MM5012, LINJÄR ALGEBRA HT 2025

ERIC AHLQVIST

Sammanfattning. Detta dokument innehåller föreläsningsanteckningar från kursen *MM5012, Linjär Algebra* vid Stockholms Universitet, hösten 2025. Anteckningarna är till stor del baserade på boken *Linear Algebra (4th edition)* av Stephen H. Friedberg, Arnold J. Insel och Lawrence E. Spence [FIS03].

Senast uppdaterad: 7 oktober 2025

Innehåll

Introduktion

Linjär algebra är ett av de mest centrala och mångsidiga matematiska områdena, och fungerar som motorn bakom många av de tekniker vi använder dagligen:

(1) **Datorgrafik och bildbearbetning:**
- *3D-modellering och rendering:* Linjär algebra används för att representera och manipulera 3D-objekt genom transformationer som rotation, skalning och translation.
- *Bildkomprimering och bildförbättring:* Algoritmer som JPEG bygger på matristransformationer (t.ex. diskret cosinustransform), och även tekniker för filtrering, brusreducering och kantdetektion bygger på linjär algebra.

(2) **Sökmotorer:**
- *Rangordning av webbsidor:* Googles PageRank-algoritm använder egenvärden och egenvektorer för att analysera länkstrukturer och bestämma sidors relevans.
- *Dokumentrepresentation:* Metoder som TF-IDF kombinerat med vektorrumsmodeller representerar texter som vektorer, vilket gör det möjligt att använda linjär algebra för att mäta likhet och relevans mellan dokument.

(3) **Maskininlärning och AI:**
- *Datarepresentation:* Data representeras som vektorer och matriser, vilket gör linjär algebra till grunden för hur modeller tränas och används.
- *Algoritmer:* Många tekniker – som linjär regression, huvudkomponentanalys (PCA), support vector machines (SVM), neurala nätverk (inklusive språkmodeller och bildigenkänning) – bygger på linjära transformationer, matrismultiplikationer och optimering.

(4) **Ingenjörskonst:**
- *Strukturanalys:* Stora linjära system används för att analysera och designa konstruktioner som broar, byggnader och flygplan.
- *Elektronik och signalbehandling:* Linjär algebra används för att analysera elektriska kretsar och för att bearbeta signaler.
- *Kontrollsystem:* Dynamiska system, som robotar och fordon, modelleras och styrs med hjälp av tillståndsmodeller baserade på matriser.

(5) **Vetenskapliga beräkningar och forskning:**
- *System av linjära ekvationer:* Centralt för nästan alla naturvetenskapliga och tekniska tillämpningar.
- *Numeriska metoder:* Algoritmer som Gauss-eliminering och LU-faktorisering är fundamentala för beräkningar inom fysik, kemi och biologi.

- *Storskaliga simuleringar:* Väderprognoser, klimatmodeller och molekylär dynamik bygger på omfattande matriskalkyler, vilket gör linjär algebra till en kärnteknik i superdatorer och högpresterande beräkningar.

Detta är bara ett axplock av de många tillämpningarna av linjär algebra och den som läser den här kursen kan vara säker på att den använder någon tillämpning av linjär algebra varje dag.

Vad är linjär algebra? Linjär algebra är teorin om linjer, plan och högdimensionella *linjära* rum och avbildningar mellan dessa. Till skillnad från många andra matematiska områden är linjär algebra en teori som vi verkligen förstår väl, vilket är en av anledningarna till dess många tillämpningar. Ett citat från William Stein lyder: *"Matematik är konsten att reducera vilket problem som helst till linjär algebra"*.

Det kanske mest typiska problemet i linjär algebra är att lösa system av linjära ekvationer, dvs. att hitta lösningar till ekvationer av formen

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \, , \end{aligned}$$

där a_{ij} och b_i är kända konstanter och x_i är variabler. Dessa ekvationer kan skrivas på matrisform som $Ax = b$ där A är en matris med koefficienterna a_{ij}, där x är en kolumnvektor med variablerna x_i och b är en kolumnvektor med konstanterna b_i:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} .$$

Matrisen A representerar en *linjär avbildning* från $\mathbb{R}^n$ till $\mathbb{R}^m$, dvs. den avbildar vektorer i $\mathbb{R}^n$ till vektorer i $\mathbb{R}^m$. Lösningen x är en vektor i $\mathbb{R}^n$ som avbildas på vektorn b i $\mathbb{R}^m$.

I den här kursen kommer vi inte längre att begränsa oss till $\mathbb{R}^n$ utan kommer att arbeta i en mer abstrakt kontext med allmänna vektorrum över en kropp k. Vi kommer att se många exempel på vektorrum, inklusive vektorrum av oändlig dimension. Under kursens gång kommer vi dock att inse att vi faktiskt kan identifiera varje vektorrum av ändlig dimension n med k^n.

1. Vektorrum

1.1. **Definitionen av ett vektorrum.** Ett vektorrum är mängd med *mycket* struktur, vilket gör att de är lätta att arbeta med, men något krångliga att definiera. Kort sagt så är ett vektorrum en *abelsk grupp* som är en *modul* över en *kropp*. Men för att förstå vad det betyder så behöver vi först definiera dessa begrepp.

Definition 1.1. En *abelsk grupp* är en mängd A tillsammans med en funktion $+\colon A \times A \to A$ (skrivs $(a, b) \mapsto a + b$) som uppfyller följande egenskaper:

(A1) `Kommutativitet:` För alla $a, b \in A$ gäller $a + b = b + a$ (dvs *abelsk*).
(A2) `Associativitet:` För alla $a, b, c \in A$ gäller $(a + b) + c = a + (b + c)$.
(A3) `Identitet:` Det finns ett element $0 \in A$ sådant att för alla $a \in A$ gäller $a + 0 = a$.
(A4) `Invers:` För varje $a \in A$ finns ett element $-a \in A$ sådant att $a+(-a) = 0$.

Heltalen $\mathbb{Z}$, reella talen $\mathbb{R}$ och komplexa talen $\mathbb{C}$ är alla exempel på abelska grupper under addition. Notera att ingen av dessa exempel är grupper under multiplikation. Om A är en abelsk grupp så är även

$$A^n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in A\},$$

med addition komponentvis (se Exempel 1.1 längre ner), en abelsk grupp för varje positivt heltal n. Dvs $\mathbb{Z}^n$, $\mathbb{R}^n$ och $\mathbb{C}^n$ är alla exempel på abelska grupper under komponentvis addition.

Sats 1.1 (Cancellationslagen)**.** *Låt A vara en abelsk grupp. Om $a, b, c \in A$ och $a + b = a + c$, så gäller att $b = c$.*

Bevis. Vi har $b = (-a) + (a + b) = (-a) + (a + c) = c$. □

Diskussionsfrågor

(1) Vilka av axiomen (A1)-(A4) använde vi i beviset?
(2) Kan du ge ett exempel på en grupp som ej är kommutativ?

Följdsats 1.2 (Unik identitet och invers)**.** *Elementet* 0 *i* $(A3)$ *och elementet* $-a$ *i* $(A4)$ *är båda unika. Det finns alltså endast ett nollelement och endast en invers till varje element.*

Notera att vi hittills inte har utnyttjat axiomet (A1), dvs. [Cancellationslagen] (Sats 1.1) och [Unik identitet och invers] (Följdsats 1.2) gäller även för icke-abelska grupper.

I vår nästa definition kommer vi att axiomatisera egenskaperna hos $\mathbb{Q}, \mathbb{R}$, $\mathbb{C}$ och $\mathbb{F}_p$.

Definition 1.2. En *kropp* är en mängd k tillsammans med två funktioner $k \times k \to k$ som vi kallar *addition* (skrivs $(a, b) \mapsto a + b$) och *multiplikation* (skrivs $(a, b) \mapsto ab$) som uppfyller följande egenskaper:

(A) k är en abelsk grupp under addition,
(B) $k \setminus \{0\}$ är en abelsk grupp under multiplikation (identiteten kallar vi 1),
(D) `Distributivitet:` För alla $a, b, c \in k$ gäller $a(b + c) = ab + ac$.

I den här kursen kommer vi mestadels att arbeta med kropparna $\mathbb{R}$, $\mathbb{C}$.

Diskussionsfrågor

(1) Är $\mathbb{Z}$ en kropp?
(2) Är $\mathbb{R}^2$ en kropp?

Nu ska vi definiera vektorrum. Poängen är att försöka kapsla in egenskaperna hos $\mathbb{R}^n$ och $\mathbb{C}^n$, men som vi kommer att se så innefattar begreppet vektorrum många fler exempel än dessa.

Definition 1.3. Ett *vektorrum* över en kropp k är en mängd V tillsammans med två funktioner: *addition* $+\colon V \times V \to V$ (skrivs $(v, w) \mapsto v + w$) och *skalärmultiplikation* $k \times V \to V$ (skrivs $(a, v) \mapsto av$), som uppfyller följande egenskaper:

(A) V är en *abelsk grupp* under addition, dvs:

(A1) `Kommutativitet:` För alla $v_1, v_2 \in V$ gäller $v_1 + v_2 = v_2 + v_1$.

(A2) `Associativitet:` För alla $v_1, v_2, v_3 \in V$ gäller $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$.

(A3) `Identitet`: Det finns ett element $0 \in V$ sådant att för alla $v \in V$ gäller $v + 0 = v$.

(A4) `Invers`: För varje $v \in V$ finns ett element $-v \in V$ sådant att $v + (-v) = 0$.

(M) V är en *modul* över k i den mening att, utöver (A) så har vi att:

(M1) `Identitet`: För alla $v \in V$ gäller $1v = v$.

(M2) `Associativitet`: För alla $a, b \in k$ och $v \in V$ gäller $a(bv) = (ab)v$.

(M3) `Distributivitet-1`: För alla $a \in k$ och $v_1, v_2 \in V$ gäller $a(v_1 + v_2) = av_1 + av_2$.

(M4) `Distributivitet-2`: För alla $a, b \in k$ och $v \in V$ gäller $(a + b)v = av + bv$.

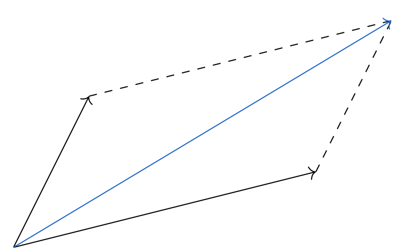
Elementen i ett vektorrum brukar kallas för *vektorer* eller *punkter*.

Ett vektorrum är alltså en *abelsk grupp som är en modul över en kropp*. Den som finner det hjälpsamt kan minnas axiomen som

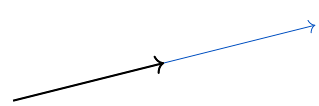
`KAIIIADD`, `K-AIIIA-DD`, eller `KAII-IADD`,

där de första fyra är axiomen för en abelsk grupp och de fyra sista är de ytterligare axiom som gör V till en modul över k, dvs. en *modul* är en abelsk grupp som uppfyller (M). Notera att de första fyra axiomen inte nämner k.

Vektoraddition i $\mathbb{R}^n$ kan visualiseras som en parallellogram, där $v_1 + v_2$ är den diagonala linjen i parallellogrammen med hörn 0, v_1 och v_2:



Skalärmultiplikation i $\mathbb{R}^n$ kan visualiseras som en förlängning/förkortning av en vektor, där av är vektorn v multiplicerad med skalären a:



Sats 1.3 (Identitetselement)**.** *Låt V vara ett vektorrum över en kropp K och låt $a \in k$ och $v \in V$. Då gäller att*

(1) $0v = 0$,
(2) $(-1)v = -v$,
(3) $a(0) = 0$,
(4) om $av = 0$ *så har vi* $a = 0$ *eller* $v = 0$.

Bevis. (1): Vi har av (M4) att $v = (1+0)v = 1v + 0v = v + 0v$ och av [Cancellationslagen] (Sats 1.1) följer det att $0v = 0$. (2): Nu ser vi att $0 = 0v = (1 + -(1))v = 1v+(-1)v = v+(-1)v$, dvs $-v = (-1)v$. (3): Vi har $a(0) = a(v+(-v)) = av+a(-v) = av + (a(-1))v = av + (-1)av = av + (-(av)) = 0$. (4): Slutligen, om $av = 0$ och $a \neq 0$ så har vi $v = 1v = (a^{-1}a)v = a^{-1}av = a^{-1}0 = 0$. □

Exempel 1.1. Det mest typiska exemplet på ett vektorrum är k^n där $n \in \mathbb{N}$ och k är en kropp, t.ex. $\mathbb{R}$ eller $\mathbb{C}$. Här är vektorerna ordnade n-tupler $v = (v_1, \dots, v_n)$ där $v_i \in k$. Vektoraddition och skalärmultiplikation definieras komponentvis:

$$v + w = (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n), \quad av = (av_1, av_2, \dots, av_n)\,.$$

Vektorrum kan se ut på olika sätt och det finns många exempel på vektorrum som inte är lika konkreta som $\mathbb{R}^n$. Här är några exempel:

Exempel 1.2. Låt k vara en kropp och M en godtycklig mängd. Då är mängden k^M av alla funktioner från M till k ett vektorrum över k. Vektoraddition och skalärmultiplikation definieras punktvis: $(f_1 + f_2)(m) = f_1(m) + f_2(m), \quad (af)(m) = af(m)$. Notera att $\mathbb{R}^n$ är ett specialfall av detta eftersom vi kan identifiera $\mathbb{R}^n$ med mängden av alla funktioner från $\{1, \dots, n\}$ till $\mathbb{R}$.

Exempel 1.3. Ett specialfall av föregående exempel är $k^{\mathbb{N}}$ som kan identifieras med mängden av *potensserier*

$$k[[x]] := \left\{ \sum_{n \in \mathbb{N}} a_n x^n : a_n \in k \right\}.$$

Vektoraddition och skalärmultiplikation definieras genom

$$\left(\sum_{n\in\mathbb{N}} a_n x^n\right) + \left(\sum_{n\in\mathbb{N}} b_n x^n\right) = \sum_{n\in\mathbb{N}} (a_n + b_n) x^n, \quad a\left(\sum_{n\in\mathbb{N}} c_n x^n\right) = \sum_{n\in\mathbb{N}} (ac_n) x^n\,.$$

Diskussionsfrågor

(1) Om $V_1, \dots V_n$ är vektorrum över en kropp k, är då $V_1 \times \cdots \times V_n$ ett vektorrum över k?
(2) Är mängden av alla 2×2 matriser med reella tal som element (addition och skalärmultiplikation komponentvis) ett vektorrum över $\mathbb{R}$?
(3) Är $\mathbb{C}$ ett vektorrum över $\mathbb{R}$?

1.2. **Delrum.** När vi betraktar delmängder av vektorrum så är det ofta användbart att veta när en delmängd är ett vektorrum i sig. Det leder oss till definitionen av ett *delrum*:

Definition 1.4. Ett *delrum* W av ett vektorrum V över en kropp k är en delmängd av V som är ett vektorrum över k med samma operationer som V, dvs

(1) W är slutet under addition: $w_1 + w_2 \in W$ om $w_1, w_2 \in W$,
(2) W är slutet under skalärmultiplikation: $aw \in W$ om $w \in W$ och $a \in k$,
(3) W tillsammans med addition och skalärmultiplikation uppfyller alla 8 axiomen för ett vektorrum.

Diskussionsfrågor

(1) Är $\{v = (v_1, v_2, v_3) \in \mathbb{R}^3 \mid 2v_1 - v_2 + 5v_3 = 0\}$ ett delrum av $\mathbb{R}^3$?
(2) Är $\{v = (v_1, v_2, v_3) \in \mathbb{R}^3 \mid 2v_1 - v_2 + 5v_3 = 1\}$ ett delrum av $\mathbb{R}^3$?
(3) Är $\{v = (v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1^2 + v_2^2 + v_3^2 = 1\}$ ett delrum av $\mathbb{R}^3$?
(4) Är $\{v = (v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1^2 + v_2^2 + v_3^2 = 0\}$ ett delrum av $\mathbb{R}^3$?
(5) För vilka $a \in \mathbb{R}$ är $\{f \colon \mathbb{R} \to \mathbb{R} \mid f(0) = a\}$ ett delrum av $\mathbb{R}^{\mathbb{R}}$ (Exempel 1.2)?

Exempel 1.4. Mängden av *polynom* med koefficienter i k,

$$k[x] = \left\{ \sum a_n x^n : a_n \in k, a_n = 0 \text{ för alla utom ändligt många } n \right\},$$

är ett delrum av $k[[x]]$ (Exempel 1.3). Mängden av *polynom av grad högst* n är ett delrum av $k[x]$:

$$k[x]_{\leq n} = \left\{ \sum_{i=0}^{n} a_i x^i : a_i \in k \right\}.$$

Exempel 1.5. Mängden av alla *kontinuerliga* funktioner $f \colon \mathbb{R} \to \mathbb{R}$ är ett delrum av $\mathbb{R}^{\mathbb{R}}$ (Exempel 1.2).

Nu ska vi visa att man faktiskt inte behöver kontrollera alla 8 axiomen för att verifiera att en delmängd är ett delrum. Det räcker med att kontrollera tre egenskaper, vilket gör det mycket enklare att avgöra om en delmängd är ett delrum eller inte.

Sats 1.4 (Delrumstest)**.** *Låt V vara ett vektorrum och $W \subseteq V$ vara en delmängd. Då är W ett delrum av V om och endast om W har följande tre egenskaper:*

(1) Innehåller nollvektorn: $0 \in W$ där 0 är nollvektorn i V,
(2) Sluten under addition: $v_1, v_2 \in W \Rightarrow v_1 + v_2 \in W$,
(3) Sluten under skalärmultiplikation: $v \in W$, $a \in K \Rightarrow av \in W$.

Bevis. Anta att W är ett delrum av V. Då håller (2) and (3) per definition och det återstår att verifiera att $0 \in W$. Låt 0_W vara nollvektorn i W. Eftersom $0_W = 0_W + 0_W$ i W (och då även i V) så har vi att $0 + 0_W = 0_W = 0_W + 0_W$ i V. Av cancellationslagen får vi att $0_W = 0$.

Å andra sidan, låt W är en delmängd av V som uppfyller (1), (2) och (3). Notera att $(A1)$, $(A2)$, $(M1)$, $(M2)$, $(M3)$ och $(M4)$ är uppfyllda i W eftersom de är uppfyllda i V. Vi behöver nu visa att $(A3)$ och $(A4)$ är uppfyllda i W. Vi ser direkt att $(A3)$ är uppfyllt eftersom $0 \in W$. Att (A4) är uppfyllt följer av (3): om $w \in W$ så har vi $(-1)w = -w \in W$. □

Diskussionsfrågor

(1) Är delmängden $\{(a_1, 1, a_3) : a_1, a_3 \in k\} \subseteq k^3$ ett delrum av k^3?
(2) Är delmängden $\{ax + bx^2 : a, b \in k\} \subseteq k[x]$ ett delrum av $k[x]$?
(3) Är unionen av två delrum ett delrum?
(4) Är snittet av två delrum ett delrum?
(5) Är mängden av alla polynom i $k[x]$ med grad exakt n ett delrum av $k[x]$?

2. Bas, linjärt oberoende och dimension

2.1. **Linjärkombination och linjärt hölje.** När man arbetar med ett konkret vektorrum är det ofta möjligt att hitta en liten mängd vektorer som kan kombineras på olika sätt för att generera hela rummet eller ett specifikt delrum. Detta görs genom att använda linjärkombinationer av vektorer. En linjärkombination är en summa av vektorer där varje vektor multipliceras med en skalär från kroppens element.

Definition 2.1. Låt V vara ett vektorrum över en kropp k. En *linjärkombination* av vektorerna $v_1, \dots, v_n \in V$ $(n \in \mathbb{N})$ är ett uttryck av formen

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n\,,$$

där $a_i \in k$ för alla i.

Exempel 2.1.

(1) Varje element i $\mathbb{R}^2$ är en linjärkombination av vektorerna $(1, 0)$ och $(0, 1)$.
(2) Varje element i $\mathbb{R}^2$ är en linjärkombination av vektorerna $(1, 2)$ och $(1, -1)$.
(3) Varje element i $k[x]$ (Exempel 1.4) är en linjärkombination av vektorerna $1, x, x^2, x^3, \dots$.

Notera att en linjärkombination byggs upp av ett *ändligt* antal vektorer men att vi inte kräver att dessa är unika, dvs att vi kan ha $v_i = v_j$ för vissa $i \neq j$.

Definition 2.2. Låt V vara ett vektorrum över en kropp k och låt $S \subseteq V$ vara en delmängd. Mängden av alla linjärkombinationer av vektorerna i S kallas det *linjära höljet* (eller *spannet*) av S och betecknas $\mathrm{Span}(S)$. Om $\mathrm{Span}(S) = V$ så sägs att S *spänner upp* eller *genererar* V. Om $S = \{v_1, \dots, v_n\}$ så skriver vi $\mathrm{Span}(v_1, \dots, v_n)$ istället för $\mathrm{Span}(\{v_1, \dots, v_n\})$. Vi definierar även $\mathrm{Span}(\emptyset) = \{0\}$ där 0 är nollvektorn i V.

Exempel 2.2. Låt $V = \mathbb{R}^3$ och låt $S = \{(1, 0, 0), (0, 1, 0)\}$. Då är $\mathrm{Span}(S)$ mängden av alla vektorer av formen $(a_1, a_2, 0)$ där $a_1, a_2 \in \mathbb{R}$. Detta är ett plan i $\mathbb{R}^3$ som innehåller z-axeln. Om vi lägger till vektorn $(0,0,1)$ i S, så att $S' = \{(1,0,0),(0,1,0),(0,0,1)\}$ så är $\operatorname{Span}(S') = \mathbb{R}^3$.

Exempel 2.3. Låt $V = k[x]$ vara mängden av alla polynom med koefficienter i en kropp k. Låt $S = \{1, x+103, x^2\}$. Då är $\operatorname{Span}(S) = K[x]_{\leq 2}$: mängden av alla polynom av grad högst 2, dvs alla polynom av formen $a_0 + a_1x + a_2x^2$ där $a_i \in k$.

Sats 2.1 (Karaktärisering av Span)**.** *Låt V vara ett vektorrum över en kropp k och låt $S \subseteq V$ vara en delmängd. Då är* $\operatorname{Span}(S)$ *ett delrum av V. Mer specifikt,* $\operatorname{Span}(S)$ *det minsta delrummet av V som innehåller S, det vill säga, om W är ett delrum av V som innehåller S, så har vi att* $\operatorname{Span}(S) \subseteq W$*.*

Bevis. Vi måste visa att $\operatorname{Span}(S)$ är ett delrum av V. För att göra detta, måste vi visa att $\operatorname{Span}(S)$ är sluten under addition och skalärmultiplikation, samt att den innehåller nollvektorn. Detta är trivialt om $S = \emptyset$ så vi antar att $S \neq \emptyset$.

- Nollvektorn: Låt $v \in S$. Då är $0 \cdot v = 0$ en linjärkombination av vektorer i S, så nollvektorn är i $\operatorname{Span}(S)$.
- Sluten under addition: Om $v = a_1v_1 + \cdots + a_nv_n$ och $w = b_1w_1 + \cdots + b_mw_m$ är linjärkombinationer av vektorer i S, så är

$$v + w = a_1v_1 + \cdots + a_nv_n + b_1w_1 + \cdots + b_mw_m$$

  en linjärkombination av vektorer i S.
- Sluten under skalärmultiplikation: Likt ovan, om $c \in k$ så har vi

$$cv = c(a_1v_1 + \cdots + a_nv_n) = ca_1v_1 + \cdots + ca_nv_n$$

  vilket är en linjärkombination av vektorer i S.

Antag nu att $S \subseteq W$ där W är ett delrum av V. Givet en linjärkombination $a_1v_1 + \cdots + a_nv_n$ där $v_i \in S$ för alla i, så är varje term $a_iv_i \in W$ eftersom W är sluten under skalärmultiplikation och W innehåller S. Därmed är summan $a_1v_1 + \cdots + a_nv_n \in W$ eftersom W är sluten under addition. Detta visar att $\operatorname{Span}(S) \subseteq W$. $\square$

Diskussionsfrågor

(1) Om $v_1 = (1,1,1)$, $v_2 = (1,4,2)$ och $v_3 = (-1,2,0)$ är vektorer i $\mathbb{R}^3$, kan varje vektor i $\mathbb{R}^3$ skrivas som en linjärkombination av v_1, v_2 och v_3? Vad är $\mathrm{Span}(v_1, v_2, v_3)$?
(2) Kan du hitta två vektorer $v_1, v_2 \in k[x]_{\leq 2}$ så att $\mathrm{Span}(v_1, v_2) = \{p(x) \in k[x] \mid p(0) = 0\}$?

2.2. **Linjärt oberoende, bas och dimension.** Givet en mängd S med vektorer som spänner upp ett vektorrum $\mathrm{Span}(S)$, kan vi fråga oss om det finns några redundanta vektorer i mängden. Om en vektor v kan skrivas som en linjärkombination av de andra vektorerna i S, så är den redundant och om vi tar bort denna vektor så förändrar vi inte spannet: $\mathrm{Span}(S \setminus \{v\}) = \mathrm{Span}(S)$. Detta leder oss till begreppet *linjärt oberoende*.

Definition 2.3. Låt V vara ett vektorrum över en kropp k och låt $S \subseteq V$ vara en delmängd. Vi säger att S är *linjärt oberoende* om den enda linjärkombinationen av vektorerna i S som ger nollvektorn är den triviala kombinationen där alla skalärer är noll, det vill säga:

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \Rightarrow a_i = 0 \text{ för alla } i\,.$$

Om S inte är linjärt oberoende så sägs den vara *linjärt beroende*.

Notera att vi endast betraktar linjärkombinationer av ett *ändligt* antal vektorer i S även om mängden S är oändlig.

Exempel 2.4. Låt $V = \mathbb{R}^3$ och låt $S = \{(1,0,0), (0,1,0), (0,0,1)\}$. Då är S linjärt oberoende eftersom den enda linjärkombinationen som ger nollvektorn är den triviala kombinationen där alla skalärer är noll. Om vi lägger till vektorn $(1,1,1)$ i S, så att $S' = \{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$ så är S' linjärt beroende eftersom $(1,1,1) = (1,0,0) + (0,1,0) + (0,0,1)$.

Definition 2.4. Låt V vara ett vektorrum över en kropp k. En (ordnad) *bas* för V är en (ordnad) delmängd $\beta \subseteq V$ som är både linjärt oberoende och spänner upp hela rummet, det vill säga $\mathrm{Span}(\beta) = V$ och β är linjärt oberoende.

Poängen med en (ändlig) *ordnad* bas för ett vektorrum V är att den hjälper oss att identifiera V med k^n. Detta kommer att bli tydligt i Kapitel 4.

Exempel 2.5. Låt $V = \mathbb{R}^3$. Då är mängden $\beta = \{(1,1,0),(0,1,1),(0,1,0)\}$ en bas för V eftersom den är linjärt oberoende och spänner upp hela rummet.

Exempel 2.6. Låt $V = k[x]$ vara mängden av alla polynom med koefficienter i en kropp k. Då är mängden $\beta = \{1, x, x^2, \dots\}$ en bas för V eftersom den är linjärt oberoende och spänner upp hela rummet. Notera att β *inte* är en bas för $k[[x]]$ eftersom elementet $1 + x + x^2 + \dots$ *ej* kan uttryckas som en *ändlig* linjärkombination av element i β.

Definition 2.5 (Standardbasen)**.** Låt k vara en kropp. Delmängden av k^n som består av vektorerna

$$e_1 = (1, 0, \dots, 0),\ e_2 = (0, 1, \dots, 0),\ \dots,\ e_n = (0, 0, \dots, 1)$$

kallas för *standardbasen* för k^n.

Definition 2.6. Ett vektorrum V är av (eller har) *ändlig dimension* om det finns en bas β för V med ändligt många element. Om inget sådant bas finns, sägs V vara av *oändlig dimension*. Vi säger att V har *uppräknelig dimension* om det finns en uppräknelig bas för V.

Sats 2.2 (Karaktärisering av bas)**.** *Låt V vara ett vektorrum över en kropp k. En delmängd $\beta \subseteq V$ är en bas för V om och endast om varje vektor i V kan uttryckas som en linjärkombination av vektorerna i β på ett unikt sätt.*

Bevis. Antag första att β utgör en bas för V. Då är $\mathrm{Span}(\beta) = V$ och varje vektor $v \in V$ kan skrivas som en linjärkombination av vektorerna i β:

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n\,.$$

Antag nu att $v = b_1 w_1 + b_2 w_2 + \cdots + b_m w_n$ är en annan linjärkombination av vektorerna i β. Genom att lägga till $0 \cdot v_1, 0 \cdot v_2$ till den andra linjär kombinationen och vice versa så kan vi anta att $n = m$ och $w_i = v_i$ för alla $1 \leq i \leq n$. Då har vi att

$$\begin{aligned} 0 &= v - v \\ &= (a_1v_1 + a_2v_2 + \cdots + a_nv_n) - (b_1v_1 + b_2v_2 + \cdots + b_nv_n) \\ &= (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n \,. \end{aligned}$$

Eftersom β är linjärt oberoende måste vi då ha $a_i = b_j$ för alla i. Därmed är varje vektor i V en unik linjärkombination av vektorerna i β.

Omvänt, antag att varje vektor i V kan skrivas som en linjärkombination av vektorerna i β på ett unikt sätt. Då är det klart att $\mathrm{Span}(\beta) = V$ och det återstår att visa att β är linjärt oberoende. Antag motsatsen, dvs att β är linjärt beroende. Då finns det skalärer $a_1, \ldots, a_n$ som inte alla är noll, och vektorer $v_1, \ldots, v_n \in \beta$ sådana att

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0\,.$$

Detta innebär att vi kan skriva nollvektorn som en linjärkombination av vektorerna i β på mer än ett sätt, vilket strider mot antagandet att varje vektor i V kan skrivas som en linjärkombination av vektorerna i β på ett unikt sätt. Därmed är β linjärt oberoende. $\square$

Notera att denna karaktärisering av bas säger att om vi har en ordnad ändlig bas $\beta = \{v_1, v_2, \ldots, v_n\}$ för ett vektorrum V, så kan vi representera varje element i V som en kolumnvektor efter att vi uttryckt den (på det unika sättet) som en linjärkombination av element i β:

$$v = a_1v_1 + a_2v_2 + a_3v_3 \cdots + a_nv_n \mapsto \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Detta oss ett sätt att identifiera V med k^n. Vi kommer att göra detta påstående mer precist i Kapitel 4 (Inverterbara avbildningar och basbyte).

Definition 2.7. Låt V vara ett vektorrum över en kropp k och antag att det finns en (ordnad) bas $\beta = \{v_1, v_2, \ldots\}$ för V. Vi definierar *koordinatvektorn* av $v \in V$ med avseende på basen β som kolumnvektorn

$$[v]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix},$$

där $v = a_1v_1 + a_2v_2 + \dots$ är den unika linjärkombinationen av vektorer i β som ger v (Sats 2.2 [Karaktärisering av bas]).

Om β är ändlig med n element: $\beta = \{v_1, v_2, \dots, v_n\}$, så kan vi identifiera $[v]_\beta$ med ett element i k^n. Om β är oändlig så kan vi identifiera $[v]_\beta$ med ett element i $\{f \in k^{\mathbb{N}} : f(n) = 0 \text{ för alla utom ändligt många } n \in \mathbb{N}\}$.

Diskussionsfrågor

(1) Vad är koordinatvektorn till $(1, 0, 0)$ i basen $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \subseteq \mathbb{R}^3$?
(2) Vad är koordinatvektorn till $(1, 0)$ i basen $\{(1, 1), (1, -1)\} \subseteq \mathbb{R}^2$?
(3) Vad är koordinatvektorn till $1 + x + x^2$ i basen $\{1, x, x^2\} \subseteq \mathbb{R}[x]$?
(4) Vad är koordinatvektorn till 1 i basen $\{1 + x, x + x^2, 1 + x^2\} \subseteq \mathbb{R}[x]$?
(5) Vad är koordinatvektorn till $i = \sqrt{-1}$ i basen $\{1 + i, 1 - i\} \subseteq \mathbb{C}$ som vektorrum över $\mathbb{R}$?

Sats 2.3 (Maximalt antal linjärt oberoende vektorer)**.** *Låt V vara ett vektorrum av ändlig dimension över en kropp k. Låt $\beta = \{v_1, \dots, v_n\}$ vara en bas för V och låt $\{w_1, \dots, w_m\} \subseteq V$ vara en en linjärt oberoende delmängd. Då gäller att $m \leq n$.*

Bevis. Låt $\beta_0 = \beta$. Eftersom $w_1 \in \mathrm{Span}(\beta)$ så finns det ett index $1 \leq i \leq n$ så att $v_i \in \mathrm{Span}(w_1, v_1, v_2, \dots, v_{i-1}, v_{i+1}, v_{i+2}, \dots, v_n)$. Låt

$$\beta_1 = \{w_1, v_1, v_2, \dots, v_{i-1}, v_{i+1}, v_{i+2}, \dots, v_n\}.$$

Då är β_1 en bas för V eftersom $\mathrm{Span}(\beta_1) = \mathrm{Span}(\beta) = V$ och β_1 är linjärt oberoende, vilket följer av att $\{v_1, v_2, \dots, v_{i-1}, v_{i+1}, v_{i+2}, \dots, v_n\}$ är linjärt oberoende och w_1 är en linjärkombination av $v_1, v_2, \dots, v_i$. Mycket riktigt, om β_1 skulle vara linjärt beroende så skulle w_1 kunna uttryckas som en linjärkombination av vektorer i β på två olika sätt, vilket av Sats 2.2 [Karaktärisering av bas] är en motsägelse.

För varje $2 \leq i \leq m$, lägg till w_i (som första vektor i en ordnad mängd likt fallet $i = 1$ ovan) till den ordnade basen β_{i-1} och upprepa argumentet ovan. Efter m steg har vi en bas β_m för V med n element, som innehåller $w_1, \dots, w_m$. Dvs $m \leq n$. $\square$

Följdsats 2.4 (Varje bas är lika stor)**.** *Låt V vara ett vektorrum av ändlig dimension över en kropp k. Då är alla baser för V av samma kardinalitet, det vill säga om β_1 och β_2 är två baser för V, så har β_1 och β_2 samma antal element.*

Definition 2.8. Låt V vara ett vektorrum av ändlig dimension över en kropp k. Dimensionen av V, betecknad $\dim V$, är antalet element i en bas för V. Om V är av oändlig dimension, så skriver vi $\dim V = \infty$.

Sats 2.5 (Delrum och dimension)**.** *Låt $W \subseteq V$ vara ett delrum av ett vektorrum V av ändlig dimension. Då gäller följande:*

(1) För varje bas $\{v_1, \ldots, v_m\} \subseteq W$ så existerar $v_{m+1}, \ldots v_n \in V$ så att $\{v_1, \ldots, v_m, v_{m+1}, \ldots v_n\}$ är en bas för V.
(2) W är av ändlig dimension och $\dim W \leq \dim V$.
(3) Om $\dim W = \dim V$ så har vi att $W = V$.

Bevis. Notera att (2) och (3) följer av (1), så det räcker att vi visar (1): Vi har att $\{v_1, \ldots, v_m\}$ är linjärt oberoende och om $\operatorname{Span}(v_1, \ldots, v_m) \neq V$ så kan vi hitta $v_{m+1} \in V$ sådan att $\{v_1, \ldots, v_m, v_{m+1}\}$ är linjärt oberoende. Itererar vi denna process så hittar vi en bas $\{v_1, \ldots, v_m, v_{m+1}, \ldots v_n\}$ för V. □

Sats 2.6 (Existens av bas)**.** *Varje vektorrum har en bas.*

Bevis. Ändlig dimension: Vi väljer först en vektor $v_1 \in V$. Om $\operatorname{Span}(v_1) \neq V$ så väljer vi $v_2 \in V \setminus \operatorname{Span}(v_1)$. Om $\operatorname{Span}(v_1, v_2) \neq V$ så väljer vi $v_3 \in V \setminus \operatorname{Span}(v_1, v_2)$ osv. Itererar vi denna process så hittar vi en bas $\{v_1, \ldots, v_n\}$ för V har ändlig dimension.

Oändlig dimension (ej del av kursen): Detta kräver att Kuratowski–Zorns lemma, vilket säger följande: Låt S vara en partiellt ordnad mängd. Om varje totalt ordnad delmängd av S har en övre gräns i S, så finns det ett maximalt element i S. Detta bör ses som ett axiom, då det visar sig vara ekvivalent med *Urvalsaxiomet (Axiom of Choice)*.

Låt S vara mängden av alla linjärt oberoende delmängder av V. Vi ordnar S med avseende på inklusion, dvs $A \leq B$ om $A \subseteq B$. Låt $A \subseteq S$ vara en totalt ordnad delmängd av S. Vi påstår att $\bigcup_{\alpha \in A} \alpha$ är ett element i S, dvs att $\bigcup_{\alpha \in A} \alpha$ är linjärt oberoende. Mycket riktigt, antag att $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$ för $v_1, \ldots, v_n \in \bigcup_{\alpha \in A} \alpha$. Eftersom A är totalt ordnad, så finns det en $\alpha \in A$ sådan att $v_i \in \alpha$ för alla i. Eftersom α är linjärt oberoende, så måste $a_i = 0$ för alla i. Detta visar att $\bigcup_{\alpha \in A} \alpha$ är linjärt oberoende. Därmed har vi en övre gräns för A i S, nämligen $\bigcup_{\alpha \in A} \alpha$. Enligt Kuratowski–Zorns lemma finns det ett maximalt element i S, dvs en maximal linjärt oberoende delmängd $B \subseteq V$. Om $\operatorname{Span} B \neq V$ så kan vi hitta $v \in V$ så att $B \cup \{v\}$ är linjärt oberoende, vilket strider mot maximaliteten av B. Därmed är $\operatorname{Span} B = V$ och B är en bas för V. $\square$

Diskussionsfrågor

(1) Kan du ge exempel på en bas för $k^{\{0,1\}}$?
(2) Betrakta funktionen $x^2 + 1 \colon \{0, 1\} \to \mathbb{R}$ som ges av $a \mapsto a^2 + 1$. Kan du skriva $x^2 + 1$ som en linjärkombination av element i din bas från (1)?
(3) Definiera $L \colon \mathbb{R}^{\{0,1\}} \to \mathbb{R}^2$ genom att skicka f till vektorn $L(f) := (f(0), f(1))$. Vad är $L(x^2 + 1)$ om $x^2 + 1$ definieras som i (2)?
(4) Om M är en matris med $n + 1$ kolumner vilka är vektorer i k^n, vad kan man säga om $\ker M$ och $\operatorname{im} M$?
(5) Hur många baser har $k[x]_{\leq 2}$?

3. Linjära avbildningar och matriser

3.1. **Linjära avbildningar.** Det här kapitlet handlar om avbildningar (funktioner) mellan vektorrum som respekterar vektorrumsstrukturen.

Definition 3.1. Låt V och W vara vektorrum över en kropp k. En *linjär avbildning* från V till W är en funktion

$$L\colon V \to W$$

som är:

(1) additiv: $L(v_1 + v_2) = L(v_1) + L(v_2)$ för alla $v_1, v_2 \in V$ och
(2) k-linjär: $L(av) = aL(v)$ för all $a \in k$ och alla $v \in V$.

Mängden av alla linjära avbildningar från V till W betecknas $\mathrm{Hom}_k(V, W)$ eller $\mathcal{L}(V, W)$.

Notera att $L\colon V \to W$ är linjär om och endast om $L(av_1 + v_2) = aL(v_1) + L(v_2)$ för alla $v_1, v_2 \in V$ och $a \in k$. Notera också att om k är en kropp så är $\mathrm{Hom}_k(V, W)$ ett vektorrum över k med vektorrumsstrukturen given av

$$(L_1 + L_2)(v) = L_1(v) + L_2(v)$$
$$(cL)(v) = cL(v)$$

för alla $L_1, L_2, L \in \mathrm{Hom}_k(V, W)$, $c \in k$ och $v \in V$.

Exempel 3.1. Varje $m \times n$-matris $A = (a_{ij})$ definierar en linjär avbildning $L_A\colon k^n \to k^m$ genom matris-vektor-multiplikation, dvs

$$L_A(v) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{pmatrix},$$

för alla $v \in k^n$. Vi kommer ofta att identifiera en matris A med den linjära avbildning L_A den definierar och skriva A istället för L_A även när vi tänker på A som en linjär avbildning på detta vis.

Diskussionsfrågor

(1) För vilka $a, b, c \in \mathbb{R}$ gäller det att funktionen $f(x) = ax^2 + bx + c$ är linjär?
(2) Är $\frac{d}{dx} \colon k[x] \to k[x]$ en linjär avbildning?
(3) Är funktionen
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + b$$
en linjär avbildning från $M_2(k)$ till k?
(4) Är funktionen
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc$$
en linjär avbildning från $M_2(k)$ till k?

Definition 3.2. Låt $L \colon V \to W$ vara en linjär avbildning. Vi definierar *kärnan* (eller *nollrummet*) som
$$\ker L := \{v \in V : L(v) = 0\}$$
och *bilden* (eller bildrummet) av L som
$$\operatorname{im} L := \{w \in W : w = L(v) \text{ för något } v \in V\}\,.$$

Sats 3.1 (Noll- och bildrumssatsen)**.** *Låt $L \colon V \to W$ vara en linjär avbildning. Då är* $\ker L$ *ett delrum av V och* $\operatorname{im} L$ *är ett delrum av W.*

Bevis. Eftersom L är additiv så är $\ker L$ och $\operatorname{im} L$ båda slutna under addition och eftersom L är k-linjär så är $\ker L$ och $\operatorname{im} L$ båda slutna under skalärmultiplikation. Det återstår att visa att $L(0) = 0$. Detta följer av att $L(-0) = L((-1)0) = -L(0)$ av k-linjäritet och därmed har vi $L(0) = L(0 - 0) = L(0) - L(0) = 0$. $\square$

Dimensionen av bildrummet kallas *rangen* av L och betecknas $\operatorname{rk} L := \dim \operatorname{im} L$.

Sats 3.2 (Injektivitetssatsen)**.** *En linjär avbildning $L \colon V \to W$ är injektiv om och endast om* $\ker L = \{0\}$.

Bevis. Det är klart att $\ker L = \{0\}$ om L är injektiv, så vi bevisar motsatsen. Antag att $\ker L = \{0\}$ och att $L(v) = L(w)$. Då har vi att $L(v - w) = 0$ så $v - w \in \ker L = \{0\}$, dvs $v = w$. □

Sats 3.3 (Dimensionssatsen)**.** *Låt $L\colon V \to W$ vara en linjär avbildning mellan vektorrum av ändlig dimension. Då gäller att*

$$\dim V = \dim \ker L + \dim \operatorname{im} L\,.$$

Bevis. Låt $\{v_1, \ldots, v_m\}$ vara en bas för $\ker L$ och utvidga denna, som i Sats 2.5, till en bas $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ för V. Vi påstår nu att $\{L(v_{m+1}), \ldots, L(v_n)\}$ en bas för $\operatorname{im} L$. Mycket riktigt, vi har att $\operatorname{Span}(L(v_{m+1}), \ldots, L(v_n)) = \operatorname{im} L$ eftersom $\operatorname{im} L$ spänns upp av $\{L(v_1), \ldots L(v_m), L(v_{m+1}), \ldots, L(v_n)\}$ och $L(v_i) = 0$ för alla $1 \leq i \leq m$. Vi behöver nu visa att $\{L(v_{m+1}), \ldots, L(v_n)\}$ är linjärt oberoende. Anta att

$$a_{m+1}L(v_{m+1}) + \cdots + a_nL(v_n) = 0\,.$$

Av linjäritet har vi då att

$$L(a_{m+1}v_{m+1} + \cdots + a_nv_n) = 0\,,$$

dvs $a_{m+1}v_{m+1}+\cdots+a_nv_n \in \ker L$. Men $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ är linjärt oberoende så $a_{m+1} = \cdots = a_n = 0$. Därmed är $\{L(v_{m+1}), \ldots, L(v_n)\}$ linjärt oberoende. Vi har därmed visat att

$$\dim \ker L + \dim \operatorname{im} L = m + (n - m) = n = \dim V\,. \qquad \square$$

Sats 3.4 (Bijektionssatsen)**.** *Låt $L\colon V \to W$ vara en linjär avbildning mellan vektorrum* av samma (ändlig) dimension*. Följande är ekvivalent:*

(1) L är surjektiv,
(2) L är injektiv,
(3) L är en bijektion.

Bevis. Av dimensionssatsen har vi att $\dim V = \dim \ker L + \dim \operatorname{im} L$. Om L är surjektiv så är $\operatorname{im} L = W$ och därmed $\dim \operatorname{im} L = \dim W = \dim V$, vilket ger $\dim \ker L = 0$ och därmed att L är injektiv. Omvänt, om L är injektiv så har vi $\dim \ker L = 0$ och därmed $\dim \operatorname{im} L = \dim V$, vilket ger att L är surjektiv. Därmed är L en bijektion. □

Övning

Låt V och W vara vektorrum och antag att β är en bas för V. Visa att varje linjär avbildning $L\colon V \to W$ är unikt bestämd av värdena $L(v)$ för alla $v \in \beta$.

3.2. **Matriser.** Linjära avbildningar kan representeras med matriser så fort man valt baser för vektorrummen i fråga. Vi börjar med att definiera vad en matris är.

Definition 3.3. En *matris* (eller mer specifikt en $m \times n$-matris) över en kropp k är en 2-dimensionell lista:

$$A = (a_{ij}) := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & & \\ a_{m1} & \cdots & & a_{mn} \end{pmatrix}.$$

Vi skriver också $A_{ij} := a_{ij}$, dvs $(a_{ij})_{st} = a_{st}$.

Definition 3.4. Låt $L\colon V \to W$ vara en linjär avbildning mellan vektorrum av ändlig dimension n respektive m. Antag att vi fixerat *ordnade* baser $\alpha = \{v_1, \ldots, v_n\} \subseteq V$ och $\beta = \{w_1, \ldots, w_m\} \subseteq W$. Vi definierar matrisen $[L]_\alpha^\beta \in M_{m\times n}(k)$ genom att sätta

$$[L]_\alpha^\beta = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & & \\ a_{m1} & \cdots & & a_{mn} \end{pmatrix} = \begin{pmatrix} [L(v_1)]_\beta & [L(v_2)]_\beta & \cdots & [L(v_n)]_\beta \end{pmatrix},$$

där $L(v_j) = \sum_i a_{ij} w_i$ för alla $1 \leq j \leq n$ och $1 \leq i \leq m$. Matrisen $[L]_\alpha^\beta$ kallas *matrisrepresentationen* av L med avseende på baserna α och β.

Exempel 3.2. Varje vektor $v \in V$ kan ses som en linjär avbildning $L_v\colon k \to V$ definierad av $L_v(a) = av$. Om vi fixerar en bas $\beta = \{v_1, \ldots, v_m\}$ för V så är matrisrepresentationen av L_v med avseende på standardbasen ε för k och basen β

givet av

$$[L_v]^\beta_\varepsilon = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix},$$

där $v = a_1 v_1 + \cdots + a_m v_m$. Dvs, matrisrepresentationen av v är precis koordinatvektorn för v med avseende på basen β: $[L_v]^\beta_\varepsilon = [v]_\beta$.

Sats 3.5 (Matrissatsen)**.** *Låt V och W vara vektorrum över en kropp k med ändlig dimension n respektive m och antag att vi fixerat* ordnade *baser $\alpha = \{v_1, \dots, v_n\} \subseteq V$ och $\beta = \{w_1, \dots, w_m\} \subseteq W$. Avbildningen som ges av*

$$[-]^\beta_\alpha \colon \operatorname{Hom}_k(V, W) \to M_{m\times n}(k)$$
$$L \mapsto [L]^\beta_\alpha$$

är en isomorfi av vektorrum över k.

Bevis. Varje linjär avbildning L bestäms unikt av hur den avbildar basvektorerna i α, eftersom om $v \in V$ är en linjärkombination $v = c_1 v_1 + \cdots + c_n v_n$ så är

$$L(v) = L(c_1 v_1 + \cdots + c_n v_n) = c_1 L(v_1) + \cdots + c_n L(v_n).$$

Å andra sidan, varje matris $A \in M_{m\times n}(k)$ bestäms unikt av sina kolumner, dvs av hur den avbildar basvektorerna i α. Vi har därmed att $[-]^\beta_\alpha$ är en bijektion.

Det återstår att visa att Φ är linjär, dvs additiv och k-linjär. För varje basvektor v_j, $L_1, L_2 \in \operatorname{Hom}_k(V, W)$ och $c \in k$ har vi att

$$(cL_1 + L_2)(v_j) := cL_1(v_j) + L_2(v_j) = \sum_i (ca^{(1)}_{ij} + a^{(2)}_{ij}) w_i.$$

Det vill säga

$$[cL_1 + L_2]^\beta_\alpha = \begin{pmatrix} ca^{(1)}_{11} + a^{(2)}_{11} & ca^{(1)}_{12} + a^{(2)}_{12} & \dots & ca^{(1)}_{1n} + a^{(2)}_{1n} \\ ca^{(1)}_{21} + a^{(2)}_{21} & \ddots & & \vdots \\ \vdots & & & \\ ca^{(1)}_{m1} + a^{(2)}_{m1} & \dots & & ca^{(1)}_{mn} + a^{(2)}_{mn} \end{pmatrix} = c[L_1]^\beta_\alpha + [L_2]^\beta_\alpha.$$

□

3.3. **Sammansättning av linjära avbildningar.** Precis som med funktioner i allmänhet kan vi definiera sammansättning av linjära avbildningar.

Definition 3.5. Låt V, W och Z vara vektorrum över en kropp k. Om $L\colon V \to W$ och $M\colon W \to Z$ är linjära avbildningar så definierar vi *sammansättningen* (eller *kompositionen*) av L och M enligt $ML(v) := M(L(v))$.

Övning

Låt V, W och Z vara vektorrum över en kropp k. Visa att om $L\colon V \to W$ och $M\colon W \to Z$ är linjära avbildningar så är $ML\colon V \to Z$ en linjär avbildning.

Vi ska nu visa att sammansättning av linjära avbildningar (representerade av matriser) motsvarar *multiplikation* av matriser.

Definition 3.6. Om $A = (a_{ij})$ är en $m \times n$-matris och $B = (b_{st})$ är en $n \times p$-matris så definieras produkten AB genom $(AB)_{it} = \sum_{s=1}^{n} a_{is}b_{st}$.

Övning

Om A och B är två matriser, sådana att AB är definierad, visa att $(AB)^T = B^T A^T$.

Sats 3.6 (Matrisen till en sammansättning)**.** *Låt V, W och Z vara vektorrum över en kropp k med ändlig dimension n, m respektive p. Antag att $\alpha = \{v_1, \dots, v_n\}$ är en bas för V, $\beta = \{w_1, \dots, w_m\}$ är en bas för W och $\gamma = \{z_1, \dots, z_p\}$ är en bas för Z. Om $L\colon V \to W$ och $M\colon W \to Z$ är linjära avbildningar så gäller att*

$$[ML]_\alpha^\gamma = [M]_\beta^\gamma [L]_\alpha^\beta .$$

Bevis. Låt $B = [L]_\alpha^\beta$ och $A = [M]_\beta^\gamma$. Vi vill visa att $[ML]_\alpha^\gamma = AB$ (titta noga på ordningen). Det räcker att visa hur ML avbildar basvektorerna i α. Vi har att $ML(v_t) = M(L(v_t)) = M\left(\sum_{s=1}^m b_{st}w_s\right) = \sum_{s=1}^m b_{st}M(w_s) = \sum_{s=1}^m \left(b_{st}\sum_{i=1}^p a_{is}z_i\right) = \sum_{i=1}^p \left(\sum_{s=1}^m a_{is}b_{st}\right) z_i$. Dvs, det *ite* elementet i $[ML]_\alpha^\gamma$ är $\sum_{s=1}^m a_{is}b_{st}$ vilket är detsamma som $(AB)_{it}$. Därmed har vi visat att $[ML]_\alpha^\gamma = AB$. □

Följdsats 3.7. *Låt V och W vara vektorrum över en kropp k med ändlig dimension n respektive m. Antag att α är en ordnad bas för V och β en ordnad bas för W. För varje $v \in V$ gäller då att*

$$[L]_{\alpha}^{\beta}[v]_{\alpha} = [L(v)]_{\beta}\,.$$

Bevis. Vi kan se vektorn v som en linjär avbildning från k till V genom att skicka $1 \mapsto v$. Om vi tar basen $\varepsilon = \{1\}$ för k som vektorrum över sig självt har vi då att $[v]_{\alpha} = [v]_{\varepsilon}^{\alpha}$ och att $[L(v)]_{\beta} = [L(v)]_{\varepsilon}^{\beta}$. Vi har därmed av [Matrisen-till-en-sammansättning] Sats 3.6 att

$$[L]_{\alpha}^{\beta}[v]_{\varepsilon}^{\alpha} = [L \circ v]_{\varepsilon}^{\beta} = [L(v)]_{\beta}\,. \qquad \square$$

4. Inverterbara avbildningar och basbyte

I detta kapitel ska vi studera hur vi kan byta bas i ett vektorrum och hur detta förhåller sig till inverterbara avbildningar. Vi kommer att se att varje basbyte motsvarar en inverterbar avbildning, och att varje inverterbar avbildning från ett vektorrum till sig självt kan representeras som ett basbyte.

4.1. **Inverterbara avbildningar.** Vi ska nu se att om en surjektiv linjär avbildning har en vänsterinvers så är även den linjär. I synnerhet är inversen (som mängdteoretisk funktion) automatiskt linjär.

Sats 4.1 (Inversen är linjär)**.** *Låt $L\colon V \to W$ vara en surjektiv linjär avbildning mellan två vektorrum över en kropp k. Antag att det finns en* funktion *$f\colon W \to V$ sådan att $f \circ L = I_V$. Då är f linjär och $L \circ f = I_W$. I synnerhet är L en bijektion.*

Bevis. Antag att vi har $w_1, w_2 \in W$ och $c \in k$. Vi vill visa att $f(cw_1 + w_2) = cf(w_1) + f(w_2)$. Eftersom L är surjektiv så finns det $v_1, v_2 \in V$ så att $L(v_1) = w_1$ och $L(v_2) = w_2$. Vi har då att

$$\begin{aligned}
f(cw_1 + w_2) &= f(cL(v_1) + L(v_2)) \\
&= f(L(cv_1 + v_2)) \\
&= (f \circ L)(cv_1 + v_2) \\
&= I_V(cv_1 + v_2) \\
&= cv_1 + v_2 \\
&= c(f \circ L)(v_1) + (f \circ L)(v_2) \\
&= cf(w_1) + f(w_2)\,.
\end{aligned}$$

Vi har därmed visat att f är linjär. Då L är surjektiv kan varje $w \in W$ skrivas som $w = L(v)$ för något $v \in V$, vilket ger att $(L \circ f)(w) = (L \circ f)(L(v)) = L((f \circ L)(v)) = L(v) = w$. Därmed är $L \circ f = I_W$. $\square$

Exempel 4.1. Låt V vara vektorrummet av alla polynom i variabeln x med koefficienter i k. Låt $L_1 = \frac{d}{dx}\colon V \to V$ vara avbildningen som skickar ett polynom till dess derivata. Då är L_1 linjär och surjektiv. Låt $L_2 = \int_0^x\colon V \to V$ vara avbildningen $L_2(p(x)) = \int_0^x p(t)\,dt$, vilken är linjär och injektiv. Då har vi att $L_1 \circ L_2 = I_V$ men $L_2 \circ L_1 \neq I_V$.

Definition 4.1. En linjär avbildning L som är både injektiv och surjektiv kallas för en *isomorfi*. Om $L\colon V \to W$ är en isomorfi så säger vi att V och W är *isomorfa* och skriver $V \cong W$.

Notera att om $L\colon V \to W$ är en isomorfi så kan vi definiera $L^{-1}\colon W \to V$ genom $L^{-1}(w) = v$ om och endast om $L(v) = w$. Då säger Sats 4.1 att L^{-1} är linjär och vi kallar L^{-1} för *inversen* till L. Vi kommer därför att använda begreppen *inverterbar* och *isomorfi* synonymt.

Exempel 4.2. Avbildningen $(-)^T\colon M_{m\times n}(k) \to M_{n\times m}(k)$ som skickar en matris $A = (a_{ij})_{i,j}$ till sitt transponat $A^T = (a_{ji})_{i,j}$:

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

är en isomorfi.

Sats 4.2 (Trivialitetssatsen)**.** *Låt V vara ett vektorrum av ändlig dimension över en kropp k. Då har vi $V \cong k^n$. Mer specifikt, vi har en bijektion mellan mängden av alla ordnade baser för V och mängden isomorfier $V \cong k^n$.*

Bevis. Låt $\beta = \{v_1, \ldots, v_n\}$ vara en ordnad bas för V. Vi definierar en avbildning $L_\beta\colon V \to k^n$ på basvektorerna genom $L_\beta(v_i) = e_i$, där e_i är standardbasvektorerna i k^n. Vi kan sedan utvidga linjärt till hela V. Dvs om $v = \sum_{i=1}^n a_i v_i$ så sätter vi $L_\beta(v) = \sum_{i=1}^n a_i e_i$. Då är L linjär per definition. Den är surjektiv eftersom standardbasen ligger i bilden och den är injektiv eftersom $L_\beta(v) = 0$ om och endast om alla $a_i = 0$, vilket är sant om och endast om $v = 0$. Alltså är L_β en isomorfi. Detta ger en avbildning $\beta \mapsto L_\beta$ från mängden av alla ordnade baser för V till mängden isomorfier $V \to k^n$. Omvänt, låt $L\colon V \to k^n$ vara en isomorfi. Då är $\beta_L := \{L^{-1}(e_1), \ldots, L^{-1}(e_n)\}$ en ordnad bas för V. Detta ger en avbildning från mängden isomorfier $V \to k^n$ till mängden av alla ordnade baser för V. Det är lätt att se att dessa två avbildningar är inverser till varandra, vilket ger en bijektion mellan mängden av alla ordnade baser för V och mängden isomorfier $V \to k^n$. □

Följdsats 4.3. *Två ändligt-dimensionella vektorrum V och W över en kropp k är isomorfa om och endast om de har samma dimension.*

Detta säger alltså att för varje positivt heltal n så finns det *upp till isomorfi* precis ett vektorrum av dimension n över en kropp k.

Diskussionsfrågor

Vilka av följande vektorrum är isomorfa?

(1) 2×2-matriser med reella koefficienter.
(2) $\mathbb{C}^2$ som vektorrum över $\mathbb{R}$, där modulstrukturen $r(z_1, z_2) = (rz_1, rz_2)$ för alla $r \in \mathbb{R}$, $(z_1, z_2) \in \mathbb{C}^2$ och komponentvis addition.
(3) $\mathbb{R}[x]_{\leq 3}$ som vektorrum över $\mathbb{R}$, dvs mängden av alla reella polynom av grad högst 3.

Lemma 4.4. *Låt $L\colon V \to W$ vara en linjär avbildning mellan vektorrum av ändlig dimension med ordnade baser $\alpha = \{v_1, \ldots, v_n\}$ respektive $\beta = \{w_1, \ldots, w_m\}$. Låt vidare $A\colon V \to k^n$ och $B\colon W \to k^m$ vara de isomorfier som bestäms av $v_i \mapsto e_i$ respektive $w_j \mapsto e_j$, där e_i är standardbasvektorerna i k^n respektive k^m. Då har vi ett kommutativt diagram*

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "A", "\cong"'] & W \arrow[d, "B", "\cong"'] \\ k^n \arrow[r, "{[L]_\alpha^\beta}"] & k^m, \end{tikzcd}$$

vilket betyder att $[L]_\alpha^\beta \circ A = B \circ L$.

Bevis. Detta är bokstavligen Följdsats 3.7 eftersom A skickar v på koordinatvektorn $[v]_\alpha$. □

Sats 4.5 (Val av bas och kommutativitet)**.** *Låt V och W vara två ändligt-dimensionella vektorrum över en kropp k och låt α och β vara ordnade baser* *för V respektive W. Låt $L\colon V \to W$ vara en linjär avbildning. Då har vi ett kommutativt diagram med isomorfier som indikerat nedan:*

$$\begin{tikzcd}
\ker L \arrow[r] \arrow[d, "\cong"] & V \arrow[r, "L"] \arrow[d, "\cong"] & \operatorname{im} L \arrow[r] \arrow[d, "\cong"] & W \arrow[d, "\cong"] \\
\ker[L]_\alpha^\beta \arrow[r] & k^n \arrow[r, "{[L]_\alpha^\beta}"] & \operatorname{im}[L]_\alpha^\beta \arrow[r] & k^m .
\end{tikzcd}$$

Bevis. Detta följer direkt av Lemma 4.4. □

Följdsats 4.6. *Låt V och W vara två ändligt-dimensionella vektorrum över en kropp k och låt α och β vara ordnade baser för V respektive W. Låt $L\colon V \to W$ vara en linjär avbildning. Då är L en isomorfi om och endast om matrisen $[L]_\alpha^\beta$ är inverterbar. I detta fall har vi $[L^{-1}]_\beta^\alpha = ([L]_\alpha^\beta)^{-1}$.*

Bevis. Av Lemma 4.4 har vi ett kommutativt diagram

$$\begin{tikzcd}
V \arrow[r, "L"] \arrow[d, "\cong"] & W \arrow[d, "\cong"] \\
k^n \arrow[r, "{[L]_\alpha^\beta}"] & k^m
\end{tikzcd}$$

vilket direkt ger oss att $[L]_\alpha^\beta$ är inverterbar om och endast om L är en isomorfi. □

4.2. **Basbyte.** Givet två ordnade baser α och β för ett ändligt-dimensionellt vektorrum V så såg vi i Kapitel 2 att α och β har samma antal element $\dim V$. Notera att om $V = W$ i Matrissatsen (Sats 3.5) så kommer identiteten $I_V\colon V \to V$ att avbildas på identitetsmatrisen om och endast om $\alpha = \beta$ som ordnade baser.

Definition 4.2. Matrisen $[I_V]_\alpha^\beta$ kallas *basbytesmatrisen* från α till β.

Basbytesmatrisen $[I_V]_\alpha^\beta$ hjälper oss alltså att uttrycka basen α i termer av basen β. Mycket riktigt, per definition har vi att

$$[I_V]_\alpha^\beta = \begin{pmatrix} [v_1]_\beta & [v_2]_\beta & \cdots & [v_n]_\beta \end{pmatrix},$$

där $\alpha = \{v_1, \ldots, v_n\}$.

Sats 4.7 (Basbyte). *Låt V vara ett vektorrum över en kropp k med ändlig dimension n. Låt α och β vara två ordnade baser för V. Då har vi att basbytesmatrisen $[I_V]_\alpha^\beta$ är inverterbar med invers $[I_V]_\beta^\alpha$ och för varje $v \in V$ gäller det att $[I_V]_\alpha^\beta[v]_\alpha = [v]_\beta$.*

Bevis. Vi har $[I_V]_\alpha^\beta[I_V]_\beta^\alpha = [I_V]_\alpha^\alpha$ vilket är identitetsmatrisen. Det sista påståendet följer direkt från Följdsats 3.7. $\square$

Diskussionsfrågor

(1) (**Jätteviktig!**) Vad är basbytesmatrisen från en given bas $\{v_1, \ldots, v_n\} \subseteq k^n$ till standardbasen för k^n?

(2) Vad är basbytesmatrisen mellan följande baser för $\mathbb{R}^3$:

$$\left\{\begin{pmatrix}3\\1\\4\end{pmatrix}, \begin{pmatrix}2\\-1\\0\end{pmatrix}, \begin{pmatrix}2\\3\\7\end{pmatrix}\right\} \quad \text{och} \quad \left\{\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right\}?$$

Tips: använd lösningen till (1).

Exempel 4.3. Låt V vara vektorrummet av alla polynom i variabeln x med koefficienter i $\mathbb{R}$ av grad högst 2. Låt $\alpha = \{2 - x + 3x^2, x + x^2, 1 + 4x^2\}$ och $\beta = \{1 - x^2, x + 1, 1\}$. Vi vill nu bestämma basbytesmatrisen $[I_V]_\alpha^\beta$. Ett sätt att göra detta är att först bestämma $[I_V]_\alpha^\varepsilon$ och $[I_V]_\beta^\varepsilon$. Vi har

$$[I_V]_\alpha^\varepsilon = \begin{pmatrix}2 & 0 & 1\\-1 & 1 & 0\\3 & 1 & 4\end{pmatrix} \quad \text{och} \quad [I_V]_\beta^\varepsilon = \begin{pmatrix}1 & 1 & 1\\0 & 1 & 0\\-1 & 0 & 0\end{pmatrix}.$$

Vi kan nu hitta $[I_V]_\alpha^\beta$ genom $[I_V]_\alpha^\beta = [I_V]_\varepsilon^\beta [I_V]_\alpha^\varepsilon = ([I_V]_\beta^\varepsilon)^{-1}[I_V]_\alpha^\varepsilon$. Vi hittar $([I_V]_\beta^\varepsilon)^{-1}$ genom radoperationer:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$
$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right)$$
$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right).$$

Vi får då

$$([I_V]_\beta^\varepsilon)^{-1}[I_V]_\alpha^\varepsilon = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 3 & 1 & 4 \end{pmatrix} = \begin{pmatrix} -3 & -1 & -4 \\ -1 & 1 & 0 \\ 6 & 0 & 5 \end{pmatrix}.$$

Vi kan kontrollera att detta stämmer genom att notera att

$$\begin{aligned} 2 - x + 3x^2 &= -3 \cdot (1 - x^2) - 1 \cdot (x + 1) + 6 \cdot (1) \\ x + x^2 &= -1 \cdot (1 - x^2) + 1 \cdot (x + 1) + 0 \cdot (1) \\ 1 + 4x^2 &= -4 \cdot (1 - x^2) + 0 \cdot (x + 1) + 5 \cdot (1)\,. \end{aligned}$$

5. Matrisoperationer och linjära ekvationssystem

Detta kapitel handlar om hur vi kan förenkla ett linjärt ekvationssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m\,, \end{aligned}$$

med hjälp av tre typer av *elementära operationer*, för att få dem i en enklare form som är lättare att lösa. Dessa operationer är:

(1) *Radbyte:* Byt plats på ekvationer.
(2) *Radmultiplikation:* Multiplicera en ekvation med en nollskild skalär.
(3) *Radaddition:* Addera en multipel av en ekvation till en annan ekvation.

Notera att dessa operationer är inverterbara och att den inversa operationen också är av denna elementära form.

5.1. **Elementära matrisoperationer.** Vi ska nu definiera de grundläggande elementära operationerna:

Definition 5.1. En *elementär matrisoperation* är en av följande operationer som kan utföras på en matris:

(1) *Rad(Kolumn)-byte:* Byt plats på två rader (kolumner) i en matris.
(2) *Rad(Kolumn)-multiplikation:* Multiplicera en rad (kolumn) i en matris med en nollskild skalär.
(3) *Rad(Kolumn)-addition:* Addera en multipel av en rad (kolumn) till en annan rad (kolumn).

Vi kallar också dessa operationer *radoperationer* eller *kolumnoperationer* beroende på om de utförs på rader eller kolumner.

Definition 5.2. En *elementär matris av typ (1), (2) eller (3)* är en matris som fås genom att utföra motsvarande typ av elementär matrisoperation på en identitetsmatris.

Notera att varje elementär matris är inverterbar och att den inversa matrisen är en elementär matris av samma typ.

Exempel 5.1 (Typ 1)**.** Matrisen

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

är en elementär matris av typ (1) eftersom vänstermultiplikation med E byter rad 2 och 3:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b & c \\ g & h & i \\ d & e & f \end{pmatrix},$$

och högermultiplikation med E byter kolumn 2 och 3:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} a & c & b \\ d & f & e \\ g & i & h \end{pmatrix}.$$

Exempel 5.2 (Typ 2)**.** Matrisen

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

är en elementär matris av typ (2) eftersom vänstermultiplikation med E multiplicerar rad 2 med 2:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{pmatrix},$$

och högermultiplikation med E multiplicerar kolumn 2 med 2:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & 2b & c \\ d & 2e & f \\ g & 2h & i \end{pmatrix}.$$

Exempel 5.3 (Typ 3)**.** Matrisen

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

är en elementär matris av typ (3) eftersom vänstermultiplikation med E adderar 2 gånger rad 3 till rad 2:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b & c \\ d+2g & e+2h & f+2i \\ g & h & i \end{pmatrix},$$

och högermultiplikation med E adderar 2 gånger kolumn 2 till kolumn 3:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & c+2b \\ d & e & f+2e \\ g & h & i+2h \end{pmatrix}.$$

Sats 5.1 (Radoperationer)**.** *Låt A vara en matris och B en matris som erhålls från A genom en rad(kolumn)operation. Då har vi att $B = EA$ ($B = AE$) där E är matrisen som fås genom att utföra den elementära matrisoperationen på identitetsmatrisen av samma dimension som antalet rader (kolumner) i A. Å andra sidan, vänster(höger)multiplikation med en elementär matris E på en matris A motsvarar att utföra den elementära matrisoperationen på A.*

Bevis. Övning. □

5.2. **Matrisrang.** Ett specialfall av matrissatsen (Sats 3.5) är att

$$M_{n\times n}(k) \cong \mathrm{Hom}_k(k^n, k^n),$$

dvs, varje $n \times n$-matris över en kropp k svarar mot en linjär avbildning från k^n till k^n, nämligen den linjära avbildning som ges av vänstermultiplikation med matrisen. Vi kan därmed tala om matriser som linjära avbildningar och vi kan tala om bildrummet, nollrummet och rangen av en matris, vilka vi betecknar som $\mathrm{im}\, A$, $\ker A$ och $\mathrm{rk}\, A$.

Sats 5.2 (Matrisrangssatsen)**.** *Låt A vara en matris över en kropp k. Då är* $\mathrm{rk}\, A$ *lika med dimensionen av rummet som spänns upp av kolumnvektorerna i A.*

Bevis. Detta följer av att bilden av standardbasen i k^n (vilken spänner upp $\operatorname{im} A$) är precis kolumnerna i A. $\square$

Sats 5.3 (Trivial form)**.** *Låt A vara en $m \times n$-matris av rang r över en kropp k. Då finns det elementära matriser $E_1, E_2, \ldots, E_l$ och $E_{1'}, \ldots, E'_{l'}$ för några $l, l' \geq 0$ sådana att*

$$E_l E_{l-1} \cdots E_1 A E_{1'} E_{2'} \cdots E_{l'} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

där 0 betecknar nollmatriser av lämplig storlek.

Bevis. Vi använder induktion över m. Basfallet $m = 1$ är trivialt. Anta nu att satsen gäller för $m-1$. Vi visar att den också gäller för m. Om $\operatorname{rk} A = 0$ så är A en nollmatris och det finns inget att bevisa. Anta nu att $\operatorname{rk} A > 0$. Då finns det ett icke-noll element i A, låt oss säga A_{ij}. Först använder vi en radoperation av typ (2) för att göra $A_{ij} = 1$ (om det inte redan är 1). Sedan använder vi rad- och kolumnoperationer av typ (1) för att flytta denna etta till index $(1, 1)$. Slutligen använder vi rad- och kolumnoperationer av typ (3) för att nollställa alla andra element i den första raden och den första kolumnen. Detta ger oss en matris av formen

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & * & * & \ldots & * \\ 0 & * & * & \ldots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \ldots & * \end{pmatrix}.$$

Nu kan vi applicera induktionsantagandet på den nedre delen av matrisen, dvs den del som inte innehåller den första raden och kolumnen. Slutresultatet blir då en matris i önskad form. $\square$

Följdsats 5.4 (Elementär generering)**.** *Varje inverterbar matris A över en kropp k kan skrivas som produkten av elementära matriser, dvs det finns elementära matriser $E_1, E_2, \ldots, E_l$ sådana att*

$$A = E_l E_{l-1} \cdots E_1 .$$

Notera att detta innebär att vi hittar inversen av A enligt följande procedur: Skriv A^{-1} som produkten av elementära matriser $E'_1, E'_2, \ldots, E'_{l'}$. Denna produkt är ej unik utan det finns många olika sätt att skriva A^{-1} på detta sätt. Vi har $A^{-1}A = I$ och $A^{-1}I = A^{-1}$. Med andra ord, om vi utför radoperationer steg för steg på A och I

samtidigt, så kommer vi ha transformerat I till A^{-1} vid det laget då vi transformerat A till I.

Följdsats 5.5. *Låt A vara en $m \times n$-matris av rang r över en kropp k. Då finns en inverterbar matris B av storlek $m \times m$ och en matris C av storlek $n \times n$ sådana att*

$$BAC = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Följdsats 5.6 (Rad- och kolumnrang)**.** *Låt A vara en $m \times n$-matris över en kropp k. Då har vi att följande tal är samma:*

(1) rk A,
(2) rk A^T,
(3) dimensionen av vektorrummet som spänns upp av kolumnerna i A,
(4) dimensionen av vektorrummet som spänns upp av raderna i A.

Följdsats 5.7 (Inverterbarhet och rang)**.** *Låt A vara en $n \times n$-matris över en kropp k. Då är A inverterbar om och endast om* rk $A = n$.

Diskussionsfrågor

(1) Vad är dimensionen av bildrummet och nollrummet till matrisen

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{pmatrix}?$$

(2) Är matrisen

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{pmatrix}$$

inverterbar?

5.3. **Lösningar till linjära ekvationssystem.** Vi ska nu se hur vi kan tillämpa våra resultat för att lösa linjära ekvationssystem.

Definition 5.3. Ett system $Ax = b$ av linjära ekvationer kallas *homogent* om $b = 0$ och *inhomogent* om $b \neq 0$.

Sats 5.8 (Lösningsmängden)**.** *Betrakta ett system $Ax = b$ där A är en $m \times n$-matris över en kropp k. Om $b = 0$ så har systemet alltid minst en lösning, nämligen $x = 0$. Om s är en lösning till systemet, så är lösningsmängden till systemet*

$$\{s + v \mid v \in \ker A\}.$$

Bevis. Om $v \in \ker A$ så har vi $A(s + v) = As + Av = b + 0 = b$. Å andra sidan, om s' är en lösning till systemet så har vi att $s - s' \in \ker A$ och $s' = s + (s' - s)$. □

Sats 5.9 (Antal lösningar till linjära ekvationssystem)**.** *Betrakta ett system $Ax = b$ där A är en $m \times n$-matris över en kropp k.*

(1) Om $\operatorname{rk} A = \operatorname{rk}(A|b) = n$ så har systemet en unik lösning.
(2) Om $\operatorname{rk} A = \operatorname{rk}(A|b) < n$ så har systemet oändligt många lösningar.
(3) Om $\operatorname{rk} A < \operatorname{rk}(A|b)$ så har systemet inga lösningar.

Bevis. Detta följer av att $\operatorname{rk} A = \operatorname{rk}(A|b)$ om och endast om b kan skrivas som en linjärkombination av kolumnerna i A, dvs $\operatorname{im} A$, och av att $\dim \ker A > 0$ om och endast om $\operatorname{rk} A < n$ (Dimensionssatsen). □

5.4. **Gausseliminering.** Vi ska nu beskriva en algoritm som kallas *Gausseliminering* för att lösa linjära ekvationssystem. Denna metod använder elementära matriser och radoperationer för att förenkla ett system till en form som är lättare att lösa.

Definition 5.4. Två linjära ekvationssystem $Ax = b$ och $A'x = b'$ sägs vara *ekvivalent*a om de har samma lösningsmängd, dvs om varje lösning till det ena systemet är en lösning till det andra systemet.

Sats 5.10 (Isomorfier och ekvivalenta system)**.** *Om $Ax = b$ är ett linjärt ekvationssystem och B är en inverterbar matris, så är $BAx = Bb$ ett ekvivalent system till $Ax = b$.*

Bevis. Övning. $\square$

Exempel 5.4. Systemen

$$\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix}1\\ 2\\ 3\end{pmatrix}$$

och

$$\begin{pmatrix}1 & 2 & 3\\ 0 & -3 & -6\\ 0 & -6 & -12\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix}1\\ -2\\ -4\end{pmatrix}$$

är ekvivalenta eftersom den andra matrisen fås genom att vänstermultiplicera den första matrisen med den inverterbara matrisen

$$B = \begin{pmatrix}1 & 0 & 0\\ -4 & 1 & 0\\ -7 & 0 & 1\end{pmatrix}.$$

Med andra ord kan vi använda elementära radoperationer för att transformera ett system till ett ekvivalent system som är lättare att lösa.

Definition 5.5. En matris A sägs vara i *(reducerad) trappstegsform* om den har följande egenskaper:

(1) Alla nollrader (om några) ligger längst ner i matrisen.
(2) Den första nollskilda elementet (pivoten) i varje rad är lika med 1 och ligger till höger om den första icke-nollkolumnen i raden ovanför.
(3) Om en kolumn innehåller en pivot, så är alla andra element i den kolumnen noll.

Exempel 5.5. Matrisen

$$A = \begin{pmatrix}1 & 0 & -2 & 0 & 2\\ 0 & 1 & 4 & 0 & 1\\ 0 & 0 & 0 & 1 & 3\end{pmatrix}$$

är i reducerad trappstegsform eftersom den uppfyller alla tre egenskaperna ovan.

Sats 5.11 (Reducerad trappstegsform)**.** *Låt A vara en $m \times n$-matris av rang r över en kropp k. Då finns det elementära matriser $E_1, E_2, \ldots, E_l$ för något $l \geq 0$ sådana att $E_l E_{l-1} \cdots E_1 A$ är i reducerad trappstegsform.*

Bevis. Vi använder induktion över m. Basfallet $m = 1$ är trivialt. Anta nu att satsen gäller för matriser med $m - 1$ rader eller färre. Vi väljer nu första nollskilda kolumnen och väljer en rad med ett icke-nollelement i denna kolumn. Använd nu en elementär matris för att flytta denna rad till den första raden. Använd sedan en elementär matris för att göra pivoten till 1 (om den inte redan är 1). Använd sedan elementära matriser för att nollställa alla andra element i denna kolumn. Detta ger oss en matris av formen

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \ldots & * \\ 0 & \cdots & 0 & 0 & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & * & \ldots & * \end{pmatrix}.$$

Nu kan vi applicera induktionsantagandet på den nedre delen av matrisen, dvs den del som inte innehåller den första raden. När vi skrivit den nedre matrisen i reducerad trappstegsform, så kan vi slutligen eliminera alla element ovanför pivoterna i den första kolumnen med hjälp av elementära matriser. Slutresultatet blir då en matris i reducerad trappstegsform. □

Definition 5.6. Metoden vi använder i beviset ovan kallas för *Gausselimination* eller *radreduktion.*

Exempel 5.6. Betrakta systemet

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Vi kan använda Gausseliminering för att förenkla detta system:

$$\begin{pmatrix}1 & 2 & 3 & | & 1\\ 4 & 5 & 6 & | & 2\\ 7 & 8 & 9 & | & 3\end{pmatrix} \xrightarrow{\text{R2}-4\text{R1}} \begin{pmatrix}1 & 2 & 3 & | & 1\\ 0 & -3 & -6 & | & -2\\ 7 & 8 & 9 & | & 3\end{pmatrix}$$

$$\xrightarrow{\text{R3}-7\text{R1}} \begin{pmatrix}1 & 2 & 3 & | & 1\\ 0 & -3 & -6 & | & -2\\ 0 & -6 & -12 & | & -4\end{pmatrix}$$

$$\xrightarrow{\text{R3}-2\text{R2}} \begin{pmatrix}1 & 2 & 3 & | & 1\\ 0 & -3 & -6 & | & -2\\ 0 & 0 & 0 & | & 0\end{pmatrix}$$

$$\xrightarrow{-3^{-1}\text{R2}} \begin{pmatrix}1 & 2 & 3 & | & 1\\ 0 & 1 & 2 & | & \frac{2}{3}\\ 0 & 0 & 0 & | & 0\end{pmatrix}$$

$$\xrightarrow{\text{R1}-2\text{R2}} \begin{pmatrix}1 & 0 & -1 & | & \frac{1}{3}\\ 0 & 1 & 2 & | & \frac{2}{3}\\ 0 & 0 & 0 & | & 0\end{pmatrix}.$$

Diskussionsfrågor

(1) Vad är dimensionen av lösningsmängden till systemet

$$\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 4 & 6 & 8\\ 3 & 6 & 9 & 12\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\end{pmatrix} = \begin{pmatrix}1\\ 2\\ 3\end{pmatrix}?$$

(2) Använd Gausseliminering för att lösa systemet

$$\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix}1\\ 2\\ 3\end{pmatrix}.$$

6. Determinanter

Innan vi definierar determinanter i full generalitet, betrakta följande exempel: Antag att vi har en 2×2-matris

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

och vi vill veta om A är inverterbar, dvs. om det finns en matris B sådan att $AB = I$, där I är identitetsmatrisen. Dvs, vi vill lösa ekvationen

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I\,.$$

Knepet är att definiera en funktion vi kallar *determinanten*:

$$\begin{aligned} \det\colon M_{2\times 2}(k) &\to k, \\ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \det(A) := ad - bc\,. \end{aligned}$$

Notera att $\det(AB) = (ae + bg)(cf + dh) - (af + bh)(ce + dg) = (ad - bc)(eh - fg) = \det(A)\det(B)$, vilket visar att determinanten har den önskade egenskapen att den är *multiplikativ*. Detta innebär att om $AB = I$ så måste $\det(A)\det(B) = \det(I) = 1$, dvs. $\det(A) \neq 0$. Å andra sidan, om $\det(A) \neq 0$, så kan vi lösa ekvationen $AB = I$ genom att sätta

$$B = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\,,$$

vilket ger oss $AB = I$. Med andra ord, A är inverterbar om och endast om $\det(A) \neq 0$.

6.1. Definition av determinanter.

Låt oss nu generalisera denna definition till $n \times n$-matriser. Vi kommer först att definiera funktioner $D_i, D_i^T\colon M_{n\times n}(k) \to k$ för varje $1 \leq i \leq n$ och sedan visa att de är lika, dvs. att $D_i = D_i^T$ och att de ej beror på i.

Definition 6.1. Låt $A = (a_{ij}) \in M_{n\times n}(k)$ vara en $n \times n$-matris. För varje $1 \leq i, j \leq n$ definierar vi $\widehat{A}_{ij}$ som den $(n-1) \times (n-1)$-matris som fås genom att ta bort rad i och kolumn j från A. För $n = 1$ definierar vi $D_1(A) = D_1^T(A) = a_{11}$.

För $n \geq 2$ definierar vi
$$D_i(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} D_1(\widehat{A}_{ij})\,,$$
och
$$D_i^T(A) = D_i(A^T)\,.$$
Slutligen definierar vi determinanten av A som
$$\det(A) = D(A) = |A| = D_1(A)\,.$$

Sats 6.1 (Multilinjäritet)**.** *Funktionen* $\det\colon M_{n\times n}(k) \to k$ *är linjär i varje rad (och varje kolumn), dvs. om* $A \in M_{n\times n}(k)$ *har formen:*
$$A = \begin{pmatrix} \overline{a}_1^T \\ \vdots \\ \overline{a}_i^T + c\overline{b}_i^T \\ \vdots \\ \overline{a}_n^T \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} + cb_{i1} & a_{i2} + cb_{i2} & \cdots & a_{in} + cb_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$
$$\left(A = \begin{pmatrix} \overline{a}_1 & \cdots & \overline{a}_i + c\overline{b}_i & \cdots & \overline{a}_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} + cb_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} + cb_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ni} + cb_{ni} & \cdots & a_{nn} \end{pmatrix} \right)$$
för något $c \in k$ *så gäller att*
$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} + c\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$
$$\left(\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix} + c\det \begin{pmatrix} a_{11} & a_{12} & \cdots & b_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & b_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & b_{ni} & \cdots & a_{nn} \end{pmatrix} \right)$$

Bevis. Vi använder induktion över n. Basfallet $n = 1$ är trivialt. Antag att $n > 1$ och att påståendet gäller för $(n-1) \times (n-1)$-matriser. Antag att A och B är identiska förutom i rad i. Om $i = 1$ så har vi

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j}(a_{1j} + cb_{1j}) \det(\widehat{A}_{1j}),$$

eftersom $\widehat{A}_{1j}$ inte ändras då vi ändrar rad 1 i A. Om $i > 1$ så har vi

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(\widehat{A}_{1j}).$$

och enligt induktionsantagandet gäller så har vi

$$\det(\widehat{A}_{1j}) = \det \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2(j-1)} & a_{2(j+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{i(j-1)} & a_{i(j+1)} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}$$
$$+ c \det \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2(j-1)} & a_{2(j+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{i(j-1)} & b_{i(j+1)} & \cdots & b_{in} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}.$$

Detta ger att

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} + c \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Det återstår att bevisa linjäritet i varje kolumn. Fallet $n = 1$ är igen trivialt. Antag att $n > 1$ och att påståendet gäller för $(n-1) \times (n-1)$-matriser. Antag att A och B är identiska förutom i kolumn l. Vi har

$$\det(A) = (a_{1l} + cb_{1l}) \det(\widehat{A}_{1l}) + \sum_{j \neq l} (-1)^{1+j} a_{1j} \det(\widehat{A}_{1j}),$$

och av induktionsantagandet gäller

$$\det(\widehat{A}_{1j}) = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1l} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2l} & \cdots & a_{2(j-1)} & a_{2(j+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nl} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}$$

$$+ c \det \begin{pmatrix} a_{11} & a_{12} & \cdots & b_{1l} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & b_{2l} & \cdots & a_{2(j-1)} & a_{2(j+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & b_{nl} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}.$$

om $j \neq l$. Det följer att

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1l} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2l} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nl} & \cdots & a_{nn} \end{pmatrix} + c \det \begin{pmatrix} a_{11} & a_{12} & \cdots & b_{1l} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & b_{2l} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & b_{nl} & \cdots & a_{nn} \end{pmatrix}$$

Vi har därmed visat att det är linjär i varje rad och varje kolumn. □

Exempel 6.1. Låt

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

Vi kan uttrycka andra kolumnen som $2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$ och verifiera multilinjariteten genom att beräkna

$$\begin{pmatrix} 1 & 4 \\ 3 & 0 \end{pmatrix} + 2 \det \begin{pmatrix} 0 & 4 \\ 5 & 0 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 5 & 3 & 0 \end{pmatrix} + 2 \det \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 4 \\ 5 & 0 & 0 \end{pmatrix}.$$

Följdsats 6.2 (Determinant med nollrad)**.** *Om A är en $n \times n$-matris med en rad eller en kolumn består enbart av nollor, så är* $\det(A) = 0$.

Bevis. Övning (skriv nollraden som nollvektorn minus nollvektorn och använd multilinjäritet). □

Sats 6.3 (Kofaktorutveckling längs en rad)**.** *Låt $A \in M_{n\times n}(k)$ vara en $n\times n$-matris. Då gäller att*
$$\det(A) = D_i(A)$$
för alla $1 \le i \le n$.

Bevis. Vi kan anta att $n \ge 2$ eftersom fallet $n = 1$ är trivialt. Fallet $i = 1$ är klart per definition så vi antar också att $i > 1$. Skriv radvektor i som en linjärkombination $\sum_{j=1}^n a_{ij}e_j$ där e_j är vektorn med 1 i position j och nollor annars. Då har vi av multilinjäritet (Sats 6.1) att
$$D_1(A) = \sum_{j=1}^n a_{ij} \det(A(ij))$$
där $A(ij)$ är matrisen vi får genom att byta ut $i$te raden i A mot basvektorn e_j. Vi behöver därför visa att $\det(A(ij)) = (-1)^{i+j}\det(\widehat{A}_{ij})$. Vi använder induktion över n. Basfallet $n = 2$ är enkelt att visa, vilket vi lämnar som övning. Antag nu att $n > 2$ och att påståendet är sant för $(n-1)\times(n-1)$-matriser. Notera först att
$$\det(\widehat{A}_{ij}) = \sum_{l=1}^{j-1}(-1)^{1+l}a_{1l}\det\left(\widehat{(\widehat{A}_{ij})}_{1l}\right) + \sum_{l=j}^{n-1}(-1)^{l}a_{1(l+1)}\det\left(\widehat{(\widehat{A}_{ij})}_{1l}\right).$$
Induktionsantagandet och Följdsats 6.2 ger oss då att
$$\det(\widehat{A(ij)}_{1l}) = \begin{cases} \det(\widehat{A}_{1l}(i(j-1))) = (-1)^{i+j}\det\left(\widehat{(\widehat{A}_{ij})}_{1l}\right) & \text{om } l < j\,, \\ 0 & \text{om } l = j\,, \\ \det(\widehat{A}_{1l}(ij)) = (-1)^{i+j+1}\det\left(\widehat{(\widehat{A}_{ij})}_{1(l-1)}\right) & \text{om } l > j\,. \end{cases}$$
och vi får då att
$$\begin{aligned}
\det(A(ij)) &= \sum_{l=1}^{n}(-1)^{1+l}a_{1l}\det(\widehat{A(ij)}_{1l}) \\
&= \sum_{l=1}^{j-1}(-1)^{1+l+i+j}a_{1l}\det\left(\widehat{(\widehat{A}_{ij})}_{1l}\right) + \sum_{l=j+1}^{n}(-1)^{l+i+j}a_{1l}\det\left(\widehat{(\widehat{A}_{ij})}_{1(l-1)}\right) \\
&= \sum_{l=1}^{j-1}(-1)^{1+l+i+j}a_{1l}\det\left(\widehat{(\widehat{A}_{ij})}_{1l}\right) + \sum_{l=j}^{n-1}(-1)^{1+l+i+j}a_{1(l+1)}\det\left(\widehat{(\widehat{A}_{ij})}_{1l}\right) \\
&= (-1)^{i+j}\det(\widehat{A}_{ij})\,.
\end{aligned}$$

Det följer nu att

$$\det(A) = \sum_{j=1}^{n} a_{ij} \det(A(ij)) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\widehat{A}_{ij}) = D_i(A)\,. \qquad \square$$

Övning

Visa att om E_1, E_2 och E_3 är elementära matriser av typ 1, 2 respektive 3, och E_2 multiplicerar en rad med skalären λ så har vi $\det(E_1) = -1$, $\det(E_2) = \lambda$ och $\det(E_3) = 1$.

Följdsats 6.4 (Determinant och ofullständig rang)**.** *Om $A \in M_{n\times n}(k)$ har rang $\mathrm{rk}(A) < n$, så är $\det(A) = 0$.*

Bevis. Antag att $n > 1$. Vi visar först att $\det(A) = 0$ om A har två identiska rader: Antag att rad i och j är identiska. Basfallet $n = 2$ är trivialt, så vi antar att $n > 2$. Antag att påståendet gäller för $(n-1) \times (n-1)$-matris med två identiska rader. Vi kan nu använda kofaktorutveckling längs en rad som ej är i eller j och av induktion är varje term i kofaktorutvecklingen lika med noll (eftersom varje sådan term involverar en determinant av en $(n-1) \times (n-1)$-matris med två identiska rader). Vi har alltså visat att $\det(A) = 0$ och vi har fullföljt induktionssteget.

Antag nu att A är en godtycklig $n \times n$-matris med $\mathrm{rk}(A) < n$. Då finns det en rad i som är en linjärkombination av de andra raderna. Av multilinjäritet (Sats 6.1) kan vi då skriva $\det(A)$ som en summa av termer där varje term involverar en determinant av en matris med två identiska rader. Eftersom varje sådan term är lika med noll så är $\det(A) = 0$. $\square$

6.2. **Multiplikativitet och inverterbarhet.** Vi ska nu titta på fler egenskaper hos determinanter och se hur determinanter förhåller sig till inverterbarhet hos matriser.

Lemma 6.5. *Låt $E, B \in M_{n\times n}(k)$ vara två matriser där E är elementär. Då gäller att*

$$\det(EB) = \det(E)\det(B)\,.$$

Bevis. Av Övning 6.1 så räcker det att visa att en radoperation av typ 1, 2 respektive 3 (typ 2 med skalär λ) på matrisen B ändrar $\det(B)$ med en faktor -1, λ respektive 1.

Fallet med typ 1 följer av induktion (alternativt av multilinjäritet). Basfallet $n = 2$ är trivialt, så vi antar att $n > 2$. Antag att påståendet gäller för $(n-1) \times (n-1)$-matris med en radoperation av typ 1. Vi kan nu använda kofaktorutveckling längs en rad som ej är den som ändras. Varje term i kofaktorutvecklingen involverar en determinant av en $(n-1) \times (n-1)$-matris med en radoperation av typ 1, vilket ger oss den önskade faktorn -1 i varje term, dvs. $\det(B') = -\det(B)$.

Fallet med typ 2 följer genom radutveckling längs den rad som multiplicerats med λ och bryta ut λ från summan som erhålls. Detta ger då $\det(B') = \lambda \det(B)$.

Fallet med typ 3 följer även det av induktion. Basfallet $n = 2$ är en enkel beräkning så vi antar att $n > 2$. Antag att påståendet gäller för $(n-1) \times (n-1)$-matris med en radoperation av typ 3. Vi kan nu använda kofaktorutveckling längs en rad som ej ändrats och som lagts till en annan rad. Varje term i kofaktorutvecklingen involverar en determinant av en $(n-1) \times (n-1)$-matris med en radoperation av typ 3, vilken av induktionsantagandet sammanfaller determinanten av samma matris utan radoperationen. Därav får vi $\det(B') = \det(B)$. □

Sats 6.6 (Multiplikativitet)**.** *Låt $A, B \in M_{n\times n}(k)$ vara två matriser. Då gäller att*

$$\det(AB) = \det(A)\det(B)\,.$$

Bevis. Om A eller B har ofullständig rang så följer resultatet av Följdsats 6.4 [Determinant och ofullständig rang]. Antag nu att A och B är inverterbara matriser. Då är A och B båda produkter av elementära matriser. Med andra ord så räcker det att visa att påståendet gäller då A är en elementär matris, men detta vet vi redan är sant av Lemma 6.5. □

Följdsats 6.7 (Kofaktorutveckling längs en kolumn)**.** *Om $A \in M_{n\times n}(k)$ så har vi $\det(A) = D_j^T(A)$ för alla $1 \le j \le n$.*

Bevis. Av Följdsats 5.6 har A samma rang som A^T och av Följdsats 6.4 [Determinant och ofullständig-rang] följer det att påståendet är sant då A ej har full rang, dvs, då A ej är inverterbar. Vi kan därför anta att A är inverterbar och därmed en produkt av elementära matriser (Följdsats 5.4 [Elementär generering]): $A = E_k \cdots E_2 E_1$. Vi har då av Sats 6.6 [Multiplikativitet hos determinanter] och Övning 3.3 att

$$\begin{aligned}\det(A) &= \det(E_k \cdots E_2 E_1) \\ &= \det(E_k) \cdots \det(E_2) \det(E_1) \\ &= \det(E_k^T) \cdots \det(E_2^T) \det(E_1^T) \\ &= \det(E_1^T E_2^T \cdots E_k^T) \\ &= \det(A^T) \,.\end{aligned}$$

□

Sats 6.8 (Inverterbarhet)**.** *Låt $A \in M_{n\times n}(k)$ vara en matris. Då gäller att*

$$A \text{ är inverterbar} \iff \det(A) \neq 0 \,.$$

Vidare har vi att om A är inverterbar så är $\det(A^{-1}) = \det(A)^{-1}$.

Bevis. Antag först att A är inverterbar. Då finns det en matris B sådan att $AB = I$. Vi har då

$$\det(AB) = \det(I) = 1 \,.$$

Eftersom determinanten är multiplikativ så har vi

$$\det(A)\det(B) = 1 \,.$$

Det följer att $\det(A) \neq 0$.

Antag nu att $\det(A) \neq 0$. Då har A full rang (Följdsats 6.4) och är därmed inverterbar (Följdsats 5.7). Det sista påståendet i satsen följer av multiplikativiteten hos determinanten. □

Vi ska nu ge en formel för inversen av en inverterbar matris med hjälp av determinanter.

Definition 6.2. Låt A vara en $n \times n$-matris och låt $\widehat{A}_{ij}$ vara matrisen som fås genom att ta bort rad i och kolumn j från A. Då definierar vi den *klassiska adjunkten* eller *komplementmatrisen* av A som

$$\operatorname{adj}(A) = \begin{pmatrix} \det(\widehat{A}_{11}) & -\det(\widehat{A}_{21}) & \cdots & (-1)^{1+n}\det(\widehat{A}_{n1}) \\ -\det(\widehat{A}_{12}) & \det(\widehat{A}_{22}) & \cdots & (-1)^{2+n}\det(\widehat{A}_{n2}) \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{n+1}\det(\widehat{A}_{1n}) & (-1)^{n+2}\det(\widehat{A}_{2n}) & \cdots & (-1)^{n+n}\det(\widehat{A}_{nn}) \end{pmatrix} .$$

Sats 6.9 (Inversformel)**.** *Om $A \in M_{n\times n}(k)$ är inverterbar så gäller att*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)\,.$$

Bevis. Om vi multiplicerar rad i i A med kolumn i i $\operatorname{adj}(A)$ så ser vi direkt att resultatet är $\det(A)$ (kofaktorutveckling längs rad i). Om vi istället multiplicerar rad i i A med kolumn j i $\operatorname{adj}(A)$ för $i \neq j$ så ser vi att resultatet är noll eftersom detta sammanfaller med kofaktorutveckling längs rad i i den matris vi får från A genom att byta ut rad j mot rad i, dvs. en matris med två lika rader. Det följer att $A \operatorname{adj}(A) = \det(A)I$. □

Följdsats 6.10 (Cramers regel)**.** *Givet ett linjärt ekvationssystem $Ax = b$ där $A \in M_{n\times n}(k)$ är inverterbar och $b \in k^n$ så ges den unika lösningen av*

$$x_j = \frac{\det(A_j)}{\det(A)}\,.$$

där A_j är matrisen som fås genom att byta ut kolumn j i A mot vektorn b.

Bevis. Sats 6.9 [Inversformel] ger oss att $x = A^{-1}b = \frac{1}{\det(A)} \operatorname{adj}(A)b$. Om vi tittar på rad i i $\operatorname{adj}(A)b$ så ser vi att detta är lika med $\det(A_i)$ enligt kofaktorutveckling längs kolumn i. □

Diskussionsfrågor

(1) Använd Cramers regel för att lösa ekvationssystemet

$$\begin{cases} 2x_1 - 3x_3 = 1 \\ x_1 + 2x_2 = 6 \end{cases}$$

(2) Vad är den klassiska adjunkten av matrisen

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

7. Egenvärden och egenvektorer

Om vi betraktar en linjär avbildning $L\colon V \to V$ från ett vektorrum till sig självt så kan vi fråga oss om det finns en ordnad bas för V i vilken matrisrepresentationen för L är diagonal. Om så är fallet så är nästa fråga hur vi i så fall hittar denna bas. För att svara på dessa frågor behöver vi först introducera begreppen egenvärden och egenvektorer.

7.1. **Egenvärden och egenvektorer.** Vi kommer nu att fokusera på linjära avbildningar från ett vektorrum till sig självt.

Definition 7.1. En linjär avbildning $L\colon V \to V$ kallas för en *operator* på V. Identitetsoperatorn på V betecknas I.

Kom ihåg att varje $n \times n$-matris kan betraktas som en operator på k^n genom matris-vektor-multiplikation och matrisrepresentationen till en operator med avseende på en bas är alltid en *kvadratisk* matris. Därför gäller varje definition och sats som rör linjära operatorer även för kvadratiska matriser.

Definition 7.2. En operator $L\colon V \to V$ kallas *diagonaliserbar* om det finns en ordnad bas α för V så att matrisen $[L]_\alpha$ är diagonal.

Notera att om en operator är diagonal i en bas så bevarar den riktningen av basvektorerna, dvs. varje basvektor v_i i α uppfyller $L(v_i) = \lambda_i v_i$ för någon skalär λ_i.

Exempel 7.1. Matrisen

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

definierar en operator på $\mathbb{R}^2$ men också på $\mathbb{C}^2$. Som operator på $\mathbb{R}^2$ är den inte diagonaliserbar eftersom den roterar alla vektorer i $\mathbb{R}^2$ 45 grader moturs kring origo (och ändrar dess längd). Alltså bevaras ingen riktning i $\mathbb{R}^2$. Som operator på $\mathbb{C}^2$ är den däremot diagonaliserbar:

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} -1+i \\ 1+i \end{pmatrix} = (1+i) \begin{pmatrix} i \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} -1-i \\ 1-i \end{pmatrix} = (1-i) \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

Detta innebär att om vi väljer basen $\beta = \{(i,1), (-i,1)\}$ för $\mathbb{C}^2$ så är matrisen för A i denna bas diagonal:

$$[A]_\beta = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}.$$

Definition 7.3. Låt $L\colon V \to V$ vara en linjär avbildning. Om $L(v) = \lambda v$ för en nollskild vektor $v \in V$ och $\lambda \in k$ så kallas λ för ett *egenvärde* till L och v kallas för en *egenvektor* till L.

Diskussionsfrågor

Betrakta matrisen A i Exempel 7.1. Vilka egenvärden har A betraktat som
(1) en operator på $\mathbb{R}^2$?
(2) en operator på $\mathbb{C}^2$?

Om vi skriver om ekvationen $L(v) = \lambda v$ som $(L - \lambda I)v = 0$, där I är identitetsavbildningen, så är mängden av egenvektorerna till L med egenvärde λ precis $\ker(L - \lambda I) \setminus \{0\}$. Det vi då först behöver undersöka är när operatorn $(L - \lambda I)$ överhuvudtaget *har* en icke-trivial kärna. Detta sker precis när $(L - \lambda I)$ ej är inverterbar, dvs när $\det([L - \lambda I]_\beta) = 0$ för varje ordnad bas β för V. För att hitta alla egenvärden λ behöver vi alltså lösa ekvationen $\det([L - xI]_\beta) = 0$, där x är en obekant.

Sats 7.1 (Basoberoende för karaktäristiskt polynom)**.** *Polynomet* $\det([L - xI]_\beta)$ *är oberoende av valet av bas* β*.*

Bevis. Givet en annan ordnad bas γ så har vi att

$$[L - xI]_\gamma = ([I]_\gamma^\beta)^{-1} [L - xI]_\beta [I]_\gamma^\beta,$$

och

$$\begin{aligned}\det([L - xI]_\gamma) &= \det(([I]_\gamma^\beta)^{-1})\det([L - xI]_\beta)\det([I]_\gamma^\beta)\\ &= \det([I]_\gamma^\beta)^{-1}\det([L - xI]_\beta)\det([I]_\gamma^\beta)\\ &= \det([L - xI]_\beta)\,.\end{aligned}$$

□

Definition 7.4. Låt $L\colon V \to V$ vara en operator och β en ordnad bas för V. Det *karakteristiska polynomet* för L är polynomet $p_L(x) = \det([L]_\beta - xI) = \det([L - xI]_\beta)$.

Vi har alltså att egenvärdena för L är precis de skalärer λ som är rötter till det karakteristiska polynomet $p_L(x)$. När vi hittat alla lösningar till ekvationen $p_L(x) = 0$ så kan vi sedan hitta de tillhörande egenvektorerna till egenvärdet λ genom att hitta kärnan till $(L - \lambda I)$.

Exempel 7.2. På k^2 har vi operatorn som i standardbasen representeras av matrisen

$$M = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}.$$

Vi har då att det karaktäristiska polynomet är

$$\det\begin{pmatrix} 1-x & 4 \\ 2 & 3-x \end{pmatrix} = (1-x)(3-x) - (2)(4) = x^2 - 4x - 5$$

vilket har rötter $\lambda_1 = 5$ och $\lambda_2 = -1$. Dessa är alltså egenvärdena till M.
För att hitta motsvarande egenvektorer så löser vi ekvationerna $(M - 5I)v = 0$ och $(M + I)v = 0$. Kärnan till

$$(M - 5I) = \begin{pmatrix} -4 & 4 \\ 2 & -2 \end{pmatrix} \text{ respektive } (M + I) = \begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix}$$

spänns upp av $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ respektive $v_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, vilka alltså är egenvektorer tillhörande egenvärdena $\lambda_1 = 5$ respektive $\lambda_2 = -1$. Inversen till matrisen $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$ ges av

$$P^{-1} = \frac{1}{\det(P)}\begin{pmatrix} -1 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$$

och vi har att

$$P^{-1}MP = P^{-1}\left(5\begin{pmatrix}1\\1\end{pmatrix} \quad -1\begin{pmatrix}2\\-1\end{pmatrix}\right) = \begin{pmatrix}5 & 0\\0 & -1\end{pmatrix}.$$

Exempel 7.3. Låt V vara vektorrummet av polynom av grad högst 2. Låt $L\colon V \to V$ vara en linjär avbildning som ges av $L(p(x)) = p'(x)$, där $p'(x)$ är derivatan av $p(x)$. Eftersom derivatan sänker graden av polynomet så finns endast ett egenvärde till L, nämligen 0, och alla konstanta, noll-skilda polynom är egenvektorer till L. Alltså är L inte diagonaliserbar. Vi kan också se detta genom att betrakta matrisen till L i basen $\{1, x, x^2\}$:

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

vars karaktäristiska polynom är $-x^3$ som endast har roten 0 med multiplicitet 3. Dvs, det enda egenvärdet till L är 0.

Sats 7.2 (Karaktärisering av diagonaliserbarhet)**.** *En operator $L\colon V \to V$ är diagonaliserbar om och endast om det finns en ordnad bas av egenvektorer till L. Om α är en sådan bas så är matrisen $[L]_\alpha^\alpha$ diagonal med egenvärdena på diagonalen.*

Bevis. Övning. $\square$

Diskussionsfrågor

(1) Hur många egenvärden kan en operator på ett n-dimensionellt vektorrum ha som mest?
(2) Är varje egenvektor till en operator unik?
(3) Är egenvärdena till en operator alltid reella?
(4) Finns det alltid en bas av egenvektorer till en operator?

8. Diagonalisering och Cayley–Hamiltons sats

Vi ska nu titta på generella kriterier för diagonaliserbarhet av linjära operatorer.

Sats 8.1 (Linjärt oberoende av egenvektorer)**.** *Om* $\lambda_1, \ldots, \lambda_n$ *är olika egenvärden till en operator* $L\colon V \to V$ *så är de tillhörande egenvektorerna linjärt oberoende.*

Bevis. Vi använder induktion över n. För $n = 1$ är påståendet uppenbart. Antag att påståendet gäller för $n-1$ och låt $v_1, \ldots, v_n$ vara egenvektorer tillhörande de olika egenvärdena $\lambda_1, \ldots, \lambda_n$. Antag att det finns skalärer $a_1, \ldots, a_n$ sådana att

$$a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0\,.$$

Låt λ_n vara egenvärdet för v_n. Om vi applicerar $L - \lambda_n I$ på båda sidor får vi:

$$a_1(\lambda_1 - \lambda_n)v_1 + a_2(\lambda_2 - \lambda_n)v_2 + \ldots + a_n(\lambda_n - \lambda_n)v_n = 0\,.$$

Eftersom $\{v_1, \ldots, v_{n-1}\}$ är linjärt oberoende och egenvärdena alla är olika så får vi att $a_1 = a_2 = \ldots = a_{n-1} = 0$. Men då måste även $a_n = 0$. Därmed är $v_1, \ldots, v_n$ linjärt oberoende. □

Följdsats 8.2 (Kriterium för diagonaliserbarhet)**.** *Om en operator* $L\colon V \to V$ *har* $n = \dim V$ *olika egenvärden så är* L *diagonaliserbar.*

Bevis. Övning. □

Diskussionsfrågor

(1) Kan du ge ett exempel på en operator som är diagonaliserbar men *inte* har $n = \dim(V)$ olika egenvärden?
(2) Kan du ge ett exempel på en operator som inte är diagonaliserbar?

Definition 8.1. Om λ är ett egenvärde till en operator L så definierar vi *egenrummet* tillhörande λ som delrummet

$$E_\lambda = \ker(L - \lambda I)\,.$$

Om det karaktäristiska polynomet $p_L(x)$ till en operator L *splittrar* i linjära faktorer:

$$p_L(x) = \pm(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_l)$$

så kan samma rot λ förekomma flera gånger, dvs. $\lambda = \lambda_i$ för mer än ett i, även om L är diagonaliserbar (t.ex. gäller detta för identitetsavbildningen).

Å andra sidan kan vi tala om dimensionen av egenrummet E_λ. Detta leder oss till att definiera två typer av *multiplicitet*:

Definition 8.2. Låt L vara en operator på ett ändligt-dimensionellt vektorrum V och låt λ vara ett egenvärde till L.

- Den *algebraiska multipliciteten* $m_a(\lambda)$ av ett egenvärde λ är det största positiva heltalet m sådant att $(x-\lambda)^m$ är en faktor av det karaktäristiska polynomet $p_L(x)$.
- Den *geometriska multipliciteten* $m_g(\lambda)$ av λ är dimensionen av egenrummet E_λ.

Exempel 8.1. När $k = \mathbb{C}$ så splittrar varje polynom i linjära faktorer (algebrans fundamentalsats).

Lemma 8.3. *Om λ är ett egenvärde till en operator L så gäller att*

$$m_a(\lambda) \geq m_g(\lambda) \geq 1\,.$$

Bevis. Låt λ vara ett egenvärde till L och välj en ordnad bas för E_λ. Denna kan utvidgas till en ordnad bas för V. I denna bas har vi att matrisen för L har formen

$$\begin{pmatrix} \lambda I & * \\ 0 & B \end{pmatrix},$$

och determinanten av $L - xI$ är alltså

$$\det(L - xI) = \det\begin{pmatrix} (\lambda - x)I & * \\ 0 & B - xI \end{pmatrix} = (\lambda - x)^{m_g(\lambda)} \det(B - xI)\,. \qquad \square$$

Vi är nu redo att ge ett nödvändigt och tillräckligt villkor för diagonaliserbarhet.

Sats 8.4. *Låt L vara en operator på ett ändligt-dimensionellt vektorrum V med* $\dim V = n$. *Antag att det karaktäristiska polynomet $p_L(x)$ splittrar i linjära faktorer. Då är följande påståenden ekvivalenta:*

(1) L är diagonaliserbar.
(2) $m_a(\lambda) = m_g(\lambda)$ för varje egenvärde λ.

Bevis. Antag att L är diagonaliserbar. Då finns en ordnad bas β för V så att $[L]_\beta$ är diagonal. Den geometriska multipliciteten av ett egenvärde λ är då antalet gånger λ förekommer på diagonalen, vilket är lika med den algebraiska multipliciteten av λ eftersom det karaktäristiska polynomet då är $\prod_\lambda(\lambda - x)$ där produkten är över diagonalelementen i $[L]_\beta$.

Å andra sidan, anta att $m_a(\lambda) = m_g(\lambda)$ för varje egenvärde λ. Eftersom det karaktäristiska polynomet har grad $\sum_\lambda m_a(\lambda) = n$ så är summan av egenrummens dimensioner lika med n och vi kan välja en ordnad bas för V bestående av vektorer från egenrummen ([Linjärt-oberoende-egenvektorer] Sats 8.1). □

Diskussionsfrågor

Vilka av följande matriser är diagonaliserbara över $\mathbb{R}$ respektive $\mathbb{C}$?

(1) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
(2) $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$
(3) $\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$

8.1. **Cayley–Hamiltons sats.** Vi ska nu bevisa en klassisk sats inom linjär algebra, nämligen Cayley–Hamiltons sats som säger att varje operator L på ett ändligt-dimensionellt vektorrum V uppfyller sin egen karaktäristiska ekvation, det vill säga $p_L(L) = 0$ (nolloperatorn) där $p_L(x)$ är det karaktäristiska polynomet för L. Detta har en rad viktiga tillämpningar, inte minst inom talteori och algebraisk geometri (se denna länk).

Innan vi bevisar denna sats behöver vi två hjälpsatser:

Lemma 8.5. *Om $W \subseteq V$ är ett invariant delrum under L, det vill säga $L(W) \subseteq W$, så gäller att det karaktäristiska polynomet till restriktionen $L|_W : W \to W$ delar det karaktäristiska polynomet till L.*

Bevis. Välj en ordnad bas β_W för W och utvidga denna till en ordnad bas β för V. I denna bas har vi att matrisen för L har formen

$$[L]_\beta = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

där $A = [L|_W]_{\beta_W}$. Därmed är

$$\begin{aligned} p_L(x) &= \det([L]_\beta - xI) \\ &= \det \begin{pmatrix} A - xI & B \\ 0 & C - xI \end{pmatrix} \\ &= \det(A - xI) \det(C - xI) \\ &= p_{L|_W}(x) \det(C - xI) \,. \end{aligned}$$ □

Lemma 8.6. *Låt $L : V \to V$ vara en operator på ett ändligt-dimensionellt vektorrum V. Låt $v \in V$ vara en icke-nollvektor och sätt $W = \mathrm{Span}(v, L(v), L^2(v), \dots)$. Om $\dim W = m$ så är $\{v, L(v), \dots, L^{m-1}(v)\}$ en bas för W och om $a_0, \dots, a_{m-1}$ är den unika lösningen till ekvationen*

$$a_0 v + a_1 L(v) + \cdots + a_{m-1} L^{m-1}(v) + L^m(v) = 0$$

så är $(-1)^m(a_0 + a_1 x + \cdots + a_{m-1} x^{m-1} + x^m) = 0$ det karaktäristiska polynomet för $L|_W$.

Bevis. Eftersom $\dim W = m$ så är $\{v, L(v), \dots, L^{m-1}(v)\}$ linjärt oberoende och därmed en bas för W. I denna bas har vi att matrisen för $L|_W$ är

$$[L|_W]_\beta = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{m-1} \end{pmatrix} .$$

Därmed är det karaktäristiska polynomet

$$p_{L|_W}(x) = \det([L|_W]_\beta - xI) = (-1)^m(a_0 + a_1 x + a_2 x^2 + \cdots + a_{m-1} x^{m-1} + x^m) \,.$$ □

Vi är nu redo att bevisa Cayley–Hamiltons sats.

Sats 8.7 (Cayley–Hamiltons sats)**.** *Låt $L\colon V \to V$ vara en operator på ett ändligt-dimensionellt vektorrum V. Om $p_L(x)$ är det karaktäristiska polynomet för L så gäller att $p_L(L) = 0$.*

Bevis. Om $V = \{0\}$ så är påståendet uppenbart. Välj en icke-nollvektor $v \in V$ och sätt $W = \operatorname{Span}(v, L(v), L^2(v), \ldots)$. Då är W ett invariant delrum under L och om $\dim W = m$ så är $\{v, L(v), \ldots, L^{m-1}(v)\}$ en bas för W. Om $a_0, \ldots, a_{m-1}$ är den unika lösningen till ekvationen

$$a_0 v + a_1 L(v) + \cdots + a_{m-1} L^{m-1}(v) + L^m(v) = 0$$

så är $(-1)^m(a_0 + a_1 x + \cdots + a_{m-1}x^{m-1} + x^m)$ det karaktäristiska polynomet för $L|_W$ vilket delar $p_L(x)$. Därmed finns ett polynom $q(x)$ så att $p_L(x) = q(x)(-1)^m(a_0 + a_1 x + \cdots + a_{m-1}x^{m-1} + x^m)$. Detta ger då att

$$p_L(L)(v) = q(L)(-1)^m(a_0 I + a_1 L + \cdots + a_{m-1}L^{m-1} + L^m)(v) = 0\,. \qquad \square$$

Exempel 8.2. Låt $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ vara den linjära avbildning som i standardbasen ε representeras av matrisen

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

Det karaktäristiska polynomet är $p_L(x) = (1-x)(3-x)$ och vi har att

$$\begin{aligned} [p_L(L)]_\varepsilon &= \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\right)\left(\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\right) \\ &= \begin{pmatrix} 0 & -2 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} 2 & -2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Övning

Visa att operatorn L på $\mathbb{R}^2$ som i standardbasen ε representeras av matrisen

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

uppfyller $p_L(L) = 0$, genom att beräkna matrisen $p_L(A) = [p_L(L)]_\varepsilon$.

9. Inre produkt och Gram–Schmidts metod

En viktig egenskap hos det euklidiska rummet $\mathbb{R}^n$ är att vi har en inre produkt:

$$(v, w) = v^T w = \sum_{i=1}^{n} v_i w_i \,.$$

Denna inre produkt hjälper oss att mäta avstånd och vinklar genom att projicera vektorer på varandra. I detta kapitel ska vi axiomatisera begreppet inre produkt så att vi kan använda begreppet för mer generella vektorrum.

9.1. **Inre produkt.** Från och med nu antar vi att k är antingen $\mathbb{R}$ eller $\mathbb{C}$. Detta för att vi kommer att använda oss av komplex konjugering; $\overline{a+bi} = a - bi$ (vilket är identiteten på $\mathbb{R}$).

Definition 9.1. En inre produkt på ett vektorrum V över k är en avbildning $\langle \cdot, \cdot \rangle \colon V \times V \to k$ som är:

(1) `Symmetrisk:` $\langle v, w\rangle = \overline{\langle w, v\rangle}$ för alla $v, w \in V$.
(2) `Linjär:` $\langle av + bu, w\rangle = a\langle v, w\rangle + b\langle u, w\rangle$ för alla $u, v, w \in V$ och $a, b \in k$.
(3) `Positivt definit:` $\langle v, v\rangle > 0$ för alla $v \in V$, $v \neq 0$.

Ett vektorrum med en inre produkt kallas för ett *inre produktrum*.

Diskussionsfrågor

(1) Hur ser linjäriteten ut i andra argumentet?
(2) Varför är $\langle v, v\rangle$ reellt för alla $v \in V$?
(3) Är den vanliga inre produkten på $\mathbb{R}^n$ en inre produkt enligt denna definition?

Exempel 9.1. Betrakta vektorrummet $k[x]$ av polynom med koefficienter i k. Vi kan definiera en inre produkt på detta rum genom

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx$$

för alla $f, g \in k[x]$, där $[a, b]$ är ett intervall i $\mathbb{R}$.

Exempel 9.2. Vektorrummet av 2×2-matriser kan ges en inre produkt genom

$$\left\langle \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \right\rangle = \operatorname{tr}\left(\begin{pmatrix} \overline{a_{11}} & \overline{a_{21}} \\ \overline{a_{12}} & \overline{a_{22}} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \right) = \sum_{i,j} \overline{a_{ij}} b_{ij} \,.$$

för alla $(a_{ij}), (b_{ij}) \in M_{2\times 2}(k)$, där A^* är det Hermiteska (komplexkonjugerade) transponatet av A och tr betecknar spåret av en matris. Denna definition fungerar inte bara för $n = 2$ utan för godtyckliga n och kallas *Frobenius inre produkt*.

Övning

Låt V vara ett inre produktrum över k. Om $v_1, v_2, v_3 \in V$ och $a \in k$, visa att följande likheter gäller:

(1) $\langle v_1, 0\rangle = \langle 0, v_1\rangle = 0$,
(2) $\langle v_1, v_1\rangle = 0$ om och endast om $v_1 = 0$,
(3) $\langle v_1, v_2 + v_3\rangle = \langle v_1, v_2\rangle + \langle v_1, v_3\rangle$,
(4) $\langle v_1, av_2\rangle = \overline{a}\langle v_1, v_2\rangle$,
(5) Om $\langle v_1, v_2\rangle = \langle v_1, v_3\rangle$ för alla $v_1 \in V$ så är $v_2 = v_3$.

Definition 9.2. Låt V vara ett inre produktrum över k och låt v, w vara vektorer i V. Vi definierar *projektionen* av v på w som

$$\operatorname{proj}_w v = \frac{\langle v, w\rangle}{\langle w, w\rangle} w \,.$$

Givet en inre produkt kan vi definiera längden eller normen av en vektor:

Definition 9.3. Låt V vara ett inre produktrum över k. Vi definierar *normen* av en vektor $v \in V$ som

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

Sats 9.1 (Normens egenskaper)**.** *Låt V vara ett inre produktrum över k. Då gäller följande egenskaper för normen:*

(1) $\|v\| \geq 0$ för alla $v \in V$, och $\|v\| = 0$ om och endast om $v = 0$.
(2) $\|av\| = |a|\|v\|$ för alla $v \in V$ och $a \in k$.

(3) (Cauchy–Schwarz olikhet) $|\langle v, w\rangle| \leq \|v\|\|w\|$ *för alla* $v, w \in V$*, med likhet om och endast om* v *och* w *är linjärt beroende.*
(4) (Triangelolikheten) $\|v+w\| \leq \|v\| + \|w\|$ *för alla* $v, w \in V$.

Bevis. De första två egenskaperna lämnar vi som övning att bevisa. För att bevisa Cauchy–Schwarz olikhet så tittar vi på $v - \operatorname{proj}_w v$:

$$\begin{aligned}
0 &\leq \|v - \operatorname{proj}_w v\| \\
&= \langle v - \operatorname{proj}_w v, v - \operatorname{proj}_w v\rangle \\
&= \langle v, v\rangle - \langle v, \operatorname{proj}_w v\rangle - \langle \operatorname{proj}_w v, v\rangle + \langle \operatorname{proj}_w v, \operatorname{proj}_w v\rangle \\
&= \|v\|^2 - \overline{\left(\frac{\langle v, w\rangle}{\langle w, w\rangle}\right)}\langle v, w\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, v\rangle + \overline{\left(\frac{\langle v, w\rangle}{\langle w, w\rangle}\right)}\frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, w\rangle \\
&= \|v\|^2 - \frac{|\langle v, w\rangle|^2}{\|w\|^2}\,,
\end{aligned}$$

vilket bevisar olikheten eftersom $\|v\|$, $\|w\|$ och $|\langle v, w\rangle|$ alla är icke-negativa reella tal.

Med hjälp av Cauchy–Schwarz olikhet kan vi nu bevisa triangelolikheten:

$$\begin{aligned}
\|v+w\|^2 &= \langle v+w, v+w\rangle \\
&= \langle v, v\rangle + \langle v, w\rangle + \overline{\langle v, w\rangle} + \langle w, w\rangle \\
&= \|v\|^2 + 2\mathrm{Re}(\langle v, w\rangle) + \|w\|^2 \\
&\leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2\,.
\end{aligned}$$

□

Definition 9.4. Två vektorer v, w i ett inre produktrum V sägs vara *ortogonala* om $\langle v, w\rangle = 0$. En delmängd av V sägs vara *ortogonal* om alla par av olika vektorer i delmängden är ortogonala. En delmängd av V sägs vara *ortonormal* om den är ortogonal och alla vektorer i delmängden har norm 1.

Exempel 9.3. Vektorerna 1 och x är ortogonala i inre produktrummet $\mathbb{R}[x]$ med inre produkten $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$ eftersom

$$\langle 1, x\rangle = \int_{-1}^{1} 1 \cdot x\,dx = \left[\frac{x^2}{2}\right]_{-1}^{1} = 0\,.$$

Delmängden $\{1, x\}$ är ortogonal och delmängden $\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x\}$ är ortonormal. T.ex. har vi

$$\left\|\sqrt{\frac{3}{2}}x\right\|^2 = \int_{-1}^{1} \left(\sqrt{\frac{3}{2}}x\right)^2 dx = \frac{3}{2}\int_{-1}^{1} x^2\, dx = \frac{3}{2}\cdot\frac{2}{3} = 1\,.$$

9.2. **Gram–Schmidts metod.** Gram–Schmidts metod är en algoritm för att omvandla en uppräknelig mängd linjärt oberoende vektorer $\{v_1, v_2, \ldots\}$ i ett *inre produktrum* till en ortogonal/ortonormal bas för $\mathrm{Span}(v_1, v_2 \ldots)$. I synnerhet, givet en ordnad uppräknelig bas β för ett inre produktrum V så kan vi använda Gram–Schmidts metod för att omvandla β till en uppräknelig ortogonal/ortonormal bas för V.

Vi börjar med en sats som beskriver hur man, givet en mängd linjärt oberoende vektorer $\{v_1, v_2, \ldots\}$ och $w \in \mathrm{Span}(v_1, v_2, \ldots)$, kan hitta den unika beskrivningen av w som en linjärkombination av element i $\{v_1, v_2, \ldots\}$.

Sats 9.2 (Projektion på bas)**.** *Låt V vara ett inre produktrum och låt β vara en ortogonal bas för V. För varje $v \in V$ så har vi att $\langle v, w\rangle = 0$ för alla utom ändligt många $w \in \beta$ och*

$$v = \sum_{w\in\beta} \frac{\langle v, w\rangle}{\langle w, w\rangle} w$$

(detta är alltså en ändlig summa).

Bevis. Av Sats 2.2 [Karaktärisering av bas] vet vi att varje $v \in V$ kan skrivas som en linjärkombination av (ändligt många) vektorer i β på ett unikt sätt:

$$v = a_1w_1 + a_2w_2 + \cdots + a_nw_n\,.$$

Vi får då att

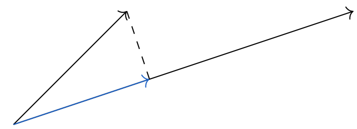
$$\frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} w_i = a_iw_i\,,$$

vilket ger att

$$\sum_{w\in\beta} \frac{\langle v, w\rangle}{\langle w, w\rangle} w = a_1w_1 + a_2w_2 + \cdots + a_nw_n = v\,. \qquad \square$$

Följdsats 9.3. *Varje ortogonal delmängd av ett inre produktrum, bestående av nollskilda vektorer, är linjärt oberoende.*

Idén bakom Gram–Schmidts metod är att, givet en bas på formen $\{v_1, v_2, \dots\}$, successivt ortogonalisera vektorerna $v_1, v_2, \dots$ genom att subtrahera projektioner på de tidigare vektorerna.



Sats 9.4 (Gram–Schmidts metod)**.** *Låt V vara ett inre produktrum och $\{v_1, v_2, \dots\}$ vara en bas för V. Definiera nu $u_1 = v_1$ och*

$$u_i = v_i - \sum_{j=1}^{i-1} \frac{\langle v_i, u_j \rangle}{\langle u_j, u_j \rangle} u_j \,.$$

Då är $\{u_1, u_2, \dots\}$ en ortogonal bas för V (och $\{\|u_1\|^{-1}u_1, \|u_2\|^{-1}u_2, \dots\}$ en ortonormal bas för V).

Bevis. Antag att vi redan ortogonaliserat $v_1, \dots, v_{n-1}$ för något positivt heltal $n \geq 2$. Vi visar nu att u_n är ortogonal mot $u_1, \dots, u_{n-1}$. Vi har

$$\langle u_n, u_j \rangle = \langle v_n, u_j \rangle - \sum_{i=1}^{n-1} \frac{\langle v_n, u_i \rangle}{\langle u_i, u_i \rangle} \langle u_i, u_j \rangle \,.$$

Eftersom $u_1, \dots, u_{n-1}$ är ortogonala så är $\langle u_i, u_j \rangle = 0$ för $i \neq j$. Vi får då

$$\langle u_n, u_j \rangle = \langle v_n, u_j \rangle - \frac{\langle v_n, u_j \rangle}{\langle u_j, u_j \rangle} \langle u_j, u_j \rangle = 0 \,.$$

Detta visar att u_n är ortogonal mot $u_1, \dots, u_{n-1}$ och därmed är $\{u_1, u_2, \dots\}$ en ortogonal bas för V. □

Övning

Använd Gram–Schmidts metod för att hitta en ortonormal bas för $k[x]_{\leq 3}$ om vi startar med basen $\{1, x, x^2, x^3\}$.

9.3. **Matrisrepresentationer av inre produkter.** Givet en inre produkt $f = \langle \cdot, \cdot \rangle$ på ett vektorrum V och en bas $\beta = \{v_1, v_2, \dots, v_n\}$ för V kan vi representera inre produkten med hjälp av en matris. Vi definierar matrisen $[f]_\beta$ genom

$$[f]_\beta = \begin{pmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \cdots & \langle v_1, v_n\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \cdots & \langle v_2, v_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1\rangle & \langle v_n, v_2\rangle & \cdots & \langle v_n, v_n\rangle \end{pmatrix}.$$

Definition 9.5. Låt A vara en matris med koefficienter i $\mathbb{C}$. Då är A

(1) *Hermitesk* eller *självadjungerad* om $A = A^*$, där A^* är det Hermiteska (komplexkonjugerade) transponatet av A, och
(2) *positivt definit* om alla dess egenvärden är positiva reella tal.

Sats 9.5 (Matrisrepresentation av en inre produkt)**.** *Låt V vara ett vektorrum av ändlig dimension och låt $\beta = \{v_1, \ldots, v_n\}$ vara en bas. Då är en avbildning $f\colon V \times V \to k$ en inre produkt på V om och endast om matrisen $[f]_\beta$ är Hermitesk och positivt definit. Å andra sidan, om A är en Hermitesk och positivt definit matris så är avbildningen*

$$\begin{aligned} f_A\colon V \times V &\to k \\ (v, w) &\mapsto [v]_\beta^* A [w]_\beta \end{aligned}$$

en inre produkt på V.

Bevis. Om f är en inre produkt så verifierar vi enkelt från definitionen att $[f]_\beta$ är Hermitesk. I kapitel 11 kommer vi se att Hermiteska matriser är diagonaliserbara med reella egenvärden och egenvektorerna är ortogonala. Om vi nu låter β vara en sådan bas av egenvektorer så är $[f]_\beta$ diagonal med egenvärdena på diagonalen, vilka enligt ovan, är på formen $\langle v_i, v_i\rangle$. Dvs alla egenvärden är reella och positiva.

Å andra sidan, om A är Hermitesk så är $[v]_\beta^* A [w]_\beta = \overline{[w]_\beta^* A [v]_\beta}$ för alla $v, w \in V$. Linjäriteten följer från matrisalgebra och positivt definit följer från att $[v]_\beta^* A [v]_\beta > 0$ för alla $v \neq 0$, eftersom varje v är en linjärkombination av egenvektorerna till A, vilka har positiva egenvärden eftersom A är Hermitesk (Övning 11). $\square$

10. Ortogonala komplementet och adjunkten till en operator

Givet en delmängd S av $\mathbb{R}^n$ så kan vi tala den delmängd som är vinkelrät mot alla vektorer i S. Detta koncept går att generalisera till godtyckliga inre produktrum.

Definition 10.1. Låt V vara ett inre produktrum och låt $S \subseteq V$ vara en delmängd. Då definierar vi det *ortogonala komplementet* till S som

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

Notera att $S^{\perp}$ inte förändras om vi byter S mot $\mathrm{Span}(S)$. Kan därför tala om delrum istället för delmängd i definitionen ovan.

Övning

Visa att $S^{\perp}$ är ett delrum av V.

Sats 10.1 (Ortogonala komplementet)**.** *Låt V vara ett inre produktrum och W ett delrum av V. För varje vektor $v \in V$ så existerar en unik uppdelning*

$$v = w + w^{\perp}$$

där $w \in W$ och $w^{\perp} \in W^{\perp}$.

Bevis. Givet en bas $\{w_1, \ldots, w_m\}$ och en bas $\{w_1^{\perp}, \ldots, w_n^{\perp}\}$ för $W^{\perp}$ så är

$$\{w_1, \ldots, w_m, w_1^{\perp}, \ldots, w_n^{\perp}\}$$

en bas för V, eftersom ortogonala vektorer är linjärt oberoende. Av Sats 9.2 [Projektion på bas] och Sats 2.2 [Karaktärisering av bas] har vi då att v, på ett unikt sätt, kan skrivas som

$$v = \sum_{j=1}^{m} \frac{\langle v, w_j \rangle}{\langle w_j, w_j \rangle} w_j + \sum_{i=1}^{n} \frac{\langle v, w_i^{\perp} \rangle}{\langle w_i^{\perp}, w_i^{\perp} \rangle} w_i^{\perp}. \qquad \square$$

Diskussionsfrågor

(1) Stämmer det att $(W^{\perp})^{\perp} = W$?
(2) Hur är $\dim(W)$ och $\dim(W^{\perp})$ relaterade?
(3) Vad är $\{0\}^{\perp}$?

Övning

Dela upp $x^2 \in \mathbb{R}[x]$ som en summa av en vektor i $W = \mathrm{Span}(1, x)$ och en i $W^{\perp}$ då vi har den inre produkten

$$\langle f(x), g(x)\rangle = \int_{-1}^{1} f(x)g(x)\, dx\, .$$

10.1. **Hermiteska adjunkten till en operator.** I detta kapitel ska vi visa att, givet inre produktrum V och W av ändlig dimension och en linjär avbildning $L\colon V \to W$, kan vi definiera en ny operator $L^*\colon W \to V$ som uppfyller att

$$\langle L(v), w\rangle = \langle v, L^*(w)\rangle$$

för alla $v \in V$ och $w \in W$. Denna operator kallas *Hermiteska adjunkten* (eller bara *adjunkten*) till L. Detta begrepp bör inte förväxlas med den *klassiska adjunkten* eller *komplementmatrisen* till en kvadratisk matris från Kapitel 6.

Definition 10.2. Låt V vara ett vektorrum över en kropp k. Då definierar vi *duala rummet* V^* som mängden av alla linjära avbildningar från V till k:

$$V^{\vee} = \mathrm{Hom}_k(V, k) = \{f\colon V \to k \mid f \text{ är linjär}\}\, .$$

(En annan notation för $V^{\vee}$ är V^*).

Sats 10.2 (Dual och dimension)**.** *Om V är ett ändligt-dimensionellt vektorrum över en kropp k så är $V^{\vee}$ också ett ändligt-dimensionellt vektorrum över k med* $\dim(V^{\vee}) = \dim(V)$*.*

Bevis. Detta följer av att om $\{v_1, \ldots, v_n\}$ är en bas för V så är

$$\{v_1^{\vee}, \ldots, v_n^{\vee}\} \subseteq V^{\vee}$$

en bas för $V^\vee$, där $v_i^\vee\colon V \to k$ är den linjära avbildning som ges av

$$v_i^\vee(v_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \qquad \square$$

Av axiomen för inre produkter följer att för varje $w \in V$ så är avbildningen $\langle -, w\rangle\colon V \to k$ given av $v \mapsto \langle v, w\rangle$ linjär. Vi kan definiera en linjär avbildning av vektorrum $V \to V^\vee$ genom att skicka w till $\langle -, \overline{w}\rangle$. Följande sats är *falsk* om vi inte antar att V är av ändlig dimension.

Sats 10.3 (Dualitetssatsen)**.** *Om V är ett inre produktrum av ändlig dimension så är avbildningen $V \to V^\vee$ given av $v \mapsto \langle -, \overline{v}\rangle$ en isomorfi av vektorrum.*

Bevis. Detta följer av att avbildningen $\langle -, \overline{v}\rangle$ är noll om och endast om $v = 0$ (av axiomen för inre produkter), dvs det enda element som avbildas på nollavbildningen är noll. Av [Injektivitetssatsen] (Sats 3.2) är avbildningen därmed injektiv. Av Sats 3.4 [Bijektionssatsen] är den också surjektiv eftersom V och $V^\vee$ har samma dimension. Alltså är avbildningen en isomorfi. $\square$

Övning

Visa att dualen till $k[x]$ är isomorf med $k^{\mathbb{N}}$, vilken i sin tur är isomorf med vektorrummet av alla formella serier $k[[x]]$.

Antag nu att vi har en linjär avbildning $L\colon V \to W$ mellan inre produktrum av ändlig dimension. Detta ger nu en avbildning $\langle L(-), w\rangle$ från V till k för varje $w \in W$, vilken vi ser är linjär genom att uttrycka den som kompositionen $\langle -, w\rangle \circ L$. Av [Dualitetssatsen] (Sats 10.3) vet vi då att det existerar en unik vektor $v' \in V$ sådan att $\langle L(-), w\rangle = \langle -, v'\rangle$. Vi kommer att definiera adjunkten L^* genom att sätta $L^*(w) = v'$:

Definition 10.3. Låt $L\colon V \to W$ vara en linjär avbildning mellan inre produktrum av ändlig dimension. Då definierar vi den *Hermiteska adjunkten* (*adjunkten*, *adjungerade avbildningen*) till L som den unika funktion $L^*\colon W \to V$ sådan att

$$\langle L(-), w\rangle = \langle -, L^*(w)\rangle$$

för alla $w \in W$.

Sats 10.4 (Adjunkten är linjär)**.** *Om $L\colon V \to W$ är en linjär avbildning mellan inre produktrum av ändlig dimension så är adjunkten $L^*\colon W \to V$ också en linjär avbildning.*

Bevis. För alla $v, w \in W$ och $a \in k$ så har vi

$$\begin{aligned}\langle -, L^*(av+w)\rangle &= \langle L(-), av+w\rangle \\ &= \overline{a}\langle L(-), v\rangle + \langle L(-), w\rangle \\ &= \overline{a}\langle -, L^*(v)\rangle + \langle -, L^*(w)\rangle \\ &= \langle -, aL^*(v) + L^*(w)\rangle\,.\end{aligned}$$

Av Övning 9.1 (5) följer då att $L^*(av+w) = aL^*(v) + L^*(w)$, dvs L^* är linjär. □

Nu ska vi se hur matrisen för L^* relaterar till matrisen för L med avseende på ortonormala baser. Anledningen till att vi behöver använda oss av *ortonormala* baser är att själva matrisoperationen $A \mapsto A^*$ går att definiera *utan* att ha valt inre produkter, medan L^* kommer att bero på valet av sådana. De inre produkterna påverkar alltså vilka ortonormala baser som finns tillgängliga.

Diskussionsfrågor

(1) (Kuggfråga) Vad är adjunkten till $z \mapsto \overline{z}$ från $\mathbb{C}$ till $\mathbb{C}$ med den vanliga inre produkten $\langle z_1, z_2\rangle = z_1\overline{z_2}$?
(2) Vad är adjunkten till $z \mapsto iz$ från $\mathbb{C}$ till $\mathbb{C}$ med den vanliga inre produkten $\langle z_1, z_2\rangle = z_1\overline{z_2}$?
(3) Vad är adjunkten till $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ (med standard inre produkt) given av matrisen
$$[L]_\varepsilon = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}?$$

Exempel 10.1. Låt $L\colon \mathbb{C}^2 \to \mathbb{C}^2$ vara den linjära avbildningen given (i standardbasen) av matrisen
$$[L]_\varepsilon = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Med den vanliga inre produkten för $\mathbb{C}^2$ har vi då att

$$\begin{aligned}\langle L(x,y),(x',y')\rangle &= \langle (ax+by,cx+dy),(x',y')\rangle \\ &= (ax+by)\overline{x'} + (cx+dy)\overline{y'} \\ &= x\overline{(\overline{a}x'+\overline{c}y')} + y\overline{(\overline{b}x'+\overline{d}y')} \\ &= \langle (x,y),(\overline{a}x'+\overline{c}y',\overline{b}x'+\overline{d}y')\rangle\,,\end{aligned}$$

dvs, adjunkten till L har matrisen

$$[L^*]_\varepsilon = \begin{pmatrix}\overline{a} & \overline{c} \\ \overline{b} & \overline{d}\end{pmatrix} = ([L]_\varepsilon)^*\,.$$

Sats 10.5 (Matris till adjunkt)**.** *Låt V och W vara inre produktrum av ändlig dimension och låt β och γ vara ortonormala baser för V respektive W. Om $L\colon V\to W$ är en linjär avbildning så gäller att*

$$[L^*]^\beta_\gamma = ([L]^\gamma_\beta)^*\,.$$

Bevis. Vi tittar på hur L^* verkar på basvektorerna i γ. Antag att $\beta = \{v_1,\dots,v_n\}$ och $\gamma = \{w_1,\dots,w_m\}$. Då har vi att

$$L(v_j) = \sum_{i=1}^n a_{ij}w_i$$

där $[L]^\gamma_\beta = (a_{ij})$. Av definitionen av adjunkten har vi då att

$$\begin{aligned}\langle v_j, L^*(w_l)\rangle &= \langle L(v_j), w_l\rangle \\ &= \left\langle \sum_{i=1}^n a_{ij}w_i, w_l\right\rangle \\ &= a_{lj}\langle w_l,w_l\rangle \\ &= a_{lj}\,,\end{aligned}$$

eftersom γ är ortonormal. Detta säger alltså att $L^*(w_l) = \sum_{j=1}^n \overline{a_{lj}}v_j$, vilket ger att $[L^*]^\beta_\gamma = (\overline{a_{ji}}) = ([L]^\gamma_\beta)^*$. □

Sats 10.6 (Egenskaper hos adjunkt)**.** *Låt V och W vara inre produktrum av ändlig dimension och låt $L\colon V\to W$ vara en linjär avbildning. Då gäller följande:*

(1) $(L^*)^* = L$.
(2) $(L+M)^* = L^* + M^*$ *för alla linjära operatorer* $M\colon V \to V$.
(3) $(cL)^* = \overline{c}L^*$ *för alla* $c \in k$.
(4) $(L \circ M)^* = M^* \circ L^*$ *för alla linjära operatorer* $M\colon V \to V$.
(5) Om L *är inverterbar så är även* L^* *inverterbar och* $(L^{-1})^* = (L^*)^{-1}$.

Bevis. Vi bevisar ett påstående i taget (och använder oss av Övning 9.1 (5)):

(1) Detta följer direkt av [Matris till adjunkt] Sats 10.5 efter att vi valt ortonormala baser.
(2) För alla $v \in V, w \in W$ har vi $\langle v, (L+M)^*(w)\rangle = \langle (L+M)(v), w\rangle = \langle L(v), w\rangle + \langle M(v), w\rangle = \langle v, L^*(w)\rangle + \langle v, M^*(w)\rangle = \langle v, (L^* + M^*)(w)\rangle$.
(3) För alla $v \in V, w \in W$ har vi $\langle v, (cL^*)(w)\rangle = \overline{c}\langle v, L^*(w)\rangle$.
(4) Detta följer av motsvarade påstående för matriser och (Sats 10.6 [Matris till adjunkt]).
(5) För alla $v, w \in V$ har vi $\langle v, w\rangle = \langle v, L(L^{-1}(w))\rangle = \langle L^*(v), L^{-1}(w)\rangle = \langle ((L^{-1})^* L^*)(v), w\rangle$, vilket ger att $(L^{-1})^* = (L^*)^{-1}$.

$\square$

11. Normala och självadjungerade avbildningar

I Kapitel 7 såg vi att en operator $L\colon V \to V$ är diagonaliserbar om och endast om det finns en ordnad bas av egenvektorer till L och i Kapitel 8 studerade vi kriterier för diagonaliserbarhet. I detta avsnitt ska vi studera när ett inre produktrum V har en *ortonormal* bas av egenvektorer till en operator $L\colon V \to V$.

Definition 11.1. En linjär operator $L\colon V \to V$ på ett inre produktrum V kallas *normal* om $LL^* = L^*L$.

Sats 11.1 (Egenskaper hos normala operatorer)**.** *Låt L vara en normal operator på ett inre produktrum V. Då gäller följande:*

(1) $\|L(v)\| = \|L^*(v)\|$ för alla $v \in V$.
(2) Operatorn $L - cI$ är normal för alla $c \in k$.
(3) Om v är en egenvektor till L med egenvärde λ så är v också en egenvektor till L^* med egenvärde $\overline{\lambda}$.
(4) Egenvektorer till L som hör till olika egenvärden är ortogonala.

Bevis. (1) Vi har att

$$\|L(v)\|^2 = \langle L(v), L(v)\rangle = \langle v, L^*L(v)\rangle = \langle v, LL^*(v)\rangle = \langle L^*(v), L^*(v)\rangle = \|L^*(v)\|^2 .$$

(2) Vi har att

$$\begin{aligned}(L-cI)(L-cI)^* &= (L-cI)(L^* - \overline{c}I)\\ &= LL^* - cL^* - \overline{c}L + |c|^2 I\\ &= (L-cI)^*(L-cI)\,.\end{aligned}$$

(3) Antag att v är en egenvektor till L med egenvärde λ. Eftersom $L - \lambda I$ är normal av (2), så har vi, av (1), att

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

vilket ger att $L^*(v) = \overline{\lambda}v$. (4) Antag att v och w är egenvektorer till L med egenvärden λ respektive μ, där $\lambda \neq \mu$. Då har vi att

$$\lambda\langle v, w\rangle = \langle L(v), w\rangle = \langle v, L^*(w)\rangle = \mu\langle v, w\rangle\,.$$

Eftersom $\lambda \neq \mu$ så måste alltså $\langle v, w\rangle = 0$.

$\square$

Sats 11.2 (Karaktärisering av normala operatorer)**.** *En operator L på ett inre produktrum över $\mathbb{C}$ (eller mer generellt sådan att $p_L(x)$ splittrar) är normal om och endast om det finns en ortonormal bas av egenvektorer till L.*

Bevis. Antag först att L är normal och att det karaktäristiska polynomet splittrar. Vi använder oss av induktion över $\dim V$. Om $\dim V = 1$ så är varje vektor i V en egenvektor till L och vi kan välja en ortonormal bas bestående av en egenvektor. Antag nu att påståendet gäller för alla inre produktrum med dimension mindre än n och låt V vara ett inre produktrum med $\dim V = n$. Eftersom $p_L(x)$ splittrar i linjära faktorer så finns ett egenvärde λ till L. Låt v vara en normerad egenvektor till L med egenvärde λ. För varje $w \in \operatorname{Span}(v)^\perp$ gäller att $L(w) \in \operatorname{Span}(v)^\perp$:

$$\langle L(w), v\rangle = \langle w, L^*(v)\rangle = \langle w, \overline{\lambda} v\rangle = 0\,.$$

Låt L' vara restriktionen av L till $\operatorname{Span}(v)^\perp$. Då är L' en operator på det inre produktrummet $\operatorname{Span}(v)^\perp$ med dimension $n-1$ och $p_L(x) = (\lambda - x)p_{L'}$. Dvs $p_{L'}$ splittrar i linjära faktorer. Av induktionsantagandet finns en ordnad ortonormal bas $\{v_2, \dots, v_n\}$ för $\operatorname{Span}(v)^\perp$ så att $[L']_{\{v_2,\dots,v_n\}}$ är diagonal. Detta bevisar att $[L]_\beta$ är diagonal där $\beta = \{v, v_2, \dots, v_n\}$ (ortonormal) eftersom

$$[L]_\beta = \begin{pmatrix} \lambda & 0 \\ 0 & [L']_{\{v_2,\dots,v_n\}} \end{pmatrix}$$

(där 0 betyder en nollkolumn/rad). Vi har alltså funnit en ortonormal bas av egenvektorer till L.

Omvänt, antag att L har en ortonormal bas β bestående av egenvektorer till L. Då är matrisen $[L]_\beta$ diagonal med egenvärdena på diagonalen. Alltså är även $[L^*]_\beta = [L]_\beta^*$ diagonal med komplexkonjugerade egenvärden till L på diagonalen. Alltså har vi att $[LL^*]_\beta = [L]_\beta [L]_\beta^* = [L]_\beta^* [L]_\beta = [L^* L]_\beta$. Dvs, $LL^* = L^* L$. $\square$

Definition 11.2. En linjär operator $L\colon V \to V$ på ett inre produktrum V kallas *självadjungerad* (eller *hermitesk*) om $L = L^*$.

Exempel 11.1. Varje matrisrepresentation av en inre produkt är självadjungerad (Se Sats 9.5).

Övning

Varje självadjungerad operator är normal med reella egenvärden.

Lemma 11.3. *Låt $L\colon V \to V$ vara en självadjungerad operator på ett inre produktrum av ändlig dimension över $\mathbb{R}$ eller $\mathbb{C}$. Då är alla egenvärden till L reella och det karaktäristiska polynomet splittrar i linjära faktorer $p_L(x)$ (även då $k = \mathbb{R}$).*

Bevis. Låt v vara en egenvektor till L med egenvärde λ. Då har vi att

$$\lambda\langle v, v\rangle = \langle L(v), v\rangle = \langle v, L(v)\rangle = \overline{\lambda}\langle v, v\rangle$$

vilket ger att $\lambda = \overline{\lambda}$ eftersom $\langle v, v\rangle > 0$. Det visar alltså att alla egenvärden är reella. Det återstår att visa att det finns tillräckligt med reella egenvärden för att splittra det karaktäristiska polynomet. Välj en ortonormal bas för V. Då är matrisen $[L]_\beta$ självadjungerad. Denna matris definierar en linjär avbildning $\mathbb{R}^n \to \mathbb{R}^n$, men också en linjär avbildning $\mathbb{C}^n \to \mathbb{C}^n$. Dessa två linjära avbildningar har samma karaktäristiska polynom eftersom båda kan beräknas med hjälp av matrisen $[L]_\beta$. Detta karaktäristiska polynom splittrar i linjära faktorer över $\mathbb{C}$ och eftersom alla egenvärden är reella så är alla linjära faktorer definierade över $\mathbb{R}$ och $p_L(x)$ splittrar i linjära faktorer över $\mathbb{R}$. $\square$

Sats 11.4 (Karaktärisering av självadjungerade operatorer)**.** *En operator $L\colon V \to V$ på ett reellt vektorrum av ändlig dimension n är självadjungerad om och endast om det finns en ortonormal bas av egenvektorer till L.*

Bevis. Existensen av en ortonormal bas av egenvektorer till L följer direkt av Lemma 11.3 i kombination med [Karaktärisering av normala operatorer] Sats 11.2.

Antag omvänt att det finns en ortonormal ordnad bas β av egenvektorer till L. Av Sats 11.1 [Egenskaper hos normala operatorer] och Övning 11 har vi $[L]_\beta = [L^*]_\beta$, dvs. $L = L^*$. $\square$

Diskussionsfrågor

(1) Vad kan vi säga om egenvärdena till LL^* då L är normal?
(2) Är matrisen

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

självadjungerad? normal? diagonaliserbar? inverterbar?

(3) Är matrisen

$$B = \begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix}$$

självadjungerad? normal? diagonaliserbar? inverterbar?

(4) Är matrisen

$$C = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$$

självadjungerad? normal? diagonaliserbar? inverterbar?

12. Unitära och ortogonala avbildningar och QR-faktorisering

I förra kapitlet tittade vi på vad det innebär att en operator L på ett inre produktrum V (över $\mathbb{C}$ respektive $\mathbb{R}$) har en uppsättning ortonormala egenvektorer som utgör en bas för V. I detta kapitel ska vi undersöka vad som händer då vi lägger till det ytterligare antagandet att alla egenvärden till L har norm 1, dvs. de ligger på enhetscirkeln i det komplexa talplanet. Som vi kommer att se är det precis dessa L som bevarar normen av vektorer.

Övning

Visa att en linjär avbildning $L\colon V \to W$ bevarar inre produkten om och endast om $L^*L = I$.

Definition 12.1. Låt V och W vara inre produktrum över en kropp k. En linjär avbildning $L\colon V \to W$ sägs vara en *isometri* om $\|L(v)\| = \|v\|$ för alla $v \in V$.

Det är klart att varje isometri är injektiv. Det är också klart att varje avbildning som bevarar inre produkten är en isometri, eftersom $\|v\| = \sqrt{\langle v, v\rangle}$. Det visar sig att även det omvända gäller, dvs. varje isometri bevarar inre produkten.

Sats 12.1 (Isometrier). *Låt $L\colon V \to W$ vara en linjär avbildning mellan inre produktrum över en kropp k. Följande egenskaper är ekvivalenta:*

(1) L bevarar inre produkten.
(2) $L^*L = I$.
(3) L är en isometri.

Bevis. Det är klart att (1) är ekvivalent med (2) och att (1) implicerar (3). Vi visar att (3) implicerar (1). Antag att (3) gäller. Då $k = \mathbb{R}$ så påstår vi att

$$\langle v, w\rangle = \frac{\|v+w\|^2 - \|v\|^2 - \|w\|^2}{2}.$$

Detta följer från att

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

När $k=\mathbb{C}$ så kan vi dela upp i realdel och imaginärdel och vi påstår att:

$$\begin{aligned}\langle v,w\rangle &= \Re\langle v,w\rangle + \Im\langle v,w\rangle \\ &= \frac{\|v+w\|^2 - \|v\|^2 - \|w\|^2}{2} + i\frac{\|v+iw\|^2 - \|v\|^2 - \|w\|^2}{2}\,.\end{aligned}$$

Vi har redan visat att den första delen ger den reella delen av $\langle v,w\rangle$. Vi visar att den andra delen ger den imaginära delen. Kom ihåg att varje komplext tal $a+ib$ uppfyller $\Im z := b = \frac{i}{2}(\overline{z}-z)$. Vi har då att

$$\begin{aligned}\|v+iw\|^2 &= \langle v+iw, v+iw\rangle \\ &= \langle v,v\rangle + i\langle v,w\rangle + i\langle w,v\rangle + \langle w,w\rangle \\ &= \langle v,v\rangle + i\langle v,w\rangle - i\overline{\langle v,w\rangle} + \langle w,w\rangle \\ &= \|v\|^2 - 2i\Im(\langle v,w\rangle) + \|w\|^2\,,\end{aligned}$$

vilket visar påståendet. □

Exempel 12.1. En linjär operator $L\colon V\to V$ som bevarar den inre produkten behöver ej vara surjektiv om V har oändlig dimension. Ett exempel ges av vektorrummet

$$\ell^2 = \{(x_1,x_2,x_3,\ldots) \mid x_i\in\mathbb{R}, \sum_{i=1}^{\infty} x_i^2 < \infty\}$$

med den inre produkten $\langle (x_1,x_2,\ldots),(y_1,y_2,\ldots)\rangle = \sum_{i=1}^{\infty} x_iy_i$ och den linjära avbildningen $L\colon \ell^2\to\ell^2$ som skickar $(x_1,x_2,x_3,\ldots)$ till $(0,x_1,x_2,x_3,\ldots)$. Denna avbildning bevarar inre produkten eftersom

$$\langle L(v),L(w)\rangle = \langle (0,x_1,x_2,x_3,\ldots),(0,y_1,y_2,y_3,\ldots)\rangle = \langle v,w\rangle$$

för alla $v,w\in\ell^2$.

Nästa sats visar dock att om V har *ändlig* dimension så är varje isometri $V\to V$ en isomorfi av inre produktrum.

Följdsats 12.2 (Unitära operatorer)**.** *Låt L vara en linjär operator på ett inre produktrum V av ändlig dimension. Följande är ekvivalent:*

(1) $LL^* = L^*L = I$.
(2) $\langle Lv, Lw\rangle = \langle v, w\rangle$ *för alla* $v, w \in V$.
(3) För varje ortonormal bas β *för* V *så är* $L(\beta)$ *också en ortonormal bas för* V.
(4) Det existerar en ortonormal bas β *för* V *sådan att* $L(\beta)$ *också är en ortonormal bas för* V.
(5) $\|Lv\| = \|v\|$ *för alla* $v \in V$.

Bevis. Att (1), (2), (5) är ekvivalenta såg vi i Sats 12.1. Vi visar att (2) implicerar (3). Anta att (2) gäller. Av (1) vet vi att L är en isomorfi och därmed är $L(\beta)$ en bas för varje bas β. Av (2) följer att $L(\beta)$ är ortonormalt, vilket ger (3). Att (3) implicerar (4) följer av Gram–Schmidt eftersom V har ändlig dimension. Antag nu att (4) gäller. Då gäller (2) för alla basvektorer i β och därmed för alla vektorer i V genom linjäritet, vilket ger (2). □

Definition 12.2. En operator L på ett inre produktrum V över $\mathbb{C}$ (över $\mathbb{R}$) sägs vara *unitär* (*ortogonal*) om den uppfyller de ekvivalenta villkoren i Sats 12.2.

Följdsats 12.3. *För en linjär operator* L *på ett inre produktrum* V *över en kropp* k *gäller följande:*

(1) Om $k = \mathbb{C}$ *så är* L *unitär om och endast om det finns en ortonormal bas för* V *bestående av egenvektorer till* L *med tillhörande egenvärden på enhetscirkeln.*
(2) Om $k = \mathbb{R}$ *så är* L *ortogonal och självadjungerad om och endast om det finns en ortonormal bas för* V *bestående av egenvektorer till* L *med tillhörande egenvärden på* $\{-1, 1\}$.

Bevis. Antag först att L är unitär (ortogonal och självadjungerad). Då har vi från Sats 11.2 (Sats 11.4) att det finns en ortonormal bas för V bestående av egenvektorer till L. Om v är en sådan egenvektor med egenvärde λ så har vi $\|v\| = \|Lv\| = \|\lambda v\| = |\lambda|\|v\|$. Eftersom $v \neq 0$ så är $\|v\| \neq 0$ och därmed måste $|\lambda| = 1$.

Omvänt, anta att det finns en ortonormal bas β för V bestående av egenvektorer till L med tillhörande egenvärden på enhetscirkeln (på $\{-1, 1\}$). Då är L diagonaliserbar med enhetsmatris som diagonalmatris i denna bas, vilket visar att L är unitär (ortogonal och självadjungerad). $\square$

Sats 12.4 (Unitära matriser)**.** *En $n \times n$-matris M uppfyller $M^*M = I$ om och endast om dess kolumner bildar en ortonormal bas för k^n.*

Bevis. Antag att M har kolumner $v_1, \dots, v_n$. Då har vi $(M^*M)_{ij} = v_i^T v_j = \langle v_i, v_j \rangle = \delta_{ij}$, där δ_{ij} är Kroneckers delta. Alltså är $M^*M = I$ om och endast om kolumnerna $v_1, \dots, v_n$ bildar en ortonormal bas för k^n (Följdsats 9.3). $\square$

Definition 12.3. En $n \times n$-matris M sägs vara *unitär* (eller *ortogonal* om $k = \mathbb{R}$) om $M^*M = I$.

Notera att $M^*M = I$ om och endast om $MM^* = I$. Alltså är M inverterbar med invers $M^{-1} = M^*$.

Följdsats 12.5. *Determinanten av en unitär (eller ortogonal) matris är av absolutbelopp 1.*

12.1. **QR-faktorisering.** I det här avsnittet ska vi visa att varje $m \times n$-matris A (med $m \geq n$) kan skrivas som produkten av en unitär (eller ortogonal) matris Q och en övre triangulär matris R, det vill säga $A = QR$:

$$\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Sats 12.6 (QR-faktorisering)**.** *Varje $m \times n$-matris A (med $m \geq n$) kan skrivas som produkten av en unitär (eller ortogonal) matris Q och en övre triangulär matris R, det vill säga $A = QR$.*

Bevis. Vi använder Gram–Schmidts metod för att konstruera Q och R. Låt A ha kolumnerna $a_1, a_2, \dots, a_n$ och låt $u_1, \dots, u_n$ vara de normerade vektorerna vi får om vi utför Gram–Schmidts metod. Utvidga dessa till en ortonormal bas för k^m: $\{u_1, \dots, u_n, u_{n+1}, \dots, u_m\}$. Definiera nu R genom $R_{ij} = \langle a_j, u_i \rangle$ för $1 \leq i \leq j \leq n$. Det följder då från [Projektion på bas] Sats 9.2 att $A = QR$. □

Följdsats 12.7. *Varje $m \times n$-matris A (med $m \leq n$) kan skrivas som produkten av en undre triangulär matris L och en unitär (eller ortogonal) matris Q, det vill säga $A = LQ$.*

Bevis. Av QR-faktorisering kan vi skriva $A^T = QR$ där Q är en unitär (eller ortogonal) matris och R är en övre triangulär matris. Tar vi transponatet igen så får vi $A = (A^T)^T = R^T Q^T$ där R^T är en undre triangulär matris och Q^T är en unitär (eller ortogonal) matris. □

Exempel 12.2. Låt
$$A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}.$$
Vi applicerar Gram–Schmidts metod: Först normaliserar vi den första kolumnen och får $u_1 = \frac{1}{\sqrt{2}}(1,0,1)^T$. Sedan ortogonaliserar vi den andra kolumnen mot u_1 (den är faktiskt redan ortogonal mot u_1):
$$\begin{aligned} u_2' &= (-1,2,1)^T - \frac{1}{2}\langle(-1,2,1),(1,0,1)\rangle(1,0,1)^T \\ &= (-1,2,1)^T \end{aligned}$$
och normerar: $u_2 = \frac{1}{\sqrt{6}}(-1,2,1)^T$. Vi får en tredje vektor, ortogonal till både u_1 och u_2 genom att ta kryssprodukten (längden är oviktig):
$$u_3' = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 2 \end{pmatrix}.$$

Vi normerar och får $u_3 = \frac{1}{\sqrt{3}}(-1, -1, 1)^T$.

$$\begin{pmatrix} 1 & -1 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{6} \\ 0 & 0 \end{pmatrix}$$

13. Singulärvärdesuppdelning

I det här kapitlet ska vi titta på generella linjära avbildningar $L\colon V \to W$ mellan inre produktrum och se hur vi kan välja baser för V respektive W för att få en så enkel beskrivning som möjligt av L.

Sats 13.1 (Singulärvärdesuppdelning)**.** *Låt $L\colon V \to W$ vara en linjär avbildning mellan ändligt-dimensionella inre produktrum över $k = \mathbb{R}$ eller $\mathbb{C}$. Då finns det ortonormala baser $\{v_1, \dots, v_n\}$ för V och $\{w_1, \dots, w_m\}$ för W samt positiva reella tal $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ (där $r = \mathrm{rk}(L)$) så att*

$$L(v_i) = \begin{cases} \sigma_i w_i & \text{om } 1 \leq i \leq r\,, \\ 0 & \text{om } r < i \leq n\,. \end{cases} \tag{13.1}$$

Å andra sidan, om (13.1) gäller så är $v_1, \dots, v_n$ egenvektorer till L^*L med tillhörande egenvärden $\sigma_1^2, \dots, \sigma_r^2, 0, 0, \dots, 0$. Alltså är $\sigma_1, \dots, \sigma_r$ unikt bestämda.

Bevis. Notera att L^*L är självadjungerad och att $\langle L^*L(v), v\rangle = \langle L(v), L(v)\rangle \geq 0$ för alla $v \in V$. Detta innebär att L^*L har icke-negativa reella egenvärden. Av Sats 11.4 finns en ortonormal bas $\{v_1, \dots, v_n\}$ för V bestående av egenvektorer till L^*L. Vi kan rangordna de positiva egenvärdena $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ och sätta $\sigma_i = \sqrt{\lambda_i}$ för $i = 1, \dots, r$. Välj nu $\{w_1, \dots, w_r\}$ genom

$$w_i = \frac{1}{\sigma_i} L(v_i)\,.$$

Då är $w_1, \dots, w_r$ ortonormala, eftersom $\|w_i\| = 1$ och $\langle w_i, w_j\rangle = \frac{1}{\sigma_i\sigma_j}\langle L(v_i), L(v_j)\rangle = \frac{1}{\sigma_i\sigma_j}\langle L^*L(v_i), v_j\rangle = \frac{\lambda_i}{\sigma_i\sigma_j}\langle v_i, v_j\rangle = \delta_{ij}$ och vi kan komplettera med $m-r$ stycken vektorer för att få en ortonormal bas $\{w_1, \dots, w_r, w_{r+1}, \dots, w_m\}$ för W. För $i > r$ har vi att $L(v_i) = 0$ eftersom v_i är en egenvektor till L^*L med egenvärde 0. Detta visar existensen av baser som uppfyller (13.1). □

Definition 13.1. De icke-negativa talen $\sigma_1, \dots, \sigma_r$ i Sats 13.1 kallas för (positiva) *singulärvärden* till L. Om r är mindre än både m och n så utvidgar vi singulärvärdena till att inkludera $\sigma_{r+1} = 0, \dots, \sigma_{\min(m,n)} = 0$. Dvs. singulärvärdena till L är de icke-negativa talen $\sigma_1, \dots, \sigma_{\min(m,n)}$, där exakt $\min(m,n) - \operatorname{rk} L$ stycken är noll.

Diskussionsfrågor

(1) Vad är singulärvärdena till den linjära avbildningen $\mathbb{C} \to \mathbb{C}$ som skickar z till iz om vi tar $\langle z_1, z_2 \rangle = z_1 \overline{z_2}$ som inre produkt på $\mathbb{C}$ som vektorrum över $\mathbb{C}$?
(2) Vad är singulärvärdena för samma avbildning om vi istället ser $\mathbb{C}$ som vektorrum över $\mathbb{R}$ med inre produkten $\langle a + ib, c + id \rangle = ac + bd$?

Övning

Visa att en linjär operator $L \colon V \to V$ uppfyller $\|L(v)\| = \|v\|$ för alla $v \in V$ om och endast om alla dess singulärvärden är lika med 1.

Exempel 13.1. Betrakta den linjära avbildningen $L \colon \mathbb{R}[x]_{\le 1} \to \mathbb{R}[x]_{\le 2}$ som ges av $f(x) \mapsto \int_0^x f(t)\,dt$. Vi har inre produkter som på båda rummen definieras av

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\,dt\,.$$

Låt oss nu bestämma singulärvärdena till L med avseende på dessa inre produkter. Vi börjar med att ta fram matrisen till L med avseende på två ortonormala baser. Med hjälp av Gram–Schmidts metod hittar vi en ortonormal bas för $\mathbb{R}[x]_{\le 1}$:

$$\beta = \{v_1 = 1\,, \quad v_2 = \sqrt{3}(2x - 1)\}\,,$$

och en ortonormal bas för $\mathbb{R}[x]_{\le 2}$:

$$\gamma = \{w_1 = 1\,, \quad w_2 = \sqrt{3}(2x - 1)\,, \quad w_3 = \sqrt{5}(6x^2 - 6x + 1)\}\,.$$

Låt $\varepsilon = \{1, x, x^2\}$. Då är basbytesmatrisen från γ till ε och dess invers lika med

$$[I]_\gamma^\varepsilon = \begin{pmatrix} 1 & -\sqrt{3} & \sqrt{5} \\ 0 & 2\sqrt{3} & -6\sqrt{5} \\ 0 & 0 & 6\sqrt{5} \end{pmatrix}, \quad [I]_\varepsilon^\gamma = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & 0 & \frac{1}{6\sqrt{5}} \end{pmatrix}.$$

Vi beräknar nu $L(v_1)$ och $L(v_2)$:

$$L(v_1) = \int_0^x 1\,dt = x \Rightarrow [L(v_1)]_\varepsilon = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$L(v_2) = \int_0^x \sqrt{3}(2t-1)\,dt = \sqrt{3}(x^2 - x) \Rightarrow [L(v_2)]_\varepsilon = \begin{pmatrix} 0 \\ -\sqrt{3} \\ \sqrt{3} \end{pmatrix}.$$

Applicerar vi $[I]_\varepsilon^\gamma$ får vi $[L(v_1)]_\gamma = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2\sqrt{3}} \\ 0 \end{pmatrix}$ och $[L(v_2)]_\gamma = \begin{pmatrix} -\frac{\sqrt{3}}{6} \\ 0 \\ \frac{\sqrt{3}}{6\sqrt{5}} \end{pmatrix}$. Alltså får vi

$$[L]_\beta^\gamma = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{6} \\ \frac{1}{2\sqrt{3}} & 0 \\ 0 & \frac{\sqrt{3}}{6\sqrt{5}} \end{pmatrix} \qquad \text{och } ([L]_\beta^\gamma)^* = \begin{pmatrix} \frac{1}{2} & \frac{1}{2\sqrt{3}} & 0 \\ -\frac{\sqrt{3}}{6} & 0 & \frac{\sqrt{3}}{6\sqrt{5}} \end{pmatrix},$$

vilket betyder att

$$([L]_\beta^\gamma)^*[L]_\beta^\gamma = \begin{pmatrix} \frac{1}{3} & -\frac{\sqrt{3}}{12} \\ -\frac{\sqrt{3}}{12} & \frac{1}{10} \end{pmatrix}.$$

Denna matris har egenvärden $\frac{13 \pm 2\sqrt{31}}{60}$ och singulärvärdena är kvadratrötterna till dessa:

$$\sigma_1 = \sqrt{\frac{13 + 2\sqrt{31}}{60}}, \quad \sigma_2 = \sqrt{\frac{13 - 2\sqrt{31}}{60}}.$$

Notera här att $13 - 2\sqrt{31} > 0$ eftersom $31 < 36 = 6^2$ och $13 - 2\sqrt{36} = 1$. Alltså är båda värdena positiva, vilket stämmer med satsen om singulärvärden.

14. Bilinjära och kvadratiska former

Vi ska nu studera ett klassiskt begrepp inom linjär algebra, nämligen bilinjära och kvadratiska former.

14.1. **Bilinjära former.** Vi börjar med att definiera bilinjära former.

Definition 14.1. Låt V vara ett vektorrum över en kropp k. En funktion

$$f\colon V \times V \to k$$

kallas för en *bilinjär form* om den är linjär i varje argument, det vill säga: för varje $v \in V$ så är funktionerna $f(-,v)\colon V \to k;\ w \mapsto f(w,v)$ och $f(v,-)\colon V \to k;\ w \mapsto f(v,w)$ linjära.

Exempel 14.1. Om $A \in M_{n\times n}(k)$ så definierar vi en bilinjär form $f\colon k^n \times k^n \to k$ genom

$$f(x,y) = x^T A y$$

för alla (kolumnvektorer) $x, y \in k^n$. Detta är en bilinjär form eftersom matris-multiplikation distrubuerar över addition och skalärmultiplikation: $(aB + C)A = aBA + CA$ och $A(aB + C) = aAB + AC$.
T.ex. om $A = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}$ så är

$$\begin{aligned} &\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \\ &(x,y) \mapsto \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &\qquad = 2x_1y_1 - x_1y_2 + 3x_2y_1 + x_2y_2 \end{aligned}$$

en bilinjär form.

Som vanligt när vi betraktar funktioner till en kropp kan vi definiera addition och skalärmultiplikation av bilinjärformer punktvis: om $f, g\colon V \times V \to k$ är bilinjärformer och $a \in k$ så definierar vi

$$(f + g)(v,w) = f(v,w) + g(v,w) \quad \text{och} \quad (af)(v,w) = af(v,w)$$

för alla $v, w \in V$. Det är lätt att se att $f + g$ och af också är bilinjärformer. Därmed bildar mängden av alla bilinjärformer på V ett vektorrum över k, vilket vi betecknar med $\mathrm{Bil}(V)$.

Sats 14.1 (Matrisrepresentation av bilinjära former)**.** *Låt V vara ett ändligdimensionellt vektorrum över en kropp k och låt $f\colon V \times V \to k$ vara en bilinjär form. Om vi väljer en bas $\beta = \{v_1, \dots, v_n\}$ för V så är avbildningen*

$$\begin{aligned} Bil(V) &\to M_{n\times n}(k) \\ f &\mapsto [f]_\beta := \big(f(v_i, v_j)\big)_{1\le i,j\le n} \end{aligned}$$

en isomorfi av vektorrum.

Bevis. □

Följdsats 14.2. *Om V är ett ändligdimensionellt vektorrum över en kropp k och β är en ordnad bas för V så har vi*

$$f(v, w) = [v]_\beta^T [f]_\beta [w]_\beta$$

för alla $v, w \in V$ och alla bilinjärformer $f\colon V \times V \to k$.

Bevis. □

Vi ska nu se hur matrisrepresentationen av en bilinjär form förändras när vi byter bas.

Sats 14.3 (Basbyte för bilinjära former)**.** *Låt V vara ett ändligdimensionellt vektorrum över en kropp k och låt $f\colon V \times V \to k$ vara en bilinjär form. Om β och γ är ordnade baser för V så gäller*

$$[f]_\gamma = ([I]_\gamma^\beta)^T [f]_\beta [I]_\gamma^\beta.$$

Bevis. Av Följdsats 14.2 räcker det att notera att de bilinjära former som definieras av $[f]_\gamma$ respektive $([I]_\gamma^\beta)^T [f]_\beta [I]_\gamma^\beta$ i basen γ är lika:

$$[v]_\gamma^T [f]_\gamma [w]_\gamma = f(v, w)\,,$$
$$[v]_\gamma^T \left(([I]_\gamma^\beta)^T [f]_\beta [I]_\gamma^\beta\right) [w]_\gamma = ([I]_\gamma^\beta [v]_\gamma)^T [f]_\beta ([I]_\gamma^\beta [w]_\gamma) = [v]_\beta^T [f]_\beta [w]_\beta = f(v, w)$$

för alla $v, w \in V$. □

14.2. **Symmetriska bilinjära former.** Vi ska nu studera bilinjära former vars matrisrepresentation är symmetrisk (i varje ordnad bas).

Definition 14.2. En bilinjär form $f\colon V \times V \to k$ kallas *symmetrisk* om $f(v, w) = f(w, v)$ för alla $v, w \in V$.

Sats 14.4 (Matrisen till en symmetrisk form)**.** *Låt V vara ett ändligdimensionellt vektorrum över en kropp k och låt $f\colon V \times V \to k$ vara en bilinjär form. Följande påståenden är ekvivalenta:*

(i) f är symmetrisk.
(ii) $[f]_\beta$ är en symmetrisk matris för någon ordnad bas β för V.
(iii) $[f]_\beta$ är en symmetrisk matris för alla ordnade baser β för V.

Bevis. Om f är symmetrisk så är $[f]_\beta$ symmetrisk för varje ordnad bas β eftersom

$$\big([f]_\beta\big)_{ij} = f(v_i, v_j) = f(v_j, v_i) = \big([f]_\beta\big)_{ji}$$

för alla $1 \leq i, j \leq n$. Dvs (i) implicerar (iii). Det är vidare klart att (iii) implicerar (ii).

Å andra sidan, om $[f]_\beta$ är symmetrisk för någon ordnad bas β så är f symmetrisk eftersom alla vektorer i V kan skrivas som linjärkombinationer av basvektorerna. Detta visar att (ii) implicerar (i). $\square$

14.3. **Kvadratiska former.** Vi ska nu definiera ett klassiskt begrepp inom linjär algebra, nämligen kvadratiska former.

Definition 14.3. En *kvadratisk form* på ett vektorrum V över en kropp k är en funktion $q\colon V \to k$ sådan att det finns en symmetrisk bilinjär form $f\colon V \times V \to k$ med

$$q(v) = f(v, v)$$

för alla $v \in V$.

Exempel 14.2. Om $A \in M_{n\times n}(k)$ så definierar vi en kvadratisk form $q\colon k^n \to k$ genom

$$q(x) = x^T A x$$

för alla (kolumnvektorer) $x \in k^n$. Detta är en kvadratisk form eftersom den är av formen $f(x,x)$ där $f\colon k^n \times k^n \to k$ är den bilinjära formen definierad av $f(x,y) = x^T Ay$.
T.ex. om $A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ så är

$$\begin{aligned} q\colon \mathbb{R}^2 &\to \mathbb{R} \\ x &\mapsto \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= 2x_1^2 - x_1x_2 - x_2x_1 + x_2^2 = 2x_1^2 - 2x_1x_2 + x_2^2 \end{aligned}$$

en kvadratisk form.

Sats 14.5. *Om q är en kvadratisk form på ett ändligdimensionellt vektorrum V över en kropp k där $2 \neq 0$ så finns en unik symmetrisk bilinjär form $f\colon V \times V \to k$ sådan att $q(v) = f(v,v)$ för alla $v \in V$, nämligen den bilinjära form f som ges av*

$$f(v,w) = \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big)$$

för alla $v, w \in V$.

Bevis. Detta följer av att $q(v+w) = f(v+w, v+w) = f(v,v) + f(v,w) + f(w,v) + f(w,w) = q(v) + 2f(v,w) + q(w)$. $\square$

15. **Repetition**

I detta kapitel ska vi lösa ett antal problem som repeterar och sammanfattar det vi gått igenom i kursen. När vi löst ett problem är det viktigt att vi *kontrollerar* svaret vi får. Vi kan t.ex. fråga oss:

(1) Är svaret rimligt givet problemet?
(2) Kan vi lösa problemet på ett annat sätt? Får vi i så fall samma svar?
(3) Vilka antaganden har vi gjort och är de giltiga?
(4) Om problemet består i att *hitta* något enligt en viss definition (t.ex. en egenvektor, egenvärde, invers till en linjär avbildning, ON-bas, singulärvärde, etc.), kan vi då *verifiera* att det vi hittat verkligen uppfyller definitionen?

Referenser

[FIS03] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence, *Linear algebra.*, 4th. ed. (international) ed., Pearson, 2003 (English).