



SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS FÖRELÄSNING 1 – HT24

MATS BOIJ

1. VEKTORRUM OCH BASER

Målet för idag.

- Repetition från SF1672 Linjär algebra när det gäller
 - Vektorrum
 - Delrum
 - Linärt oberoende och baser
 - Koordinater och basbyten
- Nya perspektiv
 - Skalärer från godtyckliga kroppar
 - Oändligdimensionella vektorrum
- Nya begrepp
 - Direkt summa $V \oplus W$

Vektorrum.

Definition 1.1. Ett *vektorrum* är en mängd V med *addition* och ett speciellt element 0 och en *multiplikation med skalärer* som uppfyller följande

- *Additionen* gör V till en *abelsk grupp* dvs
 - (kommutativ) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
 - (associativ) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
 - (identitet) $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$
 - (invers) $\forall \mathbf{x} \exists \mathbf{y}: \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} = \mathbf{0}$ (vi skriver $\mathbf{y} = -\mathbf{x}$)
- Skalärerna utgör en *kropp*¹ och *multiplikationen med skalärer* uppfyller
 - (associativ) $(ab)\mathbf{x} = a(b\mathbf{x})$
 - (distributiv vänster) $(a + b)\mathbf{x} = (a\mathbf{x}) + (b\mathbf{x})$
 - (distributiv höger) $a(\mathbf{x} + \mathbf{y}) = (a\mathbf{x}) + (a\mathbf{y})$
 - (identitet) $1\mathbf{x} = \mathbf{x}$

Exempel 1.2. Följande är vektorrum.

- $\mathbb{R}^n$ (med reella skalärer) och $\mathbb{C}^n$ (med reella eller komplexa skalärer)
- $\mathbb{R}[x]$ vilket betecknar alla polynom med koefficienter i $\mathbb{R}$.
- $\mathbb{R}[x]_{\leq n}$ vilket betecknar alla polynom av grad $\leq n$.
- $C^d([a, b])$ vilket betecknar funktioner $[a, b] \rightarrow \mathbb{R}$ vars d te derivata är kontinuerlig.

Datum: 2024-10-28.

¹dvs uppfyller de räknelagar vi kan för $\mathbb{Q}$, $\mathbb{R}$ och $\mathbb{C}$

Kroppar.

Definition 1.3. En *kropp* är en mängd k med addition och multiplikation och speciella element 0 och 1 som uppfyller

- *Additionen* gör k till en *abelsk grupp* dvs
 - $a + b = b + a$
 - $(a + b) + c = a + (b + c)$
 - $a + 0 = 0 + a = a$
 - $\forall a \exists b: a + b = b + a = 0$.
- *Multiplikationen* gör $k^\times = k \setminus \{0\}$ till en *abelsk grupp* dvs
 - $ab = ba$
 - $(ab)c = a(bc)$
 - $1 \cdot a = a$
 - $\forall a \neq 0 \exists b: ab = ba = 1$.
- Addition tillsammans med multiplikation uppfyller de *distributiva lagarna*
 - $(a + b)c = (ac) + (bc)$
 - $a(b + c) = (ab) + (ac)$

Exempel 1.4. Ni känner alla till de rationella talen $\mathbb{Q}$, de reella talen $\mathbb{R}$ och de komplexa talen $\mathbb{C}$. Det finns även ändliga kroppar. Om p är ett primtal så är heltalen modulo p , vilket betecknas $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} = \{0, 1, 2, \dots, p-1\}$, en kropp. I $\mathbb{F}_p$ identifierar vi två heltal x och y om p delar $x - y$, dvs om x och y är kongruenta modulo p . I t ex $\mathbb{F}_2$ så är $1 + 1 = 2 = 0$ och i $\mathbb{F}_3$ så är $1 + 1 = 2 = -1$ och $2 \cdot 2 = 4 = 1$.

Vi identifierar en kropp k med det 1-dimensionella vektorrummet k^1 .

Delrum.

Definition 1.5. En delmängd $W \subseteq V$ utgör ett *delrum* av V om restriktionen av addition och multiplikation med skalär till W gör W till ett vektorrum.

Sats 1.6. En delmängd $W \subseteq V$ utgör ett *delrum* av V om och endast om

- $0 \in W$,
- $\mathbf{x} + \mathbf{y} \in W$ om $\mathbf{x}, \mathbf{y} \in W$,
- $a\mathbf{x} \in W$ om $\mathbf{x} \in W$ och a är skalär.

Vi säger att W är *slutet* under addition och multiplikation med skalär.

Exempel 1.7. Varje vektorrum har det *triviala delrummet* $\{0\} \subseteq V$ och det *oäkta delrummet* $V \subseteq V$.

Exempel 1.8. Vektorrummet $k[x]$ av polynom med koefficienter i kroppen k har två delrum bestående av jämna och udda polynom

$$k[x]_{\text{even}} = \{p(x) \in k[x] : p(-x) = p(x) \ \forall x\}$$

$$k[x]_{\text{odd}} = \{p(x) \in k[x] : p(-x) = -p(x) \ \forall x\}$$

Notera att det enda polynomet som både är jämnt och udda är 0 om $1 + 1 \neq 0$ i k . Alltså är $k[x]_{\text{even}} \cap k[x]_{\text{odd}} = \{0\}$.

Exempel 1.9. Följande delmängd av kontinuerliga funktioner är ett delrum:

$$V = \{f(x) \in C^0([0, 1]) : f(1) = 0\}$$

Vad händer om vi istället kräver $f(1) = 1$?

Exempel 1.10. Vektorrummet $M_{n,n}(\mathbb{R})$ av $n \times n$ -matriser har två delrum

$$M_{n,n}^s(\mathbb{R}) = \{A \in M_{n,n}(\mathbb{R}) : A^T = A\}$$

$$M_{n,n}^{as}(\mathbb{R}) = \{A \in M_{n,n}(\mathbb{R}) : A^T = -A\}$$

av symmetriska respektive anti-symmetriska (skev-symmetriska) matriser.

Linjärt oberoende och baser.

Definition 1.11. En delmängd $\mathcal{S} \subseteq V$ är *linjärt oberoende* om för alla skalärer $a_{\mathbf{x}}$

$$\sum_{\mathbf{x} \in \mathcal{S}} a_{\mathbf{x}} \mathbf{x} = 0 \implies a_{\mathbf{x}} = 0, \forall \mathbf{x} \in \mathcal{S}$$

där summan bara kan ha ändligt många nollskilda termer.

Definition 1.12. En delmängd $\mathcal{S} \subseteq V$ *spänner upp* V om

$$V = \text{Span } \mathcal{S} = \left\{ \sum_{\mathbf{x} \in \mathcal{S}} a_{\mathbf{x}} \mathbf{x} : a_{\mathbf{x}} \neq 0 \text{ för ändligt antal } \mathbf{x} \in \mathcal{S} \right\}$$

Definition 1.13. En *bas* för V är en linjärt oberoende delmängd som spänner upp V .

Sats 1.14. Varje vektorrum V har en bas $\mathcal{B}$. Varje bas har lika många element.

Bevis. För ett bevis av det första påståendet, se extramaterialet i Appendix A nedan (använder Zorns lemma). Det andra påståendet är Corollary 2.8 i kursboken (ALA av Sadun). $\square$

Anmärkning 1.15. En bas för ett vektorrum behöver inte vara ändlig. Om basen inte är ändlig kallas V *oändligdimensionellt*.

Koordinater. Om $\mathcal{B}$ är en bas för V så kan varje vektor $\mathbf{v} \in V$ skrivas *unikt* som en linjärkombination:

$$\mathbf{v} = \sum_{\mathbf{b} \in \mathcal{B}} a_{\mathbf{b}} \mathbf{b}$$

Vi kallar $(a_{\mathbf{b}})_{\mathbf{b} \in \mathcal{B}}$ för koordinaterna till $\mathbf{v}$ i basen $\mathcal{B}$.

Speciellt, om $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ är en ändlig bas så har vi:

$$\mathbf{v} = \sum_{i=1}^n a_i \mathbf{b}_i$$

och vi låter

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

beteckna koordinatvektorn för $\mathbf{v}$ i basen $\mathcal{B}$.

Om vi har två olika baser $\mathcal{B}$ och $\mathcal{D}$ så gäller

$$[\mathbf{v}]_{\mathcal{D}} = P_{\mathcal{D}\mathcal{B}} [\mathbf{v}]_{\mathcal{B}}$$

för en matris $P_{\mathcal{D}\mathcal{B}}$ vars j te kolumn består av $[\mathbf{b}_j]_{\mathcal{D}}$, dvs koordinaterna för den j te basvektorn i $\mathcal{D}$ uttryckta i basen $\mathcal{D}$ (se Sadun, Thm. 2.9). Vidare är $P_{\mathcal{D}\mathcal{B}}$ inverterbar med invers $P_{\mathcal{B}\mathcal{D}}$ (se Sadun, Thm. 2.10).

Direkt summa.

Definition 1.16 (Yttre direkt summa). Om V och W är två vektorrum är den *direkta summan*, $V \oplus W$, det vektorrum vars underliggande mängd är $V \times W$ och där operationerna utförs komponentvis.

Definition 1.17 (Inre direkt summa). Om V och W är två delrum av ett vektorrum U är säger vi att U är den (inre) *direkta summan* av V och W om varje vektor i U kan skrivas som $\mathbf{v} + \mathbf{w}$ på ett unikt sätt där $\mathbf{v} \in V$ och $\mathbf{w} \in W$. Vi skriver $U = V \oplus W$.

Lemma 1.18. U är en inre direkt summa av V och W om och endast om:

- (1) $U = V + W$, dvs varje vektor $u \in U$ går att skriva som en summa $u = v + w$ där $v \in V$ och $w \in W$.
- (2) $V \cap W = \{\mathbf{0}\}$.

Bevis. (1) Om $U = V \oplus W$ och $\mathbf{u} \in V \cap W$ så kan vi skriva $\mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u}$ där första termen är i V och andra termen är i W . Eftersom dekompositionen är unik så är $\mathbf{u} = \mathbf{0}$.

(2) Omvänt, antag att $U = V + W$ och $V \cap W = \{\mathbf{0}\}$. Låt $\mathbf{u} \in U$ och antag att vi på två sätt kan skriva $\mathbf{u} = \mathbf{v}_1 + \mathbf{w}_1 = \mathbf{v}_2 + \mathbf{w}_2$. Då är $\mathbf{x} := \mathbf{v}_1 - \mathbf{v}_2 = \mathbf{w}_2 - \mathbf{w}_1$ ett element i $V \cap W$ och alltså är $\mathbf{x} = \mathbf{0}$. Dvs $\mathbf{v}_1 = \mathbf{v}_2$ och $\mathbf{w}_1 = \mathbf{w}_2$ så dekompositionen är unik. $\square$

Exempel 1.19 (Plan och linje). Låt $V = \mathbb{R}^3$ och $W_1 \subseteq V$ planet $x_1 = 0$ och $W_2 \subseteq V$ linjen $x_2 = x_3 = 0$. Vi har då att $V = W_1 \oplus W_2$ som en inre direkt summa i och med att $(x_1, x_2, x_3) = (0, x_2, x_3) + (x_1, 0, 0)$. På samma sätt är $V = W_1 \oplus W_2$ om W_1 är ett godtyckligt plan och W_2 en godtycklig linje som ej ligger i planet W_1 .

Exempel 1.20 (Jämna och udda polynom). Vektorrummet $\mathbb{R}[x]$ är en inre direkt summa av $\mathbb{R}[x]_{\text{even}}$ och $\mathbb{R}[x]_{\text{odd}}$ ty $p(x) = \frac{p(x)+p(-x)}{2} + \frac{p(x)-p(-x)}{2}$. Unicitet följer av $\mathbb{R}[x]_{\text{even}} \cap \mathbb{R}[x]_{\text{odd}} = \{0\}$.

Notera att en bas för $\mathbb{R}[x]_{\text{even}}$ är $\{1, x^2, x^4, \dots\}$ och en bas för $\mathbb{R}[x]_{\text{odd}}$ är $\{x, x^3, x^5, \dots\}$. Tillammans utgör dessa en bas för $\mathbb{R}[x]$.

I allmänhet gäller att om $\mathcal{B}$ en bas för V och $\mathcal{D}$ är en bas för W så är $U = V \oplus W$ om och endast om $\mathcal{B} \cap \mathcal{D} = \emptyset$ och $\mathcal{B} \cup \mathcal{D}$ är en bas för U .

Exempel 1.21 (Symmetriska och anti-symmetriska matriser). Vektorrummet $M_{n,n}(\mathbb{R})$ av $n \times n$ -matriser är en direkt summa av delrummen av symmetriska matriser $M_{n,n}^s(\mathbb{R})$ och anti-symmetriska matriser $M_{n,n}^{as}(\mathbb{R})$. Vi kan nämligen skriva varje matris A som en summa

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}$$

av en symmetrisk och en anti-symmetrisk matris, och den enda matrisen som är både symmetrisk och anti-symmetrisk är noll-matrisen.

Låt V och W vara delrum av U . Då har vi en linjär avbildning (se nästa föreläsning) $L: V \oplus W \rightarrow U$ definierad av $L((\mathbf{v}, \mathbf{w})) = \mathbf{v} + \mathbf{w}$. Följande sats förklarar varför vi har samma notation för både yttre och inre direkta summor.

Sats 1.22 (Sadun, Thm. 2.12). *Avbildningen $L: V \oplus W \rightarrow U$ ovan är en isomorfi om och endast om U är en inre direkt summa av V och W .*

Bevis. Att L är surjektiv betyder att $U = V + W$, dvs att varje vektor $\mathbf{u}$ i U går att skriva (på minst ett sätt) som en summa $\mathbf{v} + \mathbf{w}$ med $\mathbf{v} \in V$ och $\mathbf{w} \in W$. Att L är injektiv betyder att varje vektor $\mathbf{u}$ i U går att skriva på högst ett sätt som en summa $\mathbf{v} + \mathbf{w}$. $\square$

Anmärkning 1.23. Båda yttre och inre direkta summor kan generaliseras till godtyckligt antal summander om vi kräver att bara ett ändligt antal termer (element) får vara nollskilda.

$$\bigoplus_{i \in I} W_i \subseteq \prod_{i \in I} W_i \quad \text{och} \quad V = \bigoplus_{i \in I} W_i, \text{ där } W_i \subseteq V$$

BILAGA A. EXTRAMATERIAL TILL F1 (INGÅR EJ I KURSEN)

Varje vektorrum har en bas. För att se att varje vektorrum har en bas behövs Zorns lemma.

Sats A.1 (Zorns lemma). *I en icke-tom partialordnad mängd där varje kedja har en övre gräns finns minst ett maximalt element.*

Definition A.2. En *partialordning* är en relation $\leq$ som uppfyller

- $a \leq b$ och $b \leq c \implies a \leq c$ (transitivitet)
- $a \leq b$ och $b \leq a \implies a = b$ (anti-symmetri)
- $a \leq a$ (reflexivitet)

K är en *kedja*, eller en *totalordnad delmängd*, om antingen $a \leq b$ eller $b \leq a$ gäller för alla $a, b \in K$. Ett element b är en *övre gräns* för en delmängd S om $a \leq b \forall a \in S$. Ett element a är ett *maximalt* element om $a \leq b \implies a = b$.

Varje vektorrum har en bas.**Sats A.3.** Varje vektorrum V har en bas $\mathcal{B}$.

Fyll i detaljerna i detta bevis själv.

- Låt S vara mängden av linjärt oberoende delmängder i V , partiellt ordnade under inklusion.
- Kontrollera att S uppfyller villkoret i Zorns lemma.
- Låt $\mathcal{B}$ vara ett maximalt element i S .
- Kontrollera att $\mathcal{B}$ måste vara en bas för V eftersom $\mathcal{B}$ är maximalt element i S . $\square$

BILAGA B. KVOTRUM — ATT TAS UPP SENARE

Kvotrum. Detta material kommer tas upp senare. Läs gärna i förväg. Även i Sadun avsnitt 2.5.**Definition B.1.** En *sidoklass* till ett delrum $W \subseteq V$ är en delmängd på formen

$$\mathbf{x} + W = \{\mathbf{x} + \mathbf{w} : \mathbf{w} \in W\} \subseteq V$$

Lemma B.2. Sidoklasserna $\mathbf{x} + W$ och $\mathbf{y} + W$ sammanfaller om och endast om $\mathbf{x} - \mathbf{y} \in W$.*Bevis.* Om $\mathbf{x} - \mathbf{y} \in W$ och $\mathbf{w} \in W$ så är $\mathbf{x} + \mathbf{w} = \mathbf{y} + (\mathbf{x} - \mathbf{y} + \mathbf{w}) = \mathbf{y} + \mathbf{w}'$ och $\mathbf{y} + \mathbf{w} = \mathbf{x} + (-(\mathbf{x} - \mathbf{y}) + \mathbf{w}) = \mathbf{x} + \mathbf{w}''$ för några $\mathbf{w}', \mathbf{w}'' \in W$. Alltså är $\mathbf{x} + W = \mathbf{y} + W$.Omvänt om $\mathbf{x} + W = \mathbf{y} + W$ så är $\mathbf{x} + \mathbf{0} = \mathbf{y} + \mathbf{w}$ för något $\mathbf{w} \in W$ och alltså $\mathbf{x} - \mathbf{y} = \mathbf{w} \in W$. $\square$ Det följer att om $\mathbf{x} \in \mathbf{y} + W$ så är $\mathbf{x} + W = \mathbf{y} + W$. Observera att sidoklasser ej är delrum, förutom sidoklassen $\mathbf{0} + W = W$, eftersom ett delrum alltid innehåller vektorn $\mathbf{0}$.Man kan också införa en *ekvivalensrelation* på V genom att definiera att två vektorer $\mathbf{x}$ och $\mathbf{y}$ är ekvivalenta, vilket skrivs som $\mathbf{x} \sim \mathbf{y}$, om $\mathbf{x} - \mathbf{y} \in W$. *Ekvivalensklassen* till $\mathbf{x}$ betecknas $[\mathbf{x}]$ och består av alla vektorer $\mathbf{y}$ som är ekvivalenta med $\mathbf{x}$. Enligt ovan är alltså $[\mathbf{x}] = \mathbf{x} + W$. När man skriver $[\mathbf{x}]$ istället för $\mathbf{x} + W$ behöver det framgå från sammanhanget vilket delrum kvoten avser.Sidoklasserna/ekvivalensklasserna *partitionerar* V , dvs varje element $\mathbf{x}$ tillhör exakt en klass.**Definition B.3.** Om $W \subseteq V$ är ett delrum är *kvotrummet*, V/W , det vektorrum vars element är sidoklasser till W och där addition och multiplikation med skalär ges av

- $(\mathbf{x} + W) + (\mathbf{y} + W) = (\mathbf{x} + \mathbf{y}) + W$
- $a(\mathbf{x} + W) = (a\mathbf{x}) + W$.

Anmärkning B.4. Vi måste kontrollera att addition och multiplikation med skalär inte beror på val av representant $\mathbf{x}$ för sidoklassen $\mathbf{x} + W$. Om $\mathbf{x} + W = \mathbf{x}' + W$ och $\mathbf{y} + W = \mathbf{y}' + W$ så är $\mathbf{x} - \mathbf{x}' \in W$ och $\mathbf{y} - \mathbf{y}' \in W$ enligt lemmat. Alltså är $(\mathbf{x} + \mathbf{y}) + W = (\mathbf{x}' + \mathbf{y}') + (\mathbf{x} - \mathbf{x}') + (\mathbf{y} - \mathbf{y}') + W = (\mathbf{x}' + \mathbf{y}') + W$. På samma sätt är $(a\mathbf{x}) + W = a\mathbf{x}' + a(\mathbf{x} - \mathbf{x}') + W = (a\mathbf{x}') + W$.*Anmärkning B.5.* Sidoklasserna är delmängder (ej delrum) till V men kvoten V/W är **inte** en delmängd av V . Det finns däremot en linjär avbildning $V \rightarrow V/W$ som skickar $\mathbf{x}$ på $\mathbf{x} + W$.Nollan i kvoten V/W ges av sidoklassen $\mathbf{0} + W$ och $\mathbf{x} + W = \mathbf{0} + W$ betyder att $\mathbf{x} \in W$. Alltså identifieras alla element i W med $\mathbf{0}$ genom att bilda kvoten V/W .**Exempel B.6.** Betrakta linjen $L = \text{Span} \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ i $\mathbb{R}^2$. Då är $\mathbb{R}^2/L$ mängden av linjer som är parallella med L . Nollan i kvotrummet består av linjen genom origo (dvs L). Varje linje parallell med L kan skrivas som $y = 2x + m$. Vi kan alltså identifiera $\mathbb{R}^2/L$ med $\mathbb{R}$ genom $\{y = 2x + m\} \subset \mathbb{R}^2 \leftrightarrow \{m \in \mathbb{R}\}$.Om $M \neq L$ är en linje genom origo så skär varje linje parallell med L linjen M i en unik punkt. På så sätt kan vi också identifiera varje punkt i $\mathbb{R}^2/L$ med en punkt på M . Andra linjer ger andra identifikationer. Om vi betraktar $\mathbb{R}^2$ med den vanliga inre produkten så är ett naturligt val $M = L^\perp$ — linjen vinkelrät mot L .

Exempel B.7 (Kvot). Låt $V = \mathbb{R}^3$ och $W \subseteq V$ planet $x_1 = 0$. Sidoklasserna är då plan parallella med detta med $x_1 = a$ för alla olika a . Kvoten V/W blir isomorf med $\mathbb{R}$ genom att identifiera varje plan med sin första koordinat.

Liksom förut kan vi också identifiera planet $x_1 = a$ med skärningen med linjen $(t, 0, 0)$, alltså $(a, 0, 0)$. Andra linjer ger andra identifikationer.



SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS FÖRELÄSNING 2 – HT24

MATS BOIJ

2. LINJÄRA AVBILDNINGAR OCH OPERATORER

Målet för idag.

- Repetition från SF1672 Linjär algebra när det gäller
 - Linjära avbildningar
 - Matriser för linjära avbildningar
 - Basbyte för linjära avbildningar
 - Kärna och bild
- Nya perspektiv
 - Operatorer på oändligdimensionella vektorrum

Linjär avbildning.

Definition 2.1. En *linjär avbildning* är en funktion $L: V \longrightarrow W$ som uppfyller

- $L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in V.$
- $L(a\mathbf{x}) = aL(\mathbf{x}), \quad \forall \mathbf{x} \in V \text{ och alla skalärer } a.$

Anmärkning 2.2. Vi säger att L *respekterar* addition och multiplikation med skalär.

Anmärkning 2.3. En linjär avbildning $L: V \longrightarrow W$ är unikt bestämd av sina värden på en bas för V .

Exempel 2.4 (Linjära avbildningar). Några linjära avbildningar

- (1) Derivation $L = d/dx$ ger en linjär avbildning

$$\frac{d}{dx}: \mathbb{R}[x] \longrightarrow \mathbb{R}[x], \quad \text{eller} \quad \frac{d}{dx}: C^1(\mathbb{R}) \longrightarrow C(\mathbb{R})$$

- (2) $L: M_n(k) \longrightarrow M_n(k), L(A) = A + A^T$, där $M_n(k)$ är alla $n \times n$ -matriser med koefficienter i k .

- (3) Restriktion $L: C(\mathbb{R}) \longrightarrow C([0, 1])$.

Exempel 2.5 (Vektorrummet av linjära avbildningar). Summan av linjära avbildningar är linjära och vi kan multiplicera dem med skalärer. $\text{Hom}_k(V, W)$ är vektorrummet av alla linjära avbildningar $L: V \longrightarrow W$. Notera att $\text{Hom}_k(k^n, k^m)$ är vektorrummet $M_{m,n}(k)$ av $m \times n$ -matriser.

Sammansättning av linjära avbildningar är linjära $L_2 \circ L_1: V \longrightarrow U$ om $L_1: V \longrightarrow W$ och $L_2: W \longrightarrow U$. Detta ger en avbildning $\text{Hom}_k(V, W) \times \text{Hom}_k(W, U) \rightarrow \text{Hom}_k(V, U)$.

Isomorfi.

Definition 2.6. En *isomorfi* mellan vektorrum är en linjär avbildning som är *bijektiv*, dvs *injektiv* (ett-till-ett) och *surjektiv* (på).

Om $V \rightarrow W$ är en isomorfi så skriver vi $V \xrightarrow{\cong} W$ eller enbart $V \cong W$.

Sats 2.7. Ett val av en bas $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ för V ger en isomorfi

$$k^n \cong V$$

eller mer allmänt

$$k^{\oplus \mathcal{B}} = \bigoplus_{\mathbf{b} \in \mathcal{B}} k \cong V$$

om $\mathcal{B}$ inte är ändlig.

Bevis. $\{a_{\mathbf{b}}\}_{\mathbf{b} \in \mathcal{B}} \mapsto \sum_{\mathbf{b} \in \mathcal{B}} a_{\mathbf{b}} \mathbf{b}$ ger en avbildning $k^{\oplus \mathcal{B}} \rightarrow V$ som är injektiv eftersom $\mathcal{B}$ är linjärt oberoende och surjektiv eftersom $\mathcal{B}$ spänner upp V . $\square$

Exempel 2.8. Basen $\{1, x, x^2, \dots, x^n\}$ ger en isomorfi $L: k^{n+1} \rightarrow k[x]_{\leq n}$ som definieras av $L((a_0, a_1, \dots, a_n)) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$.

Exempel 2.9. Om V och W är delrum till U så har vi en linjär avbildning $L: V \oplus W \rightarrow U$ som definieras av $L((\mathbf{v}, \mathbf{w})) = \mathbf{v} + \mathbf{w}$. Avbildningen L är en isomorfi precis när U är en inre direkt summa av V och W . Att L är bijektiv betyder ju att för varje vektor $\mathbf{x} \in U$ finns ett unikt element $(\mathbf{v}, \mathbf{w}) \in V \oplus W$ sådant att $\mathbf{x} = \mathbf{v} + \mathbf{w}$.

Exempel 2.10 (Isomorfi). $V \oplus W \cong W \oplus V$ genom $(x, y) \mapsto (y, x)$.

Matriser för linjära avbildningar.

Definition 2.11. Låt $L: V \rightarrow W$ vara en linjär avbildning. Om $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ är en bas för V och $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_m\}$ är en bas för W så kan vi skriva

$$L(\mathbf{b}_j) = \sum_{i=1}^m a_{ij} \mathbf{d}_i, \quad j = 1, 2, \dots, n.$$

Detta bestämmer en matris $[L]_{\mathcal{D}\mathcal{B}} = (a_{ij})$ sådan att

$$[L(\mathbf{v})]_{\mathcal{D}} = [L]_{\mathcal{D}\mathcal{B}} [\mathbf{v}]_{\mathcal{B}}.$$

för alla $\mathbf{v} \in V$.

Se Sadun, Thm. 3.1 för en mer precis formulering.

Anmärkning 2.12. Om V och/eller W är oändligtdimensionella, och vi har baser $\mathcal{B} = \{\mathbf{b}_j\}_{j \in J}$ och $\mathcal{D} = \{\mathbf{d}_i\}_{i \in I}$ så får vi fortfarande

$$L(\mathbf{b}_j) = \sum_{i \in I} a_{ij} \mathbf{d}_i, \quad j \in J.$$

Då blir $[L]_{\mathcal{D}\mathcal{B}} = (a_{ij})$ en oändligt stor matris.

Exempel 2.13. Låt $k[x]$ vara vektorrummet av polynom med koefficienter i k . Då är $\mathcal{B} = \{1, x, x^2, \dots\}$ en bas för $k[x]$.

- Matrisen A för $\frac{d}{dx}: \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ med avseende på basen $\mathcal{B}$ ges av $A_{i,j} = j\delta_{i,j-1}$, för alla $i, j \geq 0$.

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & 0 & \dots \\ 0 & 0 & 0 & 3 & 0 & \dots \\ 0 & 0 & 0 & 0 & 4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Samma formel definierar derivation för alla kroppar k .

- Matrisen B för $ev_2: k[x] \rightarrow k = k^1$, där $ev_2(f) = f(2)$ (evaluera polynomet i $x = 2$) ges av $B_{0,j} = 2^j$ för alla $j \geq 0$.

$$[1 \quad 2 \quad 4 \quad 8 \quad \dots]$$

Observera att i dessa exempel indexerar vi matriserna med rader 0, 1, 2 och kolumner 0, 1, 2. Naturligt eftersom första basvektorn är x^0 , andra x^1 osv.

Basbyte för linjära avbildningar.

Sats 2.14 (Sadun, Thm. 3.3). Om $P_{\mathcal{D}'\mathcal{D}}$ ger basbyte på W och $P_{\mathcal{B}\mathcal{B}'}$ ger basbyte på V får vi

$$[L]_{\mathcal{D}'\mathcal{B}'} = P_{\mathcal{D}'\mathcal{D}}[L]_{\mathcal{D}\mathcal{B}}P_{\mathcal{B}\mathcal{B}'}$$

Bevis. Betrakta diagrammet

$$\begin{array}{ccccccc} V & \xrightarrow{\cong_{\mathcal{B}'}} & k^n & \xrightarrow{P_{\mathcal{B}\mathcal{B}'}} & k^n & \xleftarrow{\cong_{\mathcal{B}}} & V \\ \downarrow L & & \downarrow [L]_{\mathcal{D}'\mathcal{B}'} & & \downarrow [L]_{\mathcal{D}\mathcal{B}} & & \downarrow L \\ W & \xrightarrow{\cong_{\mathcal{D}'}} & k^m & \xleftarrow{P_{\mathcal{D}'\mathcal{D}}} & k^m & \xleftarrow{\cong_{\mathcal{D}}} & W \end{array}$$

Här betecknar $\cong_{\mathcal{B}}$ isomorfin $V \rightarrow k^{\oplus \mathcal{B}}$ som ges av basen $\mathcal{B}$ etc. □

Exempel 2.15 (Normalform för matriser). Vi kan välja bas för V och W så att matrisen för L har formen

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

Vi kommer tillbaka till hur vi lättast ser det (i samband med singularvärdesuppdelning).

Operatorer. Har vi en *operator* $L: V \rightarrow V$ kan vi välja samma bas för både källan och målet. Vi skriver då $[L]_{\mathcal{B}}$ istället för $[L]_{\mathcal{B}\mathcal{B}}$. Har vi två baser $\mathcal{B}$ och $\mathcal{D}$ så blir alltså basbytet:

$$[L]_{\mathcal{B}} = P_{\mathcal{B}\mathcal{D}}[L]_{\mathcal{D}}P_{\mathcal{D}\mathcal{B}}$$

Notera att $P_{\mathcal{D}\mathcal{B}} = (P_{\mathcal{B}\mathcal{D}})^{-1}$. Låter vi $A = [L]_{\mathcal{B}}$, $D = [L]_{\mathcal{D}}$ och $P = P_{\mathcal{B}\mathcal{D}}$ så får vi:

$$A = PDP^{-1}.$$

Vi säger att A är D konjugerat med P . Detta känner vi igen som diagonalisering när D är en diagonalmatris. Då är kolumnerna i P egenvektorer till A och vektorerna i $\mathcal{D}$ är egenvektorer till L . Vi återkommer till detta nästa föreläsning.

Kärna och bild.

Definition 2.16. Låt $L: V \rightarrow W$ vara en linjär avbildning.

- *Kärnan*, $\ker(L) \subseteq V$, ges av

$$\ker(L) = \{\mathbf{x} \in V : L(\mathbf{x}) = \mathbf{0}\}$$

- *Bilden*, $\operatorname{im}(L) \subseteq W$, ges av

$$\operatorname{im}(L) = \operatorname{Range}(L) = \{L(\mathbf{x}) : \mathbf{x} \in V\}$$

Sats 2.17 (Sadun, Thm. 3.4). $\ker(L) \subseteq V$ och $\operatorname{im}(L) \subseteq W$ är delrum.

Bevis. Vi behöver visa att $\ker(L)$ och $\operatorname{im}(L)$ är slutna under addition och multiplikation med skalär:

- Om $\mathbf{x}, \mathbf{y} \in \ker(L)$ är $L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$, så $\mathbf{x} + \mathbf{y} \in \ker(L)$.
- Om $\mathbf{x} \in \ker(L)$ och $a \in k$ är en skalär så är $L(a\mathbf{x}) = aL(\mathbf{x}) = a \cdot \mathbf{0} = \mathbf{0}$, så $a\mathbf{x} \in \ker(L)$.
- Om $\mathbf{x}, \mathbf{y} \in \operatorname{im}(L)$ finns $\mathbf{x}', \mathbf{y}' \in V$ med $L(\mathbf{x}') = \mathbf{x}$ och $L(\mathbf{y}') = \mathbf{y}$. Då är $\mathbf{x} + \mathbf{y} = L(\mathbf{x}') + L(\mathbf{y}') = L(\mathbf{x}' + \mathbf{y}') \in \operatorname{im}(L)$.
- Om $\mathbf{x} \in \operatorname{im}(L)$ finns $\mathbf{x}' \in V$ med $L(\mathbf{x}') = \mathbf{x}$ och $a\mathbf{x} = aL(\mathbf{x}') = L(a\mathbf{x}') \in \operatorname{im}(L)$ för alla skalärer $a \in k$.

□

Exempel 2.18 (Kärna och bild). Låt $k = \mathbb{R}$ eller $\mathbb{C}$.

- (a) Om $L = \frac{d}{dx} : k[x] \rightarrow k[x]$ är derivering av polynom så får vi att kärnan

$$\ker(L) = \{p(x) : p'(x) = 0\} = \{a : a \in k\}$$

består av de konstanta polynomen. Bilden är $\text{im}(L) = k[x]$ eftersom alla polynom har en primitiv funktion. Det räcker att se att baselementen ligger i bilden för att se att L är surjektiv och vi har $L\left(\frac{x^{n+1}}{n+1}\right) = x^n$, för alla $n \geq 0$.

- (b) Om $M_n(k)$ är vektorrummet av $n \times n$ -matriser med element i k och $L : M_n(k) \rightarrow M_n(k)$ ges av $L(A) = A + A^T$ får vi att kärnan

$$\ker(L) = \{A \in M_n(k) : A + A^T = 0\} = \{A \in M_n(k) : A^T = -A\}$$

är rummet av alla *skew-symmetriska* eller *anti-symmetriska* matriser. Bilden ges av

$$\text{im}(L) = \{A + A^T : A \in M_n(k)\} = \{A \in M_n(k) : A^T = A\}$$

dvs alla *symmetriska* matriser. Detta eftersom $A + A^T$ är symmetrisk, då $(A + A^T)^T = A^T + A = A + A^T$ och varje symmetrisk matris A avbildas på $2A$ genom L .

Rang och injektivitet.

Definition 2.19. *Rangen* av en linjär avbildning $L : V \rightarrow W$ är dimensionen av bilden $\text{im}(L)$, dvs

$$\text{rk}(L) = \dim \text{im}(L).$$

Anmärkning 2.20. Rangen kan vara ändlig även om V och W är oändligdimensionella.

Sats 2.21 (Dimensionssatsen, Sadun, Thm. 3.7). Låt $L : V \rightarrow W$ vara en linjär avbildning. Om $\dim V < \infty$ så är

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

Anmärkning 2.22. Mer generellt gäller att $V \cong \ker L \oplus \text{im} L$. Vi kan se det genom att välja ett komplement U till $\ker L$ i V så att $\ker L \oplus U = V$. Restriktionen av L till U är injektiv och har samma bildrum som L . Alltså är $U \cong \text{im} L$.

Anmärkning 2.23. Dimensionen av kärnan kallas ibland för “nullity” på engelska.

Sats 2.24 (Sadun, Thm. 3.5). $L : V \rightarrow W$ är *injektiv* (ett-ett) om och endast om $\ker(L) = \{0\}$.

Bevis. $L(\mathbf{x}) = L(\mathbf{y}) \iff L(\mathbf{x} - \mathbf{y}) = 0 \iff \mathbf{x} - \mathbf{y} \in \ker(L)$. □

Sats 2.25 (ungefär Sadun, Thm. 2.5). Låt V vara ett vektorrum av dimension $n < \infty$ med någon bas $\mathcal{B}$. Låt $L : V \rightarrow V$ vara en operator med matris $A = [L]_{\mathcal{B}}$. Följande påståenden är ekvivalenta

- (1a) L är injektiv.
- (1b) $\ker L = \{0\}$.
- (2a) L är surjektiv.
- (2b) $\text{rk} L = n$.
- (3) $\det A \neq 0$.
- (4) A är rad-ekvivalent med identitetsmatrisen I_n .

Bevis. (1a) $\iff$ (1b) är föregående sats. (2a) $\iff$ (2b) följer av definitionen av rangen. Gauss–Jordan-eliminering bevarar rangen så (2b) $\iff$ (4). Gauss–Jordan-eliminering bevarar determinanten upp till multiplikation med noll-skild skalär så (4) $\iff$ (3). □

Exempel 2.26 (Oändligdimensionella rum). Påståendet (1a) $\iff$ (2a) är inte sant för oändligdimensionella rum. Operatorn derivation $D = \frac{d}{dx}$ är surjektiv men inte injektiv. Om vi definierar

operatören $P: k[x] \longrightarrow k[x]$ “en primitiv funktion” på standardbasen enligt $P(x^d) = \frac{x^{d+1}}{d+1}$ så är matrisen av P :

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & \frac{1}{2} & 0 & 0 & \cdots \\ 0 & 0 & \frac{1}{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

P är injektiv men inte surjektiv (alla polynom i bilden saknar konstantterm). Sammansättningen $D \circ P$ är identiteten medan sammansättningen $P \circ D$ ges av matrisen.

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

Det kan alltså finnas ensidiga inverser som inte är inverser.



SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS FÖRELÄSNING 3 – HT24

MATS BOIJ

3. KVOTER OCH ISOMORFISATSEN. EGENVÄRDEN OCH DIAGONALISERING

Målet för idag.

- Kvotrum och kvotavbildningar
- Isomorfisatsen (generaliserar dimensionssatsen)
- Repetition från SF1672 Linjär algebra när det gäller
 - Eigenvärden och egenvektorer
 - Diagonaliserbara matriser och operatorer
 - Karakteristiska polynomet
- Nya perspektiv
 - Eigenvärden och egenvektorer för operatorer på oändligdimensionella vektorrum
 - Andra kroppar än $\mathbb{R}$

Kvotrum.

Definition 3.1. En *sidoklass* till ett delrum $W \subseteq V$ är en delmängd på formen

$$\mathbf{x} + W = \{\mathbf{x} + \mathbf{w} : \mathbf{w} \in W\} \subseteq V$$

Lemma 3.2. Sidoklasserna $\mathbf{x} + W$ och $\mathbf{y} + W$ sammanfaller om och endast om $\mathbf{x} - \mathbf{y} \in W$.

Bevis. Om $\mathbf{x} - \mathbf{y} \in W$ och $\mathbf{w} \in W$ så är $\mathbf{x} + \mathbf{w} = \mathbf{y} + (\mathbf{x} - \mathbf{y} + \mathbf{w}) = \mathbf{y} + \mathbf{w}'$ och $\mathbf{y} + \mathbf{w} = \mathbf{x} + (-(\mathbf{x} - \mathbf{y}) + \mathbf{w}) = \mathbf{x} + \mathbf{w}''$ för några $\mathbf{w}', \mathbf{w}'' \in W$. Alltså är $\mathbf{x} + W = \mathbf{y} + W$.

Omvänt om $\mathbf{x} + W = \mathbf{y} + W$ så är $\mathbf{x} + \mathbf{0} = \mathbf{y} + \mathbf{w}$ för något $\mathbf{w} \in W$ och alltså $\mathbf{x} - \mathbf{y} = \mathbf{w} \in W$. $\square$

Det följer att om $\mathbf{x} \in \mathbf{y} + W$ så är $\mathbf{x} + W = \mathbf{y} + W$. Observera att sidoklasser ej är delrum, förutom sidoklassen $\mathbf{0} + W = W$, eftersom ett delrum alltid innehåller vektorn $\mathbf{0}$.

Man kan också införa en *ekvivalensrelation* på V genom att definiera att två vektorer $\mathbf{x}$ och $\mathbf{y}$ är ekvivalenta, vilket skrivs som $\mathbf{x} \sim \mathbf{y}$, om $\mathbf{x} - \mathbf{y} \in W$. *Ekvivalensklassen* till $\mathbf{x}$ betecknas $[\mathbf{x}]$ och består av alla vektorer $\mathbf{y}$ som är ekvivalenta med $\mathbf{x}$, dvs:

$$[\mathbf{x}] = \{\mathbf{y} \in V : \mathbf{y} \sim \mathbf{x}\}$$

Enligt ovan är alltså ekvivalensklassen $[\mathbf{x}]$ detsamma som sidoklassen $\mathbf{x} + W$ till $\mathbf{x}$. När man skriver $[\mathbf{x}]$ istället för $\mathbf{x} + W$ behöver det framgå från sammanhanget vilket delrum W kvoten avser.

Sidoklasserna/ekvivalensklasserna *partitionerar* V , dvs varje element $\mathbf{x}$ tillhör exakt en klass.

Definition 3.3. Om $W \subseteq V$ är ett delrum är *kvotrummet*, V/W , det vektorrum vars element är sidoklasser till W och där addition och multiplikation med skalär ges av

- $(\mathbf{x} + W) + (\mathbf{y} + W) = (\mathbf{x} + \mathbf{y}) + W$
- $a(\mathbf{x} + W) = (a\mathbf{x}) + W$.

Anmärkning 3.4. Vi måste kontrollera att addition och multiplikation med skalär inte beror på val av representant $\mathbf{x}$ för sidoklassen $\mathbf{x} + W$. Om $\mathbf{x} + W = \mathbf{x}' + W$ och $\mathbf{y} + W = \mathbf{y}' + W$ så är $\mathbf{x} - \mathbf{x}' \in W$ och $\mathbf{y} - \mathbf{y}' \in W$ enligt lemmat. Alltså är $(\mathbf{x} + \mathbf{y}) + W = (\mathbf{x}' + \mathbf{y}') + (\mathbf{x} - \mathbf{x}') + (\mathbf{y} - \mathbf{y}') + W = (\mathbf{x}' + \mathbf{y}') + W$. På samma sätt är $(a\mathbf{x}) + W = a\mathbf{x}' + a(\mathbf{x} - \mathbf{x}') + W = (a\mathbf{x}') + W$.

Anmärkning 3.5. Sidoklasserna är delmängder (ej delrum) till V men kvoten V/W är **inte** en delmängd av V . Det finns däremot en linjär avbildning $V \rightarrow V/W$ som skickar $\mathbf{x}$ på $\mathbf{x} + W$.

Nollan i kvoten V/W ges av sidoklassen $\mathbf{0} + W$ och $\mathbf{x} + W = \mathbf{0} + W$ betyder att $\mathbf{x} \in W$. Alltså identifieras alla element i W med $\mathbf{0}$ genom att bilda kvoten V/W .

Exempel 3.6. Betrakta linjen $L = \text{Span} \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ i $\mathbb{R}^2$. Då är $\mathbb{R}^2/L$ mängden av linjer som är parallella med L . Nollan i kvotrummet består av linjen genom origo (dvs L). Varje linje parallell med L kan skrivas som $y = 2x + m$. Vi kan alltså identifiera $\mathbb{R}^2/L$ med $\mathbb{R}$ genom bijektionen mellan delmängden $\{y = 2x + m\} \subset \mathbb{R}^2$ och $\mathbb{R}$ som tar $\{y = 2x + m\}$ på m .

Om $M \neq L$ är en linje genom origo så skär varje linje parallell med L linjen M i en unik punkt. På så sätt kan vi också identifiera varje punkt i $\mathbb{R}^2/L$ med en punkt på M . Andra linjer ger andra identifikationer. Om vi betraktar $\mathbb{R}^2$ med den vanliga inre produkten så är ett naturligt val $M = L^\perp$ — linjen vinkelrät mot L .

Exempel 3.7 (Kvot). Låt $V = \mathbb{R}^3$ och $W \subseteq V$ planet $x_1 = 0$. Sidoklasserna är då plan parallella med detta med $x_1 = a$ för alla olika a . Kvoten V/W blir isomorf med $\mathbb{R}$ genom att identifiera varje plan med sin första koordinat.

Liksom förut kan vi också identifiera planet $x_1 = a$ med skärningen med linjen $(t, 0, 0)$, alltså $(a, 0, 0)$. Andra linjer ger andra identifikationer.

Kvotavbildningar och dimensionssatsen.

Definition 3.8. Låt $W \subseteq V$ vara ett delrum. *Kvotavbildningen* är avbildningen $q: V \rightarrow V/W$ som definieras av $q(\mathbf{x}) = \mathbf{x} + W$.

Sats 3.9. Kvotavbildningen q är en linjär avbildning.

Bevis. Vi har

$$(\mathbf{x} + \mathbf{y}) \mapsto (\mathbf{x} + \mathbf{y}) + W = \mathbf{x} + \mathbf{y} + W = (\mathbf{x} + W) + (\mathbf{y} + W)$$

och

$$a\mathbf{x} \mapsto a\mathbf{x} + W = a\mathbf{x} + aW = a(\mathbf{x} + W).$$

för alla vektorer $\mathbf{x}, \mathbf{y} \in V$ och skalärer $a \in k$. □

Sats 3.10 (Sadun Thm. 2.14). Om V är ändligdimensionellt så gäller $\dim(V/W) = \dim V - \dim W$.

Bevis. Välj ett komplement W' till W , dvs ett vektorrum så att $V = W + W'$ och $W \cap W' = \{0\}$. Konkret: om $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ är en bas till W , utöka denna till en bas $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m, \mathbf{b}_{m+1}, \dots, \mathbf{b}_n$ för V och låt $W' = \text{Span}\{\mathbf{b}_{m+1}, \dots, \mathbf{b}_n\}$. Då är $V = W \oplus W'$ (en inre direkt summa). Vi betraktar nu avbildningen $L: W' \rightarrow V \rightarrow V/W$, som avbildar $w' \mapsto (0, w') \mapsto [(0, w')]$, dvs $L(w') = w' + W$. Man visar att L är bijektiv genom att använda definitionen av en inre direkt summa. Det följer att $\dim V/W = \dim W' = \dim V - \dim W$. □

Isomorfisatsen och dimensionssatsen.

Sats 3.11 (Isomorfisatsen, Sadun Thm. 3.6). För en linjär avbildning $L: V \rightarrow W$ gäller att

$$\text{im}(L) \cong V/\ker(L).$$

Bevis. Vi konstruerar en avbildning $\Phi: V/\ker(L) \rightarrow \text{im}(L)$ genom att sätta $\Phi(\mathbf{x} + \ker(L)) = L(\mathbf{x})$.

- Φ är *väldefinierad* eftersom $\mathbf{x} + \ker(L) = \mathbf{y} + \ker(L)$ ger att $\mathbf{y} - \mathbf{x} \in \ker(L)$, dvs $L(\mathbf{y} - \mathbf{x}) = 0$. Alltså är $L(\mathbf{y}) = L(\mathbf{x} + (\mathbf{y} - \mathbf{x})) = L(\mathbf{x}) + L(\mathbf{y} - \mathbf{x}) = L(\mathbf{x})$.

- Φ är *linjär* eftersom $\Phi(\mathbf{x} + \ker(L) + \mathbf{y} + \ker(L)) = \Phi(\mathbf{x} + \mathbf{y} + \ker(L)) = L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}) = \Phi(\mathbf{x} + \ker(L)) + \Phi(\mathbf{y} + \ker(L))$.
- Φ är *surjektiv* ty om $\mathbf{y} \in \text{im}(L)$ så är $\mathbf{y} = L(\mathbf{x})$ för något $\mathbf{x} \in V$ och då är $\mathbf{y} = \Phi(\mathbf{x} + \ker(L))$ i bilden av Φ .
- Φ är *injektiv* ty om $\Phi(\mathbf{x} + \ker(L)) = 0$ så är $L(\mathbf{x}) = 0$, dvs $\mathbf{x} \in \ker L$ och därmed $\mathbf{x} + \ker(L) = 0 + \ker(L)$ vilket är nollvektorn i $V/\ker L$. $\square$

Korollarium 3.12 (Dimensionssatsen). Om $\dim V < \infty$ är $\dim \ker L + \dim \text{im}(L) = \dim V$.

Exempel 3.13 ($V \cong \ker(L) \oplus \text{im}(L)$). $L(A) = A + A^T$ har

$$\ker(L) = A_n = \{\text{antisymmetriska } n \times n\text{-matriser}\}$$

och

$$\text{im}(L) = S_n = \{\text{symmetriska } n \times n\text{-matriser}\}$$

Här har vi att $V = M_n(k) = \ker(L) \oplus \text{im}(L)$ som en inre direkt summa eftersom varje matris A på ett unikt sätt kan skrivas som en summa av en symmetrisk och en antisymmetrisk matris

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}.$$

Uppgift 3.14 (Uppgift 2 vid Tentamen 2018-01-10). Låt $V = \mathbb{C}[x]$ vara vektorrummet som ges av alla polynom i x med komplexa koefficienter och låt $W = \{p(x) \in V : p(1) = p(-1) = 0\}$.

- Bestäm en linjär avbildning $L: V \rightarrow U$ så att $W = \ker(L)$. (1)
- Bestäm en bas för V/W . (3)

Lösning.

- Vi kan välja $L: V \rightarrow \mathbb{C}^2$ som ges av $L(p(x)) = (p(-1), p(1))$. Då är $W = \ker(L)$ per definition
- Eftersom $V/W = V/\ker(L)$ har vi enligt isomorfinssatsen att $V/W \cong \text{im}(L)$. Eftersom L är surjektiv har vi att $V/W \cong \mathbb{C}^2$. Vi kan välja en bas genom att ta sidoklasserna till två polynom som ger linjärt oberoende bilder i $\mathbb{C}^2$. Ett exempel är 1 och x med $L(1) = (1, 1)$ och $L(x) = (-1, 1)$. Basen för V/W är då $\{1 + W, x + W\}$. □

I beviset av dimensionssatsen skapade vi en avbildning $V/\ker L \rightarrow W$. Mer allmänt gäller:

Sats 3.15 (Faktorisering genom kvot). Låt $U \subseteq V$ vara ett delrum och $L: V \rightarrow W$ en linjär avbildning. Om $U \subseteq \ker(L)$ så faktorerar L via kvotavbildningen $q: V \rightarrow V/U$, dvs det finns en unik avbildning $\Phi: V/U \rightarrow W$ sådan att $L = \Phi \circ q$. Detta sammanfattas i det kommutativa diagrammet

$$\begin{array}{ccc} V & \xrightarrow{q} & V/U \\ L \downarrow & \swarrow \Phi & \\ W & & \end{array}$$

Bevis. Definiera $\Phi(\mathbf{x} + U) = L(\mathbf{x})$. Detta är en väldefinierad avbildning precis som i beviset av isomorfinssatsen. Eftersom $\mathbf{x} + U = \mathbf{y} + U$ så är $\mathbf{y} - \mathbf{x} \in U \subseteq \ker L$ så $L(\mathbf{x} - \mathbf{y}) = 0$ vilket ger $L(\mathbf{y}) = L(\mathbf{x} + (\mathbf{y} - \mathbf{x})) = L(\mathbf{x})$. □

Eigenvärden och egenvektorer.

Definition 3.16 (Eigenvärden och egenvektorer). Om L är en operator på ett vektorrum V och $\xi \in V$ så är ξ en *egenvektor* med *eigenvärde* λ om

$$L(\xi) = \lambda \xi \quad \text{och} \quad \xi \neq 0.$$

Om λ är ett eigenvärde så är *egenrummet* $E_\lambda \subseteq V$ alla lösningar ξ till $L(\xi) = \lambda \xi$, inklusive noll-vektorn.

Anmärkning 3.17. Eigenvärdet λ är en *skalär* och måste tillhöra den kropp vektorrummet är definierat över. Egenrummet $E_\lambda = \ker(\lambda I - L)$ är en kärna till en operator och därmed ett delrum. Här betecknar I identitetsoperatoren på V .

Exempel 3.18 (Sadun, §4.4). En operator med matrisen $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ har inga egenvektorer om skalärerna är $k = \mathbb{R}$, men väl om skalärerna är $k = \mathbb{C}$ eftersom

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

Diagonalisering.

Definition 3.19. En operator L på V är *diagonaliserbar* om vi kan hitta en bas $\mathcal{B}$ för V som består av egenvektorer till L . Då blir matrisen för L med avseende på $\mathcal{B}$ en *diagonalmatrix*.

Exempel 3.20. Låt $V = \text{Span}\{1, \sin(x), \cos(x), \sin(2x), \cos(2x), \dots\} \subseteq C^\infty(\mathbb{R})$ vara vektorrummet av *trigonometriska* polynom med komplexa koefficienter. Operatoren $L = \frac{d}{dx}$ har inga reella eigenvärden och är alltså inte diagonaliserbar över $\mathbb{R}$. Däremot är L diagonaliserbar över $\mathbb{C}$: en bas av egenvektorer ges av $\{e^{i\omega x} = \cos(\omega x) + i \sin(\omega x)\}_{\omega \in \mathbb{Z}}$ och matrisen för L i denna bas ges av

$$\begin{aligned} \frac{d}{dx}(\cos(\omega x) + i \sin(\omega x)) &= -\omega \sin(\omega x) + i\omega \cos(\omega x) \\ &= i\omega(\cos(\omega x) + i \sin(\omega x)) \end{aligned}$$

dvs en diagonalmatrix med $i\omega$ på position (ω, ω) . Observera att diagonalbasen och matrisen är indexerad med $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ istället för $1, 2, 3, \dots$.

Anmärkning 3.21. L diagonaliserbar om och endast om $V = \bigoplus_\lambda E_\lambda$, dvs om varje vektor $v \in V$ går att skriva unikt som $v = \sum_\lambda v_\lambda$ där $v_\lambda \in E_\lambda$ är en egenvektor med eigenvärde λ .

Konjugerade matriser.

Definition 3.22 (Konjugerade matriser). Två kvadratiske matriser A och B är *konjugerade*, eller *similära*, om de motsvarar samma linjära operator med olika baser. Det betyder att det finns en matrix P så att

$$A = P^{-1}BP \quad \text{och} \quad B = PAP^{-1}.$$

Anmärkning 3.23. Att diagonalisera en matrix handlar om att hitta en konjugerad diagonalmatrix.

Karakteristiska polynomet.

Definition 3.24. Om A är en $n \times n$ -matrix är det *karakteristiska polynomet*

$$p_A(x) = \det(xI - A).$$

Anmärkning 3.25. Två konjugerade matriser A och $B = PAP^{-1}$ har samma determinant ty

$$\det(PAP^{-1}) = \det(P)\det(A)\det(P)^{-1} = \det(A).$$

Vi kan därför för en operator L på ett ändligdimensionellt vektorrum definiera $\det L = \det[L]_{\mathcal{B}}$ för någon bas $\mathcal{B}$ och $\det L$ beror inte på valet av bas.

Anmärkning 3.26. Två konjugerade matriser A och $B = PAP^{-1}$ har samma karakteristiska polynom ty:

$$\det(xI - PAP^{-1}) = \det(P(xI - A)P^{-1}) = \det(P)\det(xI - A)\det(P)^{-1} = \det(xI - A).$$

Vi kan därför också tala om det karakteristiska polynomet $p_L(x)$ för en operator L på ett ändlig-dimensionellt vektorrum.

Sats 3.27 (Sadun, Thm. 2.2). λ är ett egenvärde till L om och endast om $p_L(\lambda) = 0$.

Bevis. $L(\xi) = \lambda\xi \iff (\lambda I - L)(\xi) = 0 \iff \xi \in \ker(\lambda I - L)$. Att λ är ett egenvärde betyder att det finns nollskild ξ i kärnan, dvs att $p_L(\lambda) = \det(\lambda I - L) = 0$. $\square$

Anmärkning 3.28. Att de karakteristiska polynomen är lika räcker inte för att matriserna ska vara konjugerade. Exempelvis är

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{och} \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

inte konjugerade, men båda har karakteristiska polynomet $p_A(x) = x^2$.

Algebraisk och geometrisk multiplicitet.

Definition 3.29 (Algebraisk och geometrisk multiplicitet). För en linjär operator L med egenvärde λ är

- Den *algebraiska multipliciteten* hos λ är multipliciteten hos roten λ i $p_L(x)$, dvs:

$$m_a(\lambda) = \max\{m \in \mathbb{N} : (x - \lambda)^m \text{ delar } p_L(x)\}.$$

- Den *geometriska multipliciteten* hos λ är dimensionen hos egenrummet E_λ , dvs:

$$m_g(\lambda) = \dim(\ker(\lambda I - L))$$

Sats 3.30 (Sadun, Thm. 4.6). Om $\xi_1, \dots, \xi_m$ är egenvektorer till distinkta egenvärden $\lambda_1, \dots, \lambda_m$ så är $\{\xi_1, \dots, \xi_m\}$ linjärt oberoende.

Sats 3.31 (Sadun, Thm. 4.9). Låt L vara en operator på ett vektorrum V av dimension n . Då är L diagonaliserbar om och endast om $m_a(\lambda) = m_g(\lambda)$ för alla λ och om $\sum_\lambda m_a(\lambda) = n$ (t ex om $k = \mathbb{C}$).

Anmärkning 3.32. Om λ är ett egenvärde så gäller per definition att $L(E_\lambda) \subseteq E_\lambda$. Vi kan därför *restricera* (begränsa) L till en operator $L_\lambda: E_\lambda \rightarrow E_\lambda$ på delrummet E_λ . Om $m_a(\lambda) = m_g(\lambda)$ kan operatoren L_λ representeras med en diagonalmatris.



**SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS
FÖRELÄSNING 4 – HT24**

MATS BOIJ

4. SAMTIDIG DIAGONALISERING OCH JORDANS NORMALFORM

Målet för idag.

- Minimalpolynomet och Cayley–Hamiltons sats
- Samtidig diagonalisering
- Exponentialavbildningen och andra analytiska funktioner av matriser

Spår, determinant och egenvärden.

Definition 4.1. Spåret (en: *trace*) av en $n \times n$ -matris A är summan av diagonalelementen.

$$\operatorname{tr} A = \sum_{i=1}^n a_{ii}$$

Man kan visa att $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ och därmed att $\operatorname{tr}(PAP^{-1}) = \operatorname{tr}(AP^{-1}P) = \operatorname{tr}(A)$. Alltså beror inte spåret på val av bas och vi kan definiera spåret av en operator på ett ändligt dimensionellt vektorrum. Alternativt kan man visa att det karakteristiska polynomet är

$$p_A(x) = x^n - \operatorname{tr}(A)x^{n-1} + a_{n-2}x^{n-2} \cdots + a_1x + (-1)^n \det(A)$$

och eftersom det karakteristiska polynomet är invariant under konjugering så är även spåret det.

Om $k = \mathbb{C}$ så är $p_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$. Detta ger att

- Spåret är summan av egenvärdena $\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.
- Determinanten är produkten av egenvärdena $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$.

Detta kan också ses direkt genom att först överföra matrisen till övertriangulär form eftersom egenvärdena då är diagonalelementen.

Exempel 4.2. För en 2×2 -matris $A = (a_{ij})$ med egenvärden λ_1, λ_2 har vi att det karakteristiska polynomet är:

$$\begin{aligned} p_A(x) &= x^2 - \operatorname{tr}(A)x + \det(A) \\ &= x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{21}a_{12}) \\ &= x^2 - (\lambda_1 + \lambda_2)x + \lambda_1\lambda_2 \end{aligned}$$

Minimalpolynomet och Cayley–Hamiltons sats. Om $p(x) = a_d x^n + a_{d-1} x^{d-1} + \dots + a_1 x + a_0$ är ett polynom så kan vi evaluera det i en $n \times n$ -matris A på följande vis:

$$p(A) = a_d A^d + a_{d-1} A^{d-1} + \dots + a_1 A + a_0 I$$

och får en $n \times n$ -matris. Observera här att det konstanta monomet $a_0 = a_0 x^0$ blir $a_0 A^0 = a_0 I$ där I är identitetsmatrisen av storlek $n \times n$.

Definition 4.3 (Minimalpolynomet). Om A är en $n \times n$ -matris är *minimalpolynomet* $q_A(x)$ det *moniska*¹ polynom av lägsta grad så att $q_A(A) = 0$.

Anmärkning 4.4. Graden av minimalpolynomet är det lägsta heltal d så att $\{I, A, A^2, \dots, A^d\}$ är linjärt beroende eftersom $(x^d + a_d x^{d-1} + \dots + a_0)(A) = 0 \iff (A^d + \dots + a_0 I) = 0 \iff A^d$ är en linjärkombination av $\{I, A, A^2, A^3, \dots, A^{d-1}\}$. Vi noterar att $\deg q_A(x) \leq n^2$ eftersom $\{A^i\}_{i=0}^{n^2} = \{I, A, A^2, \dots, A^{n^2}\}$ är $n^2 + 1$ matriser i ett vektorrum av dimension n^2 och därmed linjärt beroende. Nästa sats ger att $\deg q_A(x) \leq n$.

Sats 4.5 (Cayley–Hamilton, LADR 8.37). $p_A(A) = 0$.

Exempel 4.6. För 2×2 -matriser har vi att $p_A(x) = x^2 - \text{tr}(A)x + \det(A)$ och Cayley–Hamiltons sats säger

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}^2 - (x_{11} + x_{22}) \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + (x_{11}x_{22} - x_{12}x_{21}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

vilket går att verifiera direkt.

Korollarium 4.7 (LADR, 8.46). $q_A(x)$ är en delare i $p_A(x)$.

Bevis. Vi utför polynomdivision av $p_A(x)$ med $q_A(x)$. Detta ger

$$p_A(x) = s(x)q_A(x) + r(x),$$

där antingen $r(x) = 0$ eller $\deg r(x) < \deg q_A(x)$. Sätter vi $x = A$ så är $r(A) = p_A(A) - s(A)q_A(A) = 0$, vilket ger $r(x) = 0$ eftersom $q_A(x)$ per definition är det polynom av lägst grad för vilket $q_A(A) = 0$. $\square$

Exempel 4.8. Låt $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ med $k = \mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$. Då har vi

$$A^2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{och} \quad A^3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I = A^2 + A.$$

Karakteristiska polynomet är $p_A(x) = x^2 + x + 1$ och eftersom $\{I, A\}$ är linjärt oberoende är detta också minimalpolynomet.

Exempel 4.9. Om vi har koefficienter i $k = \mathbb{F}_2$ får vi

$$B = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \implies B^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Det karakteristiska polynomet är $p_B(x) = x^4$ medan minimalpolynomet är $q_B(x) = x^2$.

Hur blir det med koefficienter i $\mathbb{Q}$ för samma matris?

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

Vi får $p_B(x) = x^4 - 2x^3$ och $q_B(x) = x^3 - 2x^2$.

¹med ledande koefficient ett

Anmärkning 4.10. Vi kan ordna en matris med ett givet karakteristiskt polynom $p_A(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$ genom

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}$$

Här är minimalpolynomet lika med det karakteristiska polynomet ty de första $n-k$ kolumnerna av A^k är $\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, \dots, \mathbf{e}_n$ och man ser därför lätt att $I, A, A^2, \dots, A^{n-1}$ är linjärt oberoende i $M_{n,n}$.

Sats 4.11 (LADR, 8.49). *Nollställena till minimalpolynomet $q_A(x)$ är precis egenvärdena till A .*

Bevis. Eftersom $q_A(x)$ är en delare till det karakteristiska polynomet $p_A(x)$ så måste varje nollställe till $q_A(x)$ vara ett nollställe till $p_A(x)$, dvs ett egenvärde.

Omvänt, om λ är ett egenvärde så finns en egenvektor ξ så att $A\xi = \lambda\xi$. Om $q_A(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ så är

$$\begin{aligned} 0 &= q_A(A)\xi = A^n\xi + a_{n-1}A^{n-1}\xi + \dots + a_1A\xi + a_0\xi \\ &= \lambda^n\xi + a_{n-1}\lambda^{n-1}\xi + \dots + a_1\lambda\xi + a_0\xi \\ &= q_A(\lambda)\xi. \end{aligned}$$

Eftersom $\xi \neq \mathbf{0}$ så är alltså $q_A(\lambda) = 0$. □

Bevis av Cayley–Hamiltons sats. Låt $k = \mathbb{C}$. Till att börja med kan vi se att $P_A(A) = 0$ gäller för diagonalmatriser.

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \Rightarrow p_D(D) = \begin{bmatrix} p_D(\lambda_1) & 0 & \cdots & 0 \\ 0 & p_D(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_D(\lambda_n) \end{bmatrix} = 0$$

eftersom $p_D(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$. Vi har att alla matriser i $M_n(\mathbb{C})$ kan approximeras godtyckligt nära av diagonaliserbara matriser eftersom alla matriser med distinkta nollställena till det karakteristiska polynomet är diagonaliserbara.

Eftersom $A \mapsto P_A(A)$ är en kontinuerlig funktion som är noll för alla diagonaliserbara matriser måste den vara noll för alla matriser.

Om vi skriver upp identiteten $P_A(A) = 0$ i variablerna $x_{11}, \dots, x_{nn}$ som i Exempel 4.6 är det en polynomidentitet med heltalskoefficienter. Därmed kommer den att gälla för alla kroppar k . □

Blockmatriser. En *blockmatris* är en $(m+n) \times (a+b)$ matris på formen

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$

där B och C har m rader, D och E har n rader, B och D har a kolumner och C och E har b kolumner.

Om $\mathbf{x} = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}$ är en kolonnvektor med $a+b$ kolumner så blir

$$A\mathbf{x} = \begin{bmatrix} B & C \\ D & E \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} B\mathbf{v} + C\mathbf{w} \\ D\mathbf{v} + E\mathbf{w} \end{bmatrix}$$

Om vi betraktar $\mathbb{R}^{a+b} = \mathbb{R}^a \oplus \mathbb{R}^b = V \oplus W$ och $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ som inre direkta summor och låter $L: \mathbb{R}^{a+b} \rightarrow \mathbb{R}^{m+n}$ vara den linjära avbildning som ges av A så har vi

$$L((\mathbf{v}, \mathbf{w})) = (B\mathbf{v} + C\mathbf{w}, D\mathbf{v} + E\mathbf{w})$$

där $\mathbf{v} \in \mathbb{R}^a$ och $\mathbf{w} \in \mathbb{R}^b$.

Vi säger att A är *block-övertriangulär* om $D = 0$. Då är $L((\mathbf{v}, \mathbf{0})) = (B\mathbf{v}, \mathbf{0})$ och L tar alltså $\mathbb{R}^a$ på $\mathbb{R}^m$.

Det viktigaste fallet är då $a = m$ och $b = n$. Då är A, B, E kvadratiske och L är en operator. Enkla beräkningar ger:

- (1) $\det(A) = \det(B)\det(E)$.
- (2) $\operatorname{tr}(A) = \operatorname{tr}(B) + \operatorname{tr}(E)$.
- (3) $p_A(x) = p_B(x)p_E(x)$.
- (4) Egenvärdena till A är egenvärdena av B och E tillsammans.
- (5) Eigenvektorer till B är eigenvektorer till A . Detta gäller dock ej för eigenvektorer till E .

Exempel 4.12. Betrakta

$$A = \begin{bmatrix} 1 & -2 & 7 \\ 1 & 4 & 5 \\ 0 & 0 & 8 \end{bmatrix}$$

som en övertriangulär block-matris där övre vänstra blocket har storlek 2×2 , dvs

$$B = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix}$$

i notationen ovan. Matrisen B har determinant 6 och spår 5 och alltså egenvärden $\{2, 3\}$. Alltså har A egenvärden $\{2, 3, 8\}$. Eigenvektorer till B är

- En egenvektor till $\lambda = 2$ är $(2, -1)$.
- En egenvektor till $\lambda = 3$ är $(1, -1)$.

Detta ger eigenvektorer $(2, -1, 0)$ och $(1, -1, 0)$ till A . Men $(0, 0, 1)$ är inte en egenvektor till egenvärdet $\lambda = 8$ utan lite beräkningar ger egenvektorn $(3, 7, 5)$.

Exempel 4.13. Matrisen

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

är en övertriangulär block-matris. Även om B och E är diagonaliserbara (de är diagonala!) så är inte A diagonaliserbar.

Vi säger att A är *block-diagonal* om $C = D = 0$. Då kan vi dela upp L som en operator $L|_V: V \rightarrow V$ som ges av matrisen B och en operator $L|_W: W \rightarrow W$ som ges av matrisen E . Eigenvektorer till L är då precis eigenvektorer till $L|_V$ och $L|_W$. Och A är diagonaliserbar precis när B och E är diagonaliserbara.

Operatorer och block-övertriangulära matriser. Om A är block-övertriangulär så är V *invariant* under L , dvs $L(\mathbf{v}) \in V$ för alla $\mathbf{v} \in V$. Vi får en operator $L|_V: V \rightarrow V$ som är L begränsad till V och som ges av matrisen A . T ex är en egenvektor till $L|_V$ en egenvektor till L . Däremot kan vi inte begränsa L till en operator på W eftersom $L(\mathbf{w})$ inte ligger i W för alla $\mathbf{w} \in W$ såvida inte $C = 0$.

Istället har vi en inducerad operator $U/V \rightarrow U/V$ på *kvotrummet* där $U = \mathbb{R}^{a+b} = V \oplus W$ på följande vis. Lagg först märke till att $W \rightarrow U \rightarrow U/V$ är en isomorfi eftersom W är ett komplement till V . Vi kan alltså representera varje vektor i U/V unikt som $\mathbf{w} + V$ där $\mathbf{w} \in W$. Applicera vi L på en vektor $\mathbf{w} + V$ får vi:

$$L(\mathbf{w} + V) = E\mathbf{w} + (B\mathbf{v} + C\mathbf{w}) = E\mathbf{w} + V$$

Så L inducerar en avbildning $U/V \rightarrow U/V$ som ges av matrisen E .

Samtidig diagonalisering och kommuterande operatorer.

Definition 4.14. Två operatorer L_1 och L_2 är *samtidigt diagonaliserbara* om det finns en gemensam bas av egenvektorer, dvs en bas så att *båda* operatorerna representeras av diagonalmatriser.

Anmärkning 4.15. Om L_1 och L_2 är samtidigt diagonaliserbara måste $L_1 \circ L_2 = L_2 \circ L_1$, eftersom detta gäller för diagonalmatriser.

Definition 4.16 (Kommuterande operatorer). Två operatorer L_1 och L_2 *kommuterar* om $L_1 \circ L_2 = L_2 \circ L_1$.

Sats 4.17 (Sadun, Thm. 4.10). Om $\dim V < \infty$ och L_1 och L_2 är diagonaliserbara är följande ekvivalent

- (a) L_1 och L_2 är samtidigt diagonaliserbara
- (b) L_1 och L_2 kommuterar.

Att (a) $\implies$ (b) är anmärkningen ovan.

Bevisidé för (b) $\implies$ (a). Om L_1 och L_2 kommuterar är egenrummen för operatoren L_1 är invarianta under L_2 och tvärtom: dvs om ξ är en egenvektor till L_1 med egenvärde λ , så är även $L_2(\xi)$ en egenvektor till L_1 med samma egenvärde. Detta följer av:

$$L_1(L_2(\xi)) = L_2 \circ L_1(\xi) = L_2(\lambda \xi) = \lambda L_2(\xi)$$

Om E_λ är egenrummet till L_1 med egenvärdet λ är alltså $L_2(E_\lambda) \subseteq E_\lambda$. Om vi därför väljer en egenbas för L_1 så blir matrisen för L_2 blockdiagonal i denna bas. Eftersom L_2 är diagonaliserbar måste varje block vara diagonaliserbart. $\square$

Exempel 4.18. Låt $L: V \rightarrow V$ vara en operator och $p(x)$ och $q(x)$ vara polynom. Då kommuterar $p(L)$ med $q(L)$. Detta gäller även om L inte är diagonaliserbar.

Exempel 4.19. Om $L_1, L_2: V \rightarrow V$ är kommuterande operatorer så kan vi definiera $p(L_1, L_2)$ för varje polynom $p(x, y) \in k[x, y]$.

Exponentialfunktionen och andra funktioner av matriser. Om $f(z)$ är en *analytisk* funktion² kring $z = 0$, t ex $f(z) = \exp(z)$, kan vi definiera $f(A)$ för en kvadratisk matris A genom att sätta

$$f(A) = \sum_{i=0}^{\infty} a_i A^i$$

förutsatt att serien konvergerar.

Definition 4.20 (Norm). Om $\|\cdot\|$ är den euklidiska normen på $\mathbb{R}^n$ eller $\mathbb{C}^n$ kan vi definiera normen av en $n \times n$ -matris som

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Anmärkning 4.21. Om $\|A\| < \infty$ får vi $\|A^i\| \leq \|A\|^i$ och konvergensen av $\sum_{i \geq 0} a_i A^i$ fås från konvergensen av $f(z) = \sum_{i \geq 0} a_i z^i$. Dvs, om serien för $f(z)$ konvergerar då $|z| < R$ så konvergerar serien för $f(A)$ då $\|A\| < R$.

Analytiska funktioner på diagonaliserbar matriser.

Sats 4.22. Om A är diagonaliserbar med $A = PDP^{-1}$ och $f(z) = \sum_{i \geq 0} a_i z^i$ en analytisk funktion kring $z = 0$ kan vi beräkna

$$f(A) = Pf(D)P^{-1} = P \begin{bmatrix} f(\lambda_1) & 0 & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & 0 & \cdots & 0 \\ 0 & 0 & f(\lambda_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f(\lambda_n) \end{bmatrix} P^{-1}$$

där serien är konvergent.

Bevis. För all $i \geq 0$ har vi:

$$A^i = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^i P^{-1}$$

vilket ger

$$f(A) = Pf(D)P^{-1}.$$

För en diagonalmatris $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ så är $D^i = \text{diag}(\lambda_1^i, \dots, \lambda_n^i)$ vilket ger

$$f(D) = \text{diag}(f(\lambda_1), \dots, f(\lambda_n)). \quad \square$$

²Detta betyder att Taylorserien i varje punkt konvergerar i en omgivning. Om f är en analytisk funktion definierad på hela $\mathbb{C}$ så konvergerar Taylor-serien överallt.



SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS FÖRELÄSNING 5 – HT24

MATS BOIJ

5. JORDANS NORMALFORM

Målet för idag.

- Jordans normalform
- Nilpotenta matriser
- Tillämpning: exponentialfunktionen av en matris
- Tillämpning: minimalpolynom, bevis av Cayley–Hamiltons sats
- Generaliserade egenvektorer
- Bevis av Jordans normalform

Jordans normalform. Jordans normalform säger att alla matriser nästan går att diagonalisera. Den enda skillanden är att vi får en del ettor på den så kallade superdiagonalen: diagonalen ovanför huvuddiagonalen.

Sats 5.1 (Jordans normalform, Sadun 4.13). *Låt L vara en operator på ett ändligt-dimensionellt vektorrum V . Om $p_L(x) = \prod_{i=1}^n (x - \lambda_i)$ (t ex om $k = \mathbb{C}$) finns en bas $\mathcal{B}$ så att matrisen för L blir [blockdiagonal](#)*

$$[L]_{\mathcal{B}} = \begin{bmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_r \end{bmatrix}$$

och varje block Λ_i är ett [Jordanblock](#) på formen

$$\Lambda_j = \begin{bmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

för något λ beroende på j . En matris på denna form sägs vara på [Jordans normalform](#).

Två matriser med $p_A(x) = p_B(x) = \prod_{i=1}^n (x - \lambda_i)$ är konjugerade om och endast om de har samma normalform upp till permutation av blocken.

Exempel 5.2. Här är två exempel på matriser på Jordans normalform:

$$A = \begin{bmatrix} \boxed{0} & \boxed{1} & 0 & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{4} & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & \boxed{4} & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & \boxed{4} \end{bmatrix} \quad B = \begin{bmatrix} \boxed{2} & \boxed{1} & 0 & 0 & 0 & 0 \\ \boxed{0} & \boxed{2} & \boxed{1} & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & \boxed{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{2} & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & \boxed{2} \end{bmatrix}$$

Nilpotenta matriser. Observera att ett Jordanblock av storlek $n \times n$ är på formen

$$\Lambda = \lambda I + N$$

där I är $n \times n$ -identitetsmatrisen och N är en matris som har nollor förutom 1:or på superdiagonalen.

Exempel 5.3. Om vi t ex har ett Jordanblock av storlek 4×4 får vi:

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

På samma sätt gäller för ett Jordan-block A av storlek $n \times n$ med diagonalelement λ att $(A - \lambda I)^n = 0$.

Definition 5.4. En kvadratisk matris A är *nilpotent* om $A^d = 0$ för något heltal $d \geq 0$.

Om A är övertriangulär så har A^d diagonalelementen $a_{11}^d, a_{22}^d, \dots, a_{nn}^d$. Om A är nilpotent så är alltså $a_{11} = a_{22} = \dots = a_{nn} = 0$. Man kan visa att omvändningen också är sann. Vi illustrerar detta med ett exempel (* är godtyckliga tal):

$$A = \begin{bmatrix} 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A^2 = \begin{bmatrix} 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Tillämpning: beräkna exponentialfunktionen av en matris. Om A är diagonaliserbar $A = PDP^{-1}$ kan vi beräkna

$$\exp(A) = P \exp(D) P^{-1}$$

och $\exp(D)$ är lätt att beräkna (se föregående föreläsning): om $D = \text{diag}(d_1, d_2, \dots, d_n)$ så är $\exp(D) = \text{diag}(e^{d_1}, \dots, e^{d_n})$. Samma sak gäller andra funktioner än $\exp$.

Om A inte är diagonaliserbar kan vi enligt Jordans sats ändå skriva $A = PBP^{-1}$ där B är på Jordans normalform och $\exp(A) = P \exp(B) P^{-1}$. Vi kan skriva Jordans normalform som $B = D + N$ där D är en diagonalmatris och N är en nilpotent matris. Vidare kommuterar D och N , dvs $DN = ND$. Detta ses genom att titta på varje Jordanblock $\Lambda = \lambda I + N$ och notera att λI kommuterar med varje annan matris. Slutsatsen blir då att:

$$\exp(B) = \exp(D + N) = \exp(D) \exp(N)$$

Observera att om två matriser A och B inte kommuterar gäller inte $\exp(A + B) = \exp(A) \exp(B)$ och inte heller $\exp(A) \exp(B) = \exp(B) \exp(A)$.

Vi har sett att det är lätt att beräkna $\exp(D)$. Det är även ganska lätt att beräkna $\exp(N)$: eftersom $N^n = 0$ får vi:

$$\exp(N) = I + N + \frac{N^2}{2!} + \dots + \frac{N^{n-1}}{(n-1)!}$$

Samma metod ger för en godtycklig analytisk funktion $f(z)$:

$$f(N) = f(0)I + f'(0)N + \frac{f''(0)}{2!}N^2 + \dots + \frac{f^{(n-1)}(0)}{(n-1)!}N^{n-1}$$

Tillämpning: minimalpolynom och Jordans normalform. Om $\Lambda = \lambda I + N$ är ett Jordan-block av storlek $n \times n$ så är $(\Lambda - \lambda I)^n = N^n = 0$ men $(\Lambda - \lambda I)^{n-1} = N^{n-1} \neq 0$. Från detta kan man visa:

Sats 5.5. Minimalpolynomet $q_A(x) = \prod_{\lambda} (x - \lambda)^{b_{\lambda}}$ där b_{λ} är storleken på det största Jordan-blocket med egenvärde λ .

I Exempel 5.2 så är de karaktäristiska polynomen och minimalpolynomen

$$\begin{aligned} p_A(x) &= q_A(x) = x^2(x-2)(x-4)^3, \\ p_B(x) &= (x-2)^6, \quad q_B(x) = (x-2)^3 \end{aligned}$$

Det innebär också att en matris A är diagonaliserbar precis när alla rötter har multiplicitet 1 i minimalpolynomet (se TENTA 2020-01-09, Uppg. 9).

Av Jordans normalform följer att den algebraiska multipliciteten a_{λ} hos λ är summan av storlekarna av alla Jordan-block med egenvärde λ . Alltså är det karakteristiska polynomet $p(x) = \prod_{\lambda} (x - \lambda)^{a_{\lambda}}$ och $a_{\lambda} \geq b_{\lambda}$. Detta ger ett bevis av Cayley–Hamiltons sats.

Generaliserade egenvektorer. Låt L vara en operator på ett vektorrum V .

Definition 5.6. En *generaliserad egenvektor* (Sadun: *power vector*) är en vektor $\xi \neq 0$ med

$$(L - \lambda I)^m \xi = 0, \quad \text{för något } m \geq 1.$$

Vi säger att ξ har ordning m om $(L - \lambda I)^m \xi = 0$ och $(L - \lambda I)^{m-1} \xi \neq 0$.

Definition 5.7. Det *generaliserade egenrummet* som hör till λ är

$$\tilde{E}_{\lambda} = \{\xi \in V : (L - \lambda I)^m \xi = 0, \text{ för något } m \geq 1\}.$$

Anmärkning 5.8. En generaliserad egenvektor av ordning 1 är en vanlig egenvektor. Om ξ är en generaliserad egenvektor av ordning m med egenvärde λ så är $(L - \lambda I)^{m-1} \xi$ en vanlig egenvektor med egenvärde λ . Alltså är det inga nya egenvärden som tillkommer för att vi betraktar generaliserade egenvektorer.

Exempel 5.9 (Jordanblock). Om A är ett Jordanblock av storlek $n \times n$ med egenvärde λ och $\mathbf{v} \in \text{Span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_i\}$, dvs om $\mathbf{v}$ har koordinater $(v_1, v_2, \dots, v_n)$ så är $v_{i+1} = v_{i+2} = \dots = v_n = 0$ så är $\mathbf{v}$ en generaliserad egenvektor med egenvärde λ och ordning $\leq i$ (och ordningen är exakt i om $v_i \neq 0$).

Vi har nämligen att $(A - \lambda I)\mathbf{v} \in \text{Span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{i-1}\}$ eftersom $A\mathbf{e}_i = \lambda\mathbf{e}_i + \mathbf{e}_{i-1}$.

Vi har alltså att $E_1 = \text{Span}\{\mathbf{e}_1\}$ och $\tilde{E}_1 = \mathbb{R}^n$.

Exempel 5.10 (Jordans normalform). Låt A vara på Jordans normalform, dvs:

$$A = \begin{bmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_r \end{bmatrix}.$$

Om λ är ett egenvärde så är:

- (1) den geometriska multipliciteten för λ antalet Jordanblock med egenvärde λ ;
- (2) den algebraiska multipliciteten för λ summan av storleken av Jordanblocken med egenvärde λ ;
- (3) egenrummet E_{λ} spannet av de standardvektorer $\mathbf{e}_i$ sådana att den i te raden är den första raden av ett Jordanblock med egenvärdet λ ;
- (4) det generaliserade egenrummet $\tilde{E}_{\lambda}$ spannet av de standardvektorer $\mathbf{e}_i$ sådana att den i te raden är en rad av ett Jordanblock med egenvärdet λ .

Bevis av Jordans normalform. Saduns skissar ett bevis av Jordans normalform i ett antal övningsuppgifter (§4.9, 4–12). I steg 10 är det dock ett ganska allvarligt fel.

I de följande avsnitten skissas ett något annorlunda bevis av Jordans normalform. Beviset fungerar då $k = \mathbb{C}$ eller mer allmänt när det karakteristiska polynomet faktoriseras i linjära faktorer.

- (1) Först överför vi matrisen till övertriangulär form.
- (2) Sedan gör vi den till block-diagonal form där varje block A_λ är övertriangulärt med samma egenvärde λ på alla diagonalelementen. När vi kommit så här långt har vi redan visat att det finns en bas av generaliserade egenvektorer.
- (3) Slutligen behandlar vi varje block för sig (som vi kan anta är nilpotent), och hittar en Jordanbas.

Steg 1: övertriangulär form. Vi börjar med att hitta en bas så att matrisen för en linjär operator blir övertriangulär ifall det karakteristiska polynomet kan faktoriseras i linjära faktorer.

Sats 5.11. Om $p_L(x) = \prod_{i=1}^n (x - \lambda_i)$ finns en bas så att matrisen för L blir övertriangulär med $\lambda_1, \lambda_2, \dots, \lambda_n$ på diagonalen.

$$\begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Bevisidé. • Välj $\mathbf{e}_1$ så att $L(\mathbf{e}_1) = \lambda_1 \mathbf{e}_1$ och låt $W = V / \text{Span}\{\mathbf{e}_1\}$.

- Om $\Phi: V \rightarrow W$ är kvotavbildningen faktoriseras $\Phi \circ L$ genom W eftersom $\text{Span}\{\mathbf{e}_1\} \subseteq \ker(\Phi \circ L)$ och vi har en unik avbildning $L_W: W \rightarrow W$ med $\Phi \circ L = L_W \circ \Phi$.
- Antag per induktion att påståendet stämmer för den inducerade avbildningen L_W på W . Observera att $p_L(x) = (x - \lambda_1)p_{L_W}(x)$.
- Basfallet då $\dim V = 1$ är trivialt. □

Exempel 5.12. Om $n = 2$ kan vi enligt Sats 5.11 ovan välja en bas så att matrisen för L blir

$$\begin{bmatrix} \lambda_1 & * \\ 0 & \lambda_2 \end{bmatrix}$$

Om $\lambda_1 \neq \lambda_2$ kan vi vidare välja basen så att vi får en diagonalmatris. Exempelvis

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

och mer generellt

$$\begin{bmatrix} 1 & \mu \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 & a \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & \mu \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & (\lambda_1 - \lambda_2)\mu + a \\ 0 & \lambda_2 \end{bmatrix}$$

och vi får en diagonalmatris om vi väljer $\mu = -a/(\lambda_1 - \lambda_2)$.

Steg 2: blockdiagonal form. Vi ska nu gå ett steg till och se att matrisen kan bli en blockdiagonal matris där varje block på diagonalen är övertriangulärt med samma egenvärde längs hela diagonalen.

Sats 5.13. Om $p_L(x) = \prod_{i=1}^r (x - \lambda_i)^{d_i}$ där $\lambda_1, \dots, \lambda_r$ är distinkta så finns en bas så att matrisen för L blir **blockdiagonal**

$$\begin{bmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_r \end{bmatrix} \quad \text{med} \quad \Lambda_i = \begin{bmatrix} \lambda_i & * & \cdots & * \\ 0 & \lambda_i & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

där Λ_i är en $d_i \times d_i$ -matris.

Bevis. Enligt Sats 5.11 kan vi välja en bas så att matrisen blir övertriangulär. När vi ordnar till den övertriangulära matrisen kan vi välja ordningen så att alla egenvärden som är lika hanteras tillsammans och vi får då att elementen på diagonalen är blockvis lika. Vi ska nu se att det går att ordna till så att alla element utanför dessa block blir noll.

Vi börjar med att se på hur det är med en 2×2 -matris som är övertriangulär med två olika egenvärden. Vi såg i exemplet ovan att om vi byter bas enligt

$$\begin{bmatrix} 1 & \mu \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 & a \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & \mu \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & (\lambda_1 - \lambda_2)\mu + a \\ 0 & \lambda_2 \end{bmatrix}$$

får vi en diagonalmatris om vi väljer $\mu = -a/(\lambda_1 - \lambda_2)$.

Vi kan använda denna metod successivt från vänster till höger, nedifrån och upp för att eliminera alla element ovanför blockdiagonalen. Vi börjar med understa raden från vänster och går till höger, därefter nästundersta osv. På så vis kommer de nya eliminationerna inte påverkar de nollor vi redan åstadkommit.

Detta leder till en *blockdiagonal* matris där blocken är övertriangulära med konstanta diagonalelement. $\square$

Anmärkning 5.14. I blocket Λ_i så är $\mathbf{e}_j$ en generaliserad egenvektor med egenvärde λ_i och ordning j eftersom $L\mathbf{e}_j \in \text{Span}\{\mathbf{e}_1, \dots, \mathbf{e}_{j-1}\}$. Därmed har vi en bas av generaliserade egenvektorer till L .

Anmärkning 5.15. Alltså räcker det nu att veta hur vi kan hantera dessa block. Vi noterar att $N_i := \Lambda_i - \lambda_i I$ är *nilpotent*, dvs uppfyller $N_i^m = 0$, för något m . Kan vi välja en bas så att basbytet PN_iP^{-1} uppfyller Jordans normalsats så ger detta även Jordans normalform för Λ_i eftersom

$$P\Lambda_iP^{-1} = P(\lambda_i I + N_i)P^{-1} = \lambda_i I + PN_iP^{-1}$$

Steg 3: nilpotenta operatorer. Vi har sett från Steg 2 att det räcker att bevisa Jordans normalform för nilpotenta operatorer.

Definition 5.16. En operator L är *nilpotent* om $L^m = 0$ för något positivt heltal m .

Sats 5.17. Om L är nilpotent kan vi välja en bas för V så att matrisen för L blir blockdiagonal

$$\begin{bmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_r \end{bmatrix}$$

där varje block har formen

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

Bevisidé. Det finns ett s så att $L^s = 0$ och $L^{s-1} \neq 0$. Då är $V = \ker(L^s)$ och vi börjar med att välja ett delrum W_s så att

$$V = \ker(L^s) = \ker(L^{s-1}) \oplus W_s.$$

Fortsätt sedan att välja $W_{s-1} \subseteq \ker(L^{s-1})$ så att

$$\ker(L^{s-1}) = \ker(L^{s-2}) \oplus L(W_s) \oplus W_{s-1}.$$

Det går eftersom $\ker(L^{s-2}) \subseteq \ker(L^{s-1})$, $L(W_s) \subseteq \ker(L^{s-1})$ och $\ker(L^{s-2}) \cap L(W_s) = \{0\}$.

Vi fortsätter på samma sätt tills vi har valt $W_1 \subseteq \ker(L^1)$ så att

$$\ker(L^1) = L^{s-1}(W_s) \oplus L^{s-2}(W_{s-1}) \oplus \cdots \oplus L(W_2) \oplus W_1.$$

För att se att det går att välja W_{s-k} i steget när vi skriver

$$\ker(L^{s-k}) = \ker(L^{s-k-1}) \oplus L^k(W_s) \oplus \cdots \oplus L(W_{s-k+1}) \oplus W_{s-k}$$

behöver vi se att spannet av alla utom den sista termen verkligen utgör en inre direkt summa, dvs att $\mathbf{0}$ inte kan skrivas som

$$\mathbf{0} = \mathbf{y} + L^k(\mathbf{x}_s) + \cdots + L^2(\mathbf{x}_{s-k+2}) + L(\mathbf{x}_{s-k+1})$$

där $\mathbf{x}_i \in W_i$ och $\mathbf{y} \in \ker(L^{s-k-1})$ utan att alla termer är noll. Vi får att

$$L^k(\mathbf{x}_s) + \cdots + L^2(\mathbf{x}_{s-k+2}) + L(\mathbf{x}_{s-k+1}) \in \ker(L^{s-k-1})$$

vilket innebär att

$$L^{k-1}(\mathbf{x}_s) + \cdots + L(\mathbf{x}_{s-k+2}) + \mathbf{x}_{s-k+1} \in \ker(L^{s-k})$$

men då måste alla dessa termer vara noll eftersom vi i steget innan hade

$$\ker(L^{s-k+1}) = \ker(L^{s-k}) \oplus L^{k-1}(W_s) \oplus \cdots \oplus L(W_{s-k+2}) \oplus W_{s-k+1}.$$

Välj nu baser för $W_1, W_2, \dots, W_s$ och de nollskilda bilderna av dessa potenser av L . Tillsammans ger dessa en bas $\mathcal{B}$ för V . Med denna bas, i lämplig ordning, kommer vi att få $\dim W_i$ block på formen ovan av storlek $i \times i$. Mer precist: för varje basvektor $\mathbf{w}$ i W_i , så ger $\mathbf{w}, L(\mathbf{w}), L^2(\mathbf{w}), \dots, L^{i-1}(\mathbf{w})$ en bas för ett block av storlek $i \times i$. $\square$

Exempel 5.18. Betrakta vektorrummet $V = \mathbb{C}[x, y]_{\leq 2}$ av polynom i x och y av grad ≤ 2 . Det har en bas $\{1, x, y, x^2, xy, y^2\}$. Vi har operatoren $L = \partial/\partial x$ på V . Vi får att $L^3 = 0$ och

$$\ker(L^2) = \text{Span}\{1, x, y, xy, y^2\} \quad \text{och} \quad \ker(L) = \text{Span}\{1, y, y^2\}$$

och

$$V = \ker(L^3) = \ker(L^2) \oplus \text{Span}\{x^2\},$$

$$\ker(L^2) = \ker(L) \oplus \text{Span}\{2x\} \oplus \text{Span}\{xy\},$$

$$\ker(L) = \underbrace{\ker(I)}_{=0} \oplus \text{Span}\{2\} \oplus \text{Span}\{y\} \oplus \text{Span}\{y^2\}$$

så $W_3 = \text{Span}\{x^2\}$, $W_2 = \text{Span}\{xy\}$ och $W_1 = \text{Span}\{y^2\}$. Med basen $\{2, 2x, x^2, y, xy, y^2\}$ får vi matrisen

$$\begin{bmatrix} \boxed{0} & \boxed{1} & \boxed{0} & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & \boxed{1} & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} \end{bmatrix}$$

Anmärkning 5.19. Antalet block av storlek $s \times s$ ges av dimensionen av W_s . Denna ges i sin tur enligt konstruktionen av

$$\dim W_s = \dim \ker(L^s) - \dim \ker(L^{s-1}) = (\dim V - \text{rk } L^s) - (\dim V - \text{rk } L^{s-1}) = \text{rk } L^{s-1} - \text{rk } L^s.$$

I allmänhet ges antalet block av storlek $i \times i$ av $\dim W_i$, för $i = 1, 2, \dots, s$. För $k = 1, 2, \dots, s-1$ får vi

$$\begin{aligned} \dim W_{s-k} + \dim W_{s-k+1} + \cdots + \dim W_s &= \dim \ker(L^{s-k}) - \dim \ker(L^{s-k-1}) \\ &= (\dim V - \text{rk } L^{s-k}) - (\dim V - \text{rk } L^{s-k-1}) = \text{rk } L^{s-k-1} - \text{rk } L^{s-k}. \end{aligned}$$

Om vi beräknar rangerna av $L^0, L^1, \dots, L^s$ får vi en avtagande följd där skillnaderna ger antalet Jordanblock av minst en viss storlek. Exempelvis ger $\text{rk } L^0 - \text{rk } L^1$ totala antalet Jordanblock och $\text{rk } L^1 - \text{rk } L^2$ antalet Jordanblock som är minst 2×2 .

**SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS
FÖRELÄSNING 6 – HT24**

MATS BOIJ

6. INRE PRODUKTRUM**Målet för idag.**

- Inre produktrum över $\mathbb{R}$ och $\mathbb{C}$
- Ortogonala baser och Gram–Schmidts metod
- Projektion och ortogonalt komplement

Inre produkt över $\mathbb{R}$. I många tillämpningar har vi en *metrik* som kommer från en *inre produkt*. Det gör att vi har ett sätt att mäta *längd* av vektorer och *vinklar* mellan vektorer. Det ser lite olika ut för reella och för komplexa vektorrum.

Definition 6.1. En *bilinjär* avbildning är en avbildning $\varphi(\cdot, \cdot): U \times V \longrightarrow W$ som uppfyller

- $\varphi(\mathbf{x}, \cdot): V \longrightarrow W$ är linjär för alla $\mathbf{x} \in U$.
- $\varphi(\cdot, \mathbf{y}): U \longrightarrow W$ är linjär för alla $\mathbf{y} \in V$.

En *bilinjär form* på V är en bilinjär avbildning $B: V \times V \longrightarrow k$ och den är *symmetrisk* om $\varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x}), \forall \mathbf{x}, \mathbf{y} \in V$.

Definition 6.2. En *inre produkt* på ett vektorrum V över $\mathbb{R}$ är en symmetrisk bilinjär form $\langle \cdot | \cdot \rangle: V \times V \longrightarrow \mathbb{R}$ som uppfyller $\langle \mathbf{x} | \mathbf{x} \rangle > 0$ för alla $\mathbf{x} \neq 0$.

Definition 6.3. Ett *inre produktrum* är ett vektorrum V som är utrustat med en inre produkt.

Reella inre produkter som positivt definita symmetriska matriser.

Exempel 6.4. En bilinjär form på V kan alltid representeras av en matris om $\dim V < \infty$. Med en bas $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ så är en bilinjär form $\varphi(\cdot, \cdot)$ bestämd av

$$\varphi(\mathbf{b}_i, \mathbf{b}_j) = g_{ij}, \quad 1 \leq i, j \leq n.$$

vilket ger en matris $G = (g_{ij})$. Formen är symmetrisk om G är symmetrisk. Om $\mathbf{x}, \mathbf{y} \in V$ så är

$$\varphi(\mathbf{x}, \mathbf{y}) = ([\mathbf{x}]_{\mathcal{B}})^T G [\mathbf{y}]_{\mathcal{B}}$$

Definition 6.5. Matrisen G kallas för en *metrik*. Vill vi betona basen $\mathcal{B}$ kan vi skriva $G_{\mathcal{B}}$.

Sats 6.6. En symmetrisk $n \times n$ -matris G definierar en inre produkt om och endast om G är positivt definit.

Bevis. Med $\mathbf{x} = \sum y_i \mathbf{b}_i$ är $\langle \mathbf{x} | \mathbf{x} \rangle = \mathbf{y}^T G \mathbf{y}$, där $\mathbf{y} = [\mathbf{x}]_{\mathcal{B}} = (y_1, y_2, \dots, y_n)^T \in \mathbb{R}^n$. Alltså är $\langle \mathbf{x} | \mathbf{x} \rangle > 0$ för alla $\mathbf{x} \neq \mathbf{0}$ samma sak som att G är positivt definit. $\square$

Inre produkter över $\mathbb{C}$.

Definition 6.7. En avbildning $\varphi(\cdot, \cdot): U \times V \longrightarrow W$ för komplexa vektorrum är *seskvilinjär* om

- $\varphi(\mathbf{x}, \cdot)$ är linjär för alla $\mathbf{x} \in U$
- $\varphi(\cdot, \mathbf{y})$ är antilinjär för alla $\mathbf{y} \in V$, dvs $\varphi(c_1\mathbf{x}_1 + c_2\mathbf{x}_2, \mathbf{y}) = \overline{c_1}\varphi(\mathbf{x}_1, \mathbf{y}) + \overline{c_2}\varphi(\mathbf{x}_2, \mathbf{y})$.

En seskvilinjär form på V är *konjugatsymmetrisk* om $\varphi(\mathbf{x}, \mathbf{y}) = \overline{\varphi(\mathbf{y}, \mathbf{x})}$, $\forall \mathbf{x}, \mathbf{y} \in V$.

Definition 6.8. En *inre produkt* på ett komplext vektorrum är en konjugatsymmetrisk seskvilinjär form $\langle \cdot | \cdot \rangle: V \times V \longrightarrow \mathbb{C}$ som uppfyller att $\langle \mathbf{x} | \mathbf{x} \rangle > 0$ för alla $\mathbf{x} \neq 0$.

Anmärkning 6.9. Det spelar egentligen ingen roll vilken av de två argumenten vi konjugerar så länge vi är konsekventa. I kursen SF1683 används (kanske) konventionen att det är det andra argumentet som konjugeras och notationen är (kanske) $\langle \mathbf{x}, \mathbf{y} \rangle$ i stället för $\langle \mathbf{x} | \mathbf{y} \rangle$.

Exempel 6.10.

- För $\mathbb{C}^n$ ger $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i=1}^n \overline{x_i} y_i$ en inre produkt.
- För $C^0([0, 1])$ ger $\langle f | g \rangle = \int_0^1 \overline{f(t)} g(t) dt$ en inre produkt.

Komplexa inre produkter som positivt definita från konjugatsymmetriska matriser.

Exempel 6.11. En seskvilinjär form på V kan alltid representeras av en matris om $\dim V < \infty$. Med en bas $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ definieras $(\cdot, \cdot)$ av

$$(\mathbf{b}_i, \mathbf{b}_j) = g_{ij}, \quad 1 \leq i, j \leq n.$$

Formen är konjugatsymmetrisk om $g_{ij} = \overline{g_{ji}}$, $\forall i, j$.

Definition 6.12. Om A är en matris betecknar $A^\dagger = \overline{A^T}$ konjugatet av den transponerade matrisen¹. Vi säger att A är konjugatsymmetrisk eller *Hermiteisk* om $A^\dagger = A$.

Matrisen för en konjugatsymmetrisk seskvilinjär form är alltså konjugatsymmetrisk (Hermiteisk).

Sats 6.13. En konjugatsymmetrisk $n \times n$ -matris G definierar en inre produkt om och endast om G är positivt definit.

Norm och vinklar.

Definition 6.14 (norm). *Normen* av en vektor ges av $|\mathbf{x}| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$.

Sats 6.15 (Cauchy–Schwarz olikhet). $|\langle \mathbf{x} | \mathbf{y} \rangle| \leq |\mathbf{x}| \cdot |\mathbf{y}|$ med likhet om och endast om $\mathbf{x} \parallel \mathbf{y}$.

Sats 6.16 (Triangelolikheten). $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$.

Definition 6.17 (vinkel). *Vinkeln* θ mellan $\mathbf{x}$ och $\mathbf{y}$ definieras av $\cos \theta = \frac{\langle \mathbf{x} | \mathbf{y} \rangle}{|\mathbf{x}| \cdot |\mathbf{y}|}$, för reella inre-produktrum.

Bevis för Cauchy–Schwarz olikhet. Om $\mathbf{x} \parallel \mathbf{y}$ har vi likhet. Antag att $\mathbf{x} \not\parallel \mathbf{y}$. Då är $W = \text{Span}\{\mathbf{x}, \mathbf{y}\}$ tvådimensionellt. Matrisen för inre produkten på W ges av

$$G = \begin{bmatrix} \langle \mathbf{x} | \mathbf{x} \rangle & \langle \mathbf{x} | \mathbf{y} \rangle \\ \langle \mathbf{y} | \mathbf{x} \rangle & \langle \mathbf{y} | \mathbf{y} \rangle \end{bmatrix}$$

Spåret är $\text{tr}(G) = |\mathbf{x}|^2 + |\mathbf{y}|^2 > 0$ och determinanten är $\det(G) = |\mathbf{x}|^2 |\mathbf{y}|^2 - |\langle \mathbf{x} | \mathbf{y} \rangle|^2$, så vi vill visa att determinanten är positiv. Egenvärdena är reella eftersom

$$\left(\frac{1}{2} \text{tr}(G)\right)^2 - \det(G) = \left(\frac{|\mathbf{x}|^2 + |\mathbf{y}|^2}{2}\right)^2 - |\mathbf{x}|^2 |\mathbf{y}|^2 + |\langle \mathbf{x} | \mathbf{y} \rangle|^2 = \left(\frac{|\mathbf{x}|^2 - |\mathbf{y}|^2}{2}\right)^2 + |\langle \mathbf{x} | \mathbf{y} \rangle|^2 \geq 0$$

Om det finns ett egenvärde $\lambda \leq 0$ finns en egenvektor som ger $\overline{\xi} G \xi = \lambda |\xi|^2 \leq 0$, vilket ger motsägelse. Alltså måste $\det(G) > 0$.

¹Det är också vanligt med A^* eller A^H istället för $A^\dagger$. Ibland betyder dock A^* bara det komplexa konjugatet $\overline{A}$.

(Mer allmänt så har varje Hermitesk, dvs konjugatsymmetrisk, matris reella egenvärden vilket vi kommer se senare. En Hermitesk matris är positivt definit precis när alla egenvärden är positiva.) $\square$

Bevis för triangelolikheten. Med Cauchy–Schwarz olikhet får vi

$$|\mathbf{x} + \mathbf{y}|^2 = |\mathbf{x}|^2 + \langle \mathbf{x} | \mathbf{y} \rangle + \langle \mathbf{y} | \mathbf{x} \rangle + |\mathbf{y}|^2 \leq |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2 = (|\mathbf{x}| + |\mathbf{y}|)^2. \quad \square$$

Ortogonala baser.

Definition 6.18. En bas $\mathcal{B}$ för ett inre produktrum är *ortogonal* om $\langle \mathbf{x} | \mathbf{y} \rangle = 0$ för alla $\mathbf{x} \neq \mathbf{y}$ i $\mathcal{B}$. Basen $\mathcal{B}$ är en *ortonormal* bas om dessutom $|\mathbf{x}| = 1$ för alla $\mathbf{x} \in \mathcal{B}$.

Anmärkning 6.19. Om $\mathcal{B} = \{\mathbf{b}_i\}_{i \in I}$ är en ortogonal bas är

$$\mathbf{x} = \sum_{i \in I} a_i \mathbf{b}_i \iff a_i = \frac{\langle \mathbf{b}_i | \mathbf{x} \rangle}{\langle \mathbf{b}_i | \mathbf{b}_i \rangle}, \forall i \in I$$

eftersom $\langle \mathbf{b}_i | \mathbf{x} \rangle = a_i \langle \mathbf{b}_i | \mathbf{b}_i \rangle$ och om $\mathcal{E} = \{\mathbf{e}_i\}_{i \in I}$ är ortonormal är därmed

$$\mathbf{x} = \sum_{i \in I} a_i \mathbf{e}_i \iff a_i = \langle \mathbf{e}_i | \mathbf{x} \rangle \forall i \in I.$$

Gram–Schmidts metod.

Sats 6.20. Ett ändligdimensionellt inre produktrum har en ortogonal bas.

Bevisidé. En maximal ortogonal mängd måste vara en ortogonal bas. $\square$

Vi kan också göra det algoritmiskt genom

Sats 6.21 (Gram–Schmidts metod). Om $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ är en bas för V får vi en ortogonal bas för V genom

$$\mathbf{y}_i = \mathbf{x}_i - \sum_{j=1}^{i-1} \frac{\langle \mathbf{y}_j | \mathbf{x}_i \rangle}{\langle \mathbf{y}_j | \mathbf{y}_j \rangle} \mathbf{y}_j, \quad i = 1, 2, \dots, n.$$

Anmärkning 6.22. Vi kan också göra detta om vi har en uppräknelig bas för V .

Exempel 6.23. Om $V = P = \mathbb{R}[x]$ med $\langle p | q \rangle = \int_{-1}^1 p(x)q(x) dx$ kan vi börja med $\mathcal{B} = \{1, x, x^2, \dots\}$ och få en ortogonal bas av polynom. Denna ortogonal bas (normerade så att $p(1) = 1$) kallas för Legendrepolynom.

Ortogonalt komplement.

Definition 6.24. Om $W \subseteq V$ i ett inre produktrum är det *ortogonala komplementet*

$$W^\perp = \{\mathbf{x} \in V : \langle \mathbf{x} | \mathbf{y} \rangle = 0, \forall \mathbf{y} \in W\}$$

Sats 6.25. Om $W \subseteq V$ i ett inre produktrum och $\dim W < \infty$ kan vi skriva

$$V = W \oplus W^\perp$$

Bevis. Vi kan bilda en ortonormal bas $\{\mathbf{e}_i\}_{i \in I}$ för W och får att

$$\mathbf{y} = \mathbf{x} - \sum_{i \in I} \langle \mathbf{e}_i | \mathbf{x} \rangle \mathbf{e}_i \in W^\perp \quad \square$$

Anmärkning 6.26. Om W är oändligdimensionellt kan vi inte göra så. Tag till exempel

$$V = \{f \in C(\mathbb{R}) : \int_{-\infty}^{\infty} (f(t))^2 dt < \infty\}$$

och

$$W = \{f \in V : f \text{ har kompakt stöd}\}.$$

Då är $W^\perp = \{0\}$ men $W \neq V$.

**SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS
FÖRELÄSNING 7 – HT24**

MATS BOIJ

7. HILBERTRUM OCH DUALA RUM

Målet för idag.

- Hilbertrum
- Duala rum

Hilbertrum.

Definition 7.1 (Cauchyföljd). En följd $\{\mathbf{x}_i\}_{i=0}^\infty$ i ett inre produktrum är en *Cauchyföljd* om $\forall \varepsilon > 0 \exists N \in \mathbb{N}$ så att $|\mathbf{x}_i - \mathbf{x}_j| < \varepsilon$ då $i, j \geq N$.

Definition 7.2 (Hilbertrum). Ett *Hilbertrum* är ett *fullständigt* inre produktrum V , dvs ett inre produktrum V där alla *Cauchyföljder* konvergerar: givet en Cauchyföljd $\{\mathbf{x}_i\}_{i=0}^\infty$ så finns $\mathbf{x} \in V$ så att $\lim_{i \rightarrow \infty} |\mathbf{x} - \mathbf{x}_i| = 0$.

Anmärkning 7.3. Alla ändligdimensionella inre produktrum är Hilbertrum i och med att både $\mathbb{R}$ och $\mathbb{C}$ är fullständiga.

Vi kan skapa mening i oändliga summor, *serier*, om delsummorna är Cauchyföljder.

Vi har nu två mycket viktiga Hilbertrum som ungefär på samma sätt som $\mathbb{R}^n$ och $\mathbb{C}^n$ ger en modell för oändligdimensionella Hilbertrum (med uppräknelig ortonormal bas).

Definition 7.4. Vi har följande två Hilbertrum¹

- $\ell^2(\mathbb{R}) = \{\{x_i\}_{i \geq 0} : \sum_{i \geq 0} x_i^2 < \infty\} \subseteq \prod_{i \geq 0} \mathbb{R}$ med $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i \geq 0} x_i y_i$.
- $\ell^2(\mathbb{C}) = \{\{x_i\}_{i \geq 0} : \sum_{i \geq 0} |x_i|^2 < \infty\} \subseteq \prod_{i \geq 0} \mathbb{C}$ med $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i \geq 0} \bar{x}_i y_i$.

Precis som med vektorer $\mathbf{x} = (x_1, x_2, \dots, x_n)$ i $\mathbb{R}^n$ och $\mathbb{C}^n$ så skriver vi $\mathbf{x} = (x_0, x_1, x_2, \dots) = (x_i)_{i=0}^\infty$ för vektorer i $\ell^2(\mathbb{R})$ och $\ell^2(\mathbb{C})$.

Anmärkning 7.5. Det är några saker att verifiera här (Sadun, Thm. 6.8). Låt $\mathbf{x}, \mathbf{y} \in \ell^2(\mathbb{R})$, dvs antag att serierna $\sum_{i \geq 0} x_i^2$ och $\sum_{i \geq 0} y_i^2$ är konvergenta. Då behöver vi visa:

- (1) att $\ell^2(\mathbb{R})$ är ett vektorrum: dvs att $\sum_{i \geq 0} (x_i + y_i)^2 = \sum_{i \geq 0} x_i^2 + 2x_i y_i + y_i^2$ är konvergent och att $\sum_{i \geq 0} (cx_i)^2$ är konvergent;
- (2) att den inre produkten $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i \geq 0} x_i y_i$ konvergerar.

Båda dessa påståenden följer av att $|x_i y_i| \leq \max\{x_i^2, y_i^2\} \leq x_i^2 + y_i^2$ eftersom $\sum_{i \geq 0} x_i^2$ och $\sum_{i \geq 0} y_i^2$ konvergerar.²

Datum: 2023-11-13.

¹Det förekommer att man även skriver ℓ_2 vilket Sadun gör. Sadun är dock lite inkonsekvent och skriver ℓ_2 för rum av följder men L^2 för funktionsrum.

²Olikheten mellan aritmetiskt och geometriskt medelvärde ger en starkare olikhet: $|x_i y_i| = \sqrt{x_i^2 y_i^2} \leq \frac{x_i^2 + y_i^2}{2}$.

För $\ell^2(\mathbb{C})$ blir beräkningarna liknande: $|x_i + y_i|^2 = |x_i|^2 + 2\operatorname{Re}(\overline{x_i}y_i) + |y_i|^2 \leq |x_i|^2 + 2|x_i| \cdot |y_i| + |y_i|^2$ och så vidare.

Vi har nu visat att $\ell^2(\mathbb{R})$ och $\ell^2(\mathbb{C})$ är inre produktrum. Man kan också visa att de är Hilbert-
rum, dvs fullständiga (se Sadun, 6.8, Exc. 4).

Baser i Hilbertrum.

Definition 7.6. En ortogonal bas för ett Hilbertrum H är en ortogonal mängd $\{\mathbf{x}_i\}_{i \in I}$ så att

$$H = \left\{ \sum_{i \in I} a_i \mathbf{x}_i : \text{serien konvergerar} \right\}$$

Exempel 7.7. En ortogonal bas för $\ell^2(\mathbb{R})$ ges av $\{\mathbf{e}_j\}_{j \geq 0}$ där $\mathbf{e}_j = (\delta_{ij})_{i=0}^\infty = (0, 0, \dots, 0, 1, 0, \dots)$ är följderna där j te positionen är en etta.

Anmärkning 7.8. En ortogonal bas i ett Hilbertrum är inte en bas i vanlig mening eftersom vi tillåter *oändliga linjärkombinationer* för att spänna upp Hilbertrummet. Att mängden är ortogonal innebär dock att den är linjärt oberoende precis som en vanlig bas. I ett Hilbertrum brukar en bas i vanlig mening kallas för en Hamelbas eller en algebraisk bas. Hamelbaser i oändligtdimensionella Hilbertrum är alltid överuppräknliga medan t ex $\ell^2(\mathbb{R})$ har en uppräknlig ortogonal bas.

Funktionsrum. För att hantera funktioner på intervall eller andra mångfaldar behöver vi Hilbertrummen $L^2([0, 1], \mathbb{R})$ och $L^2([0, 1], \mathbb{C})$. Den inre produkten ges av

$$\langle f | g \rangle = \int_0^1 f(t)g(t) dt \quad \text{resp.} \quad \langle f | g \rangle = \int_0^1 \overline{f(t)}g(t) dt$$

- För att detta ska vara väldefinierat måste $|f|^2$ vara *integrerbar*.
- Det räcker inte att ta bara kontinuerliga funktioner, eftersom gränsvärden av Cauchy-följder i $C([0, 1])$ inte alltid ligger i $C([0, 1])$.
- När vi lägger till alla sådana gränsvärden finner vi funktioner där $|f| = 0$ men $f \neq 0$. Vi behöver ta kvoten med detta delrum.

Exempel 7.9. Funktionerna $f_1 = x$, $f_2 = x^2$, $f_3 = x^3$, $\dots$, på intervallet $[0, 1]$ är en Cauchyföljd och konvergerar mot funktionen

$$f(x) = \begin{cases} 0 & \text{om } 0 \leq x < 1 \\ 1 & \text{om } x = 1 \end{cases}$$

Denna funktion har normen $\|f\| = \int_0^1 f(t)^2 dt = 0$ men är inte 0.

På liknande sätt går det att definiera $L^2(U, \mathbb{R})$ för andra intervall $U \subseteq \mathbb{R}$ (både öppna och slutna), inklusive obegränsade intervall som t ex $U = \mathbb{R}$. Man börjar då istället med kontinuerliga funktioner med kompakt stöd $C_0(U)$: dvs kontinuerliga funktioner $f: U \rightarrow \mathbb{R}$ som är 0 utanför något slutet begränsat intervall. Konstruktionen av $L^2(U, \mathbb{C})$ är analog.

Exempel 7.10 (Fourierserier, SF1683, Sadun Thm 8.5). En bas för $L^2([0, 1], \mathbb{R})$ ges av funktionerna

$$\{1, \cos(2n\pi x), \sin(2n\pi x)\}_{n=1,2,\dots}$$

Detta ger även en bas för $L^2([0, 1], \mathbb{C})$. En alternativ bas för $L^2([0, 1], \mathbb{C})$ är

$$\{e^{2n\pi i x}\}_{n=0,\pm 1,\pm 2,\dots}$$

Duala rum.

Definition 7.11. En *linjär form* på ett vektorrum V över en kropp k är en linjär avbildning $V \rightarrow k$.

Anmärkning 7.12. Linjära former kallas också för funktionaler och kovektorer.

Exempel 7.13. Om $\varphi: V \times V \rightarrow k$ är en bilinjär form över $\mathbb{R}$ så är $\varphi(-, \mathbf{x})$ och $\varphi(\mathbf{x}, -)$ linjära former. Över $\mathbb{C}$ är endast den andra en linjär form. Om vi har en inre produkt $\langle \cdot | \cdot \rangle$ så är alltså $\langle \mathbf{x} | \cdot \rangle$ en linjär form och vi skriver denna som $\langle \mathbf{x} |$.

Definition 7.14. För ett vektorrum V över en kropp k är *duala rummet*, eller *dualen*³,

$$V^* = \text{Hom}_k(V, k) = \{ \varphi : \varphi \text{ är en linjär form på } V \}.$$

Sats 7.15.

- (1) Om V är ett reellt inre produktrum ger $\mathbf{x} \mapsto \langle \mathbf{x} |$ en injektiv linjär avbildning $V \rightarrow V^*$.
- (2) Om V är ett komplext inre produktrum ger $\mathbf{x} \mapsto \langle \mathbf{x} |$ en injektiv antilinjär avbildning $V \rightarrow V^*$.

Om $\dim V < \infty$ är dessa (anti)isomorfier.

Bevis. Beviset är detsamma i båda fallen. Att avbildningen är (anti)linjär följer direkt av definitionen av inre produkt. Om $\langle \mathbf{x} |$ är nollfunktionen på V betyder det speciellt att $\langle \mathbf{x} | \mathbf{x} \rangle = 0$, vilket ger $\mathbf{x} = 0$. $\square$

Exempel 7.16. Om $\dim V = \infty$ blir det inte en isomorfi. Till exempel är dualen till $V = \bigoplus_{i \geq 0} k$ lika med $V^* = \prod_{i \geq 0} k$.

Exempel 7.17. Om $\dim V < \infty$ och $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ är en bas för V ges den *duala basen* av $\{\mathbf{b}_1^*, \mathbf{b}_2^*, \dots, \mathbf{b}_n^*\}$ där $\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}$. Observera att den linjära formen $\mathbf{b}_i^*$ beror på hela basen och inte bara $\mathbf{b}_i$. Om V är ett inre produktrum så har vi också basen $\{\langle \mathbf{b}_1 |, \langle \mathbf{b}_2 |, \dots, \langle \mathbf{b}_n | \}$ för V^* . Om $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ är en *ortonormal bas*, så är $\mathbf{b}_i^* = \langle \mathbf{b}_i |$ för alla i men i allmänhet så är $\mathbf{b}_i^* \neq \langle \mathbf{b}_i |$ (se nedan).

Linjära former som radvektorer. Låt $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ vara en bas för V . Detta ger en isomorfi $\mathbb{R}^n \rightarrow V$: vi identifierar $\mathbf{v} \in V$ med koordinatvektorn $[\mathbf{v}]_{\mathcal{B}} \in \mathbb{R}^n$. Det duala rummet V^* kan vi då identifiera med linjära avbildningar $\mathbb{R}^n \rightarrow \mathbb{R}^1$, dvs $(1 \times n)$ -matriser, dvs radvektorer. Om $\varphi \in V^*$ så låter vi $[\varphi]_{\mathcal{B}}$ beteckna motsvarande radvektor⁴. Om $\mathbf{v} \in V$ och $\varphi \in V^*$ så är $\varphi(\mathbf{v}) = [\varphi]_{\mathcal{B}} [\mathbf{v}]_{\mathcal{B}}$ (matrismultiplikation av en radvektor med en kolumnvektor). Den duala basvektorn $\mathbf{b}_i^*$ svarar mot radvektorn $(\mathbf{e}_i)^T$.

Exempel 7.18. Om V är ett inre produktrum och G är metriken i basen $\mathcal{B}$ så svarar den linjära formen $\langle \mathbf{v} |$ mot radvektorn

$$[\langle \mathbf{v} |]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}}^T G$$

ty $[\mathbf{v}]_{\mathcal{B}}^T G [\mathbf{w}]_{\mathcal{B}} = \langle \mathbf{v} | \mathbf{w} \rangle$. Speciellt är

$$[\langle \mathbf{b}_i |]_{\mathcal{B}} = (\mathbf{e}_i)^T G.$$

³Ibland används beteckningen $V^\vee$.

⁴I notationen för matrisen för linjära avbildningar hade vi skrivit $[\varphi]_{\mathcal{E}, \mathcal{B}}$ där $\mathcal{E} = \{\mathbf{e}_1\}$ är standardbasen för $\mathbb{R}$.

**SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS
FÖRELÄSNING 8 – HT24**

MATS BOIJ

**8. REKURSION OCH DIFFERENTIALEKVATIONER; ADJUNGERADE OPERATORER OCH
SJÄLVADJUNGERADE OPERATORER****Målet för idag.**

- Tillämpningar på egenvärden och egenvektorer
 - Rekursion
 - Ordinära differentialekvationer
- Adjungerade operatorer $L^\dagger$ och deras matriser
- Självadjungerade operatorer och deras matriser
- Egenvärden och egenvektorer till självadjungerade operatorer

Linjär rekursion.

Definition 8.1. En *linjär rekursion* (av ordning m) är en talföljd $x_0, x_1, \dots$ som ges av

$$x_n = \sum_{i=1}^m a_{m-i} x_{n-i}, \quad \forall n \geq m$$

tillsammans med begynnelsedata $x_0, x_1, \dots, x_{m-1}$. Här är $a_0, a_1, \dots, a_{m-1}$ och $x_0, x_1, \dots$ tal ur en fixerad kropp k (alternativt skulle man kunna jobba med heltalen eller en godtycklig så kallad ring).

Anmärkning 8.2. Genom att betrakta vektorerna $\mathbf{y}(n) = (x_n, x_{n-1}, \dots, x_{n-m+1})$ kan rekursionen skrivas som $\mathbf{y}(n) = A\mathbf{y}(n-1)$ för alla $n \geq m$.

$$\underbrace{\begin{bmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-m+2} \\ x_{n-m+1} \end{bmatrix}}_{\mathbf{y}(n)} = \underbrace{\begin{bmatrix} a_{m-1} & a_{m-2} & \cdots & a_1 & a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_{n-1} \\ x_{n-2} \\ \vdots \\ x_{n-m+1} \\ x_{n-m} \end{bmatrix}}_{\mathbf{y}(n-1)}$$

Begynnelsedata är alltså samlade i begynnelsevektorn $\mathbf{y}(m-1)$. Man kan visa att det karakteristiska polynomet till A är

$$p_A(x) = x^m - a_{m-1}x^{m-1} - a_{m-2}x^{m-2} - \cdots - a_1x - a_0$$

(använda radutveckling kring den första raden för att beräkna determinanten av $xI - A$).

Exempel 8.3. Fibonaccitalen ges av rekursionen

$$x_n = x_{n-1} + x_{n-2}, \quad n \geq 2$$

med begynnelsedata $x_0 = 0, x_1 = 1$. Vi kan skriva det som

$$\mathbf{y}(n) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{y}(n-1), \quad \mathbf{y}(1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

och egenvärdena ges av

$$\lambda_{\pm} = \frac{\text{tr}(A)}{2} \pm \sqrt{\left(\frac{\text{tr}(A)}{2}\right)^2 - \det(A)} = \frac{1}{2} \pm \frac{\sqrt{5}}{2}.$$

Om motsvarande egenvektorer är ξ_+ och ξ_- så kan vi skriva begynnelsevektorn

$$\mathbf{y}(1) = \begin{bmatrix} x_1 \\ x_0 \end{bmatrix} = \alpha_+ \xi_+ + \alpha_- \xi_-$$

vilket ger

$$\begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} = \mathbf{y}(n) = A^{n-1} \mathbf{y}(1) = \alpha_+ \lambda_+^{n-1} \xi_+ + \alpha_- \lambda_-^{n-1} \xi_-$$

Låt $\lambda = \lambda_+ = \frac{1+\sqrt{5}}{2}$. Då är $\lambda_- = -\lambda^{-1}$ (ty $\det A = -1 = \lambda_+ \lambda_-$) och egenvektorerna är

$$\xi_+ = \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \quad \text{och} \quad \xi_- = \begin{bmatrix} -\lambda^{-1} \\ 1 \end{bmatrix}$$

($\lambda^2 = \lambda + 1$). Konstanterna α_+ och α_- bestäms till

$$\alpha_+ = \frac{1}{\sqrt{5}}, \quad \alpha_- = \frac{-1}{\sqrt{5}}$$

och alltså

$$x_n = \alpha_+ \lambda^n + \alpha_- (-\lambda)^{-n} = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^n$$

Utveckling i diskret tid. Vid vissa tillämpningar har vi tillståndsvektorer som ändras stegvis $\mathbf{x}(0), \mathbf{x}(1), \mathbf{x}(2), \dots$ i k^n och som uppfyller relationen

$$\mathbf{x}(i) = A\mathbf{x}(i-1), \quad i \geq 1$$

för en matris A . Vi såg att linjär rekursion är ett sådant fall.

Om vi kan diagonalisera A kan vi uttrycka allt i termer egenvektorer och vi får

$$\mathbf{x}(i) = A^i \mathbf{x}(0) = \sum_j \alpha_j (\lambda_j)^i \xi_j$$

där $A\xi_i = \lambda_i \xi_i$ och $\mathbf{x}(0) = \sum_j \alpha_j \xi_j$.

Större Jordanblock. Om $n \times n$ -matrisen A inte är diagonaliserbar, men vi har att $\bigoplus \tilde{E}_\lambda = k^n$, dvs att de generaliserade egenrummen spänner upp k^n (vilket inträffar om A har n stycken egenvärden räknat med algebraisk multiplicitet, t ex $k = \mathbb{C}$) kan vi hantera de generaliserade egenrummen i Jordanblock. För ett block Λ av storlek $s \times s$ med egenvärde λ har vi att $\Lambda = \lambda I + N$ och

$$\Lambda^k = (\lambda I + N)^k = \sum_{i=0}^k \binom{k}{i} \lambda^{k-i} N^i$$

enligt binomialsatsen och mer konkret får vi

$$\begin{bmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}^i = \left(\begin{bmatrix} \lambda & & & 0 & 0 \\ 0 & \lambda & & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & & \lambda & \\ 0 & 0 & & 0 & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \right)^i$$

$$= \begin{bmatrix} \lambda^i & \binom{i}{1}\lambda^{i-1} & \dots & \binom{i}{s-2}\lambda^{i-s+2} & \binom{i}{s-1}\lambda^{i-s+1} \\ 0 & \lambda^i & \dots & \binom{i}{s-3}\lambda^{i-s+3} & \binom{i}{s-2}\lambda^{i-s+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda^i & \binom{i}{1}\lambda^{i-1} \\ 0 & 0 & \dots & 0 & \lambda^i \end{bmatrix}$$

Exempel 8.4. Om vi har $x_{n+1} = 4x_n - 4x_{n-1}$ får vi matrisen

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

med karakteristiskt polynom $p_A(x) = (x-2)^2$. Den är inte diagonaliserbar och vi bestämmer generaliserade egenvektorer som får matrisen på Jordans normalform.

$$\xi_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \ker \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \quad \text{och} \quad \xi_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{eftersom} \quad \xi_1 = (A-2I)\xi_2$$

Vi får att

$$A^n = (2I + (A-2I))^n = 2^n I^n + n2^{n-1}I^{n-1}(A-2I) = 2^n I + n2^{n-1}(A-2I)$$

med hjälp av binomialsatsen eftersom $(A-2I)^2 = 0$. Explicit ger det

$$A^n = \begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix}^n = 2^n \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + n2^{n-1} \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} (1+n)2^n & -n2^{n+1} \\ n2^{n-1} & (1-n)2^n \end{bmatrix}.$$

Om vi skriver $\mathbf{x}(0) = \alpha_1 \xi_1 + \alpha_2 \xi_2$ får vi

$$\begin{aligned} A^n \mathbf{x}(0) &= \alpha_1 A^n \xi_1 + \alpha_2 A^n \xi_2 = \alpha_1 (2^n I + n2^{n-1}(A-2I)) \xi_1 + \alpha_2 (2^n I + n2^{n-1}(A-2I)) \xi_2 \\ &= \alpha_1 2^n \xi_1 + \alpha_2 (2^n \xi_2 + n2^{n-1} \xi_1) = (2^n \alpha_1 + n2^{n-1} \alpha_2) \xi_1 + 2^n \alpha_2 \xi_2. \end{aligned}$$

Utveckling i kontinuerlig tid. I andra tillämpningar beror tillståndsvektorerna på tiden som $\mathbf{x}(t)$ och uppfyller en första ordningens linjär differentialekvation

$$\mathbf{x}'(t) = A\mathbf{x}(t)$$

med begynnelsevärden $\mathbf{x}(0) = \mathbf{x}_0$.

Lösningen kan skrivas som $\mathbf{x}(t) = \exp(At)\mathbf{x}_0$ och om vi kan diagonalisera får vi

$$\mathbf{x}(t) = \sum_j \alpha_j \exp(\lambda_j t) \xi_j$$

där $A\xi_j = \lambda_j \xi_j$ och $\mathbf{x}_0 = \sum_j \alpha_j \xi_j$.

Kan vi inte diagonalisera kan vi välja en bas av generaliserade egenvektorer (inte nödvändigtvis en Jordanbas) så att matrisen blir blockdiagonal med ett egenvärde per block och skriva $A = D + N$ där D är diagonal och N är nilpotent och D och N kommuterar och vi får formeln

$$\exp(At) = \exp(Dt) \exp(Nt) = \exp(Dt) \left(1 + Nt + \frac{N^2}{2}t^2 + \dots + \frac{N^{n-1}}{(n-1)!}t^{n-1} \right)$$

Högre ordningen. Om vi har en högre ordningens linjärt system av differentialekvationer

$$\mathbf{x}''(t) = A\mathbf{x}(t) + B\mathbf{x}'(t)$$

kan vi omformulera det som en första ordningens differentialekvation genom

$$\begin{bmatrix} \mathbf{x}'(t) \\ \mathbf{x}''(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ A & B \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}'(t) \end{bmatrix}$$

och vi får

$$\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}'(t) \end{bmatrix} = \exp \left(\begin{bmatrix} 0 & It \\ At & Bt \end{bmatrix} \right) \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{x}'(0) \end{bmatrix}$$

Adjungerad operator.

Definition 8.5. Om L är en operator på ett inre produktrum V är den *adjungerade operatorn* $L^\dagger$ unikt definierad av att

$$\langle L^\dagger(\mathbf{x}) | \mathbf{y} \rangle = \langle \mathbf{x} | L(\mathbf{y}) \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in V.$$

Anmärkning 8.6. Vi behöver kontrollera att det finns en unik sådan operator. När V är ändligdimensionellt så följer det av att titta på matrisen (nästa avsnitt). När V är oändligdimensionellt så finns inte alltid $L^\dagger$.

Exempel 8.7. Betrakta operatoren L på $\ell^2(\mathbb{R})$ definierad av $L(\mathbf{e}_i) = \mathbf{e}_{i+1}$, dvs $L((a_0, a_1, a_2, \dots)) = (0, a_0, a_1, a_2, a_3, \dots)$. Då är

$$\langle L^\dagger(\mathbf{e}_i) | \mathbf{e}_j \rangle = \langle \mathbf{e}_i | L(\mathbf{e}_j) \rangle = \langle \mathbf{e}_i | \mathbf{e}_{j+1} \rangle = \delta_{i, (j+1)}$$

Det följer att $L^\dagger(\mathbf{e}_i) = \mathbf{e}_{i-1}$, dvs $L^\dagger((a_0, a_1, a_2, \dots)) = (a_1, a_2, a_3, \dots)$. (Vi kan också notera att $L^\dagger \circ L$ är identiteten medan $L \circ L^\dagger$ inte är identiteten.)

Exempel 8.8. Betrakta operatoren L på $\ell^2(\mathbb{R})$ definierad av $L(\mathbf{e}_i) = \mathbf{e}_1 + \mathbf{e}_2 + \dots + \mathbf{e}_i$. Då är

$$\langle L^\dagger(\mathbf{e}_i) | \mathbf{e}_j \rangle = \langle \mathbf{e}_i | L(\mathbf{e}_j) \rangle = \langle \mathbf{e}_i | \mathbf{e}_1 + \mathbf{e}_2 + \dots + \mathbf{e}_j \rangle = \begin{cases} 1 & \text{om } j \geq i \\ 0 & \text{annars} \end{cases}$$

Detta ger att $L(\mathbf{e}_i) = \mathbf{e}_i + \mathbf{e}_{i+1} + \dots = (0, 0, \dots, 0, 1, 1, 1, \dots)$ vilket inte är en konvergent följd. Därför existerar inte den adjungerade operatoren.

Matrisen för den adjungerade operator.

Sats 8.9 (Sadun, Thm. 7.2). Om A är matrisen för L med avseende på en *ortonormal* bas $\{\mathbf{e}_i\}_{i \in I}$ ges matrisen $A^\dagger$ för $L^\dagger$ av

$$(A^\dagger)_{ij} = \overline{A_{ji}}.$$

Bevis.

$$(A^\dagger)_{ij} = \langle \mathbf{e}_i | L^\dagger(\mathbf{e}_j) \rangle = \overline{\langle L^\dagger(\mathbf{e}_j) | \mathbf{e}_i \rangle} = \overline{\langle \mathbf{e}_j | L(\mathbf{e}_i) \rangle} = \overline{A_{ji}} \quad \square$$

Definition 8.10. Det *Hermiteska konjugatet* $A^\dagger$ av en matris A är konjugatet av transponatet.

Alltså är $[L^\dagger]_{\mathcal{B}} = ([L]_{\mathcal{B}})^\dagger$.

Exempel 8.11. I Exempel 8.7 har vi:

$$[L]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad [L^\dagger]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Exempel 8.12. I Exempel 8.8 har vi:

$$[L]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots \\ 0 & 1 & 1 & 1 & \dots \\ 0 & 0 & 1 & 1 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad [L]_{\mathcal{B}}^\dagger = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & \dots \\ 1 & 1 & 1 & 0 & \dots \\ 1 & 1 & 1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Men $[L]_{\mathcal{B}}^\dagger$ är inte matrisen för en operator eftersom normen av varje kolumn är oändlig.

Självadjungerade operatorer och Hermiteska matriser.

Definition 8.13. En operator L på ett inre produktrum är *självadjungerad*, eller *Hermitesk*, om $L = L^\dagger$. En matris A är *Hermitesk* om $A = A^\dagger$.

Sats 8.14 (Sadun, Thm. 7.3). *En operator L är självadjungerad om och endast om matrisen för L är Hermitesk för varje val av ortonormal bas för V .*

Bevis. L är självadjungerad $\iff L = L^\dagger \iff [L]_{\mathcal{B}} = [L^\dagger]_{\mathcal{B}} = [L]_{\mathcal{B}}^\dagger$ (om $\mathcal{B}$ är ortonormal). $\square$

Anmärkning 8.15. För reella inre produktrum är $A^\dagger = A^T$ och Hermiteska matriser är symmetriska matriser.

Egenvärden och egenvektorer till självadjungerade operatorer.

Sats 8.16 (Sadun, Thm. 7.4(a)). *Egenvärden till självadjungerade operatorer är reella.*

Bevis. $L(\xi) = \lambda\xi \implies \overline{\lambda}|\xi|^2 = \langle \lambda\xi | \xi \rangle = \langle L(\xi) | \xi \rangle = \langle \xi | L(\xi) \rangle = \langle \xi | \lambda\xi \rangle = \lambda|\xi|^2 \implies \overline{\lambda} = \lambda$. $\square$

Sats 8.17 (Sadun, Thm. 7.4(b)). *Egenvektorer till en självadjungerad operator som hör till olika egenvärden är ortogonala.*

Bevis. Om $L(\xi_1) = \lambda_1\xi_1$ och $L(\xi_2) = \lambda_2\xi_2$ får vi

$$\lambda_1 \langle \xi_1 | \xi_2 \rangle = \langle L(\xi_1) | \xi_2 \rangle = \langle \xi_1 | L(\xi_2) \rangle = \lambda_2 \langle \xi_1 | \xi_2 \rangle$$

och $\langle \xi_1 | \xi_2 \rangle = 0$ om $\lambda_1 \neq \lambda_2$ $\square$

Ortogonal baser av egenvektorer.

Sats 8.18 (Sadun, Cor. 7.5). *Om L är självadjungerad och diagonaliserbar finns en ortogonal bas av egenvektorer.*

Bevis. Eftersom L är diagonaliserbar är $V = \bigoplus_{\lambda} E_{\lambda}$. Enligt förra satsen är $E_{\lambda} \perp E_{\mu}$ om $\lambda \neq \mu$. De ortogonala baserna för varje egenrum ger en ortogonal bas för V som består av egenvektorer till L . $\square$

Ortogonal diagonalisering av självadjungerade operatorer.

Sats 8.19 (Sadun, Thm. 7.6). *För en självadjungerad operator på ett ändligtdimensionellt inre produktrum över $\mathbb{C}$ finns en ortogonal bas av egenvektorer.*

Bevis. Det finns minst ett egenvärde λ . Tag $\xi \neq 0$ med $L(\xi) = \lambda\xi$ och låt $W = \text{Span}\{\xi\}$. Vi kan skriva $V = W \oplus W^\perp$ (explicit: vi har basen $\{\xi\}$ för W ; utöka till en bas $\mathcal{B}$ för hela V och gör basen ortogonal med Gram-Schmidt; andra delen av basen är nu en ortogonal bas för $W^\perp$). Vidare är $L(W) \subseteq W$ eftersom ξ en egenvektor och $L(W^\perp) \subseteq W^\perp$ eftersom

$$\mathbf{y} \in W^\perp \implies \langle \mathbf{x} | L(\mathbf{y}) \rangle = \langle L(\mathbf{x}) | \mathbf{y} \rangle = \langle \lambda \mathbf{x} | \mathbf{y} \rangle = 0 \implies L(\mathbf{y}) \in W^\perp.$$

Alternativt kan vi skriva ner matrisen för L i basen $\mathcal{B}$. I blockform:

$$[L]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & * \\ 0 & B \end{bmatrix}$$

Eftersom $[L]_{\mathcal{B}}$ är Hermitesk så är $[L]_{\mathcal{B}}^\dagger = [L]_{\mathcal{B}}$. Alltså är radvektorn märkt med $*$ lika med nollvektorn. Alltså är $[L]_{\mathcal{B}}$ block-diagonal och $L(W^\perp) \subseteq W^\perp$. Matrisen för restriktionen av L till $W^\perp$ är B .

Per induktion över $\dim V$ kan vi anta att det finns ortogonal bas $\mathcal{B}'$ av egenvektorer för L på $W^\perp$ och vi får då att $\{\xi\} \cup \mathcal{B}'$ är en ortogonal bas av egenvektorer. (Basfallet för induktionen är trivialt.) $\square$

BILAGA A. LITE EXTRAMATERIAL SOM INTE TOGS UPP PÅ FÖRELÄSNINGEN

Duala avbildningar och duala operatorer. Lite mer om detta kommer i slutet av kursen.

Definition A.1. Om $L: V \rightarrow W$ är en linjär avbildning ges den *duala avbildningen* $L^*: W^* \rightarrow V^*$ av

$$L^*(\varphi)(\mathbf{x}) = \varphi(L(\mathbf{x})), \quad \forall \mathbf{x} \in V, \forall \varphi \in W^*.$$

Sats A.2. Om A är matrisen för L med avseende på baser $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ och $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ för V respektive W är A^T matrisen för L^* med avseende på de duala baserna $\{\mathbf{f}_1^*, \mathbf{f}_2^*, \dots, \mathbf{f}_m^*\}$, respektive $\{\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*\}$.

Anmärkning A.3. Om L är en operator på ett reellt inre produktrum och $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ är en ortonormal bas kan vi se V^* som V genom den inre produkten och den duala basen är $\mathcal{B}$. Den duala operatorn L^* sammanfaller då med den adjungerade operatorn $L^\dagger$.

För ett komplext inre produktrum gäller inte samma sak eftersom identifikationen av V med V^* som ges av den inre produkten inte är en *linjär* avbildning och därför inte en isomorfi av vektorrum.



SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS FÖRELÄSNING 9 – HT24

MATS BOIJ

9. ORTOGONALA OCH UNITÄRA OPERATORER

Målet för idag.

- Ortogonala operatorer och matriser
- Unitära operatorer och matriser
- Diagonalisering av ortogonala och unitära matriser.

Operatorer som bevarar längd på reella inre produktrum.

Anmärkning 9.1. En inre produkt på ett reellt vektorrum ges av dess norm eftersom

$$\langle \mathbf{x} | \mathbf{y} \rangle = \frac{|\mathbf{x} + \mathbf{y}|^2 - |\mathbf{x} - \mathbf{y}|^2}{4}.$$

Sats 9.2 (Sadun, Thm. 7.12). För en linjär operator på ett reellt inre produktrum V är följande ekvivalent:

- (i) L bevarar normen, dvs $|L(\mathbf{x})| = |\mathbf{x}|$, $\forall \mathbf{x} \in V$.
- (ii) L bevarar $\langle \cdot | \cdot \rangle$, dvs $\langle L(\mathbf{x}) | L(\mathbf{y}) \rangle = \langle \mathbf{x} | \mathbf{y} \rangle$, $\forall \mathbf{x}, \mathbf{y} \in V$.
- (iii) $L^\dagger \circ L = I$.

Om $\dim V < \infty$ är även följande ekvivalent med ovanstående

- (iv) $L^\dagger = L^{-1}$.
- (v) $L \circ L^\dagger = I$.
- (vi) L avbildar ortonormala baser på ortonormala baser.

Definition 9.3. En operator som bevarar normen, eller ekvivalent den inre produkten, kallas för en *isometri*.

Exempel 9.4. Att normen bevaras innebär att bara nollvektorn kan avbildas på nollvektorn, så operatoren måste vara injektiv. Däremot behöver den inte vara surjektiv. Exempelvis bevarar operatoren L på $\ell^2(\mathbb{R})$ som ges av $\{a_0, a_1, a_2, \dots\} \mapsto \{0, a_0, a_1, a_2, \dots\}$ normen, men är inte surjektiv. Alltså är L en isometri som inte är surjektiv.

Ortogonala operatorer och matriser.

Definition 9.5. En operator L på ett reellt inre produktrum V är en *ortogonal operator* om

- (1) L är inverterbar; och
- (2) $L^\dagger = L^{-1}$, dvs $\langle \mathbf{x} | L(\mathbf{y}) \rangle = \langle L^{-1}(\mathbf{x}) | \mathbf{y} \rangle$ för alla $\mathbf{x}, \mathbf{y} \in V$.

En *ortogonal matris* är en matris som motsvarar en ortogonal operator med avseende på en ortonormal bas.

Anmärkning 9.6. Eftersom L är inverterbar så är $\mathbf{x} = L(L^{-1}(\mathbf{x}))$ och med $\mathbf{z} = L^{-1}(\mathbf{x})$ blir det andra villkoret ekvivalent med $\langle L(\mathbf{z}) | L(\mathbf{y}) \rangle = \langle \mathbf{z} | \mathbf{y} \rangle$, $\forall \mathbf{z}, \mathbf{y} \in V$. En ortogonal operator är alltså detsamma som en inverterbar isometri. Observera att Exempel 9.4 är en icke-inverterbar isometri: $L^\dagger L = I$ men $LL^\dagger \neq I$.

Om V är ändligtdimensionell så medför isometri inverterbar:

Sats 9.7 (Sadun, Thm. 7.12). *Följande är ekvivalent för en reell kvadratisk $(n \times n)$ -matris A :*

- (i) A är ortogonal
- (ii) $AA^T = I$
- (iii) $A^T A = I$
- (iv) *Kolumnerna i A utgör en ortonormal bas för kolonnrummet.*
- (v) *Raderna i A utgör en ortonormal bas för radrummet.*

Operatorer som bevarar längd på komplexa inre produktrum.

Anmärkning 9.8. En komplex inre produkt bestäms genom normen genom

$$\langle \mathbf{x} | \mathbf{y} \rangle = \frac{|\mathbf{x} + \mathbf{y}|^2 - |\mathbf{x} - \mathbf{y}|^2 + i|\mathbf{x} - i\mathbf{y}|^2 - i|\mathbf{x} + i\mathbf{y}|^2}{4}.$$

Sats 9.9 (Sadun, Thm. 7.12). *För en operator L på ett komplext inre produktrum V är följande ekvivalent:*

- (i) L bevarar normen, dvs $|L(\mathbf{x})| = |\mathbf{x}|$, $\forall \mathbf{x} \in V$.
- (ii) L bevarar $\langle \cdot | \cdot \rangle$, dvs $\langle L(\mathbf{x}) | L(\mathbf{y}) \rangle = \langle \mathbf{x} | \mathbf{y} \rangle$, $\forall \mathbf{x}, \mathbf{y} \in V$.
- (iii) $L^\dagger \circ L = I$

Om $\dim V < \infty$ är dessutom följande ekvivalent med ovanstående

- (iv) $L^\dagger = L^{-1}$
- (v) $L \circ L^\dagger = I$
- (vi) L avbildar ortonormala baser på ortonormala baser.

Vi kallar en operator L som bevarar normen för en *isometri*.

Unitära operatorer och unitära matriser.

Definition 9.10. En operator L ett komplext inre produktrum V är en *unitär operator* om

- (1) L är inverterbar; och
- (2) $L^\dagger = L^{-1}$, dvs $\langle \mathbf{x} | L(\mathbf{y}) \rangle = \langle L^{-1}(\mathbf{x}) | \mathbf{y} \rangle$ för alla $\mathbf{x}, \mathbf{y} \in V$.

En *unitär matris* är en matris som motsvarar en unitär operator med avseende på en ortonormal bas.

Anmärkning 9.11. När L är inverterbar så är det andra villkoret liksom tidigare ekvivalent med att L är en isometri:

$$\langle L(\mathbf{x}) | L(\mathbf{y}) \rangle = \langle \mathbf{x} | \mathbf{y} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in V.$$

Om V är ändligtdimensionellt så medför isometri inverterbar:

Sats 9.12 (Sadun, Thm. 7.12). *Följande är ekvivalent för en komplex kvadratisk $(n \times n)$ -matris A :*

- (i) A är unitär
- (ii) $AA^\dagger = I$

- (iii) $A^\dagger A = I$
- (iv) Kolonnerna i A utgör en ortonormal bas för kolonnrummet.
- (v) Raderna i A utgör en ortonormal bas för radrummet.

Exempel 9.13. Låt $V = L^2(\mathbb{R}, \mathbb{C})$ vara Hilbertrummet av kvadratintegrerbara funktioner $f: \mathbb{R} \rightarrow \mathbb{C}$. För $a \in \mathbb{R}$, betrakta operatoren

$$L_a: V \longrightarrow V, \quad L_a(f)(x) = f(x+a)$$

Då är L_a en isometri ty

$$|L_a(f)|^2 = \int_{-\infty}^{\infty} |f(x+a)|^2 dx = \int_{-\infty}^{\infty} |f(x)|^2 dx = |f|^2$$

Vi har också att $(L_a)^{-1} = L_{-a}$: operatoren L_a förskjuter funktionsgrafens a åt vänster och L_{-a} förskjuter funktionsgrafens a åt höger. Alltså är L_a inverterbar och därmed unitär. Speciellt är den adjungerade operatoren $(L_a)^\dagger = L_{-a}$ vilket också går att verifiera direkt.

Eigenvärden och egenvektorer till ortogonala och unitära operatorer.

Sats 9.14 (Sadun, Thm. 7.13). Om L är unitär (eller bara isometrisk) så har alla eigenvärden absolutbelopp 1.

Bevis. Låt λ vara ett eigenvärde och välj en egenvektor $\xi \neq 0$ med $L(\xi) = \lambda\xi$. Vi har att

$$|\xi| = |L(\xi)| = |\lambda\xi| = |\lambda| \cdot |\xi| \implies |\lambda| = 1$$

eftersom L bevarar norm och $\xi \neq 0$. □

Sats 9.15 (Sadun, Thm. 7.13). Om L är unitär (eller bara isometrisk) och ξ_1 och ξ_2 egenvektorer med distinkta eigenvärden $\lambda_1 \neq \lambda_2$ så är $\langle \xi_1 | \xi_2 \rangle = 0$.

Bevis. Eftersom L bevarar den inre produkten så är:

$$\langle \xi_1 | \xi_2 \rangle = \langle L(\xi_1) | L(\xi_2) \rangle = \langle \lambda_1 \xi_1 | \lambda_2 \xi_2 \rangle = \overline{\lambda_1} \lambda_2 \langle \xi_1 | \xi_2 \rangle.$$

Vi har att $\overline{\lambda_1} = \lambda_1^{-1}$ eftersom $|\lambda_1| = 1$. Eftersom $\lambda_1 \neq \lambda_2$ så är $\lambda_1^{-1} \lambda_2 \neq 1$ och alltså är $\langle \xi_1 | \xi_2 \rangle = 0$. □

Sats 9.16 (Sadun, Thm. 7.13). Om L är unitär och $\dim V < \infty$ finns en ortogonal bas av egenvektorer till L .

Bevis. Det finns minst ett eigenvärde λ . Välj en egenvektor $\xi \neq 0$ med $L(\xi) = \lambda\xi$. Låt $W = \text{Span}\{\xi\}$. Vi har att $L(W^\perp) \subseteq W^\perp$ eftersom L bevarar den inre produkten och $|\lambda| = 1$:

$$\mathbf{x} \in W^\perp \iff 0 = \langle \xi | \mathbf{x} \rangle = \langle L(\xi) | L(\mathbf{x}) \rangle = \langle \lambda\xi | L(\mathbf{x}) \rangle = \overline{\lambda} \langle \xi | L(\mathbf{x}) \rangle \iff L(\mathbf{x}) \in W^\perp.$$

Alltså kan vi per induktion anta att det finns en ortogonal bas $\mathcal{B}'$ för $W^\perp$ av egenvektorer till L . Därmed är $\mathcal{B} = \{\xi\} \cup \mathcal{B}'$ en ortogonal bas för V av egenvektorer till L . Basfallet för induktionen är när $\dim V = 1$ då påståendet är trivialt eftersom L innebär multiplikation med ett komplext tal av belopp 1.

För ett mer explicit argument, kan vi välja en ortogonal bas $\mathcal{B} = \{\xi, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_n\}$ vars första basvektor är ξ . Då får matrisen utseendet

$$[L]_{\mathcal{B}} = (a_{ij}) = \begin{bmatrix} \lambda & * \\ 0 & B \end{bmatrix}$$

Eftersom $[L]_{\mathcal{B}}$ är unitär så utgör kolumnerna en ortonormal bas. Detta betyder att $\overline{\lambda} a_{1j} = 0$ för $j > 1$, dvs alla elementen markerade med $*$ är 0. Det räcker nu att diagonalisera B som är en operator på $\text{Span}\{\mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_n\}$ (vilket precis är $W^\perp$). □



**SF1681 LINJÄR ALGEBRA, FORTSÄTTNINGSKURS
FÖRELÄSNING 10 – HT24**

MATS BOIJ

10. OPERATORER PÅ INREPRODUKTRUM. REPETITION OM INREPRODUKTRUM

Målet för idag.

- Exponentialavbildningen och unitära operatorer.
- Normala operatorer och ortogonalt diagonaliserbara operatorer.
- Symmetriska operatorer och kvadratiske former.
- Repetition med tentaexempel (ej i TeX-anteckningarna).

Exponentialavbildningen och unitära operatorer.

Sats 10.1 (Sadun, Thm. 7.16). *Om L är en självadjungerad (Hermiteisk) operator på ett ändlig-dimensionellt inre produktrum V är $\exp(iL)$ unitär.*

Bevis. Välj en ortonormal bas av egenvektorer $\mathcal{B}$. Då är matrisen för L diagonal $D = [L]_{\mathcal{B}} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ och egenvärdena är reella. Alltså är

$$[\exp(iL)]_{\mathcal{B}} = \text{diag}(e^{i\lambda_1}, \dots, e^{i\lambda_n})$$

vilket är en unitär matris eftersom $|e^{i\lambda_j}| = 1$. (Att $\lambda \in \mathbb{R}$ är ekvivalent med att $|\exp(i\lambda)| = 1$.) Explicit är $\overline{e^{i\lambda_j}} = e^{-i\lambda_j} = (e^{i\lambda_j})^{-1}$ och alltså $\exp(iL)^{\dagger} = \exp(iL)^{-1}$. □

Korollarium 10.2 (Sadun, Cor. 7.18). *Om A är en anti-symmetrisk reell operator så är $\exp(A)$ ortogonal.*

Bevis. $B = -iA$ är Hermiteisk så $\exp(A) = \exp(iB)$ är unitär och reell, alltså ortogonal. □

Ortogonalt diagonaliserbara operatorer (normala operatorer).

Sats 10.3 (Sadun, Thm. 7.7, Thm. 7.13). *Låt L vara en operator på ett vektorrum V som har en ortogonal bas av egenvektorer till L .*

- (a) L är självadjungerad $\iff L$ har reella egenvärden.
- (b) L är unitär $\iff L$ har egenvärden av belopp 1.

Bevis. Välj en ortonormal bas $\mathcal{B}$ av egenvektorer till L . Då är matrisen $D = [L]_{\mathcal{B}}$ diagonal med egenvärdena på diagonalen: $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Eftersom basen är ortonormal så är L självadjungerad om och endast om $D^{\dagger} = D$ och L är unitär om och endast om $D^{\dagger} = D^{-1}$. Satsen följer eftersom

$$A^{\dagger} = \text{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$$

$$A^{-1} = \text{diag}(\lambda_1^{-1}, \dots, \lambda_n^{-1})$$

och $\overline{\lambda_i} = \lambda_i$ om och endast om $\lambda_i \in \mathbb{R}$ och $\overline{\lambda_i} = \lambda_i^{-1}$ om och endast om $|\lambda_i| = 1$. □

Definition 10.4. En operator L är *normal* om $LL^\dagger = L^\dagger L$.

Exempel 10.5. Vi har sett två klasser av normala operatorer

- (1) En självadjungerad operator ($L^\dagger = L$) är normal: $L^2 = L^2$.
- (2) En unitär operator ($L^\dagger = L^{-1}$) är normal: $I = I$.

Det finns även normala operatorer som varken är självadjungerade eller unitära.

Exempel 10.6. Om L är självadjungerad (Hermiteisk) så är iL anti-Hermiteisk: $(iL)^\dagger = -iL$. Anti-hermiteska operatorer är också normala. Om L är unitär så är aL normal för alla $a \in \mathbb{C}$: $(aL)(aL)^\dagger = a\bar{a} = |a|^2 = (aL)^\dagger(aL)$.

Sats 10.7 (Sadun, exploration p. 215). En operator L på ett komplext ändligt-dimensionellt inre produktrum V är normal om och endast om det finns en ortogonal bas bestående av egenvektorer till L .

Bevis. Om det finns en ortonormal bas $\mathcal{B}$ bestående av egenvektorer så får vi att

$$A = [L]_{\mathcal{B}} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Eftersom basen är ortonormal så är $A^\dagger = [L]_{\mathcal{B}}^\dagger$ och alltså

$$A^\dagger = [L^\dagger]_{\mathcal{B}} = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n)$$

och

$$A^\dagger A = \text{diag}(|\lambda_1|^2, \dots, |\lambda_n|^2) = AA^\dagger$$

Alltså är L normal.

Omvänt, om L är normal, välj ett egenvärde λ och motsvarande egenvektor ξ . Välj en ortonormal bas $\{\xi, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ vars första basvektor är ξ . Då får matrisen utseendet

$$A = [L]_{\mathcal{B}} = (a_{ij}) = \begin{bmatrix} \lambda & * \\ 0 & B \end{bmatrix}$$

Eftersom basen är ortonormal så är $A^\dagger = [L]_{\mathcal{B}}^\dagger$ och eftersom L är normal så är $A^\dagger A = AA^\dagger$. Tittar vi på elementet i position $(1, 1)$ av dessa två $n \times n$ -matriser så får vi att normen av första kolumnen är lika med normen av första raden:

$$|\lambda|^2 = |\lambda|^2 + |a_{12}|^2 + |a_{13}|^2 + \dots + |a_{1n}|^2.$$

Alltså är $a_{1j} = 0$ för $j > 1$, dvs alla elementen markerade med $*$ är noll. Sedan argumenterar vi med induktion på n liksom i bevisen för självadjungerade och unitära operatorer. $\square$

Symmetriska operatorer.

Definition 10.8. En självadjungerad (Hermiteisk) operator på ett reellt inre produktrum kallas för en *symmetrisk* operator.

Anmärkning 10.9. Att L är symmetrisk betyder alltså att om vi väljer en ortonormal bas $\mathcal{B}$ så är $[L]_{\mathcal{B}}^T = [L]_{\mathcal{B}}$, dvs matrisen $[L]_{\mathcal{B}}$ är symmetrisk.

Exempel 10.10. En kvadratisk form f_A på $\mathbb{R}^n$ (se nästa avsnitt) kan representeras av en symmetrisk $(n \times n)$ -matris A som ger en symmetrisk operator på $\mathbb{R}^n$.

Sats 10.11 (Sadun, Thm. 7.8). För en symmetrisk operator S på ett ändligdimensionellt inre produktrum V över $\mathbb{R}$ finns en ortogonal bas av egenvektorer.

Bevis. Välj en godtycklig ortonormal bas $\mathcal{B}$. Låt $A = [S]_{\mathcal{B}}$ vara matrisen för S . Då är A reell och symmetrisk. Eftersom A är Hermiteisk, så är egenvärdena reella (Sats 8.16) och det finns en ortogonal bas av komplexa egenvektorer (Sats 8.19).

Egenrummet $E_\lambda = \ker(\lambda I - A)$ kan bestämmas genom Gauss-elimination och dimensionen (dvs den geometriska multipliciteten) ges av antalet ledande ettor i den rad-reducerade matrisen. Dimensionen för egenrummet E_λ beror alltså inte på om vi betraktar matrisen som en reell eller komplex matris. Alltså är den geometriska multipliciteten lika med den algebraiska multipliciteten för alla egenvärden och vi har en bas av reella egenvektorer. Med Gram-Schmidt kan vi

konstruera en ortogonal bas för varje egenrum och detta ger en ortogonal bas för hela V eftersom egenrummen till olika egenvärden är ortogonala för Hermiteska operatorer (Sats 8.17). $\square$

Korollarium 10.12. En ortogonal matris A är diagonaliserbar om och endast om den är symmetrisk, eller ekvivalent, endast har egenvärdena ± 1 .

Bevis. Som komplex matris är A unitär och därmed diagonaliserbar med egenvärden av absolutbelopp 1. För att den ska vara diagonaliserbar som reell matris måste alla egenvärden vara reella (dvs ± 1). Då är A även Hermitesk och alltså symmetrisk eftersom A är reell. Vi har sett att symmetriska reella matriser alltid är diagonaliserbara. $\square$

De allra flesta ortogonala matriser är alltså inte diagonaliserbara.

Exempel 10.13. Låt $\theta \in [0, 2\pi)$ och låt $L: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ vara operatoren som roterar θ radianer motsols. I standardbasen svarar då L mot den ortogonala matrisen

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

vilken har de komplexa egenvärdena $\cos \theta \pm i \sin \theta$ med belopp 1. Dessa egenvärden är reella precis när

- $\theta = 0$: då är $A = I$ och egenvärdena är 1, eller
- $\theta = \pi$: då är $A = -I$ och egenvärdena är -1 .

För alla andra vinklar θ är L ej diagonaliserbar över de reella talen.

Kvadratiske former.

Definition 10.14. En *kvadratisk form* på k^n är en funktion $q: k^n \rightarrow k$ på formen

$$q(\mathbf{x}) = f_A(x_1, x_2, \dots, x_n) := \sum_{i,j} A_{ij} x_i x_j = \mathbf{x}^T A \mathbf{x}$$

där A är en $n \times n$ -matris.

Anmärkning 10.15. En kvadratisk form är ett *homogent* polynom av grad två, dvs ett polynom där alla termer har grad två. Såvida inte $2 = 0$ i k , t ex om $k = \mathbb{F}_2$, så kan vi ordna så att A är symmetrisk genom att ersätta A med $\frac{1}{2}(A + A^T)$.

Exempel 10.16. Två kvadratiske former och motsvarande symmetriska matriser

$$\begin{aligned} q(x, y) &= x^2 + y^2, & A &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ q(x, y, z) &= x^2 + xy + y^2 - 3xz - z^2, & A &= \begin{bmatrix} 1 & \frac{1}{2} & -\frac{3}{2} \\ \frac{1}{2} & 1 & 0 \\ -\frac{3}{2} & 0 & -1 \end{bmatrix} \end{aligned}$$

Exempel 10.17. En *bilinjär form* på k^n , dvs en bilinjär avbildning $\varphi: k^n \times k^n \rightarrow k$ ger en kvadratisk form genom $q(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x})$. Notera att $\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ för någon matris A och att $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$. Omvänt, om $2 \neq 0$ i k så ger en kvadratisk form q upphov till den symmetriska bilinjära formen $\varphi(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(q(\mathbf{x} + \mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y}))$. Om $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ för en symmetrisk matris A , så är $\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$. När $2 \neq 0$ är därför följande tre begrepp ekvivalenta

- (1) symmetriska $n \times n$ -matriser A ,
- (2) kvadratiske former $q(\mathbf{x})$ på k^n , och
- (3) symmetriska bilinjära former $\varphi(\mathbf{x}, \mathbf{y})$ på k^n .

Sats 10.18. Om f_A är en kvadratisk form på $\mathbb{R}^n$ antar den ett största värde λ på enhetssfären $|\mathbf{x}| = 1$. Det finns en egenvektor ξ till $A + A^T$ med egenvärde 2λ . Speciellt, om A är symmetrisk, så är ξ en egenvektor till A med egenvärde λ . Det största värdet λ hos f_A på enhetssfären inträffar i $\frac{1}{|\xi|}\xi$.

Bevis. Enhetssfären är kompakt (sluten och begränsad) och därför har f_A ett största värde på enhetssfären. Sfären ges av ekvationen $x_1^2 + x_2^2 + \dots + x_n^2 = 1$ och enligt Lagranges metod är $f_A(\mathbf{x})$ maximal då

$$\nabla f_A(\mathbf{x}) = \lambda \nabla \left(\sum x_i^2 \right) \iff (A + A^T)\mathbf{x} = 2\lambda \mathbf{x}. \quad \square$$

BILAGA A. EXTRAMATERIAL: MATRISGRUPPER

Matrisgrupper.

Definition A.1. En *grupp* är en mängd med en binär operation $*$: $G \times G \longrightarrow G$ och ett speciellt element e som uppfyller

- (i) $(a * b) * c = a * (b * c)$, $\forall a, b, c \in G$. (*associativitet*)
- (ii) $a * e = e * a = a$, $\forall a \in G$. (*identitet*)
- (iii) $\forall a \in G, \exists b \in G$ så att $a * b = b * a = e$. (*invers*)

Till exempel är varje kropp och även heltalen en grupp under addition. Matriser är också en grupp under addition. Alla dessa exempel är *abelska grupper*: additionen är kommutativ $a + b = b + a$. Vi ska se att matriser *under multiplikation* utgör en icke-abelsk grupp.

Definition A.2. Låt k vara en kropp. Den *allmänna linjära gruppen* (*en: general linear group*) är mängden

$$\mathrm{GL}_n(k) = \{A \in M_n(\mathbb{R}) : A \text{ är inverterbar}\}$$

med matrismultiplikation som operation. Om V är ett vektorrum över k , är på samma sätt den allmänna linjära gruppen på V mängden

$$\mathrm{GL}(V) = \{L: V \longrightarrow V : L \text{ är inverterbar}\}$$

med sammansättning av linjära avbildningar som operation.

Anmärkning A.3. $\mathrm{GL}_n(k)$ är detsamma som $\mathrm{GL}(k^n)$.

Sats A.4. $\mathrm{GL}_n(k)$ och $\mathrm{GL}(V)$ är grupper.

Bevis. Det räcker att visa att $\mathrm{GL}(V)$ är en grupp.

- (1) Sammansättning av funktioner är associativ så första axiomat håller.
- (2) Identitetsavbildningen är en linjär avbildning så vi har $I \in \mathrm{GL}(V)$. Alltså $e = I$ i andra axiomat.
- (3) Per definition så har varje $L \in \mathrm{GL}(V)$ en invers funktion L^{-1} . Inversen av en linjär avbildning är också linjär så $L^{-1} \in \mathrm{GL}(V)$ och vi har att $L \circ L^{-1} = L^{-1} \circ L = I$. Om $a = L$ i tredje axiomat är alltså $b = L^{-1}$. □

Ortogonala och unitära grupper.

Definition A.5. Vi har följande delmängder av $\mathrm{GL}_n(\mathbb{R})$:

- Ortogonal matriser: $O_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A \text{ är ortogonal}\}$
- Speciella ortogonal matriser: $SO_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A \text{ är ortogonal och } \det(A) = 1\}$.

Vi har följande delmängder av $\mathrm{GL}_n(\mathbb{C})$:

- Unitära matriser: $U_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) : A \text{ är unitär}\}$
- Speciella unitära matriser: $SU_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) : A \text{ är unitär och } \det(A) = 1\}$

Observera att ortogonal och unitära matriser alltid är inverterbara: per definition är ju $A^T = A^{-1}$ resp. $A^\dagger = A^{-1}$.)

På samma sätt, om V är ett inre produktrum över $\mathbb{R}$ definierar vi $O(V)$ och $SO(V)$ och om V är ett inre produktrum över $\mathbb{C}$ definierar vi $U(V)$ och $SU(V)$. Observera att $SO(V)$ och $SU(V)$ bara går att definiera för ändligt-dimensionella inre produktrum och att determinanten inte beror på valet av bas.

Sats A.6. $O_n(\mathbb{R})$, $SO_n(\mathbb{R})$, $U_n(\mathbb{C})$, $SU_n(\mathbb{C})$ är grupper under matrismultiplikation. Om V är ett inre produktrum över $\mathbb{R}$ så är $O(V)$ och $SO(V)$ grupper under sammansättning. Om V är ett inre produktrum över $\mathbb{C}$ så är $U(V)$ och $SU(V)$ grupper under sammansättning.

Bevis. Det räcker att visa att matrismultiplikationen är definierad på dessa delmängder av $GL_n(k)$. Om A och B är ortogonala: $A^T = A^{-1}$ och $B^T = B^{-1}$, så är $(AB)^T = B^T A^T = B^{-1} A^{-1} = (AB)^{-1}$ och alltså är även AB ortogonal. För ortogonala operatorer L_1 och L_2 ser vi att $L_1 \circ L_2$ är ortogonal på samma sätt genom att använda att $(L_1 \circ L_2)^\dagger = L_2^\dagger \circ L_1^\dagger$ och $(L_1 \circ L_2)^{-1} = L_2^{-1} \circ L_1^{-1}$. $\square$

Anmärkning A.7. $O(V)$ kallas den *ortogonala gruppen* och $SO(V)$ kallas den *speciella ortogonala gruppen*. $SO_3(\mathbb{R})$ utgör gruppen av stelkroppsrotationer i rummet. $U(V)$ kallas den *unitära gruppen* och $SU(V)$ kallas den *speciella unitära gruppen*.



MATERIAL OM SINGULÄRVÄRDESUPPDELNING OCH PSEUDOINVERSER

1. SINGULÄRVÄRDESUPPDELNING

1.1. Normalform för linjär avbildning. När vi arbetar med egenvärden, egenvektorer och diagonalisering handlar det om att välja en bas så att matrisen för en operator L blir på så enkel form som möjligt. Om vi ser på linjära avbildningar $L: V \longrightarrow W$ har vi större frihet i och med att vi kan göra valen av bas för V och W oberoende av varandra. Det visar sig då att det alltid går att få matrisen på en mycket enkel form med bara ettor och nollor.

Sats 1.1. Om $L: V \longrightarrow W$ är en linjär avbildning mellan ändligdimensionella vektorrum kan vi välja baser för V och W så att matrisen för L blir en blockmatris på formen

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

där I är en identitetsmatris av storlek $r \times r$ där $r = \text{rk } L$.

Bevis. Låt $V'' = \ker L$ och välj ett komplement V' till V'' , dvs $V = V' \oplus V''$. Detta kan (till skillnad från ett ortogonalt komplement) göras på många olika sätt. Konkret kan vi börja med en bas $\mathbf{x}_{r+1}, \mathbf{x}_{r+2}, \dots, \mathbf{x}_n$ för V'' , sedan utvidga till en bas $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ för hela V och slutligen ta $V' = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r\}$.

Låt $W' = \text{im } L$ och välj ett komplement W'' till W' , dvs $W = W' \oplus W''$. Återigen kan vi välja ett komplement på många olika vis.

På grund av definitionen av V' är restriktionen av L till V' en isomorfi mellan V' och W' . Vi skriver $L' = L|_{V'}: V' \longrightarrow W'$ där alltså $L'(\mathbf{v}) = L(\mathbf{v})$ för alla $\mathbf{v} \in V'$.

Om vi nu låter $\mathbf{y}_i = L(\mathbf{x}_i)$ för $i = 1, 2, \dots, r$ så får vi en bas för W' sådan att matrisen för L' blir identitetsmatrisen. Slutligen utökar vi som tidigare till en bas $\mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ för hela W där alltså $W'' = \text{Span}\{\mathbf{y}_{r+1}, \dots, \mathbf{y}_m\}$.

Matrisen för L blir då på den önskade formen med avseende på baserna $\mathcal{B}$ och $\mathcal{B}'$. □

Notera 1.2. Satsen kan omformuleras som: det finns baser $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ och $\mathcal{B}' = \{\mathbf{y}_1, \dots, \mathbf{y}_m\}$ för V och W så att

$$L(\mathbf{x}_i) = \begin{cases} \mathbf{y}_i & \text{om } 1 \leq i \leq r \\ \mathbf{0} & \text{om } r+1 \leq i \leq n \end{cases}$$

1.2. Adjungerad avbildning. Tidigare har vi sett på adjungerade *operatorer*. Vi kan göra motsvarande även för linjära avbildningar mellan olika vektorrum. För en linjär avbildning $L: V \rightarrow W$ kan vi definiera en adjungerad avbildning $L^\dagger: W \rightarrow V$. Likasom tidigare existerar $L^\dagger$ inte alltid i det oändligt-dimensionella fallet.

Definition 1.3. Om $L: V \rightarrow W$ är en linjär avbildning mellan inre produktrum ges den *adjungerade avbildningen*, $L^\dagger: W \rightarrow V$ av

$$\langle L^\dagger(\mathbf{y}) | \mathbf{x} \rangle = \langle \mathbf{y} | L(\mathbf{x}) \rangle, \quad \forall \mathbf{x} \in V, \forall \mathbf{y} \in W.$$

Notera 1.4. Av definitionen följer att:

$$\ker L^\dagger = (\operatorname{im} L)^\perp \quad \text{och} \quad \ker L = (\operatorname{im} L^\dagger)^\perp$$

Sats 1.5. Om vi väljer *ortonormala baser* för V och W och matrisen för L är A ges matrisen för $L^\dagger$ av $A^\dagger$.

Bevis. Om $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ och $\mathcal{B}' = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ är ON-baser för V och W så är alltså $A = [L]_{\mathcal{B}', \mathcal{B}}$ och

$$A_{ij} = \langle \mathbf{f}_i | L(\mathbf{e}_j) \rangle = \langle L^\dagger(\mathbf{f}_i) | \mathbf{e}_j \rangle = \overline{\langle \mathbf{e}_j | L^\dagger(\mathbf{f}_i) \rangle} = \overline{B_{ji}}$$

där B är matrisen för $L^\dagger$, dvs $B = [L^\dagger]_{\mathcal{B}, \mathcal{B}'}$. □

1.3. Singulärvärdesuppdelning. Vi såg tidigare att vi kunde välja baserna för V och W oberoende av varandra så att matrisen för en linjär avbildning $L: V \rightarrow W$ blir på en mycket enkel form. Om vi har en avbildning mellan inre-produktrum är det naturligt att bara tillåta oss att välja ortonormala baser för V och W och då har vi inte lika stor frihet. Det visar sig dock att vi kan komma ganska nära.

Sats 1.6. Om $L: V \rightarrow W$ är en linjär avbildning mellan ändligdimensionella inre produktrum kan vi välja *ortonormala baser* för V och W så att matrisen för L med avseende på dessa baser blir en blockmatris Σ på formen

$$\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

där D är en positiv reell diagonalmatris.

Notera 1.7. Precis som för diagonalisering är baserna inte unika, men de positiva diagonalelementen, *singulärvärdena*, är unikt bestämda (inklusive deras multiplicitet). Oftast brukar man ordna singulärvärdena i fallande ordning och då blir Σ unikt bestämt. Ibland räknas även nollor på diagonalen av Σ , utanför D , som singulärvärden.

Bevis. Vi skriver $V = V' \oplus V''$ och $W = W' \oplus W''$ där

- $V' = (\ker L)^\perp$
- $V'' = \ker L$
- $W' = \operatorname{im} L$
- $W'' = (\operatorname{im} L)^\perp$

Observera att det *ortogonal* komplementet (som är meningsfullt eftersom vi har inre produktrum) är unikt till skillnad från de komplement vi valde i beviset till Sats 1.1.

Som i det beviset får vi en isomorfi $L' = L|_{V'}: V' \xrightarrow{\cong} W'$.

Om vi väljer godtyckliga ON-baser för V'' och W'' så återstår det att välja ON-baser för V' och W' så att $[L'] = D$. Observera att vi inte kan välja en ON-bas $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r$ för V' och sedan sätta $\mathbf{y}_i = L(\mathbf{x}_i)$ för $i = 1, 2, \dots, r$. Detta skulle ge en bas för W' men inte nödvändigtvis en ON-bas (såvida inte L' är unitär). Genom att betrakta $L': V' \rightarrow W'$ istället för $L: V \rightarrow W$ har vi alltså reducerat till fallet då L är inverterbar.

Bilda nu operatoren $H: V \oplus W \longrightarrow V \oplus W$ genom $H(\mathbf{x}, \mathbf{y}) = (L^\dagger(\mathbf{y}), L(\mathbf{x}))$. Med avseende på ortonormala baser för V och W får H den Hermiteska matrisen¹

$$B = \begin{bmatrix} 0 & A^\dagger \\ A & 0 \end{bmatrix}$$

om A är matrisen för L . Alltså är H självadjungerad och eftersom H kommuterar med H^2 kan vi därför välja en gemensam ortogonal bas av egenvektorer. Dessa är på formen $\begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix}$ med

$$\begin{bmatrix} 0 & A^\dagger \\ A & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} = \lambda_i \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} \quad \text{och} \quad \begin{bmatrix} A^\dagger A & 0 \\ 0 & AA^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} = \mu_i \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix}$$

det vill säga

$$\begin{cases} A^\dagger \mathbf{y}_i = \lambda_i \mathbf{x}_i \\ A \mathbf{x}_i = \lambda_i \mathbf{y}_i \end{cases} \quad \text{och} \quad \begin{cases} A^\dagger A \mathbf{x}_i = \mu_i \mathbf{x}_i \\ AA^\dagger \mathbf{y}_i = \mu_i \mathbf{y}_i \end{cases}$$

och därmed är $\mu_i = \lambda_i^2$.

Eftersom H är självadjungerad och inverterbar är λ_i (och μ_i) reella och noll-skilda. Det visar sig att hälften av de $2n$ egenvärdena är positiva och hälften är negativa: givet en egenvektor med egenvärde λ_i som ovan så är $\begin{bmatrix} \mathbf{x}_i \\ -\mathbf{y}_i \end{bmatrix}$ en egenvektor med egenvärde $-\lambda_i$ eftersom

$$\begin{bmatrix} 0 & A^\dagger \\ A & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_i \\ -\mathbf{y}_i \end{bmatrix} = \begin{bmatrix} -A^\dagger \mathbf{y}_i \\ A \mathbf{x}_i \end{bmatrix} = \begin{bmatrix} -\lambda_i \mathbf{x}_i \\ \lambda_i \mathbf{y}_i \end{bmatrix} = -\lambda_i \begin{bmatrix} \mathbf{x}_i \\ -\mathbf{y}_i \end{bmatrix}.$$

Vi kan nu anta att egenvärdena är ordnade så att de positiva egenvärdena kommer först och då kommer $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ och $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ utgöra ortogonala baser för V respektive W . Efter att normerat dessa har vi ortonormala baser och matrisen för L med avseende på dessa baser är en diagonalmatris med positiva diagonalelement $\lambda_1, \lambda_2, \dots, \lambda_n$. $\square$

Definition 1.8. En *singulärvärdesuppdelning* av den linjära avbildningen $L: V \longrightarrow W$ mellan två inreproduktrum V och W är val av ortonormala baser $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ och $\mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ för V och W så att

$$[L]_{\mathcal{B}', \mathcal{B}} = \Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

där $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ är en positiv reell diagonalmatris.

Att en singulärvärdesuppdelning existerar är Sats 1.6.

Notera 1.9. I en singulärvärdesuppdelning har vi alltså att

$$L(\mathbf{x}_i) = \begin{cases} \sigma_i \mathbf{y}_i & \text{om } 1 \leq i \leq r \\ \mathbf{0} & \text{om } r+1 \leq i \leq n \end{cases}$$

Eftersom $\{\mathbf{x}_i\}$ och $\{\mathbf{y}_i\}$ är ortonormala baser gäller även att

$$L^\dagger(\mathbf{y}_i) = \begin{cases} \sigma_i \mathbf{x}_i & \text{om } 1 \leq i \leq r \\ \mathbf{0} & \text{om } r+1 \leq i \leq m \end{cases}$$

¹Detta svarar mot att ge $V \oplus W$ den inre produkten $\langle (\mathbf{x}_1, \mathbf{y}_1) | (\mathbf{x}_2, \mathbf{y}_2) \rangle = \langle \mathbf{x}_1 | \mathbf{x}_2 \rangle + \langle \mathbf{y}_1 | \mathbf{y}_2 \rangle$.

Definition 1.10. Ett *singulärvärde* till $L: V \rightarrow W$ är ett positivt reellt tal $\sigma > 0$ så att det finns nollskilda vektorer $\mathbf{x} \in V$ och $\mathbf{y} \in W$ med

$$L(\mathbf{x}) = \sigma \mathbf{y} \quad \text{och} \quad L^\dagger(\mathbf{y}) = \sigma \mathbf{x}.$$

Vektorn $\mathbf{x}$ kallas en *högersingulär* vektor och $\mathbf{y}$ en *vänstersingulär* vektor.

Notera 1.11. Notera att varje högersingulär vektor $\mathbf{x}$ med singulärvärde σ har en *motsvarande* vänstersingulär vektor $\mathbf{y} = \frac{L(\mathbf{x})}{\sigma}$ och omvänt $\mathbf{x} = \frac{L^\dagger(\mathbf{y})}{\sigma}$.

Notera 1.12. Om $\mathbf{x}$ och $\mathbf{y}$ är höger- och vänstersingulära vektorer till σ så är $|\mathbf{x}| = |\mathbf{y}|$ ty $\sigma^2 |\mathbf{y}|^2 = \langle L\mathbf{x} | L\mathbf{x} \rangle = \langle \sigma^2 \mathbf{x} | \mathbf{x} \rangle = \sigma^2 |\mathbf{x}|^2$. Vi kan alltså skala om $\mathbf{x}$ och $\mathbf{y}$ så att de har norm 1.

Notera 1.13 (Singulära vektorer till singulärvärdet 0). Även 0 betraktas ibland som ett singulärvärde och vi kallar även nollskilda vektorer $\mathbf{x} \in \ker L$ för högersingulära och nollskilda vektorer $\mathbf{y} \in \ker L^\dagger$ för vänstersingulära. Observera dock att dessa inte svarar mot varandra på samma sätt som för positiva singulärvärden.

Notera 1.14. I en singulärvärdesuppdelning för $L: V \rightarrow W$ är alltså $\mathcal{B}$ en ortonormal bas av högersingulära vektorer $\mathbf{x}_i$ och $\mathcal{B}'$ är en ortonormal bas av vänstersingulära vektorer $\mathbf{y}_j$ och $D = \text{diag}(\sigma_1, \dots, \sigma_r)$ är de nollskilda singulärvärdena. Dessutom korresponderar de första r vänster- och högersingulära vektorerna mot varandra.

I termer av matriser har vi:

Definition 1.15 (Singulärvärdesuppdelning för matriser). Om $L: \mathbb{C}^n \rightarrow \mathbb{C}^m$ är en linjär avbildning med motsvarande $m \times n$ -matris A så är en singulärvärdesuppdelning för A en faktorisering

$$A = Y \Sigma X^\dagger$$

där

- (a) kolonnerna i X är en ortonormal bas av högersingulära vektorer,
- (b) kolonnerna i Y är en ortonormal bas av vänstersingulära vektorerna,
- (c) Σ är en $m \times n$ -matris som är noll utanför diagonalen och har singulärvärdena i fallande storlek längs diagonalen.
- (d) de första $r = \text{rk } L$ kolonnerna i X och Y , dvs som har noll-skilda singulärvärden, korresponderar mot varandra.

Notera 1.16. Faktoriseringen

$$A = Y \Sigma X^\dagger$$

är ett ortonormalt basbyte:

$$[L]_{\mathcal{E}', \mathcal{E}} = P_{\mathcal{E}', \mathcal{B}'} [L]_{\mathcal{B}', \mathcal{B}} P_{\mathcal{B}, \mathcal{E}}$$

där $P_{\mathcal{E}', \mathcal{B}'} = Y$, $[L]_{\mathcal{B}', \mathcal{B}} = \Sigma$ och $P_{\mathcal{B}, \mathcal{E}} = X^{-1} = X^\dagger$ eftersom X är unitär.

Notera 1.17. Det ortogonala komplementet till kärnan, $\ker(L)^\perp$, har en ortonormal bas av högersingulära vektorer $\mathbf{x}_1, \dots, \mathbf{x}_r$ som svarar mot nollskilda singulärvärden och bildrummet, $\text{im}(L)$, har en ortonormal bas av vänstersingulära vektorer $\mathbf{y}_1, \dots, \mathbf{y}_r$ som svarar mot nollskilda singulärvärden.

Genom konstruktionen har vi att matriserna X och Y är *ortogonala* om vi arbetar över de reella talen och *unitära* om vi arbetar över de komplexa talen.

1.4. Att bestämma singularvärdesuppdelning. Låt A vara en komplex $m \times n$ -matris. Datorprogram som MatLab bestämmer singularvärdesuppdelning till A på mer effektiva sätt men vi kan bestämma den på följande vis:

- (1) Bestäm egenvärdena $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$ till den Hermiteska $n \times n$ -matrisen $A^\dagger A$. De första $r = \text{rk} A$ är de (noll-skilda) singularvärdena i kvadrat $\mu_i = \sigma_i^2$ och de sista $n - r = \dim(\ker A)$ är noll: $\mu_{r+1} = \dots = \mu_n = 0$.
- (2) Bestäm en ortonormal bas $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ av egenvektorer till $A^\dagger A$. Detta är en ortonormal bas av högersingulära vektorer, inklusive de med singularvärde 0 (dvs de i $\ker A$).
- (3) Låt $\mathbf{y}_i = \frac{1}{\sigma_i} A \mathbf{x}_i$ för $i = 1, 2, \dots, r$. Detta är en ortonormal uppsättning av vänstersingulära vektorer för noll-skilda singularvärden.
- (4) Bestäm en ortonormal bas $\mathbf{y}_{r+1}, \dots, \mathbf{y}_m$ för $\ker A^\dagger = (\text{im} A)^\perp$.
- (5) Då har vi ortonormala baser $\mathbf{x}_1, \dots, \mathbf{x}_n$ och $\mathbf{y}_1, \dots, \mathbf{y}_m$ som ger unitära matriser $X = [\mathbf{x}_1 \dots \mathbf{x}_n]$ och $Y = [\mathbf{y}_1 \dots \mathbf{y}_m]$ så att vi får singularvärdesuppdelningen

$$A = Y \Sigma X^\dagger$$

där Σ är en $m \times n$ -matris där alla element är noll förutom de r första diagonalelementen som är $\sigma_1 \geq \dots \geq \sigma_r$.

Vi kan förstås byta ut steg (1) mot att istället ta egenvärden och egenvektorer till $AA^\dagger$. Då får vi vänstersingulära vektorer och kan bestämma de högersingulära genom $\mathbf{x}_i = \frac{1}{\sigma_i} A^\dagger \mathbf{y}_i$ för $i = 1, \dots, r$ och sedan fylla på med en bas för $\ker A$.

Exempel 1.18 (Tentamen 2018-01-10 #3). Bestäm en singularvärdesuppdelning till matrisen

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$

(Se handskrivna anteckningar till föreläsning 11.)

1.5. Projektion. En av poängerna med singularvärdesuppdelning är att vi kan bortse helt från singularvärden som är noll och motsvarande höger- och vänstersingulära vektorer. Vi kan ta steget vidare och bortse från små singularvärden. Genom att sortera singularvärdena i fallande ordning $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ kan vi se att de första termerna i summan

$$A = Y \Sigma X^\dagger = \sum_{i=1}^r \sigma_i \mathbf{y}_i \mathbf{x}_i^\dagger$$

innehåller den *viktigaste* informationen om A . Vi kan approximera A med delsummorna

$$A_s = \sum_{i=1}^s \sigma_i \mathbf{y}_i \mathbf{x}_i^\dagger, \quad s = 1, 2, \dots, r.$$

Sats 1.19. A_s är den matris av rang s som bäst approximerar A i minsta kvadratmening.

Exempel 1.20. Om vi har matrisen

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 3 & 6 & 10 \end{bmatrix}$$

så får vi singularvärdena till kvadratrötterna till egenvärdena av

$$AA^T = \begin{bmatrix} 56 & 90 \\ 90 & 145 \end{bmatrix}$$

som är $\lambda_1 \approx 200,9$ och $\lambda_2 \approx 0,09956$. Därmed är $\sigma_1 \approx 14,17$ och $\sigma_2 \approx 0,316$. Om vi bara tar med det största singularvärdet får vi approximationen

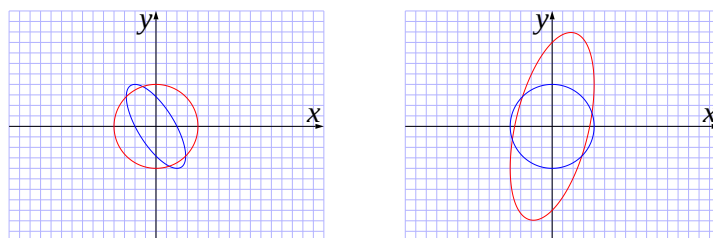
$$A_1 = Y \Sigma_1 X^T = \begin{bmatrix} 1,90139 & 3,80278 & 6,15238 \\ 3,06125 & 6,1225 & 9,90536 \end{bmatrix}$$

Om vi vill spara utrymme och singularvärdena avtar relativt snabbt kan vi använda en approximation av relativt låg rang. Detta har bland annat använts för bildkompression.

1.6. Geometrisk innebörd för operatorer på $\mathbb{R}^2$. Om vi har en inverterbar 2×2 -matris A kan vi se på hur enhetscirkeln avbildas genom multiplikation med A . Bilden kommer att bli en ellips och ekvationen för ellipsen ges av $\langle A^{-1}\mathbf{y} | A^{-1}\mathbf{y} \rangle = 1$. Vi kan också se på inversbilden av enhetscirkeln som också blir en ellips. Denna gång blir ekvationen $\langle A\mathbf{x} | A\mathbf{x} \rangle = 1$. Ellipsernas axlar ges av egenvektorer till matriserna $(A^{-1})^T A^{-1}$ respektive $A^T A$. Egenvektorena till $B = (A^{-1})^T A^{-1}$ är de samma som för dess invers, $B^{-1} = AA^T$.

Om vi till exempel tar matrisen $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ får vi bilderna enligt Figur 1 där den röda enhetscirkeln till vänster avbildas på den röda ellipsen till höger och den blå ellipsen till vänster avbildas på den blå enhetscirkeln till höger.

Det ortogonala basbytet som gör ellipsen till vänster axelparallell svarar mot valet av vänstersingulära vektorer och det ortogonala basbytet som gör ellipsen till höger axelparallell svarar mot de högersingulära vektorena.



FIGUR 1. En bild av hur singularvärdesuppdelningen ser ut för matrisen en 2×2 -matris.

2. PSEUDOINVERS

Det finns flera olika sätt att generalisera begreppet inversmatris till situationer där matriserna inte är inverterbara i vanlig mening. En sådan generalisering ges av [Moore–Penrose-inversen](#) [Dre20, Pen55].

Definition 2.1 (Moore–Penrose invers). Om A är en komplex $m \times n$ -matris ges [Moore–Penrose-inversen](#) A^+ till A av

$$A^+ = Q \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^\dagger, \quad \text{där } A = P \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} Q^\dagger$$

är singularvärdesuppdelningen av A .

Sats 2.2 (Penrose[Pen55]). A^+ är den unika lösningen till de fyra ekvationerna

$$AXA = A, \quad XAX = X, \quad (AX)^\dagger = AX \quad \text{och} \quad (XA)^\dagger = XA.$$

Bevis. Vi använder singularvärdesuppdelning av A och kan därmed anta att

$$A = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad \text{och} \quad X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}.$$

Då blir

$$AX = \begin{bmatrix} DX_{11} & DX_{12} \\ 0 & 0 \end{bmatrix} \quad \text{och} \quad XA = \begin{bmatrix} X_{11}D & 0 \\ X_{21}D & 0 \end{bmatrix}$$

och de två sista ekvationerna ger $DX_{12} = 0$, $X_{21}D = 0$ och $D_{11}X = XD_{11}$. Eftersom D är inverterbar är $X_{21} = 0$ och $X_{12} = 0$.

Nu blir

$$AXA = \begin{bmatrix} DX_{11}D & 0 \\ 0 & 0 \end{bmatrix} \quad \text{och} \quad XAX = \begin{bmatrix} X_{11}DX_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

och den första ekvationen ger $DX_{11}D = D$, vilket ger $X_{11} = D^{-1}$ och den andra ger $X_{22} = 0$. Den unika lösningen är därmed

$$X = \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

med avseende på de givna ortonormala baserna för V och W . $\square$

Notera 2.3. I och med att den vanliga inversen har dessa egenskaper sammanfaller Moore–Penrose-inversen med den vanliga inversen när den existerar.

Moore–Penrose-inversen används bland annat vid minsta-kvadratproblem. Minsta kvadratmetoden ger att vi ska välja en lösning till $A^\dagger A\mathbf{x} = A^\dagger \mathbf{b}$. Om $\ker(A) \neq \{0\}$ finns inte en unik lösning och vi kan bestämma oss för att välja den vektor i lösningsmängden som ligger närmast origo, dvs som har minst norm.

Vi har följande sats.

Sats 2.4. Till ekvationen $A\mathbf{x} = \mathbf{b}$ är $\mathbf{x} = A^+\mathbf{b}$ den minsta-kvadratlösning som har minst norm $\|\mathbf{x}\|$, dvs den vektor som minimerar $\|\mathbf{x}\|$ bland de vektorer där $\|A\mathbf{x} - \mathbf{b}\|$ är minimal.

Bevisidé. Vi vill hitta en minsta-kvadrat lösning till $AX = I$ som minimerar normen av kolonnerna i X . Vi börjar med att se att $X = A^+$ löser normalekvationen $A^\dagger AX = A^\dagger$ eftersom $AA^+ = (A^+)^\dagger A^\dagger$, vilket ger $A^\dagger (A^+)^\dagger A^\dagger = A^\dagger$ som är ekvivalent med $AA^+A = A$.

Om vi har en annan lösning kan den skrivas $A^+ + Y$ där $A^\dagger AY = 0$. Enligt konstruktionen är kolonnerna i A^+ ortogonala mot kärnan till $A^\dagger A$, vilket gör att normen av kolonnerna i $A^+ + Y$ minimeras så $Y = 0$. $\square$

Läs mer om singularvärdesuppdelning i exempelvis *Linear Algebra Done Right* [Ax15, §7.D] eller *Numerical Analysis* [Sau14]. Det finns många andra källor på nätet, exempelvis *Wikipedia* [Wik17] och dokumentationen för Matlab [Mat]. Det kan också vara intressant att se på Penroses artikel [Pen55] från 1955.

REFERENSER

- [Ax15] Sheldon Axler. *Linear algebra done right*. Undergraduate Texts in Mathematics. Springer, Cham, third edition, 2015.
- [Dre20] Arnold Dresden. The fourteenth western meeting of the american mathematical society. *Bull. Amer. Math. Soc.*, 26(9):385–396, 06 1920.
- [Mat] MathWorks. svd — documentation.
- [Pen55] R. Penrose. A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.*, 51:406–413, 1955.
- [Sau14] Timothy Sauer. *Numerical Analysis*. Pearson, 2014.
- [Wik17] Wikipedia. Singular-value decomposition — wikipedia, the free encyclopedia, 2017.

ÖVNINGAR PÅ SINGULÄRVÄRDESUPPDELNING

Uppgift 1. Låt L vara en operator på ett inre produktrum V .

- (a) Visa att $L^\dagger L$ och $LL^\dagger$ har samma egenvärden.
- (b) Visa att alla egenvärden till $L^\dagger L$ och $LL^\dagger$ är icke-negativa reella tal.
- (c) Visa att om V är ändligt-dimensionellt så finns det en självadjungerad operator K_1 så att $(K_1)^2 = L^\dagger L$ och en självadjungerad operator K_2 så att $(K_2)^2 = LL^\dagger$.
- (d) Om L i en ortonormal bas $\mathcal{B}$ ges av matrisen

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$$

bestäm matriserna för K_1 och K_2 i samma bas.

Uppgift 2. Bestäm en 2×3 -matris A så att A har singularvärdena 2 och 3 med motsvarande höger- och vänstersingulära vektorer $\frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix}^T$ och $\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \end{bmatrix}^T$, respektive $\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}^T$ och $\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \end{bmatrix}^T$.

Uppgift 3. Bestäm Moore–Penrose inversen A^+ till matrisen

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}.$$

Uppgift 4. Använd Moore–Penrose inversen för att lösa ekvationssystemet

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

i minsta kvadratmening.

Uppgift 5. Visa att om A är en reell 2×2 -matris kan vi alltid välja X eller Y som en symmetrisk matris i singularvärdesuppdelningen $A = Y\Sigma X^T$. När kan vi välja både X och Y som symmetriska matriser?

Uppgift 6 (Uppgift 3 på tentamen 2019-04-17). Bestäm det största singularvärdet och motsvarande höger- och vänstersingulära vektorer som hör till matrisen

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

Uppgift 7 (Uppgift 3 på tentamen 2018-04-05). Bestäm singularvärden och singularvektorer till matriserna

$$A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad \text{och} \quad B = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

LÖSNINGSFÖRSLAG

Lösningssförslag till uppgift 1.

- (a) Låt $\mathbf{x}$ vara en egenvektor till $L^\dagger L$ med egenvärde $\lambda \neq 0$. Då är $L\mathbf{x}$ en egenvektor till $LL^\dagger$ med egenvärde λ eftersom

$$LL^\dagger(L\mathbf{x}) = L((L^\dagger L)\mathbf{x}) = L(\lambda\mathbf{x}) = \lambda(L\mathbf{x}).$$

Notera att $L\mathbf{x} \neq 0$ eftersom $L^\dagger L\mathbf{x} \neq 0$.

Omvänt, om $\mathbf{y}$ är en egenvektor till $LL^\dagger$ med egenvärde $\lambda \neq 0$ så är $L^\dagger\mathbf{y}$ en egenvektor till $L^\dagger L$ med egenvärde λ eftersom

$$L^\dagger L(L^\dagger\mathbf{y}) = L^\dagger((LL^\dagger)\mathbf{y}) = L^\dagger(\lambda\mathbf{y}) = \lambda(L^\dagger\mathbf{y}).$$

Egenvärdet 0 förekommer hos L precis när $\text{rk} L \neq \dim V$ och analogt för $L^\dagger$. Eftersom $\text{rk} L = \text{rk} L^\dagger$ så förekommer alltså egenvärdet 0 antingen för både L och $L^\dagger$ eller för ingen av dem.

- (b) Låt $\mathbf{x}$ vara en egenvektor till $L^\dagger L$ med egenvärde λ . Då är

$$\lambda|\mathbf{x}| = \langle \mathbf{x} | \lambda\mathbf{x} \rangle = \langle \mathbf{x} | L^\dagger L\mathbf{x} \rangle = \langle L\mathbf{x} | L\mathbf{x} \rangle = |L\mathbf{x}| \geq 0.$$

Alltså är $\lambda \geq 0$.

- (c) Eftersom $L^\dagger L$ är självadjungerad så finns en ortonormal bas $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ av egenvektorer, dvs $[L^\dagger L]_{\mathcal{B}} = \text{diag}(\lambda_1, \dots, \lambda_n)$. Vi kan nu låta K_1 vara operatorn med matrisen $\text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ i basen $\mathcal{B}$, dvs $K_1(\mathbf{x}_i) = \sqrt{\lambda_i}\mathbf{x}_i$ för $i = 1, 2, \dots, n$. Eftersom $[K_1]_{\mathcal{B}}$ är reell och symmetrisk så är K_1 självadjungerad. Analogt för $LL^\dagger$.

Vi hade kunnat välja $-\sqrt{\lambda_i}$ istället för några eller alla i . Vårt val av K_1 (och K_2) har den ytterligare egenskapen att alla egenvärdena hos K_1 är icke-negativa, eller ekvivalent, att K_1 är *positivt semidefinit*, dvs $\langle K_1\mathbf{x} | \mathbf{x} \rangle \geq 0$ för alla $\mathbf{x} \in V$.

- (d) $L^\dagger L$ och $LL^\dagger$ ges av matriserna

$$A^\dagger A = \begin{bmatrix} 4 & 6 \\ 6 & 13 \end{bmatrix}, \quad \text{och} \quad AA^\dagger = \begin{bmatrix} 13 & 6 \\ 6 & 4 \end{bmatrix}.$$

Egenvärdena för båda operatorerna är 1 och 16. En egenvektorbas för $A^\dagger A$ är

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

och en egenvektorbas för $AA^\dagger$ är

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Operatorerna K_1 och K_2 har båda egenvärdena 1 och 4 och egenvektorbaserna $\{\mathbf{x}_1, \mathbf{x}_2\}$ respektive $\{\mathbf{y}_1, \mathbf{y}_2\}$. Detta ger

$$[K_1] = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}^{-1} = \frac{1}{5} \begin{bmatrix} 8 & 6 \\ 6 & 17 \end{bmatrix}$$

$$[K_2] = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}^{-1} = \frac{1}{5} \begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix}.$$

Lösningförslag till uppgift 2. För $\mathbb{R}^3$ har vi en ortonormal bas $\mathcal{B}$ bestående av de 2 höger-singulära vektorerna

$$\mathbf{x}_1 = \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{x}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

och en tredje godtycklig vektor $\mathbf{x}_3$ i det ortogonala komplementet till $\text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$.

För $\mathbb{R}^2$ har vi en ortogonal bas $\mathcal{B}'$ bestående av de 2 vänster-singulär vektorerna.

$$\mathbf{y}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{y}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

Med dessa baser har vi:

$$[A]_{\mathcal{B}', \mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

dvs

$$A = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & 0 & * \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & * \\ \frac{2}{3} & \frac{-1}{\sqrt{2}} & * \end{bmatrix}^{-1}$$

Eftersom den tredje matrisen är ortogonal så erhåller vi:

$$A = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ * & * & * \end{bmatrix} = \frac{1}{6\sqrt{5}} \begin{bmatrix} 4 & 8 + 18\sqrt{2} & 8 - 18\sqrt{2} \\ 8 & 16 - 9\sqrt{2} & 16 + 9\sqrt{2} \end{bmatrix}$$

Lösningförslag till uppgift 3. Vi beräknar singularvärdena och singularvektorer till A . Vi utgår från $AA^\dagger$ eftersom detta är en 2×2 -matris.

$$AA^\dagger = \begin{bmatrix} 21 & 42 \\ 42 & 84 \end{bmatrix}$$

Denna har egenvärdena 0 och 105. En normerad egenvektor för egenvärdet 105 är

$$\mathbf{y} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

vilket ger det enda singularvärdet $\sqrt{105}$ med vänstersingulär vektor $\mathbf{y}$ och motsvarande högersingulära vektor

$$\mathbf{x} = \frac{1}{\sqrt{105}} A^\dagger \mathbf{y} = \frac{1}{\sqrt{525}} \begin{bmatrix} 5 \\ 10 \\ 20 \end{bmatrix} = \frac{1}{\sqrt{21}} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}.$$

Detta ger singularvärdesuppdelningen

$$A = Q\Sigma P^\dagger = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & * \\ 2 & * \end{bmatrix} \begin{bmatrix} \sqrt{105} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{\sqrt{21}} \begin{bmatrix} 1 & 2 & 4 \\ * & * & * \end{bmatrix}$$

där $*$ väljs så att vi får ortonormala baser men inte behöver bestämmas för att beräkna Moore–Penrose inversen. Denna får vi som:

$$A^+ = P\Sigma^+ Q^\dagger = \frac{1}{\sqrt{21}} \begin{bmatrix} 1 & * \\ 2 & * \\ 4 & * \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{105}} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ * & * \end{bmatrix} = \frac{1}{105} \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 4 & 8 \end{bmatrix}$$

Lösningförslag till uppgift 4. Vi har att minstakvadratenlösningen med minst norm är:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A^+ \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{1}{108} \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{1}{108} \begin{bmatrix} 7 \\ 14 \\ 28 \end{bmatrix}$$

Lösningförslag till uppgift 5. Eftersom A är reell så är $A^\dagger A = A^T A$ en reell symmetrisk 2×2 -matris och har alltså en ortonormal bas bestående av reella egenvektorer. Detta ger högersingulära vektorer $\mathbf{x}_1, \mathbf{x}_2$ och motsvarande vänstersingulära vektorer $\mathbf{y}_1 = \frac{1}{\sigma_1} L(\mathbf{x}_1), \mathbf{y}_2 = \frac{1}{\sigma_2} L(\mathbf{x}_2)$ så att $X = [\mathbf{x}_1 \ \mathbf{x}_2]$ och $Y = [\mathbf{y}_1 \ \mathbf{y}_2]$ är reella ortogonala matriser.

Om $\mathbf{x}_1 = \begin{bmatrix} a \\ b \end{bmatrix}$ så finns två möjligheter för $\mathbf{x}_2$: antingen $\begin{bmatrix} b \\ -a \end{bmatrix}$ vilket gör X symmetrisk (determinant -1) eller $\begin{bmatrix} -b \\ a \end{bmatrix}$ vilket gör X anti-symmetrisk (determinant 1). På samma sätt kommer Y antingen vara symmetrisk (determinant -1) eller anti-symmetrisk (determinant 1).

Om $\det(A) > 0$ är $\det(\Sigma) > 0$ och väljer vi X symmetrisk blir även Y symmetrisk. Om $\det(A) < 0$ kommer precis en av X och Y vara symmetrisk och den andra anti-symmetrisk. Om $\det(A) = 0$ har vi högst ett singularvärde och $\mathbf{x}_2 \in \ker L$ och $\mathbf{y}_2 \in \ker L^\dagger$ och vi kan byta ut $\mathbf{y}_2$ mot $-\mathbf{y}_2$ om det behövs för att göra Y symmetrisk.

Vi kan alltså välja X och Y som symmetriska matriser precis när $\det(A) \geq 0$.

Lösningförslag till uppgift 6. Singularvärdena är kvadratrötterna till egenvärdena av $A^T A$ och AA^T . Vi får att

$$A^T A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 9 & 4 \\ 4 & 9 \end{bmatrix}$$

vars karakteristiska ekvation är $t^2 - 18t + 65 = 0$, vilket ger $t = 9 \pm \sqrt{81 - 65} = 9 \pm 4$. Det största egenvärdet är 13. Alltså är det största singularvärdet $\sigma_1 = \sqrt{13}$. Motsvarande egenvektor till $A^T A$ är $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$. Normerar vi den får vi den höger-singulära vektorn

$$\mathbf{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

och när vi multiplicerar med A får vi

$$A\mathbf{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}.$$

När vi normerar denna får vi

$$\mathbf{y} = \frac{1}{\sqrt{26}} \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$$

som är motsvarande vänstersingulära vektor.

Lösningförslag till uppgift 7. Vi beräknar $A^\dagger A$ och $AA^\dagger$ och får

$$A^\dagger A = \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{och} \quad AA^\dagger = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Därmed är båda singulärvärdena $\sqrt{2}$. Vi kan välja standardbasen för de högersingulära vektorerna och får

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{och} \quad \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix}$$

som motsvarande vänstersingulära vektorer.

Eftersom B är Hermitesk har den en ortogonal bas av egenvektorer och därmed är dessa egenvektorer singulärvektorer med beloppen av motsvarande egenvärden som singulärvärden. Det karakteristiska polynomet är $p_B(x) = x^2 - \text{tr}(B)x + \det(B) = x^2 - 2x$. Alltså är egenvärdena 0 och 2 med egenvektorer som ges av

$$\ker \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} = \text{Span} \left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\} \quad \text{respektive} \quad \ker \begin{bmatrix} -1 & i \\ -i & -1 \end{bmatrix} = \text{Span} \left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}$$



MATERIAL OM SANNOLIKHETSMATRISER OCH PERRON-FROBENIUS SATS

1. MARKOVKEDJOR OCH SANNOLIKHETSMATRISER

Många fenomen kan modelleras med *Markovkedjor* som är en typ av stokastiska processer som har stark koppling till linjär algebra. Se också §5.6 i kursboken inklusive avsnittet om slumpvandringar.

Definition 1.1. En *sannolikhetsmatris*, eller *stokastisk matris*, är en kvadratisk matris A där alla element är icke-negativa och summan i varje kolonn är 1. En *sannolikhetsvektor* är en kolonnvektor med icke-negativa element vars summa är 1.

Notera 1.2. Kolonnerna i en sannolikhetsmatris är sannolikhetsvektorer.

Notera 1.3. Låt $\mathbf{r} = [1 \ 1 \ \cdots \ 1]$ (en radvektor). Då är en icke-negativ vektor $\mathbf{p}$ en sannolikhetsvektor precis när $\mathbf{r}\mathbf{p} = 1$ och en icke-negativ matris A är en sannolikhetsmatris precis när $\mathbf{r}A = \mathbf{r}$.

Definition 1.4. En *Markovkedja* är en stokastisk process i diskret tid där sannolikheter-na att befinna sig i tillstånd i ges av $p_i(t)$ och

$$\mathbf{p}(t+1) = A\mathbf{p}(t), \quad \forall t \in \mathbb{N}$$

där A är en sannolikhetsmatris och *tillståndsvektorn* $\mathbf{p}(t) = [p_1(t) \ p_2(t) \ \cdots \ p_n(t)]^T$ är en *sannolikhetsvektor*.

Notera 1.5. Vi arbetar med kolonnvektorer $\mathbf{p}(t)$ och (vänster)-stokastiska matriser A så att $\mathbf{p}(t+1) = A\mathbf{p}(t)$. I litteraturen förekommer det att man istället arbetar med radvektorer som sannolikhetsvektorer och med höger-stokastiska matriser där varje *rad* summerar till 1. Då har man istället att $\mathbf{p}(t+1) = \mathbf{p}(t)A$ för en Markovkedja.

Vi kommer att använda radvektorer som ett hjälpmedel men inte som sannolikhetsvektorer.

Notera 1.6. Observera att eftersom A är en sannolikhetsmatris kommer $\mathbf{p}(t+1)$ att vara en sannolikhetsvektor om $\mathbf{p}(t)$ är det: elementen i $\mathbf{p}(t+1) = A\mathbf{p}(t)$ är icke-negativa och summan blir 1 i och med att

$$1 = \mathbf{r}\mathbf{p}(t) = (\mathbf{r}A)\mathbf{p}(t) = \mathbf{r}\mathbf{p}(t+1).$$

där $\mathbf{r} = [1 \ 1 \ \cdots \ 1]$. I många fall kommer Markovkedjan leda till ett stationärt tillstånd efter lång tid, så att

$$\lim_{t \rightarrow \infty} \mathbf{p}(t) = \mathbf{p}_\infty$$

Om så är fallet har vi att

$$A\mathbf{p}_\infty = \mathbf{p}_\infty$$

dvs $\mathbf{p}_\infty$ är en egenvektor med egenvärde 1.

Exempel 1.7. Om vi har Markovkedjan som definieras av

$$A = \frac{1}{3} \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{och} \quad \mathbf{p}(0) = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

får vi

$$\mathbf{p}(1) = A\mathbf{p}(0) = \frac{1}{9} \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} \quad \mathbf{p}(2) = A^2\mathbf{p}(0) = \frac{1}{27} \begin{bmatrix} 13 \\ 9 \\ 5 \end{bmatrix} \quad \mathbf{p}(3) = A^3\mathbf{p}(0) = \frac{1}{81} \begin{bmatrix} 31 \\ 41 \\ 9 \end{bmatrix},$$

och vidare

$$\mathbf{p}(10) = A^{10}\mathbf{p}(0) \approx \begin{bmatrix} 0,438489 \\ 0,411641 \\ 0,14987 \end{bmatrix}, \quad \mathbf{p}(100) = A^{100}\mathbf{p}(0) \approx \begin{bmatrix} 0,428571 \\ 0,428571 \\ 0,142857 \end{bmatrix} \approx \frac{1}{7} \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}.$$

Vi kan kontrollera att

$$A\mathbf{p}_\infty = \frac{1}{3} \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix} \frac{1}{7} \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = \mathbf{p}_\infty.$$

så det stationära tillståndet $\mathbf{p}_\infty = \lim_{t \rightarrow \infty} \mathbf{p}(t)$ är en egenvektor med egenvärde 1.

Sats 1.8. En sannolikhetsmatris har alltid minst en egenvektor med egenvärde 1.

Bevis. Låt som tidigare $\mathbf{r} = [1 \ 1 \ \cdots \ 1]$. Eftersom summan av elementen i kolonnerna är 1 är $\mathbf{r}A = \mathbf{r}$ eller ekvivalent:

$$A^T \mathbf{r}^T = (\mathbf{r}A)^T = \mathbf{r}^T$$

vilket betyder att A^T har en egenvektor $\mathbf{r}^T$ med egenvärde 1. Egenvärdena för A och A^T är dock desamma, så 1 är ett egenvärde också till A . Alltså finns en egenvektor till A med egenvärde 1. (Däremot är $\mathbf{r}^T$ oftast inte en egenvektor till A .) $\square$

Exempel 1.9. Det behöver inte vara så att $\lim_{t \rightarrow \infty} \mathbf{p}(t)$ existerar för alla $\mathbf{p}(0)$. Exempelvis får vi med

$$A = \frac{1}{3} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{och} \quad \mathbf{p}(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

att

$$\mathbf{p}(1) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{p}(2) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{p}(3) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{p}(0)$$

vilket ger $\mathbf{p}(3n) = \mathbf{p}(0)$, $\mathbf{p}(3n+1) = \mathbf{p}(1)$ och $\mathbf{p}(3n+2) = \mathbf{p}(2)$ för alla $n \geq 0$. Det finns fortfarande en egenvektor med egenvärde 1, nämligen $(\mathbf{p}(0) + \mathbf{p}(1) + \mathbf{p}(2))/3 = \frac{1}{3} [1 \ 1 \ 1]^T$, men vi får inte konvergens mot den om vi inte startar med den.

Exempel 1.10 (Google PageRank). Ett exempel på hur Markovkedjor används är i sökmotorn hos Google. Det var idéerna från Larry Page och Sergey Brin [BP98] som ledde till att Googles sökmotor blev den mest använda efter att den introducerades i slutet av 90-talet. Ett sätt att tänka på den är att se sidorna på internet som basvektorer i ett stort vektorrum. Att sidor innehåller länkar till andra sidor kan ses via en linjär avbildning som skickar en basvektor till medelvärdet av alla basvektorer som svarar till sidor som sidan har länkar till. Om vi utgår från att alla sidor har någon länk alls, skulle detta ge en sannolikhetsmatris och om det fanns sidor utan länkar skulle vi kunna skicka dessa basvektorer till sig själva och på så vis alltid få en sannolikhetsmatris som resultat. PageRank går nu ut på att beräkna egenvektorn $\mathbf{p}_\infty$ till denna sannolikhetsmatris och använda storleken på elementen i $\mathbf{p}_\infty$ för att rangordna sidorna. För att förstå hur det kommer sig att det fungerar kan vi tänka oss en bot som går runt på sidorna och slumpmässigt väljer en sida som den sida den befinner sig på länkar till att gå vidare till. Då ger vektorn $\mathbf{p}_\infty$ sannolikhetsfördelningen för var botten befinner sig efter mycket lång tid och eftersom viktiga sidor borde ha många sidor som länkar till dem och dessutom ha många andra viktiga sidor som länkar till dem borde sannolikheten att botten befinner sig på en viktig sida vara större än på en mindre viktig sida. Se <http://infolab.stanford.edu/pub/papers/google.pdf>.

Exempel 1.11. Ett nätverk med fem websidor refererar till varandra enligt följande matris.

$$A = \begin{bmatrix} 1/3 & 0 & 1/5 & 1/2 & 0 \\ 1/3 & 0 & 1/5 & 0 & 1/3 \\ 1/3 & 1/2 & 1/5 & 1/2 & 0 \\ 0 & 0 & 1/5 & 0 & 1/3 \\ 0 & 1/2 & 1/5 & 0 & 1/3 \end{bmatrix}$$

Observera att sidorna 1, 3 och 5 också länkar till sig själva. Matrisen A har en unik egenvektor med egenvärde 1 som ges av

$$\mathbf{p}_\infty = \frac{1}{45} \begin{bmatrix} 36 \\ 38 \\ 55 \\ 26 \\ 45 \end{bmatrix}$$

vilket leder till rangordningen 3, 5, 2, 1, 4 av sidorna.

2. PERRON–FROBENIUS SATS

För att exempelvis studera hur Markovkedjor beter sig behöver vi veta mer om hur det ser ut med egenvektorer och egenvärden för sannolikhetsmatriser. En viktig sats i det sammanhanget är *Perron–Frobenius sats* och det finns många olika varianter av denna sats. En av de enklare varianterna är följande:

Sats 2.1 (jämför med [Sad08, Thm. 5.1 (5)]¹). *Låt A vara en sannolikhetsmatris där alla element i A^m är positiva för något m . Då gäller följande*

- (a) *Det finns en egenvektor med egenvärde 1 där alla koordinater är positiva.*
- (b) $m_a(1) = m_g(1) = 1$.
- (c) *Om $\lambda \neq 1$ är ett egenvärde till A gäller att $|\lambda| < 1$.*

Enligt (a) och (b) finns en unik stokastisk egenvektor $\mathbf{v}$ med egenvärde 1.

- (d) *För varje stokastisk vektor $\mathbf{p}$ gäller att $\lim_{n \rightarrow \infty} A^n \mathbf{p} = \mathbf{v}$.*

¹Observera att beviset i [Sad08] är lite ofullständigt eftersom Sadun implicit antar att A är diagonaliserbar och att $m_a(1) = 1$.

Bevis. (a) Vi vet att ett egetvärde är 1 enligt Sats 1.8. Låt $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$ vara en egenvektor med $A\mathbf{x} = \mathbf{x}$. Låt $\mathbf{x}^+ = [|x_1| \ |x_2| \ \cdots \ |x_n|]^T$. Vi får att $(A\mathbf{x}^+)_j \geq (\mathbf{x}^+)_j$ för $j = 1, 2, \dots, n$, men eftersom $\sum_{j=1}^n (A\mathbf{x}^+)_j = \sum_{j=1}^n (\mathbf{x}^+)_j$ och alla termer är icke-negativa måste $A\mathbf{x}^+ = \mathbf{x}^+$.

Eftersom alla element i A^m är positiva får vi att alla koordinater till $\mathbf{x}^+ = A^m \mathbf{x}^+$ är positiva.

- (b) Om $\mathbf{y} = \mathbf{x} + \mathbf{x}^+$ inte är noll är det en egenvektor med egetvärde 1 och vi kan använda samma argument som tidigare för att se att alla koordinater till $\mathbf{y}^+$ är positiva, och därmed $\mathbf{x} = \mathbf{x}^+$ eller $\mathbf{x} = -\mathbf{x}^+$.

Därmed har vi visat att varje egenvektor med egetvärde 1 har alla koordinater av samma tecken. Om det fanns två oberoende egenvektorer med egetvärde 1 skulle det finnas en linjärkombination som ger en vektor där en av koordinaterna är noll, vilket ger motsägelse mot det vi just har visat. Alltså är $\dim E_1 = 1$ och $m_g(1) = 1$.

Om $m_a(1) > 1$ finns en generaliserad egenvektor $\mathbf{y}$ med $A\mathbf{y} = \mathbf{y} + \mathbf{x}^+$, men eftersom summan av koordinaterna bevaras genom A kan inte detta gälla. Alltså är $m_a(1) = m_g(1) = 1$.

- (c) För att se att det inte kan finnas några andra egetvärden med belopp ett eller större, även om vi betraktar komplexa egetvärden, måste vi använda villkoret att A^m bara har positiva element för något m . Eftersom egetvärdena för A^m är potenser av egetvärdena för A räcker det att betrakta fallet då A bara har positiva element. Låt $\mathbf{x}^+$ vara den positiva egenvektorn med egetvärde 1 och bilda $A_\infty = \mathbf{x}^+ [1 \ 1 \ \cdots \ 1]$. Vi har då att

$$AA_\infty = A_\infty = A_\infty A$$

och vi kan hitta ett $\varepsilon > 0$ så att $A - \varepsilon A_\infty$ bara har positiva element. Låt nu

$$B = \frac{1}{1 - \varepsilon} (A - \varepsilon A_\infty)$$

och beräkna potenserna av B som

$$\begin{aligned} B^n &= \frac{1}{(1 - \varepsilon)^n} (A - \varepsilon A_\infty)^n = \frac{1}{(1 - \varepsilon)^n} \sum_{k=0}^n (-\varepsilon)^k \binom{n}{k} A^{n-k} A_\infty^k \\ &= \frac{1}{(1 - \varepsilon)^n} \left(A^n + ((1 - \varepsilon)^n - 1) A_\infty \right) = A_\infty + \frac{1}{(1 - \varepsilon)^n} (A^n - A_\infty) \end{aligned}$$

och

$$A^n - A_\infty = (1 - \varepsilon)^n (B^n - A_\infty)$$

Eftersom B är en sannolikhetsmatris ligger elementen i B^n i intervallet $[0, 1]$ och därmed måste elementen i $A^n - A_\infty$ gå mot noll då $n \rightarrow \infty$. Om $\mathbf{y}$ är en, eventuellt komplex, egenvektor med egetvärde λ får vi att

$$(A^n - A_\infty)\mathbf{y} = (\lambda^n I - A_\infty)\mathbf{y}$$

måste gå mot noll när $n \rightarrow \infty$ och det är bara möjligt om $\lambda = 1$ eller $|\lambda| < 1$. Alltså måste övriga egetvärden ha belopp strikt mindre än 1.

- (d) Mer allmänt gäller att om A är en matris där $\lambda = 1$ har geometrisk multiplicitet 1 och egenvektor ξ_1 och $|\lambda| < 1$ för övriga egetvärden så gäller att $\lim_{n \rightarrow \infty} A^n \mathbf{w} = c_1 \xi_1$ för varje vektor $\mathbf{w}$ och någon skalär c_1 . Om A är diagonaliserbar är detta enkelt att se: om $\mathbf{w} = \sum_i c_i \xi_i$ så är $A^n \mathbf{w} = \sum_i c_i \lambda_i^n \xi_i \rightarrow c_1 \xi_1$ då $n \rightarrow \infty$.

För att visa detta för icke-diagonaliserbara A kan vi anta att A är på Jordans normalform och alltså har ett 1×1 -block med egenvärde 1 och övriga Jordan-block med egenvärde av absolutbelopp mindre än 1. Det räcker att visa att om

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

är ett $d \times d$ Jordanblock med $|\lambda| < 1$ så är $\lim_{n \rightarrow \infty} A^n = 0$. Detta följer av att

$$A^n = \begin{bmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \dots & \binom{n}{d-1}\lambda^{n-d+1} \\ 0 & \lambda^n & \binom{n}{1}\lambda^{n-1} & \dots & \binom{n}{d-2}\lambda^{n-d+2} \\ 0 & 0 & \lambda^n & \dots & \binom{n}{d-3}\lambda^{n-d+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^n \end{bmatrix}$$

eftersom $\binom{n}{k}|\lambda|^{n-k} \leq \frac{n^k}{k!}|\lambda|^{n-k} \rightarrow 0$ då $n \rightarrow \infty$. □

Notera 2.2. Om en sannolikhetsvektor $\mathbf{p}$ är en egenvektor så har den egenvärde 1. Detta följer av att $\sum_j A p_j = \sum_j p_j$. Sats 2.1 säger att om A^m är positiv så existerar det en unik sannolikhetsvektor som är en egenvektor och alla dess koordinater är positiva. Motsvarande Markovprocess har alltså ett unikt stationärt tillstånd.

En matris som i Sats 2.1 kallas *reguljär*. Vi upprepar definitionen.

Definition 2.3. En sannolikhetsmatris är *reguljär*² om alla element är positiva för någon potens av matrisen.

Exempel 2.4. Matrisen A för Markovkedjan i Exempel 1.7 uppfyller

$$A^2 = \frac{1}{9} \begin{bmatrix} 5 & 2 & 6 \\ 2 & 7 & 0 \\ 2 & 0 & 3 \end{bmatrix} \quad A^3 = \frac{1}{27} \begin{bmatrix} 9 & 16 & 6 \\ 16 & 4 & 21 \\ 2 & 7 & 0 \end{bmatrix} \quad A^4 = \frac{1}{81} \begin{bmatrix} 41 & 24 & 48 \\ 24 & 53 & 12 \\ 16 & 4 & 21 \end{bmatrix}.$$

Alltså är A reguljär och Markovprocessen har en unik stationär fördelning: den som vi fann i exemplet. Matrisen i Exempel 1.9 är inte reguljär.

Definition 2.5. En Markovkedja (eller motsvarande sannolikhetsmatris) är *irreducibel* om varje tillstånd går att nå från varje annat tillstånd efter ett (icke-fixerat) antal tidssteg.

Markovkedjan i Exempel 1.9 är irreducibel. Varje reguljär sannolikhetsmatris är irreducibel: om alla element i A^m är positiva så går varje tillstånd att nå från varje annat tillstånd efter exakt m tidssteg.

Sats 2.6. Om A är en irreducibel sannolikhetsmatris gäller följande.

- (a) Det finns en egenvektor med egenvärde 1 där alla koordinater är positiva.
- (b) Alla komplexa egenvärden med belopp 1 har algebraisk multiplicitet 1.
- (c) De komplexa egenvärden som har belopp 1 är precis de r te enhetsrötterna för något $r \in \mathbb{N}$ (kallas *perioden* av A).

Exempel 2.7. För matrisen i Exempel 1.9 är perioden 3 och egenvärdena är 1, $e^{i2\pi/3}$ och $e^{-i2\pi/3}$.

²I [Sad08, Def., s. 126] så används begreppet reguljär i en snävare betydelse: enbart matriser som har positiva element kallas reguljära.

Exempel 2.8. Följande sannolikhetsmatris är inte irreducibel

$$A = \begin{bmatrix} 0,5 & 0,2 & 0,2 \\ 0,5 & 0,8 & 0,2 \\ 0 & 0 & 0,6 \end{bmatrix}$$

men har ändå en unik egenvektor med egenvärde 1. De övriga egenvärdena är 0,3 och 0,6.

Exempel 2.9. Följande sannolikhetsmatris

$$A = \begin{bmatrix} 0,5 & 0,2 & 0 \\ 0,5 & 0,8 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

har två olika stokastiska egenvektorer med egenvärde 1

$$\mathbf{p}_1 = \frac{1}{7} \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix} \quad \text{och} \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

samt egenvärdet 0,3. Observera att $\mathbf{p}_1$ och $\mathbf{p}_2$ är icke-negativa men inte positiva.

I allmänhet gäller:

Sats 2.10 ([Sad08, Thm. 5.1 (1)&(4)]). Om A är en sannolikhetsmatris gäller

- (a) Det finns minst en sannolikhetsvektor med egenvärde 1 (men alla koordinater behöver inte vara positiva).
- (b) Om λ är ett (komplext) egenvärde, gäller $|\lambda| \leq 1$ (men det kan finnas flera med belopp 1 och den geometriska multipliciteten kan vara större än 1).

Bevis. Detta följer av Sats 2.1 genom att skriva A som en gräns av matriser med enbart positiva element. Tex kan vi välja $A_\varepsilon = \frac{1}{1+\varepsilon}(A + \varepsilon B)$ där B är $n \times n$ -matrisen där alla element är $\frac{1}{n}$. Då är $A = \lim_{\varepsilon \rightarrow 0} A_\varepsilon$. $\square$

Notera 2.11. Om ξ är en egenvektor med egenvärde λ gäller

$$\mathbf{r}\xi = (\mathbf{r}A)\xi = \mathbf{r}(\lambda\xi) = \lambda(\mathbf{r}\xi)$$

vilket ger $\lambda = 1$ eller $\mathbf{r}\xi = 0$. Egenvektorer till egenvärden $\lambda \neq 1$ är alltså aldrig sannolikhetsvektorer.

Exempel 2.12. Låt

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

Då har vi egenvärdet $\lambda = 1$ med den stokastiska egenvektorn $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ och egenvärdet $\lambda = \frac{1}{2}$ med egenvektorn $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

I beviset av [Sad08, Thm. 5.1] antas implicit att matrisen A är diagonaliserbar och att egenvärdet 1 har algebraisk multiplicitet 1. Vi såg att det senare gäller om A är reguljär (eller mer allmänt irreducibel) men A behöver inte vara diagonaliserbar även om A är reguljär:

Exempel 2.13. Låt

$$A = \frac{1}{8} \begin{bmatrix} 4 & 7 & 8 \\ 4 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Då är A^3 positiv så A är reguljär. Egenvärdena är $\lambda = 1$ och $\lambda = \frac{-1}{4}$ med algebraisk multiplicitet 2. Egenrummen är

$$E_1 = \text{Span} \left\{ \frac{1}{25} \begin{bmatrix} 16 \\ 8 \\ 1 \end{bmatrix} \right\}$$
$$E_{\frac{-1}{4}} = \text{Span} \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

Matrisen är alltså inte diagonaliserbar.

REFERENSER

- [BP98] Sergey Brim and Larry Page. The anatomy of a large-scale hypertextual web search engine, 1998. <http://infolab.stanford.edu/pub/papers/google.pdf>.
- [Sad08] Lorenzo Sadun. *Applied linear algebra*. American Mathematical Society, Providence, RI, second edition, 2008. The decoupling principle.
- [Sam57] Hans Samelson. On the Perron-Frobenius theorem. *Michigan Math. J.*, 4:57–59, 1957.
- [SC97] Laurent Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.
- [Wik17] Wikipedia. Perron–frobenius theorem — wikipedia, the free encyclopedia, 2017.

ÖVNINGAR PÅ PERRON–FROBENIUS SATS OCH SANNOLIKHETSMATRISER

Uppgift 1. Bestäm egenvektorn med egenvärde 1 till matrisen

$$A = \frac{1}{4} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 2 & 2 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

och avgör om $\lim_{t \rightarrow \infty} \mathbf{p}(t)$ existerar för alla val av sannolikhetsvektorn $\mathbf{p}(0)$ om $\mathbf{p}(t+1) = A\mathbf{p}(t)$ för alla $t \in \mathbb{N}$.

Uppgift 2. En *permutationsmatris* är en kvadratisk matris där varje rad och varje kolonn innehåller precis en etta och resten nollor.

- (a) Visa att alla potenser av en permutationsmatris också är permutationsmatriser.
- (b) Visa att permutationsmatriser är ortogonala.
- (c) Visa att om P är en permutationsmatris av storlek $n \times n$ och $p(x)$ är ett polynom med reella ickenegativa koefficienter med $p(1) = 1$ så är $p(P)$ en sannolikhetsmatris.
- (d) Vilka permutationsmatriser har en unik egenvektor med egenvärde 1?

Uppgift 3. Visa att om A och B är två kommuterande sannolikhetsmatriser och båda har en unik egenvektor (upp till multiplikation med skalär) med egenvärde 1 så är det en gemensam egenvektor.

Uppgift 4. (Uppgift 5 på tentamen 2018-01-10)

- (a) Vad är en *sannolikhetsmatris*? (1 p)
- (b) Alla sannolikhetsmatriser har 1 som ett egenvärde. Hur bevisar man det? (1 p)
- (c) Formulera någon version av Perron–Frobenius sats. (1 p)
- (d) Ge ett exempel som visar att $\lim_{n \rightarrow \infty} A^n$ inte behöver existera även om A är en irreducibel sannolikhetsmatris. (1 p)

Uppgift 5. (Seminarieuppgift 3.2, HT18)

För varje heltal $N > 0$ kan vi studera den slumpvandrande process som ges av en linjär graf, dvs en graf med hörnen $1, 2, \dots, N$ med kanter mellan efterföljande hörn. Sannolikheten att gå från hörn i till hörn $i+1$ eller hörn $i-1$ är $1/2$ utom om $i=1$ eller N då sannolikheten är 1 att gå till det enda intilliggande hörnet. Låt A_N vara den sannolikhetsmatris (stokastiska matris) som beskriver denna Markovkedja.

- (a) Använd lämplig programvara för att bestämma egenvärdena till A_N för ett antal olika N . Hur stora N fungerar?
- (b) Plotta egenvärdena och ställ upp en hypotes för hur de beror på N .
- (c) Är A_N reguljär?
- (d) Bilda nya stokastiska matriser $B_N = \frac{1}{2}A + \frac{1}{2}A^2$ och $C_N = \frac{1}{2}I + \frac{1}{2}A$. Är B_N och/eller C_N reguljära?
- (e) Hur förhåller sig egenvärdena till B_N och C_N till egenvärdena till A_N ?
- (f) Hur kan vi tolka processerna som ges av B_N och C_N i termer av processen som ges av A_N ?

Uppgift 6. (Uppgift 3 på tentamen 2019-01-09)

I en Markovkedja ges fördelningen vid tidpunkt $t = n + 1$ av fördelningen vid tidpunkt $t = n$ genom den stokastiska matrisen

$$A = \frac{1}{10} \begin{bmatrix} 3 & 2 & 4 \\ 7 & 4 & 5 \\ 0 & 4 & 1 \end{bmatrix}$$

Avgör om $\lim_{n \rightarrow \infty} A^n$ existerar och bestäm i så fall gränsvärdet. (4 p)

Uppgift 7. (Uppgift 9 på tentamen 2019-04-17)

Bestäm alla unitära stokastiska matriser (sannolikhetsmatriser) och avgör om de bildar en grupp³ under matrismultiplikation. (3 p)

³Dvs: om A och B är unitära stokastiska matriser så är även A^{-1} och AB unitära stokastiska matriser.

BILAGA A. NÅGRA MER ALLMÄNNA SATSER PÅ SAMMA TEMA (INGÅR EJ)

Perron–Frobenius sats (Sats 2.1) för sannolikhetsmatriser är ett specialfall av mer generella fenomen för matriser som inte är sannolikhetsmatriser, men där alla element är icke-negativa. För att formulera en mer generell version av Perron–Frobenius sats behöver vi införa några nya begrepp.

Definition A.1. En *icke-negativ matris* är en matris där alla element är icke-negativa. En *positiv matris* är en matris där alla element är positiva.

En *reguljär icke-negativ matris* är en icke-negativ matris där någon potens är positiv.

En icke-negativ matris är *irreducibel* om det inte finns en samtidig omordning av rader och kolonner så att matrisen blir block-övertriangulär (med mer än ett block).

Notera A.2. Denna definition av irreducibel är ekvivalent med den tidigare för en sannolikhetsmatris. Om A är en sannolikhetsmatris, låt $i \preceq j$ beteckna att tillstånd j går att nå från i . Vi grupperar tillstånden i grupper där två tillstånd tillhör samma grupp om både $i \preceq j$ och $j \preceq i$. Vi ordnar grupperna så att alla tillstånden i grupp k kan nå alla tillstånden i grupperna $1, 2, \dots, k-1$. Inom varje grupp ordnar vi tillstånden på ett godtyckligt sätt. Att ordna om tillstånden svarar mot en samtidig omordning av rader och kolonner i A och den resulterande matrisen blir block-övertriangulär där varje block svarar mot en grupp av tillstånd. Matrisen är irreducibel precis när alla tillstånd tillhör samma grupp.

Sats A.3. Om A är en irreducibel icke-negativ matris gäller följande där

$$\rho = \max\{|\lambda| : \lambda \text{ är egenvärde till } A\}.$$

- (a) A har en egenvektor med egenvärde ρ där alla koordinater är positiva.
- (b) Alla komplexa egenvärden med belopp ρ har algebraisk multiplicitet 1.
- (c) De komplexa egenvärden som har belopp ρ är enhetsrötter gånger ρ och ligger på jämna intervall kring cirkeln med radie ρ kring origo.

Denna generella variant är mycket svårare att bevisa, men det finns många olika bevis för den. Se exempelvis [Sam57] eller [SC97].

Följande sats kan användas för att kontrollera om en matris är irreducibel utan att kontrollera alla omordningar av basen.

Sats A.4. Om A är icke-negativ $n \times n$ -matris är A irreducibel om och endast om $(I + A)^{n-1}$ är positiv.

Bevis. Om A inte är irreducibel kan vi ordna om basen så att den blir block-övertriangulär och då kan inte $(I + A)^k$ vara positiv för något k .

Antag nu att A är irreducibel och att $(I + A)^{n-1}$ inte är positiv. Då kan vi hitta en icke-negativ vektor $\mathbf{x} \neq 0$ (t ex en standardbasvektor) så att $(I + A)^{n-1}\mathbf{x}$ har minst en koordinat som är noll. Vi kan byta ordning på basvektorerna så att $\mathbf{x} = [\mathbf{y} \ 0]^T$, där $\mathbf{y}$ är positiv. I denna ordning kan vi skriva $I + A$ på blockform

$$I + A = \begin{bmatrix} I + X_{11} & X_{12} \\ X_{21} & I + X_{22} \end{bmatrix}$$

där $X_{21} \neq 0$ eftersom A är irreducibel. Vi får att

$$(I + A)\mathbf{x} = \begin{bmatrix} I + X_{11} & X_{12} \\ X_{21} & I + X_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{y} + X_{11}\mathbf{y} \\ X_{21}\mathbf{y} \end{bmatrix}$$

och denna vektor har den övre delen positiv och minst en nollskild koordinat i den undre delen eftersom $X_{21} \neq 0$ och $\mathbf{y}$ är positiv. Vi upprepar detta och varje gång vi multiplicerar med $I + A$ ökar antalet nollskilda koordinater med minst ett så länge det är mindre än n . Alltså måste alla koordinater i $(I + A)^{n-1}\mathbf{x}$ vara positiva. $\square$

Notera A.5. Det finns också ett lätt kombinatoriskt bevis för Sats A.4. Som vi såg i Anmärkning A.2 så är A irreducibel precis om “vi kan gå mellan alla tillstånd”, dvs givet $1 \leq a, b \leq n$ finns det en stig av längd $k \leq n - 1$ med $i(0) = a$ och $i(k) = b$ och $a_{i(0)i(1)} > 0, a_{i(1)i(2)} > 0, \dots, a_{i(k-1)i(k)} > 0$. Om vi ersätter A med $A + I$ kan vi låta stigen ha längd exakt $n - 1$ genom att t ex lägga till $i(k) = i(k + 1) = \dots = i(n - 1) = b$ vilket är ekvivalent med att $(A + I)^{n-1}$ är positiv.

LÖSNINGSFÖRSLAG

Lösningförslag till uppgift 1. För att hitta egenvektorn gausseliminerar vi $A - I$

$$\frac{1}{4} \begin{bmatrix} -3 & 1 & 2 & 0 \\ 2 & -2 & 1 & 0 \\ 1 & 0 & -3 & 2 \\ 0 & 1 & 0 & -2 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & -7 & 6 \\ 0 & -2 & 7 & -4 \\ 1 & 0 & -3 & 2 \\ 0 & 1 & 0 & -2 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & -7 & 8 \\ 0 & 0 & 7 & -8 \\ 1 & 0 & -3 & 2 \\ 0 & 1 & 0 & -2 \end{bmatrix}$$

vilket har en unik lösning (upp till multiplikation med skalär):

$$\mathbf{p} = \frac{1}{39} \begin{bmatrix} 10 \\ 14 \\ 8 \\ 7 \end{bmatrix}$$

Vi undersöker vidare om A är reguljär

$$A^2 = \frac{1}{16} \begin{bmatrix} 5 & 3 & 5 & 4 \\ 7 & 6 & 7 & 2 \\ 2 & 3 & 3 & 6 \\ 2 & 4 & 1 & 4 \end{bmatrix}.$$

Eftersom A^2 är positiv så är A reguljär. Enligt Perron-Frobenius sats så konvergerar då alla startvektorer mot egenvektorn $\mathbf{p}$ med egenvärde 1.

Lösningförslag till uppgift 2.

- En $(n \times n)$ -matris A är en permutationsmatris precis om $A\mathbf{e}_i = \mathbf{e}_{\sigma(i)}$ för en permutation (dvs bijektion) $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$. Då är $A^m(\mathbf{e}_i) = \mathbf{e}_{\sigma^m(i)}$ och alltså är A^m permutationsmatrisen som hör till permutationen $\sigma^m = \sigma \circ \dots \circ \sigma$.
- Vi har att $\langle A\mathbf{e}_i | A\mathbf{e}_j \rangle = \delta_{\sigma(i), \sigma(j)} = \sigma_{i,j} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle$ eftersom σ är injektiv.
- Låt $p(x) = \sum_i a_i x^i$ vara ett polynom med icke-negativa koefficienter. Eftersom $I, A, A^2, \dots$ är permutationsmatriser så är $p(A) = \sum_i a_i A^i$ icke-negativ och varje kolonn summerar till $\sum_i a_i$. Om $p(1) = 1$ så är alltså $p(A)$ en permutationsmatris.
- Alla permutationsmatriser har egenvektorn $[1 \ 1 \ \dots \ 1]^T$ med egenvärdet 1. En permutationsmatris är aldrig reguljär men den är irreducibel precis när σ är en cyklisk permutation av längd n : $\{1, \sigma(1), \sigma^2(1), \dots, \sigma^{n-1}(1)\} = \{1, 2, 3, 4, \dots, n\}$. Om cykeln som 1 ingår i har en kortare längd $d < n$, dvs $\sigma^d(1) = 1$, så är $\mathbf{e}_1 + \mathbf{e}_{\sigma(1)} + \dots + \mathbf{e}_{\sigma^{d-1}(1)}$ också en egenvektor.

Lösningförslag till uppgift 3. Låt $\mathbf{p}$ vara en stokastisk egenvektor till A med egenvärde 1, dvs $A\mathbf{p} = \mathbf{p}$. Då är

$$AB\mathbf{p} = BA\mathbf{p} = B\mathbf{p}.$$

Alltså är $B\mathbf{p}$ också en egenvektor till A med egenvärde 1. Eftersom $B\mathbf{p}$ är en stokastisk vektor och A hade en unik egenvektor är alltså $B\mathbf{p} = \mathbf{p}$. Dvs, $\mathbf{p}$ är en gemensam egenvektor med egenvärde 1.

Lösningförslag till uppgift 4.

- En sannolikhetsmatris är en reell kvadratisk matris där alla element är icke-negativa och summorna av elementen i varje kolonn är ett.

- (b) Villkoret att summan i kolonnerna är 1 betyder att $\mathbf{r} = [1 \ 1 \ \dots 1]$ är en vänstereigenvektor med egenvärde 1, dvs att $\mathbf{r}A = \mathbf{r}$. Det betyder att $A^T \mathbf{r}^T = \mathbf{r}^T$, så 1 är ett egenvärde till A^T , men eftersom $\det(xI - A^T) = \det(xI - A)$ har A och A^T samma egenvärden. Därmed är 1 ett egenvärde till A .
- (c) En version av Perron–Frobenius sats är

Sats. Låt A vara en sannolikhetsmatris där alla element i A^m är positiva för något m . Då gäller följande

- (a) Det finns en egenvektor med egenvärde 1 där alla koordinater är positiva.
 (b) $m_a(1) = m_g(1) = 1$.
 (c) Om $\lambda \neq 1$ är ett egenvärde gäller att $|\lambda| < 1$.

- (d) Matrisen

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

är en irreducibel sannolikhetsmatris men vi har att

$$A^n = \begin{cases} I & \text{om } n \text{ är jämnt,} \\ A & \text{om } n \text{ är udda.} \end{cases}$$

så följderna $\{A^n\}_{n \geq 0}$ konvergerar inte.

Lösningförslag till uppgift 6. Enligt Perron–Frobenius sats kommer vi att ha konvergens om matrisen är reguljär, dvs om någon potens är positiv. I vårt fall är A^2 positiv, så vi får att

$$\lim_{n \rightarrow \infty} A^n = A_\infty = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

där $[x_1 \ x_2 \ x_3]^T$ är en egenvektor med egenvärde 1 och $x_1 + x_2 + x_3 = 1$. Vi beräknar denna egenvektor som ett element i kärnan till $A - I$, dvs genom ekvationssystemet med totalmatris

$$\left[\begin{array}{ccc|c} -7 & 2 & 4 & 0 \\ 7 & -6 & 5 & 0 \\ 0 & 4 & -9 & 0 \end{array} \right] \sim \left[\begin{array}{ccc|c} -7 & 2 & 4 & 0 \\ 0 & -4 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \sim \left[\begin{array}{ccc|c} -14 & 0 & 17 & 0 \\ 0 & -4 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

vilket ger $(x_1, x_2, x_3) = t(34, 63, 28)$ och för att $x_1 + x_2 + x_3 = 1$ måste $t^{-1} = 34 + 63 + 28 = 125$. Alltså får vi konvergens mot

$$\frac{1}{125} \begin{bmatrix} 34 & 34 & 34 \\ 63 & 63 & 63 \\ 28 & 28 & 28 \end{bmatrix} = \begin{bmatrix} 0,272 & 0,272 & 0,272 \\ 0,504 & 0,504 & 0,504 \\ 0,224 & 0,224 & 0,224 \end{bmatrix}.$$

Lösningförslag till uppgift 7. Om vi har en unitär stokastisk matris har vi att varje kolonn har element som uppfyller att $0 \leq x_i \leq 1$, för $i = 1, 2, \dots, n$ och dessutom att $x_1^2 + x_2^2 + \dots + x_n^2 = 1$ och $x_1 + x_2 + \dots + x_n = 1$. Därmed är

$$0 = 1 - 1 = (x_1 + x_2 + \dots + x_n)^2 - (x_1^2 + x_2^2 + \dots + x_n^2) = 2 \sum_{1 \leq i < j \leq n} x_i x_j.$$

Eftersom alla x_i är icke-negativa måste alla produkter vara 0, vilket ger att alla x_i utom en måste vara noll. Den sista måste vara ett. Alltså är varje kolonn en standardbasvektor. Eftersom matrisen ska vara inverterbar måste det vara en permutation av alla n standardbasvektorer.

Produkten av två sådana matriser är på nytt en sådan matris och de är alla inverterbara med inverser på samma form. Alltså bildar de en grupp under matrismultiplikation.



MATERIAL OM TENSORER OCH YTTRE ALGEBRAN

INNEHÅLL

1. Duala rum och avbildningar	1
1.1. Linjära former, duala rum och dual bas	2
1.2. Linjära former som radvektorer	3
1.3. Evalueringsavbildningen och dubbeldualen	3
1.4. Duala rum och inre produkter	3
1.5. Dual avbildning	4
2. Tensorprodukt	6
2.1. Konstruktion med hjälp av baser	6
2.2. Konstruktion med bilinjära avbildningar	8
2.3. Tensorprodukt av avbildningar och matriser	10
2.4. Konstruktion som kvot	11
Övningar på tensorprodukt	13
3. Yttre potenser och yttre algebran	15
3.1. Alternierande avbildningar	15
3.2. Yttre potensen $\wedge^\ell V$ med bas	15
3.3. Universell egenskap för yttre potenser	16
3.4. Konstruktion av $\wedge^\ell V$ som kvot	17
3.5. Yttre algebran $\wedge V$	18
3.6. Determinanten	19
3.7. Konstruktion av $\wedge^\ell V$ med alternierande avbildningar	19
3.8. Isomorfier mellan yttre produkter	19
Referenser	21
Övningar på yttre algebran	22
Lösningsförslag	23

1. DUALA RUM OCH AVBILDNINGAR

Vi har tidigare diskuterat duala rum (se anteckningar till föreläsning 7 och [Sad08, §6.3]) och duala avbildningar (appendix till föreläsning 8). I detta avsnitt gör vi detta lite mer ingående. Det går också bra att börja läsa nästa avsnitt om tensorprodukter direkt.

1.1. Linjära former, duala rum och dual bas.

Definition 1.1. En *linjär form* på ett vektorrum V över en kropp k är en linjär avbildning $\varphi: V \rightarrow k$. Det *duala rummet* eller *dualen* till V är

$$V^* = \text{Hom}_k(V, k) = \{\varphi: \varphi \text{ är en linjär form på } V\}.$$

Notera 1.2. Linjära former kallas också för funktionaler och kovektorer. Det duala rummet betecknas ibland $V^\vee$.

Exempel 1.3. Låt $V = C([0, 1], \mathbb{C})$ vara rummet av kontinuerliga funktioner $f: [0, 1] \rightarrow \mathbb{C}$. Om $a \in [0, 1]$ så är evaluering i a en linjär form på V :

$$\text{ev}_a: V \rightarrow \mathbb{C}, \quad \text{ev}_a(f) = f(a)$$

Dvs, $\text{ev}_a \in V^*$.

Definition 1.4. Om $\dim V < \infty$ och $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ är en bas för V ges den *duala basen* av $\mathcal{B}^* = \{\mathbf{b}_1^*, \mathbf{b}_2^*, \dots, \mathbf{b}_n^*\}$ där $\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}$.

Notera 1.5. Den linjära formen $\mathbf{b}_i^*$ beror på hela basen och inte bara $\mathbf{b}_i$.

Notera 1.6. Om $\mathbf{v} = \sum_{i=1}^n c_i \mathbf{b}_i$ så är $c_i = \mathbf{b}_i^*(\mathbf{v})$, dvs:

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} \mathbf{b}_1^*(\mathbf{v}) \\ \mathbf{b}_2^*(\mathbf{v}) \\ \vdots \\ \mathbf{b}_n^*(\mathbf{v}) \end{bmatrix}.$$

Sats 1.7. Om $\dim V < \infty$ så är den duala basen $\mathcal{B}^*$ en bas för V^* och om $\varphi \in V^*$ så är

$$[\varphi]_{\mathcal{B}^*} = \begin{bmatrix} \varphi(\mathbf{b}_1) \\ \varphi(\mathbf{b}_2) \\ \vdots \\ \varphi(\mathbf{b}_n) \end{bmatrix}$$

Bevis. Om $\varphi \in V^*$ och $\mathbf{v} = \sum_{i=1}^n c_i \mathbf{b}_i \in V$ så är

$$\varphi(\mathbf{v}) = \sum_{i=1}^n c_i \varphi(\mathbf{b}_i) = \sum_{i=1}^n \mathbf{b}_i^*(\mathbf{v}) \varphi(\mathbf{b}_i)$$

eftersom φ är linjära. Alltså är

$$\varphi(-) = \sum_{i=1}^n \varphi(\mathbf{b}_i) \mathbf{b}_i^*(-)$$

vilket visar satsen. □

Exempel 1.8. Låt $V = \mathbb{C}[x]_{\leq 2}$. Låt $\mathcal{B} = \{1, x, x^2\}$ och $\mathcal{B}^* = \{1^*, x^*, (x^2)^*\}$. Då är t ex

$$x^*(1 + 3x + 5x^2) = 3.$$

Om φ är den linjära formen som ges av

$$\varphi(a + bx + cx^2) = 3b + 2c$$

så är $\varphi = 3x^* + 2(x^2)^*$ och alltså:

$$[\varphi]_{\mathcal{B}^*} = \begin{bmatrix} \varphi(1) \\ \varphi(x) \\ \varphi(x^2) \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}.$$

1.2. Linjära former som radvektorer. Låt $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ vara en bas för V . Detta ger en isomorfi $\mathbb{R}^n \rightarrow V$: vi identifierar $\mathbf{v} \in V$ med koordinatvektorn $[\mathbf{v}]_{\mathcal{B}} \in \mathbb{R}^n$. Vi kan då identifiera det duala rummet V^* med linjära avbildningar $\mathbb{R}^n \rightarrow \mathbb{R}^1$, dvs $(1 \times n)$ -matriser, dvs *radvektorer*. Om $\varphi \in V^*$ så låter vi $[\varphi]_{\mathcal{B}}$ beteckna motsvarande radvektor. Den duala basvektorn $\mathbf{b}_i^*$ svarar mot radvektorn $(\mathbf{e}_i)^T$.

I notationen för matrisen för linjära avbildningar hade vi skrivit $[\varphi]_{\mathcal{E}, \mathcal{B}}$ där $\mathcal{E} = \{\mathbf{e}_1\}$ är standardbasen för $\mathbb{R}$. Om $\mathbf{v} \in V$ och $\varphi \in V^*$ så är alltså $\varphi(\mathbf{v}) = [\varphi]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}$ (matrismultiplikation av en radvektor med en kolumnvektor).

Exempel 1.9. I det tidigare exemplet har vi

$$\varphi(a + bx + cx^2) = [\varphi]_{\mathcal{B}} [a + bx + cx^2]_{\mathcal{B}} = \begin{bmatrix} 0 & 3 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 3b + 2c$$

1.3. Evalueringsavbildningen och dubbeldualen.

Definition 1.10. *Evalueringsavbildningen* är $\text{ev}: V^* \times V \rightarrow k$ där $\text{ev}(\varphi, \mathbf{v}) = \varphi(\mathbf{v})$.

Notera 1.11. Evalueringsavbildningen är *bilinjär*.

Definition 1.12. Det finns en naturlig¹ linjär avbildning

$$\gamma: V \rightarrow (V^*)^*$$

där $\gamma(\mathbf{v}): V^* \rightarrow k$ är den linjära form på V^* som ges av $\gamma(\mathbf{v})(\varphi) = \varphi(\mathbf{v})$, dvs $\gamma(\mathbf{v}) = \text{ev}(-, \mathbf{v})$.

Sats 1.13. Om $\dim V < \infty$ så är $\gamma: V \rightarrow (V^*)^*$ en isomorfi.

Bevis. En bas för $(V^*)^*$ ges av $((\mathbf{b}_i)^*)^*$. Det räcker alltså att visa att $\gamma(\mathbf{b}_i) = ((\mathbf{b}_i)^*)^*$. Vi har att

$$\gamma(\mathbf{b}_i)(\mathbf{b}_j^*) = \mathbf{b}_j^*(\mathbf{b}_i) = \delta_{ij}$$

och alltså $\gamma(\mathbf{b}_i) = ((\mathbf{b}_i)^*)^*$. □

1.4. Duala rum och inre produkter. Om V är ett inre produktrum så ger varje vektor $\mathbf{x}$ en linjär form

$$\langle \mathbf{x} | = \langle \mathbf{x} | - \rangle: V \rightarrow k$$

Om vi väljer en bas $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ så är motsvarande radvektor

$${}_{\mathcal{B}}\langle \mathbf{x} | = [{}_{\mathcal{B}}\langle \mathbf{x} |]_{\mathcal{B}} = [{}_{\mathcal{B}}\langle \mathbf{x} |]_{\mathcal{B}^*}^T = [\langle \mathbf{x} | \mathbf{b}_1 \rangle \quad \langle \mathbf{x} | \mathbf{b}_2 \rangle \quad \dots \quad \langle \mathbf{x} | \mathbf{b}_n \rangle] = ([{}_{\mathcal{B}}\langle \mathbf{x} |]^{\dagger} G_{\mathcal{B}}$$

där $G_{\mathcal{B}} = (\langle \mathbf{b}_i | \mathbf{b}_j \rangle)_{ij}$ är metriken för den inre produkten (en $n \times n$ -matris). Speciellt, om $\mathcal{B}$ är en ortonormal bas så är

$${}_{\mathcal{B}}\langle \mathbf{x} | = ([{}_{\mathcal{B}}\langle \mathbf{x} |]^{\dagger})$$

Sats 1.14. Om V är ett ändligtdimensionellt komplext inre produktrum så är

$$\alpha: V \rightarrow V^*, \quad \mathbf{x} \mapsto \langle \mathbf{x} |$$

en anti-isomorfi. Om V är ett ändligtdimensionellt reellt inre produktrum så är α en isomorfi.

Bevis. Väljer vi en ortonormal bas $\mathcal{B}$ så tar α kolonnvektorn $\mathbf{x}$ på radvektorn $\mathbf{x}^{\dagger}$ och detta är en bijektion. □

¹dvs beror inte på val av bas

Notera 1.15. Om V är ändligtdimensionellt men saknar en inre produkt kan vi välja en bas $\mathcal{B}$ och sedan definiera en isomorfi

$$\beta: V \rightarrow V^*$$

genom $\beta(\sum_i a_i \mathbf{b}_i) = \sum_i a_i \mathbf{b}_i^*$, dvs i termer av vektorer så tar β kolonnvektorn $\mathbf{x}$ på radvektorn $\mathbf{x}^T$. Denna isomorfi *beror dock på valet av bas $\mathcal{B}$* medan α bara beror på valet av inre produkt. Valet av bas $\mathcal{B}$ är dock ekvivalent med att utrusta V med en struktur som inre produktrum: vi väljer den inre produkt som gör $\mathcal{B}$ till en ON-bas. För ett reellt vektorrum är då $\alpha = \beta$. För ett komplext vektorrum är β en isomorfi och α är en anti-isomorfi (den konjugerar koordinaterna).

Notera 1.16. Om $\dim V = \infty$ så är α (och β) bara injektiva.

1.5. Dual avbildning. För alla linjära avbildningar $L: V \rightarrow W$ kan vi definiera en dual avbildning:

Definition 1.17. Om $L: V \rightarrow W$ är en linjär avbildning ges den *duala avbildningen*, $L^*: W^* \rightarrow V^*$ av

$$L^*(\phi)(\mathbf{x}) = \phi(L(\mathbf{x})), \quad \forall \mathbf{x} \in V, \forall \phi \in W^*.$$

Sats 1.18. Om A är matrisen för L med avseende på baser $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ och $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ för V respektive W är A^T matrisen för L^* med avseende på de duala baserna $\{\mathbf{f}_1^*, \mathbf{f}_2^*, \dots, \mathbf{f}_m^*\}$, respektive $\{\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*\}$.

Notera 1.19. Om $\mathbf{v} \in V$ och $\phi \in W^*$ så har vi, med $\mathbf{v}$ och ϕ uttryckta som kolonnvektorer att:

$$[L(\mathbf{v})]_{\mathcal{B}'} = [L]_{\mathcal{B}', \mathcal{B}} [\mathbf{v}]_{\mathcal{B}} = A [\mathbf{v}]_{\mathcal{B}} \quad \text{och} \quad [L^*(\phi)]_{\mathcal{B}^*} = [L^*]_{\mathcal{B}^*, \mathcal{B}'^*} [\phi]_{\mathcal{B}'^*} = A^T [\phi]_{\mathcal{B}'^*}.$$

Eftersom $(A^T \mathbf{x})^T = \mathbf{x}^T A$ kan vi också beräkna $L^*(\phi)$ med radvektorer:

$$[L^*(\phi)]_{\mathcal{B}} = [\phi]_{\mathcal{B}'} [L]_{\mathcal{B}', \mathcal{B}} = [\phi]_{\mathcal{B}'} A$$

där koordinaterna för de linjära formerna alltså är radvektorer.

Exempel 1.20. Låt $D: \mathbb{C}[x]_{\leq 2} \rightarrow \mathbb{C}[x]_{\leq 1}$ vara deriveringsoperatoren: $D(f) = f'$. Då är den duala operatoren

$$D^*: \mathbb{C}[x]_{\leq 1} \rightarrow \mathbb{C}[x]_{\leq 2}, \quad D^*(\phi) = \phi \circ D$$

Om t ex $\phi \in (\mathbb{C}[x]_{\leq 1})^*$ är den linjära formen

$$\phi: \mathbb{C}[x]_{\leq 1} \rightarrow \mathbb{C}, \quad \phi(p) = p'(0)$$

dvs $\phi(a + bx) = b$, så är

$$D^*(\phi) = \phi \circ D$$

den linjära formen $D^*(\phi) \in (\mathbb{C}[x]_{\leq 2})^*$ som ges av

$$D^*(\phi)(a + bx + cx^2) = \phi(b + 2cx) = 2c.$$

Om vi inför baserna $\mathcal{B} = \{1, x, x^2\}$ och $\mathcal{B}' = \{1, x\}$ så är

$$[D]_{\mathcal{B}', \mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad [D^*]_{\mathcal{B}^*, \mathcal{B}'^*} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Vi kan beräkna $D^*(\phi)$ med hjälp av kolonnvektorer

$$[\phi]_{\mathcal{B}'^*} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad [D^*(\phi)]_{\mathcal{B}^*} = [D^*]_{\mathcal{B}^*, \mathcal{B}'^*} [\phi]_{\mathcal{B}'^*} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

eller radvektorer

$$[\varphi]_{\mathcal{B}'} = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad [D^*(\varphi)]_{\mathcal{B}} = [\varphi]_{\mathcal{B}'} [D]_{\mathcal{B}', \mathcal{B}} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \end{bmatrix}$$

Om V och W har inre produkter så är L^* i princip samma sak som $L^\dagger$ om vi identifierar V med V^* och W med W^* . Mer precist har vi följande sats:

Sats 1.21. Låt V och W vara inre produktrum och låt $\alpha_V: V \longrightarrow V^*$ och $\alpha_W: W \longrightarrow W^*$ vara motsvarande anti-isomorfier. Om $L: V \longrightarrow W$ är en linjär avbildning så är

$$L^\dagger = \alpha_V^{-1} \circ L^* \circ \alpha_W$$

dvs, följande diagram kommuterar

$$\begin{array}{ccc} W & \xrightarrow{L^\dagger} & V \\ \downarrow \alpha_W & & \downarrow \alpha_V \\ W^* & \xrightarrow{L^*} & V^* \end{array}$$

Bevis. Om $\mathbf{w} \in W$ så har vi att

$$(\alpha_V \circ L^\dagger)(\mathbf{w}) = \langle L^\dagger(\mathbf{w}) | - \rangle_V$$

och att

$$(L^* \circ \alpha_W)(\mathbf{w}) = L^*(\langle \mathbf{w} | - \rangle_W) = \langle \mathbf{w} | L(-) \rangle_V.$$

Resultatet följer alltså av definitionen av den adjungerade operatören $L^\dagger$. $\square$

Notera 1.22. Låt $L: V \longrightarrow W$ vara en linjär avbildning och välj ortonormala baser $\mathcal{B}$ och $\mathcal{B}'$ för V och W . Om $A = [L]_{\mathcal{B}', \mathcal{B}}$ så är

$$[L^*]_{\mathcal{B}^*, \mathcal{B}'^*} = A^T, \quad \text{och} \quad [L^\dagger]_{\mathcal{B}, \mathcal{B}'} = A^\dagger$$

Matriserna är alltså desamma i det reella fallet och skiljer sig åt med komplexkonjugering i det komplexa fallet. Det senare förklaras av att α bara är en antiisomorfi.

2. TENSORPRODUKT

Den direkta summan av två vektorrum $V \oplus W$ gav en generell operation som ger $k^{m+n} \cong k^m \oplus k^n$. Vi ska nu se på en annan generell operation $\otimes$ som ger $k^{mn} \cong k^m \otimes k^n$. Denna operation kallas *tensorprodukt* av vektorrum och den hör starkt samman med bilinjära avbildningar.

Att göra en strikt definition av tensorprodukten som fungerar även för oändligdimensionella vektorrum är inte så lätt, så vi kommer att börja med en version som bara fungerar för ändligdimensionella vektorrum. Innan dess ska vi se på ett exempel som visar på vad vi vill åstadkomma med konstruktionen.

Inom fysiken och mekaniken har vi ofta tensorprodukter som bygger på $V = \mathbb{R}^3$ eller dess dual, $V^* = (\mathbb{R}^3)^*$. Vektorer skrivs efter ett val av bas $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ som $\sum_{i=1}^3 a_i \mathbf{e}_i$ och man använder ofta det förkortade skrivsättet a_i för att beteckna denna vektor. Andra storheter behöver flera index, exempelvis spänningstensorn i ett material σ_{ij} . Att vi skriver på det viset betyder att det finns ett vektorrum vars basvektorer naturligt är indexerade med två olika index. Från linjär algebra har vi också stött ihop med sådana, nämligen matriser. Dessa kan användas för att beskriva linjära avbildningar, linjära operatorer, kvadratiska former eller inre produkter. Betraktat som tensorer är de två första annorlunda än de två andra och det har att göra med hur de förändras vid basbyten.

När vi ser på en matris som matrisen för en linjär operator på V betyder det att vi har gjort ett val av bas för att representera operatoren som en matris. Gör vi ett annat val av bas får vi en annan matris och vi har lärt oss hur vi behöver ändra koordinaterna för vektorerna och matrisen med hjälp av basbytesmatriser. Om $P_{\mathcal{B}, \mathcal{D}} = (p_{ij})$ ger basbytet från $\mathcal{B}$ till $\mathcal{D}$ ändras koordinaterna a_i med avseende på basen $\mathcal{B}$ till koordinaterna a'_i med avseende på basen $\mathcal{D}$ som

$$a'_i = \sum_{j=1}^n p_{ij} a_j$$

Om vi har att matrisen (a_{ij}) representerar $L: V \rightarrow V$ med avseende på $\mathcal{B}$ får vi matrisen med avseende på $\mathcal{D}$ som

$$a'_{ij} = \sum_{k=1}^n \sum_{\ell=1}^n q_{ik} a_{k\ell} p_{\ell j}$$

där (q_{ij}) är basbytesmatrisen $P_{\mathcal{D}, \mathcal{B}}$, dvs inversmatrisen till (p_{ij}) . Vi ser att det första och det andra indexet i matrisen beter sig på olika sätt här. Vi säger att det ena är *kovariant* i och med att det beter sig på samma sätt som vektorerna med avseende på basbyte, medan det andra indexet är *kontravariant* i och med att vi behöver använda transponatet av inversmatrisen vid basbytet. Ibland väljer man att skilja på index som är kontravarianta genom att sätta dem ovanför istället för nedanför variabelnamnen.

2.1. Konstruktion med hjälp av baser. Ett sätt att definiera tensorprodukten av två vektorrum V och W är genom att välja en bas för vardera, säg $\{\mathbf{e}_i\}_{i \in I}$ för V och $\{\mathbf{f}_j\}_{j \in J}$ för W och sedan skapa $V \times W$ som ett vektorrum med bas $\{\mathbf{e}_i \otimes \mathbf{f}_j\}_{i \in I, j \in J}$, där $\mathbf{e}_i \otimes \mathbf{f}_j$ till att börja med bara är en symbol. Därefter behöver vi också kunna använda $\mathbf{x} \otimes \mathbf{y}$ även för vektorer $\mathbf{x}$ och $\mathbf{y}$ som inte tillhörde våra ursprungliga baser. Detta gör vi genom att kräva att $\otimes$ ska vara *bilinjär* och att vi ska få

$$\mathbf{x} = \sum_{i \in I} a_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{j \in J} b_j \mathbf{f}_j \quad \implies \quad \mathbf{x} \otimes \mathbf{y} = \sum_{i \in I, j \in J} a_i b_j \mathbf{e}_i \otimes \mathbf{f}_j.$$

Därmed har vi utvecklat $\mathbf{x} \otimes \mathbf{y}$ i den bas vi har definierat för $V \otimes W$.

Problemet uppstår när någon annan väljer andra baser för samma vektorrum och vi vill kunna jämföra resultaten. Då måste vi veta hur basbyten på V och W motsvarar

basbyten på $V \otimes W$. Om p_{ij} ger ett basbyte på V och q_{ij} ger ett basbyte på W betyder det att vi har andra baser $\{\mathbf{e}'_i\}_{i \in I}$ för V och $\{\mathbf{f}'_j\}_{j \in J}$ för W med

$$\mathbf{e}'_i = \sum_{k \in I} p_{ik} \mathbf{e}_k, \quad \text{och} \quad \mathbf{f}'_j = \sum_{\ell \in J} q_{j\ell} \mathbf{f}_\ell.$$

Då behöver vi ha att

$$\mathbf{e}'_i \otimes \mathbf{f}'_j = \left(\sum_{k \in I} p_{ik} \mathbf{e}_k \right) \otimes \left(\sum_{\ell \in J} q_{j\ell} \mathbf{f}_\ell \right) = \sum_{k \in I} \sum_{\ell \in J} p_{ik} q_{j\ell} \mathbf{e}_k \otimes \mathbf{f}_\ell.$$

Exempel 2.1. Om vi ser på vektorer i $V = \mathbb{R}^3$ med hjälp av koordinater med avseende på en bas $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ och representerar vektorn $\mathbf{x}$ med dess koordinater x_i där $\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3$. Från fysiken har vi mött tensorn [Kroneckers delta](#), δ_{ij} , som är en tensor som representeras av koordinaterna δ_{ij} med avseende på någon bas. Frågan är vilken och för vilket rum. Vi kan förstå det genom att se på tensorprodukten $V \otimes V$ som har bas $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{i=1, j=1}^{3,3}$ bestående av nio basvektorer. Kroneckers delta är då tensorn

$$\delta = \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3.$$

eftersom koordinaterna är ett för dessa tre basvektorer och noll för övriga $\mathbf{e}_i \otimes \mathbf{e}_j$ där $i \neq j$. Vi har lärt oss att denna ska vara invariant under basbyten (koordinatbyten) och vi ska undersöka det. Vi väljer en ny bas $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ som ges av

$$\mathbf{e}_i = \sum_{j=1}^3 p_{ij} \mathbf{f}_j$$

och får då uttrycket för Kroneckers delta i den nya basen som

$$\begin{aligned} \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 &= \left(\sum_{j=1}^3 p_{1j} \mathbf{f}_j \right) \otimes \left(\sum_{k=1}^3 p_{1k} \mathbf{f}_k \right) + \left(\sum_{j=1}^3 p_{2j} \mathbf{f}_j \right) \otimes \left(\sum_{k=1}^3 p_{2k} \mathbf{f}_k \right) \\ &\quad + \left(\sum_{j=1}^3 p_{3j} \mathbf{f}_j \right) \otimes \left(\sum_{k=1}^3 p_{3k} \mathbf{f}_k \right) \\ &= \sum_{j=1}^3 \sum_{k=1}^3 (p_{1j} p_{1k} + p_{2j} p_{2k} + p_{3j} p_{3k}) \mathbf{f}_j \otimes \mathbf{f}_k \end{aligned}$$

Här ser vi att koordinaterna med avseende på den nya basen $\{\mathbf{f}_i \otimes \mathbf{f}_j\}$ för $V \otimes V$ blir skalärprodukterna av kolonnvektorerna i basbytesmatrisen (p_{ij}) och Kroneckers delta får samma koordinater, δ_{ij} med avseende på denna bas precis om kolonnerna i basbytesmatrisen utför en ortonormal bas för $\mathbb{R}^3$, dvs om basbytet är en ortogonal transformation.

Om vi vill att Kroneckers delta ska ha samma koordinater med avseende på alla koordinatbyten kan vi inte betrakta den som en tensor i $V \otimes V$, men istället i $V^* \otimes V$. Då får vi att basbytet på V^* ges av $\mathbf{e}^*_i = \sum_{j=1}^3 q_{ij} \mathbf{f}^*_j$ där

$$\delta_{ij} = \mathbf{e}^*_i(\mathbf{e}_j) = \sum_{k=1}^3 \sum_{\ell=1}^3 q_{ik} p_{j\ell} \mathbf{f}^*_k(\mathbf{f}_\ell) = \sum_{k=1}^3 \sum_{\ell=1}^3 q_{ik} p_{j\ell} \delta_{jk} = \sum_{k=1}^3 q_{ik} p_{jk}$$

vilket säger att matrisen (q_{ij}) är inversen av transponatet av matrisen (p_{ij}) . När vi nu gör om beräkningen ovan får vi

$$\begin{aligned} \mathbf{e}_1^* \otimes \mathbf{e}_1 + \mathbf{e}_2^* \otimes \mathbf{e}_2 + \mathbf{e}_3^* \otimes \mathbf{e}_3 &= \left(\sum_{j=1}^3 q_{1j} \mathbf{f}_j^* \right) \otimes \left(\sum_{k=1}^3 p_{1k} \mathbf{f}_k \right) + \left(\sum_{j=1}^3 q_{2j} \mathbf{f}_j^* \right) \otimes \left(\sum_{k=1}^3 p_{2k} \mathbf{f}_k \right) \\ &\quad + \left(\sum_{j=1}^3 q_{3j} \mathbf{f}_j^* \right) \otimes \left(\sum_{k=1}^3 p_{3k} \mathbf{f}_k \right) \\ &= \sum_{j=1}^3 \sum_{k=1}^3 (q_{1j} p_{1k} + q_{2j} p_{2k} + q_{3j} p_{3k}) \mathbf{f}_j^* \otimes \mathbf{f}_k \\ &= \sum_{j=1}^3 \sum_{k=1}^3 \sum_{\ell=1}^3 q_{\ell j} p_{\ell k} \mathbf{f}_j^* \otimes \mathbf{f}_k \\ &= \sum_{j=1}^3 \sum_{k=1}^3 \delta_{j,k} \mathbf{f}_j^* \otimes \mathbf{f}_k = \mathbf{f}_1^* \otimes \mathbf{f}_1 + \mathbf{f}_2^* \otimes \mathbf{f}_2 + \mathbf{f}_3^* \otimes \mathbf{f}_3. \end{aligned}$$

Exempel 2.2. Om $V = \mathbb{C}[x]$ är vektorrummet av polynom med komplexa koefficienter i en variabel, x , har vi en naturlig bas som ges av $\mathbf{e}_i = x^i$, för $i \in \mathbb{N}$. Det betyder att $V \otimes V$ får en bas som ges av $\mathbf{e}_i \otimes \mathbf{e}_j$, där $i, j \in \mathbb{N}$. Vi kan tolka dessa basvektorer som polynomen $x^i y^j$ som ligger i $W = \mathbb{C}[x, y]$, dvs vektorrummet av polynom med komplexa koefficienter i två variabler, x och y .

Om vi byter bas till $\mathbf{e}'_i = (x+1)^i = \sum_{i'=0}^i \binom{i}{i'} x^{i'} = \sum_{i'=0}^i \binom{i}{i'} \mathbf{e}_{i'}$, $i \in \mathbb{N}$ på V får vi motsvarande basbyte på $W = V \otimes V$ som

$$\mathbf{e}'_i \otimes \mathbf{e}'_j = \left(\sum_{i'=0}^i \binom{i}{i'} \mathbf{e}_{i'} \right) \otimes \left(\sum_{j'=0}^j \binom{j}{j'} \mathbf{e}_{j'} \right) = \sum_{i'=0}^i \sum_{j'=0}^j \binom{i}{i'} \binom{j}{j'} \mathbf{e}_{i'} \otimes \mathbf{e}_{j'}$$

som motsvarar $(x+1)^i (y+1)^j$ i $W = \mathbb{C}[x, y]$.

Exempel 2.3. Om vi ser $V = k^n$ som vektorrummet av funktioner $f: \{1, 2, \dots, n\} \rightarrow k$ har vi en naturlig bas $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ som ges av $\mathbf{e}_i(j) = \delta_{ij}$, för $1 \leq i, j \leq n$. Vi kan nu se basvektoren $\mathbf{e}_i \otimes \mathbf{e}_j$ för $V \otimes V$ som en funktion på produktmängden, $\{1, 2, \dots, n\} \times \{1, 2, \dots, n\}$ och vi får

$$\mathbf{e}_i \otimes \mathbf{e}_j(i', j') = \delta_{ii'} \delta_{jj'} = \delta_{(i,j)(i',j')}, \quad 1 \leq i, j, i', j' \leq n.$$

Vi kan sedan skriva alla funktioner $f: \{1, 2, \dots, n\} \times \{1, 2, \dots, n\} \rightarrow k$ som linjärkombinationer av dessa basvektorer. Därmed kan vi tolka tensorprodukten som vektorrummet av funktioner på produkten, dvs

$$k^n \otimes k^n = V \otimes V = \left\{ f: \{1, 2, \dots, n\} \times \{1, 2, \dots, n\} \rightarrow k \right\}.$$

Notera 2.4. Vi kan också se vektorrummet $M_n(k)$ av $n \times n$ -matriser med element i k som funktioner $f: \{1, 2, \dots, n\} \times \{1, 2, \dots, n\} \rightarrow k$, men vi har sett tidigare att ett basbyte på V inte ger det önskade basbytet på $M_n(k)$ om vi vill att matriserna ska representera linjära operatorer på V . Däremot kan vi se $M_n(k)$ som tensorprodukten $(k^n)^* \otimes k^n$.

2.2. Konstruktion med bilinjära avbildningar. Det andra sättet vi kan konstruera tensorprodukten av två vektorrum är som ett vektorrum av bilinjära avbildningar. Fördelen med den här konstruktionen är dels att den är kortare att formulera, dels att den inte beror på något val av bas utan är *kanonisk*. Nackdelen är att den bara fungerar om vektorrummen är ändligdimensionella. Detta har att göra med att dualen av ett oändligdimensionellt vektorrum inte är isomorft med vektorrummet självt utan har en bas som har större kardinalitet än basen för vektorrummet.

Definition 2.5. En avbildning $f: V_1 \times V_2 \rightarrow W$ är *bilinjär* om

- $\mathbf{y} \mapsto f(\mathbf{x}, \mathbf{y})$ är linjär för varje $\mathbf{x} \in V_1$
- $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{y})$ är linjär för varje $\mathbf{y} \in V_2$

En avbildning $f: V_1 \times V_2 \times \cdots \times V_k \longrightarrow W$ är **multilinjär** om

- $\mathbf{x} \mapsto f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}, \dots, \mathbf{x}_n)$ är linjär när vi fixerar alla utom ett av argumenten.

Definition 2.6. Om V och W är ändligdimensionella vektorrum ges **tensorprodukten** $V \otimes W$ av

$$V \otimes W = \left\{ f: V^* \times W^* \longrightarrow k \mid f \text{ är bilinjär} \right\}$$

Notera 2.7. Varje element $(\mathbf{x}, \mathbf{y})$ i $V \times W$ ger en bilinjär avbildning $f: V^* \times W^* \longrightarrow k$ genom $f(\phi, \psi) = \phi(\mathbf{x})\psi(\mathbf{y})$. Detta element i $V \otimes W$ skrivs som $\mathbf{x} \otimes \mathbf{y}$. Eftersom V och W är ändligdimensionella kommer alla tensorer i $V \otimes W$ att vara linjärkombinationer av sådana tensorer. Att vi ser på bilinjära avbildningar från $V^* \times W^*$ och inte från $V \times W$ beror på att vi vill ha kovarians, dvs att basbyten på V och W ska leda till ett basbyte på $V \otimes W$ som använder samma basbytesmatriser och inte transponatet av deras inverser. De bilinjära avbildningarna $V \times W \longrightarrow k$ ger istället tensorprodukten av de duala rummen, $V^* \otimes W^*$.

Sats 2.8. Om $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ och $\mathcal{D} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ är baser för V och W ger $\{\mathbf{e}_i \otimes \mathbf{f}_j\}_{i=1, j=1}^{m, n}$ en bas för $V \otimes W$.

Bevis. En bilinjär avbildning $f: V^* \times W^* \longrightarrow k$ bestäms entydigt av värdet på $(\mathbf{e}_i^*, \mathbf{f}_j^*)$, $1 \leq i \leq m$, $1 \leq j \leq n$ eftersom bilinjäriteten ger

$$f\left(\sum_{i=1}^m a_i \mathbf{e}_i^*, \sum_{j=1}^n b_j \mathbf{f}_j^*\right) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j f(\mathbf{e}_i^*, \mathbf{f}_j^*)$$

Vi kan låta $\mathbf{e}_i \otimes \mathbf{f}_j$ vara den bilinjära funktion som uppfyller

$$\mathbf{e}_i \otimes \mathbf{f}_j(\mathbf{e}_{i'}^*, \mathbf{f}_{j'}^*) = \delta_{ii'} \delta_{jj'}.$$

och får att den bilinjära avbildning $f: V^* \times W^* \longrightarrow k$ som ges av $f(\mathbf{e}_i^*, \mathbf{f}_j^*) = a_{ij}$ kan skrivas som

$$f = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \mathbf{e}_i \otimes \mathbf{f}_j.$$

□

Sats 2.9. $\text{Hom}_k(V, W) \cong V^* \otimes W$.

Bevis. Välj baser för $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ och $\mathcal{D} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ för V och W . Vi får att $\{\mathbf{e}_i^* \otimes \mathbf{f}_j\}$ bildar en bas för $V^* \otimes W$. Dessa element svarar mot linjära funktioner $L_{ij}: V \longrightarrow W$ genom $L_{ij}(\mathbf{x}) = \mathbf{e}_i^*(\mathbf{x})\mathbf{f}_j$. Detta ger en bijektiv linjär avbildning eftersom varje linjär funktion kan skrivas som

$$\sum_{i,j} a_{ji} L_{ij}$$

där a_{ij} är matrisen för L med avseende på de givna baserna. □

Sats 2.10. Den naturliga avbildningen $V \times W \longrightarrow V \otimes W$ är bilinjär och varje annan bilinjär avbildning $V \times W \longrightarrow U$ faktoriseras unikt genom den.

$$\begin{array}{ccc} V \times W & \xrightarrow{L} & V \otimes W \\ & \searrow f & \downarrow \exists! F \\ & & U \end{array}$$

Bevis. Den avbildning $L: V \times W \longrightarrow V \otimes W$ som ges av $L(\mathbf{x}, \mathbf{y})(\phi, \psi) = \phi(\mathbf{x})\psi(\mathbf{y})$ är bilinjär och kan skrivas som

$$\left(\sum_{i=1}^m a_i \mathbf{e}_i, \sum_{i=1}^n b_i \mathbf{f}_i \right) \mapsto \sum_{i=1}^m \sum_{i=1}^n a_i b_i \mathbf{e}_i \otimes \mathbf{f}_i$$

efter val av baser för V och W . Om $f: V \times W \longrightarrow U$ är en bilinjär avbildning kan vi använda baser för V och W för att definiera en linjär avbildning $F: V \otimes W \longrightarrow U$ genom

$$F(\mathbf{e}_i \otimes \mathbf{f}_j) = f(\mathbf{e}_i, \mathbf{f}_j)$$

Denna avbildning uppfyller $f = F \circ L$ och den är unik i och med att om F' också uppfyller det är $F - F'$ noll på basen $\{\mathbf{e}_i \otimes \mathbf{f}_j\}$ och därmed noll som linjär avbildning. $\square$

Notera 2.11. Satsen säger att tensorprodukten har en *universell* egenskap med avseende på bilinjära avbildningar. Det går att definiera tensorprodukten med hjälp av denna egenskap och vi kan då göra det utan begränsning på dimensionerna. Vi säger då att $V \otimes W$ tillsammans med den bilinjära avbildningen $V \times W \longrightarrow V \otimes W$ är definierade av att varje bilinjär avbildning $V \times W \longrightarrow U$ faktoriserar unikt via $V \otimes W$.

Definition 2.12. Vi låter *spåravbildningen*, $\text{tr}: V^* \otimes V \longrightarrow k$, vara den linjära avbildning som kommer från den bilinjära avbildningen $ev: V^* \times V \longrightarrow k$, där $ev(\phi, \mathbf{x}) = \phi(\mathbf{x})$, för $\mathbf{x} \in V$ och $\phi \in V^*$.

2.3. Tensorprodukt av avbildningar och matriser. Om vi har två linjära avbildningar $L_1: V_1 \longrightarrow W_1$ och $L_2: V_2 \longrightarrow W_2$ kan vi använda Sats 2.10 till att konstruera en linjär avbildning

$$L_1 \otimes L_2: V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2.$$

För att göra det räcker det att tala om hur vi kan definiera en bilinjär avbildning $\Psi: V_1 \times V_2 \longrightarrow W_1 \otimes W_2$ och det kan vi göra genom $\Psi(\mathbf{x}, \mathbf{y}) = L_1(\mathbf{x}) \otimes L_2(\mathbf{y})$. Det är klart att denna blir bilinjär eftersom både L_1 och L_2 är linjära avbildningar. Därmed finns en unik linjär avbildning $L_1 \otimes L_2: V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2$ så att $(L_1 \otimes L_2)(\mathbf{x} \otimes \mathbf{y}) = L_1(\mathbf{x}) \otimes L_2(\mathbf{y})$, för alla $\mathbf{x} \in V_1$ och alla $\mathbf{y} \in V_2$.

Exempel 2.13. Om $L_1: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ och $L_2: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ ges av matriserna

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \quad \text{och} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

kan vi se på matrisen för $L_1 \otimes L_2$ genom att se hur den verkar på basen $\{\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_1 \otimes \mathbf{e}_2, \mathbf{e}_2 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2\}$ och får

$$\begin{aligned} \mathbf{e}_1 \otimes \mathbf{e}_1 &\mapsto (\mathbf{e}_1 + 4\mathbf{e}_2) \otimes \mathbf{e}_1 = 1 \cdot \mathbf{e}_1 \otimes \mathbf{e}_1 + 0 \cdot \mathbf{e}_1 \otimes \mathbf{e}_2 + 4 \cdot \mathbf{e}_2 \otimes \mathbf{e}_1 + 0 \cdot \mathbf{e}_2 \otimes \mathbf{e}_2 \\ \mathbf{e}_1 \otimes \mathbf{e}_2 &\mapsto (\mathbf{e}_1 + 4\mathbf{e}_2) \otimes (\mathbf{e}_1 + \mathbf{e}_2) = 1 \cdot \mathbf{e}_1 \otimes \mathbf{e}_1 + 4 \cdot \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_1 \otimes \mathbf{e}_2 + 4 \cdot \mathbf{e}_2 \otimes \mathbf{e}_2 \\ \mathbf{e}_2 \otimes \mathbf{e}_1 &\mapsto (2\mathbf{e}_1 + 3\mathbf{e}_2) \otimes \mathbf{e}_1 = 2 \cdot \mathbf{e}_1 \otimes \mathbf{e}_1 + 0 \cdot \mathbf{e}_1 \otimes \mathbf{e}_2 + 3 \cdot \mathbf{e}_2 \otimes \mathbf{e}_1 + 0 \cdot \mathbf{e}_2 \otimes \mathbf{e}_2 \\ \mathbf{e}_2 \otimes \mathbf{e}_2 &\mapsto (2\mathbf{e}_1 + 3\mathbf{e}_2) \otimes (\mathbf{e}_1 + \mathbf{e}_2) = 2 \cdot \mathbf{e}_1 \otimes \mathbf{e}_1 + 3 \cdot \mathbf{e}_2 \otimes \mathbf{e}_1 + 2 \cdot \mathbf{e}_1 \otimes \mathbf{e}_2 + 3 \cdot \mathbf{e}_2 \otimes \mathbf{e}_2 \end{aligned}$$

Därmed får vi matrisen för $L_1 \otimes L_2$ som

$$C = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 0 & 2 \\ 4 & 4 & 3 & 3 \\ 0 & 4 & 0 & 3 \end{bmatrix} = \begin{bmatrix} B & 2B \\ 4B & 3B \end{bmatrix}$$

och om vi istället hade valt basen i en annan ordning $\{\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_1, \mathbf{e}_1 \otimes \mathbf{e}_2, \mathbf{e}_2 \otimes \mathbf{e}_2\}$ får vi

$$C' = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 4 & 3 & 4 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 4 & 3 \end{bmatrix} = \begin{bmatrix} A & A \\ 0 & A \end{bmatrix}$$

Byter vi bas för den första faktorn till $\{\mathbf{e}'_1, \mathbf{e}'_2\}$ som är en egenbas för L_1 får vi istället matrisen

$$C = \begin{bmatrix} 5 & 5 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad \text{eller} \quad C' = \begin{bmatrix} 5 & 0 & 5 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

och vi ser att egenvärdena till $L_1 \otimes L_2$ är 5 och -1 med algebraisk multiplicitet två och geometrisk multiplicitet ett. Alltså är $L_1 \otimes L_2$ inte diagonaliserbar.

Notera 2.14. Den form av multiplikation av matriser som vi ser i exemplet kallas ibland Kroneckers tensorprodukt av matriser och i t ex Matlab är kommandot för detta `kron(A,B)`.

Notera 2.15. I allmänhet gäller att om λ är egenvärde till L_1 och μ är egenvärde till L_2 så är $\lambda\mu$ egenvärde till $L_1 \otimes L_2$ eftersom $L_1(\mathbf{x}) = \lambda\mathbf{x}$ och $L_2(\mathbf{y}) = \mu\mathbf{y}$ ger $L_1 \otimes L_2(\mathbf{x} \otimes \mathbf{y}) = \lambda\mathbf{x} \otimes \mu\mathbf{y} = \lambda\mu(\mathbf{x} \otimes \mathbf{y})$. Om både L_1 och L_2 är diagonaliserbara med egenvärden $\lambda_1, \lambda_2, \dots, \lambda_m$, respektive $\mu_1, \mu_2, \dots, \mu_n$ så är $L_1 \otimes L_2$ diagonaliserbar med egenvärden $\lambda_i\mu_j$, $1 \leq i \leq m$, $1 \leq j \leq n$.

Exempel 2.16. Om A och B är stokastiska matriser som beskriver olika oberoende Markovkedjor med m , respektive n olika tillstånd kan vi definiera den sammanslagna Markovprocessen som består av alla mn kombinationer av tillstånd. Den stokastiska matrisen som beskriver denna process är en $mn \times mn$ -matris och den ges av tensorprodukten av matriserna A och B enligt exemplet ovan. Tensorprodukten av stokastiska matriser blir på nytt en stokastisk matris.

2.4. Konstruktion som kvot. Ett ytterligare sätt att konstruera tensorprodukten är som som en kvot av ett vektorrum på följande sätt. Bilda

$$U = \left\{ f: V \times W \longrightarrow k \mid f(\mathbf{x}, \mathbf{y}) \neq 0 \text{ för ett ändligt antal } (\mathbf{x}, \mathbf{y}) \right\}.$$

som har en bas som består av de karakteristiska funktionerna $\chi_{\mathbf{x}_0, \mathbf{y}_0}$ som uppfyller

$$\chi_{\mathbf{x}_0, \mathbf{y}_0}(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & \text{om } (\mathbf{x}, \mathbf{y}) = (\mathbf{x}_0, \mathbf{y}_0), \\ 0 & \text{annars} \end{cases}$$

Poängen är att U är ett vektorrum där basvektorerna på ett naturligt sätt svarar mot elementen i $V \times W$. Vi ska nu definiera ett delrum som består av alla *relationer* vi vill ska finnas mellan tensorerna i $V \otimes W$. Låt R_1, R_2, R_3 och R_4 vara de delrum av U som ges av

$$\begin{aligned} R_1 &= \text{Span}\{a\chi_{\mathbf{x}_0, \mathbf{y}_0} - \chi_{a\mathbf{x}_0, \mathbf{y}_0} : (\mathbf{x}_0, \mathbf{y}_0) \in V \times W, a \in k\} \\ R_2 &= \text{Span}\{a\chi_{\mathbf{x}_0, \mathbf{y}_0} - \chi_{\mathbf{x}_0, a\mathbf{y}_0} : (\mathbf{x}_0, \mathbf{y}_0) \in V \times W, a \in k\} \\ R_3 &= \text{Span}\{\chi_{\mathbf{x}_0 + \mathbf{x}_1, \mathbf{y}_0} - \chi_{\mathbf{x}_0, \mathbf{y}_0} - \chi_{\mathbf{x}_1, \mathbf{y}_0} : (\mathbf{x}_0, \mathbf{y}_0), (\mathbf{x}_1, \mathbf{y}_0) \in V \times W\} \\ R_4 &= \text{Span}\{\chi_{\mathbf{x}_0, \mathbf{y}_0 + \mathbf{y}_1} - \chi_{\mathbf{x}_0, \mathbf{y}_0} - \chi_{\mathbf{x}_0, \mathbf{y}_1} : (\mathbf{x}_0, \mathbf{y}_0), (\mathbf{x}_0, \mathbf{y}_1) \in V \times W\} \end{aligned}$$

Vi kan nu definiera $R = R_1 + R_2 + R_3 + R_4$, dvs alla vektorer som kan skrivas som en summa av vektorer från de fyra delrummen. Nu kan vi definiera $V \otimes W = U/R$ med den naturliga bilinjära avbildningen $L: V \times W \longrightarrow V \otimes W$ som ges av $L(\mathbf{x}, \mathbf{y}) = \chi_{\mathbf{x}, \mathbf{y}} + R$

Denna blir bilinjär precis för att delrummen R_1, R_2, R_3 och R_4 är konstruerade så. Det återstår nu att visa att denna har samma egenskaper som det vi visat för tensorprodukten tidigare. Vi skriver som tidigare $\mathbf{x} \otimes \mathbf{y}$ för $L(\mathbf{x}, \mathbf{y}) = \chi_{\mathbf{x}, \mathbf{y}} + R$ i $V \times W$.

Definition 2.17. Om $V_1, V_2, \dots, V_n$ är ändligdimensionella vektorrum är tensorprodukten

$$V_1 \otimes V_2 \otimes \cdots \otimes V_n = \left\{ f: V_1^* \times V_2^* \times \cdots \times V_n^* \longrightarrow k \mid f \text{ är multilinjär} \right\}$$

Notera 2.18. På samma sätt som $V \otimes W$ kommer $V_1 \otimes V_2 \otimes \cdots \otimes V_n$ tillsammans med den multilinjära avbildningen $V_1 \times V_2 \times \cdots \times V_n \longrightarrow V_1 \otimes V_2 \otimes \cdots \otimes V_n$ att ha en universell egenskap med avseende på multilinjära avbildningar $V_1 \times V_2 \times \cdots \times V_n \longrightarrow U$. Dessutom kan man visa att det finns naturliga isomorfier, som exempelvis

$$(V_1 \otimes V_2) \otimes V_3 \xrightarrow{\sim} V_1 \otimes V_2 \otimes V_3 \xleftarrow{\sim} V_1 \otimes (V_2 \otimes V_3)$$

Sats 2.19. *Sammansättning av operatorer i $\text{Hom}_k(V, V) \cong V^* \otimes V$ svarar mot den naturliga avbildningen*

$$V^* \otimes V \otimes V^* \otimes V \longrightarrow V^* \otimes V$$

där spåravbildningen, tr används på de två yttersta faktorerna.

Bevis. Det är naturligare att se detta om vi skriver $\text{Hom}_k(V, V) \cong V \otimes V^*$ då vi får sammansättningen

$$\text{Hom}_k(V, V) \times \text{Hom}_k(V, V) \longrightarrow \text{Hom}_k(V, V)$$

leder till den bilinjära avbildningen

$$V \otimes V^* \times V \otimes V^* \longrightarrow V \otimes V^*$$

som på grund av den universella egenskapen ger en linjär avbildning

$$V \otimes V^* \otimes V \otimes V^* \longrightarrow V \otimes V^*$$

och det är nu de mittersta två där spåravbildningen används i och med att resultatet från den högra avbildningen går in som argument i den vänstra vid sammansättningen. $\square$

Tensorprodukten beter sig också väl tillsammans med direkt summa.

Sats 2.20. $(V_1 \oplus V_2) \otimes W = (V_1 \otimes W) \oplus (V_2 \otimes W)$.

Bevis. Vi kan definiera en avbildning $(V_1 \oplus V_2) \times W \longrightarrow (V_1 \otimes W) \oplus (V_2 \otimes W)$ genom

$$f(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = (\mathbf{x}_1 \otimes \mathbf{y}, \mathbf{x}_2 \otimes \mathbf{y})$$

På grund av den universella egenskapen kommer det att finnas en unik linjär avbildning

$$(V_1 \oplus V_2) \otimes W \longrightarrow (V_1 \otimes W) \oplus (V_2 \otimes W).$$

som stämmer ihop med f . Att detta blir en isomorfi kan vi kontrollera genom att välja baser för V_1, V_2 och W och se att basen för vänsterledet avbildas på basen för högerledet. $\square$

Definition 2.21. Delrummet $\text{Sym}^2(V) \subseteq V \otimes V$ av *symmetriska* tensorer ges av

$$\text{Sym}^2(V) = \text{Span}\{\mathbf{x} \otimes \mathbf{x} : \mathbf{x} \in V\}.$$

Delrummet $\text{Alt}^2(V) \subseteq V \otimes V$ av *alternerande* tensorer ges av

$$\text{Alt}^2(V) = \text{Span}\{\mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x} : \mathbf{x}, \mathbf{y} \in V\}.$$

Notera 2.22. De symmetriska tensorerna kan också beskrivas som de tensorer $\alpha \in V \otimes V$ så att $\tau(\alpha) = \alpha$ om $\tau: V \otimes V \longrightarrow V \otimes V$ är isomorfin som svarar mot den bilinjära avbildningen $f(\mathbf{x}, \mathbf{y}) = \mathbf{y} \otimes \mathbf{x}$. De *anti-symmetriska* tensorerna är då de tensorer $\alpha \in V \otimes V$ så att $\tau(\alpha) = -\alpha$. Om $1 + 1 \neq 0$ så är detta precis de alternerande tensorerna.

Sats 2.23. Om $1 + 1 \neq 0$ är $V \otimes V = \text{Sym}^2(V) \oplus \text{Alt}^2(V)$.

Notera 2.24. Om $1 + 1 = 0$, till exempel om $k = \mathbb{Z}_2$, så är $1 = -1$ och $\text{Alt}^2(V) \subset \text{Sym}^2(V)$ ty

$$\mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x} = \mathbf{x} \otimes \mathbf{y} + \mathbf{y} \otimes \mathbf{x} = (\mathbf{x} + \mathbf{y}) \otimes (\mathbf{x} + \mathbf{y}) - \mathbf{x} \otimes \mathbf{x} - \mathbf{y} \otimes \mathbf{y} \in \text{Sym}^2(V).$$

ÖVNINGAR PÅ TENSORPRODUKT

Uppgift 1. Låt $V = \mathbb{R}^3$ och låt $L: V \otimes V \rightarrow V$ vara den linjära avbildningen som ges av den bilinjära avbildningen $f: V \times V \rightarrow V$ som ges av kryssprodukten, dvs:

$$f(v, w) = v \times w, \quad \text{för alla } v, w \in V.$$

Bestäm en bas för $\ker L$.

Uppgift 2. Visa att $w := 1 \otimes 1 + x \otimes y \in \mathbb{C}[x] \otimes \mathbb{C}[y]$ inte är en ren tensor, dvs inte går att skriva som $w = p(x) \otimes q(y)$.

Uppgift 3. Kontrollera att avbildningen $\mathbf{x} \otimes \mathbf{y}: V^* \times W^* \rightarrow k$ som ges av

$$(\mathbf{x} \otimes \mathbf{y})(\phi, \psi) = \phi(\mathbf{x})\psi(\mathbf{y}), \quad \forall (\phi, \psi) \in V^* \times W^*$$

verkligen är bilinjär.

Uppgift 4. Visa att $\text{tr}: V^* \otimes V \rightarrow k$ svarar mot spåret av matrisen för avbildningar i $V^* \otimes V \cong \text{Hom}_k(V, V)$ oberoende av val av bas för V så länge den duala basen väljs för V^* .

Uppgift 5. Låt $\{\mathbf{e}_i\}$ och $\{\mathbf{f}_i\}$ vara baser för V där basbytet ges av matrisen $A = (a_{ij})$ med $\mathbf{f}_i = \sum_{j=1}^n a_{ji} \mathbf{e}_j$.

(a) Vilken matris $B = (b_{ij})$ ger basbytet mellan de duala baserna $\{\mathbf{e}_i^*\}$ och $\{\mathbf{f}_i^*\}$ för V^* ?

(b) Hur ser basbytet ut på $V^* \otimes V$?

Uppgift 6. Bevisa Sats 2.23, dvs att $V \otimes V = \text{Sym}^2(V) \oplus \text{Alt}^2(V)$ om $1 + 1 \neq 0$ i k .

Uppgift 7. Om V är ett reellt vektorrum kan vi bilda ett komplext vektorrum genom $V_{\mathbb{C}} = V \otimes \mathbb{C}$ där vi ser $\mathbb{C}$ som ett tvådimensionellt reellt vektorrum. Kontrollera att $V_{\mathbb{C}}$ uppfyller definitionen för ett vektorrum över $\mathbb{C}$ om vi använder multiplikation med skalär på den andra faktorn i tensorprodukten.

Uppgift 8 (Uppgift 7 på modelltentamen). Låt $L: V \rightarrow V$ vara en operator på ett reellt vektorrum som med basen $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ för V ges av matrisen

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

Med hjälp av L får vi en avbildning $L \otimes L^*: V \otimes V^* \rightarrow V \otimes V^*$ från den bilinjära avbildningen $V \times V^* \rightarrow V \otimes V^*$ som ges av $(\mathbf{x}, \phi) \mapsto L(\mathbf{x}) \otimes L^*(\phi)$, för $\mathbf{x} \in V$ och $\phi \in V^*$.

(a) Bestäm matrisen för $L \otimes L^*$ med avseende på basen $\{\mathbf{e}_i \otimes \mathbf{e}_j^*\}$. (2 p)

- (b) Bestäm $\det(L \otimes L^*)$. (Ledning: $\det(L) = 7$.) (2 p)

Uppgift 9 (Uppgift 7 på tentamen 2018-01-10). Låt $V \subseteq \mathbb{C}^3$ vara delrummet som ges av ekvationen $x_1 + x_2 + x_3 = 0$ och betrakta tensorprodukten $V \otimes V$ som ett delrum i $\mathbb{C}^3 \otimes \mathbb{C}^3$. Avgör om någon av tensorerna

$$\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_2 \otimes \mathbf{e}_1 - 2\mathbf{e}_3 \otimes \mathbf{e}_2$$

eller

$$\mathbf{e}_1 \otimes \mathbf{e}_1 - \mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2 - \mathbf{e}_3 \otimes \mathbf{e}_3$$

ligger i $V \otimes V$, där $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ är standardbasvektorerna för $\mathbb{C}^3$. (4 p)

Uppgift 10 (Uppgift 5 på tentamen 2018-04-05). Låt V och W vara ändligdimensionella vektorrum.

- (a) Definiera tensorprodukten $V \otimes W$. (1 p)
- (b) Visa att $\dim(V \otimes W) = (\dim V)(\dim W)$. (1 p)
- (c) Visa att det finns en naturlig bilinjär avbildning $\Phi: V \times W \rightarrow V \otimes W$. (1 p)
- (d) Tensorprodukten $V \otimes W$ har tillsammans med den bilinjära avbildningen från del (c) en universell egenskap med avseende på bilinjära avbildningar $V \times W \rightarrow U$, nämligen att alla sådana faktoriserar på ett entydigt sätt via Φ . Förklara vad detta innebär och skissera ett bevis för att detta gäller. (1 p)

Uppgift 11 (Uppgift 7 på tentamen 2018-04-05). För en linjär operator L på V kan vi bilda en linjär operator på $V \otimes V$ genom $I \otimes L - L \otimes I$, där I är identitetsoperatoren på V .

- (a) Bestäm rangen för $I \otimes L - L \otimes I$ om L ges av matrisen (2 p)

$$\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}.$$

- (b) Visa att $I \otimes L - L \otimes I$ är nilpotent om L är nilpotent, men att det omvända inte alltid gäller. (2 p)

3. YTTRE POTENSER OCH YTTRE ALGEBRAN

I detta avsnitt definierar vi yttre potenser $\bigwedge^\ell V$ och den yttre algebran $\bigwedge V$. På samma sätt som tensorpotensen $V^{\otimes \ell} = V \otimes V \otimes \cdots \otimes V$ har en universell egenskap för multilinjära avbildningar så har den yttre potensen en universell egenskap för *alternerande* avbildningar.

3.1. Alternerande avbildningar.

Definition 3.1. En multilinjär avbildning $f: V^\ell = V \times V \times \cdots \times V \longrightarrow U$ är

- *alternerande* om $f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\ell) = 0$ om två argument är lika, dvs $\mathbf{x}_i = \mathbf{x}_j$ för något $i \neq j$,
- *skevsymmetrisk* (eller antisymmetrisk) om $f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\ell)$ byter tecken, men i övrigt inte ändras, om två argument byter plats.

Exempel 3.2. Kryssprodukten i $\mathbb{R}^3$, dvs avbildningen $f: \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ som ges av $f(\mathbf{v}, \mathbf{w}) = \mathbf{v} \times \mathbf{w}$ är alternerande och skevsymmetrisk. Ty $\mathbf{v} \times \mathbf{v} = \mathbf{0}$ och $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$.

Exempel 3.3. Determinanten i $\mathbb{R}^n$, dvs avbildningen $f: \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ som ges av

$$f(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = \det [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]$$

är multilinjär, alternerande och skevsymmetrisk.

Exempel 3.4. Avbildningen $f: M_n \times M_n \longrightarrow M_n$ definierad av $f(A, B) = AB - BA$, dvs kommutatorn för $n \times n$ -matriser, är bilinjär, alternerande och skevsymmetrisk.

Sats 3.5. Varje alternerande avbildning är skevsymmetrisk. Om $2 \neq 0$ i kroppen av skalärer så gäller omvändningen.

Bevis. Genom att fixera alla argument utom två så erhåller vi en bilinjär avbildning $g: V \times V \longrightarrow U$ och det räcker att bevisa satsen för alla sådana g . Om g är alternerande har vi att

$$\mathbf{0} = g(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) - g(\mathbf{x}, \mathbf{x}) - g(\mathbf{y}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x})$$

för alla $\mathbf{x}, \mathbf{y} \in V$ och alltså är g skevsymmetrisk.

Om g är skevsymmetrisk har vi för alla $\mathbf{x} \in V$ att

$$g(\mathbf{x}, \mathbf{x}) = -g(\mathbf{x}, \mathbf{x}) \quad \Longleftrightarrow \quad 2g(\mathbf{x}, \mathbf{x}) = \mathbf{0}$$

Om $2 \neq 0$ följer alltså att $g(\mathbf{x}, \mathbf{x}) = 0$. □

Exempel 3.6. Låt $k = \mathbb{Z}_2$ vara kroppen med två element. Då är $-1 = 1$ och skevsymmetrisk betyder detsamma som symmetrisk. Till exempel är multiplikationsavbildningen $m: k \times k \longrightarrow k$, $m(a, b) = ab$ bilinjär och (skev)symmetrisk men inte alternerande.

3.2. Yttre potensen $\bigwedge^\ell V$ med bas. Vi börjar med en definition som bygger på valet av en bas.

Definition 3.7. Den ℓ te *yttre potensen* av ett vektorrum V med basen $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ är vektorrummet $\bigwedge^\ell V$ med följande bas. För varje delmängd $S = \{s_1, s_2, \dots, s_\ell\} \subseteq \{1, 2, \dots, n\}$ av storlek ℓ och med $s_1 < s_2 < \cdots < s_\ell$ har vi en basvektor $\mathbf{e}_S = \mathbf{e}_{s_1} \wedge \mathbf{e}_{s_2} \wedge \cdots \wedge \mathbf{e}_{s_\ell}$.

Notera 3.8. Vi har att $\dim \bigwedge^\ell V = \binom{n}{\ell}$. Om $\ell = 0$ har vi den enda basvektorn $\mathbf{e}_\emptyset$ och om $\ell = n$ har vi den enda basvektorn $\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n$. Om $\ell < 0$ eller $\ell > n$ så är $\bigwedge^\ell V = \{\mathbf{0}\}$ det triviala vektorrummet av dimension 0.

Allmänna vektorer i $\bigwedge^\ell V$ är linjärkombinationer av basvektorerna. För $\ell = 1$ så är $\mathbf{e}_{\{i\}} = \mathbf{e}_i$ och vi identifierar på så vis $\bigwedge^1 V$ med V .

Exempel 3.9. Om $V = \mathbb{R}^3$ har vi att:

- $\bigwedge^0 \mathbb{R}^3$ har dimension 1 och basen $\mathbf{e}_0$,
- $\bigwedge^1 \mathbb{R}^3$ har dimension 3 och basen $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$,
- $\bigwedge^2 \mathbb{R}^3$ har dimension 3 och basen $\mathbf{e}_1 \wedge \mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_2 \wedge \mathbf{e}_3$, och
- $\bigwedge^3 \mathbb{R}^3$ har dimension 1 och basen $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$.

För en uppsättning vektorer $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell \in V$ kommer vi nu att definiera

$$\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \dots \wedge \mathbf{v}_\ell \in \bigwedge^\ell V$$

på liknande sätt som vi definierade den rena tensorn $\mathbf{v}_1 \otimes \dots \otimes \mathbf{v}_\ell \in V^{\otimes \ell}$ men så att $\wedge$ inte bara är multilinjär utan också alternerande. Detta innebär att $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_\ell = 0$ när $\mathbf{v}_i = \mathbf{v}_j$ för något $i \neq j$.

Exempel 3.10. Om $V = \mathbb{R}^2$ och $\mathbf{v} = a\mathbf{e}_1 + b\mathbf{e}_2$ och $\mathbf{w} = c\mathbf{e}_1 + d\mathbf{e}_2$ så har vi genom multilinearitet och alternerande

$$\begin{aligned} \mathbf{v} \wedge \mathbf{w} &= (a\mathbf{e}_1 + b\mathbf{e}_2) \wedge (c\mathbf{e}_1 + d\mathbf{e}_2) \\ &= ac(\mathbf{e}_1 \wedge \mathbf{e}_1) + ad(\mathbf{e}_1 \wedge \mathbf{e}_2) + bc(\mathbf{e}_2 \wedge \mathbf{e}_1) + bd(\mathbf{e}_2 \wedge \mathbf{e}_2) \\ &= (ad - bc)\mathbf{e}_1 \wedge \mathbf{e}_2 \end{aligned}$$

eftersom $\mathbf{e}_i \wedge \mathbf{e}_i = 0$ och $\mathbf{e}_1 \wedge \mathbf{e}_2 = -\mathbf{e}_2 \wedge \mathbf{e}_1$ (alternerande implicerar antisymmetrisk).

Sats 3.11. Det finns en unik alternerande avbildning $\alpha: V^\ell \longrightarrow \bigwedge^\ell V$ sådan att

$$\alpha(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_\ell}) = \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \dots \wedge \mathbf{e}_{i_\ell}$$

om $1 \leq i_1 < i_2 < \dots < i_\ell \leq n$. Vi låter $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_\ell := \alpha(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell)$.

Bevis. Eftersom α ska vara multilinjär räcker det att bestämma vad $\alpha(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_\ell})$ är för alla $1 \leq i_1, i_2, \dots, i_\ell \leq n$. Om två index är lika måste värdet vara noll eftersom α ska vara alternerande. Om alla index är olika kan vi göra N stycken parvisa byten för något N så att indexen efteråt är ordnade. Vi måste då låta

$$\alpha(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_\ell}) = (-1)^N \mathbf{e}_S$$

där $S = \{i_1, i_2, \dots, i_\ell\}$. Heltalet N är inte unikt bestämt men dess *paritet* (udda/jämmt) är unikt och $(-1)^N = \text{sgn}(\sigma)$ är tecknet för permutationen σ på $\{1, 2, \dots, \ell\}$ som svarar mot omordningen vi gjorde. $\square$

3.3. Universell egenskap för yttre potenser. Precis som tensorprodukten $V \otimes W$ har en universell egenskap för bilinjära avbildningar och $V_1 \otimes V_2 \otimes \dots \otimes V_n$ för multilinjära avbildningar så har den yttre potensen $\bigwedge^\ell V$ en universell egenskap för alternerande avbildningar.

Sats 3.12. Om $f: V^\ell \longrightarrow U$ är en alternerande avbildning så existerar en unik linjär avbildning $F: \bigwedge^\ell V \longrightarrow U$ sådan att $f = F \circ \alpha$.

Som vanligt så sammanfattar vi detta i diagrammet:

$$\begin{array}{ccc} V^\ell & \xrightarrow{\alpha} & \bigwedge^\ell V \\ & \searrow f & \downarrow \exists! F \\ & & U \end{array}$$

Vår nuvarande definition av $\bigwedge^\ell V$ beror på ett val av bas $\mathcal{B}$ och vi borde egentligen skriva $\alpha_{\mathcal{B}}$ och $\bigwedge_{\mathcal{B}}^\ell V$. Den universella egenskapen visar att om $\mathcal{B}$ och $\mathcal{B}'$ är två baser så finns det en unik isomorfi, kompatibel med $\alpha_{\mathcal{B}}$ och $\alpha_{\mathcal{B}'}$ mellan $\bigwedge_{\mathcal{B}}^\ell V$ och $\bigwedge_{\mathcal{B}'}^\ell V$. Denna

isomorfi ges genom att uttrycka basvektorer $\mathbf{e}_S = \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_\ell}$ för den ena basen i den andra basen, dvs $\alpha_{\mathcal{B}}(\mathbf{e}_S) = \alpha_{\mathcal{B}'}(\mathbf{e}_S)$. Vi kan därför betrakta $\bigwedge_{\mathcal{B}}^\ell V$ och $\bigwedge_{\mathcal{B}'}^\ell V$ som samma vektorrum och $\alpha_{\mathcal{B}}$ och $\alpha_{\mathcal{B}'}$ som samma avbildning.

Det går också att använda den universella egenskapen för att visa att andra definitioner av $\bigwedge^\ell V$ som vi ger senare överensstämmer med vår nuvarande.

Exempel 3.13. Vi såg att kryssprodukten $f: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ är alternerande. Alltså har vi en unik linjär avbildning $F: \bigwedge^2 \mathbb{R}^3 \rightarrow \mathbb{R}^3$ sådan att $F(\mathbf{v} \wedge \mathbf{w}) = \mathbf{v} \times \mathbf{w}$. Speciellt är

$$\begin{aligned} F(\mathbf{e}_1 \wedge \mathbf{e}_2) &= \mathbf{e}_3, \\ F(\mathbf{e}_1 \wedge \mathbf{e}_3) &= -\mathbf{e}_2, \\ F(\mathbf{e}_2 \wedge \mathbf{e}_3) &= \mathbf{e}_1, \end{aligned}$$

Eftersom F tar en bas på en bas så är F en isomorfi. Detta är ett specialfall av Hodge- $\star$ avbildningen, se Anmärkning 3.24.

Flera olika par av vektorer har samma kryssprodukt, till exempel:

$$\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2 = (\mathbf{e}_1 + \mathbf{e}_2) \times (\mathbf{e}_2) = (\mathbf{e}_1 + \mathbf{e}_2) \times \left(-\frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_2\right)$$

Väljer vi två stycken linjärt oberoende vektorer $\mathbf{v}, \mathbf{w}$ i xy -planet så är deras kryssprodukt $d\mathbf{e}_3$ för något $d \neq 0$ och alltså är $\mathbf{v} \times \frac{1}{d}\mathbf{w} = \mathbf{e}_3$. Alla sådana par av vektorer $\mathbf{v}, \mathbf{w}$ ger alltså upphov till samma element $\frac{1}{d}(\mathbf{v} \wedge \mathbf{w}) = \mathbf{e}_1 \wedge \mathbf{e}_2$ i $\bigwedge^2 \mathbb{R}^3$. Elementen i $\bigwedge^2 \mathbb{R}^3$ kan betraktas som plan i $\mathbb{R}^3$ med en given orienterad yta och detta gäller mer allmänt för $\bigwedge^2 \mathbb{R}^n$.

3.4. Konstruktion av $\bigwedge^\ell V$ som kvot. Eftersom α är multilinjär så får vi av den universella egenskapen för *tensorprodukten* en linjär avbildning

$$L: V^{\otimes \ell} = V \otimes \cdots \otimes V \longrightarrow \bigwedge^\ell V$$

sådan att $L(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_\ell) = \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_\ell$. Denna avbildning är surjektiv eftersom varje baselement i $\bigwedge^\ell V$ är med i bilden. Enligt isomorfisatsen är alltså $\bigwedge^\ell V$ isomorf med $V^{\otimes \ell} / \ker L$. Man kan visa att kärnan spänns upp av alla rena tensorer där två komponenter är lika. Vi kan därför göra följande alternativa definition av $\bigwedge^\ell V$ som inte beror på någon bas och fungerar även när V är oändligt-dimensionell.

Definition 3.14. $\bigwedge^\ell V$ är kvotrummet $V^{\otimes \ell} / R$ där R är delrummet som spänns upp av alla tensorer på formen

$$\mathbf{v}_1 \otimes \mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_\ell$$

där $\mathbf{v}_i \in V$ för alla $i = 1, 2, \dots, \ell$ och $\mathbf{v}_i = \mathbf{v}_j$ för något $i \neq j$.

Exempel 3.15. Om $V = \mathbb{R}^3$ och $\ell = 2$ så har $V \otimes V$ dimension 9 och basen $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{1 \leq i, j \leq 3}$ medan den yttre potensen $\bigwedge^2 V$ har dimension 3 och basen $\{\mathbf{e}_i \wedge \mathbf{e}_j\}_{1 \leq i < j \leq 3}$. Delrummet R spänns upp av de 6 linjärt oberoende vektorerna:

$$\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2, \mathbf{e}_3 \otimes \mathbf{e}_3, (\mathbf{e}_1 + \mathbf{e}_2) \otimes (\mathbf{e}_1 + \mathbf{e}_2), (\mathbf{e}_1 + \mathbf{e}_3) \otimes (\mathbf{e}_1 + \mathbf{e}_3), (\mathbf{e}_2 + \mathbf{e}_3) \otimes (\mathbf{e}_2 + \mathbf{e}_3),$$

eller ekvivalent, de 6 linjärt oberoende vektorerna:

$$\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2, \mathbf{e}_3 \otimes \mathbf{e}_3, \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1, \mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2.$$

3.5. **Yttre algebran** $\bigwedge V$. Vi börjar med att betrakta den direkta summan

$$\bigwedge V = \bigwedge^0 V \oplus \bigwedge^1 V \oplus \bigwedge^2 V \oplus \dots$$

Notera 3.16. Om V har dimension n och bas $\mathbf{e}_1, \dots, \mathbf{e}_n$ så är:

$$\bigwedge V = \bigwedge^0 V \oplus \bigwedge^1 V \oplus \dots \oplus \bigwedge^n V.$$

med bas $\{\mathbf{e}_S\}$ där $S \subseteq \{1, 2, \dots, n\}$ är en delmängd av godtycklig storlek. Alltså är dimension för $\bigwedge V$ lika med 2^n . För $\ell = 0, 1, \dots, n$ får vi delrum $\bigwedge^\ell V = \text{Span}\{\mathbf{e}_S \mid \#S = \ell\}$ av dimension $\binom{n}{\ell}$.

Definition 3.17. Den *yttre produkten* är den bilinjära avbildningen

$$\wedge: \left(\bigwedge^\ell V \right) \times \left(\bigwedge^p V \right) \longrightarrow \bigwedge^{\ell+p} V$$

som på baserna ges av

$$\mathbf{e}_S \wedge \mathbf{e}_T = (\mathbf{e}_{s_1} \wedge \dots \wedge \mathbf{e}_{s_\ell}) \wedge (\mathbf{e}_{t_1} \wedge \dots \wedge \mathbf{e}_{t_p}) = \begin{cases} (-1)^{m(S,T)} \mathbf{e}_{S \cup T} & \text{om } S \cap T = \emptyset, \\ 0 & \text{annars,} \end{cases}$$

där

$$m(S, T) = \#\{(s, t) \in S \times T : s > t\}.$$

är antalet byten vi behöver göra för att ordna följderna $s_1, s_2, \dots, s_\ell, t_1, t_2, \dots, t_p$ där vi bara har $s_1 < s_2 < \dots < s_\ell$ och $t_1 < t_2 < \dots < t_p$.

Samma formel ger en bilinjär avbildning

$$\wedge: \left(\bigwedge V \right) \times \left(\bigwedge V \right) \longrightarrow \bigwedge V$$

som vi också kallar för den yttre produkten.

Sats 3.18. Den yttre produkten på $\bigwedge V$ uppfyller

- $(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma), \quad \forall \alpha, \beta, \gamma \in \bigwedge V$ (associativitet)
- $\alpha \wedge \beta = (-1)^{\ell p} (\beta \wedge \alpha), \quad \forall \alpha \in \bigwedge^\ell V, \beta \in \bigwedge^p V$ (graderad kommutativ)

Definition 3.19. Den *yttre algebran*² (en: *exterior algebra*) på V är vektorrummet $\bigwedge V$ med produkten $\wedge: \bigwedge V \times \bigwedge V \longrightarrow \bigwedge V$ som definierats ovan.

Notera 3.20. Om vi tar produkten av ℓ stycken basvektorer i $V = \bigwedge^1 V$ så får vi

$$\mathbf{e}_{s_1} \wedge \mathbf{e}_{s_2} \wedge \dots \wedge \mathbf{e}_{s_\ell} = (-1)^{m(s_1, s_2, \dots, s_\ell)} \mathbf{e}_{\{s_1, s_2, \dots, s_\ell\}}$$

om indexen s_1, s_2, s_ℓ är distinkta (annars blir produkten 0) och

$$m(s_1, s_2, \dots, s_\ell) = \#\{(i, j) : i < j, s_i > s_j\}.$$

Detta ger en formel för den linjära avbildningen $\alpha: V^\ell \longrightarrow \bigwedge^\ell V$ i Sats 3.11.

²Här betyder *algebra* ett vektorrum med en associativ operation ”produkt”. Texten är de komplexa talen $\mathbb{C}$ och kvaternionerna $\mathbb{H}$ algebror över $\mathbb{R}$.

3.6. Determinanten. Om vi tar produkten av $n = \dim V$ vektorer så får vi determinanten:

Sats 3.21. Om $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ är vektorer i V där $\mathbf{f}_i = \sum a_{ji} \mathbf{e}_j$ är

$$\mathbf{f}_1 \wedge \mathbf{f}_2 \wedge \dots \wedge \mathbf{f}_n = \det(A) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n.$$

Bevis. Genom att utveckla vänsterledet

$$\mathbf{f}_1 \wedge \mathbf{f}_2 \wedge \dots \wedge \mathbf{f}_n = \left(\sum a_{j1} \mathbf{e}_j \right) \wedge \left(\sum a_{j2} \mathbf{e}_j \right) \wedge \dots \wedge \left(\sum a_{jn} \mathbf{e}_j \right)$$

får vi n^n termer, men alla termer som innehåller upprepade basvektorer blir noll. Därmed blir det en summa av $n!$ termer där alla är $a_{i_1 1} a_{i_2 2} \dots a_{i_n n} \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \dots \wedge \mathbf{e}_{i_n}$. Enligt ovan kommer tecknet bero på hur många par som står i fel ordning i listan $(i_1, i_2, \dots, i_n)$ och det är precis så determinanten är definierad. $\square$

Korollarium 3.22. Om $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_\ell$ är vektorer i V så är $\mathbf{f}_1 \wedge \mathbf{f}_2 \wedge \dots \wedge \mathbf{f}_\ell = 0$ om och endast om $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_\ell\}$ är linjärt beroende.

Bevis. Om $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_\ell\}$ är linjärt beroende kan vi skriva en av dem som en linjärkombination av de övriga. Då kommer vi få ett uttryck som är en summa av termer där det finns upprepade faktorer. Varje sådan term är noll eftersom $\mathbf{f}_i \wedge \mathbf{f}_i = 0$.

Om $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_\ell\}$ är linjärt oberoende kan vi utvidga den till en bas för V genom att lägga till $\mathbf{f}_{\ell+1}, \mathbf{f}_{\ell+2}, \dots, \mathbf{f}_n$. Vi får då att

$$\mathbf{f}_1 \wedge \mathbf{f}_2 \wedge \dots \wedge \mathbf{f}_n = \det(A) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n \neq 0$$

där A är basbytesmatrisen mellan baserna. Därmed kan inte de första ℓ faktorerna ge en produkt som är noll. $\square$

3.7. Konstruktion av $\wedge^\ell V$ med alternerande avbildningar. Liksom för tensorprodukten kan vi när V är ändligdimensionellt även definiera den yttre produkten på följande vis.

Definition 3.23. Den yttre produkten $\wedge^\ell V$ är

$$\wedge^\ell V = \left\{ f: V^* \times \dots \times V^* \longrightarrow k : f \text{ är alternerande} \right\}$$

Med denna beskrivning svarar basvektorn $\mathbf{e}_{s_1} \wedge \dots \wedge \mathbf{e}_{s_\ell}$ mot följande alternerande avbildning f :

$$f(\varphi_1, \dots, \varphi_\ell) = \sum_{\sigma \in \Sigma_\ell} \text{sgn}(\sigma) \varphi_{\sigma(1)}(\mathbf{e}_{s_1}) \varphi_{\sigma(2)}(\mathbf{e}_{s_2}) \dots \varphi_{\sigma(\ell)}(\mathbf{e}_{s_\ell})$$

där summan är över alla permutation σ av $\{1, 2, \dots, \ell\}$ och $\text{sgn}(\sigma) = \pm 1$ är tecknet på permutationen.

Men kan också ge en direkt beskrivningen av den yttre produkten i $\wedge V$ med alternerande avbildningar. Om till exempel $f \in \wedge^1 V$ och $g \in \wedge^2 V$ är alternerande avbildningar som i Definition 3.23 så är $f \wedge g \in \wedge^3 V$ följande alternerande avbildning

$$(f \wedge g)(\varphi_1, \varphi_2, \varphi_3) = f(\varphi_1)g(\varphi_2, \varphi_3) - f(\varphi_2)g(\varphi_1, \varphi_3) + f(\varphi_3)g(\varphi_1, \varphi_2).$$

3.8. Isomorfier mellan yttre produkter. Låt V vara ett vektorrum av dimension n . Vi har sett att $\dim \wedge^\ell V = \binom{n}{\ell} = \binom{n}{n-\ell} = \dim \wedge^{n-\ell} V$. Finns det en naturlig isomorfi mellan vektorrummen $\wedge^\ell V$ och $\wedge^{n-\ell} V$? Svaret är nej i allmänhet men genom att välja en ordnad bas kan vi skapa en isomorfi.

Notera 3.24 (Hodge $\star$ -isomorfi). Om vi har en *ordnad* bas $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ för V får vi en isomorfi

$$\star: \bigwedge^\ell V \longrightarrow \bigwedge^{n-\ell} V$$

sådan att $\star(\mathbf{e}_S) = \pm \mathbf{e}_{\{1,2,\dots,n\} \setminus S}$ där tecknet väljs så att $\mathbf{e}_S \wedge \star(\mathbf{e}_S) = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n$.

Har vi inte en ordnad bas kan vi inte ens identifiera $\bigwedge^n V$ med k på ett naturligt sätt.

Definition 3.25. En *volymform* för ett n -dimensionellt vektorrum V är ett noll-skilt element $\omega \in \bigwedge^n V$, eller ekvivalent, en isomorfi $k \cong \bigwedge^n V$.

Notera 3.26. Om vi har en ordnad bas $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ så har vi en kanonisk volymform $\omega = \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_n$. Om $\omega^*: \bigwedge^n V \longrightarrow k$ är isomorfin som skickar ω på 1 så är

$$\omega^*(\mathbf{f}_1 \wedge \mathbf{f}_2 \wedge \dots \wedge \mathbf{f}_n) = \det \begin{bmatrix} [\mathbf{f}_1]_{\mathcal{B}} & [\mathbf{f}_2]_{\mathcal{B}} & \dots & [\mathbf{f}_n]_{\mathcal{B}} \end{bmatrix}$$

enligt Sats 3.21.

Med en volymform kan vi på ett naturligt sätt identifiera $\bigwedge^{n-\ell} V$ med dualen till $\bigwedge^\ell V$ enligt följande sats.

Sats 3.27. Om vi har en volymform ω så är multiplikationen $\bigwedge^\ell V \times \bigwedge^{n-\ell} V \longrightarrow \bigwedge^n V \cong k$ en icke-degenererad³ bilinjär form och vi kan identifiera

$$\bigwedge^\ell V \cong \left(\bigwedge^{n-\ell} V \right)^*.$$

Bevis. Vi har en linjär avbildning

$$L: \bigwedge^\ell V \longrightarrow \left(\bigwedge^{n-\ell} V \right)^*$$

som tar ett element $\alpha \in \bigwedge^\ell V$ på den linjära formen $L(\alpha): \bigwedge^{n-\ell} V \longrightarrow \bigwedge^n V \cong k$ definierad av

$$L(\alpha)(\beta) := \omega^*(\alpha \wedge \beta), \quad \forall \beta \in \bigwedge^{n-\ell} V$$

där $\omega^*: \bigwedge^n V \longrightarrow k$ är isomorfin som tar volymformen ω på 1.

För att se att L är en isomorfi väljer vi basen $\{\mathbf{e}_i\}_{i=1,\dots,n}$ för V , basen $\{\mathbf{e}_S\}_{\#S=\ell}$ för $\bigwedge^\ell V$ och basen $\{\mathbf{e}_S\}_{\#S=n-\ell}$ för $\bigwedge^{n-\ell} V$. Vi har nu att $\mathbf{e}_S \wedge \mathbf{e}_T = \mathbf{0}$ om $T \neq S^c$ och $\mathbf{e}_S \wedge \mathbf{e}_{S^c} = \pm \mathbf{e}_{\{1,2,\dots,n\}}$. Alltså är

$$L(\mathbf{e}_S) = \omega^*(\pm \mathbf{e}_{\{1,2,\dots,n\}}) \mathbf{e}_{S^c}^*$$

så L tar en bas på en bas och är därmed en isomorfi. $\square$

Exempel 3.28. Om V är ett reellt inre produktrum kan vi identifiera V med V^* (via Sats 1.14) och även $(\bigwedge^{n-\ell} V)^*$ med $\bigwedge^{n-\ell} V$. Väljer vi en ortonormal bas och volymformen $\omega = \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_n$ så ger Sats 3.27 precis Hodge $\star$ -isomorfi.

Speciellt, om $V = \mathbb{R}^3$ med den vanliga inre produkten så ger den yttre produkten en multiplikation

$$\wedge: V \times V \longrightarrow \bigwedge^2 V \cong V^* \cong V.$$

Denna multiplikation är det vi känner som *vektorprodukt* eller *kryssprodukt*.

³*icke-degenererad* betyder i det här fallet att det inte finns något element som multipliceras till noll av alla element på andra sidan.

REFERENSER

- [Cur90] Morton L. Curtis. *Abstract linear algebra*. Universitext. Springer-Verlag, New York, 1990. With revisions by Paul Place, With a preface by John Hempel.
- [Rom92] Steven Roman. *Advanced linear algebra*, volume 135 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1992.
- [Sad08] Lorenzo Sadun. *Applied linear algebra*. American Mathematical Society, Providence, RI, second edition, 2008. The decoupling principle.
- [Wik17a] Wikipedia. Exterior algebra — Wikipedia, The Free Encyclopedia, 2017.
- [Wik17b] Wikipedia. Tensor product — Wikipedia, The Free Encyclopedia, 2017.

ÖVNINGAR PÅ YTTRE ALGEBRAN

Uppgift 12. Låt $\text{Alt}^2(V) \subset V \otimes V$ vara de alternerande tensorerna (Definition 2.21). Visa att det finns en naturlig isomorfi $\bigwedge^2 V \cong \text{Alt}^2(V)$.

Uppgift 13. (a) Visa att multiplikationen i Sats 3.27 för $\ell = 1$ motsvarar Laplaceutvecklingen av determinanten av en matris efter en rad eller kolonn.
 (b) Visa att multiplikationen i Sats 3.27 för $\ell > 1$ motsvarar en generalisering av Laplaceutvecklingen av determinanten av en matris efter flera rader rad eller flera kolonner.
 (c) Skriv upp den explicita Laplaceutvecklingen av en 4×4 -determinant efter de två första raderna.

Uppgift 14. Verifiera att det är *kryssprodukten* som fås för $V = \mathbb{R}^3$ med den vanliga inre produkten som indikeras i Exempel 3.28 där vi får multiplikation

$$\wedge: V \times V \longrightarrow \bigwedge^2 V \cong V^* \cong V.$$

Uppgift 15. Genomför första steget i beviset av Sats 3.21 explicit för fallen $n = 2$ och $n = 3$, dvs utveckla uttrycken

$$(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2) \wedge (a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2)$$

och

$$(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + a_{31}\mathbf{e}_3) \wedge (a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + a_{32}\mathbf{e}_3) \wedge (a_{13}\mathbf{e}_1 + a_{23}\mathbf{e}_2 + a_{33}\mathbf{e}_3).$$

Uppgift 16 (Uppgift 7 på tentamen 2019-01-09). Låt $V = \mathbb{C}^n$ med bas $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Bestäm rangen av den linjära avbildningen

$$L: \bigwedge^2 V \longrightarrow \bigwedge^3 V$$

som ges av $L(x) = x \wedge (\mathbf{e}_1 + \mathbf{e}_2 + \dots + \mathbf{e}_n)$, för $x \in \bigwedge^2 V$. (4 p)

LÖSNINGSFÖRSLAG

Lösningförslag till uppgift 1. Vi har basen $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ för V och basen

$$\{\mathbf{e}_{11}, \mathbf{e}_{12}, \mathbf{e}_{13}, \mathbf{e}_{21}, \mathbf{e}_{22}, \mathbf{e}_{23}, \mathbf{e}_{31}, \mathbf{e}_{32}, \mathbf{e}_{33}\}$$

för $V \otimes V$ där $\mathbf{e}_{ij} = \mathbf{e}_i \otimes \mathbf{e}_j$. Vidare är

$$L(\mathbf{e}_{ij}) = L(\mathbf{e}_i \otimes \mathbf{e}_j) = f(\mathbf{e}_i, \mathbf{e}_j) = \mathbf{e}_i \times \mathbf{e}_j$$

och

$$L(\mathbf{e}_{11}) = L(\mathbf{e}_{22}) = L(\mathbf{e}_{33}) = 0$$

$$L(\mathbf{e}_{12}) = \mathbf{e}_3, \quad L(\mathbf{e}_{21}) = -\mathbf{e}_3$$

$$L(\mathbf{e}_{23}) = \mathbf{e}_1, \quad L(\mathbf{e}_{32}) = -\mathbf{e}_1$$

$$L(\mathbf{e}_{31}) = \mathbf{e}_2, \quad L(\mathbf{e}_{13}) = -\mathbf{e}_2$$

Detta ger följande matris för L

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Vi ser att matrisen har full rang (dvs 3) och därmed har kärnan dimension 6. Matrisen är redan rad-reducerad och en bas för kärnan är alltså

$$\{\mathbf{e}_{11}, \mathbf{e}_{22}, \mathbf{e}_{33}, \mathbf{e}_{23} + \mathbf{e}_{32}, \mathbf{e}_{13} + \mathbf{e}_{31}, \mathbf{e}_{12} + \mathbf{e}_{21}\}.$$

Lösningförslag till uppgift 2. Låt

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_dx^d$$

$$q(y) = b_0 + b_1y + b_2y^2 + \cdots + b_ey^e.$$

Då är

$$p(x) \otimes q(y) = \sum_{i=0}^d \sum_{j=0}^e a_ib_j(x^i \otimes y^j)$$

Eftersom $\{1, x, x^2, \dots\}$ och $\{1, y, y^2, \dots\}$ är baser så är $\{x^i \otimes y^j\}_{i \geq 0, j \geq 0}$ en bas. Om $1 \otimes 1 + x \otimes y$ är lika med $p(x) \otimes q(y)$ är deras koefficienter i denna bas lika:

$$\begin{aligned} 1 &= a_0b_0 & 0 &= a_0b_1 & 0 &= a_0b_2 & 0 &= a_ib_j \quad \forall i+j \geq 3 \\ & & 0 &= a_1b_0 & 1 &= a_1b_1 & & \\ & & & & 0 &= a_2b_0 & & \end{aligned}$$

Men $1 = a_0b_0$ och $1 = a_1b_1$ ger att $a_0, b_0, a_1, b_1 \neq 0$ vilket strider mot att $a_0b_1 = 0$ och $a_1b_0 = 0$.

Lösningförslag till uppgift 3. På grund av symmetri räcker det att kontrollera att avbildningen blir linjär om vi fixerar ett av argumenten.

$$\begin{aligned} (\mathbf{x} \otimes \mathbf{y})(\phi_1 + \phi_2, \psi) &= (\phi_1(\mathbf{x}) + \phi_2(\mathbf{x}))\psi(\mathbf{y}) = (\phi_1(\mathbf{x})\psi(\mathbf{y})) + (\phi_2(\mathbf{x})\psi(\mathbf{y})) \\ &= (\mathbf{x} \otimes \mathbf{y})(\phi_1, \psi) + (\mathbf{x} \otimes \mathbf{y})(\phi_2, \psi) \end{aligned}$$

och

$$(\mathbf{x} \otimes \mathbf{y})(a\phi, \psi) = (a\phi(\mathbf{x}))\psi(\mathbf{y}) = a\phi(\mathbf{x})\psi(\mathbf{y}) = a(\mathbf{x} \otimes \mathbf{y})(\phi, \psi)$$

vilket visar att avbildningen är linjär om vi fixerar andra argumentet till ψ .

Lösningförslag till uppgift 4. Om A är matrisen för L med avseende på en bas får vi att $\text{tr}(L) = \text{tr}(A) = \sum A_{ii}$. Detta är oberoende av val av bas eftersom $\text{tr}(PAP^{-1}) = \text{tr}(AP^{-1}P) = \text{tr}(A)$ för varje basbytesmatris P . Vi har att L svarar mot tensorn $\sum A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j^*$ och

$$\text{tr}(\sum A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j^*) = \sum A_{ij} \text{tr}(\mathbf{e}_i \otimes \mathbf{e}_j^*) = \sum A_{ij} \delta_{ij} = \sum A_{ii} = \text{tr}(A).$$

Lösningförslag till uppgift 5.

(a) Om vi skriver $\mathbf{f}_i^* = \sum b_{ji} \mathbf{e}_j^*$ och använder att $\mathbf{f}_i^*(\mathbf{f}_j) = \delta_{ij}$ får vi

$$\sum b_{ki} \mathbf{e}_k^* (\sum a_{\ell j} \mathbf{e}_\ell) = \sum_{k=1}^n \sum_{\ell=1}^n b_{ki} a_{\ell j} \mathbf{e}_k^*(\mathbf{e}_\ell) = \sum_{k=1}^n \sum_{\ell=1}^n b_{ki} a_{\ell j} \delta_{k\ell} = \sum_{k=1}^n b_{ki} a_{kj}$$

Detta är matrismultiplikation av B^T med A och resultatet ska ge identitetsmatrisen. Alltså är $B = (A^{-1})^T = (A^T)^{-1}$.

(b) På $V^* \otimes V$ får vi basbyte med

$$\mathbf{f}_i^* \otimes \mathbf{f}_j = (\sum b_{ki} \mathbf{e}_k^*) \otimes (\sum a_{\ell j} \mathbf{e}_\ell) = \sum_{k=1}^n \sum_{\ell=1}^n b_{ki} a_{\ell j} \mathbf{e}_k^* \otimes \mathbf{e}_\ell$$

dvs vi använder basbytesmatrisen A på högra faktorn och $(A^{-1})^T$ på den vänstra. När vi representerar elementen i V som kolonnvektorer och elementen i V^* som radvektorer behövs inte transponatet längre, men vi multiplicerar med basbytesmatrisen från höger istället för från vänster.

Lösningförslag till uppgift 6. Eftersom $\text{Sym}^2(V)$ och $\text{Alt}^2(V)$ är delrum av $V \otimes V$ ska vi visa att det är en inre direkt summa. Varje element i $V \otimes V$ på formen $\mathbf{x} \otimes \mathbf{y}$ kan skrivas som

$$\mathbf{x} \otimes \mathbf{y} = \frac{1}{2}(\mathbf{x} \otimes \mathbf{y} + \mathbf{y} \otimes \mathbf{x}) + \frac{1}{2}(\mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x})$$

där den första termen ligger i $\text{Sym}^2(V)$ och den andra i $\text{Alt}^2(V)$. Eftersom alla element i $V \otimes V$ kan skrivas som summor av element på formen $\mathbf{x} \otimes \mathbf{y}$ får vi att alla element i $V \otimes V$ kan skrivas som en summa av ett element från $\text{Sym}^2(V)$ och ett element från $\text{Alt}^2(V)$. Det återstår att visa att snittet är noll eller att $\dim \text{Sym}^2(V) + \dim \text{Alt}^2(V) = \dim V \otimes V$. Vi kan se att $\{\mathbf{e}_i \otimes \mathbf{e}_j + \mathbf{e}_j \otimes \mathbf{e}_i\}_{i \geq j}$ utgör en bas för $\text{Sym}^2(V)$ och $\{\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i\}_{i > j}$ utgör en bas för $\text{Alt}^2(V)$. Därmed har vi att summan av dimensionerna är $n(n+1)/2 + n(n-1)/2 = n^2$.

Lösningförslag till uppgift 7. Eftersom $V \otimes \mathbb{C}$ är ett reellt vektorrum är det en abelsk grupp med avseende på addition. Vi behöver kontrollera att multiplikationen med skalär fungerar som den ska. Genom att använda den andra faktorn får vi för varje basvektor $\mathbf{e}_j$ för V att

$$(a + ib) \cdot (\mathbf{e}_j \otimes 1) = \mathbf{e}_j \otimes (a + ib) = a\mathbf{e}_j \otimes 1 + b\mathbf{e}_j \otimes i$$

och

$$(a + ib) \cdot (\mathbf{e}_j \otimes i) = \mathbf{e}_j \otimes i(a + ib) = a\mathbf{e}_j \otimes i - b\mathbf{e}_j \otimes 1.$$

När vi multiplicerar med ytterligare ett komplext tal $c + id$ får vi

$$\begin{aligned} (c + id) \cdot ((a + ib) \cdot (\mathbf{e}_j \otimes 1)) &= (c + id) \cdot (a\mathbf{e}_j \otimes 1 + b\mathbf{e}_j \otimes i) \\ &= ac\mathbf{e}_j \otimes 1 + ade\mathbf{e}_j \otimes i + bce\mathbf{e}_j \otimes i - bde\mathbf{e}_j \otimes 1 = (ac - bd)\mathbf{e}_j \otimes 1 + (ad + bc)\mathbf{e}_j \otimes i \end{aligned}$$

som svarar mot multiplikation med $(c + id)(a + ib) = (ac - bd) + (ad + bc)i$.

Vi behöver också kontrollera de distributiva lagarna. Vi får

$$\begin{aligned} ((a+ib) + (c+id)) \cdot \mathbf{x} \otimes 1 &= ((a+ib) + (c+id)) \cdot \mathbf{x} \otimes 1 \\ &= (a+c)\mathbf{x} \otimes 1 + (b+d)\mathbf{x} \otimes i \\ &= a\mathbf{x} \otimes 1 + b\mathbf{x} \otimes i + c\mathbf{x} \otimes 1 + d\mathbf{x} \otimes i \\ &= (a+ib) \cdot \mathbf{x} \otimes 1 + (c+id) \cdot \mathbf{x} \otimes 1 \end{aligned}$$

och

$$\begin{aligned} ((a+ib) + (c+id)) \cdot \mathbf{x} \otimes i &= ((a+ib) + (c+id)) \cdot \mathbf{x} \otimes i \\ &= -(b+d)\mathbf{x} \otimes 1 + (a+c)\mathbf{x} \otimes i \\ &= -b\mathbf{x} \otimes 1 + a\mathbf{x} \otimes i - d\mathbf{x} \otimes 1 + c\mathbf{x} \otimes i \\ &= (a+ib) \cdot \mathbf{x} \otimes i + (c+id) \cdot \mathbf{x} \otimes i. \end{aligned}$$

Vidare gäller att

$$\begin{aligned} (a+bi) \cdot (\mathbf{x} + \mathbf{y}) \otimes 1 &= a(\mathbf{x} + \mathbf{y}) \otimes 1 + b(\mathbf{x} + \mathbf{y}) \otimes i \\ &= a\mathbf{x} \otimes 1 + b\mathbf{x} \otimes i + a\mathbf{y} \otimes 1 + b\mathbf{y} \otimes i \\ &= (a+bi) \cdot \mathbf{x} \otimes 1 + (a+bi) \cdot \mathbf{y} \otimes 1 \end{aligned}$$

och

$$\begin{aligned} (a+bi) \cdot (\mathbf{x} + \mathbf{y}) \otimes i &= a(\mathbf{x} + \mathbf{y}) \otimes i - b(\mathbf{x} + \mathbf{y}) \otimes 1 \\ &= a\mathbf{x} \otimes i - b\mathbf{x} \otimes 1 + a\mathbf{y} \otimes i - b\mathbf{y} \otimes 1 \\ &= (a+bi) \cdot \mathbf{x} \otimes i + (a+bi) \cdot \mathbf{y} \otimes i. \end{aligned}$$

Lösningsförslag till uppgift 8.

(a) Vi har att

$$L(\mathbf{e}_i \otimes \mathbf{e}_j^*) = L(\mathbf{e}_i) \otimes L^*(\mathbf{e}_j^*) = \left(\sum_{k=1}^3 A_{ki} \mathbf{e}_k \right) \otimes \left(\sum_{\ell=1}^3 A_{j\ell}^* \mathbf{e}_\ell^* \right) = \sum_{k=1}^3 \sum_{\ell=1}^3 A_{ki} A_{j\ell}^* \mathbf{e}_k \otimes \mathbf{e}_\ell^*$$

Om vi ordnar basen $\{\mathbf{e}_1 \otimes \mathbf{e}_1^*, \mathbf{e}_2 \otimes \mathbf{e}_1^*, \mathbf{e}_3 \otimes \mathbf{e}_1^*, \mathbf{e}_1 \otimes \mathbf{e}_2^*, \mathbf{e}_2 \otimes \mathbf{e}_2^*, \mathbf{e}_3 \otimes \mathbf{e}_2^*, \mathbf{e}_1 \otimes \mathbf{e}_3^*, \mathbf{e}_2 \otimes \mathbf{e}_3^*, \mathbf{e}_3 \otimes \mathbf{e}_3^*\}$ betyder det att vi får en blockmatris med nio 3×3 -block där varje block är en multipel av matrisen A , dvs

$$B = \begin{bmatrix} A_{11}A & A_{21}A & A_{31}A \\ A_{12}A & A_{22}A & A_{32}A \\ A_{13}A & A_{23}A & A_{33}A \end{bmatrix}$$

dvs

$$B = \begin{bmatrix} 4 & 4 & 2 & 2 & 2 & 1 & -4 & -4 & -2 \\ 2 & -2 & -2 & 1 & -1 & -1 & -2 & 2 & 2 \\ -4 & -2 & -4 & -2 & -1 & -2 & 4 & 2 & 4 \\ 4 & 4 & 2 & -2 & -2 & -1 & -2 & -2 & -1 \\ 2 & -2 & -2 & -1 & 1 & 1 & -1 & 1 & 1 \\ -4 & -2 & -4 & 2 & 1 & 2 & 2 & 1 & 2 \\ 2 & 2 & 1 & -2 & -2 & -1 & -4 & -4 & -2 \\ 1 & -1 & -1 & -1 & 1 & 1 & -2 & 2 & 2 \\ -2 & -1 & -2 & 2 & 1 & 2 & 4 & 2 & 4 \end{bmatrix}.$$

- (b) Vi kan utvidga koefficienterna till $\mathbb{C}$ där vi kan diagonalisera matrisen eller åtminstone få den på övertriangulär form. När vi sedan ser på matrisen för $L \otimes L^*$ får vi en block-undertriangulär matris varför determinanten är produkten av determinanterna för matriserna på diagonalen. Därmed blir determinanten $(\det(A))^6 = 7^6$. (Det karakteristiska polynomet är $p_A(t) = t^3 + t^2 - 5t - 7$

och eftersom det inte finns några gemensamma delare mellan $p_A(t)$ och $p'_A(t) = 3t^2 + 2t - 5$ kan vi sluta oss till att egenvärdena är distinkta i $\mathbb{C}$ och A är diagonaliserbar över $\mathbb{C}$. På så vis får vi en diagonalmatris även för $L \otimes L^*$.

Lösningsförslag till uppgift 9. Dimensionen av V är 2 och vi kan välja en bas som består av $\mathbf{e}_1 - \mathbf{e}_2$ och $\mathbf{e}_2 - \mathbf{e}_3$. Därmed spänns $V \otimes V$ upp av de fyra tensorerna

$$\begin{aligned}(\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}_1 - \mathbf{e}_2) &= \mathbf{e}_1 \otimes \mathbf{e}_1 - \mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2, \\(\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}_2 - \mathbf{e}_3) &= \mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_1 \otimes \mathbf{e}_3 - \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_3, \\(\mathbf{e}_2 - \mathbf{e}_3) \otimes (\mathbf{e}_1 - \mathbf{e}_2) &= \mathbf{e}_2 \otimes \mathbf{e}_1 - \mathbf{e}_2 \otimes \mathbf{e}_2 - \mathbf{e}_3 \otimes \mathbf{e}_1 + \mathbf{e}_3 \otimes \mathbf{e}_2, \\(\mathbf{e}_2 - \mathbf{e}_3) \otimes (\mathbf{e}_2 - \mathbf{e}_3) &= \mathbf{e}_2 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_3 - \mathbf{e}_3 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3.\end{aligned}$$

Vi har att $\mathbf{e}_1 \otimes \mathbf{e}_1$ bara förekommer i den första basvektorn, $\mathbf{e}_1 \otimes \mathbf{e}_3$ bara i den andra, $\mathbf{e}_3 \otimes \mathbf{e}_1$ bara i den tredje och $\mathbf{e}_3 \otimes \mathbf{e}_3$ bara i den fjärde. För att uttrycka $\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_2 \otimes \mathbf{e}_1 - 2\mathbf{e}_3 \otimes \mathbf{e}_2$ med hjälp av basvektorerna behöver alla koefficienter vara noll utom den andra eftersom ingen av de tre tensorerna $\mathbf{e}_1 \otimes \mathbf{e}_1$, $\mathbf{e}_3 \otimes \mathbf{e}_1$ eller $\mathbf{e}_3 \otimes \mathbf{e}_3$ förekommer i den givna tensorn. Eftersom tensorn inte är en multipel av den andra basvektorn ligger den inte i $V \otimes V$.

När vi ser på den andra tensorn behöver koefficienterna för basvektorerna vara 1, 0, 0 och -1 och vi har att

$$\mathbf{e}_1 \otimes \mathbf{e}_1 - \mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2 - \mathbf{e}_3 \otimes \mathbf{e}_3 = (\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}_1 - \mathbf{e}_2) - (\mathbf{e}_2 - \mathbf{e}_3) \otimes (\mathbf{e}_2 - \mathbf{e}_3)$$

som alltså ligger i $V \otimes V$.

Lösningsförslag till uppgift 10.

(a) Vi kan definiera

$$V \otimes W = \{f: V^* \times W^* \longrightarrow k \mid f \text{ är bilinjär}\}$$

där V^* och W^* är de duala rummen till V respektive W .

(b) För baser $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$ för V och $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ för W kan vi definiera $\mathbf{e}_i \otimes \mathbf{f}_j \in V \otimes W$ genom

$$(\mathbf{e}_i \otimes \mathbf{f}_j)(\phi, \psi) = \phi(\mathbf{e}_i)\psi(\mathbf{f}_j), \quad \forall (\phi, \psi) \in V^* \times W^*.$$

Varje element i $V \otimes W$ kan nu skrivas på ett unikt sätt som

$$f = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \mathbf{e}_i \otimes \mathbf{f}_j$$

med $a_{ij} = f(\mathbf{e}_i^*, \mathbf{f}_j^*)$ där $\{\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_m^*\}$ och $\{\mathbf{f}_1^*, \mathbf{f}_2^*, \dots, \mathbf{f}_n^*\}$ är de duala baserna. Alltså har $V \otimes W$ en bas som består av $mn = (\dim V)(\dim W)$ element.

(c) Vi kan definiera $\Phi: V \times W \longrightarrow V \otimes W$ genom $\Phi(\mathbf{x}, \mathbf{y})(\phi, \psi) = \phi(\mathbf{x})\psi(\mathbf{y})$. Detta ger för varje $(\mathbf{x}, \mathbf{y})$ en bilinjär form på $V^* \times W^*$ och om vi fixerar $\mathbf{y}$ är den linjär i $\mathbf{x}$ eftersom ϕ är linjär och om vi fixerar $\mathbf{x}$ är den linjär i $\mathbf{y}$ eftersom ψ är linjär. Ett annat sätt att definiera samma bilinjära avbildning är genom

$$\Phi(\mathbf{x}, \mathbf{y}) = \Phi\left(\sum_{i=1}^m a_i \mathbf{e}_i, \sum_{j=1}^n b_j \mathbf{f}_j\right) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathbf{e}_i \otimes \mathbf{f}_j.$$

(d) Den universella egenskapen innebär att om $\Psi: V \times W \longrightarrow U$ finns en bilinjär avbildning, så finns en entydig linjär avbildning $L: V \otimes W \longrightarrow U$ så att $\Psi = L \circ \Phi$. Vi kan bevisa det genom att definiera

$$L(\mathbf{e}_i \otimes \mathbf{f}_j) = \Psi(\mathbf{e}_i, \mathbf{f}_j), \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

Denna är väldefinierad eftersom $\{\mathbf{e}_1 \otimes \mathbf{f}_j, \mathbf{e}_m \otimes \mathbf{f}_n\}$ utgör en bas för $V \otimes W$. Den uppfyller att $\Psi = L \circ \Phi$ eftersom

$$\begin{aligned}\Psi\left(\sum_{i=1}^m a_i \mathbf{e}_i, \sum_{j=1}^n b_j \mathbf{f}_j\right) &= \sum_{i=1}^m \sum_{j=1}^n a_i b_j \Psi(\mathbf{e}_i \mathbf{f}_j) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j L(\mathbf{e}_i \otimes \mathbf{f}_j) \\ &= L\left(\sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathbf{e}_i \otimes \mathbf{f}_j\right) = L\left(\Phi\left(\sum_{i=1}^m a_i \mathbf{e}_i, \sum_{j=1}^n b_j \mathbf{f}_j\right)\right)\end{aligned}$$

Att L är entydig ser vi genom att använda $\Psi = L \circ \Phi$ på $\mathbf{e}_i \otimes \mathbf{f}_j$, vilket ger att $L(\mathbf{e}_i \otimes \mathbf{f}_j) = \Psi(\mathbf{e}_i, \mathbf{f}_j)$.

Lösningförslag till uppgift 11.

- (a) Vi väljer först en annan bas där matrisen för L blir diagonalmatris. Egenvärdena för L är 5 och -1 . Matrisen för $I \otimes L - L \otimes I$ blir då diagonalmatrisen

$$\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

om vi ordnar basen för $V \otimes V$ som $\{\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_1 \otimes \mathbf{e}_2, \mathbf{e}_2 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2\}$. Rangén är två.

- (b) Om L är nilpotent är $L^m = 0$, för något heltal m . Vi har att

$$(I \otimes L - L \otimes I)^n = \sum_{k=0}^n (-1)^k L^k \otimes L^{n-k}$$

och detta är noll om $n \geq 2n$ eftersom varje term då är noll.

Om L är nilpotent kan vi se på $(I + L)$ som är inverterbar, men vi får

$$I \otimes (I + L) - (I + L) \otimes I = I \otimes I + I \otimes L - I \otimes I - L \otimes I = I \otimes L - L \otimes I$$

som är nilpotent.

Lösningförslag till uppgift 12. Vi har att en bas för $\wedge^2 V$ ges av $\{\mathbf{e}_i \wedge \mathbf{e}_j\}_{1 \leq i < j \leq n}$ och att en bas för $\text{Alt}^2(V)$ ges av $\{\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i\}_{1 \leq i < j \leq n}$. Det är därför klart att vi kan definiera en isomorfi genom

$$\mathbf{e}_i \wedge \mathbf{e}_j \mapsto \mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i, \quad 1 \leq i < j \leq n.$$

För att se att denna isomorfi är naturlig i den mening att den inte beror på valet av bas för V ser vi vad som händer för två vektorer $\mathbf{x} = \sum x_i \mathbf{e}_i$ och $\mathbf{y} = \sum y_j \mathbf{e}_j$. Vi får

$$\mathbf{x} \wedge \mathbf{y} = \left(\sum x_i \mathbf{e}_i\right) \wedge \left(\sum y_j \mathbf{e}_j\right) = \sum_{1 \leq i < j \leq n} (x_i y_j - x_j y_i) \mathbf{e}_i \wedge \mathbf{e}_j$$

och

$$\begin{aligned}\mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x} &= \left(\sum x_i \mathbf{e}_i\right) \otimes \left(\sum y_j \mathbf{e}_j\right) - \left(\sum y_i \mathbf{e}_i\right) \otimes \left(\sum x_j \mathbf{e}_j\right) \\ &= \sum_{1 \leq i < j \leq n} (x_i y_j - x_j y_i) (\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i).\end{aligned}$$

Alltså skickar vår isomorfi $\mathbf{x} \wedge \mathbf{y}$ till $\mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x}$ för alla vektorer $\mathbf{x}$ och $\mathbf{y}$ i V och därmed beror isomorfin inte på val av bas.

Lösningförslag till uppgift 13.

- (a) Vi tänker oss en $n \times n$ -matris $A = (a_{ij})$ som en uppsättning av n kolonnvektorer $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ som alla är uttryckta i standardbasen $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Vi får då att

$$\begin{aligned}\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \dots \wedge \mathbf{x}_n &= \sum_{i=1}^n (a_{i1}a_{i2} \dots a_{in}) \mathbf{e}_{i1} \wedge \mathbf{e}_{i2} \wedge \dots \wedge \mathbf{e}_{in} \\ &= \det(A) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n\end{aligned}$$

Om vi använder multiplikationen från Sats 3.27 med $\ell = 1$ har vi

$$(\mathbf{x}_1, \mathbf{x}_2 \wedge \mathbf{x}_3 \wedge \dots \wedge \mathbf{x}_n) \mapsto \mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \dots \wedge \mathbf{x}_n = \det(A) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n \mapsto \det(A)$$

När vi utvecklar det andra argumentet får vi

$$\begin{aligned}\mathbf{x}_2 \wedge \mathbf{x}_3 \wedge \dots \wedge \mathbf{x}_n &= \sum_{i=1}^n (a_{i2}a_{i3} \dots a_{in}) \mathbf{e}_{i2} \wedge \mathbf{e}_{i3} \wedge \dots \wedge \mathbf{e}_{in} \\ &= \sum_{i_2 < i_3 < \dots < i_n} \det(A'_{i_2, i_3, \dots, i_n}) \mathbf{e}_{i_2} \wedge \mathbf{e}_{i_3} \wedge \dots \wedge \mathbf{e}_{i_n} \\ &= \det(A_{1, 2, \dots, n-1}) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_{n-1} \\ &\quad + \det(A_{1, 2, \dots, n-2, n}) \mathbf{e}_1 \wedge \mathbf{e}_3 \wedge \dots \wedge \mathbf{e}_{n-2} \wedge \mathbf{e}_n \\ &\quad + \dots + \det(A_{2, 3, \dots, n}) \mathbf{e}_2 \wedge \mathbf{e}_3 \wedge \dots \wedge \mathbf{e}_n\end{aligned}$$

där $A_{i_2, i_3, \dots, i_n}$ är delmatrisen av A som ges av de $n-1$ sista kolonnerna och raderna $i_2, \dots, i_n$. Tecknen i Laplaceutvecklingen av determinanten kommer när vi sätter samman detta med det första argumentet och

$$\mathbf{e}_k \wedge \mathbf{e}_{i_2} \wedge \mathbf{e}_{i_3} \wedge \dots \wedge \mathbf{e}_{i_n} = \begin{cases} (-1)^{k+1} \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n & \text{om } k \notin \{i_2, i_3, \dots, i_n\} \\ 0 & \text{annars} \end{cases}$$

- (b) För $\ell > 0$ kan vi använda samma princip och vi får att det första argumentet utvecklas till

$$\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \dots \wedge \mathbf{x}_\ell = \sum_{i_1 < \dots < i_\ell} \det(A'_{i_1, i_2, \dots, i_\ell}) \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \dots \wedge \mathbf{e}_{i_\ell}$$

där $A'_{i_1, \dots, i_\ell}$ är delmatrisen av A som ges av de ℓ första kolonnerna och raderna $i_1, i_2, \dots, i_\ell$. När vi multiplicerar ihop uttrycken får vi bara bidrag från de par av basvektorer $(\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \dots \wedge \mathbf{e}_{i_\ell}, \mathbf{e}_{i_{\ell+1}} \wedge \mathbf{e}_{i_{\ell+2}} \wedge \dots \wedge \mathbf{e}_{i_n})$ där $\{i_1, i_2, \dots, i_\ell\} \cup \{i_{\ell+1}, i_{\ell+2}, \dots, i_n\} = \{1, 2, \dots, n\}$ och vi får ett tecken som beror på hur många index vi behöver byta plats på för att sortera listan $i_1, i_2, \dots, i_n$.

- (c) För en 4×4 -matris

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

får vi

$$\begin{aligned}\det(A) &= \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \det \begin{bmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{bmatrix} - \det \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \det \begin{bmatrix} a_{23} & a_{24} \\ a_{43} & a_{44} \end{bmatrix} \\ &\quad + \det \begin{bmatrix} a_{11} & a_{12} \\ a_{41} & a_{42} \end{bmatrix} \det \begin{bmatrix} a_{23} & a_{24} \\ a_{33} & a_{34} \end{bmatrix} + \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \det \begin{bmatrix} a_{13} & a_{14} \\ a_{43} & a_{44} \end{bmatrix} \\ &\quad - \det \begin{bmatrix} a_{21} & a_{22} \\ a_{41} & a_{42} \end{bmatrix} \det \begin{bmatrix} a_{13} & a_{14} \\ a_{33} & a_{34} \end{bmatrix} + \det \begin{bmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \det \begin{bmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{bmatrix}\end{aligned}$$

Lösningsförslag till uppgift 14. Vi börjar med att se hur vi identifierar $\bigwedge^2 V$ med V^* och sedan med V . Enligt Sats 3.27 har vi $V \times \bigwedge^2 V \rightarrow \bigwedge^3 V \cong \mathbb{R}$ och om vi ser på hur det fungerar på baselementen har vi

$$(\mathbf{e}_1, \mathbf{e}_2 \wedge \mathbf{e}_3) \mapsto 1, \quad (\mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3) \mapsto -1, \quad \text{och} \quad (\mathbf{e}_3, \mathbf{e}_1 \wedge \mathbf{e}_2) \mapsto 1$$

medan alla andra par avbildas på noll. Det gör att vi får en identifikation

$$\mathbf{e}_2 \wedge \mathbf{e}_3 \longleftrightarrow \mathbf{e}_1^*, \quad \mathbf{e}_1 \wedge \mathbf{e}_3 \longleftrightarrow -\mathbf{e}_2^*, \quad \text{och} \quad \mathbf{e}_1 \wedge \mathbf{e}_2 \longleftrightarrow \mathbf{e}_3^*.$$

Vi använder sedan den inre produkten på $V = \mathbb{R}^3$ till att identifiera $\mathbf{e}_i$ med $\mathbf{e}_i^*$, för $i = 1, 2, 3$.

Om vi nu har vektorerna $x = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3$ och $y = y_1 \mathbf{e}_1 + y_2 \mathbf{e}_2 + y_3 \mathbf{e}_3$ får vi

$$\begin{aligned} (x, y) &\mapsto x \wedge y = (x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3) \wedge (y_1 \mathbf{e}_1 + y_2 \mathbf{e}_2 + y_3 \mathbf{e}_3) \\ &= (x_1 y_2 - x_2 y_1) \mathbf{e}_1 \wedge \mathbf{e}_2 + (x_1 y_3 - x_3 y_1) \mathbf{e}_1 \wedge \mathbf{e}_3 + (x_2 y_3 - x_3 y_2) \mathbf{e}_2 \wedge \mathbf{e}_3 \\ &\mapsto (x_1 y_2 - x_2 y_1) \mathbf{e}_3 - (x_1 y_3 - x_3 y_1) \mathbf{e}_2 + (x_2 y_3 - x_3 y_2) \mathbf{e}_1 = x \times y \end{aligned}$$

Lösningsförslag till uppgift 15.

$$\begin{aligned} (a_{11} \mathbf{e}_1 + a_{21} \mathbf{e}_2) \wedge (a_{12} \mathbf{e}_1 + a_{22} \mathbf{e}_2) &= a_{11} \mathbf{e}_1 \wedge a_{22} \mathbf{e}_2 + a_{21} \mathbf{e}_2 \wedge a_{12} \mathbf{e}_1 \\ &= (a_{11} a_{22} - a_{21} a_{12}) \mathbf{e}_1 \wedge \mathbf{e}_2 \end{aligned}$$

och

$$\begin{aligned} &(a_{11} \mathbf{e}_1 + a_{21} \mathbf{e}_2 + a_{31} \mathbf{e}_3) \wedge (a_{12} \mathbf{e}_1 + a_{22} \mathbf{e}_2 + a_{32} \mathbf{e}_3) \wedge (a_{13} \mathbf{e}_1 + a_{23} \mathbf{e}_2 + a_{33} \mathbf{e}_3) \\ &= a_{11} \mathbf{e}_1 \wedge a_{22} \mathbf{e}_2 \wedge a_{33} \mathbf{e}_3 + a_{11} \mathbf{e}_1 \wedge a_{32} \mathbf{e}_3 \wedge a_{23} \mathbf{e}_2 + a_{21} \mathbf{e}_2 \wedge a_{12} \mathbf{e}_1 \wedge a_{33} \mathbf{e}_3 \\ &\quad + a_{21} \mathbf{e}_2 \wedge a_{32} \mathbf{e}_3 \wedge a_{31} \mathbf{e}_1 + a_{31} \mathbf{e}_3 \wedge a_{12} \mathbf{e}_1 \wedge a_{23} \mathbf{e}_2 + a_{31} \mathbf{e}_3 \wedge a_{22} \mathbf{e}_2 \wedge a_{13} \mathbf{e}_1 \\ &= (a_{11} a_{22} a_{33} - a_{11} a_{32} a_{23} - a_{21} a_{12} a_{33} + a_{21} a_{32} a_{31} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13}) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3 \end{aligned}$$

Lösningsförslag till uppgift 16. Vi kan välja en ny bas för V där $\mathbf{f}_1 = \mathbf{e}_1 + \mathbf{e}_2 + \dots + \mathbf{e}_n$ och då ser vi att

$$L \left(\sum_{1 \leq i < j \leq n} a_{ij} \mathbf{f}_i \wedge \mathbf{f}_j \right) = \sum_{1 < i < j \leq n} a_{ij} \mathbf{f}_i \wedge \mathbf{f}_j \wedge \mathbf{f}_1 = \sum_{1 < i < j \leq n} a_{ij} \mathbf{f}_1 \wedge \mathbf{f}_i \wedge \mathbf{f}_j$$

Vektorerna $\mathbf{f}_1 \wedge \mathbf{f}_i \wedge \mathbf{f}_j$ för $2 \leq i < j \leq n$ är linjärt oberoende och bildar en bas för bildrummet. Alltså är rangen lika med antalet sådana basvektorer. Vi får rangen till $(n-2)(n-1)/2 = \binom{n-1}{2}$.

**MATERIAL OM KROPPSUTVIDGNINGAR OCH ÄNDLIGA KROPPAR****1. KROPPAR**

Kroppar är algebraiska objekt som generaliserar de rationella, reella och komplexa talen. Vi har mest använt oss av dessa kroppar tidigare i kursen, men mycket av det vi gjort har bara förutsatt att vi har en kropp och ibland att $1 + 1 \neq 0$ i den kroppen. Vi ska nu se att det finns andra typer av kroppar och speciellt ska vi se att det finns kroppar som bara har ett ändligt antal element, dvs *ändliga kroppar*. Först ser vi på den abstrakta definitionen.

Definition 1.1. En *kropp* är en mängd k med två binära operationer $+$ och $\cdot$ och två olika speciella element 0 och 1 som uppfyller

- (a) $(a + b) + c = a + (b + c)$, $\forall a, b, c \in k$ (*associativitet*)
- (b) $a + b = b + a$, $\forall a, b \in k$ (*kommutativitet*)
- (c) $a + 0 = 0 + a = a$, $\forall a \in k$ (*identitet*)
- (d) För alla $a \in k$ finns $b \in k$ med $a + b = 0$ (*invers*)
- (e) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$, $\forall a, b, c \in k$ (*associativitet*)
- (f) $a \cdot b = b \cdot a$, $\forall a, b \in k$ (*kommutativitet*)
- (g) $1 \cdot a = a$, $\forall a \in k$ (*identitet*)
- (h) För alla $a \neq 0$ i k finns $b \in k$ med $a \cdot b = 1$ (*invers*)
- (i) $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$, $\forall a, b, c \in k$ (*distributivitet*)

Vi skriver $-a$ för elementet b i (d) och skriver $a^{-1} = \frac{1}{a}$ för elementet b i (h).

Notera 1.2. (a)–(d) innebär att $(k, +)$ är en *abelsk grupp* med 0 som identitetselement. (e)–(h) innebär att $k^\times = k \setminus \{0\}$ är en abelsk grupp under multiplikation med 1 som identitetselement.

Definition 1.3. En *delkropp* k av en kropp K är en delmängd $k \subseteq K$ så att $0, 1 \in k$, $a + b \in k$, $ab \in k$, $-a \in k$, $c^{-1} \in k$ för alla $a, b, c \in k$ och $c \neq 0$.

Vi säger att k är *sluten* under addition, subtraktion, multiplikation och division.

Sats 1.4. Om $k \subseteq K$ är kroppar och k är en delkropp av K så är K ett vektorrum över k .

Bevis. Om vi bara multiplicerar med element från k är kraven att K är ett vektorrum över k en delmängd av kraven att K är en kropp. $\square$

Exempel 1.5. De reella talen är en delmängd av de komplexa talen och $\mathbb{C}$ är ett vektorrum av dimension 2 över $\mathbb{R}$. De rationella talen $\mathbb{Q}$ är också en delkropp av de reella talen $\mathbb{R}$ och $\mathbb{R}$ är ett oändligdimensionellt vektorrum över $\mathbb{Q}$.

Definition 1.6. En *isomorfi av kroppar* är en bijektion $\varphi: k_1 \rightarrow k_2$ som respekterar addition, multiplikation och enheter, dvs: $\varphi(a+b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$ och $\varphi(0) = 0$ och $\varphi(1) = 1$.

Det följer av övriga villkor att φ också respekterar additiva och multiplikativa inverser, dvs $\varphi(-a) = -\varphi(a)$ och $\varphi(a^{-1}) = \varphi(a)^{-1}$.

2. ÄNDLIGA KROPPAR

Definition 2.1. För heltal $n > 1$ är $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ restklasserna i $\mathbb{Z}$ vid division med n med addition och multiplikation som ges av

$$[a] + [b] = [a+b] \quad \text{och} \quad [a] \cdot [b] = [a \cdot b], \quad \forall a, b \in \mathbb{Z}.$$

Sats 2.2. Om p är ett primtal är $\mathbb{Z}_p$ en kropp.

Bevis. Heltalen $\mathbb{Z}$ är en *kommutativ ring*, dvs uppfyller alla villkoren för en kropp förutom existensen av en multiplikativ invers. Heltalen modulo p ärver också dessa egenskaper. Det återstår att visa att det finns multiplikativ inverser till nollskilda element. Multiplikation med ett nollskilt element $[a]$ ger en funktion $m_a: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$. Om $m_a([b]) = m_a([c])$ så är $[a][b-c] = 0$, dvs p delar $a(b-c)$. Eftersom p inte delar a så delar p den andra faktorn $b-c$, dvs $[b-c] = 0$, dvs $[b] = [c]$. Alltså är m_a injektiv. Eftersom $\mathbb{Z}_p$ är ändlig måste m_a också vara surjektiv och därmed finns en invers $[b]$ så att $[a][b] = m_a([b]) = [1]$. $\square$

Notera 2.3. Beviset ovan ger bara existensen av en invers, men i praktiken kan vi använda Euklides algoritm för att beräkna inversen till $[a]$ i $\mathbb{Z}_p$. Detta går relativt fort även om p är ett stort primtal, något som används i flera olika kryptosystem. Det finns också snabba algoritmer för att beräkna potenser i $\mathbb{Z}_p$. Däremot är det ett mycket svårt problem att beräkna logaritmer. Det finns alltid minst en *generator*, $[a]$, till $\mathbb{Z}_p$ så att alla nollskilda element $[b]$ i $\mathbb{Z}_p$ är potenser av $[a]$ (see Sats A.7). Att hitta m så att $[a]^m = [b]$ kallas för *diskreta logaritmproblemet* och ligger till grund för algoritmer att utbyta hemliga nycklar i kryptosystem.

I och med Sats 2.2 finns det ändliga kroppar med p element för varje primtal p . Vi kommer att se att det finns andra ändliga kroppar som inte ser ut så, men de kommer alltid att ha p^n element för något primtal p och något heltal $n > 0$. För att se det börjar vi med att se att det finns ett primtal p associerat till alla ändliga kroppar.

Sats 2.4. För varje kropp K där $1 \neq 0$ finns en minsta delkropp som är snittet av alla kroppar i K . Denna delkropp är isomorf med $\mathbb{Q}$ eller $\mathbb{Z}_p$ för något primtal p .

Bevis. Snittet av delkroppar är en ny delkropp och därmed måste snittet av alla delkroppar i K vara en kropp $k \subseteq K$ som innehåller 1 och därmed $1 + 1 + \dots + 1$. Om detta ger oändligt många element så finns alla heltal i K och därmed också alla rationella tal. Då är $k \cong \mathbb{Q}$. Om det bara finns ändligt många element på formen $1 + 1 + \dots + 1$ måste vi ha $n \cdot 1 = 1 + 1 + \dots + 1 = 0$ för något n . Om p är det minsta positiva heltal som uppfyller detta får vi att p måste vara ett primtal eller $p = 1$ eftersom $p = r \cdot s$ ger att $(r \cdot 1)(s \cdot 1) = 0$ och då måste antingen $r \cdot 1 = 0$ eller $s \cdot 1 = 0$ vilket motsäger att p var minst. Eftersom vi antagit att $1 \neq 0$ måste p vara ett primtal och elementen på formen $n \cdot 1$ bildar en delkropp isomorf med $\mathbb{Z}_p$. $\square$

Definition 2.5. Den minsta delkropp av K som innehåller 1 kallas *primkroppen* i K .

Korollarium 2.6. En ändlig kropp har ordning p^n för något primtal p och något heltal $n \geq 1$.

Bevis. Primkroppen $k \subseteq K$ är isomorf med $\mathbb{Z}_p$ för något primtal p enligt Sats 2.4. Enligt Sats 1.4 är K ett vektorrum över k och eftersom K är ändlig måste $K \cong k^n$ för något naturligt tal n . Därmed är antalet element p^n . $\square$

Definition 2.7. Om primkroppen till K är $\mathbb{Z}_p$ säger vi att K har *karakteristik* p . Om primkroppen till K är $\mathbb{Q}$ säger vi att K har *karakteristik* 0.

Följande sats är väldigt användbar när vi gör beräkningar med en kropp där karakteristiken är p .

Sats 2.8. I en kropp med karakteristik p gäller att $(a+b)^p = a^p + b^p$ för alla a, b i K .

Bevis. I binomialutvecklingen är alla koefficienter delbara med p utom den första och den sista eftersom

$$\binom{p}{i} = \frac{p!}{i!(p-i)!}$$

och p delar $p!$ men delar inte vare sig $i!$ eller $(p-i)!$ då $1 \leq i \leq p-1$. $\square$

Korollarium 2.9. I kroppen $\mathbb{Z}_p$ gäller

- (a) (Fermats lilla sats) $a^p = a$ för alla $a \in \mathbb{Z}_p$.
- (b) (Eulers sats) $a^{p-1} = 1$ för alla $a \in \mathbb{Z}_p \setminus \{0\}$

Bevis. (a) Vi använder induktion. Satsen gäller för $a = 0$. Antag att satsen gäller för a . Då är $(a+1)^p = a^p + 1^p = a + 1$ enligt den förra satsen och induktionsantagandet. Satsen gäller alltså även för $a+1$ och därmed för alla a .

(b) Om $a \neq 0$ så är a inverterbar så (b) följer från (a). $\square$

För en kropp med $q = p^n$ element gäller mer allmänt att $a^q = a$ (se Sats A.1).

3. KROPPSUTVIDGNINGAR

Vi såg i Korollarium 2.6 att varje ändlig kropp har p^n element. För att se att det faktiskt finns kroppar av alla storlekar som är p^n för något primtal p och något heltal $n > 0$ behöver vi en konstruktion. Här kommer vi att göra det med hjälp av matriser eftersom det sättet drar nytta av vad vi har lärt oss om matriser och minimalpolynom för matriser.

Notera 3.1. Alla ändliga kroppar går att representera med matriser på grund av Sats 2.4 och Sats 1.4. Vi kan välja en bas $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ för kroppen K som vektorrum över primkroppen $k = \mathbb{Z}_p$. För ett element $a \in K$ låter vi $m_a: K \rightarrow K$ vara multiplikation med a , dvs $m_a(x) = ax$. Detta är en linjär avbildning. I den valda basen ger detta en $n \times n$ -matris A_a . Avbildningarna m_a och matriserna A_a uppfyller precis samma relationer som elementen i K eftersom $m_a \circ m_b = m_{ab}$ och alltså $A_a A_b = A_{ab}$. Vi har därmed representerat K som en mängd S av $n \times n$ -matriser med koefficienter i primkroppen $k = \mathbb{Z}_p$. Mängden S med addition och multiplikation av matriser är alltså en kropp isomorf med kroppen K .

Vi ska nu se att det verkligen går att konstruera ändliga kroppar som är större än $\mathbb{Z}_p$ om vi kan hitta *irreducibla polynom* med koefficienter i $\mathbb{Z}_p$. Man kan också visa att det alltid finns irreducibla polynom (Anmärkning A.4) och att det upp till isomorfi finns en *unik* kropp med p^n element (Sats A.5).

Definition 3.2. Ett polynom $p(x)$ i $k[x]$ är *irreducibelt* om det inte kan skrivas som en produkt av två polynom av lägre grad.

Sats 3.3. Om A är en matris i $M_n(k)$ och minimalpolynomet $q_A(x)$ till A är irreducibelt är $K = \{p(A) : p(x) \in k[x]\} \subseteq M_n(k)$ en kropp med addition och multiplikation av matriser som operationer.

Bevis. En bas för K ges av $\{I, A, A^2, \dots, A^{d-1}\}$ där $d = \deg q_A(x)$. Observera att multiplikation i K är kommutativ eftersom matriserna $p_1(A)$ och $p_2(A)$ kommuterar med varandra för alla polynom $p_1(x), p_2(x) \in k[x]$. Detta gör K till en kommutativ ring. Det återstår att visa att alla element i K som inte är noll har en multiplikativ invers. Ett sådant element kan skrivas som $p(A)$ där $0 \leq \deg p(x) < d$. Om det existerar ett noll-skilt element $p(A)$ som inte är inverterbart måste det finnas ett med $p(x)$ av lägsta grad. Det är klart att $\deg p(x) > 0$ eftersom $p(x) = b$ med $b \in k \setminus \{0\}$ ger matrisen $p(A) = bI$ med invers $b^{-1}I$. Då kan vi använda polynomdivision för att skriva

$$q_A(x) = p(x)s(x) + r(x)$$

där $r(x) = 0$ eller $0 \leq \deg r(x) < \deg p(x)$. Vi har att $r(x) \neq 0$ eftersom $q_A(x)$ är irreducibelt och $0 < \deg p(x) < \deg q_A(x)$. Eftersom $\deg r(x) < \deg p(x)$ är alltså $r(A)$ inverterbart i K och vi får att

$$0 = p(A)s(A) + r(A) \iff (-r(A))^{-1}s(A)p(A) = I$$

vilket visar att $p(A)$ är inverterbart. Alltså är K en kropp. $\square$

Exempel 3.4. Polynomet $q(x) = x^2 + 1$ är irreducibelt över de reella talen och matrisen

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

har $q(x)$ som minimalpolynom.

$$K = \{p(J) : p(x) \in \mathbb{R}[x]\} = \{aI + bJ : a, b \in \mathbb{R}\}$$

är en kropp som är isomorf med $\mathbb{C}$.

Exempel 3.5. Polynomet $q(x) = x^2 + x + 1$ är irreducibelt över $\mathbb{Z}_2$ ty om det vore reducibelt så hade $q(x)$ haft en linjär faktor och därmed en rot men $q(0) = 1$ och $q(1) = 1$ i $\mathbb{Z}_2$. Matrisen

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

har $q(x)$ som minimalpolynom.

$$K = \{p(J) : p(x) \in \mathbb{Z}_2[x]\} = \{aI + bJ : a, b \in \mathbb{Z}_2\} = \left\{ \begin{bmatrix} a & b \\ b & a+b \end{bmatrix} : a, b \in \mathbb{Z}_2 \right\}$$

är en kropp med fyra element.

$$K = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\} = \{0, 1, \alpha, 1 + \alpha\}$$

med addition och multiplikation enligt

+	0	1	α	$\alpha + 1$	·	0	1	α	$\alpha + 1$
0	0	1	α	$\alpha + 1$	0	0	0	0	0
1	1	0	$\alpha + 1$	α	1	0	1	α	$\alpha + 1$
α	α	$\alpha + 1$	0	1	α	0	α	$\alpha + 1$	1
$\alpha + 1$	$\alpha + 1$	α	1	0	$\alpha + 1$	0	$\alpha + 1$	1	α

REFERENSER

- [Cur90] Morton L. Curtis. *Abstract linear algebra*. Universitext. Springer-Verlag, New York, 1990. With revisions by Paul Place, With a preface by John Hempel.
- [Gri07] Pierre Antoine Grillet. *Abstract algebra*, volume 242 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2007.
- [Wik17a] Wikipedia. Euler's totient function — Wikipedia, The Free Encyclopedia, 2017.
- [Wik17b] Wikipedia. Finite field — Wikipedia, The Free Encyclopedia, 2017.
- [Wik17c] Wikipedia. Lagrange's theorem (group theory) — Wikipedia, The Free Encyclopedia, 2017.
- [Wik17d] Wikipedia. Möbius inversion formula — Wikipedia, The Free Encyclopedia, 2017.

ÖVNINGAR PÅ ÄNDLIGA KROPPAR OCH KROPPSUTVIDGNINGAR

Uppgift 1. Beräkna inversen till 4 i $\mathbb{Z}_{11}$.

Uppgift 2. Lös om möjligt ekvationen $x^2 - 4x + 5 = 0$ i $\mathbb{Z}_{11}$ respektive i $\mathbb{Z}_{13}$.

Uppgift 3. Visa att $\mathbb{C}$ kan fås som 2×2 -matriser med koefficienter i $\mathbb{R}$ med hjälp av en godtycklig 2×2 -matris J med karakteristiskt polynom som saknar reella nollställen.

Uppgift 4. Hitta ett irreducibelt andragradspolynom med koefficienter i $\mathbb{Z}_3$ och använd det för att konstruera en kropp med nio element som 2×2 -matriser över $\mathbb{Z}_3$.

Uppgift 5 (Uppgift 8 på tentamen 2018-01-10). Låt K vara delmängden av $M_2(\mathbb{Z}_7)$ som ges av

$$K = \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} : a, b \in \mathbb{Z}_7 \right\}.$$

Visa att K bildar en kropp med 49 element och lös ekvationerna $x^2 + 2x + 2 = 0$ respektive $x^6 = 1$ i K . (4 p)

Uppgift 6 (Uppgift 1 på tentamen 2018-04-05). Låt $k = \{0, 1, a, 1 + a\}$ vara kroppen med fyra element.

- (a) Skriv upp additions- och multiplikationstabellerna för k . (2 p)
- (b) Bestäm en matris A med koefficienter i $\mathbb{Z}_2$ sådan att k kan representeras av matriserna $\{0, I, A, A + I\}$. (2 p)

Uppgift 7 (Uppgift 1 på modelltentamen 2017). Följden $\{x_i\}_{i \geq 0}$ definieras genom den linjära rekursionen $x_{i+2} = x_{i+1} + x_i$, för $i \geq 0$ och $x_0 = x_1 = 1$. Bestäm x_{2017} om räkningarna sker i $\mathbb{Z}_{11}$. (4 p)

Uppgift 8 (Uppgift 1 på tentamen 2018-05-02). Låt $k = \{0, 1, -1\}$ vara kroppen med tre element.

- (a) Bestäm alla andragradspolynom $x^2 + ax + b$ med koefficienter i k som inte har några nollställen i k . (2 p)
- (b) Bestäm en matris A sådan att matriserna $\{aI + bA : a, b \in k\}$ bildar en kropp med nio element. (2 p)

BILAGA A. UNICITET AV ÄNDLIGA KROPPAR (INGÅR EJ)

Fermats lilla sats säger att $a^p = a$ i $\mathbb{Z}_p$. Vi har följande generalisering till alla ändliga kroppar.

Sats A.1. *I en kropp K av ordning $q = p^n$ gäller att $a^q = a$ för alla element a i K .*

Bevis. $K^\times = K \setminus \{0\}$ är en abelsk grupp av ordning $q - 1$ och enligt Lagranges sats¹ har vi att $x^{q-1} = 1$ för alla $x \in K^\times$. När vi multiplicerar med x får vi att $x^q = x$ även gäller för $x = 0$, dvs för alla $x \in K$. $\square$

Sats A.2. *Polynomet $x^{p^n} - x$ är en produkt av alla irreducibla moniska polynom av grad som delar n där varje faktor förekommer precis en gång.*

Bevis. För varje irreducibelt polynom av grad d som delar n får vi en ändlig kropp av ordning p^d enligt konstruktionen ovan. Alla element i denna kropp uppfyller $a^{p^d} = a$, vilket gör att minimalpolynomet vi använde måste vara en delare i $x^{p^d} - x$ som i sin tur delar $x^{p^n} - x$ eftersom d är en delare i n . Om $q(x)$ är ett irreducibelt polynom av grad m som delar $x^{p^n} - x$ kan vi konstruera en kropp av ordning p^m där alla element uppfyller både $a^{p^n} = a$ och $a^{p^m} = a$, men det betyder att $x^{p^m} - x$ delar $x^{p^n} - x$ och det innebär att m delar n .

Om $f(x) = x^{p^n} - x$ får vi $f'(x) = -1$ och därmed kan inte $f(x) = x^{p^n} - x$ ha några upprepade faktorer. $\square$

Vi kan använda detta till att räkna antalet irreducibla polynom av grad d över $\mathbb{Z}_p$. Låt $i(p, d) = \#\{\text{moniska irreducibla polynom av grad } d \text{ i } \mathbb{Z}_p[x]\}$.

Sats A.3. $\sum_{d|n} d \cdot i(p, d) = p^n$

Bevis. Båda sidorna beräknar graden av polynomet $x^{p^n} - x$ enligt Sats A.2. $\square$

Notera A.4. Vi kan nu använda Möbius inversionsformel [Wik17d] för att beräkna $i(p, d)$ och sluta oss till att det finns irreducibla polynom av alla grader.

Sats A.5. *För varje primtalspotens $q = p^n$ finns precis en ändlig kropp av ordning q upp till isomorfi.*

Bevis. Vi har sett att det finns ett irreducibelt polynom av grad n över $\mathbb{Z}_p$ och vi kan bilda en matris med koefficienter i $\mathbb{Z}_p$ vars minimalpolynom ges av detta. Därmed har vi en kropp av ordning $q = p^n$. I denna kropp har alla irreducibla polynom över $\mathbb{Z}_p$ av grad som delar n nollställen.

Om K är en kropp av ordning $q = p^n$ uppfyller alla element i K ekvationen $x^q - x = 0$ och vi kan speciellt hitta ett element α som är ett nollställe till ett av de irreducibla polynomen av grad n . Vi ser då att $\{1, \alpha, \dots, \alpha^{n-1}\}$ utgör en bas för K som vektorrum över $k = \mathbb{Z}_p$ och därmed kan vi representera K som matriserna $\{f(A) : f(x) \in \mathbb{Z}_p[x]\}$. Därmed är K isomorf med den kropp vi konstruerat med hjälp av det irreducibla polynomet. $\square$

Definition A.6. Den kropp som har ordning $q = p^n$ betecknas $\text{GF}(q)$ (en: *Galois Field*) eller $\mathbb{F}_q$.

Sats A.7. *Den multiplikativa gruppen i en kropp av ordning q är cyklisk, dvs $K^\times = \{1, a, a^2, \dots, a^{q-2}\}$ för något a i K .*

¹En grundläggande sats inom gruppteori. Se [Wik17c]

Bevis. Alla element uppfyller $x^m = 1$ för något m och antalet element som uppfyller $x^m = 1$ är högst m eftersom det är en polynomekvation av grad m . Av dessa uppfyller högst $\varphi(m)$ inte $x^{m'} = 1$ för någon delare m' i m , där φ är Eulers φ -funktion [Wik17a].

På grund av Lagranges sats måste m vara en delare i $q - 1$ och vi får

$$q-1 \leq \sum_{m|(q-1)} \varphi(m) = q-1$$

och vi måste ha likhet i alla termer, dvs $\phi(m)$ element som uppfyller $x^m = 1$, men inte $x^{m'}$ för någon delare i m . Speciellt finns element som uppfyller $x^{q-1} = 1$, men inte $x^m = 1$ för någon delare m i $q-1$. $\square$

Notera A.8. Ett element a som i satsen kallas för en *generator*.

Uppgift 9. Använd Sats A.3 för att beräkna antalet moniska irreducibla i $\mathbb{Z}_p[x]$ av grad 1, 2, 3 respektive 4.

BILAGA B. LINJÄRA REKURSIONSFÖLJDER I ÄNDLIGA KROPPAR (INGÅR EJ)

Följande sekvens är genererad från en linjär rekursionsekvation över $\mathbb{Z}_2$ där $x_{i+10} = x_{i+7} + x_i$, för $i = 0, 1, 2, \dots$. Den är periodisk med en period på 1023 och innehåller alla sekvenser av 10 ettor och nollor utom $\{0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}$ precis en gång per period.

[illegible]

En linjär rekursionsföljd kan som vi tidigare sett hanteras genom potenser av matriser. Om

$$x_n = a_{m-1}x_{n-1} + \cdots + a_0x_{n-m}$$

för alla $n \geq m$ med begynnelsedata $x_0, x_1, \dots, x_{m-1}$ så betraktar vi vektorerna $\mathbf{y}(n) = (x_n, x_{n-1}, \dots, x_{n-m+1})$ och rekursionen kan skrivas som $\mathbf{y}(n) = A\mathbf{y}(n-1) = A^{n-m}\mathbf{y}(m)$

för alla $n \geq m$.

$$\underbrace{\begin{bmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-m+2} \\ x_{n-m+1} \end{bmatrix}}_{\mathbf{y}(n)} = \underbrace{\begin{bmatrix} a_{m-1} & a_{m-2} & \cdots & a_1 & a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_{n-1} \\ x_{n-2} \\ \vdots \\ x_{n-m+1} \\ x_{n-m} \end{bmatrix}}_{\mathbf{y}(n-1)}$$

Det *karaktelistiska polynomet* för rekursionen är det karakteristiska polynomet för A . Genom radutveckling kring den första raden får vi att

$$p_A(x) = x^m - a_{m-1}x^{m-1} - a_{m-2}x^{m-2} - \cdots - a_1x - a_0$$

Det betyder att $I, A, A^2, A^3, \dots$ uppfyller samma rekursionsekvation som $x_0, x_1, \dots$:

$$A^n = a_{m-1}A^{n-1} + \cdots + a_0A^{n-m}$$

för alla $n \geq m$.

Sats B.1. Låt $x_n = a_{m-1}x_{n-1} + \cdots + a_0x_{n-m}$ vara en linjär rekursionsekvation över $\mathbb{Z}_p$ med begynnelsedata $x_0, x_1, \dots, x_m$, dvs a_j och x_i är tal i $\mathbb{Z}_p$. Antag att det karakteristiska polynomet för rekursionen är irreducibelt. Då är $K = \text{Span}\{I, A, A^2, \dots, A^{m-1}\}$ en kropp enligt Sats 3.3 och följden kan skrivas som

$$x_i = f(A^i)$$

där $f: K \rightarrow \mathbb{Z}_p$ är en unik linjär avbildning.

Bevis. Matriserna $y_0 = I, y_1 = A, y_2 = A^2, \dots$ kan nu betraktas som tal i kroppen K och uppfyller som vi såg ovan rekursionsekvationen

$$y_n = a_{m-1}y_{n-1} + \cdots + a_0y_{n-m}$$

i K med begynnelsedata $y_i = A^i$ för $i = 0, 1, \dots, m-1$. Därmed uppfyller också $x_i = f(y_i)$ rekursionsekvationen för varje linjär funktion $f: K \rightarrow \mathbb{Z}_p$. Det återstår att visa att varje följd x_i fås på detta sätt. Det finns p^n linjära funktioner $f: K \rightarrow \mathbb{Z}_p$ och det finns lika många begynnelsedata för följden. Om en funktion ger följden som är konstant noll måste f vara noll eftersom $\{y_0, y_1, \dots, y_{n-1}\}$ är linjärt oberoende. Alltså svarar funktionerna $f: K \rightarrow \mathbb{Z}_p$ bijektivt mot följderna. $\square$

Definition B.2. Ett irreducibelt polynom av grad n i $\mathbb{Z}_p$ kallas *primitivt* om den följd som ges av $1, x, x^2, \dots$ har period $p^n - 1$.

Uppgift 10. $\{1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, \dots\}$ är början på en linjär rekursionsföljd av ordning 5 över $\mathbb{Z}_2$.

- Hur fortsätter följden?
- Vilken period har följden?

LÖSNINGSFÖRSLAG TILL ÖVNINGAR PÅ ÄNDLIGA KROPPAR OCH
KROPPSUTVIDGNINGAR

Lösningförslag till uppgift 1. Vi använder Euklides algoritm och får

$$11 = 2 \cdot 4 + 3, \quad 4 = 1 \cdot 3 + 1$$

och baklänges ger det

$$1 = 4 - 1 \cdot 3 = 4 - 1 \cdot (11 - 2 \cdot 4) = 3 \cdot 4 - 1 \cdot 11$$

vilket visar att 3 är inversen till 4 i $\mathbb{Z}_{11}$.

Om vi hade tillåtit att använda en negativ rest om den ligger närmre 0 hade det räckt med ett steg i och med att $11 = 3 \cdot 4 - 1$.

Lösningförslag till uppgift 2. Vi kan skriva om ekvationen som $(x-2)^2 = -1$. Vi behöver se om -1 är en kvadrat i $\mathbb{Z}_{11}$, respektive i $\mathbb{Z}_{13}$. I $\mathbb{Z}_{11}$ är kvadraterna 0, 1, 4, 9, 5, 3 och i $\mathbb{Z}_{13}$ är kvadraterna 0, 1, 4, 9, 3, 12, 10. Alltså är $-1 = 10$ inte en kvadrat i $\mathbb{Z}_{11}$ men $-1 = 12 = 5^2$ är en kvadrat i $\mathbb{Z}_{13}$. I det senare fallet får vi lösningarna $x = 2 \pm 5$ till ekvationen $x^2 - 4x + 5 = 0$ i $\mathbb{Z}_{13}$.

Lösningförslag till uppgift 3. Om J saknar reella egenvärden betyder det att egenvärdena är ett komplext konjugerat par. Därmed innehåller kroppen vi får som $K = \{xI + yJ : x, y \in \mathbb{R}\}$ de två talen $a \pm ib$ där $b \neq 0$. Eftersom det är ett vektorrum över $\mathbb{R}$ får vi att $(a + ib) - (a - ib) = 2ib$ tillhör kroppen och därmed också $1/(2b) \cdot 2ib = i$. Därmed är K isomorf med de komplexa talen $\mathbb{C}$.

Lösningförslag till uppgift 4. Polynomet $x^2 + 1$ är irreducibelt i över $\mathbb{Z}_3$ eftersom de två nollskilda elementen i $\mathbb{Z}_3$ båda har kvadraten 1. Vi kan därmed konstruera en kropp med nio element som 2×2 -matriserna $K = \{aI + bJ : a, b \in \mathbb{Z}_3\}$ där

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

Med andra ord är

$$K = \left\{ \begin{bmatrix} a & b \\ 2b & a \end{bmatrix} : a, b \in \mathbb{Z}_3 \right\}.$$

Lösningförslag till uppgift 5. Enligt en sats i materialet om ändliga kroppar räcker det att se att minimalpolynomet till matrisen

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

är irreducibelt. Vi har att $\{A, I\}$ är linjärt oberoende, men $A^2 = -I$. Därmed är minimalpolynomet $q_A(x) = x^2 + 1$. Detta är irreducibelt om det inte finns några linjära faktorer, dvs inte några nollställen. Vi provar alla element i $\mathbb{Z}_7$ och får att

$$q_A(0) = 1, q_A(1) = 2, q_A(2) = 5, q_A(3) = 3, q_A(4) = 3, q_A(5) = 5 \text{ och } q_A(6) = 2.$$

Eftersom det inte fanns något nollställe är $q_A(x)$ irreducibelt och K är en kropp med 49 element då det finns 7×7 möjligheter för a och b i $\mathbb{Z}_7$.

Vi kan lösa den första ekvationen genom att kvadratkomplettera och får då

$$x^2 + 2x + 2 = 0 \iff (x+1)^2 + 1 = 0$$

Eftersom $A^2 = -I$ har vi att $x = -I \pm A$ är en lösningarna till ekvationen. Därmed kan vi skriva de båda lösningarna som

$$\begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \quad \text{och} \quad \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}.$$

För att lösa den andra ekvationen kan vi konstatera att $a^6 = 1$ för varje nollskilt element a i primkroppen $k = \mathbb{Z}_7$. Detta måste vara alla lösningar eftersom $x^6 = 1$ kan ha högst sex lösningar. Vi kan också faktorisera polynomet $x^6 - 1$ som

$$x^6 - 1 = (x^3 - 1)(x^3 + 1) = (x - 1)(x^2 + x + 1)(x + 1)(x^2 - x + 1)$$

och kvadratkomplettera $x^2 + x + 1 = (x + 4)^2 - 1 = (x + 5)(x + 3)$ och $x^2 - x + 1 = (x + 3)^2 - 1 = (x + 4)(x + 2)$. Därmed ges lösningarna av $I, 2I, 3I, 4I, 5I$ och $6I$

Lösningsförslag till uppgift 6. (a) Eftersom $4 = 2^2$ måste vi ha att k är ett vektorrum över $\mathbb{Z}_2$ och därmed $1 + 1 = 0$. Detta ger additionstabellen

+	0	1	a	$a + 1$
0	0	1	a	$a + 1$
1	1	0	$a + 1$	a
a	a	$a + 1$	0	1
$a + 1$	$a + 1$	a	1	0

Eftersom $a \neq 0$ är a inverterbar inversen kan inte vara 0 eller 1, varför vi måste ha $a^2 = 1$ eller $a(a + 1) = 1$. Om $a^2 = 1$ får vi $(a + 1)^2 = a^2 + 1 = 0$, som inte stämmer eftersom $a + 1$ också ska vara inverterbara. Alltså är $a(a + 1) = 1$ och vi får $a^2 = a + 1$ och $(a + 1)^2 = a$. Multiplikationstabellen blir därmed

·	0	1	a	$a + 1$
0	0	0	0	0
1	0	0	a	$a + 1$
a	0	a	$a + 1$	1
$a + 1$	0	$a + 1$	1	a

(b) Vi behöver en matris A som uppfyller $A^2 = A + I$ och därmed kan vi ta någon av matriserna

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{eller} \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Eftersom båda dessa har $x^2 + x + 1$ som minimalpolynom kommer $K = \{0, I, A, A + I\}$ att vara isomorf med k .

Lösningsförslag till uppgift 7. Vi kan lösa rekursionen med hjälp av

$$\begin{bmatrix} x_{i+2} \\ x_{i+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{i+1} \\ x_i \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{i+1} \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}.$$

För att hantera detta försöker vi diagonalisera matrisen. Den karakteristiska ekvationen är

$$\det \left(\begin{bmatrix} x-1 & -1 \\ -1 & x \end{bmatrix} \right) = 0 \quad \Longleftrightarrow \quad x^2 - x - 1 = 0.$$

Vi kan lösa ekvationen i $\mathbb{Z}_{11}$ genom att prova de elva möjliga rötterna eller genom att använda $x = 1/2 \pm \sqrt{(1/2)^2 + 1} = 6 \pm \sqrt{6^2 + 1} = 6 \pm \sqrt{4} = 6 \pm 2$. Vi får alltså egenvärdena 4 och 8 och vi kan sluta oss till att lösningen kan skrivas som

$$x_i = \alpha 4^i + \beta 8^i$$

Begynnelsevärdena ger ekvationssystemet

$$\begin{cases} \alpha + \beta = 1 \\ 4\alpha + 8\beta = 1 \end{cases}$$

som har lösningen $\alpha = 7/4 = 7 \cdot 3 = 21 = 10$ och $\beta = -3/4 = -9 = 2$. Därmed är $x_i = 2 \cdot 8^i - 4^i$, för $i \geq 0$. Eulers sats ger att $x^{10} = 1$ för alla $x \in \mathbb{Z}_{11} \setminus \{0\}$, så vi får att $x_{2017} = x_7$. Vi får att $\{x_i\}_{i=0}^7 = \{1, 1, 2, 3, 5, 8, 2, 10\}$. Alltså är $x_{2017} = 10$.

Lösningsförslag till uppgift 8. (a) De nollställena som kan förekomma är 0, -1 och 1 och sätter vi in dessa får vi att $b \neq 0$, $1 - a + b \neq 0$ och $1 + a + b \neq 0$. Vi har två möjligheter för b , $b = -1$ och $b = 1$. För $b = -1$ får vi att $-a \neq 0$ och $a \neq 0$, vilket ger två lösningar. För $b = 1$ får vi $2 - a \neq 0$ och $2 + a \neq 0$, dvs $a \neq -1$ och $a \neq 1$. Alltså måste $a = 0$ om $b = 1$. Vi fick därmed tre möjliga polynom utan nollställena

$$x^2 - x - 1, \quad x^2 + x - 1 \quad \text{och} \quad x^2 + 1.$$

(b) Vi behöver hitta en matris som uppfyller en av de tre polnomen ovan. Om vi väljer $x^2 + 1$ kan vi välja

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

och får då att

$$\left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} : a, b \in k \right\}$$

bildar en kropp med nio element eftersom det karaktäristiska polynomet $x^2 + 1$ är irreducibelt.

Lösningsförslag till uppgift 9. Sats A.3 säger att

$$\sum_{d|n} d \cdot i(p, d) = p^n$$

där $i(p, d)$ är antalet irreducibla polynom av grad d över $\mathbb{Z}_p$. För $n = 1$ får vi $i(p, 1) = p^1 = p$. För $n = 2$ får vi $1 \cdot i(p, 1) + 2 \cdot i(p, 2) = p^2$ vilket ger $i(p, 2) = (p^2 - p)/2$. För $n = 3$ får vi $1 \cdot i(p, 1) + 3 \cdot i(p, 3) = p^3$, vilket ger $i(p, 3) = (p^3 - p)/3$. För $n = 4$ får vi $1 \cdot i(p, 1) + 2 \cdot i(p, 2) + 4 \cdot i(p, 4) = p^4$, vilket ger $i(p, 4) = (p^4 - p(p-1) - p)/4 = (p^4 - p^2)/4$.

Lösningsförslag till uppgift 10. Vi behöver se på relationer mellan kolonnerna till matrisen vi får genom att ta

$$\begin{bmatrix} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\ x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ x_4 & x_5 & x_6 & x_7 & x_8 & x_9 \\ x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

vilket efter Gausselimination ger

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

och nollrummet spänns upp av vektorn $[1 \ 0 \ 1 \ 0 \ 0 \ 1]^T$. Alltså ges rekursionen av $x_{i+5} = x_{i+2} + x_i$ och vi kan fortsätta följderna med hjälp av detta och får

1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, ...

Följden upprepas när vi kommit tillbaka till startläget och perioden blir 31. Eftersom vi vet att det bara finns 31 nollskilda element i en kropp med $2^5 = 32$ element kan perioden bli maximalt 31. Eftersom 31 är ett primtal kan det inte bli någon kortare period om inte perioden är 1 och följderna är konstant 0.