1. Løsninger og mindste kvadraters løsninger af lineære ligningssystemer

- **Ligningssystem**
- **Underdeterminerede (sætn. 1.2.1)**
- **Overdeterminerede (sætn. 5.3.1)**

Vi har m lineære ligninger og kan opstille følgende ligningssystem:

$$\begin{matrix} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_n \end{matrix}$$

Så kan vi opstille det på matrixform:

$$Ax = b$$

Med løsning på formen $c = (c_1, \dots, c_n)^T$ som sættes ind på x'ernes plads. Måden vi finder løsningen på er ved at ændre den udvidede matrix $[A|b]$ vha. rækkeoperationerne:

$$R_i \leftrightarrow R_j \qquad R_i \to sR_i \qquad R_i \to R_i + tR_j$$

Vi har så et lemma, der siger, at løsningsmængden til to systemer er ens, hvis de to systemer er ens – og de to systemer er **rækkeækvivalente**: $[A|b] \sim [B|c]$ så har $Ax = b$ og $Bx = c$ ens løsningsmængde.

Vi skal finde $[B|c]$ så løsninger er nemme at skrive ned. Vi indfører **REF**:

1. En række med 0'er ligger under alle andre.
2. Den første ikke-nul indgang i en række er 1 og ligger i søjlen til højre for første ikke-nul adgang i rækken ovenfor.
3. RREF: alle andre indgange i en søjle med pivot er 0.

Så har vi **underdeterminerede** systemer, som er ligningssystemer med flere ubekendte end ligninger. Dvs. $n > m$. Generelt er underdeterminerede systemer konsistente – dvs. der findes altid mindst én løsning. Men der kan være tilfælde, hvor de vil være inkonsistente – fx hvis RREF har to nulrækker.

Sætning 1.2.1: Et $m \times n$ homogent ligningssystem har en ikke-triviel løsning, hvis $n > m$.

Bevis: Et homogent system er altid konsistent (da. $b = 0$ - der vil altid være den trivielle løsning $x = 0$).

Ydermere må systemet have maks. m ikke-nulrækker ⇒ maks. m pivot'er.

Da der er n variabler må der altså være mindst én eller flere frie variabler, som er arbitrære – og for enhver arbitrær værdi er der en løsning.

Herudover har vi også **overdeterminerede** ligningssystemer, hvor vi har flere ligninger en ubekendte ($m > n$). Her taler vi om at ligningssystemerne ofte er inkonsistente. Vi kan dog tilnærme os en løsning. Igen har vi samme ligningssystem som fra starten, og definerer så residualet:

$$r(x) = b - Ax \quad , \quad A \in Mat_{m,n}(\mathbb{F}) \quad , \quad b \in \mathbb{R}^m \quad , \quad x \in \mathbb{R}^n$$

Så er $\hat{x} \in \mathbb{R}^n$, så $\left||r(\hat{x})|\right|^2$ er mindst mulig en mindste kvadrates løsning. Afstanden mellem b og Ax er ligeledes givet ved:

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

Hvis $\hat{x}$ er en løsning til $Ax = b$ og $p = A\hat{x}$ så er p den vektor i søjlerummet for A som er tættest på b. Vi vil med følgende sætning vise at p findes og er unik!

Sætning 5.3.1: Lad S være et underrum af $\mathbb{R}^m$. For hvert $b \in \mathbb{R}^m$ findes en unik projektion p af b på S det nærmeste punkt på b i S:

$$\left||b - y|\right| > \left||b - p|\right| \ , \ \forall y \neq p \in S$$

Ydermere gælder det, at p kun vil være tættest på b, hvis og kun hvis $b - p \in S^\perp$

Bevis: ⇒: Vi ved $\mathbb{R}^m = S \oplus S^\perp$, så vi kan skrive $b = p + z$, hvor $z \in S^\perp$. Vi kan så skrive:

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

Da $b-y \in S$ og $b-p = z \in S^{\perp}$ så kan vi bruge Pythagoras:

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

Da $p \neq y$ har vi heraf, at

$$\left|\left|b-y\right|\right| > \left|\left|b-p\right|\right|$$

$\Leftarrow$: Så, hvis vi vælger et $q \in S$ og $b-q \notin S^{\perp}$, så da $q \neq p$ har vi at $q = y$. Da y er alle elementer i S, som ikke er p, har vi igen: $\|b-q\| > \|b-p\|$

2. Vektorrum og underrum

- **Aksiomer**
- **Egenskaber**
- **Underrum**
- **Sætning 3.2.1**

Definition: En mængde V udstyret med en addition ($\boldsymbol{u}+\boldsymbol{v} \in V$) og skalarmultiplikation ($\alpha\boldsymbol{v} \in V$). Så kaldes V et vektorrum, hvis følgende aksiomer er opfyldt:

$\boldsymbol{A1}$: $\forall x, y \in V: \quad x+y=y+x \quad (kommutativ)$

$\boldsymbol{A2}$: $\forall x, y, z \in V: \quad x+(y+z)=(x+y)+z \quad (associativ)$

$\boldsymbol{A3}$: $\exists 0 \in V: \; x+0=x \quad (eksistens\ af\ neutral\ element)$

$\boldsymbol{A4}$: $\forall x \in V, \exists -x \in V: x+(-x)=0 \quad (additiv\ invers)$

$\boldsymbol{S1}$: $\forall \alpha \in \mathbb{F}, x, y \in V: \; \alpha(x+y)=\alpha x+\alpha y$

$\boldsymbol{S2}$: $\forall \alpha, \beta \in \mathbb{F}, x \in V: (\alpha+\beta)x=\alpha x+\beta x \quad (2\ distributive\ egenskaber)$

$\boldsymbol{S3}$: $\forall \alpha, \beta \in \mathbb{F}, x \in V: (\alpha\beta)x=\alpha(\beta x)$

$\boldsymbol{S4}$: $1 \cdot x = x \quad (1\ er\ et\ neutralt\ element\ for\ skalarmultiplikation)$

Herudover har vi nogle flere egenskaber ved vektorrum:

Sætning 3.1.1: V er et vektorrum og x er et element i V, så gælder:
(1) $0 \cdot x = 0$
(2) $x+y=0$ så $y=-x \quad (-x\ er\ en\ entydig\ additiv\ invers\ til\ x)$
(3) $(-1)x=-x$

Bevis:

(1) $x \overset{S4}{=} 1x = (1+0)x \overset{S2}{=} 1x + 0x \overset{S4}{=} x + 0x$

$$0 = -x + x \overset{brug\ x\ fra\ oven}{=} -x + (x+0x) \overset{A2}{=} (-x+x) + 0x$$
$$= 0 + 0x \overset{A1}{=} 0x + 0 \overset{A3}{=} 0x$$

(2) $-x \overset{A3}{=} -x + 0 \overset{brug\ x+y=0}{=} -x + (x+y) \overset{A2}{=} (-x+x)+y = 0+y \overset{A1}{=} y+0 \overset{A3}{=} y$

(3) $0 \overset{(1)}{=} 0x = \big(1+(-1)\big)x \overset{S2}{=} 1x + (-1)x \overset{S4}{=} x + (-1)x$

Vi bruger så (2) til at indse, at der må gælde: $x+(-1)x=0$ så er $(-1)x=-x$

Vi har også underrum, som er defineret ved:

Definition: En ikke-tom delmængde S af et vektorrum V kaldes et underrum, hvis følgende gælder:

$\boldsymbol{C1}$: $\forall x \in V, \alpha \in \mathbb{F}: \qquad \alpha x \in S$

$\boldsymbol{C2}$: $\forall x, y \in V: \qquad x+y \in S$

$C1$ og $C2$ er også lukkethedsegenskaber for vektor.

Vi definerer et span:

Definition: Lad $v_1, \dots, v_n$ være vektorer i et vektorrum V.

En linear kombination er en sum på formen $\alpha_1 v_1 + \cdots + \alpha_n v_n$, hvor α_i er skalarer.

Mængden af alle linear kombinationer af $v_1, \dots, v_n$ kaldes span for $v_1, \dots, v_n \rightarrow Span(v_1, \dots, v_n)$.

Vi vil nu bevise, at et span er et underrum af V.

Sætning 3.2.1: Hvis $v_1, \dots, v_n$ er elementer i et vektorrum V, så er $Span(v_1, \dots, v_n)$ et underrum til V.

Bevis: Vi bruger definitionen på et underrum. Hvis $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ er et arbitrært element i $Span(v_1, \dots, v_n)$ og β er en skalar:

$C1$: $\beta v = \beta(\alpha_1 v_1 + \cdots + \alpha_n v_n) \overset{S1}{=} (\beta\alpha_1)v_1 + \cdots + (\beta\alpha_n)v_n \in Span(v_1, \dots, v_n)$

Nu definerer vi $w = \beta_1 v_1 + \cdots + \beta_n v_n$.

$C2$: $v + w = (\alpha_1 v_1 + \cdots + \alpha_n v_n) + (\alpha_1 v_1 + \cdots + \alpha_n v_n) \overset{S2}{=} (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_n + \beta_n)v_n \in Span(v_1, \dots, v_n)$

Dermed er $Span(v_1, \dots, v_n)$ et underrum af V.

3. Lineær uafhængighed

- **Definition**
- **Lemma 1.4.2**
- **Sætning 3.3.1**
- **Lemma 1.2.1**
- **Sætning 3.4.1**

Definition: Vektorerne $v_1, \dots v_n$ i et vektorrum V er lineært uafhængige, hvis

$$c_1 v_1 + \cdots + c_n v_n = 0 \Rightarrow c_1 = 0, \dots, c_n = 0$$

Vi har et lemma, vi bruger til beviset senere:

Lemma 1.4.2: A er en $n \times n$ matrix, så gælder:

i) A er invertibel

ii) Det homogene ligningssystem $Ax = 0 \Rightarrow x = 0$

iii) A er rækkeækvivalent til I

Følgende sætning fortæller os, at hvis søjlerne i en matrix er lineært uafhængige, så er matrixen invertibel.

Sætning 3.3.1: Lad $x_1, \dots, x_n \in \mathbb{R}^n$ og $X = (x_1, \dots, x_n)$. $x_1, \dots, x_n \ lin\ uafh. \Leftrightarrow X\ er\ invertibel.$

Bevis: Vi kan skrive ligningen

$$c_1 x_1 + \cdots + c_n x_n = 0$$

Om til $Xc = 0$, hvor $c = (c_1, \dots, c_n)^T$

Ligningen har kun løsningen 0, hvis og kun hvis X er invertibel (jf. sætning, der siger, at $A\ inv. \Leftrightarrow Ax = 0 \Leftrightarrow x = 0$).

Vi skal bruge en definition…

Definition: $\{v_1, \dots, v_n\}$ udspændende mængde for V hvis og kun hvis, hver vektor i V kan skrives som en linear kombination af $v_1, \dots, v_n$.

…og et lemma…

Lemma 1.2.1: Et $m \times n$ homogent ligningssystem har en ikke-triviel løsning, hvis $n > m$.

Bevis: Et homogent system er altid konsistent (da. $b = 0$ - der vil altid være den trivielle løsning $x = 0$).

Ydermere må systemet have maks. m ikke-nulrækker ⇒ maks. m pivot'er.

Da der er n variabler må der altså være mindst én eller flere frie variabler, som er arbitrære – og for enhver arbitrær værdi er der en løsning.

…til det sidste bevis:

Sætning 3.4.1: Hvis $\{v_1, \dots, v_n\}$ er en udspændende mængde for V, så vil enhver samling af m vektorer $\{u_1, \dots, u_m\}$ i V være lineære afhængige når $m > n$.

Bevis: Lad $u_1, \dots, u_m$ være m vektorer i V. Jf. def. Af udspændende mængde, kan vi skrive:

$$u_i = a_{i1} v_1 + \cdots + a_{in} v_n = \sum_{j=1}^{n} a_{ij} v_j \quad , \quad i = 1, \dots, m$$

Vi har en linear kombination, som vi skriver lidt om på (vi indsætter ovenstående på u_i's plads):

$$c_1 u_1 + \cdots + c_m u_m = c_1 \sum_{j=1}^{n} a_{1j} v_j + \cdots + c_m \sum_{j=1}^{n} a_{mj} v_j = \sum_{i=1}^{m} \left(c_i \left(\sum_{j=1}^{n} a_{ij} v_j \right) \right)$$

(vi bytter om på sumtegnene, da kun a_{ij} afhænger af begge)

$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} c_i \right) v_j$$

Nu kigger vi på ligningssystemet (homogent):

$$\sum_{i=1}^{m} a_{ij} c_i = 0 \quad , \quad j = 1, \dots, n$$

Dette system består af m ubekendte med n ligninger (altså flere ubekendte end ligninger). Dette system har en ikke-triviel løsning $(\hat{c}_1, \dots, \hat{c}_m)^T$. Hvis vi indsætter denne løsning i den første linear kombination (hvor den midterste sum så er 0):

$$\widehat{c_1} u_1 + \cdots + \widehat{c_m} u_m = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} \widehat{c_i} \right) v_j = \sum_{j=1}^{n} 0 v_j = 0$$

Så da ovenstående giver 0 med en ikke-triviel løsning, er vektorerne $u_1, \dots, u_m$ lineært afhængige.

4. Basis for vektorrum; koordinatisering

- **Definition**
- **Sætning 3.4.1**
- **Koordinatisering**

Definition: Vektorerne $v_1, \dots, v_n$ danner en basis for vektorrummet V hvis og kun hvis:

$$(i)\ v_1, \dots, v_n \quad er\ lin\ uafh.$$

$$(ii)\ v_1, \dots, v_n \ \ udspænder\ V$$

Dvs. alle linear kombinationer for $v_1, \dots, v_n$ ligger i $Span(v_1, \dots, v_n)$.

En mængde $\{x_1, \dots, x_n\}$ er lineært uafhængige, hvis og kun hvis $c_1x_1 + \cdots + c_nx_n = 0$ med $c_i = 0, i = 1, \dots, n$. Ellers er de lineært afhængige.

Vi siger at en basis er den minimale mængde vektorer, der skal til for at udspænde et rum.

Definition: $\{v_1, \dots, v_n\}$ udspændende mængde for V hvis og kun hvis, enhver vektor i V kan skrives som en linear kombination af $v_1, \dots, v_n$.

For at bevise den næste sætning, bruger vi desuden ovenstående definition. Hvis vi i den følgende sætning antager, at vektorerne $v_1, \dots, v_n$ er basis for vektorrummet V, så er de vektorer lineært uafhængige, og sætningen siger så, at hvis man har bare $n + 1$ vektorer fra rummet, så vil de vektorer være lineært afhængige.

Sætning 3.4.1: Hvis $\{v_1, \dots, v_n\}$ er en udspændende mængde for V, så vil enhver samling af m vektorer i V være lineære afhængige. $m > n$.

Bevis: Lad $u_1, \dots, u_m$ være m vektorer i V. Jf. def. Af udspændende mængde, kan vi skrive:

$$u_i = a_{i1}v_1 + \cdots + a_{in}v_n = \sum_{j=1}^{n} a_{ij}v_j \quad , \quad i = 1, \dots, m$$

Vi har en linear kombination, som vi skriver lidt om på (vi indsætter ovenstående på u_i's plads):

$$c_1u_1 + \cdots + c_mu_m = c_1\sum_{j=1}^{n} a_{1j}v_j + \cdots + c_m\sum_{j=1}^{n} a_{mj}v_j = \sum_{i=1}^{m}\left(c_i\left(\sum_{j=1}^{n} a_{ij}v_j\right)\right)$$

(vi bytter om på sumtegnene, da kun a_{ij} afhænger af begge)

$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}c_i\right)v_j$$

Nu kigger vi på ligningssystemet (homogent):

$$\sum_{i=1}^{m} a_{ij}c_i = 0 \quad , \quad j = 1, \dots, n$$

Dette system består af m ubekendte med n ligninger (altså flere ubekendte end ligninger). Dette system har en ikke-triviel løsning $(\hat{c}_1, \dots, \hat{c}_m)^T$. Hvis vi indsætter denne løsning i den første linear kombination (hvor den midterste sum så er 0):

$$\widehat{c_1}u_1 + \cdots + \widehat{c_m}u_m = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}\widehat{c_i}\right)v_j = \sum_{j=1}^{n} 0v_j = 0$$

Så da ovenstående giver 0 med en ikke-triviel løsning, er vektorerne $u_1, \dots, u_m$ lineært afhængige.
Herudover kan det ofte være nyttigt at skifte fra en basis til en anden. Helt konkret har vi følgende definition:

Definition: Lad V være et vektorrum, og lad $E = [v_1, \dots, v_n]$ være en ordnet basis for V. Et element v fra V kan skrives som: $v = c_1 v_1 + \cdots + c_n v_n$. Så vi kan associere hver vektor v med en unik vektor $c = (c_1, \dots, c_n)^T \in R^n$. Denne vektor kaldes koordinatvektoren til v mht. E og skrives $[v]_E$. c_i'erne kaldes v's koordinater relativt til E.

Fremgangsmåden for at skifte basis i $\mathbb{R}^2$:

Fremgangsmåde: Vi ønsker at skifte basis fra $[v_1, v_2]$ til $[u_1, u_2]$, hvor de begge er ordnede baser for $\mathbb{R}^2$. Hvis vi har en vektor $x \in \mathbb{R}^2$, så kan vi skrive den ud fra V:

$$x = c_1 v_1 + c_2 v_2 \quad , \quad c = [x]_V$$

Hvor vi antager, vi kender c. x ønsker vi så at skrive ud fra U:

$$x = d_1 u_1 + d_2 u_2 \quad , \quad d = [x]_U$$

Så vi skal altså finde disse to skalarer:

$$d_1 u_1 + d_2 u_2 = c_1 v_1 + c_2 v_2$$

Hvis vi siger, at $V = (v_1, v_2)$ og $U = (u_1, u_2)$, så har vi:

$$Vc = Ud$$

Da U og V består af basis-vektorer er de lineært uafhængige, og dermed er de to matricer invertible:

$$d = U^{-1} V c$$

Vi siger så, at $S = U^{-1} V$ er transitionsmatrixen. Så får vi den nye koordinatvektor for x mht. $[u_1, u_2]$ ved at gange den gamle med S.

5. Matricer og lineære transformationer

- **Definition**
- **Egenskaber**
- **Sætning 4.2.1**
- **Lemma 4.2.3**
- **Sætning 4.2.4**

Definition: En lineær transformation $L: V \rightarrow W$ er en afbildning som respekterer lineær struktur, dvs.:

$$\forall v_1, v_2 \in V, \forall \alpha, \beta \in \mathbb{F}:$$

$$L(\alpha v) = \alpha L(v)$$

$$L(v_1 + v_2) = L(v_1) + L(v_2)$$

Dette kan sættes sammen til:

$$L(\alpha v_1 + \beta v_2) = \alpha L(v_1) + \beta L(v_2)$$

Der er nogle forskellige egenskaber for lineære transformationer. Herunder opremses tre:

Egenskaber:

$$(i) \quad L(0_V) = 0_W$$

$$(ii) \quad L(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 L(v_1) + \cdots + \alpha_n L(v_n)$$

$$(iii) \; L(-v) = -L(v) \quad , \quad \forall v \in V$$

Bevis:

$$(i) \quad L(0_V) = L(00_V) = 0L(0_V) = 0_W$$

$$(ii) \quad L(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 L(v_1) + \alpha_2 L(v_2)$$

Hvis vi antager, at det gælder for $n = k$, og så kigger på $n = k + 1$:

$$L(\alpha_1 v_1 + \cdots + \alpha_k v_k + \alpha_{k+1} v_{k+1}) \overset{IH}{=} \alpha_1 L(v_1) + \cdots + \alpha_k L(v_k) + L(\alpha_{k+1} v_{k+1})$$

Fra def. har vi, at sidste led således bliver det ønskede.

$$(iii) \; L(-v) = L\big((-1)v\big) = (-1)L(v) = -L(v)$$

Et lille eksempel til at tjekke, at en lineær transformation overholder lineær struktur – og for at lede os over i matrixrepræsentationer af lineære transformationer:

Eksempel: Vi har afbildningen $L: \mathbb{R}^2 \rightarrow \mathbb{R}^1 \quad , L(x) = x_1 + x_2$. Vi tjekker:

$$L(\alpha x + \beta y) = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) = \alpha(x_1 + x_2) + \beta(y_1 + y_2) = \alpha L(x) + \beta L(y)$$

Hvis vi så definerer $A = (1\ 1) \in Mat_{2\times 1} \in (\mathbb{F})$, så kan vi skrive den lineære transformation som:

$$Ax = (1 \;\; 1)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 + x_2 = L(x)$$

Herefter har vi en sætning, som formaliserer lineær transformation og matrixrepræsentationer heraf:

Sætning 4.2.1: Hvis L er en lineær transformation, som afbilder $\mathbb{R}^n \rightarrow \mathbb{R}^m$, så er der en $m \times n$-matrix således at $L(x) = Ax$ for hvert $x \in \mathbb{R}^n$. Den j'te søjlevektor i A er givet ved

$$a_j = L(e_j) \quad , \quad j = 1,2, \ldots, n$$

Bevis: Først definerer vi vores a_j som i sætningen og skriver

$$A = (a_{ij}) = (a_1, \ldots, a_n)$$

Vi skriver x som en linear kombination:

$$x = x_1 e_1 + \cdots + x_n e_n$$

(x er et arbitrært element i $\mathbb{R}^n$). Så skriver vi:

$$L(x) = L(x_1e_1 + \cdots + x_ne_n) \overset{ii}{=} x_1L(e_1) + \cdots + x_nL(e_n) = x_1a_1 + \cdots + x_na_n$$
$$= (a_1, \dots, a_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ax$$

Det er også muligt at finde matricer, der repræsenterer lineære transformationer fra et vektorum V til et andet W. Så arbejder vi med ordnede baser for hhv. V og W.

Lemma 4.2.3: Vi har to ordnede baser: $E = [u_1, \dots, u_n]$ og $F = [b_1, \dots, b_m]$ for $\mathbb{R}^n$ og $\mathbb{R}^m$. Vi skriver $B = (b_1, \dots, b_m)$ og har en lineær transformation: $L: \mathbb{R}^n \to \mathbb{R}^m$ og A er matricen der repræsenterer L mht. E og F, så er:

$$a_j = B^{-1}L(u_j) \quad , \quad j = 1, \dots, n$$

Dette resultat bruger vi i næste sætning, hvor vi kan finde A.

Sætning 4.2.4: A er matrixrepræsentationen for $L: \mathbb{R}^n \to \mathbb{R}^m$ med de ordnede baser: $E = [u_1, \dots, u_n]$ og $F = [b_1, \dots, b_m]$ så er den RREF af matrixen $\left(b_1, \dots, b_m \middle| L(u_1), \dots, L(u_n)\right) \sim (I|A)$.

Bevis: Hvis vi har $B = (b_1, \dots, b_m)$ Så har vi $\left(B \middle| L(u_1), \dots, L(u_n)\right)$ er rækkeækvivalent med:

$$B^{-1}\left(B \middle| L(u_1), \dots, L(u_n)\right) = \left(I \middle| B^{-1}L(u_1), \dots, B^{-1}L(u_n)\right) = (I|a_1, \dots, a_n) = (I|A)$$

6. Determinanter

- **Definition**
- **ERO**
- **Sætning 2.2.2**
- **Cramers regel**

For alle kvadratiske matricer ($n \times n$) findes der et tal, som vi kalder determinanten. Determinanter kan bl.a. bruges til at løse ligningssystemer (specielt ved $n > 3$).

Definition: For $A \in Mat_{1\times 1}$ så er determinanten $\det(A) = a_{11}$. For $n > 1$ er determinanten givet ved

$$\det(A) = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}$$

(her rækkeudvikling af 1. række). Generelt med rækkeudvikling langs *i*'te række eller søjleudvikling langs *j*'te søjle:

$$\begin{aligned}\det(A) &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \quad , \quad i, j = 1, \dots, n\end{aligned}$$

A'erne (kofaktorerne) er givet ved:

$$A_{ij} = (-1)^{i+j} \det\left(M(A)_{ij}\right)$$

Hvor igen $M(A)_{ij}$ er $(n-1) \times (n-1)$-matricen, som man får ved at fjerne rækken og søjlen, som indgang a_{ij} er i. Dvs. en slags rekursiv definition.

Vi kan sige lidt om determinanten for nogle forskellige matricer:

- $\det(A) = \det(A^T)$
- $\det(A) = 0$, hvis A indeholder en nul-række eller –søjle (ækvivalent med to ens rækker eller søjler).
- Determinanten er produktet af diagonalelementerne i en triangulær matrix.

Rækkeoperationer på en matrix gør noget forskelligt ved determinanten:

i. To rækker ombyttes: $\det(A) \to -\det(A)$
ii. En række ganges med en ikke-nul skalar: $\det(A) \to \alpha \det(A)$
iii. Lægge et multiplum af en række til en anden: $\det(A) \to \det(A)$

Som en sidenote til beviset kan det siges, at generelt gælder $\det(E) \neq 0$ for alle rækkeoperationer. Ydermere har vi, at $\det(EA) = \det(E)\det(A)$.

Sætning 2.2.2: $A \in Mat_{n\times n}$ er ikke-invertibel $\Leftrightarrow \det(A) = 0$.

Bevis: Vi kan få A på RREF ved et endeligt antal rækkeoperationer. Dvs.

$$U = E_k E_{k-1} \dots E_1 A$$

Hvis vi tager determinanten af U:

$$\det(U) = \det(E_k E_{k-1} \dots E_1 A) = \det(E_k)\det(E_{k-1}) \dots \det(E_1)\det(A)$$

Da alle $\det(E_i) \neq 0$ må $\det(A) = 0 \Leftrightarrow \det(U) = 0$. Hvis A er invertibel, vil U have 1'er på diagonalen og $\det(U) = 1$, men da $\det(U) = 0$ må A være ikke-invertibel, og U har enten nul-rækker og/eller –søjler.

Som sagt tidligere, kan determinanter også bruges til at beregne løsninger, til ligningssystemer. Først definerer vi den adjungerende matrix til brug i Cramers regel.

Definition: Den adjungerende matrix *adj A* består af kofaktorerne for *A* transponeret:

$$adj\ A = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & & & \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

A og *adj A* ganget sammen giver determinanten til *A*: $A(adj\ A) = \det(A)\, I$

Dette kan skrives om til (hvis A er invertibel → det(A)≠0): $A^{-1} = \frac{1}{\det(A)} adj\ A$

Cramers regel med bevis:

Sætning 2.3.1: $A \in Mat_{n,n}(\mathbb{F})$ som er invertibel, $b \in \mathbb{R}^n$. Så er A_i matrix A med i'te søjle skiftet ud med b. Hvis $\hat{x}$ er en unik løsning til $Ax = b$. Så er indgangene i $\hat{x}$ givet ved:

$$\widehat{x_i} = \frac{\det(A_i)}{\det(A)} \quad , \qquad i = 1,2, \dots , n$$

Bevis: Vi kan skrive $\hat{x}$ op på følgende måde:

$$\hat{x} = A^{-1}b \overset{\text{def}\, adj\ A}{=} \left(\frac{1}{\det(A)} adj\ A\right) b = \frac{1}{\det(A)} \begin{pmatrix} b_1A_{11} + b_2A_{21} + \cdots + b_nA_{n1} \\ \vdots \\ b_1A_{1n} + b_2A_{2n} + \cdots + b_nA_{nn} \end{pmatrix}$$

Og så er $\widehat{x_i}$ altså givet ved:

$$\widehat{x_i} = \frac{b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni}}{\det(A)} = \frac{\det(A_i)}{\det(A)}$$

7. Egenværdier og egenvektorer

- **Definitioner**
  - **Egenrum**
  - **Karakteristisk polynomium**
- **Sætning 6.1.1**
- **Def. på diagonalisering**
- **Sætning 6.3.2**

Definition: $A \in Mat_{n,n}(\mathbb{F}), \lambda \in \mathbb{F}$ er en egenværdi for A, hvis der findes en ikke-nul-vektor x således, at: $Ax = \lambda x$ gælder. x kaldes så for egenvektoren tilhørende egenværdien λ.

En egenværdi kan have flere egenvektorer, men en egenvektor har kun én egenværdi.

For at finde egenværdierne for en matrix A, skal vi finde det karakteristiske polynomium. Men først finder vi **egenrummet:**

Definition: Vi tager den første definition, og skriver den om:

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

Så er løsningsrummet $N(A - \lambda I)$ også kaldet egenrummet. Egenrummet består af alle egenvektorer til λ.

Dvs. vi har kun en ikke-triviel ($x \neq 0$) løsning, hvis $A - \lambda I$ er ikke-invertibel, hvilket igen vil sige, at $\det(A - \lambda I) = 0$. Det leder os frem til det **karakteristiske polynomium:**

Definition: Egenværdierne for en matrix A kan beregnes fra det karakteristiske polynomium (eller ligning):

$$p(\lambda) = \det(A - \lambda I) = 0$$

Rødderne er så egenværdierne. Typisk siger vi, at en $n \times n$ matrix har n egenværdier (talt med multiplicitet).

En lille smule om similaritet.

Definition: En matrix B er similær til A, hvis der findes en invertibel matrix S således, at:

$$B = S^{-1}AS$$

Sætning 6.1.1: For to similære matricer A og B gælder det, at deres karakteristiske polynomium er ens:

$$p_A(\lambda) = p_B(\lambda)$$

Dermed har de to matricer også samme egenværdier.

Bevis:

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

Vi kan bruge egenværdier og egenvektorere til at diagonalisere en matrix:

Definition: En $n \times n$ matrix er diagonaliserbar, hvis det findes en invertibel matrix X og en diagonal matrix D, så:

$$D = X^{-1}AX$$

Vi siger så, at X diagonaliserer A.

Følgende sætning fortæller os, hvad der skal gælde om egenvektorerne for en matrix A:

Sætning 6.3.2: En $n \times n$-matrix A er diagonaliserbar $\Leftrightarrow$ n lineært uafhængige egenvektorere for A.

Bevis: $\Leftarrow$: Vi antager A har de lineært uafhængige egenvektorer $x_1, \dots, x_n$, og lader λ_i være egenværdien tilhørende x_i. X er en matrix med x_j som j'te søjlevektor. Vi ved så: $Ax_j = \lambda_j x_j$ er den j'te søjlevektor for AX. Så:

$$AX = (Ax_1, \dots Ax_n) = (\lambda_1 x_1, \dots, \lambda_n x_n)$$
$$= (x_1, \dots, x_n)\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = XD$$

Da X består af n lineært uafhængige vektorere, er X invertibel, og vi har:

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

⇒: Hvis A er diagonaliserbar så har vi: $A = XDX^{-1} \Rightarrow AX = XD$ med

$$D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

Hvis vi igen siger, at $x_1, \dots, x_n$ er søjlevektorerne for X, så har vi igen:

$$Ax_j = \lambda_j x_j \quad (d_{jj} = \lambda_j)$$

Dvs., at x_j er egenvektorer tilhørende λ_j (til A). Da X var invertibel, består den af n lineært uafhængige søjlevektorer – dvs. A har n lineært uafhængige egenvektorer.

8. Diagonalisering

- **Egenværdier og –vektorer**
- **Definitioner**
- **Sætning 6.3.2**
- **Bemærkninger**
- **Sætning 9.2.3**

Definition: $A \in Mat_{n,n}(\mathbb{F})$, $\in \mathbb{F}$ er en egenværdi for A, hvis der findes en ikke-nul-vektor x således, at: $Ax = \lambda x$ gælder. x kaldes så for egenvektoren tilhørende egenværdien λ.

En egenværdi kan have flere egenvektorer, men en egenvektor har kun én egenværdi. For at finde egenværdierne for en matrix A, skal vi finde det karakteristiske polynomium.

Definition: Egenværdierne for en matrix A kan beregnes fra det karakteristiske polynomium (eller ligning):

$$p(\lambda) = \det(A - \lambda I) = 0$$

Rødderne er så egenværdierne. Typisk siger vi, at en $n \times n$ matrix har n egenværdier (talt med multiplicitet).

En lille smule om similaritet.

Definition: En matrix B er similær til A, hvis der findes en invertibel matrix S således, at:

$$B = S^{-1}AS$$

Sætning 6.1.1: For to similære matricer A og B gælder det, at deres karakteristiske polynomium er ens:

$$p_A(\lambda) = p_B(\lambda)$$

Dermed har de to matricer også samme egenværdier.

Bevis:

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

Vi kan bruge egenværdier og egenvektorere til at diagonalisere en matrix:

Definition: En $n \times n$ matrix er diagonaliserbar, hvis det findes en invertibel matrix X og en diagonal matrix D, så:

$$A = X^{-1}DX$$

Vi siger så, at X diagonaliserer A.

Følgende sætning fortæller os, hvad der skal gælde om egenvektorerne for en matrix A:

Sætning 6.3.2: En $n \times n$-matrix A er diagonaliserbar $\Leftrightarrow$ n lineært uafhængige egenvektorere for A.

Bevis: $\Leftarrow$: Vi antager A har de lineært uafhængige egenvektore $x_1, \dots, x_n$, og lader λ_i være egenværdien tilhørende x_i. X er en matrix med x_j som j'te søjlevektor. Vi ved så: $Ax_j = \lambda_j x_j$ er den j'te søjlevektor for AX. Så:

$$\begin{aligned} AX &= (Ax_1, \dots Ax_n) = (\lambda_1 x_1, \dots, \lambda_n x_n) \\ &= (x_1, \dots, x_n)\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = XD \end{aligned}$$

Da X består af n lineært uafhængige vektorere, er X invertibel, og vi har:

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

$\Rightarrow$: Hvis A er diagonaliserbar så har vi: $A = XDX^{-1} \Rightarrow AX = XD$ med

$$D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

Hvis vi igen siger, at $x_1, \ldots, x_n$ er søjlevektorerne for X, så har vi igen:

$$Ax_j = \lambda_j x_j \quad (d_{jj} = \lambda_j)$$

Dvs., at x_j er egenvektorere tilhørende λ_j (til A). Da X var invertibel, består den af n lineært uafhængige søjlevektorere – dvs. A har n lineært uafhængige egenvektorere.

I forbindelse med beviset kan vi nu opsummere nogle ting:

1. Hvis A er diagonaliserbar, er søjlevektorerne i diagonaliseringmatricen X egenvektorerne for A, og diagonalelementerne i D er egenværdierne for A.
2. Hvis en $n \times n$-matrix A har n forskellige egenværdier er den diagonaliserbar. Hvis **ikke**, så er den kun diagonaliserbar, hvis alle egenvektorerne er lineært uafhængige.
3. Hvis A er diagonaliserbar, kan vi faktorisere $A = XDX^{-1}$.

Til sidst vil vi vise, at det er nemmere at tage k'te potens af en diagonaliserbar matrix:

Sætning 9.2.3: A er diagonaliserbar, $x_1, \ldots, x$ er en basis bestående af egenvektorerne for A med tilsvarende egenværdier $\lambda_1, \ldots, \lambda_n$. Så gælder der, at:

1. $A^k(c_1 x_1 + \cdots + c_n x_n) = c_1 \lambda_1^k x_1 + \cdots + c_n \lambda_n^k x_n$
2. Vi siger $X = (x_1, \ldots, x_n)$, så er den k'te potens af A det samme som:

$$A^k = X \begin{pmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{pmatrix} X^1$$

Bevis: ad 1) For $k \in \mathbb{N}$:

$$A^k(c_1 x_1 + \cdots + c_n x_n) = c_1 A^k x_1 + \cdots + c_n A^k x_n = c_1 \lambda_1^k x_1 + \cdots + c_n \lambda_n^k x_n$$

Jf. definition på egenværdi og –vektor.

ad 2) Ligningen fra 1. kan skrives om til matrixform:

$$A^k X \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = X \begin{pmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Vi fjerner bare c'erne:

$$A^k X = X \begin{pmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{pmatrix} \Rightarrow A^k = X \begin{pmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{pmatrix} X^{-1}$$

X er invertibel, da den består af uafhængige søjlevektorer.

9. Indre Produkt

- **Def. Indre produkt og norm**
- **Sætning 5.4.1 (Pythagoras)**
- **Def. vektorprojektion**
- **Sætning 5.4.2 (Cauchy-Schwarz' ulighed)**
- **Sætning ---**

Definition: Et indre produkt på et vektorrum, er en operation, der tildeler et reelt tal $\langle x, y\rangle$ til hvert par af vektorer x og y. Følgende regler skal overholdes:

$$i.\ \ \langle x,x\rangle \geq 0 \quad , \ lighed\ \Leftrightarrow x = 0$$

$$ii.\ \ \langle x,y\rangle = \langle y,x\rangle \quad , \ \forall x,y \in V$$

$$iii.\ \ \langle \alpha x + \beta y, z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle \quad , \ \forall x,y,z \in V \quad og\ \ \forall \alpha,\beta$$

Et vektorrum med et indre produkt kaldes et indre produkt rum.

Helt konkret er det indre produkt defineret som $\langle x,y\rangle = x^T y$ i $\mathbb{R}^n$. I $\mathbb{C}^n$ er det indre produkt defineret som $\langle x,y\rangle = x^H y$ og der gælder desuden: $\langle x,y\rangle = \overline{\langle y,x\rangle}$

Vi vil nu definere normen, idet vi skal bruge den til senere beviser:

Definition: Hvis v er en vektor i et indre produkt rum V, så er normen (eller længden) givet ved:

$$\|v\| = \sqrt{\langle v,v\rangle} \Rightarrow \langle v,v\rangle = \|v\|^2$$

Desuden gælder der, at hvis to vektorer er ortogonale, hvis $\langle u,v\rangle = 0$ og:

$$i.\ \ \|v\| \geq 0 \quad , \ lighed \Leftrightarrow v = 0$$

$$ii.\ \ \|\alpha v\| = |\alpha|\|v\|$$

$$iii.\ \ \|v + w\| \leq \|v\| + \|w\|$$

Den sidste betingelse kaldes også trekantsuligheden.

Det vil altså sige, at et par vektorer vil opfylde pythagoras i $\mathbb{R}^n$.

Sætning 5.4.1: Hvis u og v er ortogonale vektorer i et indre produkt rum, så gælder:

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

Bevis: Vi kigger på venstresiden i ovenstående og bruger definitionen:

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

da u og v er ortogonale har vi $2\langle u,v\rangle = 0$.

Til vores næste bevis, skal vi bruge vektorprojektion:

Definition: Vektorprojektionen p af u på v er givet ved:

$$p = \frac{\langle u,v\rangle}{\langle v,v\rangle} v = \alpha\left(\frac{1}{\|v\|} v\right)$$

Nogle bemærkninger (hvis $v \neq 0$):

I. $u - p$ og p er ortogonale.

Brug def. af skalarprojektion af p og find $\langle u - p, p\rangle$

II. $u = p \Leftrightarrow u$ er et skalarmultiplum af v. Dvs. $u = \beta v$

Hvis $u = \beta v$ så er $p = \frac{\langle \beta v, v\rangle}{\langle v,v\rangle} v = \beta v = u$

Hvis $u = p$, så: $u = \alpha\left(\frac{1}{\|v\|} v\right) = \beta v, \beta = \frac{\alpha}{\|v\|}$

Til sidst har vi Cauchy-Schwarz uligheden, som bevises herunder:

Sætning 5.4.2: v og u er to vektorer i et indre produkt rum V. Så gælder:

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

Med lighed $\Leftrightarrow$ u og v er lineært afhængige.

Bevis: Hvis $v = 0$ så har vi tydeligt, at:

$$|\langle u, v\rangle| = 0 = \|u\|\|v\|$$

Hvis $v \neq 0$, så er p vektorprojektionen af u på v. Hvis vi kigger på:

$$\|u\|^2 = \|p + (u - p)\|^2 \overset{p \perp u-p}{=} \|p\|^2 + \|u - p\|^2 \Rightarrow \|p\|^2 = \|u\|^2 - \|u - p\|^2$$

Projektionen kan også skrives på en anden måde:

$$\|p\|^2 = \left\|\frac{\langle u, v\rangle}{\langle v, v\rangle} v\right\|^2 = \left|\frac{\langle u, v\rangle}{\langle v, v\rangle}\right|^2 \|v\|^2 = \frac{(\langle u, v\rangle)^2}{\|v\|^4}\|v\|^2 = \frac{(\langle u, v\rangle)^2}{\|v\|^2} = \|u\|^2 - \|u - p\|^2$$

Dette kan igen skrives om til:

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

I ovenstående har vi kun lighed, hvis $p = u$.

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

Bemærkning II siger, at hvis $u = p$ så er u et skalarmultiplum af v - dvs. u og v er lineært afhængige.

Til sidst en sætning omkring norm:

Sætning trekantsulighed: Hvis V er et indre produkt rum, så gælder:

$$\|u + v\| \leq \|u\| + \|v\|$$

Bevis: Vi prøver at skrive det ud:

$$\begin{aligned}\|u + v\|^2 &= \langle u + v, u + v\rangle = \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2 \\ &= \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \\ &\overset{Cauchy-Schwarz}{\leq} \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

10. Ortogonalt komplement og projektion

- **Def. + bemærkninger**
- **Direkte sum**
- **Sætning 5.3.1**

Definition: Y er et underrum af $\mathbb{R}^n$. Mængden af alle vektorer i $\mathbb{R}^n$ som er ortogonale på enhver vektor i Y noteres $Y^\perp$. Dvs.

$$Y^\perp = \{x \in \mathbb{R}^n | x^T y = 0 \quad for\ alle\ \ y \in Y\}$$

Denne mængde kaldes for ortogonalkomplementet af Y.

Der er nogle forskellige egenskaber ved ortogonalkomplementer:

1. $X \cap Y = \{0\}$ hvis X og U begge er ortogonale underrum til $\mathbb{R}^n$
   Hvis $x \in X \cap Y$ og $X \perp Y$ så er $x^T x = 0$ og dermed er $x = 0$
2. Hvis Y er et underrum til $\mathbb{R}^n$ så er $Y^\perp$ også et underrum til $\mathbb{R}^n$
   Tjekke lukkethedsegenskaber for underrum.
3. $(S^\perp)^\perp = S$
4. $\dim S + \dim S^\perp = n$

Definition: Hvis U og V er underrum af et vektorrum W og alle $w \in W$ kan skrives ***unikt*** som en sum $u + v$, hvor $u \in U$ og $v \in V$ så er W en **direkte sum** af U og V og vi skriver: $W = U \oplus V$

Lemma 5.2.2: Hvis $\{x_1, \dots, x_r\}$ er basis for S og $\{x_{r+1}, \dots, x_n\}$ er basis for $S^\perp$, så er $\{x_1, \dots, x_r, x_{r+1}, \dots, x_n\}$ basis for $\mathbb{R}^n$

Sætning 5.2.3: Hvis S er et underrum af $\mathbb{R}^n$, så kan vi skrive: $\mathbb{R}^n = S \oplus S^\perp$

Bevis: Hvis $S = \{0\}$ eller $S = \mathbb{R}^n$ er resultatet trivielt.

Vi siger $\{x_1, \dots, x_r\}$ er en basis for S og $\{x_{r+1}, \dots, x_n\}$ er en basis for $S^\perp$. Så kan vi skrive $x \in \mathbb{R}^n$:

$$x = c_1 x_1 + \dots + c_r x_r + c_{r+1} x_{r+1} + \dots + c_n x_n$$

Så skriver vi $u = c_1 x_1 + \dots + c_r x_r$ og så $u \in S$ og vi skriver $v = c_{r+1} x_{r+1} + \dots + c_n x_n$ og så $v \in S^\perp$. Dermed har vi, at $x = u + v$. Vi skal vise, at dette gælder entydigt, antager vi:

$$x = u' + v' \ , \ u' \in S \ , \qquad v' \in S^\perp$$

Så har vi tydeligt, at

$$u + v = u' + v' \Rightarrow u - u' = v' - v$$

Så er $u - u' \in S$ og $v' - v \in S^\perp$. Dvs.:

$$u - u' = v' - v \in S \cap S^\perp = \{0\}$$

Dermed er $u - u' = v' - v = 0$ og $u = u'$ og $v = v'$.

Sætning 5.3.1: Lad S være et underrum af $\mathbb{R}^m$ og $b \in \mathbb{R}^m$. Da er projektionen p af b på S det nærmeste punkt til b i S:

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

p er unikt. Ydermere gælder det, at p kun vil være tættest på b, hvis og kun hvis $b - p \in S^\perp$

Bevis: $\Rightarrow$: Vi ved $\mathbb{R}^m = S \oplus S^\perp$, så vi kan skrive $b = p + z$, hvor $z \in S^\perp$. Vi kan så skrive:

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

Da $b - y \in S$ og $b - p = z \in S^\perp$ så kan vi bruge Pythagoras:

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

Da $y \neq p$ har vi heraf, at

$$\left|\left|b-y\right|\right| > \left|\left|b-p\right|\right|$$

$\Leftarrow$: Så, hvis vi vælger et $q \in S$ og $b - q \notin S^{\perp}$, så da $q \neq p$ har vi at $q = y$. Da y er alle elementer i S, som ikke er p, har vi igen: $\|b-q\| > \|b-p\|$

11. Ortogonale og ortonormale baser

- **Definition**
- **Lemma 5.5.7**
- **Gram-Schmidt**

Først vil vi definere nogle grundbegreber til senere brug:

Definition: Vektorerne $v_1, \dots, v_n$ danner en basis for vektorrummet Vhvis og kun hvis:

$(i)\ v_1, \dots, v_n \quad er\ lin\ uafh.$

$(ii)\ v_1, \dots, v_n \ \ udspænder\ V$

En mængde vektorer fra Ver **ortonormale**, hvis de er:

- **ortogonale** på hinanden: $\langle v_i, v_j \rangle = 0, i \neq j$

- enhedsvektorer.

Vi har så, at en mængde af vektorer er ortonormale, hvis:

$$\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1 & , \ i = j \\ 0 & , \ i \neq j \end{cases}$$

Herudover er mængden en **ortonormal basis**, hvis vektorerne er lineært uafhængige og udspænder V.

Vi vil gerne vise og bevise Gram-Schmidt, men først vil vi vise følgende sætninger, der bruges i beviset for Gram-Schmidt.

Lemma 5.5.7: Lad S være et underrum af det indre-produkt rum V. Lad $x \in V$ og $\{x_1, \dots, x_n\}$ være en ortonormal basis for S. $p \in S$ er projektionen af $x$på S. Så gælder:

$$p - x \in S^\perp \Leftrightarrow p = \sum_{i=1}^{n} \langle x, x_i \rangle \, x_i$$

Bevis: Da $p \in S$ kan vi skrive: $p = c_1 x_1 + \dots + c_n x_n$. Ifølge 5.5.2, kan vi skrive: $p = \sum_{i=1}^{n} c_i x_i$, $c_i = \langle p, x_i \rangle$. For alle $s \in S$ har vi så:

$$\begin{aligned} p - x \in S^\perp &\Leftrightarrow \langle s, p - x \rangle = 0 \\ &\Leftrightarrow \langle c_1 x_1 + \dots + c_n x_n, p - x \rangle = 0 \quad , \quad c_i\ skalar \\ &\Leftrightarrow c_1 \langle x_1, p - x \rangle + \dots + c_n \langle x_n, p - x \rangle = 0 \quad , \qquad c_i\ skalar \\ &\Leftrightarrow \langle x_i, p - x \rangle = 0 \quad , \quad i = 1, \dots, n \\ &\Leftrightarrow \langle x_i, p \rangle = \langle x_i, x \rangle \quad , \quad i = 1, \dots, n \\ &\Leftrightarrow \langle p, x_i \rangle = \langle x, x_i \rangle \quad , \quad i = 1, \dots, n \\ &\Leftrightarrow p = \sum_{i=1}^{n} \langle x, x_i \rangle \, x_i \end{aligned}$$

Nu vil vi vise Gram-Schmidt, som bruges til at danne en ortonormal basis ud fra en anden basis. Processen går ud på, at man først gør de enkelte vektorer i basen til enhedsvektorer, og derefter retter man dem op med projektion.

Sætning 5.6.1 (Gram-Schmidt): Lad $\{x_1, \dots, x_n\}$ være en basis til det indre produkt rum V. Lad

$$u_1 = \left(\frac{1}{\|x_1\|} \right) x_1$$

og definer $u_2, \dots, u_n$ rekursivt:

$$u_{k+1} = \frac{1}{\|x_{k+1} - p_k\|} (x_{k+1} - p_k) \ , \ \ k = 1, \dots, n-1$$

hvor vi har defineret projektionen af x_{k+1} på $Span(u_1, \dots, u_k)$:

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

Så er $\{u_1, \ldots, u_k\}$ en ortonormal basis for $Span(x_1, \ldots, x_k)$. Mængden $\{u_1, \ldots, u_n\}$ er en ortonormal basis for *V*.

Bevis: *Basis:* Det ses tydeligt, at det gælder for $Span(u_1) = Span(x_1)$.

IH: Vi siger det gælder for k, så der er konstrueret en mængde $\{u_1, \ldots, u_k\}$ som er ortonormal for S_k og vi har:

$$Span(u_1, \ldots, u_k) = Span(x_1, \ldots, x_k) = S_k \quad , \quad k < n$$

Induktionsskridt: $k + 1$.

p_k er projektionen af x_{k+1} på S_k. Så skriver vi (jf. 5.5.7):

$$p_k = \langle x_{k+1}, u_1 \rangle u_1 + \cdots + \langle x_{k+1}, u_k \rangle u_k = \sum_{i=1}^{k} \langle x_{k+1}, u_i \rangle u_i$$

1) Først skal vi vise, at $\boldsymbol{x_{k+1} - p_k}$ er lin. uafh. og ikke-nul.

Da $p_k \in S_k$ kan vi skrive p_k som en linear kombination af $x_1, \ldots, x_k$:

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

Vi kigger så på:

$$x_{k+1} - p_k = x_{k+1} - c_1 x_1 - \cdots - c_k x_k$$

Og vi har her, at $x_{k+1} - p_k \in S_{k+1}$ og da $x_1, \ldots, x_{k+1}$ er lineært uafhængig da $\{x_1, \ldots, x_n\}$ er en basis, så giver ovenstående ligning ikke-nul: $x_{k+1} - p_k \neq 0$ (da den har en ikke-nul skalar).

2) Så skal vi vise, at $\boldsymbol{x_{k+1} - p_k \perp u_i}$, $\boldsymbol{i = 1, \ldots, k}$ (et krav for ONB):

Lemma 5.5.7 fortæller, at $x_{k+1} - p_k \in S_k^{\perp}$ pga. def. af p_k og dermed er

$$x_{k+1} - p_k \perp u_i \quad , \quad i = 1, \ldots, k$$

3) Til sidst skal vi vise, at vi kan danne en ny enhedsvektor ud fra $\boldsymbol{x_{k+1} - p_k}$:

Hvis vi nu siger, at:

$$u_{k+1} = \frac{x_{k+1} - p_k}{\|x_{k+1} - p_k\|}$$

Så er $\{u_1, \ldots, u_{k+1}\}$ ortonormal (enhedsvektorer og ortogonale på hinanden) og indeholdt i S_{k+1}. Da $u_1, \ldots, u_{k+1}$ er lineært uafhængige, så udgør de en ortonormal basis for S_{k+1}.

12. Ortogonale og unitære matricer

- **Def. Ortogonal matrix**
- **Sætning 5.5.5**
- **Def. Unitær matrix**
- **Schurs sætning**

Først vil vi definere nogle grundbegreber til senere brug:

Definition: En mængde vektorer er **ortonormale**, hvis de er:

- **ortogonale** på hinanden: $\langle v_i, v_j \rangle = 0 \ \ , i \neq j$

- enhedsvektorer.

Vi har så, at en mængde af vektorer er ortonormale, hvis:

$$\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1 & , \ \ i = j \\ 0 & , \ \ i \neq j \end{cases}$$

Nu vil vi definere ortogonal matricer:

Definition: En $n \times n$ matrix er en **ortogonal matrix**, hvis søjlevektorerne udgør en ortonormal mængde i $\mathbb{R}^n$.

Og en lille sætning:

Sætning 5.5.5: En $n \times n$-matrix Q er ortogonal $\Leftrightarrow Q^T Q = I$

Bevis: Det følger af definitionerne, at Q kun er ortogonal, hvis og kun hvis

$$q_i^T q_j = \langle q_i, q_j \rangle = \delta_{ij}$$

Altså, at ovenstående kun giver 1, hvis $i = j$. Da $q_i^T q_j$ også er den (i, j)'te indgang i $Q^T Q$, må alle diagonalindgangene være 1, og vi har $Q^T Q = I$.

Unitære matricer er stort set det samme, bare i det komplekse rum $\mathbb{C}^n$.

Definition: En $n \times n$ matrix er en **unitær matrix**, hvis søjlevektorerne udgør en ortonormal mængde i $\mathbb{C}^n$. Hvis U er unitær gælder $U^H U = I$.

Når vi taler om komplekse matricer, er der nogle forskellige operationer og definitioner, som er nyttige.

Først og fremmest har vi, at hvis vi kompleks konjugerer og transponerer en matrix noteres det således (med de følgende regneregler, som ligner transponeringsregnereglerne meget):

$$\bar{M}^T = M^H$$

$$(A^H)^H = A$$

$$(\alpha A + \beta B)^H = \bar{\alpha} A^H + \bar{\beta} B^H$$

$$(AC)^H = C^H A^H$$

Til sidst, siger vi, at en matrix er **hermitisk,** hvis $M = M^H$ (ligesom symmetriske matricer i $\mathbb{R}$)

Sætning 6.4.3 (Schurs sætning): For hver $n \times n$-matrix A eksisterer der en unitær matrix U således, at $U^H A U$ er øvre triangulær (noteret T).

Bevis: Vi inducerer på n.

Basis: $n = 1$. Åbenlyst korrekt.

IH: Vi antager at det gælder for en $k \times k$-matrix.

Induktionsskridt: $n = k + 1$. Så vi kigger på en $(k + 1) \times (k + 1)$-matrix A. Vi siger, at λ_1 er egenværdi til A med tilhørende **enheds**egenvektor w_1.

Vi bruger Gram-Schmidt til at finde $w_2, \dots, w_{k+1}$ således at $\{w_1, \dots, w_{k+1}\}$ er en ortonormal basis for $\mathbb{C}^{k+1}$. Så lader vi W være matricen bestående af w_i , $i = 1, \dots, k+1$ som søjlevektorer. Dermed opfylder W definitionen for en unitær matrix.

Første søjle i $W^H AW$ er så:

$$W^H A w_1 = \lambda_1 W^H w_1 = \lambda_1 e_1$$

(da e_1 er første søjle i identitsmatricen)

Så ser vi på, hvordan $W^H AW$ ser ud:

$$W^H AW = \begin{pmatrix} \lambda_1 & | & \times\times\times\times\times \\ \hline 0 & | & \\ 0 & | & M \\ 0 & | & \end{pmatrix}$$

Hvor M er en $k \times k$-matrix. Vores IH siger, at sætningen gælder for M, så vi har en unitær matrix V_1 således at $V_1^H M V_1 = T_1$ hvor T_1 er øvre triangulær. Så definerer vi:

$$V = \begin{pmatrix} 1 & | & 0000000 \\ \hline 0 & | & \\ 0 & | & V_1 \\ 0 & | & \end{pmatrix}$$

V er unitær. Så prøver vi at gange hhv. V^H og V på den første matrix:

$$V^H W^H AWV = \begin{pmatrix} 1 & | & 0000000 \\ \hline 0 & | & \\ 0 & | & V_1^H \\ 0 & | & \end{pmatrix} \begin{pmatrix} \lambda_1 & | & \times\times\times\times\times \\ \hline 0 & | & \\ 0 & | & M \\ 0 & | & \end{pmatrix} \begin{pmatrix} 1 & | & 0000000 \\ \hline 0 & | & \\ 0 & | & V_1 \\ 0 & | & \end{pmatrix} = \begin{pmatrix} \lambda_1 & | & \times\times\times\times\times \\ \hline 0 & | & \\ 0 & | & V_1^H M V_1 \\ 0 & | & \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & | & \times\times\times\times\times \\ \hline 0 & | & \\ 0 & | & T_1 \\ 0 & | & \end{pmatrix} = T$$

Da $V^H W^H AWV = (WV)^H AWV$ sætter vi $U = WV$ og ser om U er unitær:

$$U^H U = (WV)^H WV = V^H W^H WV = I$$

13. Unitær diagonalisering

- **Def. Unitær matrix**
- **Schurs sætning**
- **Spektralsætningen**

Definition: En $n \times n$ matrix er en **unitær matrix**, hvis søjlevektorerne udgør en ortonormal mængde i $\mathbb{C}^n$. Hvis U er unitær gælder $U^H U = I$ og $U^{-1} = IU^{-1} = U^H UU^{-1} = U^H$

En mængde af vektorer $\{u_1, \dots, u_n\}$ er ortonormale, hvis:

$$\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1 & , \quad i = j \\ 0 & , \quad i \neq j \end{cases}$$

Når vi taler om komplekse matricer, er der nogle forskellige operationer og definitioner, som er nyttige.

Først og fremmest har vi, at hvis vi kompleks konjugerer og transponerer en matrix noteres det således (med de følgende regneregler, som ligner transponeringsregnereglerne meget):

$$\bar{M}^T = M^H$$

$$(A^H)^H = A$$

$$(\alpha A + \beta B)^H = \bar{\alpha} A^H + \bar{\beta} B^H$$

$$(AC)^H = C^H A^H$$

Til sidst, siger vi, at en matrix er **hermitisk,** hvis $M = M^H$ (ligesom symmetriske matricer i $\mathbb{R}$)

Sætning 6.4.3 (Schurs sætning): For hver $n \times n$-matrix A eksisterer der en unitær matrix U således, at $U^H AU$ er øvre triangulær (noteret T).

Bevis: Vi inducerer på n.

Basis: $n = 1$. Åbenlyst korrekt.

IH: Vi antager at det gælder for en $k \times k$-matrix.

Induktionsskridt: $n = k + 1$. Så vi kigger på en $(k+1) \times (k+1)$-matrix A. Vi siger, at λ_1 er egenværdi til A med tilhørende **enheds**egenvektor w_1.

Vi bruger Gram-Schmidt til at finde $w_2, \dots, w_{k+1}$ således at $\{w_1, \dots, w_{k+1}\}$ er en ortonormal basis for $\mathbb{C}^{k+1}$. Så lader vi W være matricen bestående af w_i , $i = 1, \dots, k+1$ som søjlevektorer. Dermed opfylder W definitionen for en unitær matrix.

Første søjle i $W^H AW$ er så:

$$W^H A w_1 = \lambda_1 W^H w_1 = \lambda_1 e_1$$

(da e_1 er første søjle i identitsmatricen)

Så ser vi på, hvordan $W^H AW$ ser ud:

$$W^H AW = \left(\begin{array}{c|c} \lambda_1 & \times\times\times\times\times \\ \hline 0 & \\ 0 & M \\ 0 & \end{array} \right)$$

Hvor M er en $k \times k$-matrix. Vores IH siger, at sætningen gælder for M, så vi har en unitær matrix V_1 således at $V_1^H M V_1 = T_1$ hvor T_1 er øvre triangulær. Så definerer vi:

$$V = \left(\begin{array}{c|c} 1 & 0000000 \\ \hline 0 & \\ 0 & V_1 \\ 0 & \end{array} \right)$$

V er unitær. Så prøver vi at gange hhv. V^H og V på den første matrix:

$$V^HW^HAWV = \left(\begin{array}{c|c} 1 & 0000000 \\ \hline 0 & \\ 0 & V_1^H \\ 0 & \end{array}\right)\left(\begin{array}{c|c} \lambda_1 & \times\times\times\times\times \\ \hline 0 & \\ 0 & M \\ 0 & \end{array}\right)\left(\begin{array}{c|c} 1 & 0000000 \\ \hline 0 & \\ 0 & V_1 \\ 0 & \end{array}\right) = \left(\begin{array}{c|c} \lambda_1 & \times\times\times\times\times \\ \hline 0 & \\ 0 & V_1^HMV_1 \\ 0 & \end{array}\right)$$

$$= \left(\begin{array}{c|c} \lambda_1 & \times\times\times\times\times \\ \hline 0 & \\ 0 & T_1 \\ 0 & \end{array}\right) = T$$

Da $V^HW^HAWV = (WV)^HAWV$ sætter vi $U = WV$ og ser om U er unitær:

$$U^HU = (WV)^HWV = V^HW^HWV = I$$

Faktoriseringen $A = UTU^H$ kaldes Schur… Hvis A er hermitisk, så vil T være en diagonalmatrix.

Sætning 6.4.4 (spektral sætningen): Hvis A er hermitisk så eksisterer der en unitær matrix U, som diagonaliserer A.

Bevis: Vi ved fra beviset før, at der findes en unitær matrix U således at $U^HAU = T$, hvor T er øvre triangulær. Vi kigger på T:

$$T^H = (U^HAU)^H = U^HA^H(U^H)^H = U^HAU = T$$

Da T er øvre triangulær og hermitisk, *må* den også være diagonal (da man transponerer indgangene).

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & t_{nn} \end{pmatrix} \Rightarrow T^H = \begin{pmatrix} \overline{t_{11}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \overline{t_{1n}} & \cdots & \overline{t_{nn}} \end{pmatrix}$$

14. Kvadratiske former

- **Def.**
- **Sætning 6.6.1**
- **Definit**
- **Sætning 6.6.2**

Til hver kvadratisk ligning, kan der associeres en vektorfunktion $f(x) = x^T Ax$, som kaldes den kvadratiske form. Bruges specielt i optimeringsteori.

Definition: En kvadratisk ligning med to variabler x og y kan skrives på formen:

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

Dette kan skrives på matrixform:

$$(x \;\; y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (d \;\; e)\begin{pmatrix} x \\ y \end{pmatrix} + f = 0$$

Hvis vi så siger, $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ og $x = \begin{pmatrix} x \\ y \end{pmatrix}$ så kaldes:

$$x^T Ax = ax^2 + 2bxy + cy^2$$

den kvadratiske form tilhørende den kvadratiske ligning.

Korollar 6.4.5: Hvis A er en reel symmetrisk matrix, dvs. $A = A^T$, så eksisterer der en ortogonal matrix Q, som diagonaliserer A, således at $Q^T AQ = D$, hvor D er diagonal.

Givet en kvadratisk ligning kan vi oversætte og rotere den, så vi får den på "standard form". For det generelle n-dimensionelle tilfælde kan den kvadratiske form altid oversættes til en mere simpel diagonal form:

Sætning 6.6.1: Hvis A er en reel symmetrisk $n \times n$-matrix, så kan vi skifte variabler til $u = Q^T x$ således at $x^T Ax = u^T Du$ hvor D er en diagonalmatrix.

Bevis: Korollar 6.4.5 siger, at til en reel symmetrisk matrix A, findes der en ortogonalmatrix, som diagonaliserer A: $Q^T AQ = D$.

Hvis $u = Q^T x$, så er $Qu = QQ^T x \Rightarrow x = Qu$ og $x^T = u^T Q^T$

Vi tager den kvadratiske form, og indsætter ovenstående:

$$x^T Ax = u^T Q^T AQu = u^T Du$$

I forbindelse med optimering af funktioner med flere variabler, ved vi, hvis funktionen er en kvadratisk form, så er det kritiske punkt 0. Om dette er maks., min. eller saddelpunkt afhænger af egenværdierne.

Hvis vi kigger på de kvadratiske former, så siger vi, at en $f(x)$ har globalt minimum i **0**, hvis og kun hvis $x^T Ax > 0 \;\;, \;\; \forall x \neq 0$. Vi siger også, at $f(x)$ har globalt maximum i **0**, hvis og kun hvis $x^T Ax < 0 \;\;, \;\; \forall x \neq 0$. Hvis $x^T Ax$ skifter fortegn, er det et saddelpunkt. Generelt har vi følgende definition:

Definition: Vi siger, at en kvadratisk form $f(x) = x^T Ax$ er **definit**, hvis den for forskellige x ikke skifter fortegn. **En reel symmetrisk matrix A**

er **positiv definit**, hvis $x^T Ax > 0 \;\;, \;\; \forall x \neq 0$

er **negativ definit**, hvis $x^T Ax < 0 \;\;, \;\; \forall x \neq 0$

er **ikke-definit,** hvis der er fortegnskift for forskellig x

er **postiv semi-definit**, hvis $x^T Ax \geq 0 \;\;, \;\; \forall x \neq 0$

er **negativ semi-definit**, hvis $x^T Ax \leq 0 \;\;, \;\; \forall x \neq 0$

Hvis A er invertibel så er 0 det eneste stationære punkt. Jf. definition og teksten før, har vi så nu, at hvis en kvadratisk form er positiv definit er det et globalt minimum, hvis den er negativ definit er det et globalt maksimum og hvis den er ikke-definit er det et saddel-punkt. For at finde ud af, hvad et punkt er, må vi altså finde ud af, hvad matricen er. Her er næste sætning nyttig:

Sætning 6.6.2: A er en reel symmetrisk $n \times n$-matrix, så er A positiv definit, hvis og kun hvis alle A's egenværdier er positive.

Bevis: $\Rightarrow$: Hvis A er positiv definit og λ er en egenværdi til A, så for en egenvektor x tilhørende λ:

$$x^T A x = \lambda x^T x = \lambda \|x\|^2$$

Dermed har vi:

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

Og alle egenværdier er positive.

$\Leftarrow$: Hvis alle egenværdier for A er positive, så lader vi $\{x_1, \dots, x_n\}$ være en ortonormal mængde af egenvektorer for A. Hvis x er en ikke-nul vektor, så kan vi skrive den som en linear kombination:

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

Hvor $\alpha_i = x^T \quad , \quad = 1, \dots, n$ $\qquad \sum_{i=1}^{n} (\alpha_i)^2 = \|x\|^2 > 0$

Hvis vi så kigger på den kvadratiske form:

$$x^T A x = (\alpha_1 x_1 + \cdots + \alpha_n x_n)^T (\alpha_1 \lambda_1 x_1 + \cdots + \alpha_n \lambda_n x_n)$$

$$= (\alpha_1)^2 \lambda_1 x_1^T x_1 + \cdots + (\alpha_n)^2 \lambda_n x_n^T x_n = \sum_{i=1}^{n} (\alpha_i)^2 \lambda_i$$

$$\geq (\min \lambda_i) \|x\|^2 > 0$$

Og dermed er A positiv definit.

15. Lineære differentialligninger

- **Løsninger**
- **Begyndelsesværdiproblemet**
- **Putzers algoritme**

Vi betragter først et lineært differentialligningssystem:

$$\begin{array}{ccc} y_1' & = & a_{11}y_1 + \cdots + a_{1n}y_n \\ \vdots & = & \vdots \\ y_n' & = & a_{n1}y_1 + \cdots + a_{nn}y_n \end{array}$$

Hvor $y_i(t) = f_i(t)$ er en kontinuert funktion. Dette system kan skrives mere simpelt op på formen:

$$\boldsymbol{Y}' = A\boldsymbol{Y}$$

Hvor $\boldsymbol{Y}'$og $\boldsymbol{Y}$er vektorfunktioner af $t$og Aer en $n \times n$-matrix.

For det simple tilfælde $\boldsymbol{n} = \boldsymbol{1}$: $y' = ay$, har vi løsningen $y(t) = ce^{at}$, hvor $c$er en konstant.

En generel løsning for $\boldsymbol{n} > 1$ med egenværdier og egenvektorer for A er:

$$\boldsymbol{Y} = \begin{pmatrix} x_1 e^{\lambda t} \\ \vdots \\ x_n e^{\lambda t} \end{pmatrix} = e^{\lambda t}\boldsymbol{x} \quad , \quad hvor\ \boldsymbol{x}\ er\ en\ vektor$$

For at se om dette er en løsning, differentierer vi det:

$$\boldsymbol{Y}' = \lambda e^{\lambda t}\boldsymbol{x} = \lambda \boldsymbol{Y}$$

Hvis vi siger, at λ er en egenværdi til A med tilhørende egenvektor $\boldsymbol{x}$, så har vi:

$$A\boldsymbol{Y} = e^{\lambda t}A\boldsymbol{x} = \lambda e^{\lambda t}\boldsymbol{x} = \lambda \boldsymbol{Y} = \boldsymbol{Y}'$$

Dermed har vi vist, at $\boldsymbol{Y} = e^{\lambda t}\boldsymbol{x}$ er en løsning til differentialligningssystemet.

Hvis vi siger, at $\boldsymbol{Y}(t)$ er en løsning, så skal den have en bestemt værdi $\boldsymbol{Y}_0$ for $t = 0$, så siger en sætning, at ***begyndelsesværdiproblemet*** vil have en unik løsning.

$$\boldsymbol{Y}' = A\boldsymbol{Y} \quad , \quad \boldsymbol{Y}(0) = \boldsymbol{Y}_0$$

En måde at løse disse problemer på er med **matrixeksponentialet**. For $n = 1$ har vi igen, at løsningen er på formen:

$$y = e^{at}y_0 \quad , \quad y(0) = y_0$$

Dette generaliserer vi igen til $n > 1$ og prøver at indsætte e^{tA} i ovenstående løsning:

$$\boldsymbol{Y}(t) = e^{tA}\boldsymbol{Y}_0 \quad , \quad \boldsymbol{Y}(0) = e^{0A}\boldsymbol{Y}_0 = 1\boldsymbol{Y}_0 = \boldsymbol{Y}_0$$

Vi kan tjekke om det er en løsning ved at differentiere den, og se om den giver differentialligningen. Bemærk, at $\frac{d}{dt}e^{tA} = Ae^{tA}$:

$$\boldsymbol{Y}' = Ae^{tA}\boldsymbol{Y}_0 = A\boldsymbol{Y}$$

Generelt er det nemt at finde løsninger til differentialligningssystemer, hvis bare A er diagonaliserbar:

$$\boldsymbol{Y} = e^{tA}\boldsymbol{Y}_0 = Xe^{tD}X^{-1}\boldsymbol{Y}_0$$

Hvor D er diagonalmatrixen, og X er diagonaliseringsmatrixen. Så handler det blot om at finde egenværdier og egenvektorer for A. Hvis ikke A er diagonaliserbar, bruger vi Putzers algoritme til at finde matrixeksponetialet e^{tA}:

Sætning (Putzers algoritme): A er en $n \times n$-matrix i det komplekse rum med $\lambda_1, \dots, \lambda_n$ egenværdier. Så lader vi

$$P_0 = I \quad , \quad P_k = \prod_{j=1}^{k}(A - \lambda_j I) \quad , \quad k = 1, \dots, n$$

Vi definerer så:

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

Hvor $r_1(t) = e^{\lambda_1 t}$ og

$$r_k(t) = e^{\lambda_k t} \int_0^t e^{-\lambda_k s} r_{k-1}(s) ds \quad , \quad k = 2, \ldots, n$$

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

Bevis: Først ser vi på r og Q, hvis $t = 0$ indsættes:

$$r_1(0) = e^{\lambda_1 0} = 1 \quad , \quad r_k(0) = e^{\lambda_k 0} \int_0^0 e^{-\lambda_k s} r_{k-1}(s) ds = 0$$

$$Q(0) = \sum_{k=0}^{n-1} r_{k+1}(0) P_k = r_1(0) P_0 + \sum_{k=1}^{n-1} r_{k+1}(0) P_k = 1P_0 + 0 = P_0 = I$$

Det er klart, at A kommuterer med $(A - \lambda_i I)$: $A(A - \lambda_i I) = (A - \lambda_i I)A$. Dermed kommuterer A også med $P_0, \ldots, P_{n-1}$ og så med $Q(t)$ dvs.: $AQ(t) = Q(t)A$.

Så kigger vi på $k > 1$. Vi starter med at differentiere r:

$$r_k'(t) = \lambda_k e^{\lambda_k t} \int_0^t e^{-\lambda_k s} r_{k-1}(s) ds + e^{\lambda_k t} e^{-\lambda_k t} r_{k-1}(t) = \lambda_k r_k(t) + r_{k-1}(t)$$

Vi definerer $r_0(t) = 0$, så gælder dette også for $k = 1$.

Herefter differentierer vi Q:

$$Q'(t) = \sum_{k=0}^{n-1} r_{k+1}'(t) P_k = \sum_{k=0}^{n-1} \big(\lambda_{k+1} r_{k+1}(t) + r_k(t)\big) P_k$$

Til sidst skal vi se, om $Q'(t) = AQ(t)$, så vi trækker højresiden fra venstresiden og ser om det giver 0:

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

her skal vi have det karakteristiske pol frem $(A - \lambda_{k+1} I)$
så vi rykker rundt på leddene og sætter uden for ( )

$$\begin{aligned} &= \sum_{k=0}^{n-1} (-r_{k+1}(t)(A - \lambda_{k+1} I) P_k + r_k(t) P_k) \\ &= \sum_{k=0}^{n-1} (-r_{k+1}(t) P_{k+1} + r_k(t) P_k) \quad (*) \\ &= -r_n(t) P_n \\ &= 0 \end{aligned}$$

$(*)$ Her æder leddene hinanden således, at vi ender med:

$$\begin{aligned} &= (-r_1(t)P_1 + r_0(t)P_0) + (-r_2(t)P_2 + r_1(t)P_1) + \cdots + (-r_{n-1}(t)P_{n-1} + r_{n-2}(t)P_{n-2}) \\ &\qquad + (-r_n(t)P_n + r_{n-1}(t)P_{n-1}) \\ &= -r_n(t) P_n \end{aligned}$$

Ved sidste lighedstegn bruger vi Cayley-Hamilton-sætningen, da

$$P_n = p_A(A) = 0$$

$Q(t)$ er den differentiable matrixfunktion, der opfylder $\exp(0A) = I$ og $\exp(tA)' = A\exp(tA)$. Dermed kan vi bruge $Q(t)$ til at løse differentialligninger med.

16. Markovprocesser

- **Definition**
- **Sætning 6.3.3**
- **Dominerende egeværdi**
- **Sætning 6.3.4**

Definition: En stokastisk proces en sekvens af eksperimenter, hvor udfaldet afhænger af sandsynlighed. En **Markovproces** er en stokastisk proces, der opfylder følgende:

I. Mængden af mulige udfald er endelig.
II. Sandsynligheden for næste udfald afhænger udelukkende af det foregående.
III. Sandsynlighederne er konstante over tid.

En **Markovkæde** består af en række tilstandsvektorer $x_0, \dots, x_n$ som også er sandsynlighedsvektorer, hvor x_0 kaldes startvektoren. A er en $n \times n$-matrix, som vi kalder transitionsmatrixen (stokastisk matrix) for Markovprocessen. Vi har så

$$x_0\,,\ x_1 = Ax_0\,, x_2 = Ax_1\ \ , \dots$$

Dvs. mere generelt: $x_n = Ax_{n-1}$

En matrix er stokastisk, hvis indgangene er ikke-negative og summen af søjlerne er 1. Så er søjlerne også sandsynlighedsvektorer.

Sætning 6.3.3: Hvis A er en stokastisk matrix, $x_0 \in \mathbb{R}^n$ er en sandsynlighedsvektor. Vi definerer $x_k = A^k x_0\,, k = 1, \dots$, og vi antager $x_k \xrightarrow{k\to\infty} x$ (Markovkæden konvergerer mod *x*). Så gælder der om den stabile tilstandsvektor *x*:

(i) *x* er en sandsynlighedsvektor (ikke-negative indgange, som summer op til 1).
(ii) $\lambda_1 = 1$ er en egenværdi til A og *x* er en egenværdi tilhørende λ_1.

Bevis for (i): Da x_k er en sandsynlighedsvektor gælder det følgende:
Hvis $(x)_i$ er *i*'te indgang i *x* og $(x_k)_i$ er *i*'te indgang i x_k, så har vi: $(x)_i = \lim_k (x_k)_i$. Da $(x_k)_i \geq 0$ er $(x)_i \geq 0$, og *x* har altså ikke-negative indgange. Vi kontrollere om indgangene for *x* summer op til 1:

$$e^T x = e^T\left(\lim_{k\to\infty} x_k\right) = \lim_{k\to\infty} e^T x_k = \lim_{k\to\infty} e^T\left(A^k x_0\right)$$
$$= \lim_{k\to\infty}\left(e^T A^k\right)x_0 \overset{da\ A^k\ består\ af\ ss-vektorer}{=} \lim_{k\to\infty} e^T x_0 = \lim_{k\to\infty} 1 = 1$$

Hvor $e^T = (1, \dots, 1)$.

Bevis for (ii): $Ax = A(\lim_{k\to\infty} x_k) = \lim_{k\to\infty} Ax_k \overset{def}{=} \lim_{k\to\infty} x_{k+1} \overset{sætn}{=} x$

Der eksisterer altid en stabil tilstandsvektor, hvis $\lambda_1 = 1$ er en **dominerende egenværdi** til *A*. Den er dominerende, hvis $|\lambda_j| < |\lambda_1|\ \ , i = 2, \dots, n$. Vi husker på at alle indgange i *A* er positive.

Sætning 6.3.4: Hvis $\lambda_1 = 1$ er en dominerende egenværdi til *A*, så vil Markovkæden med transitionen *A* konvergere mod en stabil tilstandsvektor.

Bevis: Når A er diagonaliserbar, så er y_1 en egenvektor tilhørende $\lambda_1 = 1$ og lad $Y = (y_1, \dots, y_n)$ være den matrix, der diagonaliserer *A*. Vi definerer $E = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in Mat_{n,n}$

$$D^k = \begin{pmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{pmatrix} \xrightarrow{k\to\infty} E$$

Hvis så x_0 er startsandsynlighedsvektoren og $c = Y^{-1}x_0$, så er:

$$x_k = A^k x_0 = YD^kY^{-1}x_0 = YD^kc \xrightarrow{k\to\infty} YEc = Y(c_1e_1) = c_1\boldsymbol{y_1}$$

Så c_1y_1 er den stabile tilstandsvektor til Markovkæden. Dette passer jo også fint med, at y_1 er en egenvektor tilhørende λ_1 (c_1 er jo bare en skalar).