Numerička matematika
8. predavanje

Saša Singer

singer@math.hr

web.math.hr/~singer

PMF – Matematički odsjek, Zagreb

Sadržaj predavanja

- Metoda najmanjih kvadrata:
  - QR faktorizacija.
    - Givensove rotacije.
    - Householderovi reflektori.
  - QR faktorizacija i pivotiranje.
  - Rješenje matrične formulacije korištenjem QR faktorizacije.
  - Neprekidni problem najmanjih kvadrata.
    - Jednostavni primjer.
    - Složeniji primjer.

Informacije

Konzultacije (službeno):

- samo za NM: ponedjeljak u 13 sati (iza predavanja),
- petak, 12–14 sati, ili — po dogovoru.

Numerička matematika je u kolokvijskom razredu A2.

Ne zaboravite, "žive" su i domaće zadaće na adresi

`http://web.math.hr/nastava/unm/zadace.php`

ili, izravno

`http://degiorgi.math.hr/nm/`

Dodatni bodovi "čekaju na vas".

Informacije

Moja web stranica za Numeričku matematiku je

`http://web.math.hr/~singer/num_mat/`

Tamo su kompletna predavanja od prošle četiri godine, a stizat će i nova (kako nastaju).

Skraćena verzija skripte — 1. dio (prvih 7 tjedana):

`http://web.math.hr/~singer/num_mat/num_mat1.pdf`

Skraćena verzija skripte — 2. dio (drugih 6 tjedana):

`http://web.math.hr/~singer/num_mat/num_mat2.pdf`

Informacije — demonstratori

Kolegij "Numerička matematika" ima čak tri demonstratora:

- Mario Berljafa — termin: petak, 10–12, u Pr3, poželjna najava mailom,
- Anastasia Kruchinina — termin: ponedjeljak, 17–19, poželjna najava mailom,
- Melkior Ornik — termin: utorak, 16–18, sastanak pred oglasnom pločom, nužna najava mailom.

Demosi lijepo mole da im se najavite mailom koji dan ranije!

- Njihove mail adrese nađete na oglasnoj ploči,
- ili se javite meni.

QR faktorizacija

Definicija QR faktorizacije

Neka je zadana matrica G tipa (m, n) koja ima puni stupčani rang, tj. $\mathrm{rang}(G) = n \leq m$. Rastav matrice G tako da je

$$G = QR = Q \begin{bmatrix} R_0 \\ 0 \end{bmatrix},$$

gdje je

- Q ortogonalna matrica reda m, a
- R_0 gornja trokutasta matrica reda n, s pozitivnim dijagonalnim elementima,

zove se QR faktorizacija matrice G.

Definicija QR faktorizacije

Ako postoji, prethodna faktorizacija može se pisati i u jednostavnijoj (tzv. skraćenoj) formi.

- Prvih n stupaca matrice Q označimo s Q_0,
- a preostale stupce, koji su okomiti na Q_0, s $Q_0^\perp$.

Onda je

$$G = QR = [\ Q_0 \quad Q_0^\perp\ ] \begin{bmatrix} R_0 \\ 0 \end{bmatrix} = Q_0 R_0, \qquad Q_0^T Q_0 = I_n.$$

Ostaje samo pokazati da takva faktorizacija postoji.

Egzistencija i jedinstvenost QR faktorizacije

Teorem. Neka je $G \in \mathbb{R}^{m \times n}$, $m \geq n$, i neka je $\operatorname{rang}(G) = n$. Tada postoji jedinstvena faktorizacija oblika

$$G = Q_0 R_0,$$

pri čemu je Q_0 matrica tipa $m \times n$, s ortonormiranim stupcima, tj. vrijedi

$$Q_0^T Q_0 = I_n,$$

a R_0 je gornja trokutasta matrica s pozitivnim dijagonalnim elementima (dovoljno je fiksirati predznake na dijagonali).

Pravokutnu matricu Q_0 s ortonormiranim stupcima, također, skraćeno zovemo "ortogonalnom".

Egzistencija i jedinstvenost QR faktorizacije

Dokaz. Najjednostavniji je dokaz ovog teorema je korištenjem Gram-Schmidtove ortogonalizacije.

Ako stupce matrice

$$G = [g_1, g_2, \ldots, g_n]$$

ortogonaliziramo slijeva udesno, dobit ćemo ortonormalni niz vektora $q_1, \ldots, q_n$, koji razapinju isti potprostor kao i stupci od G.

Stavimo li

$$Q_0 = [q_1, q_2, \ldots, q_n],$$

dobili smo $m \times n$ ortogonalnu matricu.

Egzistencija i jedinstvenost QR faktorizacije

Gram–Schmidtov postupak ortogonalizacije računa i koeficijente

$$r_{ji} = q_j^T g_i$$

koji polazni stupac g_i izražavaju kao linearnu kombinaciju prvih i vektora q_j ortonormirane baze, tako da je

$$g_i = \sum_{j=1}^{i} r_{ji} q_j.$$

Koeficijenti r_{ji} su upravo elementi matrice R_0.

Iz Gram–Schmidtovog algoritma bit će jasno da se može uzeti $r_{ii} > 0$ (dovoljno je $r_{ii} \neq 0$ za fiksiranje predznaka). ■

Gram–Schmidtov postupak ortogonalizacije

U praksi se nikad ne koristi klasični Gram–Schmidtov postupak ortogonalizacije (skraćeno CGS), jer

- vektore ortogonalizira obzirom na prethodne originalne vektore g_j.
- Zbog toga je nestabilan kad su stupci od G skoro linearno zavisni.

Umjesto CGS-a, može se koristiti tzv. modificirani Gram–Schmidtov postupak (skraćeno MGS),

- koji ortogonalizira vektore obzirom na prethodno ortogonaliziranu bazu q_j, pa je mnogo stabilniji.
- No, i kod njega se može dogoditi da je izračunati Q_0 vrlo daleko od ortogonalnog, tj. $\|Q_0^T Q_0 - I\| \gg u$, kad je G vrlo loše uvjetovana.

Gram–Schmidtov algoritam

Klasični i modificirani Gram–Schmidtov algoritam:

```
za i = 1 do n radi {
  /* Nađi i-ti stupac od Q_0 i R_0 */
  q_i = g_i;
  za j = 1 do i - 1 radi {
    /* Oduzmi komponentu od q_j u smjeru g_i */
      /* kod CGS-a je */
    r_ji = q_j^T g_i;
    /* ILI */
      /* kod MGS-a je */
    r_ji = q_j^T q_i;
    q_i = q_i - r_ji q_j;
  };
```

Gram–Schmidtov algoritam (nastavak)

```
    r_ii = ‖q_i‖_2;
    ako je r_ii > 0 onda {
      q_i = q_i / r_ii;
    }
    inače {
      /* Matrica R_0 je singularna -- stani */
    };
  };
```

Napomena: $r_{ii} = 0$ je ekvivalentno s tim da je

- g_i linearna kombinacija prethodnih stupaca matrice G (linearna zavisnost, pad ranga).

Pokažite da su dvije formule za r_{ji} koje koriste CGS i MGS matematički ekvivalentne. Numerički, naravno, nisu (greške).

Gram–Schmidtov algoritam — komentari

Gram–Schmidtov algoritam daje skraćenu QR faktorizaciju $G = Q_0 R_0$. Za "punu" faktorizaciju, tj. kvadratni Q,

- fali nam ortogonalni komplement $Q_0^\perp$,

kojeg nemamo iz čega izračunati — "fale" stupci u G.

Čim je $\|q_i\|_2 \neq 0$, za dijagonalni element r_{ii} možemo uzeti bilo koji od dva predznaka

$$r_{ii} = \pm \|q_i\|_2.$$

Dakle, bilo kojim fiksiranjem predznaka na dijagonali od R_0,

- opet dobivamo jedinstvenu skraćenu QR faktorizaciju.

Drugi algoritmi

U praksi, kad želimo ortogonalan Q, koristimo

- ili Givensove rotacije,
- ili Householderove reflektore

kojima poništavamo odgovarajuće elemente u matrici G. To ponovno daje konstrukciju QR faktorizacije.

Bitna razlika među ta dva algoritma:

- Givensove rotacije poništavaju po jedan element u stupcu,
- Householderovi reflektori poništavaju sve osim jednog elementa u (skraćenom) stupcu.

Oba algoritma mogu dati skraćenu i punu QR faktorizaciju.

Givensove rotacije

Givensove rotacije

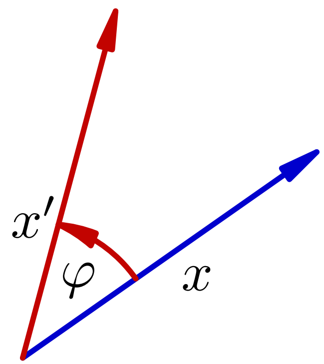
Matrica oblika

$$R(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

zove se Givensova rotacija u ravnini.

Ova transformacija rotira svaki vektor $x \in \mathbb{R}^2$ za kut φ, u smjeru obrnutom od kazaljke na satu.

Slika za $x' = R(\varphi)x$ je



Givensove rotacije u (i,j) ravnini

U $\mathbb{R}^m$, možemo definirati Givensovu rotaciju u (i,j) ravnini s

$$
R(i,j,\varphi) = \begin{array}{c c}
 & \begin{array}{ccccccccccc} & & & i & & & & j & & & \\ & & & \downarrow & & & & \downarrow & & & \end{array} \\
\begin{array}{c} \\ \\ \\ i \rightarrow \\ \\ \\ \\ j \rightarrow \\ \\ \\ \\ \end{array} &
\begin{bmatrix}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & \cos\varphi & & & & -\sin\varphi & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & \sin\varphi & & & & \cos\varphi & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{bmatrix}
\end{array}
$$

Poništavanje korištenjem Givensovih rotacija

Matrica $R(i,j,\varphi)$ je ortogonalna. Za zadani vektor $x \in \mathbb{R}^m$,

- poništavamo njegovu j-tu komponentu x_j, korištenjem rotacije $R(i,j,\varphi)$.

Množenjem matrice $R(i,j,\varphi)$ slijeva na x mijenjamo

- samo i-tu i j-tu komponentu u x,
- pa poništavanje možemo gledati samo u (i,j) ravnini.

Dobiveni sustav je

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x_i \\ x_j \end{bmatrix} = \begin{bmatrix} x_i' \\ 0 \end{bmatrix}.$$

Traže se elementi matrice rotacije $R(i,j,\varphi)$ i novi element x_i'.

Poništavanje korištenjem Givensovih rotacija

Drugi redak u matričnoj jednadžbi opisuje poništavanje

$$\sin\varphi\, x_i + \cos\varphi\, x_j = 0.$$

Ako je $x_j = 0$, nemamo što poništavati. Ako nije, dobivamo

$$\operatorname{ctg}\varphi = -\frac{x_i}{x_j}.$$

Odavde, korištenjem trigonometrijskog identiteta

$$1 + \operatorname{ctg}^2\varphi = \frac{1}{\sin^2\varphi},$$

slijedi

$$\sin^2\varphi = \frac{x_j^2}{x_i^2 + x_j^2}, \qquad \cos^2\varphi = 1 - \sin^2\varphi = \frac{x_i^2}{x_i^2 + x_j^2}.$$

Poništavanje korištenjem Givensovih rotacija

Predznake za $\sin\varphi$ i $\cos\varphi$ biramo tako da x_i' bude pozitivan. Ako stavimo

$$\sin\varphi = -\frac{x_j}{\sqrt{x_i^2 + x_j^2}}, \qquad \cos\varphi = \frac{x_i}{\sqrt{x_i^2 + x_j^2}},$$

iz prve jednadžbe dobivamo

$$x_i' = \frac{x_i}{\sqrt{x_i^2 + x_j^2}}\, x_i + \frac{x_j}{\sqrt{x_i^2 + x_j^2}}\, x_j = \frac{x_i^2 + x_j^2}{\sqrt{x_i^2 + x_j^2}}$$

$$= \sqrt{x_i^2 + x_j^2} > 0.$$

Element x_i' je norma i-te i j-te komponente polaznog vektora. Ove formule vrijede i kad je $x_j = 0$ (tada je $\varphi = 0$ ili $\varphi = \pi$).

Sustavno poništavanje

Sustavnim poništavanjem elemenata, konstruirat ćemo QR faktorizaciju matrice G.

- Postoji puno redosljeda kako napraviti nule u faktoru G.
- U sljedećem primjeru uzet je “standardni” redosljed redom, po stupcima, odozgo nadolje.

Poništavanje.

- Počinjemo s prvim stupcem i poništavamo redom elemente $g_{21}, \ldots, g_{m1}$.
- Ponovimo to isto za drugi, treći i svaki daljnji stupac od dijagonalnog mjesta nadolje.
- Time nećemo “pokvariti” već sređene nule u prethodnim stupcima.

Sustavno poništavanje - primjer

Primjer. Za jednu matricu G, tipa 4×3, to izgleda ovako.

1. stupac:

- U radnoj matrici G, redom poništavamo elemente

$$g_{i1}, \quad i = 2, \ldots, m = 4,$$

rotacijama $R(1, i, \varphi_{1i})$, koje "nabacuju" normu prvog stupca na prvi element u stupcu.

$$\begin{bmatrix} x & x & x \\ x & x & x \\ x & x & x \\ x & x & x \end{bmatrix} \to \begin{bmatrix} x & x & x \\ 0 & x & x \\ x & x & x \\ x & x & x \end{bmatrix} \to \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ x & x & x \end{bmatrix} \to \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \end{bmatrix}$$

Sustavno poništavanje - primjer (nastavak)

2. stupac:

- U radnoj matrici G, redom poništavamo elemente
$$g_{i2}, \quad i = 3, \ldots, m = 4,$$
rotacijama $R(2, i, \varphi_{2i})$, koje "nabacuju" normu drugog stupca (od dijagonale nadolje) na drugi element u stupcu.
- To neće "pokvariti" već sređene nule u prvom stupcu.
- Prvi redak se više ne mijenja.

$$\begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \end{bmatrix} \rightarrow \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & x & x \end{bmatrix} \rightarrow \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & x \end{bmatrix}$$

Sustavno poništavanje - primjer (kraj)

3. stupac:

- U radnoj matrici G, redom poništavamo elemente

$$g_{i3}, \quad i = 4, \ldots, m = 4,$$

rotacijama $R(3, i, \varphi_{3i})$, koje "nabacuju" normu trećeg stupca (od dijagonale nadolje) na treći element u stupcu.
- To neće "pokvariti" već sređene nule u prva dva stupca.
- Prva dva retka se više ne mijenjaju.

$$\begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & x \end{bmatrix} \to \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & 0 \end{bmatrix}$$

Poredak poništavanja i ocjena greške

Drugi rasporedi poništavanja.

- Za ocjenu greške zaokruživanja postoji i bolji raspored poništavanja elemenata.
- Gore opisanim algoritmom, pri “sređivanju” prvog stupca prvi redak se mijenja $m - 1$ puta, a svi ostali samo jednom.
- Poboljšanje dobivamo “ujednačavanjem”, tako da se svaki redak transformira podjednak broj puta.
- To se postiže korištenjem niza nezavisnih rotacija koje ne zahvaćaju iste retke.
- Takav raspored odvijanja rotacija, usput, još dozvoljava i paralelizaciju algoritma.

Poredak poništavanja i ocjena greške

Grafički, za jednu matricu tipa 4×3 to izgleda ovako.

Crveno i zeleno su nezavisne rotacije koje možemo istovremeno primjenjivati (samo su dvije, jer je m premalen).

$$\begin{bmatrix} x & x & x \\ x & x & x \\ x & x & x \\ x & x & x \end{bmatrix} \to \begin{bmatrix} x & x & x \\ 0 & x & x \\ x & x & x \\ 0 & x & x \end{bmatrix} \to \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & 0 & x \end{bmatrix}$$

$$\to \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & x \end{bmatrix} \to \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & 0 \end{bmatrix}$$

Parovi su: $(1, 2)$ i $(3, 4)$, $(1, 3)$ i $(2, 4)$, a na samom kraju više "ne ide".

Kako doći do Q?

Na kraju algoritma, na mjestu matrice G piše matrica R.

Do matrice Q se dolazi nakupljanjem primijenjenih rotacija

$$R(n, m, \varphi_{nm}) \cdots R(1, 2, \varphi_{12})\, G := Q^{-1} G = R.$$

Dakle, matricu G smo

- slijeva množili produktom ortogonalnih matrica, koji označimo s Q^{-1}.
- Zaključak: $G = QR$.
- Produkt ortogonalnih matrica je ortogonalna, pa je $Q^{-1} = Q^T$. Dakle, Q se lako računa iz Q^T.

Matrica Q^T dobiva se primjenom istih rotacija, samo na početnu matricu I — "što na G, to na I".

Puni i skraćeni Q

U punoj QR faktorizaciji, kvadratnu matricu Q^T, reda m, dobivamo akumulacijom rotacija slijeva, na početnu jediničnu matricu I_m, reda m,

$$Q^T = R(n, m, \varphi_{nm}) \cdots R(1, 2, \varphi_{12})\, I_m.$$

Pravokutnu matricu Q_0^T iz skraćene QR faktorizacije

- dobivamo trivijalno — radeći na prvih n stupaca!

Dovoljno je startati s matricom $I_m(1:n)$ koja sadrži prvih n stupaca jedinične matrice I_m

$$Q_0^T = R(n, m, \varphi_{nm}) \cdots R(1, 2, \varphi_{12})\, I_m(1:n).$$

A obratno? Napravimo punu QR faktorizaciju matrice Q_0^T!

Householderovi reflektori

Householderovi reflektori

Za zadani jedinični vektor $u \in \mathbb{R}^m$, matrica H definirana s

$$H = H(u) := I - 2uu^T, \quad \|u\|_2 = 1,$$

zove se Householderov reflektor.

Matrica H je

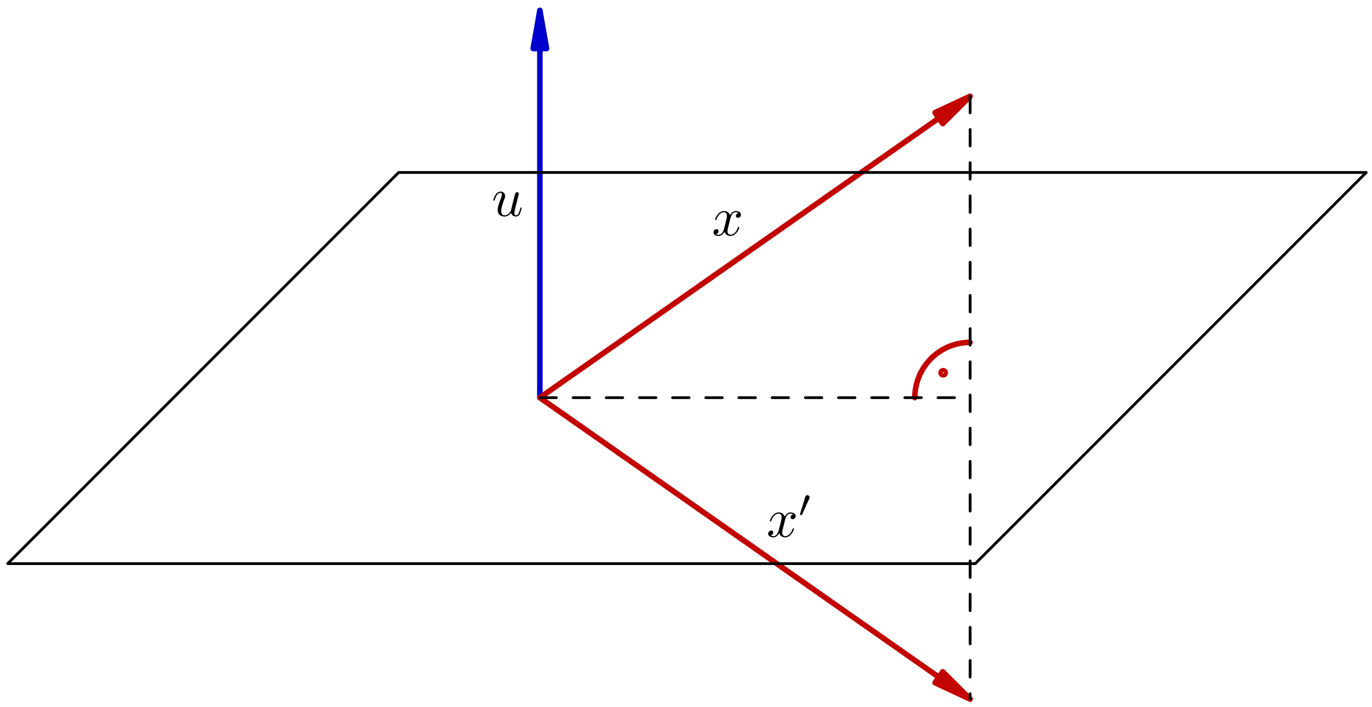
- simetrična,
- i ortogonalna, jer je

$$\begin{aligned} HH^T = H^2 &= (I - 2uu^T)\,(I - 2uu^T) \\ &= I - 4uu^T + 4u(u^T u)u^T \\ &= I - 4uu^T + 4\|u\|_2^2 uu^T = I. \end{aligned}$$

Zašto baš ime reflektor?

Promatrajmo hiperravninu koja je okomita na vektor u.

- Reflektor H sve vektore x preslikava u simetrične obzirom na tu hiperravninu.



Poništavanje Householderovim reflektorima

Ako je zadan vektor x, nađimo vektor u koji definira Householderov reflektor koji poništava sve (osim prve) komponente vektora x.

Tražimo da je

$$Hx = \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix} = c \cdot e_1.$$

Raspišimo tu jednadžbu

$$Hx = (I - 2uu^T)x = x - 2u(u^T x) = c \cdot e_1.$$

($u^T x$ je broj — skalarni produkt!)

Poništavanje Householderovim reflektorima

Premještanjem pribrojnika, dobivamo

$$u = \frac{1}{2(u^T x)} (x - c e_1),$$

pa je u linearna kombinacija od x i e_1. Preciznije, mora biti

$$u = \alpha (x - c e_1),$$

za neki broj α. Unitarne matrice čuvaju normu, pa je

$$\|x\|_2 = \|Hx\|_2 = |c|.$$

Nadalje, u mora biti jedinične norme, a paralelan s vektorom

$$\widetilde{u} = x \pm \|x\|_2 e_1.$$

Poništavanje Householderovim reflektorima

Prema tome,

$$u = \frac{\widetilde{u}}{\|\widetilde{u}\|_2}.$$

Oba izbora znakova u definiciji $\widetilde{u}$ zadovoljavaju

$$Hx = ce_1,$$

dok je $\widetilde{u} \neq 0$. U praksi se, zbog numeričke stabilnosti, koristi

$$\widetilde{u} = x + \operatorname{sign}(x_1) \|x\|_2 \, e_1,$$

jer to znači da nema kraćenja pri računanju prve komponente od $\widetilde{u}$, koja je jednaka

$$\widetilde{u}_1 = x_1 + \operatorname{sign}(x_1) \|x\|_2,$$

tj. oba su pribrojnika istog znaka.

Drugi način definicije H

Napomena. Računanje u se može izbjeći, ako definiramo

$$H(\widetilde{u}) = I - 2\,\frac{\widetilde{u}\,\widetilde{u}^T}{\widetilde{u}^T\widetilde{u}}.$$

Kako djelovati na ostale stupce?

Kad smo jednom izračunali u, ne treba računati cijelu matricu H. Dovoljno je pogledati kako ona djeluje na vektor z:

$$Hz = (I - 2uu^T)z = z - 2u(u^Tz).$$

Dakle, treba izračunati skalarni produkt u^Tz, a zatim modificirati vektor z.

QR faktorizacija korištenjem reflektora

QR faktorizacija provodi se sustavnom primjenom Householderovih reflektora na matricu G i to slijeva.

- Prvo se ponište svi elementi prvog stupca (do na prvi), a na ostale stupce se djeluje reflektorom H_1.
- Zatim se ponište elementi dijela drugog stupca od dijagonale nadolje (osim dijagonalnog elementa). To se radi "skraćenim" reflektorom H_2'.

  Ortogonalna matrica kojom smo djelovali na radnu matricu je onda

$$H_2 = \begin{bmatrix} 1 & \\ & H_2' \end{bmatrix}.$$

QR faktorizacija korištenjem reflektora

I tako redom. U k-tom koraku, za $k = 1, \ldots, n$, poništava se

- k-ti "skraćeni" stupac u radnoj matrici — od dijagonale nadolje. Ortogonalna matrica kojom djelujemo na radnu matricu ima oblik

$$H_k = \begin{bmatrix} I & \\ & H'_k \end{bmatrix},$$

s tim da je I reda $k-1$, a H'_k reda $m-k+1$.

Ako želimo formirati matricu Q, onda je

$$H_n H_{n-1} \cdots H_1 G = R, \quad \text{tj.} \quad Q^T = H_n H_{n-1} \cdots H_1,$$

ili

$$Q = (H_n H_{n-1} \cdots H_1)^T = H_1 \cdots H_{n-1} H_n.$$

QR faktorizacija i pivotiranje

Računanje QR faktorizacije

Neka je G zadana matrica tipa $m \times n$, s tim da je $m \geq n$.

Računanje QR faktorizacije matrice G

- provodimo u nizu od n koraka. Ako dozvolimo i $m < n$, broj koraka je $\min\{m, n\}$.

Na početku algoritma označimo $R^{(0)} := G$.

Opišimo kako izgleda k–ti korak algoritma, za $k = 1, \ldots, n$.

- Na početku k-tog koraka trenutna radna matrica je $R^{(k-1)}$.
- U njoj prvih $k - 1$ stupaca već ima gornjetrokutastu formu, tj. nule ispod dijagonale.
- Ti stupci se više neće mijenjati!

Računanje QR faktorizacije

Izgled radne matrice $R^{(k-1)}$ na početku k-tog koraka:

$$
\left[\begin{array}{cccc:ccc}
r_{1,1}^{(1)} & r_{1,2}^{(2)} & \cdots & r_{1,k-1}^{(k-1)} & r_{1,k}^{(k-1)} & \cdots & r_{1,n}^{(k-1)} \\
 & r_{2,2}^{(2)} & \cdots & r_{2,k-1}^{(k-1)} & r_{2,k}^{(k-1)} & \cdots & r_{2,n}^{(k-1)} \\
 & & \ddots & \vdots & \vdots & & \vdots \\
 & & & r_{k-1,k-1}^{(k-1)} & r_{k-1,k}^{(k-1)} & \cdots & r_{k-1,n}^{(k-1)} \\
 & & & & r_{k,k}^{(k-1)} & \cdots & r_{k,n}^{(k-1)} \\
 & & & & \vdots & & \vdots \\
 & & & & r_{m,k}^{(k-1)} & \cdots & r_{m,n}^{(k-1)}
\end{array}\right].
$$

Računanje QR faktorizacije

U k-tom koraku — u matrici $R^{(k-1)}$

- poništavamo sve elemente k-tog stupca ispod dijagonale, nekom ortogonalnom transformacijom Q_k.
- Tako dobivamo novu radnu matrcu $R^{(k)}$ koja ima jedan "sređeni" stupac više.

Ovu transformaciju možemo prikazati u obliku

$$R^{(k)} = Q_k R^{(k-1)}, \quad k = 1, \ldots, n.$$

Na kraju dobivamo gornju trokutastu matricu $R := R^{(n)}$.

Nije bitno kako računamo Q_k — rotacijama ili reflektorima!

Pivotiranje stupaca u QR faktorizaciji

Slično kao kod LR faktorizacije, i kod QR faktorizacije možemo koristiti pivotiranje.

- Uobičajeno se koristi pivotiranje stupaca

  $$GP = QR,$$

  gdje je P matrica permutacije.
- Ako su x_ℓ, $\ell = k, \ldots, n$, skraćeni stupci, na prvo mjesto dovodi se onaj s najvećom normom, tj. takav da je $\|x_k\|_2$ maksimalna.
- Postupak dovođenja na prvo mjesto ponavljamo u svakom koraku QR faktorizacije sa sve kraćim i kraćim stupcima.

Svrha pivotiranja

Svrha?

- Ako je matrica G bila takva da su joj stupci (skoro) linearno zavisni, onda se QR faktorizacijom s pivotiranjem određuje rang matrice G.

Teorem. Neka je $G \in \mathbb{R}^{m\times n}$ matrica ranga r. Tada postoje $n \times n$ matrica permutacije P, ortogonalna matrica Q reda m, te gornja trokutasta matrica R_0 ranga r, tipa $\min\{m,n\} \times n$, tako da vrijedi

$$GP = QR = Q \begin{bmatrix} R_0 \\ 0 \end{bmatrix}$$

i

$$\|(R_0)_{kk}\|^2 \geq \sum_{i=k}^{j} \|(R_0)_{ij}\|^2, \quad 1 \leq k \leq \min\{m,n\}, \quad k \leq j \leq n.$$

Rješenje matrične formulacije korištenjem QR faktorizacije

Korištenje QR faktorizacije

Već smo najavili da ćemo za rješenje diskretnog problema najmanjih kvadrata koristiti QR faktorizaciju.

Prisjetimo se, ako je A punog stupčanog ranga (tj. vrijedi $\mathrm{rang}(A) = m$), onda QR faktorizacija matrice A ima oblik

$$A = QR = \left[\begin{array}{cc} Q_0 & Q_0^{\perp} \end{array}\right] \begin{bmatrix} R_0 \\ 0 \end{bmatrix} = Q_0 R_0.$$

Za početak, ako minimiziramo $\|Ax - b\|_2^2$, minimizirali smo i $\|Ax - b\|_2$. Zbog unitarne invarijantnosti 2-norme, imamo

$$\min_x \|Ax - b\|_2^2 = \min_x \|Q^T(Ax - b)\|_2^2 = \min_x \|Q^T Ax - Q^T b\|_2^2.$$

Korištenje QR faktorizacije

Za Q uzmimo ortogonalnu matricu iz QR faktorizacije, pa je

$$
\begin{aligned}
\min_x \|Q^T(Ax-b)\|_2^2 &= \min_x \left\| \begin{bmatrix} R_0 \\ 0 \end{bmatrix} x - Q^T b \right\|_2^2 \\
&= \min_x \left\| \begin{bmatrix} R_0 \\ 0 \end{bmatrix} x - \begin{bmatrix} Q_0^T \\ (Q_0^\perp)^T \end{bmatrix} b \right\|_2^2 \\
&= \min_x \left\| \begin{bmatrix} R_0 x - Q_0^T b \\ 0 - (Q_0^\perp)^T b \end{bmatrix} \right\|_2^2 \\
&= \min_x \left(\|R_0 x - Q_0^T b\|_2^2 + \|(Q_0^\perp)^T b\|_2^2 \right).
\end{aligned}
$$

Korištenje QR faktorizacije

Primijetimo da samo prvi član u prethodnom minimumu ovisi o x, a drugi ne.

Budući da je R_0 kvadratna i punog ranga, onda je i regularna, pa postoji jedinstveno rješenje x linearnog sustava

$$R_0 x = Q_0^T b.$$

Time smo prvi član u kvadratu norme napravili najmanjim mogućim, jer je $\|R_0 x - Q_0^T b\|_2^2 = 0$.

Zaključak. Onda vrijedi

$$\min_x \|Ax - b\|_2 = \|(Q_0^\perp)^T b\|_2,$$

a postiže se za vektor x koji je rješenje sustava $R_0 x = Q_0^T b$.

Drugi način

Napomena. Postoji i lakši način da se dođe do prethodnog zaključka, ako znamo da su rješenja problema minimizacije

$$\min_x \|Ax - b\|_2$$

jednaka rješenju sustava normalnih jednadžbi

$$A^T Ax = A^T b.$$

Ako je $A^T A$ nesingularna, što je ekvivalentno tome da A ima puni stupčani rang, onda problem najmanjih kvadrata ima jedinstveno rješenje

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

Drugi način

Napravimo skraćenu QR faktorizaciju matrice A

$$A = Q_0 R_0.$$

Uvrštavanjem u rješenje x izlazi

$$\begin{aligned} x &= (A^T A)^{-1} A^T b = ((Q_0 R_0)^T Q_0 R_0)^{-1} (Q_0 R_0)^T b \\ &= (R_0^T Q_0^T Q_0 R_0)^{-1} R_0^T Q_0^T b = (R_0^T R_0)^{-1} R_0^T Q_0^T b \\ &= R_0^{-1} (R_0^T)^{-1} R_0^T Q_0^T b = R_0^{-1} Q_0^T b, \end{aligned}$$

pa je x, očito, rješenje sustava

$$R_0 x = Q_0^T b.$$

Primjer

Primjer. Diskretnom linearnom metodom najmanjih kvadrata nađite funkciju oblika

$$\varphi(x) = \frac{x+a}{bx+c}$$

koja aproksimira slijedeći skup podataka (točaka):

x_i	0	1	2	3	4
f_i	2.02	0.97	0.82	0.70	0.67

.

Nađite

- aproksimacije i pogreške u čvorovima x_i i
- sumu kvadrata apsolutnih grešaka S.

Primjer — linearizacija

Rješenje nađite

- korištenjem sustava normalnih jednadžbi i faktorizacije Choleskog,
- korištenjem QR faktorizacije,
- korištenjem QR faktorizacije s pivotiranjem stupaca.

Rješenje. Traženi oblik funkcije je nelinearan, pa ga treba linearizirati. To možemo napraviti na više načina.

1. Pomnožimo oblik funkcije φ s $bx + c$ i dobivamo

$$(bx + c)\varphi(x) = x + a,$$

odnosno

$$-a + b\,x\varphi(x) + c\,\varphi(x) = x.$$

Primjer — linearizacija

2. Ovu funkciju $-a + b\,x\varphi(x) + c\,\varphi(x) = x$ možemo podijeliti s $\varphi(x)$, pa dobivamo drugu linearizaciju

$$-a \cdot \frac{1}{\varphi(x)} + bx + c = \frac{x}{\varphi(x)}.$$

Primijetite da ove dvije linearizacije

- ne moraju (i neće) dati isto rješenje!

Obje pripadaju “grupi” linearizacija oblika

$$D + Eu + Fv = w,$$

pri čemu je $w = w(u, v)$.

Primjer — sustav normalnih jednadžbi

Prvo riješimo problem korištenjem sustava normalnih jednadžbi i faktorizacije Choleskog.

Za 1. slučaj metoda najmanjih kvadrata ima oblik

$$S = \sum_{i=0}^{n} \big(w_i - (-a + bu_i + cv_i)\big)^2 \to \min,$$

pri čemu su supstitucije za varijable

$$u = x\varphi(x), \quad v = \varphi(x), \quad w = x,$$

a za vrijednosti varijabli u čvorovima

$$u_i = x_i f_i, \quad v_i = f_i, \quad w_i = x_i.$$

Primjer — sustav normalnih jednadžbi

Deriviranjem po sve tri varijable izlazi

$$\frac{\partial S}{\partial a} = -2 \sum_{i=0}^{n} \big(w_i - (-a + bu_i + cv_i)\big)(-1) = 0$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=0}^{n} \big(w_i - (-a + bu_i + cv_i)\big)u_i = 0$$

$$\frac{\partial S}{\partial c} = -2 \sum_{i=0}^{n} \big(w_i - (-a + bu_i + cv_i)\big)v_i = 0.$$

Primjer — sustav normalnih jednadžbi

Odavde dobivamo simetrični, pozitivno definitni linearni sustav

$$\begin{bmatrix} (n+1) & -\sum_{i=0}^{n} u_i & -\sum_{i=0}^{n} v_i \\ -\sum_{i=0}^{n} u_i & \sum_{i=0}^{n} u_i^2 & \sum_{i=0}^{n} u_i v_i \\ -\sum_{i=0}^{n} v_i & \sum_{i=0}^{n} u_i v_i & \sum_{i=0}^{n} v_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -\sum_{i=0}^{n} w_i \\ \sum_{i=0}^{n} u_i w_i \\ \sum_{i=0}^{n} v_i w_i \end{bmatrix}.$$

Primjer — sustav normalnih jednadžbi

Kad uvrstimo zadane podatke, za 1. slučaj dobivamo linearni sustav $Mx = d$, gdje je

$$M = \begin{bmatrix} 5 & -7.39 & -5.18 \\ -7.39 & 15.2229 & 5.5513 \\ -5.18 & 5.5513 & 6.6326 \end{bmatrix}, \quad d = \begin{bmatrix} -10 \\ 21.27 \\ 7.39 \end{bmatrix}.$$

Faktorizacija Choleskog matrice M je $M = R^T R$, uz

$$R \approx \begin{bmatrix} 2.236067977 & -3.304908471 & -2.316566425 \\ & 2.073759870 & -1.014939111 \\ & & 0.485817456 \end{bmatrix}.$$

Primjer — sustav normalnih jednadžbi

Sad moramo još riješiti dva trokutasta sustava:

$$R^T y = d, \quad Rx = y.$$

Rješenja prvog, pa drugog sustava su

$$y \approx \begin{bmatrix} -4.472135955 \\ 3.129581246 \\ 0.424715923 \end{bmatrix}, \quad x \approx \begin{bmatrix} 1.768586298 \\ 1.936999050 \\ 0.874229442 \end{bmatrix}.$$

Prema tome, rješenje za parametre u 1. slučaju je

$$a = 1.768586298,$$

$$b = 1.936999050,$$

$$c = 0.874229442.$$

Primjer — sustav normalnih jednadžbi

Vrijednosti u čvorovima dobivamo tako da uvrstimo x_i u $\varphi(x)$

$$\varphi(x_i) = \frac{x_i + 1.768586298}{1.936999050x_i + 0.874229442},$$

pripadne greške su $f_i - \varphi(x_i)$, a zbroj kvadrata grešaka je

$$S = \sum_{i=0}^{4} (f_i - \varphi(x_i))^2.$$

Izračunajte sami!

Primjer — sustav normalnih jednadžbi

Za 2. slučaj treba uvesti supstitucije za varijable

$$u = -\frac{1}{\varphi(x)}, \quad v = x, \quad w = \frac{x}{\varphi(x)},$$

i vrijednosti varijabli u čvorovima

$$u_i = -\frac{1}{\varphi(f_i)}, \quad v_i = x_i, \quad w_i = \frac{x_i}{f_i}.$$

Metoda najmanjih kvadrata imat će oblik

$$S = \sum_{i=0}^{n} \big(w_i - (au_i + bv_i - c)\big)^2 \to \min .$$

Primjer — sustav normalnih jednadžbi

Pripadni linearni sustav glasi $Mx = d$, gdje je

$$M \approx \begin{bmatrix} 7.0636 & -13.7258 & -5.6666 \\ -13.7258 & 30 & 10 \\ -5.6666 & 10 & 5 \end{bmatrix}, \quad d \approx \begin{bmatrix} -19.0704 \\ 42.6467 \\ 13.7258 \end{bmatrix}.$$

Rješenje za parametre u 2. slučaju je

$$a = 1.752205717,$$

$$b = 1.938744602,$$

$$c = 0.853483129.$$

Ova rješenja se donekle razlikuju od prethodnih!

Primjer — QR faktorizacija

Riješimo sad 1. slučaj korištenjem QR faktorizacije. Uvrštavanjem točaka (x_i, f_i) u

$$-a + b\,x\varphi(x) + c\,\varphi(x) = x,$$

dobivamo

$$-a + bx_if_i + cf_i = x_i, \quad i = 0, \ldots, 4,$$

pa su A i b iz problema minimizacije jednaki

$$A = \begin{bmatrix} -1 & 0.00 & 2.02 \\ -1 & 0.97 & 0.97 \\ -1 & 1.64 & 0.82 \\ -1 & 2.10 & 0.70 \\ -1 & 2.68 & 0.67 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

Primjer — QR faktorizacija

Skraćena forma QR faktorizacije od A je $A = Q_0 R_0$, gdje je

$$Q_0 \approx \begin{bmatrix} -0.447213595 & -0.712715113 & 0.536492779 \\ -0.447213595 & -0.244965682 & -0.647620310 \\ -0.447213595 & 0.078118977 & -0.281410215 \\ -0.447213595 & 0.299938295 & -0.065005678 \\ -0.447213595 & 0.579623522 & 0.457543425 \end{bmatrix},$$

$$R_0 \approx \begin{bmatrix} 2.236067977 & -3.304908471 & -2.316566425 \\ & 2.073759870 & -1.014939111 \\ & & 0.485817456 \end{bmatrix}.$$

Uočite: $R_0 = R$ iz faktorizacije Choleskog za $M = A^T A$.

Primjer — QR faktorizacija

Desna strana linearnog sustava je $Q_0^T b$, gdje je

$$Q_0^T b \approx \begin{bmatrix} -4.472135955 \\ 3.129581246 \\ 0.424715923 \end{bmatrix}.$$

Rješenje trokutastog sustava $R_0 x = Q_0^T b$ je

$$x \approx \begin{bmatrix} 1.76858629811 \\ 1.93699905025 \\ 0.87422944194 \end{bmatrix}$$

Napravite isto sami za drugu linearizaciju.

Primjer — QR faktorizacija s pivotiranjem

Ako napravimo QR faktorizaciju s pivotiranjem stupaca, dobit ćemo $AP = Q_0 R_0$, gdje je poredak stupaca $p = [2, 3, 1]$,

$$Q_0 \approx \begin{bmatrix} 0.000000000 & 0.940989460 & 0.332153833 \\ 0.248612544 & 0.287082061 & -0.731570874 \\ 0.420334610 & 0.103390036 & -0.312927468 \\ 0.538233342 & -0.030653048 & -0.059615261 \\ 0.686888265 & -0.143155917 & 0.502991359 \end{bmatrix},$$

$$R_0 \approx \begin{bmatrix} 3.901653496 & 1.422807024 & -1.894068760 \\ & 2.146676541 & -1.157652593 \\ & & 0.268968410 \end{bmatrix}.$$

Primjer — QR faktorizacija s pivotiranjem

Dobiveno rješenje trokutastog sustava $R_0 x' = Q_0^T b$ je

$$x' \approx \begin{bmatrix} 1.93699905025 \\ 0.87422944194 \\ 1.76858629811 \end{bmatrix}.$$

Sad još treba vratiti x u "pravi poredak". Budući da je finalni pivotni vektor bio $p = [2, 3, 1]$, to odgovara matrici permutacije

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Primjer — QR faktorizacija s pivotiranjem

Pravo rješenje x dobit ćemo kao

$$x = P^T x' = \begin{bmatrix} 1.76858629811 \\ 1.93699905025 \\ 0.87422944194 \end{bmatrix}.$$

Veza QR faktorizacije i faktorizacije Choleskog

Neka je G pravokutna matrica tipa (m, n), koja ima puni stupčani rang, tj. $\text{rang}(G) = n \leq m$.

Matrica G ima jedinstvenu QR faktorizaciju (na pr. punu)

$$G = Q \begin{bmatrix} R_0 \\ 0 \end{bmatrix},$$

gdje je R_0 gornja trokutasta matrica reda n, s pozitivnim dijagonalnim elementima, a Q je unitarna matrica.

S druge strane, neka je $H := G^*G$ Gramova matrica skalarnih produkata stupaca matrice G.

Znamo da je onda H pozitivno definitna matrica. Zato H ima jedinstvenu faktorizaciju Choleskog ...

Veza QR faktorizacije i faktorizacije Choleskog

$$H = R^* R,$$

gdje je R gornja trokutasta s pozitivnom dijagonalom.

Tvrdnja. Ovi trokutasti faktori su jednaki, tj. vrijedi $R = R_0$.

Dokaz. U $H = G^* G$ uvrstimo QR faktorizaciju od G i "skratimo" $Q^* Q = I$. Jedinstvenost faktora R daje tvrdnju. ■

Ista veza (jednakost) vrijedi i za faktorizacije s pivotiranjem:

- pivotiranje stupaca po normi u QR faktorizaciji, i
- dijagonalno pivotiranje u faktorizaciji Choleskog.

Korist: ako znamo "faktor" G matrice H, i tražimo R,

- ne treba računati H, pa Choleskog, već samo QR od G.

Neprekidni problem najmanjih kvadrata

Još jednom o najmanjim kvadratima

U uvodu o aproksimaciji rečeno je da se parametri funkcije $\varphi \in \mathcal{F}$ po metodi najmanjih kvadrata, traže tako da bude

$$\min_{\varphi \in \mathcal{F}} \|e(x)\|_2,$$

pri čemu je $e(x) = f(x) - \varphi(x)$.

Da bismo mogli naći minimalnu grešku u neprekidnom slučaju, moramo definirati

- skalarni produkt za neprekidne funkcije na odgovarajućem intervalu.
- Definicija norme nije dovoljna, jer je rješenje već u diskretnom slučaju bila projekcija na potprostor, a za to nam je potreban skalarni produkt.

Definicija norme i skalarnog produkta

Neka je $w(x)$ zadana funkcija. $w(x)$ je težinska funkcija ako je

- $w(x) \geq 0$ na intervalu $[a, b]$,
- $w(x)$ može biti jednaka 0 samo u izoliranim točkama.

Težinska L_2-norma (ili samo 2-norma) funkcije u na $[a, b]$ je

$$\|u\|_2 = \left(\int_a^b w(x) \, |u(x)|^2 \, dx \right)^{1/2}.$$

Ako je ta norma konačna i za funkciju u i za funkciju v, onda možemo definirati težinski skalarni produkt

$$\langle u, v \rangle = \int_a^b w(x) \, u(x) \overline{v(x)} \, dx.$$

Definicija skalarnog produkta

Skalarni produkt $\langle u, v\rangle$ je dobro definiran (konačan), jer vrijedi Cauchy–Schwarzova nejednakost

$$|\langle u, v\rangle| \leq \|u\|_2 \cdot \|v\|_2.$$

$\langle u, v\rangle$ je skalarni produkt, jer

1. $\langle u, u\rangle \geq 0$, a jednak je 0 za one funkcije u koje su nula u svim točkama gdje je $w(x) > 0$, (v. mjera i integral)
2. vrijedi linearnost u prvom argumentu

$$\langle \alpha_1 u_1 + \alpha_2 u_2, v\rangle = \alpha_1 \langle u_1, v\rangle + \alpha_2 \langle u_2, v\rangle,$$

3. i antilinearnost/linearnost ($\mathbb{C}/\mathbb{R}$) u drugom argumentu,

$$\langle u, \beta_1 v_1 + \beta_2 v_2\rangle = \bar{\beta}_1 \langle u, v_1\rangle + \bar{\beta}_2 \langle u, v_2\rangle.$$

Ortogonalne funkcije

Napomena. Ako se radi o realnim funkcijama, onda kompleksno konjugiranje drugog argumenta izbacujemo.

U nastavku radimo samo s poljem $\mathbb{R}$, tj. s realnim funkcijama.

Za funkcije u i v reći ćemo da su ortogonalne ako vrijedi

$$\langle u, v \rangle = 0.$$

Ako su u i v ortogonalne, onda vrijedi

$$\begin{aligned} \|u + v\|^2 &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 = \|u\|^2 + \|v\|^2. \end{aligned}$$

Pitagorin poučak!

Sustavi ortogonalnih funkcija

Ako imamo sustav ortogonalnih funkcija u_k, $k = 0, \ldots, m$, za koje vrijedi

$$\langle u_i, u_j \rangle = 0, \quad \text{za} \quad i \neq j, \quad i, j = 0, \ldots, m,$$

i $u_k \not\equiv 0$ tamo gdje je $w(x) > 0$, onda vrijedi

$$\left\| \sum_{k=0}^{m} \alpha_k u_k \right\|_2^2 = \sum_{k=0}^{m} |\alpha_k|^2 \|u_k\|_2^2.$$

Prethodna jednakost znači da je ortogonalni sustav funkcija linearno nezavisan tamo gdje je $w(x) > 0$.

Ako je lijeva strana jednaka nula, mora biti i desna, a po pretpostavci je $\|u_k\|_2 > 0$, pa je jedino moguće da je $\alpha_k = 0$ za $k = 0, \ldots, m$.

Norma kvadrata greške

Ako je φ linearna funkcija, tj. $\varphi(x) = a_0\varphi_0(x) + \cdots + a_m\varphi_m(x)$, onda za normu kvadrata greške dobivamo

$$S := \|e\|_2^2 = \|f - \varphi\|_2^2 = \langle f, f\rangle - 2\langle f, \varphi\rangle + \langle \varphi, \varphi\rangle.$$

Ako uvrstimo oblik funkcije φ i definiciju skalarnog produkta, dobivamo

$$S = \int_a^b w(x) f^2(x)\, dx - 2\int_a^b w(x) f(x) \left(\sum_{j=0}^m a_j\varphi_j(x)\right) dx + \int_a^b w(x) \left(\sum_{j=0}^m a_j\varphi_j(x)\right)^2 dx.$$

Sustav normalnih jednadžbi

Kvadrat norme greške S je funkcija koeficijenata a_j.

- Radi o kvadratnoj funkciji u $m+1$ varijabli, pa je uvjet minimuma da su sve parcijalne derivacije jednake 0.

Dakle,

$$0 = \frac{\partial S}{\partial a_i} = 2 \int_a^b w(x) \left(\sum_{j=0}^m a_j \varphi_j(x) \right) \varphi_i(x) \, dx$$
$$- 2 \int_a^b w(x) f(x) \varphi_i(x) \, dx,$$

pa mora biti ...

Sustav normalnih jednadžbi

pa mora biti ...

$$\sum_{j=0}^{m} a_j \int\limits_a^b w(x)\varphi_i(x)\varphi_j(x)\,dx = \int\limits_a^b w(x)f(x)\varphi_i(x)\,dx,$$

za $i = 0, \ldots, m$. Uočimo da su odgovarajući integrali skalarni produkti, pa imamo

$$\sum_{j=0}^{m} \langle \varphi_i, \varphi_j \rangle a_j = \langle \varphi_i, f \rangle.$$

Ako označimo

$$m_{ij} = \langle \varphi_i, \varphi_j \rangle, \quad t_i = \langle \varphi_i, f \rangle, \quad i, j = 0, \ldots, m,$$

Sustav normalnih jednadžbi

pri čemu je

$$a^T = [a_0, a_1, \ldots, a_m],$$

onda problem najmanjih kvadrata možemo pisati kao sustav normalnih jednadžbi

$$Ma = t.$$

Matrica M je

- (očito) simetrična,
- ali i pozitivno definitna.

Pozitivna definitnost matrice M

Pozitivna definitnost izlazi iz definicije elemenata m_{ij}. Za svaki vektor $x \neq 0$ imamo

$$x^T M x = \sum_{i=0}^{m} \sum_{j=0}^{m} x_i x_j \langle \varphi_i, \varphi_j \rangle = \sum_{i=0}^{m} \sum_{j=0}^{m} \langle x_i \varphi_i, x_j \varphi_j \rangle$$
$$= \left\langle \sum_{i=0}^{m} x_i \varphi_i, \sum_{j=0}^{m} x_j \varphi_j \right\rangle = \left\| \sum_{i=0}^{m} x_i \varphi_i \right\|_2^2,$$

što je očito nenegativno. Nuli je jednako ako i samo ako je

$$\sum_{i=0}^{m} x_i \varphi_i \equiv 0, \quad \text{čim je } w(x) > 0.$$

Rješenje problema najmanjih kvadrata

Zaključak. Simetrične pozitivno definitne matrice su nesingularne, pa

- postoji jedinstveno rješenje problema $Ma = t$.

Nadalje, izračunati vektor a je jedinstveni minimum za problem najmanjih kvadrata, jer je

- Hesseova matrica drugih parcijalnih derivacija H pozitivno definitna, što slijedi iz

$$h_{ij} = \frac{\partial^2 S}{\partial a_i \partial a_j} = 2 \int_a^b w(x) \varphi_i(x) \varphi_j(x)\, dx = 2m_{ij},$$

tj. $H = 2M$!

Jednostavni primjer

Primjer. Neprekidnom metodom najmanjih kvadrata nađite polinom stupnja 1 koji aproksimira funkciju

$$f(x) = e^x$$

na intervalu $[-1, 1]$ uz težinsku funkciju $w(x) = 1$.

Rješenje. Treba minimizirati

$$S = \int_{-1}^{1} (e^x - a_1 x - a_0)^2 \, dx \to \min .$$

Jednostavni primjer

Deriviranjem dobivamo

$$\frac{\partial S}{\partial a_1} = -2 \int\limits_{-1}^{1} (e^x - a_1 x - a_0) x \, dx$$

$$\frac{\partial S}{\partial a_0} = -2 \int\limits_{-1}^{1} (e^x - a_1 x - a_0) \, dx,$$

pa uz oznake

$$s_k := \int\limits_{-1}^{1} x^k \, dx, \qquad t_k := \int\limits_{-1}^{1} e^x x^k \, dx,$$

treba riješiti sljedeći linearni sustav ...

Jednostavni primjer

treba riješiti sljedeći linearni sustav ...

$$
\begin{aligned}
s_2 a_1 + s_1 a_0 &= t_1 \\
s_1 a_1 + s_0 a_0 &= t_0,
\end{aligned}
$$

Izračunajmo integrale s lijeve strane

$$
s_k = \int_{-1}^{1} x^k \, dx = \frac{1}{k+1} x^{k+1} \Big|_{-1}^{1} = \frac{1 - (-1)^{k+1}}{k+1}
$$

$$
= \begin{cases} \dfrac{2}{k+1}, & \text{za } k \text{ paran}, \\ 0, & \text{za } k \text{ neparan}. \end{cases}
$$

Jednostavni primjer

Za integrale desne strane dobivamo

$$t_0 := \int_{-1}^{1} e^x \, dx = e^x \Big|_{-1}^{1} = e - e^{-1}$$

$$t_1 := \int_{-1}^{1} x e^x \, dx = \left\{ \begin{matrix} u = x & du = dx \\ dv = e^x & v = e^x \end{matrix} \right\} = x e^x \Big|_{-1}^{1} - \int_{-1}^{1} e^x \, dx$$

$$= e + e^{-1} - (e - e^{-1}) = 2e^{-1}.$$

Jednostavni primjer

Linearni sustav tada glasi:

$$\begin{aligned} \frac{2}{3} \cdot a_1 + 0 \cdot a_0 &= 2e^{-1} \\ 0 \cdot a_1 + 2 \cdot a_0 &= e - e^{-1}, \end{aligned}$$

a njegovo rješenje je

$$\begin{aligned} a_1 &= \frac{t_1}{s_2} = 3e^{-1} \approx 1.103638324, \\ a_0 &= \frac{t_0}{s_0} = \frac{e - e^{-1}}{2} = \mathrm{sh}(1) \approx 1.175201194. \end{aligned}$$

Pravac dobiven neprekidnom metodom najmanjih kvadrata je

$$p_1(x) \approx 1.103638324x + 1.175201194.$$

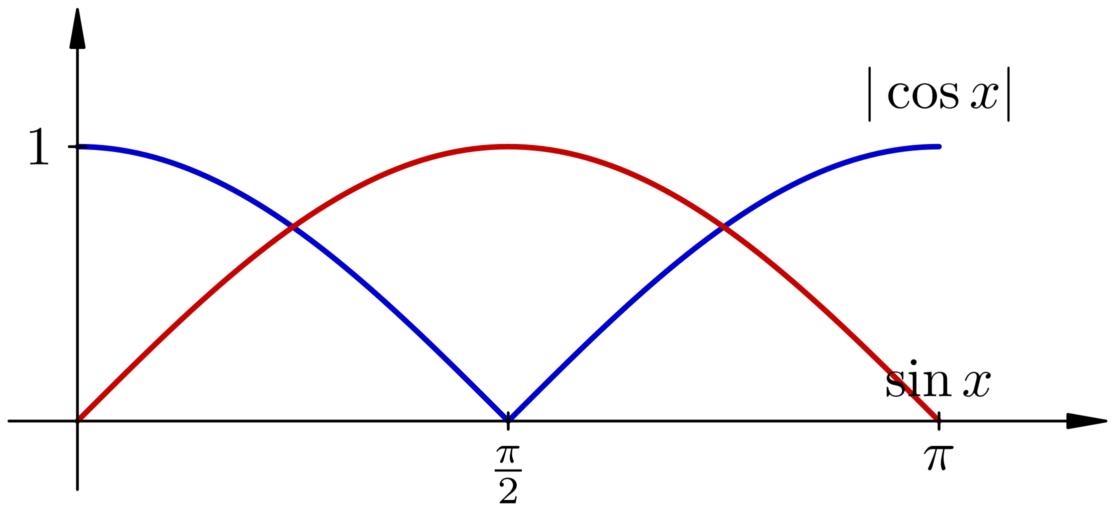
Složeniji primjer

Primjer. Neprekidnom metodom najmanjih kvadrata nađite polinome stupnjeva 1, 2 i 3 koji aproksimiraju funkciju

$$f(x) = \sin x$$

na intervalu $[0, \pi]$ uz težinsku funkciju $w(x) = |\cos x|$.

Rješenje. Skicirajmo prvo funkcije f i w.



Složeniji primjer

Budući da su obje funkcije simetrične oko točke $\frac{\pi}{2}$, polinome se isplati pisati u bazi $(x - \frac{\pi}{2})^k$.

Označimo s p_n polinom stupnja n,

$$p_n(x) = \sum_{k=0}^{n} a_{nk} \left(x - \frac{\pi}{2}\right)^k.$$

Treba minimizirati

$$S = \int_0^\pi |\cos x| \big(\sin x - p_n(x)\big)^2 \, dx \longrightarrow \min.$$

Složeniji primjer

Iz uvjeta

$$\frac{\partial S}{\partial a_{nk}} = 0, \quad k = 0, \ldots, n,$$

dobivamo linearni sustav

$$\sum_{j=0}^{n} a_{nj} \int_0^{\pi} |\cos x| \left(x - \frac{\pi}{2}\right)^{k+j} dx = \int_0^{\pi} |\cos x| \sin x \left(x - \frac{\pi}{2}\right)^{k} dx,$$

za $k = 0, \ldots, n$.

Složeniji primjer

Nađimo sad potrebne integrale:

$$\int_0^{\pi} \left(x - \frac{\pi}{2}\right)^k |\cos x| \, dx = \left\{ \begin{array}{ll} y = x - \dfrac{\pi}{2} & x = 0 \Rightarrow y = -\dfrac{\pi}{2} \\ dy = dx & x = \pi \Rightarrow y = \dfrac{\pi}{2} \end{array} \right\}$$

$$= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} y^k |\sin y| \, dy = \left\{ \begin{array}{ll} 0, & \text{za } k \text{ neparan}, \\ 2 \displaystyle\int_0^{\frac{\pi}{2}} y^k \sin y \, dy, & \text{za } k \text{ paran}. \end{array} \right\}$$

Napomena. Neparni koeficijenti su nula jer je baza pogodno odabrana, tako da koristi činjenicu da je $w(x)$ parna funkcija obzirom na $\frac{\pi}{2}$. Baza sadrži samo "parne" i "neparne" funkcije.

Složeniji primjer

Nađimo rekurziju za integral

$$s_k := 2\int\limits_0^{\frac{\pi}{2}} y^k \sin y\,dy = \left\{\begin{array}{ll} u = y^k & du = ky^{k-1}\,dy \\ dv = \sin y\,dy & v = -\cos y \end{array}\right\}$$

$$= 2\left(-y^k \cos y \,\Big|_0^{\frac{\pi}{2}} + k\int\limits_0^{\frac{\pi}{2}} y^{k-1} \cos y\,dy \right)$$

$$= \left\{\begin{array}{ll} u = y^{k-1} & du = (k-1)y^{k-2}\,dy \\ dv = \cos y\,dy & v = \sin y \end{array}\right\}$$

$$= 2ky^{k-1} \sin y \,\Big|_0^{\frac{\pi}{2}} - k(k-1)\left(2\int\limits_0^{\frac{\pi}{2}} y^{k-2} \sin y\,dy \right)$$

$$= 2k\left(\frac{\pi}{2}\right)^{k-1} - k(k-1)s_{k-2}.$$

Složeniji primjer

Još treba izračunati

$$s_0 = \int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} |\sin y|\, dy = 2\int\limits_0^{\frac{\pi}{2}} \sin y\, dy = -2\cos y\,\Big|_0^{\frac{\pi}{2}} = 2.$$

Sada je

$$\begin{aligned}
s_2 &= 4\frac{\pi}{2} - 4s_0 = 2\pi - 4, \\
s_4 &= 8\left(\frac{\pi}{2}\right)^3 - 12s_2 = \pi^3 - 24\pi + 48, \\
s_6 &= 12\left(\frac{\pi}{2}\right)^5 - 30s_4 = \frac{3}{8}\pi^5 - 24\pi^3 + 720\pi - 1440.
\end{aligned}$$

Složeniji primjer

Ostaje još izračunati integrale s desne strane:

$$t_k := 2 \int_0^{\pi} \left(x - \frac{\pi}{2} \right)^k |\cos x| \sin x \, dx$$

$$= \left\{ \begin{array}{ll} y = x - \dfrac{\pi}{2} & x = 0 \ \Rightarrow y = -\dfrac{\pi}{2} \\ dy = dx & x = \pi \ \Rightarrow y = \dfrac{\pi}{2} \end{array} \right\} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} y^k |\sin y| \cos y \, dy$$

$$= \left\{ \begin{array}{ll} 0, & \text{za } k \text{ neparan}, \\ 2 \displaystyle\int_0^{\frac{\pi}{2}} y^k \sin y \cos y \, dy = \int_0^{\frac{\pi}{2}} y^k \sin(2y) \, dy, & \text{za } k \text{ paran}. \end{array} \right\}$$

Složeniji primjer

Za parne indekse k s desne strane imamo

$$t_k := \int_0^{\frac{\pi}{2}} y^k \sin(2y)\, dy = \left\{ \begin{array}{ll} u = y^k & du = k y^{k-1}\, dy \\ dv = \sin(2y)\, dy & v = -\cos(2y)/2 \end{array} \right\}$$

$$= -\frac{1}{2}\, y^k \cos(2y) \,\Big|_0^{\frac{\pi}{2}} + \frac{k}{2} \int_0^{\frac{\pi}{2}} y^{k-1} \cos(2y)\, dy \Bigg)$$

$$= \left\{ \begin{array}{ll} u = y^{k-1} & du = (k-1) y^{k-2}\, dy \\ dv = \cos(2y)\, dy & v = \sin(2y)/2 \end{array} \right\}$$

$$= \frac{1}{2} \left(\frac{\pi}{2}\right)^k + \frac{k}{2} \cdot \frac{1}{2}\, y^{k-1} \sin(2y) \,\Big|_0^{\frac{\pi}{2}} - \frac{k-1}{2} \int_0^{\frac{\pi}{2}} y^{k-2} \sin(2y)\, dy$$

$$= \frac{1}{2} \left(\frac{\pi}{2}\right)^k - \frac{k(k-1)}{4} t_{k-2}.$$

Složeniji primjer

Još treba izračunati

$$t_0 = \int_0^{\frac{\pi}{2}} \sin(2y)\, dy = -\frac{1}{2} \cos(2y) \Big|_0^{\frac{\pi}{2}} = 1.$$

Sada je

$$t_2 = \frac{1}{2}\left(\frac{\pi}{2}\right)^2 - \frac{1}{2}\, t_0 = \frac{\pi^2}{8} - \frac{1}{2} = \frac{\pi^2 - 4}{8}.$$

Linearni sustav sada ima oblik

$$\sum_{j=0}^{n} a_{nj} s_{k+j} = t_k, \quad k = 0, \ldots, n.$$

Složeniji primjer

Za $n = 1$ sustav je:

$$\begin{aligned} s_0 a_{10} + s_1 a_{11} &= t_0 \\ s_1 a_{10} + s_2 a_{11} &= t_1, \end{aligned}$$

tj. ako uvrstimo izračunate s_k i t_k, imamo

$$\begin{aligned} 2 \cdot a_{10} + 0 \cdot a_{11} &= 1 \\ 0 \cdot a_{10} + (2\pi - 4) \cdot a_{11} &= 0. \end{aligned}$$

Rješenje tog sustava je $a_{10} = \dfrac{1}{2}$, $a_{11} = 0$, pa je aproksimacijski polinom

$$p_1(x) = \frac{1}{2}.$$

Složeniji primjer

Za $n = 2$ sustav je:

$$
\begin{aligned}
s_0 a_{20} + s_1 a_{21} + s_2 a_{22} &= t_0 \\
s_1 a_{20} + s_2 a_{21} + s_3 a_{22} &= t_1 \\
s_2 a_{20} + s_3 a_{21} + s_4 a_{22} &= t_2.
\end{aligned}
$$

Ako uvrstimo izračunate veličine, sustav glasi:

$$
\begin{aligned}
2 \cdot a_{20} + {} & \quad 0 \cdot a_{21} + {} & (2\pi - 4) \cdot a_{22} &= 1 \\
0 \cdot a_{20} + {} & (2\pi - 4) \cdot a_{21} + {} & 0 \cdot a_{22} &= 0 \\
(2\pi - 4) \cdot a_{20} + {} & \quad 0 \cdot a_{21} + {} & (\pi^3 - 24\pi + 48) \cdot a_{22} &= \frac{\pi^2 - 4}{8}.
\end{aligned}
$$

Složeniji primjer

Rješenje tog sustava je

$$a_{20} \approx 0.964909552, \quad a_{21} = 0, \quad a_{22} \approx -0.407246447,$$

pa je aproksimacijski polinom

$$p_2(x) \approx -0.407246447\left(x - \frac{\pi}{2}\right)^2 + 0.964909552.$$

Za $n = 3$ dobije se rješenje $p_3 = p_2$ (provjerite sami).

Još jedan primjer

Primjer. Ako funkciju $f(x)$ na $[0,1]$ uz $w(x) = 1$ aproksimiramo polinomom stupnja n po neprekidnoj metodi najmanjih kvadrata, matrica linearnog sustava je

$$M = \begin{bmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & \vdots & & \cdots \\ s_n & s_{n+1} & \cdots & s_{2n} \end{bmatrix},$$

pri čemu su

$$s_k := \int\limits_0^1 x^k \, dx = \frac{x^{k+1}}{k+1} \Big|_0^1 = \frac{1}{k+1}.$$

Matrica linearnog sustava je Hilbertova matrica reda $n+1$!

Komentari primjera

U posljednja dva primjera uočili smo sljedeće:

- Ako za bazu biramo funkcije $1, x, x^2, \ldots,$ matrica sustava može biti vrlo loše uvjetovana.
- U složenijem primjeru, podizanjem stupnja polinoma mijenjaju se koeficijenti polinoma p_n. Na primjer, a_0 ovisi o stupnju n.

Prethodna dva problema otklanjaju se ako se za bazu funkcija uzmu ortogonalne funkcije.