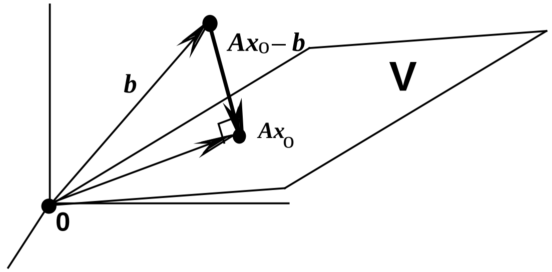
Institutionen för matematik
KTH

Minstakvadratmetoden

Komplettering till den linjära algebran
i kursen 5B1116



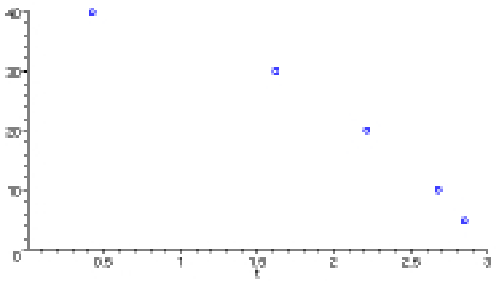
Eike Petermann/HT 2000

1. Minstakvadratmetoden

1.1. Ett exempel

Man ville bestämma ett approximativt värde på tyngdaccelerationen g: En sten slängdes från en hög byggnad och man noterade med hjälp av fotoceller placerade på höjderna 5, 10, 20, 30 och 40 meter över marken när stenen passerade dessa nivåer. Mätresultaten blev:

y, nivå i meter	40	30	20	10	5
t, tid i sek	0.42	1.62	2.22	2.67	2.85



Man bedömde att luftmotståndet inte hade någon nämnvärd inverkan på stenens rörelse varför man enligt Newtons lagar har sambandet

$$y = y_0 + v_0 t - gt^2/2$$

mellan stenens höjd över marken y [m] och falltiden t [s]. Konstanten y_0 är då stenens starthöjd, v_0 [m/s] dess initiala hastighet i lodled och g [m/s^2] den sökta tyngdaccelerationen. Dessa tre okända storheter borde då enligt mätresultaten, om dessa varit matematiskt exakta, uppfylla sambanden:

$$\begin{cases} y_0 + 0.42\, v_0 - 0.42^2/2\, g = 40 \\ y_0 + 1.62\, v_0 - 1.62^2/2\, g = 30 \\ y_0 + 2.22\, v_0 - 2.22^2/2\, g = 20 \\ y_0 + 2.67\, v_0 - 2.67^2/2\, g = 10 \\ y_0 + 2.85\, v_0 - 2.85^2/2\, g = \ \ 5 \end{cases} \qquad [1]$$

Nu är mätvärdena förstås inte exakta vilket gör det osannolikt att systemet, som ju har flera ekvationer än obekanta, har några lösningar. En kalkyl, som inte redovisas här, visar också att exempelvis de värden på y_0, v_0 och g som satisfierar de första tre sambanden inte satisfierar de två sista.

Rimligt är då att man istället letar efter värden på y_0, v_0 och g som insatta i ekvationerna ger vänsterleden värden som "så litet som möjligt" avviker från motsvarande högerled. Vad som menas med "så litet som möjligt" måste då förstås först preciseras. Flera tänkbara alternativ finns, till exempel att man bestämmer y_0, v_0 och g så att

- den största av skillnaderna mellan höger och vänster led till beloppet är så liten som möjligt,

eller

- summan av dessa skillnaders belopp är så liten som möjligt,

eller

- summan av skillnadernas kvadrater är så liten som möjligt.

Man har i alla dessa fall att lösa ett extremproblem. Vilket fall man väljer påverkar förstås väsentligt det räknearbete som behöver göras för att bestämma de "bästa" värdena på y_0, v_0 och g. Det visar sig att det sistnämnda av alternativen ger ett extremproblem som är förhållandevis lätt att lösa och vi granskar den metoden, den s.k. *minstakvadtratmetoden*, närmare här.

Litet terminologi: Man säger att man *minstakvadratanpassar* polynomet $y = y_0 + v_0 t - gt^2/2$ till de uppmätta värdena. Värdena y_0, v_0 och g kallas en *minstakvadratlösning* till systemet [1] (eller en *lösning* till [1] *i minstakvadratmening*).

Vi tillämpar först metoden på exemplet ovan, men för att de speciella koefficienterna inte skall skymma sikten för förfarandets allmängiltighet, formulerar vi om problemet i termer av lösning av linjära ekvationssystem med hjälp av matrisräkning:

1.2 Normalekvationen

På matrisform kan systemet [1] skrivas

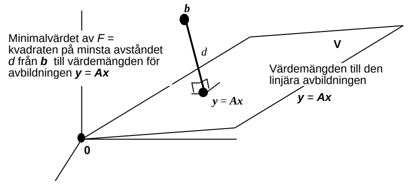
$$\boldsymbol{Ax} = \boldsymbol{b}$$

$$\text{där } \boldsymbol{A} = \begin{pmatrix} 1 & 0.42 & -0.42^2/2 \\ 1 & 1.62 & -1.62^2/2 \\ 1 & 2.22 & -2.22^2/2 \\ 1 & 2.67 & -2.67^2/2 \\ 1 & 2.85 & -2.85^2/2 \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} y_0 \\ v_0 \\ g \end{pmatrix} \text{ och } \boldsymbol{b} = \begin{pmatrix} 40 \\ 30 \\ 20 \\ 10 \\ 5 \end{pmatrix}$$

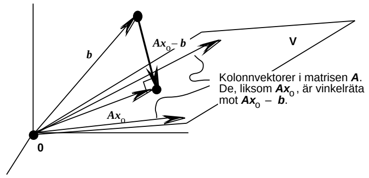
Vi vill minimera funktionen

$$F(\boldsymbol{x}) = |\boldsymbol{Ax} - \boldsymbol{b}|^2.$$

Ett analogt geometriskt problem antyds i figuren här bredvid: Det gäller att minimera avståndet från det "plan", som utgör värdemängden **V** till avbildningen $\boldsymbol{y} = \boldsymbol{Ax}$, till en punkt $\boldsymbol{b}$ som inte ligger i "planet".



För minimipunkten $\boldsymbol{x}_0$ skulle i så fall vektorn mellan $\boldsymbol{b}$ och den närmaste punkten $\boldsymbol{Ax}_0$ på värdemängden, vara vinkelrät mot varje vektor i denna värdemängd.



Eftersom **V** spänns upp av kolonnvektorerna i matrisen $\boldsymbol{A}$, så måste skalärprodukten av var och en av dem med $\boldsymbol{Ax} - \boldsymbol{b}$ vara $= 0$. Men kolonnerna i A är identiska med raderna i transponatet $\boldsymbol{A}^{\mathrm{T}}$, så det nyss sagda innebär att

$$\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{Ax} - \boldsymbol{b}) = \boldsymbol{0}, \text{ dvs}$$

$$\boxed{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}}$$

Minimipunkten $\boldsymbol{x}_0$ måste alltså satisfiera detta ekvationssystem – problemets så kallade *normalekvation*. För de fall där $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ är en inverterbar matris så gäller

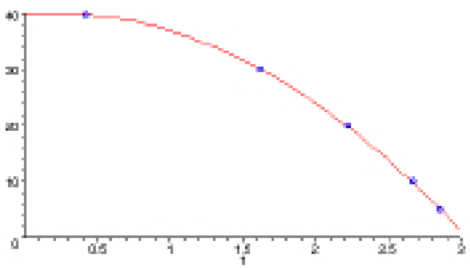
$$\boldsymbol{x}_0 = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}.$$

Vi leds alltså till påståendet:

Sats 1: (Om minstakvadratanpassning)
Om $\boldsymbol{x}_0$ är minipunkt till funktionen $F(\boldsymbol{x}) = |\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}|^2$, där $\boldsymbol{A}$ är en avbildning $\mathbf{R}^n \to \mathbf{R}^m$, så satisfierar $\boldsymbol{x}_0$ normalekvationen
$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_0 = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$$

I exemplet i avsnitt 1.1 är $\boldsymbol{A} = \begin{pmatrix} 1 & 0.42 & -0.42^2/2 \\ 1 & 1.62 & -1.62^2/2 \\ 1 & 2.22 & -2.22^2/2 \\ 1 & 2.67 & -2.67^2/2 \\ 1 & 2.85 & -2.85^2/2 \end{pmatrix}$ och $\boldsymbol{b} = \begin{pmatrix} 40 \\ 30 \\ 20 \\ 10 \\ 5 \end{pmatrix}$

Detta ger (räknehjälpmedel som räknedosa eller matematikprogram som MatLab, Maple, Matematica … underlättar!):
$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{pmatrix} 5 & 9.78 & -11.4903 \\ 9.78 & 22.9806 & -28.724976 \\ -11.4903 & -28.724976 & 37.00098508 \end{pmatrix} \text{ och } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} = \begin{pmatrix} 105 \\ 150.75 \\ -148.12875 \end{pmatrix},$$
Dvs normalekvationen är
$$\begin{pmatrix} 5 & 9.78 & -11.4903 \\ 9.78 & 22.9806 & -28.724976 \\ 1-1.4903 & -28.724976 & 37.00098508 \end{pmatrix}\begin{pmatrix} y_0 \\ v_0 \\ g \end{pmatrix} = \begin{pmatrix} 105 \\ 150.75 \\ -148.12875 \end{pmatrix},$$
med lösning:
$$\boldsymbol{x}_0 = \begin{pmatrix} y_0 \\ v_0 \\ g \end{pmatrix} = \begin{pmatrix} 5 & 9.78 & -11.4903 \\ 9.78 & 22.9806 & -28.724976 \\ -11.4903 & -28.724976 & 37.00098508 \end{pmatrix}^{-1}\begin{pmatrix} 105 \\ 150.75 \\ -148.12875 \end{pmatrix} = \begin{pmatrix} 40.0150857 \\ 1.98494802 \\ 9.963902255 \end{pmatrix}$$
vilket ger närmevärdet $g \approx 9.96$ m/s^2 för tyngdaccelerationen. Samtidigt avläser man att stenens höjd över marken i starten var ≈ 40.0 m och dess uppåtriktade hastighet i startögonblicket ≈ 1.98 m/s.
För stenens höjd som funktion av falltiden får man då sambandet:
$$y = 40 + 1.98t - 4.98t^2$$
Följande diagram visar grafen för denna funktion tillsammans med mätpunkterna:



Ett mått på anpassningens godhet är det så kallade *kvadratiska medelfelet*:
$$|\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}|/\sqrt{m},$$
där m är antalet ekvationer i systemet, vilket här kan beräknas till ≈ 0.16.

Anmärkning: Det kvadratiska medelfelet kan tolkas statistiskt. Om man antar att de mätfel man gjort, har orsakats av många små, sinsemellan oberoende störningar, så är sannolikheten för att komponenterna i $\boldsymbol{x}_0$ avviker från de ”korrekta” med mindre än medelfelet ungefär 68%. Mätningarna ovan skulle alltså innebära att det korrekta värdet av tyngdaccelerationen g med c:a 68%:s sannolikhet ligger i intervallet 9.96 ± 0.32 [m/s^2]

*1.3 Härledning av det allmänna fallet

Resonemanget ovan, som ledde till normalekvationen och sats 1 kan verka allmängiltigt, men man bör notera att det byggde på en figur där $\boldsymbol{b}$ är en punkt i $\mathbf{R}^3$, medan $\boldsymbol{b}$ i räkneexemplet ligger i $\mathbf{R}^5$, så påståendet är inte självklart riktigt. Generaliseringen till det helt allmänna fallet, då $\boldsymbol{b}$ ligger i $\mathbf{R}^m$ och $\boldsymbol{x}$ i $\mathbf{R}^n$ är ännu mindre självklar. I detta avsnitt visas att sats 1 *är* sann också i dessa fall. Eftersom det beviset inte kan bygga på någon figur, måste det göras med enbart algebraiska och/eller analytiska metoder:

1.3.1 Några hjälpsamma observationer

Först en variant av kvadreringsregeln:

Hjälpsats 1:
Om $\boldsymbol{u}$ och $\boldsymbol{v}$ är kolonnvektorer (dvs $n \times 1$-matriser)
$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ : \\ : \\ u_m \end{pmatrix} \text{ och } \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ : \\ : \\ v_m \end{pmatrix}$$
så är $|\boldsymbol{u} + \boldsymbol{v}|^2 = |\boldsymbol{u}|^2 + |\boldsymbol{v}|^2 + 2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$.

Bevis: Man har nämligen att $|\boldsymbol{u} + \boldsymbol{v}|^2 = (\boldsymbol{u} + \boldsymbol{v})^{\mathrm{T}}(\boldsymbol{u} + \boldsymbol{v}) = (\boldsymbol{u}^{\mathrm{T}} + \boldsymbol{v}^{\mathrm{T}})(\boldsymbol{u} + \boldsymbol{v}) =$
$= \boldsymbol{u}^{\mathrm{T}}\boldsymbol{u} + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} + \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u} + \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}$.
Men alla dessa produkter är skalärer (1×1-matriser) varför $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u} = (\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u})^{\mathrm{T}} = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ och $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u} = |\boldsymbol{u}|^2$, $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = |\boldsymbol{v}|^2$, alltså
$$|\boldsymbol{u} + \boldsymbol{v}|^2 = |\boldsymbol{u}|^2 + |\boldsymbol{v}|^2 + 2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}. \qquad ■$$

Sedan en observation om nollställena till $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$:

Hjälpsats 2:
$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ om och endast om $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$.

Påståendet är självklart sant om $\boldsymbol{A}^{\mathrm{T}}$ (och därmed även $\boldsymbol{A}$) är inverterbar, men poängen är att detta är riktigt för godtyckliga matriser $\boldsymbol{A}$ av godtyckligt format.

Bevis: Man har kedjan av implikationer:
$$\boldsymbol{A}\boldsymbol{x} = \mathbf{0} \qquad \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$$
$$|\boldsymbol{A}\boldsymbol{x}|^2 = 0 \qquad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = 0$$
dvs $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \mathbf{0} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} = \mathbf{0}$. ■

1.3.2 Härledningen

Vi visar följande precisering av satsen 1 ovan:

Sats 2: (Minsta-kvadrat-metoden)

För varje matris $\boldsymbol{A}$ och kolonnvektor $\boldsymbol{b}$ med samma antal rader som $\boldsymbol{A}$ har normalekvationen

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$$

(minst) en lösning.

För *varje* sådan lösning $\boldsymbol{x}_0$ gäller att

$$|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}|^2 \geq |\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}|^2,$$

dvs $\boldsymbol{x}_0$ är minimipunkt till funktionen

$$F(\boldsymbol{x}) = |\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}|^2.$$

Bevis:

1. Att normalekvationen alltid har lösningar kan inses genom ett indirekt resonemang: Anta att systemet inte har några lösningar. I så fall måste det finnas någon linjär kombination av raderna i matrisen $\boldsymbol{A}^T\boldsymbol{A}$ som är identiskt = $\boldsymbol{0}$,
   dvs $\boldsymbol{c}^T\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{0}$ för någon kolonnvektor $\boldsymbol{c}$,
   medan *samma* linjära kombination av raderna i högerledet $\boldsymbol{b}$ är $\neq 0$,
   dvs $\boldsymbol{c}^T\boldsymbol{A}^T\boldsymbol{b} \neq 0$. [1]
   Men eftersom $\boldsymbol{c}^T\boldsymbol{A}^T\boldsymbol{A} = (\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{c})^T = \boldsymbol{0}$, har vi enligt hjälpsatsen 2 att $\boldsymbol{A}\boldsymbol{c} = \boldsymbol{0}$ och därmed också $\boldsymbol{c}^T\boldsymbol{A}^T = \boldsymbol{0}$. Detta ger motsägelsen $\boldsymbol{c}^T\boldsymbol{A}^T\boldsymbol{b} = 0$ – antagandet att systemet inte har några lösningar är alltså felaktigt.
2. Låt nu $\boldsymbol{x}_0$ vara en godtycklig lösning till normalekvationen. Sätter vi $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{h}$, där $\boldsymbol{h}$ är en godtycklig vektor (av samma dimension som $\boldsymbol{x}$) så är

$$F(\boldsymbol{x}) = |\boldsymbol{A}(\boldsymbol{x}_0 + \boldsymbol{h}) - \boldsymbol{b}|^2 = |(\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}) + \boldsymbol{A}\boldsymbol{h}|^2 = \begin{bmatrix} \text{Sätt } \boldsymbol{u} = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b} \text{ och} \\ \boldsymbol{v} = \boldsymbol{A}\boldsymbol{h} \text{ i hjälpsatsen 1.} \end{bmatrix} =$$

$$= |\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}|^2 + |\boldsymbol{A}\boldsymbol{h}|^2 + 2\boldsymbol{h}^T\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}) = \begin{bmatrix} \text{Men } \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}) = \boldsymbol{0}. \end{bmatrix} =$$

$$= |\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}|^2 + |\boldsymbol{A}\boldsymbol{h}|^2 \geq |\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}|^2 = F(\boldsymbol{x}_0).$$ ■

1.4 När finns det bara en minimipunkt?

I räkneexemplet ovan är matrisen $\boldsymbol{A}^T\boldsymbol{A}$ inverterbar varför normalekvationen har en enda lösning. Ett rätt enkelt ekvivalent villkor, som dessutom i praktiken ofta är uppfyllt ges av:

[1] Om man exempelvis löser systemet med hjälp av Gausselimination – vilket inte är något annat än att man efter ett visst mönster linjärkombinerar ekvationerna i systemet till nya ekvationer – så måste förfarandet generera en ekvation av typen $\boldsymbol{0} \;\; \boldsymbol{x} = k \neq 0$, annars skulle ekvationssystemet $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ha en lösning.

Hjälpsats 3:

Den kvadratiska matrisen $\boldsymbol{A}^T\boldsymbol{A}$ är inverterbar om och endast om kolonnvektorerna i $\boldsymbol{A}$ är linjärt oberoende.

Bevis: Den kvadratiska matrisen $\boldsymbol{A}^T\boldsymbol{A}$ är inverterbar $\Leftrightarrow$ Ekvationssystemet $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ har enbart den triviala lösningen $\boldsymbol{x} = \boldsymbol{0}$ $\Leftrightarrow$ [Hjälpsats 2] $\Leftrightarrow$ Ekvationssystemet $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ har enbart den triviala lösningen $\boldsymbol{x} = \boldsymbol{0}$ $\Leftrightarrow$ Kolonnvektorerna i $\boldsymbol{A}$ är linjärt oberoende. ■

För minstakvadratmetoden innebär detta:

Sats 3: (Entydighet hos minimipunkten)

Funktionen $F(\boldsymbol{x}) = |\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}|^2$,
har en enda minimipunkt om och endast om kolonnvektorerna i $\boldsymbol{A}$ är linjärt oberoende.
Minimipunkten $\boldsymbol{x}_0$ ges av $\boldsymbol{x}_0 = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$.

En viktig typ av problem, där minimipunkten alltid är unik, är när man minstakvadratanpassar ett polynom av grad n

$$y = c_0 + c_1 t + c_2 t^2 \dots + c_n t^n$$

till mätdata $(t_1, y_1), (t_2, y_2), \dots, (t_m, y_m)$ vid *olika* ”tidpunkter” som är fler än gradtalet n, dvs $t_i \neq t_j$ om $i \neq j$ och $m > n$. Det inledande exemplet är av denna typ.

Sats 4: (Om polynomanpassning)

Kolonnvektorerna i matrisen

$$\boldsymbol{A} = \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^n \\ 1 & t_2 & t_2^2 & \dots & t_2^n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^n \end{pmatrix}$$

där $t_i \neq t_j$ om $i \neq j$ och $m > n$ är linjärt oberoende.

Bevis: Om $\boldsymbol{A}\boldsymbol{c} = \boldsymbol{0}$ för någon vektor $\boldsymbol{c} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ \vdots \\ c_n \end{pmatrix}$ så har polynomet

$$p(t) = c_0 + c_1 t + c_2 t^2 \dots + c_n t^n$$

de $m > n$ olika nollställena $t_1, t_2, \dots, t_m$. Eftersom inget polynom förutom nollpolynomet kan ha fler nollställen än sitt gradtal, så måste p vara nollpolynomet, dvs $\boldsymbol{c} = \boldsymbol{0}$. ■

1.5 Övningar

1.1 Skriv upp normalekvationerna till systemen

a. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ b. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

c. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ d. $(1, 2, 3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 1$

e. $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} x = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ f. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

g. $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ h. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$

1.2 För vilka av systemen i uppgift 1.1 finns det bara en enda minstakvadratlösning?

1.3 Ange en minstakvadratlösning för var och en av systemen i uppgift 1.1.

1.4 a. Punkterna $(x, y, z) = (1, 3, 5)t + (2, 4, 6)s$, bildar ett plan genom origo. Vilken är den vinkelräta projektionen av punkten (1, 0, 1) på detta plan?

b. Punkterna $(x, y, z) = (1, 2, 3)t$, bildar en rät linje genom origo. Vilken är den vinkelräta projektionen av punkten (1, 0, 1) på denna linje?

1.5 Låt $\boldsymbol{u}$ och $\boldsymbol{v}$ vara två vektorer i $\mathbf{R}^3$ och låt $\mathbf{L}$ vara den mängd punkter vars ortsvektorer är linjära kombinationer av $\boldsymbol{u}$ och $\boldsymbol{v}$. ($\mathbf{L}$ är ett plan om $\boldsymbol{u}$ och $\boldsymbol{v}$ inte är parallella och en linje om de är parallella och ej båda $\neq \mathbf{0}$.) Låt vidare $\boldsymbol{A}$ vara matrisen som har $\boldsymbol{u}$ och $\boldsymbol{v}$ som kolonner.

Verifiera dels att den på $\mathbf{L}$ belägna punkt $\boldsymbol{y}_0$, som ligger närmast punkten $\boldsymbol{b}$ ges av

$$\boldsymbol{y}_0 = \boldsymbol{A}\boldsymbol{x}_0, \text{ där } \boldsymbol{x}_0 \text{ är någon lösning till } \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$$

och dels att det minimala avståndet $= \sqrt{|\boldsymbol{b}|^2 - |\boldsymbol{y}_0|^2}$.

1.6 a. Vilken av punkterna på planet i uppgift 1.4a ligger närmast punkten (1, 0, 1) och vilket är det minimala avståndet?

b. Vilken av punkterna på den räta linjen i uppgift 1.4b ligger närmast punkten (1, 0, 1) och vilket är det minimala avståndet?

1.7 a. Visa att raderna i matrisen $\boldsymbol{A}$ är linjärt oberoende om och endast om matrisen $\boldsymbol{A}\boldsymbol{A}^T$ är inverterbar.

b. Visa att om raderna i matrisen $\boldsymbol{A}$ är linjärt oberoende så har normalekvationen $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ samma lösningar som det ursprungliga systemet $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

1.8 Bestäm ekvationen för den räta linje $y = kx + l$ som i minstakvadratmening bäst anpassar till punkterna (0,0), (2, 1) och (3, 8). Bestäm också medelfelet.

1.9 Bestäm ekvationen för den räta linje $y = kx + l$ som i minstakvadratmening bäst anpassar till punkterna (0, 0), (1, 1), (2, 1) och (3, 8). Bestäm också medelfelet.

1.10 Bestäm ekvationen för den parabel $y = ax^2 + bx + c$ som i minstakvadratmening bäst anpassar till punkterna (0, 0), (1, 1), (2, 1) och (3, 8). Bestäm också medel-zxxxfelet.

1.11 Bestäm ekvationen för det plan $z = ax + by + c$ som i minstakvadratmening bäst anpassar till punkterna (0, 0, 0), (1, 1, 1), (1, 2, 3) och (2, 3, 6). Bestäm också medelfelet.

1.12 Den verksamma substansen, theophyllin, hos ett läkemedel mot astma försvinner enligt en matematisk modell ur kroppen enligt sambandet

$$c(t) = Ae^{-kt},$$

där $c(t)$ och A är koncentrationerna av theophyllinet i blodet vid tiderna t och 0 och k är en konstant, som är relaterad till individens utsöndringsförmåga.

I syfte att kunna planera doceringen hos läkemedlet – koncentrationen av theophyllin måste hela tiden ligga i ett visst intervall för att läkemedlet skall ha avsedd effekt – vill man för en viss person empiriskt bestämma konstanten k. Läkemedlet injicerades i pacienten och man mätte sedan theophyllinkoncentrationerna i blodet vid olika tidpunkter. Mätningarna sammanfattas i tabellen här bredvid:

Antal timmar	Koncentration [ma/l]
1	10.0
5	5.0
9	2.5
13	1.5
15	1.0
19	0.5

Relationen ovan kan efter logaritmering skrivas

$$\ln c = \ln A - kt$$

Minstakvadratanpassa konstanterna $\ln A$ och k till de givna mätvärdena.

Svar:

1.1 a. $\begin{pmatrix} 35 & 44 \\ 44 & 56 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}$ b. $\begin{pmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

c. $\begin{pmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ d. $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

e. $14x = 4$ f. $\begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

g. $\begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ h. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

1.2 De där kolonnerna är linjärt oberoende, dvs. a, e f och h.

1.3 a. $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2/3 \\ 2/3 \end{pmatrix}$

b. T.ex. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -5/3 \\ 4/3 \\ 0 \end{pmatrix}$, (allmän lösn. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -5/3 + t \\ 4/3 - 2t \\ t \end{pmatrix}$).

c. T.ex. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -5/3 \\ 4/3 \\ 0 \end{pmatrix}$, (allmän lösn. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -5/3 + t \\ 4/3 - 2t \\ t \end{pmatrix}$).

d. T.ex. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, (allmän lösn. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 - 2s - 3t \\ s \\ t \end{pmatrix}$).

e. $x = 2/7$.

f. $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 3/2 \end{pmatrix}$ (Obs att matrisen för systemet i 1f är inverterbar, så normalekvationen har samma lösning som det givna systemet).

g. T.ex. $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1/5 \\ 0 \end{pmatrix}$ (allmän lösn. $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1/5 - 2t \\ t \end{pmatrix}$).

h. $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

1.4 *Ledning*: Lämpligt att använda resultaten från uppgifterna 1.3.a och e.
a. (2/3, 2/3, 2/3).
b. (2/7, 4/7, 6/7).

1.6 *Ledning*: Lämpligt att använda resultaten från uppgifterna 1.3.a och e.
a. (2/3, 2/3, 2/3), minimalavståndet $= \sqrt{2/3}$.
b. (2/7, 4/7, 6/7), minimalavståndet $= \sqrt{6/7}$.

1.7 *Ledningar*:
a. Tillämpa hjälpsats 1.3 på $\boldsymbol{A}$: s transponat.
b. Multiplicera relationen $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ med $\boldsymbol{A}$ från vänster och tillämpa resultatet från a-uppgiften.

1.8 $y = (33x - 13)/14$. Kvadratiska medelfelet $= 13/\sqrt{42} \approx 2.01$.
(Normalekvationen är $\begin{pmatrix} 13 & 5 \\ 5 & 3 \end{pmatrix}\begin{pmatrix} k \\ l \end{pmatrix} = \begin{pmatrix} 26 \\ 9 \end{pmatrix}$)

1.9 $y = (24x - 11)/10$. Kvadratiska medelfelet $= \sqrt{61/20} \approx 1.75$.
(Normalekvationen är $\begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix}\begin{pmatrix} k \\ l \end{pmatrix} = \begin{pmatrix} 27 \\ 10 \end{pmatrix}$)

1.10 $y = 3x^2/2 - 21x/10 + 2/5$. Kvadratiska medelfelet $= 2/\sqrt{5} \approx 0.89$.
(Normalekvationen är $\begin{pmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 4 & 6 & 4 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 77 \\ 27 \\ 10 \end{pmatrix}$.)

1.11 $z = 2y - 1/2$. Kvadratiska medelfelet = 1/2.
(Normalekvationen är $\begin{pmatrix} 6 & 9 & 4 \\ 9 & 14 & 6 \\ 14 & 6 & 4 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 16 \\ 25 \\ 10 \end{pmatrix}$.)

1.12 $\ln A \approx 2.448$ och $k \approx 0.1637$.
(Normalekvationen är $\begin{pmatrix} 6 & -62 \\ -62 & 862 \end{pmatrix}\begin{pmatrix} \ln A \\ k \end{pmatrix} = \begin{pmatrix} 4.54063 \\ -10.6976 \end{pmatrix}$)