



SJÄLVSTÄNDIGT ARBETE I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

En studie om minimalpolynom av linjära operatorer och Jordans
normalform

av

Ali Hassan Farhadi

2025 - No L3

En studie om minimalpolynom av linjära operatorer och Jordans normalform

Ali Hassan Farhadi

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2025

Förord

Jag vill rikta ett stort tack till min handledare, Yishao Zhou, för stödet, den professionella vägledningen och de värdefulla synpunkterna som har gjort mitt självständiga arbete möjligt.

Abstract

Linear algebra is used in various fields, especially in computational data processing. Linear transformations which is central to the subject, allow problems to be solved using matrices and vectors. This thesis explores linear transformations through minimal polynomials and Sheldon Axler's theorem which claims that eigenvalues are more naturally treated without determinants. We discuss minimal polynomials, their relation to eigenvalues, eigenvectors, characteristic polynomials, and Jordans normalform. The thesis begins with key concepts from basic linear algebra, including eigenvalues, eigenvectors, and diagonalization. It introduces generalized eigenvectors to study non-diagonalizable operators and proves eigenvalues exist without using determinants. The main focus is on minimal polynomials, including the Cayley–Hamilton theorem, which relates the characteristic polynomial to an annihilating polynomial. We prove the uniqueness of minimal polynomials and their connection to Jordans normalform. The thesis also compares eigenvalue treatments with and without determinants and offering an alternative method for finding eigenvalues. Finally, Jordans normalform is used in solving ordinary differential equations.

Sammanfattning

Linjär algebra används inom många områden, särskilt inom databehandling. Linjära transformationer som är det centrala gör det möjligt att lösa problem med hjälp av matriser och vektorer. Denna uppsats utforskar linjära transformationer genom minimalpolynom och undersöker Sheldon Axlers påstående att egenvärden behandlas mer naturligt utan determinanter. Vi diskuterar minimalpolynom, deras relation till egenvärden, egenvektorer, karakteristiska polynom och Jordans normalform. Uppsatsen börjar med viktiga begrepp från grundläggande linjär algebra, såsom egenvärden, egenvektorer och diagonalisering. Den introducerar generaliserade egenvektorer för att studera icke-diagonaliserbara operatorer och bevisar att egenvärden existerar utan att använda determinanter. Huvudfokus ligger på minimalpolynom, inklusive Cayley-Hamiltons sats, som relaterar det karakteristiska polynomet till ett annihilande polynom. Vi bevisar minimalpolynomens unikheter och deras koppling till Jordans normalform. Uppsatsen jämför även beräkningen av egenvärden med och utan determinanter och erbjuder en alternativ metod för att hitta egenvärden. Slutligen används Jordans normalform för att lösa ordinära differentialekvationer.

Innehåll

1	Introduktion	8
2	Invarianta delrum och direkt summa	9
2.1	Invariant delrum	10
2.2	Egenrum och diagonalisering av linjära avbildningar	13
3	Den generalized egenrumsdecompositionen	14
4	Triangulära matriser och Caylay-Hamiltons Sats	18
5	Minimalpolynom	20
5.1	Annihilerande polynom	20
5.2	Entydighet av minimalpolynom	21
5.3	Relation till det karakteristiska polynomet	22
5.4	Hur ser minimalpolynom ut?	22
6	Karakteristiska polynom och determinanter	24
7	Beräkna egenvärden utan att beräkna determinanten	27
8	Jordans normalform	31
9	Tillämpning i ODE	38
10	Diskussion	42
	Referenser	42

1 Introduktion

Linjär algebra tillämpas i många matematiska grenar och är onekligen ett av de mest tillämpade matematiska ämnen i praktiska tillämpningar exempelvis i databeräkningar. Linjära avbildningar är ett centralt begrepp i ämnet. Genom dessa kan ett analytiskt problem beräknas med matriser och vektorer.

I detta arbete försöker vi få en djupare förståelse på strukturer hos en linjär avbildning genom minimalpolynom. Ett annat mål är att förstå varför Sheldon Axler i sin bok [1] hävdar att en determinant-fri behandling av egenvärden är mer naturlig än den traditionella metod som involverar determinanter. Vi behandlar minimalpolynom och deras samband med egenvärden, egenvektorer, det karakteristiska polynomet och Jordans normalform. Uppsatsen börjar med repetition av vissa begrepp från den grundläggande kursen i linjär algebra inom Matematik II, t ex linjära operatorer, egenvärden och egenvektorer, invariant delrum, egenrum och dekomposition av linjära operatorer där den linjära operatorn är diagonaliserbar. Vi visar också att egenvärdens existens kan bevisas utan determinant. Vidare introducerar vi generaliserade egenvektorer och generaliserade egenrum för att söka svar till dekomposition av linjära operatorer som inte är diagonaliserbara. Sedan ger vi oss in i det centrala ämnet vi planerade att studera, minimalpolynom. Vi börjar med att ge en definition av annihilande polynom, följt av en genomgång av Cayley–Hamiltons sats, som visar att det karakteristiska polynomet av en kvadratisk matris eller linjär operator är ett annihilande polynom med hjälp av triangularisering av en kvadratisk matris eller linjär operator. Därav visar vi existensen av annihilande polynom till en linjär operator. Därefter bevisas minimalpolynomens unikheter samt deras relation till det karakteristiska polynomet, inklusive hur de kan representeras. Senare undersöks kopplingen mellan minimalpolynom och Jordans normalform. Vi ser att minimalpolynom till en linjär operator bär mer egenskaper än det karakteristiska polynomet eftersom det ger oss Jordans normalform, vilken spelar en stor roll i studie av linjära dynamiska system, vilket vi illustrerar i slutet av uppsatsen. I uppsatsen försöker vi även undersöka relationen mellan behandling av egenvärden med determinant och utan determinant. Där visar vi att Cayley–Hamiltons sats kan bevisas med egenskap av adjunkt matris om vi arbetar med den traditionella behandlingen av egenvärden. Vidare ges en alternativ metod för att beräkna det egenvärden utan att använda determinanter. En central aspekt är att visa hur varje kvadratisk matris kan representeras som en nästan diagonalmatris, övertriangulär matris samt att undersöka hur minimalpolynom kan användas för att beskriva denna egenskap. Slutligen appliceras Jordans normalformen på lösningen av ordinära differentialekvationer. Uppsatsen riktar sig till en målgrupp bestående av studenter som har läst matematik II–Linjär algebra samt lärarstudenter.

2 Invarianta delrum och direkt summa

I detta avsnitt repeterar vi några begrepp från linjär algebra som är återkommande i denna uppsats, såsom egenvärden och egenvektorer till linjära operatorer på ett ändligt dimensionellt vektorrum V över kroppen $\mathbb{F}$, där $\mathbb{F}$ antingen är de reella talen $\mathbb{R}$ eller de komplexa talen $\mathbb{C}$. Vi presneterar även en sammanhängande behandling av engerumuppdelning av vektorrummet för diagonaliserbara linjära operatorer. De flesta material i detta avsnitt är baserade på boken *Linear Algebra* av Friedberg, Insel och Spence (4:e upplagan, 2014) [3]. Följ vår initiala idé ger vi ett bevis på existens av egenvärden till linjära operatorer utan determinanter.

Definition 2.1 En *linjär avbildning* från ett vektorrum V till sig själv, $T : V \rightarrow V$, uppfyller villkoret

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) \text{ för alla } u, v \in V, \text{ för alla } \alpha, \beta \in \mathbb{F}.$$

Vi kommer att undersöka de strukturella egenskaperna hos en linjär operator, så som diagonaliserbarhet, standardformer som Jordans normalform, samt egenvärden och egenvektorer till T genom det så kallade minimalpolynomet, vilket är ett huvudämne i denna uppsats. Här är några antaganden i uppsatsen: dimension av vektorrummet V är n , ett ändligt positivt heltal och den underliggande kroppen $\mathbb{F}$ är antingen $\mathbb{R}$ eller $\mathbb{C}$.

Definition 2.2 Talet $\lambda \in \mathbb{C}$ kallas ett egenvärde till T om det finns nollskild vektor $u \in V$ så att

$$T(u) = \lambda u.$$

Vektorn u kallas *egenvektor* associerad med egenvärdet λ .

Nu bevisas följande sats för existensen av egenvärden utan att använda determinant.

Sats 2.1 Varje linjär operator på ett ändligt dimensionellt vektorrum över $\mathbb{C}$ har ett egenvärde.

Bevis. Låt $n > 0$ vara ett heltal och V vara ett n -dimensionellt komplext vektorrum och T en linjär operator på V . Välj en nollskild vektor $v \in V$, och definiera mängden av $n + 1$ vektorer

$$v, T(v), T^2(v), \dots, T^n(v).$$

Eftersom V har dimension n , är dessa vektorer linjärt beroende. Detta innebär att det finns en icke-trivial linjär kombination av dessa vektorer som är lika med noll, det vill säga att det finns skalärer, inte alla lika med noll, sådana att

$$c_0 v + c_1 T(v) + c_2 T^2(v) + \dots + c_n T^n(v) = 0.$$

Därmed finns ett icke-noll polynom $p(T)$ med minst grad, sådant att

$$p(T)(v) = 0.$$

Enligt algebrans fundamentala sats finns det ett tal $\lambda \in \mathbb{C}$ sådant att $p(\lambda) = 0$. Enligt faktorsatsen kan $p(z)$ skrivas som

$$p(z) = (z - \lambda)q(z),$$

för något komplext polynom $q(z)$, där $q(z)$ har lägre grad än $p(z)$.

Därmed gäller

$$0 = p(T)(v) = (T - \lambda I)q(T)(v).$$

Eftersom $q(T)(v) \neq 0$ (då q har lägre grad än p), följer att $(T - \lambda I)(q(T)(v)) = 0$. Detta innebär att $q(T)(v)$ är en egenvektor till T , associerad till egenvärdet λ . ■

2.1 Invariant delrum

Ibland är det önskvärt att förstå vilka egenskaper operatoren T har på vissa specifika lägre dimensionella delrum av V . Här kommer vi att diskutera invariant delrum av T och relationen till hela vektorrummet V . Vi påminner om att ett delrum av vektorrummet V är en delmängd av V som i sig är ett vektorrum med samma addition och multiplikation med skalär som V . Vi betecknar $\text{span}\{u_1, u_2, \dots, u_k\}$ som mängden av alla linjära kombinationer av vektorerna $u_1, u_2, \dots, u_k \in V$. Vi kan visa att det är ett delrum av V .

Definition 2.3 Ett delrum W av V kallas invariant under T om

$$TW \subseteq W,$$

det vill säga $w \in W$ medför att $T(w) \in W$ för alla $w \in W$.

Vi ger några exempel på sådana delrum till ett ändligdimensionellt vektorrummet V . Vi antar att $T : V \rightarrow V$ är en linjär operator.

Exempel 2.1 Det triviala vektorrummet $\{0\}$ är invariant eftersom om $u \in \{0\}$. Då är $u = 0$, och därmed $Tu = 0 \in \{0\}$.

Exempel 2.2 Antag att λ är ett eget värde till T . Då är $\text{span}\{v\}$ ett invariant delrum, där v är egenvektor av T associerad med λ . Detta är sant på grund av att

$$T(v) = \lambda v \in \text{span}\{v\}.$$

Vi kommer att utvidga detta i nästa avsnitt med målet att dela upp vektorrummet V som en direktsumma av egenrummen.

Exempel 2.3 Nollrummet $\mathcal{N}(T) := \{u \in V : T(u) = 0\}$ och värderummet till T , $\mathcal{R}(T) := \{v : v = T(u) \text{ för något } u \in V\}$ är invarianta delrum under T .

Mer generellt är delrummen $\mathcal{N}((\lambda I - T)^k)$ och $\mathcal{R}((\lambda I - T)^k)$ invarianta under T för alla naturliga tal k och alla $\lambda \in \mathbb{F}$. Här är den linjära avbildningen I identitetsavbildning, dvs $I(u) = u$, för $u \in V$.

Bevis. Antag att $u \in \mathcal{N}(T)$. Då gäller per definition att $T(u) = 0 \in \mathcal{N}(T)$. Därför är $\mathcal{N}(T)$ invariant under T .

Anta att $v \in \mathcal{R}(T)$. Då gäller per definition, att det finns en vektor $u \in V$ sådan att $v = T(u)$. Vi observerar att denna definition säger även att $T(v) \in \mathcal{R}(T)$. Eftersom $T(v) \in \mathcal{R}(T)$ för varje $v \in \mathcal{R}(T)$, kan vi dra slutsatsen att $\mathcal{R}(T)$ är invariant under T .

För de generella delrummen noterar vi först att

$$(\lambda I - T)^k T = T(\lambda I - T)^k.$$

Detta medför att

$$(\lambda I - T)^k T \mathcal{N}((\lambda I - T)^k) = T(\lambda I - T)^k \mathcal{N}((\lambda I - T)^k).$$

Tag nu en godtycklig vektor $u \in \mathcal{N}((\lambda I - T)^k)$. Likheten ovan ger

$$(\lambda I - T)^k T(u) = T(\lambda I - T)^k(u) = 0.$$

Alltså har vi $T(u) \in \mathcal{N}((\lambda I - T)^k)$.

Att $\mathcal{R}((\lambda I - T)^k)$ är invariant under T följs av

$$T\mathcal{R}((\lambda I - T)^k) = T(\lambda I - T)^k V = (\lambda I - T)^k TV \subseteq (\lambda I - T)^k V = \mathcal{R}((\lambda I - T)^k).$$

■

Exempel 2.4 Om delrummen $U_1, U_2 \subseteq V$ är invarianta under T är $U_1 \cap U_2$, $U_1 + U_2$ invarianta under T där summan av vektorrummen $U_1 + U_2$ definieras som

$$U_1 + U_2 = \{u : u = u_1 + u_2, u_1 \in U_1, u_2 \in U_2\}.$$

Bevis. Antag att $v \in U_1 \cap U_2$. Då $v \in U_1$ och $v \in U_2$. Så vi har

$$T(v) \in U_1, \quad T(v) \in U_2,$$

eftersom U_1, U_2 är invarianta under T . Det visar att $T(v) \in U_1 \cap U_2$.

Nu antar vi att $u_1 \in U_1, u_2 \in U_2$. Då $T(u_1) \in U_1$, $T(u_2) \in U_2$, enligt antagandet att U_1, U_2 är invarianta under T . Vi har vidare

$$T(u_1 + u_2) = T(u_1) + T(u_2) \in U_1 + U_2$$

på grund av linjäritet av T och definitionen för summan av delrum. ■

Mer allmänt kan vi definiera summan av m delrum $U_1, \dots, U_m \subseteq V$ på följande sätt: Varje vektor i $U_1 + \dots + U_m$ kan skrivas på formen

$$u_1 + \dots + u_m$$

där $u_k \in U_k$, $k = 1, \dots, m$. Ett speciellt intressant fall är att varje vektor i $U_1 + \dots + U_m$ kan representeras på ovanstående formen på bara ett entydigt sätt. Det leder till följande definition.

Definition 2.4 Summan av delrummen $U_1, \dots, U_m \subseteq V$ kallas *direkt summan*, betecknas $U_1 \oplus \dots \oplus U_m$, om varje vektor $u \in U_1 + \dots + U_m$ kan entydigt skrivas som en summa

$$u = u_1 + \dots + u_m,$$

där $u_k \in U_k$, $k = 1, \dots, m$.

Det betyder att det finns bara ett sätt att dela upp en vektor u som en summa $u_1 + \dots + u_m$ med $u_k \in U_k$. T ex $U_1 + U_2 + U_3$ är *inte* en direkt summa om

$$U_1 = \{(x, y, 0) \in \mathbb{F}^3 : x, y \in \mathbb{F}\},$$

$$U_2 = \{(0, 0, z) \in \mathbb{F}^3 : z \in \mathbb{F}\},$$

$$U_3 = \{(0, y, y) \in \mathbb{F}^3 : y \in \mathbb{F}\}.$$

Det är sant av följande resonemang. Enligt definition $U_1 + U_2 + U_3 = \mathbb{F}^3$ eftersom varje vektor $(x, y, z) \in \mathbb{F}^3$ kan skrivas som

$$(x, y, z) = \underbrace{(x, y, 0)}_{\in U_1} + \underbrace{(0, 0, z)}_{\in U_2} + \underbrace{(0, 0, 0)}_{\in U_3}.$$

Men $\mathbb{F}^3$ är inte direkt summan av U_1, U_2, U_3 eftersom

$$(0, 0, 0) = \underbrace{(0, 1, 0)}_{\in U_1} + \underbrace{(0, 0, 1)}_{\in U_2} + \underbrace{(0, -1, -1)}_{\in U_3}$$

och

$$(0, 0, 0) = \underbrace{(0, 0, 0)}_{\in U_1} + \underbrace{(0, 0, 0)}_{\in U_2} + \underbrace{(0, 0, 0)}_{\in U_3}.$$

Dvs, vektorn $(0, 0, 0)$ har två olika sätt att skrivas som summa $u_1 + u_2 + u_3$ där $u_i \in U_i$, $i = 1, 2, 3$.

Följande satser kommer att ha användning av i senare studie.

Sats 2.2 Antag att $U_1, \dots, U_m$ är delrum i V . Då gäller att summan $U_1 + \dots + U_m$ är en direkt summa om och endast om $0 = u_1 + \dots + u_m$ där $u_k \in U_k$ medför att $u_k = 0$, $k = 1, \dots, m$.

Bevis. Antag först $U_1 + \dots + U_m$ är en direkt summa. Enligt definitionen finns bara ett sätt att skriva $0 = u_1 + \dots + u_m$ där $u_k \in U_k$. Så u_k måste vara nollvektorn i U_k .

Antag nu den omvända. Vi ska visa att $U_1 + \dots + U_m$ är en direkt summa. Låt $u \in U_1 + \dots + U_m$. Då

$$u = u_1 + \dots + u_m.$$

Om representationen för u inte vore endtydig skulle det finnas $\tilde{u}_k \in U_k$ sådana att

$$u = \tilde{u}_1 + \dots + \tilde{u}_m.$$

Substrahera de två ekvationerna fås

$$0 = (u_1 - \tilde{u}_1) + \dots + (u_m - \tilde{u}_m).$$

Nu vet vi att $u_k - \tilde{u}_k \in U_k$. Enligt antagandet finns bara ett sätt att representera 0, nämligen $u_k - \tilde{u}_k = 0$, vilket är samma som $u_k = \tilde{u}_k$, för $k = 1, \dots, m$. ■

Notera att denna sats ger oss ett verktyg att kolla om en summa är en direkt summa.

Följdsats. Antag att U_1 och U_2 är delrum i V . Vi har:

(i) $U_1 + U_2$ en direkt summa $\Leftrightarrow U_1 \cap U_2 = \{0\}$.

(ii) Om $U_1 + U_2$ är en direkt summa så gäller det att $\dim(U_1 + U_2) = \dim U_1 + \dim U_2$.

Bevis. (i) " $\Rightarrow$ ": Vi ska visa att $U_1 \cap U_2 = \{0\}$. Om $u \in U_1 \cap U_2$ då är $0 = u + (-u)$ där $u \in U_1$ och $-u \in U_2$. Enligt den entydiga representationen för 0 som en summa av en vektor i U_1 och en vektor i U_2 måste vi ha $u = 0$. Det visar att $U_1 \cap U_2 = \{0\}$.

" $\Leftarrow$ ": Nu antar vi $U_1 \cap U_2 = \{0\}$. Låt $u_1 \in U_1$, och $u_2 \in U_2$ så att $0 = u_1 + u_2$. Vad vi behöver bevisa är att $u_1 = u_2 = 0$ enligt Sats 2.2. Ekvationen $0 = u_1 + u_2$ innebär $u_1 = -u_2 \in U_2$, vilket betyder u_1 tillhör både U_1 och U_2 . Således $u_1 \in U_1 \cap U_2 = \{0\}$ dvs $u_1 = 0$.

(ii) följer direkt från (i) enligt

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim U_1 \cap U_2 = \dim U_1 + \dim U_2 - \dim(\{0\}) = \dim U_1 + \dim U_2.$$

■

För den direkta summan av $U_1, \dots, U_m \subseteq V$ har vi:

Sats 2.3 Antag att $U_1, \dots, U_m$ är delrum i V . Om $U = U_1 + \dots + U_m$ är en direkt summa gäller det att

$$\dim U = \dim U_1 + \dim U_2 + \dots + \dim U_m.$$

Bevis. Låt $B_k := \{b_1^{(k)}, \dots, b_{d_k}^{(k)}\}$ vara en bas för U_k , $k = 1, \dots, m$ och $d_i := \dim U_i$, $i = 1, \dots, m$. Sätt $B := B_1 \cup \dots \cup B_m$. Observera att varje vektor $u \in U$ kan skrivas på formen

$$u = u_1 + u_2 + \dots + u_m$$

där $u_1 \in U_1, \dots, u_m \in U_m$. Då gäller även

$$u \in \text{span}(B) := \text{span}\{b_1^{(1)}, \dots, b_{d_1}^{(1)}, b_1^{(2)}, \dots, b_{d_2}^{(2)}, \dots, b_1^{(m)}, \dots, b_{d_m}^{(m)}\}.$$

eftersom u_k är linjär kombination av basvektorerna i B_k , $k = 1, \dots, m$. Då har vi $\dim \text{span}(B) \leq d_1 + \dots + d_m$. Nu ska vi visa att likheten gäller. Med andra ord ska vi visa att basvektorerna i $B_1, \dots, B_m$ tillsammans är linjärt oberoende. Låt nu

$$\sum_{k=1}^m \underbrace{\sum_{i=1}^{d_k} \beta_i^{(k)} b_i^{(k)}}_{\in U_k} = 0 \in U.$$

Eftersom U är en direkt summa av $U_1, \dots, U_m$ har vi $\sum_{i=1}^{d_k} \beta_i^{(k)} b_i^{(k)} = 0$ enligt Sats 2.2. Vi har vidare att alla koefficienter $\beta_i^{(k)} = 0$, $i = 1, \dots, d_k, k = 1, \dots, m$, vilket ska visas. ■

Följande definition är viktig i studien av uppdelning av linjära operatorer.

Definition 2.5 Om W är ett invariant delrum under operatoren T , är den naturliga restriktionen $T|_W$ av T till W en linjär operator på W , som är definierad som

$$T|_W(w) = T(w), w \in W.$$

2.2 Egenrum och diagonalisering av linjära avbildningar

Definition 2.6 Låt T vara en linjär operator på det n -dimensionella vektorrummet V och $\lambda \in \mathbb{F}$ vara ett egenvärde till T . Egenrummet associerat med egenvärdet λ , betecknat $E_\lambda(T)$, är nollrummet till $T - \lambda I$, där I är enhetsoperatoren. Vi kallar detta nollrum för egenrummet för λ .

Sats 2.4 Summan av egenrummen är en direkt summa. Mer precis gäller det att

$$E_{\lambda_1}(T) + E_{\lambda_2}(T) + \dots + E_{\lambda_k}(T)$$

är en direkt summa, där $\lambda_1, \dots, \lambda_k$ är olika egenvärden till T . Vidare gäller det att

$$\dim E_{\lambda_1}(T) + \dim E_{\lambda_2}(T) + \dots + \dim E_{\lambda_k}(T) \leq \dim(V) = n.$$

Bevis. Låt E vara summan av egenrummen. Antag att $v_i \in E_{\lambda_i}(T)$, $i = 1, \dots, k$. Vi ska visa att summan $v_1 + v_2 + \dots + v_k = 0$ medför att $v_i = 0$ för alla $i = 1, \dots, k$. Enligt Sats 2.1.1. Eftersom $\lambda_1, \dots, \lambda_k$ är olika är deras associerade egenvektorer linjärt oberoende. Alltså måste $v_i = 0$, därav

$$E_{\lambda_1}(T) + E_{\lambda_2}(T) + \dots + E_{\lambda_k}(T) = E_{\lambda_1}(T) \oplus E_{\lambda_2}(T) \oplus \dots \oplus E_{\lambda_k}(T).$$

Detta visar vidare att

$$\dim E_{\lambda_1}(T) + \dim E_{\lambda_2}(T) + \dots + \dim E_{\lambda_k}(T) = \dim(E_{\lambda_1}(T) \oplus E_{\lambda_2}(T) \oplus \dots \oplus E_{\lambda_k}(T)) \leq \dim(V) = n.$$

Den första likheten gäller enligt Sats 2.3 och den sista olikheten gäller på grund av att dimension av ett delrum i vektorrummet V inte kan överstiga $\dim V$. ■

I kursen Matematik II linjär algebra har vi lärt oss att det är ekvivalent mellan att T är en diagonaliserbar operator och att V har en bas bestående av egenvektorer till T . Med satsen ovan kan vi lägga till flera ekvivalenta villkor. Det vill säga, följande villkor är ekvivalenta:

Sats 2.5 Följande påståenden är ekvivalenta.

1. T är en diagonaliserbar operator.
2. V har en bas bestående av egenvektorer till T .
3. $V = E_{\lambda_1}(T) \oplus \dots \oplus E_{\lambda_n}(T)$.
4. $\dim(V) = \dim E_{\lambda_1}(T) + \dim E_{\lambda_2}(T) + \dots + \dim E_{\lambda_n}(T)$.

Notera att egenrummet $E_{\lambda_k}(T)$ är invariant under T (Se Exempel 2.3) och summa av delrum är också invariant (Se Exempel 2.4). Vi kan därför lägga till ett ytterligare villkor i listan: det finns ett-dimensionellt invariant delrum (under T) $V_1, V_2, \dots, V_n$ till V , sådana att

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_n.$$

där V_k ($k = 1, \dots, n$) kan vara egenrummen.

3 Den generalized egenrumsdecompositionen

Ibland kan en matris vara ”defekt”. Det innebär det att den inte har tillräckligt många egenvektorer för att täcka hela rummet V . Linjära operatorer kan sakna egenvektorer och därför är det inte alltid möjligt ha decomposition som

$$V = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_n}(T).$$

Det är här generaliserade egenvektorer kommer in.

Notera att varje linjär operator på n -dimensionellt vektorrum V kan representeras av en entydigt bestämd $n \times n$ -matris A i en given bas. Det innebär att vi kan studera den linjära operatoren T genom att analysera matrisen A i stället. Dessutom är egenvärdena för operatoren T det samma som egenvärdena för matrisen A .

Enligt definitionen har vi att det finns en vektor $v \neq 0$ så att

$$Av = \lambda v,$$

vilket vi kan skriva om som:

$$(A - \lambda I_n)v = 0 \iff v \in \mathcal{N}(A - \lambda I_n),$$

där I_n är $n \times n$ -enhetsmatrisen. Från och med detta avsnitt kommer vi att undersöka matriser och deras strukturer. Vi kommer också betrakta vektorrummet V över $\mathbb{C}$ i större utsträckning.

Definition 3.1 En vektor v kallas *generaliserade egenvektor* till A motsvarande egenvärdet λ om $v \neq 0$ och $(A - \lambda I_n)^k v = 0$ för något positivt heltal k .

Antag att $k \geq 0$ är ett heltal och $v \in \mathcal{N}(A^k)$. Då gäller att $A^k v = 0$, vilket medför att

$$A^{k+1}v = A(A^k v) = A0 = 0 \Rightarrow v \in \mathcal{N}(A^{k+1}).$$

Därför har vi bevisat följande följd av växande nollrum:

Proposition 3.1

$$\{0\} = \mathcal{N}(A^0) \subseteq \mathcal{N}(A^1) \subseteq \cdots \subseteq \mathcal{N}(A^k) \subseteq \mathcal{N}(A^{k+1}) \subseteq \cdots$$

Vi kan även se att inklusion blir likheter

$$\mathcal{N}(A^m) = \mathcal{N}(A^{m+1}) = \mathcal{N}(A^{m+2}) = \cdots$$

om för något icke-negativt heltal m $\mathcal{N}(A^m) \subseteq \mathcal{N}(A^{m+1})$.

För att bevisa detta behöver vi bara bevisa inklusion $\mathcal{N}(A^{m+k+1}) \subseteq \mathcal{N}(A^{m+k})$, eftersom vi redan har $\mathcal{N}(A^{m+k}) \subseteq \mathcal{N}(A^{m+k+1})$. Tag nu en vektor $v \in \mathcal{N}(A^{m+k+1})$. Då

$$A^{m+1}(A^k v) = A^{m+k+1}v = 0.$$

Så

$$A^k v \in \mathcal{N}(A^{m+1}) = \mathcal{N}(A^m)$$

Således $A^{m+1}v = A^m(A^k v) = 0$ dvs $v \in \mathcal{N}(A^{m+k})$, vilket medför att $\mathcal{N}(A^{m+k+1}) \subseteq \mathcal{N}(A^{m+k})$.

Proposition 3.2

$$\mathcal{N}(A^n) = \mathcal{N}(A^{n+1}) = \mathcal{N}(A^{n+2}) = \cdots$$

Bevis. Det räcker att visa att $\mathcal{N}(A^n) = \mathcal{N}(A^{n+1})$ från föregående argument. Vi bevisar det med motsägel-sesbevis. Antag att nollrummet inte slutar växa. Tillsammans med Proposition 3.1 får vi $\{0\} = \mathcal{N}(A^0)$ och $\mathcal{N}(A^k)$ är delrum av $\mathcal{N}(A^{k+1})$ men inte lika, för $k = 0, 1, \dots, n, n+1$. Delrummen i varje strikt inklusion i kedjan ökar dimension med minst 1. så $\dim \mathcal{N}(A^{n+1}) \geq \dim V + 1 = n+1$, en motsägelse eftersom dimension av delrum inte kan överstiga $\dim V = n$. ■

Med hjälp av Proposition 3.1 och Proposition 3.2 för matrisen $A - \lambda I_n$ har vi följande: *en nollskild vektor $v \in V$ är en generaliserad egenvektor till matrisen A associerad med λ om och endast om*

$$(A - \lambda I_n)^n v = 0$$

I enighet med egenrum kan vi definera generaliserade egenrum $GE_\lambda(A)$:

$$GE_\lambda(A) = \{v \in V : (A - \lambda I_n)^k v = 0 \text{ för något positivt heltal } k\}.$$

Med andra ord är $GE_\lambda(A)$ mängden av alla generaliserade egenvektorer till A samt nollvektorn. Självklart gäller att

$$E_\lambda(A) \subseteq GE_\lambda(A)$$

därför att egenvektorn är en generaliserad egenvektor om vi tar $k = 1$. För enkelhets skull skriver vi bara och GE_λ respektive E_λ om ingen tveksamhet uppstår.

Proposition 3.3 $GE_\lambda = \mathcal{N}((A - \lambda I_n)^n)$.

Bevis. Per definition $v \in \mathcal{N}(A - \lambda I_n)^n$ medför att $v \in GE_\lambda$. Därför $\mathcal{N}((A - \lambda I_n)^n) \subseteq GE_\lambda$.

Antag, å andra sidan, $v \in GE_\lambda$. Så finns ett positivt heltal k så att $v \in \mathcal{N}((A - \lambda I_n)^k)$. Enligt Proposition 3.1 och Proposition 3.2 för $A - \lambda I_n$ gäller att $v \in \mathcal{N}((A - \lambda I_n)^n) \Rightarrow GE_\lambda \subseteq \mathcal{N}(A - \lambda I_n)^n$. ■

Det generaliserade egenrummet till engenvärdet λ är, med andra ord,

$$GE_\lambda = \mathcal{N}((A - \lambda I_n)^n), \tag{1}$$

där $n = \dim(V)$.

Proposition 3.4 Vektorrummet V med dimension n är direktsumman av nollrummet till A^n och värderummet till A^n . Med andra ord

$$V = \mathcal{N}(A^n) \oplus \mathcal{R}(A^n)$$

Bevis. För att visa summan av nollrummet och värderummet är en direkt summa räcker det att visa snittet av dem är $\{0\}$ (Följsatsen till Sats 2.2), dvs

$$\mathcal{N}(A^n) \cap \mathcal{R}(A^n) = \{0\}$$

För detta mål tar vi en godtycklig vektor v i snittet. Då $A^n v = 0$ och $v = A^n u$ för någon $u \in V$. Det medför $A^{2n} u = A^n(A^n v) = 0 \Rightarrow A^n u = 0$ enligt Proposition 3.2. Så $v = A^n u = 0$, vilket är det önskade resultatet. Vi har därför

$$\dim(\mathcal{N}(A^n) \oplus \mathcal{R}(A^n)) = \dim \mathcal{N}(A^n) + \dim \mathcal{R}(A^n) = \dim V.$$

.

Den sista likheten kommer från dimensionssatsen Se [3]. Denna likheten implicerar

$$V = \mathcal{N}(A^n) \oplus \mathcal{R}(A^n).$$

■

Definition 3.2 En matris $N \in \mathbb{C}^{n \times n}$ är nilpotent om det finns ett positivt heltal k sådant att:

$$N^k = 0$$

där 0 är nollmatrisen med samma storlek som N .

Nu kommer vi att bevisa en av de huvudsatserna i denna uppsats:

Sats 3.1 Antag att V är ett komplext vectorrum av ändlig dimension. Låt $\lambda_1, \dots, \lambda_m$ vara olika egenvärden till A . Då gäller att

1. $V = GE_{\lambda_1} \oplus \dots \oplus GE_{\lambda_m}$;
2. $(A - \lambda_j I_n)|_{GE_{\lambda_j}}$ är nilpotent för alla j .
3. GE_{λ_j} är invariant under A för alla j .

Vi bevisar påstående 1 genom att först bevisa följande lemmor.

Lemma 3.1 (i) Varje generaliserad egenvektor till A motsvarar bara ett egenvärde till A .
(ii) De generaliserade egenvektorerna är linjärt oberoende.

Bevis. (i) Antag att $v \in V$ vore en generaliserad egenvektor som motsvarar två egenvärden λ och μ till A . Låt m vara det minsta positiva heltal sådant att $(A - \mu I_n)^m v = 0$. Då

$$\begin{aligned} 0 &= (A - \lambda I_n)^n v = ((A - \mu I_n) + (\mu - \lambda)I_n)^n v \\ &= \left(\sum_{k=0}^n \binom{n}{k} (\mu - \lambda)^{n-k} (A - \mu I_n)^k \right) v \end{aligned}$$

Multiplifiera $(A - \mu I_n)^{m-1}$ på båda leden av ekvationen ovan fås

$$0 = \binom{n}{0} (\mu - \lambda)^n (A - \mu I_n)^{m-1} v + \sum_{k=1}^n \binom{n}{k} (\mu - \lambda)^{n-k} (A - \mu I_n)^{k+m-1} v = (\mu - \lambda)^n (A - \mu I_n)^{m-1} v$$

eftersom $(A - \mu I_n)^m v = 0$ (Proposition 3.2). På grund av att $(A - \mu I_n)^{m-1} v \neq 0$ har vi $\mu = \lambda$, dvs varje generaliserad egenvektor till A motsvarar bara ett egenvärde.

(ii) Vi antar motsatsen. Då finns det ett minsta positivt heltal m så att det finns linjärt beroende vektorer $v_1, \dots, v_m$ av generaliserad egenvektorer associerad med olika egenvärden $\lambda_1, \dots, \lambda_m$. Observera att generaliserad egenvektor är nollskild måste $m \geq 2$. Så det finns nollskilda $\alpha_1, \dots, \alpha_m$ (på grund av att m är minimalt) så att

$$\alpha_1 v_1 + \dots + \alpha_m v_m = 0.$$

Multiplifiera $(A - \lambda_m I_n)^n$ på båda leden av ovanstående ekvationen fås

$$\alpha_1 (A - \lambda_m I_n)^n v_1 + \dots + \alpha_m (A - \lambda_m I_n)^n v_m = 0.$$

Vidare

$$(A - \lambda_m I_n)^n v_k \neq 0, \quad k = 1, \dots, m-1,$$

eftersom v_k annars skulle vara en generaliserad egenvektor associerad med två olika egenvärden λ_k och λ_m , vilket motsäger (i). Men

$$(A - \lambda_k I_n)^n ((A - \lambda_m I_n)^n v_k) = (A - \lambda_m I_n)^n ((A - \lambda_k I_n)^n v_k) = 0$$

Det visar att $(A - \lambda_m I_n)^n v_k$ är en generaliserad egenvektor associerad med egenvärdet λ_k . Så

$$(A - \lambda_m I_n)^n v_1, \dots, (A - \lambda_m I_n)^n v_{m-1}$$

är linjärt beroende mängd av $m-1$ generaliserad egenvektorer associerad med olika egenvärden, en motsägelse mot att m är minst. Då har vi bevisat (ii). ■

Lemma 3.2 Summan $GE_{\lambda_1} + \dots + GE_{\lambda_m}$ är en direkt summa.

Bevis. Betrakta $v_1 + \dots + v_m = 0$ där $v_k \in GE_{\lambda_k}$. Lemma 3.1 säger att generaliserade egenvektorer associerad med olika egenvärden till A är linjärt oberoende. Således varje $v_k = 0$. Enligt Sats 2.2 är $GE_{\lambda_1} + \dots + GE_{\lambda_m}$ en direkt summa. ■

Lemma 3.3. Det finns en bas i V som utgörs av de generaliserad egenvektorerna.

Bevis. Vi bevisar detta lemma med hjälp av matematisk induktion på n som är dimension av V . I bassteget $n = 1$ ser vi att påståendet gäller eftersom varje nollskild vektor i V är en egenvektor till A .

Antag att $n > 1$ och påståendet gäller för alla naturliga tal mindre än n . Låt λ vara ett egenvärde till A . Tillämpningen av Proposition 3.4 på $A - \lambda I_n$ ger

$$V = \mathcal{N}((A - \lambda I_n)^n) \oplus \mathcal{R}((A - \lambda I_n)^n).$$

Om $\mathcal{N}((A - \lambda I_n)^n) = V$, är varje nollskild vektor i V en generaliserad egenvektor. Och vidare i detta fall finns det en bas för V som utgörs av generaliserad egenvektorer till A . Så vi kan anta att $\mathcal{N}((A - \lambda I_n)^n) \neq V$, dvs $\mathcal{R}((A - \lambda I_n)^n) \neq \{0\}$. Notera att även $\mathcal{N}((A - \lambda I_n)^n) \neq \{0\}$ eftersom λ är ett egenvärde. Då har vi

$$0 < \dim \mathcal{R}((A - \lambda I_n)^n) < n.$$

Vidare är $\mathcal{R}((A - \lambda I_n)^n)$ invariant under A (Exempel 2.3). Låt $S = A|_{\mathcal{R}((A - \lambda I_n)^n)}$. Tillämpa nu induktionsantagandet på S , dvs att det finns en bas för $\mathcal{R}((A - \lambda I_n)^n)$ som utgörs av generaliserad egenvektorer till S , vilka är också generaliserad egenvektorer till A . Kombinera denna bas med basen för $\mathcal{N}((A - \lambda I_n)^n)$ ger en bas för V som innehåller generaliserad egenvektorer till A . ■

Bevis (Sats 3.1) Påstående 3 följer av Exempel 2.3 eftersom $GE_{\lambda_k} = \mathcal{N}((A - \lambda_k I_n)^n)$.

För att bevisa Påstående 2 antag att $k \in \{1, \dots, m\}$. Om $v \in GE_{\lambda_k}$, så gäller att

$$(A - \lambda_k I_n)^n v = 0$$

Således gäller att:

$$((A - \lambda_k I_n)|_{G(\lambda_k, A)})^n = 0.$$

Därmed är $(A - \lambda_k I_n)|_{G(\lambda_k, A)}$ nilpotent.

Återstår att bevisa påstående 1. Lemma 3.2 säger att $GE_{\lambda_1} \oplus \dots \oplus GE_{\lambda_m}$.

Slutligen kan varje vektor i V skrivas som en ändlig summa av generaliserade egenvektorer till A eftersom det finns en bas för V som består av generaliserade egenvektorer till A (Lemma 3.3). Så

$$V = GE_{\lambda_1} \oplus \dots \oplus GE_{\lambda_m}.$$

■

Från Sats 3.1 ser vi att för ett komplex vektorrum finns det en bas för V som utgörs av de generaliserade egenvektorerna till A . Vi kommer att återkomma till användningen av generaliserade egenvektorer i samband med Jordans normalform senare i arbetet.

4 Triangulära matriser och Caylay-Hamiltons Sats

Sats 4.1 (Schurs lemma) Varje kvadratisk matris A kan skrivas som

$$A = URU^*,$$

där U är en unitär matris och R är en övertriangulär matris. Här är U^* det Hermiteska transponatet av U , vilket innebär dess transponat och komplexa konjugat.

Med andra ord för varje linjär operator finns en ortonormerad bas i vilken dess matris är övertriangulär.

Bevis. Låt A vara en $n \times n$ -matris.

Vi börjar med att skriva om uppdelningen som

$$R = U^{-1}AU.$$

Vi bevisar satsen med hjälp av matematisk induktion.

När $n = 1$, vårt basfall, ska vi se om det är sant. Om vi har 1×1 -matris som en skalär (a) då är

$$(a) = (1)^{-1}(a)(1)$$

allt vi behöver.

Nu antar vi att detta är sant för matriser av storlek $(n-1) \times (n-1)$. Välj en godtycklig egenvektor med enhetslängd v_1 : $Av_1 = \lambda_1 v_1$. Då kan vi utvidga den till en ortonormerad bas $v_1, \dots, v_n$ för V . Bilda en unitär matris V_1 med kolonner $v_1, \dots, v_n$, dvs $V_1 = (v_1, \dots, v_n)$. Då har vi att

$$\tilde{R} := V_1^{-1}AV_1 = \begin{pmatrix} \lambda_1 & c \\ 0 & A_2 \end{pmatrix}$$

där c är en radvektor med $n-1$ komponenter och A_2 är en kvadratisk $(n-1) \times (n-1)$ -matris. Enligt induktionsantagande finns en $(n-1) \times (n-1)$ unitär matris U_2 så att $U_2^{-1}A_2U_2 = R_2$ är en övertriangulär matris. Vidare bildar vi matrisen

$$V_2 = \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}$$

som är en unitär matris av storlek n . Notera att

$$V_2^{-1}\tilde{R}V_2 = \begin{pmatrix} 1 & 0 \\ 0 & U_2^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & c \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & C_2 \\ 0 & U_2^{-1}A_2U_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & C_2 \\ 0 & R_2 \end{pmatrix}$$

är en övertriangulär matris. Därför får vi

$$(V_1V_2)^{-1}A(V_1V_2) = V_2^{-1}(V_1^{-1}AV_1)V_2 = V_2^{-1}\tilde{R}V_2 = R.$$

Detta är som vi vill ha eftersom V_1V_2 är unitär. ■

Nu ska vi bevisa Cayley Hamilton sats med hjälp av Schurs lemma. Men vi behöver en determinat-fri definition för det karakteristiska polynomet till A .

Definition 4.1 Multipliciteten av ett egenvärde λ till $n \times n$ -matrisen A är definierat som $\dim GE_\lambda$.

Med andra ord är multipliciteten av egenvärdet λ lika med

$$\dim \mathcal{N}((A - \lambda I_n)^n)$$

från Proposition 3.3.

Definition 4.2 Antag att V är ett n -dimensionellt vektorrum över $\mathbb{C}$. Låt $\lambda_1, \dots, \lambda_m$ vara egenvärden till A med multiplicitet $d_1, \dots, d_m$. Polynomet

$$\chi(\lambda) := (\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} \cdots (\lambda - \lambda_m)^{d_m}$$

kallas det karakteristiska polynomet till A .

Vi kommer senare att bevisa att denna definition av det karakteristiska polynomet är samma som den vi lärde oss med determinanter.

Vi använder också begreppen den algebraiska och den geometriska multipliciteten till egenvärdet λ till A , betecknas $a(\lambda)$ respektive $g(\lambda)$. Se [3]. De är

$$\begin{aligned} a(\lambda) &:= \dim \mathcal{N}((A - \lambda I_n)^n) = \dim GE_\lambda \\ g(\lambda) &:= \dim \mathcal{N}((A - \lambda I_n)^n) = \dim E_\lambda \end{aligned}$$

Det är självklart att den algebraiska multipliciteten är multiplicitet i Definition 4.1.

Enligt Sats 3.1 får vi följande resultat:

Sats 4.2 Summan av multipliciteter av alla egenvärden är lika med dimension till V : $d_1 + \cdots + d_m = n$.

Sats 4.3 (Cayley-Hamiltons Sats) Antag att $\chi(\lambda)$ är det karakteristiska polynomet till A . Då gäller att

$$\chi(A) = 0.$$

Bevis. Schurs lemma säger att det finns en unitär matris U så att $U^*AU = R$ är triangulär och vårt bevis visar att egenvärden av A ligger på diagonalen av R . Låt $\lambda_1, \dots, \lambda_m$ vara olika egenvärden med multiplicitet $d_1, \dots, d_m$. Då är

$$U^*AU = R = \begin{pmatrix} R_1 & * & \cdots & * \\ & R_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & R_m \end{pmatrix}, \text{ där } R_i = \begin{pmatrix} \lambda_i & * & \cdots & * \\ & \lambda_i & \cdots & * \\ & & \ddots & \vdots \\ & & & \lambda_i \end{pmatrix}_{d_i \times d_i}$$

Vi ser att $(R_i - \lambda_i I_{d_i})^{d_i} = 0$. Således

$$(R - \lambda_i I_n)^{d_i} = \begin{pmatrix} * & \cdots & * & \cdots & * \\ & \ddots & \vdots & & \vdots \\ & & 0 & \cdots & * \\ & & & \ddots & \vdots \\ & & & & * \end{pmatrix}_{d_i \times d_i}$$

där nollmatrisen ligger på i -te blockrad. Denna form leder till

$$(R - \lambda_1 I_n)^{d_1} (R - \lambda_2 I_n)^{d_2} \cdots (R - \lambda_m I_n)^{d_m} = 0.$$

Dvs $\chi(R) = 0$ Vidare

$$\begin{aligned} U^*p(A)U &= U^*(A - \lambda_1 I_n)^{d_1} (A - \lambda_2 I_n)^{d_2} \cdots (A - \lambda_m I_n)^{d_m} \\ &= (R - \lambda_1 I_n)^{d_1} (R - \lambda_2 I_n)^{d_2} \cdots (R - \lambda_m I_n)^{d_m} = 0 \end{aligned}$$

vilket ger $\chi(A) = 0$.

■

5 Minimalpolynom

5.1 Annihilerande polynom

Definition 5.1 Ett *annihilerande polynom* definieras som ett polynom $p(t)$ som annihilerar operatoren T , dvs $p(T) = 0$.

Om $p(t)$ är ett nollpolynom så är det klart att $p(T) = 0$. Men detta är ett trivialt fall. Vidare i arbetet förutsätter vi att det annihilerande polynomet $p(t)$ inte är ett nollpolynom. Låt oss titta på följande exempel.

För matrisen

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

ser vi att $p(\lambda) = (2 - \lambda)^2$ är ett annihilatorpolynomet genom

att först beräkna $A - 2I$:

$$A - 2I = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

och sedan beräkna kvadraten $(A - 2I)^2$:

$$(A - 2I)^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Vi ser och att det karakteristiska polynomet för A är $p(\lambda) = (2 - \lambda)^2$. $p(\lambda)$ är det karakteristiska polynomet till A . Enligt Cayley-Hamiltons sats är det karakteristiska polynomet ett annihilerande polynom.

För att definiera minimalpolynom behöver vi veta först vad ett moniskt polynom är. Ett moniskt polynom är ett polynom vars ledande koefficient, dvs högstgradskoefficient är lika med 1. Till exempel är polynomet $2 + 9z^2 + z^7$ ett moniskt polynom av grad 7.

Definition 5.2 Anta att T är en linjär operator och har ett annihilerande moniskt polynom. Det unikt bestämda annihilerande polynomet med lägsta grad kallas för minimalpolynomet till T och betecknas $m_T(t)$. Vi skriver ofta $m(t)$ om det inte finns oklarhet.

För att beräkna minimalpolynomet för en operator T , behöver vi hitta det minsta positiva heltalet m sådant att ekvationen

$$c_0I + c_1T + \dots + c_{m-1}T^{m-1} = -T^m$$

har en lösning $c_0, c_1, \dots, c_{m-1} \in \mathbb{F}$.

Exempel 5.1 Betrakta $n \times n$ identitetsmatrisen I . Då är polynomet

$$p(x) = x - 1$$

moniskt, dess ledande koefficient är lika med 1 och det är ett annihilande polynom för I eftersom

$$p(I) = I - I$$

Polynomet p har grad 1. Eftersom att minimalpolynom är ett annihilande polynom och har lägsta möjliga grad så finns det inga andra moniska annihilande polynom av lägre grad eftersom det enda moniska polynomet av grad lägre än 1 är

$$q(x) = 1$$

vilket inte är annihilande. Därför är p det minsta polynomet för I . Som tidigare kan vi studera minimalpolynom av en matris istället för en linjär operator.

5.2 Entydighet av minimalpolynom

Sats 5.1 För en $n \times n$ -matris A existerar ett entydigt minimalt polynom $m(x)$ som uppfyller att $m(A) = 0$. Enligt Cayley-Hamiltons sats är graden av $m(x)$ högst n , det vill säga $\deg(m(x)) \leq n$.

Bevis. Vi behöver endast bevisa entydigheten eftersom Cayley-Hamiltons sats ger att alla kvadratiska matriser har ett annihilande polynom. Låt p_1 och p_2 vara två olika annihilande polynom för A , båda med

grad k . Vi låter k vara den minsta graden för vilken annihilande polynom som helst för A . dvs $p_1(A) = 0$ och $p_2(A) = 0$, och inget annat polynom med grad mindre än k uppfyller denna egenskap.

Vi definierar ett polynom $d(x)$ som skillnaden mellan p_1 och p_2 :

$$d(x) = p_1(x) - p_2(x).$$

Det borde ha grad mindre än k eftersom p_1 och p_2 är moniska.

Eftersom både $p_1(A) = 0$ och $p_2(A) = 0$, får vi:

$$d(A) = p_1(A) - p_2(A) = 0 - 0 = 0.$$

Dvs $d(A) = 0$. Så $d(x)$ delad med ledande koefficienten blir ett annihilande moniskt polynom med grad strikt mindre än k . Det är omöjligt. Med andra ord $p_1(x) = p_2(x)$. Därmed är det minimalpolynomet unikt. ■

Sats 5.2 Polynomet $p(x)$ annihilerar A om och endast om $p(x)$ är en multipel av minimalpolynomet $m(x)$.

Bevis. Att varje multipel av minimalpolynomet $q(x)m(x)$ annihilerar matrisen A är uppenbart.

Antag omvänt att $p(x)$ är ett annihilande polynom. Det borde vara så att $\deg p(x) \geq \deg m(x)$. Genom att dividera $p(x)$ med minimalpolynomet $m(x)$ får vi med polynom division

$$p(x) = q(x)m(x) + r(x),$$

där resten $r(x)$ är ett polynom som har lägre grad än minimalpolynomet $m(x)$.

Insättning av matrisen A ger att $r(A) = 0$.

Således är $r(x)$ ett annihilande polynom, om $r(x)$ är nollskilt kan vi dela med dess ledande koefficient och bilda ett moniskt annihilande polynom där graden är mindre än $p(x)$. Detta motsäger $m(x)$ minimalitet. Detta är endast möjligt om $r(x)$ är nollpolynomet. Detta visar att $p(x) = q(x)m(x)$. Med andra ord gäller att $m(x)$ dividerar varje annihilande polynom. ■

5.3 Relation till det karakteristiska polynomet

Antag att $\lambda_1, \dots, \lambda_m$ är olika egenvärden till $n \times n$ -matrisen A . Låt $\chi(\lambda)$ vara det karakteristiska polynomet till A , dvs

$$\chi(\lambda) = (\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_m)^{d_m}$$

där d_i är (den algebraiska) multipliciteten till λ_i , $i = 1, \dots, m$. Låt minimalpolynomet vara

$$m_A(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_m)^{k_m},$$

där k_i , kallas index av λ_i , är så att $(A - \lambda_i I_n)^{k_i} = 0$.

Enligt sats 5.2 och Cayley Hamiltons sats delar $m_A(\lambda)$ $p(\lambda)$. Vidare har vi $m_A(\lambda) = \chi(\lambda)$ om och endast om $d_i = k_i$, $i = 1, \dots, m$. Det är ekvivalent med att det finns bara en block på formen

$$\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ & \lambda_i & 1 & \cdots & 0 \\ & & \ddots & \ddots & \vdots \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}$$

för varje λ_i .

Därför har varje egenvärde λ_i en geometrisk multiplicitet på exakt 1, vilket innebär att det endast finns en linjärt oberoende egenvektor associerad med varje λ_i . Detta leder till att A är diagonaliserbar om och endast om minimalpolynomet $m_A(\lambda)$ för A är produkt av olika linjära faktorer: $m_A(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_m)$.

Sammanfattningsvis har vi bevisat följande egenskaper hos minimalpolynomet till A .

Sats 5.3 Låt A vara en $n \times n$ -matris med olika egenvärden $\lambda_1, \dots, \lambda_m$. Då gäller följande

- $m_A(\lambda) = \chi(\lambda)$ om och endast om den geometriska multipliciteten av λ_i är lika med 1 för varje λ_i eller den algebraiska multipliciteten av λ_i är lika med index av λ_i , för varje i ;
- A är diagonaliserbar om och endast om $m_A(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_m)$

5.4 Hur ser minimalpolynom ut?

Minimalpolynomet $m(t)$ för en nilpotent matris N är det polynom av lägsta grad som annihilerar N , dvs. $m(N) = 0$, men $m(t)$ är det polynom av lägsta möjliga grad som uppfyller denna egenskap. För att förstå hur minimalpolynomet ser ut, låt oss titta på ett exempel med en nilpotent matris.

Betrakta följande nilpotenta 4×4 -matris N :

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

Denna matris har nollor på diagonalen och ettor på superdiagonalen. Matrisen är nilpotent, vilket innebär att det finns ett tal k sådant att $N^k = 0$ och $N^{k-1} \neq 0$.

Låt oss nu beräkna några potenser av N :

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^4 = 0$$

Eftersom $N^4 = 0$, är $m_N(T) = T^4$ Minimalpolynomet för matrisen N är $m_N(t) = t^4$.

Låt oss titta på den såkallade Jordanblockmatrisen J , där λ representerar egenvärdet och vi har ettor på superdiagonalen. Matrisen kan skrivas som:

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

Vi vill hitta minimalpolynomet för J ,

$$p(t) = (t - \lambda)^k$$

där k ska bestämmas. Vi kan nu beräkna $p(J)$:

$$p(J) = (J - \lambda I_k)^k = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

en nilpotent matris.

Detta ger oss minimalpolynomet för J , nämligen:

$$p(t) = (t - \lambda)^4$$

6 Karakteristiska polynom och determinanter

Här bevisar vi att den tidigare determinant-fri definitionen av det karakteristiska polynomet för A är samma som $\det(zI_n - A)$. Därefter presenterar vi ett alternativt bevis på Cayley-Hamiltons sats baserat på determinantens definition av det karakteristiska polynomet.

Sats 6.1. Det karakteristiska polynomet till A är lika med $\det(\lambda I_n - A)$. Med andra ord

$$\det(zI_n - A) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}.$$

Bevis. Enligt beviset för Schurs lemma finns det en unitär bas för vektorrummet V över $\mathbb{C}$ till vilken A har en övretriangulär matris R med λ_k på diagonalen multiplicitet d_k . Enligt räknereglererna för determinanten

$$\det(zI_n - A) = \det(zI_n - R) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}.$$

■

Nu bevisar vi Cayley-Hamiltons sats med determinanter, dvs vi använder kofaktorer av matriser, som är definierad med determinanter av mindre matriser.

Låt komponenter till matrisen A vara a_{ij} , $i, j = 1, \dots, n$. Låt matrisen M_{ij} vara den delmatris av A där rad i och kolonn j är borttagna. $\det M_{ij}$ kallas minor till a_{ij} och M_{ii} kallas principalminorer. Definiera vidare

$$c_{ij} = (-1)^{i+j} \det M_{ij}.$$

Vi säger att c_{ij} är kofaktorer till a_{ij} .

Låt C vara en matris med komponenter c_{ij} . Då kallas $C^\top$ adjunkt till matrisen A , betecknas ofta med $\text{adj}(A)$. Enligt Laplaceutveckling för beräkning av determinanter får vi

$$A[\text{adj}(A)] = [\text{adj}(A)]A = [\det A]I_n.$$

Bevis av Cayley-Hamiltons sats, alternativ 2:

Eftersom matrisen A har storlek n , kan det karakteristiska polynomet skrivas som:

$$\det(A - \lambda I_n) = p_0 + p_1\lambda + p_2\lambda^2 + \cdots + p_{n-1}\lambda^{n-1} + p_n\lambda^n$$

Det vi behöver bevisa är:

$$p_0 I_n + p_1 A + p_2 A^2 + \cdots + p_{n-1} A^{n-1} + p_n A^n = 0.$$

Eftersom storlek till kofaktorer är högst $n - 1$ har vi

$$\text{adj}(A - \lambda I_n) = B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1}.$$

Substitution i ekvationen för adj ovan:

$$(A - \lambda I_n) \cdot (B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1}) = \det(A - \lambda I_n)I_n$$

Vi får då

$$(A - \lambda I_n) \cdot (B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1}) = (p_0I + p_1A + p_2A^2 + \cdots + p_{n-1}A^{n-1} + p_nA^n)I_n.$$

Vänsterled är lika med

$$AB_0 + \lambda AB_1 + \lambda^2 AB_2 + \cdots + \lambda^{n-1} AB_{n-1} - \lambda IB_0 - \lambda^2 IB_1 - \lambda^3 IB_2 - \cdots - \lambda^n IB_{n-1}.$$

Vänsterled efter omordning fås av följande:

$$AB_0 + \lambda(AB_1 - B_0) + \lambda^2(AB_2 - B_1) + \cdots + \lambda^{n-1}(AB_{n-1} - B_{n-2}) - \lambda^n B_{n-1}.$$

Vi har högerled lika med

$$p_0I + p_1\lambda I + p_2\lambda^2 I + \cdots + p_{n-1}\lambda^{n-1}I + p_n\lambda^n I.$$

Nu återstår att matcha koefficienterna till λ^k där $k = 0, 1, \dots, n$.

$$AB_0 + \lambda(AB_1 - B_0) + \lambda^2(AB_2 - B_1) + \cdots + \lambda^{n-1}(AB_{n-1} - B_{n-2}) - \lambda^n B_{n-1}.$$

Vi har högerled:

$$p_0I + p_1\lambda I + p_2\lambda^2 I + \cdots + p_{n-1}\lambda^{n-1}I + p_n\lambda^n I.$$

Nu återstår att matcha koefficienterna:

$$\begin{aligned} AB_0 &= p_0I \\ AB_1 - B_0 &= p_1I \\ AB_2 - B_1 &= p_2I \\ &\vdots \\ AB_{n-1} - B_{n-2} &= p_{n-1}I \\ -B_{n-1} &= p_nI \end{aligned}$$

Enligt det vi skulle bevisa ser vi att vi behöver multiplicera ekvationerna ovan med A , A^2 , A^{n-1} och A^n då får vi

$$\begin{aligned} AB_0 + A^2B_1 - AB_0 + A^3B_2 - A^2B_1 + \cdots + A^nB_{n-1} - A^{n-1}B_{n-2} - A^nB_{n-1} \\ = p_0I + p_1A + p_2A^2 + \cdots + p_{n-1}A^{n-1} + p_nA^n. \end{aligned}$$

Vi kommer fram till att vänsterled kansellerar term efter term och leder oss till nollmatrisen.

$$0 = p_0I + p_1A + p_2A^2 + \cdots + p_{n-1}A^{n-1} + p_nA^n$$

■

Cayley-Hamiltons sats har många användningar. Vi demonstrerar en av dem, beräkning av inversen till en inverterbar matris.

Givet matrisen

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

kan vi beräkna A^{-1} med hjälp av Cayley-Hamiltons satsen.

Det karakteristiska polynomet för A är alltså:

$$\chi_A(\lambda) = \lambda^2 - 2\lambda + 5$$

Nu ska vi använda Cayley-Hamiltons sats för att hitta A^{-1}

Enligt Cayley-Hamiltons sats uppfyller matrisen A sitt eget karakteristiska polynom, vilket innebär att:

$$A^2 - 2A + 5I = 0$$

För att lösa ut A^{-1} kan vi skriva om ekvationen ovan. Vi har I :

$$A^2 - 2A = -5I$$

$$-\frac{1}{5}(A^2 - 2A) = I$$

Därmed kan vi uttrycka A^{-1} som

$$A^{-1} = -\frac{1}{5}(A - 2I).$$

7 Beräkna egenvärden utan att beräkna determinanten

För detta mål inför vi minimalpolynom för en vektor relaterad till matrisen A .

Definition 7.1 Minimalpolynomet för b med n (komplexa) komponenter relaterad till $n \times n$ -matrisen A är definierad som polynomet $m_{A,b}(x)$ av minsta grad sådant att $m_{A,b}(A)b = 0$.

Satsen nedan ger oss en algoritm för att hitta alla egenvärden utan att lösa ekvationen $\det(zI_n - A) = 0$.

Sats 7.1 Låt $A \in \mathbb{C}^{n \times n}$ och $B = \{b_1, \dots, b_n\}$ vara en bas för $\mathbb{C}^n$. Då är minimalpolynomet m_A den minsta gemensamma multipeln till $m_{A,b_1}, m_{A,b_2}, \dots, m_{A,b_n}$.

Bevis. Låt p vara den minsta gemensamma multipeln till $m_{A,b_1}, m_{A,b_2}, \dots, m_{A,b_n}$. Vi först bevisar $m_A(x)$ delar $p(x)$ och sedan bevisar $p(x)$ också delar $m_A(x)$.

Eftersom varje m_{A,b_i} delar $p(x)$ har vi att $p(A)b_i = 0$ för varje i . Med andra ord

$$\text{span}\{b_1, \dots, b_n\} \subset \mathcal{N}(p(A)) \Rightarrow \dim \mathcal{N}(p(A)) = n \Leftrightarrow p(A) = 0.$$

Enligt Sats 5.2 delar minimalpolynomet p .

Omvänt, ger divisionen

$$m_A = q_k m_{A,b_k} + r_k, \quad k = 1, \dots, n$$

där $\deg r_k < \deg m_{A,b_k}$. Detta ger

$$0 = m_A(A)b_k = q_k(A)m_{A,b_k}(A)b_k + r_k(A)b_k = r_k(A)b_k,$$

och minimaliteten av graden på m_{A,b_k} ger $r_k = 0$. Alltså delar $m_{A,b_k} m_A$ för alla k . Därför delar deras minsta gemensamma multipel p också m_A .

Så p och m_A delar varandra och då båda har ledande koefficient 1 gäller $m_A = p$. ■

Satsen ger oss i praktiken en metod att finna egenvärde. Det säger också att om vi hittar tillräckligt många vektorer (godtyckliga) så kan vi bilda ett minimalt polynom genom minimalpolynomen för dessa vektorer relativt A .

Exempel.

Betrakta matrisen

$$A = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 5 & 2 \\ 1 & 1 & 4 \end{pmatrix}$$

Om vi vill beräkna egenvärdena med determinantkalkyl får vi

$$\det(A - \lambda I) = \det \begin{pmatrix} 3 - \lambda & 1 & 2 \\ -1 & 5 - \lambda & 2 \\ 1 & 1 & 4 - \lambda \end{pmatrix} = 0$$

Egenvärdena får vi genom lösningarna till ekvationen $\lambda^3 - 12\lambda^2 + 44\lambda - 48 = 0$. Dessa är $\lambda = 2, 4, 6$.

Nu demonstrerar vi att egenvärdena kan fås med hjälp av minimalpolynom.

Eftersom A verkar på vektorer i $\mathbb{R}^3$, måste varje uppsättning av fyra vektorer vara linjärt beroende. Därför väljer vi en godtycklig vektor v och beräknar Av , A^2v , A^3v . För att förenkla beräkningarna väljer vi vektorn

$$v = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Beräkning av Av ger

$$Av = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}.$$

Beräkning A^2v av ger

$$A^2v = \begin{pmatrix} 10 \\ -6 \\ 6 \end{pmatrix}.$$

Beräkning av A^3v ger

$$A^3v = \begin{pmatrix} 36 \\ -28 \\ 28 \end{pmatrix}$$

Med radreducering får vi

$$\begin{pmatrix} 1 & 3 & 10 & 36 \\ 0 & -1 & -6 & -28 \\ 0 & 1 & 6 & 28 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 10 & 36 \\ 0 & 1 & 6 & 28 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

vilket ger lösningen

$$x_1 = 8s + 48t, \quad x_2 = -6s - 28t, \quad x_3 = s, \quad x_4 = t$$

om vi låter $t = 0$ och $s = 1$ får vi att $x_1 = 8, x_2 = -6, x_3 = 1, x_4 = 0$

Detta leder till följande ekvation:

$$8v - 6Av + A^2v = 0.$$

Minimalpolynomet blir således $x^2 - 6x + 8 = 0$, vilket har rötterna $x = 4$ och $x = 2$. Dessa rötter motsvarar egenvärdena $\lambda = 4$ och $\lambda = 2$.

Vi ser i detta fall att vi har missat $\lambda = 6$,

Denna metod kräver enbart radoperationer och matris-multiplikation. En fördel med metoden är att den ibland leder till en ekvation av lägre grad än det karakteristiska polynomet. En nackdel är dock att vi kan missa vissa egenvärden, eftersom minimalpolynomet inte alltid innehåller alla egenvärden, särskilt om deras algebraiska multiplicitet är större än ett.

För att beräkna nästa egenvärde väljer vi en annan vektor v alltså $v = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

Beräkning av Av, A^2v, A^3v ger

$$Av = \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix},$$

$$A^2v = \begin{pmatrix} 10 \\ 26 \\ 10 \end{pmatrix},$$

$$A^3v = \begin{pmatrix} 76 \\ 140 \\ 76 \end{pmatrix}.$$

På samma sätt får vi följande med radreducering:

$$\begin{pmatrix} 1 & 5 & 26 & 140 \\ 0 & 1 & 10 & 76 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Vi hittar lösningarna $x_1 = 24$, $x_2 = -10$, $x_3 = 1$, och $x_4 = 0$. Detta ger oss ekvationen

$$24v - 10Av + A^2v = 0.$$

Minimalpolynomet blir då

$$x^2 - 10x + 24 = 0,$$

vilket har lösningarna $x = 4$ och $x = 6$. Därmed blir egenvärdena $\lambda = 4$ och $\lambda = 6$.

Vi kan ignorera $x = 4$, eftersom vi redan har beräknat detta. Med denna metod har vi nu hittat alla de egenvärden vi behövde.

Ibland kan vi få fram minimalpolynomet via en enklare matris. Det grundas på följande observation.

Sats 7.2 Låt A och B vara två similära $n \times n$ -matriser, dvs det finns en inverterbar matris P sådana att $A = PBP^{-1}$. Då har A och B samma minimalpolynom.

Bevis. Låt p vara ett polynom

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n.$$

Då gäller

$$\begin{aligned} p(B) &= a_0I + a_1B + a_2B^2 + \cdots + a_nB^n \\ &= a_0P^{-1}P + a_1(P^{-1}AP) + a_2(P^{-1}AP)^2 + \cdots + a_n(P^{-1}AP)^n \\ &= a_0P^{-1}IP + a_1P^{-1}AP + a_2P^{-1}A^2P + \cdots + a_nP^{-1}A^nP \\ &= P^{-1}(a_0I + a_1A + a_2A^2 + \cdots + a_nA^n)P \\ &= P^{-1}p(A)P \end{aligned}$$

eller

$$p(B) = P^{-1}p(A)P$$

Om p är ett annihilande polynom för A , då gäller

$$p(A) = 0$$

och

$$p(B) = 0,$$

vilket innebär att

$$p(B) = 0.$$

Alltså är p ett annihilande polynom för A om och endast om det är ett annihilande polynom för B .

Eftersom minimalpolynomet är det moniska polynomet med lägst grad bland de annihilande polynomen för varje matris måste också deras respektive minimalpolynom vara samma. ■

Med detta resultat kan vi bestämma minimalpolynomet för en matris A genom att först hitta en matris J i Jordans normalform som är similär med A och sedan använda ovanstående för att bestämma minimalpolynomet.

Först illustrerar vi med matrisen

$$J = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

Jordan block för J är följande:

$$J_{1,2} = (2).$$

$$J_{3,2} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

$$J_{2,3} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}.$$

Vi får då minimalpolynomet för J :

$$m_A(x) = m_J(x) = (x - 2)^3(x - 3)^2.$$

8 Jordans normalform

Med hjälp av karakteristiska och minimalpolynom ska vi i detta kapitel studera Jordans normalform för en kvadratisk matris A och även finna en bas bestående av generaliserade egenvektorer till A .

Definition 8.1 Låt A vara en kvadratisk matris med egenvärden $\lambda_1, \lambda_2, \dots, \lambda_k$. En övertriangulär Jordanblock $J_k(\lambda)$ är på formen

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

För delmatriser $J_1(\lambda)$ av storlek 1×1 gäller att $J_1(\lambda) = \lambda$.

Generellt är Jordanmatrisen J en blockdiagonal matris av formen

$$J = \begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_k}(\lambda_k) \end{pmatrix},$$

där $n_1 + \cdots + n_k = n$. Skalärerna $\lambda_1 \dots \lambda_k$ kan upprepas. Om $k = n$ dvs $n_1 = \dots = n_k = 1$ skulle J bli en diagonalmatris.

Från kursen linjär algebra II vet vi att en kvadratisk matris A är diagonaliserbar över de reella eller komplexa talen om det finns en inverterbar matris X sådan att $A = XDX^{-1}$, där D är en diagonalmatris. Diagonaliserbarhet är dock inte alltid möjligt för alla kvadratiske matriser. Frågan är då hur vi kan hantera alla kvadratiske matriser. Lösningen är att använda Jordans normalform, eftersom Jordans normalform alltid existerar för en matris A över $\mathbb{C}^{n \times n}$. Vi undersöker dessa påståenden nu.

Definition 8.2 En bas för V kallas en *Jordanbas* för A om A har matrisen J i denna bas.

Sats 8.1 Låt A vara en $n \times n$ -matris. Då finns det en bas för V som är en Jordanbas för A .

Bevis. Det räcker att visa att satsen gäller för en nilpotent matris eftersom, för olika egenvärden $\lambda_1, \dots, \lambda_m$ till A ,

$$V = GE_{\lambda_1} \oplus \cdots \oplus GE_{\lambda_m},$$

där varje $(A - \lambda_k)|_{GE_{\lambda_k}}$ är nilpotent (Sats 3.1). Så någon bas till GE_{λ_k} är en Jordanbas för $(A - \lambda_k)|_{GE_{\lambda_k}}$. Alla sådana baser tillsammans utgör en bas för V som är Jordanbas för A .

Vi delar beviset i några steg. Låt N vara nilpotent.

P1: Låt vektorn $v \neq 0$. Om det finns ett positivt heltal k sådant att $N^k v = 0$ men $N^{k-1}v \neq 0$, då är $v, Nv, \dots, N^{k-1}v$ linjärt oberoende.

Vi antar motsatsen att $v, Nv, \dots, N^{k-1}v$ är linjärt beroende. Då finns det en minimal icke-trivial linjär kombination som är lika med 0.

$$\alpha_1 v + \alpha_2 Nv + \cdots + \alpha_k N^{k-1}v = 0$$

där vi kan anta att k är minimalt så att $\alpha_j \neq 0$ för alla j . Annars kan vi ta bort den motsvarande vektorn i kombinationen. Multiplicera N från vänster i denna ekvation fås

$$\alpha_1 Nv + \alpha_2 N^2 v + \cdots + \alpha_{k-1} N^{k-1}v = 0.$$

Självklart får vi en icke-trivial linjär kombination med färre vektorer som är lika med 0. Det motsäger att k är det minst. Det visar att $v, Nv, \dots, N^{k-1}v$ är linjärt oberoende.

Mer allmänt har vi:

P2: Låt vektorerna $v_k \neq 0$, $k = 1, \dots, s$. Antag att det finns positiva heltal $k_1, \dots, k_s$ sådana att $N^{k_i}v_i = 0$ men $N^{k_i-1}v_i \neq 0$. Om vektorerna $N^{k_1-1}v_1, \dots, N^{k_s-1}v_s$ är linjärt oberoende, är vektorerna i följande kedjor

$$\begin{array}{ccccccc} N^{k_1-1}v_1, & N^{k_1-2}v_1, & \dots, & Nv_1, & v_1, \\ N^{k_2-1}v_2, & N^{k_2-2}v_2, & \dots, & Nv_2, & v_2, \\ \dots & \dots & & & \\ N^{k_s-1}v_s, & N^{k_s-2}v_s, & \dots, & Nv_s, & v_s \end{array}$$

linjärt oberoende.

Beviset för P2 följer samma idé som det för P1. Antag att en minimal icke-trivial linjär kombination av vektorerna i kedjorna ovan är noll. Observera att multiplicera N från vänster i den linjära kombinationen skiftar alla vektorer ett steg till höger från kedjorna. Genom att multiplicera upprepade gånger N från vänster tills en vektor i den linjära kombinationen ovan blir noll. Då har vi fått en linjär kombination med färre vektorer som är noll, en motsägelse mot minimalitet av den ursprungliga linjära kombinationen.

Till sist bevisar vi:

P3: Antag att $N^m = 0$ för något heltal $m \geq 1$. Det finns en bas för V på formen

$$\begin{array}{ccccccc} N^{k_1-1}v_1, & N^{k_1-2}v_1, & \dots, & Nv_1, & v_1, \\ N^{k_2-1}v_2, & N^{k_2-2}v_2, & \dots, & Nv_2, & v_2, \\ \dots & \dots & & & \\ N^{k_s-1}v_s, & N^{k_s-2}v_s, & \dots, & Nv_s, & v_s, \end{array}$$

där $N^{k_i}v_i = 0$ för $1 \leq i \leq s$.

Vi presenterar ett bevis med hjälp av induktion på $\dim V = n$. Om $n = 1$ är $N = 0$ den enda nilpotenta matrisen. P3 gäller.

Antag nu $n > 1$ och P3 gäller för all vektorrum vars dimension är mindre än n .

Utan inskränkning kan vi anta att $\{0\} \subset \mathcal{R}(N) \subset V$ eftersom om $\mathcal{R}(N) = V$ skulle medföra att $\mathcal{R}(N^m) = \dots = \mathcal{R}(N) = V$, en motsägelse; Och om $N = 0$ är resultatet trivialt. Tillämpa induktionsantagandet på restriktion av N på $\mathcal{R}(N)$ kan vi hitta vektorer $u_1, \dots, u_s \in \mathcal{R}(N)$ så att

$$\begin{array}{ccccccc} N^{k_1-1}u_1, & N^{k_1-2}u_2, & \dots, & Nu_1, & u_1, \\ N^{k_2-1}u_2, & N^{k_2-2}u_2, & \dots, & Nu_2, & u_2, \\ \dots & \dots & & & \\ N^{k_s-1}u_s, & N^{k_s-2}u_s, & \dots, & Nu_s, & u_s, \end{array}$$

är en bas för $\mathcal{R}(N)$ och $N^{k_i}u_i = 0$ för $1 \leq i \leq s$.

För $1 \leq i \leq s$ väljer $v_i \in V$ så att $Nv_i = u_i$. Enligt konstruktionen och P1 är vektorerna $N^{k_1-1}u_1, \dots, N^{k_s-1}u_s$ i $\mathcal{N}(N)$ linjärt oberoende. Utvidga dessa till en bas i $\mathcal{N}(N)$ genom $w_1, \dots, w_m$. Vi påstår att

$$\begin{array}{ccccccc} N^{k_1}v_1, & N^{k_1-1}v_1, & \dots, & Nv_1, & v_1, \\ N^{k_2}v_2, & N^{k_2-1}v_2, & \dots, & Nv_2, & v_2, \\ \dots & \dots & & & \\ N^{k_s}v_s, & N^{k_s-1}v_s, & \dots, & Nv_s, & v_s, \\ w_1, & w_2, & \dots & w_m \end{array} \tag{JC}$$

är en bas för V . Att de är linjärt oberoende kan bevisas på samma sätt som P1 och P2. För att visa att de spänner upp hela V kan vi räkna på dimensioner. Vi vet att $\dim \mathcal{N}(N) = s + m$ och $\dim \mathcal{R}(N) = k_1 + \dots + k_s$.

Enligt dimensionssatsen

$$\dim V = (k_1 + 1) + (k_2 + 1) + \cdots + (k_s + 1) + m$$

vilket är antalet vektorer i den linjärt oberoende kedjan (JC). Vi har således konstruerat en bas för V som är en Jordanbas för en nilpotent matris. Då har vi bevisat Sats 8.1. ■

Låt X vara den matris vars kolonner är vektorerna i (JC). Vi kan då skriva:

$$A = XJX^{-1}$$

där J är Jordan normalform. Om A är diagonaliserbar, så blir J en diagonalmatris liknande D , men i allmänhet är J inte en diagonal matris.

Antag att en matris A är i Jordans normalform och innehåller ett antal Jordanblock längs diagonalen. Antag vidare att egenärdena i A förekommer med multiplicitet a_i och att det största Jordanblocket associerat med λ_i har storlek c_i . Antag även att storlek på Jordanblock associerat med λ_i är g_i ,

Då gäller följande villkor:

$$(1) \chi(t) = \prod_{i=1}^k (t - \lambda_i)^{a_i}$$

$$(2) m(t) = \prod_{i=1}^k (t - \lambda_i)^{d_i}$$

$$(3) \dim \mathbf{N}(A - \lambda_i I) = g_i$$

Här ser vi att minimalpolynomet ger mer information än det karakteristiska polynomet eftersom minimalpolynomet ger information om Jordans struktur hos en linjär operator.

Exempel 8.1 Låt T vara operatoren på $\mathbb{C}^5$ som med avseende på standardbasen har matrisen

$$A = \begin{pmatrix} 5 & -2 & 1 & -1 & 3 \\ 1 & 3 & 1 & -2 & 0 \\ -1 & 0 & 2 & 1 & 0 \\ 2 & -1 & 1 & 1 & 2 \\ -1 & 2 & 0 & -1 & 0 \end{pmatrix}$$

Vi ska nu bestämma en Jordanbas för T .

T :s karakteristiska polynom som redan beräknat är

$$\chi_A(t) = (t - 3)(t - 2)^4.$$

Vi ska nu beräkna minimalpolynomet för T vilket har ekvationen

$$m_T(t) = (t - 3)(t - 2)^n,$$

där exponenten n uppfyller olikheten $1 \leq n \leq 4$.

Vi kan bestämma n genom att successivt beräkna $(A - 3I)(A - 2I)^j$ för $j \geq 1$.

$$A - 3I = \begin{pmatrix} 2 & -2 & 1 & -1 & 3 \\ 1 & 0 & 1 & -2 & 0 \\ -1 & 0 & -1 & 1 & 0 \\ 2 & -1 & 1 & -2 & 2 \\ -1 & 2 & 0 & -1 & -3 \end{pmatrix}$$

$$A - 2I = \begin{pmatrix} 3 & -2 & 1 & -1 & 3 \\ 1 & 1 & 1 & -2 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 2 & -1 & 1 & -1 & 2 \\ -1 & 2 & 0 & -1 & -2 \end{pmatrix}$$

Efter lite beräkning har vi kommit fram till att $(A - 2I)(A - 3I) \neq 0$.

$$(A - 2I)^2 = \begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & -1 \end{pmatrix}$$

$$(A - 3I)(A - 2I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Det första värde för vilket denna produkt är 0 är $j = 2$, vilket betyder att $n = 2$. Vi kan dra slutsatsen att eftersom den algebraiska multipliciteten för $(A - 2I)^2$ är 2, är den största Jordan-block av storlek 2×2 .

Om för varje jordanblock där algebraiska multipliciteten sammanfaller med det geometiska har vi en diagonaliserbar jordanblock vilket gäller för $\lambda = 3$ där $\dim N(A - 3I) = 1$.

Nu Sätter vi $W_1 = N(T - 3I)$ och $W_2 = N((T - 2I)^2)$ som två invarianta delrum och då är $\mathbb{C}^5 = W_1 \oplus W_2$, och då är T fullständigt reducerad av delrummen W_1, W_2 .

Eftersom $\dim N(T - 2I) = 2$ har vi att det finns 2 Jordanblock av längd 2.

Vi kan nu ställa upp vår matris J enligt följande:

$$J = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Låt oss nu beräkna hur matrisen X ser ut.

Om $A = XJX^{-1}$, vad är då AX ?

Multiplitera båda sidor med X från höger:

$$AX = XJX^{-1}X.$$

Eftersom $X^{-1}X = I$ (identitetsmatrisen), förenklas detta till:

$$AX = XJ.$$

Vi har matrisen A :

$$A = \begin{pmatrix} 5 & -2 & 1 & -1 & 3 \\ 1 & 3 & 1 & -2 & 0 \\ -1 & 0 & 2 & 1 & 0 \\ 2 & -1 & 1 & 1 & 2 \\ -1 & 2 & 0 & -1 & 0 \end{pmatrix}$$

och matrisen J :

$$J = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Vi antar att $X = (v_1, v_2, v_3, v_4, v_5)$, där kolonnerna v_1, v_2, v_3, v_4, v_5 är egenvektorer eller generaliserade egenvektorer av A .

Eftersom $A = XJX^{-1}$ är ekvivalent med $AX = XJ$.

Detta innebär att varje kolonnvektor i X , alltså v_1, v_2, v_3, v_4, v_5 , multiplicerad med A uppfyller följande relationer, baserat på strukturen hos J :

$$A \begin{pmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \end{pmatrix} = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Sammanfattningsvis har vi:

$$Av_1 = 3v_1, \quad Av_2 = 2v_2, \quad Av_3 = 2v_3 + v_2, \quad Av_4 = 2v_4, \quad Av_5 = v_4 + 2v_5.$$

Vi har följande relationer för vektorerna:

$$Av_1 = 3v_1 \Leftrightarrow (A - 3I)v_1 = 0$$

Eftersom att $(A - 3I)v_1 = 0$ är ett endimensionellt delrum har vi endast följande egenvektor:

$$v_1 = (-1, 1, 1, 0, 1).$$

$$Av_2 = 2v_2 \Leftrightarrow (A - 2I)v_2 = 0$$

Vi får att nollrummet spänns upp av två vektorer, som vi väljer att döpa på följande sätt:

$$v_2 = (1, 1, 0, 1, 0), v_4 = (0, 1, -1, 0, 1).$$

$$Av_3 = 2v_3 + v_2 \Leftrightarrow (A - 2I)v_3 = v_2.$$

För att hitta v_3 som är en generaliserade egenvektor av v_2 , löser vi ekvationen $(A - 2I)v_3 = v_2$.

$$\begin{pmatrix} 3 & -2 & 1 & -1 & 3 \\ 1 & 1 & 1 & -2 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 2 & -1 & 1 & -1 & 2 \\ -1 & 2 & 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} v_{3,1} \\ v_{3,2} \\ v_{3,3} \\ v_{3,4} \\ v_{3,5} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Denna ger oss lösningen:

$$x_1 = x_4$$

$$x_2 = x_4 + x_5$$

$$x_3 = 1 - x_5$$

x_4, x_5 är fria variabler.

vi får då

$$v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

När vi löser $Av_5 - 2v_5 = v_4$ på samma sätt får vi

$$x_1 = 1 + x_4$$

$$x_2 = 1 + x_4 + x_5$$

$$x_3 = -1 - x_5$$

x_4, x_5 är fria variabler.

vi får då

$$v_5 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

Vi kan nu ställa upp matrisen X med vektorerna $v_1, \dots, v_5$:

$$X = \begin{pmatrix} -1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

$$A = \begin{pmatrix} -1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 2 & -1 & 1 & -1 & 2 \\ 1 & -1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 & -1 \end{pmatrix}.$$

9 Tillämpning i ODE

Det finns många tillämpningar av linjära avbildningar och dess egenvärden och egenvektorer. Valet av exempel i detta avsnitt är motiverad av att förstå dynamiska beteenden som beskrivs genom ordinära differentialekvationer. Vi lärde oss tidigt att vi kan lösa n -te ordnings linjära homogena ekvationer genom det karakteristiska ekvationen:

$$x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \cdots + a_1x'(t) + a_0x(t) = 0$$

där x är en funktion av t som ska lösas och $a_0, \dots, a_{n-1}$ är konstanter.

Denna typ av ekvationer i datorberäkningar, t ex MatLab, ofta behöver skrivas om till ett system av linjär system av första ordningens differentialekvationer. Detta görs genom att införa tillståndsvariabler:

$$x_1 = x, x_2 = x', \dots, x_n = x^{(n-1)}.$$

Då blir ett system av första ordningens ekvationer

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_{n-1}}{dt} \\ \frac{dx_n}{dt} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix}}_A \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}}_{\mathbf{x}},$$

eller ekvivalent på matrisformen

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}.$$

I många tillämpningar får vi också på linjära system där variablerna är kopplade till varandra. I sådana fall kan systemet inte skrivas som en enda n -te ordningens differentialekvation. Vi har i Matematik II Linjär algebra lärt oss att använda diagonalisering av matrisen A . Här ger vi ett exempel där systemmatrisen A inte kan diagonaliseras.

Exempel 9.1 Lös följande differentialekvationer för givet begynnelseorden $x_1(0), \dots, x_n(0)$.

Först skriver vi ekvationerna på matrisformen

$$\begin{aligned} \frac{dx_1}{dt} &= 5x_1 - 2x_2 + x_3 - x_4 + 3x_5, \\ \frac{dx_2}{dt} &= x_1 + 3x_2 + x_3 - 2x_4, \\ \frac{dx_3}{dt} &= -x_1 + 2x_3 + x_4, \\ \frac{dx_4}{dt} &= 2x_1 - x_2 + x_3 + x_4 + 2x_5, \\ \frac{dx_5}{dt} &= -x_1 + 2x_2 - x_4. \end{aligned}$$

Låt

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$$

och

$$A = \begin{pmatrix} 5 & -2 & 1 & -1 & 3 \\ 1 & 3 & 1 & -2 & 0 \\ -1 & 0 & 2 & 1 & 0 \\ 2 & -1 & 1 & 1 & 2 \\ -1 & 2 & 0 & -1 & 0 \end{pmatrix}.$$

Då har vi att

$$x' = Ax,$$

där A är den givna matrisen.

Vi inser nu att matrisen A är samma som den i Exempel 8.1. Vi har redan fått Jordan normalform till följande matris:

$$J = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Nu visar vi $\mathbf{e}^{At} = X\mathbf{e}^{Jt}X^{-1}$ genom definitionen av matrisexponentialen.

$$\mathbf{e}^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}.$$

Vi ersätter $A = XJX^{-1}$ där X är uträknad i föregående avsnitt.

$$\mathbf{e}^{At} = \sum_{k=0}^{\infty} \frac{(XJX^{-1}t)^k}{k!}.$$

Eftersom $X^{-1}X = I$, identitetsmatrisen, och X samt X^{-1} är konstanta matriser som inte beror på k , kan dessa brytas ut ur summan, vilket lämnar endast Jt i mitten.

Då får vi

$$= X \sum_{k=0}^{\infty} \frac{(Jt)^k}{k!} X^{-1}$$

Uttrycket inom parentes är definitionen av $\mathbf{e}^{Jt}$:

$$\mathbf{e}^{Jt} = \sum_{k=0}^{\infty} \frac{(Jt)^k}{k!}.$$

Slutligen har vi

$$\mathbf{e}^{At} = X\mathbf{e}^{Jt}X^{-1}.$$

För att lösa detta system av ordinära differentialekvationer med hjälp av Jordanform, börjar vi med att bestämma Jordanformen av matrisen A . Om J är Jordanformen och X är en invers matris som omvandlar A till J , så gäller:

Vi byter variabel:

$$y = X^{-1}x \Rightarrow y' = X^{-1}x' = X^{-1}XJX^{-1}x \Rightarrow y' = Jy$$

Systemet kan sedan skrivas som:

$$y' = Jy,$$

Lösningen blir då

$$y(t) = e^{Jt}y(0).$$

Vi kom fram till att Jordanformen J är:

$$J = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Då beräknas e^{Jt} :

$$e^{Jt} = \begin{pmatrix} e^{3t} & & & & \\ & e^{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}t} & & & \\ & & e^{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}t} & & \\ & & & e^{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}t} & \\ & & & & e^{2t} \end{pmatrix}.$$

För varje $e^{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}t}$ har vi att jordanblocket kan skrivas om som

$$\mathbf{J} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 2I + \mathbf{N}.$$

Vidare får vi

$$e^{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}t} = e^{\left(\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right)t}$$

Vi använder potenslagarna och skriver om

$$= e^{\left(\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right)t} = e^{\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}t} \cdot e^{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}t}$$

Eftersom $2I \cdot \mathbf{N} = \mathbf{N}(2I)$ får vi då

$$= e^{2t} \cdot e^{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}t}$$

Eftersom att vi har en nilpotent matris ($\mathbf{N}^2 = 0$) har vi enligt Taylorserien för

$$\mathbf{e}^{\mathbf{N}t} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

$$\mathbf{e}^{\mathbf{N}t} = \mathbf{I} + \mathbf{N}t.$$

Den tredje termen blir noll eftersom att vi har en nilpotent matris av grad 2.

Slutligen har vi att

$$\mathbf{e}^{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}t} = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{2t} & te^{2t} \\ 0 & e^{2t} \end{pmatrix}.$$

Vi får då samma resultat för nästa jordanblock och vi får den fullständiga matrisen

$$e^{Jt} = \begin{pmatrix} e^{3t} & 0 & 0 & 0 & 0 \\ 0 & e^{2t} & te^{2t} & 0 & 0 \\ 0 & 0 & e^{2t} & 0 & 0 \\ 0 & 0 & 0 & e^{2t} & te^{2t} \\ 0 & 0 & 0 & 0 & e^{2t} \end{pmatrix}.$$

För att beräkna detta i termer av x har vi att $x(t) = Xy(t) = Xe^{Jt}y(0)$.

Detta medför att lösningen i termer av x blir:

$$x(t) = X \begin{pmatrix} e^{3t} & 0 & 0 & 0 & 0 \\ 0 & e^{2t} & te^{2t} & 0 & 0 \\ 0 & 0 & e^{2t} & 0 & 0 \\ 0 & 0 & 0 & e^{2t} & te^{2t} \\ 0 & 0 & 0 & 0 & e^{2t} \end{pmatrix} X^{-1}x(0)$$

Vidare får vi då

$$x(t) = \begin{pmatrix} -1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} e^{3t} & 0 & 0 & 0 & 0 \\ 0 & e^{2t} & te^{2t} & 0 & 0 \\ 0 & 0 & e^{2t} & 0 & 0 \\ 0 & 0 & 0 & e^{2t} & te^{2t} \\ 0 & 0 & 0 & 0 & e^{2t} \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 2 & -1 & 1 & -1 & 2 \\ 1 & -1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 & -1 \end{pmatrix} x(0)$$

$$x(t) = e^{2t} \begin{pmatrix} 2t + e^t & -t - e^t + 1 & t & -t & 2t + e^t - 1 \\ 2t - e^t + 1 & e^t & t & -2t & t - e^t + 1 \\ 1 - e^t & -t + e^t - 1 & 1 & t & t - e^t + 1 \\ 2t & -t & t & 1 - t & 2t \\ 1 - e^t & t + e^t - 1 & 0 & -t & -t - e^t + 2 \end{pmatrix} x(0).$$

10 Diskussion

Nu reflekterar vi frågan vi ställde i introduktion. Varför är linjär algebra bättre utan determinanter? Axler har en mycket stark åsikt om det. Han skriver i sin artikel Down with Determinants¹ "*linear algebra can be done better without determinants*". Vi har försökt förstå en determinant-fri behandling av linjära operatorer. Vi får en fördjupade förståelse för linjära operatorer genom detta arbete, som i stor utsträckning är baserad på determinat-fri argument. Som en determinant-inskolad student har jag för lite kunskaper för att bedöma för- och/eller nackdelar. En fördel kan direkt tolkas som nackdel eller vice versa. Nedan är en kort diskussion om detta.

En av uppenbarliga fördelar kan jag se med determinant-fri teori är att studiera linjära operatorer utan att involvera determinanter kan sannolikt direkt generalisera till oändligtdimensionella vektorrum där determinanter inte finns. I vissa sammanhang ser jag också att beräkna egenvärden kan kosta mindre arbete i term av beräkningar. Vi illustrerade detta i ett exempel där det visserligen behöver bara Gauss-elimination. Men om vi tittar på hur determinanter beräknas i praktiken gör man ändå med Gauss-elimination för att få en triangulärmatris vars determinant är trivial.

Att beräkna egenvärde är ett centralt objekt i numeriska metoder där man strävar också efter numerisk stabilitet dvs avrundningar eller störningar i flyttal-världen är under kontroll med ett enkelt uttryck.

I själva verket beräknar Matlab rötter till n -te grads polynomekvation genom beräkning av egenvärden till en speciell matris.

$$C = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

Med induktion kan vi visa att det karakteristiska polynomet är

$$\chi_C(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

och det är också det minimalpolynom med hjälp av minimalpolynom för vektor relativ matrisen C (t ex väljer vi standardbasvektorn $e_1 = (1, 0, \dots, 0)$).

Matrisen C är redan en övre-hessenbergmatris till vilken man brukar förbehandla innan QR-metoden appliceras. Så vi kan använda denna struktur för att skapa effektiva beräkningsrutiner för beräkning av egenvärden. I denna avseende är determinanter som leder till polynomekvation helt onödiga. Det motsätter inte att determinanter gör linjär algebra sämre må man tycka.

När en diskussion handlar om beräkningar med dator är det nästan alltid så att de vi har lärt oss räkna på kurser är nästan aldrig bra metod numeriskt. Det är inte något bevis på att matematisk teori inte är viktig. Utan att kunna matematik kan många fel i datorberäkningar inte upptäcka eller förstå.

Referenser

- [1] Sheldon Axler. *Linear algebra done right*. 4 ed. Cham: Springer, 2014. ISBN: 9783319110790.
- [2] Hasse Carlsson. *Föreläsningssanteckningar, Linjär algebra II*. Version 2013. Hämtad 01/01-2025 från <https://www.math.chalmers.se/~hasse/LinalgII.pdf>. 2013.
- [3] Stephen H. Friedberg, Arnold J. Insel och Lawrence E. Spence. *Linear algebra*. Fourth edition. Harlow, Essex: Pearson, 2014. ISBN: 9781292026503.