

TMV166/186
Linjär Algebra M/TD
2009/2010
Läsvecka 6

Omfattning

- Lay, kapitel 6.1-6.6, Ortogonalitet och minsta-kvadrat metoden
- 7.1-7.2, Symmetriska matriser och kvadratiske former
- **Dugga 3:** Fredag 26/2, 13.30 - 14.10.
Omfattar kapitel 1, 2, 3, 4 och 5.

Innehåll

- Ortogonalitet
- Projektioner
- Minstakvadrat-metoden.
- Diagonalisering av symmetriska matriser
- Kvadratiske former

6.1 Skalarprodukt, norm, ortogonalitet

Viktiga begrepp

- Skalarprodukt
- Längd, norm
- Avstånd
- Ortogonalitet
- Vinklar

Mål 6.1

För betyget godkänd skall du kunna:

- beräkna skalarprodukten av två vektorer i $\mathbb{R}^n$
- tillämpa räkneregler för skalarprodukt,
- beräkna norm av en vektor i $\mathbb{R}^n$
- beräkna avståndet mellan vektorer i $\mathbb{R}^n$
- avgöra om två vektorer i $\mathbb{R}^n$ är ortogonala
- bevisa Pythagoras sats i $\mathbb{R}^n$
- förklara vad som menas med $W^\perp$ om W är ett underrum i $\mathbb{R}^n$
- tillämpa sats 6.1.3 vid problemlösning.

För högre betyg skall du dessutom kunna:

- bevisa sats 6.1.3.

Skalarprodukt, 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

skalarprodukten 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

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

Sats 6.1.1: 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:

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

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 rottecknet.

Definition:

- En **enhetsvektor** eller **normerad** vektor är en vektor $\hat{\mathbf{v}}$ sådan att $\|\hat{\mathbf{v}}\| = 1$
- 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 $1/\|\mathbf{v}\|$.
- En normerad vektor i riktningen $\mathbf{v}$ är

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

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

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

Definition:

- Två vektorer $\mathbf{u}$ och $\mathbf{v}$ i $\mathbb{R}^n$ kallas **ortogonal** (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.

Sats 6.1.2 Pythagoras sats

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

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

Definition: Ortogonal komplementet

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

Viktiga fakta:

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

Sats 6.1.3

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

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

- och

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

6.2 Ortogonal mängder

Viktiga begrepp

- Ortogonal mängd av vektorer
- Ortogonal bas
- Ortogonal projektion
- Ortonormerad mängd av vektorer
- Ortonormerad bas

Mål 6.2

För betyget godkänd skall du kunna:

- förklara vad som menas med **ortogonal bas** för ett underrum W
- tillämpa sats 6.2.5 för beräkning av koordinaterna för en vektor $y \in W$ relativt en ortogonal bas för W .
- använda projektnormalsformeln 6.2.(2) i problemlösning
- förklara vad som menas med **ortonormerad bas** för ett underrum W .
- förklara vad som menas med en **ortogonal matris**.

För högre betyg skall du dessutom kunna:

- bevisa sats 6.2.4: en ortogonal mängd av vektorer är linjärt oberoende.
- bevisa sats 6.2.7 då U är en ortogonal matris.

Definition: Ortogonal mängd

- En mängd vektorer $\{u_1, \dots, u_p\}$ kallas för en **ortogonal mängd** om vektorerna är **parvis ortogonala**.
- Alltså om alla par av vektorer (med olika index) är ortogonala mot varandra:

$$u_i \cdot u_j = 0 \quad \text{om} \quad i \neq j.$$
- En ortogonal mängd kallas **ortonormerad** om alla vektorerna i mängden dessutom är normerade.

Sats 6.2.4

- Om $S = \{u_1, \dots, u_p\}$ 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 .

Definition: ortogonal bas

- 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.

Sats 6.2.5

- Låt $\mathcal{B} = \{u_1, \dots, u_p\}$ vara en ortogonal bas för ett underrum W i $\mathbb{R}^n$. Och låt y vara en vektor i W .

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

ges då av
$$c_j = \frac{y \cdot u_j}{u_j \cdot u_j}$$

Om $\mathcal{B}$ är en **ortonormerad** bas ges koordinaterna av $c_j = y \cdot u_j$

Projektionsformeln

- Låt $u \neq 0$ vara en given vektor i $\mathbb{R}^n$. Låt y vara en annan vektor i $\mathbb{R}^n$.
- Vi söker en uppdelning av y i två ortogonala komponenter, $y = \hat{y} + z$ där $\hat{y} = \alpha u$ är parallell med u och z är ortogonalt mot u .
- Då är
$$\hat{y} = \left(\frac{y \cdot u}{u \cdot u} \right) u \quad \text{och} \quad z = y - \hat{y}$$
- Notera att $\hat{y}$ är parallell med linjen L genom origo med riktningsvektor u . Därför kan $\hat{y}$ kallas projektionen av y på L .
- Man använder ibland beteckningen $\text{proj}_{u,y}$ eller $\text{proj}_L y$ för projektionen.

Cauchys olikhet

- Låt x och y vara vektorer i $\mathbb{R}^n$.
Då gäller:

$$|x \cdot y| \leq \|x\| \|y\|$$

Bevis: $y = \hat{y} + z$ där $\hat{y}$ är projektionen av y på L .

Pythagoras sats ger: $\|y\|^2 \leq \|\hat{y}\|^2 + \|z\|^2 = \left\| \left(\frac{y \cdot x}{x \cdot x} \right) x \right\|^2 = \frac{|y \cdot x|^2}{\|x\|^2} \|x\|^2 = \frac{|y \cdot x|^2}{\|x\|^2}$

Således $\frac{|y \cdot x|}{\|x\|} \leq \|y\|$ som är ekvivalent med Cauchys olikhet.

Kommentar

- Av Cauchy-Schwartz olikhet följer att
$$-1 \leq \frac{x \cdot y}{\|x\| \|y\|} \leq 1$$
.
- Man kan därför **definiera vinkeln** θ mellan x och y genom
$$\theta = \arccos \left(\frac{x \cdot y}{\|x\| \|y\|} \right)$$
- 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".

Triangelolikheten

- För alla vektorer $\mathbf{x}$ och $\mathbf{y}$ i $\mathbb{R}^n$ gäller

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Bevis:

$$(\|\mathbf{x}\| + \|\mathbf{y}\|)^2 - \|\mathbf{x} + \mathbf{y}\|^2 = 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| - 2\mathbf{x} \cdot \mathbf{y} \geq 0$$

Sats 6.2.6

- 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.

Sats 6.2.7

- 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 $\mathbf{x} \cdot \mathbf{y} = 0$

Kommentar 1.

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

Kommentar 2

- Satsen är speciellt intressant då U är en kvadratisk $n \times n$ -matris. U är då en ortogonal matris och kolonnerna i U bildar en ortonormerad bas för $\mathbb{R}^n$.
- Om vi kallar denna bas $\mathcal{B}$ är $U = P_{\mathcal{B}}$ koordinatbytesmatris mellan standardbasen i $\mathbb{R}^n$ och $\mathcal{B}$.

$$\mathbf{x} = P_{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$$
- Satsen innebär då att längder och vinklar kan beräknas med hjälp av koordinatvektorer såväl i standardbas som i en ortonormerad bas.

Kommentar 3

- Om vi har tre vektorer $\mathbf{u}$, $\mathbf{v}$ och $\mathbf{w}$ i $\mathbb{R}^n$ så kan vi alltid (se exempel till Gram-Schmidt proceduren) välja en ON-bas $\mathcal{B}$ för $\mathbb{R}^n$ där

$$[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad [\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad [\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

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

Kommentar 4

- Satsen säger då att normer och skalärprodukter för $\mathbf{u}$, $\mathbf{v}$ och $\mathbf{w}$ är samma som normer och skalärprodukter för koordinatvektorer $[\mathbf{u}]_B$, $[\mathbf{v}]_B$, och $[\mathbf{w}]_B$.
- De senare är samma som norm och skalär produkt för,

$$\begin{bmatrix} a_1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} \text{ och } \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \text{ i } \mathbb{R}^3.$$

- I $\mathbb{R}^3$ ges den geometriska längden av normen och vinkeln mellan vektorer av

$$\theta = \arccos\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}\right)$$

- Det är därför logiskt att definiera vinkel i $\mathbb{R}^n$ med denna formel.
- 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.

6.3 Ortogonal projektioner

Viktiga begrepp

- Ortogonal projektion

Mål 6.3

För betyget godkänd skall du kunna:

- tillämpa sats 6.3.8 för att dela upp en vektor i ortogonala komponenter, en i W och den andra i $W^\perp$ då en ortogonal bas för W är känd.

För högre betyg skall du dessutom kunna:

- Inga ytterligare mål i detta avsnitt

Sats 6.3.8

Satsen om ortogonal uppdelning.

- Låt W vara ett underrum i $\mathbb{R}^n$ och låt $B = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ 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 projektnsformeln $\hat{\mathbf{y}} = \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1 + \dots + \left(\frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\right) \mathbf{u}_p$ och $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$

Kommentar

- Vi skall senare se att man alltid kan bestämma en ortogonal bas för ett underrum. Man kan därför alltid bestämma $\text{proj}_W \mathbf{y}$
- Kom ihåg! Basen som används i projektnsformeln måste vara ortogonal.

Sats 6.3.9

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 $\mathbf{v} \neq \hat{\mathbf{y}}$.

Kommentar

- 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 6.3.10

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

6.4 Gram-Schmidt proceduren

Viktiga begrepp

- Gram-Schmidt proceduren**
- QR-faktorisering ingår inte**

Mål 6.4

För betyget godkänd skall du kunna:

- tillämpa Gram-Schmidt processen för att bestämma en ortogonal bas för ett underrum W i $\mathbb{R}^n$ utgående från en annan bas för W .

För högre betyg skall du dessutom kunna:

- förklara varför Gram-Schmidt processen leder till en ortogonal bas.

Sats 6.4.11 Gram-Schmidt ortogonaliseringsprocedur

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

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1 \\ \mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 \\ \mathbf{v}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 \\ &\vdots \\ \mathbf{v}_p &= \mathbf{u}_p - \frac{\mathbf{u}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{u}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \dots - \frac{\mathbf{u}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}} \mathbf{v}_{p-1} \end{aligned}$$
- Då är $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ en ortogonal bas för W .
- Dessutom gäller att $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ för $1 \leq k \leq p$.

Kommentar

- Vi har en följd av underrum $W_1 = \text{Span}\{\mathbf{u}_1\}$, $W_2 = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, $\dots$, $W_k = \text{Span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, $\dots$, $W_p = \text{Span}\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$.
- Om vi har en ortogonal bas $\mathcal{B}_k = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ för W_k kan vi beräkna projektionen av $\mathbf{u}_{k+1}$ på W_k .
- Den nya vektorn $\mathbf{v}_{k+1}$ är "den del av $\mathbf{u}_{k+1}$ som är ortogonal mot W_k ".
- Alltså $\mathbf{v}_{k+1} = \mathbf{u}_{k+1} - \text{proj}_{W_k} \mathbf{u}_{k+1}$.

Exempel:

Vektorerna $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 3 \\ -1 \\ 3 \\ -1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 4 \end{bmatrix}$ är linjärt oberoende

Bestäm först en vektor $\mathbf{u}_4$ så att $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ är en bas för $\mathbb{R}^4$
 Bestäm sedan en ortonormerad bas $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ för $\mathbb{R}^4$ sådan att
 $\text{Span}\{\mathbf{v}_1\} = \text{Span}\{\mathbf{u}_1\}$, $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\} = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ och
 $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$

Bestäm slutligen koordinatvektorena $[\mathbf{u}_1]_{\mathcal{C}}$, $[\mathbf{u}_2]_{\mathcal{C}}$, $[\mathbf{u}_3]_{\mathcal{C}}$ och $[\mathbf{u}_4]_{\mathcal{C}}$.

6.5 Minstakvadrat-metoden

Viktiga begrepp

- Minstakvadrat-metoden

Mål 6.5

För betyget godkänd skall du kunna:

- förklara vad som menas med en minstakvadrat-lösning

För högre betyg skall du dessutom kunna:

- förklara varför minstakvadrat-lösningarna är lösningarna till den normaliserade ekvationen $A^T A \mathbf{x} = A^T \mathbf{b}$.

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

Sats 6.5.13

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

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

Sats 6.5.14

- Matrisen $A^T A$ är inverterbar om och endast om kolonnerna i A är linjärt oberoende.
- I så fall har ekvationen $A\mathbf{x} = \mathbf{b}$ exakt en minsta-kvadrat lösning.
- Denna ges av $\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$
- Notera: Satsen ger **ingen kalkylmässig fördel**, tvärtom.

6.6 Linjära modeller

Viktiga begrepp

- Kurvanpassning, modellanpassning

Mål 6.6

För betyget godkänd skall du kunna:

- tillämpa minstakvadrat-metoden för modellanpassning.

För högre betyg skall du dessutom kunna:

- tillämpa minstakvadrat-metoden för lite besvärligare modellanpassning

7.1 Diagonalisering av symmetriska matriser

Viktiga begrepp

- Symmetrisk matris
- Ortogonal diagonalisering
- Spektralsatsen

Mål 7.1

För betyget godkänd skall du kunna:

- tillämpa satserna 7.1.1 -- 7.1.3 vid problemlösning. Spektralsatsen (sats 7.1.3) är extra viktig.

För högre betyg skall du dessutom kunna:

- Inga ytterligare mål i detta avsnitt

Definition

- 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.

Sats 7.1.3: 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.

Sats 7.1.2

- 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.

Spektral uppdelning

- 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

Projektionsmatriser

- Om U har ortogonala, normerade kolonner kallas UU^T för en projektionsmatris.
- Speciellt kallas $\mathbf{u}\mathbf{u}^T$ för en projektionsmatris om $\|\mathbf{u}\| = 1$
- Projektionsformeln i sats 6.3.10 visar att $UU^T \mathbf{x}$ är projektionen av $\mathbf{x}$ på $\text{Col}(U)$.

7.2 Kvadratiska former

Viktiga begrepp

- **Kvadratisk form**
- **Huvudaxlar, principalaxlar**
- **Positivt definit, negativt definit, form**

Mål 7.2

För betyget godkänd skall du kunna:

- tillämpa sats 7.2.4 **Satsen om principalaxlar** för att skriva en kvadratisk form utan blandade termer.

För högre betyg skall du dessutom kunna:

- förklara vad som menas med positivt definit, negativt definit och indefinit kvadratisk form
- tillämpa sats 7.2.5 för klassificering av kvadratiska former.

Definition

- En **kvadratisk form** på $\mathbb{R}^n$ är en funktion $Q : \mathbb{R}^n \rightarrow \mathbb{R}$

som ges av ett uttryck

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$$

där A är en **symmetrisk** matris.

Matrisen A kallas den kvadratiska formens koefficientmatris.

Kvadratisk form och koordinatbyte

- Ett basbyte med koordinatbytesmatris $P = P_{\mathcal{B}}$ ger ett koordinat- eller **variabelbyte** $\mathbf{x} = P\mathbf{y}$ där $\mathbf{y} = [\mathbf{x}]_{\mathcal{B}}$.
- I basen $\mathcal{B}$ ges den kvadratiske formen av en annan matris:

$$\mathbf{x}^T A \mathbf{x} = (P\mathbf{y})^T A (P\mathbf{y}) = \mathbf{y}^T (P^T A P) \mathbf{y}.$$
- Matrisen $P^T A P$ är också symmetrisk.
 Den kallas för den kvadratiske formens koefficientmatris relativt basen $\mathcal{B}$.

Sats 7.2.4 Principalaxelsatsen.

- Låt A vara en symmetrisk $n \times n$ -matris och låt $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ vara en ortonormerad bas av egenvektorer till A , $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$.
- Låt P vara den ortogonala matrisen vars kolonner är vektorerna i $\mathcal{B}$
 $P = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n]$, låt D vara diagonalmatrisen
 $D = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. Då är $D = P^T A P$.
- Variabelbytet $\mathbf{x} = P\mathbf{y}$ där $\mathbf{y} = [\mathbf{x}]_{\mathcal{B}}$ överför då den kvadratiske formen $\mathbf{x}^T A \mathbf{x}$ till:

$$\mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$
- Kolonnerna i P , som alltså är egenvektorer till A , kallas den kvadratiske formens **principalaxlar**.

Definition:

- En kvadratisk form $Q(\mathbf{x})$ kallas:
- **positivt definit** om $Q(\mathbf{x}) > 0$ för alla $\mathbf{x} \neq \mathbf{0}$
- **negativt definit** om $Q(\mathbf{x}) < 0$ för alla $\mathbf{x} \neq \mathbf{0}$
- **indefinit** om $Q(\mathbf{x})$ antar såväl positiva som negativa värden.
- **positivt semidefinit** om $Q(\mathbf{x}) \geq 0$ för alla $\mathbf{x} \neq \mathbf{0}$
- **negativt semidefinit** om $Q(\mathbf{x}) \leq 0$ för alla $\mathbf{x} \neq \mathbf{0}$

Sats 7.2.5 Kvadratiske former och egenvärden.

- Låt A vara en symmetrisk $n \times n$ -matris.
 Då är den kvadratiske formen $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$:
- a. positivt definit om och endast om alla egenvärden till A är positiva,
- b. negativt definit om och endast om alla egenvärden till A är negativa,
- c. indefinit om och endast om A har såväl positiva som negativa egenvärden,
- d. positivt semidefinit om och endast om alla egenvärden till A är ≥ 0 ,
- e. negativt semidefinit om och endast om alla egenvärden till A är ≤ 0 .