Numerisk Analys, MMG410. Övningar: minstakvadratproblem

Övning

Vi vill lösa minstakvadratproblemet

$$\min_x ||Ax - b||_2^2$$

då A har ortogonala kolumner ($a_j^T a_k = 0$ då $j \neq k$). Hur förenklar denna egenskap hos A lösandet av problemet?

Svar:

Om A har ortogonala kolonner så är $A^T A = D$, där $D = diag(a_1^T a_1, a_2^T a_2, ..., a_n^T a_n)$. Antag att ingen kolonn har längden noll, då är matrisen D ickesingulär med invers $D^{-1} = diag(1/||a_1||_2^2, 1/||a_2||_2^2, ..., 1/||a_n||_2^2)$. Lösningen till minstakvadratproblemet kan skrivas (normalekvationerna) $x = (A^T A)^{-1} A^T b$ (ty om A har ortogonala kolonner och inga nollkolonner så måste den ha full kolonnrang. Varför?) Så, $x = D^{-1} A^T b$ eller $x_k = a_k^T b / ||a_k||_2^2$.

Övning

Vi vill anpassa mätpunkter (t_k, N_k) (alla $N_k > 0$) till funktionen

$$N(t) = N_0 \exp^{-\lambda t} .$$

Gör en lämplig omskriving av problemet så att parametrarna, N_0 och λ, i modellen kan bestämmas med hjälp av ett (linjärt) minstakvadratproblem.

Svar:

Problemet är att N_0 och λ ej ingår linjärt i modellen varför vi inte kan använda minx $\min_x ||Ax - b||_2^2$ direkt. Idé: logaritmera

$$\log N(t) = \log N_0 - \lambda t.$$

$x_1 := \log N_0, x_2 := \lambda, b := \log N(t)$, då

$$b = x_1 - x_2 t.$$

Då får vi

$$\min_x \left\| \begin{bmatrix} 1 & -t_1 \\ 1 & -t_2 \\ \vdots & \vdots \\ 1 & -t_m \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \log N(t_1) \\ \log N(t_2) \\ \vdots \\ \log N(t_m) \end{bmatrix} \right\|$$

När x är beräknad sätter vi $N_0 = e^{x_1}$ och $\lambda = x_2$.

Övning

Givet mätpunkterna

$$(t_k, b_k) = (-n, 0), (-(n-1), 0), ..., (-1, 0), (0, 1), (1, 0), (2, 0), ..., (n, 0)$$

anpassa en rät linje till punkterna i tvånorm. (Alla b_k-värden är 0 förutom då $t_k = 0$ när $b_k = 1$.)
Svar:
En rät linje har ekvation: $f(t) = x_1 + x_2 t$. A och b har $2n+1$ rader och ser ut som:

$$A = \begin{bmatrix} 1 & -n \\ 1 & -(n-1) \\ \vdots & \vdots \\ 1 & -1 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & n \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}.$$

På rad $n+1$ står 1 i vektor b. Vi noterar att vi kan utnyttja föregående uppgift eftersom A har ortogonala kolonner.

Lösningen blir alltså

$$x = (A^T A)^{-1} A^T b = \begin{bmatrix} 2n+1 & 0 \\ 0 & 2n(n+1)(2n+1)/6 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2n+1} \\ 0 \end{bmatrix}.$$

Den räta linjen ges alltså av $f(t) = x_1 + x_2 t = 1/(2n+1) + 0 \cdot t$ eller $f(t) = 1/(2n+1)$ (konstant oberoende av t).

Övning

Antag att $A \in \mathbb{R}^{m\times n}$ har rang n. Visa att A^TA är positivt definit.
Svar.
Vi kollar:

$$x^T(A^TA)x = (Ax)^T(Ax) = ||Ax||_2^2$$

Det sista uttrycket kan inte vara negativt (det är ju en norm). Så frågan är om det kan vara noll även om $x \neq 0$.

$$||Ax||_2 = 0 \rightarrow Ax = 0$$

vilket inte kan inträffa eftersom A har full kolonnrang enligt förutsättningarna.

Övning

Antag att $A \in \mathbb{R}^{n \times n}$ är både ortogonal och triangulär.

- a) Visa att A är diagonal.
- b) Vilka diagonalelement har A?

Svar.

- a) Antag att A är reell och undertriangulär. Eftersom A dessutom ar ortogonal gäller att $AA^T = I$. Innerprodukterna mellan första raden i A och kolonnerna i A^T blir

$$a_{1,1} \cdot a_{1,1}, a_{1,1} \cdot a_{2,1}, a_{1,1} \cdot a_{3,1}, \ldots, a_{1,1} \cdot a_{n,1}$$

eftersom första raden i A endast innehåller ett element skilt från noll (nämligen $a_{1,1}$). Men eftersom första raden i enhetsmatrisen har nollor utom i första positionen måste

$$a_{2,1} = a_{3,1} = \ldots a_{n,1} = 0.$$

Vi kan nu utnyttja induktion och upprepa resonemanget för andra kolonnen i A osv.

- b) Det måste gälla att $a_{k,k}^2 = 1$ så att $a_{k,k} = \pm 1$.

Example

n=2:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 \end{bmatrix} = \begin{bmatrix} a_{11}^2 & 0 \\ 0 & a_{22}^2 \end{bmatrix} = \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$$

Övning

Antag att den partitionerade matrisen nedan är ortogonal (A och C är kvadratiska). Visa att A och C måste vara ortogonala och att $B = 0$.

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

Övning

Antag att den partitionerade matrisen nedan är ortogonal (A och C är kvadratiska). Visa att A och C måste vara ortogonala och att $B = 0$.

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

Svar.
Detta är en form av generalisering av föregående övning. Vi kollar:

$$\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}^T \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

$$= \begin{bmatrix} A^T & 0 \\ B^T & C^T \end{bmatrix} \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \begin{bmatrix} A^T A & A^T B \\ B^T A & B^T B + C^T C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

Detta innebär att $A^T A = I$, så att A är ortogonal. $A^T B = 0$ skall vara nollmatrisen vilket, eftersom A är ortogonal och därmed ickesingulär, medför att $B = 0$. Detta medför slutligen att C är ortogonal, ty $B^T B + C^T C = I$.

Övning

Antag att x och y löser problemen

$$\min_x \|Ax - b\|_2^2 \text{ resp. } \min_y \|(A+E)y - b\|_2^2$$

y är alltså lösningen till ett stört problem. Studera minstakvadratproblemet $\min_x \|Ax - b\|_2^2$ ur störningssynpunkt då

$$\text{A} = \begin{bmatrix} 1 & 0 \\ 0 & \delta \\ 0 & 0 \end{bmatrix}, 0 < \delta << 1, b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, b_3 \neq 0.\text{och då vi stör } A$$

med E

$$E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \varepsilon \end{bmatrix}, 0 < \varepsilon << \delta << 1.$$

Svar. Vi sätter upp normalekvationerna $A^T Ax = A^T b$ och får

$$\mathrm{A}^T Ax = \begin{bmatrix} 1 & 0 \\ 0 & \delta^2 \end{bmatrix} x = A^T b = \begin{bmatrix} b_1 \\ \delta b_2 \end{bmatrix} \rightarrow x = (A^T A)^{-1} A^T b =$$
$$\begin{bmatrix} b_1 \\ b_2/\delta \end{bmatrix}$$

Vi sätter nu upp normalekvationerna fÖr det störda problemet, dvs. $(A+E)^T(A+E)y = (A+E)^T b$.

$$(A+E)^T(A+E)y = \begin{bmatrix} 1 & 0 \\ 0 & \delta^2+\varepsilon^2 \end{bmatrix} y = (A+E)^T b = \begin{bmatrix} b_1 \\ \delta b_2 + \varepsilon b_3 \end{bmatrix} \rightarrow$$

$$\mathrm{y} = (\mathrm{A} + \mathrm{E})^T (A+E)^{-1}(A+E)^T b = \begin{bmatrix} b_1 \\ \frac{\delta b_2 + \varepsilon b_3}{\delta^2+\varepsilon^2} \end{bmatrix} \approx \begin{bmatrix} b_1 \\ \frac{\delta b_2 + \varepsilon b_3}{\delta^2} \end{bmatrix}$$

Så

$$y - x \approx \frac{\varepsilon}{\delta^2} \begin{bmatrix} 0 \\ b_3 \end{bmatrix} = \|E\|_2 k_2^2(A) \begin{bmatrix} 0 \\ b_3 \end{bmatrix}$$

med $k_2(A) = \frac{1}{\delta}$.

Övning

Vi har modellen $e^{\alpha t} \approx b$ och vill bestämma α. Formulera ett ickelinjärt och ett linjärt minstakvadratproblem. Lös problemen exakt då antalet observationer är $m = 2$ och där $t_1 = 1, t_2 = 2$ samt $b_2 = 1/2$. (Den optimala lösningen beror alltså av b_1.)

Övning

Vi har modellen $e^{\alpha t} \approx b$ och vill bestämma α. Formulera ett ickelinjärt och ett linjärt minstakvadratproblem. Lös problemen exakt då antalet observationer är $m = 2$ och där $t_1 = 1, t_2 = 2$ samt $b_2 = 1/2$. (Den optimala lösningen beror alltså av b_1.)

Svar.

Ickelinjärt minstakvadratproblem:

$$\min_{\alpha} \|e^{\alpha t} - b\|_2^2$$

Linjärt minstakvadratproblem:

$$\min_{\alpha} \|\alpha t - \log(b)\|_2^2$$

För antalet observationer $m = 2$: Ickelinjärt minstakvadratproblem:

$$\min_{\alpha} (e^{\alpha t_1} - b_1)^2 + (e^{\alpha t_2} - b_2)^2$$

Linjärt minstakvadratproblem:

$$\min_{\alpha} (\alpha t_1 - \log(b_1))^2 + (\alpha t_2 - \log(b_2))^2$$

Först för antalet observationer $m = 2$ vi l¨¨¨oser ickelinjärt minstakvadratproblem. Vi vill lösa:

$$((e^{\alpha t_1} - b_1)^2 + (e^{\alpha t_2} - b_2)^2)'_\alpha = 0$$

och deriverar ickelinjärt minstakvadratproblem:

$$((e^{\alpha t_1} - b_1)^2 + (e^{\alpha t_2} - b_2)^2)'_\alpha = 2t_1 e^{\alpha t_1}(e^{\alpha t_1} - b_1) + 2t_2 e^{\alpha t_2}(e^{\alpha t_2} - b_2).$$

Då $t_1 = 1, t_2 = 2$ och $b_2 = 1/2$ vi får ekvationen:

$$0 = e^{\alpha}(e^{\alpha} - b_1) + 2e^{\alpha 2}(e^{\alpha 2} - 1/2) = e^{2\alpha} - b_1 e^{\alpha} + 2e^{4\alpha} - e^{2\alpha} = 2e^{4\alpha} - b_1 e^{\alpha}.$$

och $2e^{4\alpha} = b_1 e^{\alpha}$, logaritmerar med $\log_e$:

$$\log(2) + 4\alpha = \log(b_1) + \alpha,$$

$$4\alpha - \alpha = \log(b_1) - \log(2),$$

$$\alpha = \frac{\log(b_1) - \log(2)}{3}$$

Nu för antalet observationer $m = 2$ vi löser linjärt minstakvadratproblem. Vi vill lösa:

$$((\alpha t_1 - \log(b_1))^2 + (\alpha t_2 - \log(b_2))^2)'_\alpha = 0$$

och deriverar linjärt minstakvadratproblem:

$$((\alpha t_1 - \log(b_1))^2 + (\alpha t_2 - \log(b_2))^2)'_\alpha = 2t_1(\alpha t_1 - \log(b_1)) + 2t_2(\alpha t_2 - \log(b_2))$$

Då $t_1 = 1, t_2 = 2$ och $b_2 = 1/2$: vi får ekvationen:

$$0 = \alpha - \log(b_1) + 2(2\alpha - \log(b_2)) = 5\alpha - (log(b_1) + 2\log(b_2))$$

och $5\alpha = log(b_1) + 2\log(b_2)$,

$$\alpha = \frac{log(b_1) + 2\log(b_2)}{5} = \frac{log(b_1) + 2\log(b_2)}{5}.$$