2006–04–06

Lay 7.3 Gränser för kvadratiska former, enkelt optimeringsproblem.

SATS: (6): Om A är reell och symmetrisk: $Q(\boldsymbol{x}) = \boldsymbol{x}^T A\, \boldsymbol{x}$ kvadratisk form.

$$\lambda_{\min} \leqslant \frac{\boldsymbol{x}^T A\, \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \leqslant \lambda_{\max}$$

där $\lambda_{\min}$ och $\lambda_{\max}$ är minsta respektive största egenvärde till A.

Bevis. Bygger på ortogonal diagonalisering.

$$Q(\boldsymbol{x}) = \boldsymbol{x}^T A\, \boldsymbol{x} = \boldsymbol{y}^T D\, \boldsymbol{y}, \quad \boldsymbol{y} = P^{-1}\boldsymbol{x} = P^T \boldsymbol{x}$$

$$\boldsymbol{x}^T \boldsymbol{x} = (P\boldsymbol{y})^T (P\boldsymbol{y}) = \boldsymbol{y}^T P^T P \boldsymbol{y} = [P \text{ ortogonal}] = \boldsymbol{y}^T \boldsymbol{y}$$

$$\text{Med } D = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$$

$$Q(P\boldsymbol{y}) = \boldsymbol{y}^T D\, \boldsymbol{y} = \lambda_1\, \boldsymbol{y}_1^2 + \lambda_2\, \boldsymbol{y}_2^2 + \cdots + \lambda_n\, \boldsymbol{y}_n^2 \geqslant \lambda_{\min}\,(y_1^2 + \cdots + y_n^2) = \lambda_{\min}\, \boldsymbol{y}^T \boldsymbol{y} = \lambda_{\min}\, \boldsymbol{x}^T \boldsymbol{x}$$

Därmed har vi visat att $\boldsymbol{x}^T A\, \boldsymbol{x} \geqslant \lambda_{\min}\, \boldsymbol{x}^T \boldsymbol{x}$. När får vi likhet i denna olikhet? Jo, om alla $y_i = 0$ för alla $\lambda_i > \lambda_{\min}$, dvs för

$$\boldsymbol{x} = P\boldsymbol{y} = \sum_{\lambda_i = \lambda_{\min}} y_i\, \boldsymbol{u}_i = \lambda_{\min}\, \boldsymbol{x}$$

där $\boldsymbol{u}_i$ är en ortonormal egenvektor. Vi ser att $\boldsymbol{x}$ är egenvektor till $\lambda_{\min}$. På samma sätt visas den övre gränsen. Likhet om $\boldsymbol{x}$ är egenvektor till motsvarande $\lambda_{\max}$. □

EXEMPEL. Bestäm max och min för $Q(x) = 4\,x_1^2 + 4\,x_2^2 + 10\,x_3^2 + 8\,x_1\,x_3 + 8\,x_2\,x_3$ då $\|\boldsymbol{x}\| = 3$. För vilka $\boldsymbol{x}$ antas dessa max/min?

Lösning: $Q(\boldsymbol{x}) = \boldsymbol{x}^T A\, \boldsymbol{x}$ med

$$A = \begin{pmatrix} 4 & 2 & 4 \\ 2 & 4 & 4 \\ 4 & 4 & 10 \end{pmatrix}, \quad \text{egenvärden } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 14 \end{pmatrix}, \quad \text{egenvektorer } P = \begin{pmatrix} -2 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

Detta ger

$$2 \leqslant \frac{\boldsymbol{x}^T A\, \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \leqslant 14$$

$\boldsymbol{x}^T \boldsymbol{x} = 9$ ger

$$18 \leqslant \boldsymbol{x}^T A\, \boldsymbol{x} \leqslant 126$$

Maximum antas för

$$\boldsymbol{x} = t \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \text{ normerad så att } \|\boldsymbol{x}\| = 3$$

$$\boldsymbol{x} = \frac{\sqrt{6}}{2} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Minimum antas för (se beviset)

$$\boldsymbol{x} = y_1 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} + y_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2\,y_1 - y_2 \\ y_2 \\ y_1 \end{pmatrix} \quad \text{normerad så att } \|\boldsymbol{x}\| = 3$$

$$\|\boldsymbol{x}\|^2 = (-2\,y_1 - y_2)^2 + y_2^2 + y_1^2 \overset{!}{=} 3^2$$

dvs $\left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right)$ på ellipsen $5\,y_1^2 + 2\,y_2^2 + 4\,y_1\,y_2 = 9$.

||.

Betrakta optimeringsproblemet

(P_1) $\max \boldsymbol{x}^T A\,\boldsymbol{x}$ då $\boldsymbol{x}^T\boldsymbol{x} = 1$ och $\boldsymbol{x}^T\boldsymbol{u}_1 = 0$, där vi har ordnade egenvärden $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ med tillhörande normerade egenvektorer $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$. A är symmetrisk: $A = P\,D\,P^T$ (ortogonal diagonalisering).

$$D = \begin{pmatrix} \lambda_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \lambda_n \end{pmatrix}, \quad P = \begin{pmatrix} | & & | \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \\ | & & | \end{pmatrix}$$

SATS: (7): Lösning till (P_1) är $\boldsymbol{x} = \boldsymbol{u}_2$ och maximala värdet är $\boldsymbol{u}_2^T\,A\,\boldsymbol{u}_2 = \lambda_2$. $\boldsymbol{u}_2$ är den egenvektor som hör till näst största egenvärde.

(P_{k-1}) $\max \boldsymbol{x}^T A\,\boldsymbol{x}$ då $\boldsymbol{x}^T\boldsymbol{x} = 1, \boldsymbol{x}^T\boldsymbol{u}_1 = 0, \ldots, \boldsymbol{x}^T\boldsymbol{u}_{k-1} = 0$

SATS: (8): Lösning till (P_{k-1}) är $\boldsymbol{x} = \boldsymbol{u}_k$ och maxvärde $\boldsymbol{u}_k^T\,A\,\boldsymbol{u}_k = \lambda_k$.

Tentauppgift. Betrakta $Q(\boldsymbol{x}) = 7\,x_1^2 + 3\,x_2^2 + 3\,x_1\,x_2$

a) bestäm största värde på $Q(\boldsymbol{x})$ då $\boldsymbol{x}^T\boldsymbol{x} = 1$.

b) bestäm en vektor $\boldsymbol{u}$ med $\boldsymbol{u}^T\boldsymbol{u} = 1$ så att $Q(\boldsymbol{u})$blir maximal.

c) bestäm max $Q(\boldsymbol{x})$ under villkor $\boldsymbol{x}^T\boldsymbol{x} = 1, \boldsymbol{x}^T\boldsymbol{u} = 0$.

Lösning: $Q(\boldsymbol{x}) = \boldsymbol{x}^T A\,\boldsymbol{x}$ med

$$A = \begin{pmatrix} 7 & \frac{3}{2} \\ \frac{3}{2} & 3 \end{pmatrix}, \quad \text{egenvärden } D = \begin{pmatrix} 2.5 & 0 \\ 0 & 7.5 \end{pmatrix}, \quad \text{egenvektorer } P = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix}$$

a) $\max = \lambda_1 = \lambda_{\max} = 7.5$

b) $\boldsymbol{u} = \boldsymbol{u}_1 = \frac{1}{\sqrt{10}}\left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right)$

c) $\max = \lambda_2 = 2.5$ enligt sats 7.

"Ja, i och med detta är vi klara med Lays bok."

Potensmetoden är en numerisk metod för egenvärden/egenvektorer. Heath 4.5.1.

Från diskreta dynamiska system: $\boldsymbol{x}_{k+1} = A\,\boldsymbol{x}_k$, $\boldsymbol{x}_0$ start.

Vi vet att om $|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$ (notera: strikt olikhet i det första). Då har vi:

$$\boldsymbol{x}_k \longrightarrow c_1\,(\lambda_1)^k\,\boldsymbol{v}_1 \quad \text{då } k \to \infty$$

där $A\,\boldsymbol{v}_1 = \lambda_1\,\boldsymbol{v}_1$. Egenparet $(\lambda_1,\,\boldsymbol{v}_1)$ kallas dominerande. Om vi normerar i varje steg så får vi konvergens mot en normerad egenvektor $\boldsymbol{u}_1$.

Detta är potensmetoden.

– $\boldsymbol{x}_0$ är start

– $\boldsymbol{y}_{k+1} = A\,\boldsymbol{x}_k, \quad k = 0, 1, \ldots$

– $\boldsymbol{x}_{k+1} = \frac{\boldsymbol{y}_{k+1}}{\|\boldsymbol{y}_{k+1}\|}$ (för att undvika “overflow” i datorn).

Konvergens mot $\boldsymbol{u}_1$ (nästan alltid). Efter konvergens kan man beräkna λ_1 genom $\lambda_1 = \boldsymbol{u}_1^T\,A\,\boldsymbol{u}_1$, där $\boldsymbol{u}_1$ är en normerad egenvektor.

Detta är ett specialfall av Rayleigh-kvoten. Om $A\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$ så gäller $\boldsymbol{x}^T A\,\boldsymbol{x} = \lambda\,\boldsymbol{x}^T\boldsymbol{x} \Longrightarrow$

$$\lambda = \frac{\boldsymbol{x}^T A\,\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}$$

Singulärvärdesuppdelning (svd) Lay7.4 eller Heath 3.6

Inledning Diagonaisering av symmetrisk matris

$$A = P\,D\,P^T = \sum_{i=1}^{n} \lambda_i\,\boldsymbol{u}_i\,\boldsymbol{u}_i^T$$

Spektraluppdelning:

DEFINITION: De singulära värdena σ_1, ..., σ_n till en $m \times n$-matris är $\sigma_i = \sqrt{\lambda_i}$ där λ_i är egenvärden till $A^T A$.

Låt $r = \text{rang}(A)$. Då kan $A \in \mathbb{R}^{m \times n}$ faktoriseras:

$$A = U\,\Sigma\,V^T$$

⟨fig12⟩

$$= U_1\,\Sigma_r\,V_1^T = \sum_{i=1}^{r} \sigma_i\,\boldsymbol{u}_i\,\boldsymbol{v}_i^T$$

U och V är ortogonala, Σ_r är diagonal med $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r$ singulära värdena. Kolonnerna i U kallas vänstersingulära vektorer. Vidare gäller

$$A^T A = \left(V\,\Sigma^T U^T\right)\left(U\,\Sigma\,V^T\right) = V\,\Sigma^T\,\Sigma\,V^T$$

där

$$\Sigma^T\Sigma = \begin{pmatrix} \sigma_1^2 & & & & & & \mathbf{0} \\ & \sigma_2^2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_r^2 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ \mathbf{0} & & & & & & 0 \end{pmatrix}$$

diagonalisering av $A^T A$.

Relation till viktiga underrummen:

$$\mathrm{Col}(A) = \mathrm{Col}(U_1) \quad \text{ty } A = U_1\,\Sigma_r\,V_1^T$$

$$\mathrm{Nul}(A) = \mathrm{Col}(V_2) \quad \text{ty } A\,V_2 = \mathbf{0} \quad (A\,V = U\,\Sigma)$$

$$\mathrm{Col}(A^T) = \mathrm{Col}(V_1) \quad \text{ty } A^T = V_1\,\Sigma_r\,U_1^T$$

$$\mathrm{Nul}(A^T) = \mathrm{Col}(U_2) \quad \text{ty } A^T U_2 = \mathbf{0}$$

Minsta-kvadrat och SVD

Problem: minimera $\|A\,\boldsymbol{x} - \boldsymbol{b}\|_2$. Lösningen kan skrivas

$$\hat{\boldsymbol{x}} = V_1\,\Sigma_r^{-1}\,U_1^T\boldsymbol{b} = \sum_{i=1}^{r} \sigma_i^{-1}\,\boldsymbol{v}_i\,\boldsymbol{u}_i^T\boldsymbol{b}$$

Bevis.

$$A\,\hat{\boldsymbol{x}} = (U_1\,\Sigma_r\,V_1^T)\left(V_1\,\Sigma_r^{-1}\,U_1^T\boldsymbol{b}\right) = U_1\,U_1^T\boldsymbol{b}$$

dvs $A\,\hat{\boldsymbol{x}}$ är bästa approximation av $\boldsymbol{b}$ i $\mathrm{Col}(A) = \mathrm{Col}(U_1)$. $U_1\,U_1^T\boldsymbol{b}$ är projektionen av $\boldsymbol{b}$ på underrummet. $\square$

Överbestämt ekvationssystem $A\,\boldsymbol{x} = \boldsymbol{b}$ i minsta-kvadrat-mening löses genom

$$\hat{\boldsymbol{x}} = V_1\,\Sigma_r^{-1}\,U_1^T\boldsymbol{b}$$

$V_1\,\Sigma_r^{-1}\,U_1^T$ kallas för en *pseudoinvers*. Betecknas A^+ eller $A^\dagger$. $\hat{\boldsymbol{x}} = A^\dagger\,\boldsymbol{b}$.