Numerička analiza

4. predavanje

Autori: Saša Singer i Nela Bosner

Predavač: Nela Bosner

nela@math.hr

web.math.hr/~nela/nad.html

PMF – Matematički odjel, Zagreb

Sadržaj predavanja

- Numerička analiza problema:
  - Greške unazad i unaprijed.
  - Mjerenje greške – vektorske i matrične norme.
  - Povratna anliza greške i povratno stabilni algoritmi.
  - Miješana naprijed–nazad greška i numerički stabilni algoritmi.
- Primjer numeričke analize:
  - Numerička analiza računanja $\sqrt{x_1 + x_2}$.
- Uvjetovanost problema:
  - Pojam uvjetovanosti.
  - Mjerenje grešaka i uvjetovanost; dobro i loše uvjetovani problemi.
  - Kako dizajnirati stabilne algoritme.

Sadržaj predavanja (nastavak)

- Primjeri uvjetovanosti problema:
  - Funkcija $\ln x$.
  - Zbroj $x_1 + x_2$.
  - Rješavanje sustava $Ay = x$.
  - Rekurzija za integral.

Numerička analiza problema

Greške — ponavljanje

Pri numeričkom rješavanju nekog problema javljaju se različiti tipovi grešaka:

- greške modela — svođenje realnog problema na neki "matematički" problem,
- greške u ulaznim podacima (mjerenja i sl.),
- greške numeričkih metoda za rješavanje "matematičkog" problema,
- greške "približnog" računanja — obično su to
  - greške zaokruživanja u aritmetici računala.

Greške modela su "izvan" dosega numeričke matematike.

- Spadaju u fiziku, kemiju, biologiju, tehniku, ekonomiju, ...

Greške (nastavak)

Sljedeće tri kategorije (podaci, metoda, računanje) su vezane za “matematički” problem, i

- spadaju u domenu numeričke matematike!

O njima nešto “moramo reći”.

Skica numeričkog rješavanja nekog problema sliči algoritmu:

Posebno, ako dozvolimo da, umjesto riječi “algoritam”,

- piše i riječ “metoda”.

Zamislite da pojam “algoritam” uključuje

- metodu i stvarno računanje rezultata!

Greške (nastavak)

Sve tri vrste grešaka — podaci, metoda, računanje,

- rezultiraju nekom greškom u konačnom rezultatu!

Ta greška nas "zanima".

Uočite da greške u ulaznim podacima možemo gledati

- neovisno o metodi za rješenje problema,
- i tako dolazimo do pojma uvjetovanosti problema.

Za razliku od toga, greške metode i računanja, naravno,

- ovise o metodi, odnosno, algoritmu za rješenje problema.

Greške unazad i unaprijed

- Gledam algoritam kao preslikavanje: ulaz (domena) u izlaz (kodomena).
- Zanima me greška u rezultatu (kodomeni) — unaprijed.
  - To katkad ide, ali je, uglavnom, teško (ili daje loše ocjene). Izvod takve greške vidjet ćemo kasnije na primjeru.
- Lakše je dobiti grešku unazad, akumuliranjem greške na ulazne podatke — ista interpretacija kao i za pojedine operacije.

Uočiti da se akumulacija faktora $(1 + \varepsilon)$ prirodno radi unazad.

Greške unazad i unaprijed (nastavak)

Pretpostavimo da se u aritmetici preciznosti u vrijednost $y = f(x)$ izračuna kao $\hat{y}$. Kako možemo mjeriti kvalitetu $\hat{y}$ kao aproksimacije od y?

- U većini slučajeva bit ćemo sretni ako postignemo malu relativnu grešku u rezultatu, tj. $E_{\mathrm{rel}} \approx u$ — ali to se neće moći uvijek postići.
- Umjesto toga možemo se zapitati za koji skup podataka smo zapravo riješili problem?
- Dakle, za koji Δx imamo

$$\hat{y} = f(x + \Delta x)?$$

Općenito, bit će više takvih Δx pa će nas zanimati najveći.

Greška na ulazu – što na izlazu?

Zadatak numeričke analize je odrediti vezu između greške na ulazu i greške na izlazu.

Polazni podaci		Izlazni podaci
x_{inp}	Algoritam f $\longrightarrow$	$y_{\mathrm{out}} = f(x_{\mathrm{inp}})$
imaju grešku		imaju grešku
$\Delta x := x_{\mathrm{inp}} - x$		$\Delta y := f(x_{\mathrm{inp}}) - f(x)$

Uzimamo da su $x \in \mathcal{X}$ i $y \in \mathcal{Y}$, i da su $\mathcal{X}$ i $\mathcal{Y}$ (barem) vektorski prostori.

Kako mjeriti grešku?

Kad x_{inp} i $f(x_{\text{inp}})$ nisu brojevi, nego vektori ili matrice, grešku možemo mjeriti:

- po svakoj od komponenenata, međutim to je malo previše brojeva,
- kao neku "ukupnu ili najveću" grešku — samo jedan broj i to korištenjem vektorskih/matričnih normi.

Prisjetite se: vektorski prostor na kojem je definirana norma zove se normirani prostor.

Vektorske norme

“Vektorska” norma na vektorskom prostoru V (nad poljem F, gdje je $F = \mathbb{R}$ ili $F = \mathbb{C}$) je

- je svaka funkcija $\| \cdot \| : V \to \mathbb{R}$

koja zadovoljava sljedeća svojstva:

1. $\|x\| \geq 0$, $\forall x \in V$, a jednakost vrijedi ako i samo ako je $x = 0$,
2. $\|\alpha x\| = |\alpha| \, \|x\|, \quad \forall \alpha \in F, \forall x \in V$,
3. $\|x + y\| \leq \|x\| + \|y\|, \quad \forall x, y \in V$
   (nejednakost poznata pod imenom nejednakost trokuta).

Najpoznatije vektorske norme

Kad je $V = \mathbb{R}^n$ ili $V = \mathbb{C}^n$ (kon. dim.), najčešće se koriste sljedeće tri norme:

1. 1-norma ili ℓ_1 norma $\displaystyle \|x\|_1 = \sum_{i=1}^{n} |x_i|$,
2. 2-norma ili ℓ_2 norma ili euklidska norma
$$\|x\|_2 = (x^* x)^{1/2} = \sqrt{\sum_{i=1}^{n} |x_i|^2},$$
3. ∞-norma ili ℓ_∞ norma $\displaystyle \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$.

Samo je 2-norma izvedena iz skalarnog produkta.

Norme na prostoru funkcija

Definicija vektorskih normi u sebi ne sadrži zahtjev da je vektorski prostor V konačno dimenzionalan.

Na primjer, norme definirane na vektorskom prostoru neprekidnih funkcija na $[a, b]$ (u oznaci $C[a, b]$) definiraju se slično normama na $\mathbb{R}^n$:

1. L_1 norma $\displaystyle \|f\|_1 = \int_a^b |f(t)|\, dt$,

2. L_2 norma $\displaystyle \|f\|_2 = \left(\int_a^b |f(t)|^2\, dt \right)^{1/2}$,

3. L_∞ norma $\|f\|_\infty = \max\{|f(x)| \mid x \in [a, b]\}$.

Ekvivalentnost normi

Može se pokazati da vrijedi sljedeći teorem.

Teorem. Na svakom konačno-dimenzionalnom vektorskom prostoru V sve su norme ekvivalentne, tj. za svake dvije norme $\| \ \|_a$ i $\| \ \|_b$ postoje konstante c i C takve da je

$$c\|v\|_a \leq \|v\|_b \leq C\|v\|_a, \quad \text{za sve } v \in V.$$

Na primjer,

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

za sve $x \in \mathbb{R}^n$.

Razlika između teorije i prakse — kad je n ogroman.

Matrične norme

Zamijenimo li u definiciji vektorske norme formalno vektor matricom, dobivamo matričnu normu.

Matrična norma je svaka funkcija $\| \cdot \| : \mathbb{C}^{m \times n} \to \mathbb{R}$ koja zadovoljava sljedeća svojstva:

1. $\|A\| \geq 0$, $\forall A \in \mathbb{C}^{m \times n}$, a jednakost vrijedi ako i samo ako je $A = 0$,
2. $\|\alpha A\| = |\alpha| \, \|A\|$, $\forall \alpha \in \mathbb{R}$, $\forall A \in \mathbb{C}^{m \times n}$,
3. $\|A + B\| \leq \|A\| + \|B\|$, $\forall A, B \in \mathbb{C}^{m \times n}$.

Tome se često dodaje zahtjev konzistentnosti

4. $\|AB\| \leq \|A\| \, \|B\|$

kad god je matrični produkt AB definiran.

Matrične norme (nastavak)

Matrične norme nastaju na dva načina:

- Matricu A promatramo kao vektor s $m \times n$ elemenata i za taj vektor koristimo odgovarajuću vektorsku normu.

  Najpoznatija takva norma odgovara vektorskoj 2-normi i zove se euklidska, Frobeniusova, Hilbert–Schmidtova, ili Schurova norma

$$\|A\|_F = (\mathrm{tr}(A^* A))^{1/2} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}.$$

- operatorske norme:

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad (\text{ ili } = \max_{\|x\|=1} \|Ax\|).$$

Matrične norme (nastavak)

Uvrštavanjem odgovarajućih vektorskih normi, dobivamo

1. matrična 1-norma, "maksimalna stupčana norma"
$$\|A\|_1 = \max_{j=1,\ldots,n} \sum_{i=1}^{m} |a_{ij}|,$$
2. matrična 2-norma, spektralna norma
$$\|A\|_2 = (\rho(A^* A))^{1/2} = \sigma_{\max}(A),$$
3. matrična ∞-norma, "maksimalna retčana norma"
$$\|A\|_\infty = \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{ij}|.$$

ρ je spektralni radijus matrice (po aps. vrijednosti maksimalna svojstvena vrijednost), a σ singularna vrijednost matrice.

Matrične norme (nastavak)

Svojstva:

- I matrične norme nisu međusobno neovisne (slično kao vektorske norme) — ekvivalentnost.
- Matrična 2-norma se teško računa pa se uobičajeno procjenjuje korištenjem ostalih normi.
- Frobeniusova norma je konzistentna matrična norma.
- Svaka operatorska norma je konzistentna, i za nju vrijedi

  $$\|Ay\| \leq \|A\| \, \|y\|,$$

  za svaki vektor y, što se često koristi kod ocjena. Formula direktno izlazi iz definicije operatorske norme.
- Nužan uvjet da bi $\| \ \|$ bila operatorska norma je

  $$\|I\| = 1, \qquad I \text{ je identiteta.}$$

Greške unazad i unaprijed (nastavak)

Pretpostavimo da se u aritmetici preciznosti u vrijednost $y = f(x)$ izračuna kao $\hat{y} = f(x + \Delta x)$. Za neku normu $\| \ \|$:

- vrijednosti
$$\|\Delta x\|, \qquad \frac{\|\Delta x\|}{\|x\|},$$
zovu se apsolutna ili relativna greška unazad (engl. backward error) ili povratna greška,
- a ako definiramo $\Delta y = \hat{y} - y$, tada se vrijdenosti
$$\|\Delta y\|, \qquad \frac{\|\Delta y\|}{\|y\|},$$
zovu se apsolutna ili relativna greška unaprijed (engl. forward error)

Povratna analiza greške

Proces omeđivanja povratne greške izračunatog rješenja zove se analiza povratne greške ili povratna analiza greške (engl. backward error analysis). Motivacija za taj postupak je dvostruka:

1. Analiza interpretira greške zaokruživanja kao greške u podacima.
   - Podaci često kriju netočnosti nastale zbog prethodnih izračunavanja, zbog spremanja u računalo ili kao rezultat mjerenja.
   - Ako povratna greška nije veća od tih polaznih netočnosti, tada se izračunato rješenje ne može mnogo kritizirati jer je to rješenje koje tražimo do na "ulaznu" netočnost.

Povratna analiza greške (nastavak)

2. Privlačnost analize je ta što ona reducira problem omeđivanja greške unaprijed na primjenu teorije perturbacije za dani problem.
   - Teorija perturbacije je dobro poznata za većinu problema i važno je to da ona ovisi o problemu, a ne o pojedinoj metodi za dani problem.
   - Kada dobijemo ocjenu za povratnu grešku rješenja kod primjene određene metode, tada primjenom opće teorije perturbacije za dani problem lako dolazimo do ocjene za grešku unaprijed.

Povratna analiza greške (nastavak)

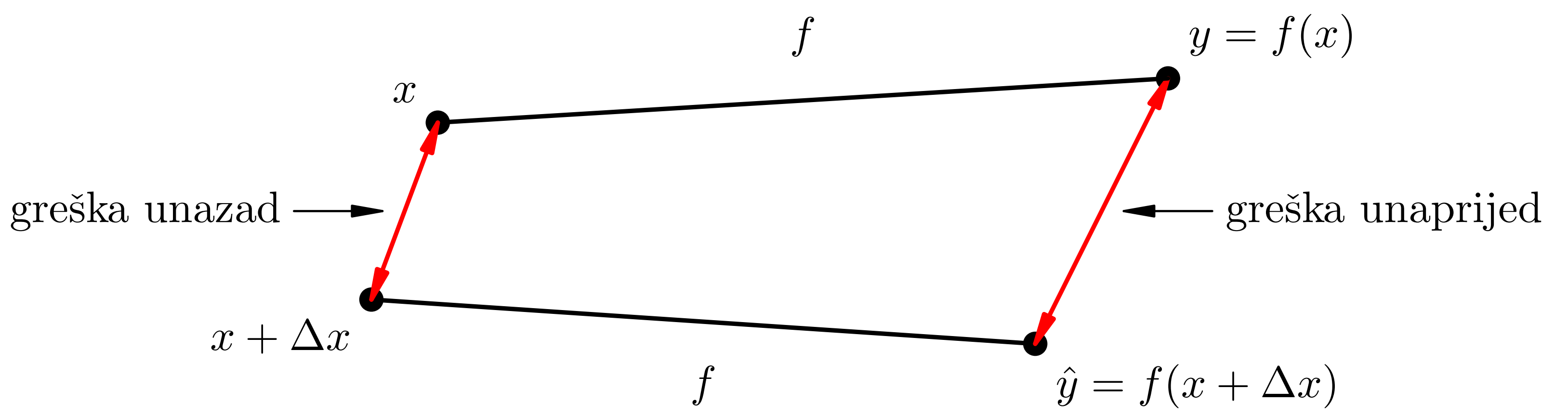


Figure 1: Greške unaprijed i unazad

Povratno stabilan algoritam

- Metoda za računanje vrijednosti $y = f(x)$ je povratno stabilna ili stabilna unazad (engl. backward stable), ako ona za svako x producira izračunati $\hat{y}$ s malom povratnom greškom, tj. vrijedi

$$\hat{y} = f(x + \Delta x) \quad \text{za malo } \Delta x.$$

- Oznaka mala ovisi o kontekstu.
- U načelu, za dani problem može postojati više metoda od kojih će neke biti povratno stabilne, a neke neće.
- Npr. sve su osnovne operacije u računalu povratno stabilne.

Miješana naprijed-nazad greška

Međutim, većina metoda za računanje funkcije $\cos(x)$ ne zadovoljava relaciju $\hat{y} = \cos(x + \Delta x)$ za malo Δx, već samo slabiji rezultat:

$$\hat{y} + \Delta y = \cos(x + \Delta x)$$

uz male Δx i Δy.

- Rezultat pisan u obliku

$$\hat{y} + \Delta y = f(x + \Delta x), \quad |\Delta y| \leq \eta |y|, \quad |\Delta x| \leq \xi |x|$$

naziva se miješana naprijed-nazad greška.

Numerički stabilan algoritam

- Ako su ξ i η mali, može se reći:
  - izračunato rješenje $\hat{y}$ se jedva razlikuje od vrijednosti $\hat{y} + \Delta y$
  - koje se dobije egzaktnim računom na ulaznoj vrijednosti $x + \Delta x$ koja se jedva razlikuje od stvarnog ulaznog podatka x.
- U tom slučaju kažemo da je algoritam numerički stabilan.
- Ova definicija uglavnom se odnosi na izračunavanja u kojima su greške zaokruživanja (osnovnih aritmetičkih operacija) dominantni oblici grešaka.
- Inače, pojam stabilnosti ima različita značenja u drugim područjima numeričke matematike.

Landauov simbol — red veličine

Neka su $g, h : \mathbb{R}^n \to \mathbb{R}^m$ funkcije, $\|\cdot\|_{\mathbb{R}^n}$ i $\|\cdot\|_{\mathbb{R}^m}$ norme i neka je $x_0 \in \mathbb{R}^n$.

Ako postoje konstante $C > 0$ i $\delta > 0$ takve da za sve x vrijedi

$$\|x - x_0\|_{\mathbb{R}^n} \leq \delta \quad \Longrightarrow \quad \|g(x)\|_{\mathbb{R}^m} \leq C\|h(x)\|_{\mathbb{R}^m},$$

onda kažemo da je

“funkcija g reda $\mathcal{O}$ od h, za x koji teži prema x_0”

i to pišemo ovako

$$g(x) = \mathcal{O}(h(x)) \quad (x \to x_0).$$

Landauov simbol (nastavak)

Primjer. Za $m = n = 1$ je

$$\begin{aligned}
\sin x &= \mathcal{O}(x) \quad (x \to a), \quad \text{za sve } a \in \mathbb{R}, \\
x^2 + 3x &= \mathcal{O}(x) \quad (x \to 0), \\
x^2 - x - 6 &= \mathcal{O}(x-3), \quad (x \to 3).
\end{aligned}$$

Primjer numeričke analize

Numerička analiza računanja $\sqrt{x^2 + y^2}$

Neka su x i y reprezentabilni u računalu, tako da vrijedi $x = f\ell(x)$ i $y = f\ell(y)$. S kolikom relativnom greškom će računalo koje koristi IEEE standard izračunati $z = \sqrt{x^2 + y^2}$?

- Da bismo riješili problem, pretpostavimo prvo da je

$$f\ell(x^2) + f\ell(y^2) < N_{\max} \quad \text{i} \quad \min\{|x|, |y|\} \geq \sqrt{N_{\min}}.$$

- Tada, možemo pisati

$$\begin{aligned} x_2 &= x \otimes x = f\ell(x^2) = x^2(1 + \varepsilon_1), \quad |\varepsilon_1| \leq u \\ y_2 &= y \otimes y = f\ell(y^2) = y^2(1 + \varepsilon_2), \quad |\varepsilon_2| \leq u. \end{aligned}$$

Numerička analiza $\sqrt{x^2+y^2}$ (nastavak)

- Umjesto

$$fl(fl(x^2)+fl(y^2)) = fl(x^2) \oplus fl(y^2),$$

  kraće pišemo $fl(x^2+y^2)$.
- Sada imamo

$$\begin{aligned} z_2 &= fl(x_2+y_2) = (x_2+y_2)(1+\varepsilon_3), \quad |\varepsilon_3| \leq u \\ z &= fl(\sqrt{z_2}) = \sqrt{z_2}(1+\varepsilon_4), \quad |\varepsilon_4| \leq u. \end{aligned}$$

Numerička analiza $\sqrt{x^2 + y^2}$ (nastavak)

- Povežimo sve te jednadžbe:

$$
\begin{aligned}
z &= (1+\varepsilon_4)\sqrt{z_2} = (1+\varepsilon_4)\sqrt{(x_2+y_2)(1+\varepsilon_3)} \\
&= (1+\varepsilon_4)\sqrt{1+\varepsilon_3}\sqrt{x^2(1+\varepsilon_1)+y^2(1+\varepsilon_2)} \\
&= (1+\varepsilon_4)\sqrt{1+\varepsilon_3}\sqrt{x^2+y^2}\sqrt{1+\frac{x^2\varepsilon_1+y^2\varepsilon_2}{x^2+y^2}} \\
&= (1+\varepsilon_4)\sqrt{1+\varepsilon_3}\sqrt{1+\varepsilon_5}\sqrt{x^2+y^2}, \\
&= \sqrt{x^2+y^2}(1+\varepsilon_z),
\end{aligned}
$$

Numerička analiza $\sqrt{x^2 + y^2}$ (nastavak)

pri čemu je

$$\varepsilon_5 = \frac{x^2\varepsilon_1 + y^2\varepsilon_2}{x^2 + y^2}, \qquad 1 + \varepsilon_z = (1 + \varepsilon_4)\sqrt{(1 + \varepsilon_3)(1 + \varepsilon_5)}.$$

- Ocijenimo prvo ε_5. Kako je

$$\frac{x^2\varepsilon_1 + y^2\varepsilon_2}{x^2 + y^2} = \frac{x^2}{x^2 + y^2}\,\varepsilon_1 + \frac{y^2}{x^2 + y^2}\,\varepsilon_2,$$

  ε_5 je konveksna suma brojeva ε_1 i ε_2, pa se nalazi u zatvorenom intervalu $[\varepsilon_1, \varepsilon_2]$ (ili $[\varepsilon_2, \varepsilon_1]$ ako je $\varepsilon_2 < \varepsilon_1$).
- To opet znači da vrijedi

$$|\varepsilon_5| \leq \max\{|\varepsilon_1|, |\varepsilon_2|\} \leq u.$$

Numerička analiza $\sqrt{x^2 + y^2}$ (nastavak)

- Sada imamo,

$$
\begin{aligned}
|\varepsilon_z| &= |(1+\varepsilon_4)\sqrt{(1+\varepsilon_3)(1+\varepsilon_5)} - 1| \\
&\le (1+u)\sqrt{(1+u)(1+u)} - 1 \\
&\le (1+u)(1+u) - 1 \le (2+u)u \\
&\le 2u + \mathcal{O}(u^2).
\end{aligned}
$$

Numerička analiza $\sqrt{x^2+y^2}$ (nastavak)

- Kod korištenja IEEE aritmetike, slučaj prekoračenja će biti dojavljen. Kod najnovijih prevodioca (npr. za programski jezik FORTRAN 90), postoji mogućnost da programer ugradi u kôd programa grananje koje u slučaju prekoračenja nastavlja s računanjem po formuli

$$\alpha = \max\{|x|, |y|\}, \quad \beta = \min\{|x|, |y|\},$$

$$z = f\ell\left(\alpha\sqrt{1 + \left(\frac{\beta}{\alpha}\right)^2}\right).$$

- Ova formula je sigurna jer nam podaci $x = f\ell(x)$ i $y = f\ell(y)$ ukazuju da su α i β reprezentabilni brojevi.

Numerička analiza $\sqrt{x^2+y^2}$ (nastavak)

- U ovom slučaju je prekoračenje moguće tek ako je

$$\alpha \leq N_{\max} < \alpha \sqrt{1 + \left(\frac{\beta}{\alpha}\right)^2} \leq \sqrt{2}\alpha.$$

- Pretpostavimo da to nije slučaj kao i da neće doći do potkoračenja međurezultata .
- Sličnim postupkom kao kod prvog algoritma možemo dobiti sljedeću ocjenu:

$$z = \left(\alpha \sqrt{1 + \left(\frac{\beta}{\alpha}\right)^2}\right)(1 + \eta_z), \quad |\eta_z| \leq 4u + \mathcal{O}(u^2).$$

Numerička analiza $\sqrt{x^2 + y^2}$ (nastavak)

- Korištenjem sofisticiranije formule, koja zahtijeva više instrukcija, dobivamo i rezultat koji je generalno nešto netočniji (iako se još uvijek radi o grešci u zadnjoj sigurnoj decimali egzaktnog rezulata).
- Međutim, dobitak je u proširenju domene funkcije koja paru reprezentabilnih brojeva (x, y) pridružuje $f\ell(\sqrt{x^2 + y^2})$.
- To je dosta važno kako bi se izbjeglo prekoračenje i prekid rada računala ili moguća relativna netočnost ako je rezultat mali broj.

Numerička analiza $\sqrt{x^2+y^2}$ (nastavak)

- Ostalo je još promotriti što se događa ako su x ili y takvi da ili x^2 ili y^2 potkoračuje, ako z postaje nula, ili ako $\sqrt{x^2+y^2}$ padne u područje subnormalnih brojeva.
  - Ako samo x^2 (y^2) potkoračuje u nulu ili subnormalni broj, rezultat će imati malu relativnu grešku, ne veću od nekoliko u.
  - Ako izračunati $f\ell(x^2+y^2)$ padne u područje subnormalnih brojeva, tada će $f\ell(x^2+y^2)$ možda nositi veliku relativnu grešku. Vađenje korijena će tu grešku približno prepoloviti, ali će ona i dalje ostati velika.
  - Potkoračenje obaju brojeva u nulu će dati maksimalnu relativnu grešku -1.

Numerička analiza $\sqrt{x^2 + y^2}$ (nastavak)

- Dakle, kad prijeti potkoračenje ili postupno potkoračenje, direktna formula neće dati točan rezultat. Što će dati druga kompliciranija formula?
- Ako je $|\alpha| \geq N_{\min}$, tada će se $f\ell(\alpha\sqrt{1 + (\beta/\alpha)^2})$ izračunati s malom relativnom greškom (kako je gore izračunato) bez obzira na to je li β subnormalan broj ili ne.
- Ako je pak α subnormalan broj, tada su oba polazna broja subnormalna i kod smještanja brojeva u računalo došlo je do gubljenja relativne točnosti. Tada će i egzaktni kvocijent β/α i izračunati kvocijent $f\ell(\beta/\alpha)$ imati veću (ili veliku) relativnu grešku pa će isto vrijediti i za z.

Numerička analiza $\sqrt{x^2 + y^2}$ (nastavak)

- Jedini lijek je da se ulazni podaci x i y, prije učitavanja, skaliraju potencijom od 2 tako da α ne bude subnormalni broj.
- Nakon računanja, z treba skalirati inverznom (negativnom) potencijom broja 2.

Uvjetovanost problema

Uvjetovanost problema

Odnos između greške unaprijed i greške unazad za dani problem u velikoj mjeri određen je uvjetovanošću problema. Neformalno rečeno, uvjetovanost problema mjeri

- osjetljivost problema na greške u ulaznim podacima.

Osnovno svojstvo uvjetovanosti:

- Ne ovisi o konkretnoj numeričkoj metodi za rješenje problema, već samo o problemu.

Svrha uvjetovanosti $=$ daje odgovor na pitanje:

- Koju točnost rezultata možemo očekivati
- pri točnom računanju
- s (malo) pomaknutim — netočnim podacima?

Model problema

Matematički model problema, zovimo ga P:

- za zadani ulaz — podatak $x \in \mathcal{X}$,
- dobivamo izlaz — rezultat $y \in \mathcal{Y}$.

Slikica modela je

$$x \longrightarrow \boxed{P} \longrightarrow y$$

Problem P interpretiramo kao računanje vrijednosti funkcije

$$f : \mathcal{X} \to \mathcal{Y},$$

gdje su $\mathcal{X}$ i $\mathcal{Y}$ odgovarajući matematički objekti. Na primjer, vektorski prostori, a vrlo često su i normirani prostori (treba nam mjera za grešku).

Primjeri problema

Primjer 1. Računanje sume dva realna broja $x_1, x_2 \in \mathbb{R}$. Tada je

$$f(x_1, x_2) = x_1 + x_2,$$

s tim da je $\mathcal{X} = \mathbb{R}^2$ i $\mathcal{Y} = \mathbb{R}$.

Primjer 2. Računanje produkta dva realna broja $x_1, x_2 \in \mathbb{R}$. Tada je

$$f(x_1, x_2) = x_1 x_2,$$

s tim da je opet $\mathcal{X} = \mathbb{R}^2$ i $\mathcal{Y} = \mathbb{R}$.

Primjeri problema (nastavak)

Primjer 3. Računanje sjecišta pravaca

$$P_1 = \{(y_1, y_2) \in \mathbb{R}^2 \mid a_{11}y_1 + a_{12}y_2 = x_1\},$$
$$P_2 = \{(y_1, y_2) \in \mathbb{R}^2 \mid a_{21}y_1 + a_{22}y_2 = x_1\}.$$

Smatramo da su koeficijenti a_{ij} i x_i, za $i, j = 1, 2$, ulazni podaci.

Ovaj problem pišemo u matričnom zapisu kao linearni sustav od dvije jednadžbe oblika $Ay = x$, gdje je

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2.$$

Primjeri problema (nastavak)

Traženo sjecište je rješenje linearnog sustava $Ay = x$.

Ako pretpostavimo da je matrica A sustava regularna, tj. $\det A \neq 0$, onda je $y = A^{-1}x$. Dakle,

$$f(x) = A^{-1}x,$$

s tim da je $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$.

Uvjetovanost problema (nastavak)

Ideja uvjetovanosti:

$$\text{greška u rezultatu} \approx \text{uvjetovanost} \cdot \text{greška u podacima}$$

Ovisi o obje vrijednosti: točnoj x i približnoj $\hat{x}$.

Napomene:

- Obično nas uvjetovanost posebno zanima za male perturbacije (greške, smetnje) podataka.
- Ako je f dovoljno glatka funkcija, možemo koristiti Taylorov razvoj u okolini točnog ulaznog podatka x
- i dobiti procjenu uvjetovanosti preko derivacija!

Mjerenje grešaka i uvjetovanost

Relativna/apsolutna uvjetovanost problema mjeri koliko je problem osjetljiv na odgovarajuće promjene polaznih podataka.

- Apsolutna greška: $\|\Delta x\|$, $\|\Delta y\|$, (svaka norma u svom prostoru), gdje je

$$\Delta x = x - \hat{x}, \qquad \Delta y = y - \hat{y}.$$

- Apsolutna uvjetovanost:

$$\kappa_{\text{abs}}(x) := \frac{\|\Delta y\|}{\|\Delta x\|}.$$

Veza s derivacijom!

Mjerenje grešaka i uvjetovanost (nastavak)

U praksi se češće koristi relativna mjera za grešku (na primjer, zbog aritmetike računala).

- Relativna greška:

$$\delta_x := \frac{\|\Delta x\|}{\|x\|}, \qquad \delta_y := \frac{\|\Delta y\|}{\|y\|}.$$

- Relativna uvjetovanost:

$$\kappa_{\text{rel}}(x) := \frac{\|\delta_y\|}{\|\delta_x\|}.$$

Problem je dobro uvjetovan ako je

- κ_{rel} što je moguće manji, za $\delta_x \to 0$.

Uvjetovanost i Taylorov teorem

Istražimo uvjetovanost problema za funkciju $f : \mathbb{R} \to \mathbb{R}$.

- Promatramo ponašanje f za male perturbacije Δx u okolini točke x. Neka je Δy pripadna perturbacija funkcijske vrijednosti $y = f(x)$, tj. $f(x + \Delta x) = y + \Delta y$.
- Neka je f još dva puta neprekidno derivabilna. Korištenjem Taylorovog polinoma stupnja 1 dobivamo

$$\begin{aligned}\Delta y &= f(x + \Delta x) - f(x) \\ &= f'(x)\Delta x + \frac{f''(x + \theta\Delta x)}{2!}(\Delta x)^2, \quad \theta \in (0, 1).\end{aligned}$$

Uvjetovanost i Taylorov teorem (nastavak)

- Za male perturbacije Δx, apsolutni oblik ove relacije je

$$\Delta y = f'(x)\,\Delta x + \mathcal{O}((\Delta x)^2),$$

  odakle slijedi da je $f'(x)$ ili $|f'(x)|$ apsolutna uvjetovanost funkcije f (za male Δx).

- Ako je $x \neq 0$ i $y \neq 0$, onda joj je relativna forma

$$\frac{\Delta y}{y} = \frac{x\,f'(x)}{f(x)}\,\frac{\Delta x}{x} + \mathcal{O}\left(\left(\frac{\Delta x}{x}\right)^2\right),$$

  pa relativnu uvjetovanost funkcije f možemo definirati kao

$$\kappa_{\mathrm{rel}}(x) = (\mathrm{cond}\, f)(x) := \left|\frac{x\,f'(x)}{f(x)}\right|.$$

Uvjetovanost i Taylorov teorem (nastavak)

- Za višedimenzionalne probleme, da bismo dobili jedan broj u prethodnoj formuli, uzmemo normu, a derivacija je gradijent (odnosno diferencijal),

$$(\mathrm{cond}\, f)(x) = \frac{\|x\| \|\nabla f(x)\|}{\|f(x)\|}.$$

Kako dizajnirati stabilne algoritme

Nema jednostavnog recepta za dizajniranje stabilnih algoritama, ali najbolji je savjet spoznaja da je numerička stabilnost najvažnija karakteristika algoritma.

1. Pokušajte izbjeći oduzmanje veličina bliskih vrijednosti koje nose greške, iako je to katkad nemoguće.
2. Minimizirajte veličinu međurezultata u odnosu na konačni rezultat. To je važno da konačni rezultat ne bi bio dobiven opasnim oduzimanjem velikih vrijednosti koje nose u sebi greške.
3. Potražite drugačije formulacije istog problema ili druge formule za isti račun.

Kako dizajnirati stabilne algoritme (nastavak)

4. Koristite prednosti jednostavnih formula za ažuriranje tipa

   $$\text{nova vrijednost} = \text{stara vrijednost} + \text{mala korekcija},$$

   ako se mala korekcija može izračunati na dovoljan broj značajnih znamenki (Eulerova metoda, Newtonova metoda, ...).
5. Koristite samo dobro uvjetovane transformacije za dobivanje rješenja. Kod matrica to znači koristiti ortogonalne matrice kad god je to moguće, jer neortogonalne matrice mogu biti loše uvjetovane.

Kako dizajnirati stabilne algoritme (nastavak)

6. Poduzmite mjere opreza da biste spriječili moguća prekoračenja granice konačne aritmetike (**overflow** i **underflow**).
7. Kod nekih računala centralna aritmetička jedinica koristi precizniju aritmetiku za operande u registrima, a zaokruživanje nastupa tek kod spremanja podatka u memoriju. To znači da nije povoljno cijepati složenije formule u više programskih linija korištenjem pomoćnih varijabla.

Primjeri uvjetovanosti problema

Funkcija $\ln x$

Relativna uvjetovanost funkcije

$$f(x) = \ln x,$$

je

$$(\text{cond}\, f)(x) = \left| \frac{x f'(x)}{f(x)} \right| = \left| \frac{1}{\ln x} \right|,$$

što je veliko za $x \approx 1$.

Pitanje: Apsolutna uvjetovanost?

Zbroj $x_1 + x_2$

Promatramo funkciju $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = x_1 + x_2$.

- Funkciju f možemo preformulirati u

$$f(x) = e^T x, \quad e = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

- Tada vrijedi

$$f(x + \Delta x) = e^T (x + \Delta x) = e^T x + e^T \Delta x = f(x) + e^T \Delta x,$$

čime dobivamo direktnu vezu između greške unaprijed i greške unazad:

$$f(x + \Delta x) - f(x) = e^T \Delta x.$$

Zbroj $x_1 + x_2$ (nastavak)

- Korištenjem Cauchy-Schwarz-Bunjakovski nejednakosti dobit ćemo odnos apsolutnih grešaka

$$|f(x + \Delta x) - f(x)| \leq \|e^T\|_2 \|\Delta x\|_2 = \sqrt{2}\|\Delta x\|_2,$$

pri čemu se gornja ograda na desnoj strani dostiže kada je Δx kolinearan sa e.

- Prema tome, $\sqrt{2}$ je apsolutna uvjetovanost.
- S druge strane imamo

$$\frac{|f(x + \Delta x) - f(x)|}{|f(x)|} \leq \frac{\sqrt{2}\|x\|_2}{|e^T x|} \cdot \frac{\|\Delta x\|_2}{\|x\|_2},$$

pri čemu se gornja ograda na desnoj strani dostiže.

Zbroj $x_1 + x_2$ (nastavak)

- S druge strane je

$$\nabla f(x) = e, \qquad \|\nabla f(x)\|_2 = \sqrt{2},$$

  odakle vidimo da je relativna uvjetovanost, kao što i prethodni izraz predlaže, jednaka

$$(\operatorname{cond} f)(x) = \frac{\sqrt{2}\|x\|_2}{|e^T x|} = \frac{\sqrt{2}\sqrt{x_1^2 + x_2^2}}{|x_1 + x_2|}.$$

- Iz ovoga možemo zaključiti da u slučaju kada su x_1 i x_2 bliski po apsolutnim vrijednostima ali suprotnih predznaka $(\operatorname{cond} f)(x_1, x_2)$ ide prema ∞, tj. zbroj je tada jako osjetljiv na perturbacije u ulaznim podacima. To smo znali i od prije!

Zbroj $x_1 + x_2$ (nastavak)

- Na primjer, za

$$x = \begin{bmatrix} 1 + 10^{-15} \\ -1 \end{bmatrix}, \quad \Delta x = \begin{bmatrix} 10^{-7} \\ 10^{-7} \end{bmatrix}$$

imamo

$$\|x\|_2 = \sqrt{2 + 2 \cdot 10^{-15} + 10^{-30}} \geq \sqrt{2}, \ \|\Delta x\|_2 = \sqrt{2} \cdot 10^{-7},$$

$$f(x) = 10^{-15}, \ f(x + \Delta x) = 2 \cdot 10^{-7} + 10^{-15},$$

što znači da je

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq 10^{-7}, \ \frac{|f(x + \Delta x) - f(x)|}{|f(x)|} = 2 \cdot 10^{8}.$$

Zbroj $x_1 + x_2$ (nastavak)

- U tom slučaju relativna uvjetovanost iznosi:

$$\begin{aligned}(\mathrm{cond}\, f)(x) &= \frac{\sqrt{2}\|x\|_2}{|e^T x|} \\ &= \frac{\sqrt{2}\sqrt{2 + 2 \cdot 10^{-15} + 10^{-30}}}{|10^{-15}|} \\ &\geq 2 \cdot 10^{15}\end{aligned}$$

- Relativna uvjetovanost dobro određuje red veličine relativne greške unaprijed.

Napomena. Valja naglasiti da kraćenje pri oduzimanju bliskih brojeva nije uvijek loša stvar, ukoliko su ulazni podaci točni.

Rješavanje sustava $Ay = x$

Promatramo funkciju $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(x) = A^{-1}x$, gdje je $A \in \mathbb{R}^{n\times n}$ regularna matrica.

- Tada vrijedi

  $$\begin{aligned} f(x + \Delta x) &= A^{-1}(x + \Delta x) = A^{-1}x + A^{-1}\Delta x \\ &= f(x) + A^{-1}\Delta x, \end{aligned}$$

  čime ponovo dobivamo direktnu vezu između greške unaprijed i greške unazad:

  $$f(x + \Delta x) - f(x) = A^{-1}\Delta x.$$

Rješavanje sustava $Ay = x$ (nastavak)

- Korištenjem konzistentnosti matrične i vektorske norme dobit ćemo odnos apsolutnih grešaka

$$\|f(x + \Delta x) - f(x)\|_2 \leq \|A^{-1}\|_2 \|\Delta x\|_2,$$

pri čemu se gornja ograda na desnoj strani dostiže kada je Δx svojstveni vektor najveće svojstvene vrijednosti od $A^{-T}A^{-1}$:

$$\begin{aligned} A^{-T}A^{-1}\Delta x &= \|A^{-1}\|_2^2 \Delta x \\ \|A^{-1}\Delta x\|_2^2 &= \Delta x^T A^{-T} A^{-1} \Delta x = \|A^{-1}\|_2^2 \|\Delta x\|_2^2. \end{aligned}$$

- Prema tome, $\|A^{-1}\|_2$ je apsolutna uvjetovanost.

Rješavanje sustava $Ay = x$ (nastavak)

- S druge strane imamo

$$
\begin{aligned}
\frac{\|f(x+\Delta x) - f(x)\|_2}{\|f(x)\|_2} &\leq \frac{\|A^{-1}\|_2 \|x\|_2}{\|A^{-1}x\|_2} \cdot \frac{\|\Delta x\|_2}{\|x\|_2} \\
&= \frac{\|A^{-1}\|_2 \|A \cdot A^{-1}x\|_2}{\|A^{-1}x\|_2} \cdot \frac{\|\Delta x\|_2}{\|x\|_2} \\
&\leq \frac{\|A^{-1}\|_2 \|A\|_2 \|A^{-1}x\|_2}{\|A^{-1}x\|_2} \cdot \frac{\|\Delta x\|_2}{\|x\|_2} \\
&= \|A^{-1}\|_2 \|A\|_2 \cdot \frac{\|\Delta x\|_2}{\|x\|_2}
\end{aligned}
$$

pri čemu se ograda na desnoj strani dostiže.

Rješavanje sustava $Ay = x$ (nastavak)

- S druge strane je
$$\nabla f(x) = A^{-1},$$
odakle vidimo da je relativna uvjetovanost, kao što i prethodni izraz predlaže, jednaka
$$\kappa_{\text{rel}}(x) = (\text{cond}\, f)(x) = \frac{\|A^{-1}\|_2 \|x\|_2}{\|A^{-1}x\|_2}.$$
- Gledamo najgoru moguću uvjetovanost, po svim vektorima x, odnosno $y = A^{-1}x$,
$$\max_{\substack{y \in \mathbb{R}^n \\ y \neq 0}} \kappa_{\text{rel}}(x) = \|A^{-1}\|_2 \max_{\substack{y \in \mathbb{R}^n \\ y \neq 0}} \frac{\|Ay\|_2}{\|y\|_2} = \|A^{-1}\|_2 \|A\|_2.$$

Rješavanje sustava $Ay = x$ (nastavak)

- Veličina
$$\kappa_2(A) = \|A^{-1}\|_2 \|A\|_2$$
naziva se uvjetovanost matrice A u normi 2.
- Uvjetovanost matrice je uvijek veća ili jednaka 1, jer je zbog konzistentnosti matrične norme
$$\kappa_2(A) = \|A^{-1}\|_2 \|A\|_2 \geq \|A^{-1}A\|_2 = \|I\|_2 = 1,$$
gdje je I identiteta.
- Taj broj nam govori koliko je rješavanje sustava osjetljivo na perturbacije desne strane.
- Broj $\kappa_2(A)^{-1}$ daje udaljenost matrice A od skupa singularnih matrica, mjerenu u 2-normi.

Rješavanje sustava $Ay = x$ (nastavak)

- Ako je $\kappa_2(A) \gg 1$ tada kažemo da je matrica A loše uvjetovana:
  - ona je onda blizu singularnosti
  - rješavanje sustava s tom matricom je osjetljivo na greške u desnoj strani sustava, kojom god metodom ga rješavali.

Rješavanje sustava $Ay = x$ (nastavak)

- Ako uključimo greške i u elementima matrice tada za uvjetovanost definiranu sa

$$\kappa_{A,x}(A,y) = \lim_{\epsilon \to 0} \sup \left\{ \frac{\|\Delta y\|}{\epsilon \|y\|} : (A + \Delta A)(y + \Delta y) = x + \Delta x, \right.$$
$$\left. \|\Delta A\| \leq \epsilon \|A\|,\ \|\Delta x\| \leq \epsilon \|x\| \right\},$$

vrijedi

$$\kappa_{A,x}(A,y) = \frac{\|A^{-1}\| \|x\|}{\|A^{-1}x\|} + \|A^{-1}\| \|A\| \leq 2\kappa(A).$$

Rješavanje konkretnog sustava – Primjer 1

Rješavamo sustav $A_1 x = b_1$, pri čemu su

$$A_1 = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

a x_1 je rješenje sustava.

Perturbirajmo sada samo drugu komponentu $b_1(2)$ perturbacijom s relativnom greškom 10^{-16}, a tako dobiven vektor desne strane označimo s $b_{1,p}$. Rješenje sustava $A_1 x = b_{1,p}$, označimo sa $x_{1,p}$. Tada je

$$b_{1,p} = \begin{bmatrix} 3 \\ 2 \cdot (1 + 10^{-16}) \end{bmatrix}, \qquad x_{1,p} = \begin{bmatrix} 1 + 4 \cdot 10^{-16} \\ 1 - 2 \cdot 10^{-16} \end{bmatrix}.$$

Rješavanje konkretnog sustava – Primjer 1 (n.)

- Relativna greška unazad:

$$\frac{\|b_{1,p} - b_1\|_2}{\|b_1\|_2} = 5.547001962252291 \cdot 10^{-17}$$

- Relativna greška unaprijed:

$$\frac{\|x_{1,p} - x_1\|_2}{\|x_1\|_2} = 3.162277660168379 \cdot 10^{-16}$$

- Uvjetovanost matrice:

$$\kappa_2(A_1) = 6.854101966249680$$

- Malom perturbacijom samo jednog elementa desne strane sustava dobili smo male greške unazad i unaprijed jer je matrica dobro uvjetovana.

Rješavanje konkretnog sustava – Primjer 2

Rješavamo sustav $A_2 x = b_2$, pri čemu su

$$A_2 = \begin{bmatrix} 1 & 10^{15}+1 \\ 1 & 10^{15} \end{bmatrix}, \quad b_2 = \begin{bmatrix} 10^{15}+2 \\ 10^{15}+1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

a x_2 je rješenje sustava.

Perturbirajmo sada samo drugu komponentu $b_2(2)$ perturbacijom s relativnom greškom 10^{-16}, a tako dobiven vektor desne strane označimo s $b_{2,p}$. Rješenje sustava $A_2 x = b_{2,p}$, označimo sa $x_{2,p}$. Tada je

$$b_{2,p} = \begin{bmatrix} 10^{15}+2 \\ (10^{15}+1)\cdot(1+10^{-16}) \end{bmatrix}, \; x_{2,p} = \begin{bmatrix} 10^{14}+1+2\cdot 10^{-1}+10^{-16} \\ 1-10^{-1}-10^{-16} \end{bmatrix}.$$

Rješavanje konkretnog sustava – Primjer 2 (n.)

- Relativna greška unazad:

$$\frac{\|b_{2,p} - b_2\|_2}{\|b_2\|_2} = 7.071067811865472 \cdot 10^{-17}$$

- Relativna greška unaprijed:

$$\frac{\|x_{2,p} - x_2\|_2}{\|x_2\|_2} = 7.071067811865489 \cdot 10^{13}$$

- Uvjetovanost matrice:

$$\kappa_2(A_2) = 1.885618083164128 \cdot 10^{30}$$

- Malom perturbacijom samo jednog elementa desne strane sustava dobili smo malu grešku unazad ali ogromnu grešku unaprijed jer je matrica jako loše uvjetovana.

Rekurzija za integral

Ispitajmo uvjetovanost problema računanja integrala

$$I_n = \int_0^1 \frac{t^n}{t+5}\, dt$$

za fiksni prirodni broj n.

U ovom obliku, problem je napisan kao preslikavanje iz $\mathbb{N}$ u $\mathbb{R}$ i ne "paše" ranijem pojmu problema.

- Domena nije $\mathbb{R}$, nego $\mathbb{N}$ (diskretan skup), pa nema smisla govoriti o neprekidnosti, derivabilnosti i sl.

Zato prvo transformiramo problem.

Rekurzija za integral (nastavak)

Nađimo vezu između I_k i I_{k-1}, s tim da I_0 znamo izračunati

$$I_0 = \int\limits_0^1 \frac{1}{t+5}\,dt = \ln(t+5)\,\Big|_0^1 = \ln\frac{6}{5}.$$

Za početak, očito vrijedi da je

$$\frac{t}{t+5} = 1 - \frac{5}{t+5},$$

Množenjem obje strane s t^{k-1} dobivamo

$$\frac{t^k}{t+5} = t^{k-1} - 5\,\frac{t^{k-1}}{t+5}.$$

Rekurzija za integral (nastavak)

Na kraju, integracijom na segmentu $[0, 1]$ izlazi

$$I_k = \int\limits_0^1 t^{k-1}\, dt - 5I_{k-1} = \frac{1}{k} - 5I_{k-1}, \quad k = 1, 2, \ldots, n.$$

Dakle, I_k je rješenje (linearne, nehomogene) diferencijske jednadžbe

$$y_k = -5y_{k-1} + \frac{1}{k}, \quad k = 1, 2, \ldots,$$

uz početni uvjet $y_0 = I_0$.

Rekurzija za integral (nastavak)

Varijacija početnog uvjeta definira niz funkcija f_k, $y_k = f_k(y_0)$.

Zanima nas relativna uvjetovanost funkcije f_n u točki $y_0 = I_0$, u ovisnosti o $n \in \mathbb{N}$.

- I_0 nije egzaktno prikaziv,
- umjesto I_0 spremi se aproksimacija $\widehat{I}_0$,
- rezultat — neka aproksimacija $\widehat{I}_n = f_n(\widehat{I}_0)$.

Indukcijom se lako dokaže da vrijedi

$$y_n = f_n(y_0) = (-5)^n y_0 + p_n,$$

gdje je p_n ovisi samo o nehomogenim članovima rekurzije, ali ne i o y_0.

Rekurzija za integral (nastavak)

Relativna uvjetovanost je

$$(\operatorname{cond} f_n)(y_0) = \left| \frac{y_0 f_n'(y_0)}{y_n} \right| = \left| \frac{y_0 (-5)^n}{y_n} \right|.$$

Iz definicije integrala: I_n monotono padaju po n, čak

$$\lim_{n \to \infty} I_n = 0.$$

Zbrajanjima dobivamo sve manje i manje brojeve!

$$(\operatorname{cond} f_n)(I_0) = \frac{I_0 \cdot 5^n}{I_n} > \frac{I_0 \cdot 5^n}{I_0} = 5^n.$$

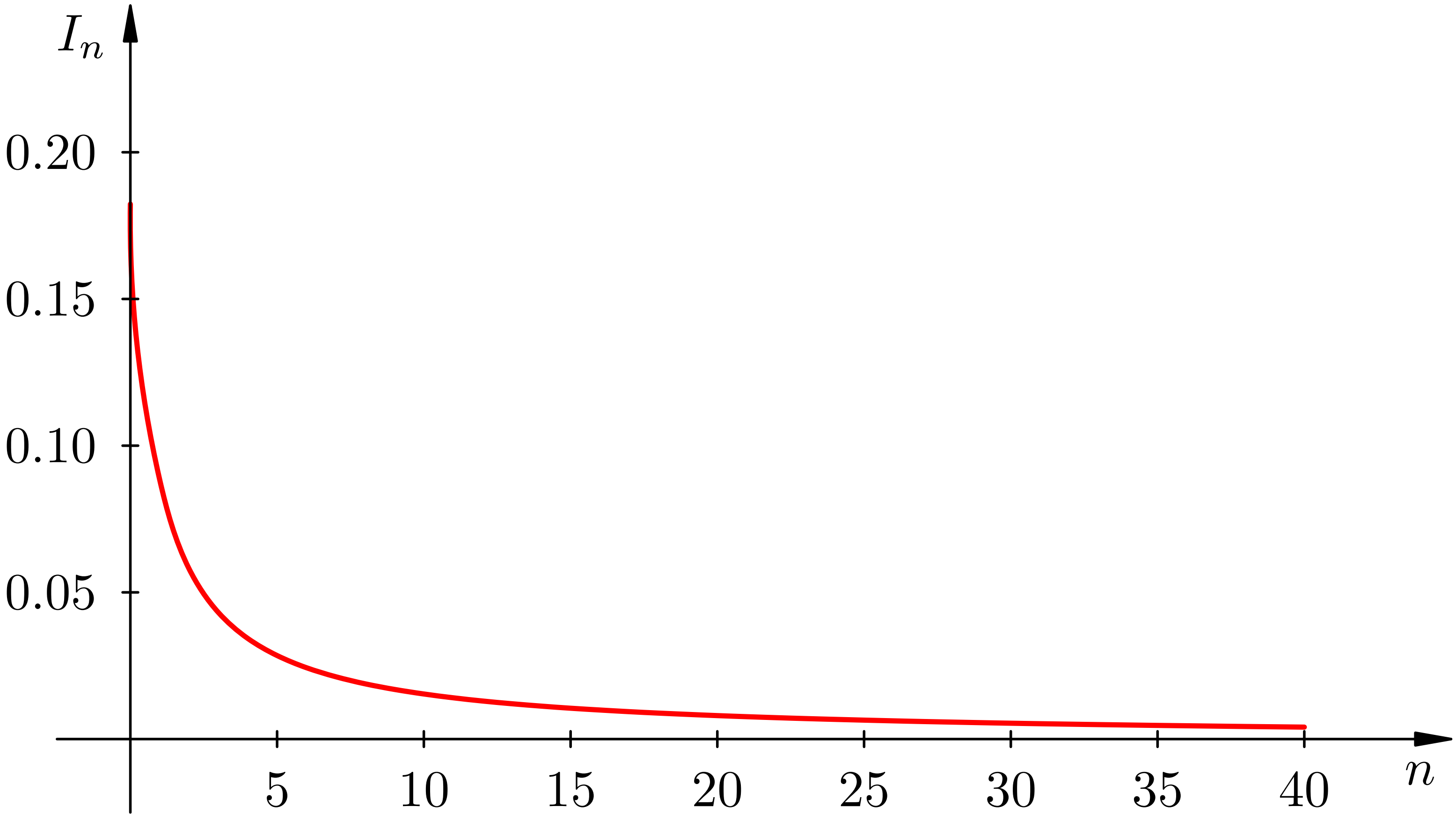
f_n je vrlo loše uvjetovana u $y_0 = I_0$, i to tim gore kad n raste.

Rekurzija unaprijed — rezultati

Pitanje: Kako se loša uvjetovanost vidi, kad stvarno računamo $f_n(I_0)$?

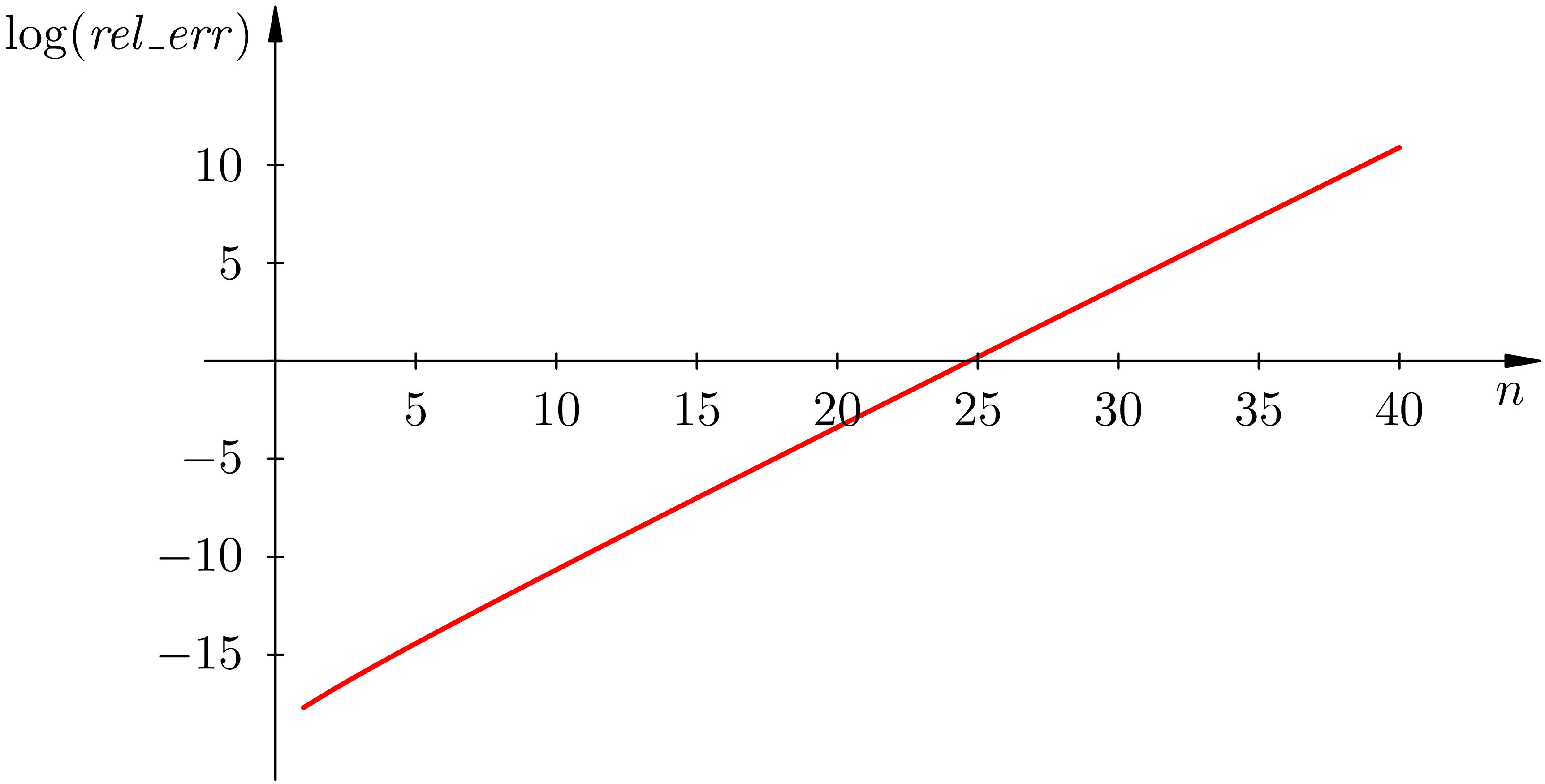
Slikice!

Točne vrijednosti integrala



egzaktne/točne vrijednosti integrala I_n

Rekurzija unaprijed za I_n



($\log_{10}$) relativne greške izračunate vrijednosti integrala I_n rekurzijom unaprijed

Rekurzija za integral (nastavak)

Može li se loša uvjetovanost izbjeći?

- Može — okretanjem rekurzije.

Treba uzeti neki $\nu > n$ i "silazno" računati

$$y_{k-1} = \frac{1}{5}\left(\frac{1}{k} - y_k\right), \quad k = \nu, \nu - 1, \ldots, n + 1.$$

Problem: kako izračunati početnu vrijednost y_ν.

Nova rekurzija definira niz funkcija g_k, koje vežu y_n i y_ν, uz $\nu > n$, tj.

$$y_n = g_n(y_\nu).$$

Rekurzija za integral (nastavak)

Relativna uvjetovanost za g_n

$$(\operatorname{cond} g_n)(y_\nu) = \left| \frac{y_\nu \, (-1/5)^{\nu-n}}{y_n} \right|, \quad \nu > n.$$

Za $y_\nu = I_\nu$, je $y_n = I_n$, a iz monotonosti I_n slijedi

$$(\operatorname{cond} g_n)(I_\nu) = \frac{I_\nu}{I_n} \cdot \left(\frac{1}{5}\right)^{\nu-n} < \left(\frac{1}{5}\right)^{\nu-n}, \quad \nu > n,$$

što je ispod 1, tj. greške se prigušuju.

Rekurzija za integral (nastavak)

Ako je $\widehat{I}_\nu$ neka aproksimacija za I_ν, onda za relativne perturbacije vrijedi

$$\left|\frac{\widehat{I}_n - I_n}{I_n}\right| = (\operatorname{cond} g_n)(I_\nu) \cdot \left|\frac{\widehat{I}_\nu - I_\nu}{I_\nu}\right| < \left(\frac{1}{5}\right)^{\nu-n} \cdot \left|\frac{\widehat{I}_\nu - I_\nu}{I_\nu}\right|.$$

Zbog linearnosti funkcije g_n, ova relacija vrijedi za bilo kakve perturbacije, a ne samo male.

- Početna vrijednost $\widehat{I}_\nu$ uopće ne mora biti blizu prave I_ν.
- Možemo uzeti $\widehat{I}_\nu = 0$, čime smo napravili relativnu grešku od 100% u početnoj vrijednosti ...

Rekurzija za integral (nastavak)

- ... a još uvijek dobivamo $\widehat{I}_n$ s relativnom greškom

$$\left|\frac{\widehat{I}_n - I_n}{I_n}\right| < \left(\frac{1}{5}\right)^{\nu - n}, \quad \nu > n.$$

- Povoljnim izborom ν, ocjenu na desnoj strani možemo napraviti po volji malom — ispod tražene točnosti ε.
- Dovoljno uzeti $\nu \geq n + \dfrac{\log(1/\varepsilon)}{\log 5}$, i $\widehat{I}_\nu = 0$ i računamo vrijednosti

$$\widehat{I}_{k-1} = \frac{1}{5}\left(\frac{1}{k} - \widehat{I}_k\right), \quad k = \nu, \nu - 1, \ldots, n + 1.$$

Rekurzija unatrag — rezultati

Pitanje: Kako se dobra uvjetovanost vidi, kad stvarno računamo $g_n(I_\nu)$?

Pokaži program i rezultate za $\varepsilon = 10^{-19}$!

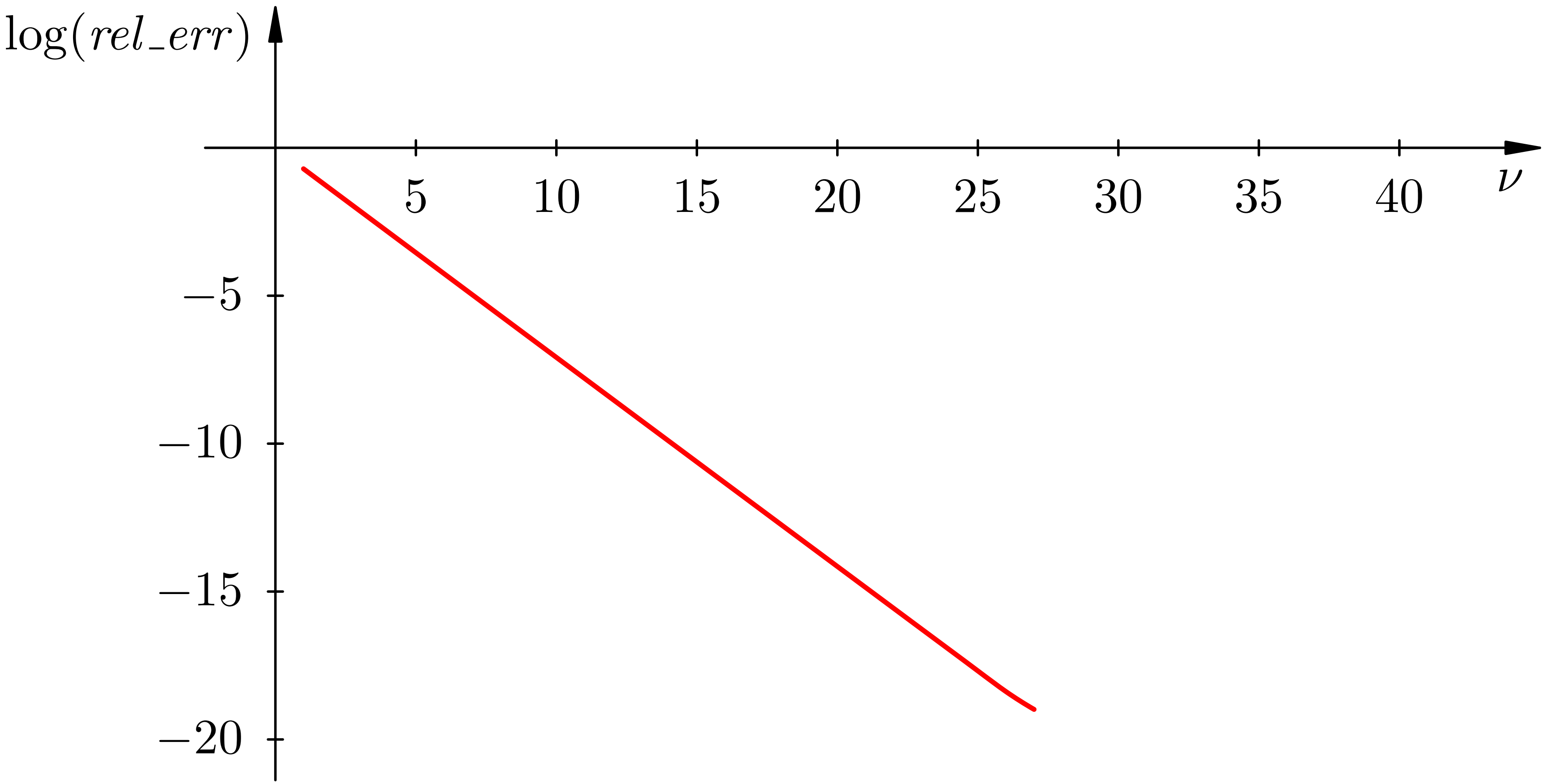
- Za ovaj ε dobijemo

$$\nu \geq n + \frac{\log(1/\varepsilon)}{\log 5} \approx n + 28.$$

  Dakle, “silazno” računamo 28 vrijednosti.

- Stvarna početna vrijednost je $\widehat{I}_\nu = 0$.

Rekurzija unazad za I_{40} — ovisno o startu ν



($\log_{10}$) relativne greške izračunate vrijednosti integrala I_{40} obratnom rekurzijom za $I_{40+\nu} = 0$