SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Transformationer på invarianta delrum och Jordans normalform

av

Peter Fogelberg

2020 - No K38

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Transformationer på invarianta delrum och Jordans normalform

Peter Fogelberg

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Gregory Arone

2020

Innehåll

1 Introduktion

Denna kandidatuppsats handlar om att undersöka hur vi kan beskriva linjära transformationer på ett vektorrum i så enkla termer som möjligt. Vi ska visa att linjära transformationer $T : V \to V$ kan delas upp i direkta summor av transformationer som verkar på invarianta delrum. Uppdraget är att hitta direkta summor av så rudimentära transformationer som möjligt. Vi vill hitta delrum där vi kan arbeta i varje delrum separat och därmed dela upp transformationen i enklare beståndsdelar. Matrisanalogin är att vi vill hitta baser för V där kvadratiska matriser har en så enkel form som möjligt. Den enklaste icke triviala matrisen är diagonalmatrisen. Därefter har vi under- och övertriangulära matriser. Vi ska visa för vilka transformationer som vi kan beskriva dess matrisrepresentation som diagonaliserbar eller som på Jordans normalform.

2 Linjära transformationer

Vi kommer i detta arbete att varva mellan att beskriva våra transformationer i termer av linjära operationer och med dess matrisframställning i en ordnad bas. När vi talar om linjära transformationer kommer vi fokusera på linjära transformationer $T : V \to V$. Detta innebär att vi arbetar med linjära transformationer som har samma vektorrum som både defintionsmängd och målmängd. En sådan linjär transformation kallas för en endomorfism. Vi kommer i detta arbete använda benämningen transformation och operation växelvis, och i detta arbete benämner de samma sak.

Matrisrepresentationen för en endomorfism utgår från en ordnad bas för V, och där vi för ett ändligdimensionellt vektorrum V över en godtycklig kropp $\mathbb{K}^n$ utgår från standardbasen $\{e_1, ..., e_n\}$. Transformationsmatrisen vid en endomorfism uttrycks i en och samma bas. Detta skiljer sig från transformationer $R : V \to Y$, då R är en transformation från vektorrummet V till Y, där $V \neq Y$. Då kommer istället transformationen R behöva uttryckas i både en bas för V och en för Y.

När vi kommer att arbeta med olika baser är detta i samband med basbyte. Basbytet uttrycker en operation $T : V \to V$ i två olika baser för V. Då arbetar vi i samma vektorrum, men basbytet kan visa olika aspekter av samma transformation. Vi antar två baser, $E = \{e_1, ..., e_n\}$ och $G = \{g_1, ..., g_n\}$. Vi låter A vara matrisrepresentationen av T i en basen E. Ett basbyte skrivs på formen $A = P^{-1}BP$. Vi har att kolumnerna i matriserna P och P^{-1} är koordinaterna för basen G men uttryckt i basen E. B är operation A men uttryckt i den *nya* basen $\{g_1, ..., g_n\}$. Det basbytet vi framförallt kommer att fokusera på i början av uppsatsen är det basbytet när basbytesmatrisen P består av egenvektorer till A. När vi i arbetet uttrycker transformationen T i en ordnad bas kommer vi konsekvent benämna matrisrepresentation som A. Denna bas behöver som sagt inte vara standardbasen.

3 Egenvektorer och egenvärden

En egenvektor $v \neq 0$ är en vektor som uppfyller sambandet att $Tv = \lambda v$, där T är en linjär operator och λ är en skalär som kallas egenvärdet till T. Vi är framförallt intresserade av egenvärden över kroppen av reella tal, alltså $\lambda \in \mathbb{R}$.

Egenvektorer är de vektorer som behåller sin riktning vid en linjär operation T och endast förändras genom en skalningsfaktor λ. Om egenvärdet λ till en egenvektor är 1 är egenvektorn oförändrad efter den linjära operationen.

Om T har en egenvektor v med tillhörande egenvärde λ, kan vi skriva om $Tv = \lambda v$ till $(\lambda I - T)v = 0$. Eftersom operationen $(\lambda I - T)$ skickar v till 0 måste v tillhöra nollrummet till $(\lambda I - T)$. Nollrummet skriver vi som $\ker(\lambda I - T)$ och eftersom $v \in \ker(\lambda I - T)$ innebär detta att om det finns en egenvektor till egenvärdet λ är nollrummet $\ker(\lambda I - T)$ icke-trivialt.

Enligt dimensionssatsen har vi att $\dim(\lambda I - T) = \text{rang}(\lambda I - T) + \dim(\ker(\lambda I - T))$. Vi vet att $(\lambda I - T)$ måste ha full rang för att vara inverterbar och eftersom nollrummet till $(\lambda I - T)$ är icke-trivialt måste $\det(\lambda I - T) = 0$. Detta ger oss möjlighet att räkna ut egenvärdena med hjälp av uträkning av determinanten och i sin tur hitta egenvektorerna till den linjära operationen.

Vi låter A vara matrisrepresentation för T i en ordnad bas. Det är ett faktum att $(\lambda I - A)$ är inverterbar om och endast om $(\lambda I - T)$ är inverterbar. Som vi tidigare nämnt är $(\lambda I - A)$ inte inverterbar om λ är ett egenvärde till A. Vi antar en godtycklig variabel x som kan anta komplexa värden och för stunden får ersätta λ. Vi skriver $(xI - A)$, där dimension på identitetsmatrisen I är samma som dimensionen på A och vi erhåller matrisen:

$$A = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{bmatrix}$$

$$xI - A = \begin{bmatrix} x - y_{11} & -y_{12} & \cdots & -y_{1n} \\ -y_{21} & x - y_{22} & \cdots & -y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -y_{n1} & -y_{n2} & \cdots & x - y_{nn} \end{bmatrix}$$

$(xI - A)$ är en matris med polynom som element i diagonalen. När vi beräknar $\det(xI - A)$ får vi ett moniskt polynom $f(x)$ där graden av $f(x)$ är densamma som dimension för A. För det tredimensionella fallet är detta lätt att se eftersom vi kan räkna ut en 3×3 matris med hjälp av Sarrus regel, då man först multiplicerar alla termer i diagonalen. Detta ger oss precis ett moniskt polynom där högsta graden är precis samma som antal termer i diagonalen. För fallet när matrisen $xI - A$ är en $n \times n$ matris kan vi med hjälp av Laplaceutveckling inse att det karakteristiska polynomet är moniskt. Om vi Laplaceutvecklar hela tiden från första raden och första kolumnen i $xI - A$ och de kofaktorsmatriser som uppstår, inser vi att får

$$(x - y_{11})(x - y_{22})...(x - y_{n-1n-1})(x - y_{nn}) + ... = x^n + ...$$

Eftersom det karakteristiska polynomet för A har grad n är x^n högsta termen i polynomet och därmed är $f(x)$ moniskt.

Om vi sätter $\det(xI - A) = 0$ vet vi att om A enbart har ett egenvärde, har vi att $\det(\lambda I - A) = 0$. Vi har också slagit fast att $f(x) = \det(xI - A)$ och vi har då att egenvärdet λ är en rot till polynomet

$f(x)$ så att $f(\lambda) = 0$. Detta polynom $f(x)$ med egenskapen att egenvärdet λ är en rot till polynomet kallas för det karakteristiska polynomet till A.

Vi är intresserade av egenvektorer eftersom dessa är sinsemellan linjärt oberoende. Uppsättningen egenvektorer kopplade till ett egenvärde för matris A bildar ett delrum till vektorummet V.

Definition: Låt λ vara ett egenvärde till transformationen $T : V \to V$. Uppsättningen egenvektor kopplat till egenvärdet λ bildar ett egenrum W som är ett delrum till V, så att $W \subseteq V$. Detta är samma sak som att skriva att egenrummet W är nollrummet till den linjära transformation $(\lambda I - T)$, så att $W = \ker(\lambda I - T)$.

Definition: Två $n \times n$ matriser A och D är similära om D representerar A i basen bestående av kolumnerna i någon matris P. Detta är samma som att skriva A på formen:

$$A = P^{-1}DP$$

Om det linjärt oberoende egenvektorerna, kopplade till ett eller flera egenvärden, bildar en bas till V säger vi att A är diagonaliserbar. Detta går även att formulera som att en matris är diagonaliserbar om vi kan skriva den kvadratiska matrisen A som $A = P^{-1}DP$, där kolumnerna i P består av egenvektorer till A och D är en diagonalmatris med egenvärdena till A i diagonalen.

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_k \end{bmatrix}$$

Om A är diagonaliserbar kan A skrivas i basen P som en blockmatris där elementen i diagonalen är diagonalmatriser med egenvärdena i diagonalen. Om A har egenvärdet λ_i som rot till det karakteristiska polynomet d_i gånger kommer diagonalelementen till D ha $\dim(\lambda_i I) = d_i$.

$$D = \begin{bmatrix} \lambda_1 I & 0 & 0 & 0 \\ 0 & \lambda_2 I & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_k I \end{bmatrix}$$

Eftersom A är similär med D är $(xI - A)$ similär med $(xI - D)$. Om vi tar $(xI - D)$ får vi för varje egenvärde λ_i

$$(xI - D) = \begin{bmatrix} (x-\lambda_1)I_1 & 0 & 0 & 0 \\ 0 & (x-\lambda_2)I_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & (x-\lambda_k)I_k \end{bmatrix}$$

där dimensionen på I_i är d_i. Eftersom determinanten av en diagonalmatris är produkten av diagonalelementen kan vi se att det karakteristiska polynomet är på formen:

$$f(x) = (x - \lambda_1)^{d_1}...(x - \lambda_k)^{d_k}$$

Om vi multiplicerar $f(x)$ med $(-1)^k$ där $k = d_1 + ... + d_k$, kan vi få det karakteristiska polynomet på formen:

$$\begin{aligned}(-1)^k f(x) &= (-1)^k (x - \lambda_1)^{d_1}...(x - \lambda_k)^{d_k} \\ &= (\lambda_1 - x)^{d_1} + ... + (\lambda_k - x)^{d_k}\end{aligned}$$

Det är framförallt den senare formen av det karakteristiska polynomet som vi kommer referera till när vi talar om det karakteristiska polynomet av en transformation eller matris. Denna form kommer hädanefter refereras till som $f(x)$.

$(xI - A)$ är similär med $(xI - D)$ och därför har A och D samma karakteristiska polynom.

Om vi skriver det karakteristiska polynomet för A på formen:

$$f(x) = (\lambda_1 - x)^{d_1}...(\lambda_i - x)^{d_i}...(\lambda_k - x)^{d_k} = (\lambda_i - x)^{d_i} g(x),$$

då är $g(\lambda_i) \neq 0$ och d_i är den högsta möjliga talet där polynomet $(\lambda_i - x)^{d_i}$ fortfarande delar $f(x)$. Då kallas d_i för den algebraiska multipliciteten för λ_i.

Den geometriska multipliciteten är dimensionen av egenrummet till egenvärdet λ_i, alltså antalet egenvektorer kopplade till det egenvärdet. För att en matris ska vara diagonaliserbar behöver den algebraiska multipliciteten och den geometriska multipliciteten för ett bestämt egenvärde vara lika.

Det kan här även vara värt att nämna att det finns transformationer som saknar reella egenvärden. Den geometriska tolkningen av detta är att inga vektorer under en transformation T behåller sin riktning. Den algebraiska tolkningen är att det karakteristiska polynomet $f(x) = 0$ för T saknar reella lösningar. I det tvådimensionella fallet saknar transformation T, som innebär en rotation $90°$ medsols eller motsols, reella egenvärden. Matrisrepresenationen A för en transformation $90°$ motsols är

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

$$xI - A = \begin{bmatrix} x & 1 \\ -1 & x \end{bmatrix}$$

$$f(x) = \det(xI - A) = x^2 + 1$$

$$f(x) = x^2 + 1 = 0$$

$f(x) = 0$ saknar reella lösningar och därmed saknar T reella egenvärden. Värt att nämna är att för vektorrummet över kroppen av komplexa tal finns det alltid komplexa egenvärden med tillhörande komplexa egenvektorer.

4 Annihilerande polynom

Vi låter polynomet $f(x)$ vara ett polynom i ringen av polynom $R[x]$ med koefficienter i kroppen av reella tal $\mathbb{R}$. Vi har att $f(x) = a_0 + a_1x + ... + a_nx^n \in R[x]$. Vi låter $f(T)$ vara en linjär avbildning på formen: $f(T) = a_0 + a_1T + ... + a_nT^n$, där $\{a_0, .., a_n\} \in \mathbb{R}$ och är skalärer till T och varje multipel av T anger hur många gånger transformationen T ska utföras.

Alla polynom som tar en operation till nollvektorn, genom att $h(T) = b_0 + b_1T + ... + b_nT^n = 0$, där $\{b_0, .., b_n\} \in \mathbb{R}$, kallas för annihilerande polynom. Dessa polynom bildar ett ideal över ringen av polynom $R[x]$. Vi noterar detta ideal som $M[x]$. Ett ideal av annihilerande polynom innebär att det annihilerande polynomen är slutna under addition och multiplikation så att:

i) $h + g$ tillhör $M[x]$ när $h, g \in M[x]$.
ii) $k \cdot g$ tillhör $M[x]$ när $k \in R[x]$ och $g \in M[x]$.

Ringen av alla polynom $R[x]$ över kroppen av reella tal, är vad som kallas ett principialidealdomän. Detta innebär att det finns ett polynom av minsta grad som som genererar alla andra polynom i $R[x]$. Eftersom $R[x]$ är ett principialidealdomän är alla ideal i $R[x]$ principialideal.

Eftersom $M[x]$ är ett ideal över ringen av polynom $R[x]$ kan vi konstatera att $M[x]$ är ett principialideal. Vi har då ett moniskt polynom m som generar hela idealet $M[x]$. Vi är intresserade av det principialideal $M[x]$ som innehåller alla polynom $h(T) = 0$. Detta ideal genereras av det moniska polynomet m som vi kallar för minimalpolynomet. Minimalpolynomet är det polynom av lägst grad som annihilerar T. När vi säger att m genererar $h(T) \in M[x]$, menas att alla polynom h som annihilerar T kan skrivas som en multipel av minimalpolynomet m. När vi skriver h som en multipel av m får vi att:

$$h(T) = m(T) \cdot g(T) = 0$$

där $g(T) \neq 0$.

Varje linjär operation $T : V \to V$ har ett unikt polynom(förutom skalärer) av lägsta grad som annihilerar T och genererar principialidealet $M[x]$.

Vi kommer att använda faktumet att minimalpolynomet delar alla annihilerande polynom, när vi visar att det karakteristiska polynomet också annihilerar T. Detta gör vi lite senare med hjälp av Cayley-Hamilton satsen.

Vi ska undersöka några viktiga egenskaper hos minimalpolynomet:

Lemma: Rötterna till det minimala polynomet är precis egenvärdena till T.

Bevis: Vi ska visa att minimalpolynomet har samma rötter som det karakteristiska polynomet, men kan skilja på multiplicitet, alltså hur många gånger ett polynom har roten c som rot. Eftersom rötterna till det karakteristiska polynomet är egenvärdena till T, ska vi visa att roten c till minimalpolynomet är precis egenvärdet λ till T.

Vi antar att:

$$m(c) = 0$$

Eftersom c är en rot till minimalpolynomet kan vi skriva minimalpolynomet som produkten av en linjär funktion och ett polynom $g(x)$

$$m(x) = (c - x)g(x)$$

Eftersom vi har brutit ut en linjär faktor måste $deg(g(x)) < deg(m(x))$ och eftersom $m(x)$ är minimalpolynomet måste $g(T) \neq 0$. Vi väljer en vektor u så att $g(T)u \neq 0$. Eftersom $g(T)$ enbart är ett polynom med T som variabel, är $g(T)$ också en linjär transformation. Vi sätter $g(T)u = v$.

$$\begin{aligned} 0 &= m(T)u \\ &= (cI - T)g(T)u \\ &= (cI - T)v \end{aligned}$$

Vi har alltså att $(cI - T)v = 0$ vilket innebär att v ligger i nollrummet till $(cI - T)$ och därmed är en egenvektor till T. Vi kan då sluta oss till att c är ett egenvärde till T, så att $c = \lambda$.

För att visa motsatsen säger vi att minimalpolynomet har samma rötter som det karakteristiska polynomet alltså att, $c = \lambda$. Vi får då för egenvektorn v kopplat till egenvärdet λ:

$$m(T)v = m(c)v$$

Eftersom minimalpolynomet annihilerar T, måste $m(T)v = m(c)v = m(\lambda)v = 0$

V.S.B.

Vi har tidigare inte vetat att rötterna till minimalpolynomet är precis rötterna till det karakteristiska polynomet vilket alltså är egenvärdet $c = \lambda$.

Om T kan representeras som en diagonalmatris i en ordnad bas har vi att minimalpolynomet faktoriseras helt till linjära funktioner med distinkta egenvärden som rötter, så att:

$$m(x) = (\lambda_1 - x)...(\lambda_k - x)$$

Detta kan vi se genom att vi antar att v är en egenvektor till T och per definition skickar någon av $(\lambda_1 - T)...(\lambda_k - T)$ v till 0 vektorn. Vi antar $(\lambda_i - T)v_i = 0$, där $1 \leq i \leq k$.

$$(\lambda_1 I - T)...(\lambda_i I - T)...(\lambda_k I - T)v_i$$

Eftersom polynom kommuterar kan vi ändra om det linjära polynomen till:

$$(\lambda_1 I - T)...(\lambda_k I - T)(\lambda_i I - T)v_i$$

$$(\lambda_i I - T)v_i = 0$$

Vilket ger att hela uttrycket blir:

$$(\lambda_1 I - T)...(\lambda_k I - T)(\lambda_i I - T)v_i = 0$$

Om T har egenvektorer som spänner upp en bas för V, vilket definierar att en tillhörande matris är diagonaliserbar, behövs inga multipler av de linjära funktionerna för att skicka v till noll. Om

T är diagonaliserbar kommer det finnas egenvektorer till varje distinkt egenvärde, med andra ord kommer det finnas en egenvektor v_i till T som tillhör nollrummet av $(\lambda_i I - T)$ och där egenvektorn v_i skickas till noll vid transformationen $(\lambda_i - T)v_i$.

Därför kommer minimalpolynomet för en diagonaliserbar operation vara på formen:

$$m(T) = (\lambda_1 - T)...(\lambda_k - T) = 0$$

Om T kan representeras som en triangulärmatris i en ordnad bas är minimalpolynomet på formen:

$$m(x) = (\lambda_1 - x)^{r_1}...(\lambda_k - x)^{r_k}$$

Detta faktum kommer inte bevisas i denna uppsats.

Vi har än så länge visat att minimalpolynomet och det karakteristiska polynomet har samma rötter och att de skrivs på samma form (förutom multipliciteter) för diagonaliserbara och triangulerbara operationer. Vi har även visat att minimalpolynomet delar alla annihilerande polynom. Vi ska nu med hjälp av Cayley-Hamiltons sats visa att det karakteristiska polynomet $f(x)$ annihilerar T så att $f(T) = 0$. Detta innebär också att minimalpolynomet delar det karakteristiska polynomet för T.

5 Cayley-Hamiltons sats

För att kunna bevisa Cayley-Hamiltons sats kommer vi att behöva använda oss av adjunkta matriser. Om vi har en $n \times n$ matris B kan vi räkna ut determinanten av B genom att hitta kofaktorsmatrisen till B. Elementen i kofaktorsmatrisen består av så kallade minorer som är determinanter som erhålls genom att stryka en rad och en kolonn. Dessa minorer eller underdeterminater är av storlek $n-1 \times n-1$. Tecknet framför minoren beror på vilken rad eller kolonn som minoren uträknas ifrån. Kofaktorsmatrisen har alltså dessa minorer som element. En adjunkt matris är sedan transponatet av kofaktorsmatrisen. Adjunktamatriser används framförallt vid uträkning av inversen till en matris. För att bevisa Cayley-Hamiltons sats använder vi sambandet $B^{-1} = \frac{adjB}{|B|}$ vilket vi skriver om som $B \cdot adjB = |B| \cdot I$. Vi kommer också utnyttja det faktum att minoren har en $n-1 \times n-1$ determinant. Vi kommer dock inte behöva ta hänsyn till tecknet framför minoren.

Vi bevisar Cayley-Hamiltons sats med hjälp av matrisrepresentation men satsen gäller även för linjära transformationer.

Sats: Låt A vara en matrisrepresentation för transformationen T över det ändligdimensionella vektorrummet V. Låt f vara det karakteristiska polynomet till A. Då uppfyller A sitt karakteristiska polynom så att $f(A) = 0$, vilket betyder att det minimalpolynomet delar det karakteristiska polynomet.

Bevis:
Vi använder:

$$B \cdot adjB = |B| \cdot I$$

Detta ger oss sambandet att matrisen B vänstermultiplicerat med sin adjungerade matris är en diagonalmatris med determinanten till B som diagonalelement.

Om vi istället utgår från matrisen $(\lambda I - A)$ där λ är någon variabel, och beräknar determinanten får vi: $|\lambda I - A| = v_0\lambda^n + v_1\lambda^{n-1} + ... + v_{n-1}\lambda + v_n = f(\lambda)$

Vi ska visa att A uppfyller sitt karakteristiska polynom och det karakteristiska polynomet även är ett annihilerande polynom:

$$f(A) = v_0A^n + v_1A^{n-1} + ... + v_{n-1}A + v_nI = 0$$

Om vi ersätter B med matrisen $(\lambda I - A)$ i sambandet $B \cdot adjB = |B| \cdot I$ får vi:

$$(\lambda I - A)adj(\lambda I - A) = |\lambda I - A| \cdot I$$

$$HL = |\lambda I - A| \cdot I = v_0\lambda^n I + v_1\lambda^{n-1}I + ... + v_{n-1}\lambda I + v_nI = f(\lambda) \cdot I$$

$$adj(\lambda I - A) = C_1\lambda^{n-1} + C_2\lambda^{n-2} + ... + C_{n-1}\lambda + C_n$$

Där $C_1, ..., C_n$ är matriser med koefficienterna till varje potens av λ som element. Eftersom $adj(\lambda I - A)$ har minorer som element kommer polynomet ha högst grad $n - 1$.

$$VL = (\lambda I - A)adj(\lambda I - A) = (\lambda I - A)(C_1\lambda^{n-1} + C_2\lambda^{n-2} + ... + C_{n-1}\lambda + C_n)$$
$$= (-C_1)\,\lambda^n + (AC_1 - C_2)\lambda^{n-1} + (AC_2 - C_3)\lambda^{n-2} + ... + (AC_{n-1} - C_n)\lambda + AC_n$$

Jämför koefficienterna mellan VL och HL, så att $VL = HL$

$$\begin{aligned} -C_1 &= v_0I \\ AC_1 - C_2 &= v_1I \\ AC_2 - C_3 &= v_2I \\ &\vdots \\ AC_{n-1} - C_n &= v_{n-1}I \\ AC_n &= v_nI \end{aligned}$$

Vi multiplicerar respektive led med $A^n, A^{n-1}, ..., A, I.$;

$$-C_1 \cdot A^n = v_0I \cdot A^n$$

$$
\begin{aligned}
(AC_1 - C_2)A^{n-1} &= v_1 I \cdot A^{n-1} \\
(AC_2 - C_3)A^{n-2} &= v_2 I \cdot A^{n-2} \\
&\vdots \\
(AC_{n-1} - C_n)A &= v_{n-1} I \cdot A \\
AC_n \cdot I &= v_n I \cdot I
\end{aligned}
$$

Vi erhåller:

$$
\begin{gathered}
-A^n C_1 + A^n C_1 - A^{n-1}C_2 + A^{n-1}C_2 - A^{n-2}C_3 + ... + A^2 C_{n-1} - AC_n + AC_n \\
= v_0 A^n + v_1 A^{n-1} + ... + v_{n-1A} + v_n I \\
VL = 0 \\
HL = f(A) \\
f(A) = 0
\end{gathered}
$$

V.S.B.

Eftersom minimalpolynomet är det polynom som genererar alla annihilerande polynom får vi att $f(T) = m(T)g(T) = 0$, där $m(T)$ är minimalpolynomet och $g(T)$ är ett tillhörande polynom. Detta säger oss att minimalpolynomet delar det karakteristiska polynomet för en transformation T. Vi har alltså fått en hel del ledtrådar och användningsområden för minimalpolynomet.

Sammanfattning minimalpolynomet och karakteristiska polynomet:

i) Minimalpolynomet för T har samma rötter som det karakteristiska polynomet för T.

ii) Minimalpolynomet delar det karakteristiska polynomet.

iii) Ett i denna uppsats obevisat faktum är att om minimalpolynomet för T är på formen:

$$m(x) = (\lambda_1 - x)^{d_1} ... (\lambda_k - x)^{d_k},$$

är T triangulerbar.

iiii)Om minimalpolynomet för T är på formen:

$$m(x) = (\lambda_1 - x)...(\lambda_k - x),$$

är T diagonaliserbar.

I nästa avsnitt vill vi undersöka linjära transformationer lite närmare genom att använda av oss av delrum. Vi ska visa att egenrummet för T är ett delrum till V och i synnerhet ett invariant delrum under transformationen T.

6 Delrum

Definition för ett icke trivialt delrum $U \subseteq V$ är att för två vektor $u_1, u_2 \in U$ och för en godtycklig skalär $t \in \mathbb{R}$ har vi att:

$$u_1 + u_2 \in U$$
$$tu \in U \text{ om } u \in U$$

Om vi förutsätter att v och u är egenvektorer till operationen T kopplat till ett distinkt egenvärde λ, och W är egenrummet kopplat till ett distinkt egenvärde har vi att $v, u \in W$. Det är klart att v och u adderat fortfarande är i W eftersom W byggs upp av linjärkombinationer av de egenvektorer som finns i W. Att en egenvektor multiplicerat med en skalär också ligger i W är uppenbart eftersom $Tv = tv = \text{spann}(v)$ och därmed också en egenvektor som ligger i W.

Vi ska visa att egenrummet W inte enbart är ett delrum till V utan också ett delrum som är invariant under en operationen T.

7 Invarianta delrum

Vi kommer att använda begreppet invariant delrum för beskriva en viss sorts delrum som är viktigt för att kunna dela upp vektorrum i mindre beståndsdelar.

Definition: Vi har operation T så att $T : V \to V$ och $U \subseteq V$ och $v \in U$. Om vi utför T på v och erhåller att $Tv \in U$, är U ett invariant delrum under transformationen T.

Ett invariant delrum betyder att ett delrum är slutet under en transformation T, i den mening att om man utför en transformationen T på en vektor i underrummet kommer denna vektor fortfarande vara kvar i samma delrum. Vi säger att vektorn är invariant under T.

Transformationen T på det invarianta delrummet U kallas här för den restriktiva operationen T på U och betecknas T_U. T_U är en operation på vektorrummet U och bibehåller målmängden så att $T_U : U \to U$. T_U är en aspekt av T men kan skilja sig dramatiskt från T eftersom den enkom arbetar i ett delrum. T_U definieras som $Tv = T_U v$ när vektorn v ligger i U.

Exempel: I det tvådimensionella fallet är det enda invarianta delrummen under transformationen T:
i) $\mathbb{R}^2$
i) 0-rummet
iii) Egenvektor av dimension 1.

Om T som agerar på $\mathbb{R}^2$ saknar reella egenvärden, har T enbart det triviala delrummen som invarianta underrum. Vi låter v vara den enda egenvektorn som bygger upp egenrummet W, så att $\text{spann}(v) = W$. Egenvektorn är uppenbart invariant eftersom om v är en egenvektor har vi att $Tv = \lambda v$, vilket betyder att $T_W = \lambda$ och därmed att $Tv \in \text{spann}(v)$.

Vi har alltså att egenrummet är ett invariant delrum till V. Vi betecknar egenrummet som är kopplat till ett distinkt egenvärde λ_i som W_i. När vi arbetar i W_i betyder den linjära operation

T_{W_i} att vi skalar alla vektorer i W_i med en skalningsfaktor λ_i på detta delrum. Vi har än så länge enbart diskuterat endimensionella egenrum men om vi har en $n \times n$ matris med åtminstone ett egenvärde λ_i kan egenrummet W_i ha dimension $1 \leq \dim(W_i) \leq n$.

För förtydligande är egenrummet W_i samma som att skriva $\ker(\lambda_i I - T)$. $\text{Ker}(\lambda_i I - T)$ är i sin tur uppsättningen av alla vektorer v som har egenskapen att $(\lambda_i I - T)v = 0$. Vi har alltså att båda benämningarna beskriver egenrummet till T så att $W_i = \ker(\lambda_i I - T)$.

Exempel: Vi antar att vi har två egenvektorer till operationen T på V där $\dim(V)$= 2. Vi antar också att egenvektorna v_1 och v_2 associeras med varsitt distinkt egenvärde $\lambda_1$och λ_2. Eftersom v_1 och v_2 är egenvektorer är dessa linjärt oberoende och en linjärkombination av dessa vektorer bildar en ny bas för planet. Vi säger att $v_1 \in W_1$ och $v_2 \in W_2$ och att W_1 och W_2 tillsammans fyller ut hela R^2. Operationen T på W_1 respektive W_2 skrivs som den restriktiva operationen T_{W_1} respektive T_{W_2}. Dessa restriktiva operationer verkar enbart på sin tillhörande egenvektor och skalar därför varje egenvektor med egenvärdet för respektive egenvektor. Vi har att $T_{W_1} = \lambda_1$ på W_1 och $T_{W_2} = \lambda_2$ på W_2.

Vi kommer fortsättningsvis enbart använda beteckningen T_i för att beteckna den restriktiva operationen T_{W_i}. Det är precis på grund av att T_i fungerar som en skalningsfaktor på de egenvektor som den tillhör som vi är intresserade av invarianta delrum. Transformationen T kan skalas ner till att arbeta i varje egenrum separat och får därmed en enklare karaktär än när den arbetar på hela V.

8 Direkt summa av invarianta delrum

Vi är intresserade av en uppdelning av vektorrummet V som en direkt summa av invarianta delrum. För att skriva V som en direkt summa av invarianta delrum så att $V = U_1 \bigoplus U_2$ behöver summan av $U_1 + U_2 = V$ och $U_1 \bigcap U_2 = \{0\}$. Att snittet är nollvektorn betyder att uppdelningen är unik och att vektorerna U_1 och U_2 är sinsemellan linjärt oberoende.

Oftast vill vi kunna dela upp V som en direkt summa av fler invarianta delrum än 2. Målet är att dela upp V som en direkt summa av så rudimentära invarianta delrum som möjligt. När vi uttrycker V som en direkt summa av fler än två invarianta delrum skriver vi

$$V = U_1 \bigoplus ... \bigoplus U_k$$

där $\dim V$=$\dim U_1$+...+$\dim U_k$ och där denna uppdelning är unik. Att uppdelningen är unik betyder att $U_i \bigcap (U_1 + ... + U_{i-1} + U_{i+1} + ... + U_k) = \{0\}$ för varje i. Detta innebär att vektorerna i U_i är linjärt oberoende till varje linjärkombination av komplementet till U_i. Eftersom vi kan göra detta för varje U_i där $1 \leq i \leq k$ och summan av dimension för varje invariant delrum tillsammans blir hela V kan vi skriva V som en unik uppdelning av invarianta delrum.

Vi arbetar i vektorrummet V över kroppen $\mathbb{R}^n$ och W_i är egenrummet kopplat till ett distinkt egenvärde λ_i till T. Om matrisrepresentationen för T är diagonaliserbar har vi tillräckligt många egenvektorer för att fylla ut hela V. Eftersom egenrummen är invarianta och egenvektorerna fyller ut hela V kan vi skriva:

$$V = W_1 \bigoplus ... \bigoplus W_k$$

där dim$V$=dimW_1+...+dimW_k=n

Dimensionen på egenrummet beror på antalet egenvektorer kopplat till ett specifikt egenvärde eftersom egenvektorerna är linjärt oberoende. Viktigt att påpeka är också att detta gäller för diagonaliserbara operationer då det finns tillräckligt av egenvektorer för att spänna upp hela V. När vi säger att en matris A, som är matrisrepresentationen av transformation T i en ordnad bas, är diagonaliserbar är detta samma som att säga att V kan delas upp i en direkt summa av egenrum. Varje W_i har en restriktiv operation T_i på egenrummet och vi säger att T är en direkt summa av dessa restriktiva operationer.

9 Matrisframställning invarianta delrum

Vi vill nu visa hur matrisrepresentation för en direkt summa av invarianta delrum ser ut. Vi låter A symbolisera operatorn T i en ordnad bas. Om vi kan skriva V som en direkt summa av invarianta delrum har vi att matrisframställningen av A som opererar på V är:

$$A = \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & A_k \end{bmatrix}$$

Där A är matrisen för T på vektorrummet V och A_i är den restriktiva operationen T_i på det invarianta delrummet U_i.

Vi vill genom denna matrisrepresentation visa att det karakteristiska polynomet $f(x)$ för A, som vi räknar ut genom $f(A) = \det(xI - A)$, delas av det karakteristiska polynomen för respektive A_i. Detta gör vi genom att determinanten för diagonalmatriser är samma som produkten av diagonalelementen.

$$\det(xI - A) = \det \begin{bmatrix} xI - A_1 & 0 & 0 & 0 \\ 0 & xI - A_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & xI - A_k \end{bmatrix}$$

Enligt determinatreglerna har vi att, $\det(xI - A) = \det(xI - A_1)\det(xI - A_2)...\det(xI - A_k)$. Vi skriver $f_i(A_i) = \det(xI - A_i)$. Eftersom A är en blockmatris med enbart element i diagonalen kan vi skriva $f(A) = f_1(A_1)f_2(A_2)...f_k(A_k)$. Vi kan alltså säga att det karakteristiska polynomet för A delas av de respektive karakteristiska polynomen till de invarianta delrummen. Eftersom A är matrisrepresentation för T i en ordnad bas gäller det också att det karakteristiska polynomet för T delas av det karakteristiska polynomet för varje T_i.

Vi undersöker kopplingen mellan minimalpolynomet till A och det individuella minimalpolynomen till varje invariant delrum av A.

$$m(A) = \begin{bmatrix} m(A_1) & 0 & 0 & 0 \\ 0 & m(A_2) & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & m(A_k) \end{bmatrix}$$

För att $m(A) = 0$, vilket är definitionen av en annihilator, måste alla $m(A_1) = m(A_2) = ... = m(A_k) = 0$. Det är endast då vi får nollmatrisen. Vi har att varje polynom som annihilerar A annihilerar också A_i, vilket säger att $m_i(A_i) \mid m(A)$. Detta gäller för varje minimalt polynom till det separata invarianta delrummen. Därför är det tydligt att det finns ett gemensamt polynom $g(x)$ mellan $m_i(x_i)$ och $m_j(x_j)$ så att $g(A_i) = g(A_j) = 0$. Eftersom vi är ute efter minimalpolynomet för A, är det den minsta gemensamma multipel av $m(A_1), m(A_2), ..., m(A_k)$ vi söker. Detta minsta gemensamma polynom är då minimalpolynomet $m(A)$.

I nästa kapitel kommer vi att visa direkta summan av invarianta delrum med hjälp av projektioner. Vi kommer att visa hur man kan beskriva direkta summan av egenrum i termer av projektioner. Detta kommer hjälpa oss när vi sedan ska beskriva hur man än mer generellt kan uttrycka V i termer av invarianta delrum och hur matrisrepresentation för dessa ser ut.

10 Projektioner

Definition: En projektion är en linjär operator E på vektorrummet V som har egenskapen att $E^2(v) = E(v)$, där v är en vektor i V.

Ett annat ord för projektion är att E är idempotent och syftar på precis egenskapen att $E^2 = E$. Detta gäller även att $E^n = E$ då vi alltid kan skriva $E^{n-2}E^2 = E^{n-2}E = ... = E^2E = E^2 = E$. Vi kan med hjälp av projektioner dela upp V i direkta summor. Vi ska nu beskriva V som en direkt summa av linjära höljet av E och dess nollrum.

Vi antar att vi har en projektion E som agerar på V. Vi har S =spann(E) och K =ker(E). Då kan vi skriva V på formen:

$$V = S \bigoplus K$$

Detta betyder att spann(E) och ker(E) utgör en bas för V där alla element i V kan skrivas som en linjärkombination av S och K alltså att $v = Ev + (I - E)v$, när $v \in V$. Om $v \in S$ har vi att $Ev = v$. Om $u \in K$ har vi att $Eu = 0$. Att $Ev = v$ när $v \in S$ beror på att när vi projicera en vektor, som redan ligger på projektionsytan kommer vektorn behålla sin längd och riktning. Om vi har en vektor $u \in K$ kan vi se att $u = Eu + (I - E)u$ där $Eu = 0$ så att enbart $u = Iu = u$.

Operatorn E är en projektion på S parallellt med K. Tvärtom är $I - E$ en projektion på K parallellt med S. Vi testar att $I - E$ också är idempotent.

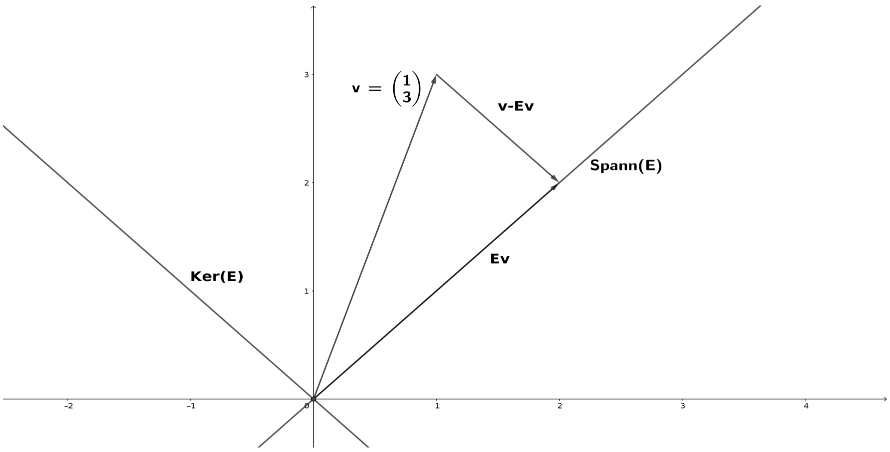
$$(I - E)^2 = I^2 - 2IE + E^2$$

$$I^2 = I \text{ och } E^2 = E \text{ vilket ger:}$$

$$I^2 - 2IE + E^2 = I - 2E + E = I - E$$

Vi får alltså att $(I - E)^2 = I - E$ vilket visar att $I - E$ idempotent och därmed en projektion.

I figur 1 kan vi se att $v = Ev + (I - E)v$ då $Ev = (2, 2)$ och $(I - E)v = (-1, 1)$ erhåller vi $v = (2-1, 2+1) = (1, 3)$, vilket överensstämmer med vektorn v i figuren. Vi kan också se att själva projektionen sker parallellt med nollrummet eftersom $v - Ev = (-1, 1)$. Varje punkt på ker(E) kommer av E projiceras ner mot origo. Det är viktigt att notera att vi också ser att hela V kan skrivas som en linjärkombination av spann(E)och ker(E) eftersom dessa är linjärt oberoende.



Figur 1: Projektion på S parallellt med K

Vi är intresserade av projektioner eftersom de ger oss ett effektivt verktyg för att kunna dela upp ett vektorrum som en direkt summa av delrum (ännu inte nödvändigtvis invarianta delrum), där varje delrum har en tillhörande projektion.

$$V = P_1 \bigoplus ... \bigoplus P_k$$

Här har varje P_i en till sig kopplad projektion E_i så att när man utför projektionen på en vektor $v \in V$ så är $E_i v \in P_i$. Vi har även att om $v_i \in P_i$ så är $E_i v_i = v_i$. $E_j v_i = 0$ om $i \neq j$ eftersom v_i ligger per definition i nollrummet till E_j. Allt detta betyder att $\text{spann}(E_i) = P_i$ eftersom E_i skickar varje vektor i V antingen till P_i eller till nollvektorn.

Eftersom V är en direkt summa av delrum P_i kan vi skriva en vektor $v \in V$ som en linjärkombination, $v = v_1 + ... + v_n$, där $v_i \in P_i$. Eftersom $E_i v = v_i$, kan vi även skriva $v = E_1 v + ... + E_k v$. Vi har alltså att en projektion E_i är en operator från V till P_i. De viktigaste slutsatserna av detta är att:

$$v = E_1 v + ... + E_k v$$

$$v = E_1 v_1 + ... + E_k v_k = v_1 + ... + v_k$$

$$I = E_1 + ... + E_k$$

$$\text{spann}(E_i) = P_i$$

Eftersom vi har en direkt summa kommer summan av komplementet till E_i vara nollrummet till denna operator:

$$\ker(E_i) = P_1 + .. + P_{i-1} + P_{i+1} + ... + P_k$$

Vi har att om vi kan skriva V som en direkt summa av delrum så att $V = P_1 \bigoplus ... \bigoplus P_k$ har vi alltid k projektioner, där varje enskild projektion är kopplat till varje separat delrum. Åt andra hållet, om vi har k projektioner som uppfyller kraven att:

$$E_i^2 = E_i$$

$$E_i E_j = 0$$

$$I = E_1 + ... + E_k$$

$$\text{spann}(E_i) = P_i$$

kan V skrivas som en direkt summa av delrummen $\{P_1, ..., P_k\}$.

Vi är framförallt intresserade av direkta summor av invarianta delrum. Ett invariant delrum är som ovanstående nämnt, ett delrum som har egenskapen att $u \in U$ också leder till att $Tu \in U$. Vi vill skapa en direkt summa av delrum med denna egenskap.

$$V = U_1 \bigoplus ... \bigoplus U_k$$
$$u = u_1 + ... + u_k$$

Där varje $u_i \in U_i$. Vi har att:

$$Tu = T_1 u_1 + ... + T_k u_k$$

Där $T_i u_i \in U_i$ och $Tu \in V$.

Vi vill kunna uttrycka direkta summan av invarianta delrum med hjälp av projektioner. För att ett delrum ska vara invariant under T, måste T kommutera med varje individuell projektion E_i. Vi ska enbart visa att om T kommuterar med E_i är U_i invariant. Vi låter beviset att om U_i är invariant kommuterar T med E_i till läsaren.

Lemma: Vi har en operation $T : V \to V$ och en projektion E där $\text{spann}(E_i) = U_i$. Om T och E_i kommuterar så att $TE_i = E_i T$, då är U_i invariant under T.

Bevis: Låt $u \in U_i$ då har vi att $E_i u = u$. Då får vi att:

$$Tu = T(E_i u)$$

Om vi förutsätter att T och E_i är kommutativ får vi att:

$$T(E_i u) = E_i(Tu)$$

$$T(u) = T(E_i u) = E_i(Tu)$$

Vi låter:

$$T(u) = v$$

Vi kan alltså skriva $E_i v = v$. Detta innebär att $T(u) = v$ ligger i spannet av E_i. Enligt förutsättning är $\text{spann}(E_i) = U_i$. Eftersom $u \in U_i$ och $Tu \in U_i$ innebär detta att U_i är invariant under T.

V.S.B.

Vi har tidigare visat om vi har delrum P_i med en tillhörande projektion E_i kan vi skriva V som en direkt summa så att $V = P_1 \bigoplus ... \bigoplus P_k$. Vi har nu visat att om projektion E kommuterar med operaton T kan vi kunna knyta en projektion E_i till varje invariant delrum U_i. Detta gör det möjligt för oss att skriva V som en direkt summa av invarianta delrum, alltså att:

$$V = U_1 \bigoplus ... \bigoplus U_k$$

Vi ska nu slutligen också visa att detta fungerar för diagonaliserbara matriser där vi arbetar i vårt välkända egenrum W som är ett invariant delrum till V under transformationen T. Vi ska visa att vi kan skriva V som en direkt summa av egenrum W_i.

Exempel: Vi har tidigare visat att:

$$v = E_1 v_1 + ... + E_k v_k$$

Om vi utför operationen T på $v \in V$ får vi:

$$Tv = TE_1 v_1 + ... + TE_k v_k$$

Detta följer direkt från definition för linjära transformationer.

Eftersom $E_i v \in W_i$ och därmed invariant får vi:

$$Tv = T_1 E_1 v_1 + ... + T_k E_k v_k$$

Vi har slagit fast att T är diagonaliserbar och därmed kommer varje restriktiv operation T_i enbart skala $E_i v_i$ med precis egenvärdet kopplat till W_i. Vi erhåller:

$$Tv = T_1 E_1 v_1 + ... + T_k E_k v_k$$

$$Tv = \lambda_1 E_1 v_1 + ... + \lambda_k E_k v_k$$

$$T = \lambda_1 E_1 + ... + \lambda_k E_k$$

Eftersom vi har k projektioner där varje projektion är kopplat till ett egenrum kan vi skriva V som en direkt summa av egenrummen till T. Vi får alltså att:

$$V = W_1 \bigoplus ... \bigoplus W_k$$

Vi har i detta kapitlet visat att vi kan använda projektioner för att kunna beskriva uppdelningen av V i olika direkta summor. Vi har egentligen inte visat något nytt då vi redan i tidigare kapitel har visat hur vi ibland kan dela upp V i invarianta delrum och i egenrum. Vi har dock behövt visa projektioner för att kunna göra än mer generella uppdelningar av V.

11 Primäruppdelningssatsen

Vi har tidigare visat att om en linjär operation T är diagonaliserbar, och därmed har minimalpolynomet på formen:

$$m(x) = (\lambda_1 - x)...(\lambda_k - x)$$

då kan vi skriva V som en direkt summa av egenrum så att $V = W_1 \bigoplus ... \bigoplus W_k$ och där vi för varje W_i har en tillhörande operation T_i. Vi vill nu med hjälp av Primäruppdelningssatsen visa att denna uppdelning även gäller för triangulerbara matriser. Vi skriver alltså minimalpolynomet på formen

$$m(x) = (\lambda_1 - x)^{d_1}...(\lambda_k - x)^{d_k}$$

Vi vill då visa att vi kan dela upp V som en direkt summa av $W_i' = \ker(\lambda_i - T)^{d_i}$, så att $V = W_1' \bigoplus ... \bigoplus W_k'$.

Detta betyder att i de fallen då vi inte har samma geometriska multiplicitet som algebraiska multiplicitet till ett bestämt egenvärde, kan vi hitta generaliserade egenvektorer som utökar egenrummet till ett generaliserat egenrum. Dessa generaliserade egenrum har för alla reella egenvärden $\dim(\ker(T - \lambda_i I)^{d_i})) = d_i$. Detta innebär att vi alltid kan hitta generaliserade egenvektor motsvarande den algebraiska multipliciteten för varje reellt egenvärde till en transformation T.

Primäruppdelningssatsen gör det möjligt att dela upp V i invarianta delrum även när det minimalpolynomet till T består av multiplar av linjära polynom över kroppen $\mathbb{R}$. Detta betyder enligt Primäruppdelningssatsen att vi kan skriva alla operationer som är linjära med multiplar och sinsemellan koprima som en direkt summa av invarianta delrum.

Sats: Om vi har en operation $T : V \to V$ där det minimala polynomet för T kan skrivas på formen $m(x) = (\lambda_1 - x)^{d_1}...(\lambda_k - x)^{d_k}$ där $\lambda_i \neq \lambda_j$, då kan vi skriva V som en direkt summa av invarianta delrum så att:

$$V = W_1' \bigoplus ... \bigoplus W_k',$$

och där varje W_i' har en restriktiv operator T_i med minimalpolynomet $(\lambda_i - x)^{d_i}$.

Bevis: Istället för att skriva ut de linjära termerna kommer vi istället skriva det linjära termerna i minimalpolynomet som $(\lambda_i - x)^{d_i} = m_i^{d_i}$. Vi skriver alltså minimalpolynomet $m(x)$ på formen:

$$m = m_1^{d_1}...m_k^{d_k}$$

Eftersom varje $m_i(x)$ är linjära irreducibla polynom kommer varje multipel av $m_i(x)$ vara relativt prima jämte resterande polynom i minimalpolynomet, alltså att $Sgd(m_i^{d_i} | m_j^{d_j}) = 1$.

Vi sätter $W_i' = \ker(\lambda_i - T)^{d_i} = \ker(m_i(T)^{d_i})$. Vi ska visa att varje $m_i^{d_i}$ är minimalpolynom till T_i som är den restriktiva operationen av T på W_i'. Vi vill hitta ett polynom h_i som agerar som en projektion E_i på W_i'. Detta hjälper oss eftersom vi tidigare har visat att om vi har k projektioner som uppfyller kraven att:

$$E_i^2 = E_i$$
$$E_iE_j = 0$$
$$I = E_1 + ... + E_k$$
$$\text{spann}(E_i) = P_i$$

Då kan vi skriva

$$V = P_1 \bigoplus ... \bigoplus P_k$$

Om vi kan hitta en projektion för varje W_i', $1 \leq i \leq k$, så att $\text{spann}(E_i) = W_i'$, då kan vi skriva $V = W_1' \bigoplus ... \bigoplus W_k'$.

Vi börjar med att skriva polynomet f_i som minimalpolynomet för T dividerat med minimalpolynomet för T_i. Vi har att:

$$f_i = \frac{m(x)}{m_i^{d_i}} = m_1^{d_1} \cdots m_{i-1}^{d_{i-1}} \cdot m_{i+1}^{d_{i+1}} \cdots m_k^{d_k}$$

$$f_j = \frac{m(x)}{m_j^{d_j}} = m_1^{d_1} \cdots m_i^{d_i} \cdots m_{j-1}^{d_{j-1}} \cdot m_{j+1}^{d_{j+1}} \cdots m_k^{d_k}$$

Eftersom f_i och f_j är relativt prima kan vi med hjälp av ett ytterligare polynom skriva:

$$f_1g_1 + ... + f_kg_k = 1$$

Låt:

$$h_i = f_ig_i$$

Detta ger oss:

$$h_1 + ... + h_k = 1$$

Vi multiplicerar h_i med h_j så att:

$$h_ih_j = f_ig_if_jg_j$$

Eftersom polynom kommuterar har vi:

$$h_ih_j = f_if_jg_ig_j$$

Eftersom f_if_j innehåller alla termer som minimalpolynomet för T innehåller är f_if_j ett annihilerande polynom till T och därmed är h_ih_j också ett annihilerande polynom för T. Vi har alltså att $h_i(T)h_j(T) = 0$.

Om vi multiplicerar $h_1(T) + ... + h_k(T) = I$ med $h_i(T)$ i båda leden får vi:

$$h_i^2(T) = h_i(T)$$

Detta eftersom $h_i(T)h_j(T) = 0$ och därmed försvinner alla dessa termer. Om vi sammanfattar egenskaperna för h_i har vi att:

$$h_1(T) + ... + h_k(T) = I$$
$$h_i(T)h_j(T) = 0$$
$$h_i^2(T) = h_i(T)$$

Detta påminner väldigt mycket om egenskaper för projektioner. Vi låter $h_i(T) = E_i$ och har enbart kvar att visa att spann$(E_i) = W_i' = \ker(m_i(T)^{d_i})$. Vi gör detta genom att anta att spann$(E_i) = W_i'$ och visar då att $m_i(T)^{d_i} = 0$ på det delrummet.

Om $v \in W_i'$ har vi $v = E_i v$.

$$\begin{aligned} m_i(T)^{d_i} v &= m_i(T)^{d_i} E_i v \\ &= m_i(T)^{d_i} f_i(T) g_i(T) v \\ &= m(T) g_i(T) v \\ &= 0 \end{aligned}$$

Vi har alltså visat att om spann$(E_i) = W_i'$, annihilerar $m_i(x)^{d_i}$ operationen T. Eftersom $m_i(T)^{d_i} = 0$ när T agerar på $v \in W_i'$ betyder det att det finns en restriktiv operator T_i på W_i' med minimalpolynom $m_i(x)^{d_i}$. Eftersom det finns en restriktiv operator T_i med minimalpolynom $m_i(x)^{d_i}$ för varje delrum W_i' vet vi att dessa delrum är invarianta. Detta säger oss också att minimalpolynomet för T_i delar minimalpolynomet för T.

Vi kan alltså skriva V som en direkt summa av invarianta generaliserade egenrum så att:

$$V = W_1' \bigoplus, ..., \bigoplus W_k'$$

V.S.B.

Primäruppdelningssatsen säger att när vi kan dela upp minimalpolynomet för operationen T i multiplar av linjära faktorer kan vi hitta ett invariant delrum med precis det polynomen som minimalpolynom. Eftersom operationen T har en restriktiv operation T_i på W_i' för varje i, kommer vi kunna skriva T som en direkt summa av restriktiva operationer. Detta är samma sak som att säga att vi kan dela upp V i en direkt summa av invarianta delrum W_i' med minimalpolynom $m_i(T_i)^{d_i}$. Över den komplexa kroppen kan minimalpolynomet delas upp så att alla irreducibla faktorer är linjära. Detta är en konsekvens av att den komplexa kroppen är algebraiskt sluten. Detta gör det möjligt att skriva alla komplexa matriser på det som kallas Jordans normalform.

Vi kommer dock fortsätta att arbeta över kroppen av de reella talen och med triangulerbara matriser. Vi kommer också att fortsätta arbeta med det generaliserade egenrummen och dess matrisrepresentation. Vi ska visa att alla triangulerbara matriser där minimalpolynomet faktoriseras ner till linjära faktorer inklusive multiplar, kan skrivas på Jordans normalform.

12 Nilpotenta operationer

En nilpotent operation är en operation N där $N^k = 0$ för $k \geqslant 0$.

Vi kommer här av formalistiska skäl utnyttja att det karakteristiska polynomet kan skrivas som $f(x) = (\lambda_1 - x)^{r_1}...(\lambda_1 - x)^{r_k} = (-1)^k(x - \lambda_1)^{r_1}...(x - \lambda_1)^{r_k}$ där $k = r_1 + ... + r_k$. Då är minimalpolynomet på formen $m(x) = (x - \lambda_1)^{d_i}...(x - \lambda_k)^{d_k}$. Vi kan i minimalpolynomet bortse från koefficienten framför polynomet eftersom minimalpolynomet annihilerar T oavsett denna koefficient. När vi skriver minimalpolynomet på denna form kan vi dela upp T i två delar, en diagonaliserbar operation och en nilpotent operation, så att $T = D + N$. På en algebraisk sluten kropp, så som den komplexa kroppen, kan alla linjära operationer skrivas $T = D + N$.

Vi utökar det linjära höljet för vår projektion så att $\text{spann}(E_i) = \ker(T - \lambda I)^{d_i}$. Eftersom $Tv = \lambda v$ är definition för en egenvektor v att vi kan skriva $(T - \lambda)v = 0$ eller $(\lambda - T)v = 0$ och beskriva samma sak. Kerneln för $(T - \lambda_i I)^{d_i}$ består alltså av samma element som $\ker(\lambda_i - T)^{d_i}$. Vi benämner $\ker(T - \lambda_i)^{d_i}$ som W' och som här betecknar det generaliserade egenrummet. Med hjälp av Primäruppdelningssatsen har vi sett att vi kan dela upp V som en direkt summa av generaliserade egenrum om minimalpolynomet är på ovanstående form. Detta betyder att det generaliserade egenrummet är invariant under operationen T.

Vi vet att:

$$I = E_1 + ... + E_k$$

Om vi utför operationen T i båda led får vi enligt definitionen för linjära transformationer:

$$T = T_1E_1 + ... + T_kE_k$$

Eftersom $\text{spann}(E_i) = W_i'$ och W_i' är invariant kan T, när det opererar på det generaliserade egenrummet, beskrivas som T_i.

$$D = \lambda_1 E_1 + ... + \lambda_k E_k$$

$$N = T - D$$

$$N = (T_1 - \lambda_1 I)E_1 + ... + (T_k - \lambda_k I)E_k$$

$$N^k = (T_1 - \lambda_1 I)^k E_1 + ... + (T_k - \lambda_k I)^k E_k$$

Eftersom projektioner är idempotent är det bara $(T_i - \lambda_i I)$ som upphöjs med k. Då är $N^k = 0$ om k är större eller lika med alla d_i så att $k \geq d_i$, för alla i.

Minimalpolynomet för T_i är $m_i(T_i) = (T_i - \lambda_i)^{d_i}$, vilket innebär att d_i är det minsta heltal som annihilerar operationen $(T_i - \lambda_i)$. Vi kallar $(T_i - \lambda_i)$ för den nilpotenta operation restriktiv till det utökade egenrummet $\ker(T_i - \lambda_i)^{d_i}$ så att $N_i = (T_i - \lambda_i)$ där $N_i^{d_i} = 0$.

Eftersom $\text{spann}(E_i) = \ker(T - \lambda_i)^{d_i}$ innebär detta att $(T - \lambda_i)^{d_i} E_i$ kan ses som att operationen $(T - \lambda_i)^{d_i}$ opererar på sitt nollrum vilket oundvikligt kommer att bli 0.

Om T är diagonaliserbar kommer minimalpolynomet för varje T_i vara $m_i(T_i) = (T_i - \lambda_i I) = N_i = 0$. Detta innebär också att den nilpotenta matrisen är nollmatrisen och att $T = D + N = D + 0 = D$, vilket är uppenbart eftersom T är diagonaliserbar.

Detta ger oss att:

$$m(N)^d = m_1^{d_1}(T_1)E_1 + ... + m_k^{d_k}(T_k)E_k.$$

där $d = max\{d_1, .., d_k\}$

En viktig slutsats är att en operation är nilpotent på sitt egenrum när vi subtraherar med det egenvärde som egenrummet är kopplat till. När en operation enbart har ett egenvärde kommer operationen subtraherat med det egenvärdet vara nilpotent över hela V. En viktig sats är att alla nilpotenta matriser kan skrivas som en blockmatris av skiftnilpotenta matriser. Detta innebär att alla nilpotenta matriser kan skrivas som en blockmatris där varje block består av en nilpotent matris som enbart har ettor i huvuddiagonalen och nollor överallt annars. Vi skriver en skiftnilpotent matris på formen:

$$N = \begin{bmatrix} S_1 & 0 & 0 & \dots & 0 \\ 0 & S_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & S_{n-1} & 0 \\ 0 & 0 & 0 & 0 & S_n \end{bmatrix}$$

där varje S_i är en skiftmatris på formen

$$S_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Vi antar att vi har en $n \times n$ matris A bestående av kolonnvektorerna $\{v_1, ..., v_n\}$. Vi låter S beteckna en skiftnilpotentmatris av samma storlek som A. Om vi låter S operera på A genom matrismultiplikation från vänster kommer varje vektor förskjutas ett steg. Detta innebär att $v_1 \to 0, v_2 \to v_1, ..., v_n \to v_{n-1}$. Vi kommer ha användning av detta då vi kommer kunna bygga kedjor av generaliserade egenvektorer med hjälp av operationer med nilpotenta matriser. Mer om detta i nästa kapitel.

13 Generaliserade egenvektorer

Om vi har en matris A som är defekt och därmed inte har tillräckligt många egenvektorer för att fylla ut hela V, kan vi använda oss av generaliserade egenvektorer. Vi förutsätter att det karakteristiska polynomet för A är en produkt av linjära faktorer och är på formen:

$$f(A) = (-1)^k (A - \lambda_1 I)^{r_1} ... (A - \lambda_k I)^{r_k}$$

och att minimalpolynomet är på formen:

$$m(A) = (A - \lambda_1 I)^{d_1} ... (A - \lambda_k I)^{d_k}$$

Som vi tidigare har nämnt kallas då $\ker(A - \lambda_i I)^{d_i}$ för det generaliserade egenrummet av A för egenvärdet λ_i och $v \in \ker(A - \lambda_i I)^{d_i}$ kallas för en generaliserad egenvektor till A. Precis som tidigare benämner vi $\ker(A - \lambda_i I)^{d_i}$ som W_i'. Om $d_i = 1$ har vi ett reguljärt egenrum bestående av egenvektorer.

Definition: En generaliserad egenvektor v har egenskapen att $((A - \lambda_i I))^k v = 0$ men att $((A - \lambda_i I))^{k-1} v \neq 0$. Dessa generaliserade egenvektorerna är sinsemellan linjärt oberoende.

Enligt Primäruppdelningssatsen kan vi skriva V som en direkt summa av generaliserade egenrum:

$$V = W_1' \bigoplus ... \bigoplus W_k'$$

Denna uppdelning är sann över det reella vektorrummet om A är kvadratisk och där A enbart har reella egenvärden och där den algebraiska multipliciteten för varje egenvärde adderat blir dimensionen för V.

Om vi utgår från det ovanstående karakteristiska polynom för A och säger att egenvärdet λ_i har algebraisk multiplicitet r_i, då är det ett faktum att $\dim(W_i') = r_i$. Detta innebär att det finns lika många generaliserade egenvektorer (inklusive egenvektorer) som den algebraiska multipliciteten för egenvärdet. För varje r_i i det karakteristiska polynomet kan vi hitta r_i generaliserade egenvektorer vilket fyller ut hela det generaliserade egenrummet till egenvärdet λ_i och där alla exponenter till de linjära termerna i det karakteristiska polynomet tillsammans får dimensionen V. Alltså $r_1 + ... + r_k = \dim(V)$. Det kan verka kontraintuitivt att $\dim(W_i') = \dim(\ker((A - \lambda_i I))^{d_i})) = r_i$ men det bygger på att $(A - \lambda_i I)$ är en nilpotent matris på W_i' och att minimalpolynomet för N_i är $m(x) = x^{d_i}$. Därmed är $m(N_i) = N_i^{d_i} = 0$. Detta betyder att nollrummet för $N_i^{d_i}$ och $N_i^{r_i}$ är lika.

För ett bestämt egenvärde har vi att:

$$\ker((A - \lambda_i I)) \subseteq \ker((A - \lambda_i I))^2 \subseteq ... \subseteq \ker((A - \lambda_i I))^{d_i}$$

Vi fortsätter att arbeta i våra individuella invarianta delrum W_i'.

Precis som tidigare ser vi att $N_i = (A_i - \lambda_i I)$ är en nilpotent matris. Detta innebär att $N_i^{d_i} = (A_i - \lambda_i I)^{d_i} = 0$. Vi skriver att minimalpolynomet för N_i är på formen $m_i(N_i) = x^{d_i} = 0$.

Vi har att:

$$\ker(N_i) \subseteq \ker(N_i^2) \subseteq ... \subseteq \ker(N_i^{d_i})$$

Om vi säger att v är en egenvektor till A kopplat till egenvärdet λ_i har vi att $N_i v = 0$. Detta innebär att $v \in \ker(N_i)$. Det är lätt att se att $v \in \ker(N_i^2)$ eftersom $N_i^2 = N_i(N_i v) = 0$. Därför innehåller kerneln av varje potens av N_i den föregående kerneln av potensen med lägre exponent.

Vi är intresserade av nilpotenta matriser eftersom dessa har egenskapen att för varje $N^t, 1 \leq t < d$ och om vi kan hitta u så att $N^t u \neq 0$, är $u, Nu, ..., N^t, ..., N^{d-1}u$ linjärt oberoende. En sådan kedja av linjärt oberoende vektorer kallas för en Jordankedja. Varje Jordankedja är det som konstituerar ett Jordanblock. Den direkta summan av Jordanblock för ett bestämt egenvärde bygger upp det generaliserade egenrummet W'.

Exempel:

$$N = \begin{bmatrix} 2 & 5 & 1 \\ 2 & 1 & 5 \\ -2 & -3 & -3 \end{bmatrix}$$

N är nilpotent eftersom

$$N^3 = \begin{bmatrix} 2 & 5 & 1 \\ 2 & 1 & 5 \\ -2 & -3 & -3 \end{bmatrix}^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Vi har att $N^2 \neq 0$ vilket innebär att $m(N) = N^3$. Vi kan nu bygga en kedja av oberoende vektorer som utgör en bas för V. Vi vill hitta en vektor som har egenskapen att v, Nv, N^2v är linjärt oberoende. Detta kan vi göra genom att först se att $\dim(\ker(N)) = 1$. Detta innebär att N har precis en egenvektor. Vi har även att dim (ker N^3)$-$dim(kerN^2) =dim (kerN^2)$-$dim(ker$N$) $= 1$. Detta betyder att det finns en ytterligare generaliserad egenvektor för varje $N^{t+1}, 0 \leq t \leq 2$. Detta visar att vi har en Jordankedja. En nilpotent matris har enbart 0 som egenvärde så för att hitta egenvektorn för N tar vi $(N-0)v_1 = Nv_1 = 0$. Vi skriver $Nv_1 = 0$

$$\begin{bmatrix} 2 & 5 & 1 \\ 2 & 1 & 5 \\ -2 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Vilket ger oss ekvationssystemet:

$$\begin{bmatrix} 2x_1 + 5x_2 + x_3 = 0 \\ 2x_1 + x_2 + 5x_3 = 0 \\ -2x_1 - 3x_3 - 3x_3 = 0 \end{bmatrix}$$

Vi löser ekvationssystem och får:

$$v_1 = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$$

Nu ska vi hitta en vektor v_2 som uppfyller kravet att $Nv_2 = v_1$. Detta beroende på att om vi vänstermultiplicerar med N i båda led får vi, $N^2v_2 = Nv_1 = 0$, och därmed har vi en vektor v_2

som ligger i kerneln av N^2. Vi skriver $Nv_2 = v_1$:

$$\begin{bmatrix} 2 & 5 & 1 \\ 2 & 1 & 5 \\ -2 & -3 & -3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$$

Vi löser ekvationssystemet och får:

$$v_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Vi fortsätter likadant för att hitta $Nv_3 = v_2$:

$$\begin{bmatrix} 2 & 5 & 1 \\ 2 & 1 & 5 \\ -2 & -3 & -3 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Vilket ger oss att:

$$v_3 = \begin{bmatrix} \frac{-3}{4} \\ \frac{1}{2} \\ 0 \end{bmatrix}$$

Vi har nu tre oberoende vektorer och där v_3 är generatorn av Jordankedjan. Jordankedjan består alltså av $v_3, Nv_3 = v_2, N^2v_3 = v_1$. Denna kedja bildar en bas för V och består av generaliserade egenvektorer till N.

$$P = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = \begin{pmatrix} -3 & 1 & \frac{-3}{4} \\ 1 & -1 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix}$$

Om vi utför ett basbyte för att se hur vår nilpotenta matris uppför sig i den nya basen som spänner upp det generaliserade egenrummet till N. Vi skriver

$$P^{-1}NP = J$$

Där P är basbytesmatrisen bestående av generaliserade egenvektorer till N och J och är den matris vi erhåller i denna bas. Vi skriver:

$$\begin{aligned} P^{-1}NP &= \begin{bmatrix} 0 & 0 & 1 \\ -2 & -3 & -3 \\ -4 & -4 & -8 \end{bmatrix} \begin{bmatrix} 2 & 5 & 1 \\ 2 & 1 & 5 \\ -2 & -3 & -3 \end{bmatrix} \begin{pmatrix} -3 & 1 & \frac{-3}{4} \\ 1 & -1 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Vi ser att N är similär med en skiftmatris i basen P. Vi har att alla nilpotenta matriser kan skrivas som en skiftmatris på sitt generaliserade egenrum. Vi kommmer i nästa kapitel att se att anledningen till att alla nilpotenta matriser kan representeras som en skiftmatris är på grund av att skiftmatrisen är på Jordans normalform. Skiftmatrisen har enbart nollor i diagonalen vilket korresponderar till att en nilpotent matris enbart har noll som egenvärde. En matris på Jordans normalform har egenvärdena till den ursprungliga matrisen i diagonalen och för nilpotenta matriser är dessa enbart noll. Vi har ettor i superdiagonalen vilket påvisar existensen av generaliserade egenvektorer kopplat till detta egenvärde. Resten av elementen i en matris på Jordanform är nollor. När vi opererar med skiftmatriser är det lätt att se att denna indikerar en kedja av generaliserade egenvektorer. Vi har tidigare visat hur $Sv_1 \to 0, Sv_2 \to v_1, Sv_3 \to v_2$ osv.

$$SP = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & 0 \\ v_2 & v_3 & 0 \\ | & | & 0 \end{bmatrix}$$

$$S^2P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & 0 & 0 \\ v_3 & 0 & 0 \\ | & 0 & 0 \end{bmatrix}$$

$$S^3P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Vi har för SP går v_1 till noll eftersom detta är en egenvektor till N. Vi har att $Sv_2 = v_1$ och att $S^2v_2 = 0$ vilket indikerar vår kedja. S^3 är nollmatrisen och därmed kommer alla vektorer multiplicerat med denna bli 0-vektorn.

Vi har påvisat att alla nilpotenta matriser kan skrivas på Jordans normalform och därmed är likartad till en skiftmatris i basen bestående av sina generaliserade egenvektorer. Vi kommer att visa att alla kvadratiska matriser med enbart reella egenvärden kan skrivas på Jordans normalform.

14 Jordans normalform

Om vi för T har det karakteristiska polynomet

$$f(x) = (-1)^k (x - \lambda_1)^{r_1} ... (x - \lambda_k)^{r_k}$$

Med minimalpolynomet för triangulerbara operationer som är på formen:

$$m(x) = (x - \lambda_1)^{d_1} ... (x - \lambda_k)^{d_k}$$

där $d_i \leq r_i$.

Om vi har lika många reella egenvärden λ_i som dimensionen på V och alla faktorer i det karakteristiska polynomet är linjära(vilket implicerar att minimalpolynomet också enbart har linjära faktorer), kan vi skriva T på Jordans normalform. I det komplexa fallet är alla linjära operationer $V \to V$ möjliga att översätta till Jordans normalform, eftersom polynom över kroppen av komplexa tal är algebraisk sluten vilket innebär att alla operationer har komplexa egenvärden. Vi fortsätter att arbeta i det reella vektorrummet V. För alla matriser över kroppen $\mathbb{R}$ med egenskapen listade ovan kan vi hitta en bas så att:

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

B är en bas bestående av generaliserade egenvektorer och J är en matris på Jordans normalform. Vi skriver att B_i är uppsättningen av generaliserade egenvektorer som tillhör W_i'. Vi har tidigare visat att varje linjär faktor $(x - \lambda_1)^{d_i}$ i minimalpolynomet, har ett kopplat generaliserat egenrum $W_i' = \ker(x - \lambda_1)^{d_i}$. Dimensionen av det generaliserade egenrummet W_i' är r_i, den algebraiska multipliciteten för λ_i. Vi har alltså alltid lika många generaliserade egenvektorer som algebraisk multiplicitet vilket hjälper oss i det fallet när den algebraiska multipliciteten inte är samma som den geometriska multipliciteten. Eftersom B i vårt fall är en bas bestående av linjärt oberoende generaliserade egenvektorer kan vi skriva $B = B_1 \bigoplus ... \bigoplus B_k$.

Jordanmatrisen är en blockdiagonalmatris där varje element i diagonalen består av det vi kallar för Jordanblock. Blocken i Jordanmatrisens diagonalelement noterar vi här som D, där varje D_i består av Jordanblocken kopplade till egenvärdet λ_i. Jordanblocken noterar vi som J_i och där varje J_i är en matris av samma dimension som längden av den Jordankedja som den representerar. Jordankedja, som vi tidigare har nämnt, är den kedja av vektorer som genereras av en generaliserad egenvektor och när vi opererar med en nilpotent operator på denna vektor får vi ytterligare en generaliserad egenvektor. I slutet av varje Jordankedja finns det alltid en egenvektor. Antalet Jordankedjor kopplat till ett egenvärde för A är alltid lika med den geometriska multipliciteten precis på grund av att varje kedja avslutas med en egenvektor. Därmed är antalet Jordanblock kopplat till ett egenvärde också lika med antalet egenvektorer. W_i' har precis dimensionen som antalet λ_i i diagonalen av Jordanmatrisen.

Om vi låter D_i vara diagonalelementet i Jordanmatrisen bestående av Jordanblocken till egenvärde λ_i har vi att:

$$J = \begin{bmatrix} D_1 & 0 & 0 & \dots & 0 \\ 0 & D_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & D_{k-1} & 0 \\ 0 & 0 & 0 & 0 & D_k \end{bmatrix}$$

Vi låter J_i beskriva våra Jordanblock som består av Jordankedjor. Varje J_i har samma dimension som Jordankedjan är lång:

$$D_i = \begin{bmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & J_{k-1} & 0 \\ 0 & 0 & 0 & 0 & J_k \end{bmatrix}$$

Om J_i är ett Jordanblock som konstitueras av en Jordankedja som strikt längre än 1 har vi:

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \lambda_i \end{bmatrix}$$

Som vi kan se består varje J_i av en diagonalmatris $\lambda_i I$ adderat med en skiftnilpotentmatris S_i.

$$J_i = \begin{bmatrix} \lambda_i & 0 & 0 & \dots & 0 \\ 0 & \lambda_i & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & & \lambda_i & 0 \\ 0 & 0 & 0 & 0 & \lambda_i \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \lambda_i \end{bmatrix}$$

Att vi kan skriva varje nilpotentmatris på skiftnilpotentform beror på att varje nilpotent matris enbart har egenvärdet 0. Det karakteristiska polynomet för en nilpotent matris är på formen $f(x) = x^r$, där r är den algebraiska multipliciteten för 0. Detta gör att vi har r generaliserade egenvektorer kopplat till egenvärdet 0, och då kan vi skriva varje nilpotent matris på Jordans normalform med nollor i diagonalen och ettor i superdiagonalen.

Om Jordankedjan är av längd 1 har vi istället ett Jordanblock av storlek 1×1:

$$J_i = \begin{bmatrix} \lambda_i \end{bmatrix}$$

Exempel på Jordans normalform:

$$J = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{bmatrix}$$

$$J_1 = \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix}, J_2 = \begin{bmatrix} \lambda_2 \end{bmatrix}, J_3 = \begin{bmatrix} \lambda_3 & 1 & 0 \\ 0 & \lambda_3 & 1 \\ 0 & 0 & \lambda_3 \end{bmatrix}, J_4 = \begin{bmatrix} \lambda_3 \end{bmatrix}$$

Varje Jordanblock J_i konstitueras av en Jordankedja som har en egenvektor i änden av kedjan. Som vi tidigare har nämnt betyder detta att t.ex. λ_3 har geometrisk multiplicitet 2 vilket gör att det finns två Jordanblock kopplat till det egenvärdet. λ_1, precis som λ_2 har geometrisk multiplicitet 1 och därmed enbart ett Jordanblock kopplat till de respektive egenvärdena. J_1 har även en generaliserad egenvektor, vilket gör att Jordanblocket, har dimension 2. Det generaliserade egenrummet W' består av alla Jordanblock kopplat till ett bestämt egenvärden. Vi har alltså att det generaliserade egenrummet W_3' består av Jordanblocken J_3 och J_4.

Om vi har en matris A med det enda egenvärdet λ och där λ har geometrisk multiplicitet r och detta är samma som den algebraisk multiplicitet, har vi tidigare sett att A är diagonaliserbar. Detta är samma som att säga att vi kan skriva A på Jordans normalform där λ uppstår r gånger i diagonalen och som består av r antal Jordanblock, varje av längd 1.

15 Koppling mellan karakteristiska polynomet och minimalpolynomet

Om vi för A har det karakteristiska polynomet

$$f(x) = (-1)^k (x - \lambda_1)^{r_1} ... (x - \lambda_1)^{r_k}$$

Med minimalpolynomet för triangulerbara operationer som är på formen:

$$m(x) = (x - \lambda_1)^{d_1} ... (x - \lambda_1)^{d_k}$$

där $d \leq r$.

Bara genom att se på det karakteristiska och minimalpolynomet kan vi få viktig information om hur matrisen A kommer att ha för Jordanmatris. Utifrån det karakteristiska polynomet får vi reda på dimensionen av det generaliserade egenrummet W_i' kopplat till egenvärdet λ_i vilket är r_i. Om $r_i = d_i$ vet vi att W_i' enbart byggs upp av ett Jordanblock. Vi kan även få veta dimensionen av det största Jordanblocket utifrån minimalpolynomet.

Om vi arbetar i det separata generaliserade egenrummet W_i' med matrisrepresentation B_i och med minimalpolynomet $m_i(x) = (x - \lambda_i)^{d_i}$. Minimalpolynomet för ett bestämt egenvärde annihilerar hela W_i'. W_i' är en direkt summa av Jordanblock, $\{J_1, ..., J_k\}$. Vi har tidigare visat att om vi kan beskriva ett vektorrummet V som en direkt summa av invarianta delrum kommer minimalpolynomet till T(som verkar på V) vara den minsta gemensamma multipel av minimalpolynomen för varje restriktiv transformation T_i(som verkar på det individuella invarianta delrummet). Detta innebär att varje minimalpolynom till T_i (som verkar på sitt generaliserade egenrum W_i'), för varje generaliserade egenrum till V, delar minimalpolynomet till T. Vår matrisrepresentation för T_i på det generaliserade egenrummen W_i' är D_i. Vi har alltså att $m(D_i) = MGM(m_1(J_1), ..., m_k(J_k))$ och att varje $m_i(J_i)$ är på formen $m_i(x) = (x - \lambda_1)^{t_i}$ där $t_i \leq d_i$. Om vi har ett polynom som annihilerar det största Jordanblocket kommer detta polynom annihilera resterande Jordanblock över W_i'. Minimalpolynomet för D_i är precis minimalpolynomet för det största Jordanblocket. Detta betyder att det största Jordanblocket i D_i är precis av dimension $d_i \times d_i$. Detta går även att se genom $(J_i - \lambda_i I)$ är en skiftnilpotent matris med minimalpolynomet $m(x) = x^{d_i}$.

$(J_i - \lambda_i I)$ är en skiftnilpotent matris på formen:

$$(J_i - \lambda_i I) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, (J_i - \lambda_i I)^{d_i} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Exempel 1:

Om vi har en matris A med ett enda egenvärde och med karakteristiska polynomet $f(A) = (-1)^3(A-\lambda)^3$. Vi låter minimalpolynomet vara $m(A) = (A-\lambda)^2$. Utifrån detta exempel kommer det första Jordanblocket vara en 2×2-Jordanmatris. Detta vet vi eftersom multipliciteten för λ som en rot till minimalpolynomet kommer konstituera storleken på det första Jordanblocket. Eftersom dimensionen på det generaliserade egenrummet måste vara lika med den algebraiska multipliciteten vet vi att det andra Jordanblocket kommer vara endimensionella och alltså vara en 1×1 matris med λ som enda element. Vi har nu två Jordanblock vilket också säger oss att vi har två egenvektorer.

Exempel 2:

Om vi istället har $f(A) = (A-\lambda)^4$ och $m(A) = (A-\lambda)^2$, kan vi enbart säga att det första Jordanblocket är av storlek 2×2. Vi vet inte om det är ett till 2×2 Jordanblock eller två 1×1 block. Detta beror på den geometriska multipliciteten för λ. Om λ enbart har två reguljära egenvektorer kommer Jordans normalformen vara två 2×2 Jordanblock. Är den geometriska multipliciteten 3 kommer det vara ett 2×2 Jordanblock och två 1×1 Jordanblock.

Exempel 3:

$$A = \begin{bmatrix} 1 & 8 & 4 & 2 \\ 0 & 2 & -2 & -2 \\ -2 & -2 & 6 & 0 \\ 2 & 2 & -2 & 4 \end{bmatrix}$$

För att hitta matrisen A på Jordans normalform måste vi först hitta alla egenvärden. Detta gör vi genom att ta $\det(A-\lambda) = 0$:

$$\det(A-\lambda) = \begin{bmatrix} 1-\lambda & 8 & 4 & 2 \\ 0 & 2-\lambda & -2 & -2 \\ -2 & -2 & 6-\lambda & 0 \\ 2 & 2 & -2 & 4-\lambda \end{bmatrix} = 0$$

När vi löser ut detta fjärdegradspolynom får vi :

$$f(x) = (x-2)^2(x-4)(x-5)$$

Vi har att $m(x) = f(x)$. Enbart utifrån minimalpolynomet kan vi se att Jordanblocket som korresponderar till egenvärdet 2 är av dimension 2. Jordanblocken som korresponderar till egenvärden 4 och 5 är av dimension 1. Vi ser direkt att Jordanformen för A är:

$$J = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

För att hitta basen B där A representeraras av en Jordanmatris behöver vi hitta alla generaliserade egenvektorer. Eftersom egenvärdena 4 och 5 enbart har reguljära egenrum kopplade till sig vet vi att det enbart är egenvärdet 2 som har ett generaliserat egenrum. Vi börjar med att hitta egenvektorn kopplat till egenvärdet 2:

$$(A-2)=\left[\begin{array}{cccc|c} -1 & 8 & 4 & 2 & 0 \\ 0 & 0 & -2 & -2 & 0 \\ -2 & -2 & 4 & 0 & 0 \\ 2 & 2 & -2 & 2 & 0 \end{array}\right]$$

När vi löser detta ekvationssystem får vi att egenvektorn v_1 kopplat till egenvärdet 2 är :

$$v_1=\begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

För att hitta den generaliserade egenvektorn skriver vi $(A-2)v_2 = v_1$ Detta ger oss uttrycket:

$$\left[\begin{array}{cccc|c} -1 & 8 & 4 & 2 & -2 \\ 0 & 0 & -2 & -2 & 0 \\ -2 & -2 & 4 & 0 & -1 \\ 2 & 2 & -2 & 2 & 1 \end{array}\right]$$

där en lösning är:

$$v_2=\begin{bmatrix} \frac{2}{3} \\ -\frac{1}{6} \\ 0 \\ 0 \end{bmatrix}$$

Eftersom vi vet att det bara finns en egenvektor kopplat till egenvärdet 2 och att $v_1 \neq v_2$, där v_1 är en egenvektor, vet vi att v_2 är en generaliserad egenvektor.

Räknar vi ut egenvektorerna för egenvärdena 4 och 5 får vi basen:

$$B=\begin{bmatrix} -2 & \frac{2}{3} & -8 & -1 \\ 0 & -\frac{1}{6} & -1 & 0 \\ -1 & 0 & -9 & -1 \\ 1 & 0 & 10 & 1 \end{bmatrix}$$

När vi har Jordanbasen B bestående av generaliserade egenvektorer till A kan vi skriva A som en Jordanmatris i basen B:

$$A = BJB^{-1} = \begin{bmatrix} -2 & \frac{2}{3} & -8 & -1 \\ 0 & -\frac{1}{6} & -1 & 0 \\ -1 & 0 & -9 & -1 \\ 1 & 0 & 10 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{2}{3} & -8 & -\frac{14}{3} & -5 \\ 0 & -6 & -1 & -6 \\ 0 & 0 & -9 & 1 \\ \frac{1}{3} & \frac{4}{3} & 10 & -2 \end{bmatrix}$$

Att skriva matrisen A på Jordans normalform är samma sak som att beskriva V som en direkt summa av generaliserade egenrum

$$V = W_1' \bigoplus W_2' \bigoplus W_3'$$

där $W_1' = \ker(A-2)$, $W_2' = \ker(A-4)$ och $W_3' = \ker(A-5)$. Detta stämmer överens med att B är en bas för V.

16 Avslutning

Vi har genom arbetets gång undersökt hur vi på enklast möjliga sätt kan beskriva en transformation utifrån invarianta delrum och rätt val av bas. Vi har beskrivit hur direkta summor av egenrum och generaliserade egenrum har en särskilt enkel matrisrepresentation. Vi har sett att med hjälp av minimalpolynomet kan vi få reda på mycket information om sammansättningen av transformationen. Det arbetet saknar är en mer genomgående bearbetning av transformationernas representation som linjära avbildningar. Linjära avbildningar kan ge ett gott stöd i förståelsen av en transformation, genom den geometriska analogin som erbjuds när man tänker i termer av linjära avbildningar. Geometriska tolkningen av linjära avbildningar är ett bra komplement jämte vektorer och matrisanalogier när vi gräver oss djupare ner i en transformations kärna. Det är något jag hoppas få arbeta mer med i framtiden.

17 Referenslista

Referenser

[1] Ray Kunze and Kenneth Hoffman. *Linear Algebra*. Second edition(1971). [PDF-fil] Hämtad 2/10-2019 från http://www.math.pku.edu.cn/teachers/anjp/textbook.pdf

[2] Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence. *Linear Algebra*. Pearson New International Edition(2014). [PDF-fil] Hämtad 1/10-2019 från https://ulissesgtz.files.wordpress.com/2019/02/stephen-h.-friedberg-arnold-j.\\-insel-lawrence-e.-spence-linear-algebra-pearson-2014.pdf

[3] Sheldon Axler. *Linear Algebra Done Right*(2015). [PDF-fil] Hämtad 1/12-2019 från https://zhangyk8.github.io/teaching/file_spring2018/linear_algebra_done_right.pdf

[4] Erik Wahlén.*The Jordan Normal Form*(2011). [Lektionsanteckningar] Hämtad 26/10-2019 från http://www.ctr.maths.lu.se/media11/MATM14/2011MATM14_vt11/jordan_.pdf

[5] Northwestern University. *Notes on Jordan Form*(2015). [Lektionsplanering] Hämtad 12/1-2020 från https://sites.math.northwestern.edu/~scanez/courses/334/notes/jordan-form.pdf