2006–03–23

Lay 6.5 Minsta kvadrat

$A\boldsymbol{x}=\boldsymbol{b}, A\in\mathbb{R}^{m\times n}, m>n.$

Antag att $\mathrm{rang}(A)=n$. Titta på homogena lösningar $A\boldsymbol{x}=\boldsymbol{0}$:

$$A\boldsymbol{x}=\boldsymbol{0}\Longleftrightarrow\boldsymbol{x}\in\mathrm{Col}(A^T)^{\perp}\tag{1}$$

ty $\boldsymbol{x}$ är vinkelrät mot alla rader i A, dvs alla kolonner i A^T, dvs $\boldsymbol{x}\in\mathrm{Col}(A^T)^{\perp}$.

(1) kan uttryckas:

$$x\in\mathrm{Nul}(A)\Longleftrightarrow x\in\mathrm{Col}(A^T)^{\perp}$$

$$\mathrm{Nul}(A)=\mathrm{Col}(A^T)^{\perp}$$

På samma sätt om vi betraktar A^T:

$$\mathrm{Nul}(A^T)=\mathrm{Col}(A)^{\perp}\tag{2}$$

Vårt problem, minsta-kvadratproblemet: minimera $\|A\boldsymbol{x}-\boldsymbol{b}\|_2$, dvs vi ska välja den vektor

$$A\hat{\boldsymbol{x}}\in\mathrm{Col}(A)$$

som är närmast $\boldsymbol{b}$. Lösningen ges av att felet $\boldsymbol{b}-A\hat{\boldsymbol{x}}\in\mathrm{Col}(A)^{\perp}$. ⟨fig6⟩

Men (2) ger då $\boldsymbol{b}-A\hat{\boldsymbol{x}}\in\mathrm{Nul}(A^T), \mathrm{dvs}\ A^T(\boldsymbol{b}-A\hat{\boldsymbol{x}})=\boldsymbol{0}\Longleftrightarrow$

$$A^TA\hat{\boldsymbol{x}}=A^T\boldsymbol{b}\tag{3}$$

(3) kallas normalekvationerna. Detta är ett *kvadratiskt* ekvationssystem med en *entydig* lösning.

Övning: Visa att A^TA är *symmetrisk* och *positivt definit* ($\Rightarrow$ reguljär, inverterbar).

Anmärkning: Men (3) är inte det bästa sättet att lösa problemet. Tyvärr har A^TA ofta ett stort konditionstal. Man kan visa att $\kappa_2(A^TA)=\kappa_2(A)^2$. ($\kappa(A)=\|A^{-1}\|\cdot\|A\|$).

Exempel. Lös i minsta-kvadrat-mening då

$$A=\begin{pmatrix}2&1\\-2&0\\2&3\end{pmatrix}\quad\boldsymbol{b}=\begin{pmatrix}-5\\8\\2\end{pmatrix}$$

Normalekvationerna:

$$\begin{pmatrix}18&8\\8&10\end{pmatrix}\hat{\boldsymbol{x}}=\begin{pmatrix}-22\\1\end{pmatrix}$$

Har lösningen:

$$\hat{\boldsymbol{x}}=\begin{pmatrix}-4.0714\\3.3571\end{pmatrix}$$

$$\boldsymbol{r}=A\hat{\boldsymbol{x}}-\boldsymbol{b}\Longrightarrow\|\boldsymbol{r}\|_2=0.2673$$

Lösning med $Q\,R$-faktorisering (MATLABs metod). Fördel: A^TA med stort kondtionstal kommer inte in i räkningen. Minsta kvadratlösningen ges av

$$R\,\hat{\boldsymbol{x}} = Q^T\boldsymbol{b} \tag{4}$$

där $A = Q\,R$ är $Q\,R$-faktorisering av A.

Bevis. Matematiskt ekvivalent med normalekvationerna (fast inte *numeriskt* ekvivalent som beräkningsmetod betraktat.).

$$A^TA\,\hat{\boldsymbol{x}} = A^T\boldsymbol{b} \underset{A=Q\,R}{\Longleftrightarrow} (Q\,R)^T(Q\,R)\,\hat{\boldsymbol{x}} = (Q\,R)^T\boldsymbol{b} \Longleftrightarrow$$

Q är ortogonal: vi vill använda $Q^TQ = I$.

$$\Longleftrightarrow R^TQ^TQ\,R\,\hat{\boldsymbol{x}} = R^TQ^T\boldsymbol{b} \Longleftrightarrow R^TR\,\hat{\boldsymbol{x}} = R^TQ^T\boldsymbol{b} \Longleftrightarrow$$

R (och därmed R^T) är reguljär, inverterbar:

$$R\,\hat{\boldsymbol{x}} = Q^T\boldsymbol{b} \qquad \square$$

EXEMPEL

$$A = \begin{pmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{pmatrix} \underset{\text{Schmidt}}{\overset{\text{Gram-}}{=\!=}} \left(\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \quad \frac{1}{\sqrt{42}}\begin{pmatrix} -1 \\ 4 \\ 5 \end{pmatrix} \right) R = Q\,R$$

$$R = Q^TA = \begin{pmatrix} \frac{6}{\sqrt{3}} & \frac{4}{\sqrt{3}} \\ 0 & \frac{14}{\sqrt{42}} \end{pmatrix} = \frac{1}{\sqrt{3}} = \begin{pmatrix} 6 & 4 \\ 0 & \sqrt{14} \end{pmatrix}$$

Nollan kan tjäna som kontroll att man räknat rätt. R ska ju vara uppåt triangulär.

$$Q^T\boldsymbol{b} = \frac{1}{\sqrt{3}}\begin{pmatrix} -11 \\ \frac{47}{\sqrt{14}} \end{pmatrix}$$

$R\,\hat{\boldsymbol{x}} = Q^T\boldsymbol{b}$ är ett uppåt triangulärt system.

$$\Longrightarrow \hat{\boldsymbol{x}} = \begin{pmatrix} -4.0714 \\ 3.3571 \end{pmatrix}$$

Lay 6.6 Linjära modeller i statistik.

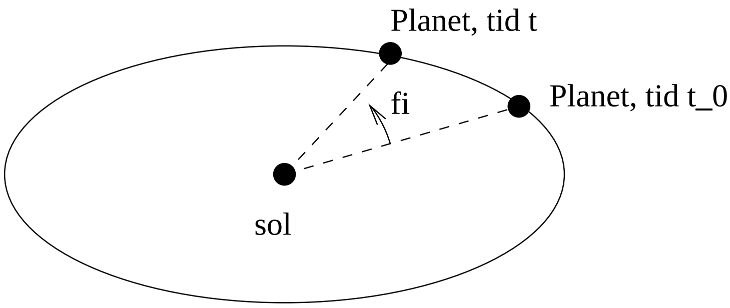
$$X\,\boldsymbol{\beta} = \boldsymbol{y}$$

X är designmatrix, $\boldsymbol{\beta}$ är parametrar, $\boldsymbol{y}$ är en observation. ⟨fig7⟩ Anpassar några mätpunkter till en linjär ekvation $\beta_0 + \beta_1\,x = y$ i minsta kvadratmening. Normalekvation $X^TX\,\hat{\boldsymbol{\beta}} = X^T\boldsymbol{y}$.

⟨fig8⟩. Kvadratisk funktion anpassas. Detta är fortfarande en linjär modell, fast funktionen är kvadratisk.

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 \end{pmatrix}$$

EXEMPEL på linjär modell: Keplers ekvation.



Figur 1.

$$\varphi - e \sin \varphi = a(t - t_0)$$

e och a är parametrar som ska bestämmas. Observationer (t_i, φ_i), $i = 1, ..., m$. Överbestämt system om $m > 2$.

$$a(t_i - t_0) + e \sin \varphi_i = \varphi_i$$

på matrisform:

$$\begin{pmatrix} t_1 - t_0 & \sin \varphi_1 \\ t_2 - t_0 & \sin \varphi_2 \\ \vdots & \vdots \\ t_m - t_0 & \sin \varphi_n \end{pmatrix} \begin{pmatrix} a \\ e \end{pmatrix} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_m \end{pmatrix}$$

I matlab:

Indata:

```
» t0 = _____
» t = [ . . . . . ]';
» fi = [ . . . . . ]';
```

Beräkning (MATLAB kör QR-faktorisering):

```
» x = [ t-t0 sin(fi) ] \fi;
» a = x(1);
» e = x(2);
```

Bonusuppgift B12: Anpassa en ellipsbana till mätpunkter, numeriskt känsligt. ⟨fig10⟩.

Bonusuppgift B8: Anpassning av en periodisk modell, dyngsmedeltemperaturen över året i Göteborg. När kommer våren?

Bonusuppgift B6: Avgör om punkter i rymden ligger på en cirkelperiferi. Kolla om punkterna ligger i ett plan, kolla om de ligger på en cirkelperifieri. QR-faktorisering.

Lay 6.7 Skalärproduktrum.

DEFINITION: En skalärprodukt (innerprodukt) i ett linjärt rum V är en funktion

$$\langle\,\cdot\,,\cdot\,\rangle\colon\quad V\times V\rightarrow\mathbb{R}$$

så att följande regler gäller för $\boldsymbol{u}\in V$, $\boldsymbol{v}\in V$, $\boldsymbol{w}\in V$ och $c\in\mathbb{R}$.

1. $\langle\boldsymbol{u},\boldsymbol{v}\rangle=\langle\boldsymbol{v},\boldsymbol{u}\rangle$ (symmetri)
2. $\langle\boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle=\langle\boldsymbol{u},\boldsymbol{w}\rangle+\langle\boldsymbol{v},\boldsymbol{w}\rangle$
3. $\langle c\,\boldsymbol{u},\boldsymbol{v}\rangle=c\,\langle\boldsymbol{u},\boldsymbol{v}\rangle$
4. $\langle\boldsymbol{u},\boldsymbol{u}\rangle\geqslant 0$ med likhet om endast om $\boldsymbol{u}=\boldsymbol{0}$.

EXEMPEL. Om A är en symmetrisk positivt definit matris så är

$$\langle x,y\rangle=x^TA\,y$$

en skalärprodukt.

Bevis.

1. $\langle y,x\rangle=y^TA\,x=(A^T y)^Tx=(A\,y)^Tx=$ [symmetri hos vanlig skalärprodukt] $=x^T(A\,y)=x^TA\,y=\langle x,y\rangle$
2. $\langle x+y,z\rangle=(x+y)^T A\,z=x^TA\,z+y^T A\,z=\langle x,z\rangle+\langle y,z\rangle$.
3. $\langle c\,x,y\rangle=(c\,x)^TA\,y=c\,x^TA\,y=c\,\langle x,y\rangle$
4. $\langle x,x\rangle=x^T A\,x\geqslant 0$ för alla x med likhet om och endast om $x=\boldsymbol{0}$. □

EXEMPEL: I $\mathcal{C}^0[a,b]$ (mängden kontinuerliga funktioner på $[a,b]$) är

$$\langle f,g\rangle=\int_a^b f(t)\,g(t)\,\mathrm{d}t$$

en skalärprodukt.

DEFINITION: Vinkeln mellan u och v definieras som

$$\theta=\arccos\frac{\langle u,v\rangle}{\|u\|\cdot\|v\|}$$

ÖVNING: Bestäm vinkeln mellan $f(t)=t$ och $g(t)=t^2$ i $\mathcal{C}^0[0,1]$ med skalärprodukten ovan.

Norm: $\|u\|=\sqrt{\langle u,u\rangle}$.

LÖSNING:

$$\langle f,g\rangle=\int_0^1 t\cdot t^2\,\mathrm{d}t=\frac{1}{4}$$

$$\|f\|^2=\int_0^1 t^2\,\mathrm{d}t=\frac{1}{3}$$

$$\|g\|^2=\int_0^1 t^4=\frac{1}{5}$$

$$\cos\theta=\frac{1/4}{\sqrt{\frac{1}{15}}}=\frac{\sqrt{15}}{4}$$

Svar: $\theta=\arccos\left(\frac{\sqrt{15}}{4}\right)$.

FRÅGA: Vad händer med vinkeln mellan $f=t^n$ och $g=t^{n+1}$ om $n\rightarrow\infty$?