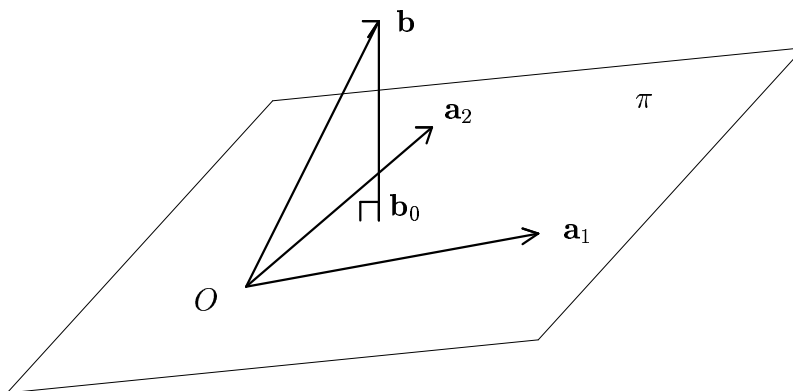


Överbestämda ekvationssystem

Antag att vi har 3 ekvationer och 2 obekanta, så $A\mathbf{x} = \mathbf{b}$ där $A = (\mathbf{a}_1 \ \mathbf{a}_2)$ är en 3×2 -matris. Då är

$$A\mathbf{x} = (\mathbf{a}_1 \ \mathbf{a}_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$$

en linjärkombination av kolonnerna \mathbf{a}_1 och \mathbf{a}_2 i A . Så $A\mathbf{x} = \mathbf{b}$ har lösning om och endast om \mathbf{b} är en linjärkombination av kolonnerna i A . Alla linjärkombinationer av \mathbf{a}_1 och \mathbf{a}_2 är ett plan π i \mathbb{R}^3 .



Antag nu att vi har en vektor \mathbf{b} som inte ligger i planet π och låt \mathbf{b}_0 vara ortogonala projektionen av \mathbf{b} på π . Då är en vektor \mathbf{x}_0 sådan att $A\mathbf{x}_0 = \mathbf{b}_0$ så nära en lösning till $A\mathbf{x} = \mathbf{b}$ man komma i den meningen att

$$|A\mathbf{x}_0 - \mathbf{b}| \leq |A\mathbf{x} - \mathbf{b}|$$

för alla $\mathbf{x} \in \mathbb{R}^2$.

Hur bestämmer man \mathbf{x}_0 ? Sätt $\mathbf{r} = A\mathbf{x}_0 - \mathbf{b}$. Då är \mathbf{r} ortogonal mot varje vektor i π och speciellt mot \mathbf{a}_1 och \mathbf{a}_2 . Det ger

$$\begin{cases} \mathbf{a}_1^t \mathbf{r} = 0 \\ \mathbf{a}_2^t \mathbf{r} = 0 \end{cases} \text{ d v s } A^t \mathbf{r} = \mathbf{0}.$$

Men

$$0 = A^t \mathbf{r} = A^t (A\mathbf{x}_0 - \mathbf{b}) = A^t A\mathbf{x}_0 - A^t \mathbf{b},$$

d v s $A^t A\mathbf{x}_0 = A^t \mathbf{b}$, så \mathbf{x}_0 är en lösning till $A^t A\mathbf{x} = A^t \mathbf{b}$.

Detta kan (ganska enkelt) bevisas i det allmänna fallet (se boken) och vi har:

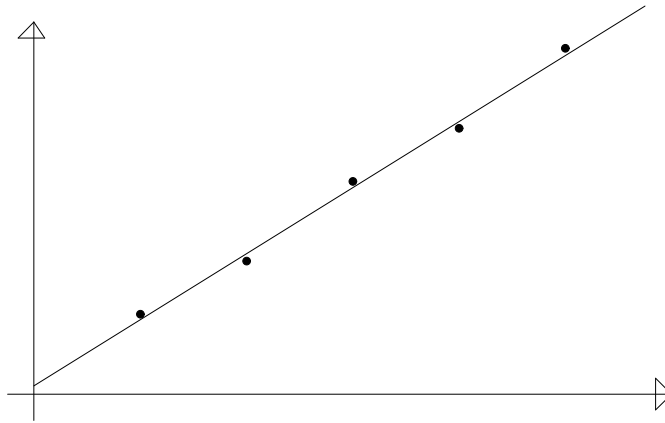
Sats 1 Residualvektorn $\mathbf{r} = A\mathbf{x} - \mathbf{b}$ blir minimal om \mathbf{x} satisfierar (normal-ekvationen) $A^t A\mathbf{x} = A^t \mathbf{b}$.

Anmärkning 1 Denna 'lösning' kallas för en minstakvadratlösning.

Exempel 1 Typexempel på ett överbestämt system är när man har en mängd data som man vill anpassa så bra som möjligt till någon typ av funktion. Antag att vi har mätdata

x	1	2	3	4	5
y	3	5	8	10	13

och så vill vi hitta en rät linje $y = kx + m$ som stämmer överens 'så bra som möjligt' med dessa data.



Vi får ett ekvationssystem $A\mathbf{x} = \mathbf{b}$ med 5 ekvationer och 2 obekanta:

$$\begin{cases} 3 = 1k + m \\ 5 = 2k + m \\ 8 = 3k + m \\ 10 = 4k + m \\ 13 = 5k + m \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 8 \\ 10 \\ 13 \end{pmatrix}.$$

Vi bestämmer normalekvationen $A^t A\mathbf{x} = A^t \mathbf{b}$ och får

$$A^t A = \begin{pmatrix} 55 & 15 \\ 15 & 5 \end{pmatrix} \text{ och } A^t \mathbf{b} = \begin{pmatrix} 142 \\ 39 \end{pmatrix}.$$

Löser vi denna så får vi lösningen $(5/2 \ 3/10)^t$, dvs $k = 5/2$ och $m = 3/10$ som är linjen som är ritad i bilden.

Basbyten

Vi rekapitulerar att en uppsättning B av vektorer kallas för en bas om varje vektor kan skrivas som en unik linjärkombination av vektorerna i B . Standardexemplet är basen

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

som kallas just standardbasen för \mathbb{R}^n .

Låt nu $\mathbf{f}_1, \dots, \mathbf{f}_n$ vara en bas för \mathbb{R}^n , vilken som helst. Vi tänker oss \mathbf{f}_i som en vektor med koordinater i standardbasen. Varje vektor $\mathbf{x} \in \mathbb{R}^n$ kan skrivas som

$$\mathbf{x} = x_1 \mathbf{f}_1 + x_2 \mathbf{f}_2 + \dots + x_n \mathbf{f}_n,$$

för några tal x_1, x_2, \dots, x_n som kallas för \mathbf{x} koordinater i basen $(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n)$. Denna likhet kan skrivas på matrisform som

$$\mathbf{x} = (\mathbf{f}_1 \ \mathbf{f}_2 \ \dots \ \mathbf{f}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = F\mathbf{x}_F.$$

Med andra ord är kolumnerna i F koordinaterna för basvektorerna $\mathbf{f}_1, \dots, \mathbf{f}_n$ (i standardbasen) och \mathbf{x}_F är vektorn som består av \mathbf{x} koordinater i basen F . Från den här likheten får vi genom att multiplicera med F^{-1} (som existerar eftersom kolumnerna i F är linjärt oberoende) att

$$\mathbf{x}_F = F^{-1}\mathbf{x}.$$

Så man får alltså koordinaterna \mathbf{x}_F i basen F genom att multiplicera de gamla koordinaterna med F^{-1} .

Exempel 2 Vektorerna $\mathbf{f}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ och $\mathbf{f}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ är inte parallella så de utgör en bas för \mathbb{R}^2 . Sätt

$$F = (\mathbf{f}_1 \ \mathbf{f}_2) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Då är

$$F^{-1} = \frac{1}{1-4} \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}.$$

Koordinaterna för standardvektorerna \mathbf{e}_1 och \mathbf{e}_2 i basen $(\mathbf{f}_1, \mathbf{f}_2)$ ges då av

$$(\mathbf{e}_1)_F = F^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 2/3 \end{pmatrix}$$

och

$$(\mathbf{e}_2)_F = F^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix}.$$

Antag nu mer allmänt att vi har två godtyckliga baser $F = (\mathbf{f}_1 \ \mathbf{f}_2 \ \dots \ \mathbf{f}_n)$ och $G = (\mathbf{g}_1 \ \mathbf{g}_2 \ \dots \ \mathbf{g}_n)$. Vi kan uttrycka basvektorerna i F i de i G :

$$\begin{cases} \mathbf{f}_1 = p_{11}\mathbf{g}_1 + p_{21}\mathbf{g}_2 + \dots + p_{n1}\mathbf{g}_n \\ \vdots \\ \mathbf{f}_n = p_{1n}\mathbf{g}_1 + p_{2n}\mathbf{g}_2 + \dots + p_{nn}\mathbf{g}_n \end{cases}$$

för några tal p_{ij} . Detta blir på matrisform

$$F = (\mathbf{f}_1 \ \mathbf{f}_2 \ \dots \ \mathbf{f}_n) = (\mathbf{g}_1 \ \mathbf{g}_2 \ \dots \ \mathbf{g}_n) \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} = GP.$$

Låt \mathbf{x} vara en godtycklig vektor. Då är

$$\mathbf{x} = F\mathbf{x}_F \text{ och } \mathbf{x} = G\mathbf{x}_G,$$

där \mathbf{x}_F är koordinaterna för \mathbf{x} i basen F och \mathbf{x}_G de i basen G . Vi får

$$G\mathbf{x}_G = \mathbf{x} = F\mathbf{x}_F = GP\mathbf{x}_F.$$

Detta ger att

$$\mathbf{x}_G = P\mathbf{x}_F,$$

ty båda vektorerna är \mathbf{x} koordinater i basen G . Vi sammanfattar i följande sats:

Sats 2 Låt $F = (\mathbf{f}_1 \mathbf{f}_2 \dots \mathbf{f}_n)$ och $G = (\mathbf{g}_1 \mathbf{g}_2 \dots \mathbf{g}_n)$ vara två baser för \mathbb{R}^n . Antag att de är relaterade genom $F = GP$. Låt \mathbf{x} vara en vektor med koordinaterna \mathbf{x}_F respektive \mathbf{x}_G i baserna F respektive G . Då gäller

- $\mathbf{x}_G = P\mathbf{x}_F$.
- $\mathbf{x}_F = P^{-1}\mathbf{x}_G$.
- $G = FP^{-1}$
- $P = G^{-1}F$

Exempel 3 Låt $\mathbf{f}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ och $\mathbf{g}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\mathbf{g}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ vara två baser för \mathbb{R}^2 . Dessa är relaterade genom

$$F = (\mathbf{f}_1 \mathbf{f}_2) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = G \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Antag att \mathbf{x} har koordinaterna $\mathbf{x}_F = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ i basen F , dvs

$$\mathbf{x} = 2\mathbf{f}_1 + 3\mathbf{f}_2.$$

Då är

$$\mathbf{x}_G = P\mathbf{x}_F = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

Kontroll i standardbasen ger:

$$\begin{aligned} 2\mathbf{f}_1 + 3\mathbf{f}_2 &= 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 3 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix} \\ 3\mathbf{g}_1 - 2\mathbf{g}_2 &= 3 \begin{pmatrix} -1 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix}. \end{aligned}$$

ON-matriser

Definition 1 En linjär avbildning från \mathbb{R}^n till \mathbb{R}^n är **isometrisk** om den bevarar längder, dvs om

$$|\mathbf{Ax}| = |\mathbf{x}|,$$

för alla vektorer $\mathbf{x} \in \mathbb{R}^n$. Motsvarande matris säges vara en **ON-matris**.

Exempel 4 Exempel på isometriska avbildningar i \mathbb{R}^2 är rotationer kring origo och spegling genom origo. Båda dessa typer av avbildningar bevarar uppenbarligen längden hos en vektor.

Man kan visa (se boken avsnitt 6.2) att ON-matriser inte bara bevarar längder utan skalärprodukter i allmänhet, dvs

$$\mathbf{Ax} \cdot \mathbf{Ay} = \mathbf{x} \cdot \mathbf{y},$$

för alla vektorer \mathbf{x} och \mathbf{y} . Detta ger att om A är en ON-matris och man tar en ON-bas $\mathbf{e}_1, \dots, \mathbf{e}_n$ så kommer också $\mathbf{Ae}_1, \dots, \mathbf{Ae}_n$ att vara en ON-bas, ty alla kommer fortfarande vara av längd 1 och skalärprodukten mellan olika element kommer att vara 0. Men vi vet sedan tidigare att

$$A = (\mathbf{Ae}_1 \quad \mathbf{Ae}_2 \quad \dots \quad \mathbf{Ae}_n)$$

så vi får:

Sats 3 Den kvadratiska matrisen A är en ON -matris om och endast om dess kolonner utgör en ON -bas.

Exempel 5 Betrakta matrisen

$$A = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix}.$$

Vi ser att alla kolonner har längd 1 och att de är parvis ortogonala. Alltså är A en ON -matris.

Ett alternativt sätt att uttrycka att kolonnerna i en matris A är ortonormerade är att säga att $A^t A = I$, ty när man beräknar $A^t A$ så tar man skalärprodukt mellan kolonnerna i A . Denna produkt blir identitetsmatrisen om och endast om skalärprodukten av en kolonn med sig själv (dvs längden i kvadrat) blir 1 och skalärprodukten mellan olika kolonner blir 0. Men om $A^t A = I$ så måste $A^t = A^{-1}$, så vi har följande viktiga egenskap hos ON -matriser som gör det en barnlek att beräkna dess invers.

Sats 4 Den kvadratiska matrisen A är en ON -matris om och endast $A^t = A^{-1}$.