1 Løsning og mindste kvadraters løsninger af lineære ligningssystemer

Def. Linære ligningssystemer

Et lineært ligningssystem er et system af *m* ligninger i *n* ubekendte, hvor disse kan skrives som:

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$
$$\ldots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

Disse systemer kan også skrives på matrix form således:

$$A\underline{x} = \underline{b}$$

Hvor løsningsmængden for systemer er de $\underline{x}$ for hvilket systemet er konsistent. Hvis systemet ingen løsninger har, er løsningsmængden tom – det siges at systemet er inkonsistent.

Teorem 1.2.1 [L]

Et $m \times n$ homogent system af lineære ligninger har en ikke-triviel løsning hvis $n > m$.

Bevis Teorem 1.2.1 [L]

Et homogent system er altid konsistent ($\underline{x} = \underline{0}$ er altid en løsning). REF formen af matricen har højst m ikke-nul rækker. Derved er der også maksimalt m pivoter. Siden der er n ubekendte og $n > m$ så må der være en eller flere frie variabler. De frie variabler kan derfor blive givet arbitrære værdier og for enhver værdi af de frie variable er der en løsning til systemet.

Def. Mindste kvadrat

Mindste kvadrater er en metode vi bruger til at lave et best fit af en given mængde data, således at man kan strukturere disse som eksempelvis en ret linje eller noget tilsvarende.
Givet et $m \times n$ lineært ligningssystem $Ax = b$ med $m > n$ (overdetermineret), så kan vi ikke generelt forvente at finde et $x \in \mathbb{R}^n$ for hvilket $Ax = b$. Derfor hvis $b \in \mathbb{R}^m$, så for ethvert $x \in \mathbb{R}^n$ kan vi forme et residual:

$$r(x) = b - Ax$$

Distancen mellem b og Ax er givet ved:

$$\|b - Ax\| = \|r(x)\|$$

Hvor vi ønsker at finde en vektor $x \in \mathbb{R}^n$ for hvilket $\|r(x)\|$ vil være minimal. En vektor $\hat{x}$ der gør dette siges at være en mindste kvadraters løsning til systemet $Ax = b$. Det følgende teorem garanterer at en sådan tætteste vektor p ikke blot eksisterer, men er unik:

Teorem 5.3.1 [L]

Lad $S \subset \mathbb{R}^m$ være et underrum og $b \in \mathbb{R}^m$. Da er projektionen p af b på S unik og tættest på b. Altså:

$$\|b-y\| > \|b-p\| \forall y \in S \wedge y \neq p$$

En vektor $p \in S$ vil desuden være tættest på en vektor $b \in \mathbb{R}^m \Leftrightarrow b-p \in S^{\perp}$

Bevis Teorem 5.3.1

Vi ved at $\mathbb{R}^m = S \oplus S^{\perp}$. Vi kan derfor skrive $b \in \mathbb{R}^m$ unikt som summen:

$$b = p + z$$

Hvor $p \in S$ og $z \in S^{\perp}$. Lad $y \in S \wedge y \neq p$, da gælder der
(fra b trækkes p, igen lægges p til og y trakkes fra):

$$\|b-y\|^2 = \|(b-p)+(p-y)\|^2$$

Da $b-p = z \in S^{\perp}$ og $p-y \in S$ giver Pythagoras' lov (lav eventuelt tegning):

$$\|b-y\|^2 = \|b-p\|^2 + \|p-y\|^2$$

hvor $\|p-y\|^2 > 0$ da $p \neq y$ hvilket giver

$$\|b-y\|^2 > \|b-p\|^2$$

For at vise at $b-p \in S^{\perp}$ hvis p er vektoren tættest på b, antag at $q \in S \wedge b-q \notin S^{\perp}$, da er $q \neq p$, og så følger samme argument som overstående (med $y = q$)

$$\|b-q\|^2 > \|b-p\|^2$$

Proposition 5.2.4 [N]

Systemet $A^T Ax = A^T b$ er konsistent, og z er en løsning til dette system $\Leftrightarrow$ z er en mindste kvadraters løsning til $Ax = b$.

Bevis Proposition 5.2.4

Lad $p = P_{Sø(A)}(b)$

$$\begin{aligned} Az = p &\Leftrightarrow b - Az \in Sø(A)^{\perp} \\ &\Leftrightarrow b - Az \in N(A^T) \\ &\Leftrightarrow A^T(b-Az) = 0 \\ &\Leftrightarrow A^T b - A^T Az = 0 \\ &\Leftrightarrow A^T Az = A^T b \end{aligned}$$

Det vil sige z er en løsning til $A^T Az = A^T b \Leftrightarrow Az = p$ hvor p er mindste kvadraters løsning til $Ax = b$.

Teorem 2.2.9 [N]

Lad V være et F vektorrum, lad $v_1,...,v_n \in V$ og lad $S = span(v_1,...,v_n)$. Et element $v \in S$ kan udtrykkes entydigt som en lineær kombination af $v_1,...,v_n \Leftrightarrow v_1,...,v_n$ er lineært uafhængige.

Bevis Teorem 2.2.9

Da $v \in span(v_1,...,v_n)$ kan vi skrive $v = a_1v_1 + ... + a_nv_n$ med $a_1,...,a_n \in \mathrm{F}$.

Lad os antage at v også kan skrives som $v = b_1v_1 + ... + b_nv_n$.

Så er

$$\begin{aligned} 0 &= v - v \\ &= (a_1v_1 + ... + a_nv_n) - (b_1v_1 + ... + b_nv_n) \\ &= (a_1 - b_1)v_1 + ... + (a_n - b_n)v_n \end{aligned}$$

Hvis $v_1,...,v_n$ er lineært uafhængige vil ligningen $0 = c_1v_1 + ... + c_nv_n \Rightarrow c_1 = 0,...,c_n = v_n$. Dette vil sige at $a_i - b_i = 0 \Rightarrow a_i = b_i$ så v kan skrives unikt som en lineær kombination af $v_1,...,v_n$.

Hvis $v_1,...,v_n$ er lineært afhænge vil ligningen $0 = c_1v_1 + ... + c_nv_n$ have en ikke-triviel løsning $\hat{c}_1,...,\hat{c}_n$ så $0 = \hat{c}_1v_1 + ... + \hat{c}_nv_n$ og vi får at

$$\begin{aligned} v &= v + 0 \\ &= a_1v_1 + ... + a_nv_n + \hat{c}_1v_1 + ... + \hat{c}_nv_n \\ &= (a_1 + \hat{c}_1)v_1 + ... + (a_n + \hat{c}_n)v_n \end{aligned}$$

Hvor $a_i + \hat{c}_i \neq a_i$ for et eller andet $i = 1,...,n$. Så v kan ikke skrives unikt hvis $v_1,...,v_n$ er lineært afhængige.

2 Vektorrum og underrum

Def. Underum

Et ikke-tomt delmængde S af et F vektorum V kaldes et underum hvis:
(tilsammen kaldet lukkethedsegenskaber)

$C1: \quad \forall y \in S, a \in \mathrm{F} : ay \in S$ (skalamultiplikation)

$C2: \quad \forall y, w \in S : y + w \in S$ (vekteraddition)

Def. Linear transformation

Lad V, W være F vektorum.

En lineær transformation $L : V \to W$ er en afbildning som respekterer lineære strukturer, det vil sige følgende er opfyldt:

$$\forall v_1, v_2 \in V, \forall a_1, a_2 \in \mathrm{F} : \quad L(a_1 v_1 + a_2 v_2) = a_1 L(v_1) + a_2 L(v_2)$$

Def. Ker

Lad $L : V \to W$ være en linear transformation. Kernen af L er da

$$Ker(L) = \{ v \in V \mid L(v) = 0_W \}$$

Def. Billede

Lad $L : V \to W$ være en linear transformation, lad $S \subset V$ være et underrum, da er billedemængden

$$L(S) = \{ w \in W \mid w = L(v), v \in S \}$$

Teorem 4.1.1 [L]

Hvis $L : V \to W$ er en lineær transformation og S et underrum af V . Så er

1. $Ker(L)$ et underrum af V
2. $L(S)$ et underrum af W

Bevis teorem 4.1.1

For 1:

$Ker(L)$ er en ikke-tom mængde da $0_V \in V$ og $0_W = L(0_V)$.

C1 (skalarmultiplikation): Lad $v \in Ker(V), a \in \mathrm{F}$, da gælder $\quad L(av) = aL(v) = a0_w = 0_w \in Ker(L)$

C2 (vektoradition): $v, w \in Ker(L)$, da gælder $\quad L(v + w) = L(v) + L(w) = 0_W + 0_W = 0_W \in Ker(L)$

$Ker(L)$ er da et underum.

For 2: $L(S)$ er ikke-tom da $0_W = L(0_V) \in L(S)$.

C1: Lad $a \in \mathrm{F}, w = L(v)$ for et $v \in S$, da gælder $aw = aL(v) = L(av)$. Da $av \in S$ så $L(av) \in L(S)$

C2: Lad $w_1 = L(v_1), w_2 = L(v_2)$ for et $v_1, v_2 \in S$, da gælder $w_1 + w_2 = L(v_1) + L(v_2) = L(v_1 + v_2)$ og da $v_1 + v_2 \in S$ så $L(v_1 + v_2) \in L(S)$

$L(S)$ er da et underrum.

Def. Span

Lad V være et F vektorrum, $v_1, \ldots, v_n \in V$ udspænder V ($V = span(v_1, \ldots, v_n)$) hvis og kun hvis enhver vektor $v \in V$ kan skrives som en linear kombination af $v_1, \ldots, v_n \in V$

Teorem 3.4.1 [L]

Hvis $span(v_1, \ldots, v_n) = V$, så er ethvert sæt af $m > n$ vektorer tilhørende V være lineært afhængige.

Bevis teorem 3.4.1

Det muligt at opskrive u_i for $i = 1, \ldots, m$ som en linear kombination af $v_1, \ldots, v_n$.

$$u_i = \sum_{j=1}^{n} a_{ij} v_j$$

For en vektor $x \in V$, da kan x skrives som en linear kombination af u_i.

$$\begin{aligned} x &= x_1 u_1 + \ldots + x_m u_m \\ &= \sum_{i=1}^{m} x_i u_i \\ &= \sum_{i=1}^{m} x_i \sum_{j=1}^{n} a_{ij} v_j \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_i a_{ij} \right) v_j \end{aligned}$$

Lad os prøve at løse det homogene ligningssystem $\sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_i a_{ij} \right) = 0$. Dette er et ligningssystem med m variable (ubekendte) og n ligninger. Da $m > n$ er der flere ubekendte end der er ligninger. Da eksisterer der en ikke-triviel løsning $c = (c_1, \ldots, c_m)^T$. Dette betyder at følgende gælder

$$c_1 u_1 + \ldots + c_m u_m = \sum_{i=1}^{m} c_i u_i = \sum_{j=1}^{n} 0 v_j = 0$$

Derved er $u_1, \ldots, u_m$ lineært afhængige.

Teorem 2.2.9 [N]

Lad V være et F vektorrum, lad $v_1,...,v_n \in V$ og lad $S = span(v_1,...,v_n)$. Et element $v \in S$ kan udtrykkes entydigt som en lineær kombination af $v_1,...,v_n \Leftrightarrow v_1,...,v_n$ er lineært uafhængige.

Bevis Teorem 2.2.9

Da $v \in span(v_1,...,v_n)$ kan vi skrive $v = a_1v_1 + ... + a_nv_n$ med $a_1,...,a_n \in \mathrm{F}$.

Lad os antage at v også kan skrives som $v = b_1v_1 + ... + b_nv_n$.

Så er

$$\begin{aligned} 0 &= v - v \\ &= (a_1v_1 + ... + a_nv_n) - (b_1v_1 + ... + b_nv_n) \\ &= (a_1 - b_1)v_1 + ... + (a_n - b_n)v_n \end{aligned}$$

Hvis $v_1,...,v_n$ er lineært uafhængige vil ligningen $0 = c_1v_1 + ... + c_nv_n \Rightarrow c_1 = 0,...,c_n = v_n$. Dette vil sige at $a_i - b_i = 0 \Rightarrow a_i = b_i$ så v kan skrives unikt som en lineær kombination af $v_1,...,v_n$.

Hvis $v_1,...,v_n$ er lineært afhænge vil ligningen $0 = c_1v_1 + ... + c_nv_n$ have en ikke-triviel løsning $\hat{c}_1,...,\hat{c}_n$ så $0 = \hat{c}_1v_1 + ... + \hat{c}_nv_n$ og vi får at

$$\begin{aligned} v &= v + 0 \\ &= a_1v_1 + ... + a_nv_n + \hat{c}_1v_1 + ... + \hat{c}_nv_n \\ &= (a_1 + \hat{c}_1)v_1 + ... + (a_n + \hat{c}_n)v_n \end{aligned}$$

Hvor $a_i + \hat{c}_i \neq a_i$ for et eller andet $i = 1,...,n$. Så v kan ikke skrives unikt hvis $v_1,...,v_n$ er lineært afhængige.

3 Lineær uafhængighed

Def. Lineær uafhængighed

Lad V være et F vektorrum, og lad $v_1,\ldots,v_n \in V$

Et sæt $v_1,\ldots,v_n$ af vektorer er lineært uafhængige $\Leftrightarrow c_1v_1+\ldots+c_nv_n=0 \Rightarrow c_1=0,\ldots,c_n=0$

Hvis der findes en ikke-triviel løsning til ligningssystemet kaldes $v_1,\ldots,v_n$ afhængige og en af vektorerne kan skrives som en linearkombination af de andre.

Def. Span

Lad V være et F vektorrum, $v_1,\ldots,v_n \in V$ udspænder V ($V = span(v_1,\ldots,v_n)$) hvis og kun hvis enhver vektor $v \in V$ kan skrives som en linear kombination af $v_1,\ldots,v_n \in V$

Teorem 3.3.1 [L]

Lad $x_1,\ldots,x_n \in \mathbb{R}^n$ og lad $X=(x_1,\ldots,x_n)$. Så gælder der

X er singulær $\Leftrightarrow x_1,\ldots,x_n$ er lineært afhængige

Bevis Teorem 3.3.1

Vi har en ligning

$$c_1x_1+\ldots+c_nx_n=0$$

Der kan skrives på matrixform:

$$Xc=0$$

Vi ved at hvis X er invertibel, så har ligningen $Xc=0$ kun løsningen $c=0$, da
$Xc=0 \Rightarrow X^{-1}Xc=X^{-1}0 \Rightarrow c=0$.
Vi ved også at hvis ligningen $Xc=0$ kun har løsningen $c=0$ så er X invertibel. (Sætning 1.4.8 [N])

Teorem 3.4.1 [L]

Hvis $span(v_1,...,v_n)=V$, så er ethvert sæt af $m>n$ vektorer tilhørende V være lineært afhængige.

Bevis teorem 3.4.1

Det muligt at opskrive u_i for $i=1,...,m$ som en linear kombination af $v_1,...,v_n$.

$$u_i = \sum_{j=1}^{n} a_{ij} v_j$$

For en vektor $x \in V$, da kan x skrives som en linear kombination af u_i.

$$\begin{aligned} x &= x_1u_1 + ... + x_m u_m \\ &= \sum_{i=1}^{m} x_i u_i \\ &= \sum_{i=1}^{m} x_i \sum_{j=1}^{n} a_{ij} v_j \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_i a_{ij} \right) v_j \end{aligned}$$

Lad os prøve at løse det homogene ligningssystem $\sum_{j=1}^{n}\left(\sum_{i=1}^{m} x_i a_{ij}\right)=0$. Dette er et ligningssystem med m variable (ubekendte) og n ligninger. Da $m>n$ er der flere ubekendte end der er ligninger. Da eksisterer der en ikke-triviel løsning $c=(c_1,...,c_m)^T$. Dette betyder at følgende gælder

$$c_1u_1 + ... + c_m u_m = \sum_{i=1}^{m} c_i u_i = \sum_{j=1}^{n} 0 v_j = 0$$

Derved er $u_1,...,u_m$ lineært afhængige.

Teorem 2.2.9 [N]

Lad V være et F vektorrum, lad $v_1,...,v_n \in V$ og lad $S = span(v_1,...,v_n)$. Et element $v \in S$ kan udtrykkes entydigt som en lineær kombination af $v_1,...,v_n \Leftrightarrow v_1,...,v_n$ er lineært uafhængige.

Bevis Teorem 2.2.9

Da $v \in span(v_1,...,v_n)$ kan vi skrive $v = a_1v_1 + ... + a_nv_n$ med $a_1,...,a_n \in \mathrm{F}$.

Lad os antage at v også kan skrives som $v = b_1v_1 + ... + b_nv_n$.

Så er

$$\begin{aligned} 0 &= v - v \\ &= (a_1v_1 + ... + a_nv_n) - (b_1v_1 + ... + b_nv_n) \\ &= (a_1 - b_1)v_1 + ... + (a_n - b_n)v_n \end{aligned}$$

Hvis $v_1,...,v_n$ er lineært uafhængige vil ligningen $0 = c_1v_1 + ... + c_nv_n \Rightarrow c_1 = 0,...,c_n = v_n$. Dette vil sige at $a_i - b_i = 0 \Rightarrow a_i = b_i$ så v kan skrives unikt som en lineær kombination af $v_1,...,v_n$.

Hvis $v_1,...,v_n$ er lineært afhænge vil ligningen $0 = c_1v_1 + ... + c_nv_n$ have en ikke-triviel løsning $\hat{c}_1,...,\hat{c}_n$ så $0 = \hat{c}_1v_1 + ... + \hat{c}_nv_n$ og vi får at

$$\begin{aligned} v &= v + 0 \\ &= a_1v_1 + ... + a_nv_n + \hat{c}_1v_1 + ... + \hat{c}_nv_n \\ &= (a_1 + \hat{c}_1)v_1 + ... + (a_n + \hat{c}_n)v_n \end{aligned}$$

Hvor $a_i + \hat{c}_i \neq a_i$ for et eller andet $i = 1,...,n$. Så v kan ikke skrives unikt hvis $v_1,...,v_n$ er lineært afhængige.

4 Basis for vektorrum; koordinatisering

Def. Ordnet basis

Lad V være et F vektorrum, og lad $v_1,...,v_n \in V$

$\{v_1,...,v_n\}$ er en ordnet basis for V hvis

1. $v_1,...,v_n$ er lineært uafhængige
2. $v_1,...,v_n$ udspænder V

Normalt er rækkefølgen irrelevant, men i visse tilfælde kan det være nødvendigt at have dem ordet (Hvilket det bl.a. er for koordinatisering)

Teorem 3.1.3 [N]

Lad V være et F vektorrum, $V \neq \{0\}$, som er udspændt af endelig mange vektorer $v_1,...,v_n$. Da har V en basis indeholdt i $\{v_1,...,v_n\}$

Bevis Teorem 3.1.3 [N]

Induktion i n. Altså induktion i antallet af vektorer i den udspændende mængde.
Basis $n=1$:
Vi ved at $V \neq \{0\}$ og at $V = span(v_1)$, derfor er $v_1 \neq 0$ så $\{v_1\}$ er lineært uafhængige og derfor en basis for V.

Induktionshypotese
Vi antager at udsagnet gælder for et vektorrum udspændt af $n-1$ vektorer.

Induktionsskridt
Vi vil gerne vise at det gælder for et vektorrum udspændt af n vektorer.
Lad $V = span(v_1,...,v_n)$.

Hvis $v_1,...,v_n$ er lineært uafhængige, så er $\{v_1,...,v_n\}$ en basis for V.

Hvis $v_1,...,v_n$ er lineært afhængige, så kan én af dem skrives som en lineærkombination af de andre. Efter omnummerering kan vi antage at $v_n = c_1v_1 + ... + c_{n-1}v_{n-1}$. Men så er $V = span(v_1,...,v_{n-1})$ og ifølge induktionshypotesen har V en basis indeholdt i $\{v_1,...,v_{n-1}\}$.

I begge tilfælde har V en basis indeholdt i $\{v_1,...,v_n\}$.

Def. Koordinatisering

Lad $\nu = \{v_1,...,v_n\}$ være en ordnet basis for V. Lad $v \in V$; v kan skrives entydigt som en lineær kombination af $v_1,...,v_n, v = c_1v_1 + ... + c_nv_n$.

Der findes således et entydigt element $\begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix} \in \mathrm{F}^n$, som angiver koordinaterne for v mht. V. Dette kaldes koordinatvektoren for v mht. V og vi skriver $[v]_v = \begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix}$

Lemma 3.3.2 [N]

Koordinatisering bevarer lineær struktur, dvs:

1. $[v+w]_v = [v]_v + [w]_v$
2. $[rv]_v = r[v]_v$

Bevis Lemma 3.3.2

1:

Lad $v, w \in v$, skriv $v = c_1v_1 + \dots + c_nv_n$, $w = d_1v_1 + \dots + d_nv_n$.

Da er $v + w = (c_1 + d_1)v_1 + \dots + (c_n + d_n)v_n$ og så gælder der

$$[v+w]_v = \begin{bmatrix} c_1 + d_n \\ \dots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \dots \\ d_n \end{bmatrix} = [v]_v + [w]_v$$

2:

Lad $v \in v$, skriv $v = c_1v_1 + \dots + c_nv_n$ da er $rv = rc_1v_1 + \dots + rc_nv_n$ og

$$[rv]_v = \begin{bmatrix} rc_1 \\ \dots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix} = r[v]_v$$

Proposition 3.3.7 [N]

Lad $U_b = \{u_1, \dots, u_n\}, V_b = \{v_1, \dots, v_n\}$ være to ordnede baser for et F vektorrum W.

Lad $K \in Mat_{n,n}(\mathrm{F})$ være givet ved søjleform som

$$\left[[u_1]_{V_b} \cdots [u_n]_{V_b} \right]$$

Der gælder

1. K er invertibel
2. $\forall w \in W : [w]_{V_b} = K[w]_{U_b}$
3. K er den entydige matrix i $Mat_{n,n}(\mathrm{F})$ således at (2) gælder.

Bevis Proposition 3.3.7

1:

Antag at $K\begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix} = 0$ så

$$\begin{aligned} 0 &= x_1\left[u_1\right]_{V_b} + \dots + x_n\left[u_n\right]_{V_b} \\ &= \left[x_1u_1 + \dots + x_nu_n\right]_{V_B} \end{aligned}$$

Idet at koordinatisering bevarer lineær struktur.

Så $x_1u_1 + \dots + x_nu_n = 0$ og da $u_1, \dots, u_n$ er lineært uafhængige, er den eneste løsning til $Kx = 0$ løsningen $x = 0$. Da $x = 0$ er den enste løsning er K invertibel.

2:

Vi udtrykker u_i som en lineær kombination af $v_1, \dots, v_n$

$$u_i = k_{1i}v_1 + \dots + k_{ni}v_n \text{ hvor}$$

$$\begin{bmatrix} k_{1i} \\ \dots \\ k_{ni} \end{bmatrix} = \left[u_i\right]_{V_b} \text{ som er den i'te søjle i } K$$

Skriv $w = c_1u_1 + \dots + c_nu_n$ så

$$[w]_{U_b} = \begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix}$$

Vi indsætter u_i

$$\begin{aligned} w &= c_1\left(k_{11}v_1 + \dots + k_{n1}v_n\right) + \dots + c_n\left(k_{1n}v_1 + \dots + k_{nn}v_n\right) \\ &= \left(c_1k_{11} + \dots + c_nk_{1n}\right)v_1 + \dots + \left(c_1k_{n1} + \dots + c_nk_{nn}\right)v_n \end{aligned}$$

så

$$[w]_{V_b} = \begin{bmatrix} k_{11}c_1 + \dots + k_{1n}c_n \\ \vdots \\ k_{n1}c_1 + \dots + k_{nn}c_n \end{bmatrix} = Kc = K\left[w\right]_{U_b}$$

3:

Antag at $\left[w\right]_{V_b} = K'\left[w\right]_{U_b}$ for alle $w \in W$

Dette gælder specielt for u_i for $i = 1,...,n$

$$\left\{1'te\, søjle\, i\, K\right\} = \left[u_i\right]_{V_b} = K'\left[u_i\right]_{U_b} = K'e_i = \left\{i'te\, søjle\, i\, K'\right\}$$

K og K' har altså samme søjler og derfor er $K' = K$.

Note for bevis:

$\left[u_i\right]_{U_b} = e_i = \begin{bmatrix} 0 \\ ... \\ 1\left(i'te\ plads\right) \\ ... \\ 0 \end{bmatrix}$ da u_i udtrykket i sin egen base er $1 \cdot u_i$.

5 Matricer og lineære transformationer

Def. Linear transformation

Lad V, W være F vektorrum.

En lineær transformation $L: V \to W$ er en afbildning som respekterer lineære strukturer, det vil sige følgende er opfyldt:

$$\forall v_1, v_2 \in V, \forall a_1, a_2 \in \mathrm{F}: \quad L(a_1 v_1 + a_2 v_2) = a_1 L(v_1) + a_2 L(v_2)$$

Sætning 4.2.1 [N]

Lad $L: \mathrm{F}^n \to \mathrm{F}^m$ være en lineær transformation. Definer $M(L) \in Mat_{m,n}(\mathrm{F})$ ved

$$M(L) = [L(e_1), \ldots, L(e_n)]$$

Så er $L(x) = M(L)x$ for alle $x \in \mathrm{F}^n$. $M(L)$ er den entydige matrix med denne egenskab.

Bevis Sætning 4.2.1

Lad $x \in \mathrm{F}^n$, x kan da skrives som $x = x_1 e_1 + \ldots + x_n e_n$.

Vi beregner der lineær transformation på denne

$$\begin{aligned} L(x) &= L(x_1 e_1 + \ldots + x_n e_n) \\ &= x_1 L(e_1) + \ldots + x_n L(e_n) \\ &= [L(e_1), \ldots, L(e_n)] \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix} \\ &= M(L)x \end{aligned}$$

Antag at der nu findes en anden matrix $M'(L) \in Mat_{m,n}(\mathrm{F})$ med $L(x) = M'(L)x$. Så gælder der at

$$\{i'te\, søjle\, i\, M(L)\} = L(e_i) = M'(L)e_i = \{i'te\, søjle\, i\, M'(L)\}$$

Søjlerne i $M(L)$ og $M'(L)$ er ens, og da er $M'(L) = M(L)$. $M(L)$ er da den entydige matrix med denne egenskab.

Diagram

Vi vil gerne se at dette diagram kommuterer:

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "\theta_{\mathcal{V}}"] & W \arrow[d, "\theta_{\mathcal{W}}"] \\ \mathbb{F}^n \arrow[r, "L_{M_{\mathcal{W},\mathcal{V}}(L)}"] & \mathbb{F}^m \end{tikzcd}$$

Sætning 4.2.4 [N]

Lad V,W være F vektorrum. Lad $V_b = \{v_1,..,v_n\}$ og $W_b = [w_1,...,w_n]$ være ordnede baser for V og W. Lad $L: V \to W$ være en linear transformation. Definer

$$M_{W_b,V_b}(L) = \left[\left[L(v_1)_{W_b} \right], ..., \left[L(v_n) \right]_{W_b} \right]$$

Så er $\left[L(v) \right]_{W_b} = M_{W_b,V_b}(L)[v]_{V_b}$ for alle $v \in V$, og at $M_{W_b,V_b}(L)$ er den entydige matrix med denne egenskab.

Bevis Sætning 4.2.2

Lad $v \in V$ så kan v skrives som $v = x_1 v_1 + ... + x_n v_n$ og derved er

$$[v]_{V_b} = \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix}$$

Lav nu den lineære transformation på v.

$$\begin{aligned} L(v) &= L(x_1 v_1 + ... + x_n v_n) \\ &= x_1 L(v_1) + ... + x_n L(v_n) \end{aligned}$$

Lad os nu se på koordinatiseringen af denne

$$\begin{aligned} \left[L(v) \right]_{W_b} &= \left[x_1 L(v_1) + ... + x_n L(v_n) \right]_{W_b} \\ &= x_1 \left[L(v_1) \right]_{W_b} + ... + x_n \left[L(v_n) \right]_{W_b} \\ &= \left[\left[L(v_1) \right]_{W_b}, ..., \left[L(v_n) \right]_{W_b} \right] \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix} \\ &= M_{W_b,V_b}(L)[v]_{V_b} \end{aligned}$$

Lad os nu antage der findes en anden matrix $M'_{W_b,V_b}(L) \in Mat_{m,n}(\mathrm{F})$ hvor $[L(v)]_{W_b} = M'_{W_b,V_b}(L)[x]_{V_b}$

Dette gælder specielt for $v_1,\dots,v_n$ så

$$\{i'te\,søjle\,i\,M_{W_b,V_b}(L)\} = [L(v_i)]_{W_b} = M'_{W_b,V_b}(L)[v_i]_{V_b} = M'_{W_b,V_b}(L)e_i = \{i'te\,søjle\,i\,M'_{W_b,V_b}(L)\}$$

Så søjlerne i $M_{W_b,V_b}(L)$ og $M'_{W_b,V_b}(L)$ er ens og derfor er $M'_{W_b,V_b}(L) = M_{W_b,V_b}(L)$. Derfor er $M_{W_b,V_b}(L)$ den entydige matrix med denne egenskab.

Def. Ker

Lad $L:V \to W$ være en linear transformation. Kernen af L er da

$$Ker(L) = \{v \in V \mid L(v) = 0_W\}$$

Def. Billede

Lad $L:V \to W$ være en linear transformation, lad $S \subset V$ være et underrum, da er billedemængden

$$L(S) = \{w \in W \mid w = L(v), v \in S\}$$

Teorem 4.1.1 [L]

Hvis $L:V \to W$ er en lineær transformation og S et underrum af V. Så er

3. $Ker(L)$ et underrum af V
4. $L(S)$ et underrum af W

Bevis teorem 4.1.1

For 1:

$Ker(L)$ er en ikke-tom mængde da $0_V \in V$ og $0_W = L(0_V)$.

C1 (skalarmultiplikation): Lad $v \in Ker(V), a \in \mathrm{F}$, da gælder $L(av) = aL(v) = a0_w = 0_w \in Ker(L)$

C2 (vektoradition): $v,w \in Ker(L)$, da gælder $L(v+w) = L(v) + L(w) = 0_W + 0_W = 0_W \in Ker(L)$

$Ker(L)$ er da et underum.

For 2:

$L(S)$ er ikke-tom da $0_W = L(0_V) \in L(S)$.

C1: Lad $a \in \mathrm{F}, w = L(v)$ for et $v \in S$, da gælder $aw = aL(v) = L(av)$. Da $av \in S$ så $L(av) \in L(S)$

C2: Lad $w_1 = L(v_1), w_2 = L(v_2)$ for et $v_1, v_2 \in S$, da gælder $w_1 + w_2 = L(v_1) + L(v_2) = L(v_1 + v_2)$ og da $v_1 + v_2 \in S$ så $L(v_1 + v_2) \in L(S)$

$L(S)$ er da et underrum.

6 Determinanter

Def. Determinant

Til enhver $n \times n$ matrix er det muligt at associere en skalar $\det(A)$.

For en 1×1 matrix er $\det(A) = a_{11}$.

For e $n \times n$ matrix er $\det(A) = a_{11}A_{11} + ... + a_{1n}A_{1n}$ hvor $A_{ij} = (-1)^{i+j}\det(M_{ij})$ og M_{ij} er matricen A med den i'te række og j'te søjle slettet.

A_{ij} kaldes kofaktor for A og M_{ij} kaldes minorer af A

Det er ligegyldigt om hvilken række eller søjle vi udvikler fra

Teorem 2.1.2 [L]

Lad $A \in Mat_{n,n}(\mathrm{F})$, da gælder at $\det(A) = \det(A^T)$

Bevis Teorem 2.1.2

Dette er et induktionsbevis over n
Basis 1×1:
Dette gælder trivialt da $A = A^T \Rightarrow \det(A) = \det(A^T)$

Induktionshypotese:
Vi antager at det gælder for alle $k-1 \times k-1$ matricer.

Induktionsskridt:
Vi skal vise det gælder for alle $k \times k$ matricer. Lad $A \in Mat_{k,k}(\mathrm{F})$ da er

$$\det(A) = a_{11}\det(M_{11}) + a_{12}\det(M_{12}) - + ... \pm a_{1n}\det(M_{1n})$$

Her er alle M_{ij} matricer $k-1 \times k-1$ og vi kan bruge induktionshypotesen:

$$\det(A) = a_{11}\det(M_{11}{}^T) + a_{12}\det(M_{12}{}^T) - + ... \pm a_{1n}\det(M_{1n}{}^T)$$

Da $a_{11}\det(M_{11}{}^T) + a_{12}\det(M_{12}{}^T) - + ... \pm a_{1n}\det(M_{1n}{}^T)$ blot er $\det(A^T)$ udviklet efter første søjle:

$$\det(A) = \det(A^T)$$

Teorem 2.2.2 [L]

$A \in Mat_{n,n}(\mathbb{F})$ er singulær $\Leftrightarrow \det(A) = 0$

Bevis Teorem 2.2.2

Det er muligt at rækkereducere A til RREF H ved endeligt mange rækkeoperationer, så

$$H = E_k \dots E_1 A$$

Vi tager nu determinanten på begge sider

$$\begin{aligned}\det(H) &= \det(E_k \dots E_1 A) \\ &= \det(E_k)\dots\det(E_1)\det(A)\end{aligned}$$

Da $\det(E_i) \neq 0$ for $i = 1,\dots,k$ fordi E_i er en elementær rækkeoperation, så gælder der at $\det(H) = 0 \Leftrightarrow \det(A) = 0$. Hvis A er singulær så vil H have en nul-række og vi kan udvikle langs denne og få $\det(H) = 0$. Hvis A er invertibel så vil $H = I$ og da er $\det(H) = 1$.

Teorem 2.3.1 [L]

(Cramers regel)

Lad $A \in Mat_{n,n}(\mathbb{F})$ være invertibel og lad $b \in \mathbb{F}^n$. Lad A_i være matricen der fås ved erstatte den i'te søjle i A med b. Den entydige løsning $\hat{x}$ til systemet $Ax = b$ er givet ved;

$$\hat{x}_i = \frac{\det(A_i)}{\det(A)}, i = 1,\dots,n$$

Bevis Teorem 2.3.1

$A^{-1}b$ er den entydige løsning til $Ax = b$. Så vi har

$$\hat{x} = A^{-1}b = \frac{1}{\det(A)} adj(A) b$$

Og for $i = 1,\dots,n$

$$\begin{aligned}x_i &= \frac{1}{\det(A)}\left(i\text{'te række i } adj(A)\right)b \\ &= \frac{1}{\det(A)}\left(A_{1i}b_1 + \dots + A_{ni}b_n\right) \\ &= \frac{\det(A_i)}{\det(A)}\end{aligned}$$

7 Egenværdier og egenvektorer

Def. Egenværdi og egenvektor

Lad $A \in Mat_{n,n}(\mathrm{F}), \lambda \in \mathrm{F}$ er en egenværdi for A hvis der findes $x \in \mathrm{F} \setminus \{0\}$, så

$$Ax = \lambda x$$

x kaldes da for en egenvektoren til A associeret til λ.

Def. Egenrum

Ved omskrivning af definitionen ovenfor får vi

$$(A - \lambda I)x = 0$$

Løsningsrummet $N(A - \lambda I)$ kaldes for egenrummet til matricen A. Egenrummet består af alle egenvektorer til en given egenværdi.

Def. Similaritet

Lad $A, B \in Mat_{n,n}(\mathrm{F})$. B siges at være similar til A hvis der eksisterer en matrix $S \in Mat_{n,n}(\mathrm{F})$ så:
$B = S^{-1}AS$

Teorem 6.1.1 [L]

Lad $A, B \in Mat_{n,n}(\mathrm{F})$ være similare. Da er $p_A = p_B$

Bevis Teorem 6.1.1

Da A og B er similare kan B skrives som $B = S^{-1}AS$.
Lad os beregne $p_B(\lambda)$

$$\begin{aligned}
p_B(\lambda) &= \det(B - \lambda I) \\
&= \det\left(S^{-1}AS - \lambda I\right) \\
&= \det\left(S^{-1}(A - \lambda I)S\right) \\
&= \det\left(S^{-1}\right)\det(A - \lambda I) \\
&= \det\left(S^{-1}S\right)\det(A - \lambda I) \\
&= \det(I)\det(A - \lambda I) \\
&= \det(A - \lambda I) \\
&= p_A(\lambda)
\end{aligned}$$

Def. Diagonaliserbar

Lad $A \in Mat_{n,n}(\mathrm{F})$. A er diagonaliserbar hvis der eksisterer en invertibel matrix $V \in Mat_{n,n}(\mathrm{F})$ så

$$D = V^{-1}AV$$

Hvor D er diagonal.

Teorem 6.3.2 [L]

Lad $A \in Mat_{n,n}(\mathrm{F})$. A er diagonaliserbar $\Leftrightarrow A$ har n linear uafhængige egenvektorer

Bevis Teorem 6.3.2

$\Leftarrow$

Lad $x_1, \ldots, x_n$ være linear uafhængige egenvektorer til A med tilhørende egenværdi λ_i. Lad $X = [x_1, \ldots, x_n]$ da er

$$\begin{aligned} AX &= [Ax_1, \ldots, Ax_n] \\ &= [\lambda_1 x_1, \ldots, \lambda_n x_n] \\ &= [x_1, \ldots, x_n]\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \\ &= XD \end{aligned}$$

Da X har n lineært uafhængige søjlevektorer, så er X invertibel. Så

$$AX = XD \Rightarrow X^{-1}AX = X^{-1}XD \Rightarrow X^{-1}AX = D$$

A er altså diagonaliserbar.

$\Rightarrow$

A er nu diagonaliserbar og derfor

$$X^{-1}AX = D \Rightarrow \left(XX^{-1}\right)AX = XD \Rightarrow AX = XD$$

Lad $X = [x_1, \ldots, x_n], D = \begin{bmatrix} d_{11} & & \\ & \ddots & \\ & & d_{nn} \end{bmatrix}$ så

$$XD = [x_1, \ldots, x_n]\begin{bmatrix} d_{11} & & \\ & \ddots & \\ & & d_{nn} \end{bmatrix} = [x_1 d_{11}, \ldots, x_n d_{nn}]$$

Da $AX = XD$ (vi kan gange med e_1 på begge sider for at pille første søjle ud) må $d_{11},...,d_{nn}$ må være egenværdier for A med tilhørende egenvektorer $x_1,...,x_n$. Fordi X er invertibel er $x_1,...,x_n$ linear uafhængige.

8 Diagonalisering

Def. Egenværdi og egenvektor

Lad $A \in Mat_{n,n}(\mathrm{F}), \lambda \in \mathrm{F}$ er en egenværdi for A hvis der findes $x \in \mathrm{F} \setminus \{0\}$, så

$$Ax = \lambda x$$

x kaldes da for en egenvektoren til A associeret til λ.

Def. Diagonaliserbar

Lad $A \in Mat_{n,n}(\mathrm{F})$. A er diagonaliserbar hvis der eksisterer en invertibel matrix $V \in Mat_{n,n}(\mathrm{F})$ så

$$D = V^{-1}AV$$

Hvor D er diagonal.

Teorem 6.3.2 [L]

Lad $A \in Mat_{n,n}(\mathrm{F})$. A er diagonaliserbar $\Leftrightarrow A$ har n linear uafhængige egenvektorer

Bevis Teorem 6.3.2

$\Leftarrow$

Lad $x_1, \ldots, x_n$ være linear uafhængige egenvektorer til A med tilhørende egenværdi λ_i. Lad $X = [x_1, \ldots, x_n]$ da er

$$\begin{aligned} AX &= [Ax_1, \ldots, Ax_n] \\ &= [\lambda_1 x_1, \ldots, \lambda_n x_n] \\ &= [x_1, \ldots, x_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \\ &= XD \end{aligned}$$

Da X har n lineært uafhængige søjlevektorer, så er X invertibel. Så

$$AX = XD \Rightarrow X^{-1}AX = X^{-1}XD \Rightarrow X^{-1}AX = D$$

A er altså diagonaliserbar.

$\Rightarrow$

A er nu diagonaliserbar og derfor

$$X^{-1}AX = D \Rightarrow (XX^{-1})AX = XD \Rightarrow AX = XD$$

Lad $X = [x_1, ..., x_n], D = \begin{bmatrix} d_{11} & & \\ & \ddots & \\ & & d_{nn} \end{bmatrix}$ så

$$XD = [x_1, ..., x_n] \begin{bmatrix} d_{11} & & \\ & \ddots & \\ & & d_{nn} \end{bmatrix} = [x_1 d_{11}, ..., x_n d_{nn}]$$

Da $AX = XD$ (vi kan gange med e_1 på begge sider for at pille første søjle ud) må $d_{11}, ..., d_{nn}$ må være egenværdier for A med tilhørende egenvektorer $x_1, ..., x_n$. Fordi X er invertibel er $x_1, ..., x_n$ linear uafhængige.

Def. Unitær

En matrix $U \in Mat_{n,n}(\mathbb{C})$ er unitær, hvis søjlerne i U udgør en ortonormalbasis for $\mathbb{C}^n$.

Def. Hermitisk matrix

En matrix $A \in Mat_{n,n}(\mathbb{C})$ er hermite'sk hvis $A = A^H = (\bar{A})^T$

Teorem 11.1.9 [N]

(Schurs teorem)

Lad $A \in Mat_{n,n}(\mathbb{C})$. Der findes en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så $U^H AU$ er en øvretrekantsmatrix.

Bevis Teorem 11.1.9

Beviser er ved induktion over n.

Basis $n = 1$

Trivielt idet alle 1×1 matricer er øvretriangulære.

Induktionshypotese

Vi antager at udsagnet gælder for alle matricer i $Mat_{n-1,n-1}(\mathbb{C})$.

Induktionsskridt

Lad $A \in Mat_{n,n}(\mathbb{C})$ og lad $\lambda_1 \in \mathbb{C}$ være en egenværdi for A, med tilsvarende egenvektor $v_1 \in \mathbb{C}^n$.

Vi kan arrangere at $\|v_1\| = 1$. $\{v_1\}$ kan udvides til en basis for $\mathbb{C}^n$, og ved at anvende Gram-schmidt på denne basis, får vi en ortonormalbasis $\{v_1, ..., v_n\}$ for $\mathbb{C}^n$.

Lad $U_1 = [v_1, ..., v_n] \in Mat_{n,n}(\mathbb{C})$. U er da unitær ved konstruktion.

Vi har da (vi piller først søjle af $U_1^H AU_1$ ud)

$$U_1^H AU_1 e_1 = U_1^H Av_1 = \lambda_1 U_1^H v_1 = \lambda_1 U_1^H U_1 e_1 = \lambda_1 e_1$$

Vi kan altså skrive

$$U_1^H A U_1 = \begin{bmatrix} \lambda_1 & \underline{r} \\ \underline{0} & A_1 \end{bmatrix}$$

Hvor $\underline{0}$ er en søjlevektor og $\underline{r}$ er en rækkevektor med $n-1$ indgange.
Ifølge induktionshypotesen så eksisterer der en matrix $C \in Mat_{n-1,n-1}(\mathbb{C})$ så $C^H A_1 C$ er øvre en øvretrekantsmtrix.
Definer nu

$$U_2 = \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix}$$

Matricen U_2 er unitær da de sidste $n-1$ søjler udgør en ortonormal mængde, og hver for sig er ortogonale på den første søjle, som er en enhedsvektor. Så er produktet $U = U_1 U_2$ også unitær.

Vi vil vise at $U^H A U$ er øvre-triangulær. Da blokstørelserne i matricerne er ens kan vi bruge blok matrix multiplikation.

$$\begin{aligned} U^H A U &= U_2^H \left(U_1^H A U_1 \right) U_2 \\ &= \begin{bmatrix} 1 & 0^T \\ 0 & C^H \end{bmatrix} \begin{bmatrix} \lambda_1 & r \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & r \\ 0 & C^H A_1 \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & rC \\ 0 & C^H A_1 C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & rC \\ 0 & B \end{bmatrix} \end{aligned}$$

Denne matrix er øvre triangulær da B (fra induktionshypotesen) er det.

Teorem 11.1.10 [N]

(Spektralsætning for hermite'ske matricer)

Lad $A \in Mat_{n,n}(\mathbb{C})$ være hermite'sk. Så kan A diagonaliseres, dvs der findes en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så

$$U^H AU$$

Er diagonal med reelle diagonalindgange.

Bevis Teorem 11.1.10

Ifølge Schurs sætning findes der en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så $U^H AU = T$ er en øvretrekantsmatrix. T er hermite'sk da:

$$\begin{aligned} T^H &= \left(U^H AU\right)^H \\ &= U^H A^H U \\ &= U^H AU \\ &= T \end{aligned}$$

Men

$$T = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & t_{nn} \end{bmatrix}, \; T^H = \begin{bmatrix} \overline{t}_{11} & 0 & 0 \\ \vdots & \ddots & 0 \\ \overline{t}_{1n} & \cdots & \overline{t}_{nn} \end{bmatrix}$$

Da må $t_{i,j} = 0$ for $i \neq j$ og da $t_{i,i} = \overline{t}_{i,i}$ må $t_{i,i}$ være reel. T er da en diagonalmatrix med reelle indgange.

9 Indre produkt

Def. Indre produkt

Et indre produkt på et reelt vektorrum V er en funktion $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$, således at for alle $x, y, z \in V, a, b \in \mathbb{R}$:

1. $\langle x, x \rangle \geq 0$ med lighed $\Leftrightarrow v = 0$
2. $\langle x, y \rangle = \langle y, x \rangle$
3. $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$

Def. Norm og ortogonalitet

Lad V være et vektorrum med indre produkt $\langle\ ,\ \rangle$.

Normen, også kaldt længden af $x \in V$ er da defineret til

$$\|x\| = \sqrt{\langle x, x \rangle}$$

$x, y \in V$ er ortogonale hvis

$$\langle x, y \rangle = 0$$

Teorem 6.1.4 [N]

(Pythagoras)

Lad V være et reelt vektorrum med indre produkt $\langle\ ,\ \rangle$, og lad $u, v \in V$ være ortogonale. Da gælder

$\|u + v\|^2 = \|u\|^2 + \|v\|^2$

Bevis Teorem 6.1.4

Vi beregner $\|u + v\|^2$ ved brug af indre produkt $\langle\ ,\ \rangle$:

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v \rangle \\
&= \langle u, u \rangle + \langle v, v \rangle + 2\langle u, v \rangle \\
&= \langle u, u \rangle + \langle v, v \rangle \\
&= \|u\|^2 + \|v\|^2
\end{aligned}$$

$2\langle u, v \rangle$ forsvinder da $u \perp v$.

Def. Vektorprojektion

Lad V være et reelt vektorrum med indre produkt $\langle\ ,\ \rangle$, og lad $u, v \in V$ med $v \neq 0$ så er

$$p = \frac{\langle u, v\rangle}{\langle v, v\rangle} v$$

Vektorprojektionen af u på v.

Hjælpelemma 6.1.6 [N] (ingen bevis)

Lad V være et reelt vektorrum med indre produkt $\langle\ ,\ \rangle$, og lad $u, v \in V$ med $v \neq 0$ og lad p være vektorprojektionen af u på v.

1. $u - p, v$ er ortogonale
2. $u = p \Leftrightarrow u$ er et skalarmultiplum af v

Teorem 6.1.7 [N]

(Cauchy-Schwarz uligheden)

Lad V være et reelt vektorrum med indre produkt $\langle\ ,\ \rangle$, og lad $u, v \in V$. Da gælder

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

Og lighed gælder hvis, og kun hvis, u, v er lineært afhængige.

Bevis Teorem 6.1.7

Hvis $v = 0$, da er $|\langle u, v\rangle| = 0 = \|u\| \cdot \|v\|$ og u, v er klart lineært afhængige.

Hvis $v \neq 0$, så lad p være vektorprojektionen af u på v. $u - p, v$ er da ortogonale (ifølge lemma 6.1.6) og derfor er $u - p, p$ det også. Pythagoras giver da:

$$\|u\|^2 = \|u - p\|^2 + \|p\|^2$$

Hvilket giver at hvis man løser i forhold til $\|p\|^2$ og indsætter definitionen af $\|p\|^2$

$$\|u\|^2 - \|u - p\|^2 = \|p\|^2 = \frac{(\langle u, v\rangle)^2}{\|v\|^2}$$

Som kan løses i forhold til $\langle u, v\rangle$:

$$(\langle u, v\rangle)^2 = \|u\|^2 \cdot \|v\|^2 - \|u - p\|^2 \cdot \|v\|^2 \leq \|u\|^2 \cdot \|v\|^2$$

Hvilket da giver

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

Lighed gælder hvis, og kun hvis $\|u-p\|^2 = 0 \Leftrightarrow u = p$. Lemma 6.1.6 siger at så er u et skalarmultiplum af V , og derfor lineært afhængige.

10 Ortogonalkomplement og projektion

Def. Ortogonalkomplement

Lad $Y \subset \mathbb{R}^n$ være et underrum

$$Y^{\perp} = \left\{ x \in \mathbb{R}^n \mid x^T y = 0 \ \forall y \in Y \right\}$$

Kaldes ortogonalkomplementet til Y.

Hjælpelemma 5.1.16 [N]

$A \in Mat_{n,n}(\mathbb{R})$ så er

$$Sø(A)^{\perp} = N\left(A^T\right)$$

Teorem 5.1.17.1 [N]

Lad $S \subset \mathbb{R}^n$ være et underrum.

Da er $S^{\perp}$ også et underrum med $\dim\left(S^{\perp}\right) = n - \dim(S)$

Bevis Teorem 5.1.17.1

Hvis $S = \{0\}$, så er $S^{\perp} = \mathbb{R}^n$ og $\dim\left(S^{\perp}\right) = n - \dim(s) = n - 0 = n$.

Hvis $S \neq \{0\}$, lad $x_1, \ldots, x_r \in R^n$ være en basis for S. Lad $X = \left[x_1, \ldots, x_r\right]$. Da er $S = Sø(X)$. Derfor må $S^{\perp} = Sø(X)^{\perp}$ Lemma 5.1.16 siger da at $S^{\perp} = Sø(X)^{\perp} = N\left(X^T\right)$. $N\left(X^T\right)$ er et underrum (ifølge proposition 2.1.8).

Dimensionen udregnes:

$$\begin{aligned}
\dim\left(N\left(X^T\right)\right) &= antal\ søjler\ i\ X^T - rank\left(X^T\right) \\
&= n - rank(X) \\
&= n - r \\
&= n - \dim(S)
\end{aligned}$$

Teorem 5.1.17.2 [N]

Hvis $S \neq \{0\}, S \neq \mathbb{R}^n$, og $\{x_1, \dots, x_r\}$ er en basis for S og $\{x_{r+1}, \dots, x_n\}$ er en basis for $S^\perp$, så er $\{x_1, \dots, x_n\}$ en basis for $\mathbb{R}^n$.

Bevis Teorem 5.1.17.2

Ifølge sætning 3.1.4 er det nok at vise at $x_1, \dots, x_n$ er uafhængige. Antag at

$$c_1x_1 + \dots + c_rx_r + c_{r+1}x_{r+1} + \dots + c_nx_n = 0$$

Lad da $y = c_1x_1 + \dots + c_rx_r, z = c_{r+1}x_{r+1} + \dots + c_nx_n$. Så $y \in S, z \in S^\perp$. Så har vi

$$y + z = 0 \Rightarrow y = -z \Rightarrow y, z \in S \cap S^\perp = \{0\}$$

Dette betyder altså at:

For y: $c_1x_1 + \dots + c_rx_r = 0 \Rightarrow c_1 = 0, \dots, c_r = 0$

For w: $c_{r+1}x_{r+1} + \dots + c_nx_n = 0 \Rightarrow c_{r+1} = 0, \dots, c_n = 0$

Tilsammen giver dette:

$c_1x_1 + \dots + c_rx_r + c_{r+1}x_{r+1} + \dots + c_nx_n = 0 \Rightarrow c_1 = 0, \dots, c_n = 0$ og derved er $x_1, \dots, x_n$ uafhængige.

Teorem 5.5.7 [L]

Lad S være et underrum af et indreproduktrum V med indre produkt $\langle \ , \ \rangle$.

Lad $\{x_1,\ldots,x_n\}$ være en ortonormalbasis for S og lad $x \in V$.

Hvis

$$p = \sum_{i=1}^{n} c_i x_i$$

Hvor

$$c_i = \langle x_i, x \rangle$$

Så $p - x \in S^{\perp}$.

Bevis Teorem 5.5.7 [Ændret lidt]

Vi vil gerne vise at $\langle p-x, y \rangle = 0 \ \ \forall y \in S$

Lad $y = a_1 x_1 + \ldots + a_n x_n$ da:

$$\begin{aligned}
\langle p-x, y \rangle &= \langle p-x, a_1x_1 + \ldots + a_n x_n \rangle \\
&= \left\langle p-x, \sum_{i=1}^{n} a_i x_i \right\rangle \\
&= \sum_{i=1}^{n} a_i \langle p-x, x_i \rangle \\
&= \sum_{i=1}^{n} a_i \left(\langle p, x_i \rangle - \langle x, x_i \rangle \right) \\
&= \sum_{i=1}^{n} a_i \left(\left\langle \sum_{j=1}^{n} c_j x_j, x_i \right\rangle - c_i \right) \\
&= \sum_{i=1}^{n} a_i \left(\sum_{j=1}^{n} c_j \langle x_j, x_i \rangle - c_i \right) \\
&= \sum_{i=1}^{n} a_i \left(c_1 - c_i \right) \\
&= 0
\end{aligned}$$

11 Ortogonale og ortonormale baser

Def. Ortogonal mængde

Lad V være et indre-produkt rum. Lad $v_1,...,v_n \in V \setminus \{0\}$. Hvis $\langle v_i, v_j \rangle = 0, i \neq j$ er $v_1,...,v_n$ en ortogonal mængde.

Teorem 5.5.1 [L]

Lad V være et indre-produkt rum. Lad $\{v_1,...,v_n\} \subset V$ være en ortogonal mængde, så er $v_1,...,v_n$ lineært uafhængige.

Bevis Teorem 5.5.1

Antag at

$$c_1 v_1 + ... + c_n v_n = 0$$

For $j = 1,...,n$ tag det indre produkt på begge sider af udregningen

$$c_1 \langle v_1, v_j \rangle + ... + c_n \langle v_n, v_j \rangle = 0$$

Da $\langle v_i, v_j \rangle = 0$ når $j \neq i$ fås

$$c_j \langle v_j, v_j \rangle = 0$$
$$c_j \|v_j\|^2 = 0$$

Da v_j er en del af en ortogonal mængde er $v_j \neq 0$, hvilket medfører at $\|v_j\|^2 > 0$. Derved må $c_j = 0$ og $v_1,...,v_n$ er derved lineært uafhængige.

Def. Basis

Lad V være et F vektorrum, og lad $v_1,...,v_n \in V$

$v_1,...,v_n$ er en basis hvis

1. $v_1,...,v_n$ er lineært uafhængige
2. $v_1,...,v_n$ udspænder V

Def. Ortonormal mængde

En ortogonal mængde $\{u_1,...,u_n\}$ er ortonormal $\Leftrightarrow \langle u_i, u_j \rangle = \delta_{ij}$ hvor

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Teorem 5.6.1 [L]

(Gram-Schmidt processen)

Lad V være et indre-produktrum med indre produkt $\langle \ , \ \rangle$.

Lad $\{x_1,\ldots,x_n\}$ være en basis for V.

Definer

$$u_1 = \frac{1}{\|x_1\|} x_1$$

Og definer $u_2,\ldots,u_n$ rekursivt ved

$$u_{k+1} = \frac{1}{\|x_{k+1} - p_k\|}\left(x_{k+1} - p_k\right) \text{ for } k = 1,\ldots,n-1$$

Hvor

$$p_k = \langle x_{k+1}, u_1 \rangle u_1 + \ldots + \langle x_{k+1}, u_k \rangle u_k$$

Altså den ortogonale projektion af x_{k+1} på $span(u_1,\ldots,u_k)$.

Da er $\{u_1,\ldots,u_k\}$ en ortonormal basis for $span(x_1,\ldots,x_k)$ for $k = 1,\ldots,n$. Specielt er $\{u_1,\ldots,u_n\}$ en ortonormal basis for V.

Bevis Teorem 5.6.1

Lad $S_k = span(x_1,\ldots,x_k)$ for $k = 1,\ldots,n$.

Induktion i k:

Basis $k = 1$:

Det er klart at u_1 er en enhedsvektor med samme retning som x_1 og at $span(u_1) = span(x_1) = S_1$.

Induktionshypotese:

Antag at $\{u_1,\ldots,u_k\}$ er en ortonormal basis for $span(x_1,\ldots,x_k) = S_k$ for $k < n$

Induktionsskridt:

Lad p_k være projektionen af x_k på $S_k = span(u_1,\ldots,u_k)$. Sætning 6.4.14 giver da

$$p_k = \langle x_{k+1}, u_1 \rangle u_1 + \ldots + \langle x_{k+1}, u_k \rangle u_k$$

Da $p_k \in S_k$ kan den skrives som en linear kombination af $x_1,\ldots,x_k$.

$$p_k = c_1 x_1 + \ldots + c_k x_k$$

og

$$x_{k+1} - p_k = x_{k+1} + c_1 x_1 + \ldots + c_k x_k \neq 0$$

Da højresiden er en ikke-triviel ($c'er$ er ikke alle 0) linearkombination af $x_1,...,x_{k+1}$. Desuden er

$$x_{k+1} - p_k \in span(x_1,...,x_{k+1}) = S_{k+1}$$

Ifølge sætning 6.4.14 så er $x_{k+1} - p_k \in S_k^\perp$ og derved er $x_{k+1} - p_k \perp u_i$ for $i = 1,...,k$.

Lad nu $u_k = \frac{1}{\|x_{k+1} - p_k\|}(x_{k+1} - p_k)$. Så er $\{u_1,...,u_{k+1}\}$ en ortonormal mængde og er indeholdt i S_{k+1} .

Da $u_1,...,u_{k+1}$ er $k+1$ uafhængige elementer i rummet S_{k+1} af dimension $k+1$ udgør de en basis, og $\{u_1,...,u_{k+1}\}$'er en ortonormal basis for S_{k+1} .

Induktionsskridtet er taget og resultatet dermed bevist.

12 Ortogonale og unitære matricer

Def. Ortonormalt mængde

En ortogonal mængde $\{u_1,\ldots,u_n\}$ er ortonormal $\Leftrightarrow \langle u_i,u_j\rangle = \delta_{ij}$ hvor

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Def. Ortogonalmatrix

En matrix $Q \in Mat_{n,n}(\mathbb{R})$ er en ortogonalmatrix hvis søjlerne i Q udgør en ortonormalbasis for $\mathbb{R}^n$

Teorem 5.5.5 [L]

$Q \in Mat_{n,n}(\mathbb{R})$ er en ortogonalmatrix $\Leftrightarrow Q^T Q = I$

Bevis Teorem 5.5.5

Lad os beregne $Q^T Q$ i,j'te indgang.

$$\left(Q^T Q\right)_{ij} = q_i^T q_j = \langle q_i, q_j\rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Derfor gælder det at $Q \in Mat_{n,n}(\mathbb{R})$ er en ortogonalmatrix $\Leftrightarrow Q^T Q = I$.

Def. Unitær

En matrix $U \in Mat_{n,n}(\mathbb{C})$ er unitær, hvis søjlerne i U udgør en ortonormalbasis for $\mathbb{C}^n$.

Def. Hermitisk matrix

En matrix $A \in Mat_{n,n}(\mathbb{C})$ er hermite'sk hvis $A = A^H = \left(\bar{A}\right)^T$

Teorem 11.1.9 [N]

(Schurs teorem)

Lad $A \in Mat_{n,n}(\mathbb{C})$. Der findes en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så $U^H AU$ er en øvretrekantsmatrix.

Bevis Teorem 11.1.9

Beviser er ved induktion over n.

Basis $n = 1$

Trivielt idet alle 1×1 matricer er øvretriangulære.

Induktionshypotese

Vi antager at udsagnet gælder for alle matricer i $Mat_{n-1,n-1}(\mathbb{C})$.

Induktionsskridt

Lad $A \in Mat_{n,n}(\mathbb{C})$ og lad $\lambda_1 \in \mathbb{C}$ være en egenværdi for A, med tilsvarende egenvektor $v_1 \in \mathbb{C}^n$.

Vi kan arrangere at $\|v_1\| = 1$. $\{v_1\}$ kan udvides til en basis for $\mathbb{C}^n$, og ved at anvende Gram-schmidt på denne basis, får vi en ortonormalbasis $\{v_1, ..., v_n\}$ for $\mathbb{C}^n$.

Lad $U_1 = [v_1, ..., v_n] \in Mat_{n,n}(\mathbb{C})$. U er da unitær ved konstruktion.

Vi har da (vi piller først søjle af $U_1^H A U_1$ ud)

$$U_1^H A U_1 e_1 = U_1^H A v_1 = \lambda_1 U_1^H v_1 = \lambda_1 U_1^H U_1 e_1 = \lambda_1 e_1$$

Vi kan altså skrive

$$U_1^H A U_1 = \begin{bmatrix} \lambda_1 & \underline{r} \\ \underline{0} & A_1 \end{bmatrix}$$

Hvor $\underline{0}$ er en søjlevektor og $\underline{r}$ er en rækkevektor med $n-1$ indgange.

Ifølge induktionshypotesen så eksisterer der en matrix $C \in Mat_{n-1,n-1}(\mathbb{C})$ så $C^H A_1 C$ er øvre en øvretrekantsmtrix.

Definer nu

$$U_2 = \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix}$$

Matricen U_2 er unitær da de sidste $n-1$ søjler udgør en ortonormal mængde, og hver for sig er ortogonale på den første søjle, som er en enhedsvektor. Så er produktet $U = U_1 U_2$ også unitær.

Vi vil vise at $U^H AU$ er øvre-triangulær. Da blokstørelserne i matricerne er ens kan vi bruge blok matrix multiplikation.

$$\begin{aligned} U^H AU &= U_2^H \left(U_1^H A U_1\right) U_2 \\ &= \begin{bmatrix} 1 & 0^T \\ 0 & C^H \end{bmatrix} \begin{bmatrix} \lambda_1 & r \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & r \\ 0 & C^H A_1 \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & rC \\ 0 & C^H A_1 C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & rC \\ 0 & B \end{bmatrix} \end{aligned}$$

Denne matrix er øvre triangulær da B (fra induktionshypotesen) er det.

Teorem 11.1.10 [N]

(Spektralsætning for hermite'ske matricer)

Lad $A \in Mat_{n,n}(\mathbb{C})$ være hermite'sk. Så kan A diagonaliseres, dvs der findes en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så

$$U^H AU$$

Er diagonal med reelle diagonalindgange.

Bevis Teorem 11.1.10

Ifølge Schurs sætning findes der en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så $U^H AU = T$ er en øvretrekantsmatrix. T er hermite'sk da:

$$\begin{aligned} T^H &= \left(U^H AU\right)^H \\ &= U^H A^H U \\ &= U^H AU \\ &= T \end{aligned}$$

Men

$$T = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & t_{nn} \end{bmatrix}, \; T^H = \begin{bmatrix} \overline{t}_{11} & 0 & 0 \\ \vdots & \ddots & 0 \\ \overline{t}_{1n} & \cdots & \overline{t}_{nn} \end{bmatrix}$$

Da må $t_{i,j} = 0$ for $i \neq j$ og da $t_{i,i} = \overline{t}_{i,i}$ må $t_{i,i}$ være reel. T er da en diagonalmatrix med reelle indgange.

13 Unitær diagonalisering

Def. Unitær

En matrix $U \in Mat_{n,n}(\mathbb{C})$ er unitær, hvis søjlerne i U udgør en ortonormalbasis for $\mathbb{C}^n$.

Def. Hermitisk matrix

En matrix $A \in Mat_{n,n}(\mathbb{C})$ er hermite'sk hvis $A = A^H = \left(\overline{A}\right)^T$

Teorem 11.1.9 [N]

(Schurs teorem)

Lad $A \in Mat_{n,n}(\mathbb{C})$. Der findes en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så $U^H AU$ er en øvretrekantsmatrix.

Bevis Teorem 11.1.9

Beviser er ved induktion over n.

Basis $n = 1$

Trivielt idet alle 1×1 matricer er øvretriangulære.

Induktionshypotese

Vi antager at udsagnet gælder for alle matricer i $Mat_{n-1,n-1}(\mathbb{C})$.

Induktionsskridt

Lad $A \in Mat_{n,n}(\mathbb{C})$ og lad $\lambda_1 \in \mathbb{C}$ være en egenværdi for A, med tilsvarende egenvektor $v_1 \in \mathbb{C}^n$.

Vi kan arrangere at $\|v_1\| = 1$. $\{v_1\}$ kan udvides til en basis for $\mathbb{C}^n$, og ved at anvende Gram-schmidt på denne basis, får vi en ortonormalbasis $\{v_1, \ldots, v_n\}$ for $\mathbb{C}^n$.

Lad $U_1 = [v_1, \ldots, v_n] \in Mat_{n,n}(\mathbb{C})$. U er da unitær ved konstruktion.

Vi har da (vi piller først søjle af $U_1^H AU_1$ ud)

$$U_1^H AU_1 e_1 = U_1^H Av_1 = \lambda_1 U_1^H v_1 = \lambda_1 U_1^H U_1 e_1 = \lambda_1 e_1$$

Vi kan altså skrive

$$U_1^H AU_1 = \begin{bmatrix} \lambda_1 & \underline{r} \\ \underline{0} & A_1 \end{bmatrix}$$

Hvor $\underline{0}$ er en søjlevektor og $\underline{r}$ er en rækkevektor med $n-1$ indgange.

Ifølge induktionshypotesen så eksisterer der en matrix $C \in Mat_{n-1,n-1}(\mathbb{C})$ så $C^H A_1 C$ er øvre en øvretrekantsmatrix.

Definer nu

$$U_2 = \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix}$$

Matricen U_2 er unitær da de sidste $n-1$ søjler udgør en ortonormal mængde, og hver for sig er ortogonale på den første søjle, som er en enhedsvektor. Så er produktet $U = U_1 U_2$ også unitær.

Vi vil vise at $U^H AU$ er øvre-triangulær. Da blokstørelserne i matricerne er ens kan vi bruge blok matrix multiplikation.

$$\begin{aligned} U^H AU &= U_2^H \left(U_1^H AU_1\right)U_2 \\ &= \begin{bmatrix} 1 & 0^T \\ 0 & C^H \end{bmatrix} \begin{bmatrix} \lambda_1 & r \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & r \\ 0 & C^H A_1 \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & rC \\ 0 & C^H A_1 C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & rC \\ 0 & B \end{bmatrix} \end{aligned}$$

Denne matrix er øvre triangulær da B (fra induktionshypotesen) er det.

Teorem 11.1.10 [N]

(Spektralsætning for hermite'ske matricer)

Lad $A \in Mat_{n,n}(\mathbb{C})$ være hermite'sk. Så kan A diagonaliseres, dvs der findes en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så

$$U^H AU$$

Er diagonal med reelle diagonalindgange.

Bevis Teorem 11.1.10

Ifølge Schurs sætning findes der en unitær matrix $U \in Mat_{n,n}(\mathbb{C})$ så $U^H AU = T$ er en øvretrekantsmatrix. T er hermite'sk da:

$$\begin{aligned} T^H &= \left(U^H AU\right)^H \\ &= U^H A^H U \\ &= U^H AU \\ &= T \end{aligned}$$

Men

$$T = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & t_{nn} \end{bmatrix}, T^H = \begin{bmatrix} \overline{t}_{11} & 0 & 0 \\ \vdots & \ddots & 0 \\ \overline{t}_{1n} & \cdots & \overline{t}_{nn} \end{bmatrix}$$

Da må $t_{i,j} = 0$ for $i \neq j$ og da $t_{i,i} = \overline{t}_{i,i}$ må $t_{i,i}$ være reel. T er da en diagonalmatrix med reelle indgange.

14 Kvadratiske former

Def. Kvadratisk form

En kvadratisk form er en funktion $f:\mathbb{R}^n \to \mathbb{R}$ givet ved $f(x) = x^T Ax$, hvor $A \in Mat_{n,n}(\mathbb{R})$.

Grunden til den kaldes en kvadratisk form er:

Hvis

$$x = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}, A = \begin{bmatrix} a_{ij} \end{bmatrix}$$

Så

$$\begin{aligned} f\left(\begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}\right) &= \begin{bmatrix} x_1 \dots x_n \end{bmatrix} \begin{bmatrix} a_{11} & & a_{1n} \\ & & \\ a_{n1} & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 \dots x_n \end{bmatrix} \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \dots \\ a_{n1}x_1 + \dots + a_{nn}x_n \end{bmatrix} \\ &= x_1\left(a_{11}x_1 + \dots + a_{1n}x_n\right) + \dots + x_n\left(a_{n1}x_1 + \dots + a_{nn}x_n\right) \end{aligned}$$

Vi ser at i alle led indgår der mindst to x'er, hvilket er grunden til at den kaldes en kvadratisk form.

Eksempel

$$2xy - y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} y \\ x - y \end{bmatrix} = 2xy - y^2$$

Teorem 11.3.3

(Hovedaksesætningen)

Lad $A \in Mat_{n,n}(\mathbb{R})$ være symmetrisk, og lad $\lambda_1, \dots, \lambda_n$ være A's egenværdier, talt med multiplicitet.

Der findes en rotationsmatrix $R \in Mat_{n,n}(\mathbb{R})$ og en tilsvarende rotation af koordinater $u = R^T x$ så

$$x^T Ax = \lambda_1 u_1^2 + \dots + \lambda_n u_n^2$$

For alle $x \in \mathbb{R}^n$

Bevis teorem 11.3.3

Ifølge spektralsætningen så findes der en ortonormal basis $\{v_1,...,v_n\}$ for $\mathbb{R}^n$ bestående af A's egenvektorer. Så $[v_1,...,v_n]$ er en ortonormal matrix.

Lad

$$R = \begin{cases} [v_1,...,v_n] & \text{hvis} \det([v_1,...,v_n]) = 1 \\ [-v_1,...,v_n] & \text{hvis} \det([v_1,...,v_n]) = -1 \end{cases}$$

R er da en rotationsmatrix, hvis søjler er egenvektorer for A og som udgør en ortonormal basis for $\mathbb{R}^n$.
Vi har da

$$R^T AR = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

Hvor $\lambda_1,...,\lambda_n$ er egenværdierne til $\pm v_1,...,v_n$, så

$$A = R \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} R^T$$

Vi har da for alle $x \in \mathbb{R}^n$

$$\begin{aligned} x^T Ax &= x^T R \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} R^T x \\ &= u^T \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} u \qquad (\text{hvor}\, u = R^T x) \\ &= \lambda_1 u_1^2 + ... + \lambda_n u_n^2 \end{aligned}$$

Def. Positiv definit

Lad $A \in Mat_{n,n}(\mathbb{R})$

A er positiv definit hvis:
$\forall x \in \mathbb{R}^n \setminus \{0\} : x^T Ax > 0$

Teorem 6.6.2 [L]

Lad $A \in Mat_{n,n}(\mathbb{R})$ være symmetrisk.

A er positiv definit $\Leftrightarrow$ A's egenværdier alle er positive.

Bevis Teorem 6.6.2

$\Rightarrow$

Lad λ være en egenværdi for A med tilhørende egenvektor x. Da

$$\begin{aligned} x^T Ax &= x^T \lambda x \\ &= \lambda x^T x \\ &= \lambda \|x\|^2 \end{aligned}$$

Dette medfører

$$\lambda = \frac{x^T Ax}{\|x\|^2} > 0$$

Det vil sige at enhver egenværdi er positiv.

$\Leftarrow$

Da A symmetrisk kan A ifølge spektralsætningen diagonaliseres og der eksisterer en ortonormalbasis $\{x_1, \ldots, x_n\}$ for $\mathbb{R}^n$ hvor $x_1, \ldots, x_n$ er A's egenvektorer.

Lad $x \in \mathbb{R}^n \setminus \{0\}$ så x kan skrives som en linear kombination af $x_1, \ldots, x_n$

$$x = a_1 x_1 + \ldots + a_n x_n$$

Hvor $a_i = \langle x, x_i \rangle$ for $i = 1, \ldots, n$ og $\sum_{i=1}^{n} a_i^2 = \|x\|^2$ (bare beregn $\langle x, x \rangle = \langle a_1 x_1 + \ldots + a_n x_n, a_1 x_1 + \ldots + a_n x_n \rangle$)

$$\begin{aligned} x^T Ax &= (a_1 x_1 + \ldots + a_n x_n)^T A(a_1 x_1 + \ldots + a_n x_n) \\ &= (a_1 x_1 + \ldots + a_n x_n)^T (a_1 Ax_1 + \ldots + a_n Ax_n) \\ &= (a_1 x_1^T + \ldots + a_n x_n^T)(a_1 \lambda_1 x_1 + \ldots + a_n \lambda_n x_n) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_j a_i \lambda_i x_j^T x_i \\ &= a_1^2 \lambda_1 + \ldots + a_n^2 \lambda \\ &= \sum_{i=1}^{n} a_i^2 \lambda_i \\ &= \|x\|^2 \sum_{i=1}^{n} \lambda_i \\ &\geq \|x\|^2 \min(\lambda_i) \\ &> 0 \end{aligned}$$

(Husk at $x_i^T x_j = 0 \quad hvis\, i \neq j$

15 Lineære differentialligninger

Def. Lineære differentialligninger

Et generelt differentialligniningsystem skrives som

$$\begin{aligned} x_1' &= f_1(x_1,\ldots,x_n) \\ &\vdots \\ x_n' &= f_n(x_1,\ldots,x_n) \end{aligned}$$

Et lineær differentialtligningssystem er på formen er et system hvor afbildningen f_i er en lineær transformation. Den har da en standard matrix repræsentation A. I ikke-matrix form:

$$\begin{aligned} x_1' &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + \ldots + a_{nn}x_n \end{aligned}$$

Hvor x_i er en differentiabel funktion.
Dette kan skrives kompakt på matrixform

$$x' = Ax$$

Hvis $n = 1$ har vi ligningen

$$x' = ax$$

Som vil blive løst i næste bevis.

Lemma 10.1.1 [N]

Lad $\lambda, a \in \mathbb{C}$ ligningen

$$x' = \lambda x$$

Har en entydig løsning $z : \mathbb{R} \to \mathbb{C}$ med $z(0) = a$ givet ved $z(t) = e^{\lambda t} a$

Bevis 10.1.1

Hvis $\lambda = c + id$

$$\begin{aligned} e^{\lambda t} &= e^{(c+id)t} \\ &= e^{ct} e^{idt} \\ &= e^{ct}(\cos dt + i \sin dt) \end{aligned}$$

$$
\begin{aligned}
\frac{d}{dt}\left(e^{\lambda t}\right) &= \lambda e^{ct}\left(\cos dt + i \sin dt\right) + e^{ct}\left(-d \sin dt + id \cos dt\right) \\
&= e^{ct}\left(c + id\right)\left(\cos dt + i \sin dt\right) \\
&= e^{ct}\left(\lambda\right)\left(e^{idt}\right) \\
&= \lambda e^{ct} e^{idt} \\
&= \lambda e^{(c+id)t} \\
&= \lambda e^{\lambda t}
\end{aligned}
$$

Så $e^{\lambda t}$ opfyldet $x' = \lambda x$ men har startværdi 1. Vi ganger derfor $e^{\lambda t}$ med a som er en konstant, som går igennem hele beregningen ovenfor. Vi finder derfor ud af at $z(t) = \lambda e^{\lambda t} a$ er en løsning til $x' = \lambda x$ med startværdi $z(0) = a$.

Er denne så entydig? Vi antager at der eksisterer en funktion $y(t)$ som er en løsning til $x' = \lambda x$ med startværdi $y(0) = a$.

Vi kigger på funktionen $e^{-\lambda t} y(t)$ og vil se hvordan denne ændrer sig, vi differentierer altså

$$
\begin{aligned}
\frac{d}{dt} e^{-\lambda t} y(t) &= -\lambda e^{-\lambda t} y(t) + e^{-\lambda t} y'(t) \\
&= e^{-\lambda t}\left(y'(t) - \lambda y(t)\right) \\
&= e^{-\lambda t}\left(\lambda y(t) - \lambda y(t)\right) \\
&= 0
\end{aligned}
$$

Så funktionen $e^{-\lambda t} y(t)$ er altså konstant. Derfor må vi indsætte hvilken som helst værdi for t og se hvad den giver:

$$
e^{-\lambda t} y(t) = e^{-\lambda \cdot 0} y(0) = y(0) = a
$$

Denne ligning løser vi i forhold til $y(t)$

$$
\begin{aligned}
e^{-\lambda t} y(t) &= a \\
y(t) &= a \frac{1}{e^{-\lambda t}} \\
y(t) &= a e^{\lambda t} \\
&= z(t)
\end{aligned}
$$

Altså er $z(t)$ den entydige løsning til $x' = \lambda x$ med startværdi $z(0) = a$.

Teorem 12.2.2

(Putzers algoritme)

Lad $A \in Mat_{n,n}(\mathbb{C})$. Lad $\lambda_1, \ldots, \lambda_n$ være egenværdier for A, talt med multiplicitet.

Lad $P_0 = I$ og for $k = 1, \ldots, n$, $P_k = \prod_{j=1}^{k}(A - \lambda_j I)$ og definer

$$Q(t) = \sum_{k=0}^{n-1} r_{k+1}(t) P_k$$

Hvor $r_1(t) = e^{\lambda_1 t}$ og definer r_k induktivt for $k = 2, \ldots, n$ ved

$$r_k = e^{\lambda_k t} \int_0^t e^{-\lambda_k s} r_{k-1}(s)\, ds$$

Så gælder $Q(0) = I$ og $Q'(t) = AQ(t) = Q(t)A$.

Bevis Teorem 12.2.2

Vi ser at

$$r_1(t) = e^{\lambda_1 \cdot 0} = 1$$

$$k > 1 : r_k(t) = e^{\lambda_k \cdot 0} \int_0^0 e^{-\lambda_k s} r_{k-1}(s)\, ds = 0$$

Så

$$Q(0) = \sum_{k=0}^{n-1} r_{k+1}(0) P_k = P_k = I$$

Vi ser at A kommuterer med $A - \lambda_i I$ for $i = 1, \ldots, n$ så den kommuterer med $P_0, \ldots, P_{n-1}$ og derfor også med $Q(t)$. Dette giver os at $Q(t)A = AQ(t)$.

Vi vil nu gerne vise at $Q'(t) = AQ(t)$. Først vil vi differentiere r_k for $k > 1$.

$$\begin{aligned}
\frac{d}{dt} r_k(t) &= \lambda_k e^{\lambda_k t} \int_0^t e^{-\lambda_k s} r_{k-1}(s)\, ds + e^{\lambda_k t}\left(e^{-\lambda_k t} r_{k-1}(t)\right) \\
&= \lambda_k e^{\lambda_k t} \int_0^t e^{-\lambda_k s} r_{k-1}(s)\, ds + r_{k-1}(t) \\
&= \lambda_k r_k(t) + r_{k-1}(t)
\end{aligned}$$

Så vi har derfor at

$$
\begin{aligned}
Q'(t) &= \sum_{k=0}^{n-1} r_{k+1}'(t) P_k \\
&= \sum_{k=0}^{n-1} \left(\lambda_{k+1} r_{k+1}(t) + r_k(t) \right) P_k
\end{aligned}
$$

Og vi har at

$$
\begin{aligned}
Q'(t) - AQ(t) &= \sum_{k=0}^{n-1} \left(\lambda_{k+1} r_{k+1}(t) + r_k(t) \right) P_k - A \sum_{k=0}^{n-1} r_{k+1}(t) P_k \\
&= \sum_{k=0}^{n-1} -r_{k+1}(t) \left(A - \lambda_{k+1} I \right) P_k + r_k(t) P_k \\
&= \sum_{k=0}^{n-1} -r_{k+1}(t) P_{k+1} + r_k(t) P_k \\
&= -r_n(t) P_n \\
&= 0
\end{aligned}
$$

Det sidste gælder da Cayley Hamilton sætningen siger at $(A - \lambda_1 I)\ldots(A - \lambda_n I) = P_n = 0$. Da $Q'(t) - AQ(t) = 0 \Rightarrow Q'(t) = AQ(t)$ og sætningen er bevist.

16 Markov Processor

Def. Markov Process

En Markov Process er en sekvens af hændelser med følgende egenskaber

1. Sættet af tilstande er endeligt
2. Den næste tilstand afhænger kun af den forrige tilstand
3. Sandsynlighederne ved hændelsesskift er konstante

Def. Sansynelighedsvektor

En vektor $x \in \mathbb{R}^n$ med ikke-negative indgange og sådan at summen af dens indgange er 1 kaldes en sandsynlighedsvektor.

Def. Stokastisk Matrix

En matrix $A \in Mat_{n,n}(\mathbb{R})$ er en stokastisk matrix hvis hver af dens søjler hver er en sandsynlighedsvektor.

Eksempel

$x_0 = \begin{bmatrix} 0,8 \\ 0,2 \end{bmatrix}$ er en sandsynlighedsvektor. Den fortæller at 80% kører Lamboghini og 20% kører Fait.

$A = \begin{bmatrix} 0,1 & 0,2 \\ 0,9 & 0,8 \end{bmatrix}$ fortæller at 10% af dem der kører Lamboghini gør det også ved en hændelse, mens 20% af dem der kører Fait skifter til Lamboghini. 90% skifter fra Lamboghini til Fait og 80% af dem der kører Fait bliver ved med det.

$$
\begin{aligned}
x_1 &= Ax_0 \\
&= \begin{bmatrix} 0,1 & 0,2 \\ 0,9 & 0,8 \end{bmatrix} \begin{bmatrix} 0,8 \\ 0,2 \end{bmatrix} \\
&= \begin{bmatrix} 0,1 \cdot 0,8 + 0,2 \cdot 0,2 \\ 0,9 \cdot 0,8 + 0,8 \cdot 0,2 \end{bmatrix} \\
&= \begin{bmatrix} 0,08 + 0,04 \\ 0,72 + 0,16 \end{bmatrix} \\
&= \begin{bmatrix} 0,12 \\ 0,88 \end{bmatrix}
\end{aligned}
$$

Lemma 10.2.4 [N]

Lad $e = \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix} \in \mathbb{R}^n$.

1. Lad $x \in \mathbb{R}^n$ med ikke-negative indgange, x er en sandsynlighedsvektor $\Leftrightarrow e^T x = 1$
2. Lad $A \in Mat_{n,n}(\mathbb{R})$ være en matrix med ikke-negative indgange. A er stokastisk $\Leftrightarrow e^T A = e^T$

Bevis Lemma 10.2.4

1:

$e^T x = x_1 + \dots + x_n = 1$ hvis og kun hvis x er en sandsynlighedsvektor.

2:

Skriv $A = [a_1, \dots, a_n]$. Så:

$$\begin{aligned} e^T A &= \left[e^T a_1, \dots, e^T a_n \right] \\ &= [1, \dots, 1] \\ &= e^T \end{aligned}$$

Hvis og kun hvis $a_1, \dots, a_n$ er sandsynlighedsvektorer.

Lemma 10.2.7 [N]

Lad $A \in Mat_{n,n}(\mathbb{R})$ være en stokastisk matrix.

Så er 1 en egenværdi for A.

Bevis Lemma 10.2.7

Lad os beregne $A^T e$

$$\begin{aligned} A^T e &= \left(e^T A \right)^T \\ &= \left(e^T \right)^T \\ &= 1e \end{aligned}$$

Vi ser at e er en egenvektor til $\lambda = 1$ for A^T. Derfor er $p_{A^T}(1) = 0$. Men:

$$\begin{aligned} p_A(1) &= \det(A - I) \\ &= \det\left((A - I)^T \right) \\ &= \det\left(A^T - I \right) \\ &= p_{A^T}(1) \end{aligned}$$

Så $p_A(1) = 0$ og derfor er 1 en egenværdi til A.

Teorem 6.3.4 [L]

Hvis A er en stokastisk matrix med dominant egenværdi $\lambda_1 = 1$ så vil Markov rækken konvergere mod en steady-state vektor x.

Bevis Teorem 6.3.4

Hvis A er diagonaliserbar:

Lad y_1 være egenvektoren til λ_1. Lad $Y = [y_1, \ldots, y_n]$ være en matrix der diagonaliserer A

(Så $A = YDY^{-1}$), så når $k \to \infty$

$$D^k = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_n^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix} \to \begin{bmatrix} 1 & & \\ & 0 & \\ & & \ddots & \\ & & & 0 \end{bmatrix} = E$$

Og hvis x_0 er en sandsynlighedsvektor og $c = Y^{-1}x_0$ så:

$$x_k = A^k x_0 = YD^kY^{-1}x_0 = YD^k c \to YEc = Y\left(c_1\underline{e}_1\right) = c_1\underline{y}_1$$

Så $c_1\underline{y}_1$ bliver steady-state vektoren.

Hvis A ikke er diagonaliserbar kan dette stadig bevises, men dette er uden for pensum.