

Linjär Algebra M/TD Läsvecka 6

Omfattning och innehåll

6.1 Skalarprodukt, norm (längd) och ortogonalitet i \mathbb{R}^n . Ortogonal komplementet till ett underrum i \mathbb{R}^n .

6.2 Ortogonal mängd 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 .

tmv165/185 V6, Vt06 bild 1

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.

7.1 Diagonalisering av symmetriska matriser.

tmv165/185 V6, Vt06 bild 2

Mål

Att kunna beräkna skalärprodukt och norm och behärska räkneregler för dessa.

Kunna bevisa Pythagoras sats i \mathbb{R}^n .

Att veta vad som menas med W^\perp om W är ett underrum 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 $c_j = \mathbf{y} \cdot \mathbf{u}_j$.

Att veta vad som menas med en ortogonal matris och kunna bevisa sats 7.

tmv165/185 V6, Vt06 bild 3

Mål forts.

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 $A^T A \mathbf{x} = A^T \mathbf{b}$.

Att kunna satserna 1 – 3 i kapitel 7.1, Spektralsatsen (sats 3) är extra viktig.

Att veta vad som menas med spektral uppdelning av en matris och projektionsmatris.

tmv165/185 V6, Vt06 bild 4

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

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

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

Med matrisbeteckningar kan detta skrivas

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

Lay 6.1, tmv165/185 V6, Vt06 bild 5

sats 1, Räknelagar för skalärprodukten.

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

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$
- $\mathbf{u} \cdot \mathbf{u} \geq 0$, och $\mathbf{u} \cdot \mathbf{u} = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$

Lay 6.1, tmv165/185 V6, Vt06 bild 6

Definition:

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

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

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

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 rot-tecknet.

Lay 6.1, tmv165/185 V6, Vt06 bild 7

En **enhetsvektor** eller **normerad** vektor är en vektor \mathbf{v} sådan att $\|\mathbf{v}\| = 1$. Att **normera** en vektor \mathbf{v} innebär att skapa en normerad vektor $\hat{\mathbf{v}}$ med samma riktning som \mathbf{v} .

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

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

Lay 6.1, tmv165/185 V6, Vt06 bild 8

Definition:

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

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

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Lay 6.1, tmv165/185 V6, Vt06 bild 9

Definition:

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

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

Lay 6.1, tmv165/185 V6, Vt06 bild 10

Sats 2, Pythagoras sats

Två vektorer \mathbf{u} och \mathbf{v} i \mathbb{R}^n är ortogonala mot varandra om och endast om

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Lay 6.1, tmv165/185 V6, Vt06 bild 11

Det är föga överraskande att Pythagoras sats gäller också i \mathbb{R}^n eftersom definitionen av normen ser ut som den gör. Om \mathbf{e}_i och \mathbf{e}_j är två av vektorerna i standardbasen så har vi direkt från definitionen att $\|x_i\mathbf{e}_i + x_j\mathbf{e}_j\|^2 = \|x_i\mathbf{e}_i\|^2 + x_j\|\mathbf{e}_j\|^2$. Detta bevisar naturligtvis inte satsen generellt, Man måste ge ett algebraiskt bevis, liksom bokens.

Låt W vara ett underrum i \mathbb{R}^n .

Om en vektor \mathbf{z} är ortogonal mot **varje** vektor $\mathbf{u} \in W$ så säger vi att \mathbf{z} är ortogonal mot W .

Mängden av alla vektorer \mathbf{z} som är ortogonala mot W kallas W :s ortogonala komplement. Denna mängd betecknas W^\perp .

Viktiga fakta:

- a. $\mathbf{x} \in W^\perp$ om och endast om \mathbf{x} är ortogonal mot alla vektorer i en mängd som spänner W .
- b. W^\perp är ett underrum i \mathbb{R}^n .

Lay 6.1, tmv165/185 V6, Vt06 bild 12

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

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

Ortogonal 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 vektorerna i S eftersom dessa spänner upp W .

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

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

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

En enkel kontroll visar att $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 1 \\ 5 \\ 0 \end{bmatrix}$ och $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ är ortogonala mot både

\mathbf{u}_1 och \mathbf{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.

Sats 3

Låt A vara en $m \times n$ -matris. Då gäller:

$$(\text{Row}(A))^\perp = \text{Nul}(A) \text{ och } (\text{Col}(A))^\perp = \text{Nul}(A^T)$$

Lay 6.1, tmv165/185 V6, Vt06 bild 13

Med $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 0 & 1 \\ -1 & 3 \end{bmatrix}$ är $\text{Col}(A) = W$ i föregående kommentar. Kalkylen ovan

visar att $(\text{Col}(A))^\perp$ 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

$(\text{Col}(A))^\perp = \text{Nul}(A^T)$ i detta exempel.

Eftersom alla underrum i \mathbb{R}^n kan erhållas som $\text{Col}(A)$ 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 .

En mängd vektorer $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ kallas för en **ortogonal mängd** om vektorerna är **parvis ortogonala**.

Alltså om alla par av vektorer (med olika index) är ortogonala mot varandra, $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ om $i \neq j$.

En ortogonal mängd kallas **ortonormerad** om alla vektorerna i mängden dessutom är normerade.

Lay 6.2, tmv165/185 V6, Vt06 bild 14

Sats 4

Om $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ 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 .

Lay 6.2, tmv165/185 V6, Vt06 bild 15

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.

Lay 6.2, tmv165/185 V6, Vt06 bild 16

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

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 0, \quad \mathbf{u}_1 \cdot \mathbf{u}_3 = 0 \quad \text{och} \quad \mathbf{u}_2 \cdot \mathbf{u}_3 = 0.$$

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

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

$\mathcal{B} = \{\frac{1}{\sqrt{6}}\mathbf{u}_1, \frac{1}{3}\mathbf{u}_2, \frac{1}{\sqrt{21}}\mathbf{u}_3\}$ är en ON-bas för H .

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

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

ges då av

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

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

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

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

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

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

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

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

Vi kan nu kontrollera att $\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3$ vilket visar sig stämma. Nu vet vi att $\mathbf{y} \in H$ 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 $[\mathbf{y}]_{\mathcal{B}}$ beräknas, med satsen eller genom att lösa ekvationssystem, och vi får

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

Projektionsformeln (sid 386)

Låt $\mathbf{u} \neq \mathbf{0}$ vara en given vektor i \mathbb{R}^n . Låt \mathbf{y} vara en annan vektor i \mathbb{R}^n .

Vi söker en uppdelning av \mathbf{y} i två ortogonala komponenter, $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ där $\hat{\mathbf{y}} = \alpha \mathbf{u}$ är parallell med \mathbf{u} och \mathbf{z} är ortogonal mot \mathbf{u} .

Då är

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

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

Lay 6.2, tmv165/185 V6, Vt06 bild 18

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 enkelt ta den gamla formeln och undersöka om den även i \mathbb{R}^n leder till en vektor $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$ som är ortogonal mot \mathbf{u} . En enkel kalkyl visar att $\mathbf{z} \cdot \mathbf{u} = (\mathbf{y} - \hat{\mathbf{y}}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - \hat{\mathbf{y}} \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - \left(\frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \right) (\mathbf{u} \cdot \mathbf{u}) = 0$

En intressant konsekvens av projektionsformeln och Pythagoras sats är **Cauchy-Schwartz olikhet** $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$.

Bevis: Projektionsformeln ger att $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \left(\frac{\mathbf{y} \cdot \mathbf{x}}{\mathbf{x} \cdot \mathbf{x}} \right) \mathbf{x} + \mathbf{z}$, där $\hat{\mathbf{y}}$ och \mathbf{z} är ortogonala mot varandra.

Pythagoras sats ger då att $\|\mathbf{y}\|^2 = \|\hat{\mathbf{y}}\|^2 + \|\mathbf{z}\|^2 = \left(\frac{\mathbf{y} \cdot \mathbf{x}}{\mathbf{x} \cdot \mathbf{x}} \right)^2 \|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 \geq \left(\frac{\mathbf{y} \cdot \mathbf{x}}{\mathbf{x} \cdot \mathbf{x}} \right)^2 \|\mathbf{x}\|^2 = \left(\frac{\mathbf{y} \cdot \mathbf{x}}{\|\mathbf{x}\|^2} \right)^2 \|\mathbf{x}\|^2$. Alltså $\|\mathbf{y}\|^2 \geq \frac{(\mathbf{y} \cdot \mathbf{x})^2}{\|\mathbf{x}\|^2}$. Multiplikation med $\|\mathbf{x}\|^2$ och rotutdragning ger den önskade olikheten.

Av Cauchy-Schwartz olikhet följer att $-1 \leq \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1$.

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

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

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"..

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.

Lay 6.2, tmv165/185 V6, Vt06 bild 19

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

- $\|U\mathbf{x}\| = \|\mathbf{x}\|$
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ om och endast om $\mathbf{x} \cdot \mathbf{y} = 0$

Lay 6.2, tmv165/185 V6, Vt06 bild 20

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 \mathbf{u} och \mathbf{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 $U[\mathbf{u}]_{\mathcal{B}} = \mathbf{u}$ och $U[\mathbf{v}]_{\mathcal{B}} = \mathbf{v}$.

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.

Sats 8 Satsen om ortogonal uppdelning.

Låt W vara ett underrum i \mathbb{R}^n och låt $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ vara en ortogonal bas för W . Då kan varje vektor \mathbf{y} i \mathbb{R}^n skrivas, på exakt ett sätt, som en summa

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

där $\hat{\mathbf{y}}$ är en vektor i W och \mathbf{z} är en vektor i W^\perp .

Vektorn $\hat{\mathbf{y}}$ kallas projektionen av \mathbf{y} på W och betecknas även $\text{proj}_W \mathbf{y}$.

Vektorn $\hat{\mathbf{y}}$ beräknas med projektionsformeln

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

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

Lay 6.3, tmv165/185 V6, Vt06 bild 21

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.

Sats 9 Satsen om bästa approximation.

Låt W vara ett underrum i \mathbb{R}^n och \mathbf{y} en godtycklig vektor i \mathbb{R}^n . Låt $\hat{\mathbf{y}}$ vara projektionen av \mathbf{y} på W .

Då är $\hat{\mathbf{y}}$ den punkt i W som ligger närmast \mathbf{y} .

Med andra ord gäller att

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

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

Lay 6.3, tmv165/185 V6, Vt06 bild 22

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

sats 10

Låt W vara ett underrum i \mathbb{R}^n och låt $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ vara en ortonormerad bas för W .

Då är

$$\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1) \mathbf{u}_1 + \dots + (\mathbf{y} \cdot \mathbf{u}_p) \mathbf{u}_p$$

Om $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_p]$, så gäller

$$\text{proj}_W \mathbf{y} = UU^T \mathbf{y} \text{ för alla } \mathbf{y} \in \mathbb{R}^n$$

Lay 6.3, tmv165/185 V6, Vt06 bild 23

De olika sätten att beräkna matrisprodukt ger:

$$\begin{bmatrix} \mathbf{y} \cdot \mathbf{u}_1 \\ \vdots \\ \mathbf{y} \cdot \mathbf{u}_p \end{bmatrix} = U^T \mathbf{y} \text{ och } \text{proj}_W \mathbf{y} = U \begin{bmatrix} \mathbf{y} \cdot \mathbf{u}_1 \\ \vdots \\ \mathbf{y} \cdot \mathbf{u}_p \end{bmatrix}$$

Sats 11 Gram-Schmidt ortogonaliseringsprocedur.

Låt $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ vara en bas för ett underrum W i \mathbb{R}^n . Sätt

$$\mathbf{v}_1 = \mathbf{u}_1$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2$$

\vdots

$$\mathbf{v}_p = \mathbf{u}_p - \frac{\mathbf{u}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{u}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \dots - \frac{\mathbf{u}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}} \mathbf{v}_{p-1}$$

Då är $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ en ortogonal bas för W . Dessutom gäller att

$$\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \text{ för } 1 \leq k \leq p$$

.

Lay 6.4, tmv165/185 V6, Vt06 bild 24

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

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

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

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

Kolumnerna 1, 2 och 4 är pivotkolumner i trappstegsmatrisen. Dessa kolumner i matrisen A ger en bas för $\text{Col}(A)$.

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

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

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

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

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

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

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

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

Sats 8 ger att den sökta projektionen är

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

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

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

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

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

Det skulle kunna vara så att vektorerna $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ faktiskt inte är en ortogonal bas för $\text{Col}(A)$. 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. \mathbf{z} skall vara ortogonal mot $\text{Col}(A)$ och alltså ligga i $\text{Nul}(A^T)$.

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

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.

Definition:

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

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

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

Lay 6.5, tmv165/185 V6, Vt06 bild 25

Sats 13

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

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

Lay 6.5, tmv165/185 V6, Vt06 bild 26

Med en **symmetrisk matris** menas en matris A sådan att $A^T = A$.
Obs! En symmetrisk matris måste vara kvadratisk, men det räcker inte
OBS! I kapitlet menas alltid att A är en reell matris.

Lay 7.1, tmv165/185 V6, Vt06 bild 27

Sats 3: Spektralsatsen för reella symmetriska matriser

Om A är en reell, symmetrisk, $n \times n$ -matris så gäller:

- A har n reella egenvärden om man räknar multiplicitet.
- För varje egenvärde är egenrummets dimension samma som egenvärdets multiplicitet som rot till karakteristiska ekvationen.
- Egenrummen är parvis ortogonala, alltså egenvektorer som hör till olika egenvärden är ortogonala.
- A kan diagonaliseras med en ortogonal matris.

Lay 7.1, tmv165/185 V6, Vt06 bild 28

Sats 1 = Sats 3 c.

Lay 7.1, tmv165/185 V6, Vt06 bild 29

Sats 2

A är en reell, symmetrisk, $n \times n$ -matris om och endast om A kan diagonaliseras med en ortogonal matris.

Skillnaden på sats 2 och 3d. är att sats 2 säger att symmetriska matriser och **inga andra** kan diagonaliseras med ortogonal matris.

Lay 7.1, tmv165/185 V6, Vt06 bild 30

Om $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ är en ON-bas av egenvektorer till den symmetriska matrisen A och $\lambda_1 \dots \lambda_n$ är motsvarande egenvärden så är

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

Detta kallas en spektral uppdelning av A

Lay 7.1, tmv165/185 V6, Vt06 bild 31

Om U har ortogonala kolonner kallas UU^T för en projektionsmatris.

Speciellt kallas $\mathbf{u}\mathbf{u}^T$ för en projektionsmatris om $\|\mathbf{u}\| = 1$

Projektionsformeln i sats 10 kap.6 visar att $UU^T \mathbf{x}$ är projektionen av \mathbf{x} på $\text{Col}(U)$.

Lay 7.1, tmv165/185 V6, Vt06 bild 32