SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Cayley-Hamiltons sats

av

Leo Gustafsson

2020 - No K14

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Cayley-Hamiltons sats

Leo Gustafsson

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Alan Sola

2020

Innehåll

Kapitel 1

Inledning

Detta arbete kommer att behandla Cayley-Hamiltons sats inom den linjära algebran och hur den satsen kan bevisas.

För att kunna göra det är det först en bakgrundsdel där det relevanta gås igenom vad det gäller definitioner, begrepp, satser och bevis om: matriser, determinanter, linjära avbildningar, egenvärden och egenvektorer och diagonaliserbarhet av matriser samt hur man räknar på detta.
Efter att bakgrunden har gåtts igenom är tanken att vi har fått en större kunskap om de olika delarna inom den linjära algebran så att vi kan knyta ihop dom för att ta oss an själva huvudsatsen och beviset.

Det Cayley-Hamiltons sats säger är att om vi sätter in en kvadratisk matris i sin egna karakteristiska ekvation så får vi ut nollmatrisen. I bevisdelen går vi igenom att det stämmer för en matris av dimension 1 till och med 3. Vi utgår ifrån att vi har en övertriangulär blockmatris som innehåller mindre kvadratiska matriser av maximal storlek 1×1 eller 2×2 med eller utan egenvärden längs med diagonalen av den övertriangulära blockmatrisen. Sedan delar vi upp beviset i olika fall beroende på dimension och undersöker de enskilda bitarna av en matris med hjälp av det karakteristiska polynomet och invarianta delrummen. (Axler 1997, ss. 182-200.)

För att ge en djupare insikt om Cayley-Hamiltons sats uppkomst så redovisas den historiska bakgrunden om de två huvudpersonerna Arthur Cayley(1821-1895) och Sir William Rowan Hamilton(1805-1865) som ligger bakom denna sats.

Kapitel 2

Bakgrund

2.1 Matriser

Det finns många typer av matriser som är uppbyggda på olika sätt och i denna del ska vi gå igenom hur några av dom är uppbyggda, vilka egenskaper de har samt räkneregler.

Definition av en matris

"*Med en matris, eller mer precist en matris av storlek $m \times n$, där m och n är två positiva heltal, menas mn reella tal ordnade i ett rektangulärt schema med m rader och n kolonner.*"(Bøgvad & Vaderlind 2017, s.13)
Vi kan beskriva en matris M som:

$$M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Varje tal i en matris kallas för ett element och har en speciell plats och betecknas: a_{ij}. Som vi ser ovan så befinner sig talet a_{11} på platsen där $i = 1$ och $j = 1$, alltså första platsen för både kolonn och rad. En matris kan vi således skriva såsom: $M = (a_{ij})$, där $1 \leq i \leq m$ och $1 \leq j \leq n$. (Bøgvad & Vaderlind 2017, s.13)

Nollmatrisen

Nollmatrisen är en matris av storlek $m \times n$ som har egenskapen att alla element är lika med noll. Vi skriver det som: $M = (a_{ij}) = 0$, då $1 \leq i \leq m$ och $1 \leq j \leq n$. Vi har dock inte en och samma nollmatris som är unik utan nollmatrisen beror på hur många rader m och hur många kolonner n vi har i matrisen. (Bøgvad & Vaderlind 2017, s. 14)

Kvadratisk matris

I de kvadratiska matriserna är antalet rader av samma antal som antalet kolonner, dvs $m = n$. De betecknas då: $M_{n \times n}$. Exempel:
$M = \begin{pmatrix} a_{11} \end{pmatrix}$ och $M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. (Bøgvad & Vaderlind 2017, s.14)

Enhetsmatrisen (identitetsmatrisen)

Enhetsmatrisen är en matris på formen $M_{n \times n}$ likt ovan nämnda exempel och kallas Id eller E_n. Skillnaden är att vi har ettor längs elementen på diagonalen på matrisen och resten av elementen består av nollor.
Exempel: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ och $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (Bøgvad & Vaderlind 2017, s.14)

Radmatris Kallas också radvektorer och består av endast en rad $m = 1$ och n stycken kolonner. $M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_n \end{pmatrix}$

Kolonnmatris Kallas också kolonnvektor och består av en kolonn n och m stycken rader: $M = m \times 1$. $M = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_n \end{pmatrix}$ (Bøgvad & Vaderlind 2017, s.14)

Invers matris

Definition 1.3
Om vi har en kvadratisk matris $M_{n \times n}$ så kallas matrisen $A_{n \times n}$ av samma storlek som M för dess invers om följande gäller: $MA = AM = Id$, där även

Id är av samma kvadratiska storlek som M och A. Det skulle vara logiskt att beteckna M^{-1} för A om den existerar. Men skulle det kunna finnas andra matriser än A som skulle fungera som M^{-1}? Vi undersöker detta med nedanstående sats. (Bøgvad & Vaderlind 2017, s. 28)

Sats 1.6

Vi har att om $MA = AM = Id$ och $MC = CM = Id$, då är $A = C$.

Bevis

Vi börjar med att multiplicera $MC = Id$ med A från vänster och vi får $A(MC) = (AM)C = IdC = C$, där vi ser att $AM = Id$ och använder associativitet enligt sats 1.4.

Om en invers existerar så är den alltså entydig. Invers kan dock saknas. (Bøgvad & Vaderlind 2017, s. 29)

Matristransponering

Definition

Transponatet definieras som matrisen vi får om vi byter plats på rader och kolonner. Om vi har en matris M så betecknas transponatet som M^t. (Bøgvad & Vaderlind 2017, s. 27 f.)

Exempel: $M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$, då är $M^t = \begin{pmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{pmatrix}$

Triangulära matriser

Övertriangulära matrier har endast 0:or under huvuddiagonalen och undertriangulära matriser är av samma form fast tvärtom, där det endast är 0:or på elementen över huvuddiagonalen. Sedan har vi diagonala matrisen som endast har element längs med huvuddiagonalen. Nedan har vi tre exempel i 3×3-fallet:

$$M_1 = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ 0 & m_{22} & m_{23} \\ 0 & 0 & m_{33} \end{pmatrix}, M_2 = \begin{pmatrix} m_{11} & 0 & 0 \\ m_{21} & m_{22} & 0 \\ m_{31} & m_{32} & m_{33} \end{pmatrix}, M_3 = \begin{pmatrix} m_{11} & 0 & 0 \\ 0 & m_{22} & 0 \\ 0 & 0 & m_{33} \end{pmatrix}.$$

Här är M_1 en övertriangulär matris och M_2 är undertriangulär matris och M_3 är en diagonalmatris. En diagonalmatris kan vi konstatera är både uppåt och nedåt triangulär (Bøgvad & Vaderlind 2017, s. 50f)

Övertriangulär blockmatris

En övertriangulär blockmatris är en kvadratisk matris $M_{n\times n}$ på formen

$$M = \begin{pmatrix} A_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & A_m \end{pmatrix}$$

med kvadratiska matriser $A_{n\times n}$ längs med matrisens diagonal. Vi har nollor under diagonalen och $*$ är godtyckliga element ovanför diagonalen.
Exempel:

$M = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 2 & 4 & 5 \\ 0 & 2 & 2 & 4 & 5 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix}$, där M har tre blockmatriser längs med diagonalen. Vi får då: $M = \begin{pmatrix} A_1 & & * \\ & A_2 & \\ 0 & & A_3 \end{pmatrix}$,
där vi har $A_1 = \begin{pmatrix} 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ och $A_3 = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$.

Det vi även allmänt kan konstatera är att varje övertrianglär matris är en blockmatris men med endast 1×1-matriser längs med huvuddiagonalen. Sedan är även varje kvadratisk matris en blockmatris eftersom vi kan ha hela matrisen som en blockmatris. (Axler 1997, s. 183 f.)

Räkneregler matriser

Först går vi igenom regler för addition av (flera) matriser och vi börjar med definitionen.

Definition
Vi har två matriser av samma form: A och B, där $A = (a_{ij})$ och $B = (b_{ij})$ där $1 \le i \le m$ och $1 \le j \le n$. Om vi adderar dessa två matriser så får vi matrisen $A + B$ definieras som matrisen med elementen $a_{ij} + b_{ij}$ där $1 \le i \le m$

och $1 \leq j \leq n$.

Exempel: $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ och $B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, addition av A och B ger oss: $A + B = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$. Med andra ord så adderas talen som ligger på samma plats i matriserna A och B med varandra och vi får en ny matris $A + B$.

Sats 1.1

Vi antar att vi har tre matriser av samma form $m \times n$. Då gäller följande räkneregler.

(i): $A + B = B + A$, kommutativa lagen,
(ii): $A + (B + C) = (A + B) + C$, associativa lagen,
(iii): $A + 0 = A$, för en nollmatris av samma typ som A.

Bevis för de här räknereglerna är desamma som egenskaperna för tal och enkla definitioner av de, exempel: $a+b = b+a, a+(b+c) = (a+b)+c, a+0 = a$. (Bøgvad & Vaderlind 2017, s.15)

Sedan går vi igenom matriser och räkneregler när man ska multiplicera en matris med ett tal. Vi antar även här att matrisen A och B är av samma form och λ, μ är reella tal.

Sats 1.2

(i): $\lambda(A + B) = \lambda A + \lambda B$,
(ii): $(\lambda\mu)A = \lambda(\mu A)$,
(iii): $(\lambda\mu)A = \lambda(\mu A)$,

Beviset för denna sats hänvisas till Bøgvad & Vaderlind (2017, s. 16).

Till sist kan det vara bra att gå igenom multiplikation av fler än en matris och räkneregler för dessa:

Sats 1.4

(i): $AId = IdA = A$; $A0 = 0$; $0A = 0$,

(ii): $(A+B)C = AC + BC$ och $C(A+B) = CA + CB$ (distributiva lagen),
(iii):$(\lambda A)B = A(\lambda B) = \lambda(AB)$,
(iv): $A(BC) = (AB)C$ (associativa lagen).

Bevis för dessa fyra räkneregler gås noggrant igenom av Bøgvad och Vaderlind (2017, s. 20 f.) men de är för omfattande för att vara med i denna uppsats.

Nu vet vi mer om hur matriser ser ut och några av deras räkneregler. Därför är det dags att gå vidare till nästa del, nämligen determinanter.

2.2 Determinanter

I detta avsnitt går vi kort igenom determinanter och hur de är uppbyggda. Vi går även igenom de enklaste räknereglerna för 1×1, 2×2 och 3×3-matriser samt några olika viktiga klasser av matriser. Determinanten används för att studera egenskaper hos kvadratiska matriser.
För att definiera determinanter så betraktar vi kvadratiska matriser $M_{n \times n}$.

$$\det M = \begin{vmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \cdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{vmatrix}$$

Exempel

1×1-fallet: Vi har en matris med endast ett element $M = (m_{11})$ då definieras determinanten som: $\det M = m_{11}$

2×2-fallet: Om vi har en matris av storlek 2×2: $M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$, då definieras determinanten av M som: $\det M = (m_{11}m_{22}) - (m_{12}m_{21})$.

3×3-fallet: Om vi har en matris på formen 3x3 så får vi determinanten genom utveckling efter rad 1. (Bøgvad & Vaderlind 2017, s. 37 ff.)

$$\begin{vmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{vmatrix} = m_{11} \begin{vmatrix} m_{22} & m_{23} \\ m_{32} & m_{33} \end{vmatrix} - m_{12} \begin{vmatrix} m_{21} & m_{23} \\ m_{31} & m_{33} \end{vmatrix} + m_{13} \begin{vmatrix} m_{21} & m_{22} \\ m_{31} & m_{32} \end{vmatrix} =$$

$$= m_{11}(m_{22}m_{33} - m_{23}m_{32}) - m_{12}(m_{21}m_{33} - m_{23}m_{31}) + m_{13}(m_{21}m_{32} - m_{22}m_{31}).$$

Vid beräkning av en determinant medelst radutveckling eller kolonnutveckling så behöver man veta att varannat element i matrisen med början från m_{11} har ett positivt $(+)$ eller negativt$(-)$-tecken framför sig, där m_{11} är positivt och m_{12} således blir negativt. Alltså på platsen (ij) i matrisen blir tecknet framför elementet $(-1)^{(i+j)}$.

Exempel: $\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$. För att det ska bli klarare går vi igenom mer i determinanter av godtycklig storlek här nedan. (Bøgvad & Vaderlind 2017, s. 42 ff.)

Determinanter av godtycklig storlek

Definition
Låt T vara en kvadratisk matris, $T \in M_{n \times n}(\mathbb{R})$. Om vi har $n = 1$ så definieras determinanten som: $T = T_{11}$.
Om $n \geq 2$ så definieras determinanten rekursivt som:
$\det T = \sum_{j=1}^{n} (-1)^{1+j} T_{1j} \cdot \det(\widehat{T}_{1j})$, där $(\widehat{T}_{1j})$ är den matris vi får om rad (i) och kolonn (j) stryks med utgång från den positionen i matrisen vill beräkna vår determinant ifrån. $(\widehat{T}_{1j})$ är därmed en matris av storlek $(n-1) \times (n-1)$.

Kofaktorn till $\det T$ betecknas som $(-1)^{i+j} \cdot \det(\widehat{T}_{ij})$ och beskriver om vi utvecklar efter rad eller kolonn i vår beräkning av determinanten till T, se nedan exempel. Vi kan förenkla genom att beteckna kofaktorn som: $u_{ij} = (-1)^{i+j} \cdot \det(\widehat{T}_{ij})$. Om vi utvecklar determinanten till T efter rad 1 får vi: $\det(T) = T_{11}u_{11} + T_{12}u_{12} + T_{13}u_{13} + \cdots + T_{1n}u_{1n}$. (Friedberg et al. 2014, s. 209 f.)

Exempel: Vi har matrisen $T = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 2 & 2 \\ 4 & 2 & 2 \end{pmatrix}$ och vi vill beräkna dess determinant längs rad 1, då får vi:

$$\det(T) = T_{11}(-1)^{1+1} \cdot \det(\widehat{T}_{11}) + T_{12}(-1)^{1+2} \cdot \det(\widehat{T}_{12}) + T_{13}(-1)^{1+3} \cdot \det(\widehat{T}_{13})$$

$$= (1)(-1)^2 \cdot \det \begin{vmatrix} 2 & 2 \\ 2 & 2 \end{vmatrix} + (2)(-1)^3 \cdot \det \begin{vmatrix} 3 & 2 \\ 4 & 2 \end{vmatrix} + (5)(-1)^4 \cdot \det \begin{vmatrix} 3 & 2 \\ 4 & 2 \end{vmatrix}$$

$$= 1 \begin{vmatrix} 2 & 2 \\ 2 & 2 \end{vmatrix} - 2 \begin{vmatrix} 3 & 2 \\ 4 & 2 \end{vmatrix} + 5 \begin{vmatrix} 3 & 2 \\ 4 & 2 \end{vmatrix} = 1(0) - 2(-2) + 5(-2) = 0 + 4 - 10 = -6$$

Elementära radoperationer

(i): Vi kan byta plats på två rader i en determinant men då skiftar determinantens tecken.
(ii): Om vi multiplicerar en rad i matrisen M med en skalär $\lambda \neq 0$ så blir den nya determinantens värde lika med: $\lambda \det M$
(iii): Vi kan addera eller subtrahera en rad till en annan utan att determinantens värde ändras.

Vi använder dessa metoder vid räkning för att få matrisen på en reducerad trappstegsform och att den även ska vara uppåt triangulär då det förenklar vår beräkning av en determinant. (Bøgvad & Vaderlind 2017, s. 106 f)

Metod för beräkning av determinanter:

Sats 4.8: Antag att vi har en övertriangulär kvadratisk matris av storlek $M_{n \times n}$. Då är determinanten lika med produkten av matrisens element längs huvuddiagonalen. Alltså får vi: $\det M = m_{11}m_{22}...m_{nn}$

$$M = \begin{pmatrix} m_{11} & m_{12} & ... & m_{1n} \\ 0 & m_{22} & . & . \\ . & . & . & . \\ 0 & 0 & ... & m_{nn} \end{pmatrix}$$

Bevis: Vi utvecklar matrisen M efter första kolonn och då stryks den första raden och kolonnen i matrisen M. Då får vi den nya determinanten $\det M = m_{11} \det M'$, där M' blir den nya uppåt triangulära matrisen efter utvecklingen efter kolonn 1. M' har storlek $(n-1) \times (n-1)$ då det försvinner en rad och kolonn efter utvecklingen. Sedan fortsätter vi att utveckla determinanten av M och vi ser att m_{22} är det nya översta elementet uppe i vänstra hörnet av M' och därför utvecklar vi efter den och får: $\det M = m_{11} m_{22} M''$ där M'' blir den nya uppåt triangulära matrisen efter utveckling. M'' har storlek $(n-2) \times (n-2)$ då det försvinner en rad och kolonn efter utveckling. Om vi forsätter med utveckling efter elementet i det övre vänstra hörnet i den nya matrisen som har bildats så kommer vi tillslut få det resultat vi ville ha i och med sats 4.8. (Bøgvad & Vaderlind 2017, s. 105f.)

När är en matris inverterbar?

Sats 4.2
Låt T vara en en 2×2-matris, $T \in M_{2\times 2}(\mathbb{R})$. Då är T inverterbar om och endast om $\det T \neq 0$ och $T^{-1} = \frac{1}{\det T} \begin{pmatrix} T_{22} & -T_{12} \\ -T_{21} & T_{11} \end{pmatrix}$

Bevis
Om $\det T \neq 0$, så kan vi definiera en matris $A = \frac{1}{\det T} \begin{pmatrix} T_{22} & -T_{12} \\ -T_{21} & T_{11} \end{pmatrix}$

Vi ska nu kontrollera att $AT = Id$. Vi får:

$$AT = \frac{1}{\det T} \begin{pmatrix} T_{22} & -T_{12} \\ -T_{21} & T_{11} \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} =$$

$$= \frac{1}{\det T} \begin{pmatrix} T_{11}T_{22} - T_{12}T_{21} & 0 \\ 0 & -T_{21}T_{12} + T_{11}T_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Efter beräkning får vi ut identitetsmatrisen $Id = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, där $\det \neq 0$ och därmed är A en invers till T.
Om vi istället antar att T är inverterbar och T^{-1} existerar så är $\det \neq 0$. Då kan vi ta determinanten av båda sidorna i $TT^{-1} = id$. Då får man $\det T \det T^{-1} = \det id = 1$, så $\det T \neq 0$. (Friedberg et al. 2014, s. 200 f.)

Allmänna n x n-fallet En $n \times n$-matris A är inverterbar om och endast om $\det A \neq 0$. Detta är i sin tur ekvivalent med att $\text{rank}A = n$. Om A är inverterbar, så är även tranponatet A^t inverterbart och vi har $\det A = \det A^t$ samt $(A^t)^{-1} = (A^{-1})^t$. För bevis se följdsats 4.7 och sats 4.8. i Friedberg et al. (2014, s. 223 f.).

2.3 Vektorrum

Definition (Axler 1997, s. 9 ff.)
Ett vektorrum över $\mathbb{R}$ är en mängd V som är utrustad med två operationer, nämligen addition och multiplikation med skalär.

Addition tar två element i V och producerar ett nytt element i V: $V \times V \to V$
Multiplikation tar ett element i $\mathbb{R}$ och ett i V och producerar ett nytt element i V: $\mathbb{R} \times V \to V$

Om de två operationerna uppfyller följande axiom så är V ett vektorrum:

(i): $x + y = y + x, (\forall x, y \in V)$, kommutativa lagen
(ii): $x + (y + z) = (x + y) + z, (\forall x, y, z \in V)$, associativa lagen
(iii): Det existerar ett element $0_v \in V$ sådant att: $x + 0_v = x$, $(\forall x \in V)$, additiv identitet
(iv): För varje $x \in V$ finns ett $y \in V$ sådant att: $x + y = 0_V$, additiv invers
(v): $1x = x, \forall x \in V$, multiplikativ identitet
(vi): $(ab)x = a(bx)$, $\forall a, b \in \mathbb{R}$ och $\forall x \in V$ associativa lagen
(vii): $a(x + y) = ax + ay$ och $(a + b)x = ax + by$ $\forall x, y \in V$ och $\forall a, b \in \mathbb{R}$ (distributiva egenskapen)

Exempel på reellt vektorrum $\mathbb{R}$

Vi låter $\mathbb{R}^n$ vara vår underliggande mängd. Det är alltså mängden av n-tupler som är listor av n st element av reella tal i $\mathbb{R}$.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, x_j \in \mathbb{R}$$

Multiplikation med skalär a definieras genom:

$$ax = \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}, x_j \in \mathbb{R}, a \in \mathbb{R}$$

Exempel: $2A = 2\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}$

Vi multiplicerar således skalären a med koordinaterna för x:s positioner tills vi kommer till den $n:te$ koordinaten.

Addition definieras genom:

$$x + y = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, x_j \in \mathbb{R}, y_j \in \mathbb{R}$$

Vi adderar således motsvarande koordinater för x och y tills vi kommer till den $n:te$ koordinaten.

Exempel: $A = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, A + B = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$. (Axler 1997, s. 4 ff.)

Vektorrum av polynom:

Låt oss nu studera ett andra exempel där vi har ett vektorrum: $V = P_n(\mathbb{R})$

som är mängden av polynom: $p = a_0 + a_1x + ... + a_nx^n$, där koefficienterna a_j tas från $\mathbb{R}$, dvs $a_j \in \mathbb{R}$, och graden som detta polynomet har är $\leq n$, där n är ett icke negativt heltal.

Ett polynom kan uppfattas som en funktion av x, då får vi funktionen av p och q som: $(p+q)(x) = p(x) + q(x)$, $p, q \in P(\mathbb{R})$.
Sedan har vi multiplikation med skalär på $P(\mathbb{R})$, där ap är polynomet definierad av:
$(ap)(x) = a(px)$, om $a \in \mathbb{R}$ och $p \in P(\mathbb{R})$. (Axler 1997, s. 10 f)

Exempel:
Vi har: $p(x) = 1 - x^2 + 10x$ och $q(x) = x^3 + x^2 + 2$

$$(p+q)(x) = x^3 + 10x + 3$$

$$(ap)(x) = a(1 - x^2 + 10x) = a - ax^2 + 10ax$$

Delrum

Ett vektorrum kan man bryta upp i mindre bitar för att förstå det bättre. Då kommer vi till delmängder som har fått överta en speciell struktur av vektorrum.

Definition

Vi har ett vektorrum V över $\mathbb{R}$. En delmängd U av V, $U \subseteq V$, säges vara ett delrum till V om U är ett vektorrum över $\mathbb{R}$ med följande operationer som U ärvt från V.
(i): $x + y \in U$, $x, y \in U$. (slutet under addition)
(ii): $ax \in U$, $x \in U$ och $a \in \mathbb{R}$. (slutet under multiplikation med skalär). (Axler 1997, s. 13 f.)

Linjärt beroende

Definition
Låt $S = \{x_1, ..., x_n\} \subset V$ och skalärer $c_1, ..., c_n$ där alla inte är lika med noll. Om vi kan uttrycka någon av vektorerna som en linjärkombination av de andra vektorerna så är de linjärt beroende.:

$$c_1x_1 + \cdots + c_{n-1}x_{n-1} = x_n$$

Linjärt oberoende

Om vi väljer ut en mängd av vektorer $S = \{x_1, ..., x_n\} \subset V$ och skalärer $c_1, ..., c_n$ så är de linjärt oberoende om de uppfyller att:
$c_1x_1 + \cdots + c_nx_n = 0_V$, har den enda lösningen: $c_1 = \cdots = c_n = 0$.
Ingen av vektorerna kan därmed skrivas som en linjärkombination av de övriga vektorerna. (Friedberg et al. 2014, s. 36 f.)

Linjära höljet

Definition

Låt $S = \{x_1, ..., x_n\} \subset V$. Det linjära höljet betecknas som $span(S)$ och definieras som en delmängd till V som består av alla möjliga linjärkombinationer av vektorerna $\{x_1, ..., x_n\}$.

Det linjära höljet är ett delrum till V, alltså $spanS \subseteq V$, då det är slutet under multiplikation med skalärer och under addition i V. För bevis se sid. 30 ff. i Friedberg et al. (2014).

Exempel Vi har att vektorn $x = \begin{pmatrix} 2 \\ -5 \\ 0 \end{pmatrix}$ är en linjärkombination av $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ och $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ i $\mathbb{R}^3$ eftersom $\begin{pmatrix} 2 \\ -5 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-5) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

Baser:

Definition: En icke-tom mängd B är en bas till ett vektorum V om B utgörs av linjärt oberoende vektorer $\{x_1, ..., x_n\}$ och $span(B) = span\{x_1, ..., x_n\} = V$. (Friedberg et al. 2014, s. 43.)

2.4 Linjära avbildningar

Definition

Antag att vi har två vektorrum V och W över $\mathbb{R}$. En linjär avbildning är då en funktion eller transformation mellan de två rummen, $T : V \to W$ med egenskaperna.

(i): $T(x + y) = Tx + Ty,\ \forall x, y \in V$;
(ii): $T(cx) = c(Tx),\ \forall c \in \mathbb{R}$ och $\forall x \in V$ (Axler 1997, s. 38)

Nollrum och Bildrum

Definition

Låt $T : V \to W$ vara en linjär avbildning, där V och W är två vektorrum. Nollrummet $N(T)$ är mängden av alla vektorer $x \in V$ så att T avbildar dem på nollvektorn. dvs $N =\{x \in V : T(x) = 0_w\}$.
Vi definierar bildrummet till T som mängden av alla $T(x)$ för $x \in V$ och betecknar det med $R(T)$.

Om $T : V \to W$ är en linjär avbildning, då är nollrummet $N(T)$ och bildrummet $R(T)$ delrum till V respektive W. För bevis se sid 68 i Friedberg et al. (2014).

Låt $T : V \to W$ vara en linjär avbildning. Om vi har en bas till V som vi kallar $\gamma = \{x_1, ..., x_n\}$ så får vi $R(T) = span(T(\gamma)) = span(\{T(x_1), ..., T(x_n)\})$.

Exempel: Vi definierar en avbildning $T : P_3(\mathbb{R}) \to P_2(\mathbb{R})$ genom $T(p) = p'$, derivatan av p. Då är T linjär enligt deriveringsreglerna. Vi har också standardbasen för $P_3(\mathbb{R}) = \{1, x, x^2, x^3\}$. Om vi nu studerar bilderna av basvektorerna i standardbasen för T.

$$T(1) = 0$$

$$T(x) = 1$$
$$T(x^2) = 2x$$
$$T(x^3) = 3x^2$$

Vi har $N(T) = span(\{1\})$ och $R(T) = span(\{1, x, x^2\})$. (Friedberg et al. 2014, s. 67 ff.)

Dimensionssatsen

Låt $T : V \to W$ vara en linjär avbildning. Då gäller: $\dim N(T) + \dim R(T) = \dim V$. För bevis se sid 70 i Friedberg et al. (2014).

Matrisen för en linjär avbildning

Definition
Vi tittar nu återigen på en linjär avbildning mellan de två vektorummen.
$T : V \to W$ där vi antar $(x_1, ...x_n)$ är en bas av linjärt oberoende vektorer som tillhör V och linjärt oberoende vektorer $(y_1, ..., y_m)$ tillhör W. Vi kan skriva.
$\gamma = \{x_1, ...x_n\} \subset V$, vi kallar ena basen för γ.
$\beta = \{y_1, ...y_m\} \subset W$, vi kallar den andra basen för β.

Varje vektor i V har koordinater relativt γ, alltså $[x]_\gamma = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$, alltså en kolonnvektor med n stycken tal.

Varje vektor i W har koordinater relativt β, alltså $[y]_\beta = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} \in \mathbb{R}^m$, alltså en kolonnvektor med m stycken tal.

Definiera T på matrisformen $m \times n$ relativt γ och β. Där vi gått från basen γ i $\mathbb{R}^n$ till basen β i $\mathbb{R}^m$. Vi tillämpar först linjär avbildning på första elementet i basen γ. Varje element i y har koordinatframställning relativt basen β. Så vi ställer upp den som en kolonn $[T(x_1)]_\beta$. Vi gör detta för varje element i basen γ tills vi når $[T(x_n)]_\beta$. Vi ställer upp matrisen T av n stycken kolonnvektorer av längd m.

$$\Big[T\Big]_\gamma^\beta = ([T(x_1)]_\beta | \cdots | [T(x_n)]_\beta)$$

.

Exempel:
Vi har en linjär avbildning som definieras av $p(x) :\to (x + x^2)p''(x)$ i ett

fyrdimensionellt rum $T : P_3(\mathbb{R}) \to P_3(\mathbb{R})$.

Hur får vi då fram matrisen för den linjära avbildningen T relativt standardbasen $S = \{1, x, x^2, x^3\}$? Eftersom vi nu har samma vektorrum kan vi använda en och samma bas.

$T(1) = (x + x^2)0 = 0$, vi får första kolonnvektorn $[T(1)]_S = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

$T(x) = (x + x^2)0 = 0$, vi får andra kolonnvektorn $[T(x)]_S = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

$T(x^2) = (x + x^2)2 = 2x + 2x^2$, vi får tredje kolonnvektorn $[T(x^2)]_S = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 0 \end{pmatrix}$

$T(x^3) = (x+x^2)6x = 6x^2+6x^3$, vi får fjärde kolonnvektorn $[T(x^3)]_S = \begin{pmatrix} 0 \\ 0 \\ 6 \\ 6 \end{pmatrix}$

Vi har nu fått fram matrisen T för den linjära avbildningen $[T]_S = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 6 \end{pmatrix}$

(Friedberg et al. 2014, s. 79 ff.)

2.5 Egenvärden och egenvektorer

Definition

Antag att vi har en linjär avbildning: $T : V \to V$ där V är ett vektorrum. Då är vektorn $x \in V, x \neq 0_v$ en egenvektor till den linjära avbildningen T och λ egenvärdet om följande gäller.

$$T(x) = \lambda x, \lambda \in \mathbb{R}$$

Mängden av egenvektorer till ett specifikt egenvärde λ bildar ett delrum som också kallas för egenvärdet λ:s egenrum och det betecknas med E_λ.

Det är inte klart att varje linjär avbildning har en egenvektor eller egenvärde. Därför behöver vi en metod för att hitta dessa. (Friedberg et al. 2014, s. 246)

Bestämning av egenvärden och egenvektorer:

Ekvationen $T(x) = \lambda x$ är då ekvivalent med $(T - \lambda Id)x = 0$ och vi börjar med sats 5.2.

Sats 5.2

Låt $T \in M_{n\times n}$. Då är λ ett egenvärde om och endast om det är en rot till den karakteristiska ekvationen $\det([T]_\beta - \lambda Id) = 0$

Bevis

λ är ett egenvärde till T om och endast om det existerar en egenvektor $x \neq 0$, så att: $Tx = \lambda x \iff (T - \lambda Id)(x) = 0$.
Då är λ egenvärde om och endast om: $(T - \lambda Id)$ inte är inverterbar, dvs determinantens värde är 0. Då får vi icke-triviala lösningar till ett homogent ekvationssystem med parameterlösning. Detta resultat är ekvivalent med påståendet: $\det(T - \lambda Id) = 0$ vilket kallas den karakteristiska ekvationen. När vi utvecklar determinanten får vi det karakteristiska polynomet. (Friedberg et al. 2014, s.247 ff.)

Karakteristiska polynomet

Definition
Låt matrisen $T \in M_{nxn}(\mathbb{R})$. Då är det karakteristiska polynomet av T $q(\lambda) = \det(T - \lambda Id)$.

Sats 5.2 ovan innebär att egenvärden för matrisen T är nollställena till det karakteristiska polynomet och för att bestämma dessa nollställen så löser vi $q(\lambda) = \det(T - \lambda Id) = 0$. (Friedberg et al. 2014, s.247 ff.)

Exempel: Räkna ut det karakteristiska polynomet av matrisen $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$q(\lambda) = \det\left(\begin{vmatrix} a & b \\ c & d \end{vmatrix} - \lambda \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\right) = \det\begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = 0$$

Utveckling av determinanten ger: $(a-\lambda)(d-\lambda)-bc = \lambda^2-(a+d)\lambda+ad-bc = 0$

Sats 5.2 ger oss därmed möjlighet att finna egenvärden till en linjär avbildning eller matris förutsatt att vi kan använda oss utav det karakteristiska polynomet för att finna nollställena till polynomet. Nästa steg är att hitta egenvektorerna till ett givet egenvärde λ på ett systematiskt vis och det gör vi med nedanstående sats.

Sats 5.4

Låt $T : V \to V$ och λ är ett egenvärde till T. Då är vektorn $x \in V$ egenvektor av T motsvarande till egenvärdet λ om och endast om $x \in N(T - \lambda Id)$ där $x \neq \vec{0}$(nollvektorn). (Friedberg et al. 2014, s. 250 ff.)

Exempel: Vi har matrisen $T = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$ och vi vill hitta egenvärdena λ genom att beräkna det karakteristiska polynomet.

$$q(\lambda) = \det\begin{vmatrix} 1 - \lambda & 3 \\ 4 & 2 - \lambda \end{vmatrix} = 0$$

Efter utveckling får vi: $\lambda^2 - 3\lambda - 10 = 0$

dvs två egenvärden: $\lambda_1 = -2$ och $\lambda_2 = 5$ som är rötterna till $q(\lambda)$.

Hur finner vi då egenvektorerna?

Vi bestämmer en bas för nollrummet $N(T - \lambda Id)$ genom att använda de två rötterna till det $q(\lambda)$ och bestämma nollrummet genom att först stoppa in $\lambda_1 = -2$ och lösa ekvationssystemet för att sedan göra samma sak med $\lambda_2 = 5$ och lösa det nya ekvationssystemet.

När $\lambda_1 = -2$ får vi:

$E_{-2} = \left(\begin{array}{cc|c} 3 & 3 & 0 \\ 4 & 4 & 0 \end{array}\right) \sim T = \left(\begin{array}{cc|c} 3 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right) \sim T = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$.Vi får då parameterlösningen:

$$\begin{bmatrix} x_1 = -x_2 \\ x_2 = s \end{bmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} s \\ -s \end{pmatrix} = s \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Vi får då egenvektorn x_1 vilket ligger i egenrummet som tillhör egenvärdet $\lambda_1 = -2$.

$$E_{-2} = Span\left(x_1\right) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Testa om definitionen stämmer, dvs: $T(x) = \lambda x$:

$\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (-2) \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, vilket den gör.

När $\lambda_2 = 5$ får vi:

$E_5 = \left(\begin{array}{cc|c} -4 & 3 & 0 \\ 4 & -3 & 0 \end{array}\right) \sim T = \left(\begin{array}{cc|c} -4 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right)$.Vi får då parameterlösningen:

$$\begin{bmatrix} -4x_1 = -3x_2 \\ x_2 = s \end{bmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3s \\ 4s \end{pmatrix} = s \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Vi får då egenvektorn x_2 vilket ligger i egenrummet som tillhör egenvärdet $\lambda_2 = 5$.

$$E_5 = Span\left(x_2\right) = Span \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Testa om definitionen stämmer, dvs: $T(x) = \lambda x$:

$\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = (5) \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, vilket den gör.

Vi har därmed hittat två egenvektorer ur vår ursprungsmatris T och observera att de utgör en bas för $\mathbb{R}^2$ och därmed är matrisen T *diagonaliserbar*. Vad menas då med att en matris T är *diagonaliserbar*? Detta går vi igenom i nästa avsnitt. (Friedberg et al. 2014, s. 250 ff.)

Invarianta delrum:

Anta att $T : V \to V$ där $V = U_1 \oplus \cdots \oplus U_k$ där varje U_k är ett delrum till V.

Om U är ett delrum av V och $T : V \to V$ så är U ett invariant delrum av T om vektorn $x \in U$ medför $Tx \in U$. $Tx \in U$ är ekvivalent med att U är invariant under T om $T|_U$ är en operator på delrummet U.

I $T|_U$ betraktar vi T som verkar på vektorer i delrummet U.

I allmänhet skickar $T|_U$ vektorer i U till vektorer i V. Det är om U är invariant som vi har: $T : U \to U$, T skickar element i U till element i U.

Exempel 1:

Om $T : V \to V$ så är nollrummet till T, $N(T)$ ett invariant delrum till T eftersom om vi har: $x \in N(T)$ så får vi att $Tx = 0$ och därav är $Tx \in N(T)$, eftersom $0_x \in N(T)$

Exempel 2:

$[T]_\beta = \begin{pmatrix} 1 & 3 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, där $U = \text{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ är invariant delrum till $[T]_\beta$.

$$\begin{pmatrix} 1 & 3 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2a \end{pmatrix} \in \text{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Exempel 3:

$$[T]_\beta = \begin{pmatrix} 1 & 3 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \text{ där } U_1 = \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \text{ ej är invariant delrum till } [T]_\beta.$$

$$\begin{pmatrix} 1 & 3 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ b \\ 0 \end{pmatrix} = \begin{pmatrix} 3b \\ 5b \\ 0 \end{pmatrix} \notin \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

Obs: $U_2 = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ är invariant delrum till $[T]_\beta$

$$\begin{pmatrix} 1 & 3 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} b \\ c \\ 0 \end{pmatrix} = \begin{pmatrix} b+3c \\ 4b+5c \\ 0 \end{pmatrix} \in \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

Hur beskriver vi då ett invariant delrum av V med dimension 1?
Ett endimensionellt vektorrum U till V består av alla multiplar ax av någon fix vektor $x \neq 0$ i V. Om U invariant, så måste Tx tillhöra U, dvs $Tx = ax$ för något a. Alltså är isåfall x egenvektor till T. (Axler 1997, s. 74 f.)

2.6 Diagonaliserbarhet av matriser

Definition
En matris $T : V \to V$ sägs vara diagonaliserbar om det finns en bas γ i V sådan att matrisen för T i basen γ är en diagonalmatris. Observera att isåfall är vektorerna i γ egenvektorer till T och diagonalelementen i matrisen för T är motsvarande egenvärden. En matris $T : V \to V$ är diagonaliserbar om det finns en inverterbar matris Q sådan att $D = Q^{-1}TQ$ eller $T = QDQ^{-1}$, där D =diag($\lambda_1 ... \lambda_k$). Kolonnerna i Q motsvarar egenvektorerna till T, alltså $[T]_\gamma$ och D motsvarar egenvärdena till $[T]_\gamma$. (Friedberg et al. 2014, ss. 245-269)

Sats 5.5

Egenvektorer hörande till olika egenvärden är linjärt oberoende.

Bevis

Induktion över k. Dim$V = k$. (Friedberg et al. 2014, s.261.)

Följdsats 5.5

Om matrisen $T : V \to V$ har distinkta egenvärden och dessa är lika många som dim$V$ så är T diagonaliserbar.

Bevis:

Matrisen T har k stycken distinka egenvärden $(\lambda_1, \lambda_2, ..., \lambda_k)$. För varje j får vi en egenvektor x_j med tillhörande egenvärde λ_k. Från sats 5.5 ovan så får vi att egenvektorerna $\{x_1, ..., x_j\}$ är linjärt oberoende och vi får även att dim$V$ är lika med antalet linjärt oberoende egenvektorerna. Som därmed bildar en bas i V. Då får vi att T är diagonaliserbar. (Friedberg et al. 2014, s. 261 f.)

Exempel: Vi tar föregående exempel. $T = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$
Med tillhörande egenvärden $\lambda_1 = -2$ och $\lambda_2 = -5$, vi får $D = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}$ och egenvektorerna $\{x_1, x_2\} = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\}$ som utgör en bas till vektorrummet R^2 då de båda är linjärt oberoende egenvektorer till T och bildar $Q = \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix}$. Egenvektorerna och egenvärden uppfyller ekvationen: $T(x) = \lambda x$. (Friedberg et al. 2014, s. 245 ff.)

Geometrisk Multiplicitet

Sats 5.7

Vi har en linjär avbildning $T : V \to V$ som har ett egenvärde λ med multiplicitet k då kommer dimensionen av egenrummet till detta egenvärde ej att överstiga dimensionen på λ:s multiplicitet. $(1 \leq \dim(E_\lambda) \leq k)$. För bevis se sid 264 i Friedberg et al. (2014)

Kapitel 3

Cayley-Hamiltons sats

3.1 Formulera satsen

Cayley-Hamiltons sats:

Låt T vara en kvadratisk matris för en linjär avbildning av ett n-dimensionellt reellt vektorrum V. Vi låter $q(\lambda) = \det(T - \lambda Id)$ beteckna det karateristiska polynomet av T. Det satsen säger är att om vi sätter in T i sin egna karakteristiska ekvation så får vi ut nollavbildningen. Därmed uppfyller T sin egen karakteristiska ekvation, $q(T) = 0$. (Axler Sheldon 1997, s. 195)

Vad är då $q(T)$?

Vi har ett polynom på formen:

$q_j(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, där $a \in \mathbb{R}$,n positivt heltal och graden av detta polynom är $\leq n$

Om vi har en matrismultiplikation, där A, B är $n \times n$-matriser, så vet vi redan vad AB är.
Speciellt vet vi vad $A^2 = AA$ är. Då vet vi vad A^n är, där n är ett positivt heltal.
Vi vet även vad aA är för matris om $a \in \mathbb{R}$.
Det vi gör nu är att applicera polynomet definierat ovan på vår matris T som är på formen $n \times n$, vi får då:

$q_j(T) = a_n T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 Id$, där $a \in \mathbb{R}$, n positivt heltal och graden av detta polynom är $\leq n$

Exempel: Vi har beräknat fram det karakteristiska polynomet: $q(x) = x^2 + 1$ av den kvadratiska matrisen T.

För att beräkna $q(T)$ så ersätter vi x^2 med matrisen $T^2 = TT$. Termen 1 i polynomet tolkas här som identitetsmatrisen $Id = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Exempel:

Vi betraktar en kvadratisk 2×2 matris. För att tydliggöra detta så beräknar vi den karakteristiska ekvationen $q_T(\lambda) = 0$, för att sedan sätta in T på λ:s plats.

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$q_T(\lambda) = \det(T - \lambda Id)$$

$$= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix}$$

Efter beräkning av determinanten så får vi det karakteristiska polynomet:

$$q_T(\lambda) = \lambda^2 - a\lambda - d\lambda + ad - bc$$

Vi sätter in ursprungsmatrisen T på λ:s plats och beräkna vänsterledet för att se om vi får ut nollmatrisen:

$$q_T(T) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 - a \begin{pmatrix} a & b \\ c & d \end{pmatrix} - d \begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Utveckling av matriserna ger oss:

$$\begin{pmatrix} a^2+bc & ab+bd \\ ac+dc & bc+d^2 \end{pmatrix} - \begin{pmatrix} a^2 & ab \\ ac & ad \end{pmatrix} - \begin{pmatrix} ad & bd \\ cd & d^2 \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} =$$
$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Efter räkning med matriserna får vi ut nollmatrisen. Oavsett vad vi har för kvadratisk matris så kommer Cayley-Hamiltons sats att gälla och det ska vi nu bevisa.

3.2 Historisk bakgrund

Vi ska i detta kapitel gå igenom den historiska bakgrunden av de tre huvudaktörerna bakom denna kända sats som heter Cayley-Hamiltons sats.

Sir William Rowan Hamilton (1805-1865)

Vi börjar med den irländske matematikern Sir William Rowan Hamilton. Hamilton föddes år 1805 i Dublin där han växte upp med en pappa som advokat. Han började tidigt att visa ett stort intresse för skolan då han var begåvad och främst för språk och matematik i synnerhet där hans stora intresse var astronomi och matematiken som den innehöll. Han skaffade snart ett eget teleskop för att fortsätta sin utforskning av rymden. Han tittade inte bara på stjärnorna utan han började med att observera planeternas och stjärnornas rörelser där han skrev ner matematiska kurvor och ytor för detta. Det som la grunden till den tidiga matematiska erfarenheten var att Hamilton läste och studerade Sir Isaac Newtons verk, 'Newton's principa'. I den står det mycket om rörelselager, gravitationslagar och himlakroppars rörelser etc. Han läste även Mécanique céleste(1799) som är ett verk som fullständigt går igenom hur planetsystemet är uppbyggt på Newtons hypoteser om gravitationen och mycket mer. Hamilton fick mycket inspiration från dessa två herrar och deras vetenskapliga verk. Vid 18 års ålder började Hamilton på Trinity College i Dublin där han studerade matematik och fick en stabil matematisk grund att stå på. Hamilton blev sedermera professor i astronomi på Trinity College vid 22 års ålder och en Royal Astromer of Ireland vilket betydde att han bl.a blev föreståndare av Dunsinkobservatoriet. Det här var en börja ny upptäcksålder för grunden till algebra. Under 1830-talet var det många professorer vid universtitetet i Cambridge i England som arbetade med nya metoder inom algebra för att rättfärdiga de negativa och imaginära talen och svara mot kritik mot just detta. Under denna tid jobbade Hamilton också mycket med algebra och dess grunder och även han forskning tog stöd från universitetet av Cambridge.

Hamilton blev känd och mest ihågkommen för sina matematiska resultat inom den matematiska fysiken och den geometriska optiken och dynamiken. Hamiltons bidrag är att han kom på kvaternioner vilket ingår i den matematiska fysiken och även abstrakt algebra, vilket han själv trodde skulle bli hans viktigaste gärning inom den matematiska grenen. Upptäcken banade vägen till en bok om ämnet med namnet 'Lectures on quaternions' (1853). Han hade rätt då det gav de komplexa siffrorna legitimitet. Dock så uppfyllde inte hans formel för kvaternioner den kommutativa lagen för multiplikation av vanlig aritmetik för reella tal. Hans upptäckt av denna formel som är av mer abstrakt algebra banade vägen för dåtidens matematiker i mitten av 1800-talets med nya upptäckter såsom: oktonion, Clifford-algebra, vektorrum, linjära associativa algebror m.m. av andra matematiker vid namnen: Cayley, Sylvester, Clifford, Gibbs och många fler.

Upptäckterna av de här nya icke-standardiserade algebraiska formerna gjorde upphov till att den icke-euklidiska geometrin tog form. Så deras arbete med algebran gjorde stor påverkan på matematiken. Hamilton presenterade 1827 den matematiska forskningen 'Theory of a system of rays', vilket är ett vetenskapligt arbete om optik och introducerar de centrala delarna av den matematiska fysiken och den karakteristiska funktionen på optiska system. Hamilton tyckte själv att det var av stor vikt att uttrycka en funktion så enkelt som möjligt vilket gör den kraftfullt och attraktiv. Ekvationen tyckte han själv utgjorde behovet av den matematikens optik. Under åren 1830-1832 publicerade Hamilton ytterligare tre stycken nya utgåvor till 'Theory of a system of rays'. Han använde sig av samma idéer när han forskade om dynamiska system och dess rörelsesystem i konfigurationsutrymmet. Hamiltons arbete inom optiken och dynamiken inspirerade De Broglie och Schrödinger att komma på och formulera vågmekanisk formation vilket var revolutionerande inom kvantmekaniken.

Under dessa år var Hamilton väldigt fokuserad på sina studier och upptäckter inom optiken men matematikern Graves tog hjälp av Hamilton och försökte skapa logaritmer av de negativa och komplexa talen. Eftersom Hamilton hade invändningar mot just den tidens användning av de komplexa och negativa talen så gick Hamilton igång på det och började arbeta. Han fick läsa om Argands arbete om det komplexa talplanet. Argands arbete blev intressant för Hamilton för att Argand representerade de komplexa talen som punkter i talplanet. Enligt tolkning så skulle man då kunna utföra elementära operationer om man ser de komplexa talen som vektorer i planet. Detta gjorde Hamilton nyfiken att fortsätta sina studier om detta och den tredimensionella rymden i talplanet. För om de komplexa talen nu kan generaliseras och definieras så innebär det också att vektorerna kan förskjutas och roteras i den tredimensionella rymden.

Hamilton hoppades även här på att få fram ett kraftfullt redskap just som han gjorde med den karakteristiska funktionen för optiken för att beskriva rörelserna av olika kroppar i rymden då det var ett av hans stora intressen. Rörelserna kropparna utför i rymden kan beskrivas som de spatiala vektorerna som beskriver acceleration, tröghet m.m. Hamilton jobbade väldigt länge med det här, då han till slut kom på lösningen då han gick över bron som går över Royal Canal i Dublin år 1843: Formeln nedan hämtad från: Mactutor (1998):

$$i^2 = j^2 = k^2 = ijk = -1$$

Hamilton hann med mycket under sina år som matematiker och han var först med att behandla komplexa tal i ordnade par av reella tal och få det publicerat. Hamilton sysslade inte endast med matematik under sin levnadsperiod utan han var även engagerad i framförallt poesi under sin fritid vid sidan om matematiken. Hamiltons arbete om bl.a. den tredimensionella rymden av de komplexa talen och de algebraiska uttrycken ledde till upptäckten av den icke-kommunikativa lagarna. De viktigaste publicerade handlingarna som Hamilton har gjort finns i tre utgåvor mellan 1831-1867 och finns förvarade i sin hemstad Dublin där han gick i college på Trinitys bibliotek och även på Irlands nationella bibliotek. (Ewald. 1996, ss. 362-369)

Arthur Cayley (1821-1895)

Arthur Cayley studerade på Trinity College i Cambridge. Cayley fick mottaga ett stipendium för sina insatser där men valde senare den juridiska vägen och började studera till advokat på Lincolns Inn i London 1849. Cayley stannade där till 1863.

Under dessa år producerade Cayley närmare 300 forskningspapper om matematik och framförallt teorin om invarianter. Han jobbade nära en annan känd matematiker i algebra vid namn (James Joseph Sylvester) i dessa forskningsfrågor. Cayleys forskning var väldigt stor och betydande för den matematiska historien då han totalt fyllde 13 böcker med forskning om matematiska områden inom främst algebra såsom: den icke-euklidiska geometrin, representation av komplexa tal, Riemannytor, teorin om invarianter, n-dimensionell geometri, teorin av matriser, determinanter, algebraisk geometri, teorin om grupper. (Ewald. 1996, s. 542 ff.)

T.ex. så publicerade Cayley 'On a theory of determinants' år 1843 där han utvidgade idéen om en 2-dimensionell determinant till en determinant av godtycklig storlek n. År 1844 reste även Cayley runt med sin vän Edmund Venables där de bland annat besökte de Schweiziska alperna, Italien och Frankrike för att få internationell erfarenhet och bredda sina matematiska kunskaper. Cayley hann även under sina utlandsresor få sina forskning publicerad i bland annat 'Journal de Mathématiques Pures et Appliquées' och i 'Journal für die reine und angewandte Mathematik'. Forskningen Cayley fick publicerad under åren 1845-1846 är till stor del grunden för den invarianta teorin. Eftersom Cambridge stipendiet innebar begränsad tjänstgöring, så var det inget alternativ för Cayley utan han valde då att studera till advokat på Lincolns inn vilket nämnts tidigare i texten. Under åren som Cayley utbildade sig till advokat så åkte han till Dublin för att höra när William Rowan Hamilton föreläste om kvaternioner. Där träffade han George Salmon som han utbytte idéer om matematik under många år framöver. Cayley blev även en god vän till Hamilton även om de var oeniga om kvaternionernas betydelse för geometrin. Cayley's arbete där han utvecklade bland annat teorion om invarians och den n-dimensionella geometrin har varit betydelsefull för fysiken där den har använts för att studera rumtid och kontinuum. Kvantmekaniken fick anvädning av Cayleys arbete om matriser då de låg som grunden till ett arbete om kvantmekanik som Werner Heisenberg gjorde år 1925. Cayley kom också på att euklidisk- och icke-euklidisk geometri är specifika typer av geometri och så för han även samman metrisk- och projektiv geometri då de är beroende av storlekarna på vinklarna och längderna av linjerna. Mactutor (2014)

År 1858 publicerade Cayley 'Memoir of the theory of matrices' och den innehåller den första definitionen av en matris som är abstrakt. I denna forskningen definierar Cayley addition-, multiplikations-, multiplikation med skalär samt invers av matriser. Cayley gav en väldigt uttrycklig definition av inversen till en matris i termer av determinanten av matrisen. Sedan gav Cayley beviset på att när du har en kvadratisk matris av storlek 2×2 så uppfyller den sin egen karakteristiska ekvation vilket sedermera också fick ett eget namn nämligen Cayley-Hamiltons Sats. Varför heter den då även Hamilton i satsen? Hamilton undersökte och forskade om kvaternioner och då bevisade han det i 4×4-fallet. Mactutor (1996)

Så han berörde flera av de andra stora matematikernas forskningsfrågor och bland annat William Rowan Hamiltons forskning om framförallt algebra. Cayley återvände sedan dit allting startade. Alltså till Trinity College i Cambridge där han arbetade som professor i matematik från år 1863 till sin död. (Ewald. 1996, s. 542 ff.)

Ferdinand Georg Frobenius (1849-1917)

Varken Hamilton eller Cayley gav det allmänna beviset för Cayley-Hamiltons sats i $n \times n$ -fallet, utan det var en tysk matematiker vid namn Georg Frobenius (1849-1917) som gjorde det år 1878. Mactutor (1996)
Frobenius växte upp i Berlin i Tyskland, där han även tog sin doktorsexamen 1870. Han gick på seminarier som leddes av bland annat Weierstrass som ansåg att Frobenius var väldigt begåvad. År 1874 utnämndes Frobenius till matematikprofessor av universitet i Berlin. Mellan åren 1875-1892 bodde Frobenius i Zurich där han var professor vid Eidgenössische Polytechnikum. Sedan blev ordförandeposten på universitetet för matematiken i Berlin ledig och Frobenius antogs som ny ordförande med hjälp ifrån bland annat Weierstrass. Frobenius valdes för att hans arbeten i matematik hade väldigt hög kvalite och att det finns 15 ämnen som Frobenius hade givit stora bidrag till redan då, bland annat: Teorin om linjära differentiella ekvationer, algebraiska lösningar av ekvationer, vars koefficienter är rationella funktioner för en variabel, linjära former med heltalskoefficienter m.m listan kan göras lång. Frobenius var ordförande i Berlin i 25 år. Mactutor (2000)

3.3 Bevis

Vi väljer en bas β för V enligt sats 9.4 nedan så att $[T]_\beta$ bildar en övertriangulär blockmatris, se 9.18. Vi har att blockmatriserna $(A_1, ..., A_j)$ längs med diagonalen av $[T]_\beta$ antingen är 1×1-matris eller en 2×2-matris utan reella egenvärden. Där alla element under $(A_1, ..., A_j)$ är lika med noll och elementen ovan dem är godtyckliga element.

9.18

$$[T]_\beta = \begin{pmatrix} A_1 & * & * \\ \vdots & \ddots & * \\ 0 & \cdots & A_j \end{pmatrix}$$

Vi behöver sedan använda oss utav tre satser: 9.1, 9.4 och 9.19 för att ta oss vidare i beviset.

Sats 9.1

Antag att $T : V \to V$ och att A är matrisen av T med avseende på någon bas av V, då är egenvärdena till T samma som egenvärdena till A.

Bevis

Låt $\{x_1, ..., x_n\}$ vara en bas för V så att $A = [T]_\beta$ är matrisen för T relativt β. Låt $\lambda \in \mathbb{R}$.

Vi behöver visa att λ är egenvärde till T om och endast om λ är egenvärde till matrisen A. Vi antar då först att λ är egenvärde till T.

Låt $x \in V$, så att $Tx = \lambda x$, där $x \neq \vec{0}$
Vi skriver $x = c_1x_1 + \cdots + c_nx_n$, där $(c_1, \cdots, c_n \in \mathbb{R})$

Vi låter $[x]_\beta$ representera koordinaterna av vektorn x med hänsyn till basen $\beta = \{x_1, \cdots, x_n\}$.

$$[x]_\beta = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

vi har då att:

$$A[x]_\beta = [T]_\beta [x]_\beta = [Tx]_\beta = [\lambda x]_\beta = \lambda [x]_\beta$$

Detta visar att λ är egenvärde till A, där den andra likheten kommer från sats 3.12 (Axler 1997, s. 52)

Anta nu istället att λ är egenvärde till A. Vi låter $[x]_\beta$ representera koordinaterna av vektorn x med avseende på basen $\beta = \{x_1, \cdots, x_n\}$.

$$[x]_\beta = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Vi låter $[x]_\beta$ vara vara en nollskild $n \times 1$-matris så att: $A[x]_\beta = \lambda [x]_\beta$

Vi definierar x som ovan alltså: $x = c_1 x_1 + \cdots + c_n x_n$. Vi får då:

$$[Tx]_\beta = [T]_\beta [x]_\beta = A[x]_\beta = \lambda [x]_\beta = [\lambda x]_\beta$$

Vilket gör att vi får $Tx = \lambda x$ där λ är egenvärde till $[T]_\beta$ och beviset är klart och egenvärdena till $[T]_\beta$ är detsamma som egenvärdena till matrisen A. (Axler 1997, s. 182 f.)

Sats 9.4

Låt $T : V \to V$ linjär avbildning, $\dim V = n$. Då finns en bas β så att $[T]_\beta$ bildar en övertriangulär blockmatris enligt matrisen T nedan

$$[T]_\beta = \begin{pmatrix} A_1 & * & * \\ & \ddots & * \\ 0 & & A_j \end{pmatrix}$$

Bevis

Fall 1 Om dim$V = 1$, så finns det ett fall:

$T(x) = ax$, för något $a \in \mathbb{R}$. Eftersom en linjära avbildning på $\mathbb{R}$ är multiplikation med fixt tal.

Om vi har bas x_1: får vi: $T(x_1) = ax_1$ så
$[T]_\beta = (a)$ alltså ett 1×1-block.

Fall 2: Om dim$V = 2$, så finns det två fall:

a) egenvärde finns
b) egenvärde saknas

Fall a: Antag att $x_1 \neq \vec{0}$ så att $T(x_1) = ax_1$

Vi vill nu bilda en bas β så att $\beta = \{x_1, x_2\}$, där x_1 och x_2 är linjärt oberoende.

$T(x_1) = ax_1 + 0 \cdot x_2$ enligt antagandet.
$T(x_2) = bx_1 + dx_2$, så vi får:

$$\Big[T\Big]_\beta = ([T(x_1)]_\beta | [T(x_2)]_\beta) = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

Det karakteristiska polynomet i fall a) blir således
$[T]_\beta = det([T]_\beta - \lambda[Id]_\beta) = \det \begin{pmatrix} a-\lambda & b \\ 0 & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda)$, dvs två egenvärden.

$[T]_\beta = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$, där $A_1 = a$ och $A_2 = d$ och $*$ är ett godtyckligt element. Vi ser även att vi får två stycken 1×1-block.

Fall b: Antag att egenvärden saknas, välj därför någon bas $\beta = \{x_1, x_2\}$

$$[T]_\beta = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (A_1)$$

Exempel på matris utan reella egenvärden: $A_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
Om nu T inte har egenvärden så kan vi välja en bas $\beta = (x_1, x_2)$ till V så att $[T]_\beta$ ej har några egenvärden enligt sats 9.1

Fall 3 Om dim$V = 3$, så finns det två fall.

a) Egenvärden finns där i har 3 stycken 1×1-block.
b) Ett styck 1×1-block med egenvärde samt ett styck 2×2-block utan egenvärde. (Axler 1997, s. 184 ff.)

Fall a:

Vi vill nu bilda en bas β så att $\beta = \{x_1, x_2, x_3\}$, där x_1, x_2 och x_3 är linjärt oberoende.

$$T(x_1) = ax_1 + 0 \cdot x_2 + 0 \cdot x_3$$

$$T(x_2) = bx_1 + ex_2 + 0 \cdot x_3$$

$$T(x_3) = cx_1 + fx_2 + ix_3$$

$$\Big[T\Big]_\beta = ([T(x_1)]_\beta | [T(x_2)]_\beta | [T(x_3)]_\beta) = \begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & i \end{pmatrix}$$

Det karakteristiska polynomet blir:
$[T]_\beta = det([T]_\beta - \lambda[Id]_\beta) = \det \begin{pmatrix} a-\lambda & b & c \\ 0 & e-\lambda & f \\ 0 & 0 & i-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda)(i-\lambda)$,
dvs tre egenvärden.

$$[T]_\beta = \begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & i \end{pmatrix} = \begin{pmatrix} A_1 & * & * \\ 0 & A_2 & * \\ 0 & 0 & A_3 \end{pmatrix}, \text{ där } A_1 = (a),\ A_2 = (e) \text{ och } A_3 = (i).$$
Vi får tre stycken 1×1-block.

Fall b

Vi har ett 1×1-block med egenvärde samt ett 2×2-block utan egenvärden. Vi får att $V = U_1 \oplus U_2$, där U_1 är av dimension 1 och har egenvärde och U_2 är av dimension 2 och är utan egenvärde. Både delrummen U_1 och U_2 är invariant under $[T]_\beta$. Om U_1 invariant sågäller det att vektorn $\begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix} \in [T]_\beta$ och om U_2 invariant så gäller det att vektorn på linjärkombinationen $\begin{pmatrix} 0 \\ s \\ t \end{pmatrix} \in [T]_\beta$.

$[T]_\beta = \begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix} = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$, där $A_1 = (a)$, $A_2 = \begin{pmatrix} e & f \\ h & i \end{pmatrix}$. Vi får ett 1×1-block med egenvärde och ett 2×2-block utan egenvärden enligt sats 9.1. Sen kan vi självklart byta plats på dem två blocken. (Axler 1997, s. 182 ff.)

Sats 9.19:

Det karakteristiska polynomet av en blockmatris $[T]_\beta$ är produkten av blockens $\{A_1, ..., A_j\}$ karakteristiska polynom om blocken har storlek 1×1 eller 2×2.

Explicit:

$$q_j(x) = \begin{cases} (x-a), \quad \text{om} \quad A_j = 1 \times 1\text{matris } \big(a\big) \\ (x-a)(x-d) - bc, \quad \text{om} \quad A_j = 2 \times 2\text{matris} \quad = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{cases}$$

Det karakteristiska polynomet till den övertriangulära blockmatrisen $[T]_\beta$ blir således: $q_1(x) \cdots q_j(x)$, där graden av det karakteristiska polynomet till $[T]_\beta$ är lika med dimensionen av vektorrummet V.

Rötterna till det karakteristiska polynomet av $[T]_\beta$ blir således T:s egenvärden.

Vi ska nu visa att $q(T) = 0$ för ett V med dimension $1, 2$ eller 3.

Det räcker då med att visa att: $q_1(T) \cdots q_j(T)|U_j = 0$ (för $j = 1, ..., 3$), dvs produkten av de karakteristiska polynomen $\{A_1, \cdots, A_j\}$ längs med diagonalen av $[T]_\beta$. U_j är här de enskilda invarianta delrummen till V.

Vi har att de karakteristiska polynomen $q = q_1 \cdots q_j$ enligt sats 9.19 ovan som verkar på $V = U_1 \oplus \cdots \oplus U_j$, där vi har att varje delrum till V betecknas som U_j enligt matrisen 9.18 ovan.

Fall 1: dimV=1

Här finns det endast 1 stycken fall då vi har en 1×1-matris.

Vi ska nu visa att: $q_1(T)|U_1 = 0$

Vi har matrisen $[T]_\beta = (a)$, där U_1 är invariant delrum. Observera att på grund av matrisen blockform räcker det med att undersöka vektorn på formen (s).

$$[T]_\beta x = (a)(s)$$

Vi får det karakteristiska polynomet:

$$q_1(\lambda) = a - \lambda$$

$$q_1(\lambda)|U_1 = (a - \lambda)(s) = ((a - \lambda)s)$$

Sätt in $\lambda = a$

$q_1(a)|U_1 = (a - a)(s) = ((a - a)s)$= (0), vi får ut nollmatrisen.
Alltså är $q_1(T)|U_1 = 0$

Fall 2: dim V=2

$$[T]_\beta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Här finns det två stycken fall som vi ska undersöka.

Fall a: Egenvärden saknas

Vi ska visa att: $q_1(T)|U_1 = 0$
Vi har matrisen $[T]_\beta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ utan egenvärden. Där $\dim U_1 = 2$.

Detta fall har vi redan visat att det blir nollmatrisen i början av Cayley-Hamiltons sats.
Vilket även det ger oss 0 och uppfyller sats 9.21, alltså $q_1(T)|U_1 = 0$.

Fall b: Egenvärden finns

Antag vi har matrisen: $[T]_\beta = \begin{pmatrix} a & * \\ 0 & d \end{pmatrix}$ med två egenvärden.
Där $\dim U_1 = 1$ och $\dim U_2 = 1$. Alltså två stycken 1×1-block.

Vi ska nu visa att: $q_1(T)q_2(T)|U_2 = 0$

$$q_1(\lambda) = (a - \lambda)$$
$$q_2(\lambda) = (d - \lambda)$$

Vi undersöker först $q_1([T]_\beta)$:

$$q_1([T]_\beta) = \left(a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - [T]_\beta \right) = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} - \begin{pmatrix} a & * \\ 0 & d \end{pmatrix} = \begin{pmatrix} 0 & * \\ 0 & a-d \end{pmatrix}$$

Sedan undersöker vi $q_1(T)|U_1$. Observera att på grund av matrisens blockform räcker det med att undersöka vektorer på formen $\begin{pmatrix} s \\ 0 \end{pmatrix}$
Vi får $q_1([T]_\beta)x = q_1([T]_\beta) \begin{pmatrix} s \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & * \\ 0 & a-d \end{pmatrix} \begin{pmatrix} s \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Slutsats: $q_1([T]_\beta)|U_1 = 0$, nolloperatorn på U_1.

Vi undersöker sedan $q_1([T]_\beta)q_2([T]_\beta)|U_2$

$$q_1(\lambda)q_2(\lambda) = (a-\lambda)(d-\lambda)$$
$$q_2([T]_\beta) = \begin{pmatrix} d-a & * \\ 0 & 0 \end{pmatrix}$$

Vi kan se att $q_1([T]_\beta)q_2([T]_\beta) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Vi undersöker $q_1([T]_\beta)q_2([T]_\beta)|U_2$. Observera att på grund av matrisens blockform räcker det med att undersöka vektorer på formen $\begin{pmatrix} 0 \\ s \end{pmatrix}$

Vi får $q_1([T]_\beta)q_2([T]_\beta)x = \begin{pmatrix} 0 & * \\ 0 & a-d \end{pmatrix} \begin{pmatrix} d-a & * \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ s \end{pmatrix}$

Vi undersöker dom två matriserna till höger:

$$q_2([T]_\beta)|U_2 = \begin{pmatrix} d-a & * \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Observera att $*$ måste vara nollelementet. Slutsats: $q_1(T)q_2(T)|U_2 = 0$

Fall 3: dim V=3

Här har vi fallet då $\dim V = 3$, alltså de olika 3×3-fallen:

$$[T]_\beta = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

Fall a:

Vi har tre stycken 1×1-block. $A_1 = (a)$, $A_2 = (e)$ och $A_3 = (i)$.

Exempel

$A_1 = (a)$ och $A_2 = (e)$ och $A_3 = (i)$ vi får $[T]_\beta = \begin{pmatrix} A_1 & * & * \\ 0 & A_2 & * \\ 0 & 0 & A_3 \end{pmatrix}$, där * är ospecifierade element.

Därmed har vi tre egenvärden längs med diagonalen av matrisen och således tre stycken 1×1-block som vi har visat ovan nämligen A_1, A_2 och A_3. Där vi har delrummen $\dim U_1 = 1$, $\dim U_2 = 1$ och $\dim U_3 = 1$. Alltså $V = U_1 \oplus U_2 \oplus U_3$.

Idén: Det räcker med att visa att: $q_1(T)q_2(T)q_3(T)|U_3 = 0$. Eftersom det karakteristiska polynomet q är en produkt av polynomen q_1,q_2 och q_3 och $\mathbb{R}^3$ ges av summan av de invarianta delrummen $U_1 \oplus U_2 \oplus U_3$ så räcker det med att undersöka varje enskilt delrum och om det blir nolloperatorn så har vi bevisat det.

Vi tittar först på fallet då: $q_1(T)|U_1$

$$q_1(\lambda) = (a - \lambda)$$

$$q_1([T]_\beta) = \left(a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - [T]_\beta \right) = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} - \begin{pmatrix} a & * & * \\ 0 & e & * \\ 0 & 0 & i \end{pmatrix} = \begin{pmatrix} 0 & * & * \\ 0 & a-e & * \\ 0 & 0 & a-i \end{pmatrix}$$

Vi undersöker sedan: $q_1(T)_\beta|U_1$. Observera att på grund av matrisens blockform räcker det med att undersöka vektorer på formen $\begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix}$

$$q_1([T]_\beta)x = \begin{pmatrix} 0 & * & * \\ 0 & a-e & * \\ 0 & 0 & a-i \end{pmatrix} \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Slutsats: $q_1(T)|U_1 = 0$, nolloperatorn.

Vi tittar sedan på fallet då: $q_2(T)|U_2$

$$q_2(\lambda) = (e - \lambda)$$

$$q_2([T]_\beta) = \left(e \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - [T]_\beta \right) = \begin{pmatrix} e & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & e \end{pmatrix} - \begin{pmatrix} a & * & * \\ 0 & e & * \\ 0 & 0 & i \end{pmatrix} = \begin{pmatrix} e-a & * & * \\ 0 & 0 & * \\ 0 & 0 & e-i \end{pmatrix}$$

Vi undersöker sedan: $q_2(T)_\beta|U_2$. Observera att på grund av matrisens block-form räcker det med att undersöka vektorer på formen $\begin{pmatrix} 0 \\ s \\ 0 \end{pmatrix}$

$$q_1([T]_\beta)x = \begin{pmatrix} e-a & * & * \\ 0 & 0 & * \\ 0 & 0 & e-i \end{pmatrix} \begin{pmatrix} 0 \\ s \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Observera att $*$-elementet på rad 1 måste vara nollelementet. Slutsats: $q_2(T)|U_2 = 0$. Nolloperatorn.

Vi kan här se att:

$$q_1(T)q_2(T) = \begin{pmatrix} 0 & * & * \\ 0 & a-e & * \\ 0 & 0 & a-i \end{pmatrix} \begin{pmatrix} e-a & * & * \\ 0 & 0 & * \\ 0 & 0 & e-i \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (a-i)(e-i) \end{pmatrix}$$

Vi får därmed att $q_1(T)q_2(T)|U_1 = 0$ och $q_1(T)q_2(T)|U_2 = 0$.

Vi tittar sedan på fallet då: $q_3(T)|U_3$

$$q_3(\lambda) = (i - \lambda)$$

$$q_3([T]_\beta) = \left(i \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - [T]_\beta \right) = \begin{pmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{pmatrix} - \begin{pmatrix} a & * & * \\ 0 & e & * \\ 0 & 0 & i \end{pmatrix} = \begin{pmatrix} i-a & * & * \\ 0 & i-e & * \\ 0 & 0 & 0 \end{pmatrix}$$

Vi undersöker sedan: $q_3(T)_\beta|U_3$. Observera att på grund av matrisens block-form räcker det med att undersöka vektorer på formen $\begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix}$

$$q_3([T]_\beta)x = \begin{pmatrix} i-a & * & * \\ 0 & i-e & * \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Observera att $*$-elementen i kolonn 3 måste vara nollelementet. Slutsats: $q_3(T)|U_3 = 0$. Nolloperatorn.

Vi undersöker till sist fallet då $q_1(T)q_2(T)q_3(T)|U_3$

$$q_1(\lambda)q_2(\lambda)q_3(\lambda) = (a-\lambda)(e-\lambda)(i-\lambda)$$

$q_1([T]_\beta)q_2([T]_\beta)q_3([T]_\beta) =$

$$= \begin{pmatrix} 0 & * & * \\ 0 & a-e & * \\ 0 & 0 & a-i \end{pmatrix} \begin{pmatrix} e-a & * & * \\ 0 & 0 & * \\ 0 & 0 & e-i \end{pmatrix} \begin{pmatrix} i-a & * & * \\ 0 & i-e & * \\ 0 & 0 & 0 \end{pmatrix}$$

Vi undersöker nu $q_1([T]_\beta)q_2([T]_\beta)q_3([T]_\beta)|U_3$. Observera att på grund av matrisens blockform räcker det med att undersöka vektorer på formen: $\begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix}$

$q_1([T]_\beta)q_2([T]_\beta)q_3([T]_\beta)x =$

$$= \begin{pmatrix} 0 & * & * \\ 0 & a-e & * \\ 0 & 0 & a-i \end{pmatrix} \begin{pmatrix} e-a & * & * \\ 0 & 0 & * \\ 0 & 0 & e-i \end{pmatrix} \begin{pmatrix} i-a & * & * \\ 0 & i-a & * \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

det räcker med att beräkna de två matriserna till höger och vi får ut nollmatrisen. Observera att $*$-elementen i kolonn 3 då måste vara nollelementen.

Slutsats: Vi har undersökt varje bit av matrisen $[T]_\beta$ som verkar på de enskilda delrummen (U_1, U_2, U_3) och vi får att: $q_1(T)q_2(T)q_3(T)|U_3 = 0$.

Fall b:

$$\begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix}$$

Vi har att $A_1 = (a)$ alltså ett 1×1-block med egenvärde och A_2 som då är ett 2×2-block utan egenvärde. Vi kan även skifta plats så att A_1 istället är på A_2:s plats vice versa, det blir samma fall.

$A_1 = (a)$ och $A_2 = \begin{pmatrix} e & f \\ h & i \end{pmatrix}$.Vi får $[T]_\beta = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix}$, där
* är ospecifierade element.

Där vi har delrummen $\dim U_1 = 1$ och $\dim U_2 = 2$. Alltså $V = U_1 \oplus U_2$.

Idén: Det räcker med att visa att $q_1(T)q_2(T)|U_2 = 0$. Eftersom det karakteristiska polynomet q ges av produkten av polynomen q_1 och q_2 och $\mathbb{R}^3$ ges av de invarianta delrummen $U_1 \oplus U_2$ så räcker det med att undersöka varje enskilt delrum och om det blir nolloperatorn har vi bevisat det.

Vi undersöker först fallet: $q_1(T)|U_1$

$$q_1(\lambda) = (a - \lambda)$$

$$q_1([T]_\beta) = \left(a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - [T]_\beta \right) = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} - \begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a-e & -f \\ 0 & -h & a-i \end{pmatrix}$$

Vi kollar på: $q_1([T]_\beta)|U_1$. Observera att på grund av matrisens blockform räcker det med att undersöka vektorer på formen: $\begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix}$

$q_1([T]_\beta)x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix} \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Vi får därmed att: $q_1(T)|U_1 = 0$.

Vi undersöker sedan fallet då: $q_2(T)|U_2$

$$q_2(\lambda) = \det \begin{pmatrix} e-\lambda & f \\ h & i-\lambda \end{pmatrix} = ei - e\lambda - i\lambda + \lambda^2 - fh$$

$q_2([T]_\beta) = ei \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - e \begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix} - i \begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix} + \begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{pmatrix}^2 - fh \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} (ei - ae - ai + a^2 - fh) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, observera här att vi fick användning av 2×2-fallet utan egenvärden från ursprungsfallet då vi får ut nollmatrisen.

Vi ska nu undersöka $q_2(T)|U_2$. Observera att på grund av matrisens blockform räcker det med att undersöka vektorer på formen av linjärkombinationen: $\begin{pmatrix}0\\s\\0\end{pmatrix}+\begin{pmatrix}0\\0\\t\end{pmatrix}$

$$q_2([T]_\beta)x = \begin{pmatrix}(ei-ae-ai+a^2-fh) & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}0\\s\\t\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

Vi har nu sett att: $q_2(T)|U_2 = 0$

Slutsats: $q_1(T)q_2(T)|U_2 = 0$ då vi har undersökt båda fallen ($q_1(T)|U_1$ och $q_2(T)|U_2$) och konstaterat att de båda blir nolloperatorn.

Vi får ej glömma att vi kan byta plats på A_1 och A_2 så att vi istället får $\dim U_1 = 2$ och $\dim U_2 = 1$.

Vi har nu visat att:

$$q_1(T)|U_1 = 0$$
$$q_1(T)q_2(T)|U_2 = 0$$
$$q_1(T)q_2(T)q_3(T)|U_3 = 0$$

Denna metod kan vi sedan fortsätta med oavsett hur stor den kvadratiska matrisen är. Det kan man bland annat göra med hjälp av induktion på j:

Antag:

$$q_1(T)|U_1 = 0$$
$$q_1(T)q_2(T)|U_2 = 0$$
$$\vdots$$
$$q_1(T)_{j-1}(T)|U_{j-1} = 0$$

Detta visar att: $q_1(T) \cdots q_j(T)|U_j$ följer.

(Axler 1997, ss. 182-200.)

Kapitel 4

Litteratur

Axler, Sheldon Jay (1997). *Linear algebra done right.* 2. ed. New York: Springer

Bøgvad, Rikard & Vaderlind, Paul (2017).*Linjär algebra: grundkurs.* Upplaga 1 Stockholm: Liber

Ewald, W.B. (red.) (2000[1996]). *From Kant to Hilbert: a source book in the foundations of mathematics.* Vol. 1. Oxford: Clarendon.

Friedberg, Stephen H, Insel, Arnold J Spence, Lawrence E. (2014). *Linear-Linear Algebra: Pearson New International Edition* [Elektronisk resurs]. Pearson

Mactutor (1996), Hämtad 2019-11-19 från:
http://mathshistory.st-andrews.ac.uk/HistTopics/Matrices_and_determinants.html

Mactutor (2014), Hämtad 2019-11-19 från:
http://mathshistory.st-andrews.ac.uk/Biographies/Cayley.html

Mactutor (1998), Hämtad 2019-11-19 från:
http://mathshistory.st-andrews.ac.uk/Biographies/Hamilton.html

Mactutor (2000), Hämtad 2020-03-12 från:
http://mathshistory.st-andrews.ac.uk/Biographies/Frobenius.html