Numerička matematika

7. predavanje

Saša Singer i Nela Bosner

PMF - Matematički odsjek, Zagreb

Sadržaj predavanja

Metoda najmanjih kvadrata

Minimizacija 2-norme vektora pogreške

Neka je funkcija f

- zadana na diskretnom skupu točaka (čvorova) $x_0, \ldots, x_n$.

Uzmimo da točaka $x_0, \ldots, x_n$ ima mnogo više nego nepoznatih parametara $a_0, \ldots, a_m$ aproksimacijske funkcije φ, tj. $n \gg m$.

U diskretnoj metodi najmanjih kvadrata, aproksimacijska funkcija

$$\varphi(x) = \varphi(x; a_0, \ldots, a_m)$$

određuje se tako da 2-norma vektora pogrešaka u čvorovima aproksimacije bude najmanja moguća, tj. minimizira se

$$S = \sum_{k=0}^{n} \big(f(x_k) - \varphi(x_k; a_0, \ldots, a_m)\big)^2 \to \min .$$

Sustav normalnih jednadžbi

Ovdje se S gleda kao funkcija nepoznatih parametara

$$S = S(a_0, \ldots, a_m) : \mathbb{R}^{m+1} \to \mathbb{R}.$$

Uočimo da uvijek vrijedi $S \geq 0$, bez obzira na to kakvi su parametri, jer se radi o zbroju kvadrata.

- Funkcija S se minimizira kao funkcija više varijabli $a_0, \ldots, a_m$.
- Pretpostavljamo da je S dovoljno glatka funkcija, kao funkcija parametara a_k, pa je nužni uvjet ekstrema

$$\frac{\partial S}{\partial a_k} = 0, \quad k = 0, \ldots, m.$$

Ovaj pristup vodi na tzv. sustav normalnih jednadžbi.

Diskretna metoda najmanjih kvadrata za pravac

Primjer. Zadane su točke $(x_0, f_0), \ldots, (x_n, f_n)$, koje treba, po diskretnoj metodi najmanjih kvadrata, aproksimirati pravcem

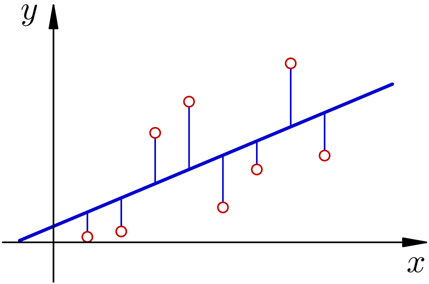
$$\varphi(x) = a_0 + a_1 x.$$

Suma kvadrata grešaka ove aproksimacije u čvorovima je (to je izraz kojeg minimiziramo)

$$\begin{aligned} S = S(a_0, a_1) &= \sum_{k=0}^{n} (f_k - \varphi(x_k))^2 \\ &= \sum_{k=0}^{n} (f_k - a_0 - a_1 x_k)^2 \to \min . \end{aligned}$$

Diskretna metoda najmanjih kvadrata za pravac

Slika zadanih točaka u ravnini i pravca koji ih aproksimira:



Uočiti da se greška u svakoj točki "mjeri" u smjeru osi y, pa je

$$S = S(a_0, a_1) = \sum_{k=0}^{n} (f_k - a_0 - a_1 x_k)^2 \to \min .$$

Može se gledati i "okomita" udaljenost do pravca $\to$ problem "potpunih" najmanjih kvadrata (engl. "total least squares").

Diskretna metoda najmanjih kvadrata za pravac

Parcijalne derivacije po parametrima a_0 i a_1 su:

$$0 = \frac{\partial S}{\partial a_0} = -2\sum_{k=0}^{n}(f_k - a_0 - a_1 x_k),$$

$$0 = \frac{\partial S}{\partial a_1} = -2\sum_{k=0}^{n}(f_k - a_0 - a_1 x_k)x_k.$$

Dijeljenjem s -2 i sređivanjem po nepoznanicama a_0, a_1, dobivamo linearni sustav

$$a_0(n+1) + a_1 \sum_{k=0}^{n} x_k = \sum_{k=0}^{n} f_k$$

$$a_0 \sum_{k=0}^{n} x_k + a_1 \sum_{k=0}^{n} x_k^2 = \sum_{k=0}^{n} f_k x_k.$$

Diskretna metoda najmanjih kvadrata za pravac

Uvedemo li standardne skraćene oznake

$$s_\ell = \sum_{k=0}^{n} x_k^\ell, \quad t_\ell = \sum_{k=0}^{n} f_k x_k^\ell, \quad \ell \geq 0,$$

onda linearni sustav možemo napisati kao

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

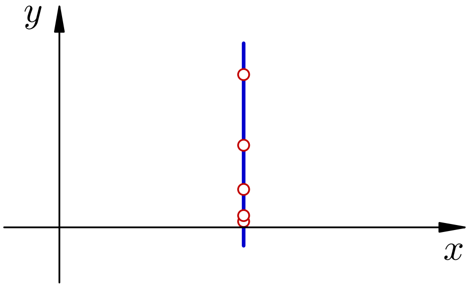
Matrica ovog sustava je (Gramova matrica)

$$M = \begin{bmatrix} n+1 & \sum_{k=0}^{n} x_k \\ \sum_{k=0}^{n} x_k & \sum_{k=0}^{n} x_k^2 \end{bmatrix} = X^T X, \qquad X = \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}.$$

Uz uvjet da imamo barem dvije različite točke x_k, stupci od X su linearno nezavisni, pa je M regularna.
U tom slučaju, postoji jedinstveno rješenje sustava.

Diskretna metoda najmanjih kvadrata za pravac

Slika situacije u kojoj problem najmanjih kvadrata za pravac nema rješenja, s tim da je $n \geq m = 1$, tj. imamo barem dvije točke podataka:



Ako imamo više različitih podataka u jednoj jedinoj točki x_0,

- aproksimacijski pravac (očito) postoji i jedinstven je,
- ali je okomit na x-os (jednadžba je $x = x_0$),
- pa njegova jednadžba nema oblik $y = a_0 + a_1 x$.

Minimalnost rješenja?

Je li dobiveno rješenje zaista minimum?

- To nije teško pokazati, korištenjem drugih parcijalnih derivacija — dovoljan uvjet minimuma je pozitivna definitnost Hesseove matrice (v. kasnije, za opći slučaj).

Provjera je li to minimum, može i puno lakše, jer se radi o zbroju kvadrata. Onda,

- S predstavlja paraboloid s otvorom prema gore, u varijablama a_0, a_1, tj. nad (a_0, a_1)-ravninom,
- pa je očito da takav paraboloid ima minimum.

Zbog toga se nikad ni ne provjerava je li dobiveno rješenje minimum za S.

Najmanji kvadrati za polinome

Za funkciju φ mogli bismo uzeti i polinom višeg stupnja,

$$\varphi(x) = a_0 + a_1 x + \cdots + a_m x^m.$$

Međutim, tu postoji opasnost — za malo veće m ($m \approx 10$)

- dobiveni sustav je (skoro sigurno) vrlo loše uvjetovan,
- pa dobiveni rezultati mogu biti jako pogrešni.

U praksi se to nikada direktno ne radi (na ovaj način), čim je $m \geq 2, 3$.

Ako se koriste aproksimacije polinomima viših stupnjeva,

- onda se to radi korištenjem tzv. ortogonalnih polinoma (vidjeti kasnije).

Najmanji kvadrati za opće linearne funkcije

Linearni model diskretnih najmanjih kvadrata je potpuno primjenjiv na opću linearnu funkciju

$$\varphi(x) = a_0\varphi_0(x) + \cdots + a_m\varphi_m(x),$$

gdje su $\varphi_0, \ldots, \varphi_m$ poznate (zadane) funkcije.

- Linearnost se ovdje odnosi na linearnu ovisnost aproksimacijske funkcije φ o parametrima $a_0, \ldots, a_m$.

Zadatak. Zadane su točke $(x_0, f_0), \ldots, (x_n, f_n)$, koje treba, po diskretnoj metodi najmanjih kvadrata, aproksimirati funkcijom oblika

$$\varphi(x) = a_0\varphi_0(x) + a_1\varphi_1(x).$$

Rješenje. Ide sasvim analogno (pogledati u skripti).

Što s nelinearnim funkcijama?

Što ako φ nelinearno ovisi o parametrima?

- Dobivamo nelinearni sustav jednadžbi, koji se relativno teško rješava.
- Problem postaje ozbiljan optimizacijski problem, koji se može približno rješavati.
- Metode koje se najčešće koriste su metode pretraživanja ili Levenberg–Marquardt metoda.

Postoji i drugi pristup.

- Katkad se jednostavnim transformacijama problem može transformirati u linearni problem najmanjih kvadrata.
- Rješenja lineariziranog i nelinearnog problema nisu jednaka, jer je i greška (nelinearno) transformirana!

Primjer nelinearnih najmanjih kvadrata

Primjer. Zadane su točke $(x_0, f_0), \ldots, (x_n, f_n)$, koje, po diskretnoj metodi najmanjih kvadrata, treba aproksimirati funkcijom oblika

$$\varphi(x) = a_0 e^{a_1 x}.$$

Uočite da φ nelinearno ovisi o parametru a_1.

Direktni pristup problemu vodi na minimizaciju

$$\begin{aligned} S = S(a_0, a_1) &= \sum_{k=0}^{n} (f_k - \varphi(x_k))^2 \\ &= \sum_{k=0}^{n} (f_k - a_0 e^{a_1 x_k})^2 \to \min . \end{aligned}$$

Primjer nelinearnih najmanjih kvadrata

Parcijalnim deriviranjem po varijablama a_0 i a_1 dobivamo

$$0 = \frac{\partial S}{\partial a_0} = -2 \sum_{k=0}^{n} (f_k - a_0 e^{a_1 x_k})\, e^{a_1 x_k},$$

$$0 = \frac{\partial S}{\partial a_1} = -2 \sum_{k=0}^{n} (f_k - a_0 e^{a_1 x_k}) a_0\, x_k e^{a_1 x_k}.$$

To je nelinearan sustav jednadžbi, kojeg ne znamo riješiti!

S druge strane, ako logaritmiramo relaciju

$$\varphi(x) = a_0 e^{a_1 x},$$

dobivamo

$$\ln \varphi(x) = \ln(a_0) + a_1 x.$$

Primjer nelinearnih najmanjih kvadrata

Moramo logaritmirati još i zadane vrijednosti funkcije f u točkama x_k. Uz supstitucije

$$h(x) = \ln f(x), \quad h_k = h(x_k) = \ln f_k, \quad k = 0, \dots, n,$$

i

$$\psi(x) = \ln \varphi(x) = b_0 + b_1 x,$$

gdje je

$$b_0 = \ln a_0, \quad b_1 = a_1,$$

dobivamo linearni problem najmanjih kvadrata

$$\begin{aligned} \widetilde{S} = \widetilde{S}(b_0, b_1) &= \sum_{k=0}^{n} (h_k - \psi(x_k))^2 \\ &= \sum_{k=0}^{n} (h_k - b_0 - b_1 x_k)^2 \to \min . \end{aligned}$$

Primjer nelinearnih najmanjih kvadrata

Grafički, to je pravac u tzv. lin–log skali. Iz rješenja b_0 i b_1, lako izlaze a_0 i a_1

$$a_0 = e^{b_0}, \quad a_1 = b_1.$$

Napomene uz linearizaciju:

- Pri linearizaciji smo pretpostavili da je $f_k > 0$, da bismo mogli logaritmirati.
- Ovako dobiveno rješenje uvijek daje pozitivan a_0, tj. linearizirani $\varphi(x)$ će uvijek biti veći od 0.
- Kad su neki $f_k \leq 0$, korištenjem translacije svih podataka treba dobiti $f_k + \text{translacija} > 0$, pa onda linearizirati.
- Pokušajte korektno formulirati takvu linearizaciju!
- Ako su svi $f_k < 0$, onda napravimo supstituciju $f \mapsto -f$.

Tipične linearizacije — opća potencija

Kratki pregled nekih funkcija koje se često koriste i njihovih standardnih linearizacija u problemu najmanjih kvadrata.

(a) Funkcija

$$\varphi(x) = a_0 x^{a_1}$$

linearizira se logaritmiranjem (u nekoj izabranoj bazi)

$$\begin{aligned} \psi(x) &= \log \varphi(x) = \log(a_0) + a_1 \log x, \\ h_k &= \log f_k, \qquad k = 0, \ldots, n. \end{aligned}$$

Uvedimo oznake

$$b_0 = \log(a_0), \quad b_1 = a_1.$$

Tipične linearizacije — opća potencija

Linearizirani problem najmanjih kvadrata glasi

$$\widetilde{S} = \widetilde{S}(b_0, b_1) = \sum_{k=0}^{n} (h_k - b_0 - b_1 \log(x_k))^2 \to \min,$$

Grafički, to odgovara pravcu u tzv. log–log skali.

U ovom slučaju, da bismo mogli provesti linearizaciju,

- mora biti i $x_k > 0$ i $f_k > 0$.

Tipične linearizacije — 1 / linearna funkcija

(b) Funkcija

$$\varphi(x) = \frac{1}{a_0 + a_1 x}$$

linearizira se na sljedeći način

$$\psi(x) = \frac{1}{\varphi(x)} = a_0 + a_1 x,$$

$$h_k = \frac{1}{f_k}, \qquad k = 0, \ldots, n.$$

Pripadni linearni problem najmanjih kvadrata je

$$\widetilde{S} = \widetilde{S}(a_0, a_1) = \sum_{k=0}^{n} (h_k - a_0 - a_1 x_k)^2 \to \min .$$

Tipične linearizacije — x / linearna funkcija

(c) Funkciju

$$\varphi(x) = \frac{x}{a_0 + a_1 x}$$

možemo linearizirati na više načina.

1. način:

$$\psi(x) = \frac{1}{\varphi(x)} = a_0 \, \frac{1}{x} + a_1,$$

$$h_k = \frac{1}{f_k}, \qquad k = 0, \ldots, n.$$

Pripadni linearni problem najmanjih kvadrata je

$$\widetilde{S} = \widetilde{S}(a_0, a_1) = \sum_{k=0}^{n} (h_k - a_0 \, \frac{1}{x_k} - a_1)^2 \to \min .$$

Tipične linearizacije — x / linearna funkcija

2. način:

$$\psi(x) = \frac{x}{\varphi(x)} = a_0 + a_1 x,$$

$$h_k = \frac{x_k}{f_k}, \qquad k = 0, \ldots, n.$$

Pripadni linearni problem najmanjih kvadrata je

$$\widetilde{S} = \widetilde{S}(a_0, a_1) = \sum_{k=0}^{n} (h_k - a_0 - a_1 x_k)^2 \to \min .$$

Tipične linearizacije — još jedan primjer

(d) Funkcija

$$\varphi(x) = \frac{1}{a_0 + a_1 e^{-x}}$$

linearizira se stavljanjem

$$\psi(x) = \frac{1}{\varphi(x)} = a_0 + a_1 e^{-x},$$

$$h_k = \frac{1}{f_k}, \qquad k = 0, \ldots, n.$$

Pripadni linearni problem najmanjih kvadrata je

$$\widetilde{S} = \widetilde{S}(a_0, a_1) = \sum_{k=0}^{n} (h_k - a_0 - a_1 e^{-x_k})^2 \to \min .$$

Primjer

Primjer. Uvaženi znanstvenik dr. Zurić, Ulica astronoma 69, dobio je ideju da se Zemlja giba oko Sunca po eliptičnoj orbiti, sa Suncem u jednom fokusu.

Nakon niza opažanja i mjerenja (uz dosta računa), dobio je slijedeće podatke

x [$^\circ$]	0	45	90	135	180
r [10^6 km]	147	148	150	151	152

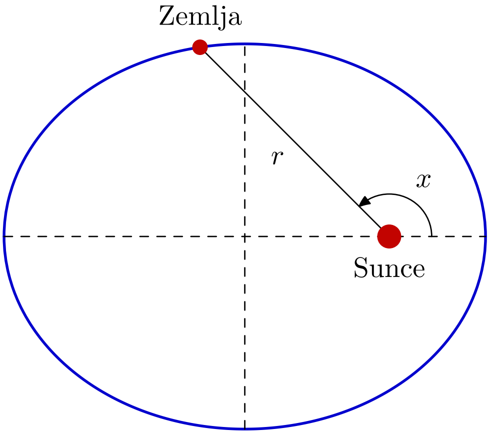
,

u kojima je

- r udaljenost od Zemlje do Sunca (u 10^6 km),
- a x je kut između spojnice Zemlja–Sunce i glavne osi elipse (u stupnjevima). [$\text{radijani} = \pi \cdot \text{stupnjevi}/180$]!

Zemlja na elipsi oko Sunca

Skica problema:



Podsjetnik: Ako elipsa ima veliku os a i malu os b, onda je $e = \sqrt{a^2 - b^2}$ i $\varepsilon = e/a$.

Primjer (nastavak)

Dr. Zurić, naravno, zna da se elipsa može opisati jednadžbom

$$r(x) = \frac{\rho}{1 + \varepsilon \cos x},$$

gdje je ε numerički ekscentricitet elipse, a $\rho = b^2/a$ je tzv. "srednja" udaljenost elipse od fokusa.

Pomognite mu da nađe ρ i ε, diskretnom linearnom metodom najmanjih kvadrata, nakon preuređenja ove jednadžbe.

Rješenje. Pomnožimo jednadžbu s nazivnikom funkcije, pa dobivamo

$$r(1 + \varepsilon \cos x) = \rho,$$

odnosno,

$$-\varepsilon\, r \cos x + \rho = r.$$

Primjer (nastavak)

Relaciju $-\varepsilon\, r \cos x + \rho = r$ gledamo kao funkciju oblika

$$au + b = v,$$

gdje je

$$u = r \cos x, \quad v = r, \quad a = -\varepsilon, \quad b = \rho.$$

Zatim se primijeni linearna metoda najmanjih kvadrata za pravac, s nepoznatim koeficijentima a i b.

Prema tome, treba minimizirati

$$S = \sum_{i=0}^{4} (v_i - au_i - b)^2 \to \min .$$

Primjer (nastavak)

Deriviranjem izlazi

$$\frac{\partial S}{\partial a} = -2\sum_{i=0}^{4}(v_i - au_i - b)u_i = 0$$

$$\frac{\partial S}{\partial b} = -2\sum_{i=0}^{4}(v_i - au_i - b) = 0.$$

Nakon sređivanja, uz $n + 1 = 5$, dobivamo linearni sustav

$$\begin{bmatrix} \sum_{i=0}^{4} u_i^2 & \sum_{i=0}^{4} u_i \\ \sum_{i=0}^{4} u_i & 5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{4} u_i v_i \\ \sum_{i=0}^{4} v_i \end{bmatrix}.$$

Primjer (nastavak)

Kad se radi “na ruke”, traženi podaci se obično slože u tablicu

i	$v_i = r_i$	$u_i = r_i \cos x_i$	u_i^2	$u_i v_i$
0	147	147	21609	21609
1	148	104.6518036	10952	15488.46693
2	150	0	0	0
3	151	−106.7731240	11400.5	−16122.74172
4	152	−152	23104	−23104
$\sum$	748	−7.1213204	67065.5	−2129.27479

Brojevi na dnu tablice su poznati elementi linearnog sustava.

Primjer (nastavak)

Linearni sustav za nepoznate koeficijente je

$$\begin{bmatrix} 67065.5 & -7.1213204 \\ -7.1213204 & 5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -2129.27479 \\ 748 \end{bmatrix}.$$

Rješenje tog sustava je

$$a = -1.58663722 \cdot 10^{-2}, \quad b = 149.5774021,$$

pa je

$$\varepsilon = 1.58663722 \cdot 10^{-2}, \quad \rho = 149.5774021.$$

Ove vrijednosti su vrlo blizu pravih, kad se sjetimo značenja:

- ε je ekscentricitet Zemljine orbite,
- ρ je "srednja" udaljenost od Zemlje do Sunca.

Izbor aproksimacijske funkcije

Kad smo jednom odabrali oblik aproksimacijske funkcije pitamo se — jesmo li dobro izabrali njezin oblik?

Kad nađemo aproksimaciju, moramo pogledati graf pogreške.

- Ako on "jednoliko" oscilira oko nule, i
- te oscilacije izgledaju kao "slučajne", a ne "sistematske",

onda je aproksimacijska funkcija dobro odabrana.

Bitno: Metoda najmanjih kvadrata smanjuje slučajne greške (recimo, kod mjerenja). To joj je osnovna svrha u statistici!

- Postoji tzv. Gauss–Markovljev teorem koji opravdava metodu najmanjih kvadrata.

Primjer uklanjanja slučajne greške

Primjer. Eksperimentalni podaci uzeti su tako da se egzaktne y-koordinate točaka na pravcu

$$y(x) = 4x + 3$$

za $x = 0, 1, \ldots, 100$, perturbiraju za

- uniformno distribuirani slučajni broj, između -1 i 1.

Tako se dobiju početni podaci (x_i, f_i), gdje je $x_i = i$, a

$$f_i = 4x_i + 3 + (\text{slučajna perturbacija između } -1 \text{ i } 1),$$

za $i = 0, \ldots, 100$.

Primjer uklanjanja slučajne greške

Prvih nekoliko podataka izgleda ovako:

x_i	$y(x_i)$	f_i
0	3	3.481757957246973
1	7	7.905987449877890
2	11	11.93107009769001 5
3	15	15.49513187608454 9
4	19	18.68144135301999 8
5	23	22.98482020710819 4

Kad se metodom najmanjih kvadrata za pravac $\varphi(x) = ax + b$ izračunaju parametri, dobijemo

$$a = 3.99598, \quad b = 3.20791.$$

Primjer uklanjanja slučajne greške

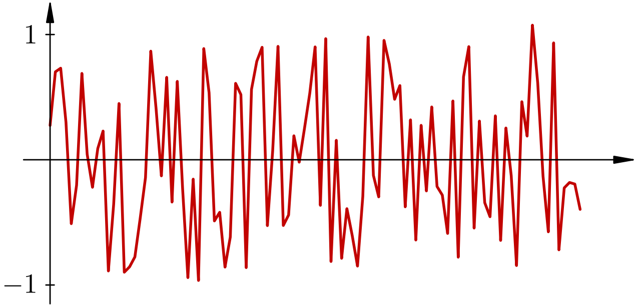
Pogledajmo što su aproksimacije $\varphi(x_i)$ za vrijednosti f_i u prvih nekoliko podataka:

x_i	$y(x_i)$	$\varphi(x_i)$
0	3	3.207905163100534
1	7	7.203881519200112
2	11	11.199857875299690
3	15	15.195834231399269
4	19	19.191810587498847
5	23	23.187786943598425

Uočite da su greške $\varphi(x_i)$ obzirom na $y(x_i)$ znatno manje nego polazne greške f_i obzirom na $y(x_i)$.

Primjer uklanjanja slučajne greške

Pogledajmo kako se ponaša greška $f_i - \varphi(x_i)$ u svim točkama.



Greška izgleda skoro kao slučajna varijabla koja oscilira oko nule, što znači da smo dobro odabrali aproksimaciju, i smanjili slučajnu grešku u podacima.

Krivac za "skoro slučajni" = ulazni podaci imaju sistematsku grešku na početku (uglavnom su > 0) i zato je b prevelik!

Demo primjeri

Diskretnom metodom najmanjih kvadrata aproksimiraju se

- izmjereni podaci za viskoznost 40% etilnog alkohola, u ovisnosti o temperaturi.

Primjer pokazuje način izbora aproksimacijske funkcije i

- različita rješenja, ako problem lineariziramo, ili ako ga ne lineariziramo
  (rješenja nelinearnih problema $=$ Nelder–Mead metoda).

Cijeli demo je na sljedećim stranicama.

Demo primjer — viskoznost “votke”

Promatramo kako ovisi

- viskoznost 40% etilnog alkohola o temperaturi.

Treba naći:

- “zgodan” oblik aproksimacijske funkcije i
- parametre za dobru aproksimaciju.

Za aproksimaciju koristimo metodu najmanjih kvadrata.

Ovaj primjer pokazuje:

- način izbora aproksimacijske funkcije,
- i različita rješenja — koja dobivamo kad problem lineariziramo, odnosno, kad ga ne lineariziramo.

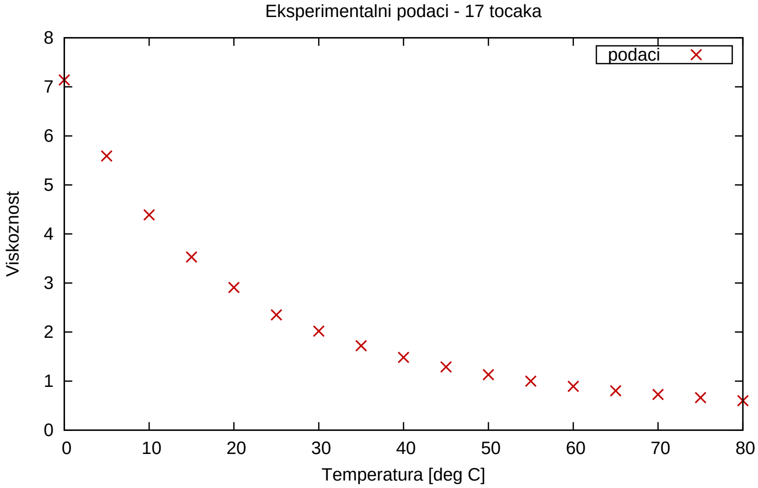
Eksperimentalno izmjereni podaci — tablica

Eksperimentalno su izmjereni sljedeći podaci (17 točaka):

Temperatura	Viskoznost
0	7.14
5	5.59
10	4.39
15	3.53
20	2.91
25	2.35
30	2.02
35	1.72
40	1.482

Temperatura	Viskoznost
45	1.289
50	1.132
55	0.998
60	0.893
65	0.802
70	0.727
75	0.663
80	0.601

Eksperimentalno izmjereni podaci — slika



Prva aproksimacija — oblik i parametri

Izmjereni podaci imaju

- “eksponencijalno” padajući oblik, a ne polinomni!

Nelinearni model:

$$f(x) = e^{ax+b}$$

ima parametre

$$a = -3.848637 \cdot 10^{-2}, \quad b = 1.911946\,.$$

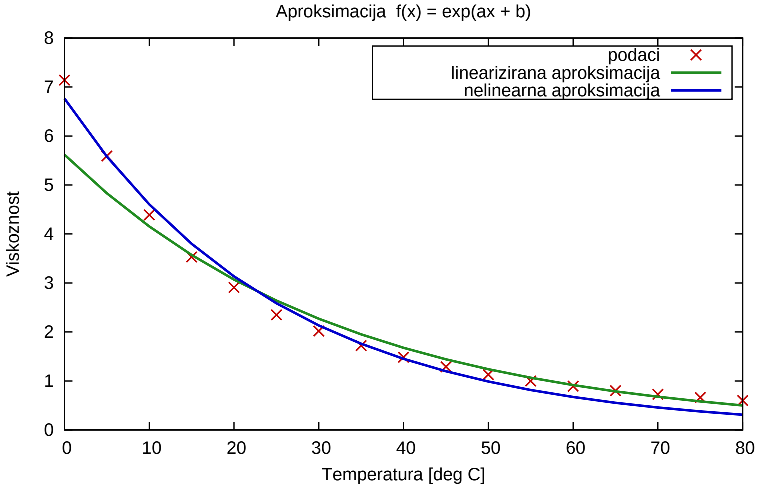
Linearizirani model:

$$\ln f(x) = ax + b$$

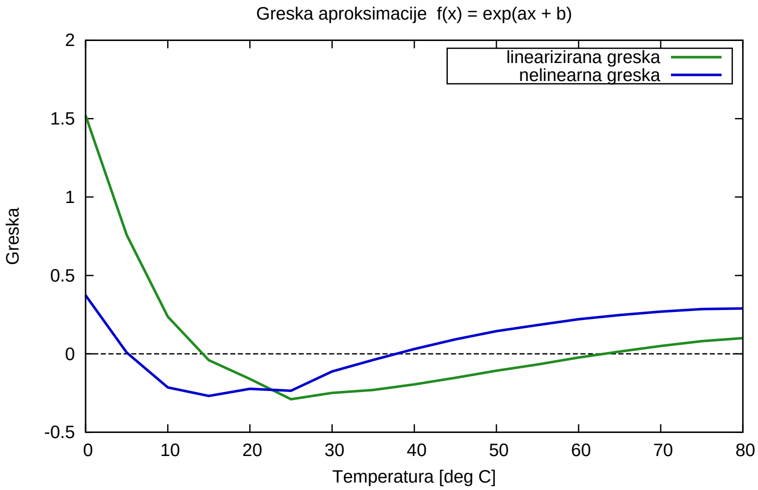
ima parametre

$$a = -3.022676 \cdot 10^{-2}, \quad b = 1.726233\,.$$

Prva aproksimacija



Greške prve aproksimacije



Druga aproksimacija — oblik i parametri

Dobivene greške nisu “slučajne”, pa nismo baš pogodili oblik.

- Ponašanje upućuje na “popravak” kvadratnim članom.

Nelinearni model:

$$f(x) = e^{ax^2+bx+c}$$

ima parametre

$$a = 2.487568 \cdot 10^{-4}, \quad b = -4.977411 \cdot 10^{-2}, \quad c = 1.962208\,.$$

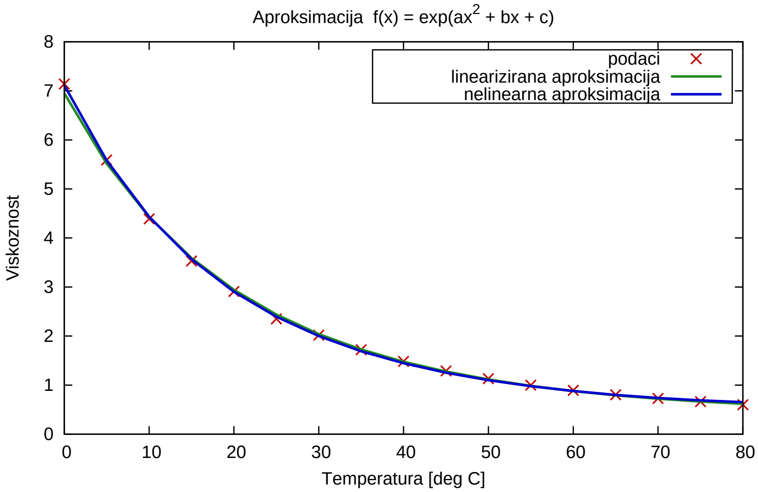
Linearizirani model:

$$\ln f(x) = ax^2 + bx + c$$

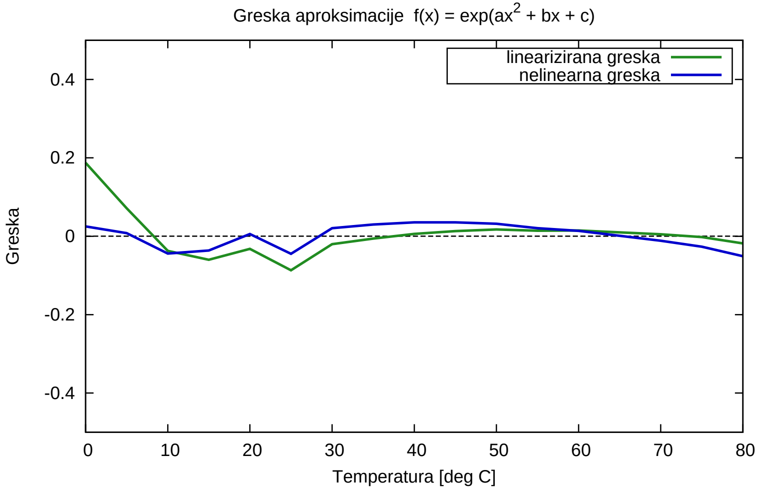
ima parametre

$$a = 2.128853 \cdot 10^{-4},\ b = -4.725758 \cdot 10^{-2},\ c = 1.939119\,.$$

Druga aproksimacija



Greške druge aproksimacije



Matrična formulacija

Diskretni linearni problem najmanjih kvadrata najčešće se rješava u matričnom obliku.

Da bismo formirali matrični zapis linearnog problema najmanjih kvadrata, zgodno je preimenovati nepoznanice,

- tako da matricu,
- vektor desne strane i
- nepoznanice u linearnom sustavu

pišemo u uobičajenoj formi:

- standardno su nepoznanice $x_1, \ldots, x_m$,
- a ne $a_0, \ldots, a_m$.

Matrična formulacija — oznake

Pretpostavimo da skup podataka (t_k, y_k), za $k = 1, \ldots, n$, želimo aproksimirati linearnom funkcijom

$$\varphi(t) = x_1\varphi_1(t) + \cdots + x_m\varphi_m(t).$$

Funkcija φ je neka linearna kombinacija izabranih funkcija baze $\varphi_1, \ldots, \varphi_m$.

Želimo pronaći parametre x_j tako da zadani podaci (t_k, y_k) zadovoljavaju

$$y_k = \sum_{j=1}^{m} x_j\varphi_j(t_k), \quad k = 1, \ldots, n.$$

Primijetite da to nije uvijek moguće, jer je podataka, uobičajeno, znatno više nego parametara ($n \gg m$), pa je zapravo “$=$”$\to$“$\approx$”

Matrična formulacija — preodređeni sustav

Uz oznake

$$a_{kj} = \varphi_j(t_k), \quad b_k = y_k,$$

prethodne jednadžbe možemo napisati kao

$$y_k = \sum_{j=1}^{m} a_{kj} x_j = \begin{bmatrix} a_{k1} & a_{k2} & \cdots & a_{km} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}, \quad k=1,\ldots,n,$$

što daje konačni matrični oblik jednadžbi

$$Ax = b.$$

Oprez: ovdje je A matrica tipa $n \times m$, a ne $m \times n$, kao inače!
Budući da je matrica A "visoka i tanka" ($n \geq m$), imamo

- preodređen sustav linearnih jednadžbi.

Matrična formulacija — min norme reziduala

Taj sustav ne mora uvijek imati rješenje, tj. može se dogoditi, a gotovo uvijek se i događa u praksi, da je

$$r := b - Ax \neq 0, \quad \text{za svaki} \quad x \in \mathbb{R}^m.$$

Postavlja se pitanje: Što je onda "najbolje" rješenje x ovog sustava $Ax = b$? Prirodni odgovor:

- onaj vektor $x \in \mathbb{R}^m$ za kojeg dobivamo "najmanji" rezidual $r = r(x)$.

Naravno, "najmanji" se mjeri u nekoj normi na prostoru $\mathbb{R}^n$.

Najčešće, x određujemo tako da se minimizira Euklidska norma reziduala $r = b - Ax$, tj. tražimo rješenje problema

$$\min_x \|r\|_2 = \min_x \|Ax - b\|_2, \quad A \in \mathbb{R}^{n \times m}, \quad b \in \mathbb{R}^n.$$

Matrična formulacija — min norme reziduala

Problem "najmanjih kvadrata"svodi se na minimizaciju norme $\|Ax - b\|_2$. To vidimo iz sljedeće analize:

Minimizacija norme $\|Ax - b\|_2$ ekvivalentna je minimizaciji kvadrata norme $\|Ax - b\|_2^2$, a on je dalje jednak

$$
\begin{aligned}
\|Ax - b\|_2^2 &= \sum_{i=1}^{n} (Ax - b)_i^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} x_j - b_i \right)^2 \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \varphi_j(t_i) x_j - y_i \right)^2 \\
&= \sum_{i=1}^{n} (\varphi(t_i) - y_i)^2 = S(x_1, \ldots, x_n),
\end{aligned}
$$

funkciji S koja se minimizira u problemu najmanjih kvadrata.

Komentar

Ako smo dobro izabrali bazne funkcije φ_j, onda je razumno pretpostaviti da su one linearno nezavisne na zadanim podacima, tj. stupci matrice A su linearno nezavisni, pa

- matrica A ima puni stupčani rang, tj. $\mathrm{rang}(A) = m$.

Pokazat ćemo da, uz taj uvjet, problem najmanjih kvadrata uvijek ima jedinstveno rješenje.

S druge strane, ako je $\mathrm{rang}(A) < m$, onda

- rješenje x sigurno nije jedinstveno,
- jer mu možemo dodati bilo koji vektor iz nul-potprostora od A, a da se rezidual ne promijeni.

Za početak, nećemo pretpostaviti nikakva specijalna svojstva matrice A, tj. tvrdnje u nastavku vrijede za bilo koji problem.

Rješenje problema najmanjih kvadrata

Teorem. Skup svih rješenja problema $\min_x \|r\|_2$ označimo s

$$\mathcal{S} = \{\, x \in \mathbb{R}^m \mid \|Ax - b\|_2 = \min \,\}.$$

Tada je $x \in \mathcal{S}$, tj. x je rješenje problema najmanjih kvadrata, ako i samo ako vrijedi sljedeća relacija ortogonalnosti

$$A^T(b - Ax) = 0,$$

koju obično nazivamo sustav normalnih jednadžbi i pišemo u obliku

$$A^T A x = A^T b.$$

Rješenje problema najmanjih kvadrata

Dokaz. Pretpostavimo da je $x \in \mathcal{S}$, tj. da x minimizira normu reziduala. Treba pokazati da x zadovoljava sustav normalnih jednadžbi.

Kao što smo vidjeli problem $\min_x \|Ax - b\|_2$ ekvivalentan je minimizaciji funkcije

$$S(x) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} x_j - b_i \right)^2 .$$

Nužan uvjet za postojanje ekstrema je

- $\nabla S(x) = 0$, odnosno $\frac{\partial S(x)}{\partial x_k} = 0$ za $k = 1, \ldots, m$,

Rješenje problema najmanjih kvadrata

Zbog toga računamo:

$$
\begin{aligned}
\frac{\partial S(x)}{\partial x_k} &= \sum_{i=1}^{n} 2 \left(\sum_{j=1}^{m} a_{ij} x_j - b_i \right) a_{ik} \\
&= 2 \sum_{i=1}^{n} a_{ik} (Ax - b)_i = 2 \sum_{i=1}^{n} (A^T)_{ki} (Ax - b)_i \\
&= 2 (A^T (Ax - b))_k = 0.
\end{aligned}
$$

Prema tome je

$$
\nabla S(x) = 0 \quad \Longleftrightarrow \quad A^T (Ax - b) = 0.
$$

Obrat. Pretpostavimo da x zadovoljava normalne jednadžbe, i tada imamo stacionarnu točku funkcije $S(x)$. Moramo još provjeriti Hesseovu matricu $H(x) = \left[\frac{\partial^2 S(x)}{\partial x_k \partial x_\ell} \right]$

Rješenje problema najmanjih kvadrata

Za drugu derivaciju vrijedi

$$\frac{\partial^2 S(x)}{\partial x_k \partial x_\ell} = 2\sum_{i=1}^{n} a_{ik} a_{i\ell} = 2\sum_{i=1}^{n} (A^T)_{ki} a_{i\ell} = 2(A^T A)_{k\ell}.$$

Dakle, možemo zaključiti da je

$$H(x) = 2A^T A.$$

Matrica $A^T A$ simetrična i pozitivno semidefinitna, jer za svaki vektor $y \in \mathbb{R}^m$ vrijedi

$$y^T A^T A y = (y^T A^T)(Ay) = (Ay)^T (Ay) = \|Ay\|_2^2 \geq 0.$$

To znači da je i $H(x)$ pozitivno semidefinitna matrica.

- Za slučajeve kade je $H(x)$ pozitivno definitna (A punog stupčanog ranga) dokazali smo teorem, pri čemu
  - normalne jednadžbe imaju jedinstveno rješenje x (matrica sustava je regularna)
  - $S(x)$ zaista poprima strogi globalni minimum.

Rješenje problema najmanjih kvadrata

- Za slučajeve kade je $H(x)$ nije regularna, neka je $r = b - Ax$ rezidual od x i vrijedi $A^T r = 0$. Za bilo koji vektor $\hat{x} \in \mathbb{R}^m$, njegov rezidual $\hat{r}$ ima oblik

$$\hat{r} = b - A\hat{x} = (r + Ax) - A\hat{x} = r - A(\hat{x} - x).$$

Ako označimo $e = \hat{x} - x$, onda je $\hat{r} = r - Ae$, pa imamo

$$\begin{aligned} \|\hat{r}\|_2^2 &= \hat{r}^T\hat{r} = (r - Ae)^T(r - Ae) \\ &= r^T r - r^T Ae - (Ae)^T r + (Ae)^T Ae \\ &= \|r\|_2^2 - (A^T r)^T e - e^T(A^T r) + \|Ae\|_2^2 = \{A^T r = 0\} \\ &= \|r\|_2^2 + \|Ae\|_2^2 = \|r\|_2^2 + \|A(\hat{x} - x)\|_2^2. \end{aligned}$$

Zbog $\|Ae\|_2^2 \geq 0$, odavde odmah slijedi da je

$$\|\hat{r}\| \geq \|r\|, \quad \text{za svaki} \quad \hat{x} \in \mathbb{R}^m,$$

pa x minimizira normu reziduala, tj. vrijedi $x \in \mathcal{S}$.

Struktura skupa svih rješenja

Napomena. Trenutno, još uvijek ne znamo je li $\mathcal{S} \neq \emptyset$.

Iz zadnje relacije

$$\|\hat{r}\|_2^2 = \|r\|_2^2 + \|Ae\|_2^2 = \|r\|_2^2 + \|A(\hat{x} - x)\|_2^2$$

odmah dobivamo i preciznu strukturu skupa $\mathcal{S}$ svih rješenja problema najmanjih kvadrata.

Korolar. Neka je $x \in \mathcal{S}$. Onda je $\hat{x} \in \mathcal{S}$ ako i samo ako je $\hat{x} - x \in \mathcal{N}(A)$, tj. skup $\mathcal{S}$ je linearna mnogostrukost u $\mathbb{R}^m$

$$\mathcal{S} = x + \mathcal{N}(A).$$

Dokaz. Očito je $\hat{x} \in \mathcal{S}$, ako i samo ako vrijedi $\|\hat{r}\|_2 = \|r\|_2$. No, iz prethodne relacije vidimo da je to ekvivalentno s $Ae = A(\hat{x} - x) = 0$, odnosno, $\hat{x} - x \in \mathcal{N}(A)$. ∎

Usput (kao dodatak dokazu teorema), onda iz $\hat{r} = r - Ae$ slijedi $\hat{r} = r$, pa i $\hat{r}$ zadovoljava normalne jednadžbe.

Egzistencija rješenja — uvijek postoji!

Prethodni teorem, zapravo, kaže da je skup $\mathcal{S}$ rješenja problema minimizacije $\|r\|_2$ jednak skupu rješenja sustava normalnih jednadžbi

$$A^T A x = A^T b.$$

Odavde odmah slijedi egzistencija rješenja, tj. $\mathcal{S} \neq \emptyset$, jer znamo da vrijedi sljedeće:

- Matrica $A^T A$ je simetrična i pozitivno semidefinitna.
- Sustav normalnih jednadžbi uvijek ima rješenje, jer je

  $$A^T b \in \mathcal{R}(A^T) = \mathcal{R}(A^T A)$$

  (v. teorem Kronecker–Capelli). Da $\mathcal{R}(A^T) = \mathcal{R}(A^T A)$ vrijedi, provjerite sljedeće
  - $\mathcal{R}(A) \subseteq \mathbb{R}^n$, i kad primijenimo A^T imamo $\mathcal{R}(A^T A) \subseteq \mathcal{R}(A^T)$.
  - Neka je $z \in \mathcal{R}(A^T)$, tada je $z = A^T y$. Kako je $\mathbb{R}^n = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$ tada y možemo raspisati kao $y = Ax + w$, gdje je $w \in \mathcal{N}(A^T)$. Tada je $z = A^T A x + A^T w = A^T A x \in \mathcal{R}(A^T A)$, odnosno $\mathcal{R}(A^T) \subseteq \mathcal{R}(A^T A)$.

Jedinstvenost rješenja

Dobivamo čak i jače — zaključak o jedinstvenosti rješenja.

Teorem. Problem najmanjih kvadrata ima jedinstveno rješenje, ako i samo ako vrijedi bilo koja od sljedećih tvrdnji:

- A ima puni stupčani rang, tj. vrijedi $\mathrm{rang}(A) = m$,
- stupci matrice A su linearno nezavisni,
- $A^T A$ je pozitivno definitna matrica.

Dokaz. Iz korolara o strukturi skupa $\mathcal{S}$ svih rješenja problema najmanjih kvadrata, vidimo da je

- skup $\mathcal{S}$ jednočlan, ako i samo ako
- matrica A ima trivijalan nul-potprostor $\mathcal{N}(A)$, tj. vrijedi $\dim \mathcal{N}(A) = 0$, odnosno, $x \neq 0 \implies Ax \neq 0$.

Jedinstvenost rješenja

Trivijalnost $\mathcal{N}(A)$ ekvivalentna je sljedećim zaključcima.

1. tvrdnja — koristimo $\dim \mathcal{N}(A) = 0$.

- Iz teorema o rangu i defektu za matricu A, tipa $n \times m$, to je ekvivalentno s $\mathrm{rang}(A) = m$. Usput, onda je $n \geq m$.

2. i 3. tvrdnja — koristimo $x \neq 0 \Longrightarrow Ax \neq 0$.

- Po definiciji, stupci $\{a_1, a_2, \ldots, a_m\}$ matrice A su linearno nezavisni, ako i samo ako za svaki skup skalara $x_1, x_2, \ldots x_m$ kod kojih je bar jedan $x_k \neq 0$ vrijedi $\sum_{j=1}^{m} x_j a_j \neq 0$ To je ekvivalentno tvrdnji da za svaki vektor $x = [\ x_1 \quad x_2 \quad \cdots x_m\ ]^T \neq 0$ vrijedi $Ax \neq 0$.
- Znamo da je $A^T A$ simetrična i pozitivno semidefinitna. Onda je gornja implikacija ekvivalentna pozitivnoj definitnosti matrice $A^T A$, jer za $x \neq 0$ vrijedi

$$x^T A^T A x = \|Ax\|_2^2 > 0.$$

Karakterizacija jedinstvenog rješenja

Zadnja tvrdnja — preko pozitivne definitnosti matrice $A^T A$, odgovara činjenici da sustav normalnih jednadžbi

$$A^T A x = A^T b$$

ima jedinstveno rješenje, ako i samo ako je $A^T A$ regularna matrica (pozitivno definitna matrica je regularna).

U tom slučaju, jedinstveno rješenje problema najmanjih kvadrata je

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

a pripadni rezidual najmanje 2–norme je

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

Geometrijska interpretacija općeg rješenja

Desna strana b može se napisati preko reziduala kao

$$b = Ax + r,$$

pri čemu je, očito, $Ax \in \mathcal{R}(A)$. Nadalje, iz sustava normalnih jednadžbi odmah vidimo da je

$$A^T(b - Ax) = A^T r = 0,$$

što znači da je za svaki $Av \in \mathcal{R}(A)$

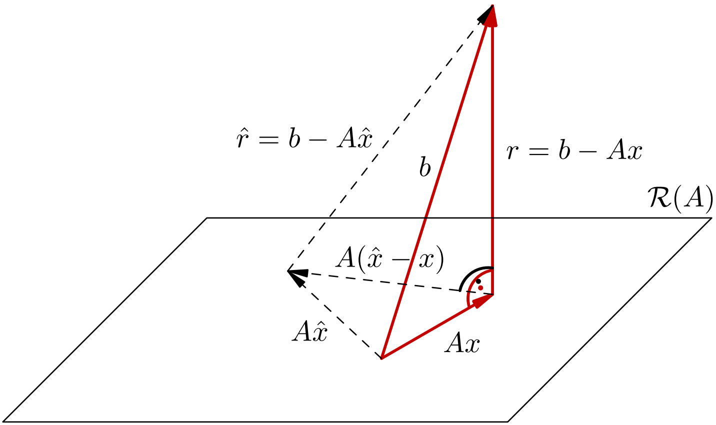
$$(Av)^T r = (v^T A^T) r = v^T (A^T r) = v^T 0 = 0$$

tj. $r \perp \mathcal{R}(A)$, pa vektor r mora biti okomit i na Ax. To daje geometrijsku interpretaciju rješenja problema najmanjih kvadrata.

Geometrijska interpretacija općeg rješenja

Rješenje problema najmanjih kvadrata dobivamo

- ortogonalnom projekcijom vektora b na potprostor $\mathcal{R}(A)$.



Struktura skupa rješenja i linearni sustav

Označimo s $\mathcal{P}(b)$ ortogonalnu projekciju vektora b na $\mathcal{R}(A)$.

- Taj vektor $\mathcal{P}(b)$ je sigurno jedinstven (kao projekcija).

Prema slici, za rješenje x mora vrijediti $Ax = \mathcal{P}(b)$. Preciznije,

- sva rješenja $x \in \mathcal{S}$ problema najmanjih kvadrata su,
- upravo, sva rješenja linearnog sustava $Ax = \mathcal{P}(b)$.

Zato skup $\mathcal{S}$ ima istu strukturu — linearna mnogostrukost, kao i opće rješenje linearnog sustava s matricom A, s tim da

- ovdje znamo da je $\mathcal{S}$ neprazan, tj. postoji $x_0 \in \mathcal{S}$.

Analogno, jedinstvenost rješenja x svodi se na

- jedinstvenost rješenja linearnog sustava s matricom A, tj. trivijalnost nul-potprostora $\mathcal{N}(A)$.

Numeričke metode i jedinstvenost rješenja

Numeričke metode za računanje rješenja problema najmanjih kvadrata imaju smisla samo kad je

- objekt kojeg računamo jedinstven.

Znamo da linearni problem najmanjih kvadrata uvijek ima rješenje. Osim toga, rješenje je jedinstveno, ako i samo ako

- matrica A ima puni stupčani rang.

U nastavku, tražimo efikasne i točne numeričke metode za računanje rješenja. Promatramo samo one probleme

- u kojima imamo garantiranu jedinstvenost rješenja.

Prije toga — završni komentar o jedinstvenosti.

Osiguranje jedinstvenosti u općem problemu

Ako matrica A nema puni rang po stupcima, tj. ako je $\mathrm{rang}(A) < m$, onda

- rješenje problema najmanjih kvadrata nije jedinstveno.

U tom slučaju, jedinstvenost rješenja se dobiva dodatnim uvjetom — pored $\|Ax - b\|_2 \to \min$, još tražimo

- rješenje $x \in \mathcal{S}$ koje ima najmanju 2-normu, tj. dodatni uvjet je $\|x\|_2 \to \min$.

Geometrijska interpretacija:

- Na $\|x\|_2 \to \min$ možemo gledati kao na vektor oblika $x_0 + \sum_{i=1}^{d} y_i n_i$ kojem minimiziramo normu, pri čemu vektori n_i, $i = 1, \ldots, d$ čine bazu od $\mathcal{N}(A)$ (znamo da je jezgra matrice vektorski prostor), a $x_0 \in \mathcal{S}$.
- To znači da rješavamo problem $\|x_0 + Ny\|_2 \to \min$, pri čemu je $N = [n_1 \ \cdots \ n_d]$. Na analogan način kao i kod linearnog problema najmanjih kvadrata zaključujemo da rješenje problema $x = x_0 + Ny$, koje je oblika kao rezidual najmanjih kvadrata, mora bit okomito na $\mathcal{R}(N) = \mathcal{N}(A)$, tj.
$$x \perp \mathcal{N}(A).$$

Osiguranje jedinstvenosti u općem problemu

Alternativno, skup rješenja u ovom slučaju čini skup

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

za bilo koje rješenje x polaznog problema najmanjih kvadrata.

- Ako je $x \perp \mathcal{N}(A)$, onda je

$$\|x + z\|_2^2 = \|x\|_2^2 + \|z\|_2^2, \quad z \in \mathcal{N}(A),$$

  pa je x jedinstveno rješenje problema najmanjih kvadrata koje ima minimalnu 2-normu.

Ovaj pristup ima smisla u primjenama, na pr., u statistici.

Jedinstveno rješenje — ponavljanje

Odsad nadalje, pretpostavljamo (ako drugačije nije rečeno) da matrica A ima puni stupčani rang. Posebno, to znači da

- A ima više redaka nego stupaca, $n \geq m$, i
- stupci od A su linearno nezavisni, $\mathrm{rang}(A) = m$.

Matrica $A^T A$ je pozitivno definitna, a iz sustava normalnih jednadžbi

$$A^T A x = A^T b$$

dobivamo jedinstveno rješenje problema najmanjih kvadrata

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

Pripadni rezidual najmanje 2–norme je

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

Računanje rješenja problema najmanjih kvadrata

Treba još pronaći način kako jednostavno izračunati rješenje. Jasno je da se matrica $A^T A$ ne invertira, nego se rješava linearni sustav

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

Ovaj pozitivno definitni sustav normalnih jednadžbi možemo riješiti tako da iskoristimo faktorizaciju Choleskog za $A^T A$.

Prednosti/nedostaci ove metode = “množenje + Cholesky”:

- ukupan broj aritmetičkih operacija za rješenje je $nm^2 + \frac{1}{3} m^3 + O(m^2)$, što je brzo,
- ali, rješavanje na ovaj način nije naročito točno.

Može se koristiti za mali broj parametara m, ako ne tražimo jako točno rješenje (često u praksi — za “mala” mjerenja).

Korištenje QR faktorizacije

Opet, neka A ima puni stupčani rang. Promatramo problem minimizacije

$$\|Ax - b\|_2 \to \min .$$

Prisjetite se, za proizvoljnu ortogonalnu matricu Q^T vrijedi da čuva skalarni produkt — onda i kvadrat norme, pa i normu.

Dakle, rješenje problema minimizacije možemo zapisati kao

$$\|Ax - b\|_2 = \|Q^T(Ax - b)\|_2 = \|Q^T Ax - Q^T b\|_2 \to \min .$$

Pitanje. Kako naći pogodan Q^T, tako da, iz problema ili "sustava" s matricom $Q^T A$, lako izračunamo rješenje x?

Odgovor. Korištenjem QR faktorizacije — tako da $Q^T A$ bude gornja trokutasta matrica R, tj. faktorizacijom $A = QR$!

Definicija QR faktorizacije

Napomena. U ovom dijelu mijenjamo oznake $m \leftrightarrow n$, na uobičajene za matrice: m = broj redaka, a n = broj stupaca.

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 napisati i u jednostavnijoj — tzv. skraćenoj formi.

- Prvih n stupaca matrice Q označimo s Q_0, tako da matrica Q_0 ima isti tip kao i G,
- a preostale stupce, koji su okomiti na Q_0, označimo s $Q_0^\perp$.

Onda je

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

Ostaje samo pokazati da takva faktorizacija postoji.

Napomena. Ako je $m > n$, onda $Q_0^\perp$ možemo izabrati na više načina $\Longrightarrow$ "puna" QR faktorizacija sigurno nije jedinstvena.

Egzistencija i jedinstvenost QR faktorizacije

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

$$G = Q_0 R_0,$$

gdje 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 u R).

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

Egzistencija i jedinstvenost QR faktorizacije

Dokaz. Najjednostavniji dokaz ide tako da stupce matrice

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

ortonormiramo korištenjem Gram-Schmidtovog postupka.

1. korak: Zbog $g_1 \neq 0$, definiramo

$$q_1' = g_1, \qquad q_1 = \frac{q_1'}{\|q_1'\|_2}.$$

j-ti korak: Već imamo ortonormirane vektore $q_1, \ldots, q_{j-1}$, koji razapinju isti potprostor kao i stupci $g_1, \ldots, g_{j-1}$ matrice G. Onda definiramo novi vektor q_j' i normiramo ga

$$q_j' = g_j - \sum_{i=1}^{j-1} \langle g_j, q_i \rangle \, q_i, \qquad q_j = \frac{q_j'}{\|q_j'\|_2}.$$

Egzistencija i jedinstvenost QR faktorizacije

Stupci matrice G su linearno nezavisni, što osigurava $q_j' \neq 0$. Stavljanjem

$$Q_0 = [\, q_1 \; q_2 \; \dots \; q_n \,]$$

dobivamo $m \times n$ ortogonalnu matricu (ortonormirani stupci).

Uz oznaku za skalarne produkte i norme iz prethodne formule

$$r_{ij} = \langle g_j, q_i \rangle = q_i^T g_j, \qquad r_{jj} = \|q_j'\|_2,$$

koeficijenti r_{ij} su, upravo, elementi tražene matrice R_0. Polazni stupac g_j možemo napisati kao linearnu kombinaciju prvih j vektora q_i ortonormirane baze, u obliku

$$G(:,j) = g_j = \sum_{i=1}^{j} r_{ij} q_i = [\, q_1 \; q_2 \; \dots \; q_n \,] \begin{bmatrix} r_{1j} \\ \vdots \\ r_{jj} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = Q_0 \cdot R_0(:,j).$$

tj. $G = Q_0 R_0$.

Gram–Schmidtov postupak ortogonalizacije

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

- koji ortogonalizira originalne vektore g_i.
- Zbog toga jer je nestabilan kad su stupci od G skoro linearno zavisni, tj. kad je G loše uvjetovana.

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

- koji ortogonalizira modificirane vektore g_j, koji su već sami ortogonalni na prethodne vektore q_i, pa je mnogo stabilniji.
- No, i kod njega se može dogoditi da je izračunati Q_0 vrlo daleko od ortogonalnog, tj. $\|Q_0^T Q_0 - I_n\| \gg u$, kad je G vrlo loše uvjetovana.

Gram–Schmidtov algoritam

Klasični i modificirani Gram–Schmidtov algoritam:

```
za j = 1 do n radi {
  /* Nađi j-ti stupac od Q_0 i R_0 */
  q'_j = g_j;
  za i = 1 do j - 1 radi {
    /* Oduzmi komponentu od g_j u smjeru q_i */
    /* kod CGS-a je */
    r_ij = q_i^T * g_j;
    /* kod MGS-a je */
    r_ij = q_i^T * q'_j;
    q'_j = q'_j - r_ij * q_i;
  };
```

Gram–Schmidtov algoritam (nastavak)

```
  r_jj = ‖q′_j‖₂;
  ako je r_jj > 0 onda {
    q_j = q′_j / r_jj;
  }
  inače {
    /* Matrica R_0 je singularna - stani */
  };
};
```

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

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

Pokažite da su dvije formule za r_{ij}, ona iz CGS i ona iz MGS, matematički ekvivalentne. Numerički, naravno, nisu (greške).

Gram–Schmidtov algoritam — komentari

Gram–Schmidtov algoritam daje skraćenu QR faktorizaciju $G = Q_0 R_0$. Ako je $m > n$ (matrica Q_0 nije kvadratna), onda nam za "punu" faktorizaciju, tj. za kvadratni Q,

- fali ortogonalni komplement $Q_0^{\perp}$,

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

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

$$r_{jj} = \pm \|q_j'\|_2.$$

Dakle, bilo kojim fiksiranjem predznaka na dijagonali od R_0,

- opet dobivamo jedinstvenu skraćenu QR faktorizaciju.

Drugi algoritmi

U praksi, kad želimo ortogonalni Q, koristimo

- ili Givensove rotacije,
- ili Householderove reflektore,

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

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 punu i skraćenu QR faktorizaciju.

Metoda najmanjih kvadrata

Diskretni problem najmanjih kvadrata

Matrična formulacija linearnog problema najmanjih kvadrata

QR faktorizacija

Givensove rotacije

Householderovi reflektori

QR faktorizacija i pivotiranje

Givensove rotacije

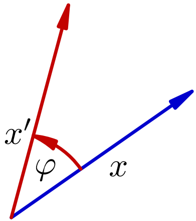
Matrica oblika

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

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 = u pozitivnom smjeru.

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

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

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 jednadžbi 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'.

Za $x_i = x_j = 0$, mora biti $x_i' = 0$ i možemo uzeti $R(i, j, \varphi) = I$.

Poništavanje korištenjem Givensovih rotacija

U nastavku uzimamo da je $x_i^2 + x_j^2 > 0$, tj. bar jedan nije nula.

Drugi redak u matričnoj jednadžbi opisuje poništavanje

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

Ako je $x_j = 0$ (tj. nemamo što poništavati), onda je $\sin\varphi = 0$.
U suprotnom, izlazi

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

Odavde, korištenjem trigonometrijskog identiteta

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

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

$$\begin{aligned} x_i' &= \frac{x_i}{\sqrt{x_i^2 + x_j^2}}\, x_i + \frac{x_j}{\sqrt{x_i^2 + x_j^2}}\, x_j = \frac{x_i^2 + x_j^2}{\sqrt{x_i^2 + x_j^2}} \\ &= \sqrt{x_i^2 + x_j^2} > 0. \end{aligned}$$

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 redoslijeda kako napraviti nule u matrici G.
- U sljedećem primjeru uzet je “standardni” redoslijed: redom, po stupcima ($\rightarrow$), odozgo nadolje ($\downarrow$) u stupcu.

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

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

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

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

Sustavno poništavanje — primjer (nastavak)

2. stupac:

- U radnoj matrici G, redom poništavamo elemente

$$g_{i2}, \quad i = 3, \ldots, m \ \ (m = 4),$$

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 (i stupac) 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

$$g_{i3}, \quad i = 4, \dots, m \ \ (m = 4),$$

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 (i stupca) 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 primjene rotacija, usput, još dozvoljava i paralelizaciju algoritma.

Nezavisne rotacije — paralelno poništavanje

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 $(2, 4)$. Na samom kraju, u zadnja dva retka, više "ne ide" paralelno. Plave 0 su konačne.

Kako doći do Q?

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

Do matrice Q dolazi se nakupljanjem primijenjenih rotacija, na pr.

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

Matricu G smo slijeva pomnožili

- produktom ortogonalnih matrica, kojeg označimo s Q^{-1}.
- Produkt ortogonalnih matrica je opet ortogonalna, pa i regularna. Isto vrijedi i za njezin inverz $(Q^{-1})^{-1} = Q$.
- Zaključak: $G = QR$, gdje je Q ortogonalna.
- Ako znamo $Q^{-1} = Q^T$, onda se Q lako računa iz Q^T.

Matrica $Q^{-1} = Q^T$ dobiva se primjenom istih rotacija, samo na početnu matricu I_m, reda m — “što na G, to na I_m”.

Alternativa za Q, puni i skraćeni Q

Kvadratnu matricu Q, u punoj QR faktorizaciji, možemo dobiti i bez transponiranja — akumulacijom produkta

- inverznih rotacija zdesna, na početnu matricu I_m.

Inverzna rotacija = rotacija za suprotni kut (tj. $\varphi \mapsto -\varphi$).
Na pr.,

$$
\begin{aligned}
Q &= \big(R(n, m, \varphi_{nm}) \cdots R(1, 2, \varphi_{12})\, I_m\big)^T \\
&= I_m\, R(1, 2, -\varphi_{12}) \cdots R(n, m, -\varphi_{nm}).
\end{aligned}
$$

Pravokutnu matricu Q_0 iz skraćene QR faktorizacije dobivamo tako da uzmemo prvih n stupaca od završne matrice Q,

$$
Q_0 = Q(1:n).
$$

A obratno? Napravimo punu QR faktorizaciju matrice Q_0!

Metoda najmanjih kvadrata

Diskretni problem najmanjih kvadrata

Matrična formulacija linearnog problema najmanjih kvadrata

QR faktorizacija

Givensove rotacije

Householderovi reflektori

QR faktorizacija i pivotiranje

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

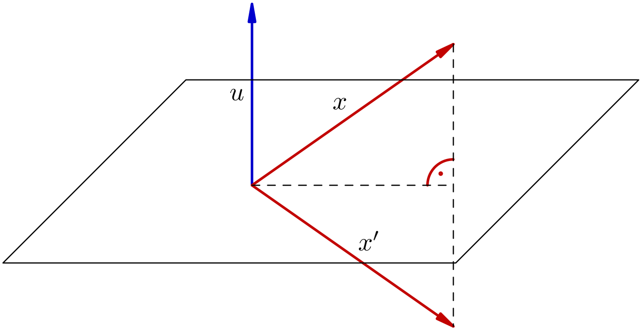
- simetrična,
- i ortogonalna, jer je

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

Zašto baš ime reflektor?

Promatrajmo hiperravninu koja je okomita na vektor u.

- Reflektor H sve vektore x preslikava u simetrične obzirom na tu hiperravninu, $x' = Hx$.



Usput, $P(u) := I - uu^T$ je projektor na tu hiperravninu.

Poništavanje Householderovim reflektorima

Neka je zadan vektor x. Treba naći Householderov reflektor H koji poništava sve komponente vektora x, osim prve. Dakle, treba naći jedinični vektor u koji definira takav reflektor H.

Tražimo da je

$$Hx = \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix} = c \cdot e_1, \quad c \in \mathbb{R}.$$

Za početak, H je unitarna matrica, pa čuva normu vektora,

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

odakle slijedi da je $c = \pm\|x\|_2$.

Poništavanje Householderovim reflektorima

Napišimo traženu jednadžbu preko nepoznatog vektora u

$$Hx = (I - 2uu^T)x = x - 2u(u^T x) = \pm \|x\|_2 \, e_1.$$

Uočimo da je $u^T x$ broj = skalarni produkt dva vektora.

Zanemarimo na trenutak mogućnost da je $u^T x = 0$. Onda je

$$u = \frac{1}{2(u^T x)} (x \mp \|x\|_2 \, e_1).$$

Obzirom na to da $u^T x$ ne znamo, možemo zaključiti da je

$$u = \alpha(x \mp \|x\|_2 \, e_1), \quad \text{za neki } \alpha \in \mathbb{R},$$

tj. da je u paralelan s vektorom $\widetilde{u} = x \mp \|x\|_2 e_1$. Konačno, konstantu α nalazimo normiranjem vektora $\widetilde{u}$ na normu 1.

Poništavanje Householderovim reflektorima

Prema tome,

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

ako je $\widetilde{u} \neq 0$. U protivnom, stavljamo $u = 0$, odnosno, $H = I$. Uočite da izbor predznaka $(\mp)$ u definiciji $\widetilde{u}$ daje

$$Hx = c \cdot e_1 = \pm \|x\|_2 \, e_1 .$$

Ako je $x = 0$, onda je $\widetilde{u} = 0$, a to znači da je $H = I$. Tada možemo uzeti bilo koji u, jer je $H \cdot 0 = 0$ za svaki H.

U protivnom je i $\|x\|_2 \neq 0$, pa barem jedan izbor predznaka $(\mp)$ u formuli za $\widetilde{u}$ daje $\widetilde{u} \neq 0$. Ako jedan izbor daje $\widetilde{u} = 0$, onda je $x = \pm \|x\|_2 \, e_1$, pa i nemamo što raditi. Dakle, oba reflektora su korektna.

Izbor predznaka za u

Ranija mogućnost da je $u^T x = 0$ ne pravi nikakve poteškoće. Onda je $Hx = x = \pm \|x\|_2\, e_1$ i samo tada bar jedan izbor predznaka daje $\widetilde{u} = 0$. Dobiveni $H(u)$ je i tada korektan!

Za $x \neq 0$, u praksi se, zbog numeričke stabilnosti, često koristi

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

zato da nema kraćenja pri računanju prve komponente od $\widetilde{u}$, tj. da oba pribrojnika budu istog znaka,

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

To znači da je $c = -\mathrm{sign}(x_1) \|x\|_2$, što ponekad zbunjuje u praksi.

- Na primjer, za $x = e_1$ dobivamo $Hx = -e_1$!

Stvarno, uz pažljiviji poredak računanja, to nije potrebno!

Drugi način definicije H, primjena H na vektor

Napomena. Normiranje $\widetilde{u} \mapsto u$ se, formalno, može izbjeći, ako definiramo

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

Kako djelovati s H na ostale stupce (ili neki vektor)?

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

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

Dakle, treba izračunati skalarni produkt $u^T z$, a zatim modificirati vektor z (to je neki stupac radne matrice).

Ako koristimo $H(\widetilde{u})$, onda $\widetilde{u}^T z$ stalno treba dijeliti s $\|\widetilde{u}\|_2^2$.

QR faktorizacija korištenjem reflektora

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

- Prvo se reflektorom H_1 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'.

  Na cijelu radnu matricu onda djelujemo reflektorom

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

  U pripadnom vektoru u, prva komponenta je $u_1 = 0$.

QR faktorizacija korištenjem reflektora

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

- reflektorom H'_k se poništava k-ti "skraćeni" stupac u radnoj matrici — od dijagonale nadolje.
- Na cijelu radnu matricu onda djelujemo reflektorom

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

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

Ako želimo formirati ortogonalnu 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

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

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 — radna matrica

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

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

Računanje QR faktorizacije — k-ti korak

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

Kod QR faktorizacije, također, možemo koristiti pivotiranje, slično kao kod LU faktorizacije ili faktorizacije Choleskog.

- Uobičajeno se koristi pivotiranje (= zamjene) stupaca

  $$GP = QR,$$

  gdje je P matrica permutacije.

Pivotiranje stupaca u k-tom koraku algoritma ($k = 1, \ldots, n$):

- Ako su x_ℓ, za $\ell = k, \ldots, n$, skraćeni stupci (od k-tog do m-tog reda), na "prvo" (tj. k-to) mjesto dovodi se onaj s najvećom normom, tj. tako da $\|x_k\|_2$ bude maksimalna.
- Zamjene se rade s cijelim stupcima, a ne sa skraćenim!

U zadnjem koraku više nema zamjena (samo jedan stupac).

Svrha pivotiranja

Svrha?

- Ako matrica G ima (skoro) linearno zavisne stupce, onda se QR faktorizacijom s pivotiranjem numerički određuje rang matrice G — “rez” kad dijagonala u R_0 “padne”.

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

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

Veza QR faktorizacije i faktorizacije Choleskog

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

Matrica G ima jedinstvenu skraćenu QR faktorizaciju, pa je puni QR oblika

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

gdje je R_0 jedinstvena 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.