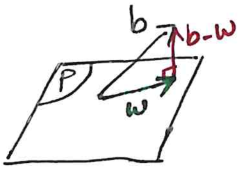
Kapitel 6 - Minstakvadrat-metoden

Låt A vara en $m\times n$-matris och b en vektor i $\mathbb{R}^m$. Antag att det linjära systemet $Ax=b$ är olösbart där x är en okänd vektor i $\mathbb{R}^n$. Man vill bestamma en vektor x där det euklidiska avståndet $\|b-Ax\|$ är minimalt. I så fall kallas x minsta kvadratlösning, $b-Ax$ minstakvadrat felvektor och $\|b-Ax\|$ minstakvadratfel.

Vi illustrerar problemet i $\mathbb{R}^3$.

Låt A vara en $3\times n$-matris och $b\in\mathbb{R}^3$. Antag att kolumner av A bestämmer ett plan P. Eftersom $Ax=b$ är olösbar, tillhör b inte planet.



Vi vill approximera $\vec{b}$ med en vektor w som ligger i planet sådana att felvektorn $b-w$ har minsta normen $\|b-w\|$.

$\|b-w\|$ är minimalt när $w=\text{proj}_P b$.

Eftersom w är i planet P så är w i span(kolumner av A) så det finns $\hat{x}$ i $\mathbb{R}^3$ så att $A\hat{x}=w=\text{proj}_P b$.

Med hjälp av matematiska resultat, kan ekvationen $Ax=\text{proj}_P b$ skrivas som $A^TAx=A^Tb$. $A^TAx=A^Tb$ kallas normal ekvation och är alltid lösbar.

<u>Sats:</u> Mängden av minstakvadrat lösningar till $Ax=b$ är lika med mängden av alla lösningar till ekvationen $A^TAx=A^Tb$.

Ex: Bestäm minstakvadratlösningar till följande linjärt system:

$$\begin{cases} x_1 - x_2 = 4 \\ 3x_1 + 2x_2 = 1 \\ -2x_1 + 4x_2 = 3 \end{cases}$$

Lösning: Matrisekvation:

$$\underbrace{\begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}}_{b}$$

Normal ekvation: $A^T A x = A^T b$

$$A^T = \begin{pmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{pmatrix} \Rightarrow A^T A = \begin{pmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} 14 & -3 \\ -3 & 21 \end{pmatrix}$$

och $A^T b = \begin{pmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \end{pmatrix}$

normal ekvation: $\begin{pmatrix} 14 & -3 \\ -3 & 21 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \end{pmatrix}$

$\det(A^T A) = 14 \cdot 21 - 9 = 285 \neq 0$ så är $A^T A$ inverterbar.

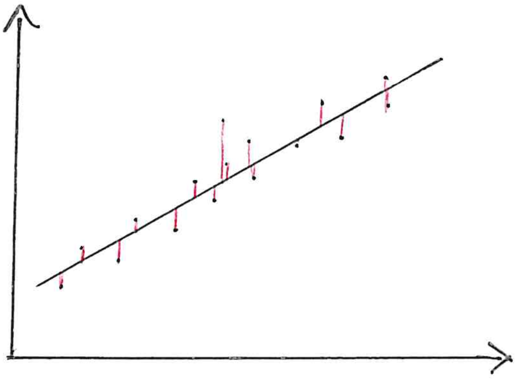
Alltså är lösningen till normal ekvationen: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\frac{1}{285} \begin{pmatrix} 21 & 3 \\ 3 & 14 \end{pmatrix}}_{(A^T A)^{-1}} \underbrace{\begin{pmatrix} 1 \\ 10 \end{pmatrix}}_{b} = \frac{1}{285} \begin{pmatrix} 51 \\ 143 \end{pmatrix}$

Sats: Låt A vara en $m \times n$-matris. Följande påstående är ekvivalenta

1. A har linjärt oberoende kolumner.
2. $A^T A$ är inverterbar.
3. $Ax = b$ har en unik lösning för alla $b \in \mathbb{R}^n$.

Om påståendena stämmer så är lösningen $\hat{x} = (A^T A)^{-1} A^T b$.

Antag att vi har mätdata (x_i, y_i) och vi vill anpassa en rät linje $y = mx + k$ genom alla punkter (x_i, y_i) i planet.



antag att vi har n punkter $(x_1, y_1), \ldots, (x_n, y_n)$. Om punkterna ligger på linjen $y = kx + m$ så är $y_i = kx_i + m$ för alla (x_i, y_i).

Alltså (m, k) är lösning till systemet:

$$\begin{cases} y_1 = kx_1 + m \\ y_2 = kx_2 + m \\ \vdots \\ y_n = kx_n + m \end{cases} \iff \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}_{b} = \underbrace{\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}}_{A} \begin{pmatrix} k \\ m \end{pmatrix}$$

Men eftersom det finns ingen linjen som passar genom punkterna (x_i, y_i) i figuren ovanför så är ekvationsystemet olösbar.

Linjen med bästa uppskattning ges genom att lösa systemet med minsta kvadrat metoden. Dvs med minsta kvadrat metoden får man en linje så att summan av alla avstånden i kvadrat är minimal.

Eftersom kolumnerna av A är linjärt oberoende (alla x_i är olika) så systemet har en unik minstakvadratlösning $\hat{x} = (A^T A)^{-1} A^T b$.

Ex: Bestäm den räta linjen som i minsta-kvadratmening är bäst anpassad till följande data:

x	0	1	2	3
y	1	3	4	4

Låt $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}$ $\quad y = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 4 \end{pmatrix}$

$x = (A^T A)^{-1} A^T y$ där $x = \begin{pmatrix} k \\ m \end{pmatrix}$

$$A^T = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix} = A^T A$$

$$(A^T A)^{-1} = \frac{1}{14 \cdot 4 - 6 \cdot 6} \begin{pmatrix} 4 & -6 \\ -6 & 14 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 4 & -6 \\ -6 & 14 \end{pmatrix}$$

$(A^T A)^{-1} \cdot A^T$:

$$\frac{1}{20} \begin{pmatrix} 4 & -6 \\ -6 & 14 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} -6 & -2 & 2 & 6 \\ 14 & 8 & 2 & -4 \end{pmatrix} \cdot \frac{1}{20} \Rightarrow (A^T A)^{-1} A^T = \frac{1}{20} \begin{pmatrix} -6 & -2 & 2 & 6 \\ 14 & 8 & 2 & -4 \end{pmatrix}$$

$$(A^T A)^{-1} A^T \cdot y = \frac{1}{20} \begin{pmatrix} -6 & -2 & 2 & 6 \\ 14 & 8 & 2 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 4 \\ 4 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 20 \\ 30 \end{pmatrix} = \begin{pmatrix} 1 \\ 1{,}5 \end{pmatrix} \Rightarrow k = 1 \text{ och } m = 1{,}5$$

Alltså $y = x + 1{,}5$