Numerička analiza

5. 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

- Teorija linearnih sustava:
  - Rješenje linearnog sustava.
  - Relativna greška unazad.
  - Teorija perturbacije linearnih sustava.
- Rješavanje linearnih sustava: direktne metode:
  - Gaussove eliminacije.
  - Pivotiranje i pivotni rast.
  - Simetrične pozitivno definitne matrice i faktorizacija Choleskog.

Teorija linearnih sustava

Općenito o linearnim sustavima — teorija

Neka je $\mathbb{F} = \mathbb{R}$ polje realnih brojeva (može i $\mathbb{F} = \mathbb{C}$).

Zadani su:

- (pravokutna) matrica $A \in \mathbb{F}^{m \times n}$ i vektor $b \in \mathbb{F}^m$.

Tražimo rješenje linearnog sustava

$$Ax = b.$$

Teorem Kronecker–Capelli kaže da linearni sustav $Ax = b$

- ima rješenje $x \in \mathbb{F}^n$ — ako i samo ako je rang matrice A, u oznaci r, jednak rangu proširene matrice $\widehat{A} = [A \mid b]$,
- rješenje sustava je jedinstveno ako je $r = n$.

Znamo čak i malo više!

Linearni sustavi — teorija (nastavak)

Opće rješenje sustava $Ax = b$ (ako postoji) ima oblik

$$x = x_p + \mathcal{N}(A),$$

gdje je

- x_p jedno partikularno rješenje polaznog sustava $Ax = b$,
- $\mathcal{N}(A)$ je nul-potprostor od A, ili opće rješenje pripadnog homogenog sustava $Ax = 0$.

Iz teorema o rangu i defektu za matricu A

$$r + \dim \mathcal{N}(A) = n,$$

slijedi tvrdnja o jedinstvenosti rješenja:

- $\dim \mathcal{N}(A) = 0 \iff r = n$.

Linearni sustavi — od teorije prema praksi

U praksi se najčešće rješavaju linearni sustavi $Ax = b$ kod kojih je matrica A kvadratna i regularna.

- A kvadratna — znači da je $m = n$ (A je reda n).
- A regularna — znači, na primjer, $\det A \neq 0$.

Iz teorema Kronecker–Capelli onda izlazi da

- rješenje x takvog sustava postoji i jedinstveno je.

$\Longrightarrow$ ima smisla promatrati algoritme za računanje rješenja.

Nema smisla računati nešto što (možda) i ne postoji, ili nije jedinstveno (koje od mnogo rješenja računamo).

Oprez: To što unaprijed znamo da je A regularna

- ne znači da to vrijedi i numerički!

Linearni sustavi — od teorije prema praksi (n.)

Obzirom da ćemo sustav rješavati nekim algoritmom na računalu (u aritmetici konačne preciznosti) neminovno će se u dobivenom rješenju akumulirati greške.

- Već smo spomenuli da je kod numeričke analize algoritma često lakše dobiti grešku unazad akumuliranjem greške u ulaznim podacima.
- Ukoliko želimo znati i grešku u samom rješenju, tj. grešku unaprijed, potrebno je koristiti rezultate teorije perturbacija.

Uloga reziduala

Kad smo računali računalom, umjesto pravog rješenja sustava x, dobili smo približno, $\hat{x}$.

Veličinu

$$r = b - A\hat{x},$$

zovemo rezidual rješenja, i njegova norma je najčešće jedina veličina koju možemo izmjeriti a koja nam govori o odstupanju aproksimacije od pravog rješenja.

Rezidual pravog rješenja x je $r = 0$!

Međutim, ako je rezidual

- velik, onda sigurno nismo blizu pravom rješenju,
- ali rezidual može biti malen, a da izračunato rješenje sustava nije niti blizu pravom.

Uloga reziduala (nastavak)

Odnos između greške i reziduala možemo vidjeti iz sljedeće analize:

$$
\begin{aligned}
\frac{\|x-\hat{x}\|}{\|x\|} &= \frac{\|A^{-1}A(x-\hat{x})\|}{\|x\|} \leq \frac{\|A^{-1}\|\|Ax-A\hat{x}\|}{\|x\|} \\
&= \frac{\|A\|\|A^{-1}\|\|b-A\hat{x}\|}{\|A\|\|x\|} = \kappa(A)\frac{\|r\|}{\|A\|\|x\|} \\
&\leq \kappa(A)\frac{\|r\|}{\|Ax\|} = \kappa(A)\frac{\|r\|}{\|b\|}
\end{aligned}
$$

gdje smo koristili konzistentnost operatorske norme, i činjenice da je $r = b - A\hat{x}$ i $Ax = b$.

Uloga reziduala (nastavak)

- Veličinu $\frac{\|r\|}{\|A\|\|x\|}$ ili $\frac{\|r\|}{\|b\|}$ nazivamo relativnom normom reziduala.
- Vidimo da u slučaju loše uvjetovanih matrica rezidual može biti mali po normi, a greška svejedno može biti velika.

Uloga reziduala (nastavak)

Reziduali se obično koriste za poboljšavanje netočnog rješenja linearnog sustava.

To se obično provodi u tri koraka.

- Izračuna se rezidual $r = b - A\hat{x}$, pri čemu je $\hat{x}$ izračunato rješenje sustava.
- Riješi se sustav $Ad = r$, gdje je d korekcija.
- Korekcija se doda izračunatom rješenju

  $$y = \hat{x} + d,$$

  što bi trebalo dati bolje rješenje.

Relativna greška unazad

Teorem (Rigal i Gaches). Greška unazad (po normi) koja se definira kao

$$\begin{aligned}\eta_{E,f}(\hat{x}) := \min\{\varepsilon :& (A+\Delta A)\hat{x} = b + \Delta b, \\ & \|\Delta A\| \leq \varepsilon \|E\|,\ \|\Delta b\| \leq \varepsilon \|f\|\}\end{aligned}$$

dana je sa

$$\eta_{E,f}(\hat{x}) = \frac{\|r\|}{\|E\| \|\hat{x}\| + \|f\|},$$

gdje je $r = b - A\hat{x}$.

Ako su $E = A$ i $f = b$, tada se $\eta_{E,f}(\hat{x})$ naziva relativnom greškom unazad (po normi).

Relativna greška unazad (nastavak)

Dokaz. Krenimo od jednakosti

$$A\hat{x} + \Delta A\hat{x} = b + \Delta b,$$

i preuredimo je tako da sve greške budu na jednoj strani:

$$r = b - A\hat{x} = \Delta A\hat{x} - \Delta b.$$

Uzimajući normu u gornjem izrazu, koristeći konzistentnost operatorske norme i ograde na norme grešaka $\|\Delta A\| \leq \varepsilon \|E\|$, $\|\Delta b\| \leq \varepsilon \|f\|$ dobivamo

$$\|r\| \leq \|\Delta A\| \|\hat{x}\| + \|\Delta b\| \leq \varepsilon (\|E\| \|\hat{x}\| + \|f\|).$$

Relativna greška unazad (nastavak)

Dakle, za neku ogradu greške ε vrijedi da je

$$\varepsilon \geq \frac{\|r\|}{\|E\| \|\hat{x}\| + \|f\|},$$

pa to mora vrijediti i za minimalnu ogradu $\eta_{E,f}(\hat{x})$.

S druge strane, ako je norma $\| \ \|$ jednaka 2-normi, tada perturbacije

$$\Delta A_{min} = \frac{\|E\| \|\hat{x}\|}{\|E\| \|\hat{x}\| + \|f\|} r \frac{\hat{x}^T}{\hat{x}^T \hat{x}}, \quad \Delta b_{min} = -\frac{\|f\|}{\|E\| \|\hat{x}\| + \|f\|} r$$

zadovoljavaju jednakost

$$(A + \Delta A_{min}) \hat{x} = b + \Delta b_{min}.$$

Relativna greška unazad (nastavak)

Inače, za općenitu operatorsku normu definiramo

$$\Delta A_{min} = \frac{\|E\| \|\hat{x}\|}{\|E\| \|\hat{x}\| + \|f\|} r z^T,$$

gdje je z vektor sa svojstvima

$$z^T \hat{x} = \max_{y \neq 0} \frac{|z^T y|}{\|y\|} \|\hat{x}\| = 1.$$

Takav vektor postoji prema Teoremu o dualnosti. Norme tih perturbacija su jednake

$$\|\Delta A_{min}\| = \frac{\|r\|}{\|E\| \|\hat{x}\| + \|f\|} \|E\|, \quad \|\Delta b_{min}\| = \frac{\|r\|}{\|E\| \|\hat{x}\| + \|f\|} \|f\|.$$

Relativna greška unazad (nastavak)

Zaključujemo onda da je zbog minimalnosti

$$\eta_{E,f}(\hat{x}) \leq \frac{\|r\|}{\|E\|\|\hat{x}\| + \|f\|},$$

Teorija perturbacije linearnih sustava

Teorija perturbacije linearnih sustava bavi se ocjenom (po elementima i/ili po normi) koliko se najviše promijeni rješenje sustava x, ako se malo promijene elementi A i/ili b.

Problem. Neka je

$$Ax = b,$$

$A \in \mathbb{F}^{n \times n}$ regularna, a b zadani vektor.

Zanima nas koliko će se najviše promijeniti rješenje x ovog problema, ako perturbiramo A, odnosno, b.

- Pojednostavnimo problem i pretpostavimo da je A fiksna matrica, a dozvoljene su perturbacije samo vektora b.

Koliko je dobro uvjetovan linearni sustav?

Za prethodni problem

- ulazni podaci su elementi od A i b — njih $n^2 + n$,
- rezultat je vektor $x \in \mathbb{F}^n$, a
- pripadna funkcija problema je $f : \mathbb{F}^n \to \mathbb{F}^n$ je

$$x = f(b) := A^{-1}b.$$

Iz prethodnog predavanja znamo da je $\nabla f(b) = A^{-1}$ (samo, umjesto x, pišemo b), a relativna uvjetovanost problema je jednaka

$$\kappa_{\mathrm{rel}}(b) = \frac{\|b\|_2 \|A^{-1}\|_2}{\|A^{-1}b\|_2} = \frac{\|Ax\|_2 \|A^{-1}\|_2}{\|x\|_2}.$$

Koliko je dobro uvjetovan linearni sustav? (n.)

Ako gledamo najgoru moguću uvjetovanost, po svim vektorima b,

$$\kappa(A) = \max_{\substack{b\in\mathbb{F}^n \\ b\neq 0}} \kappa_{\mathrm{rel}}(b) = \max_{\substack{x\in\mathbb{F}^n \\ x\neq 0}} \frac{\|Ax\|_2}{\|x\|_2} \, \|A^{-1}\|_2 = \|A\|_2\|A^{-1}\|_2.$$

Dakle, vrijedi

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta b\|}{\|b\|}.$$

Uvjetovanost matrice $\kappa(A)$ stalno se pojavljuje kao ocjena greški vezanih uz linearne sustave — VAŽAN PARAMETAR.

Primjer: Hilbertova matrica

Primjer. Kod aproksimacije polinomima javljaju se linearni sustavi oblika

$$H_n x = b,$$

gdje je H_n Hilbertova matrica reda n, tj.

$$H_n = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \cdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}.$$

Primjer: Hilbertova matrica (nastavak)

Da bismo ispitali točnost rješenja, stavimo desnu stranu

$$b(i) := \sum_{j=1}^{n} H_n(i,j) = \sum_{j=1}^{n} \frac{1}{i+j-1}, \quad i = 1, \ldots, n,$$

tako da je rješenje sustava $x^T = [1, 1, \ldots, 1]$.

Što možemo očekivati od rješenja takvog sustava?

Pogled na Frobeniusovu normu matrice A kaže da ona nije naročito velika, tj.

$$\|H_n\|_F = \sqrt{\sum_{i=1}^{n}\sum_{i=1}^{n} \left|\frac{1}{i+j-1}\right|} \leq \sqrt{\sum_{i=1}^{n}\sum_{i=1}^{n} 1} = n.$$

Primjer: Hilbertova matrica (nastavak)

Međutim. . . ne treba gledati samo normu matrice!!!

Uvjetovanost Hilbertovih matrica je vrlo visoka:

n	$\kappa_2(H_n)$
2	$1.928 \cdot 10^{1}$
3	$5.241 \cdot 10^{2}$
4	$1.551 \cdot 10^{4}$
5	$4.766 \cdot 10^{5}$
6	$1.495 \cdot 10^{7}$
7	$4.754 \cdot 10^{8}$
8	$1.526 \cdot 10^{10}$

n	$\kappa_2(H_n)$
9	$4.932 \cdot 10^{11}$
10	$1.603 \cdot 10^{13}$
11	$5.231 \cdot 10^{14}$
12	$1.713 \cdot 10^{16}$
13	$5.628 \cdot 10^{17}$
14	$1.853 \cdot 10^{19}$

n	$\kappa_2(H_n)$
15	$6.117 \cdot 10^{20}$
16	$2.022 \cdot 10^{22}$
17	$6.697 \cdot 10^{23}$
18	$2.221 \cdot 10^{25}$
19	$7.376 \cdot 10^{26}$
20	$2.452 \cdot 10^{28}$

Primjer: Hilbertova matrica — $n = 2, 5$

Za sustav s Hilbertovom matricom, u `extended` točnosti, umjesto svih jedinica u rješenju, dobivamo:

Red 2

$x(1) = 1.000000000000000 \qquad x(2) = 1.000000000000000$

Red 5

$x(1) = 1.000000000000000 \qquad x(4) = 0.999999999999990$

$x(2) = 0.999999999999999 \qquad x(5) = 1.000000000000005$

$x(3) = 1.000000000000007$

Primjer: Hilbertova matrica — $n = 10$

$$
\begin{array}{ll}
x(1) = 1.0000000000003436 & x(6) = 0.9999999294831902 \\
x(2) = 0.9999999999710395 & x(7) = 1.0000001151701616 \\
x(3) = 1.0000000006068386 & x(8) = 0.9999998890931838 \\
x(4) = 0.9999999945453735 & x(9) = 1.0000000580638087 \\
x(5) = 1.0000000258066880 & x(10) = 0.9999999872591526
\end{array}
$$

Uvjetovanost: $\approx 1.603 \cdot 10^{13}$.

Primjer: Hilbertova matrica — $n = 15$

$$x(1) = 1.000000005406387 \qquad x(9) = 1.095291944304200$$
$$x(2) = 0.999999069805858 \qquad x(10) = 0.879782036388407 0$$
$$x(3) = 1.000003979094857 3 \qquad x(11) = 1.099467144423633 3$$
$$x(4) = 0.999925752566044 7 \qquad x(12) = 0.950810251115830 0$$
$$x(5) = 1.000754345227162 1 \qquad x(13) = 1.010602710894005 0$$
$$x(6) = 0.995323419079559 7 \qquad x(14) = 1.001234684115326 1$$
$$x(7) = 1.018864367456238 3 \qquad x(15) = 0.999225202937702 3$$
$$x(8) = 0.948714254434183 8$$

Uvjetovanost: $\approx 6.117 \cdot 10^{20}$.

Primjer: Hilbertova matrica — $n = 20$

$$
\begin{aligned}
x(1) &= 1.0000000486333029 & x(11) &= 231.3608002738048500\\
x(2) &= 0.9999865995557111 & x(12) &= -60.5143391625873562\\
x(3) &= 1.0008720556363132 & x(13) &= -57.6674972682886125\\
x(4) &= 0.9760210562677670 & x(14) &= 5.1760567992057506\\
x(5) &= 1.3512820600312678 & x(15) &= 8.7242780841976215\\
x(6) &= -2.0883247796748707 & x(16) &= 210.1722288687690970\\
x(7) &= 18.4001541798146106 & x(17) &= -413.9544667202651170\\
x(8) &= -63.8982130462650081 & x(18) &= 349.7671855031355400\\
x(9) &= 161.8392478869777220 & x(19) &= -142.9134532513063250\\
x(10) &= -254.7902985140752950 & x(20) &= 25.0584794423327874
\end{aligned}
$$

Uvjetovanost $\approx 2.452 \cdot 10^{28}$.

Primjer: Uvjetovanost Hilbertovih matrica

Može se pokazati da za uvjetovanost Hilbertove matrice H_n vrijedi formula

$$\kappa_2(H_n) \approx \frac{\left(\sqrt{2}+1\right)^{4n+4}}{2^{15/4}\sqrt{\pi n}} \quad \text{za} \quad n \to \infty.$$

Dakle, iako Hilbertove matrice imaju "idealna" svojstva,

- simetrične, pozitivno definitne (čak totalno pozitivne),

njihova uvjetovanost katastrofalno brzo raste!

"Krivci" za to su elementi inverza H_n^{-1}.

Primjer: Inverz Hilbertove matrice

Recimo, H_5^{-1} izgleda ovako:

$$H_5^{-1} = \begin{bmatrix} 25 & -300 & 1050 & -1400 & 630 \\ -300 & 4800 & -18900 & 26880 & -12600 \\ 1050 & -18900 & 79380 & -117600 & 56700 \\ -1400 & 26880 & -117600 & 179200 & -88200 \\ 630 & -12600 & 56700 & -88200 & 44100 \end{bmatrix}.$$

A kako tek izgledaju elementi H_{20}^{-1}?

Primjer: Inverz Hilbertove matrice (nastavak)

Elementi inverza H_n^{-1} Hilbertove matrice mogu se eksplicitno izračunati u terminima binomnih koeficijenata

$$(H_n^{-1})_{ij} = (-1)^{i+j}\,(i+j-1)\cdot$$
$$\cdot \binom{n+i-1}{n-j}\binom{n+j-1}{n-i}\binom{i+j-2}{i-1}^2.$$

Lako se vidi da ovi elementi vrlo brzo rastu za malo veće n.

Pogledajte

http://mathworld.wolfram.com/HilbertMatrix.html

Još malo o perturbacijama linearnih sustava

Ocjenu koliko se najviše promijenilo rješenje sustava $Ax = b$ možemo dobiti i direktno i to po elementima/normi

- ako perturbiramo samo A ili samo b,
- ako perturbiramo i A i b.

Pretpostavimo da smo perturbirali samo A. Umjesto sustava $Ax = b$ tada rješavamo

$$(A + \Delta A)(x + \Delta x) = b.$$

Također, možemo pretpostaviti da za operatorsku normu perturbacije vrijedi

$$\|\Delta A\| \leq \varepsilon \|A\|.$$

Komentar. Ako je ε točnost računanja, tolika perturbacija napravljena je već pri spremanju elemenata matrice u računalo.

Perturbacija matrice A

Oduzimanjem $Ax = b$ od $(A + \Delta A)(x + \Delta x) = b$ dobivamo

$$A\,\Delta x + \Delta A\,(x + \Delta x) = 0.$$

Množenjem slijeva s A^{-1} i sređivanjem dobivamo

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

Uzimanjem norme lijeve i desne strane, a zatim ocjenjivanjem odozgo, dobivamo

$$\begin{aligned} \|\Delta x\| &\leq \|A^{-1}\|\,\|\Delta A\|\,\|x + \Delta x\| \leq \varepsilon\,\|A^{-1}\|\,\|A\|\,\|x + \Delta x\| \\ &\leq \varepsilon\,\kappa(A)\,(\|x\| + \|\Delta x\|), \end{aligned}$$

pri čemu je $\kappa(A) = \|A\|\,\|A^{-1}\|$ uvjetovanost matrice A.

Perturbacija matrice A (nastavak)

Premještanjem na lijevu stranu svih pribrojnika koji sadrže Δx dobivamo

$$(1 - \varepsilon\,\kappa(A))\,\|\Delta x\| \leq \varepsilon\,\kappa(A)\,\|x\|.$$

Ako je $\varepsilon\,\kappa(A) < 1$, a to znači i $\|\Delta A\|\,\|A^{-1}\| < 1$, onda je

$$\|\Delta x\| \leq \frac{\varepsilon\,\kappa(A)}{1 - \varepsilon\,\kappa(A)}\,\|x\|,$$

tj.

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\varepsilon\,\kappa(A)}{1 - \varepsilon\,\kappa(A)},$$

što pokazuje da je pogreška u rješenju približno proporcionalna uvjetovanosti matrice A.

Perturbacija vektora b

Pretpostavimo sad da, umjesto sustava $Ax = b$ rješavamo

$$A(x + \Delta x) = b + \Delta b.$$

Ponovno pretpostavljamo da je operatorska norma perturbacije vektora b

$$\|\Delta b\| \leq \varepsilon \|b\|.$$

Oduzimanjem $Ax = b$ od $A(x + \Delta x) = b + \Delta b$ izlazi

$$A\,\Delta x = \Delta b.$$

Množenjem slijeva s A^{-1} dobivamo

$$\Delta x = A^{-1}\,\Delta b.$$

Perturbacija vektora b (nastavak)

Uzimanjem norme lijeve i desne strane, a zatim ocjenjivanjem odozgo, dobivamo

$$\begin{aligned} \|\Delta x\| &\leq \|A^{-1}\| \, \|\Delta b\| \leq \varepsilon \, \|A^{-1}\| \, \|b\| \leq \varepsilon \, \|A^{-1}\| \, \|Ax\| \\ &\leq \varepsilon \, \|A^{-1}\| \, \|A\| \, \|x\| \leq \varepsilon \, \kappa(A) \, \|x\|, \end{aligned}$$

tj.

$$\frac{\|\Delta x\|}{\|x\|} \leq \varepsilon \, \kappa(A),$$

što pokazuje da je pogreška u rješenju, ponovno, proporcionalna uvjetovanosti matrice A.

Sada možemo generalizirati rezultat ako perturbiramo istovremeno i A i b.

Perturbacija matrice A i vektora b

Teorem. Neka je $Ax = b$ i

$$(A + \Delta A)\,(x + \Delta x) = b + \Delta b,$$

gdje je

$$\|\Delta A\| \leq \varepsilon\,\|E\|, \quad \|\Delta b\| \leq \varepsilon\,\|f\|,$$

pri čemu je E neka matrica, a f neki vektor. Također, neka je

$$\varepsilon\,\|A^{-1}\|\,\|E\| < 1.$$

Tada za $x \neq 0$ vrijedi ocjena

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\varepsilon}{1 - \varepsilon\,\|A^{-1}\|\,\|E\|} \left(\frac{\|A^{-1}\|\,\|f\|}{\|x\|} + \|A^{-1}\|\,\|E\| \right).$$

Perturbacija matrice A i vektora b (nastavak)

Komentar: Uobičajeno se za E uzima A, jer je to pogreška koju napravimo spremanjem matrice A u računalo. Jednako tako, za f se uzima b. U tom slučaju je

$$
\begin{aligned}
\frac{\|\Delta x\|}{\|x\|} &\leq \frac{\varepsilon}{1 - \varepsilon\,\|A^{-1}\|\,\|A\|} \left(\frac{\|A^{-1}\|\,\|b\|}{\|x\|} + \|A^{-1}\|\,\|A\| \right) \\
&= \frac{\varepsilon}{1 - \varepsilon\,\kappa(A)} \left(\frac{\|A^{-1}\|\,\|Ax\|}{\|x\|} + \kappa(A) \right) \\
&\leq \frac{\varepsilon}{1 - \varepsilon\,\kappa(A)} \left(\frac{\|A^{-1}\|\,\|A\|\,\|x\|}{\|x\|} + \kappa(A) \right) \\
&= \frac{2\varepsilon\,\kappa(A)}{1 - \varepsilon\,\kappa(A)}.
\end{aligned}
$$

Perturbacija matrice A i vektora b (nastavak)

Dokaz. Provodi se na sličan način kao za pojedinačne perturbacije.

Ako od $(A + \Delta A)\,(x + \Delta x) = b + \Delta b$ oduzmemo $Ax = b$, dobivamo

$$A\,\Delta x = \Delta b - \Delta A\,x - \Delta A\,\Delta x.$$

Množenjem s A^{-1} slijeva i sređivanjem dobivamo

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

Uzimanjem norme lijeve i desne strane, a zatim ocjenjivanjem odozgo, dobivamo

Perturbacija matrice A i vektora b (nastavak)

$$\begin{aligned}\|\Delta x\| &\leq \|A^{-1}\|\|\Delta b\| + \|A^{-1}\|\|\Delta A\|(\|x\| + \|\Delta x\|)\\ &\leq \varepsilon\|A^{-1}\|\|f\| + \varepsilon\|A^{-1}\|\|E\|(\|x\| + \|\Delta x\|)\end{aligned}$$

Premještanjem na lijevu stranu svih pribrojnika koji sadrže Δx dobivamo

$$(1 - \varepsilon\|A^{-1}\|\|E\|)\|\Delta x\| \leq \varepsilon(\|A^{-1}\|\|f\| + \|A^{-1}\|\|E\|\|x\|).$$

Budući da je $\varepsilon\|A^{-1}\|\,\|E\| < 1$ tada vrijedi

$$\|\Delta x\| \leq \frac{\varepsilon}{1 - \varepsilon\|A^{-1}\|\|E\|}(\|A^{-1}\|\|f\| + \|A^{-1}\|\|E\|\|x\|).$$

Rješavanje linearnih sustava: direktne metode

Kako naći rješenje? — Inverz matrice

Teorija (1). Možemo naći inverz matrice A, tj. matricu A^{-1}

- i slijeva pomnožiti sustav $Ax = b$ matricom A^{-1}.

Dobivamo

$$x = A^{-1}b.$$

Samo je pitanje: kako ćemo izračunati inverz A^{-1}?

Zaključak: Lakši problem sveli smo na teži — u prijevodu, pali smo s konja na magarca.

Zašto? Jednostavno, zato što je

- j-ti stupac inverza $=$ rješenje sustava $Ax = e_j$.

Dakle, za n stupaca od A^{-1} treba riješiti n linearnih sustava. A krenuli smo od jednog (sustava)! Ne tako!

Kako naći rješenje? — Cramerovo pravilo

Teorija (2). Iz linearne algebre znate za Cramerovo pravilo:

- j-ta komponenta rješenja sustava je

$$x_j = \frac{\det A_j}{\det A}, \quad j = 1, \ldots, n,$$

- pri čemu je matrica A_j jednaka matrici A,
- osim što je j-ti stupac u A_j zamijenjen desnom stranom b.

Treba još "samo" izračunati determinante — i to $n + 1$ njih.
A kako ćemo to?

"Klasični" odgovor: pa... recimo, Laplaceovim razvojem.

Jao, jao... Bilo kako, samo ne tako!

Kako naći rješenje? — Zaboravite Cramera

Zašto? Laplaceovim razvojem dobijemo

- "samo" $n!$ pribrojnika u svakoj determinanti,
- a svaki pribrojnik je produkt od n faktora.

Prava "sitnica". I tako to, još $n + 1$ puta ...

Zaključak: Ako determinante računamo na ovaj način,

- složenost Cramerovog pravila za rješavanja linearnog sustava je eksponencijalna u n (dokažite to!)
- i nikad se ne koristi kao metoda numeričkog rješavanja.

Zaboravite determinante i Cramera — finale

Komentar: Determinante možemo računati i puuuno brže,

- tako da matricu svedemo na trokutasti oblik,
- postupkom sličnim Gaussovim eliminacijama.

Naime, determinanta trokutaste matrice (gornje ili donje) je

- produkt dijagonalnih elemenata,

pa se lako računa!

No, isti postupak eliminacija koristimo i za

- rješavanje "cijelog" linearnog sustava $Ax = b$,
- i to samo jednom, a ne $n + 1$ puta.

Dakle, Cramerovo pravilo se ne isplati ni kad ovako računamo determinante.

Kako naći rješenje? — Gaussove eliminacije!

Najjednostavnija metoda za rješavanje linearnih sustava su

- Gaussove eliminacije, odnosno
- slične metode svođenja na trokutastu formu.

Ideja: Sustav $Ax = b$ se ekvivalentim transformacijama svodi na sustav oblika

$$Rx = y,$$

gdje je

- R trokutasta matrica (recimo, gornja),

iz kojeg se lako, tzv. povratnom supstitucijom, nalazi rješenje.

Oznaka R — “right” (desna) = gornja trokutasta matrica.

Gaussove eliminacije — komentari

Par komentara, prije detaljnog opisa metode.

Gaussove eliminacije su metoda direktnog transformiranja linearnog sustava $Ax = b$, zajedno s desnom stranom b.

Možemo ih implementirati i tako da se desna strana b ne transformira istovremeno kad i matrica A.

- Tada se formiraju dvije matrice L i R takve da je $A = LR$, gdje je R gornja trokutasta matrica iz Gaussovih eliminacija, a L je donja trokutasta matrica.
- Tako implementirane Gaussove eliminacije zovemo LR (ili LU) faktorizacija matrice A — standard u praksi.
- Ovaj pristup je posebno zgodan kad imamo više desnih strana za isti A.

Gaussove eliminacije — komentari (nastavak)

U praksi se koriste za “opće”, ali ne pretjerano velike matrice (n u tisućama), ili za sustave s tzv. “vrpčastom” strukturom.

Složenost: polinomna i to kubna, tj. $O(n^3)$, što je sporo za još veće sustave. Za njih se koriste iterativne metode.

Mnogi sustavi imaju specijalna svojstva koja koristimo za brže i/ili točnije rješenje. Na primjer,

- za simetrične, pozitivno definitne matrice koristi se “simetrična” LR faktorizacija, tzv. faktorizacija Choleskog,
- za dijagonalno dominantne sustava ne treba pivotiranje,
- za vrpčaste, posebno tridijagonalne matrice, algoritam se drastično skraćuje (v. kubična spline interpolacija).

Gaussove eliminacije — algoritam

Općenito, k–ti korak izgleda ovako, za $k = 1, \ldots, n-1$:

- iz proširene matrice $[A^{(k)} \mid b^{(k)}]$ dobivamo novu proširenu matricu $[A^{(k+1)} \mid b^{(k+1)}]$,
- tako da poništimo sve elemente ispod dijagonale u k–tom stupcu matrice $A^{(k)}$.

Relacije za nove elemente koje treba izračunati u matrici $A^{(k+1)}$ i vektoru $b^{(k+1)}$ su

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}, \qquad b_i^{(k+1)} = b_i^{(k)} - m_{ik} b_k^{(k)},$$

za $i, j = k+1, \ldots, n$, a multiplikatori m_{ik} su

$$m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \qquad i = k+1, \ldots, n.$$

Gaussove eliminacije — algoritam (nastavak)

Prvih k redaka u $[A^{(k+1)} \mid b^{(k+1)}]$ ostaju isti kao u $[A^{(k)} \mid b^{(k)}]$.

Konačno, ako su svi $a_{ii}^{(i)} \neq 0$, za $i = 1, \ldots, n-1$, završni linearni sustav $[A^{(n)} \mid b^{(n)}]$, ekvivalentan polaznom, je

$$\left[\begin{array}{cccc:c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & \vdots & \vdots \\ & & & a_{nn}^{(n)} & b_n^{(n)} \end{array}\right].$$

Dobili smo gornju trokutastu matricu $R = A^{(n)}$ (nule u donjem trokutu matrice R ne pišemo).

Gaussove eliminacije — algoritam (nastavak)

Uz pretpostavku da je $a_{nn}^{(n)} \neq 0$, ovaj se linearni sustav lako rješava povratnom supstitucijom

$$
\begin{aligned}
x_n &= \frac{b_n^{(n)}}{a_{nn}^{(n)}}, \\
x_i &= \frac{1}{a_{ii}^{(i)}} \left(b_i^{(i)} - \sum_{j=i+1}^{n} a_{ij}^{(i)} x_j \right), \quad i = n-1, \ldots, 1.
\end{aligned}
$$

Pitanje: Ako je A kvadratna i regularna,

- moraju li svi elementi $a_{ii}^{(i)}$ biti različiti od nule?

To je nužno (i dovoljno) da algoritam "prođe" u ovom obliku.

Gaussove eliminacije — primjedba na algoritam

Odgovor: Ne!

Primjer. Linearni sustav $Ax = b$ s proširenom matricom

$$[A \mid b] = \left[\begin{array}{cc:c} 0 & 1 & 1 \\ 1 & 0 & 1 \end{array}\right]$$

je regularan ($\det A = -1$), sustav ima jedinstveno rješenje $x_1 = x_2 = 1$,

- a ipak ga ne možemo riješiti Gaussovim eliminacijama,
- ako ne mijenjamo poredak jednadžbi.

Zaključak: Moramo dozvoliti promjenu poretka jednadžbi.

Gaussove eliminacije s pivotiranjem

Pitanje: Dozvolimo li promjene poretka jednadžbi — tzv. "pivotiranje" u stupcu kojeg sređujemo,

- može li se Gaussovim eliminacijama s pivotiranjem riješiti svaki sustav kojemu je matrica kvadratna i regularna?

Odgovor: Ako dozvolimo pivotiranje

- zamjenom "ključne" jednadžbe i bilo koje druge koja ima ne-nula element (u tom stupcu),
- Gaussovim eliminacijama rješiv je svaki regularni kvadratni linearni sustav.

Objašnjenje: Ako u prvom stupcu nemamo ne-nula elemenata, matrica je singularna. Isto vrijedi i za svaki sljedeći korak (Laplaceov razvoj determinante!).

Gaussove eliminacije — kako pivotiramo?

Pitanje: Kako vršiti pivotiranje, tj. zamjene jednadžbi?

- Zamjenom “ključne” jednadžbe i bilo koje druge koja ima ne-nula element (u tom stupcu)?

Odgovor: Tu je ključna razlika između egzaktnog i približnog računanja (kad imamo greške zaokruživanja).

- U teoriji — kod egzaktnog računanja, dovoljno je naći bilo koji ne-nula element (u tom stupcu).
- U praksi — kad računamo približno, to može dovesti do potpuno pogrešnog rezultata.

Jedna jedina operacija može upropastiti rezultat!

- Postoji puno bolja strategija za pivotiranje, kojom se to (barem dijelom) može izbjeći.

Gaussove eliminacije — primjer

Primjer. Zadan je linearni sustav

$$
\begin{aligned}
0.0001\, x_1 + x_2 &= 1 \\
x_1 + x_2 &= 2.
\end{aligned}
$$

Matrica sustava je regularna, pa postoji jedinstveno rješenje

$$x_1 = 1.\dot{0}00\dot{1}, \quad x_2 = 0.\dot{9}99\dot{8}.$$

Riješimo taj sustav "računalom" koje ima 4 decimalne znamenke mantise i 2 znamenke eksponenta.

Uočiti: Broj 0.0001 je "mali", ali nije nula. Po teoriji,

- možemo ga uzeti kao prvi (ili bilo koji) ne-nula element.

Gaussove eliminacije — primjer (nastavak)

Sustav zapisan u takvom “računalu” pamti se kao

$$
\begin{aligned}
1.000 \cdot 10^{-4}\, x_1 + 1.000 \cdot 10^{0}\, x_2 &= 1.000 \cdot 10^{0} \\
1.000 \cdot 10^{0} \quad x_1 + 1.000 \cdot 10^{0}\, x_2 &= 2.000 \cdot 10^{0}.
\end{aligned}
$$

Množenjem prve jednadžbe s -10^4 i dodavanjem drugoj, dobivamo novu drugu jednadžbu

$$
(1.000 \cdot 10^{0} - 1.000 \cdot 10^{4})\, x_2 = 2.000 \cdot 10^{0} - 1.000 \cdot 10^{4}.
$$

Oduzimanje u računalu se vrši tako da manji eksponent postane jednak većem, a mantisa se denormalizira. Dobivamo

$$
\begin{aligned}
1.000 \cdot 10^{0} = 0.100 \cdot 10^{1} = 0.010 \cdot 10^{2} &= 0.001 \cdot 10^{3} \\
&= 0.000|1 \cdot 10^{4}.
\end{aligned}
$$

Gaussove eliminacije — primjer (nastavak)

Za zadnju jedinicu nema mjesta u mantisi, pa je mantisa postala 0, tj. prvi broj “nema” utjecaja na rezultat. Slično se dogodi i s desnom stranom (i 2 je “zanemariv” prema 10^4).

Dakle, nova druga jednadžba glasi

$$-1.000 \cdot 10^4 \, x_2 = -1.000 \cdot 10^4.$$

Rješenje ove jednadžbe je očito $x_2 = 1.000 \cdot 10^0$. Uvrštavanjem u prvu jednadžbu, dobivamo

$$\begin{aligned} 1.000 \cdot 10^{-4} \, x_1 &= 1.000 \cdot 10^0 - 1.000 \cdot 10^0 \, \cdot \, 1.000 \cdot 10^0 \\ &= 0.000 \cdot 10^0, \end{aligned}$$

pa je $x_1 = 0$, što nije niti približno točan rezultat.

Gaussove eliminacije — primjer (nastavak)

Razlog za ogromnu relativnu grešku (100%):

- prvu jednadžbu množimo velikim brojem -10^4 (po apsolutnoj vrijednosti) i dodajemo drugoj,
- što "uništava" drugu jednadžbu.

Drugim riječima, utjecaj polazne druge jednadžbe

- postaje zanemariv u novoj drugoj jednadžbi.

U polaznoj drugoj je moglo pisati "bilo što"!

Gaussove eliminacije s pivotiranjem — primjer

Promijenimo li poredak jednadžbi, dobivamo

$$\begin{aligned} 1.000 \cdot 10^{0} \quad x_1 + 1.000 \cdot 10^{0}\, x_2 &= 2.000 \cdot 10^{0} \\ 1.000 \cdot 10^{-4}\, x_1 + 1.000 \cdot 10^{0}\, x_2 &= 1.000 \cdot 10^{0}. \end{aligned}$$

Množenjem prve jednadžbe s -10^{-4} i dodavanjem drugoj, dobivamo novu drugu jednadžbu

$$(1.000 \cdot 10^{0} - 1.000 \cdot 10^{-4})\, x_2 = 1.000 \cdot 10^{0} - 2.000 \cdot 10^{-4}.$$

Ovdje nema oduzimanja — drugi broj s 10^{-4} je “zanemariv” prema 1. Dakle, nova druga jednadžba sad glasi

$$1.000 \cdot 10^{0}\, x_2 = 1.000 \cdot 10^{0}.$$

Gaussove eliminacije s pivotiranjem — primjer

Ponovno dobivamo rješenje $x_2 = 1.000 \cdot 10^0$. Međutim, uvrštavanjem u prvu jednadžbu dobivamo

$$\begin{aligned} 1.000 \cdot 10^0 \, x_1 &= 2.000 \cdot 10^0 - 1.000 \cdot 10^0 \cdot 1.000 \cdot 10^0 \\ &= 1.000 \cdot 10^0, \end{aligned}$$

pa je $x_1 = 1.000 \cdot 10^0$, što je točan rezultat — korektno zaokruženo egzaktno rješenje na četiri decimalne znamenke!

Razlog za vrlo malu relativnu grešku:

- prvu jednadžbu sad množimo malim brojem -10^{-4} (po apsolutnoj vrijednosti) i dodajemo drugoj,
- što nema utjecaja na drugu jednadžbu — tj. ovdje nema "uništavanja" jednadžbi.

Gaussove eliminacije s pivotiranjem — primjer

Kao i ranije, u koraku eliminacije,

- (bivša) druga jednadžba nema utjecaja na (bivšu) prvu.

Međutim, nakon zamjene

- prva jednadžba (bivša druga) ostaje netaknuta u prvom koraku eliminacije i uredno utječe na rješenje.

Zaključak: Sigurno nije dovoljno uzeti

- prvi (bilo koji) ne-nula element u stupcu

kao ključni element za eliminacije,

- jer možemo dobiti potpuno pogrešan rezultat.

Gaussove eliminacije s pivotiranjem — primjer

Primjer. Usporedimo izračunata rješenja sustava

$$\begin{aligned} \varepsilon\, x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2, \end{aligned}$$

za $\varepsilon = 10^{-1}, \ldots, 10^{-25}$, Gaussovim eliminacijama bez zamjena i sa zamjenom poretka jednadžbi, u aritmetici računala.

Računanjem u najvećoj mogućoj preciznosti (`extended`) dobivamo sljedeću tablicu.

U tablici je x_2 naveden samo jednom — jer ga obje metode izračunaju jednako (i točno)!

U prvom stupcu pišu samo eksponenti p, pri čemu je $\varepsilon = 10^p$.

Gaussove eliminacije s pivotiranjem — primjer

p	x_1 bez pivotiranja	x_1 s pivotiranjem	x_2
-4	1.000100010001000**00**	1.000100010001000**10**	0.999899989998999990
-5	1.000010000100000**200**	1.000010000100000**100**	0.999989999899999900
-6	1.0000010000**099609**	1.0000010000**100000**	0.99999899999900000
-7	1.0000000**9999978538**	1.0000001**0000001000**	0.99999989999999000
	$\vdots$	$\vdots$	
-17	**0.99746599868666408**	**1.00000000000000001**	0.99999999999999999
-18	**0.97578195523695399**	**1.00000000000000000**	1.00000000000000000
-19	**1.08420217248550443**	**1.00000000000000000**	1.00000000000000000
-20	**0.00000000000000000**	**1.00000000000000000**	1.00000000000000000

Gaussove eliminacije s parcijalnim pivotiranjem

Pivotni element uobičajeno se bira korištenjem parcijalnog pivotiranja

- pivotni element je po apsolutnoj vrijednosti najveći u "ostatku" stupca — na glavnoj dijagonali ili ispod nje,

tj. ako je u k-tom koraku

$$|a_{rk}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^{(k)}|,$$

onda ćemo zamijeniti r-ti i k-ti redak i početi korak eliminacije elemenata k-tog stupca.

Gaussove eliminacije s parcijalnim pivotiranjem

Motivacija: elementi “ostatka” linearnog sustava koje treba izračunati u matrici $A^{(k+1)}$ u k-tom koraku transformacije su

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}, \quad b_i^{(k+1)} = b_i^{(k)} - m_{ik} b_k^{(k)},$$

za $i, j = k+1, \ldots, n$, a multiplikatori m_{ik} su

$$m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad i = k+1, \ldots, n.$$

Ako je multiplikator m_{ik} velik, u aritmetici pomičnog zareza može doći do kraćenja najmanje značajnih znamenki $a_{ij}^{(k)}$, tako da izračunati $a_{ij}^{(k+1)}$ može imati veliku relativnu grešku.

Gaussove eliminacije s parcijalnim pivotiranjem

Sasvim općenito, ideja pivotiranja je minimizirati korekcije elemenata pri prijelazu s $A^{(k)}$ na $A^{(k+1)}$. Dakle, multiplikatori trebaju biti što manji.

Za multiplikatore kod parcijalnog pivotiranja vrijedi

$$|m_{ik}| \leq 1, \quad i = k+1, \ldots, n.$$

U praksi, parcijalno pivotiranje funkcionira izvrsno, ali matematičari su konstruirali primjere kad ono "nije savršeno".

Gaussove eliminacije s potpunim pivotiranjem

Osim parcijalnog pivotiranja, može se provoditi i potpuno pivotiranje. U k-tom koraku, bira se maksimalni element u cijelom "ostatku" matrice $A^{(k)}$, a ne samo u k-tom stupcu.

Ako je u k-tom koraku

$$|a_{rs}^{(k)}| = \max_{k \leq i,j \leq n} |a_{ij}^{(k)}|,$$

onda ćemo zamijeniti r-ti i k-ti redak, s-ti i k-ti stupac i početi korak eliminacije elemenata k-tog stupca.

Oprez: zamjenom s-tog i k-tog stupca zamijenili smo ulogu varijabli x_s i x_k.

Ovo nisu jedine mogućnosti pivotiranja kod rješavanja linearnih sustava.

Veza Gaussovih eliminacija i LR faktorizacije

Može se pokazati da je

- matrica R dobivena LR faktorizacijom jednaka
- matrici R dobivenoj Gaussovim eliminacijama.

Neka je

- $A^{(k)}$ matrica na početku k-tog koraka Gaussovih eliminacija,
- a $A^{(k+1)}$ matrica dobivena na kraju tog koraka.

Onda se $A^{(k+1)}$ može matrično napisati kao produkt

$$A^{(k+1)} = M_k A^{(k)},$$

pri čemu je...

Veza Gaussovih eliminacija i LR faktorizacije

$$M_k = \left[\begin{array}{c|ccccc} I_{k-1} & & & & & \\ \hline & 1 & & & & \\ & -m_{k+1,k} & 1 & & & \\ & -m_{k+2,k} & & \ddots & & \\ & \vdots & & & \ddots & \\ & -m_{n,k} & & & & 1 \end{array}\right]$$

a m_{ik} su odgovarajući multiplikatori u k-tom koraku.

Na kraju eliminacija, nakon $n-1$ koraka, dobijemo gornju trokutastu matricu $\widetilde{R}$,

$$\widetilde{R} := A^{(n)} = M_{n-1}\, M_{n-2} \cdots M_1\, A.$$

Veza Gaussovih eliminacija i LR faktorizacije

Sve matrice M_k su regularne, jer su M_k donje trokutaste s 1 na dijagonali, pa postoje njihovi inverzi. Onda se A može napisati kao

$$A = M_1^{-1}\, M_2^{-1} \cdots M_{n-1}^{-1}\, \widetilde{R} := \widetilde{L}\widetilde{R},$$

gdje je

$$\widetilde{L} = \begin{bmatrix} 1 & & & & \\ m_{21} & 1 & & & \\ \vdots & m_{32} & \ddots & & \\ \vdots & \vdots & & \ddots & \\ m_{n1} & m_{n2} & \cdots & m_{n,n-1} & 1 \end{bmatrix}.$$

Iz jedinstvenosti LR faktorizacije slijedi da je $\widetilde{R} = R$.

LR faktorizacija

U praksi se linearni sustavi najčešće rješavaju korištenjem LR faktorizacije — A faktoriziramo kao

$$A = LR,$$

pri čemu je

- L donja trokutasta matrica s jedinicama na dijagonali,
- R gornja trokutasta.

Matrica L je regularna, jer je $\det L = 1$, pa regularnost matrice A povlači i regularnost matrice R, jer je

$$\det A = \det L \cdot \det R = \det R.$$

LR faktorizacija (nastavak)

Ako znamo LR faktorizaciju od A, onda linearni sustav $Ax = b$ postaje

$$LRx = b.$$

Uz oznaku $y = Rx$, sustav $LRx = b$ svodi se na dva sustava

$$Ly = b, \qquad Rx = y.$$

Prednost LR faktorizacije:

- rješavaju se dva jednostavna sustava,
- desna strana b ne transformira se istovremeno s matricom A, pa promjena desne strane košta samo $O(n^2)$ operacija.

LR faktorizacija (nastavak)

Oba sustava se lako rješavaju:

- prvi $Ly = b$ — supstitucijom unaprijed

$$y_1 = b_1,$$
$$y_i = b_i - \sum_{j=1}^{i-1} \ell_{ij}\, y_j, \quad i = 2, \ldots, n,$$

- drugi $Rx = y$ — povratnom supstitucijom

$$x_n = \frac{y_n}{r_{nn}},$$
$$x_i = \frac{1}{r_{ii}} \left(y_i - \sum_{j=i+1}^{n} r_{ij}\, x_j \right), \quad i = n-1, \ldots, 1.$$

Egzistencija i jedinstvenost LR faktorizacije

Ostaje vidjeti uz koje je uvjete $r_{ii} \neq 0$ za $i = 1, \ldots, n$.

Teorem. Postoji jedinstvena LR faktorizacija matrice A ako i samo ako su vodeće glavne podmatrice $A_k := A(1:k, 1:k)$, $k = 1, \ldots, n-1$, regularne.

Ako je A_k singularna za neki k, faktorizacija može postojati, ali nije jedinstvena.

Dokaz. Ide indukcijom po dimenziji matrice: vidi Dodatak.

Parcijalno pivotiranje u LR faktorizaciji

Veza LR faktorizacije i Gaussovih eliminacija upućuje nas da pivotiranje vršimo na isti način kao kod Gaussovih eliminacija.

Ako koristimo parcijalno pivotiranje, onda se LR faktorizacija tako dobivene matrice — permutiranih redaka, zapisuje kao

$$PA = LR,$$

pri čemu je P matrica permutacije.

Matrica permutacije P u svakom retku i stupcu

- ima točno jednu jedinicu, a ostalo su nule.

P je uvijek regularna matrica — pokažite to!

Parcijalno pivotiranje u LR faktorizaciji (nast.)

Ako znamo "permutiranu" faktorizaciju $PA = LR$, kako ćemo riješiti linearni sustav $Ax = b$?

Najjednostavnije je lijevu i desnu stranu (slijeva) pomnožiti s P, pa dobivamo

$$PAx = LRx = Pb.$$

Oprez: kad permutiramo, istovremeno zamjenjujemo retke

- u obje "radne matrice" — $(L - I)$ i R,

tj. permutiramo dosadašnje multiplikatore i jednadžbe.

Potpuno pivotiranje u LR faktorizaciji

Ako koristimo potpuno pivotiranje, dobivamo LR faktorizaciju matrice koja ima permutirane retke i stupce obzirom na A, tj.

$$PAQ = LR,$$

gdje su P i Q matrice permutacije.

Skica rješenja. Q je unitarna, pa iz $PA = LRQ^T$, uz pokratu $Q^T x = z$, imamo

$$PAx = LR(Q^T x) = LRz = Pb.$$

Dakle, jedina razlika obzirom na parcijalno pivotiranje je

- da na kraju treba "izokretati" rješenje z da se dobije x, tj. $x = Qz$.

Parcijalno vs. potpuno pivotiranje

Možemo li i na temelju čega reći da je potpuno pivotiranje "bolje" od parcijalnog? Tradicionalno to se čini na temelju pivotnog rasta.

Pivotni rast (ili "faktor rasta") je omjer

- najvećeg (po apsolutnoj vrijednosti) elementa u svim koracima eliminacije,
- i najvećeg elementa u originalnoj matrici

$$\rho_n = \frac{\max_{i,j,k} |a_{ij}^{(k)}|}{\max_{i,j} |a_{ij}|}.$$

Intuitivno je jasno da nije dobro da elementi rastu po apsolutnoj vrijednosti, jer bi to moglo dovesti do gubitka točnosti.

Pivotni rast

Koliki je pivotni rast kod parcijalnog pivotiranja?

Korištenjem relacija za poništavanje elemenata

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}, \quad |m_{ik}| \leq 1,$$

za parcijalno pivotiranje vrijedi

$$|a_{ij}^{(k+1)}| \leq |a_{ij}^{(k)}| + |a_{kj}^{(k)}| \leq 2 \max_{i,j} |a_{ij}^{(k)}|.$$

Prethodna ocjena, za n koraka algoritma daje pivotni rast ρ_n^p

$$\rho_n^p \leq 2^{n-1}.$$

Pivotni rast (nastavak)

Već je J. H. Wilkinson primijetio da se taj pivotni rast može dostići za sve matrice oblika

$$\begin{bmatrix} 1 & & & & 1 \\ -1 & 1 & & & 1 \\ -1 & -1 & \ddots & & 1 \\ -1 & -1 & \ddots & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix}.$$

Eksponencijalno rastu elementi posljednjeg stupca.

Ovo je samo "umjetno" konstruirani primjer, a u praksi je takvih matrica izrazito malo, pa se parcijalno pivotiranje ponaša mnogo bolje od očekivanog.

Pivotni rast (nastavak)

Za potpuno pivotiranje pivotni rast ρ_n^c može se ograditi odozgo s

$$\rho_n^c \leq n^{1/2}\Big(2 \cdot 3^{1/2} \cdots n^{1/(n-1)}\Big)^{1/2} \approx c\, n^{1/2}\, n^{(\log n)/4},$$

ali ta ograda nije dostižna. Ovo je dokazao J. H. Wilkinson, šezdesetih godina prošlog stoljeća.

Dugo se mislilo da vrijedi

$$\rho_n^c \leq n.$$

Međutim, nađeni su primjeri matrica kad to ne vrijedi.

Kontraprimjer (konstruiran 1991. godine), matrice reda 13 ima pivotni rast $\rho_n^c = 13.0205$.

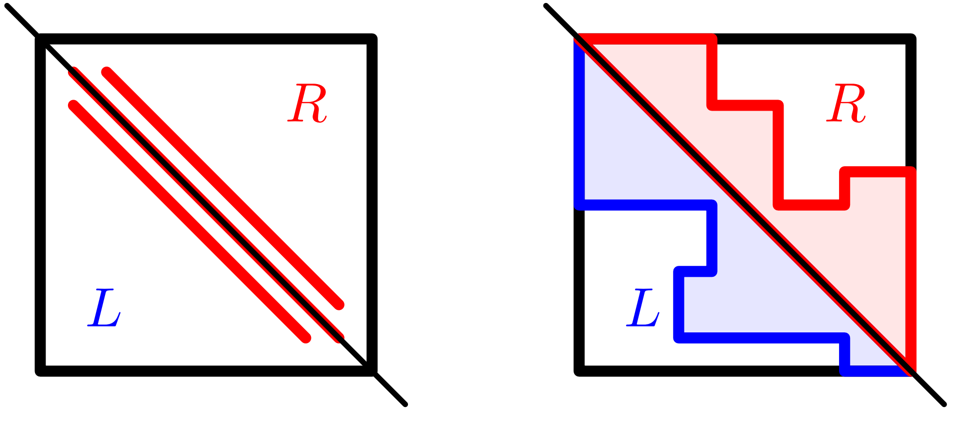
Struktura LR faktorizacije

Ako matrica A koja ulazi u LR faktorizaciju ima nekakvu strukturu, pitanje je kad će se ta struktura sačuvati.

To je posebno bitno za

- sustave gdje je A takva da se bitna informacija o njoj može spremiti u bitno manje od n^2 elemenata.

Ako ne pivotiramo, onda se čuvaju, recimo, sljedeće forme:



Kad ne moramo pivotirati?

Odgovor. Postoje tipovi matrica kad ne moramo pivotirati. Na primjer, to su:

- dijagonalno dominantne matrice po stupcima, tj. matrice za koje vrijedi
$$|a_{jj}| > \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}|,$$
- dijagonalno dominantne matrice po recima,
- simetrične pozitivno definitne matrice — o njima malo kasnije.

Simetrične pozitivno definitne matrice

Za simetrične/hermitske pozitivno definitne matrice radi se "simetrizirana varijanta" LR faktorizacije

- jer je 2 puta brža nego obična LR faktorizacija,
- čuva strukturu matrice A – čak i kad računamo u strojnoj aritmetici, množenjem faktora uvijek dobivamo simetričnu matricu.

"Simetrizirana LR" faktorizacija zove se faktorizacija Choleskog.

Prisjećanje. Matrica je hermitska ako je

$$A = A^*.$$

Ako je $A \in \mathbb{R}^{n\times n}$, onda je hermitska matrica isto što i simetrična, tj. $* = T$.

Simetrične pozitivno definitne matrice (nast.)

Pozitivna definitnost matrice se ne vidi odmah. Uobičajeno se unaprijed, iz prirode problema zna da je neka matrica pozitivno definitna.

Matrica $A \in \mathbb{F}^{n \times n}$ je pozitivno definitna ako je

$$x^* A x > 0 \quad \text{za svaki} \quad x \neq 0, \quad x \in \mathbb{F}^n.$$

Ekvivalentni uvjeti za pozitivnu definitnost:

- sve svojstvene vrijednosti od A su pozitivne, tj. vrijedi

  $$\lambda_k(A) > 0, \quad \text{za} \quad k = 1, \ldots, n,$$

  gdje λ_k označava k-tu najveću svojstvenu vrijednost;

Simetrične pozitivno definitne matrice (nast.)

Ekvivalentni uvjeti (nastavak):

- sve vodeće glavne minore od A su pozitivne, tj. vrijedi

$$\det(A_k) > 0, \quad \text{za} \quad k = 1, \ldots, n,$$

gdje je $A_k = A(1:k, 1:k)$ vodeća glavna podmatrica od A reda k.

Digresija. Katkad se puno lakše vidi da neka matrica nije pozitivno definitna. Pokažite da nisu pozitivno definitne matrice koje

- na dijagonali imaju negativan element ili nulu.

Faktorizacija Choleskog

Iz ekvivalentnog uvjeta odmah izlazi da za hermitsku/simetričnu pozitivno definitnu matricu uvijek može napraviti LR faktorizacija bez pivotiranja (v. Teorem)!

Tvrdnja. Za hermitsku/simetričnu pozitivno definitnu matricu A, LR faktorizaciju možemo napisati u obliku

$$A = LDL^*.$$

Ta faktorizacija se obično zove LDL^* faktorizacija.

U LR faktorizaciji matrice A faktor R rastavi se na

$$R = DM^*$$

D dijagonalna, M^* gornjetrokutasta s jedinicama na dijagonali.

Faktorizacija Choleskog (nastavak)

Da se dobije takva faktorizacija,

- dijagonalni elementi R stave se na dijagonalu od D,
- svaki redak u R podijeli se s dijagonalnim elementom u tom retku da se dobije M^*.

Dakle,

$$A = LDM^*, \quad M \ \text{donjetrokutasta, regularna.}$$

Zbog hermitičnosti/simetrije vrijedi

$$A = A^* = (LDM^*)^* = MDL^*,$$

pa je $LDM^* = MDL^* \ldots$

Faktorizacija Choleskog (nastavak)

(nastavak)...

$$LDM^* = MDL^*.$$

Množenjem slijeva s L^{-1} i zdesna s L^{-*} dobivamo

$$DM^*L^{-*} = L^{-1}MD.$$

Na lijevoj strani imamo produkt gornjetrokutastih matrica, a na desnoj strani donjetrokutastih, pa su ti produkti dijagonalne matrice.

Te dijagonalne matrice su jednake D (jer i M i L imaju na dijagonali jedinice), pa je

$$L^{-1}MD = D \implies MD = LD \implies M = L.$$

Faktorizacija Choleskog (nastavak)

Nadalje, D ima pozitivne elemente, jer bi inače postojao vektor x takav da je

$$(Ax, x) = (LDL^*x, x) = (DL^*x, L^*x) := (Dy, y) \leq 0.$$

Dakle, D možemo rastaviti na

$$D = \Delta \cdot \Delta$$

gdje je Δ dijagonalna i $\Delta_{ii} = \sqrt{D_{ii}}$.

Tada LDL^* faktorizaciju možemo napisati u obliku:

$$A = LDL^* = (L\Delta)(\Delta L^*) = (L\Delta)(\Delta^* L^*) = (L\Delta)(L\Delta)^*.$$

Faktorizacija Choleskog (nastavak)

Uz oznaku $R := (L\Delta)^*$ dobivamo faktorizaciju Choleskog

$$A = R^* R.$$

Digresija. Mnogi slovom L označavaju $L := L\Delta$, pa se u literaturi faktorizacija Choleskog može naći napisana kao

$$A = LL^*.$$

Oprez: taj L nema jedinice na dijagonali!

Kad znamo da postoji, Faktorizacija Choleskog se može i direktno izvesti, znajući da je $A = R^* R$.

Algoritam

Ograničimo se na realni slučaj. Tada je

$$a_{ij} = \sum_{k=1}^{i} r_{ki}\, r_{kj}, \quad i \le j,$$

pa dobivamo sljedeću rekurziju za elemente:

$$\text{za } j = 1, \ldots, n:$$

$$r_{ij} = \left(a_{ij} - \sum_{k=1}^{i-1} r_{ki}\, r_{kj} \right) \Big/ r_{ii}, \quad i = 1, \ldots, j-1,$$

$$r_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} r_{kj}^2 \right)^{1/2}.$$

Za $j = 1$ računamo samo r_{11}.
Greške zaokruživanja $\Rightarrow$ moguć negativan izraz pod korijenom.

Rješenje linearnog sustava

Kad imamo faktorizaciju Choleskog $A = R^T R$, onda se rješenje linearnog sustava $Ax = b$ svodi na dva rješavanja trokutastih sustava

$$R^T y = b, \quad Rx = y,$$

koje lako rješavamo:

- sustav $R^T y = b$ — supstitucijom unaprijed

$$y_1 = b_1 / r_{11}$$

$$y_i = \left(b_i - \sum_{j=1}^{i-1} r_{ji}\, y_j \right) \bigg/ r_{ii}, \quad i = 2, \ldots, n,$$

Rješenje linearnog sustava (nastavak)

- sustav $Rx = y$ — supstitucijom unatrag

$$x_n = y_n / r_{nn}$$

$$x_i = \left(y_i - \sum_{j=i+1}^{n} r_{ij}\, x_j \right) \Big/ r_{ii}, \quad i = n-1, \ldots, 1.$$

Za razliku od LR faktorizacije, ovdje u obje supstitucije imamo dijeljenja.
Zbog toga se često koristi LDL^T oblik faktorizacije:

- u algoritmu nema računanja drugih korijena;
- rješavaju se tri linearna sustava:

$$Lz = b, \quad Dy = z, \quad L^T x = y.$$

- L ima jediničnu dijagonalu, pa štedimo n dijeljenja.

Može li LDL^T za simetrične matrice?

Može li se LDL^T faktorizacija napraviti za bilo koju simetričnu matricu A (uz dozvolu da D ima i negativne elemente)?

To ne vrijedi! Kontraprimjer:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

Pomaže li simetrična permutacija redaka/stupaca? Ne!

Poopćenje na indefinitne matrice dobivamo tako da dozvolimo dijagonalne blokove reda 2 u matrici D.

Pivotiranje u faktorizaciji Choleskog

I kod faktorizacije Choleskog možemo koristiti pivotiranje.

- Da bismo očuvali simetriju radne matrice, pivotiranje mora biti "simetrično", tj. radimo istovremene zamjene redaka i stupaca u A
$$A \to P^T A P,$$
gdje je P matrica permutacije.
- "Simetrična zamjena" $\Rightarrow$ dijagonalni element zamjenjuje mjesto s dijagonalnim!
- Standardni izbor pivotnog elementa u k-tom koraku je
$$a_{rr}^{(k)} = \max_{k \leq i \leq n} a_{ii}^{(k)},$$
što odgovara potpunom pivotiranju u Gaussovim eliminacijama.

Pivotiranje u faktorizaciji Choleskog (nastavak)

Ovim postupkom dobivamo faktorizaciju Choleskog

$$P^T A P = R^T R,$$

a za elemente matrice R vrijedi

$$r_{kk}^2 \geq \sum_{i=k}^{j} r_{ij}^2, \quad j = k+1, \ldots, n, \quad k = 1, \ldots, n.$$

Posebno, to znači da R ima nerastuću dijagonalu

$$r_{11} \geq \cdots \geq r_{nn} > 0.$$