Рада Мутавџић Ђукић

3 Елементи теорије грешака

3.1 Појам и врсте грешака

Посматрајмо проблем приближног одређивања величине $y = A(x)$, где је оператор A такав да се решење не може експлицитно написати или тачно израчунати (нпр. $y = \int_a^b x(t)dt$, при чему овај интеграл који није елементарно израчунљив):

- можемо заменити функцију x полиномом или неком другом функцијом $\bar{x}$ чији интеграл знамо да израчунамо или
- можемо заменити интеграл бесконачном сумом - оператор A се замењује блиским оператором $\bar{A}$ да би се израчунала вредност $\bar{y} = \bar{A}(\bar{x})$.

Грешком се оцењује колико је приближно решење $\bar{y}$ блиско тачном решењу y.

Разликујемо следеће типове грешака:

- *неотклоњива грешка* (настаје због недостатка математичког модела или грешке улазних података);
- *грешка методе* (настаје јер се оператор или улазне величине замењују приближним);
- *рачунска грешка или грешка заокруживања*, која ће бити детаљније описана у наредној подсекцији.

3.2 Приближни бројеви

Задавање приближних бројева

Наводимо различите начине задавања приближних бројева.

1. *Запис у фиксном зарезу* је одређен природним бројевима n_1 и n_2 тако да се број записује са n_1 цифара испред и n_2 цифара иза децималне тачке.

   Нпр. ако је $n_1 = 4$ и $n_2 = 6$, запис броја 12.4453 у фиксном зарезу је $\boxed{0012|445300}$.

2. *Запис у покретном зарезу* - положај децималне (бинарне) тачке није фиксиран, већ се он у односу на прву цифру записа одређује задавањем експонента,

   $$a = p \cdot 10^q,\ (a = p \cdot 2^q), \quad \text{где је } |p| < 1 \textit{ мантиса и } q \in \mathbb{Z} \textit{ експонент.}$$

   Број цифара мантисе t и број цифара експонента e су фиксирани.

   Нпр. ако је $t = 7$ и $e = 2$ записи у покретном зарезу броја 12.4453 могу бити $\boxed{1244530|02}$ или $\boxed{0124453|03}$.

   Да би запис у покретном зарезу био јединствен, дефинише се *нормализовани* запис броја у покретном зарезу - онај у коме је прва цифра мантисе различита од нуле (у нашем примеру први запис је нормализован).

   Уклико број има више од t цифара, његов нормализован запис представља само приближну вредност датог броја.

Апсолутна и релативна грешка

Ако је a тачна вредност неке велижине, а $\bar{a}$ њена приближна вредност, онда су

$$|a-\bar{a}| \quad \textit{апосолутна грешка} \qquad \text{и} \qquad \frac{|a-\bar{a}|}{a} \quad \textit{релативна грешка}$$

и границе апосултне и релативне грешке су редом

$$|a-\bar{a}| \leqslant \Delta(\bar{a}) \qquad \text{и} \qquad \frac{|a-\bar{a}|}{a} \leqslant \delta(\bar{a}).$$

Процентуална грешка је $\delta(\bar{a}) \cdot 100$, а промилна $\delta(\bar{a}) \cdot 1000$.

Пошто тачна вредност броја a обично није позната, у пракси се као граница релативне грешке користи количник

$$\delta(\bar{a}) = \frac{\Delta(\bar{a})}{|\bar{a}|}.$$

Значајне и сигурне цифре приближног броја

Значајне цифре неког броја су све цифре његовог записа, полазећи од прве ненула цифре са леве стране. То значи, ако је у запису броја

$$\bar{a} = \alpha_1 \cdot 10^n + \alpha_2 \cdot 10^{n-1} + \cdot + \alpha_m \cdot 10^{n-m+1} \tag{3.1}$$

при чему је $\alpha_1 \neq 0$, онда су све цифре $\alpha_1, \alpha_2, \ldots, \alpha_m$ значајне.

Пример 3.1. У броју $\bar{a} = 0.00443100$ све цифре изузев прве три нуле су значајне - а те нуле нису значајне јер број може да се написе и без њих, нпр. нормализован запис у покретном зарезу је $\bar{a} = 0.443100 \cdot 10^{-2}$. Последње две нуле јесу значајне, јер указују на тачност са којом је број дат.

Значајна цифра α_k броја $\bar{a}$ у изразу (3.1) је *сигурна цифра*, ако апсолутна грешка броја није већа од декадне јединице која одговара тој цифри (10^{n-k+1}) помножене задатим тежинским фактором ω, тј. треба да важи

$$\Delta(\bar{a}) \leqslant \omega \cdot 10^{n-k+1}, \quad 0 < \omega \leqslant 1. \tag{3.2}$$

При томе, ако је:

- $\omega \leqslant \frac{1}{2}$, цифра је *сигурна у ужем смислу*,
- $\frac{1}{2} < \omega \leqslant 1$, цифра је *сигурна у ширем смислу.*

Ако је цифра α_k сигурна, онда су и све цифре $\alpha_1, \ldots, \alpha_{k-1}$ сигурне.

Пример 3.2. Ако се зна да је $\Delta(\bar{a}) = 0.5 \cdot 10^{-5}$ апсолутна грешка приближног броја $\bar{a} = 0.00443100$, онда су сигурне цифре $4, 4, 3$. Последње три цифре $1, 0, 0$ нису сигурне јер је обзиром на апсолутну грешку $0.00442600 \leqslant a \leqslant 0.00443600$, па се последње три цифре броја a и $\bar{a}$ могу разликовати.

Приликом одбацивања цифара које нису сигурне, последња цифра се мења тако а буде сигурна у ужем смислу:

последња сигурна цифра α_k се неће мењати ако је:

- $\alpha_{k+1} < 5$ или
- $\alpha_{k+1} = 5$, $\alpha_m = 0$ за $m > k+1$, а α_k је парно.

У осталим случајевима α_k се повећава за 1 (приметимо да је ово стандардно заокруживање бројева). У нашем примеру, после одбацивања цифара које нису сигурне је $\bar{a} = 0.00443$.

Пример 3.3. Заокружимо број $\bar{a} = 12.336$, $\Delta(\bar{a}) = 0.03$ тако да му све цифре буду сигурне у ужем смислу.

Како је $\Delta(\bar{a}) = 0.03 < 0.05 = \frac{1}{2} \cdot 10^{-1}$ и $n = 1$, из $n - k + 1 = -1$ добијамо да је $k = 3$, па су сигурне цифре броја $\bar{a}$ у ужем смислу $1, 2, 3$.

Заокружимо број $\bar{a}$ на број са три цифре, тј. $\bar{a}_1 = 12.3$. Грешка која се јавља при овом заокруживању је $\Delta_1 = 0.036$, па је приближан број $\bar{a}_1$ дат са грешком $\Delta(\bar{a}_1) = \Delta(\bar{a}) + \Delta_1 = 0.03 + 0.036 = 0.066 > 0.05$, то броју $\bar{a}_1$ последња цифра 3 није сигурна у ужем смислу.

Заокружимо сада број $\bar{a}$ на број са две сигурне цифре, тј. на број $\bar{a}_2 = 12$. Грешка која се јавља при овом заокруживању је 0.336, па је приближан број $\bar{a}_2$ дат са грешком $\Delta(\bar{a}_2) = \Delta(\bar{a}) + \Delta_2 = 0.336 + 0.03 = 0.366 < 0.5$, па су у заокруженом броју $\bar{a}_2 = 12$ обе цифре сигурне у ужем смислу.

Између броја сигурних цифара и релативне грешке постоји следећа веза

$$\frac{\omega}{(\alpha_1+1)10^k} < \delta(\bar{a}) \leqslant \frac{\omega}{\alpha_1 10^{k-1}}, \quad 0 < \omega \leqslant 1, \tag{3.3}$$

где је k број сигурних цифара броја $\bar{a}$, а α_1 његова прва сигурна цифра.

Стабилност нумеричког алгоритма

Рачунање са приближним бројевима утиче на грешку коначног резултата.

Нумерички алгоритам је *стабилан*, ако се рачунска грешка не акумулира. Иначе је *нетабилан* и тада се јавља велика грешка коначног резултата.

Често је узрок нестабилности нумеричких алгоритама губитак сигурних цифара, до кога долази одузимањем блиских бројева.

Пример 3.4. Мањи корен квадратне једначине $x^2-140x+1=0$ је $a=70-\sqrt{4899}$. У запису са 4 сигурне цифре је $\sqrt{4899}\approx 69.99$, па је приближна вредност овог корена $\bar{a}=0.01$. Добијени резултат је са само једном сигурном цифром и релативна грешка је $\delta(\bar{a})=1=100\%$ (примена десне неједнакости у (3.3) за $\omega=1$), што значи да је коришћени алгоритам нестабилан.

Стабилан алгоритам за рачунање овог корена је

$$a=\frac{70^2-4899}{70+\sqrt{4899}}=\frac{1}{70+\sqrt{4899}}\approx\frac{1}{70+69.99}=0.007143,$$

омогућава добијање резултата на 4 сигурне цифре, са релативном грешком $\delta(\bar{a})=1.4\cdot 10^{-4}$.

3.3 Грешке приближних вредности функција

Нека је y функција параметара $\mathbf{a} = (a_1, \ldots, a_n) \in D$, $y = y(\mathbf{a})$ и нека је $\bar{y}$ приближна вредност за y. Апсолутна и релативна грешка величине $\bar{y}$ су редом

$$A(\bar{y}) = \sup_{\mathbf{a} \in D} |y(\mathbf{a}) - \bar{y}| \qquad \text{и} \qquad r(\bar{y}) = \frac{A(\bar{y})}{|\bar{y}|}.$$

Ако је област D n-димензиони правоугаоник $|a_k - \bar{a}_k| \leqslant \Delta(\bar{a}_k)$, $k = 1, \ldots, n$, где је $\bar{y} = y(\bar{\mathbf{a}})$, $\bar{\mathbf{a}} = (\bar{a}_1, \ldots, \bar{a}_n)$ и функција y је непрекидно диференцијабилна, на основу Лагранжове теореме (о средњој вредности функције) важи

$$y(a_1, \ldots, a_n) - \bar{y} = \sum_{k=1}^{n} \frac{\partial y}{\partial a_k}(c_1, \ldots, c_n)(a_k - \bar{a}_k), \qquad c_k = \bar{a}_k + \theta(a_k - \bar{a}_k),\ k = 1, \ldots, n.$$

Следи да је

$$A(\bar{y}) \leqslant \sum_{k=1}^{n} \sup_{\mathbf{a} \in D} \left| \frac{\partial y}{\partial a_k}(\mathbf{a}) \right| \Delta(\bar{a}_k). \tag{3.4}$$

У пракси се уместо ове оцене користи *линеарна оцена апсолутне грешке функције*

$$\Delta(\bar{y}) = \sum_{k=1}^{n} \left| \frac{\partial y}{\partial a_k}(\bar{\mathbf{a}}) \right| \Delta(\bar{a}_k), \tag{3.5}$$

при чему важи

$$\Delta(\bar{y}) + \varepsilon_1(\rho) \leqslant A(\bar{y}) \leqslant \Delta(\bar{y}) + \varepsilon_2(\rho), \quad \rho = \sqrt{\Delta^2(\bar{a}_1) + \cdots + \Delta^2(\bar{a}_n)},\ \varepsilon_j = o(\rho),\ j = 1, 2,$$

што значи да је оцена (3.5) задовољавајућа за мале апсолутне грешке аргумента.

Ако су уместо горњих граница апсолутне грешке $\Delta(\bar{a}_k)$ дате дорње границе релативне грешке $r(\bar{a}_k) = \Delta(\bar{a}_k)/|\bar{a}_k|$, имамо

$$\delta(\bar{y}) = \sum_{k=1}^{n} \left| \frac{\bar{a}_k}{\bar{y}} \frac{\partial y}{\partial a_k}(\bar{\mathbf{a}}) \right| r(\bar{a}_k). \tag{3.6}$$

Пример 3.5. Одредити грешку вредности функције $y = x^{10}$ за $\bar{a} = 1$ и $\Delta(\bar{a}) = 10^{-3}$.

Прво налазимо

$$\bar{y} = 1, \quad \frac{\partial y}{\partial x}(1) = 10,$$

па на основу формуле (3.5) добијамо

$$\Delta(\bar{y}) = \left| \frac{\partial y}{\partial x}(1) \right| \Delta(\bar{a}) = 10^{-3} = 0.01.$$

Пример 3.6. Одредити број сигурних цифара у вредности функције $z = \cos \frac{y}{x}$, за $\bar{x} = 2.345$ и $\bar{y} = 3.4567$, ако су све написане цифре сигурне у ужем смислу.

Како су дате цифре сигурне у ужем смислу, то су границе апсолутних грешака аргумената $\Delta(\bar{x}) = 0.5 \cdot 10^{-3}$ и $\Delta(\bar{y}) = 0.5 \cdot 10^{-4}$. Нађимо још парцијалне изводе дате функције:

$$\frac{\partial z}{\partial x}(\bar{x}, \bar{y}) = \frac{\bar{y}}{\bar{x}^2} \cdot \sin \frac{\bar{y}}{\bar{x}} \approx 0.63, \quad \frac{\partial z}{\partial y}(\bar{x}, \bar{y}) = -\frac{1}{\bar{x}} \cdot \sin \frac{\bar{y}}{\bar{x}} \approx -0.42.$$

Коначно, граница апсолутне грешке приближне вредности функције је према формули (3.5)

$$\Delta(\bar{z}) = 0.63 \cdot 0.5 \cdot 10^{-3} + 0.42 \cdot 0.5 \cdot 10^{-4} = 0.34 \cdot 10^{-3} < 0.5 \cdot 10^{-3},$$

па је приближна вредност функције $\bar{z} = \sin \frac{\bar{x}}{\bar{y}} = 0.096573... \approx 0.097$ дата са две сигурне цифре (након заокруживања на две значајне цифре добијамо $\bar{\bar{z}} = 0.097$ са грешком заокруживања $\Delta \leqslant 0.5 \cdot 10^{-3}$, па је за приближну вредност $\bar{\bar{z}}$ граница апсолутне грешке $\Delta(\bar{z}) + \Delta \leqslant 10^{-3}$).

Из оштег израза за грешку функције се могу оценити грешке које настају при стандардним операцијама са приближним бројевима.

1. Линеарна оцена апсолутне грешке збира или разлике једнака је збиру апсолутних грешака аргумента.

2. Линеарна оцена релативне грешке производа или количника једнака је суми релативних грешака аргумената.

3.4 Обратан (инверзан) проблем грешке

Обртан проблем грешке представља налажење допустивих грешака аргумената, при којима грешка функције не прелази дозвољену вредност. Задатак је једнозначно решив само за функцију једног аргумента $y = y(x)$. Ако је та функција диференцијабилна, онда је

$$y - \bar{y} = y'(c)(x - \bar{x}), \qquad c = \bar{a} + \theta(a - \bar{a}), \quad 0 \leqslant \theta \leqslant 1,$$

па је за $y'(c) \neq 0$

$$x - \bar{x} = \frac{y - \bar{y}}{y'(c)}.$$

Граница апсолутне грешке је приближно одређена релацијом

$$\Delta(\bar{x}) = \frac{\Delta(\bar{y})}{|y'(\bar{x})|}, \qquad y'(\bar{x}) \neq 0.$$

Ако је y функција више оромењљивих, $y = y(x_1, \ldots, x_n)$, онда се задавањем грешке функције задаје само једна веза између n непознатих $\Delta(\bar{x}_1), \ldots, \Delta(\bar{x}_n)$. Ако је задата линеарна оцена апсолутне грешке функције (3.5), додатни услови које апсолутне грешке аргумената треба да задовољавају дефинишу се на један од следећих начина.

(*i*) *Принцип једнаких утицаја*: $\left|\frac{\partial y}{\partial x_1}(\bar{\mathbf{x}})\right| \Delta(\bar{x}_1) = \cdots = \left|\frac{\partial y}{\partial x_n}(\bar{\mathbf{x}})\right| \Delta(\bar{x}_n)$. Тада је

$$\Delta(\bar{y}) = n\left|\frac{\partial y}{\partial x_k}(\bar{\mathbf{x}})\right| \Delta(\bar{x}_k), \quad \text{па је} \quad \Delta(\bar{x}_k) = \frac{\Delta(\bar{y})}{n\left|\frac{\partial y}{\partial x_k}(\bar{\mathbf{x}})\right|}, \quad k = 1, \ldots, n.$$

(*ii*) *Принцип једнаких апсолутних грешака*: $\Delta(\bar{x}_1) = \cdots = \Delta(\bar{x}_n)$. Тада је

$$\Delta(\bar{y}) = \Delta(\bar{x}_k) \sum_{j=1}^{n} \left|\frac{\partial y}{\partial x_j}(\bar{\mathbf{x}})\right|, \quad \text{па је} \quad \Delta(\bar{x}_k) = \frac{\Delta(\bar{y})}{\sum_{j=1}^{n}\left|\frac{\partial y}{\partial x_j}(\bar{\mathbf{x}})\right|}, \quad k = 1, \ldots, n.$$

Овај принцип се може применити само ако су променљиве исте димензије.

(*iii*) *Принцип једнаких релативних грешака*: $\delta(\bar{x}_1) = \cdots = \delta(\bar{x}_n)$. Тада је

$$\Delta(\bar{y}) = \frac{\Delta(\bar{x}_k)}{|\bar{x}_k|} \sum_{j=1}^{n} \left|\bar{x}_j \frac{\partial y}{\partial x_j}(\bar{\mathbf{x}})\right|, \quad \text{па је} \quad \Delta(\bar{x}_k) = \frac{\Delta(\bar{y})|\bar{x}_k|}{\sum_{j=1}^{n}\left|\bar{x}_j \frac{\partial y}{\partial x_j}(\bar{\mathbf{x}})\right|}, \quad k = 1, \ldots, n.$$

Одредити са коликом тачношћу треба наћи променљиве x, y и z да би се величина $f(x, y, z) = \frac{xy + \sqrt{z}}{x + 2z}$ израчунала са тачношћу 10^{-3}, ако су приближне вредности аргумената $\bar{x} = 2.16$, $\bar{y} = 1.12$ и $\bar{z} = 1.44$, при чему се користи принцип једнаких утицаја на грешку.

Из услова задатка треба да важи

$$\Delta(\bar{f}) = \left|\frac{\partial f}{\partial x}(\bar{x}, \bar{y}, \bar{z})\right| \Delta(\bar{x}) + \left|\frac{\partial f}{\partial y}(\bar{x}, \bar{y}, \bar{z})\right| \Delta(\bar{y}) + \left|\frac{\partial f}{\partial z}(\bar{x}, \bar{y}, \bar{z})\right| \Delta(\bar{z}) < 10^{-3},$$

при чему је

$$\left|\frac{\partial f}{\partial x}(\bar{x}, \bar{y}, \bar{z})\right| \Delta(\bar{x}) = \left|\frac{\partial f}{\partial y}(\bar{x}, \bar{y}, \bar{z})\right| \Delta(\bar{y}) = \left|\frac{\partial f}{\partial z}(\bar{x}, \bar{y}, \bar{z})\right| \Delta(\bar{z}),$$

одакле добијамо

$$\Delta(\bar{x}) \leqslant \frac{10^{-3}}{3\left|\frac{\partial f}{\partial x}(\bar{x}, \bar{y}, \bar{z})\right|} < 0.00418, \quad \Delta(\bar{y}) \leqslant \frac{10^{-3}}{3\left|\frac{\partial f}{\partial y}(\bar{x}, \bar{y}, \bar{z})\right|} < 0.00077, \quad \Delta(\bar{z}) \leqslant \frac{10^{-3}}{3\left|\frac{\partial f}{\partial z}(\bar{x}, \bar{y}, \bar{z})\right|} < 0.00164.$$

3.5 Условљеност проблема

За пресликавање $y = f(x)$ занима нас како се мале промена улазних вредности x одражавају на промене излазних вредности y, тј. осетљивост пресликавања f у некој датој тачки на мале промене x. Степен те условљености мери *фактор условљености* или *кондициони број* пресликавања f у тачки x. При томе претпостављамо да се функција израчунава тачно, па фактор условљености не зависи од алгоритма којим се функција израчунава.

Нека је функција f диференцијабилна (у некој околини тачке x). Тада за промену Δy на основу Тејлорове формуле важи

$$\Delta y \approx f'(x)\Delta x,$$

па за релативне грешке важи

$$\frac{\Delta y}{y} \approx \frac{xf'(x)}{f(x)} \cdot \frac{\Delta x}{x.}$$

Ова апроксимација постаје тачна када је f линеарна функција или када $\Delta x \to 0$. На основу ове апроксимације, условљеност функције f у тачки x је

$$(\operatorname{cond} f)(x) = \left|\frac{xf'(x)}{f(x)}\right|$$

и мери колико је пута релативна промена y већа од релативне промене x. Што је овај број мање (већи), проблем је боље (слабије) условљен. Ако је $x = 0$ и $y \neq 0$ фактор условљености дефинишемо као $|f'(x)/f(x)|$, ако је $y = 0$ и $x \neq 0$, тада је $\operatorname{cond} f(x) = |xf'(x)|$, а ако је $x = y = 0$, тада је $(\operatorname{cond} f)(x) = |f'(x)|$.

Претпоставимо сада да $f : x = (x_1, \ldots, x_m) \mapsto y = (y_1, \ldots, y_n)$, тј.

$$y_i = f_i(x_1, \ldots, x_m), \quad i = 1, \ldots, n.$$

Јакобијева матрица пресликавања f је

$$J = \left\|\frac{\partial f}{\partial x}\right\| = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}.$$

Да бисмо добили јединствен фактор условљености, морамо дефинисати неку норму матрице.

3.6 Норме матрице и вектора

Норма вектора $x = (x_1, \ldots, x_n)$ је број $\|x\|$ за који важи

- $\|x\| > 0$ за $x > 0$ и $\|x\| = 0 \Leftrightarrow x = \vec{0}$,
- $\|c \cdot x| = c \cdot \|x\|$, $c \in \mathbb{R}$,
- $\|x + y\| \leqslant \|x\| + \|y\|$.

Норма матрице $A = [a_{ij}]_{n\times n} \in \mathbb{R}^{n\times n}$ је број $\|A\|$ за који важи

- $\|A\| > 0$ за $A > 0$ и $\|A\| = 0 \Leftrightarrow A = \vec{0}$,
- $\|c \cdot A| = c \cdot \|A\|$, $c \in \mathbb{R}$,
- $\|A + B\| \leqslant \|A\| + \|B\|$,
- $\|A \cdot B\| \leqslant \|A\| \cdot \|B\|$.

Ако је $|a_{ij}| < \|A\|$ норма је канонска.

Норма матрице је *сагласна* са нормом вектора, ако за сваки вектор x важи

$$\|Ax\| \leqslant \|A\| \cdot \|x\|.$$

Специјално, за $\|x\| = 1$ треба да важи $\|Ax\| \leqslant \|A\|$. Најмања матрична норма за коју ово важи је

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

и за њу се каже да је норма матрице *индукована нормом вектора* или *операторска норма*.

Наводимо индуковане норме које се често користе: униформна или бесконачна норма, 1- норма и 2- норма или *еуклидска норма* су редом

$$\|x\|_\infty = \max_{1\leqslant j\leqslant n} |x_j|, \qquad \|A\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|,$$

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|, \qquad \|A\|_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}|,$$

$$\|x\|_2 = \left[\sum_{j=1}^{m} |x_j|^2\right]^{1/2}, \qquad \|A\|_2 = \max_{1\leqslant i\leqslant n} \lambda_i^{1/2},\ \lambda_i \text{ су сопствене вредности матрице} A^T A.$$

Вратимо се фактору условљености функције $f : \mathbb{R}^n \to \mathbb{R}^m$. Из Тејлорове формуле сада имамо

$$\Delta y_i \approx \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} \Delta x_j,$$

па је

$$|\Delta y_i| \leqslant \max_j |\Delta x_j| \max_i \left| \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} \right|, \quad \text{тј.} \quad \|\Delta y\|_\infty \leqslant \|\Delta x\|_\infty \cdot \|J\|_\infty.$$

За релативне промене онда важи

$$\frac{\|\Delta y\|_\infty}{\|y\|_\infty} \leqslant \frac{\|x\|_\infty \|J\|_\infty}{\|y\|_\infty} \cdot \frac{\|\Delta x\|_\infty}{\|x\|_\infty},$$

па је фактор условљености

$$(\operatorname{cond} f)(x) = \frac{\|x\|_\infty \|J\|_\infty}{\|y\|_\infty}.$$

На крају, посматрајмо условљеност регуларне матрице $A \in \mathbb{R}^{n\times n}$ која ће бити важна приликом решавања система једначина $AX = b$.

Дефиниција 3.1. Кондициони број несингуларне матрице A је $k(A) = \|A\| \cdot \|A\|^{-1}$.

Важи да је $k(A) \geqslant 1$ и ако је $k(A) \gg 1$, тада је матрица лоше условљена.

Ако је $A(x + \Delta x) = b + \Delta b$, тада важи

$$\frac{\|\Delta x\|}{\|x\|} \leqslant k(A) \cdot \frac{\|\Delta b\|}{\|b\|}.$$

Заиста, $b = Ax$ и $\Delta x = A^{-1}(b + \Delta b) - x = A^{-1}\Delta b$, па је

$$\|b\| \leqslant \|A\| \cdot \|x\| \quad \text{и} \quad \|\Delta x\| \leqslant \|A^{-1}\| \cdot \|\Delta b\|,$$

чијим множењем следи резултат.

3.7 Итеративне методе за решавање система једначина

Посматрајмо систем једначина $Ax = b$. Ако је $A = M - K$, за неку регуларну матрицу M, тада је $Mx - Kx = b$, па је $x = M^{-1}Kx + M^{-1}b = Tx + c$, где је $T = M^{-1}K$ и $x = M^{-1}b$. Следи да је решење x фиксна тачка пресликавања $f(x) = Tx + c$, па дефинишемо итеративну методу

$$x^{(k+1)} = Tx^{(k)} + c, \quad k = 0, 1, \dots \tag{3.7}$$

Теорема 3.1. Да би итеративни процес (3.7) 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 3.2. Итеративни процес (3.7) 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).