NTNU – Trondheim
Norwegian University of
Science and Technology

MA2002 - Bachelorprosjekt i matematiske fag

Algebraens fundamentalteorem

---

av
Maren Mortensdatter Isachsen

veileder
Øyvind Solberg

May 31, 2021

Contents

1 Introduksjon

Algebraens fundamentalteorem lyder som følgende:

Alle polynomer med komplekse koeffisienter har en kompleks rot.

I denne oppgaven skal vi se på et bevis for fundamentalteorem utledet av Harm Derksen. [1]. Beviset baserer seg på lineær algebra og går ut på å se på $n \times n$-matriser som kommuterer og vise at de har en felles egenvektor. Gjennom å vise dette ønsker vi å bevise at det karakteristiske polynomet har en egenverdi, og dermed eksisterer det et nullpunkt.

Oppgaven er delt opp i fem seksjoner, hvor den første er introduksjonen. Den andre seksjonen gjengir regneregler vi har for matriser. I den tredje seksjonen blir preliminære resultater presentert. Disse resultatene er ikke direkte knyttet opp til beviset av algebraens fundamentalteorem, men vi kommer til å bruke de som støtte. Harm Derksen sitt bevis blir presentert i seksjon fire. Avslutningsvis så kommer det til å bli presentert to alternative bevis for fundamentalteorem, samt konsekvenser av det.

2 Regneregler for matriser

I oppgaven kommer jeg til å bruke en del regneregeler for matriser som er lurt å kunne. La C og D være to vilkårlige matriser, da gjelder:

$$\begin{aligned}
(C+D)^{\mathsf{T}} &= C^{\mathsf{T}} + D^{\mathsf{T}} \\
(CD)^{\mathsf{T}} &= D^{\mathsf{T}} C^{\mathsf{T}} \\
(aC)^{\mathsf{T}} &= aC^{\mathsf{T}} \\
(C^{T})^{\mathsf{T}} &= C \\
\overline{C+D} &= \overline{C} + \overline{D} \\
\overline{aC} &= \overline{a}\overline{C} \\
\overline{ab} &= \overline{a}\overline{b} \\
(CD)^{*} &= D^{*} C^{*}
\end{aligned}$$

3 Preliminære resultater

Før vi ser på Harm Derksens bevis for algebraens fundamental teorem skal vi se på noen preliminære resultater som blir nyttige senere. Det første resultatet vi ser på angår polynomer av odde grad og deres nullpunkt. Resultatet kan ikke vises algebraisk, og vi kommer derfor til å bruke skjæringssetningen fra analysen.

Lemma 3.1. *Alle polynom av odde grad med reelle koeffisienter har et nullpunkt.*

Bevis. Vi har et vilkårlig polynom

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0,$$

hvor $a_0, ..., a_n \in \mathbb{R}$, $a_n \neq 0$ og n er odde.

Vi ser på hva som skjer med nullpunktene til polynomet $P(x)$ dersom vi deler eller ganger med en ikke-null konstant k.

1. $\forall k \in \mathbb{R}$: $P(x) = 0 \implies P(x) \cdot k = 0$
2. $\forall k \in \mathbb{R}, k \neq 0$: $P(x) \cdot k = 0 \implies P(x) \cdot k \cdot k^{-1} = 0 \implies P(x) = 0$

Nullpunktene vil ikke endre seg, og dermed kan vi uten å miste generalitet anta at $P(x)$ er et monisk polynom. At polynomet er monisk vil si at $a_n = 1$, og vi kan sette $a_n = 1$ for resten av dette beviset.

Vi kan skrive $P(x) = x^n R(x)$, hvor

$$R(x) = 1 + a_{n-1}\frac{1}{x} + a_{n-2}\frac{1}{x^2} + \cdots + a_1 \frac{1}{x^{n-1}} + a_0 \frac{1}{x^n}.$$

Grensene til $R(x)$ blir da

$$\lim_{x \to \infty} R(x) = \lim_{x \to -\infty} R(x) = 1.$$

Siden n er odde får vi

$$\lim_{x\to\infty} P(x) = \lim_{x\to\infty} x^n = +\infty$$

og

$$\lim_{x\to-\infty} P(x) = \lim_{x\to-\infty} x^n = -\infty.$$

Fordi grensene til $P(x)$ går mot $\pm\infty$, må det $\exists a \in \mathbb{R}$ s.a. $P(a) > 0$ og $P(-a) < 0$. Siden $P(x)$ er reell og kontinuerlig i det lukkende intervallet $[-a, a]$ kan vi bruke skjærringssetningen. Setningen sier at det eksisterer en $\lambda \in [-a, a]$ s.a. $P(\lambda) = 0$ og $P(x)$ har et nullpunkt.

■

Det andre resultatet som blir presentert er at alle komplekse tall har en kvadratrot. Dette kommer vi til å bruke senere i oppgaven når vi vil faktorisere komplekse polynomer.

Lemma 3.2. *Alle komplekse tall har en kvadratrot.*

Bevis. Vi har et vilkårlig komplekst ikke-null tall $z = \alpha + \beta i$ med $\alpha, \beta \in \mathbb{R}$. Vi vil vise at z har er en kvadratrot, dvs. at det finnes et komplekst tall w s.a.

$$w^2 = z.$$

Vi setter $\gamma = \sqrt{\alpha^2 + \beta^2}$. Siden både α og β er reelle tall vil γ være et positivt reelt tall. Vi kan skrive om på likningen

$$\begin{gathered} \gamma = \sqrt{\alpha^2 + \beta^2} \\ \Longleftrightarrow \\ \gamma^2 = \alpha^2 + \beta^2 \\ \Longleftrightarrow \\ \gamma^2 - \alpha^2 = \beta^2 \\ \Longleftrightarrow \\ (\gamma + \alpha)(\gamma - \alpha) = \beta^2. \end{gathered}$$

Siden vi vet at β^2 er positiv vil $(\gamma + \alpha)$ og $(\gamma - \alpha)$ begge enten være positive eller negative. Vi vet at $\gamma \geq 0$, og ser på tilfellende $\alpha \geq 0$ og $\alpha < 0$.

1. $\alpha \geq 0$: $\gamma + \alpha \geq 0$. Fordi $(\gamma + \alpha)$ og $(\gamma - \alpha)$ enten er positive eller negative, må $\gamma - \alpha \geq 0$.

2. $\alpha < 0$: $-\alpha \geq 0 \implies \gamma - \alpha \geq 0$. Fordi $(\gamma + \alpha)$ og $(\gamma - \alpha)$ enten er positive eller negative, må $\gamma + \alpha \geq 0$.

Leddene $(\gamma + \alpha)$ og $(\gamma - \alpha)$ vil alltid være positive.

Vi ser på hva som skjer dersom vi setter $w = \sqrt{\frac{\gamma+\alpha}{2}} + \sqrt{\frac{\gamma-\alpha}{2}}i$.

$$
\begin{aligned}
w^2 &= \left(\sqrt{\frac{\gamma+\alpha}{2}} + \sqrt{\frac{\gamma-\alpha}{2}}i\right)^2 \\
&= \left(\sqrt{\frac{\gamma+\alpha}{2}} + \sqrt{\frac{\gamma-\alpha}{2}}i\right)\left(\sqrt{\frac{\gamma+\alpha}{2}} + \sqrt{\frac{\gamma-\alpha}{2}}i\right) \\
&= \left(\sqrt{\frac{\gamma+\alpha}{2}}\right)^2 + 2\left(\sqrt{\frac{\gamma+\alpha}{2}}\right)\left(\sqrt{\frac{\gamma-\alpha}{2}}i\right) + \left(\sqrt{\frac{\gamma-\alpha}{2}}i\right)^2 \\
&= \frac{\gamma+\alpha}{2} + 2\left(\sqrt{\frac{(\gamma+\alpha)(\gamma-\alpha)}{2^2}}\right)i + \frac{\gamma-\alpha}{2}i^2 \\
&= \frac{\gamma+\alpha}{2} - \frac{\gamma-\alpha}{2} + 2\left(\sqrt{\frac{\beta^2}{2^2}}\right)i \\
&= 2\frac{\alpha}{2} + 2\frac{\beta}{2}i \\
&= \alpha + \beta i
\end{aligned}
$$

Siden $w^2 = \left(\sqrt{\frac{\gamma+\alpha}{2}} + \sqrt{\frac{\gamma-\alpha}{2}}i\right)^2 = \alpha + \beta i = z$, så har z en kvadratrot.

■

3.1 Hermitiske matriser

I denne seksjonen skal vi se på to resultater om hermitiske matriser. En matrise A er *hermitisk* når

$$A = \overline{A^T},$$

dvs. at matrisen ikke endres når den transponeres og kompleks konjugeres. Vi bruker notasjonen A^* for $\overline{A^\intercal}$. Det første resultatet vi skal se på er at de hermitiske $n \times n$-matrisene danner et vektorrom over $\mathbb{R}$.

Lemma 3.3. *De hermitiske $n \times n$-matrisene danner et $\mathbb{R}$-vektorrom,* $\mathrm{Herm}_n(\mathbb{C})$.

Bevis. Siden mengden av komplekse $n \times n$-matriser, $M_n(\mathbb{C})$, er et vektorrom over $\mathbb{C}$, er det følgelig også et vektorrom over $\mathbb{R}$, siden $\mathbb{C}$ er et vektorrom over $\mathbb{R}$. For at $\mathrm{Herm}_n(\mathbb{C})$ skal være et $\mathbb{R}$-vektorrom, er det nok å vise at $\mathrm{Herm}_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) | A = A^*\} \subseteq M_n(\mathbb{C})$. Det er tre krav som oppfylles for at en delmengde U skal være et underrom av $\mathbb{R}$,

1. $0 \in U$ (nullmatrisen er med i U)
2. Hvis $V, W \in U$, så er $V + W \in U$
3. Hvis $V \in U$ og $k \in \mathbb{R}$, så er $kV \in U$

Vi ser om delmengden $\mathrm{Herm}_n(\mathbb{C})$ oppfyller kravene. La $V, W \in \mathrm{Herm}_n(\mathbb{C})$ og $k \in \mathbb{R}$.

1. $0^* = \overline{0^\intercal} = 0$
2. $(V+W)^* = V^* + W^* = V + W$
3. $(kV)^* = kV^* = kV$

Matrisene er fortsatt med i delmengden $\mathrm{Herm}_n(\mathbb{C})$ og $\mathrm{Herm}_n(\mathbb{C})$ er et underrom av $M_n(\mathbb{C})$.

■

Det siste resultatet i denne seksjonen presenterer dimensjonen til $\mathrm{Herm}_n(\mathbb{C})$.

Lemma 3.4. *Dimensjonen til $\mathbb{R}$-vektorrommet $V = \mathrm{Herm}_n(\mathbb{C})$ er n^2.*

Bevis Vi begynner å se på $n = 1$. Vi har en 1×1-matrise

$$C = \begin{bmatrix} a \end{bmatrix}.$$

For at C skal ligge i V må den være hermitisk, altså $C = C^*$. Da vil

$$\begin{bmatrix} a \end{bmatrix} = \begin{bmatrix} \overline{a} \end{bmatrix},$$

som vil si $a \in \mathbb{R}$. Vektorrommet av alle 1×1 hermitiske matriser vil være utspent av en vektor $v = \begin{bmatrix} 1 \end{bmatrix}$. Når $n = 1$, vil $\dim V = 1$.

Vi ser videre på $n = 2$ og en 2×2-matrise

$$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

For at C skal ligge i vektorrommet V må $C = C^*$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{bmatrix}.$$

Da vil $a, d \in \mathbb{R}$ og $b \in \mathbb{C}$ for $c = \overline{b}$, og alle 2×2-matriser i V kan beskrives på følgende form:

$$\begin{bmatrix} a & b \\ \overline{b} & d \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \mathrm{Re}(b) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \mathrm{Im}(b) \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

og, siden de fire matrisene er lineært uavhengige utgjør de en basis. Dimensjonen til V blir 4 når $n = 2$. Vi ser at formelen for $\dim V = n^2$ stemmer for $n = 1, 2$.

Vi ser nå på en $k \times k$-matrise i $\mathrm{Herm}_k(\mathbb{C})$

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ c_{k1} & \cdots & \cdots & c_{kk} \end{bmatrix}$$

For at C skal ligge i vektorrommet V må $C = C^*$. Dvs.

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ c_{k1} & \cdots & \cdots & c_{kk} \end{bmatrix} = \begin{bmatrix} \overline{c_{11}} & \overline{c_{21}} & \cdots & \overline{c_{k1}} \\ \overline{c_{12}} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \overline{c_{1k}} & \cdots & \cdots & \overline{c_{kk}} \end{bmatrix}$$

Alle c_{ii} vil være lik sin kompleks konjurgerte og er derfor reelle. For $c_{ji} = \overline{c_{ij}}$ kan alle $k \times k$-matriser i V kan beskrives på følgende måte

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ \overline{c_{12}} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \overline{c_{1k}} & \cdots & \cdots & c_{kk} \end{bmatrix}$$

Siden diagonalen består av reelle tall så blir diagonalen til de hermetiske matrisene spent ut av k matriser

$$\begin{bmatrix} 1 & & & & \\ & 0 & & & \\ & & \cdots & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & & \\ & 1 & & & \\ & & \cdots & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix}, \cdots, \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & \cdots & & \\ & & & 1 & \\ & & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & \cdots & & \\ & & & 0 & \\ & & & & 1 \end{bmatrix}.$$

For alle c_{ij} på plass (i, j), hvor $i \neq j$, så finnes $\overline{c_{ij}}$ på (j, i). Matrisen kan derfor utspennes ved å se på elementene over diagonalen, hvor hvert element utspennes derfor av to matriser, den som har 1 på plass (i, j) og (j, i), hvor $i \neq j$, og den som har i på plass (i, j) og $-i$ på (j, i), hvor $i \neq j$. Siden det finnes to matriser for hvert element over diagonalen, vil antallet matriser som spenner ut elementene over diagonalen være

$$\sum_{i=1}^{k-1} 2k = +2\frac{k(k-1)}{2} = k^2 - k.$$

Og, siden matrisene som spenner ut diagonalen og matrisene som spenner ut alle de andre elementene er lineært uavhengig, så utgjør de en basis for $\mathrm{Herm}_n(\mathbb{C})$. Det vil si at $\dim V$ kan beskrives på følgende måte

$$k + k^2 - k = k^2$$

Siden $\dim V = k^2$ for $\mathrm{Herm}_k(\mathbb{C})$, kan konkludere med at dimensjonen til et vektorrommet $V = \mathrm{Herm}_n(\mathbb{C})$ er n^2.

■

3.2 Skjev-symmetriske matriser

Vi skal også se på $n \times n$ skjev-symmestriske matriser senere i beviset. En matrise A er *skjev-symmetrisk* når

$$A = -A^\intercal,$$

dvs. at matrisen ikke endres når den skifter fortegn og transponeres. Vi bruker notasjonen $\mathrm{Skew}_n(\mathbb{C})$ for alle skjev-symmetriske matriser over $\mathbb{C}$. Vi skal se på to resultater om skjev-symmertiske matriser, hvor det første er at de $\mathrm{Skew}_n(\mathbb{C})$ danner et vektorrom over $\mathbb{C}$.

Lemma 3.5. *De skjev-symmertiske $n \times n$-matrisene danner et $\mathbb{C}$-vektorrom,* $\mathrm{Skew}_n(\mathbb{C})$.

Bevis. Siden mengden av komplekse $n \times n$-matriser, $M_n(\mathbb{C})$, er et vektorrom over $\mathbb{C}$ er det nok å vise at $\mathrm{Skew}_n(\mathbb{C}) = \{M_n(\mathbb{C}) \mid A = -A^\intercal\}$ er et ekte underrom. Det er tre krav som må oppfylles for at en delmengde U skal være et underrom av $\mathbb{C}$,

1. $0 \in U$ (nullmatrisen er med i U)
2. Hvis $V, W \in U$, så er $V + W \in U$
3. Hvis $V \in U$ og $k \in \mathbb{C}$, så er $kV \in U$

Vi ser om delmengden Skew($\mathbb{C}$) oppfyller kravene. La $V, W \in$ Skew($\mathbb{C}$) og $k \in \mathbb{C}$.

1. $-0^\intercal = 0$
2. $-(V+W)^\intercal = -(V^\intercal + W^\intercal) = -V^\intercal - W^\intercal = V + W$
3. $-(kV)^\intercal = -kV^\intercal = kV$

Matrisene er fortsatt med i delmengden $\mathrm{Skew}_n(\mathbb{C})$ og $\mathrm{Skew}_n(\mathbb{C})$ er et underrom av $M_n(\mathbb{C})$.

■

Det andre resultatet vi skal se på er dimensjonen til $\mathrm{Skew}_n(\mathbb{C})$.

Lemma 3.6. *Dimensjonen til det $\mathbb{C}$-vektorrommet $V = \mathrm{Skew}_n(\mathbb{C})$ er $\frac{n(n-1)}{2}$.*

Bevis. Vi begynner å se på $n = 1$. Vi har en 1×1-matrise,

$$C = \begin{bmatrix} a \end{bmatrix}.$$

For at C skal ligge i vektorrommet V må $C = -C^\intercal$. Dvs.

$$\begin{bmatrix} a \end{bmatrix} = \begin{bmatrix} -a \end{bmatrix}$$

og, siden ingen tall er lik den negative av seg selv, må C være null-matrisen. Dimensjonen til V blir da 0.

For $n = 2$ har vi en 2×2-matrise

$$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

For at C skal ligge i vektorrommet V må $C = -C^\intercal$. Dvs.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -a & -c \\ -b & -d \end{bmatrix}.$$

Akkurat som for $n = 1$ må $a = d = 0$. For $b = -c$, vil alle 2×2-matriser i V uttrykkes på følgende form:

$$\begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} = b \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Matrisen utgjør en basis for V når $n = 2$, og $\dim V = 1$. Vi ser at formelen for $\dim V = \frac{n(n-1)}{2}$ stemmer for $n = 1, 2$.

For $n = k$ har vi en $k \times k$-matrise

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1k} \\ c_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ c_{k1} & \cdots & \cdots & c_{kk} \end{bmatrix}.$$

For at C skal ligge i vektorrommet V må $C = -C^{\intercal}$. Dvs.

$$\begin{bmatrix} c_{11} & c_{12} & \dots & c_{1k} \\ c_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ c_{k1} & \cdots & \cdots & c_{kk} \end{bmatrix} = \begin{bmatrix} -c_{11} & -c_{12} & \dots & -c_{1k} \\ -c_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ -c_{k1} & \cdots & \cdots & -c_{kk} \end{bmatrix}$$

For skjev-symmetriske matriser er $c_{ii} = 0$, og $c_{ij} = -c_{ji}$. Alle $k \times k$-matriser i V kan beskrives på følgende måte:

$$\begin{bmatrix} 0 & c_{12} & \dots & c_{1k} \\ -c_{12} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ -c_{1k} & \cdots & \cdots & 0 \end{bmatrix}$$

En $n \times n$-matise vil ha n^2 elementer. Siden diagonalen til skjevsymmetriske matriser er null vil matrisen ha $n^2 - n$ elementer. Men, siden $c_{ji} = -c_{ij}$, vil C bare ha halvparten så mange frie elementer. Matrisen C kan derfor utspennes av matrisene som har null på alle plasser unntatt 1 på plass (i, j) og -1 på plass (j, i) for $i < j$. Det er $\frac{n(n-1)}{2}$ slike matriser. Et direkte argument gir at disse matrisene er lineært uavhengige. Dette gir

$$\dim V = \frac{n(n-1)}{2}$$

■

3.3 Karakteristisk polynom av matriser

Til sist, før vi ser på algebraens fundamentalteorem, skal vi se på karaterisktiske polynomer. Det karaterisktiske polynomet $P(x)$ til en $n \times n$-matrise M er definert ved $P(x) = \det(xI - M)$.

Lemma 3.7. *Polynomet $P_n(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ er det karakteristiske polynomet til matrisen M_n for $n \geq 2$, hvor*

$$M_n = \begin{bmatrix} 0 & 0 & 0 & \dots & -a_n \\ 1 & 0 & 0 & \dots & -a_{n-1} \\ 0 & 1 & 0 & \dots & -a_{n-2} \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & -a_1 \end{bmatrix}$$

Bevis. Vi beviser dette ved induksjon.

Vi sjekker om induksjonsantagelsen stemmer. For $n = 2$ vil

$$P_2(x) = x^2 + a_1x + a_2$$

og

$$M_2 = \begin{bmatrix} 0 & -a_2 \\ 1 & -a_1 \end{bmatrix}$$

Vi finner

$$\det(xI_2 - M_2) = \begin{vmatrix} x & a_2 \\ -1 & x + a_1 \end{vmatrix} = x(x + a_1) - (-1)a_0 = x^2 + a_1x + a_0 = P_2(x),$$

og induksjonsantagelsen stemmer.

Vi antar nå at påstanden stemmer for $n - 1$. Altså så vil det karakteristiske polynomet til matrisen

$$M_{n-1} = \begin{bmatrix} 0 & 0 & 0 & \dots & -a_{n-1} \\ 1 & 0 & 0 & \dots & -a_{n-2} \\ 0 & 1 & 0 & \dots & -a_{n-3} \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & -a_1 \end{bmatrix}$$

være

$$P_{n-1}(x) = x^{n-1} + x^{n-2}a_1 + \cdots + a_{n-2}x + a_{n-1}.$$

Det betyr at

$$P_{n-1}(x) = \det(xI_{n-1} - M_{n-1}).$$

Vi sjekker for n, hvor

$$P_n(x) = x^n + x^{n-1}a_1 + \cdots + xa_{n-1} + a_n$$

og

$$M_n = \begin{bmatrix} 0 & 0 & 0 & \dots & -a_n \\ 1 & 0 & 0 & \dots & -a_{n-1} \\ 0 & 1 & 0 & \dots & -a_{n-2} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & -a_2 \\ 0 & 0 & \dots & 1 & -a_1 \end{bmatrix}$$

Vi noterer oss at når vi regner ut determinanten vil fortegnet til hvert av leddene i utregningen alternere,

$$\begin{vmatrix} \overbrace{11}^{+} & \overbrace{12}^{-} & \overbrace{13}^{+} & \cdots & 1n \\ 21 & 22 & \cdots & & 2n \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ n1 & n2 & \cdots & & nn \end{vmatrix}$$

og det $n-te$ elementet vil ha fortegn $(-1)^{n+1}$.

Vi finner

$$\det(xI_n - M_n) = \begin{vmatrix} x & 0 & 0 & \dots & a_n \\ -1 & x & 0 & \dots & a_{n-1} \\ 0 & -1 & x & \dots & a_{n-2} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & -1 & x & a_2 \\ 0 & 0 & \dots & -1 & x+a_1 \end{vmatrix}$$

ved hjelp av utvikling langs første rad.

$$\det(xI_n - M_n) = x \underbrace{\begin{vmatrix} -1 & x & 0 & \dots & a_{n-1} \\ 0 & -1 & x & \dots & a_{n-2} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & -1 & x & a_2 \\ 0 & 0 & \dots & -1 & x+a_1 \end{vmatrix}}_{(1)} + (-1)^{n+1}a_n \underbrace{\begin{vmatrix} -1 & x & \dots & 0 \\ \vdots & \vdots & \vdots & 0 \\ \dots & 0 & -1 & x \\ 0 & \dots & 0 & -1 \end{vmatrix}}_{(2)}$$

Determinanten til (1) vil fra induksjonsantagelsen være lik $P_{n-1} = x^{n-1} + x^{n-2}a_1 + \cdots + a_{n-2}x + a_{n-1}$.

Siden (2) er en øvre triangulær matrise vet vi fra MA1201 at determinanten er produktet av alle elementene langs diagonalen. Den blir da $(-1)^{n-1}$ fordi den er en $n-1 \times n-1$-matrise ettersom vi fjernet en rad og en kolonne fra en $n \times n$-matrise.

Da får vi at

$$\begin{aligned}
\det(xI_n - M_n) &= x \cdot P_{n-1} + (-1)^{n+1} a_n (-1)^{n-1} \\
&= x(x^{n-1} + x^{n-2}a_1 + \cdots + a_{n-2}x + a_{n-1}) + (-1)^{2n} a_n \\
&= x^n + x^{n-1}a_1 + \cdots + a_{n-2}x^2 + a_{n-1}x + a_n = P_n(x)
\end{aligned}$$

og resultatet stemmer ved induksjon.

$\blacksquare$

4 Beviset for algebraens fundamentalteorem

Ved hjelp av de preliminære resultatene i kapittel 3 og den lineære algebraen som er undervist i MA1201, har vi alle redskapene vi trenger for å bevise algebraens fundamentalteorem. Beviset er som nevnt utledet av Harm Derksen, og går ut på å bevise at den tilhørende matrisen til et karakteristisk polynom har en egenvektor. Dersom det finnes en egenvektor, finnes det en tilhørende egenverdi, og det karakteristiske polynomet har et nullpunkt. For å vise at den tilhørende matrisen har en egenvektor, skal vi se på K-vektorrom, hvor K enten er $\mathbb{R}$ eller $\mathbb{C}$, med kommuterende matriser og vise at de har en felles egenvektor. La oss begynne.

La K være enten $\mathbb{R}$ eller $\mathbb{C}$ og la d og r være positive heltall. Se på det følgende utsagnet.

$P(K, d, r)$: Anta så at $A_1, A_2, ..., A_r$ er $n \times n$-matriser over K som kommuterer med hverandre og at d ikke deler n. Da har $A_1, A_2, ..., A_r$ en felles egenvektor.

Merk at vi må ha $d > 1$ for at d ikke deler n.

Det første lemmaet i beviset går ut på å bevise at dersom det stemmer for grunntilfellet $r = 1$, så stemmer det for alle $r \geq 1$.

Lemma 4.1. *Dersom $P(K, d, 1)$ stemmer, så vil $P(K, d, r)$ stemme for alle $r \geq 1$.*

Bevis. Vi beviser dette ved induksjon på r.

Anta at $P(K, d, r-1)$ stemmer. Anta videre at $A_1, A_2, ..., A_r$ er $n \times n$-matriser som kommuterer med hverandre og at d ikke deler n. Ved å ta induksjon på n kan vi bevise at $A_1, A_2, ..., A_r$ har en felles egenvektor.

Vi ser på $n = 1$. Siden vi har at $d > 1$, så er kravet om at $d \nmid n$ oppfylt og vi kan sjekke induksjonsgrunnlaget.

Vi vil vise at 1×1-matrisene $A_1, A_2, \ldots, A_r$ har en felles egenvektor v. For at en matrise $A = \begin{bmatrix} a \end{bmatrix}$ skal ha en slik egenvektor må den ha en egenverdi λ s.a. $Av = \lambda v$. Vi ser på determinanten til A for å finne egenverdien λ.

$$\det(A - \lambda I) = 0$$
$$\Longleftrightarrow$$
$$a - \lambda = 0 \Longrightarrow a = \lambda$$

Siden A har egenverdi a, vil egenvektoren v være

$$\begin{bmatrix} a \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = a \begin{bmatrix} 1 \end{bmatrix},$$

og dette vil være tilfellet for alle 1×1-matriser og det eksister en felles egenvektor, v, for $A_1, A_2, \ldots, A_r$. Induksjonsgrunnlaget stemmer og vi kan anta at matrisene $A_1, A_2, ..., A_r$ har en felles egenvektor for $n = 1$. Anta at $P(K, d, r)$ stemmer for alle $m \times m$-matriser, hvor $m < n$. Siden $P(K, d, 1)$ stemmer, har A_r en egenverdi λ i K.

Vi har lineær transformasjonen

$$T \colon K^n \xrightarrow{A_r - \lambda I} K^n$$

Vi setter $W = \text{Ker}(A_r - \lambda I)$ og $Z = \text{im}(A_r - \lambda I)$.

Vi vil sjekke om W og Z er stabile under $A_1, A_2, \ldots, A_{r-1}$, dvs. at $A_i(W) \subseteq W$ og at $A_i(Z) \subseteq Z$ for $i \in \{1, 2, \ldots, r-1\}$. Vi sjekker først W.

For en $w \in \text{Ker}(A_r - \lambda I)$ så vil

$$(A_r - \lambda I)w = 0 \tag{1}$$

For at W skal være stabil så må $(A_r - \lambda I)A_i(w) = 0$ for $i \in \{1, 2, \ldots, r-1\}$ og $w \in W$. Vi har at:

$$\begin{aligned}
(A_r - \lambda I)A_i(w) =& ((A_r - \lambda I)A_i)w \\
=& (A_r A_i - \lambda I A_i)w \\
=& (A_i A_r - \lambda A_i I)w \\
=& A_i \underbrace{((A_r - \lambda I)w)}_{=0 \text{ fra likning 1}} \\
=& 0
\end{aligned}$$

Siden $(A_r - \lambda I)A_i(w) = 0$, så medfører det at $w \in \text{Ker}(A_r - \lambda I) \implies A_i(\text{Ker}(A_r - \lambda I)) \subseteq \text{Ker}(A_r - \lambda I)$ og W er stabil for alle A_i hvor $i \in \{1, 2, \ldots, r-1\}$.

Vi sjekker nå om Z er stabil. For at bildet skal være stabilt må $A_i(z) \in Z$ for alle $z \in Z$ og for $i \in \{1, 2, \ldots, r-1\}$. Vi har at $z = (A_r - \lambda I)v$ for en $v \in K^n$. Dette gir:

$$\begin{aligned}
A_i(z) =& A_i(A_r - \lambda I)v \\
=& (A_i A_r - \lambda A_i I)v \\
=& (A_r A_i - \lambda I A_i)v \\
=& (A_r - \lambda I)A_i(v) \in Z
\end{aligned}$$

Siden $A_i(z) \in Z$, er Z stabil for alle A_i, hvor $i \in \{1, 2, \ldots, r-1\}$.

Da vet vi at både V og Z er stabile under A_i, hvor $i = \{1, 2, \ldots, r-1\}$. Fra induksjonsantagelsen har vi at $P(K, d, r-1)$ stemmer, som betyr at matrisene $A_1, A_2, \ldots, A_{r-1}$ en felles egenvektor. Til slutt i dette beviset skal vi se på de to tilfellende når $\dim Z \geq 1$ og $\dim Z = 0$.

Vi ser først på $\dim Z \geq 1$. Nullitet + rank teoremet sier at $\dim W + \dim Z = n$. Både $\dim Z < n$ og $\dim W < n$ fordi $\dim W + \dim Z = n$. Siden vi allerede har antatt at $P(K, d, r)$ stemmer for alle $m \times m$-matriser, hvor $m < n$, og

$$d \nmid n \implies d \nmid \dim W \vee d \nmid \dim Z.$$

kan vi med induksjon på n anta at $A_1, \ldots, A_r$ allerede har en felles egenvektor i W eller i Z.

Det gjenværende tilfellet er når $\dim Z = 0$, da vil $\dim W = n$. Siden $P(K, d, r-1)$ stemmer fra induksjonsantagelsen vår kan vi anta at $A_1, \ldots, A_{r-1}$ har en felles egenvektor, v. Siden W er egenrommet til A_r for egenverdien λ, så er $A_r v = \lambda v$. Dette gir at v er en felles egenvektor til $A_1, \ldots, A_r$.

■

Nå som vi vet at det holder å bevise påstanden for grunntilfellet $r = 1$, ønsker vi å se på kommutative matriser av et odde dimensjons $\mathbb{R}$-vektorrom, og sjekke om påstanden gjelder her.

Lemma 4.2. *$P(\mathbb{R}, 2, r)$ holder for alle r, m.a.o dersom $A_1, A_2, \ldots, A_r$ er kommutative matriser av et odde dimensjons $\mathbb{R}$-vektorrom, så har de en felles egenvektor.*

Bevis. Lemma 4.1 forteller oss at det er nok å bevise at dette stemmer for $P(\mathbb{R}, 2, 1)$. Dvs. at vi må vise at en matrise A som ligger i et odde dimensjons $\mathbb{R}$-vektorrom har en egenvektor.

Vi ser på den karakteristiske likningen til A og sjekker om det finnes egenverdier ved å ta determinanten $\det(xI - A)$. Siden dimensjonen n er odde, vil det karakteristiske polynomet $P(x)$ til A være av odde grad og vi vet fra Lemma 3.1 at det eksisterer en λ s.a. $P(\lambda) = 0$. Da blir λ en ekte egenverdi til A, og det finnes en egenvektor v s.a.

$$\begin{gathered}(A - \lambda I)v = 0 \\ \iff \\ Av = \lambda v.\end{gathered}$$

■

I det neste lemmaet ønsker vi å se om påstanden holder for alle odde dimensjons $\mathbb{C}$-vektorrom.

Lemma 4.3. *$P(\mathbb{C}, 2, 1)$ holder. Dvs. at alle $n \times n$-matriser over $\mathbb{C}$ med odde n har en egenvektor.*

Bevis. I dette beviset skal vi se på $\mathbb{R}$-vektorrommet V som består av de hermitiske $n \times n$-matrisene $\mathrm{Herm}_n(\mathbb{C})$. Matrisene er beskrevet i avsnitt 3.1.

Anta at A er en $n \times n$-matrise over $\mathbb{C}$ hvor n er odde. La V være $\mathbb{R}$-vektorrommet $\mathrm{Herm}_n(\mathbb{C})$. Vi ønsker å definerer to kommutative linære avbildninger $L_1, L_2 \colon V \to V$ med

$$L_1(B) = \frac{AB + BA^*}{2}$$

og

$$L_2(B) = \frac{AB - BA^*}{2i}.$$

For at L_1 og L_2 skal være kommutative må vi sjekke at $L_1(L_2) = L_2(L_1)$. Da må vi først verifisere at både $L_1(B)$ og $L_2(B)$ ligger i V for alle $B \in V$.

For at $L_1(B)$ skal ligge i V så må $\overline{L_1(B)}^{\intercal} = L_1(B)$ for en B i V.

$$\begin{aligned}
\overline{L_1(B)}^{\intercal} &= \overline{\left(\frac{AB+BA^*}{2}\right)}^{\intercal} \\
&= \overline{\left(\frac{1}{2}AB+\frac{1}{2}BA^*\right)}^{\intercal} \\
&= \overline{\left(\frac{1}{2}AB\right)^{\intercal}+\left(\frac{1}{2}BA^*\right)^{\intercal}} \\
&= \overline{\left(\frac{1}{2}AB\right)}^{\intercal}+\overline{\left(\frac{1}{2}BA^*\right)}^{\intercal} \\
&= \frac{1}{2}(AB)^*+\frac{1}{2}(BA^*)^* \\
&= \frac{1}{2}B^*A^*+\frac{1}{2}AB^* && \text{(fordi } B \text{ ligger i } V \text{ er } B=B^*) \\
&= \frac{1}{2}BA^*+\frac{1}{2}AB \\
&= \frac{1}{2}(AB+BA^*) \\
&= L_1(B)
\end{aligned}$$

Den lineære avbildingen $L_1(B)$ er hermitisk og vil derfor ligge i V.

For at $L_2(B)$ skal ligge i V så må $\overline{L_2(B)}^{\intercal} = L_2(B)$ for en B i V.

$$\begin{aligned}
\overline{L_2(B)}^{\intercal} &= \overline{\left(\frac{AB-BA^*}{2i}\right)}^{\intercal} \\
&= \overline{\left(\frac{1}{2i}AB-\frac{1}{2i}BA^*\right)}^{\intercal} \\
&= \overline{\left(\frac{1}{2i}AB\right)^{\intercal}-\left(\frac{1}{2i}BA^*\right)^{\intercal}} \\
&= \overline{\left(\frac{1}{2i}AB\right)}^{\intercal}-\overline{\left(\frac{1}{2i}BA^*\right)}^{\intercal} \\
&= -\frac{1}{2i}(AB)^*+\frac{1}{2i}(BA^*)^* \\
&= -\frac{1}{2i}B^*A^*+\frac{1}{2i}AB^* \\
&= -\frac{1}{2i}BA^*+\frac{1}{2i}AB && \text{(fordi } B \text{ ligger i } V \text{ er } B=B^*) \\
&= \frac{1}{2i}(AB-BA^*) \\
&= L_2(B)
\end{aligned}$$

Den linære avbildingen $L_2(B)$ er hermitisk og vil derfor ligge i V.

Vi har sjekket at $L_1, L_2 \colon V \to V$, og vil nå sjekke om de kommuterer. Dvs. at $L_1(L_2(B)) = L_2(L_1(B))$. Vi ser først på $L_1(L_2(B))$.

$$\begin{aligned}
L_1(L_2(B)) &= \frac{AL_2(B) + L_2(B)A^*}{2} \\
&= \frac{1}{2i}\left(\frac{A(AB - BA^*) + (AB - BA^*)A^*}{2}\right) \\
&= \frac{1}{2i}\left(\frac{A^2B - ABA^* + ABA^* - BA^{*2}}{2}\right) \\
&= \frac{A^2B - BA^{*2}}{4i}
\end{aligned}$$

Vi ser så på $L_2(L_1(B))$.

$$\begin{aligned}
L_2(L_1(B)) &= \frac{AL_1(B) - L_1(B)A^*}{2i} \\
&= \frac{1}{2}\left(\frac{A(AB + BA^*) - (AB + BA^*)A^*}{2i}\right) \\
&= \frac{1}{2}\left(\frac{A^2B + ABA^* - ABA^* - BA^{*2}}{2i}\right) \\
&= \frac{A^2B - BA^{*2}}{4i}
\end{aligned}$$

Siden $L_1L_2(B) = L_2L_1(B)$ for alle $B \in V$, så kommuterer de linære avbildingene L_1 og L_2.

Noter. *Fra Lemma 3.4 vet vi at* $\dim V = n^2$ *for vektorrommet* $V = \mathrm{Herm}_n(\mathbb{C})$*. Vi har antatt at* n *er odde, som sikrer at* $\dim V$ *er odde, og at* n^2 *ikke kan deles på 2.*

Påstanden $P(\mathbb{R}, 2, 2)$ fra Lemma 4.2 impliserer at L_1 og L_2 har en felles egenvektor B, si $L_1(B) = \lambda B$ og $L_2(B) = \mu B$ med $\lambda, \mu \in \mathbb{R}$. Da er

$$(L_1 + iL_2)(B) = \left(\tfrac{AB+BA^*}{2}\right) + i\left(\tfrac{AB-BA^*}{2i}\right) = AB = (\lambda + \mu i)(B),$$

og hvilken som helst ikke-null kolonnevektor av B vil gi en egenvektor for matrisen A med egenverdi $(\lambda + \mu i)$.

■

Videre så sjekker vi at påstanden også holder for alle 2^k-faktorer og alle r.

Lemma 4.4. *$P(\mathbb{C}, 2^k, r)$ holder for alle k og r.*

Bevis. I dette beviset ser vi på $\mathbb{C}$-vektorrommet $\mathrm{Skew}_n(\mathbb{C})$, som er mengden av skjevsymmetriske $n \times n$-matriser som beskrevet i avsnitt 3.2.

Vi beviser Lemma 4.4 med induksjon på k.

Vi sjekker at grunntilfellet $k = 1$ stemmer. Fra Lemma 4.3 vet vi at $P(\mathbb{C}, 2, 1)$ holder. Lemma 4.1 forteller at dersom $P(\mathbb{C}, 2, 1)$ holder, så holder $P(\mathbb{C}, 2, r)$ og induksjonsantagelsen stemmer.

Videre kan vi anta at $P(\mathbb{C}, 2^l, r)$ holder for $1 < l < k$. Vi vil sjekke at $P(\mathbb{C}, 2^k, r)$ stemmer, og fra Lemma 4.1 vet vi at det holder å bevise for $P(\mathbb{C}, 2^k, 1)$.

Anta at A er en $n \times n$-matrise over $\mathbb{C}$, hvor n er delelig på 2^{k-1}, men ikke på 2^k. La V være det $\mathbb{C}$-vektorrommet $\mathrm{Skew}_n(\mathbb{C})$. Vi ønsker å definerer $L_1, L_2 \colon V \to V$ med

$$L_1(B) = AB + BA^\intercal$$

og

$$L_2(B) = ABA^\intercal.$$

For at L_1 og L_2 skal være kommutative må vi sjekke at $L_1(L_2) = L_2(L_1)$. Da må vi først verifisere at både $L_1(B)$ og $L_2(B)$ ligger i V for alle $B \in V$.

For at $L_1(B)$ skal være i V må $-L_1(B)^\intercal = L_1(B)$. Vi har at:

$$\begin{aligned}
-L_1(B)^\intercal &= -(AB + BA^\intercal)^\intercal \\
&= -(AB)^\intercal - (BA^\intercal)^\intercal \\
&= -B^\intercal A^\intercal - AB^\intercal && \text{(Fordi } B \text{ ligger i } V \text{ så vil } B^\intercal = -B) \\
&= -(-B)A^\intercal - A(-B) \\
&= BA^\intercal + AB \\
&= L_1(B).
\end{aligned}$$

Matrisen $L_1(B)$ er skjevsymmetrisk og vil ligge i V for B i V.

Vi sjekker om $L_2(B)$ er i V på samme måte. Vi har at:

$$\begin{aligned}
-L_2(B)^\intercal &= -(ABA^\intercal)^\intercal \\
&= -AB^\intercal A^\intercal && \text{(Fordi } B \text{ ligger i } V \text{ så vil } B^\intercal = -B) \\
&= -A(-B)A^\intercal \\
&= ABA^\intercal \\
&= L_2(B)
\end{aligned}$$

Matrisen $L_2(B)$ er skjevsymmetrisk og vil ligge i V for B i V.

Nå har vi verifisert at $L_1, L_2 \colon V \to V$. Nå vil vi sjekke om de er kommuterer. Vi begynner med å sjekke $L_1(L_2(B))$.

$$\begin{aligned} L_1(L_2(B)) &= AL_2(B) + L_2(B)A^\intercal \\ &= A(ABA^\intercal) + (ABA^\intercal)A^\intercal \\ &= A^2BA^\intercal + ABA^{\intercal 2} \end{aligned}$$

Vi sjekker $L_2(L_1(B))$.

$$\begin{aligned} L_2(L_1(B)) &= AL_1(B)A^\intercal \\ &= A(AB + BA^\intercal)A^\intercal \\ &= A^2BA^\intercal + ABA^{\intercal 2} \end{aligned}$$

De lineære avbildningene L_1 og L_2 kommuterer.

Fra Lemma 3.6 vet vi at $\dim V = \frac{n(n-1)}{2}$, og vi må sjekke at 2^{k-1} ikke deler $\dim V$. Vi begynner med å anta at $2^{k-1} \mid \dim V$. Da vil

$$\begin{gathered} 2^{k-1}m_1 = \dim V = \tfrac{n(n-1)}{2} \\ \Longleftrightarrow \\ 2^k m_1 = n(n-1) \end{gathered}$$

Vi vet at $2^{k-1} \mid n$, dvs. at $n = 2^{k-1}m_2$, hvor m_2 må være odde fordi om $m_2 = 2m_3$ ville $n = 2^{k-1}2m_3 = 2^k m_3$ og 2^k ville delt n, noe vi vet fra antagelsen vår at den ikke gjør. Siden $n = 2^{k-1}m_2$ vil n være et partall, og $n-1$ vil være et oddetall. Da vil

$$\begin{gathered} 2^k m_1 = 2^{k-1}m_2(n-1) \\ \Longleftrightarrow \\ 2m_1 = m_2(n-1) \end{gathered}$$

være en selvmotsigelse, siden du får to oddetall ganget sammen blir et partall, og vi har verifisert at $2^{k-1} \nmid \dim V$.

Fra $P(\mathbb{C}, 2^{k-1}, 2)$ har L_1 og L_2 en felles egenvektor B, si $L_1(B) = \lambda B$ og $L_2(B) = \mu B$. Da følger det at

$$\lambda B = AB + BA^\intercal \iff BA^\intercal = AB - \lambda B \tag{2}$$

og dermed er

$$\mu B = ABA^\intercal = A\underbrace{(AB - \lambda B)}_{\text{fra Likning 2}}. \tag{3}$$

Skriver vi om på Likning 3, får vi

$$(A^2 - \lambda A - \mu I)B = 0 \tag{4}$$

La v være en ikke-null kolonne av B. Da får vi

$$(A^2 - \lambda A - \mu I)v = 0 \tag{5}$$

Fra Lemma 3.2 vet vi at alle andregrads polynomer har røtter og vi vet at vi kan uttrykke $x^2 - \lambda x - \mu = (x-\alpha)(x-\beta)$. Vi har da at:

$$(A - \alpha I)w = 0 \tag{6}$$

hvor $w = (A - \beta I)v$. Dersom $w = 0$, så er v en egenvektor til A med egenverdi β. Dersom $w \neq 0$, så vil w være en egenvektor til A med egenverdi α. Siden A har en egenvektor følger påstanden og $P(\mathbb{C}, 2^k, r)$ stemmer.

■

I det siste lemmaet før vi ser på algebraens fundamentalteorem vil vi se at kommuterende $n \times n$-matriser over $\mathbb{C}$ har en felles egenvektor.

Teorem 4.1. *Dersom $A_1, A_2, \ldots, A_r$ er kommuterende $n \times n$-matriser over $\mathbb{C}$, så har de en felles egenvektor.*

Bevis. Vi vil finne et positivt heltall k s.a. 2^k ikke deler n. Dersom $2^k > n$, så vil ikke 2^k dele n.

$$\begin{gathered} 2^k > n \\ \Longleftrightarrow \\ \log(2^k) > \log(n) \\ \Longleftrightarrow \\ k \cdot \log(2) > \log(n) \\ \Longleftrightarrow \\ k > \tfrac{\log(n)}{\log(2)} \end{gathered}$$

For alle $k > \frac{\log(n)}{\log(2)}$ så vil 2^k ikke dele n. Vi vet fra Lemma 4.4 at $P(\mathbb{C}, 2^k, r)$ holder, så påstanden følger.

■

Da har vi kommet til det siste resultatet, som er et algebraens fundamentalteorem. Resultatet blir presenter i et Korollar som følger av Teorem 4.1.

Korollar 4.1 (Algebraens fundamental teorem). *Dersom $P(x)$ er et ikke-konstant polynom med komplekse koffesienter, eksisterer det en λ i $\mathbb{C}$ s.a. $P(\lambda) = 0$.*

Bevis. Ved samme argumentasjon som i Lemma 3.1, så holder det å bevise dette for et monisk polynom. La

$$P_n(x) = x^n + x^{n-1}a_1 + \ldots + a_{n-1}x + a_n$$

Fra Lemma 3.7 vet vi at polynomet $P_n(x)$ er det karakteristiske polynomet til matrisen M_n, hvor

$$M_n = \begin{bmatrix} 0 & 0 & 0 & \dots & -a_n \\ 1 & 0 & 0 & \dots & -a_{n-1} \\ 0 & 1 & 0 & \dots & -a_{n-2} \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & -a_1 \end{bmatrix}$$

Fra Lemma 4.1 vet vi at M_n har en egenvektor, som vil si at det den har en egenverdi $\lambda \in \mathbb{C}$ s.a $P(\lambda) = 0$.

$\blacksquare$

5 Alternative bevis og konsekvenser

Det finnes flere måter å bevise algebraens fundamentalteorem på. Nå vil jeg presentere et par av disse bevisene, samt at jeg vil se på en direkte konsekvens av teoremet.

5.1 Liouvilles teorem og algebraens fundamentalteorem - Bevis 2

For det andre beviset av algebraens fundamentalteorem skal vi bruke Liouvilles teorem som sier følgende:

Teorem 5.1. *Anta at $f(x)\colon \mathbb{C} \to \mathbb{C}$ er hel og at $\big|f(x)\big|$ er begrenset for alle verdier $x \in \mathbb{C}$. Da er f(x) en konstant.*

Det at en funksjon $f(x)$ er hel, vil si at den er definert i hele det komplekse planet. En funksjon $f(x)$ vil være begrenset dersom hele settet av verdiene er begrenset. Det vil si at det eksisteren en M slik at $\big|f(x)\big| \leq M$. I denne oppgaven beviser jeg ikke Liouvilles teorem, men beviset kan leses på i boken The fundamental theorem of algebra (side 70).

Korollar 5.1. *Korollar 4.1 følger av Teorem 5.1*

Bevis. Anta at $P(x)$ er et komplekst polynom, og la $f(x) = \frac{1}{P(x)}$. Dersom $P(x)$ ikke har en kompleks rot, vil $f(x)$ være en hel funksjon.

Siden $\big|P(x)\big| \to \infty$ når $|x| \to \infty$, finnes det $M, r > 0$ s.a. $\big|P(x)\big| > M$ for $|x| > r$. Dermed følger det at for $|x| > r$, så er $f(x) = \frac{1}{P(x)} < \frac{1}{M}$. Dersom $f(x)$ var hel, måtte den vært kontinuerlig for å være definert i hele $\mathbb{C}$ og ville dermed være begrenset til den kompakte mengden $|x| \leq r$. Det følger også at dersom $f(x)$ var hel, ville den vært begrenset i hele $\mathbb{C}$. Siden $f(x)$ da er hel og begrenset, vil det fra Teorem 5.1 følge at $f(x)$ er en konstant. Da vil også $P(x)$ være en konstant, noe som er en motsigelse. Derfor, kan ikke $f(x)$ være hel, og $P(x)$ må være null for minst en $x \in \mathbb{C}$.

■

5.2 Avansert kalkulus og algebraens fundamentalteorem - Bevis 3

Det siste beviset i denne oppgaven er utelukkende basert på avansert kalkulus. For å kunne bevise fundamentalteoremet trenger vi det neste resultatet.

Lemma 5.1. *Om $f\colon D \to \mathbb{R}$ er kontinuerlig der D er en lukket og avgrenset (kompakt) delmengde av $\mathbb{R}^2$, da har $f(x, y)$ en minimums- og maksimumsverdi på D.*

Dette er den todimensjonale versjonen av ekstremalverdisetningen fra elementær kalkulus, som sier at om $f\colon [a, b] \to \mathbb{R}$ er kontinuerlig, da har $f(x)$ et minimum og et maximum på $[a, b]$. Mer generelt gjelder dette teoremet for kontinuerlige funksjoner $f\colon \mathbb{R}^n \to \mathbb{R}^n$ på kompakte definisjonsområder for alle $n \geq 1$. Vi kommer ikke til å føre bevis for Lemma 5.1 i denne oppgaven, men bevis kan finnes i tekser om avansert kalkulus.

Vi begynner nå å se på beviset. Det er bygd opp av to lemmaer som til slutt settes sammen til det tredje beviset.

Lemma 5.2. *La $f(x) \in \mathbb{C}[x]$. Da vil $\left|f(x)\right|$ ta en minimumsverdi for en $x_0 \in \mathbb{C}$.*

Bevis. Når $|x| \to \infty$, vil $\left|f(x)\right| \to \infty$. Siden $\left|f(x)\right|$ er stor for en stor $|x|$ vil den største nedre skranken m av $\left|f(x)\right|$ for $z \in \mathbb{C}$ også være den største nedre skranken i en tilstrekkelig stor disk $|z| < r$. Siden $\left|f(x)\right|$ er en kontinuerlig reell funksjon, følger det av Lemma 3.5.1 at $f(x)$ vil oppnå sin minimumsverdi på denne disken.

■

Lemma 5.3. *Anta at $f(x) \in \mathbb{C}[x]$ hvor $f(x)$ er ikke-konstant. Om $f(x_0) \neq 0$, så er $\left|f(x_0)\right|$ ikke minimumsverdien av $\left|f(x)\right|$.*

Bevis. La $f(x)$ være et ikke-konstant kompleks polynom og anta at x_0 er et punkt med $f(x_0) \neq 0$. Gjør et variabelskifte av $x + x_0$ for x. Dette skifter x_0 til origo, slik at vi kan anta at $f(0) \neq 0$. Vi multipliser deretter $f(x)$ med $f(0)^{-1}$ slik at $f(0) = 1$. Vi må da vise at 1 ikke er minimumsverdien til $\left|f(x)\right|$.

La k være den laveste ikke-null potensen av x som forekommer i $f(x)$. Da kan $f(x)$ antas å ha formen

$$f(x) = 1 + ax^k + x^{k+1}g_1(x),$$

for et polynom $g_1(x)$.

La nå α være den k-te roten av $-a^{-1}$. Gjør den endelige endringen av variabel αx for x. Nå er $f(x)$ på formen

$$f(x) = 1 - x^k + x^{k+1}g_2(x),$$

for et polynom $g_2(x)$.

For små positive reelle x får vi fra trekantulikhet

$$\left|f(x)\right| \leq \left|1 - x^k\right| + x^{k+1}\left|g_2(x)\right|$$

Men, for små positive reelle verdier av x er $x^k < 1$, så ulikeheten kan skrives på formen

$$\left|f(x)\right| \leq 1 - x^k + x^{k+1}\left|g_2(x)\right| = 1 - x^k(1 - x\left|g_2(x)\right|$$

For små reelle x, er $x\left|g(x)\right|$ liten, så x_0 kan bli valgt slik at $x_0\left|f(x_0)\right| < 1$. Det følger da at $x_0^k(1 - x_0\left|f(x_0)\right|) < 1 = \left|f(0)\right|$, som vil si at $\left|f(0)\right|$ ikke er minimumsverdien.

■

Ved å nå kombinere de to siste lemmaene, får vi vårt tredje og siste bevis for algebraens fundamentalteorem.

Korollar 5.2. *Korollar 4.1 følger av Lemma 5.2 og 5.3.*

Bevis. La $f(x)$ være et ikke-konstant kompleks polynom. Fra Lemma 5.2 har $|f(x)|$ en minimumsverdi for en $x_0 \in \mathbb{C}$. Så fra Lemma 5.3 følger det at $|f(x_0)| = 0$, og derav $f(x_0) = 0$ for ellers ville det ikke vært minimumsverdien. Derfor har $f(x)$ en kompleks rot.

■

5.3 Linearisering av komplekse polynomer

Det aller siste vi skal se på i denne oppgaven er en direkte konsekvens fra Korollar 4.1.

Korollar 5.3. *Et komplekst polynom kan faktoriseres til lineære faktorer.*

Bevis. La $f(x) \in \mathbb{C}[x]$ og bruk induksjon på graden n. Dersom $\deg f(x) = 1$ stemmer korollaret, siden da er $f(x)$ i seg selv en lineær faktor. Anta at det stemmer for $\deg f(x) = n - 1$. Vi lar $\deg f(x) = n$. Fra Korollar 4.1 eksisterer det en rot α_1, og $(x - \alpha_1)$ deler $f(x)$. Dermed er $f(x) = (x-\alpha_1)g(x)$, hvor $\deg g(x) < n$. Fra induksjonshypotesen så kan $g(x)$ faktoriseres til lineære faktorer, og følgelig kan $f(x)$ også faktoriseres, og $f(x)$ kan skrives på følgende måte

$$f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \dots (x - \alpha_n).$$

■

References

[1] Harm Derksen, **The Fundamental Theorem of Algebra and Linear Algebra**, The American Mathematical Monthly, www.jstor.org/stable/3647746, [hentet: 30.09.2020]

[2] Benjamin Fine og Gerhard Rosenberger, **The fundamental theorem of algebra**, Springer, 1997.