3 Елементи теорије грешака - допуна

3.1 Веза сигурних цифара и релативне грешке

Нека је x тачна вредност неке величине, а $\bar{x}$ њена приближна вредност. *Апсолутна и релативна грешка* приближног броја $\bar{x}$ су редом бројеви

$$\Delta(\bar{x}) = |x - \bar{x}| \quad \text{и} \quad \delta(\bar{x}) = \frac{|x - \bar{x}|}{|x|}.$$

За разлику од апсолутне грешке, релативна грешка је инваријантна у односу на скалирање:

$$1 - \frac{\bar{x}}{x} = 1 - \frac{k \cdot \bar{x}}{k \cdot x}, \quad k \in \mathbb{R} \setminus \{0\}.$$

Нека су x и $\bar{x}$ записани у децималном облику. Број *сигурних* цифара нам говори на колико се позиција x и $\bar{x}$ поклапају: кажемо да $\bar{x}$ има k сигурних цифара (у ужем смислу) броја x, ако апсолутна грешка $|x-\bar{x}|$ има нуле на првих k децималних места, почев од водећег ненула децималног места броја x, а након тих нула је број од 0 до 4 или 50000..., тј.

$$|x - \bar{x}| \leqslant \overset{1}{\boxed{0}}\overset{2}{\boxed{0}} \cdots \overset{k}{\boxed{0}}\overset{k+1}{\boxed{5}}\boxed{0}\boxed{0}\boxed{0} \cdots$$

Ако запишемо број x у облику

$$x = \pm a_1.a_2a_3 \ldots \times 10^n, \quad a_1 \neq 0,$$

тада број $\bar{x}$ има k сигурних цифара, ако важи

$$|x - \bar{x}| \leqslant 5 \times 10^{n-k}.$$

Из записа броја x видимо да важи

$$a_1 \times 10^n \leqslant |x| \leqslant (a_1 + 1) \times 10^n,$$

па важи

$$\frac{|x - \bar{x}|}{|x|} \leqslant \frac{5 \times 10^{n-k}}{a_1 \times 10^n} = \frac{5}{a_1} 10^{-k} \leqslant 5 \cdot 10^{-k}.$$

Обрнуто, нека је граница релативне грешке највише 5×10^{-k}. Тада важи

$$|x - \bar{x}| \leqslant 5 \cdot |x| \cdot 10^{-k} \leqslant 5(a_1 + 1) \cdot 10^{n-k} \leqslant 5 \cdot 10^{n-k+1},$$

па $\bar{x}$ има бар $k - 1$ сигурну цифру броја x.

Коначно, приближан број сигурних цифара је

$$-\log_{10} \delta(\bar{x}).$$

4 Нумеричко решавање система линеарних једначина

Посматрајмо систем линеарних једначина

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

где су коефицијенти a_{ij} и слободни чланови b_i $(i, j = 1, \dots, n)$ тачни или приближни бројеви. Овај систем се може написати и у матричном облику $Ax = b$, где је

$$
A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.
$$

Ако је $b = \vec{0}$ систем је хомоген и има нетривијално решење ако и само ако је $\det A = 0$. Иначе је систем нехомоген и има јединствено решење ако и само ако је $\det A \neq 0$. Тада се решење система може наћи из Крамерових формула

$$
x_i = \frac{\Delta_i}{\Delta}, \quad i = 1, \dots, n,
$$

где је Δ детерминанта система, а Δ_i су детерминанте променљивих x_i. Овај приступ захтева много рачунских опеарација - за (квадратни) систем димензије n број рачунских операције је $O((n+1)!)$. Сличан проблем се јавља и ако се решење тражи преко инверзне матрице $x = A^{-1}b$.

Нумеричке методе за решавање система линеарних једначина се деле на *тачне* и *итеративне*. *Тачна метода* даје решење након коначног броја аритметичких операција. *Итеративна метода* даје решење као граничну вредност низа узастопних апроксимација, где важну улогу има и брзина конвергенције итеративног процеса.

Ако су a_{ij} и b_i тачни бројеви (или приближни са најмање m тачних децимала), да би решење било тачно са m децимала треба рачунати са $m+1$ децимала (а за велики број непознатих са $m+2$ или $m+3$ децимале).

4.1 Гаусов метод са избором главног елемента

Гаусов метод се састоји у елиминисању једне по једне непознате из система, док матрица система не буде троугаона - директан ход. Обрнут ход је израчунавање непознатих x_i. Укупан број потребних аритметичких операција је $O(n^3)$.

Проблем са Гаусовим методом може настати ако је коефицијент уз елемент којим елимишемо елементе испод њега близак нули: тада може доћи до губитка сигурних цифара јер делимо бројем блиским нули. Зато се приликом нумеричког решавања система линеарних једначина Гаусовим методом често захтева да елемент матрице A на главној дијагонали којим се елиминишу елементи испод њега буде по апсолутној вредности што већи и он се назива *пивот*. У зависности од начина избора пивота разликујемо Гаусов метод са

- *парцијалним пивотирањем* (захтева се да је пивот по апсолутној вредности већи од елемената испод њега у истој колони - дозвољена је замена врста), или
- *тоталним пивотирањем* (захтева се да је пивот највећи по апсолутној вредности у његовој врсти и свим врстама после његове врсте почев од његове колоне - дозвољена је замена врста и замена колона, па се може променити редослед променљивих).

Видимо да ће тотално пивотирање повећати стабилност, на рачун доста комплекснијег алгоритама (врши се дводимензионо уместо једнодимензионог претраживања). Зато овде радимо само Гаусов метод са парцијалним пивотирањем, који најћешће буде довољно добар.

4.2 Итеративне методе за решавање система једначина

У пракси је често матрица система A ретка, што значи да има много нула, при чему ненула елементи могу бити правилно распоређени. Зато се алгоритми за решавање система прилагођавају да искористе такву структуру матрице, што се најлакше постиже ако се матрица A користи само као фактор. То нас мотивише да дефинишемо низ итерација $x^{(0)}, x^{(1)}, \ldots$ које ће конвергирати ка тачном решењу $x = A^{-1}b$, при чему се матрица A користи само као фактор.

Посматрајмо систем једначина $Ax = b$. Свако растављање $A = M - K$, за неку регуларну матрицу M, дефинише итеративну методу на следећи начин:

$$Mx - Kx = b \quad \Rightarrow \quad \underline{x} = M^{-1}Kx + M^{-1}b = \underline{Tx + c}, \quad T = M^{-1}K, \quad x = M^{-1}b.$$

Следи да је решење x фиксна тачка пресликавања $f(x) = Tx + c$, па дефинишемо итеративну методу

$$x^{(k+1)} = Tx^{(k)} + c, \quad k = 0, 1, \ldots \tag{4.1}$$

Теорема 4.1. Да би итеративни процес (4.1) конвергирао ка решењу система $Ax = b$ за све почетне векторе $x^{(0)}$ и све десне стране b потребно је и довољно да су све сопствене вредности матрице T по апсолутној вредности мање од 1.

Ипак, у пракси нам је значајније наредно тврђење.

Теорема 4.2. Итеративни процес (4.1) конвергира ка решењу система $Ax = b$ за све почетне векторе $x^{(0)}$ и све десне стране b ако је

$$\|T\| < 1,$$

при чему је $\|\cdot\|$ произвољна операторска норма (матрична норма индукована векторском нормом).

4.2.1 Критеријум заустављања и комплексност итеративне методе

Када посматрамо комплексност итеративне методе, занима нас колико итерација треба спровести да би се постигла жељена тачност ε.

Нека су k-тој итерацији грешка решења и грешка остатка

$$e_k = x - x^{(k)} \quad \text{и} \quad r_k = b - Ax^{(k)} = Ae_k,$$

редом. Циљ је наћи k тако да важи

$$\|e_k\| \leqslant \varepsilon,$$

али како не знамо тачно решење x овај критеријум није практичан. Уместо тога, критеријум заустављања ће бити

$$\|r_k\| \leqslant \varepsilon_r, \quad \text{за неко } \varepsilon_r > 0.$$

Ако нам тада треба процена грешке, важиће

$$\|e_k\| = \|A^{-1}r_k\| \leqslant \|A^{-1}\| \cdot \|r_k\| \leqslant \|A^{-1}\| \cdot \varepsilon_r.$$

Претпоставимо сада да је дата граница релативне грешке остатка ε. Како важи

$$\frac{\|e_k\|}{\|x\|} \leqslant k(A) \cdot \frac{\|r_k\|}{\|b\|}, \quad k(A) = \|A\| \cdot \|A^{-1}\|,$$

ако нађемо k за које важи

$$\frac{\|r_k\|}{\|b\|} \leqslant \varepsilon,$$

граница релативне грешке решења је

$$\frac{\|e_k\|}{\|x\|} \leqslant k(A) \cdot \varepsilon.$$

Посматрајмо сада итеративни процес (4.1). Из неједнакости

$$\|x^{(k)} - x\| \leqslant \|T\| \cdot \|x^{(0)} - x\|,$$

добијамо

$$-\log_{10} \frac{\|x^{(k)} - x\|}{\|x\|} \geqslant -k \log_{10} \|T\| - \log_{10} \frac{\|x^{(0)} - x\|}{\|x\|}.$$

Ако приметимо да је приближан број сигурних цифара у k-тој итерацији

$$z_k \approx -\log_{10} \frac{\|x^{(k)} - x\|}{\|x\|},$$

добијамо да је број итерација потребних да се постигне $z_k = m$ сигурних цифара, под претпоставком да почетна вредност $x^{(0)}$ има $z_0 = 0$ сигурних цифара

$$k \leqslant \frac{m}{-\log_{10} \|T\|}.$$

Важи и генерална оцена

$$\|x^{(k+1)} - x\| \leqslant \frac{\|T\|}{1 - \|T\|} \|x^{(k+1)} - x^{(k)}\|.$$

4.2.2 Гаус-Зајделова метода

Гаус-Зајделова метода је специјални случај претходне методе, при чему је $A = L + U$, за регуларну доње-троугаону матрицу L и строго горње-троугаону матрицу U. Тада је $(L + U)x = b$, одакле је $x = -L^{-1}Ux + L^{-1}b$, па је овај метеод дефинисан итерацијом (4.1) за $T = -L^{-1}U$ и $c = L^{-1}b$.

Иако се Гуас-Зајделова метода може применити на сваку матрицу A са ненула елементима на главној дијагонали, конвергенција је гарантована само ако је матрица $A = [a_{ij}]_{i,j=1}^{n}$:

- *строго дијагонално доминантна*, тј. $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ за свако i, или
- симетрична и *позитивно дефинитна*, тј. $v^T A v > 0$ за сваки ненула вектор $v \in \mathbb{R}^n$.

Рада Мутавџић Ђукић