.tex
5.9 Adamsova metoda četvrtog reda xnumeri.tex
5.10 Algoritam zasnovan na diferencnoj metodi xnumeri.tex
5.11 Milnova metoda xnumeri.tex
6.1 Metoda konačnih razlika xnumerj.tex

Fajlovi (gsview): xnumera.tex 11 pages xnumerb.tex 8 pages xnumerc.tex 9 pages xnumerd.tex 16 pages xnumere.tex 16 pages xnumerf.tex 6 pages xnumerg.tex 11 pages xnumerh.tex 13 pages xnumeri.tex 12 pages xnumerj.tex 7 pages. $\Sigma = 10$ files & 109 pages