Minsta -kvadratmetoden: vad handlar den om?

Minstakvadratmetoden

Ceres är den första asteroiden som upptäcktes av en slump på nyårsdagen 1801 av Giuseppe Piazzi. Han sökte efter en stjärna vid namn Mayer 87. Men istället för stjärnan fann Piazzi ett stjärnliknande objekt som han först trodde var en komet.

Han fortsatte att studera asteroidens bana, men den 11 februari hindrades han att fortsätta på grund av sjukdom. Han hade då inte hunnit berättat för någon om sin upptäckt och kort därefter skymd solen dess bana.

För att återfinna asteroiden utvecklade Carl Friedrich Gauss **minstakvadratmetoden**, och utifrån Piazzis 24 observationer, som täckte ungefär 9° av Ceres omloppsbana, och antagandet att banan var ett kägelsnitt tog han fram en trolig omloppsbana för Ceres.

Modellproblem. Betrakta nedanstående tabell:

x	1	2	3
y	2	5	6

En matrisformulering är

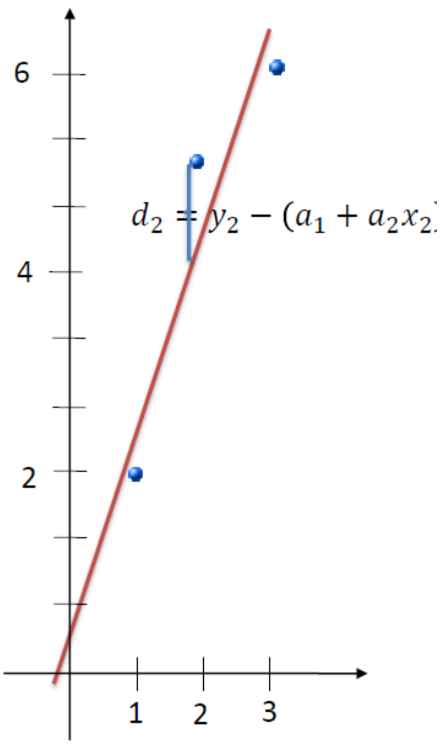
$$\begin{cases} y_1 = a_1 + a_2 x_1 \\ y_2 = a_1 + a_2 x_2 \\ y_3 = a_1 + a_2 x_3 \end{cases} \Leftrightarrow \begin{cases} 2 = a_1 + a_2 1 \\ 5 = a_1 + a_2 2 \\ 6 = a_1 + a_2 3 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}$$

Obs! att systemet är olösbart.

Istället: Bestäm den räta linjen $y = a_1 + a_2 x$ som i minstakvadrat-mening bäst ansluter till data, d v s bestäm a_1 och a_2 så att

$$\sum_{i=1}^{3} \left(y_i - (a_1 + a_2 x_i)\right)^2$$

blir så liten som möjligt.

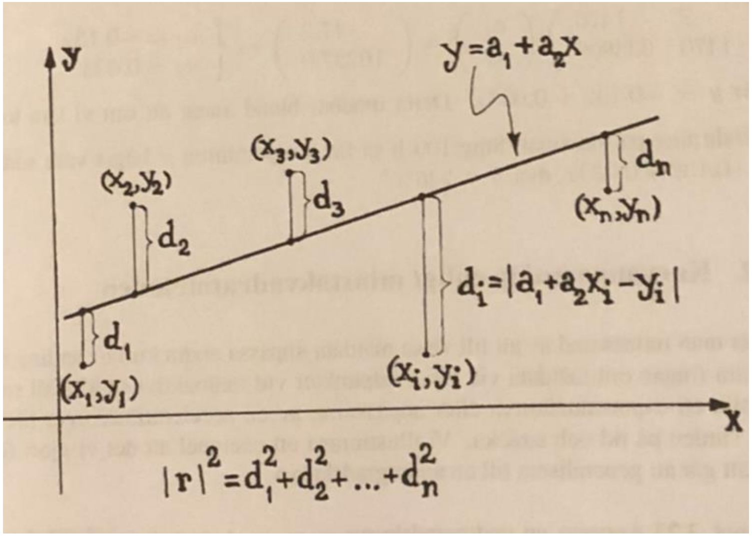


Minsta -kvadratmetoden: vad handlar den om?

Vi får alltså den linje som har egenskapen att summan av kvadraterna på avstånden d_i i y-led mellan mätpunkterna och motsvarande punkt på linje blir minimal.

Man brukar därför säga att man erhåller den linje som är bäst i **minstakvadratmening.**

I statistiken kallas den linje man får med minstakvadratmetoden för **regressionslinjen.**



Vi kommer att jobba här med ett givet (olösbart) linjärt system $A\vec{x}=\vec{b}$ av m ekvationer och n obekanta för att bestäma en vektor $\vec{x}$ som minimerar $\left|A\vec{x}-\vec{b}\right|$.

- Ett sådant $\vec{x}$ kallas en **minsta-kvadratlösning till systemet**
- $A\vec{x}-\vec{b}=\vec{r}$ kallas **residulavektorn i minsta-kvadratmening**
- $\left|A\vec{x}-\vec{b}\right|$ **minsta-kvadratfelet**

Minsta -kvadratmetoden: vad handlar den om?

Sats 3.2

Låt A vara en $m \times n$ - matris, $\vec{x}$ en $n \times 1$ - matris och $\vec{b}$ en $m \times 1$ - matris.

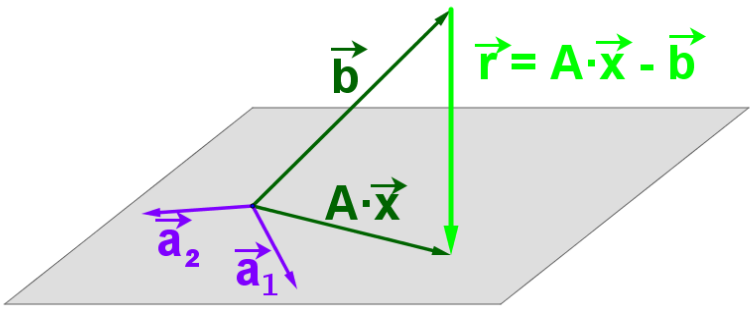
Minsta – kvadratlösningarna till systemet $A\vec{x} = \vec{b}$, dvs. de $\vec{x}$ för vilka residualvektorn $\vec{r} = A\vec{x} - \vec{b}$ blir så kort som möjligt $|\vec{r}| = |A\vec{x} - \vec{b}|$, fås genom att lösa **normalekvationerna**

$$A^t A\vec{x} = A^t\vec{b}$$

Exempel:

$$\underset{A}{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}} \cdot \underset{\vec{x}}{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}} = \underset{\vec{b}}{\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}}$$

$$x_1 \underset{\vec{a}_1}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} + x_2 \underset{\vec{a}_2}{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}} = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$$



Hur tar man normalekvationen fram ?

Vi vet att kortaste avståndet är vinkelrät mot planet, dvs.

- Vektorerna $\vec{a}_1$ och $\vec{a}_2$ spänner upp planet.
- $\vec{a}_1 \bullet \vec{r} = 0$ och $\vec{a}_2 \bullet \vec{r} = 0$ eller i matrisform $\vec{a}_1{}^t \vec{r} = 0 \Leftrightarrow [1 \quad 1 \quad 1] \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = [0]$ och

  $\vec{a}_2{}^t \bullet \vec{r} = 0 \Leftrightarrow [1 \quad 2 \quad 3] \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = [0] \Rightarrow A^t \vec{r} = 0$
- $A^t \vec{r} = [\vec{r} = A\vec{x} - \vec{b}] = A^t (A\vec{x} - \vec{b}) = A^t A\vec{x} - A^t \vec{b} = 0 \Rightarrow A^t A\vec{x} - A^t \vec{b} = 0 \Leftrightarrow A^t A\vec{x} = A^t \vec{b}$
- $A^t A\vec{x} = A^t \vec{b}$ kallas **normalekvation**

Minsta -kvadratmetoden: vad handlar den om?

Vi söker lösningen till $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$. Någon lösning till detta system finns inte (kontrollera). Systemet är överbestämt. Vi ska lösa ekvationen med hjälp av minstakvadratmetoden.

Vi ska lösa normalekvationen: $A^t A\vec{x} = A^t \vec{b} \Rightarrow$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 13 \\ 30 \end{bmatrix}$$

Totalmatrisen ger då

$$\left(\begin{array}{cc|c} 3 & 6 & 13 \\ 6 & 14 & 30 \end{array}\right) \sim [rad1 \cdot (-2)] \sim \left(\begin{array}{cc|c} -6 & -12 & -26 \\ 6 & 14 & 30 \end{array}\right) \sim \begin{bmatrix} \\ rad2 + rad1 \end{bmatrix} \sim \left(\begin{array}{cc|c} -6 & -12 & -26 \\ 0 & 2 & 4 \end{array}\right) \sim$$

$$\sim \begin{bmatrix} \\ rad2 \cdot 6 \end{bmatrix} \sim \left(\begin{array}{cc|c} -6 & -12 & -26 \\ 0 & 12 & 24 \end{array}\right) \begin{bmatrix} rad1 + rad2 \\ \end{bmatrix} \sim \left(\begin{array}{cc|c} -6 & 0 & -2 \\ 0 & 12 & 24 \end{array}\right) \sim \begin{bmatrix} rad1/(-6) \\ rad2/12 \end{bmatrix} \sim \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & 2 \end{array}\right)$$

$$\Rightarrow \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & 2 \end{array}\right) \Rightarrow \begin{cases} x_1 = \frac{1}{3} \\ x_2 = 2 \end{cases}$$

- Residualen : $\vec{r} = A\vec{x} - \vec{b} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{3} \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \frac{1}{3}\left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 6 \end{bmatrix} - \begin{bmatrix} 6 \\ 15 \\ 18 \end{bmatrix}\right) = \frac{1}{3}\left(\begin{bmatrix} 7 \\ 13 \\ 19 \end{bmatrix} - \begin{bmatrix} 6 \\ 15 \\ 18 \end{bmatrix}\right) = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

- Minsta – kvadratfelet : $|\vec{r}| = \sqrt{\left(\frac{1}{3}\right)^2 + \left(\frac{-2}{3}\right)^2 + \left(\frac{1}{3}\right)^2} = \frac{1}{3}\sqrt{6}$

- Kontroll: $\vec{r} \perp A\vec{x} \Rightarrow \vec{r} \bullet A\vec{x} = 0$, $A\vec{x} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{3} \\ 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 7 \\ 13 \\ 19 \end{bmatrix}$ och $\vec{r} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ ger att

  $\vec{r} \bullet A\vec{x} = \frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \bullet \left(\frac{1}{3}\begin{pmatrix} 7 \\ 13 \\ 19 \end{pmatrix}\right) = \frac{1}{9}(7 - 26 + 19) = 0$ alltså $\vec{r} \perp A\vec{x}$ och detta medför att vara beräkningar är korrekta

Minsta -kvadratmetoden: vad handlar den om?

Exempel: Vid ett försök har ett antal stäckor uppmätts vid några tidpunkter

t	1	2	3	4	5	6
s	1,1	3,2	6,9	12,8	21,3	31,0

Vi antar att vi kan anpassa mätdata m.h.a. andragradskurva $s = a + bt + ct^2$. Bestäm s som funktion av t, alltså bestäm $s(t)$.

Lösning:

Insättning av rspetkive data i $s = a + bt + ct^2$ ger oss följande ekvationssystem:

$$\begin{cases} 1{,}1 = a + b + c \\ 3{,}2 = a + 2b + 4c \\ 6{,}9 = a + 3b + 9c \\ 12{,}8 = a + 4b + 16c \\ 21{,}3 = a + 5b + 25c \\ 31{,}0 = a + 6b + 36c \end{cases} \Leftrightarrow \left[A \cdot \vec{x} = \vec{b}\right] \Leftrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1{,}1 \\ 3{,}2 \\ 6{,}9 \\ 12{,}8 \\ 21{,}3 \\ 31{,}0 \end{bmatrix}$$ alltså:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1{,}1 \\ 3{,}2 \\ 6{,}9 \\ 12{,}8 \\ 21{,}3 \\ 31{,}0 \end{bmatrix}$$

Normalekvationerna $A^t A\vec{x} = A^t \vec{b}$ ges då av

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 9 & 16 & 25 & 36 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 9 & 16 & 25 & 36 \end{bmatrix} \cdot \begin{bmatrix} 1{,}1 \\ 3{,}2 \\ 6{,}9 \\ 12{,}8 \\ 21{,}3 \\ 31{,}0 \end{bmatrix} \Leftrightarrow$$

$$\begin{bmatrix} 6 & 21 & 91 \\ 21 & 91 & 441 \\ 91 & 441 & 2275 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 73{,}6 \\ 371{,}9 \\ 1929{,}3 \end{bmatrix} \Leftrightarrow \ldots \Leftrightarrow \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1{,}28 \\ -1{,}15857 \\ 1{,}024 \end{bmatrix} \Rightarrow s = 1{,}28 - 1{,}15857 \cdot t + 1{,}024 \cdot t^2$$

svar: $s = 1{,}28 - 1{,}15857 \cdot t + 1{,}024 \cdot t^2$

Minsta -kvadratmetoden: vad handlar den om?

Exempel: (3.67 i boken) Ett radioaktivt preparat och följande intensiteter uppmätts vid följande tidpunkter:

t	1	2	3	4	5
I	188	147	139	130	100

Bestäm I_0 och λ i sambandet $I = I_0 \cdot e^{-\lambda \cdot t}$

Lösning:

Börja med att skriva om $I = I_0 \cdot e^{-\lambda \cdot t}$ på lijär form (använd logaritmlagarna)

$$I = I_0 \cdot e^{-\lambda \cdot t} \Leftrightarrow \left[I > 0, I_0 \cdot e^{-\lambda \cdot t} > 0\right] \Leftrightarrow \ln I = \ln\left(I_0 \cdot e^{-\lambda \cdot t}\right) \Leftrightarrow \ln I = \ln\left(I_0\right) + \ln\left(e^{-\lambda \cdot t}\right) \Leftrightarrow$$

$$\Leftrightarrow \ln I = \ln\left(I_0\right) - \lambda \cdot t \ln(e) \Leftrightarrow \ln I = \ln\left(I_0\right) - \lambda \cdot t$$

Insättning av respektive data i $\ln I = \ln\left(I_0\right) - \lambda \cdot t$ ger följande ekvationssystem:

$$\begin{cases} \ln 188 = \ln I_0 - \lambda \cdot 1 \\ \ln 147 = \ln I_0 - \lambda \cdot 2 \\ \ln 139 = \ln I_0 - \lambda \cdot 3 \\ \ln 130 = \ln I_0 - \lambda \cdot 4 \\ \ln 100 = \ln I_0 - \lambda \cdot 5 \end{cases} \Leftrightarrow \left[A \cdot \vec{x} = \vec{b}\right] \Leftrightarrow \begin{bmatrix} 1 & -1 \\ 1 & -2 \\ 1 & -3 \\ 1 & -4 \\ 1 & -5 \end{bmatrix} \cdot \begin{bmatrix} \ln I_0 \\ \lambda \end{bmatrix} = \begin{bmatrix} \ln 188 \\ \ln 147 \\ \ln 139 \\ \ln 130 \\ \ln 100 \end{bmatrix}$$ alltså:

$$A = \begin{bmatrix} 1 & -1 \\ 1 & -2 \\ 1 & -3 \\ 1 & -4 \\ 1 & -5 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} \ln I_0 \\ \lambda \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} \ln 188 \\ \ln 147 \\ \ln 139 \\ \ln 130 \\ \ln 100 \end{bmatrix}$$

Normalekvationerna $A^t A\vec{x} = A^t \vec{b}$ ges då av

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -2 & -3 & -4 & -5 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & -2 \\ 1 & -3 \\ 1 & -4 \\ 1 & -5 \end{bmatrix} \cdot \begin{bmatrix} \ln I_0 \\ \lambda \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -2 & -3 & -4 & -5 \end{bmatrix} \cdot \begin{bmatrix} \ln 188 \\ \ln 147 \\ \ln 139 \\ \ln 130 \\ \ln 100 \end{bmatrix}$$

Fortsätt lösa på egen hand, du kan behöva miniräknare eller dator ☺.

(Testa! Jmf med svaren i boken)

Minsta -kvadratmetoden: vad handlar den om?

Exempel: Bestäm den andragradskurva som bäst anpassar sig till punkterna $(-1,\ 1)$, $(0,0)$, $(1,1)$, $(2,8)$ i minsta – kvadratmening. (gammal tentamensuppgift)

"Minimalt" lösningsförslag:

Med $y = ax^2 + bx + c$ får vi följande fyra villkor

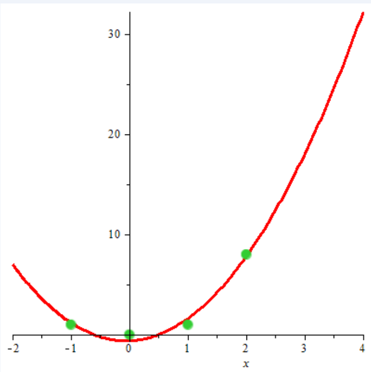
$$\begin{cases} a - b + c = 1 \\ c = 0 \\ a + b + c = 1 \\ 4a + 2b + c = 8 \end{cases}$$

I matrisform $$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 8 \end{bmatrix}$$

Vi får $A^t A = \begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix}$ och $A^t b = \begin{bmatrix} 34 \\ 16 \\ 10 \end{bmatrix}$

Normalekvationen $A^t Ax = A^t b$ ger ekvationssystemet $\begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 34 \\ 16 \\ 10 \end{bmatrix}$.

Detta ekvationssystem kan lösas med exempelvis Gausselimination. Det ger $a = 2$, $b = \frac{1}{5}$ och $c = \frac{-3}{5}$. Minsta kvadratlösningen är alltså $y = 2x^2 + \frac{1}{5}x - \frac{3}{5}$.



Svar : Minsta kvadratlösningen är alltså $y = 2x^2 + \frac{1}{5}x - \frac{3}{5}$.