

1 Увод

Нумеричка математика је област математике која се бави развојем метода за приближно решавање математичких проблема када тачна решења нису лако доступна (што је у пракси скоро увек случај). Док се класична математика фокусира на испитивање услова под којима решења постоје и на анализу њихових особина, нумеричка математика се бави практичним аспектима њиховог проналажења са задатом тачношћу. Један од кључних изазова нумеричке математике дакле није само долазак до решења, већ и процена његове тачности и поузданости. Будући да се нумеричке методе увек користе апроксимацијама, важно је разумети колико се добијено решење разликује од стварног и који све фактори утичу на прецизност коначног резултата.

Колико је неко решење квалитетно говори грешка која представља његово одступање од тачног резултата. Узроци грешке могу бити различити, па се у односу на порекло разликују:

- неотклоњива грешка,
- грешка методе и
- рачунска грешка.

Пример 1 Наизменична (АС) струја мења свој смер и амплитуду у правилним интервалима. Напон у АС колима често се моделира са $V(t) = V_{max} \sin(\pi \nu t)$, где је V_{max} максималан напон у колу, а ν фреквенција алтернације. Претпоставимо да желимо да одредимо количину напона у неком струјном колу у тренутку $\hat{t}$. Треба да имамо на уму:

- Дати модел не описује физички процес идеално, јер у пракси увек постоји одређени шум (неотклоњива грешка)
- $\sin(\cdot)$ се на компјутеру/дигитрону рачуна приближно, на пример коришћењем коначног броја сабирака Тејлоровог реда (грешка методе)
- Компјутер/дигитрон располаже са коначном количином меморије за запис аргумената и међу-резултата, који се често не могу представити тачно унутар ових ограничења (грешка рачуна)

На наредним страницама биће речи о рачунским грешкама. У даљем току курса, при представљању различитих нумеричких метода, осврнућемо се и на грешке које оне могу произвести. Будући да је неотклоњива грешка пре свега повезана са математичким моделирањем, о њој нећемо даље дискутовати.

1.1 Приближне вредности

Како рачунар располаже са тачно одређеном, коначном количином меморије за репрезентацију сваког броја, многи бројеви се не могу тачно сачувати. Из овог разлога принуђени смо да чувамо и у рачуну користимо само њихове **приближне вредности**.

Дефиниција 1 Нека је x тачна вредност броја, а x^* његова приближна вредност. Апсолутна грешка приближне вредности x^* дефинише се са

$$\Delta x^* = |x - x^*|.$$

Како нам је x , па и самим тим Δx^* у пракси често непознато, уводимо кориснији појам границе апсолутне грешке $Ax^* \geq \Delta x^*$, одакле је

$$x^* - Ax^* \leq x \leq x^* + Ax^*.$$

Релативна грешка приближне вредности x^* дефинисана је са

$$\delta x^* = \frac{\Delta x^*}{|x|}.$$

Као и у случају апсолутне грешке, често је смисленије причати о граници релативне грешке. Иако се она теоријски дефинише са $Rx^* \geq \delta x^* = \Delta x^*/|x|$, пошто је x најчешће непознато, у пракси ћемо тражити $Rx^* \geq \Delta x^*/|x^*|$. За овако дефинисану границу важи

$$(1 - Rx^*)x^* \leq x \leq (1 + Rx^*)x^*.$$

Пример 2 Познато је да је $22/7$ добра и једноставна апроксимација броја π . Како је

$$|\pi - \frac{22}{7}| = |3.1415... - 3.1428...| \leq 0.0014,$$

можемо одредити границе апсолутне и релативне грешке ове апроксимације са

$$Ax^* = 1.4 \cdot 10^{-3} \quad Rx^* = 4.5 \cdot 10^{-4} \geq \frac{0.0014}{22/7}.$$

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

$$x^* = \pm(\alpha_1 + \dots + \alpha_k 10^{1-k} + \dots + \alpha_m 10^{1-m}) \cdot 10^n,$$

цифра α_1 различита од нуле, онда су све цифре $\alpha_1, \dots, \alpha_m$ значајне. За значајну цифру броја се каже да је **сигурна цифра** ако апсолутна грешка броја није већа од декадног чиниоца који одговара тој цифри, тј. α_k је сигурна цифра ако је

$$Ax^* = w \cdot 10^{n-k+1}, \quad w \in (0, 1].$$

Ако је $w \leq 1/2$, цифра је сигурна у ужем смислу, а ако је $w > 1/2$, сигурна је у ширем смислу. Ако је цифра α_k сигурна, онда су сигурне и $\alpha_1, \dots, \alpha_{k-1}$.

Подсетимо се правила заокруживања. Нека је $x^* = \pm(\beta_1 + \dots + \beta_m 10^{1-m}) \cdot 10^n$ број који се добија заокруживањем броја $x = \pm(\alpha_1 + \dots + \alpha_k 10^{1-k} + \dots) \cdot 10^n$ на m значајних цифара. Тада је $\beta_k = \alpha_k$, за $k < m$, док се цифра β_m одређује тако да:

- Ако је $\alpha_{m+1} < 5$, онда је $\beta_m = \alpha_m$.
- Ако је $\alpha_{m+1} > 5$, или $\alpha_{m+1} = 5$ и постоји цифра различита од нуле у даљем запису броја x , онда је $\beta_m = \alpha_m + 1$.
- Ако је $\alpha_{m+1} = 5$, и нема цифара различитих од нуле у даљем запису броја x , онда је $\beta_m = \alpha_m$, уколико је α_m парна цифра, а $\alpha_m + 1$ у супротном.

На основу датих правила није тешко показати да ће све цифре броја x^* , добијеног заокруживањем броја x , бити сигурне у ужем смислу.

Напомена: Цифре приближне вредности које нису сигурне нема много сврхе чувати, јер њиховим одстрањивањем (заокруживањем броја) неће доћи до значајног увећавања апсолутне грешке.

Обратимо пажњу на везу која постоји између броја сигурних цифара приближне вредности x^* и границе релативне грешке. Нека је x^* приближна вредност са m сигурних цифара. Из дефиниције знамо да је $Ax^* = w \cdot 10^{n-m+1}$, где је $w \in (0, 1]$. На основу овога је

$$Rx^* = \frac{w \cdot 10^{n-m+1}}{|x^*|} = \frac{w \cdot 10^{n-m+1}}{(\alpha_1 + \alpha_2 10^{-1} + \dots) \cdot 10^n} = \frac{w}{(\alpha_1 + \alpha_2 10^{-1} + \dots) 10^{m-1}}.$$

Имајући у виду да је $\alpha_1 + \alpha_2 10^{-1} + \dots \geq 1$ и начин на који је ограничено w , можемо закључити да је

$$Rx^* \leq \frac{1}{10^{m-1}}.$$

Пример 3 Нека је $x = \frac{3000}{7}$. Ако је $x_1^* = 428.55$, а x_2^* и x_3^* су добијени заокруживањем броја x на две, односно шест значајних цифара, одредити границе апсолутне и релативне грешке, као и број сигурних цифара ових приближних вредности.

Решење. Како је $x = 428.\overline{571428}$, применом правила заокруживања долазимо до

$$x_2^* = 4.3 \cdot 10^2 \text{ и } x_3^* = 428.571.$$

Како је

$$|x - x_1^*| \approx 0.0214, \quad |x - x_2^*| \approx 1.4285, \quad |x - x_3^*| \approx 0.0004,$$

можемо изабрати следеће вредности за границе апсолутних и релативних грешака.

i	1	2	3
Ax_i^*	$0.3 \cdot 10^{-1}$	$0.2 \cdot 10^1$	$0.5 \cdot 10^{-3}$
Rx_i^*	$7 \cdot 10^{-5}$	$5 \cdot 10^{-3}$	$2 \cdot 10^{-6}$

На основу дефиниције, јасно је да x_1^* има четири, x_2^* две, а x_3^* шест сигурних цифара.

1.2 Грешка приближних вредности функција

У претходном делу било је речи о приближним вредностима и грешкама које описују њихово одступање од тачних вредности. Природно питање које се следеће намеће јесте шта се дешава са грешком када се приближне вредности користе као аргументи функција.

Дефиниција 2 Нека је задата функција $f: \mathbb{R}^n \rightarrow R$. Нека је $x = (x_1, \dots, x_n)$ вектор тачних, а $x^* = (x_1^*, \dots, x_n^*)$ вектор приближних вредности. Апсолутна и релативна грешка функције f за приближну вредност x^* дефинисане су са

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

Границе апсолутне и релативне грешке, у ознаци $Af(x^*)$ и $Rf(x^*)$ дефинишу се као у случају приближних вредности.

Уколико се x^* подразумева, претходне вредности ћемо краће означавати са $\Delta f^*, \delta f^*, Af^*$ и Rf^* . Ако су нам за улазне аргументе познате границе апсолутних грешака Ax_i^* , јасно је да се за границу апсолутне грешке може узети

$$Af^* = \sup_{x \in G} |f(x) - f(x^*)|, \quad G = \{x : |x_i - x_i^*| \leq Ax_i^*, i = 1, \dots, n\}.$$

Под претпоставком да је f непрекидно диференцијабилна, на основу теореме о средњој вредности знамо да за сваки пар (x, x^*) постоји $\theta \in (0, 1)$ такво да за $\hat{x} = x + \theta(x^* - x)$ важи

$$\Delta f^* = |f(x) - f(x^*)| = \left| \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\hat{x}) \Delta x_i^* \right|,$$

па се за границу апсолутне грешке може узети и

$$Af^* = \sum_{i=1}^n \sup_{\hat{x} \in G} \left| \frac{\partial f}{\partial x_i}(\hat{x}) \right| Ax_i^*.$$

Како је често незгодно тачно одредити $\sup_{\hat{x} \in G} \left| \frac{\partial f}{\partial x_i}(\hat{x}) \right|$, у пракси ћемо се задовољити такозваном **линеарном оценом границе апсолутне грешке**, датом са

$$Af^* = \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x^*) \right| Ax_i^*.$$

Напомена: Иако за овако дефинисану границу не мора нужно да важи $Af^* \geq \Delta f^*$, може се показати да је наведена оцена задовољавајућа за мале апсолутне грешке аргумената.

Одавде се директно изводи и **линеарна оцена границе релативне грешке**, која је дата са

$$Rf^* = \frac{Af^*}{|f(x^*)|} = \frac{1}{|f(x^*)|} \sum_{i=1}^n \left| \frac{x_i^*}{x_i^*} \right| \cdot \left| \frac{\partial f}{\partial x_i}(x^*) \right| Ax_i^* = \frac{1}{|f(x^*)|} \sum_{i=1}^n \left| x_i^* \frac{\partial f}{\partial x_i}(x^*) \right| Rx_i^*.$$

Искористимо претходно наведене оцене и испитајмо какву грешку можемо очекивати код примене стандардних операција $(+, -, \cdot, \div)$ на приближне вредности. Како бисмо проценили грешку која настаје при сабирању/одузимању приближних вредности, дефинишимо функцију $s : \mathbb{R}^n \rightarrow \mathbb{R}$ са

$$s(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n, \quad a_i \in \{-1, 1\}.$$

Лако се израчунава да су линеарне оцене граница апсолутне и релативне грешке за функцију s

$$As^* = \sum_{i=1}^n |a_i| Ax_i^* = \sum_{i=1}^n Ax_i^* \quad \text{и} \quad Rs^* = \frac{\sum_{i=1}^n Ax_i^*}{|s(x^*)|} = \frac{\sum_{i=1}^n |x_i^*| Rx_i^*}{|s(x^*)|}.$$

Није тешко показати да је за функцију s испуњено $\sum_{i=1}^n Ax_i^* = \sup_{x \in G} |s(x) - s(x^*)|$, па су овако дефинисане As^* и Rs^* заиста границе грешки, односно важи $As^* \geq \Delta s^*$ и $Rs^* \geq \delta s^*$, при чему се једнакости могу достићи. Из ових израза можемо закључити следеће:

- Граница апсолутне грешке збира/разлике једнака је збиру граница апсолутних грешака аргумената. Одавде је јасно да ће доминантан утицај на крајњу грешку имати бројеви задати са најлошијом прецизношћу (малим бројем децимала). Из овог разлога у реализацији рачуна бројеви задати са већом прецизношћу могу се заокружити на једну децималу више од најмањег броја децимала без великог губитка прецизности крајњег резултата.
- Релативна грешка може бити висока ако је крајњи резултат мали по апсолутној вредности - на пример када се одузимају два броја блиских вредности. Овај проблем се некад може избећи алгебарским манипулацијама израза.

Како бисмо проценили грешку која настаје при множењу/дељењу приближних вредности, дефинишимо функцију $p : \mathbb{R}^n \rightarrow \mathbb{R}$ са

$$p(x_1, \dots, x_n) = x_1^{a_1} \dots x_n^{a_n}, \quad a_i \in \{-1, 1\}.$$

Имајући у виду да је

$$\frac{\partial p}{\partial x_i}(x) = \frac{p(x)}{x_i^{a_i}} a_i x_i^{a_i-1} = a_i \frac{p(x)}{x_i},$$

линеарне оцене граница апсолутне и релативне грешке за функцију p су

$$Ap^* = \sum_{i=1}^n \left| a_i \frac{p(x^*)}{x_i^*} \right| Ax_i^* = |p(x^*)| \sum_{i=1}^n \frac{Ax_i^*}{|x_i^*|} = |p(x^*)| \sum_{i=1}^n Rx_i^* \quad \text{и}$$

$$Rp^* = \frac{Ap^*}{|p(x^*)|} = \sum_{i=1}^n Rx_i^*.$$

Видимо да је однос који је постојао између апсолутних грешки у случају сабирања/одузимања, код множења/дељења прешао на релативну грешку.

- Граница релативне грешке производа/количника једнака је збиру релативних грешака аргумената. У овом случају, доминантан утицај на релативну грешку ће имати чиниоци са великом релативном грешком, односно малим бројем сигурних цифара. За спровођење рачуна ће дакле често бити довољно задржавати једну више сигурну цифру (код бројева који их имају више) од најмањег броја сигурних цифара.
- За разлику од случаја сабирања/одузимања, у овом случају је апсолутна грешка та која може претрпети велико увећање и то када је крајњи резултат велик по модулу.

Пример 4 Одредити површину ваљка као и границе апсолутне и релативне грешке, ако је

$$r = 3.7 \pm 0.5 \cdot 10^{-1}, \quad H = 8.2 \pm 0.3 \cdot 10^{-1}, \quad \pi = 3.142 \pm 0.5 \cdot 10^{-3}.$$

Подразумеван начин за читање горе наведених једначина је $x = x^* \pm Ax^*$. Површина ваљка задата је формулом $P = 2r^2\pi + 2r\pi H = 2r\pi(r + H)$, па је очигледна процена тачне површине

$$P^* = 2r^*\pi^*(r^* + H^*) = 2 \cdot 3.7 \cdot 3.142 \cdot (3.7 + 8.2) = 276.68452.$$

На основу линеарних оцена граница грешки долазимо до

$$RP^* = \frac{0.05}{3.7} + \frac{0.0005}{3.142} + \frac{0.05 + 0.03}{3.7 + 8.2} = 0.0204 \quad \text{и} \quad AP^* = |P^*|RP^* = 5.6431.$$

До алтернативне процене можемо доћи заокруживањем π на 3 значајне цифре у производу (јер постоје чиниоци са само 2 значајне цифре). На овај начин се долази до приближног резултата $P^{**} = 276.5084$, чије су границе грешки

$$RP^{**} = \frac{0.05}{3.7} + \frac{0.0005 + 0.0002}{3.14} + \frac{0.05 + 0.03}{3.7 + 8.2} = 0.0205 \quad \text{и} \quad AP^{**} = |P^{**}|RP^{**} = 5.6571.$$

Додатно, крајњи резултат можемо заокружити тако да су му све цифре сигурне. Обзиром да је $1 < AP^{**} \leq 10$, последња сигурна је цифра десетица, па је $P^{***} = 28 \cdot 10^1 = 280$, док су нове границе грешке

$$AP^{***} = AP^{**} + |280 - 276.5084| = 9.1487 \quad \text{и} \quad RP^{***} = \frac{AP^{***}}{P^{***}} = 0.0323.$$

Примећујемо да се у овом случају граница релативне грешке није значајно увећала и поред примењених заокруживања.

1.3 Обратни проблем грешке

Обратни проблем грешке подразумева одређивање граница грешака аргумената које гарантују да ће грешка функције остати прихватљиве величине ε . Проблем се једнозначно може решити само у случају функције једног аргумента. На основу линеарне оцене границе апсолутне грешке

$$Af^* = \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x^*) \right| Ax_i^*,$$

видимо да за функцију једног аргумента ($n = 1$) важи $Af^* = |f'(x^*)|Ax^*$, па закључујемо

$$Af^* \leq \varepsilon \iff Ax^* \leq \frac{\varepsilon}{|f'(x^*)|}.$$

У случају када функција има више аргумената, за одређивање тражених граница неопходно је наметнути додатне услове које границе апсолутних грешки аргумената треба да задовоље. Ови услови познати су и као принципи, а најчешће коришћени су:

1. Принцип једнаких утицаја

Овај принцип подразумева да је испуњено

$$\left| \frac{\partial f}{\partial x_1}(x^*) \right| Ax_1^* = \dots = \left| \frac{\partial f}{\partial x_n}(x^*) \right| Ax_n^*.$$

На основу ове претпоставке закључујемо

$$Af^* = n \left| \frac{\partial f}{\partial x_k}(x^*) \right| Ax_k^* \leq \varepsilon \iff Ax_k^* \leq \frac{\varepsilon}{n \left| \frac{\partial f}{\partial x_k}(x^*) \right|}, \quad k = 1, \dots, n.$$

2. Принцип једнаких апсолутних грешака

Ако претпоставимо да су све границе апсолутних грешки једнаке, односно $Ax_1^* = \dots = Ax_n^* = Ax^*$, закључујемо

$$Af^* = Ax^* \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x^*) \right| \leq \varepsilon \iff Ax_k^* \leq \frac{\varepsilon}{\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x^*) \right|}, \quad k = 1, \dots, n.$$

3. **Принцип једнаких релативних грешака** Ако претпоставимо да су све границе релативних грешки једнаке, односно $Rx_1^* = \dots = Rx_n^* = Rx^*$, закључујемо

$$Af^* = Rx^* \sum_{i=1}^n \left| x_i^* \frac{\partial f}{\partial x_i}(x^*) \right| \leq \varepsilon \iff Ax_k^* = |x_k^*| Rx^* \leq \frac{|x_k^*| \varepsilon}{\sum_{i=1}^n \left| x_i^* \frac{\partial f}{\partial x_i}(x^*) \right|}, \quad k = 1, \dots, n.$$

Пример 5 Дата је функција $f(x_1, x_2, x_3) = x_1^2/x_2 - x_3$ и приближне вредности $x_1^* = 4.5227$, $x_2^* = 13.37$ и $x_3^* = 2.391$. Користећи принцип једнаких апсолутних вредности, одредити колике требају бити границе апсолутних грешки наведених приближних вредности како би важило $Af^* \leq 0.04$.

Решење.

$$\frac{\partial f}{\partial x_1}(x) = \frac{2x_1}{x_2}, \quad \frac{\partial f}{\partial x_2}(x) = -\left(\frac{x_1}{x_2}\right)^2, \quad \frac{\partial f}{\partial x_3}(x) = -1$$

$$Ax_k^* \leq \frac{\varepsilon}{\left| \frac{2x_1^*}{x_2^*} \right| + \left(\frac{x_1^*}{x_2^*}\right)^2 + 1} \approx \frac{0.04}{1.791} \approx 0.0223$$

Па је довољно захтевати $Ax_k^* = 0.02$, $k = 1, 2, 3$.