1. Vektorrum

Definitionsmängderna och värdemängderna för de linjära avbildningar som studeras i denna kurs utgörs av vektorrum. I detta kapitel introduceras grundbegrepp såsom vektorrum, underrum, summor och direkta summor av underrum, linjärt oberoende, linjära höljen, baser och dimension. Vi kommer främst att studera avbildningar mellan ändligtdimensionella vektorrum, varvid matrisrepresentationer av olika typ är viktiga. Först definieras begreppet vektorrum:

Definition 1.1. *Låt K vara kroppen R av de reella talen eller kroppen C av de komplexa talen. En mängd V säges utgöra ett vektorrum (linjärt rum) över K om V har följande egenskaper.*

(i) Till varje ordnat par (u, v) av element i V finns ordnat ett element $u + v$ i V (V slutet under "addition"). Sammansättningen $+$ ("addition") har egenskaperna:

1. $u + v = v + u$, *för alla* $u, v \in V$,
2. $(u + v) + w = u + (v + w)$, *för alla* $u, v, w \in V$,
3. *I mängden V finns ett element 0 (nollelement) sådant att $u + 0 = u$ för alla $u \in V$,*
4. *För varje $v \in V$ existerar $w \in V$ sådant att $v + w = 0$ (w additiv invers till v).*

(ii) Till varje tal $a \in K$ och varje $v \in V$ finns ordnat ett element $a \cdot v$ i V, ofta betecknat av, (V sluten under "multiplikation" med skalär). Denna "multiplikation" har egenskaperna:

1. $a(bv) = (ab)v$, *för varje* $v \in V$ *och alla* $a, b \in K$,
2. $1 \cdot v = v$, *för alla* $v \in V$.

(iii) För alla $u, v \in V$ och alla skalärer $a, b \in K$ gäller:

1. $a(u + v) = au + av$,
2. $(a + b)u = au + bu$.

Om den underliggande skalärkroppen $K = R$ säger vi att V är ett reellt vektorrum, om $K = C$ säger vi att V är ett komplext vektorrum. Elementen i V kallas vektorer. I fortsättningen betecknar V ett vektorrum över skalärkroppen $K (= R$ eller $C)$.

Anmärkning 1.2. Nollelementet i Definition 1.1(i)3 är entydigt bestämt, ty om $0, 0' \in V$ och båda är nollelement erhålls:

$$0' = 0' + 0 = 0 + 0' = 0,$$

där första likheten gäller då 0 är ett nollelement, den andra ges av 1.1(i)1 och den sista likheten gäller då $0'$ är ett nollelement.

Varje vektor $v \in V$ har en entydig additiv invers i $1.1(i)4$, ty om $w, w' \in V$ är additiva inverser till $v \in V$ gäller:

$$w = w + 0 = w + (v + w') = (w + v) + w' = (v + w) + w' = 0 + w' = w' + 0 = w',$$

med stöd av axiomen i Definition 1.1. På grund av entydigheten inför vi beteckningen $-v$ för inversen till $v \in V$. Med beteckningen $u - v$ avses vektorn $u + (-v)$.

Sats 1.3. *För varje $v \in V$ och varje $a \in K$ gäller:*

$$(i)\ 0v = 0,$$
$$(ii)\ a0 = 0,$$
$$(iii)\ (-1)v = -v.$$

Bevis: Hemuppgift.

Exempel 1.4. Exempel på vektorrum. (Se föreläsningsanteckningar).

Underrum

Definition 1.5. *Låt V vara ett vektorrum över K och U en icke-tom delmängd av V, $U \subseteq V$. Då är U ett underrum av V om det gäller att:*

$$(i)\ u + v \in U, \quad \textit{för alla } u, v \in U,$$
$$(ii)\ au \in U, \quad \textit{för varje } u \in U \textit{ och för varje } a \in K.$$

Genom att välja $a = 0$ i $1.5(ii)$ fås med stöd av $1.3(i)$ att nollelementet 0 i V tillhör U. För varje $u \in U$ gäller att $-u \in U$, ty $1.5(ii)$ med $a = -1$ ger att $(-1)u \in U$ och $1.3(iii)$ ger att $(-1)u = -u$. Vidare gäller $1.1(i)1$, $1.1(i)2$, $1.1(ii)$ och $1.1(iii)$ eftersom $U \subseteq V$.

Därmed har vi att varje underrum U av V är ett vektorrum över K med de operationer $(+, \cdot)$ som definierats i V.

För att utreda om en delmängd U av vektorrummet V är ett underrum bör man alltså kontrollera tre saker: 1) att $0 \in U$, 2) att U är sluten under addition och 3) att U är sluten under multiplikation med skalär ur K.

Exempel 1.6. (a) $U = \{0\}$ och $U = V$ är de triviala underrummen av V.

(b) Mängden $U = \{(x_1, x_2, 0) : x_1, x_2 \in K\}$ är ett underrum av K^3.

(c) Mängden $M = \{(x_1, x_2) : x_{1,2} \in R,\ x_1 + x_2 = 2\}$ är inte ett underrum av R^2, ty $0 = (0, 0) \notin M$.

(d) Mängden $\mathcal{P}_n(K)$ definieras som mängden av alla polynom av gradtal högst n med koefficienter i K. Då är $\mathcal{P}_n(K)$ ett underrum av $\mathcal{P}(K)$ med de operationer som definierats i Exempel $1.4(c)$.

För två delmängder U_1 och U_2 av en mängd V gäller det att $U_1 \cup U_2$ är den minsta delmängd av V som innehåller både U_1 och U_2. Om V är ett vektorrum och U_1, U_2 är underrum kan man fråga sig vilket är det minsta underrum av V som innehåller både U_1 och U_2. Svaret på frågan ges i nästa avsnitt, men vi kan konstatera att unionen av två underrum är i allmänhet inte ett underrum. Vi har nämligen:

Sats 1.7. *Antag att U_1 och U_2 är underrum av vektorrummet V. Unionen av U_1 och U_2 är ett underrum av V om och endast om $U_1 \subseteq U_2$ eller $U_2 \subseteq U_1$.*

Bevis: Se föreläsningsanteckningar.

Däremot gäller:

Sats 1.8. *Snittet av ett godtyckligt antal underrum av V är ett underrum.*

Bevis: Hemuppgift.

Summor och direkta summor av underrum

Definition 1.9. *Antag att $U_1, ..., U_m$ är underrum av vektorrummet V. Summan av $U_1, ..., U_m$ betecknas $U_1 + ... + U_m$ och definieras genom*

$$U_1 + ... + U_m = \{u_1 + ... + u_m : u_1 \in U_1, ..., u_m \in U_m\}.$$

Antag att $V = U_1 + ... + U_m$. Då kan varje element $v \in V$ uttryckas som $v = u_1 + ... + u_m$ med en eller flera kombinationer av element $u_j \in U_j$, $j = 1, ..., m$. Om denna framställning är entydig för varje $v \in V$ säger vi att V är direkta summan av $U_1, ..., U_m$. Detta betecknas med

$$V = U_1 \oplus U_2 \oplus ... \oplus U_m.$$

Anmärkning 1.10. Om $U_1, ..., U_m$ är underrum av V så är även summan $U_1 + ... + U_m$ ett underrum av V. Vidare gäller att $U_1, ..., U_m$ är delmängder av $U_1 + ... + U_m$, (välj turvis alla u_j:n utom ett till noll). Å andra sidan, om U är ett underrum av V som innehåller $U_1, ..., U_m$ så innehåller U även $U_1 + ... + U_m$, ty alla ändliga summor av element i U, speciellt $u_1 + ... + u_m$ då $u_j \in U_j$, ingår i U. **Därmed gäller det att $U_1 + ... + U_m$ är det minsta underrummet av V som innehåller $U_1, ..., U_m$.**

Exempel 1.11. (a) Låt $V = K^n$ och låt $E_j = \{(0, ..., 0, x, 0, ..., 0) : x \in K\}$, $j = 1, ..., n$. Underrummen E_j består av de n-tiplar i K^n för vilka alla utom eventuellt den j:te koordinaten är lika med 0. Då gäller:

$$K^n = E_1 \oplus ... \oplus E_n.$$

(b) Låt U_e vara underrummet av $\mathcal{P}(K)$ bestående av alla polynom p av formen

$$p(z) = a_0 + a_2 z^2 + ... + a_{2m} z^{2m},$$

och låt U_o vara underrummet bestående av alla polynom av formen

$$p(z) = a_1 z + a_3 z^3 + ... + a_{2m+1} z^{2m+1},$$

där m är ett icke-negativt heltal och $a_0, ..., a_{2m+1} \in K$. Då gäller:

$$\mathcal{P}(K) = U_o \oplus U_e.$$

(c) Betrakta följande underrum av R^3; $U_1 = \{(x, y, 0) \in R^3 : x, y \in R\}$, $U_2 = \{(0, 0, z) \in R^3 : z \in R\}$ och $U_3 = \{(0, y, y) \in R^3 : y \in R\}$. En godtycklig vektor $(x, y, z) \in R^3$ kan framställas som $(x, y, z) = (x, y, 0) + (0, 0, z) + (0, 0, 0)$. Därmed gäller $R^3 = U_1 + U_2 + U_3$. Men

$$\begin{aligned} (0,0,0) &= (0,1,0) + (0,0,1) + (0,-1,-1) \\ &= (0,0,0) + (0,0,0) + (0,0,0), \end{aligned}$$

så framställningen av $(0, 0, 0)$ är inte entydig varför R^3 inte kan framställas som direkta summan av U_1, U_2, U_3.

Om ett vektorrum V kan skrivas som summan av ett givet antal underrum så räcker det faktiskt att kontrollera om nollelementet har en entydig framställning för att bedömma om summan är en direkt summa, vi har nämligen:

Sats 1.12. *Antag att $U_1, ..., U_n$ är underrum av vektorrummet V. Då är $V = U_1 \oplus ... \oplus U_n$ om och endast om både (a) och (b) gäller:*

(a) $V = U_1 + ... + U_n$,

(b) den enda framställningen $0 = u_1 + ... + u_n$, där $u_j \in U_j$, ges av $u_1 = u_2 = ... = u_n = 0$.

Bevis: Se föreläsningsanteckningar.

Nästa sats ger ett enkelt villkor för när V kan framställas som direkta summan av två underrum.

Sats 1.13. *Antag att U och W är underrum av vektorrummet V. Då är $V = U \oplus W$ om och endast om $V = U + W$ och $U \cap W = \{0\}$.*

Bevis: Se föreläsningsanteckningar.

Anmärkning 1.14. Om V är summan av flera än två underrum så räcker det inte att kontrollera om de parvisa snitten är lika med $\{0\}$, jämför Exempel 1.11 (c).

I resterande delen av kapitel 1 behandlas begrepp som är centrala i teorin för ändligtdimensionella vektorrum, nämligen linjärt oberoende, linjära höljet, baser och dimension.

Linjärt oberoende

Definition 1.15. *Vektorerna $v_1, \ldots, v_n$ i ett vektorrum V över kroppen K är linjärt beroende om det finns tal $a_1, \ldots, a_n$ i K som alla inte är noll och som är sådana att*

$$a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0.$$

Att vektorerna $v_1, \ldots, v_n$ inte är linjärt beroende innebär att ovanstående ekvation är uppfylld endast för $a_1 = \ldots = a_n = 0$. Vi säger i detta fall att vektorerna är linjärt oberoende.

Exempel 1.16. (a) Polynomen x och x^2 i $\mathcal{P}(R)$ är linjärt oberoende, ty om $a_1, a_2 \in R$ har vi att $a_1x + a_2x^2 = 0$ för alla $x \in R$ om och endast om $a_1 = a_2 = 0$.

(b) Polynomen $3x - 2x^2$, x, x^2 är däremot linjärt beroende, ty $1(3x - 2x^2) + (-3)x + 2x^2 = 0$ för varje $x \in R$.

(c) Vektorn v_1 i V är linjärt oberoende om och endast om $v_1 \neq 0$.

(d) Om $M = \{v_1, \ldots, v_n\}$ är en mängd av linjärt oberoende vektorer i V så består varje icke-tom delmängd av M av linjärt oberoende vektorer (hemuppgift).

Linjära höljet

Definition 1.17. *En linjärkombination av vektorerna $v_1, \ldots, v_n$ i vektorrumet V är en vektor av formen*

$$a_1v_1 + a_2v_2 + \ldots + a_nv_n, \tag{1.1}$$

där $a_1, \ldots, a_n \in K$. Mängden av alla linjärkombinationer av vektorerna $v_1, \ldots, v_n$ kallas linjära höljet av $v_1, \ldots, v_n$ och betecknas med $[v_1, \ldots, v_n]$. Med andra ord gäller:

$$[v_1, \ldots, v_n] = \{a_1v_1 + \ldots + a_nv_n : a_1, \ldots, a_n \in K\}.$$

Anmärkning 1.18. Om vektorerna u och w kan skrivas som linjärkombinationer av vektorerna $v_1, \ldots, v_n$ så är också $u + w$ och au för $a \in K$ linjärkombinationer av $v_1, \ldots, v_n$. Då dessutom $0 \in [v_1, \ldots, v_n]$, så är $[v_1, \ldots, v_n]$ ett underrum av V.

Exempel 1.19. Givet polynomen $1, x, ..., x^n$ i $\mathcal{P}(R)$. Då gäller

$$[1, x, ..., x^n] = \mathcal{P}_n(R),$$

där $\mathcal{P}_n(R)$ betecknar underrummet av polynom med reella koefficienter och gradtal högst n.

Låt $v_1, ..., v_n$ vara vektorer i V. Då gäller $v_j \in [v_1, ..., v_n]$, $j = 1, ..., n$, vilket inses genom att välja $a_j = 1$ och $a_i = 0$ för $i \neq j$ i (1.1). Å andra sidan måste varje underrum som innehåller $v_1, ..., v_n$ även innehålla $[v_1, ..., v_n]$ på grund av slutenheten med avseende på addition och multiplikation med skalär. Då är $[v_1, ..., v_n]$ **det minsta underrummet i V som innehåller vektorerna** $v_1, ..., v_n$.

Definition 1.20. *Låt $v_1, ..., v_n$ vara vektorer i vektorrummet V. Om $[v_1, ..., v_n] = V$ säger vi att vektorerna $v_1, ..., v_n$ spänner upp (genererar, alstrar) vektorrummet V. Ett vektorrum V är ändligtdimensionellt (av ändlig dimension) om det spänns upp av en ändlig mängd av vektorer. Ett vektorrum som inte kan spännas upp av en ändlig mängd av vektorer är oändligtdimensionellt (av oändlig dimension).*

Exempel 1.21. (a) Vektorrummet K^n är av ändlig dimension, ty

$$K^n = [(1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1)].$$

(b) Vektorrummet $\mathcal{P}_n(K)$ är av ändlig dimension, ty

$$\mathcal{P}_n(K) = [1, z, z^2, ..., z^n].$$

(c) Vektorrummet $\mathcal{P}(K)$ är av oändlig dimension, ty givet en godtycklig ändlig mängd $M = \{p_1, ..., p_n\}$ av polynom i $\mathcal{P}(K)$, där m må beteckna det maximala gradtalet av polynomen i M, så gäller det för polynomet $z^{m+1} \in \mathcal{P}(K)$ att $z^{m+1} \notin [p_1, ..., p_n]$.

Följande hjälpresultat är användbart i fortsättningen:

Lemma 1.22. *Om vektorerna $v_1, ..., v_n$ är linjärt beroende i vektorrummet V och $v_1 \neq 0$, så finns det ett index j, $2 \leq j \leq n$, sådant att både (a) och (b) gäller:*

(*a*) $v_j \in [v_1, ..., v_{j-1}]$,

(*b*) $[v_1, ..., v_{j-1}, v_{j+1}, ..., v_n] = [v_1, ..., v_n]$.

Bevis: Se föreläsningsanteckningar.

Vi avslutar detta avsnitt med två viktiga resultat.

Sats 1.23. *Antag att V är ett ändligtdimensionellt vektorrum och att $V = [v_1, ..., v_n]$, där $v_1, ..., v_n$ är vektorer i V. Om vektorerna $u_1, ..., u_m \in V$ är linjärt oberoende så gäller det att $m \leq n$.*

Bevis: Antag att $V = [v_1, ..., v_n]$ och att vektorerna $u_1, ..., u_m \in V$ är linjärt oberoende. Vi visar att $m \leq n$ genom att stegvis införa ett u_j och stryka ett v_j ur en ordnad mängd.

Steg 1: Den ordnade mängden $M_0 = \{v_1, ..., v_n\}$ spänner upp V. Vi tillfogar $u_1 \neq 0$ till mängden och erhåller $M' = \{u_1, v_1, ..., v_n\}$, som då blir en mängd av linjärt beroende vektorer. Då ger Lemma 1.22 att vi kan stryka ett av v_j:na och erhåller då en ordnad mängd M_1 med n vektorer som spänner upp V.

$\vdots$

Steg j: Den ordnade mängden M_{j-1} med n vektorer från Steg j–1 spänner upp V, så genom att tillfoga vektorn $u_j \neq 0$ till M_{j-1}, precis efter $u_1, ..., u_{j-1}$ får vi en ordnad mängd M' av linjärt beroende vektorer. Enligt Lemma 1.22 finns det en vektor i M' som tillhör linjära höljet av vektorerna som föregår vektorn. Eftersom $u_1, ..., u_j$ är linjärt oberoende måste denna vektor vara en av v_j :na. Vi stryker denna vektor v_j ur mängden M' och erhåller den ordnade mängden M_j med n stycken vektorer av vilka $u_1, ..., u_j$ är de första vektorerna i mängden. Enligt Lemma 1.22 spänner vektorerna i M_j upp V.

Efter **Steg m** har vi tillfogat alla $u_j : n$ och proceduren avstannar. Om vi i något Steg j kunde lägga till ett u_j men inget v_j skulle finnas att stryka skulle vi få en motsägelse, ty då skulle mängden $\{u_1, ..., u_j\}$ vara en mängd av linjärt beroende vektorer. Därmed måste proceduren fortgå till Steg m och $m \leq n$.

Sats 1.24. *Varje underrum U av ett ändligtdimensionellt vektorrum V är ändligtdimensionellt.*

Bevis: Antag att V är av ändlig dimension och att U är ett underrum av V.

Steg 1: Om $U = \{0\}$ så är U ändligtdimensionellt och proceduren avstannar. Om $U \neq \{0\}$, så väljer vi $v_1 \in U$ sådant att $v_1 \neq 0$.

Steg j: $(j \geq 2)$ Om $U = [v_1, ..., v_{j-1}]$ så är U ändligtdimensionell och proceduren avstannar. Om $U \neq [v_1, ..., v_{j-1}]$, så väljs $v_j \in U$ sådant att $v_j \notin [v_1, ..., v_{j-1}]$.

Efter Steg j, $(j \geq 2)$, gäller $v_j \notin [v_1, ..., v_{j-1}]$, $v_{j-1} \notin [v_1, ..., v_{j-2}], ..., v_2 \notin [v_1]$ och $v_1 \neq 0$. Då ger Lemma 1.22 att $v_1, ..., v_j$ är linjärt oberoende i U (och därmed även i V). Eftersom V är ändligtdimensionellt finns det vektorer $w_1, ..., w_m$ så att $V = [w_1, ..., w_m]$. Med stöd av Sats 1.23 får vi att ovanstående procedur måste avstanna i Steg $n+1$ med $U = [v_1, ..., v_n]$ för något n sådant att $n \leq m$. Därmed är U ändligtdimensionell.

Baser

Definition 1.25. *Vektorerna $v_1, ..., v_n \in V$ utgör en bas för vektorrummet V om både (a) och (b) gäller:*

(a) vektorerna $v_1, ..., v_n$ är linjärt oberoende,

(b) $V = [v_1, ..., v_n]$.

Exempel 1.26. De naturliga baserna för K^n och $\mathcal{P}_n(K)$ utgörs av $(1, 0, ..., 0), ..., (0, ..., 0, 1)$ respektive $1, z, ..., z^n$.

Observera att en bas för ett vektorrum är ingalunda entydigt bestämd. Däremot, om vi har fixerat en bas för V, så gäller det att varje vektor i V har en entydig framställning som linjärkombination av basvektorerna.

Sats 1.27. *Vektorerna $v_1, ..., v_n \in V$ är en bas för vektorrummet V om och endast om varje vektor $v \in V$ har en entydig framställning*

$$v = a_1 v_1 + ... + a_n v_n, \tag{1.2}$$

där $a_1, ..., a_n \in K$. Talen $a_1, ..., a_n$ kallas koordinaterna för v med avseende på basen $v_1, ..., v_n$.

Bevis: Se föreläsningsanteckningar.

En bas för ett ändligtdimensionellt vektorrum kan alltid erhållas ur en mängd vektorer som spänner upp vektorrummet:

Sats 1.28. *Antag att vektorerna $v_1, ..., v_n \in V$, $V \neq \{0\}$, spänner upp vektorrummet V. Då finns det en delmängd $\{u_1, ..., u_m\}$ av $\{v_1, ..., v_n\}$ sådan att vektorerna $u_1, ..., u_m$ utgör en bas för V.*

Bevis: Om $n = 1$ är saken klar med $u_1 = v_1 \neq 0$. Antag att $n \geq 2$ och att $[v_1, ..., v_n] = V \neq \{0\}$. Definiera den ordnade mängden $B = \{v_1, ..., v_n\}$, för vilken gäller att $[B] = V$.

Steg 1: Om $v_1 = 0$, stryk v_1 ur B. Gör ingen ändring i B om $v_1 \neq 0$.

Steg j: ($j \geq 2$) Låt D beteckna den delmängd av $\{v_1, ..., v_{j-1}\}$ som ej strykits efter Steg j–1. Om $D = \emptyset$ och $v_j = 0$, stryk v_j ur B. Om $D \neq \emptyset$ och $v_j \in [D]$, stryk v_j ur B.

Stoppa proceduren efter Steg n. För vår slutliga mängd $B = \{u_1, ..., u_m\}$ gäller $[u_1, ..., u_m] = V$, ty vi har strykit vektorer som låg i linjära höljet av de kvarvarande. Vidare gäller för varje u_j i vår slutliga mängd $B = \{u_1, ..., u_m\}$ att $u_j \neq [u_1, ..., u_{j-1}]$. Då ger Lemma 1.22 att vektorerna $u_1, ..., u_m$ är linjärt oberoende. Därmed har vi erhållit en bas $u_1, ..., u_m$ för V.

Korollarium 1.29. *Varje ändligtdimensionellt vektorrum V, $V \neq \{0\}$, har en bas.*

Bevis: Antag att V är ändligtdimensionellt och $\neq \{0\}$. På basen av Definition 1.20 existerar det vektorer $v_1, ..., v_n \in V$ sådana att $[v_1, ..., v_n] = V$. Med stöd av Sats 1.28 har då V en bas.

Sats 1.30. *Varje linjärt oberoende mängd av vektorer i ett ändligtdimensionellt vektorrum V kan utvidgas till en bas för V.*

Bevis: Antag att $v_1, ..., v_m$ är linjärt oberoende vektorer i ett ändligtdimensionellt vektorrum V. Det finns då vektorer $w_1, ..., w_n$ sådana att $[w_1, ..., w_n] = V$. Definiera den ordnade mängden $B = \{v_1, ..., v_m\}$.

Steg 1: Lämna B oförändrad om $w_1 \in [v_1, ..., v_m]$. Sätt $B = \{v_1, ..., v_m, w_1\}$ om $w_1 \notin [v_1, ..., v_m]$.

Steg j: Lämna B oförändrad om $w_j \in [B]$. Lägg till w_j som sista element i B om $w_j \notin [B]$.

Efter varje steg ger Lemma 1.22 att vektorerna i B är linjärt oberoende. Efter Steg n gäller det att $w_j \in [B]$, $j = 1, ..., n$. Men då gäller det att $[B] = V$, och vektorerna i mängden B erhållen efter Steg n är en bas för V.

Med hjälp av föregående sats erhåller vi följande uppdelning av ett ändligtdimensionellt vektorrum i en direkt summa av två underrum:

Sats 1.31. *Antag att vi har ett ändligtdimensionellt vektorrum V och ett underrum U av V. Då finns det ett underrum W av V* sådant att $V = U \oplus W$.

Bevis: Om $V = \{0\}$ så gäller $V = \{0\} \oplus \{0\}$. Antag att $V \neq \{0\}$ är ändligtdimensionellt. Låt U vara ett underrum av V. Om $U = \{0\}$ så är $V = U \oplus V$ och om $U = V$ så är $V = U \oplus \{0\}$. Antag då att $U \neq \{0\}, V$. Då är U ändligtdimensionellt (Sats 1.24) och har därmed en bas $u_1, ..., u_m$ (Koroll. 1.29). Denna bas är linjärt oberoende i V och kan utökas till en bas $u_1, ..., u_m, w_1, ..., w_n$ för V (Sats 1.30). Definiera $W = [w_1, ..., w_n]$. Enligt Sats 1.13 bör vi visa att $V = U + W$ och att $U \cap W = \{0\}$. Antag att $v \in V$. Då finns det $a_1, ..., a_m, b_1, ..., b_n \in K$ sådana att

$$v = (a_1 u_1 + ... + a_m u_m) + (b_1 w_1 + ... + b_n w_n) = u + w,$$

Alltså $v = u + w$ där $u \in U$ och $w \in W$, så $v \in U + W$. Därmed gäller $V = U + W$. Antag nu att $v \in U \cap W$. Då finns $a_1, ..., a_m, b_1, ..., b_n \in K$ sådana att

$$v = a_1 u_1 + ... + a_m u_m = b_1 w_1 + ... + b_n w_n,$$

så

$$a_1 u_1 + \ldots + a_m u_m - b_1 w_1 - \ldots - b_n w_n = 0 .$$

Eftersom $u_1, \ldots, u_m, w_1, \ldots, w_n$ är linjärt oberoende har vi $a_1 = \ldots = a_m = b_1 = \ldots = b_n = 0$. Därmed är $v = 0$ och $U \cap W = \{0\}$. Alltså gäller $V = U \oplus W$ och satsen är bevisad.

Dimension

Vi har behandlat ändligtdimensionella vektorrum utan att definiera begreppet dimension. Det i förra avsnittet införda begreppet bas ger en lämplig grund för definitionen. Det gäller nämligen att:

Sats 1.32. *Om vektorerna $v_1, \ldots, v_n$ och $u_1, \ldots, u_m$ utgör två baser för det ändligtdimensionella vektorrummet V så är $m = n$.*

Bevis: Vi har att $V = [v_1, \ldots, v_n]$ och att $u_1, \ldots, u_m$ är linjärt oberoende. Då ger Sats 1.23 att $m \leq n$. Det gäller även att $V = [u_1, \ldots, u_m]$ och att $v_1, \ldots, v_n$ är linjärt oberoende, så $n \leq m$ med stöd av Sats 1.23. Därmed gäller $m = n$.

Definition 1.33. *Dimensionen av ett ändligtdimensionellt vektorrum V, $V \neq \{0\}$, definieras som antalet vektorer i en bas för V och betecknas med* $\dim V$. Vidare defineras $\dim \{0\} = 0$.

Exempel 1.34. Med stöd av de naturliga baserna för K^n och $\mathcal{P}_n(K)$, jämför Exempel 1.26, har vi då att $\dim K^n = n$ och $\dim \mathcal{P}_n(K) = n + 1$.

Om U är ett underrum av ett ändligtdimensionellt vektorrum V så är även U ändligtdimensionellt med stöd av Sats 1.24. Om $U = \{0\}$ så är $0 = \dim U \leq \dim V$. Om $U \neq \{0\}$ ger Korollarium 1.29 att U har en bas $u_1, \ldots, u_m$. Eftersom denna bas är linjärt oberoende i V så ger Sats 1.30 att basen för U kan utvidgas till en bas $v_1, \ldots, v_n$ för V, där $m \leq n$. Därmed har vi bevisat:

Sats 1.35. *Antag att V är ett ändligtdimensionellt vektorrum och att U är ett underrum av V. Då gäller det att* $\dim U \leq \dim V$.

Antag att V är ändligtdimensionellt med $\dim V = n \geq 1$ och att vektorerna $v_1, \ldots, v_n \in V$ spänner upp V. Enligt Sats 1.28 finns det då en delmängd $u_1, \ldots, u_m$ av $v_1, \ldots, v_n$ som utgör en bas för V och eftersom $\dim V = n$ så måste $m = n$ gälla och vi har bevisat satsen:

Sats 1.36. *Antag att V är ett ändligtdimensionellt vektorrum med* $\dim V = n \geq 1$ *och att vektorerna $v_1, \ldots, v_n \in V$ spänner upp V. Då är vektorerna $v_1, \ldots, v_n$ en bas för V.*

En linjärt oberoende mängd av vektorer för ett ändligtdimensionellt vektorrum utgör en bas om antalet vektorer sammanfaller med dimensionen:

Sats 1.37. *Antag att* V *är ett ändligtdimensionellt vektorrum med* $\dim V = n \geq 1$ *och att vektorerna* $v_1, ..., v_n \in V$ *är linjärt oberoende. Då är vektorerna* $v_1, ..., v_n$ *en bas för* V.

Bevis: Antag att V är ändligtdimensionellt med $\dim V = n \geq 1$ och att vektorerna $v_1, ..., v_n \in V$ är linjärt oberoende. Enligt Sats 1.30 kan mängden $\{v_1, ..., v_n\}$ utvidgas till en bas $u_1, ..., u_m$ för V. Eftersom $\dim V = n$ så måste $m = n$ gälla och $v_1 = u_1, ..., v_n = u_m$. Därmed är satsen är bevisad.

Avslutningsvis bevisas två satser, den första ger en formel för dimensionen av summan av två underrum av ett ändligtdimensionellt vektorrum och den andra ett resultat om framställningen av V som en direkt summa av underrum. Satserna bevisas på föreläsningen.

Sats 1.38. *Om* U_1 *och* U_2 *är underrum av ett ändligtdimensionellt vektorrum* V, *så gäller*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2). \tag{1.3}$$

Sats 1.39. *Antag att vektorrummet* V *är ändligtdimensionellt med* $\dim V \geq 1$ *och att* $U_1, ..., U_n$ *är underrum av* V *sådana att*

$$V = U_1 + ... + U_n \tag{1.4}$$

och

$$\dim V = \dim U_1 + ... + \dim U_n. \tag{1.5}$$

Då gäller det att $V = U_1 \oplus ... \oplus U_n$.

Appendix A

Vi ger här axiomen för en talkropp K. I denna kurs är K den reella eller den komplexa talkroppen (R eller C).

Definition *En kropp är en mängd K som är försedd med två operationer kallade addition och multiplikation och som uppfyller axiomen (i), (ii) och (iii).*

(i) Till varje ordnat par (x, y) av element i K finns ordnat ett element $x+y$ i K (K slutet under "addition"). Sammansättningen $+$ ("addition") har egenskaperna:

1. $x + y = y + x$, *för alla* $x, y \in K$,
2. $(x + y) + z = x + (y + z)$, *för alla* $x, y, z \in K$,
3. *I mängden K finns ett element 0 (nollelement) sådant att $x + 0 = x$ för alla $x \in K$,*
4. *För varje $x \in K$ existerar $y \in K$ sådant att $x + y = 0$ (y additiv invers till x).*

(ii) Till alla tal $x, y \in K$ finns ordnat ett element $x \cdot y$ i K, ofta betecknat xy, (K sluten under "multiplikation" med skalär). Denna "multiplikation" har egenskaperna:

1. $x(yz) = (xy)z$, *för alla* $x, y, z \in K$,
2. $xy = yx$, *för alla* $x, y \in K$,
3. *K har ett element $1 \neq 0$ sådant att $1 \cdot x = x$, för varje $x \in K$,*
4. *Om $x \in K$ och $x \neq 0$ så finns ett element $y \in K$ sådant att $x \cdot y = 1$.*

(iii) För alla $x, y, z \in K$ gäller:

1. $x(y + z) = xy + xz$.

2. Linjära avbildningar

Exempel 2.1. Låt $\mathrm{Mat}(m,n,K)$ beteckna mängden av $m \times n$ matriser med element i $K = R$ eller C. Då bestämmer $A \in \mathrm{Mat}(m,n,K)$ en avbildning $T : K^n \mapsto K^m$ genom tilldelningen $x \mapsto Ax$, där vektorerna i K^n och K^m tolkas som kolonnvektorer. Egenskaperna för matrisaddition och multiplikation av matris med skalär ger att för alla $x, y \in K^n$ och alla $a \in K$ gäller:

$$T(x+y) = A(x+y) = Ax + Ay = T(x) + T(y),$$
$$T(ax) = A(ax) = aAx = aT(x).$$

Ovanstående egenskaper tar vi som definition på en linjär avbildning.

Definition 2.2. *En linjär avbildning (linjär transformation) från vektorrummet V till vektorrummet W är en funktion $T : V \mapsto W$ som uppfyller*

$$T(ax + by) = aT(x) + bT(y)$$

för alla $x, y \in V$ och alla skalärer $a, b \in K$. Vi kallar V definitionsrummet för T. Mängden av alla linjära avbildningar från V till W betecknas med $\mathcal{L}(V,W)$.

Anmärkning 2.3. Det är klart att $T(0) = 0$. Man betecknar ofta $T(x)$ med Tx.

Exempel 2.4. Exempel på linjära avbildningar. (Se föreläsningsanteckningar).

Vektorrummet $\mathcal{L}(V,W)$

Om V och W är två vektorrum så kan $\mathcal{L}(V,W)$ göras till ett vektorrum genom att definiera addition och multiplikation med skalär på $\mathcal{L}(V,W)$.

För $S, T \in \mathcal{L}(V,W)$ definerar vi avbildningen $S + T \in \mathcal{L}(V,W)$ genom

$$(S+T)(x) = S(x) + T(x), \quad \text{för varje } x \in V.$$

Man verifierar lätt att $S+T \in \mathcal{L}(V,W)$, (ty om $x, y \in V$, $a, b \in K$ gäller: $(S+T)(ax+by) = S(ax+by)+T(ax+by) = \ldots = a(S(x)+T(x)) + b(S(y)+T(y)) = a((S+T)(x)) + b((S+T)(y))$. Vidare är det lätt att verifiera att axiomen för addition är uppfyllda. Nollelementet ges av funktionen $0(x) = 0$ för varje $x \in V$.

För $T \in \mathcal{L}(V,W)$ och $a \in K$ definerar vi avbildningen $aT \in \mathcal{L}(V,W)$ genom

$$(aT)(x) = a(T(x)), \quad \text{för varje } x \in V.$$

Man verifierar även nu lätt att $aT \in \mathcal{L}(V, W)$ och att axiomen för multiplikation med skalär är uppfyllda, likväl som de distributiva axiomen. Därmed är $\mathcal{L}(V, W)$ ett vektorrum.

Produkter av linjära avbildningar

Antag att U, V och W är vektorrum över K. Låt $T \in \mathcal{L}(U, V)$ och $S \in \mathcal{L}(V, W)$. Då definieras produkten $ST \in \mathcal{L}(U, W)$ genom

$$(ST)(u) = S(T\,u) \text{ för varje } u \in U.$$

Därmed utgör produkten ST den vanliga sammansättningen $S \circ T$ av funktionerna S och T. Om $u, v \in U$ och $a, b \in K$ så gäller: $(ST)(a\,u + b\,v) = S(T(a\,u + b\,v)) = S(a\,T(u) + b\,T(v)) = a\,S(T\,u) + b\,S(T\,v) = a\,(ST)(u) + b\,(ST)(v)$. Därmed har vi att $ST \in \mathcal{L}(U, W)$.

Följande egenskaper för produkter är lätta att verifiera:

1. **Associativitet:** $(T_1T_2)T_3 = T_1(T_2T_3)$ såsnart de linjära avbildningarna T_1, T_2 och T_3 är väldefinierade.
2. **Produkt med identitet:** $TI = T$ och $IT = T$ om $T \in \mathcal{L}(V, W)$. I den första ekvationen är I den identiska avbildningen på V och i den andra ekvationen den identiska avbildningen på W.
3. **Distributivitet:** $(S_1 + S_2)T = S_1T + S_2T$ och $S(T_1 + T_2) = ST_1 + ST_2$ såsnart $T, T_1, T_2 \in \mathcal{L}(U, V)$ och $S, S_1, S_2 \in \mathcal{L}(V, W)$.

I allmänhet är produkten inte kommutativ. Låt exempelvis $T \in \mathcal{L}(\mathcal{P}(R), \mathcal{P}(R))$ svara mot att derivera polynom och låt $S \in \mathcal{L}(\mathcal{P}(R), \mathcal{P}(R))$ svara mot att multiplicera ett polynom med x^2. Då är $(ST)(p) = S(T\,p) = S(p') = x^2\,p'(x)$, men $(TS)(p) = T(S\,p) = T(x^2\,p(x)) = 2x\,p(x) + x^2\,p'(x)$.

Nollrum och värderum

Definition 2.5. *Nollrummet $N(T)$ för en linjär avbildning $T \in \mathcal{L}(V, W)$ definieras som mängden av alla element i V vilka av T avbildas på nollelementet i W,*

$$N(T) = \{v \in V : T\,v = 0\}.$$

Exempel 2.6. Antag att $T \in \mathcal{L}(\mathcal{P}(R), \mathcal{P}(R))$ definieras av $T(p) = p'$ för varje $p \in \mathcal{P}(R)$, där p' betecknar derivatan av p. Då är $N(T)$ mängden av konstanta funktioner på R.

Sats 2.7. *Om $T \in \mathcal{L}(V, W)$, så är $N(T)$ ett underrum av V.*

Bevis: Se föreläsningsanteckningar.

Definition 2.8. *En linjär avbildning $T : V \mapsto W$ kallas injektiv om för alla $u, v \in V$ gäller att $Tu = Tv$ medför att $u = v$.*

För att försäkra sig om att en linjär avbildning är injektiv krävs att enbart nollelementet i V avbildas på nollelementet i W:

Sats 2.9. *Låt $T \in \mathcal{L}(V, W)$. Då är T injektiv om och endast om $N(T) = \{0\}$.*

Bevis: Se föreläsningsanteckningar.

Definition 2.10. *Värderummet $R(T)$ för en linjär avbildning $T \in \mathcal{L}(V, W)$ definieras som delmängden av W bestående av alla vektorer Tv, $v \in V$,*

$$R(T) = \{Tv : v \in V\}.$$

Sats 2.11. *Om $T \in \mathcal{L}(V, W)$ så är $R(T)$ ett underrum av W.*

Bevis: Se föreläsningsanteckningar.

Definition 2.12. *En linjär avbildning $T \in \mathcal{L}(V, W)$ kallas surjektiv om $R(T) = W$.*

Exempel 2.13. Avbildningen $T \in \mathcal{L}(\mathcal{P}(R), \mathcal{P}(R))$ som definieras av $T(p) = p'$ för varje $p \in \mathcal{P}(R)$, där p' betecknar derivatan av p, är surjektiv men inte injektiv, ty för varje polynom $q \in \mathcal{P}(R)$ finns det ett polynom $p \in \mathcal{P}(R)$ sådant att p avbildas på q och vidare avbildas även $p + r \in \mathcal{P}(R)$ på q då r är ett godtyckligt konstant polynom.

Följande resultat är av central betydelse:

Sats 2.14. *(Dimensionssatsen) Om V är ett ändligtdimensionellt vektorrum och $T \in \mathcal{L}(V, W)$ så är $R(T)$ ett ändligtdimensionellt underrum av W och*

$$\dim V = \dim N(T) + \dim R(T).$$

Bevis: Se föreläsningsanteckningar.

Korollarium 2.15. *Om V, W är ändligtdimensionella vektorrum med $\dim V > \dim W$ så är ingen linjär avbildning $T \in \mathcal{L}(V, W)$ injektiv.*

Bevis: Antag att V och W är ändligtdimensionella med $\dim V > \dim W$ och antag att $T \in \mathcal{L}(V, W)$. Med stöd av Sats 2.14, (och Sats 1.35 "dimensionen för ett underrum U till ett ändligtdimensionellt vetorrum W är inte större än dimensionen för W"), erhålls då:

$$\begin{aligned}\dim N(T) &= \dim V - \dim R(T)\\ &\geq \dim V - \dim W\\ &> 0.\end{aligned}$$

Alltså är $\dim N(T) > 0$ och då är T inte injektiv (Sats 2.9).

Korollarium 2.16. *Om V, W är ändligtdimensionella vektorrum med* $\dim V < \dim W$ *så är ingen linjär avbildning $T \in \mathcal{L}(V, W)$ surjektiv.*

Bevis: Antag att V och W är ändligtdimensionella med $\dim V < \dim W$ och antag att $T \in \mathcal{L}(V, W)$. Med stöd av Sats 2.14 gäller då:

$$\begin{aligned}\dim R(T) &= \dim V - \dim N(T)\\ &\leq \dim V\\ &< \dim W.\end{aligned}$$

Alltså är $\dim R(T) < \dim W$ och då kan inte $R(T) = W$ gälla. Därmed är avbildningen T inte surjektiv.

Vi tillämpar ovanstående korollarier på linjära ekvationssystem:

Exempel 2.17. Antag att vi har givet två positiva heltal m och n samt $m \cdot n$ skalärer $a_{j,k} \in K$ där $j = 1, ..., m$ och $k = 1, ..., n$. Vi definierar $T \in \mathcal{L}(K^n, K^m)$ genom

$$T(x_1, ..., x_n) = (\sum_{k=1}^{n} a_{1,k}\, x_k, ..., \sum_{k=1}^{n} a_{m,k}\, x_k),$$

för varje vektor $x = (x_1, ..., x_n) \in K^n$. Ekvationen $Tx = 0$, (där 0 är nollelementet i K^m), kan då skrivas som ett **homogent system av linjära ekvationer:**

$$\begin{aligned}a_{1,1}\, x_1 + a_{1,2}\, x_2 + ... + a_{1,n}\, x_n &= 0\\ &\vdots\\ a_{m,1}\, x_1 + a_{m,2}\, x_2 + ... + a_{m,n}\, x_n &= 0.\end{aligned}$$

Skalärerna $a_{j,k} \in K$ är bekanta och vi söker lösningar $x = (x_1, ..., x_n) \in K^n$ till ovanstående ekvationssystem, där benämningen homogen avser att högra ledet i varje ekvation utgörs av skalären $0 \in K$.

Vi har alltså m ekvationer och n obekanta. Ekvationssystemet har givetvis den triviala lösningen $x_1 = ... = x_n = 0$, men har vi möjligen icke-triviala lösningar? Svaret på frågan ges av nollrummet för T; vi har lösningar $x \neq 0$ om och endast om $N(T) \neq \{0\}$. Detta inträffar då T inte är injektiv (Sats 2.9). Om $n > m$ ger korollarium 2.15 att T inte är injektiv. Alltså: ett homogent system av linjära ekvationer med flera obekanta än ekvationer, $n > m$, har icke-triviala lösningar.

Betrakta nu ekvationen $Tx = c$ där $0 \neq c = (c_1, ..., c_m) \in K^m$. Vi kan skriva ekvationen som ett **inhomogent system av linjära ekvationer**:

$$
\begin{aligned}
a_{1,1}\,x_1 + a_{1,2}\,x_2 + \ldots + a_{1,n}\,x_n &= c_1 \\
&\vdots \\
a_{m,1}\,x_1 + a_{m,2}\,x_2 + \ldots + a_{m,n}\,x_n &= c_m.
\end{aligned}
$$

Igen är skalärerna $a_{j,k} \in K$ bekanta och vi söker lösningar $x = (x_1, ..., x_n) \in K^n$ till ekvationssystemet, där benämningen inhomogen avser att minst en av skalärerna $c_j \in K$ är olika noll.

Nu frågar vi oss om det för varje val av skalärer $c_1, ..., c_m \in K$ existerar en lösning till ekvationssystemet. Detta gäller om $R(T) = K^m$, det vill säga om T är surjektiv. Om $n < m$ ger Korollarium 2.16 att T inte är surjektiv. Alltså: Ett ekvationssystem med flera ekvationer än obekanta, $n < m$, kan inte lösas för varje val av högersida $c_1, ..., c_m \in K$.

Matrisframställning av linjära avbildningar

Vi noterade i Exempel 2.1 att varje $m \times n$ matris definierar en linjär avbildning från K^n till K^m. Vi skall nu visa att varje linjär avbildning mellan två ändligtdimensionella vektorrum V och W kan framställas med en matris som beror av valet av baser för de två vektorrummen.

Antag att $v_1, ..., v_n$ är en bas för vektorrummet V och att avbildningen $T : V \mapsto W$ är linjär. Vektorerna $T\,v_j$, $j = 1, ..., n$, bestämmer värdet på $T\,v \in W$ för varje $v \in V$, ty givet den entydiga framställningen $v = c_1\,v_1 + \ldots + c_n\,v_n$ så har vi

$$T\,v = c_1\,T\,v_1 + \ldots + c_n\,T\,v_n.$$

Antag nu att $w_1, ..., w_m$ är en bas för vektorrummet W. För varje k, $k = 1, ..., n$, kan då $T\,v_k$ framställas på ett entydigt sätt,

$$T\,v_k = a_{1,k}\,w_1 + \ldots + a_{m,k}\,w_m, \tag{2.1}$$

där $a_{j,k}$ är skalärer ur K. Den linjära avbildningen $T \in \mathcal{L}(V, W)$ bestäms av skalärerna $a_{j,k}$, eftersom T bestäms av sina värden $T\,v_k$ på basen $v_1, ..., v_n$. Vi kan då representera T med matrisen

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}. \tag{2.2}$$

Matrisen (2.2) kallas matrisen för T med avseende på baserna $v_1, ..., v_n$ och $w_1, ..., w_m$ och betecknas

$$\mathcal{M}(T, (v_1, ..., v_n), (w_1, ..., w_m)),$$

eller kortare $\mathcal{M}(T)$ om det av sammanhanget är klart vilka baser för V och W som avses. Vi ser att den k:te kolonnen i (2.2) innehåller koordinaterna för framställningen av $T v_k$ i basen $w_1, ..., w_m$.

Vid avbildningar från K^n till K^m avses de naturliga baserna, (se Exempel 1.26), ifall inte annat anges. Vi återgår till att studera linjära ekvationssystem:

Exempel 2.18. Antag att vi har givet två positiva heltal m och n samt $m \cdot n$ skalärer $a_{j,k} \in K$ där $j = 1, ..., m$ och $k = 1, ..., n$. Vi definierar $T \in \mathcal{L}(K^n, K^m)$ genom

$$T(x_1, ..., x_n) = (\sum_{k=1}^{n} a_{1,k}\, x_k, ..., \sum_{k=1}^{n} a_{m,k}\, x_k),$$

för varje vektor $x = (x_1, ..., x_n) \in K^n$. Nu antar vi att K^n och K^m är försedda med sina naturliga baser $v_1, ..., v_n$ respektive $w_1, ..., w_m$. För $k = 1, ..., n$ erhålls då

$$\begin{aligned} T\, v_k &= (a_{1,k} \cdot 1, ..., a_{m,k} \cdot 1) \\ &= a_{1,k}\, w_1 + a_{2,k}\, w_2 + ... + a_{m,k}\, w_m. \end{aligned}$$

Därmed ges $\mathcal{M}(T)$ av (2.2). Låt oss beteckna den k:te kolonnen i $\mathcal{M}(T)$ med a_k, $k = 1, ..., n$.

Betrakta nu det inhomogena linjära ekvationssystemet $T x = c$, där $c = (c_1, ..., c_m) \in K^m$ och $x = (x_1, ..., x_n) \in K^n$. Ekvationssystemet kan skrivas i formen

$$x_1\, a_1 + ... + x_n\, a_n = c.$$

Därmed har systemet en lösning om $c \in [a_1, ..., a_n] = R(T)$.

Ifall $n \geq m$ kan systemet lösas för varje $c \in K^m$ om $K^m \subseteq [a_1, ..., a_n]$. Om $n < m$ såg vi redan i Exempel 2.17 att systemet inte kan lösas för alla c, ty $K^m \not\subseteq [a_1, ..., a_n]$.

Antag att $T x = c$, alltså att $x \in K^n$ är en lösning för ett givet $c \in K^m$. Då är även $x + x_0$ en lösning för $x_0 \in N(T)$, ty $T(x + x_0) = T(x) + T(x_0) = c + 0 = c$. Antag att $T y = c$ för $y \in K^n$. Då är $y = x + (y - x)$ och $T(y - x) = T(y) - T(x) = c - c = 0$, så $y - x \in N(T)$. Därmed är alla lösningar av formen $x + x_0$, där $x_0 \in N(T)$. Lösningen är alltså entydigt bestämd om och endast om $N(T) = \{0\}$.

Exempel 2.19. Rotation i R^2. Se föreläsningsanteckningar.

Antag nu att vi har baserna $v_1, ..., v_n$ och $w_1, ..., w_m$ för vektorrummen V respektive W. Varje $T \in \mathcal{L}(V, W)$ har då en matris $\mathcal{M}(T)$ med avseende på de givna baserna. För $S, T \in \mathcal{L}(V, W)$ har vi för $k = 1, ..., n$ att

$$T\, v_k = a_{1,k}\, w_1 + ... + a_{m,k}\, w_m \quad \text{och} \quad S\, v_k = b_{1,k}\, w_1 + ... + b_{m,k}\, w_m.$$

Alltså gäller:

$$(S+T)(v_k) = S\,v_k + T\,v_k = (a_{1,k}+b_{1,k})\,w_1 + \ldots + (a_{m,k}+b_{m,k})\,w_m,$$

och vi har därmed att

$$\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T). \tag{2.3}$$

För $c \in K$ och $T \in \mathcal{L}(V,W)$ gäller

$$(cT)(v_k) = c(T\,v_k) = c\,a_{1,k}\,w_1 + \ldots + c\,a_{m,k}\,w_m,$$

och vi har då att

$$\mathcal{M}(cT) = c\,\mathcal{M}(T). \tag{2.4}$$

Om vi med $\mathrm{Mat}(m,n,K)$ betecknar alla $m\times n$ matriser med element i K så är det lätt att kontrollera att axiomen för ett vektorrum i Definition 1.1 är uppfyllda med de vanliga räkneoperationerna för matriser. Nollelementet i $\mathrm{Mat}(m,n,K)$ ges av $m \times n$ matrisen där alla element är noll.

Betrakta nu $S \in \mathcal{L}(U,V)$ och $T \in \mathcal{L}(V,W)$, där $v_1,\ldots,v_n$ är en bas för V, $w_1,\ldots,w_m$ är en bas för W och $u_1,\ldots,u_p$ är en bas för U. Låt matriserna för S och T ges av

$$\mathcal{M}(T) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \quad \text{och} \quad \mathcal{M}(S) = \begin{pmatrix} b_{1,1} & \cdots & b_{1,p} \\ \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,p} \end{pmatrix}.$$

För $k \in \{1,\ldots,p\}$ gäller då

$$\begin{aligned}(TS)(u_k) &= T(\sum_{r=1}^{n} b_{r,k}\,v_r) = \sum_{r=1}^{n} b_{r,k} T\,v_r \\ &= \sum_{r=1}^{n} b_{r,k}(\sum_{j=1}^{m} a_{j,r}\,w_j) \\ &= \sum_{j=1}^{m}(\sum_{r=1}^{n} a_{j,r}\,b_{r,k})\,w_j.\end{aligned}$$

Tydligen ges $\mathcal{M}(TS)$ av en $m \times p$ matris vars element i position (j,k) är $\sum_{r=1}^{n} a_{j,r}\,b_{r,k}$. Därmed gäller det att

$$\mathcal{M}(TS) = \mathcal{M}(T)\,\mathcal{M}(S). \tag{2.5}$$

Antag nu att $v_1,\ldots,v_n$ är en bas för vektorrummet V. För $v \in V$ har vi en entydig framställning:

$$v = b_1\,v_1 + \ldots + b_n\,v_n, \quad b_j \in K.$$

Matrisen för v betecknas $\mathcal{M}(v,(v_1,\ldots,v_n))$ eller enbart $\mathcal{M}(v)$ och definieras som den $n \times 1$ kolonnvektor som består av koordinaterna för v:

$$\mathcal{M}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Nästa sats ger oss "receptet" på hur man givet $T \in \mathcal{L}(V, W)$ och $v \in V$, där V och W är ändligtdimensionella, beräknar koordinaterna för Tv i basen för W.

Sats 2.20. *Antag att $T \in \mathcal{L}(V, W)$ och att $v_1, \dots, v_n$ och $w_1, \dots, w_m$ är baser för V respektive W. Då gäller:*

$$\mathcal{M}(Tv) = \mathcal{M}(T)\,\mathcal{M}(v)$$

för varje $v \in V$.

Bevis: Se föreläsningsanteckningar.

Inverterbarhet

Definition 2.21. *Låt V, W vara vektorrum och låt $T \in \mathcal{L}(V, W)$. Då är T inverterbar om det finns en linjär avbildning $S \in \mathcal{L}(W, V)$ sådan att $ST = I$, där I är identiska avbildningen på V, och $TS = I$, där I är identiska avbildningen på W. En linjär avbildning $S \in \mathcal{L}(W, V)$ som uppfyller $ST = I$ och $TS = I$ kallas för inversen till avbildningen $T \in \mathcal{L}(V, W)$.*

Anmärkning 2.22. Om S och S' är inverser till T så gäller

$$S = SI = S(TS') = (ST)S' = IS' = S',$$

varför inversen är entydigt bestämd och vi betecknar den vanligen med T^{-1}. Alltså gäller $TT^{-1} = I$ och $T^{-1}T = I$.

Vi har följande karakterisering av inverterbara avbildningar:

Sats 2.23. *En linjär avbildning är inverterbar om och endast om den är injektiv och surjektiv.*

Bevis: Se föreläsningsanteckningar.

Definition 2.24. *Vektorrummen V och W kallas isomorfa om det finns en inverterbar linjär avbildning $T \in \mathcal{L}(V, W)$.*

Ändligtdimensionella vektorrum av olika dimension kan inte vara isomorfa, vi har nämligen att:

Sats 2.25. *Två ändligtdimensionella vektorrum är isomorfa om och endast om de är av samma dimension.*
Bevis: Se föreläsningsanteckningar.

Korollarium 2.26. *Varje ändligtdimensionellt vektorrum V med* $\dim V = n$ *är isomorft med* K^n.

Bevis: Följer direkt ur Sats 2.25.

Exempel 2.27. Vektorrummen $\mathcal{P}_n(K)$ och K^{n+1} är isomorfa.

Om $v_1, \ldots, v_n$ är en bas för V och $w_1, \ldots, w_m$ är en bas för W, så tillordnas varje $T \in \mathcal{L}(V, W)$ en matris $\mathcal{M}(T) \in \mathrm{Mat}(m, n, K)$. Med fixerade baser är då $\mathcal{M}$ en avbildning, $\mathcal{M} : \mathcal{L}(V, W) \mapsto \mathrm{Mat}(m, n, K)$.

Sats 2.28. *Antag att $v_1, \ldots, v_n$ och $w_1, \ldots, w_m$ är baser för V respektive W. Då är $\mathcal{M}$ en inverterbar linjär avbildning från $\mathcal{L}(V, W)$ till* $\mathrm{Mat}(m, n, K)$.

Bevis: Se föreläsningsanteckningar.

Den naturliga basen för $\mathrm{Mat}(m, n, K)$ ges av alla sinsemellan olika $m \times n$ matriser med ett element lika med 1 och de övriga lika med 0. Det finns $m \cdot n$ sådana matriser, så $\dim \mathrm{Mat}(m, n, K) = m \cdot n$.

Sats 2.29. *Om V och W är ändligtdimensionella vektorrum, $V, W \neq \{0\}$, så är även vektorrummet $\mathcal{L}(V, W)$ ändligtdimensionellt och*

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W).$$

Bevis: Sätt $\dim V = n$ och $\dim W = m$. Sats 2.28 ger att $\mathcal{L}(V, W)$ och $\mathrm{Mat}(m, n, K)$ är isomorfa och de har då med stöd av Sats 2.25 samma dimension, $m \cdot n = \dim \mathcal{L}(V, W)$.

Definition 2.30. *En linjär avbildning $T : V \mapsto V$ kallas för en operator. Beteckningen $\mathcal{L}(V)$ står för vektorrummet av alla operatorer på V, med andra ord är $\mathcal{L}(V) = \mathcal{L}(V, V)$.*

På oändligtdimensionella vektorrum räcker inte enbart injektivitet eller enbart surjektivitet för att garantera inverterbarhet för en operator (se Exempel 2.13). På ändligtdimensionella vektorrum är sakförhållandet ett annat:

Sats 2.31. *Antag att V är ett ändligtdimensionellt vektorrum. Om $T \in \mathcal{L}(V)$ så är följande påståenden ekvivalenta:*

(a) T är *inverterbar*;

(b) T är *injektiv*;

(c) T är *surjektiv*.

Bevis: Se föreläsningsanteckningar.

För en linjär avbildning mellan vektorrum av lika dimension ger den kvadratiska matrisen svarande mot avbildningen ett kriterium för inverterbarhet. En kvadratisk matris A är ju som bekant inverterbar om det finns en kvadratisk matris B av samma dimension och $AB = BA = I$, varvid B är entydigt bestämd och betecknas med $B = A^{-1}$.

Sats 2.32. *Antag att $T \in \mathcal{L}(V, W)$, där $\dim V = \dim W = n \geq 1$, och låt $v_1, \ldots, v_n$ och $w_1, \ldots, w_n$ vara baser för V respektive W. Matrisen för T med avseende på de givna baserna är $\mathcal{M}(T)$. Då är T inverterbar om och endast om $\mathcal{M}(T)$ är inverterbar. Matrisen för den inversa avbildningen $T^{-1} \in \mathcal{L}(W, V)$ ges av*

$$\mathcal{M}(T^{-1}) = \mathcal{M}(T)^{-1}.$$

Bevis: Se föreläsningsanteckningar.

Den identiska avbildningen $I \in \mathcal{L}(V, V)$ är inverterbar, ty den är injektiv och surjektiv. Då är följande korollarium en följd av Sats 2.32:

Korollarium 2.33. *Om $u_1, \ldots, u_n$ och $v_1, \ldots, v_n$ är baser för vektorrummet V, så är matrisen för den identiska avbildningen, $\mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$ inverterbar och*

$$\mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))^{-1} = \mathcal{M}(I, (v_1, \ldots, v_n), (u_1, \ldots, u_n)).$$

Med hjälp av ovanstående korollarium kan vi nu utreda sambandet mellan matriserna för en operator T vid basbyte:

Sats 2.34. *Antag att $T \in \mathcal{L}(V)$ samt att $u_1, \ldots, u_n$ och $v_1, \ldots, v_n$ är baser för V. Inför beteckningen $A = \mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$. Då gäller:*

$$\mathcal{M}(T, (u_1, \ldots, u_n), (u_1, \ldots, u_n)) = A^{-1}\, \mathcal{M}(T, (v_1, \ldots, v_n), (v_1, \ldots, v_n))\, A. \tag{2.6}$$

Bevis: Se föreläsningsanteckningar.

Definition 2.35. *Om A och B är $n \times n$ matriser och Q är en inverterbar $n \times n$ matris sådan att*

$$B = Q^{-1} A\, Q, \tag{2.7}$$

så är A och B likartade (similära) och (2.7) är en likhetstransformation.

Sats 2.34 ger därmed vid handen att matriserna för en operator i olika baser är likartade.

3. Tabellerade fakta om komplexa tal och polynom

Detta kapitel är en sammanställning av resultat från (komplexa) analysen. Inga nya begrepp inom linjär algebra presenteras och vi redogör inte för bevisen av satserna. Kapitlet utgör närmast en källa för referens för senare kapitel.

Komplexa tal

Låt i beteckna den imaginära enheten, $i^2 = -1$. Som bekant är då $z = a + b\,i$, där $a, b \in R$, ett komplext tal med realdelen $\operatorname{Re} z = a$ och imaginärdelen $\operatorname{Im} z = b$. För varje $z \in C$ är då $z = \operatorname{Re} z + (\operatorname{Im}\ z)\,i$.

Det komplexa konjugatet av $z = a + b\,i$ betecknas $\overline{z}$ och definieras av

$$\overline{z} = a - b\,i.$$

Absoluta beloppet av $z = a + b\,i$ betecknas $|z|$ och definieras av

$$|z| = \sqrt{a^2 + b^2}.$$

För alla $z, w \in C$ gäller det att:

$$\begin{aligned}
\operatorname{Re}(w + z) &= \operatorname{Re} w + \operatorname{Re} z,\\
\operatorname{Im}(w + z) &= \operatorname{Im} w + \operatorname{Im} z,\\
z + \overline{z} &= 2 \operatorname{Re} z,\\
z - \overline{z} &= 2\,(\operatorname{Im} z)\,i,\\
z\,\overline{z} &= |z|^2,\\
\overline{w + z} &= \overline{w} + \overline{z},\\
\overline{w\,z} &= \overline{w}\,\overline{z},\\
\overline{(\overline{z})} &= z,\\
|w\,z| &= |w|\,|z|.
\end{aligned}$$

Polynom

En avbildning $p : K \mapsto K$ är ett polynom med koefficienter i K om det finns $a_0, \ldots, a_m \in K$ sådana att

$$p(z) = a_0 + a_1 z + \ldots + a_m z^m, \tag{3.1}$$

för alla $z \in K$. Som tidigare definieras $\mathcal{P}(K)$ som mängden av alla polynom med koefficienter i K. Om p kan skrivas i formen (3.1) med $a_m \neq 0$ säger vi att gradtalet för p är m.

Ett tal $\lambda \in K$ kallas för ett nollställe för $p \in \mathcal{P}(K)$ om $p(\lambda) = 0$.

Sats 3.1. *Antag att $p \in \mathcal{P}(K)$ är ett polynom av gradtal $m \geq 1$. Då är $\lambda \in K$ ett nollställe för p om och endast om det finns ett polynom $q \in \mathcal{P}(K)$ av gradtal $m - 1$ sådant att*

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

för alla $z \in K$.

Korollarium 3.2. *(a) Antag att $p \in \mathcal{P}(K)$, $p \not\equiv 0$, är ett polynom av gradtal $m \geq 0$. Då har p högst m nollställen i K.*
(b) Antag att $a_0, \ldots, a_m \in K$. Om

$$a_0 + a_1 z + \ldots + a_m z^m = 0$$

för alla $z \in K$, så är $a_0 = \ldots = a_m = 0$.

Sats 3.3. *Antag att $p, q \in \mathcal{P}(K)$, med $p \not\equiv 0$. Då finns det polynom $s, r \in \mathcal{P}(K)$ sådana att*

$$q(z) = s(z)\, p(z) + r(z)$$

för alla $z \in K$ och gradtalet för r är mindre än gradtalet för p.

Sats 3.4. *Algebrans fundamentalsats. Varje icke-konstant polynom $p \in \mathcal{P}(C)$ har ett nollställe i C.*

Korollarium 3.5. *Om $p \in \mathcal{P}(C)$ är ett icke-konstant polynom, så har p en entydig faktorisering av formen*

$$p(z) = c\,(z - \lambda_1)\ldots(z - \lambda_m),$$

där $c, \lambda_1, \ldots, \lambda_m \in C$.

Sats 3.6. *Antag att* $p \in \mathcal{P}(R)$. *Om* $\lambda \in C$ *är ett nollställe för* p, *så är även* $\overline{\lambda}$ *ett nollställe för* p.

Sats 3.7. *Låt* $\alpha, \beta \in R$. *Då har vi en faktorisering av formen*

$$x^2 + \alpha\, x + \beta = (x - \lambda_1)(x - \lambda_2),$$

med $\lambda_1, \lambda_2 \in R$, *om och endast om* $\alpha^2 \geq 4\,\beta$.

Sats 3.8. *Om* $p \in \mathcal{P}(R)$ *är ett icke-konstant polynom, så har* p *en entydig faktorisering av formen*

$$p(z) = c\,(x - \lambda_1)...(x - \lambda_m)(x^2 + \alpha_1\, x + \beta_1)...(x^2 + \alpha_M\, x + \beta_M),$$

där $c, \lambda_1, ..., \lambda_m \in R$ *och* $(\alpha_1, \beta_1), ..., (\alpha_M, \beta_M) \in R^2$ *med* $\alpha_j^2 < 4\,\beta_j$ *för varje* j.

Bevisen till ovanstående satser kan studeras i t.ex. kursboken av Axler.

4. Egenvärden och egenvektorer för linjära operatorer

I detta kapitel studeras linjära operatorer, det vill säga linjära avbildningar T med värderummet i definitionsrummet, $T \in \mathcal{L}(V,V) = \mathcal{L}(V)$. Fokuseringen är nu på ändligtdimensionella vektorrum, ty en stor del av de resultat som presenteras gäller ej i oändligtdimensionella vektorrum.

Invarianta underrum

Antag att vi har ett ändligtdimensionellt vektorrum V och en operator $T \in \mathcal{L}(V)$. För att utreda vilka egenskaper T har kan det vara fördelaktigt att studera egenskaperna för restriktionen av T till en delmängd av V. Antag att V kan uppdelas i direkta summan

$$V = U_1 \oplus ... \oplus U_n. \tag{4.1}$$

Låt $T|_{U_j}$ beteckna restriktionen av T till underrummet U_j, $j = 1, ..., n$. Om uppdelningen i (4.1) har egenskapen att varje $T|_{U_j}$ är en operator på U_j, $T|_{U_j} \in \mathcal{L}(U_j)$, så är de ändliga produkterna av operatorn $T|_{U_j}$ väldefinierade. Detta är en önskvärd och viktig egenskap.

Definition 4.1. *Antag att $T \in \mathcal{L}(V)$, där V är ett ändligtdimensionellt vektorrum. Om U är ett underrum av V sådant att $u \in U \Rightarrow T(u) \in U$ säger vi att U är invariant under T. Med andra ord är U invariant under T om $R(T|_U) \subseteq U$.*

Exempel 4.2. (a) De triviala underrummen $\{0\}$ och V av vektorrummet V är invarianta under varje $T \in \mathcal{L}(V)$.
(b) Nollrummet $N(T)$ och värderummet $R(T)$ är invarianta under T.
(c) Låt $T \in \mathcal{L}(\mathcal{P}_{10}(R))$ vara derivataoperatorn $Tp = p'$. Då är $\mathcal{P}_5(R)$ ett invariant underrum under T, ty $p \in \mathcal{P}_5(R) \Rightarrow Tp = p' \in \mathcal{P}_5(R)$.

Man kan fråga sig om ett ändligtdimensionellt vektorrum alltid har icke-triviala invarianta underrum under en given operator T. Exempel 4.2 besvarar inte denna fråga, ty om T är inverterbar så sammanfaller $N(T)$ och $R(T)$ med de triviala underrummen $\{0\}$ respektive V. Vi skall senare se att ändligtdimensionella komplexa vektorrum av dimension större än 1 och ändligtdimensionella reella vektorrum av dimension större än 2 alltid har invarianta underrum som är icke-triviala.

Egenvärden och egenvektorer

Vi övergår nu till att undersöka de enklaste möjliga icke-triviala invarianta underrummen, nämligen sådana som har dimensionen 1. Underrum av dimension 1 till vektorrummet V är lätta att beskriva. För ett godtyckligt $u \in V$, sådant att $u \neq 0$, är

$$U = \{a\,u : a \in K\} \tag{4.2}$$

ett underrum av V med $\dim U = 1$. Vidare är varje 1-dimensionellt underrum av V av formen (4.2). Antag att $0 \neq u \in V$ och att underrummet U som ges av (4.2) är invariant under $T \in \mathcal{L}(V)$. Eftersom det då gäller att $T\,u \in U$, så finns det ett $\lambda \in K$ sådant att

$$T\,u = \lambda\,u. \tag{4.3}$$

Omvänt, om $T \in \mathcal{L}(V)$, $0 \neq u \in V$ och det finns ett $\lambda \in K$ sådant att $T\,u = \lambda\,u$, så är underrummet $U = \{a\,u : a \in K\}$ invariant under T, ty för $a \in K$ gäller $T(a\,u) = a\,T\,u = a\,\lambda\,u \in U$.

Definition 4.3. *En skalär $\lambda \in K$ är ett egenvärde till $T \in \mathcal{L}(V)$ om det existerar en vektor $u \neq 0$ i V sådan att $T\,u = \lambda\,u$. Om λ är ett egenvärde till T och $u \in V$ är en vektor sådan att $T\,u = \lambda\,u$, så kallas u en egenvektor till T svarande mot egenvärdet λ.*

Anmärkning 4.4. Med den ovangivna definitionen är $u = 0$ en egenvektor till varje egenvärde λ. Vidare är mängden av egenvektorer svarande mot ett givet egenvärde λ alltid olika mängden $\{0\}$.

Utredningarna kring formel (4.3) ger då att vektorrummet V har ett invariant underrum U av dimension 1 under $T \in \mathcal{L}(V)$ om och endast om T har ett egenvärde $\lambda \in K$.

Sats 4.5. *Egenvektorerna till $T \in \mathcal{L}(V)$ svarande mot egenvärdet $\lambda \in K$ bildar underrummet $N(T - \lambda\,I)$ av V.*

Bevis: Antag att $T \in \mathcal{L}(V)$ och att $\lambda \in K$ är ett egenvärde till T. Då gäller för $u \in V$ att

$$T\,u = \lambda\,u \Leftrightarrow (T - \lambda\,I)\,u = 0 \Leftrightarrow u \in N(T - \lambda\,I).$$

Beaktande definitionen på egenvärden och egenvektorer samt satserna 4.5, 2.9 och 2.31 erhåller vi följande karakteriseringar:

Korollarium 4.6. *Antag att $T \in \mathcal{L}(V)$ och att $\lambda \in K$. Då gäller:*

a) λ är ett egenvärde till T om och endast om $T - \lambda I$ inte är injektiv,

b) λ är ett egenvärde till T om och endast om $T - \lambda I$ inte är surjektiv,

c) λ är ett egenvärde till T om och endast om $T - \lambda I$ inte är inverterbar,

där V är ändligtdimensionell i fallen b) och c).

Exempel 4.7. Exempel på egenvärden och egenvektorer. (Se föreläsningsanteckningar).

Sats 4.8. *Antag att $T \in \mathcal{L}(V)$ och att $\lambda_1, \ldots, \lambda_n \in K$ är olika egenvärden svarande mot T. Låt $v_1, \ldots, v_n$ vara egenvektorer som alla är olika noll och svarar mot de ovangivna egenvärdena. Då är $v_1, \ldots, v_n$ linjärt oberoende vektorer i V.*

Bevis: Se föreläsningsanteckningar.

Korollarium 4.9. *Varje operator $T \in \mathcal{L}(V)$ har högst* $\dim V$ *olika egenvärden.*

Bevis: Låt $T \in \mathcal{L}(V)$ och antag att $\lambda_1, \ldots, \lambda_m \in K$ är olika egenvärden för T med korresponderande egenvektorer $v_1, \ldots, v_m$ alla olika noll. Då är egenvektorerna linjärt oberoende i V (Sats 4.8) och därmed gäller $m \leq \dim V$ (Sats 1.23).

Exempel 4.10. Antag att A är en $n \times n$ matris med elementen $a_{i,j} \in K$. Ett tal $\lambda \in K$ kallas ett egenvärde till A om det finns en $n \times 1$ matris (kolonnvektor) $x \neq 0$ sådan att

$$A\,x = \lambda\,x. \tag{4.4}$$

Om $\lambda \in K$ är ett egenvärde till A kallas alla $n \times 1$ kolonnvektorer x som uppfyller (4.4) för egenvektorer till A svarande mot egenvärdet λ.

Om $\lambda \in K$ är ett egenvärde till A gäller

$$(A - \lambda\,I)\,x = 0, \tag{4.5}$$

där I är $n \times n$ enhetsmatrisen. Med andra ord är λ ett egenvärde till A om och endast om det homogena ekvationssystemet (4.5) har icke-triviala lösningar $x \neq 0$, vilket inträffar då och endast då

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

där $\det(A - \lambda\,I)$ betecknar determinanten av matrisen $A - \lambda\,I$, som är ett polynom $p(\lambda)$ av gradtal n. Polynomet $p(\lambda)$, med koefficienter i K, kallas karakteristiska polynomet till A. Nollställena i K för $p(\lambda)$ utgör mängden av egenvärden till A i K.

Om exempelvis $A = \begin{pmatrix} 7 & 8 \\ 1 & 5 \end{pmatrix}$ har vi det karakteristiska polynomet

$$\begin{aligned} p(\lambda) &= \det(A - \lambda\,I) = \begin{vmatrix} 7-\lambda & 8 \\ 1 & 5-\lambda \end{vmatrix} = (7-\lambda)(5-\lambda) - 8 = \lambda^2 - 12\,\lambda + 27 \\ &= (\lambda - 3)(\lambda - 9), \end{aligned}$$

med nollställena $\lambda_1 = 3$ och $\lambda_2 = 9$, som då utgör egenvärdena till A i R (och C). För $\lambda = 3$ kontrollerar man lätt att $(2,-1)^T$ är en egenvektor, och för $\lambda = 9$ är $(4,1)^T$ en egenvektor.

Om $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, så har A inga reella egenvärden, ty $p(\lambda) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$, varför $\lambda = \pm i$ är egenvärdena till A i C.

(**Exempel.** Överkurs. Numeriska närmevärden för egenvärdena till en $n \times n$ matris A beräknas behändigt i Matlab med kommandot

$$>> \mathrm{eig}(A)$$

Kommandot

$$>> [E, L] = \mathrm{eig}(A)$$

ger egenvärdena till A i diagonalmatrisen L och motsvarande egenvektorer i kolonnerna för E.)

Antag att A och B är likartade $n \times n$ matriser. Då finns det en inverterbar $n \times n$ matris Q så att $B = Q^{-1} A Q$. Om nu $\lambda \in K$ och $x \neq 0$ så gäller det att

$$\begin{aligned} B x = \lambda x &\Leftrightarrow (Q^{-1} A Q) x = \lambda x \\ &\Leftrightarrow A (Q x) = \lambda (Q x) \\ &\Leftrightarrow A y = \lambda y, \quad y = Q x. \end{aligned}$$

Därmed har likartade matriser samma egenvärden. Sats 2.34 ger då vid handen att om $T \in \mathcal{L}(V)$ så är egenvärdena för matrisen för T, $\mathcal{M}(T)$, oberoende av valet av bas för vektorrummet V.

Nu bör vi utreda om egenvärdena för operatorn T sammanfaller med egenvärdena för $\mathcal{M}(T)$. Detta vore önskvärt med tanke på numerisk bestämning av närmevärden till egenvärdena för T.

Sats 4.11. *Antag att $T \in \mathcal{L}(V)$ och att $\mathcal{M}(T)$ är matrisen för T med avseende på någon bas för V. Då har T och $\mathcal{M}(T)$ samma egenvärden.*

Bevis: Se föreläsningsanteckningar.

Operatorpolynom

Om $T \in \mathcal{L}(V)$ är en operator på V så är även produkten TT en operator på V. För heltaliga $m \geq 0$ definierar vi

$$\begin{aligned} &T^0 = I, \ T^1 = T, \\ &T^m = T T^{m-1}, \ m \geq 2. \end{aligned}$$

Vi har då att $T^m \in \mathcal{L}(V)$ för $m \geq 0$. Om $T \in \mathcal{L}(V)$ är inverterbar har vi att $T^{-1} \in \mathcal{L}(V)$. Vi definierar då för $m > 0$ operatorn $T^{-m} \in \mathcal{L}(V)$ genom

$$T^{-m} = (T^{-1})^m, \ m > 0.$$

Med hjälp av ovanstående definitioner och induktion verifieras potenslagarna

$$T^m T^n = T^{m+n} \text{ och } (T^m)^n = T^{m\,n},$$

för $T \in \mathcal{L}(V)$, där m, n är godtyckliga heltal om T är inverterbar och $m, n \geq 0$ om T saknar invers.

Antag att $T \in \mathcal{L}(V)$ och att $p \in \mathcal{P}(K)$ är polynomet givet av

$$p(z) = a_0 + a_1 z + \ldots + a_m z^m,$$

för alla $z \in K$. Då definieras operatorpolynomet $p(T) \in \mathcal{L}(V)$ genom

$$p(T) = a_0 I + a_1 T + \ldots + a_m T^m.$$

För fixt $T \in \mathcal{L}(V)$ är avbildningen från $\mathcal{P}(K)$ till $\mathcal{L}(V)$ given av $p \mapsto p(T)$ för alla $p \in \mathcal{P}(K)$ linjär,

$$(a\,p + b\,q)(T) = a\,p(T) + b\,q(T),$$

för $p, q \in \mathcal{P}(K)$ och $a, b \in K$, (Hemuppgift).

Om $p, q \in \mathcal{P}(K)$ så definieras polynomet $pq \in \mathcal{P}(K)$ genom

$$(pq)(z) = p(z)\,q(z),$$

för varje $z \in K$. Om $T \in \mathcal{L}(V)$ så gäller

$$(pq)(T) = p(T)\,q(T),$$

för alla $p, q \in \mathcal{P}(K)$, (Hemuppgift). Vidare kommuterar operatorpolynom, ty

$$p(T)\,q(T) = (pq)(T) = (qp)(T) = q(T)\,p(T),$$

för alla $p, q \in \mathcal{P}(K)$.

Uppåt triangulära matriser

Om K är skalärkroppen C och $V \neq \{0\}$ är ett ändligtdimensionellt vektorrum så har vi följande viktiga resultat:

Sats 4.12. *Låt $T \in \mathcal{L}(V)$, där $V \neq \{0\}$ är ett ändligtdimensionellt komplext vektorrum. Då har operatorn T ett egenvärde.*

Bevis: Se föreläsningsanteckningar.

Anmärkning 4.13. Föregående sats gäller ej om V är ett komplext vektorrum av oändlig dimension. Exempelvis om $T \in \mathcal{L}(C^\infty)$ definieras genom

$$T(z_1, z_2, \ldots) = (0, z_1, z_2, \ldots),$$

för varje $z = (z_1, z_2, \ldots) \in C^\infty$, så har T inga egenvärden i C. (Hemuppgift).

Betrakta den kvadratiska $n \times n$ matrisen

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

vars **diagonal** består av elementen $a_{1,1}, a_{2,2}, ..., a_{n,n}$. En matris kallas **uppåt triangulär** om alla element nedanför diagonalen är 0. Speciellt är varje 1×1 matris uppåt triangulär.

En uppåt triangulär matris för vilken vi inte känner till, eller inte vill ange, elementen ovanför diagonalen betecknas med

$$\begin{pmatrix} a_{1,1} & & * \\ & \ddots & \\ 0 & & a_{n,n} \end{pmatrix}.$$

Ett viktigt mål i linjär algebra är att utreda, givet en operator $T \in \mathcal{L}(V)$, om det existerar en bas för V med avseende på vilken matrisen för T, $\mathcal{M}(T)$, är av enkel form. Man eftersträvar vanligen en framställning $\mathcal{M}(T)$ med många element lika med 0. I detta avseende kan en uppåt triangulär matrisframställning $\mathcal{M}(T)$ anses vara tämligen enkel. Följande sats påvisar ett samband mellan uppåt triangulära matriser och invarianta underrum.

Sats 4.14. *Antag att $T \in \mathcal{L}(V)$ och att $v_1, ..., v_n$ är en bas för V. Då är följande påståenden ekvivalenta:*

(*a*) *Matrisframställningen $\mathcal{M}(T)$ med avseende på $v_1, ..., v_n$ är uppåt triangulär;*

(*b*) *$T v_k \in [v_1, ..., v_k]$ för varje $k = 1, ..., n$;*

(*c*) *$[v_1, ..., v_k]$ är invariant under T för varje $k = 1, ..., n$.*

Bevis: Se föreläsningsanteckningar.

Om V är ett komplext ändligtdimensionellt vektorrum kan vi garantera att en operator $\mathcal{L}(V)$ har en uppåt triangulär matrisframställning $\mathcal{M}(T)$ med avseende på någon bas för V:

Sats 4.15. *Antag att V är ett ändligtdimensionellt komplext vektorrum och att $T \in \mathcal{L}(V)$. Då finns det en bas $v_1, ..., v_n$ för V så att $\mathcal{M}(T)$ är uppåt triangulär.*

Bevis: Se föreläsningsanteckningar.

Vi har tidigare sett (Sats 2.32) att $T \in \mathcal{L}(V)$ är inverterbar om och endast om $\mathcal{M}(T)$ är inverterbar. Speciellt lätt är detta att kontrollera om $\mathcal{M}(T)$ är uppåt triangulär med avseende på någon bas för V:

Sats 4.16. *Antag att $\mathcal{M}(T)$ för $T \in \mathcal{L}(V)$ är uppåt triangulär med avseende på någon bas för V. Då är T inverterbar om och endast om alla diagonalelement i $\mathcal{M}(T)$ är olika noll.*

Bevis: Se föreläsningsanteckningar.

I allmänhet kan man inte exakt bestämma egenvärdena för en operator $T \in \mathcal{L}(V)$ om $\dim V > 4$. Ett viktigt undantag är dock fallet då $\mathcal{M}(T)$ är uppåt triangulär:

Sats 4.17. *Antag att $\mathcal{M}(T)$ för $T \in \mathcal{L}(V)$ är uppåt triangulär med avseende på någon bas för V. Då ges alla egenvärden till T av diagonalelementen för $\mathcal{M}(T)$.*

Bevis: Se föreläsningsanteckningar.

De **diagonala matriserna** har nollelement både nedanför och ovanför diagonalen, och är därmed enkla specialfall av uppåt triangulära matriser. Vi betecknar en diagonalmatris med

$$\begin{pmatrix} a_{1,1} & & 0 \\ & \ddots & \\ 0 & & a_{n,n} \end{pmatrix},$$

där diagonalen $a_{1,1}, \ldots, a_{n,n} \in K$. För en diagonal matrisframställning av en operator gäller:

Sats 4.18. *Antag att $T \in \mathcal{L}(V)$ och att $v_1, \ldots, v_n$ är en bas för V. Då är $\mathcal{M}(T)$ en diagonalmatris om och endast om $v_1, \ldots, v_n$ är egenvektorer till T.*

Bevis: Antag att $T \in \mathcal{L}(V)$ och att $v_1, \ldots, v_n$ är en bas för V. Då gäller:

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \Leftrightarrow T v_1 = \lambda_1 v_1, \ldots, T v_n = \lambda_n v_n$$

$$\Leftrightarrow v_1, \ldots, v_n \text{ är egenvektorer till } T.$$

En operator T på ett ändligtdimensionellt komplext vektorrum har minst ett egenvärde, men det behöver inte finnas någon bas för V sådan att $\mathcal{M}(T)$ svarande mot basen är diagonal, vilket följande exempel belyser:

Exempel 4.19. (Se föreläsningsanteckningarna).

Om en operator har tillräckligt många olika egenvärden på ett ändligtdimensionellt vektorrum V, så är $\mathcal{M}(T)$ diagonal med avseende på någon bas för V:

Sats 4.20. *Antag att $T \in \mathcal{L}(V)$ har $\dim V = n$ olika egenvärden. Då har V en bas $v_1, \ldots, v_n$ sådan att $\mathcal{M}(T)$ med avseende på basen är diagonal.*

Bevis: Antag att $T \in \mathcal{L}(V)$ har n olika egenvärden $\lambda_1, ..., \lambda_n$. Låt $v_1, ..., v_n$ vara egenvektorer olika noll svarande mot egenvärdena. Då är $v_1, ..., v_n$ linjärt oberoende (Sats 4.8). Men då är $v_1, ..., v_n$ en bas för V (Sats 1.37). Med avseende på denna bas är $\mathcal{M}(T)$ diagonal (Sats 4.18).

Anmärkning 4.21. Sats 4.20 kan inte omvändas. Låt exempelvis $T \in \mathcal{L}(K^3)$ vara given av

$$T(z_1, z_2, z_3) = (2\,z_1, 3\,z_2, 3\,z_3).$$

Då är $\mathcal{M}(T)$ med avseende på den naturliga basen för K^3 en diagonalmatris med diagonalen $a_{1,1} = 2, a_{2,2} = a_{3,3} = 3$. Men T har enbart $2(\neq 3 = \dim K^3)$ olika egenvärden, $\lambda_1 = 2, \lambda_2 = 3$.

Vi avslutar detta avsnitt med en sats som ger flera ekvivalenta villkor för att en operator på ett ändligtdimensionellt vektorrum V skall ha en diagonal matrisframställning $\mathcal{M}(T)$ med avseende på någon bas för V.

Sats 4.22. *Antag att V är ett ändligtdimensionellt vektorrum och att $T \in \mathcal{L}(V)$. Låt $\lambda_1, ..., \lambda_m$ vara olika egenvärden för T. Då är följande påståenden ekvivalenta:*

a) $\mathcal{M}(T)$ *är diagonal med avseende på någon bas för V;*

b) V har en bas bestående av egenvektorer till T;

c) det finns endimensionella underrum $U_1, ..., U_n$ av V, alla invarianta under T, sådana att $V = U_1 \oplus ... \oplus U_n$;

d) $V = N(T - \lambda_1 I) \oplus ... \oplus N(T - \lambda_m I)$;

e) $\dim V = \dim N(T - \lambda_1 I) + ... + \dim N(T - \lambda_m I)$.

Bevis: Se föreläsningsanteckningar.

Invarianta underrum på reella vektorrum

Vi har visat att varje operator på ett komplext ändligtdimensionellt vektorrum har minst ett egenvärde. Detta gäller ej för operatorer på ändligtdimensionella reella vektorrum, (se Exempel 4.7). Därmed kan en operator på ett reellt vektorrum sakna invarianta underrum av dimension 1, men då måste det finnas ett invariant underrum av dimension 2:

Sats 4.23. *Varje operator T på ett ändligtdimensionellt reellt vektorrum $V \neq \{0\}$ har ett invariant underrum av dimensionen 1 eller 2.*

Bevis: Se föreläsningsanteckningar.

Definition 4.24. *Antag att U och W är underrum av vektorrummet V och att*

$$V = U \oplus W.$$

Varje vektor $v \in V$ har då en entydig framställning:

$$v = u + w, \quad u \in U, w \in W.$$

Då betecknar $P_{U,W} \in \mathcal{L}(V)$ projektionen på U med nollrummet W, och definieras genom

$$P_{U,W}(v) = u.$$

Om $v_1 = u_1 + w_1$, $v_2 = u_2 + w_2$, $u_j \in U, w_j \in W$, så gäller det för $a, b \in K$ att $P_{U,W}(a\,v_1 + b\,v_2) = a\,u_1 + b\,u_2 = a\,P_{U,W}(v_1) + b\,P_{U,W}(v_2)$, så $P_{U,W}$ är linjär.

Om $v = u + w$ och $u \neq 0$ så är $P_{U,W}(v) \neq 0$. Om $u = 0$ så har vi att $v \in W$ och att $P_{U,W}(v) = 0$. Därmed är $N(P_{U,W}) = W$. Klart är att $R(P_{U,W}) = U$.

Eftersom $P_{W,U}(v) = w$ har vi att:

$$v = P_{U,W}(v) + P_{W,U}(v),$$

för alla $v \in V$.

Vidare gäller $P_{U,W}^2 = P_{U,W}$, ty $P_{U,W}^2(v) = P_{U,W}(P_{U,W}(v)) = P_{U,W}(u) = u = P_{U,W}(v)$.

Med hjälp av den ovandefinierade projektionen kan vi bevisa följande resultat för reella vektorrum av udda dimension:

Sats 4.25. *Varje operator T på ett reellt vektorrum V av udda dimension har ett egenvärde.*

Bevis: Se föreläsningsanteckningar.

5. Vektorrum med skalär produkt

Då vi definierade abstrakta vektorrum generaliserde vi den bekanta linjära strukturen (addition och skalär multiplikation) från R^2 och R^3. Nu tar vi fasta på de i R^2 och R^3 bekanta egenskaperna längd av vektorer och vinkeln mellan vektorer. Dessa begrepp kan generaliseras till abstrakta vektorrum genom införandet av en skalär produkt. Vi antar i detta kapitel att $V \neq \{0\}$ är ett ändligtdimensionellt vektorrum över talkroppen K.

Skalära produkter

Längden av en vektor x i R^2 eller R^3 kallas **normen** av x och betecknas $\|x\|$. Som bekant generaliseras begreppet längd av vektor till R^n genom

$$\|x\| = \sqrt{x_1^2 + \ldots + x_n^2},$$

där $x = (x_1, \ldots, x_n) \in R^n$. Normen är inte en linjär avbildning från R^n till R.

Linjära egenskaper kan införas med hjälp av **skalära produkten** $x \cdot y$ av vektorerna $x, y \in R^n$, vilken definieras genom

$$x \cdot y = x_1 y_1 + \ldots + x_n y_n,$$

för varje $x, y \in R^n$. Skalära produkten är därmed en avbildning från $R^n \times R^n$ till R. Tydligen är $0 \leq x \cdot x = \|x\|^2$, med skalära produkten lika med noll om och endast om $x = 0$. Om vi fixerar $y \in R^n$ så är avbildningen $T(x) = x \cdot y$ linjär, $T \in \mathcal{L}(R^n, R)$. Vidare kommuterar skalära produkten, $x \cdot y = y \cdot x$ för alla $x, y \in R^n$.

Ovanstående egenskaper kan tas till grund för en generalisering av skalära produkten till ett reellt vektorrum. Vi vill dock att den definierade skalära produkten skall vara meningsfylld även på abstrakta komplexa vektorrum, så vi bör studera begreppet norm på C^n.

För komplexa tal $z = a + b\,i,\ a, b \in R$, definieras **belopp** och **konjugat** av

$$|z| = \sqrt{a^2 + b^2},$$
$$\overline{z} = a - b\,i,$$

och ett samband mellan begreppen ges av

$$|z|^2 = z\,\overline{z}.$$

För $z = (z_1, \ldots, z_n) \in C^n$ definieras normen genom

$$\|z\| = \sqrt{|z_1|^2 + \ldots + |z_n|^2},$$

varmed $\|z\| \geq 0$ och

$$\|z\|^2 = z_1 \overline{z_1} + \ldots + z_n \overline{z_n}. \tag{5.1}$$

Om vi tolkar (5.1) som skalära produkten av z och z borde vi då definiera skalära produkten av w och z i C^n genom

$$w_1 \overline{z_1} + \ldots + w_n \overline{z_n}.$$

Om w och z byter roller borde vi då ha att skalära produkten av z och w är lika med det komplexa konjugatet av skalära produkten av w och z. Vi är nu redo att definiera skalära produkten på ett vektorrum V över K.

Definition 5.1. *En skalär produkt på V är en funktion som avbildar varje ordnat par $(u, v) \in V \times V$ på ett tal $\langle u, v\rangle \in K$ och som har följande egenskaper för varje $u, v, w \in V$ och varje $a \in K$:*

(i) $\langle v, v\rangle \geq 0$,

(ii) $\langle v, v\rangle = 0$, *om och endast om* $v = 0$,

(iii) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$,

(iv) $\langle a\, u, v\rangle = a\, \langle u, v\rangle$,

(v) $\langle u, v\rangle = \overline{\langle v, u\rangle}$.

Ett vektorrum V försett med en skalär produkt kallas ett euklidiskt rum om $K = R$ och ett unitärt rum om $K = C$.

Anmärkning 5.2. Om V är ett euklidiskt rum, så verifierar man lätt att följande två egenskaper gäller: $\langle u, v\rangle = \langle v, u\rangle$ och $\langle u, a\, v\rangle = a\, \langle u, v\rangle$.

Vidare gäller i ett vektorrum V med skalär produkt att $\langle 0, u\rangle = \langle u, 0\rangle = 0$, för alla $u \in V$, $\langle u, a\, v\rangle = \overline{a}\, \langle u, v\rangle$, och $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$.

Exempel 5.3. (a) Det viktigaste exemplet på ett vektorrum med skalär produkt är K^n med den euklidiska skalära produkten

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\rangle = w_1 \overline{z_1} + \ldots + w_n \overline{z_n}.$$

K^n antas vara försedd med ovanstående skalärprodukt om inget annat anges.

(b) Vi kan förse K^n med andra skalära produkter. Om exempelvis $c_1, \ldots, c_n \in R$ är positiva tal så definierar

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\rangle = c_1\, w_1 \overline{z_1} + \ldots + c_n\, w_n \overline{z_n}$$

en skalär produkt på K^n.

(c) På $\mathcal{P}_n(K)$ kan vi definiera en skalär produkt genom

$$\langle p, q\rangle = \int_0^1 p(z)\, \overline{q(z)}\; dz.$$

Normer

Definition 5.4. *Låt V vara ett vektorrum med skalär produkt. För $v \in V$ betecknas normen av v med $\|v\|$ och definitionen ges av*

$$\|v\| = \sqrt{\langle v, v\rangle}\,.$$

Anmärkning 5.5. Vi noterar att $\|v\| = 0$ om och endast om $v = 0$, (ty $\langle v, v\rangle = 0 \Leftrightarrow v = 0$). Vidare gäller för varje $a \in K$ och varje $v \in V$ att $\|a\,v\| = |a|\,\|v\|$, (ty $\|a\,v\|^2 = \langle a\,v, a\,v\rangle = a\,\langle v, a\,v\rangle = a\,\overline{a}\,\langle v, v\rangle = |a|^2\,\|v\|^2$).

Exempel 5.6. (a) Om $V = K^n$ med den euklidiska skalära produkten får vi för $(z_1, ..., z_n) \in K^n$ att

$$\|(z_1, ..., z_n)\| = \sqrt{|z_1|^2 + ... + |z_n|^2}\,.$$

(b) Om $V = \mathcal{P}_n(K)$ försedd med skalärprodukten i Exempel 5.3 (c) erhåller vi för $p \in \mathcal{P}_n(K)$ normen

$$\|p\| = \sqrt{\int_0^1 |p(z)|^2\, dz}\,.$$

Definition 5.7. *Vektorerna $u, v \in V$ säges vara ortogonala om $\langle u, v\rangle = 0$.*

Vi noterar att $\langle u, v\rangle = 0$ om och endast om $\langle v, u\rangle = 0$, och att 0 är ortogonal mot varje vektor i V. Vidare är 0 den enda vektor som är ortogonal mot sig själv.

För $V = R^2$ är följande sats bekant från högstadiet:

Sats 5.8. Pythagoras sats: *Antag att u, v är ortogonala vektorer i V. Då gäller:*

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

Bevis: Antag att $u, v \in V$ är ortogonala. Då gäller:

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

Lemma 5.9. *Antag att $u, v \in V$ och att $v \neq 0$. Då finns det en vektor $w \in V$ och en skalär $a \in K$ så att v och w är ortogonala och*

$$u = av + w,$$

med

$$a = \frac{\langle u, v \rangle}{\|v\|^2} \quad \text{och} \quad w = u - av.$$

Bevis: Se föreläsningsanteckningar.

Följande sats, (som 1821 bevisades för skalära produkten i Exempel 5.3(a) av den franska matematikern Augustin-Louis Cauchy, och 1886 för skalära produkten i Exempel 5.3 (c) av den tyska matematikern Herman Schwarz), introducerar en mycket viktig olikhet i matematiken:

Sats 5.10. Cauchy-Schwarz olikhet: *Antag att $u, v \in V$. Då gäller*

$$|\langle u, v \rangle| \leq \|u\| \, \|v\|, \tag{5.2}$$

där vi har likhet om och endast om den ena av u och v är en skalär gånger den andra.

Bevis: Se föreläsningsanteckningar.

Längden av en sida i en triangel är alltid mindre än summan av längderna av de två övriga sidorna. Generaliseringen till ett abstrakt vektorrum med skalär produkt och norm ges av:

Sats 5.11. Triangelolikheten: *Antag att $u, v \in V$. Då gäller*

$$\|u + v\| \leq \|u\| + \|v\|, \tag{5.3}$$

där vi har likhet om och endast om den ena av u och v är en icke-negativ multipel av den andra.

Bevis: Se föreläsningsanteckningar.

I en parallellogram är summan av kvadraterna på längderna av diagonalerna lika med summan av kvadraterna på längderna av de fyra sidorna:

Sats 5.12. Parallellogramidentiteten: *Antag att $u, v \in V$. Då gäller*

$$\|u + v\|^2 + \|u - v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right). \tag{5.4}$$

Bevis: Hemuppgift.

Ortonormerade baser

Definition 5.13. *Vektorerna $e_1, ..., e_n \in V$ säges vara ortonormerade om de har normen 1 och är parvis ortogonala:*

$$\langle e_j, e_k \rangle = \begin{cases} 0, & \text{om } j \neq k; \\ 1, & \text{om } j = k, \end{cases}$$

där $j, k \in \{1, ..., n\}$.

Exempel 5.14. Den naturliga basen i K^n består av ortonormerade vektorer med avseende på den euklidiska skalära produkten.

Om en vektor i V uttrycks som en linjärkombination av ortonormerade vektorer är det lätt att bestämma normen av vektorn:

Sats 5.15. *Om vektorerna $e_1, ..., e_n \in V$ är ortonormerade så är*

$$\|a_1\, e_1 + ... + a_n\, e_n\|^2 = |a_1|^2 + ... + |a_n|^2, \tag{5.5}$$

för alla $a_1, ..., a_n \in K$.

Bevis: Se föreläsningsanteckningar.

Med hjälp av föregående sats bevisar vi enkelt att:

Korollarium 5.16. *Varje mängd av ortonormerade vektorer i V består av linjärt oberoende vektorer.*

Bevis: Antag att $e_1, ..., e_n \in V$ är ortonormerade vektorer och att $a_1, ..., a_n \in K$ är sådana att

$$a_1\, e_1 + ... + a_n\, e_n = 0\,.$$

Då ger Sats 5.15 att $|a_1|^2 + ... + |a_n|^2 = 0$, vilket betyder att $a_1 = ... = a_n = 0$, och beviset är klart.

Definition 5.17. *Vektorerna $e_1, ..., e_n \in V$ är en ortonormerad bas för V om vektorerna är ortonormerade och utgör en bas för V.*

Exempel 5.18. (a) Den naturliga basen i K^n är en ortonormerad bas med avseende på den euklidiska skalära produkten.

(b) Om $e_1, ..., e_n$ är ortonormerade vektorer i V, $\dim V = n$, så utgör vektorerna en bas till V, ty de är linjärt oberoende (Koroll. 5.16) och antalet vektorer sammanfaller med dimensionen för V (Sats 1.37).

I allmänhet kan det vara besvärligt att bestämma koordinaterna för basframställningen av en vektor $v \in V$. Om basen för V är ortonormerad förenklas detta problem avsevärt, vilket är en viktig motivering till användandet av ortonormerade baser:

Sats 5.19. *Antag att vektorerna $e_1, ..., e_n$ är en ortonormerad bas för V. Då är*

$$v = \langle v, e_1 \rangle e_1 + ... + \langle v, e_n \rangle e_n \tag{5.6}$$

och

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + ... + |\langle v, e_n \rangle|^2, \tag{5.7}$$

för alla $v \in V$.

Bevis: Låt $v \in V$ och antag att $e_1, ..., e_n$ är en ortonormerad bas för V. Då finns det skalärer $a_1, ..., a_n \in K$ så att

$$v = a_1 e_1 + ... + a_n e_n .$$

Genom att bilda skalära produkten med e_j i båda leden av ovanstående ekvation erhåller vi för $j = 1, ..., n$ att

$$\langle v, e_j \rangle = \langle a_1 e_1 + ... + a_n e_n, e_j \rangle = \sum_{i=1}^{n} \langle a_i e_i, e_j \rangle = \sum_{i=1}^{n} a_i \langle e_i, e_j \rangle = a_j \langle e_j, e_j \rangle = a_j,$$

så (5.6) gäller. Att (5.7) gäller följer då ur (5.5), och beviset är klart.

Då föregående resultat ger en god orsak till att använda ortonormerade baser, är vi ställda inför problemet att utreda om en sådan bas alltid existerar och hur en ortonormerad bas kan konstrueras. Följande sats och korollarium besvarar frågorna. I beviset av nästa resultat använder vi en algoritm uppkallad efter den danska matematikern Jörgen Gram (1850-1916) och den tyska matematikern Erhard Schmidt (1876-1959).

Sats 5.20. Gram-Schmidt: *Antag att vektorerna $v_1, ..., v_n$ är linjärt oberoende i V. Då finns det ortonormerade vektorer $e_1, ..., e_n \in V$ sådana att*

$$[v_1, ..., v_j] = [e_1, ..., e_j]$$

för $j = 1, ..., n$.

Bevis: Se föreläsningsanteckningar.

Nu kan vi besvara frågan om existensen av ortonormerade baser:

Korollarium 5.21. *Varje ändligtdimensionellt vektorrum med en skalär produkt har en ortonormerad bas.*

Bevis: Låt $v_1, ..., v_n$ vara en bas för V. Då är basvektorerna linjärt oberoende i V och vi kan därmed hitta ortonormerade vektorer $e_1, ..., e_n \in V$ sådana att $[e_1, ..., e_n] = [v_1, ..., v_n]$, (Sats 5.20). Men eftersom $[v_1, ..., v_n] = V$ och $e_1, ..., e_n$ är linjärt oberoende (Koroll. 5.16), så är vektorerna $e_1, ..., e_n$ en bas för V.

Nästa fråga vi bör ställa oss är, givet en mängd av ortonormerade vektorer i V, kan vi alltid utvidga mängden till en ortonormerad bas för V?

Korollarium 5.22. *Varje mängd av ortonormerade vektorer i ett ändligtdimensionellt vektorrum V kan utvidgas till en ortonormerad bas för V.*

Bevis: Antag att $e_1, ..., e_m$ är ortonormerade vektorer i V. Då är vektorerna linjärt oberoende i V (Koroll. 5.16) och vi kan därmed utvidga mängden till en bas $e_1, ..., e_m, v_1, ..., v_n$ för V (Sats 1.30). Genom att tillämpa Gram-Schmidts ortonormeringsprocedur, (se beviset av Sats 5.20), på basvektorerna erhåller vi vektorerna

$$e_1, ..., e_m, f_1, ..., f_n\,, \qquad (*)$$

som är ortonormerade. Gram-Schmidts algoritm lämnar $e_1, ..., e_m$ oförändrade, eftersom de redan är ortonormerade (kolla att detta gäller!). Eftersom vektorerna i $(*)$ är ortonormerade är de linjärt oberonde (Koroll. 5.16) och vidare är $[e_1, ..., e_m, f_1, ..., f_n] = [e_1, ..., e_m, v_1, ..., v_n] = V$ (Sats 5.20), så $e_1, ..., e_m, f_1, ..., f_n$ är en ortonormerad bas för V, och därmed är korollariet bevisat.

I föregående kapitel visade vi att för varje operator T på ett ändligtdimensionellt komplext vektorrum V kan vi hitta en bas för V så att matrisframställningen $\mathcal{M}(T)$ av operatorn med avseende på denna bas är uppåt triangulär, (Sats 4.15). Följande korollarium gäller på både reella och komplexa ändligtdimensionella vektorrum med skalär produkt:

Korollarium 5.23. *Antag att $T \in \mathcal{L}(V)$. Om $\mathcal{M}(T)$ är uppåt triangulär med avseende på någon bas för V, så är $\mathcal{M}(T)$ uppåt triangulär med avseende på någon ortonormerad bas för V.*

Bevis: Antag att $\mathcal{M}(T)$ är uppåt triangulär med avseende på basen $v_1, ..., v_n$ för V. Då är $[v_1, ..., v_j]$ invariant under T för varje $j = 1, ..., n$, (Sats 4.14). Vi tillämpar Gram-Schmidts ortogonaliseringsförfarande på basen och erhåller en ortonormerad bas $e_1, ..., e_n$ för V. Eftersom $[e_1, ..., e_j] = [v_1, ..., v_j]$, $j = 1, ..., n$ (Sats 5.20), så är $[e_1, ..., e_j]$ invariant under T för $j = 1, .., n$, och då ger Sats 4.14 att $\mathcal{M}(T)$ med avseende på den ortonormerade basen $e_1, ..., e_n$ är uppåt triangulär.

Nästa resultat, (vars första bevis gavs av den tyska matematikern Issai Schur år 1909), följer lätt ur föregående korollarium och Sats 4.15.

Korollarium 5.24. *Antag att $T \in \mathcal{L}(V)$, där V är ett ändligtdimensionellt komplext vektorrum. Då är $\mathcal{M}(T)$ uppåt triangulär med avseende på någon ortonormerad bas för V.*

Bevis: Sats 4.15 ger att $\mathcal{M}(T)$ är uppåt triangulär med avseende på någon bas för V. Då ger Korollarium 5.23 att det finns en ortonormerad bas för V sådan att $\mathcal{M}(T)$ med avseende på denna bas är uppåt triangulär.

Ortogonala projektioner

Definition 5.25. *Antag att U är en delmängd av vektorrummet V. Då definieras ortogonala komplementet av U som mängden av alla vektorer i V som är ortogonala mot varje vektor i U,*

$$U^{\perp} = \{v \in V : \langle v, u\rangle = 0 \text{ för alla } u \in U\}\,.$$

Anmärkning 5.26. Det gäller alltid att $U^{\perp}$ är en icke-tom mängd, ty $0 \in U^{\perp}$. Antag att $v, w \in U^{\perp}$ och att $a \in K$. Då gäller för varje $u \in U$: $\langle v + w, u\rangle = \langle v, u\rangle + \langle w, u\rangle = 0$ och $\langle a v, u\rangle = a \langle v, u\rangle = 0$, så $U^{\perp}$ är ett underrum av V. Vidare gäller det trivialt att $V^{\perp} = \{0\}$ och $\{0\}^{\perp} = V$.

Antag att $U \subseteq W$, där U, W är delmängder av V. Om $w \in W^{\perp}$ så gäller för varje $v \in W$ att $\langle w, v\rangle = 0$. Eftersom U är en delmängd av W gäller då för alla $u \in U$ att $\langle w, u\rangle = 0$, så $w \in U^{\perp}$. Alltså $U \subseteq W \Rightarrow W^{\perp} \subseteq U^{\perp}$.

Varje ändligtdimensionellt underrum av ett vektorrum V med skalär produkt ger upphov till en naturlig framställning av V som en direkt summa:

Sats 5.27. *Om U är ett ändligtdimensionellt underrum av ett vektorrum V, så är*

$$V = U \oplus U^{\perp}\,.$$

Bevis: Se föreläsningsanteckningar.

Anmärkning 5.28. Ovanstående sats gäller även om V är oändligtdimensionell.

En viktig konsekvens av Sats 5.27 är

Korollarium 5.29. *Om U är ett ändligtdimensionellt underrum av ett vektorrum V, så är*

$$U = (U^{\perp})^{\perp}.$$

Bevis: 1) Tag $u \in U$. Då är $\langle u, v\rangle = 0$ för alla $v \in U^{\perp}$, så $u \in (U^{\perp})^{\perp}$ och $U \subseteq (U^{\perp})^{\perp}$.
2) Välj godtyckligt $v \in (U^{\perp})^{\perp}$. Med stöd av Sats 5.27 gäller

$$v = u + w, \quad \text{där } u \in U, \ w \in U^{\perp}.$$

Då gäller: $0 = \langle v, w\rangle = \langle u + w, w\rangle = \langle u, w\rangle + \langle w, w\rangle = \|w\|^2$. Alltså är $w = 0$ och $v = u \in U$, så $(U^{\perp})^{\perp} \subseteq U$, och beviset är klart.

Definition 5.30. *Antag att U är ett ändligtdimensionellt underrum av vektorrummet V. Då ger Sats 5.27 att varje vektor $v \in V$ på ett entydigt sätt kan framställas som*

$$v = u + w,$$

där $u \in U$ och $w \in U^{\perp}$. Vi definierar nu operatorn $P_U \in \mathcal{L}(V)$, kallad ortogonala projektionen av V på U, genom

$$P_U(v) = u,$$

där v och u ges av framställningen ovan.

Anmärkning 5.31. Med stöd av Definition 4.24 har vi då att $P_U = P_{U,U^{\perp}}$. Då ger utredningarna på sida 34 att P_U är linjär, $P_U \in \mathcal{L}(V)$. Vidare har P_U följande egenskaper:

(*i*) $R(P_U) = U$;
(*ii*) $N(P_U) = U^{\perp}$;
(*iii*) $v - P_U v \in U^{\perp}$ *för varje* $v \in V$;
(*iv*) $P_U^2 = P_U$;
(*v*) $\|P_U v\| \leq \|v\|$ *för varje* $v \in V$.

Egenskaperna $(i), (ii)$ och (iv) har vi visat på sida 34.
Om $v = u + w$, med $u \in U, w \in U^{\perp}$, så gäller: $v - P_U v = (u + w) - u = w \in U^{\perp}$, så (iii) gäller.
Uppdelningen $v = u + w$, $u \in U, w \in U^{\perp}$ och $P_U v = u$, ger att $P_U v$ är ortogonal mot w. Då ger Sats 5.8 att $\|v\|^2 = \|P_U v + w\|^2 = \|P_U v\|^2 + \|w\|^2 \geq \|P_U v\|^2$, så (v) gäller.

Sats 5.32. *Om U är ett ändligtdimensionellt underrum av ett vektorrum V och $e_1, \ldots, e_n$ är en ortonormerad bas för U, så gäller för varje $v \in V$ att*

$$P_U v = \langle v, e_1\rangle e_1 + \ldots + \langle v, e_n\rangle e_n.$$

Bevis: Tag $v \in V$. Med den givna ortonormerade basen för U har vi framställningen

$$v = (\langle v, e_1 \rangle e_1 + \ldots + \langle v, e_n \rangle e_n) + (v - \langle v, e_1 \rangle e_1 - \ldots - \langle v, e_n \rangle e_n) .$$

Den första parentesen är en vektor i U. I beviset av Sats 5.27 visade vi att den andra parentesen är en vektor i $U^{\perp}$, så $P_U\, v$ är tydligen lika med den första parentesen, och satsen är bevisad.

Ett minimeringsproblem: Givet ett ändligtdimensionellt underrum U av ett vektorrum V och en vektor $v \in V$, önskar vi finna en vektor $u \in U$ som minimerar avståndet $\|v - u\|$.

Följande sats visar att ovanstående minimeringsproblem alltid har en lösning i U som är entydigt bestämd och given av $P_U\, v$.

Sats 5.33. *Antag att U är ett ändligtdimensionellt underrum av ett vektorrum V och att $v \in V$. Då gäller*

$$\|v - P_U\, v\| \leq \|v - u\|$$

för alla $u \in U$. Likhet gäller om och endast om $u = P_U\, v$.

Bevis: Se föreläsningsanteckningar.

Exempel 5.34. Vi skall lösa ett minimeringsproblem av ovanbeskrivet slag genom att utnyttja satserna 5.32 och 5.33. Antag att vi söker ett polynom u med reella koefficienter och av gradtal högst 5, som på intervallet $[-\pi, \pi]$ approximerar funktionen $\sin x$ så gott som möjligt då vi mäter felet med

$$\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \, dx .$$

För att lösa problemet låter vi $V = C[-\pi, \pi]$, det reella vektorrummet av reellvärda kontinuerliga funktioner på intervallet $[-\pi, \pi]$ med skalära produkten

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\, g(x)\, dx .$$

Då är $v \in V$ given av $v(x) = \sin x$ och U är det ändligtdimensionella underrummet bestående av alla polynom med reella koefficienter och av gradtal högst 5. Vi bör alltså hitta $u \in U$ så att $\|u - v\|$ minimeras. Lösningen presenteras på föreläsning!

6. Operatorer på vektorrum med skalär produkt

Till varje linjär avbildning på ett vektorrum med skalär produkt kan man associera en adjungerad avbildning. Med hjälp av denna avbildning kan vi studera egenskaper hos flera viktiga klasser av operatorer. Vi antar i detta kapitel genomgående att $V \neq \{0\}$ är ett ändligtdimensionellt vektorrum över talkroppen K.

Linjära funktionaler och adjungerade avbildningar

Definition 6.1. *En linjär funktional på ett vektorrum V över skalärkroppen K är en linjär avbildning $\varphi : V \to K$. Alltså $\varphi \in \mathcal{L}(V, K)$.*

Exempel 6.2. a) Funktionen $\varphi : K^2 \to K$ definierar en linjär funktional på K^2 om

$$\varphi(z_1, z_2) = 3z_1 - 8z_2 \,.$$

b) På vektorrummet $\mathcal{P}_n(R)$ är $\varphi : \mathcal{P}_n(R) \to R$ definierad av

$$\varphi(p) = \int_0^1 p(x) \sin x \, dx, \quad \forall p \in \mathcal{P}_n(R),$$

en linjär funktional.

Om V är ett vektorrum med skalär produkt och $v \in V$ så definierar

$$\varphi(u) = \langle u, v \rangle, \quad u \in V,$$

en linjär funktional på V. Den linjära funktionalen i Exempel 6.2 a) kan skrivas som

$$\varphi(z_1, z_2) = \langle (z_1, z_2), (3, -8) \rangle,$$

ifall K^2 är försedd med den euklidiska skalära produkten. Överraskande är kanske att varje linjär funktional på ett vektorrum med skalär produkt kan framställas på ovanbeskrivna sätt:

Sats 6.3. *Antag att φ är en linjär funktional på V. Då finns det en entydig vektor $v \in V$ sådan att*

$$\varphi(u) = \langle u, v \rangle$$

för varje $u \in V$. Om $e_1, \ldots, e_n$ är en ortonormerad bas för V, så ges $v \in V$ av $v = \overline{\varphi(e_1)}\, e_1 + \ldots + \overline{\varphi(e_n)}\, e_n$.

Bevis: Se föreläsningsanteckningar.

Definition 6.4. *Antag att V, W är ändligtdimensionella vektorrum med skalär produkt. Låt $T \in \mathcal{L}(V, W)$. Den adjungerade avbildningen till T, betecknad T^*, är en funktion från W till V, $T^* : W \to V$, definierad på följande vis. Fixera $w \in W$.* Betrakta den linjära funktionalen $\varphi \in \mathcal{L}(V, K)$ *definierad för alla $v \in V$ av $\varphi(v) = \langle T v, w \rangle$. Definiera T^*w att vara den entydigt bestämda vektor i V för vilken gäller att $\varphi(v) = \langle v, T^*w \rangle$ för alla $v \in V$ (Sats 6.3). Med andra ord, T^*w är den entydigt bestämda vektor i V för vilken gäller att*

$$\langle T v, w \rangle = \langle v, T^*w \rangle$$

för alla $v \in V$.

Exempel 6.5. Exempel på bestämning av en adjungerad avbildning. (Se föreläsningsanteckningar).

Anmärkning 6.6. Den adjungerade avbildningen T^* som bestämdes i föregående exempel är linjär. Detta gäller allmänt, ty antag att $T \in \mathcal{L}(V, W)$ och att vi fixerar $w_1, w_2 \in W$. Då gäller för $a_1, a_2 \in K$ och $v \in V$ att

$$\begin{aligned}
\langle T v, a_1 w_1 + a_2 w_2 \rangle &= \langle T v, a_1 w_1 \rangle + \langle T v, a_2 w_2 \rangle \\
&= \overline{a}_1 \langle T v, w_1 \rangle + \overline{a}_2 \langle T v, w_2 \rangle \\
&= \overline{a}_1 \langle v, T^*w_1 \rangle + \overline{a}_2 \langle v, T^*w_2 \rangle \\
&= \langle v, a_1 T^*w_1 + a_2 T^*w_2 \rangle,
\end{aligned}$$

varför $T^*(a_1 w_1 + a_2 w_2) = a_1 T^*w_1 + a_2 T^*w_2$, och T^* är linjär, $T^* \in \mathcal{L}(W, V)$.

Sats 6.7. *Antag att U, V och W är ändligtdimensionella vektorrum med skalär produkt. Då gäller:*

(*i*) $(S + T)^* = S^* + T^*$ *för alla* $S, T \in \mathcal{L}(V, W)$;

(*ii*) $(aT)^* = \overline{a}T^*$ *för alla* $a \in K$ *och* $T \in \mathcal{L}(V, W)$;

(*iii*) $(T^*)^* = T$ *för alla* $T \in \mathcal{L}(V, W)$;

(*iv*) $I^* = I$, *där I är den identiska avbildningen på V*;

(*v*) $(ST)^* = T^*S^*$ *för alla* $T \in \mathcal{L}(V, W)$, $S \in \mathcal{L}(W, U)$.

Bevis: Hemuppgift.

Följande sats ådagalägger sambanden mellan nollrum och värderum för en linjär avbildning och dess adjungerade avbildning:

Sats 6.8. *Antag att $T \in \mathcal{L}(V, W)$. Då gäller följande påståenden:*

(a) $N(T^*) = R(T)^\perp$;

(b) $R(T^*) = N(T)^\perp$;

(c) $N(T) = R(T^*)^\perp$;

(d) $R(T) = N(T^*)^\perp$.

Bevis: Låt $w \in W$. Då gäller:

$$\begin{aligned} w \in N(T^*) &\Leftrightarrow T^*w = 0 \\ &\Leftrightarrow \langle v, T^*w\rangle = 0 \ \textit{för alla}\ v \in V \\ &\Leftrightarrow \langle Tv, w\rangle = 0 \ \textit{för alla}\ v \in V \\ &\Leftrightarrow w \in R(T)^\perp, \end{aligned}$$

så påståendet (a) är bevisat. Tag de ortogonala komplementen av båda leden i (a) och vi får att (d) gäller. Låt T och T^* byta roller i (a) och (d), varmed vi får att (c) respektive (b) gäller, och beviset är klart.

Korollarium 6.9. *Antag att $T \in \mathcal{L}(V)$. Då är T inverterbar om och endast om T^* är inverterbar.*

Bevis: T inverterbar $\Leftrightarrow N(T) = \{0\} \Leftrightarrow R(T^*)^\perp = \{0\} \Leftrightarrow R(T^*) = V \Leftrightarrow T^*$ är inverterbar.

Definition 6.10. *Antag att A är en $m \times n$ matris med matriselementen $a_{i,j} \in K$, $i = 1, ..., m$, $j = 1, ..., n$. Då definieras det konjugerade transponatet av A som den $n \times m$ matris A^* vars matriselement $b_{i,j}$ ges av $b_{i,j} = \overline{a}_{j,i}$, $i = 1, ..., n$, $j = 1, ..., m$.*

Exempel 6.11.

$$A = \begin{pmatrix} 1 & i \\ 3 & 2-i \\ 1+2i & 4 \end{pmatrix}, \quad A^* = \begin{pmatrix} 1 & 3 & 1-2i \\ -i & 2+i & 4 \end{pmatrix}.$$

Nästa sats visar hur vi kan bestämma $\mathcal{M}(T^*)$ om vi känner $\mathcal{M}(T)$ och vektorrummen i fråga är försedda med **ortonormerade baser**. Om baserna ej är ortonormerade behöver sambandet mellan $\mathcal{M}(T)$ och $\mathcal{M}(T^*)$ ej gälla.

Sats 6.12. *Antag att $T \in \mathcal{L}(V, W)$. Om $e_1, ..., e_n$ är en ortonormerad bas för V och $f_1, ..., f_m$ är en ortonormerad bas för W, så gäller:*

$$\mathcal{M}(T^*, (f_1, ..., f_m), (e_1, ..., e_n)) = \mathcal{M}(T, (e_1, ..., e_n), (f_1, ..., f_m))^*.$$

Bevis: Se föreläsningsanteckningar.

Självadjungerade operatorer

Definition 6.13. *En operator $T \in \mathcal{L}(V)$ kallas självadjungerad om $T = T^*$.*

Exempel 6.14. a) Om $T \in \mathcal{L}(K^2)$ har matrisframställningen

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 3 \\ c & 8 \end{pmatrix}$$

med avseende på den naturliga basen (som är ortonormerad), så är T självadjungerad om och endast om $c = 3$, ty då gäller: $\mathcal{M}(T^*) = \mathcal{M}(T)^* = \mathcal{M}(T)$, och därmed är $T = T^*$.

b) Antag att $S, T \in \mathcal{L}(V)$ är självadjungerade. Om $e_1, ..., e_n$ är en ortonormerad bas för V gäller då: $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T) = \mathcal{M}(S^*) + \mathcal{M}(T^*) = \mathcal{M}(S^* + T^*) = \mathcal{M}((S+T)^*)$, så $S + T = (S+T)^*$, dvs. $S+T$ är självadjungerad. Om $a \in R$ erhålls: $\mathcal{M}(aT) = a\,\mathcal{M}(T) = a\,\mathcal{M}(T^*) = \mathcal{M}(aT^*) = \mathcal{M}((\overline{a}T)^*) = \mathcal{M}((aT)^*)$, så $aT = (aT)^*$, dvs. aT är självadjungerad om $a \in R$.

Den adjungerade operatorn T^* kan jämföras med komplex konjugering i C. Ett komplext tal z är reellt om och endast om $z = \overline{z}$. Självadjungerade operatorer, $T = T^*$, uppvisar en analogi med reella tal. Denna analogi reflekteras i några viktiga egenskaper för sjävadjungerade operatorer.

Sats 6.15. Varje egenvärde till en självadjungerad operator $T \in \mathcal{L}(V)$ är reellt.

Bevis: Se föreläsningsanteckningar.

Anmärkning 6.16. Om $K = R$, så är varje egenvärde definitionsmässigt reellt, varför ovanstående sats är av intresse endast då $K = C$.

Exempel 6.17. Antag att A är en $n \times n$ matris med matriselement i $K = C$. Antag vidare att $A = A^*$. Då finns det en operator $T \in \mathcal{L}(K^n)$ som med avseende på den naturliga basen för K^n, som är ortonormerad med avseende på den euklidiska skalära produkten, har matrisframställningen $\mathcal{M}(T) = A$. Med stöd av Sats 6.12 gäller då att $\mathcal{M}(T^*) = \mathcal{M}(T)^* = A^* = A$, så $T = T^*$, och T har enbart reella egenvärden (Sats 6.15). Men då har även A enbart reella egenvärden (Sats 4.11). Speciellt om alla matriselement för A är reella gäller $A = A^* = A^T$, så en reell symmetrisk $n \times n$ matris har reella egenvärden.

Följande sats gäller ej på reella vektorrum med skalär produkt. Betrakta exempelvis $T \in \mathcal{L}(R^2)$ definierad genom $T(x,y) = (-y,x)$. Då gäller det att $\langle T(x,y),(x,y)\rangle = \langle(-y,x),(x,y)\rangle = -x\,y + x\,y = 0$, för alla $(x,y) \in R^2$, fastän $T \neq 0$.

Sats 6.18. *Antag att V är ett komplext vektorrum och att $T \in \mathcal{L}(V)$. Om $\langle Tv, v\rangle = 0$ för alla $v \in V$, så är $T = 0$.*

Bevis: Se föreläsningsanteckningar.

Följande korollarium belyser vidare analogin mellan reella tal och självadjungerade operatorer:

Korollarium 6.19. *Antag att V är ett komplext vektorrum och att $T \in \mathcal{L}(V)$. Då är T självadjungerad om och endast om*

$$\langle Tv, v\rangle \in R$$

för varje $v \in V$.

Bevis: Se föreläsningsanteckningar.

Sats 6.18 gäller ej på reella vektorrum, (se sista raderna på sida 48), men om vi antar att T är självadjungerad så gäller påståendet för reella vektorrum:

Sats 6.20. *Antag att T är en självadjungerad operator på ett vektorrum V och att*

$$\langle Tv, v\rangle = 0$$

för alla $v \in V$. Då är $T = 0$.

Bevis: Se föreläsningsanteckningar.

Normala operatorer

Definition 6.21. *Antag att V är ett vektorrum med skalär produkt. Operatorn $T \in \mathcal{L}(V)$ kallas normal om*

$$TT^* = T^*T,$$

det vill säga om T kommuterar med sin adjungerade avbildning.

Exempel 6.22. a) Om T är självadjungerad så är T normal, ty $TT^* = TT = T^*T$.

b) En normal operator behöver ej vara självadjungerad. Låt exempelvis $T \in \mathcal{L}(K^2)$, där K^2 är försedd med den naturliga basen (ortonormerad). Antag vidare att

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

Då är $\mathcal{M}(T) \neq \mathcal{M}(T)^*$, så $T \neq T^*$. Enkla matriskalkyler visar, med stöd av (2.5) och Sats 6.12, att

$$\begin{aligned}\mathcal{M}(TT^*) &= \mathcal{M}(T)\,\mathcal{M}(T^*) = \mathcal{M}(T)\,\mathcal{M}(T)^* = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \\ &= \mathcal{M}(T)^*\mathcal{M}(T) = \mathcal{M}(T^*)\,\mathcal{M}(T) = \mathcal{M}(T^*T),\end{aligned}$$

så $TT^* = T^*T$ och därmed är T normal.

Följande resultat är en enkel karakterisering av normala operatorer:

Sats 6.23. *En operator $T \in \mathcal{L}(V)$ är normal om och endast om*

$$\|Tv\| = \|T^*v\|$$

för alla $v \in V$.

Bevis: Låt $T \in \mathcal{L}(V)$. Notera att $T^*T - TT^*$ är självadjungerad, $(T^*T - TT^*)^* = (T^*T)^* - (TT^*)^* = T^*(T^*)^* - (T^*)^*T^* = T^*T - TT^*$, så vi kan tillämpa Sats 6.20 i den andra av nedanstående ekvivalenser:

$$\begin{aligned} T \text{ är normal } &\Leftrightarrow T^*T - TT^* = 0 \\ &\Leftrightarrow \langle (T^*T - TT^*)\,v, v\rangle = 0,\ \forall v \in V, \\ &\Leftrightarrow \langle T^*T\,v, v\rangle = \langle TT^*v, v\rangle,\ \forall v \in V, \\ &\Leftrightarrow \langle T\,v, T\,v\rangle = \langle T^*v, T^*v\rangle,\ \forall v \in V, \\ &\Leftrightarrow \|T\,v\|^2 = \|T^*v\|^2,\ \forall v \in V, \end{aligned}$$

och satsen är bevisad.

Anmärkning 6.24. Föregående sats ger att $N(T) = N(T^*)$ för varje normal operator T, ty $v \in N(T) \Leftrightarrow T\,v = 0 \Leftrightarrow \|T\,v\| = 0 \Leftrightarrow \|T^*v\| = 0 \Leftrightarrow T^*v = 0 \Leftrightarrow v \in N(T^*)$.

Korollarium 6.25. *Antag att $T \in \mathcal{L}(V)$ är normal. Om $v \in V$ är en egenvektor till T med egenvärdet $\lambda \in K$, så är v även en egenvektor till T^* med egenvärdet $\overline{\lambda}$.*

Bevis: Antag att $v \in V$ är en egenvektor till T med egenvärdet λ. Då gäller $(T - \lambda I)\,v = 0$.

$$\begin{aligned} T \text{ är normal } \Rightarrow &(T - \lambda I)(T - \lambda I)^* = (T - \lambda I)(T^* - \overline{\lambda} I) \\ &= TT^* - \overline{\lambda}\,T\,I - \lambda\,I\,T^* + \lambda\overline{\lambda}\,I \\ &= T^*T - \overline{\lambda}\,I\,T - \lambda\,T^*I + \overline{\lambda}\lambda\,I \\ &= (T^* - \overline{\lambda} I)(T - \lambda I) = (T - \lambda I)^*(T - \lambda I), \end{aligned}$$

så $(T - \lambda I)$ är en normal operator. Då ger Sats 6.23 att:

$$0 = \|(T - \lambda I)\,v\| = \|(T - \lambda I)^*v\| = \|(T^* - \overline{\lambda} I)\,v\|,$$

så $(T^* - \overline{\lambda} I) v = 0$, varmed v är en egenvektor till T^* med egenvärdet $\overline{\lambda}$.

Eftersom varje självadjungerad operator är normal, så gäller följande resultat speciellt för självadjungerade operatorer:

Korollarium 6.26. *Om $T \in \mathcal{L}(V)$ är normal, så är egenvektorerna till T svarande mot olika egenvärden ortogonala.*

Bevis: Antag att $T \in \mathcal{L}(V)$ är normal och att $\lambda_1, \lambda_2 \in K$ är olika egenvärden till T med korresponderande egenvektorer $u_1, u_2 \in V$. Därmed gäller $T u_1 = \lambda_1 u_1$ och $T u_2 = \lambda_2 u_2$. Vidare ger Korollarium 6.25 att $T^* u_2 = \overline{\lambda}_2 u_2$. Då gäller:

$$\begin{aligned}(\lambda_1 - \lambda_2) \langle u_1, u_2 \rangle &= \langle \lambda_1 u_1, u_2 \rangle - \langle u_1, \overline{\lambda}_2 u_2 \rangle \\ &= \langle T u_1, u_2 \rangle - \langle u_1, T^* u_2 \rangle \\ &= 0,\end{aligned}$$

vilket betyder att $\langle u_1, u_2 \rangle = 0$, emedan $\lambda_1 \neq \lambda_2$.

Spektralsatsen

En operator $T \in \mathcal{L}(V)$ har som bekant en diagonal matrisframställning $\mathcal{M}(T)$ om och endast om V har en bas av egenvektorer till T. När är en dylik bas ortonormerad? Spektralsatsen ger svaret: Om $K = C$ finns det en ortonormerad bas av egenvektorer om och endast om T är normal, om $K = R$ om och endast om T är självadjungerad. Vi delar upp spektralsatsen i en komplex och i en reell version.

Sats 6.27. Komplexa spektralsatsen: *Antag att V är ett komplext vektorrum med skalärprodukt och att $T \in \mathcal{L}(V)$. Då har V en ortonormerad bas av egenvektorer till T om och endast om T är normal.*

Bevis: Se föreläsningsanteckningar.

Exempel 6.28. I Exempel 6.22 b) visade vi att operatorn $T \in \mathcal{L}(C^2)$, där C^2 är försedd med den naturliga basen, och där

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix},$$

är en normal operator. Bestäm som hemuppgift en ortonormerad bas till C^2 bestående av egenvektorer till T!

För att bevisa den reella spektralsatsen behövs två hjälpresultat.

Lemma 6.29. *Antag att $T \in \mathcal{L}(V)$ är självadjungerad. Om $\alpha, \beta \in R$ och uppfyller $\alpha^2 < 4\beta$, så är operatorn*

$$T^2 + \alpha T + \beta I$$

inverterbar.

Bevis: Antag att $\alpha, \beta \in R$ och att $\alpha^2 < 4\beta$. Låt $v \in V$ och antag att $v \neq 0$. Då gäller:

$$\begin{aligned}
\langle (T^2 + \alpha T + \beta I)\, v, v\rangle &= \langle T^2 v, v\rangle + \alpha\,\langle T\,v, v\rangle + \beta\,\langle v, v\rangle \\
&= \langle T\,v, T\,v\rangle + \alpha\,\langle T\,v, v\rangle + \beta\,\|v\|^2 \\
&\geq \|T\,v\|^2 - |\alpha|\,\|T\,v\|\,\|v\| + \beta\,\|v\|^2 \\
&= (\|T\,v\| - \frac{|\alpha|\,\|v\|}{2})^2 + (\beta - \frac{\alpha^2}{4})\,\|v\|^2 \\
&> 0\,,
\end{aligned}$$

där den första olikheten erhålls med hjälp av Cauchy-Schwartz olikhet, ty då $\langle T\,v, v\rangle \in R$ (Korollarium 6.19) så gäller

$$\alpha\,\langle T\,v, v\rangle \geq -|\alpha|\,|\langle T\,v, v\rangle| \geq -|\alpha|\,\|T\,v\|\,\|v\|\,.$$

Sista olikheten ger att $(T^2 + \alpha T + \beta I)\, v \neq 0$, så $N(T^2 + \alpha T + \beta I) = \{0\}$ och därmed är $T^2 + \alpha T + \beta I$ inverterbar.

Följande Lemma är av intresse för reella vektorrum med skalär produkt:

Lemma 6.30. *Antag att $T \in \mathcal{L}(V)$ är självadjungerad. Då har T ett egenvärde.*

Bevis: Antag att V är ett reellt vektorrum. (Klart att påståendet gäller om V är komplext). Sätt $n = \dim V$ och låt $v \in V$, $v \neq 0$. Då är vektorerna

$$v, T\,v, \ldots, T^n v$$

linjärt beroende ($n+1$ st. vektorer). Därmed finns det reella tal $a_0, \ldots, a_n$ sådana att $a_j \neq 0$ för något $j \in \{1, \ldots, n\}$, och

$$0 = a_0 v + a_1 T\,v + \ldots + a_n T^n v\,.$$

Vi bildar polynomet

$$a_0 + \ldots + a_n x^n = c(x^2 + \alpha_1 x + \beta_1)\ldots(x^2 + \alpha_M x + \beta_M)(x - \lambda_1)\ldots(x - \lambda_m)\,,$$

där $c \neq 0$ och $c, \alpha_j, \beta_j, \lambda_j \in R$, med $\alpha_j^2 < 4\beta_j$ och $m + M \geq 1$. Då erhålls:

$$\begin{aligned}
0 &= (a_0 I + a_1 T + \ldots + a_n T^n)\, v \\
&= c(T^2 + \alpha_1 T + \beta_1 I)\ldots(T^2 + \alpha_M T + \beta_M I)(T - \lambda_1 I)\ldots(T - \lambda_m I)\, v\,.
\end{aligned}$$

Varje $T^2 + \alpha_j T + \beta_j I$ är inverterbar (Lemma 6.29). Vidare är $c \neq 0$, så det måste gälla att

$$0 = (T - \lambda_1 I)...(T - \lambda_m I)\, v\,.$$

Därmed är $(T - \lambda_j I)$ icke-injektiv för något j, och då är $\lambda_j \in R$ ett egenvärde till T.

Nu kan vi bevisa den reella spektralsatsen:

Sats 6.31. Reella spektralsatsen: *Antag att V är ett reellt vektorrum med skalär produkt och att $T \in \mathcal{L}(V)$. Då har V en ortonormerad bas av egenvektorer till T om och endast om T är självadjungerad.*

Bevis: Se föreläsningsanteckningar.

För T självadjungerad, eller mera allmänt om T är normal då $K = C$, erhåller vi en uppdelning av V i invarianta underrum $N(T - \lambda_j I)$ där T är en enkel multiplikation (med λ_j) på varje underrum:

Sats 6.32. *Antag att $T \in \mathcal{L}(V)$ är självadjungerad (eller att $K = C$ och $T \in \mathcal{L}(V)$ är normal). Låt $\lambda_1, ..., \lambda_m$ beteckna de olika egenvärdena till T. Då gäller:*

$$V = N(T - \lambda_1 I) \oplus ... \oplus N(T - \lambda_m I),$$

och varje vektor i $N(T - \lambda_j I)$ är ortogonal mot varje vektor i de övriga underrummen $N(T - \lambda_l I)$, $l \neq j$.

Bevis: Spektralsatserna 6.27 och 6.31 ger att V har en bas bestående av ortonormerade egenvektorer till T. Då ger Sats 4.22*d*) att uppdelningen av V i ovanstående direkta summa är möjlig. Påståendet om ortogonaliteten följer ur Korollarium 6.26.

Definition 6.33. *Antag att A är en $n \times n$ matris. Om A har reella element och*

$$A^T A = A\,A^T = I\,,$$

så kallas A en ortogonal matris. Om A har komplexa element och

$$A^* A = A\,A^* = I\,,$$

så kallas A en unitär matris. Om

$$A\,A^* = A^* A\,,$$

så kallas A en normal matris, och speciellt om matrisen har reella element gäller $A\,A^T = A^T A$.

Anmärkning 6.34. För en ortogonal matris A och en unitär matris B gäller: a) $A^{-1} = A^T$ och $B^{-1} = B^*$; b) A^T och B^* är en ortogonal respektive unitär matris; c) Kolonnerna för A och B

bildar vardera en ortonormerad mängd av vektorer; d) Raderna för A och B bildar vardera en ortonormerad mängd av vektorer.

Exempel 6.35. Antag att A är en $n \times n$ matris med komplexa (reella) element som uppfyller $A A^* = A^* A$ ($A = A^T$). Låt C^n (R^n) vara försedd med den euklidiska skalära produkten och den naturliga ortonormerade basen $v_1, \ldots, v_n$. Då finns det en operator $T \in \mathcal{L}(C^n)$ ($T \in \mathcal{L}(R^n)$) sådan att $\mathcal{M}(T, (v_1, \ldots, v_n)) = A$. Nu gäller $\mathcal{M}(T T^*) = \mathcal{M}(T)\, \mathcal{M}(T^*) = \mathcal{M}(T)\, \mathcal{M}(T)^* = A A^* = A^* A = \mathcal{M}(T^* T)$ ($\mathcal{M}(T^*) = \mathcal{M}(T)^* = A^* = A^T = A = \mathcal{M}(T)$), så T är normal (självadjungerad). Då ger den komplexa (reella) spektralsatsen att det finns en ortonormerad bas $e_1, \ldots, e_n$ till C^n (R^n) bestående av egenvektorer till T. Då är $\mathcal{M}(T, (e_1, \ldots, e_n)) = D$ en diagonal matris med egenvärden till T i diagonalen. Eftersom A och D är likartade matriser (se sida 20) finns det en inverterbar matris Q sådan att $A = Q^{-1} D\, Q$. Med stöd av Sats 2.34 erhålls då

$$Q = \mathcal{M}(I, (v_1, \ldots, v_n), (e_1, \ldots, e_n))$$

$$= \begin{pmatrix} \langle v_1, e_1 \rangle & \cdots & \langle v_n, e_1 \rangle \\ \vdots & & \vdots \\ \langle v_1, e_n \rangle & \cdots & \langle v_n, e_n \rangle \end{pmatrix}$$

$$= \begin{pmatrix} \overline{e_1^1} & \cdots & \overline{e_1^n} \\ \vdots & & \vdots \\ \overline{e_n^1} & \cdots & \overline{e_n^n} \end{pmatrix},$$

där $e_j = (e_j^1, \ldots, e_j^n)$, $j = 1, \ldots, n$. Då ser vi att $Q\, Q^* = I$ ($Q\, Q^T = I$), så $Q^* = Q^{-1}$ ($Q^T = Q^{-1}$) och därmed gäller $Q^* Q = I$ ($Q^T Q = I$), så Q är unitär (ortogonal). Då gäller uppdelningen

$$A = Q^* D\, Q \quad (A = Q^T D\, Q)\ ,$$

där kolonnerna till Q^* (Q^T) ges av $e_1, \ldots, e_n$. En matris A som kan uppdelas på ovanstående sätt kallas **unitärt diagonaliserbar (ortogonalt diagonaliserbar)**. Denna omständighet föreligger om och endast om A är normal (symmetrisk), på grund av spektralsatserna.

(Överkurs: Om vi för en normal (symmetrisk) matris A vill beräkna uppdelningen $A = Q^* D\, Q$ ($A = Q^T D\, Q$) så kan detta göras i t.ex. matlab med kommandot

$$>> [U, D] = \mathrm{schur}(A)$$

varvid D är den eftersökta diagonala matrisen av egenvärden och $U = Q^*$ ($U = Q^T$), så $A = U\, D\, U^*$ ($A = U\, D\, U^T$).)

Exempel 6.36. Antag att $K = C$ och att A är en $n \times n$ matris som är normal, $A\,A^* = A^*A$. Bestäm ett behändigt sätt att beräkna potenserna A^m, för $m = 1, 2, \ldots$.

Enligt Exempel 6.35 är A unitärt diagonaliserbar. Genom att bestämma egenvärden och egenvektorer till A erhåller vi en ortonormerad bas $e_1, \ldots, e_n$ för C^n bestående av egenvektorer till A svarande mot egenvärdena $\lambda_1, \ldots, \lambda_n$. Då har vi uppdelningen:

$$A = Q^* D\, Q, \quad \text{där } D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ och } Q^* = \begin{pmatrix} | & & | \\ e_1 & \cdots & e_n \\ | & & | \end{pmatrix}.$$

För $m = 1$ gäller då $A^1 = Q^* D^1\, Q$. Antag att $A^m = Q^* D^m\, Q$ för $m \geq 1$. Då gäller:

$$\begin{aligned} A^{m+1} &= A\,A^m = (Q^* D\, Q)(Q^* D^m Q) = (Q^* D)(Q\,Q^*)(D^m Q) = Q^* D^{m+1} Q \\ &= Q^* \begin{pmatrix} \lambda_1^{m+1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{m+1} \end{pmatrix} Q. \end{aligned}$$

Induktion ger då att ovanstående formel gäller.

Exempel 6.37. Varje år flyttar 1/10 av populationen i området B till området A och 2/10 av populationen i området A flyttar till B. Totala individantalet antas hållas konstant över åren. Vid år $k = 0$ har vi z_0 och y_0 individer i området A respektive området B. Hur många individer har vi i området A respektive B efter k år?

För $k = 0, 1, \ldots$ har vi sambanden

$$\begin{aligned} y_{k+1} &= (9/10) y_k + (2/10) z_k \\ z_{k+1} &= (1/10) y_k + (8/10) z_k, \end{aligned}$$

vilka kan skrivas i matrisform,

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} \frac{9}{10} & \frac{2}{10} \\ \frac{1}{10} & \frac{8}{10} \end{pmatrix} \begin{pmatrix} y_k \\ z_k \end{pmatrix}.$$

Om vi betecknar

$$M = \begin{pmatrix} \frac{9}{10} & \frac{2}{10} \\ \frac{1}{10} & \frac{8}{10} \end{pmatrix}$$

så ser vi enkelt att

$$\begin{pmatrix} y_k \\ z_k \end{pmatrix} = M^k \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}, \quad k = 1, 2, \ldots.$$

Notera att M inte är symmetrisk (ej självadjungerad), så det finns ingen bas för R^2 bestående av ortonormerade egenvektorer till M, (reella spektralsatsen).

Däremot finns det en bas för R^2 bestående av egenvektorer till M, ty $\lambda_1 = 7/10$ och $\lambda_2 = 1$ är egenvärdena till M med motsvarande linjärt oberoende egenvektorer $v_1 = (-1, 1)$ och $v_2 = (2, 1)$. Om vi definierar

$$S = \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} \text{ och } D = \begin{pmatrix} \frac{7}{10} & 0 \\ 0 & 1 \end{pmatrix}$$

så följer det då ur demonstrationsuppgift 2) 14.2.2002 att

$$M = S\,D\,S^{-1} \text{ och } M^k = S\,D^k S^{-1},\ k = 1, 2, \ldots.$$

Alltså gäller:

$$\begin{aligned}
\begin{pmatrix} y_k \\ z_k \end{pmatrix} &= \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (\frac{7}{10})^k & 0 \\ 0 & 1^k \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \\
&= (y_0 + z_0) \cdot 1^k \cdot \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} + (y_0 - 2z_0) \cdot (\frac{7}{10})^k \cdot \begin{pmatrix} 1/3 \\ -(1/3) \end{pmatrix} \\
&\longrightarrow (y_0 + z_0) \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}, \text{ då } k \to \infty.
\end{aligned}$$

I det långa loppet kommer således B att bestå av ungefär 2/3 och A av ungefär 1/3 av den totala populationen.

Normala operatorer på reella vektorrum med skalär produkt

Den komplexa spektralsatsen ger en fullständig beskrivning av normala operatorer på komplexa vektorrum och den reella spektralsatsen ger motsvarande karakterisering av självadjungerade operatorer på reella vektorrum. För att få en komplett bild av normala operatorer som inte är självadjungerade inleder vi våra studier i reella vektorrum av dimension 2.

Lemma 6.38. *Antag att V är ett 2-dimensionellt reellt vektorrum och att $T \in \mathcal{L}(V)$. Då är följande påståenden ekvivalenta:*

(a) *T är normal men ej självadjungerad;*

(b) *Om basen för V är ortonormerad så är matrisframställningen av T av formen*

$$\mathcal{M}(T) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \text{ där } b \neq 0;$$

(c) *Det finns en ortonormerad bas för V så att matrisframställningen av T är av formen*

$$\mathcal{M}(T) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \text{ med } b > 0.$$

Bevis: Antag att (a) gäller. Låt e_1, e_2 vara en ortonormerad bas för V. Antag att

$$\mathcal{M}(T, (e_1, e_2)) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Då är $\|T e_1\|^2 = \langle a e_1 + b e_2, a e_1 + b e_2 \rangle = a^2 + b^2$ och $\|T^* e_1\|^2 = \ldots = a^2 + c^2$. Då T är normal gäller $\|T e_1\| = \|T^* e_1\|$ (Sats 6.23), så $b^2 = c^2$ och därmed gäller $c = b$ eller $c = -b$. Men $c = b$ kan inte gälla, ty då vore T självadjungerad. Därmed gäller $c = -b$. Då $\mathcal{M}(T^*) = \mathcal{M}(T)^*$ erhåller vi

$$\mathcal{M}(TT^*) = \mathcal{M}(T)\,\mathcal{M}(T^*) = \mathcal{M}(T)\,\mathcal{M}(T)^* = \begin{pmatrix} a^2+b^2 & ab-bd \\ ab-bd & b^2+d^2 \end{pmatrix},$$

$$\mathcal{M}(T^*T) = \mathcal{M}(T^*)\,\mathcal{M}(T) = \mathcal{M}(T)^*\mathcal{M}(T) = \begin{pmatrix} a^2+b^2 & -(ab-bd) \\ -(ab-bd) & b^2+d^2 \end{pmatrix}.$$

Eftersom T är normal måste ovanstående matriser vara lika. Alltså: $bd = ab$ måste gälla. Nu är $b \neq 0$, ty annars vore T självadjungerad, (jämför $\mathcal{M}(T)$ med $c = -b = 0$). Då är $a = d$, och $(a) \Rightarrow (b)$.

Antag nu att (b) gäller. Låt e_1, e_2 vara en ortonormerad bas för V. Vi vet att $\mathcal{M}(T)$ är av den form som anges i 6.38 (b), med $b \neq 0$. Om $b > 0$ så gäller (c). Om $b < 0$ så betraktar vi istället den ortonormerade basen $e_1, -e_2$ för V med $\mathcal{M}(T) = \begin{pmatrix} c & -d \\ d & c \end{pmatrix}$. Då har vi $T e_1 = c e_1 + d(-e_2) = c e_1 - d e_2$. Men eftersom framställningen av $T e_1$ i basen e_1, e_2 är entydig, så gäller $c = a$ och $-d = b < 0$, så $d > 0$ och $(b) \Rightarrow (c)$.

Antag slutligen att (c) gäller. Då har $\mathcal{M}(T)$ formen given i 6.38 (c) med avseende på någon ortonormerad bas för V. Eftersom $\mathcal{M}(T) \neq \mathcal{M}(T)^*$, så är T inte självadjungerad. Vidare gäller:

$$\mathcal{M}(TT^*) = \mathcal{M}(T)\,\mathcal{M}(T)^* = \begin{pmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{pmatrix} = \mathcal{M}(T)^*\mathcal{M}(T)$$
$$= \mathcal{M}(T^*T),$$

så $TT^* = T^*T$ och därmed är T normal. Då har vi att $(c) \Rightarrow (a)$, och beviset är klart.

Om vi delar in en matris D i delmatriser och namnger dessa, kan vi betrakta D som en **blockmatris**. Om exempelvis

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

så kan vi skriva D som blockmatrisen

$$D = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

där blocken ges av

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \; B = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}, \; C = \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix}$$

och 0 betecknar nollmatrisen av dimension 3×2. En **block diagonalmatris** är en kvadratisk matris av formen

$$D = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

där $A_1, \ldots, A_m$ är kvadratiska matriser vars diagonalelement ligger på diagonalen för D. Om Exempelvis

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

så är D en block diagonalmatris med blocken

$$A_1 = (4),\ A_2 = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix},\ A_3 = \begin{pmatrix} 2 & -5 \\ 5 & 2 \end{pmatrix}.$$

Om A och B är block diagonala matriser av formen

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix} \text{ och } B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix},$$

där A_j har samma dimension som B_j, $j = 1, \ldots, m$, så är AB block diagonal av formen

$$AB = \begin{pmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{pmatrix}.$$

Sats 6.39. *Antag att operatorn $T \in \mathcal{L}(V)$ är normal och att U är ett underrum av V som är invariant under T. Då gäller:*

(a) $U^\perp$ är invariant under T;

(b) U är invariant under T^*;

(c) $(T|_U)^* = (T^*)|_U$;

(d) $T|_U$ är en normal operator på U;

(e) $T|_{U^\perp}$ är en normal operator på $U^\perp$.

Bevis: Först visar vi att (a) gäller. Låt $e_1, \ldots, e_m$ vara en ortonormerad bas för U. Utvidga basen till en ortonormerad bas $e_1, \ldots, e_m, f_1, \ldots, f_n$ för V (Korollarium 5.22). Då U är invariant under T gäller det att $Te_j \in [e_1, \ldots, e_m]$. Därmed är matrisen för T med avseende på basen $e_1, \ldots, e_m, f_1, \ldots, f_n$ av formen

$$\mathcal{M}(T) = \begin{array}{c} \\ e_1 \\ \vdots \\ e_m \\ f_1 \\ \vdots \\ f_n \end{array} \!\!\begin{array}{c} \begin{array}{cccccc} e_1 & \ldots & e_m & f_1 & \ldots & f_n \end{array} \\ \left(\begin{array}{cc} A & B \\ & \\ 0 & C \end{array}\right) \end{array},$$

där A är en $m \times m$ matris, 0 betecknar en $n \times m$ nollmatris, B en $m \times n$ matris och C en $n \times n$ matris. Låt $a_{i,j}$ och $b_{i,j}$ beteckna elementen i A respektive B. Då gäller, med stöd av att $\mathcal{M}(T^*) = \mathcal{M}(T)^*$, att

$$\sum_{j=1}^{m} \|T e_j\|^2 = \sum_{j=1}^{m} \sum_{i=1}^{m} |a_{i,j}|^2,$$

$$\sum_{j=1}^{m} \|T^* e_j\|^2 = \sum_{j=1}^{m} (\sum_{i=1}^{m} |a_{j,i}|^2 + \sum_{i=1}^{n} |b_{j,i}|^2).$$

Då T är normal gäller $\|T e_j\| = \|T^* e_j\|$ för alla j, (Sats 6.23), vilket tillsammans med ovanstående ekvationer ger att $|b_{i,j}| = 0$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Därmed är B nollmatrisen av dimension $m \times n$, vilket betyder att $T f_j \in [f_1, \ldots, f_n]$, $j = 1, \ldots, n$. Eftersom $[f_1, \ldots, f_n] = U^\perp$ så erhåller vi då för $u \in U^\perp$ att $T u = T(a_1 f_1 + \ldots + a_n f_n) = a_1 T f_1 + \ldots + a_n T f_n \in U^\perp$, så $U^\perp$ är invariant under T. Därmed gäller (a).

Då B är nollmatrisen har $\mathcal{M}(T^*) = \mathcal{M}(T)^*$ ett $n \times m$ block av nollor i nedre vänstra hörnet, vilket ger att $T^* e_j \in [e_1, \ldots, e_m]$, så (b) gäller.

För att bevisa (c) låter vi $S = T|_U$. Fixera $v \in U$. Då gäller

$$\langle S u, v \rangle = \langle T u, v \rangle = \langle u, T^* v \rangle,$$

för alla $u \in U$. Eftersom $T^* v \in U$ (med stöd av (b)), så gäller $S^* v = T^* v$, alltså $(T|_U)^* = (T^*)|_U$, och (c) är bevisad.

Egenskap (c) och att T är normal ger: $(T|_U)(T|_U)^* = (T|_U)(T^*|_U) = (T^*|_U)(T|_U) = (T|_U)^*(T|_U)$, så (d) gäller.

Eftersom (d) gäller för varje underrum U som är invariant under T, och $U^\perp$ är ett underrum som enligt (a) är invariant under T, får vi att (e) gäller. Därmed är beviset klart.

Vi skall nu bevisa avsnittets huvudresultat som säger att en normal operator som inte är självadjungerad på ett reellt vektorrum har en "nästan" diagonal matrisframställning med avseende på någon ortonormerad bas:

Sats 6.40. *Antag att V är ett reellt vektorrum med skalär produkt och att $T \in \mathcal{L}(V)$. Då är T normal om och endast om det finns en ortonormerad bas för V med avseende på vilken $\mathcal{M}(T)$ är en block diagonal matris där varje block är en 1×1 matris eller en 2×2 matris av formen*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \tag{6.1}$$

med $b > 0$.

Bevis: 1) Antag att det finns en ortonormerad bas för V sådan att $\mathcal{M}(T)$ är en block diagonal matris där varje block är en 1×1 matris eller en 2×2 matris av formen (6.1). Då kommuterar varje sådant block med sitt transponat, så om

$$\mathcal{M}(T) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

så gäller det att

$$\begin{aligned}\mathcal{M}(TT^*) &= \mathcal{M}(T)\,\mathcal{M}(T)^* \\ &= \begin{pmatrix} A_1A_1^* & & 0 \\ & \ddots & \\ 0 & & A_mA_m^* \end{pmatrix} \\ &= \begin{pmatrix} A_1^*A_1 & & 0 \\ & \ddots & \\ 0 & & A_m^*A_m \end{pmatrix} \\ &= \mathcal{M}(T)^*\mathcal{M}(T) \\ &= \mathcal{M}(T^*T)\,.\end{aligned}$$

Då är $TT^* = T^*T$, så T är normal.

2) Antag att T är normal. Om $\dim V = 1$ finns det en ortonormerad bas med den önskade egenskapen. Om $\dim V = 2$ och T är självadjungerad ger den reella spektralsatsen att det finns en ortonormerad bas så att $\mathcal{M}(T)$ är diagonal, om T är normal men ej självadjungerad ger Lemma 6.38 att det finns en ortonormerad bas så att $\mathcal{M}(T)$ är av form (6.1).

Antag nu att $\dim V > 2$ och att resultatet gäller för reella vektorrum av lägre dimension än V. Med stöd av Sats 4.23 har V ett underrum U av dimension 1 eller 2 som är invariant under T. Om $\dim U = 1$ väljer vi en vektor i U med normen 1. Denna vektor utgör då en ortonormerad bas för U och $\mathcal{M}(T|_U)$ är en 1×1 matris. Om vi inte har ett invariant underrum U av dimension 1 väljer vi U med dimesionen 2. Då är $T|_U$ normal (Sats 6.39 (d)). Men $T|_U$ är inte självadjungerad, ty då skulle $T|_U$ (och T) ha en egenvektor $v \neq 0$ (Lemma 6.30), och därmed skulle $[v]$ vara ett invariant underrum av dimension 1 under T. Då kan vi hitta en ortonormerad bas för U så att $\mathcal{M}(T|_U)$ har formen (6.1) (Lemma 6.38). Med stöd av Sats 6.39 är nu $U^\perp$ invariant under T och $T|_{U^\perp}$ är en normal operator på $U^\perp$. Vårt induktionsantagande ger då att $U^\perp$ har en ortonormerad bas sådan att $\mathcal{M}(T|_{U^\perp})$ är av önskad form. Genom att sammanslå de ortonormerade baserna för U och $U^\perp$ får vi då en ortonormerad bas för V sådan att $\mathcal{M}(T)$ är av önskad form,

$$\mathcal{M}(T) = \begin{pmatrix} \mathcal{M}(T|_U) & 0 \\ 0 & \mathcal{M}(T|_{U^\perp}) \end{pmatrix},$$

en block diagonal matris där blocken är block diagonala matriser vars block består av 1×1 matriser eller 2×2 matriser av form (6.1). Därmed är satsen bevisad.

Positiva operatorer

Definition 6.41. *En operator $T \in \mathcal{L}(V)$ kallas positiv om T är självadjungerad och*

$$\langle Tv, v\rangle \geq 0$$

för alla $v \in V$.

Exempel 6.42. a) Låt U vara ett ändligtdimensionellt underrum av ett vektorrum V och P_U må vara ortogonala projektionen av V på U. Vi har att $V = U \oplus U^{\perp}$. Tag $v_1, v_2 \in V$. Då är $v_1 = u_1 + w_1$ och $v_2 = u_2 + w_2$, där $u_1, u_2 \in U$ och $w_1, w_2 \in U^{\perp}$. Då gäller: $\langle P_U v_1, v_2\rangle = \langle u_1, u_2 + w_2\rangle = \langle u_1, u_2\rangle + \langle u_1, w_2\rangle = \langle u_1, u_2\rangle = \langle u_1, u_2\rangle + \langle w_1, u_2\rangle = \langle u_1 + w_1, u_2\rangle = \langle v_1, P_U v_2\rangle$. Därmed är $P_U = P_U^*$, och P_U är självadjungerad. Om $v \in U^{\perp}$ så är $v = 0 + v$ och $\langle P_U v, v\rangle = \langle 0, v\rangle = 0$. Om $v \in U$ så är $v = v + 0$ och $\langle P_U v, v\rangle = \langle v, v\rangle = \|v\|^2 \geq 0$. Därmed är $\langle P_U v, v\rangle \geq 0$ för alla $v \in V$ och P_U är en positiv operator.

b) Om $T \in \mathcal{L}(V)$ är självadjungerad och $\alpha^2 < 4\beta$, där $\alpha, \beta \in R$, så är $T^2 + \alpha T + \beta I$ en positiv operator. (jämför beviset av Lemma 6.29).

Definition 6.43. *En operator $S \in \mathcal{L}(V)$ kallas en kvadratrot av operatorn $T \in \mathcal{L}(V)$ om $S^2 = T$.*

Exempel 6.44. Antag att $S, T \in \mathcal{L}(K^3)$. Vi definierar S och T genom $T(z_1, z_2, z_3) = (z_3, 0, 0)$ och $S(z_1, z_2, z_3) = (z_2, z_3, 0)$. Då är S en kvadratrot av T eftersom

$$S^2(z_1, z_2, z_3) = S(z_2, z_3, 0) = (z_3, 0, 0) = T(z_1, z_2, z_3).$$

Vi har tidigare påpekat analogin mellan reella tal och självadjungerade operatorer. Följande sats, som är huvudresultatet i detta avsnitt, påvisar en viss analogi mellan positiva operatorer och icke-negativa reella tal:

Sats 6.45. *Låt $T \in \mathcal{L}(V)$. Då är följande påståenden ekvivalenta:*

- *(a) T är en positiv operator;*
- *(b) T är självadjungerad och alla egenvärden till T är reella och icke-negativa;*
- *(c) T har en positiv kvadratrot;*
- *(d) T har en självadjungerad kvadratrot;*
- *(e) det finns en operator $S \in \mathcal{L}(V)$ sådan att $T = S^*S$.*

Bevis: Se föreläsningsanteckningar.

Varje icke-negativt reellt tal har en entydigt bestämd icke-negativ reell kvadratrot. En liknande egenskap gäller för positiva operatorer:

Sats 6.46. *Varje positiv operator $T \in \mathcal{L}(V)$ har en entydigt bestämd positiv kvadratrot.*

Bevis: Antag att $T \in \mathcal{L}(V)$ är positiv och låt $\lambda_1, ..., \lambda_m$ beteckna de olika egenvärdena till T. Sats 6.45 ger att $\lambda_j \geq 0$, $j = 1, ..., m$. Då T är självadjungerad ger Sats 6.32 att

$$V = N(T - \lambda_1 I) \oplus ... \oplus N(T - \lambda_m I) \,. \tag{6.2}$$

Antag att $S \in \mathcal{L}(V)$ är en positiv kvadratrot av T. Låt α vara ett egenvärde till S, Sats 6.45 (b) ger att $\alpha \geq 0$. För $v \in N(S - \alpha I)$ gäller $S v = \alpha v$, så

$$T v = S^2 v = S(\alpha v) = \alpha^2 v, \tag{6.3}$$

vilket betyder att $v \in N(T - \alpha^2 I)$. Därmed är α^2 ett egenvärde till T och $\alpha^2 = \lambda_j$ för något j, vilket ger att $\alpha = \sqrt{\lambda_j}$, (ty $\alpha \geq 0$). Ur (6.3) följer då att:

$$N(S - \sqrt{\lambda_j}\, I) \subseteq N(T - \lambda_j I) \,. \tag{6.4}$$

Nu är de enda tänkbara egenvärdena till S givna av $\sqrt{\lambda_1}, ..., \sqrt{\lambda_m}$. Eftersom S är självadjungerad gäller med stöd av Sats 6.32 att

$$V = N(S - \sqrt{\lambda_1}\, I) \oplus ... \oplus N(S - \sqrt{\lambda_m}\, I) \,. \tag{6.5}$$

Då följer det av (6.2), (6.5) och (6.4) att $N(S - \sqrt{\lambda_j}\, I) = N(T - \lambda_j I)$ måste gälla för alla j. På varje $N(T - \lambda_j I)$ är då operatorn S en multiplikation med $\sqrt{\lambda_j}$, så S är entydigt bestämd av T, och beviset är klart.

Definition 6.47. *Om T är en positiv operator så betecknar vi den entydigt bestämda positiva kvadratroten av T med $\sqrt{T}$.*

Exempel 6.48. Antag att A är en $n \times n$ matris med matriselement i K. Om $A = A^*$ och $x^* A\, x \geq 0$ för alla $x \in K^n$ så kallas A **positivt semidefinit.**

Antag att A är positivt semidefinit och låt $V = K^n$ vara försedd med den naturliga ortonormerade basen och den euklidiska skalära produkten. Då finns det en operator $T \in \mathcal{L}(V)$ sådan att $\mathcal{M}(T) = A$. Vidare gäller: $\mathcal{M}(T) = A = A^* = \mathcal{M}(T)^* = \mathcal{M}(T^*)$, så T är självadjungerad. Med beaktande av att $\mathcal{M}(v) = v$ för alla $v \in V$ då $V = K^n$ är försedd med den naturliga basen, erhåller vi

$$\langle T v, v \rangle = \langle \mathcal{M}(T v), v \rangle = \langle \mathcal{M}(T)\mathcal{M}(v), v \rangle = \langle A v, v \rangle = v^* A v \geq 0$$

för alla $v \in V$. Då är T en positiv operator med reella och icke-negativa egenvärden. Eftersom $\mathcal{M}(T)$ (och A) har samma egenvärden gäller det då att **alla egenvärden till en positivt semidefinit matris är reella och icke-negativa.**

Isometri

Definition 6.49. *En operator $S \in \mathcal{L}(V)$ kallas en isometri om*

$$\|S\,v\| = \|v\|$$

för alla $v \in V$. Om $K = C$ $(K = R)$ kallas S en unitär (ortogonal) operator.

Anmärkning 6.50 Varje isometri är inverterbar, ty om $v \neq 0$ så gäller $0 < \|v\| = \|S\,v\|$ och därmed är $S\,v \neq 0$, så $N(S) = \{0\}$ och S är injektiv, (Sats 2.31).

Exempel 6.51 a) Varje unitär (ortogonal) matris är en isometri, ty $\|A\,x\|^2 = \langle A\,x, A\,x\rangle = (A\,x)^* A\,x = x^* A^* A\,x = x^* x = \|x\|^2$ $(\|A\,x\|^2 = \langle A\,x, A\,x\rangle = (A\,x)^T A\,x = x^T A^T A\,x = x^T x = \|x\|^2)$.

b) Operatorn S som roterar varje vektor i R^2 vinkeln θ moturs runt origo är en isometri. Om vi har den naturliga basen för R^2 så visade vi i Exempel 2.19 att matrisen för S ges av

$$\mathcal{M}(S) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Sats 6.52. *Antag att $S \in \mathcal{L}(V)$. Då är följande påståenden ekvivalenta:*

(a) S är en isometri;

(b) $\langle S\,u, S\,v\rangle = \langle u, v\rangle$ för alla $u, v \in V$;

(c) $S^* S = I$;

(d) om $e_1, ..., e_n$ är ortonormerade vektorer, så är även $S\,e_1, ..., S\,e_n$ ortonormerade i V;

(e) det finns en ortonormerad bas $e_1, ..., e_n$ för V sådan att $S\,e_1, ..., S\,e_n$ är ortonormerade;

(f) S^* är en isometri;

(g) $\langle S^*\,u, S^*\,v\rangle = \langle u, v\rangle$ för alla $u, v \in V$;

(h) $S\,S^* = I$;

(i) om $e_1, ..., e_n$ är ortonormerade vektorer, så är även $S^* e_1, ..., S^* e_n$ ortonormerade i V;

(j) det finns en ortonormerad bas $e_1, ..., e_n$ för V sådan att $S^* e_1, ..., S^* e_n$ är ortonormerade.

Bevis: Antag att (a) gäller. Om V är ett reellt vektorrum så gäller det för varje $u, v \in V$ att

$$\langle u, v\rangle = (\|u + v\|^2 - \|u - v\|^2)/4,$$

och om V är ett komplext vektorrum gäller

$$\langle u, v\rangle = (\|u + v\|^2 - \|u - v\|^2 + \|u + i\,v\|^2\,i - \|u - i\,v\|^2\,i)/4,$$

där i är den imaginära enheten. Kolla att detta stämmer! Om V är reellt gäller då

$$\begin{aligned}\langle S\,u, S\,v\rangle &= (\|S\,u + S\,v\|^2 - \|S\,u - S\,v\|^2)/4\\ &= (\|S(u+v)\|^2 - \|S(u-v)\|^2)/4\\ &= (\|u+v\|^2 - \|u-v\|^2)/4\\ &= \langle u, v\rangle\,.\end{aligned}$$

På motsvarande sätt verifieras $\langle S\,u, S\,v\rangle = \langle u, v\rangle$ om V är ett komplext vektorrum (kolla!). Alltså: $(a) \Rightarrow (b)$.

Antag nu att (b) gäller. Då har vi att

$$\langle (S^*S - I)\,u, v\rangle = \langle S\,u, S\,v\rangle - \langle u, v\rangle = 0$$

för alla $u, v \in V$. Speciellt om $v = (S^*S - I)\,u$, ser vi att $S^*S - I = 0$, så $S^*S = I$ och $(b) \Rightarrow (c)$.

Antag att (c) gäller. Låt $e_1, \ldots, e_n$ vara ortonormerade vektorer i V. Då gäller

$$\langle S\,e_j, S\,e_k\rangle = \langle S^*Se_j, e_k\rangle = \langle e_j, e_k\rangle,$$

vilket ger att $S\,e_1, \ldots, S\,e_n$ är ortonormerade, så $(c) \Rightarrow (d)$.

Antag att (d) gäller. Vektorrummet V har en ortonormerad bas $e_1, \ldots, e_n$ (Korollarium 5.21). Då gäller (e), så $(d) \Rightarrow (e)$.

Antag att (e) gäller. Låt $e_1, \ldots, e_n$ vara en ortonormerad bas för V sådan att $S\,e_1, \ldots, S\,e_n$ är ortonormerade vektorer. Om $v \in V$ så gäller

$$\begin{aligned}\|S\,v\|^2 &= \|S(\langle v, e_1\rangle\,e_1 + \ldots + \langle v, e_n\rangle\,e_n)\|^2\\ &= \|\langle v, e_1\rangle\,S\,e_1 + \ldots + \langle v, e_n\rangle\,S\,e_n\|^2\\ &= |\langle v, e_1\rangle|^2 + \ldots + |\langle v, e_n\rangle|^2\\ &= \|v\|^2\,,\end{aligned}$$

där den första och sista likheten ges av Sats 5.19. Därmed ser vi att $(e) \Rightarrow (a)$.

Då har vi visat att påståendena (a) till (e) är ekvivalenta. Genom att ersätta S med S^* i (a) till (e) erhåller vi att påståendena (f) till (j) är ekvivalenta. Vi bör slutligen upprätta en ekvivalens mellan dessa två kedjor av ekvivalenser. Vi visar att $(c) \Leftrightarrow (h)$. Med stöd av demonstrationsuppgift 3) b) 20.1.2005 gäller det att $S^*S = I$ om och endast om $S\,S^* = I$, så $(c) \Leftrightarrow (h)$ och beviset är klart.

Anmärkning 6.53. Föregående sats visar att skalära produkten inte påverkas av en isometri, egenskap (b).

Egenskaperna (d) och (i) antyder att om S är en isometri och $e_1, \ldots, e_n$ är en ortonormerad bas för V, så är kolonnerna respektive raderna i $\mathcal{M}(S)$ ortonormerade, vilket betyder att $\mathcal{M}(S)$ är en unitär (ortogonal) matris om $K = C$ $(K = R)$.

Vidare ger (c) och (h) att varje isometri är en normal operator, vilket vi kan utnyttja i följande satser för att ge en fullständig beskrivning av isometrier.

Sats 6.54. *Antag att V är ett komplext vektorrum och att $S \in \mathcal{L}(V)$. Då är S en isometri om och endast om det finns en ortonormerad bas $e_1, ..., e_n$ för V bestående av egenvektorer till S vars korresponderande egenvärden $\lambda_1, ..., \lambda_n$ uppfyller $|\lambda_j| = 1$, $j = 1, ..., n$.*

Bevis: Se föreläsningsanteckningar.

Varje isometri på ett reellt vektorrum byggs upp av delar som svarar mot multiplikation med 1 eller -1, eller mot rotation i ett 2-dimensionellt underrum:

Sats 6.55. *Antag att V är ett reellt vektorrum och att $S \in \mathcal{L}(V)$. Då är S en isometri om och endast om det finns en ortonormerad bas $e_1, ..., e_n$ för V med avseende på vilken $\mathcal{M}(S)$ är en block diagonalmatris sådan att varje block är en 1×1 matris bestående av skalären 1 eller -1, eller en 2×2 matris av formen*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{6.6}$$

där $\theta \in (0, \pi)$.

Bevis: 1) Antag att S är en isometri. Eftersom S då är normal finns det en ortonormerad bas $e_1, ..., e_n$ för V sådan att $\mathcal{M}(S)$ är block diagonal, där varje block är antingen en 1×1 matris eller en 2×2 matris av formen

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \tag{6.7}$$

där $b > 0$ (Sats 6.40). Om ett 1×1 block på diagonalen av $\mathcal{M}(S)$ består av skalären λ, så svarar då mot detta block en egenvektor e_j som är en basvektor och $S\,e_j = \lambda\, e_j$. Eftersom S är en isometri måste $|\lambda| = 1$ gälla, $(1 = \|e_j\| = \|S\,e_j\| = |\lambda|\,\|e_j\|)$. Men då $\lambda \in R$ är $\lambda = 1$ eller $\lambda = -1$.

Betrakta nu en 2×2 matris av formen (6.7) som ett diagonalblock för $\mathcal{M}(S)$. Då finns det basvektorer e_j, e_{j+1} så att

$$S\,e_j = a\,e_j + b\,e_{j+1},$$

vilket ger att

$$1 = \|e_j\|^2 = \|S\,e_j\|^2 = ... = a^2 + b^2\,.$$

Då $0 < b \leq 1$ finns det ett tal $\theta \in (0, \pi)$ sådant att $a = \cos\theta$ och $b = \sin\theta$, varvid (6.7) är av formen (6.6).

2) Antag att det finns en ortonormerad bas för V med avseende på vilken $\mathcal{M}(S)$ har den i påståendet beskrivna block diagonala formen. Då kan vi dela upp V i direkta summan

$$V = U_1 \oplus ... \oplus U_m,$$

där varje U_j är ett underrum av V av dimensionen 1 eller 2, beroende av dimensionen på blocket A_j i matrisen $\mathcal{M}(S)$. Två vektorer hörande till olika underrum U_j är ortogonala, (varje U_j spänns upp av en eller två av basvektorerna i den ortonormerade basen för V). Vidare är $S|_{U_j} \in \mathcal{L}(U_j)$

en isometri på U_j, ty om $\dim U_j = 1$ och $u \in U_j$ så är $S\,u = u$ eller $S\,u = -u$, eller om $\dim U_j = 2$ och $u \in U_j$, så är $S\,u$ en rotation av u, och därmed gäller $\|S\,u\| = \|u\|$ för $u \in U_j$. Tag $v \in V$. Då har v en entydig framställning

$$v = u_1 + \ldots + u_m,$$

där $u_j = U_j$. Eftersom $\langle S\,u_i, S\,u_j\rangle = 0$ och $\langle u_i, u_j\rangle = 0$, om $i \neq j$, så kan vi tillämpa Pythagoras sats på ovanstående framställning av v,

$$\begin{aligned}\|S\,v\|^2 &= \|S(u_1 + \ldots + u_m)\|^2 = \|S\,u_1 + \ldots + S\,u_m\|^2 \\ &= \|S\,u_1\|^2 + \ldots + \|S\,u_m\|^2 = \|u_1\|^2 + \ldots + \|u_m\|^2 \\ &= \|v\|^2,\end{aligned}$$

så S är en isometri på V och beviset är klart.

Anmärkning 6.56. Ur ovanstående sats följer att om S är en isometri på ett reellt vektorrum av udda dimension, så är 1 eller -1 ett egenvärde till S.

Polära uppdelningen

Vi skall nu utvidga analogin mellan C och $\mathcal{L}(V)$. I denna analogi kan vi tänka oss att $z \in C$ svarar mot en operator T och $\bar{z}$ mot den adjungerade operatorn T^*. Självadjungerade operatorer, $T = T^*$, svarar mot reella tal $z = \bar{z}$ och positiva operatorer mot icke-negativa reella tal.

En speciell delmängd av C utgörs av enhetscirkeln $|z| = 1$, med andra ord mängden av z sådana att $\bar{z}\,z = 1$. I vår analogi svarar detta mot $T^*T = I$, så enhetscirkeln i C är analog med mängden av isometrier på V.

Varje komplext tal $z \neq 0$ kan skrivas i polär form

$$z = \left(\frac{z}{|z|}\right)|z| = \left(\frac{z}{|z|}\right)\sqrt{z\,\bar{z}}\ (= e^{i\theta}\,r),$$

där $z/|z|$ är ett tal på enhetscirkeln i C och $\sqrt{z\,\bar{z}} > 0$. Vår analogi kan utvidgas till att varje operator $T \in \mathcal{L}(V)$ kan framställas som produkten av en isometri och en positiv operator $\sqrt{T^*T}$.

Sats 6.57. Polära uppdelningen: *Om $T \in \mathcal{L}(V)$ så finns det en isometri $S \in \mathcal{L}(V)$ sådan att*

$$T = S\,\sqrt{T^*T}.$$

Bevis: Antag att $T \in \mathcal{L}(V)$. Om $v \in V$ så gäller

$$\begin{aligned}\|T\,v\|^2 &= \langle T\,v, T\,v\rangle = \langle T^*Tv, v\rangle \\ &= \langle \sqrt{T^*T}\,\sqrt{T^*T}\,v, v\rangle = \langle \sqrt{T^*T}\,v, \sqrt{T^*T}\,v\rangle \\ &= \|\sqrt{T^*T}\,v\|^2,\end{aligned}$$

ty T^*T är en positiv operator och har därmed en positiv kvadratrot $\sqrt{T^*T}$ (Sats 6.46). Därmed gäller

$$\|T\,v\| = \|\sqrt{T^*T}\,v\| \tag{6.8}$$

för alla $v \in V$. Definiera nu $S_1 \in \mathcal{L}(R(\sqrt{T^*T}), R(T))$ genom

$$S_1(\sqrt{T^*T}\,v) = T\,v\,. \tag{6.9}$$

Vi skall utvidga S_1 till en isometri $S \in \mathcal{L}(V)$ sådan att $T = S\sqrt{T^*T}$.

Är S_1 väldefinierad? Låt $v_1, v_2 \in V$ och antag att $\sqrt{T^*T}\,v_1 = \sqrt{T^*T}\,v_2$. Då gäller det att

$$\begin{aligned}\|T\,v_1 - T\,v_2\| &= \|T(v_1 - v_2)\| = \|\sqrt{T^*T}\,(v_1 - v_2)\| \\ &= \|\sqrt{T^*T}\,v_1 - \sqrt{T^*T}\,v_2\| \\ &= 0\,,\end{aligned}$$

där den andra likheten följer ur (6.8). Därmed är $T\,v_1 = T\,v_2$, så S_1 är väldefinierad. Att S_1 är en linjär avbildning är enkelt att kontrollera. Från (6.9) ser vi att S_1 är en surjektiv avbildning på $R(T)$. Vidare ger (6.8) och (6.9) att att $\|S_1\,u\| = \|u\|$ för alla $u \in R(\sqrt{T^*T})$, och därmed är $N(S_1) = \{0\}$, så S_1 är injektiv. Då ger dimensionssatsen att

$$\begin{aligned}\dim R(\sqrt{T^*T}) &= \dim R(S_1) + \dim N(S_1) \\ &= \dim R(T)\,,\end{aligned}$$

vilket betyder att $\dim R(\sqrt{T^*T})^\perp = \dim R(T)^\perp$, (ty $R(\sqrt{T^*T}) \oplus R(\sqrt{T^*T})^\perp = V = R(T) \oplus R(T)^\perp$) . Om $R(T)^\perp = \{0\}$ och $R(T) = V$, så är satsen bevisad. Antag därför att $R(T)^\perp \neq \{0\}$. Vi kan då välja ortonormerade baser $e_1, ..., e_m$ och $f_1, ..., f_m$ för $R(\sqrt{T^*T})^\perp$ respektive $R(T)^\perp$. Definiera nu $S_2 \in \mathcal{L}(R(\sqrt{T^*T})^\perp, R(T)^\perp)$ genom

$$S_2(a_1\,e_1 + ... + a_m\,e_m) = a_1\,f_1 + ... + a_m\,f_m\,.$$

Då är $\|S_2 w\| = \|w\|$ för alla $w \in R(\sqrt{T^*T})^\perp$.

Varje $v \in V$ kan entydigt skrivas i formen

$$v = u + w\,, \tag{6.10}$$

där $u \in R(\sqrt{T^*T})$ och $w \in R(\sqrt{T^*T})^\perp$. Vi definierar $S \in \mathcal{L}(V)$ för alla $v \in V$ genom

$$S\,v = S_1 u + S_2 w\,,$$

där u och w ges av (6.10). För varje $v \in V$ gäller:

$$S(\sqrt{T^*T}\,v) = S_1(\sqrt{T^*T}\,v) = T\,v\,,$$

så $T = S\sqrt{T^*T}$. Återstår att visa att S är en isometri. Om v ges av (6.10) så tillämpar vi Pythagoras sats på $S_1 u \in R(T)$ och $S_2 w \in R(T)^\perp$ och erhåller

$$\begin{aligned}\|S\,v\|^2 &= \|S_1 u + S_2 w\|^2 = \|S_1 u\|^2 + \|S_2 w\|^2 \\ &= \|u\|^2 + \|w\|^2 \\ &= \|v\|^2\,,\end{aligned}$$

så S är en isometri och satsen är bevisad.

Antag att $T = S\sqrt{T^*T}$ är den polära uppdelningen av $T \in \mathcal{L}(V)$. Då finns det en ortonormerad bas för V sådan att $\mathcal{M}(S)$ är diagonal om $K = C$ eller en block diagonal med block av dimension 1×1 eller 2×2 om $K = R$. Vidare finns det en ortonormerad bas för V sådan att $\mathcal{M}(\sqrt{T^*T})$ är diagonal. Men observera att det inte behöver finnas en ortonormerad bas som samtidigt ger dessa trevliga matrisframställningar för både S och $\sqrt{T^*T}$.

Exempel 6.58. Antag att A är en $n \times n$ matris och att $K = C$. Låt $V = C^n$ vara försedd med den naturliga ortonormerade basen och den euklidiska skalära produkten. Då gäller $\mathcal{M}(v) = v$ för alla $v \in V$. Det finns då en operator $T \in \mathcal{L}(V)$ sådan att $\mathcal{M}(T) = A$. Då ger Sats 6.57 att

$$T = S\sqrt{T^*T},$$

där $S \in \mathcal{L}(V)$ är en isometri och $\sqrt{T^*T}$ är en positiv operator. Vi har alltså att

$$A = \mathcal{M}(T) = \mathcal{M}(S)\,\mathcal{M}(\sqrt{T^*T}).$$

För $v \in V$ gäller $\langle \mathcal{M}(\sqrt{T^*T})\,v, v\rangle = \langle \mathcal{M}(\sqrt{T^*T})\,\mathcal{M}(v), v\rangle = \langle \mathcal{M}(\sqrt{T^*T}\,v), v\rangle = \langle \sqrt{T^*T}\,v, v\rangle \geq 0$. Vidare är $\mathcal{M}(\sqrt{T^*T})^* = \mathcal{M}((\sqrt{T^*T})^*) = \mathcal{M}(\sqrt{T^*T})$, ty $\sqrt{T^*T}$ är självadjungerad. Därmed är $\mathcal{M}(\sqrt{T^*T})$ en positivt semidefinit matris.

För $\mathcal{M}(S)$ gäller $\mathcal{M}(S)^*\mathcal{M}(S) = \mathcal{M}(S^*)\,\mathcal{M}(S) = \mathcal{M}(S^*S) = \mathcal{M}(I) = I = \mathcal{M}(I) = \mathcal{M}(S\,S^*) = \mathcal{M}(S)\,\mathcal{M}(S)^*$, så $\mathcal{M}(S)$ är en unitär matris.

Vi ser därmed att varje $n \times n$ matris A har en polär framställning

$$A = U\,P$$

där U är en unitär matris (en isometri) och P är en positivt semidefinit matris. Om A är reell så är U och P reella, och U är därmed en ortogonal matris.

Singulärvärdesuppdelningen

Definition 6.59. *Antag att $T \in \mathcal{L}(V)$. Singulära värdena till T definieras som egenvärdena till den positiva operatorn $\sqrt{T^*T}$, där varje egenvärde λ upprepas* $\dim N(\sqrt{T^*T} - \lambda\, I)$ *gånger.*

Anmärkning 6.60. a) Singulära värdena till $T \in \mathcal{L}(V)$ är reella och icke-negativa, ty $\sqrt{T^*T}$ är en positiv operator.

b) Varje $T \in \mathcal{L}(V)$ har $\dim V$ singulära värden, ty eftersom $\sqrt{T^*T}$ är självadjungerad ger spektralsatserna att V har en ortonormerad bas av egenvektorer till $\sqrt{T^*T}$. Vidare ger Sats 4.22 (e) att

$$\dim V = \dim N(\sqrt{T^*T} - \lambda_1\, I) + \ldots + \dim N(\sqrt{T^*T} - \lambda_m\, I),$$

där $\lambda_1, ..., \lambda_m$ är de olika egenvärdena till $\sqrt{T^*T}$.

Exempel 6.61. Antag att K^4 är försedd med den naturliga ortonormerade basen och låt $T \in \mathcal{L}(V)$ vara given av

$$T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4).$$

För $(w_1, w_2, w_3, w_4) \in K^4$ gäller

$$\begin{aligned}\langle T(z_1, z_2, z_3, z_4), (w_1, w_2, w_3, w_4)\rangle &= 3z_1\overline{w}_2 + 2z_2\overline{w}_3 - 3z_4\overline{w}_4 \\ &= \langle (z_1, z_2, z_3, z_4), (3w_2, 2w_3, 0, -3w_4)\rangle,\end{aligned}$$

så $T^*(w_1, w_2, w_3, w_4) = (3w_2, 2w_3, 0, -3w_4)$. Då är

$$\begin{aligned}T^*T(z_1, z_2, z_3, z_4) &= T^*(0, 3z_1, 2z_2, -3z_4) \\ &= (9z_1, 4z_2, 0, 9z_4).\end{aligned}$$

Sätt $S(z_1, z_2, z_3, z_4) = (3z_1, 2z_2, 0, 3z_4)$. Då är $S^2 = T^*T$ (Kolla). Vidare är $\mathcal{M}(S)$ en diagonal matris med diagonalelementen $3, 2, 0, 3$. Då är $\mathcal{M}(S^*) = \mathcal{M}(S)^* = \mathcal{M}(S)$, så S är en självadjungerad operator med reella icke-negativa egenvärden $0, 2, 3$. Då är S en positiv operator (Sats 6.45) och därmed är S den entydigt bestämda positiva kvadratroten till T^*T, alltså $S = \sqrt{T^*T}$.

Man ser lätt att $\dim N(\sqrt{T^*T} - 3I) = 2$, $\dim N(\sqrt{T^*T} - 2I) = 1$ och $\dim N(\sqrt{T^*T} - 0I) = 1$. Därmed ges de singulära värdena till T av $3, 3, 2, 0$.

Varje operator $T \in \mathcal{L}(V)$ har en enkel framställning med hjälp av de singulära värdena för T och två ortonormerade baser för V:

Sats 6.62. **Singulärvärdesuppdelningen:** *Antag att $T \in \mathcal{L}(V)$ har de singulära värdena $s_1, ..., s_n$. Då finns det två ortonormerade baser $e_1, ..., e_n$ och $f_1, ..., f_n$ till V sådana att*

$$Tv = s_1\langle v, e_1\rangle f_1 + ... + s_n\langle v, e_n\rangle f_n \tag{6.11}$$

för alla $v \in V$.

Bevis: Se föreläsningsanteckningar.

Anmärkning 6.63. Antag att $T \in \mathcal{L}(V)$ och att $s_1, ..., s_n$ är de singulära värdena till T. Låt $e_1, ..., e_n$ och $f_1, ..., f_n$ vara de ortonormerade baserna för V som ger framställningen (6.11). Då gäller, eftersom $Te_j = s_j\langle e_j, e_j\rangle f_j = s_jf_j$, att

$$\mathcal{M}(T, (e_1, ..., e_n), (f_1, ..., f_n)) = \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix}.$$

Med andra ord kan vi för varje $T \in \mathcal{L}(V)$ hitta två vanligen olika ortonormerade baser för V sådana att matrisframställningen med avseende på dessa är en diagonalmatris med egenvärdena till $\sqrt{T^*T}$ på diagonalen.

Exempel 6.64. Antag att A är en $n \times n$ matris med element i K. Då är A^*A en positivt semidefinit matris med reella och icke-negativa egenvärden. De singulära värdena till A definieras som kvadratrötterna av egenvärdena till A^*A.

Med stöd av Exempel 6.58 har A en polär uppdelning

$$A = Q\,P\,,$$

där P är positivt semidefinit och Q är en unitär matris (en isometri). Eftersom P är självadjungerad, och därmed normal, ger Exempel 6.35 att P har en unitär uppdelning

$$P = W^*D\,W\,,$$

där W är en unitär matris och D är en reell diagonalmatris med egenvärden till P på diagonalen. Då är $A = (Q\,W^*)\,D\,W$. Sätt $U = Q\,W^*$, vilket ger framställningen

$$A = U\,D\,W\,. \tag{6.12}$$

Nu är $U\,U^* = (Q\,W^*)\,(W\,Q^*) = I = (W\,Q^*)\,(Q\,W^*) = U^*U$, så U är en unitär matris. Vidare gäller det att

$$A^*A = (W^*D\,U^*)\,(U\,D\,W) = W^*D^2W\,,$$

vilket ger att kolonnerna i W^* är ortonormerade egenvektorer till A^*A med motsvarande icke-negativa egenvärden i diagonalen för D^2. Därmed är D en diagonal matris med de singulära värdena till A på diagonalen, och (6.12) ger en singulärvärdesuppdelning av matrisen A. (Om A är en reell matris kan ordet unitär bytas till ordet ortogonal och symbolen * kan bytas till reell transponering ^{T}).

(**Exempel.** Överkurs. För att numeriskt bestämma en singulärvärdesuppdelning av en $n \times n$ matris A kan vi använda Matlab. Kommandot

$$>>\ \ [U, D, W] = \mathrm{svd}(A)$$

ger en diagonalmatris D med de singulära värdena till A i avtagande storleksordning på diagonalen. Matriserna U och W är unitära (ortogonala i det reella fallet), och

$$A = U\,D\,W^*\,,$$

där man bör observera den endast beteckningsmässiga avvikelsen från (6.12).)

Appendix B

I detta appendix framläggs kortfattat och utan bevis Jordans normalform för komplexa matriser.

Definition. *En uppåt triangulär $k \times k$ matris $J_k(\lambda)$ av formen*

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

kallas ett Jordanblock. $J_1(\lambda)$ är en 1×1 matris bestående av skalären λ.

En $n \times n$ Jordan matris J är en block diagonalmatris av formen

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

där $n_1 + \dots + n_k = n$. Skalärerna $\lambda_1, \dots, \lambda_k$ behöver inte vara olika. Om $k = n$ så är J en diagonalmatris.

Sats. Jordans normalform. *Låt A vara en $n \times n$ matris med element i C. Då finns det en inverterbar $n \times n$ matris S sådan att*

$$A = SJS^{-1} = S \begin{pmatrix} J_{n_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{n_k}(\lambda_k) \end{pmatrix} S^{-1}, \qquad (*)$$

där $n_1 + \dots + n_k = n$ och $\lambda_1, \dots, \lambda_k$ är egenvärden till A som inte behöver vara olika. Om A är en reell matris med enbart reella egenvärden, så kan S väljas reell.

Anmärkning. a) Antalet k av Jordanblock i $(*)$ sammanfaller med antalet linjärt oberoende egenvektorer till J.

b) Matrisen J är diagonaliserbar om och endast om $k = n$.

c) Om λ är ett egenvärde till A och $\dim N(A - \lambda I) = m$, så finns det exakt m stycken Jordanblock $J_{n_j}(\lambda)$ associerade till egenvärdet λ.

För en detaljerad beskrivning av Jordans normalform hänvisas till läroböcker i Linjär algebra. I kursboken av Axler utreds Jordans normalform för operatorer på komplexa vektorrum i kapitel 8.