bild 1

Linjär Algebra M/TD Läsvecka 6

Omfattning och innehåll
6.1 Skalärprodukt, norm (längd) och ortogonalitet i $\mathbb{R}^n$. Ortogonala komplementet till ett underrum i $\mathbb{R}^n$.
6.2 Ortogonala mängder av vektorer i $\mathbb{R}^n$. Ortogonal bas, ortonormerad bas för underrum i $\mathbb{R}^n$. Ortogonal projektion på vektor, projektionsformeln.
6.3 Ortogonal projektion på underrum i $\mathbb{R}^n$.

bild 2
Omfattning och innehåll forts.
6.4 Gram-Schmidt proceduren för att bestämma ortogonal bas för underrum i $\mathbb{R}^n$.
6.5, 6.6 Minstakvadrat-metoden med tillämpningar.

bild 3
Mål
Att kunna beräkna skalärprodukt och norm och behärska räknereglerna för dessa.
Kunna bevisa Pythagoras sats i $\mathbb{R}^n$.
Att veta vad som menas med $W^\perp$ om W är ett underrunm i $\mathbb{R}^n$ och kunna bevisa sats 3.
Att kunna bevisa att en ortogonal mängd av vektorer är linjärt oberoende.
Att veta vad som menas med en ortonormerad bas och kunna bevisa att koordinaterna ges av $c_j = \boldsymbol{y} \cdot \boldsymbol{u}_j$.

bild 4
Mål forts.
Att veta vad som menas med en ortogonal matris och kunna bevisa sats 7.
Att kunna dela upp en vektor i ortogonala komponenter, en i W och den andra i $W^\perp$ och kunna bevisa sats 8.
Att kunna bestämma en ortonormerad bas för underrum i $\mathbb{R}^n$.
Att veta vad som menas med en minstakvadrat-lösning och att kunna tillämpa minstakvadrat-metoden för modellanpassning.
Att kunna förklara varför minstakvadrat-lösningarna är lösningar till den normaliserade ekvationen $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$.

bild 5

Skalärprodukt, dot product, i $\mathbb{R}^n$

Om $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ och $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ så är

skalärprodukten av $\boldsymbol{u}$ och $\boldsymbol{v}$

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Med matrisbeteckningar kan detta skrivas

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v}$$

bild 6

sats 1, Räknelagar för skalärprodukten.

Låt $\boldsymbol{u}$, $\boldsymbol{v}$ och $\boldsymbol{w}$ vara vektorer i $\mathbb{R}^n$ och låt c vara en skalär. Då gäller:

a. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$

b. $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$

c. $(c\boldsymbol{u}) \cdot \boldsymbol{v} = c(\boldsymbol{u} \cdot \boldsymbol{v})$

d. $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$, och $\boldsymbol{u} \cdot \boldsymbol{u} = 0 \Leftrightarrow \boldsymbol{u} = \boldsymbol{0}$

bild 7

Definition:

Normen, längden av en vektor $\boldsymbol{v}$ är skalären $\|\boldsymbol{v}\|$ som ges av:

$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$

Notera att $\|\boldsymbol{v}\|^2 = \boldsymbol{v} \cdot \boldsymbol{v}$.

Detta är mycket användbart t.ex. då man vill bevisa att två längder är lika eftersom man då kan utnyttja räkneregler för skalärprodukt och slipper rottecknet.

bild 8

En **enhetsvektor** eller **normerad** vektor är en vektor $\boldsymbol{v}$ sådan att $\|\boldsymbol{v}\| = 1$

Att **normera** en vektor $\boldsymbol{v}$ innebär att skapa en normerad vektor $\hat{\boldsymbol{v}}$ med samma riktning som $\boldsymbol{v}$.

Eftersom $\|c\boldsymbol{v}\| = |c| \|\boldsymbol{v}\|$ kan alla vektorer $\boldsymbol{v}$ (utom $\mathbf{0}$) normeras genom att man multiplicerar med $1/\|\boldsymbol{v}\|$.

$$\hat{\boldsymbol{v}} = \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}$$

bild 9

Definition:

Låt $\boldsymbol{u}$ och $\boldsymbol{v}$ vara vektorer i $\mathbb{R}^n$.

Avståndet mellan $\boldsymbol{u}$ och $\boldsymbol{v}$ betecknas $\operatorname{dist}(\boldsymbol{u}, \boldsymbol{v})$ och ges av

$$\operatorname{dist}(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|$$

bild 10

Definition:

Två vektorer $\boldsymbol{u}$ och $\boldsymbol{v}$ i $\mathbb{R}^n$ kallas **ortogonala** mot varandra om $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

Notera att nollvektorn är ortogonal mot alla vektorer. Notera också att ingen annan vektor har denna egenskap.

bild 11

Sats 2, Pythagoras sats

Två vektorer $\boldsymbol{u}$ och $\boldsymbol{v}$ i $\mathbb{R}^n$ är ortogonala mot varandra om och endast om $\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$

Kommentar

Det är föga överraskande att Pythagoras sats gäller också i $\mathbb{R}^n$ eftersom definitionen av normen ser ut som den gör. Om $\boldsymbol{e}_i$ och $\boldsymbol{e}_j$ är två av vektorerna i standardbasen så har vi direkt från definitionen att $\|x_i\boldsymbol{e}_i + x_j\boldsymbol{e}_j\|^2 = \|x_i\boldsymbol{e}_i\|^2 + x_j \|\boldsymbol{e}_j\|^2$.

Detta bevisar naturligtvis inte satsen generellt, Man måste ge ett algebraiskt bevis, likt bokens.

bild 12
Låt W vara ett underrum i $\mathbb{R}^n$.
Om en vektor $\boldsymbol{z}$ är ortogonal mot **varje** vektor $\boldsymbol{u} \in W$ så säger vi att $\boldsymbol{z}$ är ortogonal mot W.
Mängden av alla vektorer $\boldsymbol{z}$ som är ortogonala mot W kallas W:s ortogonala komplement. Denna mängd betecknas $W^\perp$.
Viktiga fakta:

a. $\boldsymbol{x} \in W^\perp$ om och endast om $\boldsymbol{x}$ är ortogonal mot alla vektorer i en mängd som spänner W.

b. $W^\perp$ är ett underrum i $\mathbb{R}^n$.

Kommentar
Låt W vara underrummet i $\mathbb{R}^4$ som spänns upp av vektorerna

$$S = \{\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \ \boldsymbol{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 3 \end{bmatrix}\}, \text{ alltså } W = \text{Span}(S).$$

Ortogonala komplementet $W^\perp$ till W är alla vektorer som är ortogonala mot alla vektorer i W. Men det räcker att veta att de är ortogonala mot alla vekterna i S eftersom dessa spänner upp W.

Vi har således att $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in W^\perp \Leftrightarrow \boldsymbol{x} \cdot \boldsymbol{u}_1 = 0$, och $\boldsymbol{x} \cdot \boldsymbol{u}_2 = 0 \Leftrightarrow$

$$\begin{cases} 1x_1 + 2x_2 + 0x_3 - 1x_4 & = & 0 \\ 2x_1 - 1x_2 + 1x_3 + 3x_4 & = & 0 \end{cases} \Leftrightarrow \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} 5 & 0 & 2 & 5 \\ 0 & -5 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 5 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

forts. kommentar

En enkel kontroll visar att $\boldsymbol{v}_1 = \begin{bmatrix} -2 \\ 1 \\ 5 \\ 0 \end{bmatrix}$ och $\boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ är ortogonala mot både $\boldsymbol{u}_1$ och $\boldsymbol{u}_2$.
Vi ser också att $\dim W + \dim W^\perp = 4$ vilket är rimligt. I själva verket ger dessa två kontroller att vi vet att $W^\perp$ är korrekt bestämt.

bild 13
Sats 3
Låt A vara en $m \times n$-matris. Då gäller:

$$(\mathrm{Row}\,(A))^\perp = \mathrm{Nul}\,(A) \;\text{ och }\; (\mathrm{Col}\,(A))^\perp = \mathrm{Nul}\left(A^T\right)$$

Kommentar

Med $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 0 & 1 \\ -1 & 3 \end{bmatrix}$ är $\mathrm{Col}(A) = W$ i föregående kommentar. Kalkylen ovan visar att $(\mathrm{Col}\,(A))^\perp$ är alla lösningar till $\begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.
Alltså $(\mathrm{Col}\,(A))^\perp = \mathrm{Nul}\left(A^T\right)$ i detta exempel.

Eftersom alla underrum i $\mathbb{R}^n$ kan erhållas som $\mathrm{Col}\,(A)$ för någon matris A följer av rangsatsen att $\dim W + \dim W^\perp = n$.
Därför räcker det vid kalkyler alltid att kontrollera att $\dim W + \dim W^\perp = n$ och att de vektorer som spänner upp W är ortogonala mot de man anser spänner upp $W^\perp$.

bild 14
En mängd vektorer $\{\boldsymbol{u}_1, \; \cdots \; \boldsymbol{u}_p\}$ kallas för en **ortogonal mängd** om vektorerna är **parvis ortogonala**.
Alltså om alla par av vektorer (med olika index) är ortogonala mot varandra, $\boldsymbol{u}_i \cdot \boldsymbol{u}_j = 0$ om $i \neq j$.
En ortogonal mängd kallas **ortonormerad** om alla vektorerna i mängden dessutom är normerade.

bild 15

Sats 4

Om $S = \{\boldsymbol{u}_1, \; \cdots \; \boldsymbol{u}_p\}$ är en ortogonal mängd av vektorer i $\mathbb{R}^n$ så är S en linjärt oberoende mängd av vektorer.

S är då en bas för underrummet i $\mathbb{R}^n$ som spänns upp av S.

bild 16

Definition: Med en **ortogonal bas** för ett underrum W i $\mathbb{R}^n$ menas en bas $\mathcal{B}$ som också är en ortogonal mängd.

En bas kallas **ortonormerad** om den är en ortogonal bas och alla vektorer i basen är normerade.

Kommentar

Med $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}$, $\boldsymbol{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix}$ och $\boldsymbol{u}_3 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 4 \end{bmatrix}$ har vi att

$\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = 0, \; \boldsymbol{u}_1 \cdot \boldsymbol{u}_3 = 0$ och$\boldsymbol{u}_2 \cdot \boldsymbol{u}_3 = 0$.

$\mathcal{S} = \{\boldsymbol{u}_\infty, \; \boldsymbol{u}_\in, \; \boldsymbol{u}_\ni\}$ är således en ortogonal mängd och alltså, enligt sats 4, en ortogonal bas för $H = \text{Span}\{\boldsymbol{u}_1, \; \boldsymbol{u}_2, \; \boldsymbol{u}_3\}$.

$\|\boldsymbol{u}_1\| = \sqrt{6}, \; \|\boldsymbol{u}_2\| = \sqrt{9} = 3, \; \|\boldsymbol{u}_3\| = \sqrt{21}.$

$\mathcal{B} = \{\frac{\infty}{\sqrt{/}}\boldsymbol{u}_\infty, \; \frac{\infty}{\ni}\boldsymbol{u}_\in, \; \frac{\infty}{\sqrt{\in\infty}}\boldsymbol{u}_\ni\}$ är en ON-bas för H.

bild 17

Sats 5 Låt $\mathcal{B} = \{\boldsymbol{u}_1, \; \cdots, \; \boldsymbol{u}_p\}$ vara en ortogonal bas för ett underrum W i $\mathbb{R}^n$. Låt $\boldsymbol{y}$ vara en vektor i W.

Koordinaterna för $\boldsymbol{y}$ relativt basen $\mathcal{B}$, $[\boldsymbol{y}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$

ges då av

$$c_j = \frac{\boldsymbol{y} \cdot \boldsymbol{u}_j}{\boldsymbol{u}_j \cdot \boldsymbol{u}_j}$$

Om $\mathcal{B}$ är en ortonormerad bas ges koordinaterna av

$$c_j = \boldsymbol{y} \cdot \boldsymbol{u}_j$$

Kommentar

Låt $\mathcal{S}$, $\mathcal{B}$ och H vara enligt föregående exempel och låt $\boldsymbol{y} = \begin{bmatrix} 5 \\ -2 \\ -3 \\ -5 \end{bmatrix}$.

Att $\boldsymbol{y} \in H$ är inte självklart men jag har valt $\boldsymbol{y}$ så att det är uppfyllt.

Nu har vi att $\boldsymbol{y} \cdot \boldsymbol{u}_1 = 6,\ \boldsymbol{y} \cdot \boldsymbol{u}_2 = -18$ och $\boldsymbol{y} \cdot \boldsymbol{u}_3 = -21$.

Enligt satsen är $[\boldsymbol{y}]_{\mathcal{S}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$ där $c_1 = \frac{6}{6},\ c_2 = \frac{-18}{9}, c\text{‘}\ c_3 = \frac{-21}{21}$

Alltså är $[\boldsymbol{y}]_{\mathcal{S}} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$.

Vi kan nu kontrollera att $\boldsymbol{y} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + c_3\boldsymbol{u}_3$ vilket visar sig stämma. Nu vet vi att $\boldsymbol{y} \in H$ och att kalkylerna stämmer.
Naturligtvis skall samma resultat erhållas genom att man löser ekvationssystemet

$$\begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & -1 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ -3 \\ -5 \end{bmatrix}.$$

Kontrollen ovan kan enkelt utföras genom matrismultiplikationen

$$\begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & -1 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ -3 \\ -5 \end{bmatrix}.$$

På samma sätt kan $[\boldsymbol{y}]_{\mathcal{B}}$ beräknas, med satsen eller genom att lösa ekvationssystem, och vi får

$$[\boldsymbol{y}]_{\mathcal{B}} = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{-2}{3} \\ \frac{-1}{\sqrt{21}} \end{bmatrix}$$

bild 18
Projektionsformeln (sid 386)
Låt $\boldsymbol{u} \neq \mathbf{0}$ vara en given vektor i $\mathbb{R}^n$. Låt $\boldsymbol{y}$ vara en annan vektor i $\mathbb{R}^n$.
Vi söker en uppdelning av $\boldsymbol{y}$ i två ortogonala komponenter, $\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z}$ där $\hat{\boldsymbol{y}} = \alpha \boldsymbol{u}$ är parallell med $\boldsymbol{u}$ och $\boldsymbol{z}$ är ortogogonal mot $\boldsymbol{u}$.
Då är

$$\hat{\boldsymbol{y}} = \left(\frac{\boldsymbol{y} \cdot \boldsymbol{u}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right) \boldsymbol{u} \quad \text{och} \quad \boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{y}}$$

Notera att $\hat{\boldsymbol{y}}$ är parallell med linjen L genom origo med riktningsvektor $\boldsymbol{u}$. Därför är $\hat{\boldsymbol{y}}$ projektionen av $\boldsymbol{y}$ på L. Man använder ibland beteckningen $\text{proj}_{\boldsymbol{u}} \boldsymbol{y}$ eller $\text{proj}_L \boldsymbol{y}$ för projektionen.

Kommentar
Denna projektionsformel ser exakt likadan ut som den vi härledde i inledande kursen. Då gav vi geometriska argument för att inse att den är korrekt. Nu kan vi helt enkel ta den gamla formeln och undersöka om den även i $\mathbb{R}^n$ leder till en vektor $\boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{y}}$ som är ortogonal mot $\boldsymbol{u}$. En enkel kalkyl visar att $\boldsymbol{z} \cdot \boldsymbol{u} = (\boldsymbol{y} - \hat{\boldsymbol{y}}) \cdot \boldsymbol{u} = \boldsymbol{y} \cdot \boldsymbol{u} - \hat{\boldsymbol{y}} \cdot \boldsymbol{u} = \boldsymbol{y} \cdot \boldsymbol{u} - \left(\frac{\boldsymbol{y} \cdot \boldsymbol{u}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right)(\boldsymbol{u} \cdot \boldsymbol{u}) = 0$

Kommentar
En intressant konsekvens av projektionsformeln och Pythagoras sats är **Cauchy-Schwartz olikhet** $|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|$.
Bevis: Projektionsformeln ger att $\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z} = \left(\frac{\boldsymbol{y} \cdot \boldsymbol{x}}{\boldsymbol{x} \cdot \boldsymbol{x}}\right) \boldsymbol{x} + \boldsymbol{z}$, där $\hat{\boldsymbol{y}}$ och $\boldsymbol{z}$ är ortogonala mot varandra.
Pythagoras sats ger då att $\|\boldsymbol{y}\|^2 = \|\hat{\boldsymbol{y}}\|^2 + \|\boldsymbol{z}\|^2 = \left(\frac{\boldsymbol{y} \cdot \boldsymbol{x}}{\boldsymbol{x} \cdot \boldsymbol{x}}\right)^2 \|\boldsymbol{x}\|^2 + \|\boldsymbol{z}\|^2 \geq \left(\frac{\boldsymbol{y} \cdot \boldsymbol{x}}{\boldsymbol{x} \cdot \boldsymbol{x}}\right)^2 \|\boldsymbol{x}\|^2 = \left(\frac{\boldsymbol{y} \cdot \boldsymbol{x}}{\|\boldsymbol{x}\|^2}\right)^2 \|\boldsymbol{x}\|^2$. Alltså $\|\boldsymbol{y}\|^2 \geq \frac{(\boldsymbol{y} \cdot \boldsymbol{x})^2}{\|\boldsymbol{x}\|^2}$. Multiplikation med $\|\boldsymbol{x}\|^2$ och rotutdragning ger den önskade olikheten.

Kommentar
Av Cauchy-Schwartz olikhet följer att $-1 \leq \dfrac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|} \leq 1$.

Man kan därför *definiera* **vinkeln θ mellan $\boldsymbol{x}$ och $\boldsymbol{y}$** genom

$$\theta = \arccos\left(\frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|}\right)$$

Vi skall se senare att denna definition är rimlig, den stämmer överens med den vinkel man får om man betraktar planet som de två vektorerna ligger i som ett plan i åskådliga rummet och där använder ”den vanliga vinkeln”..

bild 19

Sats 6. En $m \times n$-matris U har ortonormerade kolonner om och endast om $U^T U = I$

Definition: En kvadratisk inverterbar matris U kallas **ortogonal** om $U^{-1} = U^T$

Sats 6'. En $n \times n$-matris är ortogonal om och endast om den har ortonormerade kolonner.

bild 20

Sats 7. Låt U vara en $m \times n$-matris med ortonormerade kolonner och låt $\boldsymbol{x}$ och $\boldsymbol{y}$ vara vektorer i $\mathbb{R}^n$. Då gäller:

a. $\|U\boldsymbol{x}\| = \|\boldsymbol{x}\|$

b. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$

c. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = 0$ om och endast om $\boldsymbol{x} \cdot \boldsymbol{y} = 0$

Kommentar

Satsen innebär att om en avbildningsmatris U har ortonormerade kolonner så bevaras både längd och vinklar mellan vektorer.

Om U dessutom är kvadratisk och alltså en ortogonalmatris kan vi se U som basbytesmatris $U = P_{\mathcal{B}}$ från basen $\mathcal{B}$ som ges av kolonnerna i U till standardbasen.

Om vi har två vektorer $\boldsymbol{u}$ och $\boldsymbol{v}$ i $\mathbb{R}^n$ så kan vi alltid välja en ON-bas $\mathcal{B}$ för $\mathbb{R}^n$

där den ena har koordinater $\begin{bmatrix} a_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ och den andra har koordinater $\begin{bmatrix} b_1 \\ b_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Vi låter sedan U vara den ortogonala matrisen vars kolonner är vektorerna i $\mathcal{B}$. Då är alltså $U\,[\boldsymbol{u}]_{\mathcal{B}} = \boldsymbol{u}$ och $U\,[\boldsymbol{v}]_{\mathcal{B}} = \boldsymbol{v}$.

Satsen säger då att längd och vinkel i $\mathbb{R}^n$ är samma som längd och vinkel i $\mathbb{R}^2$, alltså samma som i ”det vanliga åskådliga euklidiska planet ”.

Så länge vi bara arbetar med två eller tre vektorer kan vi därför rita figurer och resonera som om det mångdimensionella är två- eller tredimensionellt.

bild 21
Sats 8 Satsen om ortogonal uppdelning.
Låt W vara ett underrum i $\mathbb{R}^n$ och låt $\{\boldsymbol{u}_1, \cdots \boldsymbol{u}_p\}$ vara en ortogonal bas för W. Då kan varje vektor $\boldsymbol{y}$ i $\mathbb{R}^n$ skrivas, på exakt ett sätt, som en summa

$$\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z}$$

där $\hat{\boldsymbol{y}}$ är en vektor i W och $\boldsymbol{z}$ är en vektor i $W^{\perp}$.
Vektorn $\hat{\boldsymbol{y}}$ kallas projektionen av $\boldsymbol{y}$ på W och betecknas även $\mathrm{proj}_W \boldsymbol{y}$.
Vektorn $\hat{\boldsymbol{y}}$ beräknas med projektionsformeln

$$\hat{\boldsymbol{y}} = \left(\frac{\boldsymbol{y} \cdot \boldsymbol{u}_1}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}\right) \boldsymbol{u}_1 + \cdots + \left(\frac{\boldsymbol{y} \cdot \boldsymbol{u}_p}{\boldsymbol{u}_p \cdot \boldsymbol{u}_p}\right) \boldsymbol{u}_p \quad \text{och} \quad \boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{y}}$$

Vi skall senare se att man alltid kan bestämma en ortogonal bas för ett underrum. Man kan därför alltid bestämma $\mathrm{proj}_W \boldsymbol{y}$.

Kommentar
Bokens formulering är något annorlunda. I själva verket haltar den något eftersom vi ännu inte har sett att man alltid kan hitta en ortogonal bas för ett givet underrum i $\mathbb{R}^n$. Vi vet bara att man alltid kan hitta en bas. Det är inte särskilt svårt att argumentera för att man kan välja basen ortogonal men det följer av Gram-Schmidt proceduren att det är så. Kanske den borde kommit före sats 8 istället för efter. Vid problemlösning startar man oftast med att först använda Gram-Schmidt sedan sats 8 så som jag gör i exemplet i kommentaren till Gram-Schmidt proceduren.

bild 22
Sats 9 Satsen om bästa approximation.
Låt W vara ett underrum i $\mathbb{R}^n$ och $\boldsymbol{y}$ en godtycklig vektor i $\mathbb{R}^n$. Låt $\hat{\boldsymbol{y}}$ vara projektionen av $\boldsymbol{y}$ på W.
Då är $\hat{\boldsymbol{y}}$ den punkt i W som ligger närmast $\boldsymbol{y}$.
Med andra ord gäller att

$$\|\boldsymbol{y} - \hat{\boldsymbol{y}}\| < \|\boldsymbol{y} - \boldsymbol{v}\|$$

för alla $\boldsymbol{v}$ i W med $\boldsymbol{v} \neq \hat{\boldsymbol{y}}$.

Kommentar
Rita en tredimensionell figur där $\boldsymbol{y}$, $\hat{\boldsymbol{y}}$ och $\boldsymbol{v}$ utgår från samma punkt och spänner upp en tetrader. Rita så att $\boldsymbol{y}$ är hypotenusa i en rätvinklig triangel med kateter $\hat{\boldsymbol{y}}$ och $\boldsymbol{y}-\hat{\boldsymbol{y}}$. De två återstående tetraederkanterna är $\boldsymbol{y}-\boldsymbol{v}$ och $\hat{\boldsymbol{y}}-\boldsymbol{v}$. Dessa bildar också en rätvinklig triangel med $\boldsymbol{y}-\hat{\boldsymbol{y}}$ som en katet. Denna är alltid kortare än hypotenusan $\boldsymbol{y}-\boldsymbol{v}$

bild 23
sats 10
Låt W vara ett underrum i $\mathbb{R}^n$ och låt $\{\boldsymbol{u}_1, \cdots \boldsymbol{u}_p\}$ vara en ortonormerad bas för W.
Då är
$$\text{proj}_W \boldsymbol{y} = (\boldsymbol{y}\cdot\boldsymbol{u}_1)\,\boldsymbol{u}_1 + \cdots + (\boldsymbol{y}\cdot\boldsymbol{u}_p)\,\boldsymbol{u}_p$$
Om $U = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_p \end{bmatrix}$, så gäller
$$\text{proj}_W \boldsymbol{y} = UU^T\boldsymbol{y} \text{ för alla } \boldsymbol{y} \in \mathbb{R}^n$$

Kommentar
De olika sätten att beräkna matrisprodukt ger:
$$\begin{bmatrix} \boldsymbol{y}\cdot\boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{y}\cdot\boldsymbol{u}_p \end{bmatrix} = U^T\boldsymbol{y} \text{ och } \text{proj}_W \boldsymbol{y} = U \begin{bmatrix} \boldsymbol{y}\cdot\boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{y}\cdot\boldsymbol{u}_p \end{bmatrix}$$

bild 24
Sats 11 Gram-Schmidt ortogonaliseringsprocedur.
Låt $\{\boldsymbol{u}_1, \cdots \boldsymbol{u}_p\}$ vara en bas för ett underrum W i $\mathbb{R}^n$. Sätt
$$\boldsymbol{v}_1 = \boldsymbol{u}_1$$
$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1$$
$$\boldsymbol{v}_3 = \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_3\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2$$
$$\vdots$$
$$\boldsymbol{v}_p = \boldsymbol{u}_p - \frac{\boldsymbol{u}_p\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_p\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2 - \cdots - \frac{\boldsymbol{u}_p\cdot\boldsymbol{v}_{p-1}}{\boldsymbol{v}_{p-1}\cdot\boldsymbol{v}_{p-1}}\boldsymbol{v}_{p-1}$$
Då är $\{\boldsymbol{v}_1, \cdots \boldsymbol{v}_p\}$ en ortogonal bas för W. Dessutom gäller att
$$\text{Span}\{\boldsymbol{v}_1, \cdots \boldsymbol{v}_k\} = \text{Span}\{\boldsymbol{u}_1, \cdots \boldsymbol{u}_k\} \quad \text{för } 1 \le k \le p$$
.

Kommentar

Låt $A = \begin{bmatrix} 1 & -1 & 1 & -3 \\ 2 & 3 & 7 & 1 \\ 0 & 2 & 2 & 1 \\ -1 & -1 & -3 & 2 \end{bmatrix}$ och $\boldsymbol{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$.

Vi skall bestämma den ortogonala projektionen av $\boldsymbol{y}$ på Col(A).

Vi måste först bestämma en bas för Col(A). Sedan en ortogonal bas för Col(A). Och slutligen tillämpa sats 10 för att bestämma projektionen.

Radoperationer ger att $A = \begin{bmatrix} 1 & -1 & 1 & -3 \\ 2 & 3 & 7 & 1 \\ 0 & 2 & 2 & 1 \\ -1 & -1 & -3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Kolonnerna 1, 2 och 4 är pivotkolonner i trappstegsmatrisen. Dessa kolonner i matrisen A ger en bas för Col(A).

Alltså $S = \{\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} -1 \\ 3 \\ 2 \\ -1 \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 2 \end{bmatrix}\}$ är den sökta basen.

Nästa steg är alltså att skapa en ortogonal bas $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ utgående från S.

Enligt Gram-Schmidt proceduren sätter vi då $\boldsymbol{v}_1 = \boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}$.

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \boldsymbol{\cdot} \boldsymbol{v}_1}{\boldsymbol{v}_1 \boldsymbol{\cdot} \boldsymbol{v}_1}\boldsymbol{v}_1 = \begin{bmatrix} -1 \\ 3 \\ 2 \\ -1 \end{bmatrix} - \frac{6}{6}\begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\boldsymbol{v}_3 = \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3 \boldsymbol{\cdot} \boldsymbol{v}_1}{\boldsymbol{v}_1 \boldsymbol{\cdot} \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_3 \boldsymbol{\cdot} \boldsymbol{v}_2}{\boldsymbol{v}_2 \boldsymbol{\cdot} \boldsymbol{v}_2}\boldsymbol{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 2 \end{bmatrix} - \frac{-3}{6}\begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} - \frac{9}{9}\begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ -2 \\ 3 \end{bmatrix}$$

Vi kan istället välja $\boldsymbol{v}_3 = \begin{bmatrix} -1 \\ 2 \\ -2 \\ 3 \end{bmatrix}$, det påverkar inte ortogonaliteten.

forts. kommentar
Det är nu lämpligt att kontrollera att vi har en ortogonal bas.
$\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0, \boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = 0, \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 = 0$ Som önskat.

Vi kan nu beräkna projektionen av $\boldsymbol{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ på $\mathrm{Col}(A)$.

Sats 8 ger att den sökta projektionen är

$$\hat{\boldsymbol{y}} = \left(\frac{\boldsymbol{y} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right) \boldsymbol{v}_1 + \left(\frac{\boldsymbol{y} \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\right) \boldsymbol{v}_2 + \left(\frac{\boldsymbol{y} \cdot \boldsymbol{v}_3}{\boldsymbol{v}_3 \cdot \boldsymbol{v}_3}\right) \boldsymbol{v}_3 =$$

$$\left(\tfrac{1}{6}\right) \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} + \left(\tfrac{6}{9}\right) \begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix} + \left(\tfrac{9}{18}\right) \begin{bmatrix} -1 \\ 2 \\ -2 \\ 3 \end{bmatrix} = \tfrac{1}{3} \begin{bmatrix} -5 \\ 6 \\ 1 \\ 4 \end{bmatrix}$$

Vi får nu den mot $\mathrm{Col}(A)$ ortogonala komponenten,

$$\boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} - \tfrac{1}{3} \begin{bmatrix} -5 \\ 6 \\ 1 \\ 4 \end{bmatrix} = \tfrac{1}{3} \begin{bmatrix} 8 \\ 0 \\ 8 \\ 8 \end{bmatrix}$$

Vi kan nu kontrollera att $\boldsymbol{z}$ är ortogonal mot vektorerna i $\mathcal{B}$. $\boldsymbol{v}_1 \cdot \boldsymbol{z} = 0, \boldsymbol{v}_2 \cdot \boldsymbol{z} = 0, \boldsymbol{v}_3 \cdot \boldsymbol{z} = 0$.

Det skulle kunna vara så att vektorerna $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ faktiskt inte är en ortogonal bas för $\mathrm{Col}(A)$. I så fall har vi gjort något räknefel i början av lösningen.. Det skulle vi inte upptäcka med de kontroller som gjorts. Men ännu en kontroll kan göras. $\boldsymbol{z}$ skall vara ortogonal mot $\mathrm{Col}(A)$ och alltså ligga i $\mathrm{Nul}(A^T)$.

Vi finner att $A^T \boldsymbol{z} = \frac{8}{3} \begin{bmatrix} 1 & 2 & 0 & -1 \\ -1 & 3 & 2 & -1 \\ 1 & 7 & 2 & -3 \\ -3 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

Nu vet vi (nästan säkert) att uppgiften är korrekt löst. Försök lista ut vilken kombination av felaktigheter som inte skulle upptäckts genom de kontroller som utförts. Det finns ännu en kontroll som kan göras och som slutgiltigt garanterar att lösningen är korrekt.

bild 25

Definition:

Om A är en $m \times n$-matris och $\boldsymbol{b} \in \mathbb{R}^n$ så är
en **minstakvadrat-lösning** till ekvationen $A\boldsymbol{x} = \boldsymbol{b}$
en vektor $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ sådan att

$$\|A\hat{\boldsymbol{x}} - \boldsymbol{b}\| \leq \|A\boldsymbol{x} - \boldsymbol{b}\| \text{ för alla } \boldsymbol{x} \in \mathbb{R}^n.$$

Minstakvadrat-felet är $\|A\hat{\boldsymbol{x}} - \boldsymbol{b}\|$ då $\hat{\boldsymbol{x}}$ är minsta-kvadrat lösningen.
Ofta använder man istället *kvadratiska medelfelet* som är $\|A\hat{\boldsymbol{x}} - \boldsymbol{b}\| / \sqrt{n}$ då n är antalet ekvationer i systemet

bild 26

Sats 13

Minstakvadrat-lösningarna till ekvationen $A\boldsymbol{x} = \boldsymbol{b}$ är samma som lösningarna till den (konsistenta) normaliserade ekvationen

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