Teil I

Lineare Algebra und Analytische Geometrie

Vorlesung

Wintersemester 1999/2000

Volker Mehrmann

Übung/Seminar: **Matthias Pester**
Uwe Schrader
Andreas Steinbrecher

Lineare Algebra und Analytische Geometrie I

Vorlesung
TU Chemnitz

Volker Mehrmann, Chemnitz
Rh 41/614, Tel.: 531 8367
email: mehrmann@mathematik.tu-chemnitz.de

Übung/Seminar:

Matthias Pester, Rh 41/617, Tel.: 531 2656
Uwe Schrader, Rh 41/611, Tel.: 531 2708
Andreas Steinbrecher, Rh 41/612, Tel.: 531 3953

Organisatorische Details:

Vorlesung:

- Vermittlung des Stoffes

Seminar:

- Nutzung von mathematischer Software
- lernen, „in Mathematik“ zu reden und anderen zu erklären
- wichtige Inhalte, die sonst nicht im Stoff vorkommen
- Beweistechniken

Übung:

- Diskussion und Besprechung der Übungsaufgaben.
- Mathematik lernt man am besten durch selbermachen.

Kapitel 0

Motivation

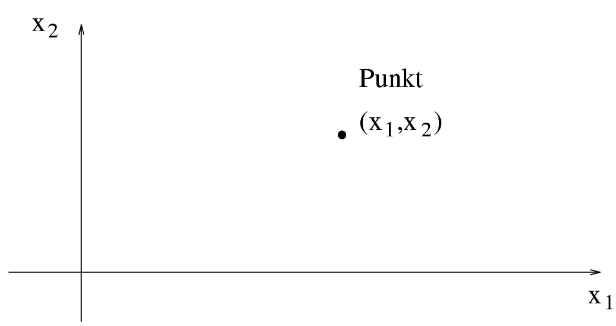
In diesem Grundkurs Lineare Algebra beschäftigen wir uns mit einem Themenkreis, der einige wesentliche Gesichtspunkte der Mathematik umfasst. Er liefert

- die Sprache und das Handwerkszeug für viele Bereiche der Mathematik, aber auch inzwischen aller Ingenieurwissenschaften, Naturwissenschaften, Wirtschaftswissenschaften, Informatik
- die Grundlage für die abstrakte moderne Mathematik, die in der Lage ist, abstrahiert von einem realen Problem Fortschritte im mathematischen Kalkül zu schaffen, aber dann auch diese wiederum auf die Praxis anzuwenden.

Ein einfaches Beispiel aus der Wirtschaft:

Beispiel 0.1 Ein Betrieb produziert zwei Produkte P_1, P_2. Produkt P_i kostet a_i DM an Rohstoffen und b_i DM an Arbeitslohn. Damit kann ein Gewinn von q_i erzielt werden, für $i = 1, 2$.
Insgesamt stehen a DM an Kapital und b Arbeitslohneinheiten zur Verfügung.

Jedes denkbare Produktionsprogramm ist von der Form x_1 Einheiten von P_1 und x_2 Einheiten von P_2.
Man kann geometrisch jedes Produktionsprogramm als Zahlenpaar x_1, x_2 darstellen.



Es sind natürlich nur solche Produktionsprogramme erlaubt, die man mit den vorhandenen Ressourcen auch erzielen kann, d.h.,

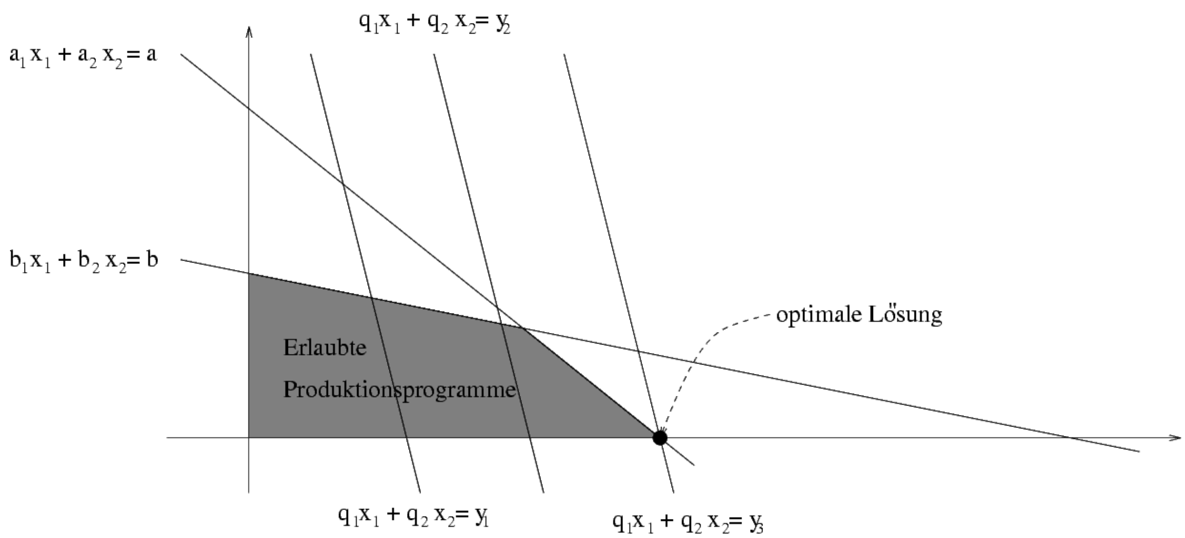
$$\begin{aligned} a_1x_1 + a_2x_2 &\leq a, \\ b_1x_1 + b_2x_2 &\leq b. \end{aligned}$$

Ziel der Aufgabe ist die Gewinnmaximierung, d.h., man sucht ein Maximum der Funktion

$$\Phi(x_1, x_2) = q_1 x_1 + q_2 x_2.$$

Wie kann man dieses Maximum finden?

Beobachtung:



Wenn $q_1 x_1 + q_2 x_2 = y$ ist, so hat man den Gewinn y. Für feste y_i sind das parallele Geraden. Verschiebt man also diese Parallelen, bis man an die Ecke mit dem maximalen y kommt, so hat man das Problem gelöst.

$\Longrightarrow$ „Lineare Programmierung“, Allgemeine Theorie linearer Gleichungen und Ungleichungen.

Ein Beispiel aus der Mechanik.

Beispiel 0.2 Gleichgewichtslage

Eine Masse m sei mit Hilfe von Federn im dreidimensionalen Raum aufgehängt. Das Gleichgewicht sei im Punkt

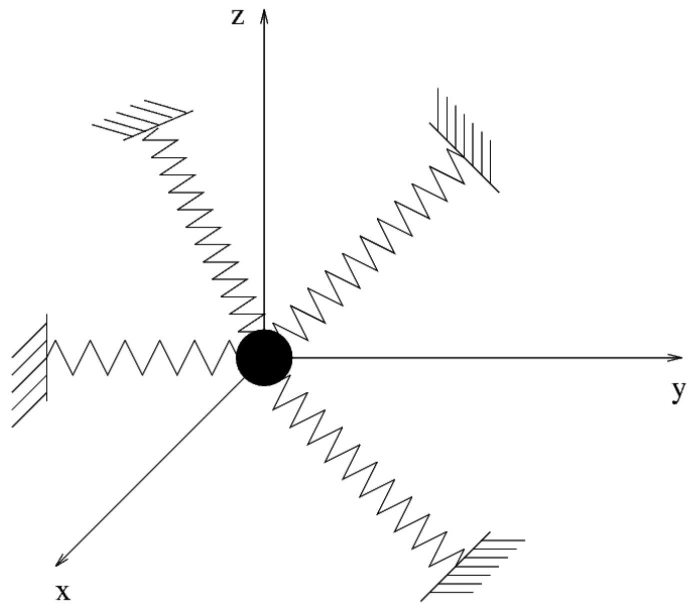
$$(x, y, z) = (0, 0, 0).$$

Ist das Gleichgewicht stabil? Um das zu entscheiden, betrachten wir $\triangle V$, die Veränderung der potentiellen Energie, wenn m von $(0, 0, 0)$ aus in einen anderen Punkt $(\tilde{x}, \tilde{y}, \tilde{z})$ gebracht wird. Abhängig von den Größen der Federkonstanten ergibt sich

$$\triangle V = a_0 x^2 - a_1 xy + a_2 xz + a_3 y^2 - a_4 yz + a_5 z^2,$$

z.B.

$$\triangle V = x^2 - 4xy + 2xz + 3y^2 - 2yz + 4z^2.$$



Durch quadratische Ergänzung bekommen wir

$$\begin{aligned}\triangle V &= (x-2y+z)^2 - y^2 + 2yz + 3z^2 \\ &= (x-2y+z)^2 - (y-z)^2 + 4z^2.\end{aligned}$$

Wir erhalten lauter Quadrate, aber eines davon mit negativem Vorzeichen. Damit kann $\triangle V < 0$ sein, z.B. für $(x, y, z) = (2, 1, 0)$. Damit ist das Gleichgewicht für diese Federkonstanten instabil.

$\Longrightarrow$ „Polynomielle Gleichungen“, „Summen von Quadraten“.

Beispiele: siehe C. Blatter, Lineare Algebra für Ingenieure, Mechaniker und Naturwissenschaftler, VDI Verlag, Zürich, 1989.

Man kann noch viele weitere Beispiele anführen, und wir werden noch viele im Text haben, aber die Beispiele sind für uns die Motivation, nicht das Ziel. Wir wollen eine allgemeine Theorie entwickeln, die nicht nur für ein spezielles Problem, sondern für viele Probleme gleichermaßen anwendbar ist. Dazu brauchen wir eine abstrakte Sprache, „(Lineare) Algebra“, und einen mathematischen Kalkül.
Damit werden wir dann sofort loslegen und das wird teilweise sehr losgelöst sein von irgendwelchen konkreten Objekten. Aber wir werden immer wieder Beispiele und reale Objekte betrachten, und unsere Theorie darauf anwenden.

Kapitel 1

Mathematische Strukturen

Wir wollen zuerst ein paar Grundlagen mathematischer Strukturen einführen und uns etwas vertraut damit machen.

Definition 1.1 *Ein* <u>*kommutativer Ring mit Eins-Element* $(R, +, \cdot)$</u> *ist eine Menge* R *mit zwei „Operationen“*

$$\begin{array}{rcll} (a,b) & \to & a+b & (\text{„Addition“) und} \\ (a,b) & \to & a\cdot b & (\text{„Multiplikation“}), \end{array}$$

für die folgende Gesetze gelten:

$$\text{Add.}\left\{\begin{array}{cll} (\text{Ass }+) & (a+b)+c = a+(b+c) \quad \forall a,b,c, \in R & (\text{Assoziativgesetz}), \\ (\text{Komm }+) & a+b = b+a \quad \forall a,b \in R & (\text{Kommutativgesetz}), \\ (\text{Null}) & \exists \text{ ein } 0 \in R \text{ mit } 0+a = a+0 = a \quad \forall a \in R & \\ & & (\text{Existenz eines Null-Elements}), \\ (\text{Inv }+) & \forall a \in R\ \exists a' \in R \text{ mit } a + a' = 0 & \\ & & (\text{Existenz eines inversen Elements,} \\ & & \text{wir schreiben } -a \text{ anstatt } a'.), \end{array}\right.$$

$$\text{Mult.}\left\{\begin{array}{cll} (\text{Ass }\cdot) & (a\cdot b)\cdot c = a\cdot(b\cdot c) \quad \forall a,b,c \in R & (\text{Assoziativgesetz}), \\ (\text{Komm }\cdot) & a\cdot b = b\cdot a \quad \forall a,b \in R & (\text{Kommutativgesetz}), \\ (\text{Eins}) & \exists 1 \in R \text{ mit } 1\cdot a = a\cdot 1 = a \quad \forall a \in R & \\ & & (\text{Existenz eines Eins-Elements}), \end{array}\right.$$

$$(\text{Distr}) \qquad (a+b)\cdot c = a\cdot c + b\cdot c \quad \forall a,b,c \in R \qquad (\text{Distributivgesetz}).$$

Definition 1.2 *(i) Es sei* R *ein kommutativer Ring mit Eins-Element und* $r \in R$*. Dann heißt* r <u>*invertierbar*</u>*, falls es ein* $\tilde{r} \in R$ *mit* $r \cdot \tilde{r} = 1$ *gibt. Wir schreiben dann* r^{-1} *oder* $\frac{1}{r}$ *für* $\tilde{r}$*.*

(ii) Ein kommutativer Ring mit Eins-Element heißt <u>*Körper*</u>*, wenn* $0 \neq 1$ *und zusätzlich das weitere Gesetz gilt:*

$$(\text{Inv }\cdot) \qquad \text{Jedes Element } r \in R \text{ mit } r \neq 0 \text{ ist invertierbar.}$$

Beispiel 1.3 Bekannte Mengen

$$\begin{array}{lcll}
\mathbb{N} & = & \{1, 2, 3, ...\} & \text{die natürlichen Zahlen,} \\
\mathbb{N}_0 & = & \mathbb{N} \cup \{0\}, & \\
\mathbb{Z} & = & \{..., -2, -1, 0, 1, 2, ...\} & \text{die ganzen Zahlen,} \\
\mathbb{Q} & = & \left\{\frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{N}\right\} & \text{die rationalen Zahlen,} \\
\mathbb{R} & & & \text{die reellen Zahlen.}
\end{array}$$

Mit der bekannten Addition und Multiplikation sind $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ kommutative Ringe mit Eins-Element und $\mathbb{Q}, \mathbb{R}$ sind sogar Körper.
$\mathbb{N}, \mathbb{N}_0$ passen nicht in diese Definitionen. Warum nicht? Welche Gesetze gelten nicht?

Beispiel 1.4 Der kleinste Körper $\mathbb{F}_2$.

$\mathbb{F}_2 = (\{0, 1\}, +, \cdot)$, wobei $+$ und $\cdot$ wie folgt definiert sind:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Die Multiplikation ist die übliche. Die Addition geht „modulo“ 2, das heißt, man nimmt die übliche Addition und verwendet immer den ganzzahligen Rest nach Division durch 2 als Ergebnis:

$$\begin{aligned}
1+1+1+1+1+1 &= 0 \quad (6 : 2 = 3 \ \underline{\text{Rest } 0}), \\
1+1+1 &= 1 \quad (3 : 2 = 1 \ \underline{\text{Rest } 1}).
\end{aligned}$$

Kann es Körper mit weniger als 2 Elementen geben?

Beispiel 1.5 Sei

$$\begin{aligned}
V &= \{v = a + \sqrt{2}\, b \mid a, b \in \mathbb{Q}\}, \\
v_1 + v_2 &= (a_1 + \sqrt{2}\, b_1) + (a_2 + \sqrt{2}\, b_2) \\
&= (a_1 + a_2) + \sqrt{2}\,(b_1 + b_2), \\
v_1 \cdot v_2 &= (a_1 + \sqrt{2}\, b_1) \cdot (a_2 + \sqrt{2}\, b_2) \\
&= a_1 a_2 + \sqrt{2}\, a_1 b_2 + \sqrt{2}\, a_2 b_1 + 2 b_1 b_2 \\
&= (a_1 a_2 + 2 b_1 b_2) + \sqrt{2}\,(a_1 b_2 + a_2 b_1).
\end{aligned}$$

Ist $\{V, +, \cdot\}$ ein Körper (oder nur ein „Ring“)?

$$\frac{1}{v} = \frac{1}{a + \sqrt{2}\, b} = \frac{a - \sqrt{2}\, b}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} - \sqrt{2}\,\frac{b}{a^2 - 2b^2}.$$

Da $\sqrt{2} \notin \mathbb{Q}$, ist $a^2 - 2b^2 \neq 0$ für alle $v \in V$, $v \neq 0$. Damit ist $\frac{1}{v} \in V \quad \forall v \in V, v \neq 0$.
$\Rightarrow \{V, +, \cdot\}$ ist ein Körper!

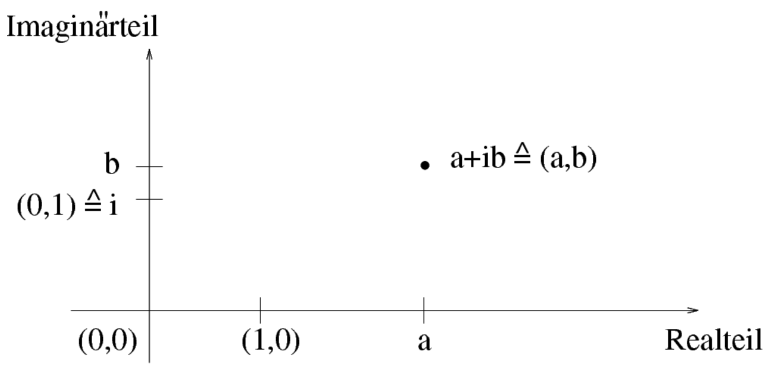
Beispiel 1.6 Komplexe Zahlen

Sei $\mathbb{C} = \{z = a+ib | a, b \in \mathbb{R}\}$, wobei $i = \sqrt{-1}$ die imaginäre Einheit ist, mit den Operationen

$$\begin{aligned} z_1 + z_2 &= (a_1 + ib_1) + (a_2 + ib_2) &= (a_1 + a_2) + i(b_1 + b_2), \\ z_1 \cdot z_2 &= (a_1 + ib_1) \cdot (a_2 + ib_2) &= (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2). \end{aligned}$$

Für $z = a + ib$ heißt a Realteil und b Imaginärteil von z.

Null-Element	$0 = 0 + i0$	$(0,0)$
Eins-Element	$1 = 1 + i0$	$(1,0)$
imaginäre Einheit	$i = 0 + i1$	$(0,1)$



Die konjugiert komplexe Zahl zu $z = a + ib$ ist die Zahl $\bar{z} = a - ib$.

$\mathbb{C}$ ist ein Körper, denn das inverse Element zu $z \neq 0$ ist

$$\frac{1}{z} = \frac{1}{a+ib} = \frac{a-ib}{(a+ib)(a-ib)} = \frac{\bar{z}}{z\bar{z}} = \frac{a-ib}{a^2+b^2} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2} \in \mathbb{C},$$

da $a^2 + b^2 > 0$ für $(a,b) \neq (0,0)$.

Übung: Andere Darstellung komplexer Zahlen: $z = re^{i\varphi} = r(\cos\varphi + i\sin\varphi)$.

Mengen und Abbildungen

Symbole:

$\in$	Element	$1 \in \mathbb{N}$
$\subset$	Teilmenge	$\mathbb{N} \subset \mathbb{Z}$
$\cap$	Durchschnitt	$\mathbb{N} \cap \mathbb{N}_0 = \mathbb{N}$
$\cup$	Vereinigung	$\mathbb{N} \cup \mathbb{N}_0 = \mathbb{N}_0$
$\setminus$	Mengendifferenz	$\mathbb{N}_0 \setminus \mathbb{N} = \{0\}$
$\times$	kartesisches Produkt	$\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3 = \{(a,b,c) \| a,b,c \in \mathbb{R}\}$

Definition 1.7 *Seien X, Y zwei Mengen. Eine <u>Abbildung f von X nach Y</u>,*

$$f : X \to Y,$$

ist eine Vorschrift, die jedem $x \in X$ genau ein Element $y = f(x) \in Y$ zuordnet. Für die <u>Zuordnung einzelner Elemente</u> schreiben wir $x \mapsto y$.

Beispiel 1.8 Sei $X = Y = \mathbb{R}$.

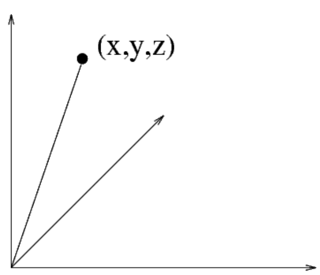
$$a)\ f : X \to Y, \quad x \mapsto x^3$$

$$b)\ f : X \to Y, \quad x \mapsto \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$$

Beispiel 1.9 <u>Euklidische Norm</u> oder <u>Euklidische Länge</u>

$$X = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3, \quad Y = \mathbb{R}$$

$$\|\cdot\|_2 : \ X \to Y, \quad (x,y,z) \mapsto \sqrt{x^2 + y^2 + z^2}$$



$\|(x,y,z)\|_2 \triangleq$ "Abstand des Punktes (x,y,z) vom Ursprung"

Definition 1.10

(a) *Sei A eine Menge. Dann ist*

$$\begin{aligned} Id_A \ : \ & A \to A \\ & a \mapsto a \end{aligned}$$

die Identitätsabbildung.

(b) *Seien X, Y Mengen und $A \subset X$, $B \subset Y$. Sei $f : X \to Y$ eine Abbildung. Dann heißt*

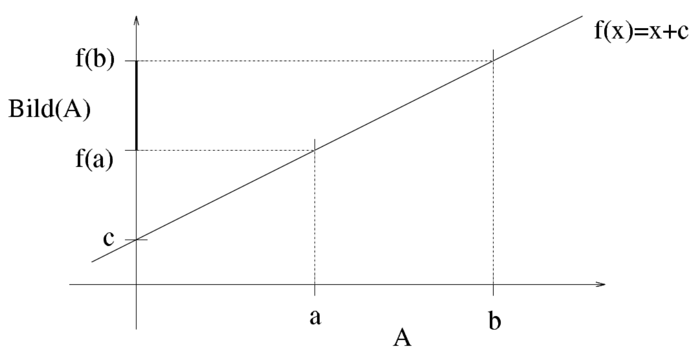
$$f(A) = \mathrm{Bild}(A) := \{f(x) \mid x \in A\}$$

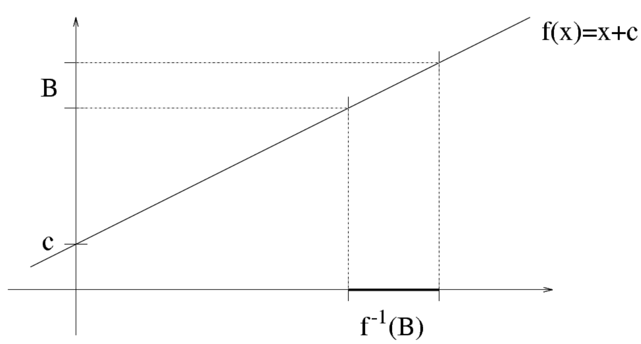
die Bildmenge von A und

$$f^{-1}(B) := \{x \mid f(x) \in B\}$$

das Urbild von B.

Beispiel 1.11





Definition 1.12 *Eine Abbildung $f : X \to Y$ heißt injektiv, wenn keine zwei Elemente von X auf dasselbe Element in Y abgebildet werden.*
Sie heißt surjektiv oder Abbildung auf Y, wenn jedes $y \in Y$ von der Form $f(x)$ ist.
Sie heißt bijektiv, wenn sie injektiv und surjektiv ist.

Beispiel 1.13

a) Sei $X = Y = \mathbb{R}$.
Ist $f(x) = x^2$ injektiv, surjektiv?
Ist $f(x) = 2x + 3$ injektiv, surjektiv?

b) Sei $X = Y = \mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$.
Ist $f(x) = x^2$ injektiv, surjektiv?

Merke! Zu einer Abbildung gehören immer die Mengen, auf denen sie operiert.

Definition 1.14 *Sind $f : X \to Y$ und $g : Y \to Z$ Abbildungen, so ist die zusammengesetzte Abbildung $g \circ f$ definiert durch*

$$\begin{aligned} g \circ f \quad : \quad & X \to Z \\ & x \mapsto g(f(x)). \end{aligned}$$

Ist f bijektiv, so heißt die Abbildung $f^{-1} : Y \to X$, für die $f^{-1} \circ f = Id_X$, die Umkehrabbildung von f.

Beispiel 1.15

$$X = \left[0, \frac{\pi}{2}\right], \;\; Y = [0,1], \;\; Z = [-1,0], \qquad \begin{aligned} f : \; & X \to Y \\ & x \mapsto \sin x, \end{aligned} \qquad \begin{aligned} g : \; & Y \to X \\ & y \mapsto -y. \end{aligned}$$

$$\begin{aligned} g \circ f \quad : \quad & X \to Z \\ & x \mapsto -\sin x, \\ f^{-1} \quad : \quad & Y \to Z \\ & x \mapsto \arcsin x, \\ g^{-1} \quad : \quad & Z \to Y \\ & z \mapsto -z. \end{aligned}$$

Definition 1.16 *Seien X, Y Mengen, $A \subset X$, $f : X \to Y$. Dann heißt*

$$\begin{aligned} f\big|_A \quad : \quad & A \to Y \\ & a \mapsto f(a) \end{aligned}$$

die Einschränkung von f auf A.

Beispiel 1.17

$$Y = X = \mathbb{R}, \;\; A = \left[0, \frac{\pi}{2}\right], \qquad \begin{aligned} f \quad : \quad & X \to Y \\ & x \mapsto \sin x. \end{aligned}$$

f ist nicht injektiv, aber $f\big|_A$ ist injektiv.

Kapitel 2

Matrizen

Definition 2.1 *Sei $\{R,+,\cdot\}$ ein kommutativer Ring mit Eins-Element und $n,m \in \mathbb{N}_0$. Ein Feld*

$$A=[a_{ij}]=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m}\\ a_{21} & a_{22} & \cdots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nm}\end{bmatrix}$$

mit $a_{ij}\in R$, $i=1,\ldots,n$, $j=1,\ldots,m$, heißt $\underline{n\times m\text{-Matrix mit Koeffizienten in } R}$ oder $(n\times m\text{-})$ $\underline{\text{Matrix über } R}$.
Dabei gelten folgende Bezeichnungen:

- $R^{n,m}$: *Menge aller $n\times m$-Matrizen über R,*
- a_{ij} : *der i,j-te $\underline{\text{Koeffizient}}$ oder $\underline{\text{Eintrag}}$,*
- $[a_{i1},\ldots,a_{im}]$: *die i-te $\underline{\text{Zeile}}$ von A (das ist eine $1\times m$-Matrix),*
- $\begin{bmatrix} a_{1j}\\ \vdots\\ a_{nj}\end{bmatrix}$: *die j-te $\underline{\text{Spalte}}$ von A (das ist eine $n\times 1$-Matrix),*
- 0 : *die $\underline{\text{Nullmatrix}}$, d.h., die Matrix in $R^{n,m}$, bei der alle Einträge 0 sind,*
- I_n : *die $\underline{\text{Einheitsmatrix}}$ in $R^{n,n}$, d.h., die Matrix mit den Einträgen*

$$\delta_{ij}=\begin{cases} 1, & i=j\\ 0, & \text{sonst}\end{cases},$$

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

- E_{ij} : *die Matrix in* $R^{n,m}$*, die in der Position* (i,j) *den Eintrag* 1 *und in allen anderen Postionen den Eintrag* 0 *hat, z.B.*

$$E_{11} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Operationen mit Matrizen

Addition von Matrizen

Wir können Matrizen gleicher Größe addieren:

$$\begin{array}{cclcl} + & : & (R^{n,m} \times R^{n,m}) & \to & R^{n,m} \\ & & (A,B) & \mapsto & A+B=C=[c_{ij}], \\ & & & & c_{ij} = a_{ij} + b_{ij},\ \ i=1,\ldots,n,\ \ j=1,\ldots,m. \end{array}$$

Eigenschaften der Matrizenaddition:

Seien $A, B, C \in R^{n,m}$, $A = [a_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}]$ und setze $\tilde{A} = [-a_{ij}]$. Dann gilt:

$$\begin{array}{crcl} (\text{Ass } +) & (A+B)+C & = & A+(B+C), \\ (\text{Komm } +) & A+B & = & B+A, \\ (\text{Null}) & A+0 & = & 0+A \ = \ A, \\ (\text{Inv } +) & A+\tilde{A} & = & \tilde{A}+A \ = \ 0. \end{array}$$

Skalarmultiplikation

Wir können Matrizen mit Elementen aus R multiplizieren.

$$\begin{array}{cclcl} \cdot & : & (R^{n,m} \times R) & \to & R^{n,m} \\ & & (A,r) & \mapsto & r \cdot A = [r \cdot a_{ij}]. \end{array}$$

Eigenschaften der Skalarmultiplikation:

Seien $A, B \in R^{n,m}$, $r, s \in R$. Dann gilt:

a) $(r \cdot s)A = r(sA)$,
b) $(r+s)A = rA + sA$,
c) $r(A+B) = rA + rB$,
d) $1 \cdot A = A$,
e) $A + (-1)A = 0$,
f) $A = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} a_{ij} E_{ij}$.

Multiplikation von Matrizen

Sei $A = [a_{ij}] \in R^{n,m}$, $B = [b_{ij}] \in R^{m,s}$.
Setze $A \cdot B = C = [c_{ij}] \in R^{n,s}$, $c_{ij} = \sum\limits_{k=1}^{m} a_{ik} b_{kj}$.

<u>Technik:</u>

$$\begin{array}{cc} & \begin{bmatrix} b_{11} & \cdots & \left[\begin{matrix} b_{1j} \\ \vdots \\ b_{mj} \end{matrix}\right] & \cdots & b_{1s} \\ \vdots & & & & \vdots \\ b_{m1} & \cdots & & \cdots & b_{ms} \end{bmatrix} \\ \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ [a_{i1} & \cdots & a_{im}] \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} & \begin{bmatrix} & & & \\ & & \downarrow & \\ & \longrightarrow & c_{ij} & \\ & & & \\ & & & \end{bmatrix} \end{array}$$

Im folgenden lassen wir den Multiplikationspunkt meistens weg.

<u>Eigenschaften der Matrizenmultiplikation:</u>

Lemma 2.2 *Seien $A = [a_{ij}] \in R^{n,m}$, $\tilde{A} = [\tilde{a}_{ij}] \in R^{n,m}$, $B = [b_{ij}] \in R^{m,s}$, $\tilde{B} = [\tilde{b}_{ij}] \in R^{m,s}$, $C = [c_{ij}] \in R^{s,t}$, $r \in R$. Dann gilt:*

a)	*(Ass $\cdot$)*	$(A \cdot B) \cdot C = A \cdot (B \cdot C)$,
b)	*(Distr 1)*	$(A + \tilde{A})B = AB + \tilde{A}B$,
c)	*(Distr 2)*	$A(B + \tilde{B}) = AB + A\tilde{B}$,
d)	(I_n, I_m)	$I_n A = A I_m = A$,
e)		$(r \cdot A)B = r(AB) = A(rB)$.

Beweis: a) Sei $D = [d_{ij}] = (A \cdot B) \cdot C$, $\tilde{D} = [\tilde{d}_{ij}] = A \cdot (B \cdot C)$. Es gilt

$$\begin{aligned} d_{ij} &= \sum_{l=1}^{s} \left(\sum_{k=1}^{m} a_{ik}\, b_{kl} \right) c_{lj} \\ &= \sum_{l=1}^{s} \sum_{k=1}^{m} (a_{ik}\, b_{kl})\, c_{lj} && \text{! Distributivität in } R \\ &= \sum_{l=1}^{s} \sum_{k=1}^{m} a_{ik}\, (b_{kl}\, c_{lj}) \\ &= \sum_{k=1}^{m} a_{ik} \left(\sum_{l=1}^{s} b_{kl}\, c_{lj} \right) \\ &= \tilde{d}_{ij}. \end{aligned}$$

b)-e) Übung! $\square$

Definition 2.3 *Eine Matrix $A \in R^{n,n}$ heißt* <u>*invertierbar*</u>*, wenn es ein $\tilde{A} \in R^{n,n}$ gibt mit $\tilde{A}A = A\tilde{A} = I_n$. Man schreibt dann $\tilde{A} = A^{-1}$, die* <u>*inverse Matrix*</u> *von A.*

Lemma 2.4 *Seien $A, B \in R^{n,n}$ invertierbar. Dann ist AB invertierbar und es gilt*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Beweis:

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n,\\ (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n.\end{aligned}$$

□

Lemma 2.5 *Falls $A \in R^{n,n}$ invertierbar ist, gibt es genau eine Inverse von A.*

Beweis: Angenommen, es gibt zwei verschiedene Matrizen B, $\tilde{B}$, so dass

$$\begin{aligned}AB = I_n, &\quad A\tilde{B} = I_n,\\ BA = I_n, &\quad \tilde{B}A = I_n.\end{aligned}$$

$$\begin{aligned}&\Longrightarrow \quad AB - A\tilde{B} = A(B - \tilde{B}) = 0\\ &\Longrightarrow \quad \underbrace{BA}_{I_n}(B - \tilde{B}) = B\,0 = 0\\ &\Longrightarrow \quad B - \tilde{B} = 0\\ &\Longrightarrow \quad B = \tilde{B}.\end{aligned}$$

□

Beispiel 2.6

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

A ist invertierbar genau dann, wenn $ad - bc \neq 0$.

Bemerkung 2.7

a) Nicht alle Matrizen in $R^{n,n}$ sind invertierbar, siehe Beispiel 2.6, z.B.

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.$$

b) Die Matrizenmultiplikation ist i.a. nicht kommutativ. Z.B.

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix},$$
$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

c) Aus $A \cdot B = 0$ folgt nicht $A = 0$ oder $B = 0$, z.B.

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

Definition 2.8 *Sei $A = [a_{ij}] \in R^{n,m}$. Dann heißt die Matrix $B = [b_{ij}] \in R^{m,n}$ mit $b_{ij} = a_{ji}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, die transponierte Matrix zu A. Wir schreiben $B = A^T$.*

Eigenschaften der Transponierten:

Lemma 2.9 *Seien $A, \tilde{A} \in R^{n,m}$, $B \in R^{m,s}, r \in R$. Dann gilt*

a) $(A + \tilde{A})^T = A^T + \tilde{A}^T$,
b) $(rA)^T = rA^T$,
c) $(AB)^T = B^T A^T$,
d) $(A^T)^T = A$.
e) *Falls $n = m$ und A invertierbar, so gilt $(A^{-1})^T = (A^T)^{-1}$.*

Beweis: a), b), d) sind offensichtlich.

c) Sei $A \cdot B = C = [c_{ij}]$ mit $c_{ij} = \sum\limits_{k=1}^{m} a_{ik}\, b_{kj}$ und $A^T = [a'_{ij}]$, $B^T = [b'_{ij}]$, $C^T = [c'_{ij}]$. Es gilt

$$\begin{aligned} c'_{ij} = c_{ji} &= \sum_{k=1}^{m} a_{jk}\, b_{ki} \\ &= \sum_{k=1}^{m} a'_{kj}\, b'_{ik} \\ &= \sum_{k=1}^{m} b'_{ik}\, a'_{kj}. \end{aligned}$$

$\Longrightarrow \quad C^T = B^T A^T.$

e) $A^{-1}A = I_n \implies (A^{-1}A)^T = I_n^T = I_n \implies A^T(A^{-1})^T = I_n \implies (A^{-1})^T = (A^T)^{-1}$, da die Inverse nach Lemma 2.5 eindeutig ist.

□

Spezielle Klassen von Matrizen

Definition 2.10 *Sei $A \in R^{n,n}$.*

a) *A heißt symmetrisch, falls $A = A^T$.*

b) *A heißt obere Dreiecksmatrix, falls $a_{ij} = 0 \ \forall\, i = 2, \ldots, n,\ j = 1, \ldots, i-1$.*

c) *A heißt untere Dreiecksmatrix, falls A^T obere Dreiecksmatrix ist.*

d) *A heißt Diagonalmatrix, falls A obere und untere Dreiecksmatrix ist.*

e) *A heißt Permutationsmatrix, falls in jeder Zeile und in jeder Spalte genau ein Eintrag 1 ist und alle anderen Einträge 0 sind.*

Definition 2.11

a) *Eine additive Gruppe $\{G, +\}$ ist eine Menge G mit einer Operation + , die bezüglich der Operation + abgeschlossen ist, und für die die Gesetze (Ass +), (Null) und (Inv +) aus der Definition 1.1 gelten. Falls auch noch (Komm +) gilt, so heißt $\{G, +\}$ kommutative additive Gruppe.*

b) *Eine multiplikative Gruppe $\{G, \cdot\}$ ist eine Menge G mit einer Operation · , die bezüglich der Operation · abgeschlossen ist, für die die Gesetze (Ass ·) und (Eins) aus Definition 1.1 gelten und in der jedes Element $g \in G$ invertierbar ist. Falls auch noch (Komm ·) gilt, so heißt $\{G, \cdot\}$ kommutative multiplikative Gruppe.*

Korollar 2.12 *$R^{n,m}$ ist bezüglich der Matrizenaddition eine kommutative additive Gruppe.*

Beweis: Siehe „Eigenschaften der Matrizenaddition“. □

Korollar 2.13 *Die Menge $GL_n(R)$ der invertierbaren Matrizen in $R^{n,n}$ ist bezüglich der Matrixmultiplikation eine (nicht kommutative) multiplikative Gruppe.*

Beweis: Das Einselement ist I_n. Der Rest ist bereits bewiesen, siehe „Eigenschaften der Matrizenmultiplikation“. □

Korollar 2.14

a) *Die Menge der invertierbaren oberen Dreiecksmatrizen in $R^{n,n}$ ist bezüglich der Matrixmultiplikation eine (nicht kommutative) multiplikative Gruppe. (Analog für untere Dreiecksmatrizen).*

b) *Die Menge der nichtsingulären Diagonalmatrizen bildet eine kommutative multiplikative Gruppe.*

Beweis:

a) Es seien $A = [a_{ij}]$, $B = [b_{ij}]$ invertierbare obere Dreiecksmatrizen in $R^{n,n}$. Wir müssen zunächst beweisen, dass $A \cdot B$ wieder eine obere Dreiecksmatrix ist. Sei $C = A \cdot B = [c_{ij}]$. Für $i > j$ gilt

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{n} a_{ik}\, b_{kj} \\ &= \sum_{k=1}^{j} a_{ik}\, b_{kj} \qquad (\text{da } b_{kj} = 0 \text{ für } k > j) \\ &= 0. \qquad\qquad\quad (\text{da } a_{ik} = 0 \text{ für } i > k) \end{aligned}$$

Die Gültigkeit von (Ass $\cdot$) und (Eins) ist klar. Nun müssen wir noch zeigen, dass A^{-1} eine obere Dreiecksmatrix ist (Existenz ist klar nach Voraussetzung). Wie bekommen wir A^{-1}? Wir suchen $C = [c_{ij}]$, so dass $AC = I$, d.h., für alle $j = 1, \ldots, n$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ & & a_{nn} \end{bmatrix} \begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix} = \begin{bmatrix} \delta_{1j} \\ \vdots \\ \delta_{nj} \end{bmatrix}. \qquad \left(\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & \text{sonst} \end{cases} \right)$$

$$\begin{aligned} a_{nn}\, c_{nj} &= \delta_{nj} & \Longrightarrow \quad c_{nj} &= \frac{\delta_{nj}}{a_{nn}}, \\ a_{n-1,n-1}\, c_{n-1,j} + a_{n-1,n}\, c_{nj} &= \delta_{n-1,j} & \Longrightarrow \quad c_{n-1,j} &= \frac{1}{a_{n-1,n-1}} (\delta_{n-1,j} - a_{n-1,n}\, c_{nj}), \\ & \vdots \end{aligned}$$

Für $j = 1, \ldots, n:$

$$c_{nj} = \frac{\delta_{nj}}{a_{nn}}, \tag{2.15}$$

$$c_{ij} = \frac{1}{a_{ii}} \left(\delta_{ij} - \sum_{k=i+1}^{n} a_{ik}\, c_{kj} \right), \qquad i = n-1, \ldots, 1.$$

Formel für die Inverse einer oberen Dreiecksmatrix (Rückwärts-Einsetzen)

Die Existenz von A^{-1} liefert $a_{ii} \neq 0$, $i = 1, \ldots, n$.

Zeigen nun per Induktion, rückwärts, dass C obere Dreiecksmatrix ist.
Sei $j < n$. Dann:

I.A.: $\quad c_{nj} = \dfrac{\delta_{nj}}{a_{nn}} = 0.$

I.V.: $\quad$ Für l mit $j+2 \leq l \leq n$ sei $c_{kj} = 0$ für $k = l, \ldots, n$.

I.S.: $\quad c_{l-1,j} = \dfrac{1}{a_{l-1,l-1}} \left(\delta_{l-1,j} - \sum_{k=l}^{n} a_{l-1,k}\, c_{kj} \right) = \dfrac{\delta_{l-1,j}}{a_{l-1,l-1}} = 0.$

b) Abgeschlossenheit des Produktes, (Ass $\cdot$) und (Eins) sind klar. Seien $A = [a_{ij}]$, $B = [b_{ij}]$ Diagonalmatrizen. $C = A^{-1}$ existiert nach Voraussetzung. Aus Formel (2.15) folgt $c_{ij} = \delta_{ij}/a_{ii}$, $i, j = 1, \ldots, n$. Also ist C Diagonalmatrix. Weiterhin gilt

$$A \cdot B = \operatorname{diag}(a_{11}\, b_{11}, \ldots, a_{nn}\, b_{nn}) = \operatorname{diag}(b_{11}\, a_{11}, \ldots, b_{nn}\, a_{nn}) = B \cdot A.$$

$\square$

Bemerkung 2.16 In (2.15) haben wir eine Formel für die Inverse einer invertierbaren oberen Dreiecksmatrix erhalten, die auch gleich einen <u>rekursiven</u> Algorithmus liefert. Analoges gilt natürlich für untere Dreiecksmatrizen.
Beispiel:

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}. \qquad A^{-1} = \begin{bmatrix} 1 & -2 & 4 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Übung: Allgemeine Formel für die Inverse von Block-Dreiecksmatrizen.

Es sei

$$A = \begin{array}{cl} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] & \begin{array}{l} \scriptstyle k \\ \scriptstyle n-k \end{array} \\ \begin{array}{cc} \scriptstyle k & \scriptstyle n-k \end{array} & \end{array} .$$

A_{11}, A_{22} seien invertierbar. Zeige:

$$A^{-1} = \begin{array}{cl} \left[\begin{array}{c|c} A_{11}^{-1} & -A_{11}^{-1}\, A_{12}\, A_{22}^{-1} \\ \hline 0 & A_{22}^{-1} \end{array}\right] & \begin{array}{l} \scriptstyle k \\ \scriptstyle n-k \end{array} \\ \begin{array}{cc} \scriptstyle k & \scriptstyle n-k \end{array} & \end{array} .$$

Satz 2.17 *Die Menge der Permutationsmatrizen in $R^{n,n}$ bildet eine multiplikative Gruppe. Ist $A \in R^{n,n}$ eine Permutationsmatrix, so gilt $A^{-1} = A^T$.*

Beweis: Seien $A = [a_{ij}], B = [b_{ij}] \in R^{n,n}$ Permutationsmatrizen, $C = A \cdot B = [c_{ij}]$ mit

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj} = [a_{i1}, \ldots, a_{in}] \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix}.$$

Da es nur genau ein Element a_{ik} gibt, welches von 0 verschieden (nämlich $= 1$) ist, und genau ein Element b_{kj}, welches von 0 verschieden ($= 1$) ist, so gibt es in jeder Zeile und in jeder Spalte von C genau ein von Null verschiedenes Element ($= 1$), nämlich dort, wo $a_{ik} = b_{kj} = 1$ ist.
Sei $A \cdot A^T = C = [c_{ij}]$. Dann gilt:

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, a_{jk} = \delta_{ij}.$$

□

Kapitel 3

Die Treppennormalform und der Gauß'sche Algorithmus

Wir haben bereits gesehen, wie wir die Inverse von Dreiecksmatrizen berechnen können. Wir würden auch gerne für volle Matrizen solche Formeln haben, aber das geht nicht so einfach. Um dies zu erreichen, versuchen wir, die Matrix erst auf eine Dreiecksform zu bringen, und zwar durch Multiplikation mit Matrizen, deren Inverse wir leicht berechnen können. Diese sogenannten Elementarmatrizen führen elementare Operationen aus:

- Vertauschung zweier Zeilen (Spalten),
- Multiplikation einer Zeile (Spalte) mit einem Skalar,
- Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

Es sei $P_{ij} \in R^{n,n}$ für $1 \leq i < j \leq n$ die Permutationsmatrix

$$P_{ij} := \begin{array}{cc}
\left[\begin{array}{ccc|c|ccc|c|ccc}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\ \hline
& & & 0 & & & & 1 & & & \\ \hline
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\ \hline
& & & 1 & & & & 0 & & & \\ \hline
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{array}\right] & \begin{array}{c} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array} \\
\begin{array}{ccccccccccc} & & & \uparrow & & & & \uparrow & & & \\ & & & i & & & & j & & & \end{array} &
\end{array} .$$

Ist $A \in R^{n,m}$, so werden durch die Multiplikation $P_{ij}A$ die Zeilen i und j in A vertauscht.

Beachte: $P_{ij} = P_{ij}^T = P_{ij}^{-1}$.

Es sei $M_i(\lambda) \in R^{n,n}$ für $1 \leq i \leq n$, $\lambda \in R\backslash\{0\}$ die Matrix

$$M_i(\lambda) := \begin{array}{cc} \left[\begin{array}{ccc|c|ccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ \hline & & & \lambda & & & \\ \hline & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right] & \begin{array}{c} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{array} \\ \begin{array}{c} \uparrow \\ i \end{array} & \end{array} .$$

Ist $A \in R^{n,m}$, so wird durch die Multiplikation $M_i(\lambda)A$ die i-te Zeile von A mit λ multipliziert.

Beachte: $M_i(\lambda)^{-1} = M_i\left(\dfrac{1}{\lambda}\right)$.

Es sei $G_{ij}(\lambda) \in R^{n,n}$ für $1 \leq i < j \leq n$, $\lambda \in R$ die Matrix

$$G_{ij}(\lambda) := \begin{array}{cc} \left[\begin{array}{ccc|c|ccc|c|ccc} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ \hline & & & 1 & & & & & & & \\ \hline & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ \hline & & & \lambda & & & & 1 & & & \\ \hline & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{array}\right] & \begin{array}{c} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array} \\ \begin{array}{ccc} \uparrow & \qquad & \uparrow \\ i & & j \end{array} & \end{array} .$$

Es sei $A \in R^{n,m}$. Durch die Multiplikation $G_{ij}(\lambda)A$ wird das λ-fache der i-ten Zeile von A zur j-ten Zeile addiert. Durch die Multiplikation $G_{ij}^T(\lambda)A$ wird das λ-fache der j-ten Zeile zur i-ten Zeile von A addiert.

Beachte: $[G_{ij}(\lambda)]^{-1} = G_{ij}(-\lambda)$.

Satz 3.1 *Sei K ein Körper, $A \in K^{n,m}$. Dann gibt es Elementarmatrizen $S_1, \ldots, S_t \in K^{n,n}$, so dass $S_t \cdots S_1 A$ in Treppennormalform ist, d.h.,*

$$S_t \cdots S_1 A = \left[\begin{array}{c|cc|cc|cc|c|cc}
 & 1 & * & 0 & * & 0 & * & & 0 & * \\ \cline{2-3}
 & & & 1 & & 0 & & & \vdots & \\ \cline{4-5}
0 & & 0 & & & 1 & & \ddots & 0 & \\ \cline{6-7}
 & & & & 0 & & 0 & & 1 & \\ \cline{9-10}
 & & & & & & & & & 0
\end{array}\right].$$

Insbesondere, falls $n = m$ und A invertierbar ist, so ist $S_t \cdots S_1 A = I$, d.h., $A^{-1} = S_t \cdots S_1$ oder $A = S_1^{-1} \cdots S_t^{-1}$.

Beweis: Den Beweis führen wir konstruktiv mit Hilfe des Gauß'schen Algorithmus.

Ist A die Null-Matrix, dann setzen wir $t = 1$, $S_1 = I_n$ und sind fertig.

Nun sei A von der Null-Matrix verschieden. Sei j_1 die erste Spalte von $A^{(1)} := A = \left[a_{ij}^{(1)}\right]$, die nicht aus lauter Nullen besteht, und sei $a_{i1,j1}^{(1)}$ das erste Element in der j_1-ten Spalte, welches nicht 0 ist. Bilde

$$\tilde{A}^{(1)} := M_1\left(\frac{1}{a_{i_1,j_1}^{(1)}}\right) P_{1,i_1}\, A^{(1)} = \begin{array}{c} \left[\begin{array}{c|c|c} & 1 & \\ & \tilde{a}_{2,j1}^{(1)} & \\ 0 & \vdots & * \\ & \tilde{a}_{n,j_1}^{(1)} & \end{array}\right] \\ \quad j_1 \end{array} = \left[\tilde{a}_{i,j}^{(1)}\right].$$

Multipliziere von links mit

$$G_{1,n}\left(-\tilde{a}_{n,j_1}^{(1)}\right) \cdots G_{1,2}\left(-\tilde{a}_{2,j_1}^{(1)}\right).$$

So gilt mit $S_1 := G_{1,n}\left(-\tilde{a}_{n,j_1}^{(1)}\right) \cdots G_{1,2}\left(-\tilde{a}_{2,j_1}^{(1)}\right) M_1\left(\frac{1}{a_{i_1,j_1}^{(1)}}\right) P_{1,i_1}$

$$S_1 A^{(1)} = \begin{array}{c} \left[\begin{array}{c|c|c} 0 & 1 & * \\ \hline & 0 & \\ 0 & \vdots & A^{(2)} \\ & 0 & \end{array}\right] \\ j_1 \end{array}.$$

Setze $A^{(2)} = \left[a_{ij}^{(2)}\right]$, $i = 2, \ldots, n$, $j = j_1 + 1, \ldots, m$ (das heißt, wir behalten die Indizes aus der „großen“ Matrix in der kleineren bei).
Für $k = 2, \ldots, s$ seien die Matrizen S_k rekursiv definiert durch

$$S_k = \left[\begin{array}{c|c} I_{k-1} & 0 \\ \hline 0 & \tilde{S}_k \end{array}\right],$$

$$\tilde{S}_k A^{(k)} = \left[\begin{array}{c|c|c} 0 & 1 & * \\ \hline & 0 & \\ 0 & \vdots & A^{(k+1)} \\ & 0 & \end{array}\right],$$

wobei $\tilde{S}_k$ analog zu S_1 konstruiert wird: $\tilde{S}_k$ entsteht, indem ich die erste Spalte j_k von $A^{(k)}$ finde, die nicht aus lauter Nullen besteht, und das erste Element $a^{(k)}_{i_k,j_k}$ der j_k-ten Spalte, welches ungleich 0 ist. Dann ist

$$\tilde{S}_k = G_{k,n}\Big(-\tilde{a}^{(k)}_{n,j_k}\Big) \cdots G_{k,k+1}\Big(-\tilde{a}^{(k)}_{k+1,j_k}\Big) M_k\left(\frac{1}{a^{(k)}_{i_k,j_k}}\right) P_{k,i_k},$$

wobei

$$\tilde{A}^{(k)} := M_k\left(\frac{1}{a^{(k)}_{i_k,j_k}}\right) P_{k,i_k} A^{(k)} = \Big[\tilde{a}^{(k)}_{ij}\Big].$$

Beachte: Wenn $A^{(k)}$ die Null-Matrix ist, ist nichts mehr zu tun, und wir setzen $s = k - 1$. Nach höchstens $\min\{n, m\}$ Schritten bricht dieser Prozess ab.

Dann folgt, dass

$$\begin{aligned} R &= R^{(1)} = \Big[r^{(1)}_{ij}\Big] \\ &:= S_s \cdots S_2 S_1 A^{(1)} \end{aligned}$$

$$= \begin{array}{c} \begin{array}{cccccccccc} & j_1 & & j_2 & & j_3 & & & j_s & \end{array} \\ \left[\begin{array}{c|cc|cc|cc|c|cc} & 1 & * & * & & * & & & * & \\ \cline{2-3} & & & 1 & * & * & * & & \vdots & \\ \cline{4-5} 0 & & & & & 1 & & \ddots & * & * \\ \cline{6-7} & & 0 & & 0 & & 0 & & 1 & \\ \cline{9-10} & & & & & & & & & 0 \end{array}\right] \end{array}. \tag{3.2}$$

Ist $s = 1$, dann haben wir die Treppennormalform bereits erreicht. Ist $s > 1$, bleiben nur noch die $*$ über den Einsen auszuräumen. Dazu bilden wir für $k = 2, \ldots, s$ rekursiv

$$\begin{aligned} S_{s+k-1} &= G^T_{1,k}\Big(-r^{(k-1)}_{1,j_k}\Big) \cdots G^T_{k-1,k}\Big(-r^{(k-1)}_{k-1,j_k}\Big), \\ R^{(k)} &:= S_{s+k-1} R^{(k-1)} =: \Big[r^{(k)}_{ij}\Big]. \end{aligned}$$

Setze $t = s + s - 1$. Aus der Konstruktion folgt dann, dass $R^{(s)} = S_t \cdots S_1 A$ in Treppennormalform ist. $\square$

Definition 3.3 *Die Positionen der Einsen in der Treppennormalform heißen* <u>*Pivotpositionen*</u>*.*

Beispiel 3.4

$$
\begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 2 & 0 & 1 \\ 0 & 2 & 0 & 2 \end{bmatrix}
\underset{M_1\left(\frac{1}{2}\right)}{\overset{j_1=2}{\longrightarrow}}
\left[\begin{array}{c|ccc} 0 & 1 & \frac{1}{2} & \frac{3}{2} \\ 0 & 2 & 0 & 1 \\ 0 & 2 & 0 & 2 \end{array}\right]
\underset{G_{1,3}(-2)}{\longrightarrow}
\left[\begin{array}{c|ccc} 0 & 1 & \frac{1}{2} & \frac{3}{2} \\ 0 & 2 & 0 & 1 \\ 0 & 0 & -1 & -1 \end{array}\right]
$$

$$
\underset{G_{1,2}(-2)}{\longrightarrow}
\left[\begin{array}{c|c|cc} 0 & 1 & \frac{1}{2} & \frac{3}{2} \\ \hline 0 & 0 & -1 & -2 \\ 0 & 0 & -1 & -1 \end{array}\right]
\underset{M_2\left(\frac{1}{-1}\right)}{\longrightarrow}
\left[\begin{array}{c|c|cc} 0 & 1 & \frac{1}{2} & \frac{3}{2} \\ \hline 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & -1 \end{array}\right]
$$

$$
\underset{G_{2,3}(1)}{\longrightarrow}
\left[\begin{array}{c|c|c|c} 0 & 1 & \frac{1}{2} & \frac{3}{2} \\ \cline{2-2} 0 & 0 & 1 & 2 \\ \cline{3-3} 0 & 0 & 0 & 1 \end{array}\right]
\underset{G^T_{1,2}\left(-\frac{1}{2}\right)}{\longrightarrow}
\left[\begin{array}{c|c|c|c} 0 & 1 & 0 & \frac{1}{2} \\ \cline{2-2} 0 & 0 & 1 & 2 \\ \cline{3-3} 0 & 0 & 0 & 1 \end{array}\right]
$$

$$
\underset{G^T_{2,3}(-2)}{\longrightarrow}
\left[\begin{array}{c|c|c|c} 0 & 1 & 0 & \frac{1}{2} \\ \cline{2-2} 0 & 0 & 1 & 0 \\ \cline{3-3} 0 & 0 & 0 & 1 \end{array}\right]
\underset{G^T_{13}\left(-\frac{1}{2}\right)}{\longrightarrow}
\left[\begin{array}{c|c|c|c} 0 & 1 & 0 & 0 \\ \cline{2-2} 0 & 0 & 1 & 0 \\ \cline{3-3} 0 & 0 & 0 & 1 \end{array}\right].
$$

Korollar 3.5 *Falls $A \in K^{n,n}$ invertierbar ist, so reicht es, den Gauß'schen Algorithmus bis zur Form (3.2) auszuführen. Dann ist R eine obere Dreiecksmatrix und es gilt $A = S_1^{-1} \cdots S_s^{-1} R$. Weiterhin gilt, dass $S_1^{-1} \cdots S_s^{-1}$ die Form $P \cdot L$ hat, wobei P eine Permutationsmatrix und L eine invertierbare untere Dreiecksmatrix ist.*

Beweis: Falls A invertierbar ist, so hat (3.2) die Form

$$
R = \begin{bmatrix} 1 & * & \cdots & * \\ & 1 & \ddots & \vdots \\ & & \ddots & * \\ & & & 1 \end{bmatrix} = [r_{ij}],
$$

welches eine obere Dreiecksmatrix mit Einsen auf der Hauptdiagonale ist.

$$
R = S_s \cdots S_1 A \qquad \Longrightarrow \qquad A = S_1^{-1} \cdots S_s^{-1} \cdot R.
$$

Wir schauen uns nun die S_i näher an. Jedes S_i hat die Form

$$
S_i = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & s_{i,i} & & & \\ & & & s_{i+1,i} & 1 & & \\ & & & & & \ddots & \\ & & & s_{n,i} & & & 1 \end{bmatrix} P_{i,j_i}
$$

mit $j_i \geq i$.

Also folgt, dass

$$S_s \cdots S_1 =$$

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & s_{n,n} \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & s_{n-1,n-1} & \\ & & & s_{n,n-1} & 1 \end{bmatrix} P_{n-1,j_{n-1}} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{n-2,n-2} & & \\ & & & s_{n-1,n-2} & 1 & \\ & & & s_{n,n-2} & 0 & 1 \end{bmatrix}$$

$$P_{n-2,j_{n-2}} \cdots \begin{bmatrix} 1 & & & & \\ & s_{22} & & & \\ & s_{32} & 1 & & \\ & \vdots & & \ddots & \\ & s_{n,2} & & & 1 \end{bmatrix} P_{2,j_2} \begin{bmatrix} s_{11} & & & & \\ s_{21} & 1 & & & \\ s_{31} & & 1 & & \\ \vdots & & & \ddots & \\ s_{n,1} & & & & 1 \end{bmatrix} P_{1,j_1}$$

mit $j_i \geq i$ für alle $j = 1, \ldots, n-1$. Es gilt aber, dass durch die Multiplikation mit $P_{n-1,j_{n-1}}$ in

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{n-2,n-2} & & \\ & & & s_{n-1,n-2} & 1 & \\ & & & s_{n,n-2} & 0 & 1 \end{bmatrix}$$

höchstens die letzten beiden Zeilen vertauscht werden, also kann ich schreiben

$$P_{n-1,j_{n-1}} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{n-2,n-2} & & \\ & & & s_{n-1,n-2} & 1 & \\ & & & s_{n,n-2} & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{n-2,n-2} & & \\ & & & \tilde{s}_{n-1,n-2} & 1 & \\ & & & \tilde{s}_{n,n-2} & & 1 \end{bmatrix} P_{n-1,j_{n-1}}.$$

(Durch die Multiplikation AP_{ij} werden die Spalten i und j in A vertauscht.)
Analog gilt

$$P_{k,j_k} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{l,l} & & \\ & & & s_{l+1,l} & 1 & \\ & & & \vdots & & \ddots & \\ & & & s_{n,l} & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{l,l} & & \\ & & & \tilde{s}_{l+1,l} & 1 & \\ & & & \vdots & & \ddots & \\ & & & \tilde{s}_{n,l} & & & 1 \end{bmatrix} P_{k,j_k}$$

für $k = 2, \ldots, n-1$, $l = 1, \ldots, k-1$.

Es folgt per Induktion, dass

$$S_s \cdots S_1 = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & s_{n-1,n-1} & \\ & & & s_{n,n-1} & s_{n,n} \end{bmatrix} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{n-2,n-2} & & \\ & & & \tilde{s}_{n-1,n-2} & 1 & \\ & & & \tilde{s}_{n,n-2} & & 1 \end{bmatrix} \cdots$$

$$\cdots \begin{bmatrix} 1 & & & & \\ & s_{22} & & & \\ & \tilde{s}_{32} & 1 & & \\ & \vdots & & \ddots & \\ & \tilde{s}_{n2} & & & 1 \end{bmatrix} \begin{bmatrix} s_{11} & & & \\ \tilde{s}_{21} & 1 & & \\ \vdots & & \ddots & \\ \tilde{s}_{n,1} & & & 1 \end{bmatrix} \cdot P_{n-1,j_{n-1}} \cdots P_{1,j_1}$$

die Form $\tilde{L} \cdot \tilde{P}$ hat, wobei $\tilde{L}$ eine untere Dreiecksmatrix (vgl. Korollar 2.14) und $\tilde{P}$ eine Permutationsmatrix (vgl. Satz 2.17) ist. Also ist $S_1^{-1} \cdots S_s^{-1} = \tilde{P}^{-1}\tilde{L}^{-1} = P \cdot L$. □

Korollar 3.6 *Es sei $\mathbb{K}$ ein Körper und $A, B \in \mathbb{K}^{n,m}$ in Treppennormalform. Falls es eine invertierbare Matrix $Z \in \mathbb{K}^{n,n}$ mit $A = ZB$ gibt, so gilt $A = B$, d.h., die Treppennormalform ist invariant unter Multiplikation mit nichtsingulären Matrizen von links.*

Beweis: Es seien a_i, b_i, $i = 1, \ldots, m$, die Spalten von A, B. Weiterhin seien $(1, j_1), \ldots, (s, j_s)$ die Pivotpositionen von B.

Wir zeigen mit vollständiger Induktion über r, $1 \leq r \leq s$:
Es gilt

$$Z = \left[\begin{array}{c|c} I_r & * \\ \hline 0 & Z_{n-r} \end{array}\right],$$

wobei Z_{n-r} invertierbar ist, und die ersten $j_{r+1} - 1$ Spalten von A und B stimmen überein. (Wir setzen $j_{s+1} := m + 1$.)

I.A.: Es gilt $b_k = 0$ für $1 \leq k \leq j_1 - 1$. Da $A = Z \cdot B$, gilt auch $a_k = 0$, $1 \leq k \leq j_1 - 1$. Weiter ist $b_{j_1} = [1, 0, \ldots, 0]^T$, und da Z invertierbar ist, ist auch $a_{j_1} \neq 0$. Da A in TNF ist, folgt, dass auch $a_{j1} = [1, 0, \ldots, 0]^T$. Weiterhin folgt, dass

$$Z = \left[\begin{array}{c|c} 1 & * \\ \hline 0 & \\ \vdots & Z_{n-1} \\ 0 & \end{array}\right].$$

Damit ist auch $a_k = b_k$, $k = j_1 + 1, \ldots, j_2 - 1$.

I.V.: Die Aussage gelte für ein r, $1 \leq r \leq s - 1$.

I.S.: Wir betrachten die Pivotposition $(r+1, j_{r+1})$. Da B in TNF ist, folgt

$$b_{j_{r+1}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow {\scriptstyle r+1}. \\ \\ \\ \\ \end{array}$$

Wegen $a_{j_{r+1}} = Zb_{j_{r+1}}$ und der Invertierbarkeit von Z_{n-r} folgt wie in der Induktionsannahme $a_{j_{r+1}} = b_{j_{r+1}}$ und

$$Z = \begin{array}{cc} \left[\begin{array}{ccc|c|c} 1 & & & & \\ & \ddots & & & * \\ & & 1 & & \\ \hline & & & 1 & * \\ \hline & & & & \\ & & & & Z_{n-(r+1)} \\ & & & & \end{array}\right] & \begin{array}{l} \\ {\scriptstyle r} \\ \\ {\scriptstyle 1} \\ \\ {\scriptstyle n-r-1} \\ \\ \end{array} \\ \begin{array}{ccc} {\scriptstyle r} & \qquad {\scriptstyle 1} & \qquad {\scriptstyle n-r-1} \end{array} & \end{array} \quad ,$$

und die ersten $j_{r+2} - 1$ Spalten von A und B sind gleich.

$\square$

Kapitel 4

Der Rang einer Matrix und die Lösung linearer Gleichungssysteme

Mit Hilfe des Gauß'schen Algorithmus lassen sich nicht nur die Inversen von Matrizen bestimmen, sondern auch die Lösung von linearen Gleichungssystemen. Dabei hat der Gauß'sche Algorithmus gewisse Nachteile auf dem Rechner (in endlicher Arithmetik), aber er ist eines der schnellsten Verfahren zur Lösung dieses zentralen Problems der linearen Algebra (und der Praxis in fast allen Anwendungen).
Bevor wir zu linearen Gleichungssystemen kommen, wollen wir noch eine wichtige Größe einführen.

Definition 4.1 *Die Anzahl s der Zeilen mit Pivotelementen (d.h., der Zeilen, die nicht gleich 0 sind) in der Treppennormalform einer Matrix $A \in K^{n,m}$ heißt* Rang *von A, wir schreiben*

$$\text{Rang}\,(A) = s.$$

Satz 4.2 *Sei K ein Körper und $A \in K^{n,m}$. Dann gilt:*

(1) *Es gibt nichtsinguläre Matrizen $Q \in K^{n,n}$ und $Z \in K^{m,m}$, so dass*

$$QAZ = \sum_{i=1}^{r} E_{ii} = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$$

genau dann, wenn $\text{Rang}\,(A) = r$.

(2) $\text{Rang}\,(A) = \text{Rang}\,(A^T)$.

(3) *Ist $A = B \cdot C$ mit $B \in K^{n,s}$ und $C \in K^{s,m}$, so gilt*

$$\begin{aligned} \text{Rang}\,(A) &\leq \text{Rang}\,(B), \\ \text{Rang}\,(A) &\leq \text{Rang}\,(C). \end{aligned}$$

(4) *Falls $Q \in K^{n,n}$ und $Z \in K^{m,m}$ invertierbar sind, so gilt* $\text{Rang}\,(A) = \text{Rang}\,(QAZ)$.

(5) *Es gibt Matrizen $B \in K^{n,s}$, $C \in K^{s,m}$, so dass $A = BC$ genau dann, wenn* $\text{Rang}\,(A) \leq s$.

(6) *Für quadratische Matrizen $A \in K^{n,n}$ sind folgende Aussagen äquivalent.*

  i) *A ist invertierbar.*
  ii) *Es gibt ein $B \in K^{n,n}$ mit $A \cdot B = I$.*
  iii) *Es gibt ein $\tilde{B} \in K^{n,n}$ mit $\tilde{B} \cdot A = I$.*
  iv) $\text{Rang}\,(A) = n$.
  v) $\text{Rang}\,(A^T) = n$.

Beweis:

(3a) Wir zeigen zunächst nur die erste Ungleichung von (3).

Sei Q eine Matrix, so dass QB in TNF ist,

$$QB = \begin{array}{l} \left[\begin{array}{c|cc|cc|cc|c|cc} 0 & 1 & * & 0 & * & 0 & * & & 0 & * \\ 0 & 0 & & 1 & * & 0 & * & & 0 & * \\ 0 & 0 & & 0 & & 1 & * & & \vdots & * \\ 0 & 0 & & 0 & & 0 & & \ddots & 0 & * \\ 0 & 0 & & 0 & & 0 & & & 1 & * \\ 0 & 0 & & 0 & & 0 & & & 0 & 0 \end{array}\right] \end{array} \begin{array}{l} \\ \\ \\ \\ \leftarrow r_B \\ \\ \end{array},$$

mit $\text{Rang}\,(B) = r_B$.
Wenn wir $QA = QB \cdot C$ bilden, so folgt, dass $QB \cdot C$ höchstens r_B Zeilen hat, die nicht Null sind, d.h., in der TNF von A sind höchstens r_B Zeilen von Null verschieden, d.h., $\text{Rang}\,(A) \leq \text{Rang}\,(B)$.

(1) Sei $\text{Rang}\,(A) = r$. Dann gibt es eine nichtsinguläre Matrix Q, so dass QA in Treppennormalform mit r Pivots ist:

$$QA = \left[\begin{array}{c|cc|cc|cc|c|cc} 0 & 1 & * & 0 & * & 0 & * & & 0 & * \\ 0 & 0 & & 1 & * & 0 & * & & 0 & * \\ 0 & 0 & & 0 & & 1 & * & & \vdots & * \\ 0 & 0 & & 0 & & 0 & & \ddots & 0 & * \\ 0 & 0 & & 0 & & 0 & & & 1 & * \\ 0 & 0 & & 0 & & 0 & & & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ \\ \\ \leftarrow r \\ \\ \end{array}$$

Betrachte $(QA)^T = A^T Q^T$. Dann gibt es eine Permutation P, so dass PA^TQ^T die Form

$$PA^TQ^T = \begin{array}{c} \left[\begin{array}{ccc|c} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ \hline & * & & 0 \end{array}\right] \\ \begin{array}{cc} \quad r \quad & \quad m-r \end{array} \end{array} \begin{array}{l} r \\ \\ n-r \end{array} =: \left[\begin{array}{c|c} I_r & 0 \\ \hline V & 0 \end{array}\right]$$

hat. Setze $Y = \left[\begin{array}{c|c} I_r & 0 \\ \hline -V & I_{n-r} \end{array}\right]$. Es folgt, dass $YPA^TQ^T = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. Mit $Z = P^TY^T$ ergibt sich $QAZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

Für die Umkehrung zeigen wir zunächst eine Folgerung aus (3a). Es sei $Z \in K^{m,m}$ eine invertierbare Matrix. Dann gilt wegen (3a)

$$\text{Rang}\,(A) = \text{Rang}\,(AZZ^{-1}) \leq \text{Rang}\,(AZ) \leq \text{Rang}\,(A),$$

folglich gilt überall das Gleichheitszeichen, und somit gilt:

Ist $Z \in K^{m,m}$ eine invertierbare Matrix, dann ist $\text{Rang}\,(A) = \text{Rang}\,(AZ)$.

Es seien nun also $Q \in K^{n,n}$ und $Z \in K^{m,m}$ invertierbare Matrizen, so dass $QAZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. Wegen des gerade Bewiesenen ist dann

$$r = \text{Rang}\,(QAZ) = \text{Rang}\,(QA).$$

Nach Korollar 3.6 folgt $\text{Rang}\,(A) = r$.

(2) Folgt direkt aus (1).

(3b) Die zweite Ungleichung von (3) folgt unter Verwendung von (2) und (3a).

(4) $\text{Rang}\,(QAZ) = \text{Rang}\,(AZ) = \text{Rang}\,(Z^TA^T) = \text{Rang}\,(A^T) = \text{Rang}\,(A)$.

(5) Sei $A = BC, \quad B \in K^{n,s},\ C \in K^{s,m}$, so folgt aus (3), dass $\text{Rang}\,(A) \leq \text{Rang}\,(C) \leq s$, da C natürlich höchstens Rang s haben kann.

Für die Umkehrung verwenden wir (1): Ist $r \leq s$ der Rang von A, dann gibt es nichtsinguläre Matrizen Q, Z, so dass $QAZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. Wir schreiben

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = {}_n\underset{s}{\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}} \underset{m}{\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}}_s = \tilde{B} \cdot \tilde{C}.$$

Es folgt, dass

$$A = \underbrace{Q^{-1}\tilde{B}}_{B}\underbrace{\tilde{C}Z^{-1}}_{C} =: BC$$

die gewünschte Zerlegung ist.

(6) $(i) \Longrightarrow (ii), (i) \Longrightarrow (iii)$ laut Definition der Invertierbarkeit.

$(iv) \Longleftrightarrow (v)$ wegen (2).

$(ii) \Longrightarrow (iv)$ wegen (3), analog $(iii) \Longrightarrow (iv)$.

$(iv) \Longrightarrow (i)$, denn nach (1) gibt es bei $\text{Rang}\,(A) = n$ invertierbare Matrizen Q und Z mit $A = QIZ = QZ$, also ist A invertierbar.

□

Bemerkung 4.3 *Eine Zerlegung* $A = B \cdot C$ *mit* $A \in K^{n,m}, B \in K^{n,r}, C \in K^{r,m}$*, wobei* $r = \text{Rang}\,(A)$*, nennt man im allgemeinen* Vollrangzerlegung*. Diese Zerlegung spielt eine große Rolle bei der Lösung von Optimierungsproblemen.*

Das Lösen von Gleichungen ist eine der wichtigsten Anwendungen der linearen Algebra. Insbesondere führen fast alle wissenschaftlich-technischen Probleme im Endeffekt auf die Lösung linearer Gleichungssysteme. Heutzutage sind Systeme mit bis zu 10 000 000 Gleichungen und Unbekannten lösbar (Autoindustrie, Flugtechnik, Optimierung, ...).

Definition 4.4

(i) *Ein* lineares Gleichungssystem *über* K *hat die Form*

$$Ax = b \tag{4.5}$$

mit $A = [a_{ij}] \in K^{n,m}$, $b = [b_i] \in K^{n,1}$, $x = [x_j] \in K^{m,1}$*. Das sind* n *Gleichungen in* m *Unbekannten:*

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1m}x_m &= b_1 \\ &\vdots \\ a_{n1}x_1 + \ldots + a_{nm}x_m &= b_n \end{aligned} \tag{4.6}$$

(ii) *Jedes* $x \in K^{m,1}$*, welches (4.5) erfüllt, heißt* Lösung *des linearen Gleichungssystems (4.5).*

(iii) *Die Menge aller Lösungen des linearen Gleichungssystems (4.5) heißt* Lösungsmenge *von (4.5). Die Lösungsmenge von (4.5) ist also eine Teilmenge von* $K^{m,1}$*.*

(iv) *Falls* $b = 0$*, so heißt das Gleichungssystem* homogen*. Zu jedem Gleichungssystem der Form (4.5) gibt es das zugeordnete homogene System*

$$Ax = 0. \tag{4.7}$$

Lemma 4.8 *Ist* x *eine Lösung von* $Ax = b$ *und* Z *die Menge aller Lösungen des zugeordneten homogenen Systems* $Ax = 0$*, so ist*

$$L = \{x + z | z \in Z\} \tag{4.9}$$

die Lösungsmenge von $Ax = b$*.*

Beweis: Sei $z \in Z$, so gilt

$$A(x + z) = Ax + Az = b + 0 = b.$$

Umgekehrt: Ist $\tilde{x}$ eine Lösung von $Ax = b$, dann müssen wir ein $z \in Z$ finden, so dass $\tilde{x} = x + z$ ist. Wir zeigen, dass $z = \tilde{x} - x$ diese Anforderung erfüllt, denn es ist Lösung des homogenen Systems:

$$A(\tilde{x} - x) = A\tilde{x} - Ax = b - b = 0.$$

□

Äquivalente Umformungen

Lemma 4.10 *Ist $T \in R^{n,n}$ invertierbar, so sind die Lösungsmengen von $Ax = b$ und $TAx = Tb$ identisch.*

Beweis: Sei x Lösung von $Ax = b$ dann gilt $TAx = Tb$. Also ist x auch Lösung von $TAx = Tb$. Umgekehrt sei y eine Lösung von $(TA)x = Tb$, d.h. $(TA)y = Tb$, also auch

$$T^{-1}(TA)y = T^{-1}(Tb)$$

und

$$Ay = b\,,$$

womit die Behauptung bewiesen ist. □

Wir können also nach Lemma 4.10 die Matrix $A \in R^{n,m}$ auf TNF bringen, ohne die Lösungsmenge zu ändern

$$TAx = \left[\begin{array}{c|cc|cc|cc|c|cc} & 1 & * & 0 & * & 0 & * & & 0 & * \\ & & & 1 & & 0 & & & \vdots & \\ 0 & & & & & 1 & & & & \\ & & 0 & & 0 & & & \ddots & 0 & \\ & & & & & & 0 & & 1 & \\ & & & & & & & & & 0 \end{array}\right] x = \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \end{bmatrix} = Tb$$

Es gibt eine Umordnung der Komponenten von x, welche durch eine Permutationsmatrix gegeben ist, so dass unser Gleichungssystem die Form

$$\underbrace{TAP^\top}_{C}\underbrace{Px}_{y} = \left[\begin{array}{ccc|c} 1 & & & \\ & \ddots & & * \\ & & 1 & \\ \hline & 0 & & 0 \end{array}\right] \begin{bmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_r \\ \tilde{b}_{r+1} \\ \vdots \\ \tilde{b}_n \end{bmatrix} = \underbrace{Tb}_{d} \tag{4.11}$$

hat (*vgl. Satz 4.2, (1)*).
Dabei ist $y = [y_1, \ldots, y_m]^\top = Px,\ C = TAP^\top,\ d = Tb$.

Wie sieht die Permutationsmatrix aus?

$$P^\top = \left[\begin{array}{c|c|c|c|c}
0 & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} & & & \\ \hline
1\;\;0\;\cdots\;0 & 0 & & \cdots & 0 \\ \hline
0 & & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} & & \\ \hline
0\;\;1\;\cdots\;0 & 0 & & \cdots & 0 \\ \hline
\vdots & & & \ddots & \\ \hline
0\;\;0\;\cdots\;1 & 0 & & \cdots & 0 \\ \hline
0 & & & & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix}
\end{array}\right]$$

Lemma 4.12 *Das Gleichungssystem $Ax = b$ hat mindestens eine Lösung, genau dann wenn*

$$\text{Rang}\, A = \text{Rang}\, [A, b]. \tag{4.13}$$

Beweis: Nach Lemma 4.10 reicht es, die Aussage für das Gleichungssystem (4.11) zu zeigen, denn Rang C = Rang A und Rang $[C, d]$ = Rang $[A, b]$.

Falls Rang (C) = Rang $[C, d]$, so ist $\tilde{b}_{r+1} = \cdots \tilde{b}_m = 0$ und damit ist

$$[y_1, \cdots, y_m]^\top = \left[\tilde{b}_1, \cdots \tilde{b}_r, 0, \cdots, 0\right]^\top$$

eine Lösung. (Es kann noch mehr Lösungen geben.)
Falls Rang $(A) <$ Rang $[A, b]$, so ist eines der $\tilde{b}_{r+1}, \cdots, \tilde{b}_m$ nicht gleich 0 und damit hat das System keine Lösung. □

Lemma 4.14 *Ein Verfahren zur Lösung des Gleichungssystems $Ax = b$.*

(1) Sei $A' = [A, b]$ die erweiterte Koeffizientenmatrix.
Transformiere A und A' auf TNF B und B'.

(2) Ist Rang $(A) <$ *Rang(A'), so besitzt das Gleichungssystem keine Lösung.*

(3) Ist $r =$ Rang (A) = Rang (A'), *so bilde das assoziierte Gleichungssystem*

$$Cy = d \quad \textit{mit} \quad C = QAP^\top,\ y = Px,\ d = Qb$$

Eine spezielle Lösung ist

$$\tilde{y} = \begin{bmatrix} d_1 \\ \vdots \\ d_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.15}$$

(4) Bestimme alle Lösungen des homogenen Systems $Cy = 0$ durch Rückwärtseinsetzen

$$\hat{y} = \begin{bmatrix} -\sum\limits_{j=r+1}^{m} c_{1j} y_j \\ \vdots \\ -\sum\limits_{j=r+1}^{m} c_{rj} y_j \\ y_{r+1} \\ \vdots \\ y_m \end{bmatrix} \qquad \text{für beliebige} \quad y_{r+1}, \cdots, y_m \in K. \tag{4.16}$$

Die Lösungsmenge ist:

$$\mathcal{L} = \left\{ x = P^\top y \mid y = \tilde{y} + \hat{y},\ \tilde{y} \text{ wie in (4.15)},\ \hat{y} \text{ wie in (4.16)} \right\}. \tag{4.17}$$

Wir können also folgende Übersicht über die Lösbarkeit des linearen Gleichungssystems

$$Ax = b, \quad A \in K^{n,m},\ b \in K^{n,1},\ x \in K^{m,1}$$

gewinnen. Sei $r = \text{Rang}\,(A)$, $r' = \text{Rang}\,([A, b])$.

Tabelle 4.18

$r < r'$	$r = m = r'$	$r = r' < m$
keine Lösung	*genau eine Lösung* $x = P^\top \tilde{y}$	*viele Lösungen, Lösungsmenge* $\mathcal{L}$

Beispiel 4.19 $\quad K = \mathbb{Q}$

$$\begin{array}{rcrcrcrcr} 1\,x_1 & + & 2\,x_2 & + & 2\,x_3 & + & 1\,x_4 & = & 1 \\ & & 1\,x_2 & + & & & 3\,x_4 & = & 0 \\ 1\,x_1 & + & & & 3\,x_3 & + & & = & 2 \\ 2\,x_1 & + & 3\,x_2 & + & 5\,x_3 & + & 4\,x_4 & = & 3 \\ 1\,x_1 & + & 1\,x_2 & + & 3\,x_3 & + & 3\,x_4 & = & 2 \end{array}$$

$$A = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 3 \\ 1 & 0 & 3 & 0 \\ 2 & 3 & 5 & 4 \\ 1 & 1 & 3 & 3 \end{bmatrix},\ x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},\ b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \\ 2 \end{bmatrix}$$

Gauß'scher Algorithmus

$$[A,b] \rightsquigarrow \left[\begin{array}{rrrr|r} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & -2 & 1 & -1 & 1 \\ 0 & -1 & 1 & 2 & 1 \\ 0 & -1 & 1 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{rrrr|r} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 1 & 5 & 1 \end{array}\right]$$

$$\rightsquigarrow \left[\begin{array}{rrrr|r} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{rrrr|r} 1 & 0 & 2 & -5 & 1 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\rightsquigarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & -15 & -1 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Es gilt: $\operatorname{Rang}(A) = \operatorname{Rang}([A,b])$. Deshalb gibt es Lösungen.
Es ist keine Permutation notwendig:

$$\tilde{x} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \hat{x} = \begin{bmatrix} 15x_4 \\ -3x_4 \\ -5x_4 \\ x_4 \end{bmatrix}$$

$$\mathcal{L} = \left\{ x \in \mathbb{Q}^4 : \ x = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 15 \\ -3 \\ -5 \\ 1 \end{bmatrix} \mid x_4 \in \mathbb{Q} \right\}$$

Kapitel 5

Die Determinante

Wir haben bereits eine wichtige (Funktion) Größe kennengelernt, den Rang einer Matrix. Jetzt wollen wir eine weitere Funktion betrachten.

Definition 5.1 *Eine Permutation der Zahlen* $\{1, \cdots, n\}$ *ist eine bijektive Abbildung*

$$\sigma : \{1, \cdots, n\} \to \{1, \cdots, n\}.$$

Die Menge aller Permutationen von $\{1, \cdots, n\} \to \{1, \cdots, n\}$ *bezeichnen wir mit* S_n.

Sei $v = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix}$ und P_σ eine Permutationsmatrix, so ist $P_\sigma v = w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ und

$$\sigma(i) = w_i, \ i = 1, \cdots, n$$

gibt eine Permutation an. Für $n = 3$ gibt es folgende Permutationen:

123	213	312
132	231	321

Satz 5.2 *Die Anzahl der möglichen Permutationen von* $\{1, \cdots, n\}$ *ist* $n\,! = 1 \cdot 2 \cdots (n-1) \cdot n$.

Beweis: Vollständige Induktion.

I.A.: Für $n = 1$ gibt es nur eine Permutation

$$1\,! = 1.$$

I.V.: Die Behauptung sei richtig für $n = k$.

I.S.: Sei $n = k+1$, dann können wir für jede Permutation von $\{1, \cdots, k\}$, wovon es $k!$ Stück gibt, die Zahl $k+1$ an jede beliebige Stelle setzen, also gibt es

$$(k+1) \cdot (k!) = (k+1)!$$

Permutationen.

$\square$

Definition 5.3 *Das* <u>*Signum*</u> *einer Permutation ist definiert durch*

$$\mathrm{sgn}\,(\sigma) = \left\{ \begin{array}{rl} 1, & \textit{bei gerader} \\ -1, & \textit{bei ungerader} \end{array} \right\} \textit{Anzahl von Paaren } (i,j) \textit{ mit } i > j \textit{ und } \sigma(i) < \sigma(j)\,.$$

Beispiel 5.4 $\{1,2,3,4,5\}$, $\sigma_1 = 1\ 3\ 5\ 4\ 2$, $\sigma_2 = 1\ 3\ 4\ 5\ 2$

$$\sigma_1 : \begin{array}{c|ccccc} i & 1 & 2 & 3 & 4 & 5 \\ \hline \sigma(i) & 1 & 3 & 5 & 4 & 2 \end{array} \qquad \sigma_2 : \begin{array}{c|ccccc} i & 1 & 2 & 3 & 4 & 5 \\ \hline \sigma(i) & 1 & 3 & 4 & 5 & 2 \end{array}$$

4 Paare, $\mathrm{sgn}\,(\sigma_1) = 1$

3 Paare, $\mathrm{sgn}\,(\sigma_2) = -1.$

Definition 5.5 *Sei R ein kommutativer Ring mit Eins und $A = [a_{ij}] \in R^{n,n}$. Dann heißt*

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}\,(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \tag{5.6}$$

die <u>*Determinante*</u> *von A. Dies ist eine Abbildung*

$$\begin{aligned} \det &: R^{n,n} \to R \\ &: A \longmapsto \det(A) \end{aligned}$$

Beispiel 5.7 $A = [a_{ij}] \in R^{3,3}$, (Regel von Sarrus)

$$\begin{array}{lll} \sigma_1 = 1\,2\,3 & \sigma_3 = 2\,1\,3 & \sigma_5 = 3\,1\,2 \\ \sigma_2 = 1\,3\,2 & \sigma_4 = 2\,3\,1 & \sigma_6 = 3\,2\,1 \\ \mathrm{sgn}\,(\sigma_1) = 1 & \mathrm{sgn}\,(\sigma_3) = -1 & \mathrm{sgn}\,(\sigma_5) = 1 \\ \mathrm{sgn}\,(\sigma_2) = -1 & \mathrm{sgn}\,(\sigma_4) = 1 & \mathrm{sgn}\,(\sigma_6) = -1 \end{array}$$

$$\begin{aligned} \det(A) \;=\; & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned}$$

oder für $\quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in R^{2,2}$

$$\begin{array}{ll} \sigma_1 = 1\,2 & \sigma_2 = 2\,1 \\ \operatorname{sgn}\sigma_1 = 1 & \operatorname{sgn}\sigma_2 = -1 \end{array}$$

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Lemma 5.8 *Sei $A \in R^{n,n}$.*

i) Ist A obere oder untere Dreiecksmatrix, so ist $\det(A) = a_{11} \cdots a_{n,n}$.

ii) Hat A eine Zeile oder Spalte von Nullen, so ist $\det(A) = 0$.

iii) Die Determinante einer Permutationsmatrix ist gleich dem Signum der zugehörigen Permutation.

Beweis:

i) Sei $\sigma \in S_n$, $\sigma \neq 1\,2 \cdots n$, so gibt es ein i mit $i > \sigma(i)$. Also gilt für eine obere Dreiecksmatrix

$$a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} = 0, \quad \text{da } a_{i,\sigma(i)} = 0\,.$$

Also bleibt nur

$$\det(A) \;=\; \underbrace{\operatorname{sgn}(1\,2\,3 \cdots n)}_{=1, \text{ da } 0 \text{ Paare}} \cdot a_{11} \cdots a_{nn} \;=\; a_{11} \cdots a_{nn}$$

Für untere Dreiecksmatrizen gilt der Beweis analog.

ii) Falls A eine Nullzeile hat, so gilt $a_{k,l} = 0,\ l = 1, \cdots, n$, also ist in jedem der Produkte

$$a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

mindestens ein Faktor gleich 0. Für Nullspalten gilt das analog.

iii) Es gibt natürlich nur ein Produkt, welches ungleich Null ist; das ist $a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$. Alle anderen sind 0 und damit folgt die Behauptung.

□

Beispiel 5.9 Die Determinanten der Elementarmatrizen $P_{ij}, M_i(\lambda), G_{ij}(\lambda)$ sind:

$$\begin{array}{lcr} \det P_{ij} & = & -1, \\ \det M_i(\lambda) & = & \lambda, \\ \det G_{ij}(\lambda) & = & 1. \end{array}$$

Rechenregeln mit Determinanten

Lemma 5.10 *Sei $A = [a_{ij}] \in K^{n,n}, K$ Körper.*

(a) Hat A zwei gleiche Zeilen, so ist $\det A = 0$.

(b) Die Multiplikation einer Zeile von A mit $\lambda \in K$ führt zur Multiplikation der Determinante mit λ.

$$\det(M_j(\lambda) \cdot A) = \lambda \det A = \det(M_j(\lambda)) \cdot \det A. \tag{5.11}$$

(c) Wenn ich das Vielfache einer Zeile zu einer anderen addiere, so ändert sich die Determinante nicht.

$$\begin{aligned} \det(G_{ij}(\lambda) \cdot A) &= \det A = \det(G_{ij}(\lambda)) \cdot \det A \\ \det(G_{ij}^\top(\lambda) \cdot A) &= \det A = \det\left(G_{ij}^\top(\lambda)\right) \cdot \det A \end{aligned} \tag{5.12}$$

(d) Wenn ich zwei Zeilen von A vertausche, so erhalte ich das Negative der Determinante.

$$\det(P_{ij} \cdot A) = -\det A = \det P_{ij} \cdot \det A \tag{5.13}$$

Beweis: Wir verwenden die Notation: $\prod_{i=1}^{n} a_{ij} = a_{1j} \cdots a_{nj}$.

(a) Da A zwei gleiche Zeilen hat, gibt es ein Indexpaar i, i' mit

$$a_{i,j} = a_{i',j} \quad \forall j = 1, \ldots, n\,.$$

Für jedes $\sigma \in S_n$ setze

$$\sigma'(l) = \begin{cases} \sigma(l) & l \neq i, i' \\ \sigma(i) & l = i' \\ \sigma(i') & l = i \end{cases} .$$

Dann ist $\sigma' \in S_n$ und $\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\sigma')$, aber $\prod_{l=1}^{n} a_{l,\sigma(l)} = \prod_{l=1}^{n} a_{l,\sigma'(l)}$.

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{l=1}^{n} a_{l,\sigma(l)} \right) \\ &= -\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma') \left(\prod_{l=1}^{n} a_{l,\sigma'(l)} \right) \\ &= -\sum_{\sigma' \in S_n} \operatorname{sgn}(\sigma') \left(\prod_{l=1}^{n} a_{l,\sigma'(l)} \right) \\ &= -\det A \end{aligned}$$

Somit gilt $\det(A) = 0$, da $\det(A) \in K$.

(b) Sei $\tilde{A} = M_j(\lambda) \cdot A = [\tilde{a}_{li}]$. Dann gilt: $\tilde{a}_{li} = \begin{cases} a_{li}, & l \neq j \\ \lambda a_{li}, & l = j. \end{cases}$

Also

$$\begin{aligned}
\det \tilde{A} &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \tilde{a}_{1,\sigma(1)} \cdots \tilde{a}_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{l=1}^{n} \tilde{a}_{l,\sigma(l)} \right) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{\substack{l=1 \\ l \neq j}}^{n} a_{l,\sigma(l)} \right) \cdot \underbrace{\tilde{a}_{j,\sigma(j)}}_{\lambda a_{j,\sigma(j)}} \\
&= \lambda \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{l=1}^{n} a_{l,\sigma(l)} = \lambda \cdot \det A.
\end{aligned}$$

(c) Sei $\tilde{A} = G_{ij}(\lambda) A = [\tilde{a}_{ls}]$. Dann gilt: $\tilde{a}_{ls} = \begin{cases} a_{ls}, & l \neq j \\ a_{ls} + \lambda a_{is}, & l = j \end{cases}$

$$\begin{aligned}
\det \tilde{A} &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{l=1}^{n} \tilde{a}_{l,\sigma(l)} \right) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{\substack{l=1 \\ l \neq j}}^{n} a_{l,\sigma(l)} \right) \cdot (a_{j,\sigma(j)} + \lambda a_{i,\sigma(j)}) \\
&= \left[\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{l=1}^{n} a_{l,\sigma(l)} \right] + \lambda \left[\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{\substack{l=1 \\ l \neq j}}^{n} a_{l,\sigma(l)} \right) a_{i,\sigma(j)} \right] \\
&= \det A + \lambda \det \hat{A},
\end{aligned}$$

wobei $\hat{A}$ zwei gleiche Zeilen hat und nach (a) folgt $\det \hat{A} = 0$, also

$$\det \tilde{A} = \det A \cdot 1 = \det A \cdot \underbrace{\det G_{ij}(\lambda)}_{1}.$$

Der Beweis für $G_{ij}^{\top}(\lambda)$ erfolgt analog.

(d)

$$P_{ij} = \begin{bmatrix}
1 &&&&&&&&& \\
& \ddots &&&&&&&& \\
&& 1 &&&&&&& \\
&&& 0 & \cdots & 1 &&&& \\
&&&& 1 &&&&& \\
&&& \vdots & \ddots & \vdots &&&& \\
&&&&& 1 &&&& \\
&&& 1 & \cdots & 0 &&&& \\
&&&&&&& 1 && \\
&&&&&&&& \ddots & \\
&&&&&&&&& 1
\end{bmatrix} = G_{ij}(1) G_{ij}^{\top}(-1) G_{ij}(1) M_j(-1),$$

denn

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

Also folgt:

$$\begin{aligned} \det(P_{ij}A) &= \det\left[G_{ij}(1)G_{ij}^\top(-1)G_{ij}(1)M_j(-1)A\right] \\ &= \det\left[G_{ij}(1)M_j(-1)A\right] = \det\left[M_j(-1)A\right] = -\det A. \end{aligned}$$

□

Satz 5.14 *Sei K ein Körper mit $A, B \in K^{n,n}$, so gilt*

(a) $\det(A \cdot B) = \det A \cdot \det B$,

(b) $\det A^\top = \det A$.

Beweis:

(a) Nach Satz 3.1 (TNF) läßt sich A als Produkt von Elementarmatrizen und einer Matrix in TNF schreiben, d.h., $A = S_1^{-1} \cdots S_t^{-1} \tilde{A}$, wobei $\tilde{A}$ entweder die Einheitsmatrix ist, oder mindestens eine Nullzeile hat, falls Rang $(A) < n$.

Falls $\tilde{A}$ eine Nullzeile hat, so hat auch $S_t \cdots S_1 A \cdot B = \tilde{A}B$ eine Nullzeile.
Wegen $\det A = \pm\lambda \det \tilde{A}$ gilt für den Fall, daß $\tilde{A}$ eine Nullzeile hat,

$$\det \tilde{A} = 0 = \det A = \det \tilde{A}B = \det AB.$$

Wir können uns also auf den Fall $\det A \neq 0$ beschränken, d.h., $\tilde{A} = I_n$. Dann ist

$$\begin{aligned} \det A \cdot \det B &= \det(S_1^{-1} \cdots S_t^{-1}) \det B = \det S_1^{-1} \cdots \det S_t^{-1} \cdot \det B\,, \\ \det AB &= \det(S_1^{-1} \cdots S_t^{-1} B) = \det S_1^{-1} \cdots \det S_t^{-1} \cdot \det B\,. \end{aligned}$$

(b) Mit $A = S_1^{-1} \cdots S_t^{-1} \tilde{A}$ gilt

$$\begin{aligned} \det(A^\top) &= \det(S_1^{-1} \cdots S_t^{-1} \tilde{A})^\top = \begin{cases} \det(S_t^{-\top} \cdots S_1^{-\top}), & \tilde{A} = I \\ 0, & \text{sonst}, \end{cases} \\ &= \begin{cases} \det S_t^{-\top} \cdots \det S_1^{-\top} \\ 0 \end{cases} \\ &= \begin{cases} \det S_t^{-1} \cdots \det S_1^{-1} \\ 0 \end{cases} \\ &= \det A\,. \end{aligned}$$

□

Beispiel 5.15

$$
\begin{aligned}
\det\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}
&= 2\cdot 3\cdot 4\cdot \det\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & \frac{3}{2} & 2 & \frac{5}{2} \\ 1 & \frac{4}{3} & \frac{5}{3} & 2 \\ 1 & \frac{5}{4} & \frac{3}{2} & \frac{7}{4} \end{bmatrix} \\
&= 24\cdot\det\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -\frac{1}{2} & -1 & -\frac{3}{2} \\ 0 & -\frac{2}{3} & -\frac{4}{3} & -2 \\ 0 & -\frac{3}{4} & -\frac{3}{2} & -\frac{9}{4} \end{bmatrix} \\
&= \left(-\tfrac{1}{2}\right)\left(-\tfrac{2}{3}\right)\left(-\tfrac{3}{4}\right)\cdot 24\cdot\det\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{bmatrix} = 0\,,
\end{aligned}
$$

da drei gleiche Zeilen.

Beispiel 5.16 $$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = aei + bfg + cdh - ceg - afh - bdi\,.$$

Definition 5.17 *Sei $A \in R^{n,n}$. Dann heißt die Matrix $A(s,t) \in R^{n-1,n-1}$, die durch das Streichen der s-ten Zeile und der t-ten Spalte von A entsteht, Minor von A.*
Man bilde die Matrix $B = [b_{ij}]$ mit $b_{ij} = (-1)^{i+j}\det A(j,i),\ \ i,j = 1,\dots,n$.
B heißt Adjungierte von A und wird als $\mathrm{adj}\,(A)$ bezeichnet.

Satz 5.18 *Für eine Matrix $A \in R^{n,n}$, ihre Determinante und ihre Adjungierte gilt:*

$$\det(A)\cdot I = A\cdot \mathrm{adj}\,(A) = \mathrm{adj}\,(A)\cdot A.$$

Beweis: Sei $B = [b_{ij}] = \underbrace{\mathrm{adj}(A)}_{=[c_{ij}]}\cdot A$.

$$
\begin{aligned}
b_{ij} &= \sum_{k=1}^{n} c_{ik}a_{kj} \\
&= \sum_{k=1}^{n} (-1)^{i+k}\det A(k,i)\cdot a_{kj}\,.
\end{aligned}
$$

Nun gilt aber $\det A(k,i) = (-1)^{k+i}\det \tilde{A}_{k,i}$, wobei

$$
\tilde{A}_{k,i} = \begin{bmatrix}
a_{11} & \cdots & a_{1,i-1} & 0 & a_{1,i+1} & \cdots & a_{1n} \\
\vdots & & & \vdots & & & \vdots \\
a_{k-1,1} & \cdots & a_{k-1,i-1} & 0 & a_{k-1,i+1} & \cdots & a_{k-1,n} \\
a_{k1} & \cdots & a_{k,i-1} & 1 & a_{k,i+1} & \cdots & a_{kn} \\
a_{k+1,1} & \cdots & a_{k+1,i-1} & 0 & a_{k+1,i+1} & \cdots & a_{k+1,n} \\
\vdots & & & \vdots & & & \vdots \\
a_{n1} & \cdots & a_{n,i-1} & 0 & a_{n,i+1} & \cdots & a_{nn}
\end{bmatrix},
$$

denn durch $(k-1)$ Zeilenvertauschungen und $(i-1)$ Spaltenvertauschungen kann man $\tilde{A}_{ki}$ auf die Form $\left[\begin{array}{c|c} 1 & * \\ \hline 0 & A(k,i) \end{array}\right]$ bringen und dafür gibt die Übung

$$\begin{aligned} \det \tilde{A}_{k,i} &= (-1)^{i-1+k-1} \det \left[\begin{array}{c|c} 1 & * \\ \hline 0 & A(k,i) \end{array}\right] \\ &= (-1)^{i+k} \det A(k,i). \end{aligned}$$

Also folgt

$$\begin{aligned} b_{ij} &= \sum_{k=1}^{n} (-1)^{i+k} \det A(k,i) a_{kj} \\ &= \sum_{k=1}^{n} (-1)^{i+k} (-1)^{i+k} \det \tilde{A}_{k,i} a_{kj} \\ &= \sum_{k=1}^{n} \det \begin{bmatrix} a_{11} & \cdots & a_{1,i-1} & 0 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ & & & 0 & & & \\ a_{k1} & \cdots & a_{k,i-1} & a_{kj} & a_{k,i+1} & \cdots & a_{kn} \\ & & & 0 & & & \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & & a_{n,i-1} & 0 & a_{n,i+1} & \cdots & a_{nn} \end{bmatrix} \\ &= \det \left[\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1,i1} & a_{1j} & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & a_{nj} & a_{n,i+1} & \cdots & a_{nn} \end{array}\right] \qquad \text{(s. Übung)} \\ &= \begin{cases} 0, & i \neq j \\ \det A, & i = j \end{cases} \\ &= \delta_{ij} \cdot \det A \end{aligned}$$

Der Beweis für die andere Aussage $A \cdot \operatorname{adj}(A) = \det A \cdot I$ ist analog. □

Korollar 5.19 *Sei $A \in K^{n,n}$ und $\det A \neq 0$ in K so ist A invertierbar und es gilt*

(a) $$A^{-1} = \frac{1}{\det A} \operatorname{adj}(A). \tag{5.20}$$

(b) $$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A(i,j) \tag{5.21}$$

(Laplace-Entwicklung der Determinante nach der i-ten Zeile),

(c) $$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A(i,j) \tag{5.22}$$

(Laplace-Entwicklung der Determinante nach der j-ten Spalte).

Beweis:

(a) Klar aus Satz 5.18, denn falls $\frac{1}{\det A}$ existiert, so existiert $A^{-1} = \frac{1}{\det A}\mathrm{adj}\,(A)$.

(b) Folgt aus $A \cdot \mathrm{adj}\, A = \det A \cdot I$.

(c) Für beliebiges j gilt nach Satz 5.18

$$\det A \cdot \delta_{jj} = \sum_{k=1}^{n} (-1)^{j+k} \det A(k,j) a_{k,j}\,.$$

□

Beispiel 5.23

$$\begin{array}{cl}
\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} & = \quad 1 \cdot (+1) \det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} \\
\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix} & \quad\; +2 \cdot (-1) \det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} \\
& \quad\; +3 \cdot (+1) \det \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} \\
& = (45-48) - 2 \cdot (36-42) + 3 \cdot (32-35) \\
& = -3 + 12 - 9 = 0
\end{array}$$

$\longrightarrow$ *nicht invertierbar.*

$$\begin{array}{cl}
\det \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{bmatrix} & = 1(+1) \det \begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix} \\
& \quad +2(-1) \det \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix} \\
& \quad +3(+1) \det \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} \\
& = 2 + 0 - 3 \;=\; -1.
\end{array}$$

$\longrightarrow$ *invertierbar*

Korollar 5.24 ***Cramer'sche Regel***

(nur von theoretischem Wert, nie auf Rechner anwenden für große n)

Sei K ein Körper, $A \in K^{n,n}, b \in K^{n,1}$ und $\mathrm{Rang}\, A = \mathrm{Rang}\,[A,b] = n$.
Dann ist die Lösung des Gleichungssystems

$$Ax = b$$

gegeben durch

$$x = A^{-1}b = \frac{1}{\det A}\mathrm{adj}\,(A)\cdot b \tag{5.25}$$

also

$$x_i = \frac{\det[a_1,\cdots,a_{i-1},b,a_{i+1},\cdots,a_n]}{\det A}, \qquad i = 1,\cdots,n\,. \tag{5.26}$$

Man muß also, um x zu berechnen, $n+1$ Determinanten berechnen. Das geht viel billiger mit dem Gauß'schen Algorithmus. (Man berechnet übrigens in der Praxis Determinanten auch mit dem Gauß'schen Algorithmus.)

Beispiel 5.27

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

$$\begin{aligned}
\det\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix}\begin{matrix}\leftarrow \\ \\ \\ \\ \end{matrix} &= 1\cdot\det\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} - 1\cdot\det\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \\
&= 2\cdot(4-1-1) - (4+1-2-1) = 2,
\end{aligned}$$

$$\begin{aligned}
\det\begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}\begin{matrix}\leftarrow \\ \\ \\ \\ \end{matrix} &= 2\cdot\det\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} - 1\cdot\det\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \\
&= 2\cdot(4+1-2-1) - (4-1) = 1,
\end{aligned}$$

$$\begin{aligned}
\det\begin{array}{c} \downarrow\qquad\quad\;\; \\ \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \end{array} &= 2\cdot\det\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} - 1\cdot\det\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \\
&= 2\cdot(4-2-2) - (2-2-1) = 1,
\end{aligned}$$

$$\begin{aligned}
\det\begin{array}{c} \downarrow\qquad\quad\;\; \\ \begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array} &= 2\cdot\det\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} - 1\cdot\det\begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\
&= 2\cdot(4+1-1-2) - (2+1-1) = 2,
\end{aligned}$$

$$\begin{aligned}\det A &= 2\cdot\det\begin{bmatrix}2&1&0\\1&2&1\\0&1&2\end{bmatrix} - 1\cdot\det\left[\begin{array}{c|cc}1&1&0\\\hline 0&2&1\\0&1&2\end{array}\right]\\ &= 2\cdot(8-2-2)-3=5,\end{aligned}$$

$$x=\begin{bmatrix}\frac{2}{5}\\\frac{1}{5}\\\frac{1}{5}\\\frac{2}{5}\end{bmatrix},\qquad \text{Probe: }\begin{bmatrix}2&1&0&0\\1&2&1&0\\0&1&2&1\\0&0&1&2\end{bmatrix}\begin{bmatrix}\frac{2}{5}\\\frac{1}{5}\\\frac{1}{5}\\\frac{2}{5}\end{bmatrix}=\begin{bmatrix}\frac{5}{5}\\\frac{5}{5}\\\frac{5}{5}\\\frac{5}{5}\end{bmatrix}=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}.$$

Mit Gauß:

$$\left[\begin{array}{cccc|c}2&1&0&0&1\\1&2&1&0&1\\0&1&2&1&1\\0&0&1&2&1\end{array}\right]\rightsquigarrow\left[\begin{array}{cccc|c}2&1&0&0&1\\0&\frac{3}{2}&1&0&\frac{1}{2}\\0&1&2&1&1\\0&0&1&2&1\end{array}\right]\rightsquigarrow\left[\begin{array}{cccc|c}2&1&0&0&1\\0&\frac{3}{2}&1&0&\frac{1}{2}\\0&0&\frac{4}{3}&1&\frac{2}{3}\\0&0&1&2&1\end{array}\right]$$
$$\rightsquigarrow\left[\begin{array}{cccc|c}2&1&0&0&1\\0&\frac{3}{2}&1&0&\frac{1}{2}\\0&0&\frac{4}{3}&1&\frac{2}{3}\\0&0&0&\frac{5}{4}&\frac{1}{2}\end{array}\right]$$

Rückwärts einsetzen:

$$\begin{aligned}&x_4=\tfrac{2}{5},\\&x_3=\tfrac{3}{4}\left(\tfrac{2}{3}-\tfrac{2}{5}\right)=\tfrac{3}{4}\cdot\tfrac{4}{15}=\tfrac{1}{5},\\&x_2=\tfrac{2}{3}\left(\tfrac{1}{2}-\tfrac{1}{5}\right)=\tfrac{2}{3}\cdot\tfrac{3}{10}=\tfrac{1}{5},\\&x_1=\tfrac{1}{2}\left(1-\tfrac{1}{5}\right)=\tfrac{2}{5}.\end{aligned}$$

Man kann natürlich auch andere Matrixfunktionen außer der Determinante betrachten. Eine für einige Anwendungen in der Kombinatorik wichtige Funktion ist die *Permanente* einer Matrix

$$\underline{\operatorname{per}(A)}=\sum_{\sigma\in S_n}\prod_{j=1}^{n}a_{j,\sigma(j)}. \tag{5.28}$$

Man läßt also einfach die Vorzeichen weg und erhält das Ergebnis wie bei der Determinante.

Beispiel 5.29

$$\operatorname{per}\begin{bmatrix}1&2&3\\1&2&3\\1&2&3\end{bmatrix}=6+6+6+6+6+6=36.$$

Kapitel 6

Eigenwerte von Matrizen

Neben Rang und Determinante einer Matrix gibt es noch weitere wichtige Größen, die eine Rolle bei Matrizen spielen. Dazu gehören das charakteristische Polynom und dessen Nullstellen, die Eigenwerte der Matrix. Wir haben schon in den Übungen die Menge der Polynome eingeführt:

$$\mathcal{P} = \left\{ \sum_{i=0}^{n} a_i x^i \mid a_i \in R \right\}, \text{ wobei } R \text{ ein Ring ist.}$$

Definition 6.1 *Sei $A \in R^{n,n}$, R ein kommutativer Ring mit Eins-Element. Dann heißt das Polynom*

$$P_A(\lambda) = \det(\lambda I_n - A) \tag{6.2}$$

das charakteristische Polynom von A.

Warum ist das überhaupt ein Polynom?

$$P_A(\lambda) = \det(\lambda I_n - A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} (\lambda \delta_{j,\sigma(j)} - a_{j,\sigma(j)})$$

ist eine Summe von Produkten von Termen der Form $(\lambda \delta_{j,\sigma(j)} - a_{j,\sigma(j)})$ und damit ein Polynom. Es ist klar, dass einer dieser Terme die Form $\prod\limits_{j=1}^{n} (\lambda - a_{jj})$ hat, dies ist ein Polynom vom Grad n. Somit kann man $P_A(\lambda)$ schreiben als

$$P_A(\lambda) = \prod_{j=1}^{n} (\lambda - a_{jj}) + \sum_{\substack{\sigma \in S_n \\ \sigma \neq 12 \cdots n}} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} (\lambda \delta_{j,\sigma(j)} - a_{j,\sigma(j)}) \,.$$

In jedem Term des zweiten Summanden gibt es mindestens zwei Faktoren, die kein λ enthalten. Damit ist der Grad des zweiten Summanden höchstens $n-2$. Es folgt

$$P_A(\lambda) = \lambda^n - \left(\sum_{j=1}^{n} a_{jj} \right) \lambda^{n-1} + q(\lambda)$$

mit $q(\lambda)$ vom Grad $\leq n-2$.

Definition 6.3 *Sei $A \in R^{n,n}$, R ein kommutativer Ring mit Eins-Element. Dann heißt die Summe*

$$\operatorname{tr}(A) = \operatorname{spur}(A) = \sum_{j=1}^{n} a_{jj} \tag{6.4}$$

die Spur *der Matrix A. (Englisch:* trace*).*

Damit folgt

$$P_A(\lambda) = \lambda^n - \operatorname{tr}(A)\lambda^{n-1} + q(\lambda) \qquad \text{mit } \operatorname{grad}(q) \leq n-2.$$

Man hat also zu jeder Matrix ein Polynom. Umgekehrt kann man auch zu jedem Polynom vom Grad n mit führendem Term λ^n eine Matrix konstruieren, die dieses als charakteristisches Polynom hat.

Lemma 6.5 *Sei $p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$, so ist $p(\lambda)$ charakteristisches Polynom von*

$$A_p = \begin{bmatrix} 0 & & & & -a_n \\ 1 & 0 & & & \\ & \ddots & \ddots & & \vdots \\ & & \ddots & 0 & -a_2 \\ & & & 1 & -a_1 \end{bmatrix}. \tag{6.6}$$

Beweis: mittels vollständiger Induktion:
I.A.: Für $n = 1$, d.h. $p(\lambda) = \lambda + a_1$, gilt mit $A_p = [-a_1]$

$$P_{A_p}(\lambda) = \det(\lambda I_1 - A_p) = \lambda + a_1 = p(\lambda).$$

I.V.: Für Polynome kleineren Grades als n sei die Behauptung bewiesen.

I.S.:

$$\det(\lambda I_n - A_p) = \det \begin{array}{c} \downarrow \\ \begin{bmatrix} \lambda & & & a_n \\ -1 & \ddots & & \vdots \\ & \ddots & \lambda & a_2 \\ & & -1 & \lambda + a_1 \end{bmatrix} \end{array}$$

$$= \lambda \cdot \underbrace{\det \begin{bmatrix} \lambda & & & a_{n-1} \\ -1 & \ddots & & \vdots \\ & \ddots & \lambda & a_2 \\ & & -1 & \lambda + a_1 \end{bmatrix}}_{\text{nach I.V.}} + \det \begin{bmatrix} 0 & 0 & \cdots & 0 & a_n \\ -1 & \lambda & \cdots & 0 & a_{n-2} \\ & \ddots & \ddots & \vdots & \vdots \\ & & -1 & \lambda & a_2 \\ & & & -1 & \lambda + a_1 \end{bmatrix}$$

$$= \lambda \cdot (\lambda^{n-1} + a_1\lambda^{n-2} + \cdots + a_{n-2}\lambda + a_{n-1})$$
$$+ (-1)^{(n-1)+1} a_n \cdot \det \underbrace{\begin{bmatrix} -1 & \lambda & & \\ & \ddots & \ddots & \\ & & -1 & \lambda \\ & & & -1 \end{bmatrix}}_{n-2}$$

$$
\begin{aligned}
&= \lambda^n + a_1\lambda^{n-1} + \cdots a_{n-2}\lambda^2 + a_{n-1}\lambda + (-1)^{2n-2}a_n \\
&= p(\lambda).
\end{aligned}
$$

□

Definition 6.7 *Die zu einem Polynom $p(\lambda)$ konstruierte Matrix aus (6.6) heißt* Begleitmatrix zu $p(\lambda)$.

Definition 6.8 *Sei K ein Körper und seien $A, B \in K^{n,n}$. A und B heißen* ähnlich, *falls es ein invertierbares $Z \in K^{n,n}$ gibt, so dass*

$$A = Z^{-1}BZ.$$

Satz 6.9 *Wenn zwei Matrizen $A, B \in K^{n,n}$ ähnlich sind, so besitzen sie das gleiche charakteristische Polynom.*

Beweis: Sei $A = Z^{-1}BZ$, dann ist

$$
\begin{aligned}
\det(\lambda I - A) &= \det(\lambda I - Z^{-1}BZ) \\
&= \det\left(Z^{-1}(\lambda Z Z^{-1} - B)Z\right) \\
&\overset{\text{Satz 5.14}}{=} \det Z^{-1} \det(\lambda I - B) \det Z \\
&\overset{\text{Satz 5.14}}{=} \det\left(Z^{-1}Z\right) \det(\lambda I - B) \\
&= \det(\lambda I - B).
\end{aligned}
$$

□

Die Umkehrung von Satz 6.9 gilt nicht immer (aber wenn das charakteristische Polynom in Linearfaktoren zerfällt, vgl. Satz 13.12).

Ein fundamentaler Satz der linearen Algebra ist der folgende Satz von Cayley und Hamilton.

Satz 6.10 ***Satz von Cayley und Hamilton***
Sei K ein Körper und $A \in K^{n,n}$ mit dem charakteristischen Polynom $P_A(\lambda)$. Dann erfüllt A die Gleichung

$$0 = P_A(A) = A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI_n\,. \tag{6.11}$$

Beweis: Betrachte die Matrix

$$\operatorname{adj}(\lambda I - A).$$

Alle Einträge sind Polynome vom Grad $\leq n-1$, es sind Minoren von $\lambda I - A$.
Wir können also $\operatorname{adj}(\lambda I - A)$ schreiben als

$$\operatorname{adj}(\lambda I - A) = \sum_{i=0}^{n-1} B_i\lambda^i \quad \text{mit } B_i \in K^{n,n}.$$

Nach Satz 5.18 folgt

$$\begin{aligned}
(\lambda I - A)\mathrm{adj}\,(\lambda I - A) \;&=\; (\lambda I - A)\left(\sum_{i=0}^{n-1} B_i\lambda^i\right) = P_A(\lambda)I. \\
\Longrightarrow\; &\underbrace{\sum_{i=0}^{n-1} B_i\lambda^{i+1}}_{\sum_{i=1}^{n} B_{i-1}\lambda^i} - \sum_{i=0}^{n-1} AB_i\lambda^i = \begin{bmatrix} P_A(\lambda) & & 0 \\ & \ddots & \\ 0 & & P_A(\lambda) \end{bmatrix}
\end{aligned}$$

mit $P_A(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$.

Der Koeffizientenvergleich (für gleiche Potenzen von λ) ergibt

$$\begin{array}{lr|rcl l rcl}
\lambda^n & : A^n\cdot & B_{n-1} & = & I & \Longrightarrow & A^nB_{n-1} & = & A^n \\
\lambda^{n-1} & : A^{n-1}\cdot & B_{n-2} - AB_{n-1} & = & a_1 I & \Longrightarrow & A^{n-1}B_{n-2} - A^nB_{n-1} & = & a_1A^{n-1} \\
 & \vdots & & & & & & & \\
\lambda^1 & : A\cdot & B_0 - AB_1 & = & a_{n-1}I & \Longrightarrow & AB_0 - A^2B_1 & = & a_{n-1}A \\
\lambda^0 & : & -AB_0 & = & a_nI & \Longrightarrow & -AB_0 & = & a_nI
\end{array}$$

Addieren ergibt

$$A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI = 0.$$

□

Definition 6.12 *Sei K ein Körper und sei $A \in K^{n,n}$. Falls $u \in K^{n,1}$, $u \neq 0$, und $\lambda \in K$ die Gleichung*

$$Au = \lambda u \tag{6.13}$$

erfüllen, so heißt u Eigenvektor von A zum zugehörigen Eigenwert λ.

Da Eigenvektoren immer $\neq 0$ sind, folgt aus $\lambda_1 u = \lambda_2 u$, dass $\lambda_1 = \lambda_2$. Der Eigenwert zu u ist also eindeutig.

Satz 6.14 *Zu $A \in K^{n,n}$ gibt es einen Eigenvektor u mit Eigenwert λ genau dann, wenn λ Nullstelle des charakteristischen Polynoms $P_A(\lambda)$ ist.*

Beweis:

$$\begin{aligned}
P_A(\lambda) = 0 \;&\Longleftrightarrow\; \det(\lambda I - A) = 0 \\
&\Longleftrightarrow\; \text{das homogene Gleichungssystem } (\lambda I - A)u = 0 \\
&\qquad\;\; \text{hat mindestens eine nichttriviale Lösung.} \\
&\Longleftrightarrow\; \exists\, u \in K^{n,1}, u \neq 0 \text{ mit } \lambda u = Au.
\end{aligned}$$

□

Beispiel 6.15 *Eigenwerte und Eigenvektoren von $A \in K^{2,2}$*

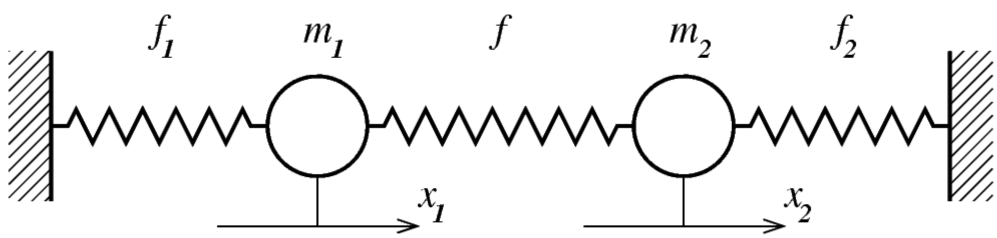
$$\begin{aligned}
\det(\lambda I - A) &= \det \begin{bmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{bmatrix} \\
&= \lambda^2 - \underbrace{(a_{11} + a_{22})}_{\text{spur}\,(A)} \lambda + \underbrace{a_{11}a_{22} - a_{12}a_{21}}_{\det(A)} = 0
\end{aligned}$$

$$\lambda_{1,2} = \frac{\text{spur}\,(A)}{2} \pm \frac{1}{2}\sqrt{\text{spur}\,(A)^2 - 4\det(A)}$$

<u>Eigenvektoren</u>

$$\begin{aligned}
\begin{bmatrix} \lambda_1 - a_{11} & -a_{12} \\ -a_{21} & \lambda_1 - a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= 0 \\
\begin{bmatrix} \lambda_2 - a_{11} & -a_{12} \\ -a_{21} & \lambda_2 - a_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= 0
\end{aligned}$$

Eigenwerte in der Technik



f_1, f, f_2 - Federkonstanten

Bewegungsgleichungen im Gleichgewicht

$$\begin{aligned}
m_1 \frac{d^2 x_1}{dt^2} &= -f_1 x_1 + f(x_2 - x_1) \\
m_2 \frac{d^2 x_2}{dt^2} &= -f_2 x_2 - f(x_2 - x_1)
\end{aligned}$$

Führe Geschwindigkeiten $v_1 = \dfrac{dx_1}{dt}, v_2 = \dfrac{dx_2}{dt}$ ein. Dann gilt

$$\begin{aligned}
\frac{dv_1}{dt} &= \frac{1}{m_1}\left(-f_1 x_1 + f(x_2 - x_1)\right) = -\frac{f_1 + f}{m_1} x_1 + \frac{f}{m_1} x_2, \\
\frac{dv_2}{dt} &= \frac{1}{m_2}\left(-f_2 x_2 - f(x_2 - x_1)\right) = \frac{f}{m_2} x_1 - \frac{f + f_2}{m_2} x_2\,, \\
\frac{dx_1}{dt} &= v_1\,, \\
\frac{dx_2}{dt} &= v_2\,.
\end{aligned}$$

$$y := \begin{bmatrix} x_1 \\ x_2 \\ v_1 \\ v_2 \end{bmatrix} \Longrightarrow \frac{dy}{dt} = Ay = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{f_1+f}{m_1} & \frac{f}{m_1} & 0 & 0 \\ \frac{f}{m_2} & -\frac{f+f_2}{m_2} & 0 & 0 \end{bmatrix} y\,.$$

<u>Ansatz:</u> $y = e^{\lambda t} z$ wobei $z \in \mathbb{C}^4$.

$$\begin{aligned} \frac{dy}{dt} &= \lambda e^{\lambda t} z = A e^{\lambda t} z \qquad \Big| \cdot e^{-\lambda t} \\ &\Longrightarrow \lambda z = Az \\ &\Longrightarrow (\lambda I - A) z = 0 \\ &\Longrightarrow \lambda \text{ ist Eigenwert und } z \text{ Eigenvektor.} \end{aligned}$$

Die Eigenwerte λ sind die Eigenfrequenzen des Systems.
Konkret wählen wir $f_1 = f_2 = f = 1$, $m_1 = m_2 = 1$ und erhalten:

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

$$\begin{aligned} \det(\lambda I - A) &= \det \begin{bmatrix} \lambda & 0 & -1 & 0 \\ 0 & \lambda & 0 & -1 \\ 2 & -1 & \lambda & 0 \\ -1 & 2 & 0 & \lambda \end{bmatrix} \leftarrow \\ &= \lambda \det \begin{bmatrix} \lambda & 0 & -1 \\ -1 & \lambda & 0 \\ 2 & 0 & \lambda \end{bmatrix} - \det \begin{bmatrix} 0 & \lambda & -1 \\ 2 & -1 & 0 \\ -1 & 2 & \lambda \end{bmatrix} \\ &= \lambda(\lambda^3 + 2\lambda) - (-4 + 1 - 2\lambda^2) \\ &= \lambda^4 + 4\lambda^2 + 3 = \omega^2 + 4\omega + 3 \qquad \text{mit } \ \omega = \lambda^2 \end{aligned}$$

$$\omega = \frac{-4 \pm \sqrt{16 - 12}}{2} = \frac{-4 \pm \sqrt{4}}{2} = -2 \pm 1\,,$$
$$\omega_1 = -1,\ \omega_2 = -3\,,$$
$$\lambda_{1,2} = \pm i,\ \lambda_{3,4} = \pm\sqrt{3} i\,.$$

Kapitel 7

Vektorräume

Definition 7.1 *Sei K ein Körper. Ein Tripel $(V,+,\cdot)_K$, bestehend aus einer Menge V, einer Abbildung $+$ (Addition) und einer Abbildung $\cdot$ (skalare Multiplikation),*

$$+: \begin{array}{l} V \times V \longrightarrow V \\ (x,y) \mapsto x+y \end{array}, \qquad \cdot: \begin{array}{l} K \times V \longrightarrow V \\ (\lambda,x) \mapsto \lambda \cdot x \end{array},$$

heißt <u>K-Vektorraum</u>*, falls folgende Axiome erfüllt sind.*

(i)	$\forall x,y,z \in V:$	$(x+y)+z$	$=$	$x+(y+z)$	*Ass. + in V*
(ii)	$\forall x,y \in V:$	$x+y$	$=$	$y+x$	*Komm. + in V*
(iii)	$\exists\, 0 \in V : \forall x \in V:$	$0+x =$	x	$= x+0$	*Null in V*
(iv)	$\forall x \in V\ \exists\, y \in V:$	$y+x$	$=$	0	*Add-Inv. in V*
(v)	$\forall \lambda,\mu \in K, x \in V:$	$\lambda\cdot(\mu\cdot x)$	$=$	$(\lambda\cdot\mu)x$	*Ass. der Skalarmult.*
(vi)	$\forall x \in V:$	$1\cdot x$	$=$	x	*Eins der Skalarmult.*
(vii)	$\forall \lambda \in K, x,y \in V:$	$\lambda(x+y)$	$=$	$\lambda x+\lambda y$	*Distr.*
(viii)	$\forall \lambda,\mu \in K, x \in V:$	$(\lambda+\mu)x$	$=$	$\lambda x+\mu x$	*Distr.*

V muss also eine kommutative Gruppe bezüglich $+$ sein.

Beispiel 7.2

a) $K = \mathbb{C}$, $V = \mathbb{C}^n = \{(x_1,\cdots,x_n) \mid x_i \in \mathbb{C}\}$.

b) $K = \mathbb{R}$, $V = \mathcal{C}([-1,1],\mathbb{R}) = \{f : [-1,1] \longrightarrow \mathbb{R} \mid f \text{ stetig }\}$.

c) $K = \mathbb{R}$, $V = \mathbb{C}$.

d) $K = \mathbb{R}$, $V = \{0\}$.

e) $K = \mathbb{F}_2$, $V = \mathbb{F}_2^n$.

f) $K = \mathbb{R}$, $V = \mathbb{R}^{n,m}$.

Korollar 7.3

(a) In einem Vektorraum gibt es genau ein $0 \in V$.

(b) Das additive Inverse ist eindeutig.

(c) $0 \cdot v = 0\,, \quad \forall v \in V.$

(d) $\lambda \cdot 0 = 0\,, \quad \forall \lambda \in K.$

(e) $(-\lambda)v = -\lambda v = \lambda(-v)\,, \quad \forall \lambda \in K,\, v \in V.$

Beweis:

(a) Angenommen, $\exists\, 0, 0'$, so folgt nach Axiomen (ii), (iii)

$$0 = 0 + 0' = 0' + 0 = 0'\,.$$

(b) Wenn $a + x = 0$ und $b + x = 0$ gilt, so folgt

$$\begin{array}{rcll} a & = & a + 0 & \text{(iii)} \\ & = & a + (b + x) & \\ & = & a + (x + b) & \text{(ii)} \\ & = & (a + x) + b & \text{(i)} \\ & = & 0 + b & \\ & = & b & (iii) \end{array}$$

(c)
$$\begin{array}{rcll} 0 \cdot v & = & (0 + 0) \cdot v & \\ & = & 0 \cdot v + 0 \cdot v & \text{(viii)} \\ 0 & = & 0 \cdot v + 0 \cdot v - 0 \cdot v = 0 \cdot v. & \end{array}$$

(d)
$$\lambda \cdot 0 \overset{\text{(iii)}}{=} \lambda(0 + 0) \overset{\text{(vii)}}{=} \lambda \cdot 0 + \lambda \cdot 0 \qquad \Longrightarrow \qquad \lambda \cdot 0 = 0.$$

(e)
$$\begin{array}{l} \lambda v + (-\lambda)v = (\lambda - \lambda)v = 0 \cdot v = 0\,, \\ \lambda v + \lambda(-v) = \lambda(v - v) = \lambda \cdot 0 = 0\,. \end{array}$$

□

Definition 7.4 *Sei* V *ein* K*-Vektorraum und* $U \subset V$. *Dann heißt* U *Untervektorraum von* V *(kurz: Unterraum), falls* U *abgeschlossen bezüglich* $+$ *und* $\cdot$ *ist und selbst ein Vektorraum.*

Beispiel 7.5

a)
$$\begin{array}{rcl} V & = & \mathcal{C}([-1, +1], \mathbb{R})\,, \\ U & = & \mathcal{C}^1([-1, +1], \mathbb{R}) \;=:\; \{f \in \mathcal{C}([-1, 1], \mathbb{R}) \mid f \text{ differenzierbar in } [-1, 1]\}\,. \end{array}$$

b)
$$\begin{array}{rcl} V & = & \mathbb{R}^4\,, \\ U & = & \mathbb{R}^3 \quad = \{(x, y, z, 0) \mid x, y, z \in \mathbb{R}\}\,. \end{array}$$

Definition 7.6 *Seien V_1, V_2 Vektorräume über einem Körper K. Dann bezeichnen wir eine Abbildung*

$$f : V_1 \longrightarrow V_2,$$

die folgende Axiome erfüllt,

$$\begin{array}{lll} \text{(i)} & f(v+w) = f(v) + f(w), & \forall v, w \in V_1, \\ \text{(ii)} & f(\lambda v) = \lambda f(v), & \forall v \in V_1,\ \lambda \in K, \end{array}$$

als Vektorraum-Homomorphismus*,* K-lineare Abbildung *oder auch* lineare Transformation*.*

Ist f auch noch bijektiv, so heißt f Vektorraum-Isomorphismus*.*

Falls $V_1 = V_2$, so heißt ein Homomorphismus Endomorphismus *und ein Isomorphismus* Automorphismus*.*

Zwei Vektorräume V_1, V_2 heißen isomorph*, falls es einen Isomorphismus $f : V_1 \longrightarrow V_2$ gibt,*

$$V_1 \cong V_2\,.$$

Korollar 7.7

(a) *Ist $f : V_1 \longrightarrow V_2$ ein Vektorraumhomomorphismus, so gilt*

$$\begin{array}{lll} f(0) & = 0 & \textit{und} \\ f(-v) & = -f(v), & \forall v \in V_1\,. \end{array}$$

(b) *Ein Vektorraumhomomorphismus $f : V_1 \longrightarrow V_2$ ist ein Isomorphismus genau dann, wenn es einen Homomorphismus $g : V_2 \longrightarrow V_1$ gibt mit*

$$(f \circ g)(v_2) = v_2 \quad \text{und} \quad (g \circ f)(v_1) = v_1, \quad \forall v_1 \in V_1,\ v_2 \in V_2\,.$$

Beweis:

(a)
$$\begin{array}{llll} f(0) & = f(0 \cdot 0) & = 0 \cdot f(0) & = 0 \\ f(-v) & = (-1)f(v) & = -f(v) & \forall v \in V_1. \end{array}$$

(b) Sei f bijektiv und sei $g = f^{-1}$ die zu f inverse Abbildung. Seien $v, w \in V_1$, mit $v = g(\tilde{v}),\ w = g(\tilde{w}),\ \tilde{v}, \tilde{w} \in V_2, \tilde{v} = f(v), \tilde{w} = f(w)$ und $\lambda \in K$. Dann gilt

$$\begin{aligned} g(\tilde{v} + \tilde{w}) &= g(f(v) + f(w)) = g(f(v+w)) \\ &= (g \circ f)(v+w) = v + w = g(\tilde{v}) + g(\tilde{w}), \\ g(\lambda \tilde{v}) &= g(\lambda f(v)) = g(f(\lambda v)) = (g \circ f)(\lambda v) \\ &= \lambda v = \lambda \cdot g(\tilde{v}). \end{aligned}$$

$\Rightarrow g$ ist ein Homomorphismus. Die Umkehrung ist klar.

$\square$

Beispiel 7.8

(a) $$\mathbb{R}^2 \cong \mathbb{R}^{2,1} \cong \mathbb{C}\ ,$$

$$\mathbb{R}^2 = \{(x,y) \mid x,y \in \mathbb{R}\}, \quad \mathbb{R}^{2,1} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x,y \in \mathbb{R} \right\}, \quad \mathbb{C} = \{x + iy \mid x,y \in \mathbb{R}\}.$$

Isomorphismen:

$$f: \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{R}^{2,1} \\ (x,y) \mapsto \begin{bmatrix} x \\ y \end{bmatrix} \end{array} \qquad g: \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{C} \\ (x,y) \mapsto x + iy. \end{array}$$

(b) $V = \mathbb{R}^{n,1}$, $A \in \mathbb{R}^{n,n}$,

$$A: \begin{array}{l} V \longrightarrow V \\ x \mapsto Ax\,, \end{array} \qquad \begin{array}{l} A(x+y) = Ax + Ay, \\ A(\lambda x) = \lambda Ax\,. \end{array}$$

A beschreibt einen Isomorphismus (Automorphismus) genau dann, wenn A^{-1} existiert.

c) $V = \mathbb{R}^{n,1}$, $\quad W = \mathbb{R}^{m,1}$, $\quad B \in \mathbb{R}^{m,n}$,

$$B: \begin{array}{l} V \longrightarrow W \\ x \mapsto Bx \end{array} \quad .$$

B ist ein Isomorphismus genau dann, wenn $m = n$ und B^{-1} existiert.

Korollar 7.9 *Bei einem Isomorphismus ist* 0 *das einzige Element, welches auf* 0 *abgebildet wird.*

Beweis: folgt aus der Bijektivität. □

Beispiel 7.10 $V = \{0\}$.

$$f: \begin{array}{l} V \longrightarrow V \\ v \mapsto v \end{array}$$

ist ein Isomorphismus (Automorphismus).

Sind Vektorräume isomorph, so unterscheiden wir sie im allgemeinen nicht mehr, z. B.

$$\mathbb{R}^n = \mathbb{R}^{n,1}.$$

Kapitel 8

Basen und Dimension von Vektorräumen

Sei V ein Vektorraum über einem Körper K und seien $v_1, \ldots, v_r \in V$ und $\lambda_1, \ldots, \lambda_r \in K$. Man nennt die Summe

$$\sum_{j=1}^{r} \lambda_j v_j \in V \tag{8.1}$$

eine Linearkombination der Vektoren $v_1, \ldots, v_r$.

Definition 8.2 *Sei V ein Vektorraum über K und seien $v_1, \ldots, v_r \in V$. Die Menge*

$$\mathcal{L}(v_1, \ldots, v_r) : \;=\; \{\lambda_1 v_1 + \ldots + \lambda_r v_r \mid \lambda_i \in K\,, i = 1, \ldots, r\} \in V \tag{8.3}$$

heißt die lineare Hülle von $v_1, \ldots, v_r$ oder das Erzeugnis von $v_1, \ldots, v_r$. Spezialfall ist, dass die lineare Hülle der leeren Menge der Nullvektor aus V ist, $\mathcal{L}(\emptyset) := \{0\}$.

Korollar 8.4 *Seien $v_1, \ldots, v_r \in V$, so ist $\mathcal{L}(v_1, \ldots, v_r)$ ein Untervektorraum von V.*

Beweis: Wir müssen zeigen, dass es ein Vektorraum ist. Dazu müssen wir die Axiome überprüfen. Hier nur exemplarisch

$$\begin{aligned}
\sum_{j=1}^{r} \lambda_j v_j + \sum_{j=1}^{r} \mu_j v_j &= \sum_{j=1}^{r} (\lambda_j + \mu_j) v_j &&\in \mathcal{L}(v_1, \ldots, v_r)\,, \\
\lambda \left(\sum_{j=1}^{r} \lambda_j v_j \right) &= \sum_{j=1}^{r} (\lambda \lambda_j) v_j &&\in \mathcal{L}(v_1, \ldots, v_r)\,.
\end{aligned}$$

□

Definition 8.5 *Sei V ein Vektorraum über K. Eine Menge von r Vektoren $v_1, \ldots, v_r \in V$ heißt linear unabhängig, wenn gilt*

$$\sum_{j=1}^{r} \lambda_j v_j = 0 \quad \Rightarrow \quad \lambda_j = 0, \;\; j = 1, \ldots, r.$$

$\emptyset$ ist linear unabhängig.

Korollar 8.6 *Vektoren $v_1, \ldots, v_r \in V$ sind genau dann linear unabhängig, wenn keines der v_i, $i = 1, \ldots, r$, als Linearkombination der übrigen geschrieben werden kann.*

Beweis: Seien $v_1, \ldots, v_r$ linear unabhängig. Angenommen, es gäbe ein j mit

$$v_j = \sum_{\substack{i=1\\ i\neq j}}^{r} \lambda_i v_i.$$

Dann gilt mit $\lambda_j = (-1)$

$$\sum_{i=1}^{r} \lambda_i v_i = 0.$$

Dies ist ein Widerspruch.

Für die Umkehrung sei keines der v_i Linearkombination der anderen.

Angenommen, $v_1, \ldots, v_r$ wären linear abhängig. Dann gibt es $\lambda_1, \ldots, \lambda_r \in K$ und mindestens ein $\lambda_i \neq 0$, so dass

$$\begin{array}{rcl} 0 = \sum\limits_{j=1}^{r} \lambda_j v_j & = & \sum\limits_{\substack{j=1\\ j\neq i}}^{r} \lambda_j v_j + \lambda_i v_i. \\ \lambda_i \neq 0 \quad \Rightarrow \quad v_i & = & \sum\limits_{\substack{j=1\\ j\neq i}}^{r} \left(\frac{-\lambda_j}{\lambda_i}\right) v_j. \end{array}$$

Dies ist ein Widerspruch. □

Definition 8.7 *Sei V ein Vektorraum über einem Körper K. Eine Menge $\{v_1, \ldots, v_n\} \subset V$ von Vektoren heißt Basis von V, wenn*

a) $v_1, \ldots, v_n$ *linear unabhängig sind,*

b) $\mathcal{L}(v_1, \ldots, v_n) = V$.

Lemma 8.8 *Sei $\{v_1, \ldots, v_n\}$ eine Basis von V. Dann gibt es zu jedem $v \in V$ genau ein Element*

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \in K^{n,1} \cong K^n, \qquad \textit{so dass}$$

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n \,. \tag{8.9}$$

Beweis: Da $V = \mathcal{L}(v_1, \ldots, v_n)$, so gibt es zu jedem $v \in V$ Skalare $\lambda_1, \ldots, \lambda_n \in K$ oder

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \in K^{n,1}, \quad \text{so dass} \quad v = \sum_{j=1}^{n} \lambda_j v_j.$$

Es bleibt also nur zu zeigen, dass es genau einen solchen Vektor gibt. Sei $\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} \in K^{n,1}$ ein weiteres Element, so dass

$$v = \sum_{j=1}^{n} \mu_j v_j \, .$$

Dann folgt

$$\begin{aligned} 0 \; = v - v \; &= \sum_{j=1}^{n} \lambda_j v_j - \sum_{j=1}^{n} \mu_j v_j \\ &= \sum_{j=1}^{n} (\lambda_j - \mu_j) v_j \; , \end{aligned}$$

und wegen der linearen Unabhängigkeit der v_j folgt

$$\lambda_j - \mu_j = 0 \, , \quad \forall j = 1, \ldots, n.$$

$\square$

Die Darstellung von v bezüglich der Basis $\{v_1, \ldots, v_n\}$ mit den Skalaren $\lambda_1, \ldots, \lambda_n$ ist somit eindeutig. Man kann also jeden Vektor v bezüglich $v_1, \ldots, v_n$ darstellen, die Größen $\lambda_1, \ldots, \lambda_n \in K$, die zu v gehören, heißen *Koordinaten* von v.

Der Vektor $\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \in K^n$ heißt *Koordinatenvektor von v*, oder auch *Basisdarstellung von v*.

Die wichtigste Basis in K^n ist die sogenannte *kanonische Basis aus den Einheitsvektoren*

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow i, \\ \\ \\ \\ \end{matrix} \qquad i = 1, \ldots, n. \tag{8.10}$$

Dies sind gerade die Spalten der Einheitsmatrix in $K^{n,n}$.

Beispiel 8.11

Sei $V = \mathbb{R}^n, K = \mathbb{R},$ Sei $v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ n \end{bmatrix} \in V$.

So hat v eine Basisdarstellung in der kanonischen Basis.

$$v = 1e_1 + 2e_2 + \ldots + ne_n = \sum_{j=1}^{n} je_j = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ n \end{bmatrix}.$$

Der Koordinatenvektor ist dann v selbst.

Wähle eine andere Basis,

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad v_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Zeige, dass es eine Basis ist.

$$\sum_{j=1}^{n} \lambda_j v_j = 0 \Rightarrow \begin{bmatrix} \lambda_1 + \lambda_2 \\ \lambda_2 + \lambda_3 \\ \vdots \\ \lambda_{n-1} + \lambda_n \\ \lambda_n \end{bmatrix} = 0, \quad \text{bzw.} \quad \underbrace{\begin{bmatrix} 1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 1 \end{bmatrix}}_{A} \begin{bmatrix} \lambda_1 \\ \vdots \\ \vdots \\ \lambda_n \end{bmatrix} = 0\,.$$

Es gilt $\det A = 1$ und damit ist $\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = 0.$

Also ist $v_1, \ldots, v_n$ Basis. Außerdem hat damit für alle $v \in \mathbb{R}^n$ das Gleichungssystem

$$\begin{bmatrix} 1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ \vdots \\ v_n \end{bmatrix}$$

eine Lösung und damit ist $\mathcal{L}(v_1, \ldots, v_n) = \mathbb{R}^n$.

Wir suchen nun einen Vektor $\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$, so dass $\sum_{j=1}^{n} \lambda_j v_j = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix}$ oder, anders ausgedrückt,

$$A \cdot \lambda := \begin{bmatrix} 1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix}.$$

Rückwärtseinsetzen ergibt

$$\left.\begin{array}{lclcl} \lambda_n & = & n & & \\ \lambda_{n-1} & = & (n-1)-n & = & -1 \\ \lambda_{n-2} & = & (n-2)-\lambda_{n-1} & = & n-1 \\ \lambda_{n-3} & = & (n-3)-\lambda_{n-2} & = & -2 \\ \lambda_{n-4} & = & (n-4)-\lambda_{n-3} & = & n-2 \\ \vdots & & & & \\ \lambda_1 & = & 1-\lambda_2 & = & \dots \end{array}\right\} \qquad \begin{array}{lcl} \lambda_{n-2k} & = & n-k, \\ \lambda_{n-(2k-1)} & = & -k \end{array}$$

Korollar 8.12

(a) *Wir können die Basisdarstellung über die Lösung eines Gleichungssystems erhalten.*

(b) *Alle Basen sind gleichberechtigt. Oft wird eine basisfreie Form bevorzugt; zum konkreten Rechnen braucht man eine konkrete Basis.*

Bemerkung 8.13 *In vielen Lehrbüchern werden*

$$K^n = \{(\lambda_1, \dots, \lambda_n) \mid \lambda_i \in K,\, i = 1, \dots, n\} \qquad \textit{und}$$

$$K^{n,1} = \left\{ \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \mid \lambda_i \in K,\, i = 1, \dots, n \right\}$$

immer unterschieden. Diese beiden Mengen sind, wie gezeigt, isomorph. In Zukunft werden wir sie nicht mehr unterscheiden, also immer K^n für $K^{n,1}$ verwenden.

Satz 8.14 ***<u>Basisergänzungssatz</u>***
Sei V ein Vektorraum über dem Körper K und seien $v_1, \dots, v_r \in V, w_1, \dots, w_s \in V$. Wenn $v_1, \dots, v_r$ linear unabhängig sind und $\mathcal{L}(v_1, \dots, v_r, w_1, \dots, w_s) = V$, dann kann man $v_1, \dots, v_r$ durch eventuelle Hinzunahme von Vektoren aus $\{w_1, \dots, w_s\}$ zu einer Basis ergänzen.

Beweis: Induktion nach s. ($s = 0, r = 0$ waren zugelassen, da $\mathcal{L}(\emptyset) = \{0\}$.)

I.A.: Falls $s = 0$ ist, so gibt es nichts zu zeigen, da dann schon $v_1, \dots, v_r$ eine Basis bilden.

I.V.: Sei die Behauptung richtig für $s = k$.

I.S.: Sei $s = k + 1$. Falls schon $\{v_1, \dots, v_r\}$ eine Basis ist, so ist nichts zu zeigen. Sei also $\mathcal{L}(v_1, \dots, v_r) \neq V$. Dann gibt es mindestens ein $j \in \{1, \dots, s\}$, so dass

$$w_j \notin \mathcal{L}(v_1, \dots, v_r),$$

denn sonst wäre $\mathcal{L}(v_1, \dots, v_r) = \mathcal{L}(v_1, \dots, v_r, w_1, \dots, w_s) = V$.
Dann sind aber $v_1, \dots, v_r, w_j$ linear unabhängig, denn aus

$$\sum_{i=1}^{r} \lambda_i v_i + \lambda w_j = 0$$

folgt $\lambda = 0$, (sonst wäre $w_j \in \mathcal{L}(v_1, \dots, v_r)$) und damit $\lambda_1, \dots, \lambda_r = 0$, da $v_1, \dots, v_r$ linear unabhängig sind. Nach Induktionsvoraussetzung kann man nun $v_1, \dots, v_r, w_j$ durch geeignete Hinzunahme von den $s - 1 = k$ Vektoren w_l, $l = 1, \dots, s$, $l \neq j$, zu einer Basis ergänzen.

□

Satz 8.15 ***Steinitz'scher Austauschsatz***
Sind $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_m\}$ Basen eines Vektorraumes V über dem Körper K, so gibt es zu jedem v_i ein w_j, so dass aus $v_1, \dots, v_n$ wieder eine Basis entsteht, wenn man v_i durch w_j ersetzt.

Beweis: Sei $i \in \{1, \dots, n\}$ fest gewählt, dann gibt es ein $j \in \{1, \dots, m\}$, so dass

$$w_j \notin \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)\,,$$

anderenfalls wäre $V = \mathcal{L}(w_1, \dots, w_m)$ Teilmenge von

$$\mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n),$$

und damit echte Teilmenge von $\mathcal{L}(v_1, \dots, v_n) = V$, welches aber nicht sein kann, da die v_i linear unabhängig sind.
Für dieses j sind $v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_n$ linear unabhängig, denn aus

$$\sum_{\substack{l=1 \\ l \neq i}}^{n} \lambda_l v_l + \mu w_j = 0$$

folgt erst einmal $\mu = 0$, da $w_j \notin \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$, und dann natürlich auch $\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_n = 0$. Wir fügen nun v_i wieder hinzu und erhalten

$$\mathcal{L}(v_1, \dots, v_n, w_j) = V.$$

Dann ist nach dem Basisergänzungssatz entweder $\{v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_n\}$ eine Basis oder es wird durch Hinzufügen von v_i zu einer Basis. Das letztere ist nicht möglich, denn wir können w_j aus $v_1, \dots, v_n$ linear konstruieren. Dann wären aber $v_1, \dots, v_n, w_j$ linear abhängig. Also bilden

$$v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n, w_j$$

eine Basis. □

Korollar 8.16 *Jeder aus endlich vielen Vektoren erzeugte Vektorraum besitzt eine Basis und je zwei Basen desselben Vektorraumes haben gleich viele Elemente.*

Beweis: Seien $v_1, \dots, v_n \in V$, so dass $\mathcal{L}(v_1, \dots, v_n) = V$. Sei $\{v_1, \dots, v_r\}$ eine linear unabhängige Teilmenge von $\{v_1, \dots, v_n\}$, so folgt aus dem Basisergänzungssatz, dass es eine Basis gibt.

Angenommen, es gibt zwei Basen $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_m\}$ von V und $n \neq m$. Ohne Beschränkung der Allgemeinheit sei $n > m$. Dann können wir mit Satz 8.15 nacheinander alle Vektoren von $\{v_1, \dots, v_n\}$ durch Vektoren aus $\{w_1, \dots, w_m\}$ austauschen. Dann kommt aber wegen $n > m$ irgendwann ein Vektor doppelt vor und das kann wegen der linearen Unabhängigkeit der Elemente einer Basis nicht sein. Dies ist ein Widerspruch zur Annahme. □

Definition 8.17 *Die Anzahl der Basiselemente eines Vektorraumes V über einem Körper heißt* Dimension von V, *kurz* $\dim V$.

Beispiel 8.18 Es gilt $\dim K^n = n$, denn $e_1, \dots, e_n$ ist eine Basis.

Korollar 8.19 *Seien $v_1, \dots, v_r \in V$ und $r > \dim V$. Dann sind $v_1, \dots, v_r$ linear abhängig.*

Beweis: Sei $\{w_1, \dots, w_n\}$ eine Basis von V, dann gilt

$$\mathcal{L}(w_1, \dots, w_n) = \mathcal{L}(w_1, \dots, w_n, v_1, \dots, v_r) = V.$$

Wären $v_1, \dots, v_r$ linear unabhängig, so könnten wir nach dem Basisergänzungssatz $v_1, \dots, v_r$ zu einer Basis von V ergänzen, die mindestens r Elemente hat. Dies ist ein Widerspruch. □

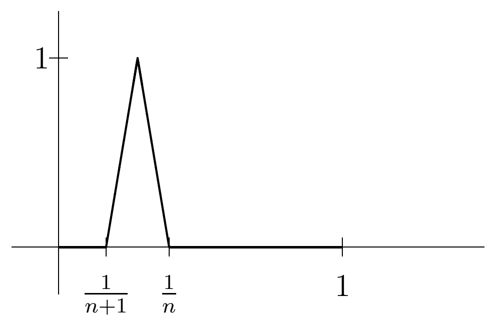
Definition 8.20 *Sei V ein Vektorraum über einem Körper K. Falls es keine endliche Basis $\{v_1, \dots, v_n\}$ von V gibt, so heißt V* unendlichdimensionaler Vektorraum, $\dim V = \infty$.

Beispiel 8.21 Sei $V = \mathcal{C}([0,1], \mathbb{R})$ der Vektorraum der stetigen reellwertigen Funktionen auf $[0,1]$, so ist

$$\dim V = \infty.$$

Sei für $n = 1, 2, \dots$

$$f_n(x) = \begin{cases} 0, & x < \dfrac{1}{n+1}, \\ 0, & \dfrac{1}{n} < x, \\ 2n(n+1)x - 2n, & \dfrac{1}{n+1} \leq x \leq \dfrac{1}{2}\left(\dfrac{1}{n} + \dfrac{1}{n+1}\right), \\ -2n(n+1)x + 2n + 2, & \dfrac{1}{2}\left(\dfrac{1}{n} + \dfrac{1}{n+1}\right) < x \leq \dfrac{1}{n}. \end{cases}$$



Jede Linearkombination $\sum\limits_{j=1}^{k} \lambda_j f_j$ hat an der Stelle $\frac{1}{2}(\frac{1}{j} + \frac{1}{j+1})$ den Wert λ_j, also kann 0 nur vorkommen, wenn alle $\lambda_j = 0$ sind. Somit sind $f_1, \dots, f_k$ für alle k linear unabhängig. Aus Korollar 8.19 folgt, dass es keine endliche Basis geben kann.

Korollar 8.22 *Sei V ein Vektorraum über einem Körper K und $\dim V = n$ endlich.*

(a) Falls $U \subset V$ ein Untervektorraum ist, so ist $\dim U \leq n$.

(b) Sei U ein Untervektorraum von V und $\dim U < \dim V \Rightarrow U \neq V$.

Beweis:

(a) Sind $v_1, \dots, v_r \in U \subset V$ linear unabhängig, dann ist $r \leq \dim V$. Also gibt es eine größte linear unabhängige Teilmenge $\{v_1, \dots, v_r\} \subset U$. Dann gilt $\mathcal{L}(v_1, \dots, v_r) = U$, denn für jedes $u \in U$ sind $v_1, \dots, v_r, u$ linear abhängig, also ist

$$u = \sum_{j=1}^{r} \lambda_j v_j \in \mathcal{L}(v_1, \dots, v_r).$$

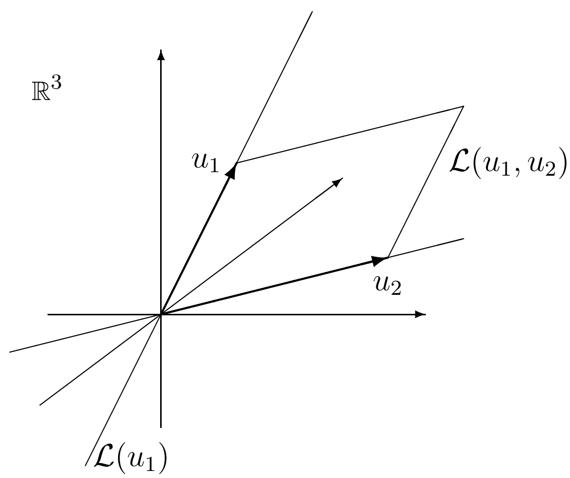
Dann ist also $v_1, \dots, v_r$ eine Basis von U, und U ist damit endlichdimensional.

(b) Eine Basis von U ist Basis von V genau dann, wenn $U = V$. Nach dem Basisergänzungssatz können wir eine Basis $\{u_1, \dots, u_r\}$ von U zu einer Basis von V ergänzen. Falls $\dim U < \dim V$, so kommt mindestens ein Vektor dazu und damit $U \neq V$. □

Beispiel 8.23

$$V = \mathbb{R}^3 = \mathcal{L}\,(e_1, e_2, e_3)\,,$$
$$\dim V = 3\,.$$

Nimm zwei beliebige linear unabhängige Vektoren $u_1, u_2 \in V$ und bilde $\mathcal{L}(u_1, u_2)$ dann bildet $\mathcal{L}(u_1, u_2)$ eine Fläche im Raum und $\mathcal{L}(u_1)$ eine Gerade.



Definition 8.24 *Sind U_1, U_2 Untervektorräume eines Vektorraums V über einem Körper K, so heißt*

$$U_1 + U_2 = \{x + y \mid x \in U_1,\ y \in U_2\} \subset V$$

die <u>*Summe*</u> *von U_1 und U_2. Falls $U_1 \cap U_2 = \{0\}$, so heißt es die* <u>*direkte Summe*</u> *von U_1 und U_2 und wir schreiben $U_1 \dot{+} U_2$.*

Korollar 8.25 *Sei V ein Vektorraum über einem Körper K und seien $U_1, U_2 \subset V$ Untervektorräume. Dann gilt*

(a) $U_1 + U_2$ *ist ein Untervektorraum,*

(b) $U_1 + U_1 = U_1$,

(c) $U_1 + \{0\} = U_1$,

(d) $U_1 \subset U_1 + U_2$,

(e) $U_1 + U_2 = U_1 \;\Leftrightarrow\; U_2 \subset U_1$.

Beweis: (Übung) □

Satz 8.26 <u>***Dimensionsformel***</u>
Sind U_1, U_2 endlichdimensionale Unterräume des Vektorraumes V über dem Körper K, so gilt

$$\dim(U_1 \cap U_2) + \dim(U_1 + U_2) = \dim U_1 + \dim U_2\,.$$

Beweis: Wir verwenden den Basisergänzungssatz. Sei $v_1, \ldots, v_r$ eine Basis von $U_1 \cap U_2$. Ergänze sie zu einer Basis $v_1, \ldots, v_r, w_1, \ldots, w_s$ von U_1 und einer Basis $v_1, \ldots, v_r, x_1, \ldots, x_t$ von U_2.

Wir zeigen: Dann bilden $v_1, \ldots, v_r, w_1, \ldots, w_s, x_1, \ldots, x_t$ eine Basis von $U_1 + U_2$.

Da $\mathcal{L}(v_1, \ldots, v_r, w_1, \ldots, w_s, x_1, \ldots, x_t) = U_1 + U_2$, müssen wir nur zeigen, dass die Vektoren linear unabhängig sind. Sei also

$$\lambda_1 v_1 + \ldots + \lambda_r v_r + \mu_1 w_1 + \ldots + \mu_s w_s + \varphi_1 x_1 + \ldots \varphi_t x_t = 0.$$

Dann folgt, dass

$$\sum_{i=1}^{t} \varphi_i x_i \in U_1 \cap U_2,$$

denn aus U_2 ist es per Konstruktion, und da

$$\sum_{i=1}^{t} \varphi_i x_i = -\left(\sum_{i=1}^{r} \lambda_i v_i\right) - \left(\sum_{i=1}^{s} \mu_i w_i\right),$$

ist es auch aus U_1. Also folgt, dass es $\alpha_1, \ldots, \alpha_r$ gibt, so dass

$$\sum_{i=1}^{t} \varphi_i x_i = \sum_{i=1}^{r} \alpha_i v_i,$$

da $v_1, \ldots, v_r$ Basis von $U_1 \cap U_2$.

Aus der linearen Unabhängigkeit von $v_1, \ldots, v_r, x_1, \ldots, x_t$ folgt dann $\varphi_i = 0,\ i = 1, \ldots, t$, und $\alpha_i = 0,\ i = 1, \ldots, r$, und dann $\lambda_i = 0,\ i = 1, \ldots, r$ und $\mu_i = 0,\ i = 1, \ldots, s$.

Dann folgt $\dim U_1 \cap U_2 = r$, $\dim U_1 = r + s$, $\dim U_2 = r + t$ und $\dim(U_1 + U_2) = r + s + t$.

□

Bemerkung: Nach Lemma 8.8 ist jeder Vektor v eines endlichdimensionalen Vektorraumes V durch seinen Koordinatenvektor aus $K^{n,1}$ bezüglich einer gegebenen Basis eindeutig bestimmt. Da das ebenso für die Basisvektoren v_j einer beliebigen anderen Basis gilt, kann man jede Basis $\{v_1, \ldots, v_n\}$ auch als Matrix $B \in K^{n,n}$ ihrer Koordinatenvektoren bezüglich der zuerst gewählten Basis schreiben.

Basiswechsel

Manchmal ist es notwendig, zwischen Basen zu wechseln. Seien $\{v_1, \ldots, v_n\}$ und $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$ Basen des K-Vektorraumes V. Sei

$$v_i = \sum_{j=1}^{n} p_{ji}\tilde{v}_j = \begin{bmatrix} \tilde{v}_1, \ldots, \tilde{v}_n \end{bmatrix} \begin{bmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{bmatrix}, \qquad \text{und}$$

$$\tilde{v}_i = \sum_{j=1}^{n} q_{ji}v_j = \begin{bmatrix} v_1, \ldots, v_n \end{bmatrix} \begin{bmatrix} q_{1i} \\ \vdots \\ q_{ni} \end{bmatrix}, \qquad i = 1, \ldots, n,$$

bzw.

$$v_i = \tilde{B} \cdot p_i, \quad \tilde{v}_i = B \cdot q_i \,.$$

Die Elemente $p_{ij}, q_{ij},\ i, j = 1, \ldots, n,$ sind eindeutig bestimmt, da $\{v_1, \ldots, v_n\}$, $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$ Basen sind (Lemma 8.8).

Setze $P = [p_{ij}]$, $Q = [q_{ij}]$. Dann sind P und Q invers zueinander, d. h.

$$PQ \;=\; QP \;=\; I\,,$$

denn es gilt

$$\begin{aligned} v_i &= \sum_{j=1}^{n} p_{ji}\tilde{v}_j \;=\; \sum_{j=1}^{n} p_{ji} \sum_{k=1}^{n} q_{kj}v_k \\ &= \sum_{j,k=1}^{n} (q_{kj}p_{ji})v_k \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} q_{kj}p_{ji} \right) v_k. \end{aligned}$$

Da $\{v_1, \ldots, v_n\}$ eine Basis ist, muss gelten

$$\sum_{j=1}^{n} q_{kj}p_{ji} = \delta_{ki} \quad \Longleftrightarrow \quad QP = I.$$

Analog folgt aus

$$\tilde{v}_i = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} p_{kj} q_{ji} \right) \tilde{v}_k, \quad \text{dass} \quad PQ = I \,.$$

Sei $v = \sum\limits_{i=1}^{n} \lambda_i v_i = \sum\limits_{i=1}^{n} \mu_i \tilde{v}_i$, so gilt

$$v = \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{n} p_{ji} \tilde{v}_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \lambda_i p_{ji} \right) \tilde{v}_j = \sum_{j=1}^{n} \mu_j \tilde{v}_j .$$

Also folgt

$$\begin{bmatrix} \tilde{v}_1, \ldots, \tilde{v}_n \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \vdots \\ \mu_n \end{bmatrix} = \begin{bmatrix} \tilde{v}_1, \ldots, \tilde{v}_n \end{bmatrix} \begin{bmatrix} \sum\limits_{i=1}^{n} \lambda_i p_{1i} \\ \vdots \\ \sum\limits_{i=1}^{n} \lambda_i p_{ni} \end{bmatrix},$$

oder

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \text{d.h.} \quad \mu_j = \sum_{i=1}^{n} \lambda_i p_{ji},$$

da $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$ eine Basis ist und deshalb die μ_j eindeutig bestimmt sind.

Das heißt, die Koordinatenumrechnung erfolgt durch Vormultiplikation mit P,

$$\mu = P \cdot \lambda \,.$$

Für die Basisvektoren gilt jedoch

$$\begin{bmatrix} v_1, \ldots, v_n \end{bmatrix} = \begin{bmatrix} \tilde{v}_1, \ldots, \tilde{v}_n \end{bmatrix} \cdot \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix},$$

$$B = \tilde{B} \cdot P \,.$$

P heißt <u>*Basisübergangsmatrix*</u>.

Kapitel 9

Produkte, Längen und Normen

In diesem Abschnitt wollen wir einige wichtige Begriffe betrachten, die bei Vektorräumen von großer Bedeutung sind.

Definition 9.1 *Sei V ein Vektorraum über einem Körper K ($K = \mathbb{R}$ oder $K = \mathbb{C}$). Eine Abbildung*

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow K$$

heißt Skalarprodukt *oder* inneres Produkt*, wenn sie folgende Bedingungen erfüllt:*

(1) Für jedes $w \in V$ ist $\langle \cdot, w \rangle : V \longrightarrow K$ linear, das heißt

(a) $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle \quad \forall v_1, v_2 \in V,$ (Linearität)

(b) $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle \quad \forall v \in V,\ \lambda \in K$ (Homogenität)

(2) $\langle v, w \rangle = \overline{\langle w, v \rangle}\,,$ (Symmetrie)

(3) $\langle v, v \rangle \geq 0 \quad \forall v \in V$ *und* $\langle v, v \rangle = 0$ *genau dann, wenn* $v = 0$. (positive Definitheit)

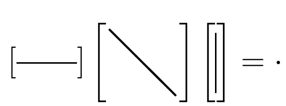
Falls der Körper $K = \mathbb{R}$ ist, so ist in (2) $\langle v, w \rangle = \langle w, v \rangle$.

Beispiel 9.2

(i) $K = \mathbb{R},\ V = \mathbb{R}^n$.
$\langle v, w \rangle = w^\top \cdot v = \sum\limits_{i=1}^{n} v_i w_i,$ (Standardskalarprodukt)

(ii) $K = \mathbb{R},\ V = \mathbb{R}^n\,,\ D = \operatorname{diag}(d_1, \ldots, d_n) = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix},\ d_1, \cdots, d_n > 0.$

$\langle v, w \rangle_D = w^\top D v = \sum\limits_{i=1}^{n} w_i d_i v_i\,.$ (gewichtetes Skalarprodukt)



Überprüfung der Bedingungen:

$$\text{(1a)}\quad \langle v+w,z\rangle_D = \sum_{i=1}^{n}(v_i+w_i)d_iz_i = \sum_{i=1}^{n} v_id_iz_i + \sum_{i=1}^{n} w_id_iz_i = \langle v,z\rangle_D + \langle w,z\rangle_D.$$

$$\text{(1b)}\quad \langle \lambda v,w\rangle_D = \sum_{i=1}^{n} \lambda v_id_iw_i = \lambda\langle v,w\rangle_D\,.$$

$$\text{(2)}\quad \langle v,w\rangle_D = \sum_{i=1}^{n} v_id_iw_i \qquad (K=\mathbb{R}),$$
$$= \sum_{i=1}^{n} w_id_iv_i = \langle w,v\rangle_D.$$

$$\text{(3)}\quad \langle v,v\rangle_D = \sum_{i=1}^{n} v_id_iv_i = \sum_{i=1}^{n} v_i^2d_i \geq 0,\ \text{da } d_i>0,\ v_i^2\geq 0, \qquad (K=\mathbb{R}).$$

Falls $v \neq 0$, so gibt es ein $v_j \neq 0$ und damit ist $\sum\limits_{i=1}^{n} v_i^2 d_i > 0$.

(iii) $K=\mathbb{C}\,,\ V=\mathbb{C}^n$.
$\langle v,w\rangle = \overline{w}^\top \cdot v = \sum_{i=1}^{n} \overline{w}_iv_i$.
$\overline{w}^\top$ wird oft auch als w^* oder w^H (hermitesch) bezeichnet.

Definition 9.3 *Falls V Vektorraum über $K=\mathbb{R}$, so heißt V* euklidischer Vektorraum. *Dann bezeichnet für $v \in V$*

$$\|v\|_2 = \sqrt{\langle v,v\rangle}$$

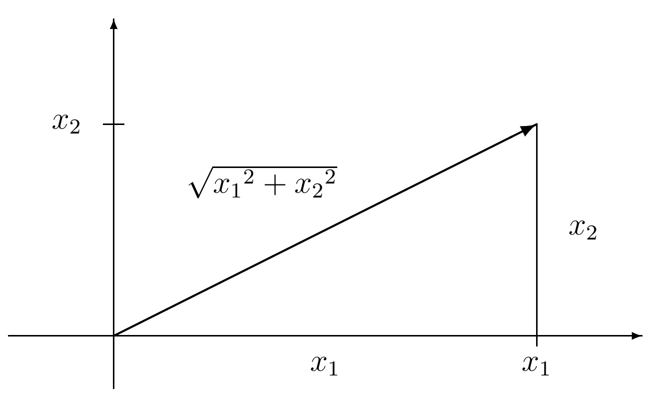
die (euklidische) Länge *oder* euklidische Norm *von v, und*

$$\varphi(v,w) = \arccos\frac{\langle v,w\rangle}{\|v\|_2\|w\|_2}, \qquad \cos\varphi(v,w)\|v\|_2\,\|w\|_2 = \langle v,w\rangle$$

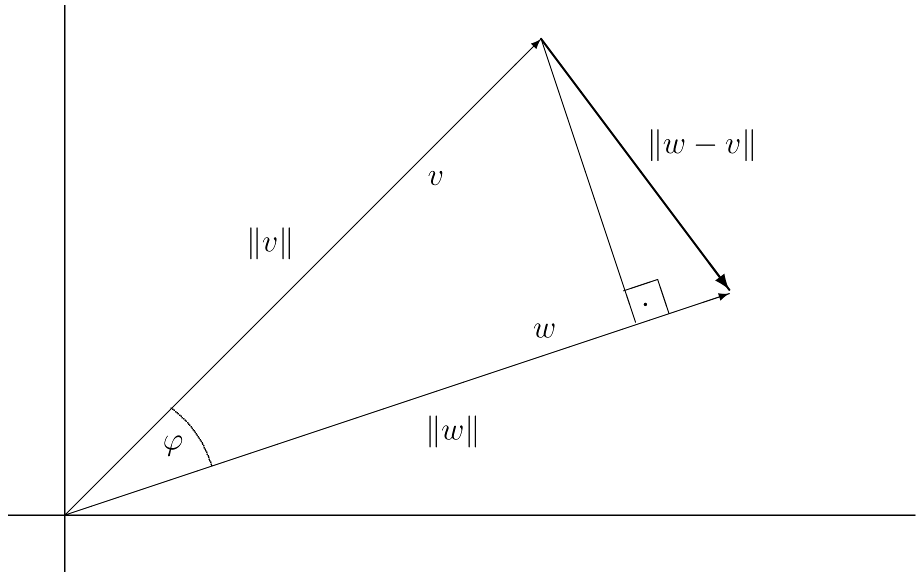
für $v,w \in V\backslash\{0\}$ den Winkel *zwischen v und w.*

Beachte: Es gilt stets $\left|\dfrac{\langle v,w\rangle}{\|v\|_2\|w\|_2}\right| \leq 1$. Der Beweis folgt später (Satz 9.7).

Beispiel 9.4 $\quad V=\mathbb{R}^2,\ x=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix},\quad \|x\|_2 = \sqrt{\langle x,x\rangle} = \sqrt{{x_1}^2+{x_2}^2}.$



Pythagoras



$$\|w-v\|_2{}^2 = \langle w-v, w-v\rangle = -2\langle w, v\rangle + \|v\|_2{}^2 + \|w\|_2{}^2$$

<u>Kosinussatz der Trigonometrie</u>

$$\|w-v\|_2{}^2 = \|v\|_2{}^2 + \|w\|_2{}^2 - 2\|v\|_2\,\|w\|_2 \cdot \cos\varphi$$

$$\boxed{\langle w, v\rangle = \|v\|_2 \cdot \|w\|_2 \cdot \cos\varphi}$$

$$\cos\varphi = \frac{\langle w, v\rangle}{\|v\|_2\,\|w\|_2}$$

Definition 9.5 *Ist V ein euklidischer Vektorraum und $x \in V$, so heißt eine Abbildung*

$$\|\cdot\| : V \longrightarrow \mathbb{R},$$

die folgende Eigenschaften erfüllt, eine <u>Norm in V</u>:

$$\begin{array}{rlll} (i) & \|x\| \geq 0, & \forall x \in V, & \\ (ii) & \|x\| = 0 \;\Leftrightarrow\; x = 0, & & \\ (iii) & \|\lambda x\| = |\lambda| \cdot \|x\|\,, & \forall x \in V\,, \lambda \in \mathbb{R}, & \\ (iv) & \|x+y\| \leq \|x\| + \|y\| & \forall x, y \in V, & \text{(Dreiecksungleichung).} \end{array}$$

Beispiel 9.6 $V = \mathbb{R}^n$.

$$\begin{array}{rll} a) & \|v\|_2 = \sqrt{\langle v, v\rangle} & \underline{\text{euklidische Norm}} \\ b) & \|v\|_\infty = \max\limits_{i=1,\cdots,n} |v_i| & \underline{\text{Maximumnorm}} \\ c) & \|v\|_1 = \sum\limits_{i=1}^{n} |v_i| & \underline{\text{Summennorm}} \end{array}$$

Satz 9.7 ***<u>Cauchy-Schwarz'sche Ungleichung</u>***

In jedem euklidischen Vektorraum V gilt

$$|\langle x, y\rangle| \;\leq\; \|x\|_2 \cdot \|y\|_2, \quad \forall x, y \in V. \tag{9.8}$$

Beweis: Für $y = 0$ ist die Behauptung trivial.
Sei $y \neq 0$ und setze $\lambda := \frac{\langle x, y\rangle}{\|y\|_2^{\,2}}$.

Dann gilt

$$\begin{aligned} 0 \;\leq\; \langle x - \lambda y, x - \lambda y\rangle &= \langle x, x\rangle - 2\lambda\langle x, y\rangle + \lambda^2\langle y, y\rangle \\ &= \|x\|_2^{\,2} - 2\frac{\langle x, y\rangle^2}{\|y\|_2^{\,2}} + \frac{\langle x, y\rangle^2}{\|y\|_2^{\,2}} \\ &= \|x\|_2^{\,2} - \frac{\langle x, y\rangle^2}{\|y\|_2^{\,2}}. \end{aligned}$$

Daraus folgt $\langle x, y\rangle^2 \leq \|x\|_2^{\,2}\|y\|_2^{\,2}$ und damit (9.8). □

Eine analoge Ungleichung gilt auch in komplexen Räumen.

Im $\mathbb{R}^3$ gibt es noch weitere wichtige Produkte, welche eine große Bedeutung in der Physik haben (Vektorprodukt, Spatprodukt).

Definition 9.9 Das Vektorprodukt *in $\mathbb{R}^3$ ist definiert durch*

$$u \times v = \begin{bmatrix} u_2v_3 & - & u_3v_2 \\ u_3v_1 & - & u_1v_3 \\ u_1v_2 & - & u_2v_1 \end{bmatrix}, \qquad \forall\, u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix},\; v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3,$$

wobei u_i, v_i die Koordinaten von u, v in einer ONB von $\mathbb{R}^3$ sind.

Lemma 9.10 *Es seien $u, v, w \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$. Dann gilt*

(a) $u \times v = -v \times u$,

(b) $(u + v) \times w = u \times w + v \times w$,

(c) $(\lambda u) \times v = \lambda(u \times v)$,

(d) $\langle u, u \times v\rangle = \langle v, u \times v\rangle = 0$,

(e) $\|u \times v\|_2^{\,2} = \|u\|_2^{\,2} \cdot \|v\|_2^{\,2} - \langle u, v\rangle^2$.

Beweis: Übung. □

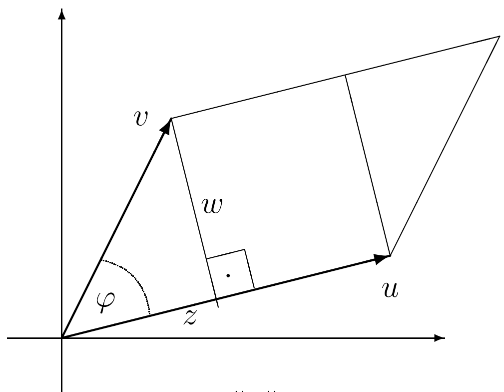
Es folgt, dass das Vektorprodukt senkrecht auf seinen beiden Faktoren steht. Wenn wir (e) noch umformen, erhalten wir

$$\begin{aligned} \|u \times v\|_2^{\,2} &= \|u\|_2^{\,2} \cdot \|v\|_2^{\,2} - \|u\|_2^{\,2} \cdot \|v\|_2^{\,2} \cos^2(\varphi(u, v)) \\ &= \|u\|_2^{\,2} \cdot \|v\|_2^{\,2} \sin^2(\varphi(u, v))\,, \end{aligned}$$

also

$$\|u \times v\|_2 = \|u\|_2 \cdot \|v\|_2 \cdot \sin(\varphi(u,v))\,. \tag{9.11}$$

Die Länge des Vektorproduktes von u und v ist also gleich dem Flächeninhalt des von u, v aufgespannten Parallelogramms.

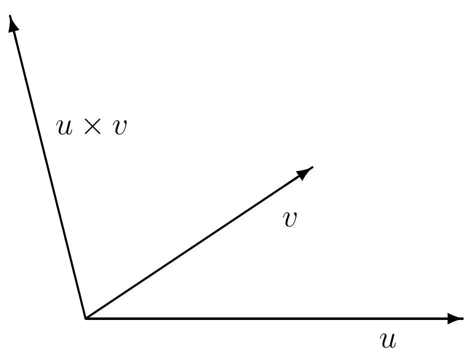


Beachte:
z, w sind Längen, während u, v Vektoren sind.

$$w \;=\; \|v\|_2 \cdot \sin\varphi$$

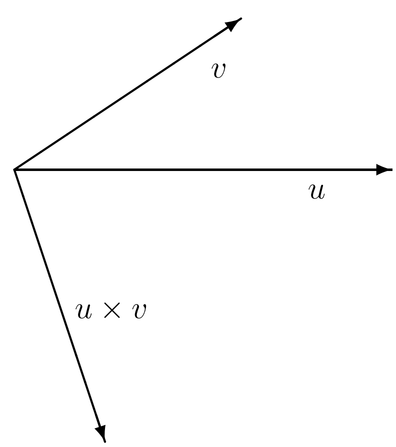
$$\begin{aligned}
\text{Fläche:}\quad 2\left(z \cdot \frac{w}{2}\right) + (\|u\|_2 - z)\cdot w &= zw + \|u\|_2 \cdot w - zw \\
&= \|u\|_2 \cdot w \\
&\quad\ (\text{Grundseite} \cdot \text{Höhe}) \\
&= \|u\|_2 \cdot \|v\|_2 \cdot \sin(\varphi(u,v)).
\end{aligned}$$

Da $u \times v$ senkrecht auf u und v steht, gibt es zwei Möglichkeiten:



Rechte-Hand-Regel

oder



Welche Variante sich ergibt, hängt von der Wahl der Basisvektoren ab. In dem meistens verwendeten rechtsorientierten kartesischen Koordinatensystem gilt die Rechte-Hand-Regel.

Kapitel 10

Orthogonale Vektoren und Matrizen

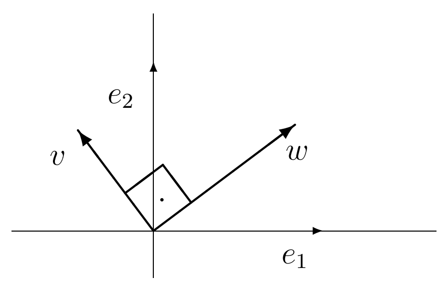
Im letzten Kapitel haben wir Skalarprodukte eingeführt und auch die Winkel zwischen Vektoren über

$$\cos\varphi = \frac{\langle w, v\rangle}{\|v\| \cdot \|w\|}\,,$$

für Vektoren $v, w \in \mathbb{R}^n$.

Der Kosinus ist 0 für $\varphi \in [0, \pi]$ genau dann, wenn der Winkel $\frac{\pi}{2}$ ist, also 90°.

Dann stehen die Vektoren v, w senkrecht aufeinander.



Definition 10.1 *Zwei Vektoren $v, w \in V$ heißen* orthogonal *bezüglich eines Skalarproduktes $\langle\cdot,\cdot\rangle$, falls*

$$\langle v, w\rangle = 0, \quad \textit{d.h., der Winkel } \varphi(v, w) = \frac{\pi}{2}.$$

Eine Basis $\{v_1, \ldots, v_n\}$ von V heißt Orthonormalbasis *(ONB), falls gilt*

$$\langle v_i, v_j\rangle = \delta_{ij}, \quad i, j = 1, \ldots, n.$$

Beispiel 10.2

(a) $V = \mathbb{R}^n, \quad \langle v, w\rangle = v^\top w,$

$$\begin{aligned} v_i &= e_i, & i &= 1, \ldots, n\,, \\ e_i^\top e_j &= \delta_{ij}, & i, j &= 1, \ldots, n\,. \end{aligned}$$

(b) $V = \mathbb{R}^2\,, \quad \langle v, w\rangle = v^\top w,$

$$v_1 = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \; v_2 = \begin{bmatrix} -\sin\alpha \\ \cos\alpha \end{bmatrix}, \quad \alpha \in \left[0, \frac{\pi}{2}\right],$$

$$v_1^\top v_1 = v_2^\top v_2 = \cos^2\alpha + \sin^2\alpha = 1\,,$$

$$v_1^\top v_2 = -\sin\alpha\cos\alpha + \sin\alpha\cos\alpha = 0.$$

Definition 10.3

(a) Eine Matrix $Q = [q_1, \ldots, q_n] \in \mathbb{R}^{n,n}$ heißt orthogonal*, falls die Spalten $q_i,\; i = 1, \ldots, n$, bezüglich des Standard-Skalarproduktes $\langle u, v\rangle = u^\top v$ eine Orthonormalbasis von $\mathbb{R}^n$ bilden.*

(b) Eine Matrix $Q = [q_1, \ldots, q_n] \in \mathbb{C}^{n,n}$ heißt unitär*, falls die Spalten $q_i,\; i = 1, \ldots, n$, bezüglich des Skalarproduktes $\langle u, v\rangle = \bar{v}^\top u = v^* u$ eine Orthonormalbasis von $\mathbb{C}^n$ bilden.*

Satz 10.4

(a) $Q \in \mathbb{R}^{n,n}$ ist orthogonal genau dann, wenn $Q^{-1} = Q^\top$.

(b) $Q \in \mathbb{C}^{n,n}$ ist unitär genau dann, wenn $Q^{-1} = \overline{Q}^\top = Q^*$.

Beweis:

(a)
$$\begin{aligned} Q^\top Q &= \begin{bmatrix} q_1^\top \\ \vdots \\ q_n^\top \end{bmatrix} \cdot [q_1, \cdots, q_n] \\ &= [q_i^\top q_j] = [\langle q_i, q_j\rangle] \\ &= [\delta_{ij}] \quad \Longleftrightarrow \quad Q^\top = Q^{-1}\,. \end{aligned}$$

(b) (*Übung*)

□

Beispiel 10.5

(a) Jede Permutationsmatrix P ist orthogonal, denn $P^\top = P^{-1}$.

(b) Jede Matrix der Form

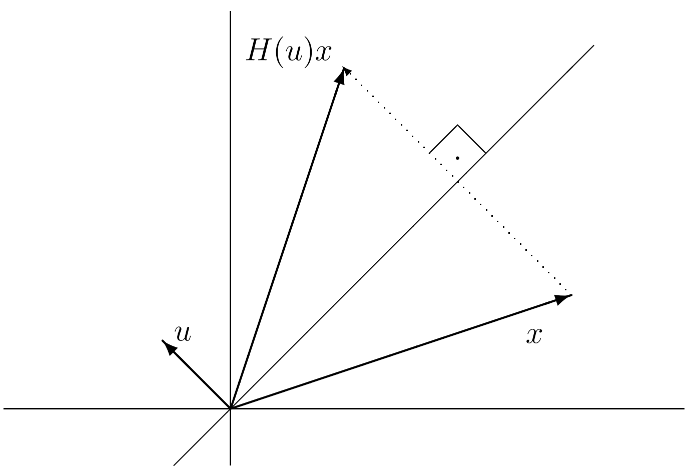
$$H(u) := I - 2\frac{uu^\top}{u^\top u}, \quad u \in \mathbb{R}^n\,,$$

ist orthogonal, denn

$$
\begin{aligned}
H(u)^\top H(u) &= \left(I - \frac{2uu^\top}{u^\top u}\right)^\top \left(I - \frac{2uu^\top}{u^\top u}\right) \\
&= \left(I - \frac{2uu^\top}{u^\top u}\right) \left(I - \frac{2uu^\top}{u^\top u}\right) = H(u)^2 \\
&= I - \frac{4uu^\top}{u^\top u} + 4\frac{u(u^\top u)u^\top}{(u^\top u)(u^\top u)} \\
&= I - \frac{4uu^\top}{u^\top u} + 4\frac{uu^\top}{u^\top u} = I.
\end{aligned}
$$

$I - \dfrac{2uu^\top}{u^\top u}$ heißt <u>*Householder- oder Spiegelungsmatrix*</u>

Die Multiplikation von $H(u)$ mit einem Vektor $x \in \mathbb{R}^n$ bewirkt eine Spiegelung von x an der durch u beschriebenen Hyperebene (Geraden in $\mathbb{R}^2$), $\left\{y \in \mathbb{R}^n \mid u^\top y = 0\right\}$.



$$\mathbb{R}^2, \quad u = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\Longrightarrow \quad I - \frac{2uu^\top}{u^\top u} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Beachte, dass $H(u) \cdot (H(u)x) = (H(u))^2 \cdot x = x$.

(c) Jede Matrix der Form

$$
R_{ij}(\alpha) = \begin{array}{c}
\left[\begin{array}{c|c|c|c|c}
\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} & & & & \\
\hline
 & \cos\alpha & & \sin\alpha & \\
\hline
 & & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} & & \\
\hline
 & -\sin\alpha & & \cos\alpha & \\
\hline
 & & & & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}
\end{array}\right]
\begin{array}{c} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array} . \\
\begin{array}{ccccc} & \uparrow & & \uparrow & \\ & i & & j & \end{array}
\end{array}
$$

ist orthogonal, denn

$$R_{ij}(\alpha)^\top R_{ij}(\alpha) = \left[\begin{array}{c|c|c|c|c}
\begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} & & & & \\ \hline
 & \cos^2\alpha+\sin^2\alpha & & 0 & \\ \hline
 & & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} & & \\ \hline
 & 0 & & \cos^2\alpha+\sin^2\alpha & \\ \hline
 & & & & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix}
\end{array}\right]$$

$R_{ij}(\alpha)$ heißt Rotationsmatrix. Die Multiplikation von $R_{ij}(\alpha)$ mit $x \in \mathbb{R}^n$ beschreibt eine Drehung in der (i,j)-Koordinatenebene.

Definition 10.6 *Ist M ein Untervektorraum eines euklidischen Vektorraumes V, z.B. mit Standard-Skalarprodukt, so heißt*

$$M^\perp := \{v \in V \mid \langle v,u\rangle = 0 \quad \forall u \in M\}$$

das orthogonale Komplement von M.

Korollar 10.7 *$M^\perp$ ist ein Untervektorraum von V.*

Beweis: $M^\perp \neq \emptyset$, da $\langle 0,v\rangle = 0\,, \forall v \in M$. Sind $x,y \in M^\perp$, so folgt aus der Linearität von $\langle\cdot,\cdot\rangle$, dass auch $x+y$ und λx für alle $\lambda \in \mathbb{R}$ orthogonal zu M sind. □

Lemma 10.8 *Ist $\{u_1,\ldots,u_n\}$ eine ONB eines euklidischen oder komplexen Vektorraumes V, so gilt für jedes $v \in V$*

$$v = \sum_{i=1}^n \langle v,u_i\rangle u_i\,.$$

Z.B. ist $\langle\cdot,\cdot\rangle$ das Standardskalarprodukt, falls $K=\mathbb{R}$ und das komplexe Skalarprodukt, falls $K=\mathbb{C}$.

Beweis: Jedes $v \in V$ lässt sich schreiben als $\sum\limits_{i=1}^n \lambda_i u_i$. Dann gilt für alle $j = 1,\ldots,n$

$$\begin{aligned}\langle v,u_j\rangle &= \langle \sum_{i=1}^n \lambda_i u_i, u_j\rangle \\ &= \sum_{i=1}^n \lambda_i \langle u_i,u_j\rangle = \lambda_j.\end{aligned}$$

□

Beispiel 10.9 $V = \mathbb{R}^3,\ \langle x, y\rangle = y^\top x$

$$u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

$$\begin{aligned} v &= \sum_{i=1}^{3} \langle v, u_i\rangle u_i \\ &= \left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}^\top \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}\right) \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + \left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}^\top \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}\right) \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + \\ &\quad \left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}^\top \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \tfrac{3}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + \tfrac{1}{\sqrt{2}} \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \end{aligned}$$

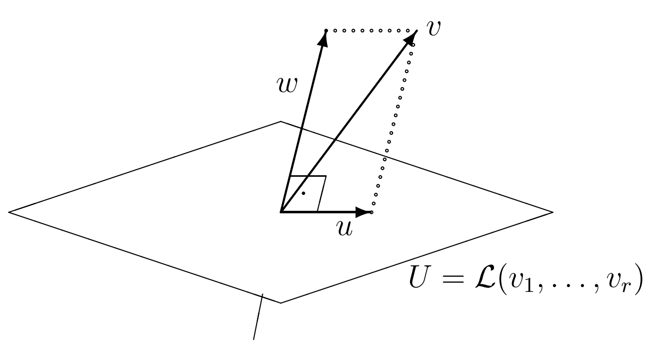
$$\begin{aligned} \mathcal{L}(u_1, u_2)^\perp &= \mathcal{L}(u_3), \\ \mathcal{L}(u_1)^\perp &= \mathcal{L}(u_2, u_3), \\ \mathcal{L}(u_2)^\perp &= \mathcal{L}(u_1, u_3), \qquad \textit{usw.} \end{aligned}$$

Wir haben in dem Beispiel gesehen, dass es sehr leicht ist, das orthogonale Komplement zu bekommen, wenn man eine ONB hat.

Lemma 10.10 *Ist $\{v_1, \ldots, v_r\}$ ein Orthonormalsystem in einem Vektorraum V und $U = \mathcal{L}(v_1, \ldots, v_r)$, so lässt sich jedes $v \in V$ auf eindeutige Weise als $v = u + w$ mit $u \in U$ und $w \in U^\perp$ schreiben, und zwar ist*

$$u = \sum_{i=1}^{r} \langle v, v_i\rangle v_i$$

und folglich $w = v - \sum_{i=1}^{r} \langle v, v_i\rangle v_i$.



Beweis: Man kann v höchstens auf eine Weise als Summe $v = u + w$ mit $u \in U$ und $w \in U^\perp$ schreiben, denn aus

$$v = u + w = \tilde{u} + \tilde{w}$$

mit $u, \tilde{u} \in U$, $w, \tilde{w} \in U^\perp$ folgt

$$\begin{array}{l} (u - \tilde{u}) + (w - \tilde{w}) = 0 \quad \text{und} \\ \langle u - \tilde{u}, w - \tilde{w} \rangle = \langle u, w \rangle - \langle \tilde{u}, w \rangle - \langle u, \tilde{w} \rangle + \langle \tilde{u}, \tilde{w} \rangle = 0\,, \end{array}$$

also $\langle u - \tilde{u}, u - \tilde{u} \rangle = 0$ (mit $(w - \tilde{w}) = -(u - \tilde{u})$) und damit

$$u - \tilde{u} = 0 \quad \text{also} \quad w - \tilde{w} = 0\,.$$

Das gilt für jeden Untervektorraum U von V, selbst wenn V nicht endlichdimensional wäre. Dass überhaupt so ein w, u existiert, folgt schon aus der Darstellung

$$\begin{array}{ll} w & = v - \sum\limits_{i=1}^{r} \langle v, v_i \rangle \, v_i\,, \\ u & = \sum\limits_{i=1}^{r} \langle v, v_i \rangle \, v_i\,, \end{array}$$

aber selbst wenn wir das nicht wüssten, könnten wir ja ansetzen

$$u = c_1 v_1 + \ldots + c_r v_r$$

und

$$w = v - u\,.$$

w ist aus $U^\perp$ genau dann, wenn $\langle w, v_i \rangle = 0,\ i = 1, \ldots, r$, d.h.

$$\begin{array}{lll} 0 & = \langle v - u, v_i \rangle & = \langle v, v_i \rangle - \langle u, v_i \rangle \\ & & = \langle v, v_i \rangle - c_i \quad \Longrightarrow \quad c_i = \langle v, v_i \rangle. \end{array}$$

□

Wie können wir überhaupt ein Orthonormalsystem erhalten?
Das geht mit dem **Gram-Schmidt-Verfahren**.

Satz 10.11 *Seien $v_1, \ldots, v_r$ linear unabhängige Vektoren in einem Vektorraum V. So bilden die durch die Rekursionsformel*

$$\begin{array}{lll} u_1 & = & \dfrac{v_1}{\|v_1\|_2}\,, \\ u_{k+1} & = & \dfrac{v_{k+1} - \sum\limits_{i=1}^{k} \langle v_{k+1}, u_i \rangle u_i}{\left\| v_{k+1} - \sum\limits_{i=1}^{k} \langle v_{k+1}, u_i \rangle u_i \right\|_2}, \qquad k = 1, \ldots, r-1\,, \end{array} \tag{10.12}$$

gegebenen Vektoren $u_1, \ldots, u_r$ ein Orthonormalsystem, bei dem jeweils

$$\mathcal{L}(u_1, \ldots, u_l) = \mathcal{L}(v_1, \ldots, v_l)\,, \quad l = 1, \ldots, r\,.$$

Insbesondere gilt: Falls die Vektoren $v_1, \ldots, v_n$ eine Basis von V bilden, so ist $\{u_1, \ldots, u_n\}$ eine ONB von V.

Beweis: Mittels vollständiger Induktion.

I.A.: Wenn man nur einen Vektor v_1 hat ($r = 1$), so erhält man ein Orthonormalsystem, indem man einfach v_1 normiert.

$$u_1 = \frac{v_1}{\|v_1\|_2}\,.$$

I.V.: Seien $v_1, \ldots, v_k$ linear unabhängig und $u_1, \cdots, u_k$, wie in (10.12) konstruiert, bilden ein ONS mit $\mathcal{L}(u_1, \ldots, u_k) = \mathcal{L}(v_1, \ldots, v_k)$.

I.S.: Betrachte nun $\mathcal{L}(v_1, \ldots, v_{k+1})$. Nach Lemma 10.10 können wir $v_{k+1} \in \mathcal{L}(v_1, \ldots, v_{k+1})$ schreiben als $z_1 + z_2$ mit $z_1 \in \mathcal{L}(v_1, \ldots, v_k) = \mathcal{L}(u_1, \ldots, u_k)$ und $z_2 \in \mathcal{L}(v_1, \ldots, v_k)^\perp$. Dann gilt aber mit der Formel aus Lemma 10.10, dass

$$z_2 = v_{k+1} - z_1 = v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, u_i \rangle u_i\,.$$

Wenn wir z_2 normieren, erhalten wir damit $u_{k+1} = \dfrac{z_2}{\|z_2\|_2}$ und das ergibt dann 10.12.

□

Wir können also aus jeder Basis eine Orthonormalbasis machen, indem wir das Gram-Schmidt-Verfahren anwenden.

Korollar 10.13 *Sei $\{v_1, \ldots, v_s\}$ eine Menge linear unabhängiger Vektoren im Vektorraum $V = \mathbb{R}^n$. Bilde die Matrix $A = [v_1, \ldots, v_s] \in \mathbb{R}^{n,s}$. Sei $Q = [q_1, \ldots, q_n] \in \mathbb{R}^{n,n}$ eine orthogonale Matrix, so dass*

$$Q^\top A = \underbrace{\begin{bmatrix} r_{11} & \cdots & r_{1s} \\ 0 & \ddots & \vdots \\ \vdots & & r_{ss} \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}}_{R}\,, \tag{10.14}$$

so bilden die Spalten $q_1, \ldots, q_s$ ein ONS mit $\mathcal{L}(q_1, \ldots, q_j) = \mathcal{L}(v_1, \ldots, v_j)$, $\forall j = 1, \ldots, s$.

Beweis: Dass die Spalten ein ONS bilden, ist klar. Wir müssen also nur zeigen, dass

$$\mathcal{L}(q_1,\ldots,q_j) = \mathcal{L}(v_1,\ldots,v_j)$$

ist. Da Q orthogonal ist, folgt

$$\begin{aligned}
[v_1,\ldots,v_s] = A &= QR \\
&= [q_1,\ldots,q_n]R \;=\; [q_1,\ldots,q_s]\begin{bmatrix} r_{11} & \ldots & r_{1s} \\ & \ddots & \vdots \\ & & r_{ss}\end{bmatrix}. \\
\Rightarrow \quad v_1 &= q_1 r_{11}, \quad \ldots, \\
v_j &= \sum_{l=1}^{j} q_l r_{lj}, \qquad j=1,\ldots,s,
\end{aligned}$$

und damit

$$\mathcal{L}(v_1,\ldots,v_j) = \mathcal{L}(q_1,\ldots,q_j)\,, \quad j=1,\ldots,s.$$

$\square$

Definition 10.15

(a) Die Zerlegung eines Vektorraums V als $V = U \oplus W$, $W = U^\perp$

$$V = \{u + w \mid u \in U, w \in U^\perp\}$$

heißt die orthogonale Summe *und wird im folgenden als $U \dot{\oplus} W$ bezeichnet.*

(b) Eine Zerlegung einer Matrix $A \in \mathbb{R}^{n,s}$ mit vollem Spaltenrang in $A = Q \cdot R$, mit $Q \in \mathbb{R}^{n,n}$ orthogonal, und $R \in \mathbb{R}^{n,s}$ (rechteckige) obere Dreiecksmatrix, heißt QR-Zerlegung.

Wie erhalten wir so eine QR-Zerlegung? Das Gram-Schmidt-Verfahren liefert diese sofort, denn aus der Formel (10.12) und $A = QR$ erhalten wir

$$\begin{aligned}
v_1 &= u_1 r_{11} && \text{mit } r_{11} = \|v_1\|_2 \\
v_k &= \sum_{j=1}^{k} u_j r_{jk} && \text{mit } r_{jk} = \langle v_k, u_j\rangle\,,\ j=1,\ldots,k, \quad \text{wobei} \\
& && r_{kk} = \langle v_k, u_k\rangle = \left\| v_k - \sum_{i=1}^{k-1}\langle v_k, u_i\rangle u_i \right\|_2,
\end{aligned}$$

d.h., die Koeffizienten von R sind aus (10.12) zu erhalten, und die Matrix Q ergibt sich aus $\{u_1,\ldots,u_s,u_{s+1},\ldots,u_n\}$, wobei $u_1,\ldots,u_s$ aus dem Gram-Schmidt-Verfahren und die noch fehlenden Vektoren $u_{s+1},\ldots,u_n$ durch Ergänzung zu einer Orthonormalbasis bestimmt werden.

Mit Hilfe einer QR-Zerlegung lassen sich sehr leicht Lösungen beliebiger linearer Gleichungssysteme berechnen, die QR-Zerlegung ist so etwas ähnliches wie die Treppennormalform. Sie ist die genaueste Methode, lineare Gleichungssysteme zu lösen, falls A linear unabhängige Spalten hat.

Beispiel 10.16 Sei $s \leq n$, $A = [a_1, \ldots, a_s] \in \mathbb{R}^{n,s}$ und seien $a_1, \ldots, a_s$ linear unabhängig. Betrachte das Gleichungssystem

$$Ax = b \quad \text{mit} \quad x \in \mathbb{R}^{s,1},\, b \in \mathbb{R}^{n,1}\,.$$

Sei $A = Q \cdot R = Q \cdot \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ eine QR-Zerlegung mit $R_1 \in \mathbb{R}^{s,s}$. So folgt aus $Q^\top A \cdot x = Q^\top b$

$$\begin{bmatrix} R_1 \\ 0 \end{bmatrix} x = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix} \text{ mit } \begin{array}{ll} \tilde{b}_1 \in \mathbb{R}^{s,1}, & = \mathbb{R}^s \\ \tilde{b}_2 \in \mathbb{R}^{n-s,1} & = \mathbb{R}^{n-s} \end{array}\,.$$

Damit folgt: $\begin{array}{rcl} R_1 x & = & \tilde{b}_1,\ . \\ 0 & = & \tilde{b}_2. \end{array}$

Das heißt, das System ist lösbar $\quad\Leftrightarrow\quad \tilde{b}_2 = \begin{bmatrix} q_{s+1}^\top \\ \vdots \\ q_n^\top \end{bmatrix} b \;=\; 0\,.$

Wenn das System lösbar ist, so folgt

$$x = R_1^{-1}\tilde{b}_1\,.$$

Lemma 10.17 *Sei $x \in \mathbb{R}^n$ und $Q \in \mathbb{R}^{n,n}$ orthogonal, dann gilt*

$$\|x\|_2 = \|Qx\|_2\,,$$

das heißt, die Multiplikation mit orthogonalen Matrizen erhält die Länge.

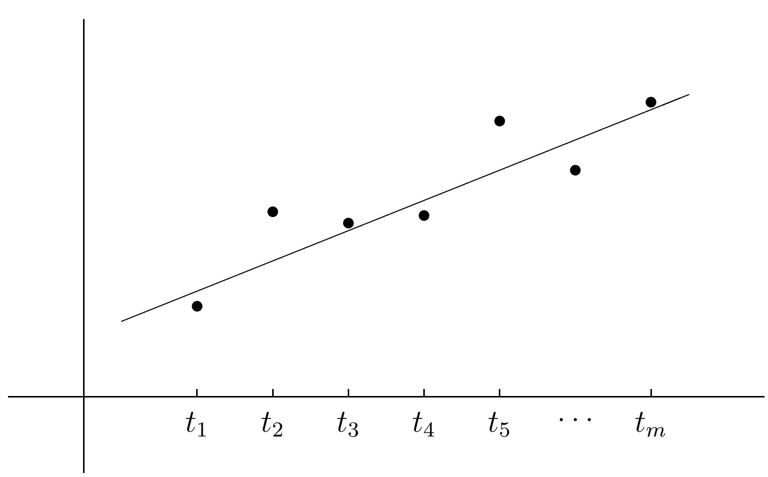
Beweis:

$$\|x\|_2 = \sqrt{\langle x, x\rangle} = \sqrt{x^\top x} = \sqrt{x^\top Q^\top Q x} = \|Qx\|_2\,.$$

□

Beispiel 10.18 Lineare Regression.

Gegeben sind Punkte (t_i, y_i), $i = 1, \ldots, m$. Suche eine Gerade, die diese Punkte am besten repräsentiert.



Gerade: $\quad y = at + b.$

Es gibt natürlich für $m > 2$ im allgemeinen keine Gerade, die genau durch die Punkte geht. Wenn es das gäbe, so hätten wir

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} t_1 \\ \vdots \\ t_m \end{bmatrix} a + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} b$$

oder

$$\begin{array}{c} \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix} \\ \begin{matrix} v_1 & v_2 \end{matrix} \end{array} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \qquad \text{oder} \qquad [v_1, v_2] \begin{bmatrix} a \\ b \end{bmatrix} = y\,.$$

Gram-Schmidt'sches Orthonormalisierungsverfahren:

$$\begin{aligned}
u_1 &= \frac{v_1}{\|v_1\|_2}, \\
r_{11} &= \|v_1\|_2 = \sqrt{\sum_{j=1}^m t_j^2}\,, \\
r_{12} &= \langle v_2, u_1 \rangle \;=\; \frac{1}{\|v_1\|_2} \langle v_2, v_1 \rangle \;=\; \frac{\sum\limits_{j=1}^m t_j}{\sqrt{\sum\limits_{j=1}^m t_j^2}}\,, \\
u_2 &= \frac{v_2 - \langle v_2, u_1 \rangle u_1}{\|v_2 - \langle v_2, u_1 \rangle u_1\|_2} \;=\; \frac{v_2 - r_{12} u_1}{\|v_2 - r_{12} u_1\|_2}\,, \\
v_2 - r_{12} u_1 &= v_2 - \frac{v_1}{\|v_1\|_2} \cdot \frac{1}{\|v_1\|_2} \langle v_2, v_1 \rangle \;=\; \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} t_1 \\ \vdots \\ t_m \end{bmatrix} \cdot \frac{1}{\sum\limits_{j=1}^m t_j^2} \cdot \sum_{j=1}^m t_j \\
&= \begin{bmatrix} 1 \; - \; t_1 \dfrac{\sum\limits_{j=1}^m t_j}{\sum\limits_{j=1}^m t_j^2} \\ \vdots \\ 1 \; - \; t_m \dfrac{\sum\limits_{j=1}^m t_j}{\sum\limits_{j=1}^m t_j^2} \end{bmatrix}, \\
r_{22} &= \langle v_2, u_2 \rangle \;=\; \left\langle v_2, \frac{1}{\|v_2 - r_{12} u_1\|_2} (v_2 - r_{12} u_1) \right\rangle .
\end{aligned}$$

$[u_1, u_2]$ sind die ersten beiden Spalten der orthogonalen Matrix Q der QR-Zerlegung von $A = [v_1, v_2] = Q \cdot R$, d.h.,

$$\begin{bmatrix} u_1, u_2 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = y$$

ist zu lösen, um a, b zu bestimmen. In Analogie zu Beispiel 10.16 erhalten wir

$$\begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} u_1^\top \\ u_2^\top \end{bmatrix} \cdot y = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix}.$$

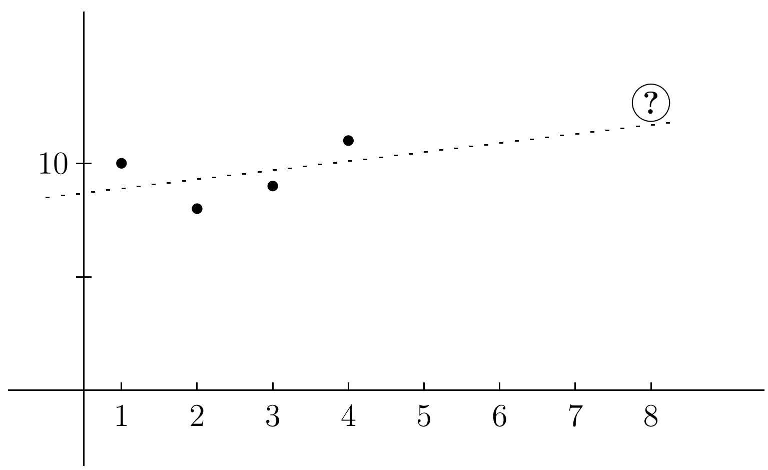
Konkretes Beispiel: In den 4 Quartalen eines Jahres erzielt eine Firma Gewinne von

$$10.000.000,\ 8.000.000,\ 9.000.000,\ 11.000.000\ \text{DM},$$

es soll das Firmenergebnis im letzten Quartal des nächsten Jahres geschätzt werden.

Vermutung: Der Gewinn verhält sich im wesentlichen wie eine Gerade.

$$10^6 \begin{bmatrix} 10 \\ 8 \\ 9 \\ 11 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} a + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} b$$



$$10^6 \begin{bmatrix} 10 \\ 8 \\ 9 \\ 11 \end{bmatrix} = \begin{array}{c} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \\ \;v_1 \quad v_2 \end{array} \begin{bmatrix} a \\ b \end{bmatrix}$$

In diesem Fall liefern die obigen Formeln das Gleichungssystem

$$\begin{bmatrix} \sqrt{30} & \frac{1}{3}\sqrt{30} \\ 0 & \frac{1}{3}\sqrt{6} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{30}} & \frac{3}{\sqrt{30}} & \frac{4}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{6}} \end{bmatrix} \cdot \begin{bmatrix} 10.000.000 \\ 8.000.000 \\ 9.000.000 \\ 11.000.000 \end{bmatrix} = \begin{bmatrix} 9.700.000 \cdot \frac{\sqrt{30}}{3} \\ 8.500.000 \cdot \frac{\sqrt{6}}{3} \end{bmatrix}$$

und man erhält die Lösung

$$\begin{aligned} a &= 400.000 \\ b &= 8.500.000 \\ y &= a \cdot t + b \\ y(8) &= a \cdot 8 + b = 11.700.000\ \text{DM}. \end{aligned}$$

Kapitel 11

Lineare Abbildungen und Matrizen

Wir haben bereits lineare Abbildungen verschiedenster Art eingeführt, zur Erinnerung: für zwei Vektorräume V, W über K ist

$$f : V \to W \quad \text{ein Homomorphismus,}$$

falls

$$\begin{aligned} f(v_1 + v_2) &= f(v_1) + f(v_2)\,, & \forall\, v_1, v_2 \in V, \\ f(\lambda v) &= \lambda f(v)\,, & \forall\, v \in V,\ \lambda \in K. \end{aligned}$$

Lemma 11.1 *Die linearen Abbildungen (Homomorphismen) $f : K^m \to K^n$ sind genau die Abbildungen der Form*

$$\begin{aligned} f_A \ :\ & K^m \to K^n \\ & x \mapsto Ax \end{aligned}$$

mit $A \in K^{n,m}$.

Beweis: Daß f_A linear ist, haben wir schon gezeigt (Kapitel 2).

Sei nun f eine lineare Abbildung von $K^m \to K^n$. Sei $\{e_1, \cdots, e_m\}$ die kanonische Basis von K^m und $\{\tilde{e_1}, \cdots, \tilde{e}_n\}$ die kanonische Basis von K^n. Dann gilt wegen $f(e_j) \in K^n$, daß $f(e_j) = \sum\limits_{i=1}^{n} a_{ij}\tilde{e}_i$ mit eindeutig bestimmten $a_{ij} \in K$. Setze $A = [a_{ij}] \in K^{n,m}$, so folgt $f(v) = A \cdot v$, denn

$$\begin{aligned} f(v) &= f\left(\sum_{k=1}^{m} \lambda_k e_k\right) \\ &= \sum_{k=1}^{m} \lambda_k f(e_k) = \sum_{k=1}^{m} \lambda_k \sum_{i=1}^{n} a_{ik}\tilde{e}_i\,, \end{aligned}$$

$$\begin{aligned} Av &= A \cdot \sum_{k=1}^{m} \lambda_k e_k \;=\; \sum_{k=1}^{m} \lambda_k (Ae_k) \\ &= \sum_{k=1}^{m} \lambda_k \left(\begin{bmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{bmatrix}\right) = \sum_{k=1}^{m} \lambda_k \left(\sum_{i=1}^{n} a_{ik}\tilde{e}_i\right). \end{aligned}$$

□

Sind V, W beliebige K-Vektorräume und gilt $\dim V = m$, $\dim W = n$, mit den Basen $\{v_1, \cdots, v_m\}$ von V und $\{w_1, \cdots, w_n\}$ von W, so können wir auch die linearen Abbildungen von $V \to W$ durch Matrizen beschreiben. Sei $f : V \to W$ linear und $f(v_j) = \sum_{i=1}^{n} a_{ij} w_i$, so setze $A = [a_{ij}] \in K^{n,m}$. Damit ist A die Matrixdarstellung von f bezüglich der Basen $\{v_1, \cdots, v_m\}, \{w_1, \cdots, w_n\}$.

<u>A hängt von der Wahl der Basen ab</u>.

Lemma 11.2 *Seien V, W K-Vektorräume mit Basen $\{v_1, \cdots, v_m\}, \{\tilde{v}_1, \cdots, \tilde{v}_m\}$ von V und $\{w_1, \cdots, w_n\}, \{\tilde{w}_1, \cdots, \tilde{w}_n\}$ von W. Seien P_V, P_W die Basisübergangsmatrizen in V und W, d.h.,*

$$\begin{array}{rcll} [v_1, \cdots, v_m] & = & [\tilde{v}_1, \cdots, \tilde{v}_m] \cdot P_V, & P_V \in K^{m,m}, \\ [w_1, \cdots, w_n] & = & [\tilde{w}_1, \cdots, \tilde{w}_n] \cdot P_W, & P_W \in K^{n,n}. \end{array}$$

Sei f eine lineare Abbildung von V nach W und $F = [f_{ij}]$ die Matrixdarstellung von f bezüglich der Basen $\{v_1, \cdots, v_m\}, \{w_1, \cdots, w_n\}$, d.h.

$$f(v_j) = \sum_{i=1}^{n} f_{ij} w_i,$$

so ist $P_W F P_V^{-1}$ die Matrixdarstellung von f bezüglich der Basen $\{\tilde{v}_1, \cdots, \tilde{v}_m\}, \{\tilde{w}_1, \cdots, \tilde{w}_n\}$.

Beweis: Sei $P_W = [p_{ij}]$, $P_V^{-1} = [q_{ij}]$. Dann gilt

$$\begin{array}{rcll} [\tilde{v}_1, \cdots, \tilde{v}_m] & = & [v_1, \cdots, v_m] \cdot P_V^{-1}, & \tilde{v}_j = \sum_{i=1}^{m} q_{ij} v_i, \\ [w_1, \cdots, w_n] & = & [\tilde{w}_1, \cdots, \tilde{w}_n] \cdot P_W, & w_j = \sum_{i=1}^{n} p_{ij} \tilde{w}_i, \end{array}$$

Für $j = 1, \ldots, m$ gilt:

$$\begin{aligned} f(\tilde{v}_j) &= f\left(\sum_{i=1}^{m} q_{ij} v_i\right) = \sum_{i=1}^{m} q_{ij} f(v_i) \\ &= \sum_{i=1}^{m} q_{ij} \left(\sum_{k=1}^{n} f_{ki} w_k\right) \\ &= \sum_{i=1}^{m} q_{ij} \left(\sum_{k=1}^{n} f_{ki} \left(\sum_{l=1}^{n} p_{lk} \tilde{w}_l\right)\right) \\ &= \left(\sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{l=1}^{n} q_{ij} f_{ki} p_{lk}\right) \tilde{w}_l \\ &= \sum_{l=1}^{n} \underbrace{\left(\sum_{k=1}^{n} \sum_{i=1}^{m} p_{lk} f_{ki} q_{ij}\right)}_{z_{lj}} \tilde{w}_l \ . \end{aligned}$$

Das heißt, $f([\tilde{v}_1, \ldots, \tilde{v}_m]) = [\tilde{w}_1, \ldots, \tilde{w}_n]\, Z$, mit $Z = P_W F P_V^{-1}$. $\square$

Definition 11.3 *Seien V, W K-Vektorräume und $f : V \to W$ eine lineare Abbildung und $F \in K^{n,m}$ die Matrixdarstellung. Dann heißt*

$$\begin{array}{lcl} \text{Kern}\,(f) & := & \{v \in V \mid f(v) = 0\} \quad \textit{bzw.} \\ \text{Kern}\,(F) & := & \{v \in K^m \mid Fv = 0\} \end{array}$$

<u>der Kern von f</u> (F) *oder* <u>Nullraum von f</u> (F) *und*

$$\begin{array}{lcl} \text{Bild}\,(f) & := & \{f(v) \in W \mid v \in V\} \quad \textit{bzw.} \\ \text{Bild}\,(F) & := & \{Fv \in K^n \mid v \in K^m\} \end{array}$$

<u>das Bild von f</u> (F).

Lemma 11.4 *Sei $f : V \to W$ eine lineare Abbildung zwischen den K-Vektorräumen V, W. So gilt:*

(a) Kern (f) *ist Untervektorraum von V und* Bild (f) *ist Untervektorraum von W.*

(b) f *ist injektiv genau dann, wenn* Kern $(f) = \{0\}$.

(c) Sei $v \in V$, so gilt

$$\{u \in V \mid f(u) = f(v)\} = \{v + \tilde{v} \mid \tilde{v} \in \text{Kern}\,(f)\} =: \underline{v + \text{Kern}\,(f)}.$$

Beweis:

(a) Übungsaufgabe

(b) f injektiv $\Longleftrightarrow (f(u) = f(v) \Longrightarrow u = v)$.

„$\Rightarrow$" Sei $v \in \text{Kern}\,(f) \Longrightarrow f(v) = 0$. Da $f(0) = 0$ ist und f injektiv, folgt $v = 0$.

„$\Leftarrow$" Seien $u, v \in V$ mit $f(u) = f(v)$.

$$\begin{array}{rcl} \Longrightarrow f(u - v) = 0 & \Longrightarrow & u - v \in \text{Kern}\,(f) \\ & \Longrightarrow & u = v, \text{ da Kern}\,(f) = \{0\}. \end{array}$$

(c) Sei $v \in V$.

$$\begin{array}{rcl} \{u \in V \mid f(u) = f(v)\} & = & \{u \in V \mid f(u - v) = 0\} \\ & = & \{u \in V \mid u - v \in \text{Kern}\,(f)\} \\ & = & \{v + \tilde{v} \mid \tilde{v} = (u - v) \in \text{Kern}\,(f)\}\,. \end{array}$$

$\square$

Satz 11.5 *Seien V, W K-Vektorräume,* $\dim V = m$, $\dim W = n$ *und sei $f : V \to W$ linear. Dann gibt es eine eindeutige Zahl $r = dim(\text{Bild}\,(f))$, mit $0 \le r \le \min(m, n)$, und Basen $\{v_1, \cdots, v_m\}$ von V und $\{w_1, \cdots, w_n\}$ von W mit*

$$f(v_i) = \begin{cases} w_i, & \textit{für } 1 \le i \le r \\ 0, & \textit{für } i \ge r + 1. \end{cases}$$

Es gilt

$$r =: \underline{\text{rang}\,(f)} = \text{Rang}\,(F),$$

wobei F die Matrixdarstellung von f bezüglich beliebiger Basen von V, W ist.

Beweis: Es ist klar, daß aus der (eindeutigen) Existenz von r folgt, daß $\dim(\text{Bild}\,(f)) = r$. Sei $\{w_1, \cdots, w_r\}$ Basis von Bild (f) und seien $v_1, \cdots, v_r \in V$, so daß

$$f(v_i) = w_i, \quad i = 1, \cdots, r.$$

Weiter sei $\{u_1, \cdots, u_k\}$ eine Basis von Kern (f). Wir zeigen, daß $\{v_1, \cdots, v_r, u_1, \cdots, u_k\}$ Basis von V ist. Dann kann man mit dem Basisergänzungssatz $\{w_1, \cdots, w_r\}$ zu einer Basis von W ergänzen.

Sei $v \in V$, so ist $f(v) \in \text{Bild}\,(f)$, also gibt es $\lambda_1, \cdots, \lambda_r \in K$ mit $f(v) = \sum_{i=1}^{r} \lambda_i w_i = \sum_{i=1}^{r} \lambda_i f(v_i)$. Also ist

$$f\left(v - \sum_{i=1}^{r} \lambda_i v_i\right) = 0, \qquad \text{und damit} \qquad \left(v - \sum_{i=1}^{r} \lambda_i v_i\right) \in \text{Kern}\,(f),$$

$$v - \sum_{i=1}^{r} \lambda_i v_i = \sum_{i=1}^{k} \mu_i u_i,$$

$$v = \sum_{i=1}^{k} \mu_i u_i + \sum_{i=1}^{r} \lambda_i v_i \quad \Longrightarrow \quad v \in \mathcal{L}(u_1, \cdots, u_k, v_1, \cdots, v_r).$$

Es fehlt noch die lineare Unabhängigkeit. Es gelte

$$\sum_{i=1}^{r} \lambda_i v_i + \sum_{i=1}^{k} \mu_i u_i = 0\,.$$

Die Anwendung von f ergibt

$$\begin{aligned} & \sum_{i=1}^{r} \lambda_i f(v_i) + \underbrace{\sum_{i=1}^{k} \mu_i f(u_i)}_{=0} \;=\; \sum_{i=1}^{r} \lambda_i f(v_i) = 0 \\ \Longrightarrow \quad & \lambda_1, \cdots, \lambda_r = 0 \\ \Longrightarrow \quad & \mu_1, \cdots, \mu_k = 0. \end{aligned}$$

Noch zu zeigen, daß $r = \text{Rang}\ (F)$.

Die Matrixdarstellung F von f bezüglich $\{v_1, \cdots, v_m\}$, $\{w_1, \cdots, w_n\}$ ist

$$F = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right], \qquad \text{in diesem Fall gilt somit Rang}\ (F) = r\,.$$

Sei $\{\tilde{v}_1, \cdots, \tilde{v}_m\}$ Basis von V und $\{\tilde{w}_1, \cdots, \tilde{w}_n\}$ Basis von W und sei $\tilde{F}$ die Matrixdarstellung von f bezüglich dieser Basen. Seien P_V, P_W die Basisübergangsmatrizen. Dann gilt nach Lemma 11.2, daß

$$\tilde{F} = P_W F P_V^{-1}$$

wobei F die Matrixdarstellung von f bezüglich $\{v_1, \cdots, v_m\}, \{w_1, \cdots, w_n\}$ ist. Da aber F Rang r hat, folgt, daß auch $\tilde{F}$ Rang r hat. $\square$

Definition 11.6 *Seien V, W K-Vektorräume. Die Menge aller linearen Abbildungen von V nach W bezeichnen wir mit $L(V, W)$.*

Wenn man $L(V, W)$ als Teilmenge aller Abbildungen $A(V, W) = \{f : V \to W \mid f \text{ Abbildung }\}$ auffaßt, so wird $A(V, W)$ durch die Operationen

$$\begin{array}{rlcll} + : & (f+g)(v) & = & f(v) + g(v) & \forall f, g \in A(V,W),\ \forall v \in V, \\ \cdot : & (\lambda \cdot f)(v) & = & \lambda \cdot f(v) & \forall f \in A(V,W),\ \forall \lambda \in K,\ \forall v \in V, \end{array}$$

zu einem K-Vektorraum. $L(V, W)$ ist ein Unterraum von $A(V, W)$.

Satz 11.7 *Seien V, W K-Vektorräume, mit $\{v_1, \cdots, v_m\}$ Basis von V und $\{w_1, \cdots, w_n\}$ Basis von W. Ist $f \in L(V, W)$, so sei $mat(f) \in K^{n,m}$ die Matrixdarstellung von f bezüglich der Basen $\{v_1, \cdots, v_m\}$ und $\{w_1, \cdots, w_n\}$. Dann ist*

$$mat : L(V, W) \to K^{n,m}$$

ein Vektorraumisomorphismus.

Beweis: Zeige zuerst, daß mat eine lineare Abbildung ist (also ein Homomorphismus). Seien $f, g \in L(V, W)$ mit

$$\begin{aligned} f(v_j) &= \sum_{i=1}^{n} f_{ij} w_i \quad \text{und} \\ g(v_j) &= \sum_{i=1}^{n} g_{ij} w_i. \end{aligned}$$

Dann gilt $mat(f) = F = [f_{ij}]$ und $mat(g) = G = [g_{ij}]$. Es folgt, daß

$$\begin{aligned} (f+g)(v_j) &= f(v_j) + g(v_j) \\ &= \sum_{i=1}^{n} (f_{ij} + g_{ij}) w_i \,. \end{aligned}$$

Also gilt $mat(f + g) = F + G$. Sei nun $f \in L(V, W)$, $\lambda \in K$, so folgt

$$\begin{aligned} (\lambda \cdot f)(v_j) = \lambda \cdot f(v_j) &= \lambda \left(\sum_{i=1}^{n} f_{ij} w_i \right) \\ &= \sum_{i=1}^{n} (\lambda f_{ij}) w_i \end{aligned}$$

$$\Longrightarrow \quad mat(\lambda \cdot f) = \lambda F = \lambda \cdot mat(f).$$

Wir müssen noch die Bijektivität zeigen.

Sei $f \in \text{Kern}\,(mat)$, so gilt $mat(f) = 0 \in K^{n,m}$ und damit $f(v_j) = 0$ für alle $1 \leq j \leq m$. Also folgt $f(v) = 0$, $\forall v \in V$, und damit $f = 0$ in $L(V, W) \Rightarrow$ mat ist injektiv (Lemma 11.4). Ist $A = [a_{ij}] \in K^{n,m}$, so definiere $f \in L(V, W)$ durch

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

Dann ist $mat(f) = A$ und damit ist mat surjektiv. □

Wir erhalten auch sehr leicht eine Basis von $L(V, W)$, wenn wir Basen in V und W haben.

Sei $\{E_{ij},\ i = 1, \dots, n, j = 1, \dots, m\}$ die kanonische Basis in $K^{n,m}$.
Definiere für $i = 1, \dots, n,\ j = 1, \dots, m$, φ_{ij} durch $\varphi_{ij} = mat^{-1}(E_{ij}) \in L(V, W)$, so folgt

$$\varphi_{ij}(v_l) = \begin{cases} w_i, & \text{falls} \quad j = l \\ 0, & \text{falls} \quad j \neq l \end{cases}$$

und damit erhält man über $\varphi_{ij}\left(\sum_{l=1}^{m} \alpha_l v_l\right) = \alpha_j w_i$ eine Basis von $L(V, W)$.

Definition 11.8 *Sei V ein K-Vektorraum, dann nennt man $V^* := L(V, K)$ den* <u>Dualraum</u> *zu V.*

Ist $\{v_1, \cdots, v_m\}$ eine Basis von V, so definiere $\varphi_1, \cdots, \varphi_m \in V^*$ durch

$$\varphi_i(v_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} . \tag{11.9}$$

Dann ist $\{\varphi_1, \cdots, \varphi_m\}$ eine Basis von V^*, sie heißt <u>duale Basis</u> *zu $\{v_1, \cdots, v_m\}$.*

Es gilt

$$\varphi_i\left(\sum_{j=1}^{m} \alpha_j v_j\right) = \alpha_i \quad \text{mit} \quad \alpha_1, \cdots, \alpha_m \in K, \tag{11.10}$$

d.h., φ_i ordnet jedem $v \in V$ das Körperelement zu, welches der Koeffizient von v_i ist, wenn man v als Linearkombination von $v_1, \cdots, v_m$ darstellt. Also

$$v = \sum_{i=1}^{m} \varphi_i(v) v_i\,,\ \forall v \in V, \tag{11.11}$$

aber auch

$$\left(\sum_{i=1}^{m} \beta_i \varphi_i\right)(v_j) = \beta_j \quad \text{mit} \quad \beta_1, \cdots, \beta_m \in K \tag{11.12}$$

und damit

$$f = \sum_{j=1}^{m} f(v_j) \varphi_j\,,\ \forall f \in V^*.$$

Beispiel 11.13 $V = \mathbb{R}^3, W = \mathbb{R}$.

$$\begin{aligned} V^* = L(V,W) &= \{f : \mathbb{R}^3 \to \mathbb{R} \mid f \text{ linear}\} \\ &\cong \{A \in \mathbb{R}^{1,3}\} = \{[x,y,z] \mid x,y,z \in \mathbb{R}\}\,. \end{aligned}$$

- $\{e_1, e_2, e_3\}$ Basis von V,

  $$\varphi_i(e_j) = \delta_{ij}, \text{ zugehörige Matrizen: } \begin{aligned} \phi_1 &= [\,1,\ 0,\ 0\,], \\ \phi_2 &= [\,0,\ 1,\ 0\,], \\ \phi_3 &= [\,0,\ 0,\ 1\,]. \end{aligned}$$

- $\left\{ v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ Basis von V.

  Es muß gelten $\varphi_i(v_j) = \delta_{ij}$. Seien ϕ_1, ϕ_2, ϕ_3 Matrixdarstellung von $\varphi_1, \varphi_2, \varphi_3$ bezüglich der kanonischen Basis, so gilt

  $$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} [v_1, v_2, v_3] = I_3\,,$$

  d.h., die ϕ_i sind die Zeilen der Inversen von $[v_1, v_2, v_3]$.

  $$\begin{aligned} \phi_1 &= [\ \ 1, -1, \ \ 0\,], \\ \phi_2 &= [\ \ 0, \ \ 1, -1\,], \\ \phi_3 &= [\ \ 0, \ \ 0, \ \ 1\,]. \end{aligned}$$

  $$\varphi_i(v_j) = \phi_i \cdot v_j = \delta_{ij}.$$

Kapitel 12

Äquivalenzrelationen und Quotientenräume

Definition 12.1 *Sei M eine Menge. Eine Teilmenge R von $M \times M$ heißt* Äquivalenzrelation*, falls folgende Eigenschaften für beliebige $u, v, w \in M$ gelten:*

(a) Reflexivität: $(u, u) \in R$,

(b) Symmetrie: $(u, v) \in R \implies (v, u) \in R$,

(c) Transitivität: $(u, v) \in R$, $(v, w) \in R \implies (u, w) \in R$.

Falls $(u, v) \in R$, so schreiben wir $u \sim v$, u äquivalent v.

Beispiel 12.2

a) Sei $M = \mathbb{K}^{m,n}$, $\quad A \sim B \iff \exists P, Q$ invertierbar: $A = PBQ$.

   (i) $A \sim A$, mit $P = I_m, Q = I_n$.

   (ii) $A \sim B \Longrightarrow B \sim A$, denn $A = PBQ \Longrightarrow B = P^{-1}AQ^{-1}$.

   (iii) $A \sim B,\ B \sim C$,
   $A = PBQ,\ B = WCY,\ P, Q, W, Y$ invertierbar,
   $\Longrightarrow A = (PW)C(YQ)$.

b) Sei K ein Körper, V ein K-Vektorraum und $U \subset V$ ein Untervektorraum.
Sei für $v, w \in V$: $\quad v \sim w \iff v - w \in U$.

   (i) $v \sim v$, denn $v - v = 0 \in U$,

   (ii) $v \sim w \implies v - w \in U \implies w - v = -(v - w) \in U$,

   (iii) $v \sim w, w \sim z \Longrightarrow v - w \in U,\ w - z \in U \implies v - z = (v - w) + (w - z) \in U$.

c) $M = K^{n,n}$, $\quad A \sim B \iff \exists\, P$ invertierbar: $A = P^{-1}BP$.

   (i) $A \sim A$ mit $P = I$.

   (ii) $A \sim B \implies A = P^{-1}BP \implies B = PAP^{-1}$.

(iii) $A \sim B,\ B \sim C \implies A = P^{-1}BP,\ B = Q^{-1}CQ,$
$A = (QP)^{-1}C(QP).$

Definition 12.3 *Sei M eine Menge und R eine Äquivalenzrelation auf M. Eine Teilmenge $T \subset M$ heißt* Äquivalenzklasse *bezüglich R, falls gilt:*

(a) $T \neq \emptyset$,

(b) $u, v \in T \implies u \sim v$,

(c) $u \in T$, $v \in M$ *und* $u \sim v \implies v \in T$.

Lemma 12.4 *Ist R eine Äquivalenzrelation auf M, so gehört jedes $u \in M$ zu genau einer Äquivalenzklasse. Das heißt insbesondere, dass für zwei verschiedene Äquivalenzklassen T_1, T_2 gilt:*

$$T_1 \cap T_2 = \emptyset\,.$$

Beweis: Sei $u \in M$. Dann ist $T_u = \{w \in M \mid u \sim w\}$ eine Äquivalenzklasse, denn

(a) $u \in T_u \implies T_u \neq \emptyset$,

(b) $w_1, w_2 \in T_u \implies w_1 \sim u, w_2 \sim u \overset{\text{Sym./Trans.}}{\implies} w_1 \sim w_2$,

(c) $w \in T_u,\ z \in M,\ w \sim z,$
$u \sim w, w \sim z \overset{\text{Trans.}}{\implies} u \sim z.$

Also gehört jedes $u \in M$ zu einer Klasse. Seien T_u, T_v zwei Klassen. Falls $T_u \cap T_v \neq \emptyset$, so gibt es $w \in T_u \cap T_v$. Sei $u_1 \in T_u \implies u_1 \sim w$. Da $w \in T_v$ und $w \sim u_1$, so folgt $u_1 \in T_v \implies T_u \subset T_v$. Analog folgt $T_v \subset T_u \implies T_u = T_v$. □

Jede Äquivalenzrelation R auf einer Menge M liefert eine Zerlegung von M in disjunkte Äquivalenzklassen $T_i,\ i \in I$, d.h.,

$$M = \bigcup_{i \in I} T_i\,, \qquad T_i \cap T_j = \emptyset,\ i \neq j\,.$$

Dies heißt Klasseneinteilung von M.

Umgekehrt definiert jede Zerlegung einer Menge M in disjunkte Teilmengen eine Klasseneinteilung.

Definition 12.5

a) Sei R eine Äquivalenzrelation auf einer Menge M. Die Menge

$$M/R = \{T_i \mid i \in I, T_i \text{ Äquivalenzklasse bzgl. } R\}$$

heißt Quotientenmenge*. M/R ist eine Menge von Teilmengen von M.*

b) *Sei T eine Äquivalenzklasse, so heißt jedes $t \in T$ ein* Repräsentant *dieser Klasse.*

Zu M/R gibt es eine natürliche Abbildung

$$\Pi : \; M \to M/R,$$

die jedem $u \in M$ die Äquivalenzklasse zuordnet, zu der u gehört. Π ist offensichtlich surjektiv.

Beispiel 12.6 Sei V ein Vektorraum über dem Körper K und U ein Untervektorraum von V. Betrachte die Relation aus Beispiel 12.2 b).
Die Äquivalenzklassen sind alle Teilräume der Form

$$v + U, \quad v \in V.$$

Die Quotientenmenge bezüglich dieser Relation bezeichnen wir als V/U. Jede dieser Teilmengen hat v als Repräsentanten.

In den nächsten Kapiteln werden wir Repräsentanten für die Relationen in Beispiel 12.2 a) und c) studieren.

Wir wollen nun die Äquivalenzrelation für Unterräume U des Vektorraums V über K, gegeben durch

$$u \sim v \quad \text{wenn} \quad u - v \in U$$

genauer betrachten.

Definition 12.7 *Sei V ein Vektorraum über dem Körper K. $M \subset V$ heißt* affiner Unterraum *von V, falls $M = \emptyset$ oder $M = v + U$, wobei U ein Untervektorraum von V ist.*

Lemma 12.8

(a) *Seien $v, \tilde{v} \in V$ und $U, \tilde{U}$ Untervektorräume von V, so gilt:*

$$v + U = \tilde{v} + \tilde{U} \iff U = \tilde{U} \text{ und } v - \tilde{v} \in U.$$

(b) *Sei $v \in V$, $\quad U$ Untervektorraum von V und $M = v + U$, so ist*

$$U = \{m_1 - m_2 \;\mid\; m_1, m_2 \in M\}.$$

Beweis:

(a) „$\Longrightarrow$“ $\quad v+U=\tilde{v}+\tilde{U}$. Da $0\in\tilde{U}$, so folgt

$$\tilde{v}=v+u \text{ für ein } u\in U \quad\Rightarrow\quad \tilde{v}-v\in U.$$

Sei $\tilde{u}\in\tilde{U}$. $\Rightarrow \exists\, u\in U$ mit

$$v+u=\tilde{v}+\tilde{u}, \text{ also } \tilde{u}=v-\tilde{v}+u\in U \quad\Rightarrow\quad \tilde{U}\subset U.$$

Umgekehrt folgt genauso $\tilde{U}\supset U$.

$$\Rightarrow\ \tilde{U}=U.$$

„$\Longleftarrow$“ $\quad m\in v+U\Rightarrow m=v+u$ für ein $u\in U$. Sei $v-\tilde{v}=u_1\in U$

$$\Rightarrow\ v=u_1+\tilde{v}\ \Rightarrow\ m=u+u_1+\tilde{v}.$$

Da $u+u_1\in U$ und $U=\tilde{U}$, so folgt $m\in\tilde{v}+\tilde{U}$. Also ist $v+U\subset\tilde{v}+\tilde{U}$. Umgekehrt zeigt man analog $\tilde{v}+\tilde{U}\subset v+U$.

(b) Sei $u\in U$. Setze $m_1=v+u\in M$ und $m_2=v+0\in M$

$$\Rightarrow\ u=m_1-m_2\,.$$

Umgekehrt sei $m_1=v+u\in M,\ m_2=v+\tilde{u}\in M$ mit $u,\tilde{u}\in U$.

$$\Rightarrow m_1-m_2=v+u-v-\tilde{u}=u-\tilde{u}\in U.$$

□

Man definiert für affine Räume

$$\dim(v+U):=\dim U \quad\text{und}\quad \dim\emptyset=-1.$$

Geometrische Interpretation: Ist M affiner Teilraum von V, so ist das,

falls $\dim M=0$, ein Punkt,
$\dim M=1$, eine Gerade,
$\dim M=2$, eine Ebene.

Falls $\dim M=n-1$, $\dim V=n$, so heißt M Hyperebene.

Satz 12.9 *Sei $M\subset K^m$ für Körper K. M ist affiner Unterraum von K^m genau dann, wenn M die Lösungsmenge eines linearen Gleichungssystems in m Unbekannten ist.*

Beweis: Sei M Lösungsmenge von $Ax = b$ mit $A \in K^{n,m}$ und sei

$$U = \{u \in K^m \mid Au = 0\},$$

so haben wir schon gezeigt, dass U ein Untervektorraum ist. Damit ist

$$M = y + U \quad \text{für eine Lösung } y \text{ von } Ay = b.$$

Für die Umkehrung sei $v \in K^m$ und $U \subset K^m$ ein Untervektorraum, $M = v + U$.
Sei $\{u_1, \ldots, u_r\}$ Basis von U und $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$ Basis von K^m.
Sei $\{\varphi_1, \ldots, \varphi_m\}$ duale Basis zu $\{u_1, \ldots, u_m\}$, d.h. $\varphi_i(u_j) = \delta_{ij}$. Dann gilt

$$U = \{w \in K^m \mid \varphi_i(w) = 0 \text{ für } r+1 \leq i \leq m\}, \text{ denn:}$$

Für $u = \sum_{j=1}^{r} \lambda_j u_j$ gilt

$$\varphi_i(u) = \sum_{j=1}^{r} \varphi_i(u_j) = 0 \text{ für } i > r.$$

Umgekehrt, sei $w \in K^m$ mit $\varphi_i(w) = 0$ für $i > r$ und $w = \sum_{j=1}^{m} \lambda_j u_j$, so folgt

$$\varphi_i(w) = \lambda_i \Rightarrow \lambda_i = 0 \text{ für } i > r,$$

oder $w \in U$.

Sei $a_i \in K^{1,m}$ die Matrixdarstellung von φ_i bezüglich der kanonischen Basen in K^m, K, und

$$A = \begin{bmatrix} a_{r+1} \\ \vdots \\ a_m \end{bmatrix} \in K^{m-r,r},$$

so ist U Lösungsmenge von $Ax = 0$, dargestellt in der kanonischen Basis.

D.h., es gilt

$$\text{Kern}\,(A) = \bigcap_{i=r+1}^{m} \text{Kern}\,(\varphi_i) = U.$$

Setze $b = Av$, so ist $M = v + U$ Lösungsmenge von $Ax = b$.

Es bleibt der Fall $M = \emptyset$. Dann wähle

$$A = 0 \in K^{1,m}, \; b \neq 0 \in K.$$

$\square$

Definition 12.10 *Sei V Vektorraum über K und U Unterraum von V. Betrachte die Äquivalenzklassen V/U, d.h., die Menge aller affinen Teilräume der Form $v+U$ mit $v \in V$. V/U heißt* Quotientenraum von V modulo U.

Wie schon vorher definiert, gibt es eine surjektive Abbildung

$$\begin{aligned} \Pi \ &: V \longrightarrow V/U \qquad \text{mit Kern}\,(\Pi) = U. \\ &: v \longmapsto v + U \end{aligned}$$

Es ist klar, dass $\dim V/U = \dim V - \dim U$.

Weitere Eigenschaften.

Satz 12.11

(a) Sei $\{v_1, \ldots, v_s\}$ Basis von V, so bilden die verschiedenen Elemente von $\{v_1 + U, \ldots, v_s + U\}$ eine Basis von V/U.

(b) Ist $V = U + \mathcal{L}(v_1, \ldots, v_t)$, so ist $V/U = \mathcal{L}(v_1 + U, \ldots, v_t + U)$.

(c) Sind $v_1, \ldots, v_s \in V$ und sind $v_1 + U, \ldots, v_s + U$ linear unabhängig in V/U, so sind $v_1, \ldots, v_s$ linear unabhängig in V.

(d) Seien $v_1, \ldots, v_s \in V$ und seien $v_1 + U, \ldots, v_s + U$ linear unabhängig in V/U. Sind $u_1, \ldots, u_r$ linear unabhängig in U, so sind $u_1, \ldots, u_r, v_1, \ldots, v_s$ linear unabhängig in V.

(e) Sei $\{u_1, \ldots, u_r\}$ Basis von U, $v_1, \ldots, v_s \in V$. $\{u_1, \ldots, u_r, v_1, \ldots, v_s\}$ ist Basis von V genau dann, wenn $\{v_1 + U, \ldots, v_s + U\}$ Basis von V/U ist.

Beweis: (a) …(d) Übung.

(e) „$\Longrightarrow$“ Wegen (b) ist $\mathcal{L}(v_1 + U, \ldots, v_s + U) = V/U$.
Da $\{u_1, \ldots, u_r, v_1, \ldots, v_s\}$ eine Basis von V bilden, gilt

$$v_i - v_j \notin U, \qquad i, j = 1, \ldots, s,\ i \neq j.$$

Damit sind alle Elemente von $\{v_1 + U, \ldots, v_s + U\}$ voneinander verschieden. Aus (a) folgt nun die lineare Unabhängigkeit von $v_1 + U, \ldots, v_s + U$.

„$\Longleftarrow$“ Nach (d) sind $u_1, \ldots, u_r, v_1, \ldots, v_s$ linear unabhängig in V.
$\mathcal{L}(u_1, \ldots, u_r, v_1, \ldots, v_s) = V$, da $\dim V = r + s$. Also folgt die Behauptung.

□

Kapitel 13

Eigenwerte und Eigenvektoren

Definition 13.1 *Seien $n, m \in \mathbb{N}$. Zwei Matrizen $A, B \in K^{n,m}$ heißen* <u>äquivalent</u>*, falls es invertierbare Matrizen P, Q gibt, $P \in K^{n,n}, Q \in K^{m,m}$, so dass*

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

Wir wissen bereits, dass die Ränge von A und B gleich sind, wenn A und B äquivalent sind. Aus Satz 11.5 und Satz 4.2 wissen wir, dass A und damit auch B äquivalent ist zu $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, wobei $r = \text{Rang}\,(A) = \text{Rang}\,(B)$. Betrachte nun die Menge

$$M_r = \left\{ A \in K^{n,m} \mid A \text{ äquivalent zu } \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \right\} \text{ mit } 0 \leq r \leq \min(m,n).$$

Dann folgt

$$\begin{aligned} K^{n,m} &= \bigcup_{r=0}^{\min(m,n)} M_r \\ &= M_0 \cup M_1 \cup \ldots \cup M_{\min(m,n)}\,. \end{aligned}$$

Die Matrix $N_r = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in K^{n,m}$ heißt <u>Normalform</u> von $A \in M_r$ bezüglich der Relation Äquivalenz. Wir wollen nun die Normalform unter einer anderen Äquivalenzrelation betrachten, nämlich unter der Ähnlichkeit für Matrizen in $K^{n,n}$ (s. Def. 6.8). Dies ist im allgemeinen ein sehr schweres Problem, wir werden uns daher auf $K = \mathbb{C}$ beschränken.

<u>Zur Erinnerung:</u> Ein Endomorphismus $f : V \to V$ (V Vektorraum über K) ist eine lineare Abbildung (Homomorphismus) von V in sich. Ist $\{v_1, \ldots, v_n\}$ eine Basis von V, so besitzt f eine Matrixdarstellung bezüglich $\{v_1, \ldots, v_n\}$. Aus Lemma 11.2 folgt, dass alle Matrixdarstellungen von f ähnlich sind, denn wenn $\{v_1, \ldots, v_n\}, \{\tilde{v}_1, \ldots, \tilde{v}_n\}$ Basen von V sind mit der Basisübergangsmatrix P und F die Matrixdarstellung von f bezüglich $\{v_1, \ldots, v_n\}$, Z die Matrixdarstellung von f bezüglich $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$, so folgt

$$Z = PFP^{-1}.$$

Wir können dann auch wieder $K^{n,n}$ darstellen als disjunkte Vereinigung von Mengen

$$J_i = \{A \in K^{n,n}, A \text{ ähnlich zu } A_i\}$$

und dann versuchen, eine Normalform zu bestimmen.

Zur Erinnerung: Ähnliche Matrizen haben das gleiche charakteristische Polynom (Satz 6.9), d.h., unabhängig von der Matrixdarstellung gibt es zu einem Endomorphismus f ein *charakteristisches Polynom* $P_f(x) = P_F(x)$, wobei F eine Matrixdarstellung von f ist.

Definition 13.2 *Sei $f : V \to V$ ein Endomorphismus, V ein Vektorraum über dem Körper K. Ein Vektor $v \in V \setminus \{0\}$ heißt Eigenvektor von f mit Eigenwert $\lambda \in K$, falls $f(v) = \lambda v$.*

Zur Erinnerung: Ist $A \in K^{n,n}$, so ist v Eigenvektor mit Eigenwert λ, falls $Av = \lambda v, \lambda \in K$.

A definiert einen Endomorphismus und es gilt, dass v, λ Eigenvektor und Eigenwert dieser linearen Abbildung sind. Die Eigenwerte des Endomorphismus sind Nullstellen des charakteristischen Polynoms, daraus folgt: Falls $\dim V = n$, so hat f höchstens n Eigenwerte.

Lemma 13.3 *Sei $f : V \to V$ ein Endomorphismus des Vektorraums V und sei $\{v_1, \ldots, v_n\}$ Basis von V. $v_1, \ldots, v_n$ sind Eigenvektoren von f genau dann, wenn die Matrixdarstellung von f bezüglich $\{v_1, \ldots, v_n\}$ eine Diagonalmatrix ist.*

Beweis: Sei $\lambda_i \in K$ Eigenwert zum Eigenvektor v_i für $1 \leq i \leq n$, also $f(v_i) = \lambda_i v_i$. Dann ist die Matrixdarstellung F von f bezüglich $\{v_1, \ldots, v_n\}$ die Diagonalmatrix

$$F = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Umgekehrt sei diese Matrixdarstellung F von f bezüglich $\{v_1, \ldots, v_n\}$ als Diagonalmatrix gegeben. Dann gilt $f(v_i) = \lambda_i v_i\,,\ i = 1, \ldots, n$. Da $\{v_1, \ldots, v_n\}$ eine Basis ist, sind $v_1, \ldots, v_n \neq 0$, also Eigenvektoren. □

Definition 13.4 *Sei $f : V \to V$ ein Endomorphismus eines Vektorraums V über einem Körper K. Falls es eine Basis $\{v_1, \ldots, v_n\}$ von V aus Eigenvektoren zu f gibt, so heißt f diagonalisierbar. Entsprechend heißt eine Matrix $A \in K^{n,n}$ diagonalisierbar, falls sie ähnlich zu einer Diagonalmatrix ist.*

Beispiel 13.5

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & -2 \\ 0 & 2 & 0 \\ 4 & 0 & 5 \end{bmatrix} \text{ ist diagonalisierbar.} \\ B &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ ist nicht diagonalisierbar,} \end{aligned}$$

denn sei $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $P^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \cdot \dfrac{1}{ad - bc}$, so folgt

$$\begin{aligned} PBP^{-1} &= \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \cdot \frac{1}{ad - bc} \\ &= \begin{bmatrix} -ca & a^2 \\ -c^2 & ca \end{bmatrix} \cdot \frac{1}{ad - bc} . \end{aligned}$$

Das kann nur diagonal sein, falls $a = 0, c = 0$. Dann ist aber P nicht invertierbar.

Satz 13.6 *Sei $f : V \to V$ ein Endomorphismus eines Vektorraums V über dem Körper K. Sei v_i Eigenvektor von f mit Eigenwert $\lambda_i \in K$ für $1 \leq i \leq r$. Sind die Zahlen λ_i paarweise verschieden, d.h., $\lambda_i \neq \lambda_j$ für $i \neq j$, so sind $v_1, \ldots, v_r$ linear unabhängig.*

Beweis: Induktion nach r.

I.A.: $r = 1$ ist klar, da Eigenvektoren von Null verschieden sind.

I.V.: Behauptung sei richtig für $1 \leq k \leq r - 1$.

I.S.: Sei $r > 1$ und $\sum\limits_{i=1}^{r} \alpha_i v_i = 0$ mit $\alpha_1, \ldots, \alpha_r \in K$.
Dann gilt

$$\begin{aligned} 0 = f(0) &= f\left(\sum_{i=1}^{r} \alpha_i v_i \right) = \sum_{i=1}^{r} \alpha_i f(v_i) \\ &= \sum_{i=1}^{r} \alpha_i \lambda_i v_i \ , \end{aligned}$$

und da

$$\begin{aligned} 0 &= \lambda_r \cdot 0 = \sum_{i=1}^{r} \alpha_i \lambda_r v_i, \quad \text{so folgt} \\ 0 &= \sum_{i=1}^{r} \alpha_i \lambda_i v_i - \sum_{i=1}^{r} \alpha_i \lambda_r v_i \\ &= \sum_{i=1}^{r-1} \alpha_i (\lambda_i - \lambda_r) v_i \, . \end{aligned}$$

Nach I.V. sind $v_1, \ldots, v_{r-1}$ linear unabhängig, also folgt $\alpha_i(\lambda_i - \lambda_r) = 0 \Longrightarrow \alpha_i = 0$ für $i = 1, \ldots, r - 1$. Dann folgt aus

$$0 = \sum_{i=1}^{r} \alpha_i v_i = \alpha_r v_r$$

auch $\alpha_r = 0$, da $v_r \neq 0$.

$\square$

Korollar 13.7

a) *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums V über einem Körper K, $\dim V = n$. Besitzt f n paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n$, so ist f diagonalisierbar.*

b) *Sei $A \in K^{n,n}$. Besitzt A n paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n$, so ist A diagonalisierbar.*

Beispiel 13.8 Die Umkehrung gilt nicht, denn z. B. λI_n hat n mal den Eigenweert λ, ist aber diagonalisierbar. Ob eine Matrix diagonalisierbar ist, hängt, da die Eigenwerte Nullstellen des charakterischen Polynoms sind, stark vom Körper K ab.
Sei $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{R}^{2,2}$, so ist $P_A(\lambda) = \lambda^2 + 1$ und es gibt keine Eigenwerte in $\mathbb{R}$. Wenn wir A aber als Element von $\mathbb{C}^{2,2}$ auffassen, so ist A ähnlich zu $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$.

Definition 13.9 *Sei $A \in K^{n,n}$ und $\lambda \in K$ Eigenwert von A.*

a) *Dann heißt $E_\lambda = \text{Kern}\,(\lambda I - A)$ der* <u>Eigenraum</u> *zu λ.*
Die Dimension von E_λ, $g(\lambda) = \dim E_\lambda$, heißt <u>geometrische Vielfachheit</u> *von λ.*

b) *Sei $P_A(x)$ das charakteristische Polynom von A. Dann ist $x - \lambda$ ein Teiler von $P_A(x)$. Mit $a(\lambda)$ bezeichnen wir die größte Zahl, so dass $(x-\lambda)^{a(\lambda)}$ ein Teiler von $P_A(x)$ ist. $a(\lambda)$ heißt* <u>algebraische Vielfachheit</u> *von λ.*

Lemma 13.10 *Sei $A \in K^{n,n}$ und λ Eigenwert von A, so gilt*

$$a(\lambda) \geq g(\lambda).$$

Beweis: Sei $r = g(\lambda)$ und $\{v_1, \ldots, v_r\}$ eine Basis von E_λ. Nach dem Basisergänzungssatz können wir $v_{r+1}, \ldots, v_n$ bestimmen, so dass $\{v_1, \ldots, v_n\}$ eine Basis von K^n ist. Sei f_A die zu A gehörende lineare Abbildung und sei B die Matrixdarstellung von f_A in der Basis $\{v_1, \ldots, v_n\}$. B und A sind ähnlich, also $P_A(x) = P_B(x)$.
Da $f_A(v_i) = \lambda v_i$ für $1 \leq i \leq r$ und $f_A(v_j) = \sum_{i=1}^{n} b_{ij} v_i$, so hat B die Form

$$B = \begin{array}{cc} \left[\begin{array}{c|c} \lambda I_r & B_{12} \\ \hline 0 & B_{22} \end{array}\right] & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} r & n-r \end{array} & \end{array} .$$

Also folgt $P_A(x) = P_B(x) = (x-\lambda)^r \cdot P_{B_{22}}(x)$. Insbesondere ist damit $a(\lambda) \geq g(\lambda) = r$. $\square$

Beispiel 13.11

$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ hat offensichtlich den Eigenwert 0 mit algebraischer Vielfachheit 2.

$$\text{Kern}\,(0 \cdot I - A) = \text{Kern}\,(-A) = \text{Kern}\,(A) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$$

hat Dimension 1 (denn $Ax = \begin{bmatrix} x_2 \\ 0 \end{bmatrix} = 0 \Longleftrightarrow x_2 = 0$).

Damit erhalten wir nun **einen der wichtigsten Sätze über Eigenwerte**.

Satz 13.12 *Sei V endlichdimensionaler Vektorraum über dem Körper K und $f : V \to V$ ein Endomorphismus. Dann sind die folgenden Bedingungen äquivalent:*

(a) f ist diagonalisierbar.

(b) (i) $P_f(x)$ zerfällt in Linearfaktoren, d.h., $P_f(x) = \prod_{i=1}^{n}(x - \lambda_i)$.
(ii) Für alle Eigenwerte λ von f gilt $a(\lambda) = g(\lambda)$.

(c) Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f, so gilt

$$V = \bigoplus_{i=1}^{r} E_{\lambda_i} = E_{\lambda_1} \dot{+} E_{\lambda_2} \dot{+} \ldots \dot{+} E_{\lambda_r}\,.$$

Beweis:
„*(a)* $\Longrightarrow$ *(c)*“
Sei $\{v_1, \ldots, v_n\}$ Basis von V aus Eigenvektoren von f. Fasse davon jeweils alle Eigenvektoren zum gleichen Eigenwert λ_i zusammen als $v_{i_1}, \ldots, v_{i_l}$. Da die v_j eine Basis bilden, sind sie linear unabhängig, und damit gilt

$$E_{\lambda_i} = \mathcal{L}(v_{i_1}, \ldots, v_{i_l}) \;\text{ und }\; V = E_{\lambda_1} \dot{+} \ldots \dot{+} E_{\lambda_r}.$$

„*(c)* $\Longrightarrow$ *(a)*“
Wähle Basen $\{v_{i_1}, \ldots, v_{i_{l_i}}\}$ aus E_{λ_i}. Dann gilt $f(v_{i_j}) = \lambda_i v_{i_j}$ für $j = 1, \ldots, l_i$.
Es folgt, dass

$$\{v_{1_1}, \ldots, v_{1_{l_1}}, v_{2_1}, \ldots, v_{2_{l_2}}, \ldots, v_{r_1}, \ldots, v_{r_{l_r}}\}$$

eine Basis aus Eigenvektoren von V ist, da $E_{\lambda_i} \cap E_{\lambda_j} = \{0\}$, falls $i \neq j$. Denn, seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f und sei $\{v_{i_j},\, j = 1, \ldots, g(\lambda_i)\}$ eine Basis von E_{λ_i}, $i = 1, \ldots, r$. So sind nach Satz 13.6 die v_{i_j} linear unabhängig und die Gesamtdimension ist n.

„*(c)* $\Longrightarrow$ *(b)*“
Aus (c) folgt, dass

$$\sum_{i=1}^{r} g(\lambda_i) = n,$$
$$\sum_{i=1}^{r} g(\lambda_i) \leq \sum_{i=1}^{r} a(\lambda_i) = \text{grad}\,(P_f) = n\,,$$

also zerfällt P_f in Linearfaktoren und $a(\lambda_i) = g(\lambda_i), i = 1, \ldots, r$.

„*(b)* $\Longrightarrow$ *(c)*“ folgt trivialerweise.

□

Beispiel 13.13 Sei $f = f_A$ mit $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix}$.

$$\begin{aligned} P_A(x) &= \det \begin{bmatrix} x & -1 & 0 \\ 1 & x-2 & 0 \\ 1 & -1 & x-1 \end{bmatrix} \\ &= (x-1)(x(x-2)+1) \\ &= (x-1)(x^2-2x+1) \\ &= (x-1)^3 \end{aligned}$$

Damit zerfällt P_A in Linearfaktoren. Wir müssen noch $g(1)$ und $a(1)$ bestimmen. Natürlich ist $a(1) = 3$.

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

hat Rang 1, also $\dim(\text{Kern}\,(B)) = 2$ und damit $g(1) < a(1)$. f ist nicht diagonalisierbar.

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

A hat damit 3 verschiedene Eigenwerte und ist diagonalisierbar.

Als Abschluss dieses Kapitels wollen wir ein paar wichtige Eigenschaften von Eigenwerten betrachten.

Satz 13.14 *Seien $A, B \in K^{n,n}$.*

(i) $A^\top$ hat dieselben Eigenwerte wie A.

(ii) A hat den Eigenwert 0 genau dann, wenn A singulär ist.

(iii) AB und BA haben dieselben Eigenwerte.

Beweis:

(i) $\det(A - \lambda I) = \det(A - \lambda I)^\top = \det(A^\top - \lambda I)$.

(ii) $\det(A - 0I) = \det A$.

(iii) Da $\det(AB) = \det A \cdot \det B = \det(BA)$ ist, so ist 0 Eigenwert von AB genau dann, wenn 0 Eigenwert von BA ist.

Sei $\lambda \neq 0$ ein Eigenwert von AB, d.h., $ABv = \lambda v$ für $v \neq 0$.
Setze $z = Bv \Longrightarrow Az = \lambda v \neq 0 \Longrightarrow z \neq 0$.
$BAz = BABv = B\lambda v = \lambda Bv = \lambda z$.
Also ist $z \neq 0$ Eigenvektor von BA mit Eigenwert λ und damit λ auch Eigenwert von BA.

Das gleiche Argument für Eigenwert $\mu \neq 0$ von BA liefert dann, dass jeder Eigenwert von BA auch Eigenwert von AB ist. Das komplettiert den Beweis.

□

Zwar kann man nicht jede komplexe Matrix diagonalisieren, aber jede Matrix ist ähnlich (sogar unitär-ähnlich) zu einer Dreiecksmatrix.

Satz 13.15 ***<u>Satz von Schur</u>***
Sei $A \in \mathbb{C}^{n,n}$, so gibt es eine unitäre Matrix $U \in \mathbb{C}^{n,n}(U^H U = I)$, so dass

$$U^{-1}AU = U^H AU = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \end{bmatrix} \in \mathbb{C}^{n,n}$$

Beweis: Per Induktion.

I.A.: $n = 1$ klar.

I.V.: Die Behauptung sei richtig für $A \in \mathbb{C}^{n-1,n-1}$.

I.S.: Sei $A \in \mathbb{C}^{n,n}$ und sei $v \in \mathbb{C}^n \setminus \{0\}$ Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$. Dann existiert eine unitäre Matrix U_0, so dass $U_0^H v = e_1 \cdot \|v\|_2$, z.B. mit Spiegelung.

$$\begin{array}{rrcl} & Av & = & \lambda v \\ \Longrightarrow & \underbrace{U_0^H A U_0}_{B=[b_{ij}]} \underbrace{U_0^H v}_{e_1\|v\|_2} & = & \lambda \underbrace{U_0^H v}_{e_1\|v\|_2} \\ \Longrightarrow & \begin{bmatrix} b_{11} \\ \vdots \\ b_{n1} \end{bmatrix} \|v\|_2 & = & \lambda e_1 \|v\|_2 \\ \Longrightarrow & b_{21} = \cdots = b_{n1} = 0,\ b_{11} = \lambda\,. & & \end{array}$$

Also $U_0^H A U_0 = B = \left[\begin{array}{c|c} \lambda & * \\ \hline 0 & B_1 \end{array}\right]$. Nach I.V. gibt es U_1 unitär, so dass

$$U_1^H B_1 U_1 = B_2 \quad \text{obere Dreiecksmatrix.}$$

Setze $U = U_0 \left[\begin{array}{c|c} 1 & \\ \hline & U_1 \end{array}\right]$, so folgt

$$\begin{aligned} U^H A U &= \begin{bmatrix} 1 & 0 \\ 0 & U_1^H \end{bmatrix} B \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} \\ &= \left[\begin{array}{c|c} \lambda & * \\ \hline & B_2 \end{array}\right] \qquad \text{obere Dreiecksmatrix .} \end{aligned}$$

□

Dieser Satz gilt im Reellen nicht, weil es nicht unbedingt reelle Eigenwerte geben muss.

Beispiel 13.16 Für $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ gibt es keine reelle orthogonale Matrix U, so dass $U^\top A U = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}$, denn

$$U \text{ hätte die Form } \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \quad \text{oder} \quad \begin{bmatrix} c & s \\ s & -c \end{bmatrix}, \quad (c = \cos\varphi, s = \sin\varphi)$$

$$\Longrightarrow \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} -s & c \\ -c & -s \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

oder

$$\Longrightarrow \begin{bmatrix} c & s \\ s & -c \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} c & s \\ s & -c \end{bmatrix} = \begin{bmatrix} c & s \\ s & -c \end{bmatrix} \begin{bmatrix} s & -c \\ -c & -s \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Anderes Argument:

$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} = \lambda^2 + 1$ hat Nullstellen $-i, i$, aber $U^\top A U$ ist reell, also sind r_{11}, r_{22} reell. Das ist ein Widerspruch.

Kapitel 14

Nilpotente Matrizen / Endomorphismen

Definition 14.1 *Ein Endomorphismus $f : V \to V$ (V Vektorraum über Körper K) heißt* nilpotent, *falls $f^j = \underbrace{f \circ f \circ \ldots \circ f}_{j\text{-mal}} \equiv 0$ für ein $j \in \mathbb{N}$.*

Das kleinste j, für das $f^j \equiv 0$, heißt Nilpotenzindex.

Ist F eine Matrixdarstellung von f, so heißt F nilpotent, *falls $F^j = 0$ für $j \in \mathbb{N}$.*

Beispiel 14.2

$$F = \begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix} \in K^{n,n} \quad \text{ist nilpotent vom Nilpotenzindex } n.$$

$$F^i = \begin{array}{c} \left[\begin{array}{c|c} 0 & \begin{matrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{matrix} \\ \hline \underbrace{\quad 0 \quad}_{i} & 0 \end{array}\right] \end{array} \begin{matrix} \\ \\ \}\, i \end{matrix}, \qquad F^{n-1} = \begin{bmatrix} & & 1 \\ & 0 & \\ & & \end{bmatrix}.$$

Lemma 14.3

(a) *Sei $A \in K^{n,n}$ obere Dreiecksmatrix, so ist A nilpotent genau dann, wenn alle Diagonalelemente 0 sind. Der Nilpotenzindex ist höchstens n, kann aber kleiner sein.*

(b) *$A \in K^{n,n}$ ist nilpotent und diagonalisierbar genau dann, wenn $A = 0$.*

(c) *Ist $A \in K^{n,n}$ nilpotent, so hat A nur den Eigenwert 0.*

Beweis:

(a) Für eine Dreiecksmatrix A gilt stets

$$A^j = \begin{bmatrix} a_{11}^j & & * \\ & \ddots & \\ 0 & & a_{nn}^j \end{bmatrix}.$$

Also kann A nur nilpotent sein, falls alle $a_{ii} = 0$ sind, $i = 1, \cdots, n$.
Umgekehrt, ist A strikte obere Dreiecksmatrix, so folgt per Induktion, dass A^j die Form

$$\left[\begin{array}{c|c} 0 & \\ \hline \underbrace{\;0\;}_{j} & 0 \end{array}\right]\left.\vphantom{\begin{array}{c}0\\0\end{array}}\right\} j$$

hat, und damit $A^n = 0$.

(b) $A = 0 \implies A$ ist nilpotent vom Index 1 und diagonalisierbar.
Sei A diagonalisierbar und nilpotent, dann gibt es eine invertierbare Matrix P mit

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

und da

$$(P^{-1}AP)^j = P^{-1}(A^j)P = \begin{bmatrix} \lambda_1^j & & \\ & \ddots & \\ & & \lambda_n^j \end{bmatrix},$$

so folgt aus $A^j = 0$, dass $\begin{bmatrix} \lambda_1^j & & \\ & \ddots & \\ & & \lambda_n^j \end{bmatrix} = 0 \implies \lambda_k = 0$ für alle $k = 1, \ldots, n$.
$\implies A = P0P^{-1} = 0$.

(c) Angenommen, es gäbe einen anderen Eigenwert $\lambda \neq 0$ und zugehörigen Vektor $v \neq 0$, mit $Av = \lambda v$. Da A nilpotent ist, gibt es ein $j \in \mathbb{N}$ mit

$$\begin{aligned} A^j = 0,\ A^{j-1} \neq 0 \quad &\implies \quad 0 = A^j v = A^{j-1}\lambda v = \lambda^j v \\ &\implies \quad \lambda = 0 \ \text{ oder } \ v = 0. \end{aligned}$$

(Widerspruch) $\square$

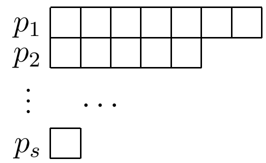
Wie konstruiert man nilpotente Matrizen?

Definition 14.4 *Eine* <u>*Partition*</u> *von n in s Teile ist eine Folge natürlicher Zahlen*

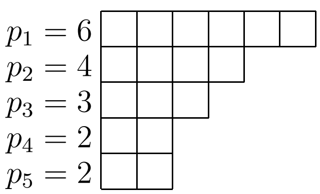
$$p_1 \geq p_2 \geq p_3 \geq \ldots \geq p_s$$

mit $\sum_{i=1}^{s} p_i = n$.

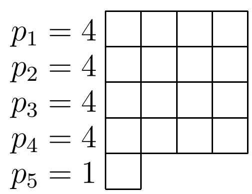
Die Darstellung erfolgt am besten durch Kästchendiagramme:



Beispiel 14.5 $n = 17,\ s = 5.$



oder:



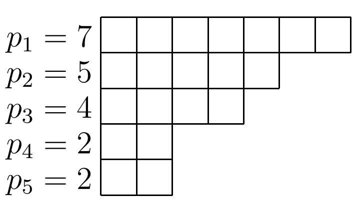
.

Es gibt sehr viele verschiedene Partitionen, z.B. für $n = 50$: 204 226, für $n = 10$: 42.

Definition 14.6 *Ist $p = (p_1, \ldots, p_s)$ eine Partition, so erhält man die zu p* <u>*duale*</u> *Partition, $p^* = (q_1, \ldots, q_t)$, indem man für q_j die Anzahl derjenigen p_i mit $p_i \geq j$ setzt.*

Dann sieht man sofort, dass q_j gerade die Anzahl der Kästchen in der j-ten Spalte ist und außerdem $t = p_1,\ s = q_1,\ (p^*)^* = p$.

Beispiel 14.7 $n = 20, s = 5.$



$$q_1 = 5,\ q_2 = 5,\ q_3 = 3,\ q_4 = 3,\ q_5 = 2,\ q_6 = 1,\ q_7 = 1\,.$$

$$\begin{aligned}
(7)^* &= (1,1,1,1,1,1,1), \\
(7,5)^* &= (2,2,2,2,2,1,1), \\
(5,4)^* &= (2,2,2,2,1), \\
(4,2,2)^* &= (3,3,1,1).
\end{aligned}$$

Zu jeder Partition p ordnen wir eine spezielle nilpotente Matrix zu:

$$N(p) = \left[\begin{array}{cccc} \boxed{\begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{matrix}} & & & \\ & \boxed{\begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{matrix}} & & \\ & & \ddots & \\ & & & \boxed{\begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{matrix}} \end{array}\right] \begin{array}{l} \}p_1 \\ \\ \}p_2 \\ \vdots \\ \}p_s \end{array} =: \begin{bmatrix} N(p_1) & & & \\ & N(p_2) & & \\ & & \ddots & \\ & & & N(p_s) \end{bmatrix}.$$

Satz 14.8

(a) Seien $p, \tilde{p}$ Partitionen von n. Die Matrizen $N(p)$ und $N(\tilde{p})$ sind nur dann ähnlich, wenn $p = \tilde{p}$.

(b) Sei $A \in K^{n,n}$ mit Nilpotenzindex m. Mit $q_i := \mathrm{Rang}\,(A^{i-1}) - \mathrm{Rang}\,(A^i)$ ergibt sich als $q = (q_1, \ldots, q_m)$ eine Partition von n und A ist ähnlich zu $N(p)$ mit $p = q^*$.

Beweis:

(a) Sei $p = (p_1, \ldots, p_s)$ eine Partition von n und $N(p)$ die zugehörige nilpotente Matrix. Da $N(p)$ Nullspalten hat, gibt es Indizes $k_1, \ldots, k_s$:

$$N(p)e_{k_i} = 0, \qquad k_i = \sum_{j=1}^{i-1} p_i + 1.$$

Ansonsten gilt für $j \notin \{k_1, \ldots, k_s\}$: $N(p)e_j = e_{j-1}$. Numeriere die Einheitsvektoren mit Doppelindizes analog zur Partition

$$\begin{array}{lll} e_{11} = e_1, & \ldots, & e_{1,p_1} = e_{p_1} \\ e_{21} = e_{p_1+1}, & \ldots, & e_{2,p_2} = e_{p_1+p_2} \\ \vdots & & \\ e_{s1} = e_{k_s}, & \ldots, & e_{s,p_s} = e_n. \end{array}$$

e_{11}					e_{16}
e_{21}					e_{26}
e_{31}			e_{34}		
e_{41}			e_{44}		
$\vdots$					
e_{k_i}					
$\vdots$					
e_{81}	e_{82}				

Dann gilt für $l \geq 0$

$$\mathrm{Bild}\,(N(p)^l) = \mathcal{L}(\{e_{ij} \mid 1 \leq i \leq s, 1 \leq j \leq p_i - l\}),$$

und da alle e_{ij} linear unabhängig sind, erhalten wir eine Basis von Bild $(N(p)^l)$ und

$$\text{Rang}\,(N(p)^l) = n - \sum_{i=1}^{l} q_i \quad \text{mit } (q_1, \ldots, q_m) = p^* .$$

$\Longrightarrow$ Für $l \geq 1$ gilt

$$\text{Rang}\,(N(p)^{l-1}) - \text{Rang}\,(N(p)^l) = n - \sum_{i=1}^{l-1} q_i - n + \sum_{i=1}^{l} q_i = q_l.$$

Damit bestimmt $N(p)$ eindeutig die Partition $(q_1, \ldots, q_m)$ und damit auch p. Sind $p, \tilde{p}$ Partitionen von n und $N(p) = N(\tilde{p})$ ähnlich, so gilt $\text{Rang}\,(N(p)^l) = \text{Rang}\,(N(\tilde{p})^l)\ \forall l$, also $p^* = \tilde{p}^*$, und damit $p = \tilde{p}$.

(b) Wir müssen zeigen, dass es eine Matrix P gibt, so dass

$$P^{-1}AP = N(p)$$

ist, wobei p die zu q duale Partition von n ist. Sei m der Nilpotenzindex von A. Setze

$$S_i = \text{Kern}\,(A^i).$$

Dann gilt $\{0\} = S_0 \subset S_1 \subset \ldots \subset S_m = K^n$, denn:
Da $A^0 = I_n$, so folgt $S_0 = \{0\}$, und da $A^m = 0$, so folgt $S_m = V$. Sei $0 \leq i < m$ und $v \in S_i$, also

$$A^i v = 0 \implies A^{i+1} v = A(A^i v) = A0 = 0.$$

Also gilt $S_i \subset S_{i+1}$. Beachte, dass alle S_i Unterräume sind. Setze

$$\begin{aligned} q_i \ = \ \dim S_i - \dim S_{i-1} \ &= \ (n - \text{Rang}\,(A^i)) - (n - \text{Rang}\,(A^{i-1})) \\ &= \ \text{Rang}\,(A^{i-1}) - \text{Rang}\,(A^i). \end{aligned}$$

Zeige, dass $(q_1, \ldots, q_m)$ eine Partition von n ist.

Da S_i Unterräume sind, können wir den Quotienten S_i/S_{i-1} bilden, d.h., wir bilden die Menge aller Äquivalenzklassen in S_i bezüglich der Relation für $u, \tilde{u} \in S_i$:

$$u \sim \tilde{u} \quad \Longleftrightarrow \quad u - \tilde{u} \in S_{i-1}.$$

Es folgt, dass S_i/S_{i-1} einen Vektorraum bildet, und dessen Dimension ist natürlich

$$\dim S_i/S_{i-1} = \dim S_i - \dim S_{i-1} = q_i.$$

Wir brauchen nun den folgenden Sachverhalt: Sei eine Abbildung

$$z: \ S_{i+1}/S_i \to S_i/S_{i-1}$$

definiert durch $z(x + S_i) = Ax + S_{i-1}$ für $x \in S_{i+1}$.

Wir zeigen, dass z wohldefiniert, linear und injektiv für $1 \leq i \leq m-1$ ist.

Da z auf Äquivalenzklassen definiert ist, müssen wir zeigen, dass z nicht von der Wahl des Repräsentanten abhängt. Seien $x, y \in S_{i+1}$ mit $x - y \in S_i$. Wir müssen zeigen, dass $Ax \in S_i$ und $Ax - Ay \in S_{i-1}$.

$$x \in S_{i+1} \implies A^{i+1}x = 0 \implies A^i(Ax) = 0 \implies Ax \in \text{Kern}\,(A^i) = S_i,$$

$$A^{i-1}(Ax - Ay) = A^i x - A^i y = A^i(x - y) = 0,$$

denn $x - y \in S_i$. Damit ist z wohldefiniert.

Seien $x, y \in S_{i+1}$ und $\lambda \in K$.

$$\begin{aligned} z((x + S_i) + (y + S_i)) &= z(x + y + S_i) \\ &= A(x + y) + S_{i-1} \\ &= Ax + Ay + S_{i-1} \\ &= (Ax + S_{i-1}) + (Ay + S_{i-1}) \\ &= z(x + S_i) + z(y + S_i), \end{aligned}$$

$$\begin{aligned} z(\lambda(x + S_i)) &= z(\lambda x + S_i) \\ &= A\lambda x + S_{i-1} \\ &= \lambda Ax + S_{i-1} \\ &= \lambda(Ax + S_{i-1}) \\ &= \lambda z(x + S_i). \end{aligned}$$

Also ist z linear. Für die Injektivität sei $x + S_i \in \text{Kern}\,(z)$.

$$\begin{array}{ll} 0 & = z(x + S_i) = Ax + S_{i-1}, \qquad (0 = 0 + S_i \in S_{i+1}/S_i), \\ \Rightarrow & Ax \in S_{i-1} \\ \Rightarrow & A^{i-1}(Ax) = 0 \\ \Rightarrow & A^i x = 0 \\ \Rightarrow & x \in S_i. \end{array}$$

Wegen der Injektivität von z gilt

$$q_{i+1} = \dim S_{i+1}/S_i \leq \dim S_i/S_{i-1} = q_i.$$

Außerdem gilt $S_{m-1} \subsetneqq S_m$, da m der Nilpotenzindex von A ist, also ist $q_m \geq 1$. Also ist $q = (q_1, \ldots, q_m)$ eine Partition, denn

$$q_{i+1} = \dim S_{i+1}/S_i = \dim S_{i+1} - \dim S_i$$

$$\sum_{i=1}^{m} q_i = \sum_{i=1}^{m} (\dim S_i - \dim S_{i-1}) = \dim S_m - \dim S_0 = n.$$

Wir werden nun eine Basis $\{v_{ij} \mid 1 \leq j \leq m, 1 \leq i \leq q_j\}$ konstruieren, so dass

$$\begin{array}{ll} Av_{ij} = v_{i,j-1} & \text{für } \ j > 1, \\ Av_{i1} = 0\,. & \end{array}$$

Dazu verwenden wir Induktion rückwärts: $j = m$.

Sei $\{v_{1,m} + S_{m-1}, \ldots, v_{q_m,m} + S_{m-1}\}$ eine beliebige Basis von S_m/S_{m-1} mit

$$v_{1,m}, \ldots, v_{q_m,m} \in S_m = K^n.$$

Für $1 \leq j < m$ sei $\{v_{1,j+1} + S_j, \ldots, v_{q_{j+1},j+1} + S_j\}$ eine Basis von S_{j+1}/S_j mit

$$v_{i,j+1} \in S_{j+1}.$$

Setze nun $v_{ij} = Av_{i,j+1}, \ 1 \leq i \leq q_{j+1}$.

Wir haben gezeigt, dass $v_{1,j} + S_{j-1}, \ldots, v_{q_{j+1},j} + S_{j-1}$ linear unabhängig sind, und wir ergänzen diese zu einer Basis von S_j/S_{j-1} durch Elemente

$$v_{q_{j+1}+1,j} + S_{j-1}, \ldots, v_{q_j,j} + S_{j-1}, \qquad \text{wobei } v_{q_{j+1}+1,j}, \ldots, v_{q_j,j} \in S_j.$$

Damit haben wir induktiv $v_{1j}, \ldots, v_{q_j,j} \in S_j$ konstruiert, so dass

$$\{v_{1,j} + S_{j-1}, \ldots, v_{q_j,j} + S_{j-1}\}$$

eine Basis von S_j/S_{j-1} bilden. Nach Konstruktion ist

$$\begin{array}{ll} Av_{i,j} = v_{i,j-1} & \text{für } j > 1, \ \text{ und} \\ Av_{i,1} = 0, & \text{denn } v_{i,1} \in S_1 = \text{Kern}\,(A). \end{array}$$

Nach Satz 12.11 (e) folgt, dass die $\{v_{ij}\}$ eine Basis bilden, und damit ist die aus $\{v_{ij}\}$ gebildete Matrix P invertierbar und mit der zu q dualen Partition folgt nach Konstruktion

$$P^{-1}AP = N(p).$$

□

Korollar 14.9 *Sei $A \in K^{n,n}$, dann sind die folgenden Aussagen äquivalent:*

(a) *A ist nilpotent.*

(b) *A ist ähnlich zu $N(p)$ für eine Partition p.*

(c) $\det(\lambda I - A) = \lambda^n$.

(d) $A^n = 0$.

Auf der Menge aller nilpotenten Matrizen bildet Ähnlichkeit eine Äquivalenzrelation und die Repräsentanten der Äquivalenzklassen sind $N(p)$.

Beispiel 14.10 $V = \mathbb{R}^3$.

$$A = \begin{bmatrix} -1 & 1 & -2 \\ -3 & 3 & -6 \\ -1 & 1 & -2 \end{bmatrix} \qquad \boxed{\text{Rang }(A) = 1} \qquad \boxed{\dim(\text{Kern }(A)) = 2}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 3 & 3 & 2 \\ 1 & 1 & 1 \end{bmatrix}}_{P} \underbrace{\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]}_{N(p)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -2 \\ 0 & -1 & 3 \end{bmatrix}}_{P^{-1}}.$$

$A^2 = 0 \;\Rightarrow\; A$ hat den Nilpotenzindex 2.

$$\begin{aligned} S_0 &= \{0\}, \\ S_1 &= \text{ Kern }(A) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right), \\ S_2 &= \text{ Kern }(A^2) = \mathbb{R}^3 = V \end{aligned}$$

Wähle Element von $V \setminus S_1 = \mathbb{R}^3 \setminus S_1$, z. B. $v_{12} = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$,

also ist $\{v_{12} + S_1\}$ eine Basis von V/S_1.

$$\begin{aligned} v_{11} &= Av_{12} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \\ q_2 &= \dim S_2/S_1 = \dim V/S_1 = \dim V - \dim S_1 = 1\,, \\ q_1 &= \dim S_1/S_0 = \dim S_1 - \dim S_0 = 2\,. \end{aligned}$$

Wir ergänzen $v_{11} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ durch $v_{21} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ zu einer Basis von S_1.

$$\begin{array}{c|c|c|} \cline{2-3} p_1 & v_{11} & v_{12} \\ \cline{2-3} p_2 & v_{21} & \multicolumn{1}{c}{} \\ \cline{2-2} \end{array} \qquad \text{Partition,}$$

$$P = [v_{11}, v_{12}, v_{21}] = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 3 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

Kapitel 15

Das charakteristische Polynom und das Minimalpolynom

Wir beginnen dieses Kapitel mit einigen Eigenschaften von Polynomen. Aus dem Seminar ist bereits der Begriff des größten gemeinsamen Teilers (ggT) von Polynomen bekannt. Wir betrachten eine Folgerung aus dem Euklidischen Algorithmus.

Lemma 15.1 *Sei K ein Körper, $p_1, \dots, p_n \in K[x]$ nicht alle Null, $d = \operatorname{ggT}(p_1, \dots, p_n)$. Dann gibt es Polynome $z_1, \dots, z_n \in K[x]$, so dass*

$$d = \sum_{i=1}^{n} p_i z_i.$$

Wie in $\mathbb{Z}$ können wir auch für Polynome den Begriff des kleinsten gemeinsamen Vielfachen (kgV) einführen:

Definition 15.2 *Sei K ein Körper, $p_1, \dots, p_n \in K[x]$, $p_1, \dots, p_n \neq 0$.*
Ein Polynom $m \in K[x], m \neq 0$, heißt <u>kleinstes gemeinsames Vielfaches</u> (kgV) *von $p_1, \dots, p_n$, wenn*

(i) $p_i \mid m,\ i = 1, \dots, n,$

(ii) *aus $p_i \mid m'$, $i = 1, \dots, n$, folgt $m \mid m'$.*

Wir erinnern daran, dass in $K[x]$ alle Elemente $a \in K \setminus \{0\}$ Einheiten sind, d.h., Teiler des Einspolynoms. Wir wiederholen aus dem Seminar den folgenden Begriff.

Definition 15.3 *Ein Polynom $p \in K[x]$, $p \neq 0$, heißt* <u>irreduzibel</u> *(oder* <u>unzerlegbar</u>*), wenn p nicht Einheit ist und in jeder Zerlegung $p = q \cdot z$ einer der beiden Faktoren $q, z \in K[x]$ Einheit ist, d.h., p hat keine echten Teiler.*

Lemma 15.4 *Sei $p \in K[x]$ ein irreduzibles Polynom, und für zwei Polynome $q_1, q_2 \in K[x]$ gelte $p \mid q_1 \cdot q_2$. Dann folgt $p \mid q_1$ oder $p \mid q_2$.*

Beweis: Sei $p \nmid q_1$. Dann ist $d = \text{ggT}\,(p, q_1)$ Einheit, denn sonst wäre d ein echter Teiler von p. Wegen Lemma 15.1 gibt es Polynome z, z_1, so dass

$$p \cdot z + q_1 \cdot z_1 = 1.$$

Daraus folgt

$$p \cdot z \cdot q_2 + q_1 \cdot q_2 \cdot z_1 = q_2.$$

Da $p \mid q_1 \cdot q_2$, ist p ein Teiler der linken Seite dieser Gleichung, und somit $p \mid q_2$. □

Aus einer Äquivalenzklasse zueinander assoziierter Polynome wollen wir im weiteren der Einfachheit halber jeweils nur einen Vertreter betrachten, nämlich das Polynom, das als höchsten Koeffizienten die 1 hat. Wir nennen es <u>normiert</u>.

Satz 15.5 *Für jedes normierte Polynom $p \in K[x]$, $p \neq 1$, gibt es eine eindeutige Zerlegung in irreduzible normierte Polynome.*

Beweis: Die Existenz beweisen wir an dieser Stelle nicht. Für die Eindeutigkeit nehmen wir an, es gibt ein Polynom p mit nicht eindeutiger Zerlegung, d.h., es existieren irreduzible normierte Polynome $q_1, \dots, q_r$ und $z_1, \dots, z_s$, so dass

$$p = q_1 \cdot \ldots \cdot q_r = z_1 \cdot \ldots \cdot z_s\,. \tag{15.6}$$

Wir führen nun eine Induktion über s durch.

I.A.: $s = 1$. Das heißt $z_1 = q_1 \cdot \ldots \cdot q_r$ und daraus folgt $r = 1$ und $z_1 = q_1$, denn sonst wäre z_1 kein irreduzibles Polynom.

I.V.: Für $s = k$ sei bewiesen, dass $r = k$ und $z_i = q_i,\ i = 1, \dots, k$.

I.S.: Aus (15.6) und Lemma 15.4 folgt, dass $q_r \mid z_i$ ist für ein $i \in \{1, \dots, k+1\}$. Da z_i irreduzibel ist, muss deshalb $q_r = z_i$ gelten. Damit ist (15.6) äquivalent zu

$$q_1 \cdot \ldots \cdot q_{r-1} = z_1 \cdot \ldots \cdot z_{i-1} \cdot z_{i+1} \cdot \ldots \cdot z_{k+1}.$$

Nach I.V. erhalten wir die Behauptung.

□

Beispiel 15.7 $K = \mathbb{R}$. *Das Polynom*

$$p(x) = (x^2 + 1)(x + 1) = x^3 + x^2 + x + 1$$

hat die unzerlegbaren Faktoren $(x^2 + 1)$, $(x + 1)$.

Falls $K = \mathbb{C}$, so hat $p(x)$ die unzerlegbaren Faktoren $(x - i)$, $(x + i)$, $(x + 1)$.

Wir wissen aus dem Satz von Cayley/Hamilton, dass für

$$P_A(x) = \det(xI - A) \quad \text{für jedes } A \in K^{n,n}$$

gilt:

$$P_A(A) = 0.$$

Definition 15.8 *Sei $m_A(x)$ das normierte Polynom (höchster Koeffizient 1) kleinsten Grades, für welches*

$$m_A(A) = 0,$$

so heißt $m_A(x)$ das Minimalpolynom *zu A.*

Es ist klar, dass

$$\text{Grad } m_A(x) \leq \text{ Grad } P_A(x).$$

Lemma 15.9 *Für das Minimalpolynom $m_A(x)$ und das charakteristische Polynom $P_A(x)$ einer Matrix A gilt: $m_A \mid P_A$.*

Beweis: Aufgrund des Divisionsalgorithmus mit Rest gibt es Polynome $r(x), s(x)$, so dass

$$P_A(x) = r(x)m_A(x) + s(x),$$

wobei entweder $s(x) = 0$ oder $\deg(s(x) < \deg(m_a(x))$. Angenommen, es wäre $s(x) \neq 0$, dann wäre

$$s(A) = P_A(A) - r(A)m_A(A) = 0.$$

Das ist ein Widerspruch dazu, dass m_A Minimalpolynom ist. Also ist $s(x) = 0$ und folglich $m_A \mid P_A$. □

Bemerkung 15.10 *Analog kann man für jedes Polynom $q(x)$ mit $q(A) = 0$ zeigen, dass $m_A \mid q$.*

Lemma 15.11 *Das Minimalpolynom einer Matrix A ist eindeutig bestimmt.*

Beweis: Sei $p \neq m_A$ ein normiertes Polynom mit Koeffizienten in K, für das $p(A) = 0$ und p hat den gleichen Grad wie m_A.
Setze $q = p - m_A$, so gilt

$$q(A) = p(A) - m_A(A) = 0.$$

Da aber p und m_A den gleichen größten Koeffizienten haben, ist der Grad von q kleiner als der von m_A. (Widerspruch) □

Satz 15.12 *Sei $A \in K^{n,n}$, so teilt $P_A(x)$ das Polynom m_A^n.*

Beweis: Setze

$$\begin{aligned} P_A(x) &= \sum_{i=0}^{n} \alpha_i x^i, \quad \alpha_n = 1\,, \\ m_A(x) &= \sum_{i=0}^{r} \beta_i x^i, \quad \beta_r = 1\,. \end{aligned}$$

Betrachte die Matrizenfolge

$$\begin{aligned} B_r &= I, \\ B_{r-1} &= A + \beta_{r-1} I, \\ &\vdots \\ B_1 &= A^{r-1} + \beta_{r-1} A^{r-2} + \ldots + \beta_1 I, \end{aligned}$$

d.h.,

$$\begin{aligned} B_r &= I, \\ B_{i-1} &= \beta_{i-1} I + A B_i, \qquad i = r, \ldots, 2. \end{aligned}$$

So gilt

$$\begin{aligned} -AB_1 &= -A^r - A^{r-1}\beta_{r-1} - \ldots - A\beta_1 \\ &= \beta_0 I - \underbrace{m_A(A)}_{=0} = \beta_0 I\,. \end{aligned}$$

Setze $B(x) = x^{r-1} B_r + x^{r-2} B_{r-1} + \ldots + x^1 B_2 + B_1$ (Matrixpolynom), so folgt

$$\begin{aligned} (xI - A)B(x) &= x^r B_r + x^{r-1} B_{r-1} + \ldots + x^2 B_2 + x B_1 \\ &\qquad - x^{r-1} A B_r - \ldots - x^2 A B_3 - x A B_2 - A B_1 \\ &= x^r I + \beta_{r-1} x^{r-1} I + \ldots + \beta_2 x^2 I + \beta_1 x I + \beta_0 I \\ &= m_A(x) I\,. \end{aligned}$$

$$\begin{aligned} &\Longrightarrow \det(xI - A) \cdot \det B(x) = m_A(x)^n \\ &\Longrightarrow P_A(x) \text{ teilt } m_A(x)^n. \end{aligned}$$

□

Definition 15.13 *Sei $p(x)$ ein Polynom mit Koeffizienten in K. Jeder Teiler $q(x)$ von $p(x)$, für den es keine echten Teiler mit Koeffizienten in K gibt, heißt* <u>irreduzibler Faktor</u> *in p.*

Korollar 15.14 *Sei $A \in K^{n,n}$, so haben $P_A(x)$ und $m_A(x)$ dieselben irreduziblen Faktoren.*

Beweis: Sei $z(x)$ irreduzibler Faktor von $m_A(x)$. Da m_A P_A teilt und z Teiler von m_A ist, folgt, dass z auch P_A teilt.
Sei $z(x)$ irreduzibler Faktor von $P_A(x)$, so gilt, dass $z(x)$ Teiler von $(m_A)^n$ ist. Also folgt aus der Irreduzibilität und Satz 15.5, dass $z(x)$ auch m_A teilt. □

Korollar 15.15 *Das Minimalpolynom der Matrix*

$$J(\lambda, n) := \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} = \lambda I_n + N(p) \in K^{n,n},\ p = (n),$$

ist gleich dem charakteristischen Polynom und hat die Form

$$P_{J(\lambda,n)}(x) = m_{J(\lambda,n)}(x) = (x - \lambda)^n.$$

Beweis: Das charakteristische Polynom hat per Definition diese Form.
$m_{J(\lambda,n)}(x)$ teilt $P_{J(\lambda,n)}(x)$, also $m_{J(\lambda,n)}(x) = (x - \lambda)^j$ (wegen Satz 15.5).
Nach Definition des Minimalpolynoms gilt

$$\begin{aligned} &0 = m_{J(\lambda,n)}(J(\lambda, n)), \qquad \text{aber} \\ &m_{J(\lambda,n)}(J(\lambda, n)) = (J(\lambda, n) - \lambda I)^j = N(n)^j \end{aligned}$$

Da aber $N(n)$ nilpotent vom Index n ist, folgt $j = n$. □

Definition 15.16 *Seien $A_1, \ldots, A_k$ quadratische Matrizen mit $A_i \in K^{p_i,p_i}$. So bezeichnen wir mit $\bigoplus_{i=1}^{k} A_i$ die Matrix*

$$\bigoplus_{i=1}^{k} A_i := \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix}.$$

Dies heißt <u>direkte Summe</u> von Matrizen (und entspricht natürlich einer direkten Summe von Vektorräumen).
Eine Matrix der Form

$$A = \bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il}) =: \bigoplus_{i=1}^{k} J(\lambda_i, (p_i))$$

mit den Partitionen $p_i = (p_{i1}, \ldots, p_{i,s_i})$, $i = 1, \ldots, k$, heißt <u>Jordan'sche Normalform</u>.

Beispiel 15.17

$$A = \left[\begin{array}{cc|ccc} J(1,5) & & & & \\ & J(1,3) & & & \\ \hline & & J(2,3) & & \\ & & & J(2,2) & \\ & & & & J(2,1) \end{array}\right] = J(1,(5,3)) \bigoplus J(2,(3,2,1))$$

$$= \left[\begin{array}{cccccccc|cccccc}
1 & 1 & & & & & & & & & & & & \\
 & 1 & 1 & & & & & & & & & & & \\
 & & 1 & 1 & & & & & & & & & & \\
 & & & 1 & 1 & & & & & & & & & \\
 & & & & 1 & & & & & & & & & \\
 & & & & & 1 & 1 & & & & & & & \\
 & & & & & & 1 & 1 & & & & & & \\
 & & & & & & & 1 & & & & & & \\
\hline
 & & & & & & & & 2 & 1 & & & & \\
 & & & & & & & & & 2 & 1 & & & \\
 & & & & & & & & & & 2 & & & \\
 & & & & & & & & & & & 2 & 1 & \\
 & & & & & & & & & & & & 2 & \\
 & & & & & & & & & & & & & 2
\end{array}\right],$$

$$A = \begin{bmatrix} J(1,1) & & \\ & J(2,1) & \\ & & J(3,1) \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}.$$

Lemma 15.18 *Seien* $A_1, \ldots, A_k$, *mit* $A_i \in K^{k_i,k_i}$, $i = 1, \ldots, k$, *mit den Minimalpolynomen* $m_{A_i}(x)$. *Dann ist für* $A = \bigoplus_{i=1}^{k} A_i$

m_A *das normierte kleinste gemeinsame Vielfache von* $m_{A_1}(x), \ldots, m_{A_k}(x)$.

Beweis: Per Induktion nach k.

I.A.: $k = 1$ ist klar.
$k = 2$:

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \qquad m_A(A) = \begin{bmatrix} m_A(A_1) & 0 \\ 0 & m_A(A_2) \end{bmatrix} = 0$$

$\Longrightarrow$ $m_A(A_1) = 0$ und $m_A(A_2) = 0$.
$\Longrightarrow$ m_{A_1} teilt m_A und m_{A_2} teilt m_A (nach Bemerkung 15.10).
$\Longrightarrow$ m_A ist gemeinsames Vielfaches von m_{A_1} und m_{A_2}.
Sei q ein anderes gemeinsames Vielfaches von m_{A_1}, m_{A_2}.

$$q(A) = \begin{bmatrix} q(A_1) & \\ & q(A_2) \end{bmatrix} = 0\,.$$

m_A ist Minimalpolynom $\Longrightarrow$ m_A teilt q (aufgrund der Bemerkung 15.10). Also ist m_A das normierte kleinste gemeinsame Vielfache von m_{A_1} und m_{A_2}.

I.V.: Sei die Behauptung richtig für $k = l$ Blöcke.

I.S.: Sei $k = l + 1$, so können wir A schreiben als

$$A = \begin{bmatrix} B_1 & \\ & A_{l+1} \end{bmatrix}, \quad \text{mit} \quad B_1 = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_l \end{bmatrix}.$$

Nach I.V. ist m_{B_1} normiertes kgV von $m_{A_1}, \ldots, m_{A_l}$ und nach I.A. ist m_A normiertes kgV von m_{B_1} und $m_{A_{l+1}}$, und damit normiertes kgV von $m_{A_1}, \ldots, m_{A_{l+1}}$.

$\square$

Beispiel 15.19

$$A = \left[\begin{array}{cc|ccc|c|cc} 1 & 1 & & & & & & \\ & 1 & & & & & & \\ \hline & & 1 & 1 & & & & \\ & & & 1 & 1 & & & \\ & & & & 1 & & & \\ \hline & & & & & 1 & & \\ \hline & & & & & & 1 & 1 \\ & & & & & & & 1 \end{array}\right], \qquad P_A(x) = (x-1)^8,$$

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad P_{A_1} = (x-1)^2, \quad m_{A_1} = (x-1)^2\,,$$

$$\text{denn } (A_1 - I_2)^1 \neq 0, \quad \text{aber } (A_1 - I_2)^2 = 0.$$

$$A_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_{A_2} = (x-1)^3, \quad m_{A_2} = (x-1)^3,$$

$$A_3 = \begin{bmatrix} 1 \end{bmatrix}, \qquad P_{A_3} = (x-1), \quad m_{A_3} = (x-1),$$

$$A_4 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad P_{A_4} = (x-1)^2, \quad m_{A_4} = (x-1)^2\,.$$

Normiertes kleinstes gemeinsames Vielfaches von $(x-1)^2, (x-1)^3, (x-1), (x-1)^2$ ist $(x-1)^3$ und das ist tatsächlich das Minimalpolynom, da

$$\begin{aligned} (A - I_8)^1 &\neq 0\,, \\ (A - I_8)^2 &\neq 0\,, \quad . \\ (A - I_8)^3 &= 0 \end{aligned}$$

Satz 15.20 *Sei* $A = \bigoplus_{i=1}^{k} A_i$*, mit* $A_i = \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il})$ *eine Matrix in Jordan'scher Normalform, wobei* $p_i = (p_{i1}, \ldots, p_{i,s_i})$ *für alle* i *eine Partition ist und* $\lambda_i,\ i = 1, \ldots, k,$ *paarweise verschieden sind.*
So ist

$$m_A = \prod_{i=1}^{k} (x - \lambda_i)^{p_{i1}}.$$

Beweis: Nach Lemma 15.18 ist m_A kleinstes gemeinsames Vielfaches von $m_{A_1}, \ldots, m_{A_k}$. Jedes m_{A_i} ist kleinstes gemeinsames Vielfaches von

$$q_{i1} = m_{J(\lambda_i, p_{i1})},\ \ldots,\ q_{i,s_i} = m_{J(\lambda_i, p_{i,s_i})}.$$

Da p_i eine Partition ist, gilt $p_{i1} \geq p_{i2} \geq \ldots \geq p_{i,s_i}$.

$$q_{ij} = m_{J(\lambda_i, p_{ij})} = (x - \lambda_i)^{p_{ij}}.$$

$\Longrightarrow$ Kleinstes gemeinsames Vielfaches von $q_{i1}, \ldots, q_{i,s_i}$ ist q_{i1}.

Also ist m_A kleinstes gemeinsames Vielfaches von $q_{11}, \ldots, q_{k1}$, und da alle λ_i paarweise verschieden sind, folgt

$$m_A = \prod_{i=1}^{k} (x - \lambda_i)^{p_{i1}}.$$

□

Kapitel 16

Die Jordan'sche Normalform

Wir wollen nun einen der zentralen Sätze der Linearen Algebra beweisen.

Satz 16.1

(a) Sei $A \in K^{n,n}$, so ist A ähnlich zu einer Jordan'schen Normalform genau dann, wenn m_A in Linearfaktoren zerfällt.

(b) Seien A und B zwei Jordan'sche Normalformen,

$$\begin{aligned} A &= \bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il}) & \qquad B &= \bigoplus_{i=1}^{r} \bigoplus_{l=1}^{t_i} J(\mu_i, c_{il}) \\ &= \bigoplus_{i=1}^{k} J(\lambda_i, (p_i)), & &= \bigoplus_{i=1}^{r} J(\mu_i, (c_i)), \end{aligned}$$

mit Partionen $p_i = (p_{i1}, \ldots, p_{i,s_i})$, $c_i = (c_{i1}, \ldots, c_{i,t_i})$.

A und B sind ähnlich genau dann, wenn

$$\begin{aligned} & r = k, \\ & \{\lambda_1, \ldots, \lambda_k\} = \{\mu_1, \ldots, \mu_k\}, \\ & p_i = (p_{i1}, \ldots, p_{i,s_i}) = (c_{i1}, \ldots, c_{i,t_i}) = c_i, \qquad i = 1, \ldots, k. \end{aligned}$$

Dazu brauchen wir einige Hilfssätze.

Lemma 16.2 *Sei $A \in K^{n,n}$, und für $\lambda \in K$ sei $q(x) = (x - \lambda)^m$, so dass $q(A) = 0$. Dann ist A ähnlich zu einer Matrix*

$$\bigoplus_{l=1}^{t} J(\lambda, p_l),$$

wobei $(p_1, \ldots, p_t)$ eine Partition von n ist und alle $p_i \leq m$.

Beweis: Sei $B = A - \lambda I$.

$$0 = q(A) = (A - \lambda I)^m = B^m.$$

B ist nilpotent mit Nilpotenzindex $\leq m$. Nach Satz 14.8 gibt es dann eine invertierbare Matrix P, so dass

$$P^{-1}(A - \lambda I)P = N(p) \quad \text{für eine Partition } \; p = (p_1, \ldots, p_t) \text{ von } n \text{ mit } p_i \leq m, \ \forall i.$$

$$\Rightarrow \ P^{-1}AP \ = \ \lambda I_n + N(p) \ = \ \bigoplus_{l=1}^{t} J(\lambda, p_l)$$

□

Definition 16.3 *Sei V ein K-Vektorraum und $f : V \to V$ linear. Ein Unterraum U von V heißt* <u>invarianter Unterraum</u> *von V zu f, falls $f(U) \subset U$, d.h.,*

$$f(u) \in U \quad \forall u \in U.$$

Beispiel 16.4 Sei $A \in K^{n,n}$, λ Eigenwert und v Eigenvektor zu λ, so ist $\mathcal{L}(v)$ invarianter Unterraum zu A, denn

$$A(\mu v) = (\lambda\mu)v \in \mathcal{L}(v), \ \forall \mu \in K.$$

Lemma 16.5 *Sei V endlichdimensionaler K-Vektorraum und $a : V \to V$ linear. Sei $q(x)$ ein Polynom mit Koeffizienten in K und $q(a) = 0$, d.h., für jede Matrixdarstellung A von a gilt $q(A) = 0$. Sei $q(x) = q_1(x) \cdot \ldots \cdot q_r(x)$, wobei alle $q_i(x)$ paarweise teilerfremd sind.*

Sei $V_i = \text{Kern}\,(q_i(a))$, so ist V_i invarianter Unterraum zu a und

$$V = \bigoplus_{i=1}^{r} V_i.$$

Beweis: Setze für $j = 1, \ldots, r$

$$z_j = \prod_{i \neq j} q_i, \ \text{ d.h., } \ q = z_j \cdot q_j.$$

Dann gilt $1 \sim \text{ggT}\,(z_1, \ldots, z_r) = w$, denn: Es gilt $w \mid z_1$, also existiert ein $i \neq 1$, so dass $w \mid q_i$. Es gilt auch $w \mid z_i$, also existiert ein $j \neq i$, so dass $w \mid q_j$. Da q_i und q_j teilerfremd sind, ist w Einheit. Wegen Lemma 15.1 gibt es Polynome $y_1, \ldots, y_r$ mit Koeffizienten in K, so dass

$$\sum_{j=1}^{r} z_j y_j = 1\,,$$

denn 1 ist größter gemeinsamer Teiler von $z_1, \ldots, z_r$.

Also folgt

$$\sum_{j=1}^{r} z_j(a)y_j(a) = 1.$$

Sei $v \in V$ und $v_j = z_j(a)y_j(a)(v)$. Dann gilt

$$\begin{aligned} q_j(a)(v_j) &= q_j(a)z_j(a)y_j(a)(v) \\ &= q(a)y_j(a)(v) = 0, \qquad \text{da } q(a) = 0, \end{aligned}$$

also ist $\underline{v_j \in \operatorname{Kern} q_j(a) = V_j}$.

$$\begin{aligned} v &= \underbrace{\left(\sum_{j=1}^{r} z_j(a)y_j(a)\right)}_{1}(v) \\ &= \sum_{j=1}^{r} (z_j(a)y_j(a)(v)) = \sum_{j=1}^{r} v_j. \end{aligned}$$

Wir haben damit schon gezeigt, dass jedes $v \in V$ als $\sum\limits_{j=1}^{r} v_j$ mit $v_j \in V_j$ geschrieben werden kann. Wir müssen noch zeigen, dass diese Darstellung eindeutig ist. Sei

$$v = \sum_{j=1}^{r} b_j = \sum_{j=1}^{r} v_j, \qquad v_j, b_j \in V_j.$$

Dann gilt

$$\begin{aligned} v_j &= z_j(a)y_j(a)(v) \\ &= z_j(a)y_j(a)\left(\sum_{i=1}^{r} b_i\right) \\ &= z_j(a)y_j(a)(b_j), \quad \text{da } b_i \in \operatorname{Kern}(q_i) \text{ und für } i \neq j: \; q_i \text{ Faktor von } z_j \\ &= \underbrace{\left(\sum_{i=1}^{r} (z_i(a)y_i(a))\right)}_{1}(b_j) = b_j\,. \end{aligned}$$

Also gilt: $V = \bigoplus\limits_{j=1}^{r} V_j\,.$

Wir müssen dann nur noch zeigen, dass V_i invarianter Unterraum ist.
Sei $v \in V_i$. Da $q_i(a) \cdot a = a \cdot q_i(a)$, so folgt

$$q_i(a)(a(v)) = (a \cdot q_i(a))(v) = a(0) = 0 \quad \Rightarrow \quad a(v) \in V_i\,.$$

□

Beweis: (Satz 16.1).

(a) „$\Longleftarrow$“. Sei m_A das Minimalpolynom von A und zerfalle in Linearfaktoren,

$$m_A = \prod_{i=1}^{r} (x - \lambda_i)^{m_i}, \qquad \text{mit } \lambda_i \neq \lambda_j \text{ für } i \neq j\,.$$

Setze $\mu_i(x) = (x - \lambda_i)^{m_i} \Longrightarrow m_A = \prod_{i=1}^{r} \mu_i$, und die μ_i sind paarweise teilerfremd.
Setze $V_i = \text{Kern}\,(\mu_i(f_A))$, wobei $f_A(x) = Ax$.
Nach Lemma 16.5 sind $V_1, \ldots, V_r$ invariante Unterräume zu f_A und

$$V = \bigoplus_{i=1}^{r} V_i.$$

Sei $f_i : V_i \to V_i$ die lineare Abbildung, die durch $f_i(v) = f_A(v)$ für $v \in V_i$ definiert ist, und A_i eine Matrixdarstellung von f_i bzgl. einer Basis von V_i. So ist

$$\begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{bmatrix}$$

eine Matrixdarstellung von f_A bzgl. einer Basis von V.
Es gilt $\mu_i(f_i) = 0$. Dann folgt mit Lemma 16.2, dass

$$A_i \quad \text{ähnlich zu} \quad \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il})$$

und damit folgt diese Richtung.

(a) „$\Longrightarrow$“. A sei ähnlich zu Jordan'scher Normalform.
$\Rightarrow P_A$ zerfällt in Linearfaktoren. $\Rightarrow m_A$ zerfällt in Linearfaktoren, denn $m_A \mid P_A$ (Lemma 15.9).

(b) „$\Longleftarrow$“ ist klar, denn dann ist $A = B$.

(b) „$\Longrightarrow$“ A ähnlich B, und B in Jordan'scher Normalform.
Wir wissen aus der Änlichkeit, dass die charakteristischen Polynome gleich sind, also

$$k = r\,,$$
$$\{\lambda_1, \ldots, \lambda_k\} = \{\mu_1, \ldots, \mu_k\}\,,$$
$$\sum_{l=1}^{s_i} p_{il} = \sum_{l=1}^{t_i} c_{il}\,, \qquad i = 1, \ldots, k\,.$$

Nun folgt aber gerade aus der Tatsache, dass wir eine direkte Summe haben,

$$\begin{array}{rcl} q_{il} & = & \text{Rang}\,(A - \lambda_i I)^{l-1} - \text{Rang}\,(A - \lambda_i I)^{l} \\ d_{il} & = & \text{Rang}\,(B - \lambda_i I)^{l-1} - \text{Rang}\,(B - \lambda_i I)^{l} \end{array},$$

wobei q die duale Partition zu p und d die duale Partition zu c ist. Aber da A und B ähnlich sind, so sind diese Ränge und damit die Partitionen q_i und d_i alle gleich. Folglich sind auch die Partitionen p_i und c_i alle gleich. $\square$

Korollar 16.6

(a) *Jedes $A \in \mathbb{C}^{n,n}$ ist ähnlich zu einer Jordan'schen Normalform.*

(b) *Hat $A \in \mathbb{R}^{n,n}$ nur reelle Eigenwerte, so ist A ähnlich zu einer Jordan'schen Normalform.*

(c) *Ist $A \in K^{n,n}$ diagonalisierbar, so ist A ähnlich zu einer Jordan'schen Normalform.*

(d) *Sei $A \in \mathbb{C}^{n,n}$, so ist $A = D + N$ mit D diagonalisierbar, N nilpotent und $DN = ND$.*

(e) *A ist ähnlich zu $A^\top$.*

Beweis:

(a) Da jedes Polynom über $\mathbb{C}$ in Linearfaktoren zerfällt, so folgt die Behauptung aus Satz 16.1(a).

(b) Falls alle Eigenwerte von A reell sind, so zerfällt m_A als Teiler von P_A in Linearfaktoren, also folgt die Behauptung aus Satz 16.1(a).

(c) A ist ähnlich zu einer Diagonalmatrix. Diese ist eine Jordan'sche Normalform.

(d) Nach (a) ist A ähnlich zu einer Jordan'schen Normalform, d.h.,

$$\begin{aligned}
PAP^{-1} &= \bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il}) \\
&= \bigoplus_{i=1}^{k} (\lambda_i I_{n_i}) + \bigoplus_{i=1}^{k} \left(\bigoplus_{l=1}^{s_i} N(p_{il}) \right), \\
D &= P^{-1} \left(\bigoplus_{i=1}^{k} \lambda_i I_{n_i} \right) P, \\
N &= P^{-1} \left(\bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} N(p_{il}) \right) P,
\end{aligned}$$

(natürlich in der gleichen Ordnung).

Es bleibt zu zeigen, dass $DN = ND$. (Übung)

(e)
$$\begin{aligned}
A &= D + N, \\
A^\top &= D^\top + N^\top = D + N^\top.
\end{aligned}$$

Es ist klar, dass $PAP^{-1} = A^\top \;\Leftrightarrow\; PNP^{-1} = N^\top$.

Weiterhin folgt aus der Gleichheit der Minimal- und charakteristischen Polynome und Satz 16.1(b), dass wir die Behauptung nur für einen Block zeigen müssen.

$$\text{Also } N = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}, \quad N^\top = \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}.$$

Die Behauptung folgt mit

$$P = \begin{bmatrix} & & 1 \\ & \diagup & \\ 1 & & \end{bmatrix}.$$

$\square$

Beispiel 16.7

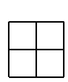
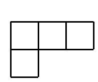
$$\begin{aligned}
A &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & -1 & 2 \end{bmatrix}, \\
P_A &= \det \begin{bmatrix} \lambda & -1 \\ 1 & \lambda - 2 \end{bmatrix} \cdot \det \begin{bmatrix} \lambda & -1 \\ 1 & \lambda - 2 \end{bmatrix} \\
&= (\lambda^2 - 2\lambda + 1)^2 = (\lambda - 1)^4 \,, \\
(A - I)^1 &= \left[\begin{array}{rr|rr} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ \hline -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{array}\right], \\
(A - I)^2 &= \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} = 0 \,, \\
\Longrightarrow m_A &= (\lambda - 1)^2 \,.
\end{aligned}$$

Rang $(A - I_4)^0 = 4$,
Rang $(A - I_4)^1 = 2$.

Alle Eigenwerte sind reell. Nach Korollar 16.6(b) ist A ähnlich zu Jordan'scher Normalform.

$$\begin{aligned}
n = \sum_{l=1}^{s} p_i &= 4 \\
p_1 + p_2 &= 4
\end{aligned}$$

Es bleiben als Partitionen nur

$p = (2, 2)$  oder $p = (3, 1)$  .

Die duale Partition zu p ist $q = p^*$, und diese erhalten wir aus

$$\begin{aligned}
q_j &= \operatorname{Rang}\,(A - 1\,I_4)^{j-1} - \operatorname{Rang}\,(A - 1\,I_4)^j\,, \\
q_1 &= 4 - 2 = 2\,, \\
q_2 &= 2 - 0 = 2\,,
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \qquad & q = p^* = (2,2), \\
\Rightarrow \qquad & p = (2,2)\,.
\end{aligned}$$

Also

$$P^{-1}AP = \left[\begin{array}{cc|cc} 1 & 1 & & \\ & 1 & & \\ \hline & & 1 & 1 \\ & & & 1 \end{array}\right],$$

$$P = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Inhaltsverzeichnis

Index

Numbers written in italic refer to the page where the corresponding entry is described, the ones underlined to the definition, the rest to the pages where the entry is used.