
Lineare Algebra und analytische Geometrie

Benjamin Gufler

Hinweis

Bei dem Text handelt es sich um eine von Benjamin Gufler erstellte persönliche Mitschrift zu den Vorlesungen »Lineare Algebra und Analytische Geometrie I« (Wintersemester 2001/2002) und »Lineare Algebra und Analytische Geometrie II« (Sommersemester 2002) von Prof. Dr. Peter Gritzmann, TU München.

Sie wurde inhaltlich nicht vom Dozenten überprüft und ist daher weder als Vorlesungsausarbeitung noch als Grundlage für Prüfungen oder sonstige Zwecke von diesem autorisiert.

Inhaltsverzeichnis

| | | |
|----------|--------------------------------------------------------------------|-----------|
| 1 | Einführung und Überblick | 1 |
| 1.1 | Lineare Gleichungssysteme | 1 |
| 1.2 | Differenzgleichungen | 7 |
| 2 | Das Eliminationsverfahren von Gauß | 9 |
| 2.1 | Elementare Umformungen & Zeilenstufenform | 9 |
| 2.2 | Lösen von Gleichungssystemen in Zeilenstufenform | 11 |
| 2.3 | Überführen eines Gleichungssystems in Zeilenstufenform | 12 |
| 2.4 | Praktische Aspekte des Eliminationsverfahrens | 15 |
| 3 | Grundbegriffe der Algebra | 17 |
| 3.1 | Gruppen | 17 |
| 3.2 | Ringe und Körper | 22 |
| 3.3 | Polynome | 26 |
| 4 | Vektorräume | 33 |
| 4.1 | (Unter-)Vektorräume und lineare Hülle | 33 |
| 4.2 | Lineare (Un-)Abhängigkeit, Basis, Dimension | 38 |
| 4.3 | Summen von Vektorräumen | 44 |
| 5 | Lineare Abbildungen und Matrizen | 47 |
| 5.1 | Lineare Abbildungen | 47 |
| 5.2 | Die Geometrie linearer Abbildungen und Gleichungssysteme | 50 |
| 5.3 | Lineare Abbildungen und Matrizen | 56 |
| 6 | Determinanten | 67 |
| 6.1 | Einführung | 67 |
| 6.2 | Permutation | 69 |
| 6.3 | Multilineare Abbildungen, Determinantenformen | 72 |
| 6.4 | Determinanten von Endomorphismen und Matrizen | 75 |
| 6.5 | Berechnung von Determinanten | 78 |
| 6.6 | Anwendungen | 80 |
| 7 | Eigenwerte und Normalformen | 87 |
| 7.1 | Einführung | 87 |
| 7.2 | Diagonalisierbarkeit | 88 |
| 7.3 | Die Jordansche Normalform | 94 |
| 7.4 | Anwendungen | 101 |

| | | |
|----------|--------------------------------------------------------------|------------|
| 8 | Euklidische und unitäre Räume | 105 |
| 8.1 | Einführung | 105 |
| 8.2 | Skalarprodukte und hermitesche Formen | 108 |
| 8.3 | Orthogonalität und Orthonormalität | 111 |
| 8.4 | Adjungierte Abbildungen und normale Endomorphismen | 118 |
| 8.5 | Orthogonale und unitäre Abbildungen | 125 |
| 8.6 | Anwendungen | 132 |

Kapitel 1

Einführung und Überblick

1.1 Lineare Gleichungssysteme

Beispiel

In einer Produktionsanlage laufen zwei Produktionsprozesse P_1 und P_2 , die jeweils (beide) die Güter G_1, G_2 erzeugen. Pro Zeiteinheit werden in P_1 zwei Einheiten G_1 produziert und eine Einheit G_2 . Bei P_2 treten -1 Einheiten G_1 und eine Einheit G_2 auf. Man beachte, dass in P_2 G_1 verbraucht wird (entsorgt?).

Wie sind die Produktionsprozesse zu »fahren«, wenn man genau eine Einheit G_1 und fünf Einheiten G_2 produzieren will?

Modellierung:

| | P_1 | P_2 | Bedarf |
|-------|-------|-------|--------|
| G_1 | 2 | -1 | 1 |
| G_2 | 1 | 1 | 5 |

Gleichungssystem:

$$2x_1 - 1x_2 = 1$$

$$1x_1 + 1x_2 = 5$$

- (i) Wie kann man lineare Gleichungssysteme lösen?
→ Eliminationsverfahren von Gauß

- (ii) Wie kann man die Menge aller Lösungen beschreiben?
→ Vektoren, Vektorräume, lineare Unabhängigkeit, Basis, Dimension
Geometrie: Basis, Dimension

- (iii) Was passiert, wenn man Lösungen über verschiedene Zahlenbereiche betrachtet?
→ z.B. $\mathbb{R}, \mathbb{Q}, \mathbb{Z}, \{0, 1\}$
Körper, \mathbb{R}, \mathbb{Q} (Ringe, \mathbb{Z} , Gruppen)
Zahlentheorie: \mathbb{Z}
Kombinatorik: $\{0, 1\}$

Betrachte diese Frage für das Beispielgleichungssystem.

$$\begin{cases} 2x_1 - x_2 = 1 & (i) \\ x_1 + x_2 = 5 & (ii) \end{cases}$$

$$\begin{cases} 2x_1 - x_2 = 1 & (i) \\ -2x_1 - 2x_2 = -10 & (-2)(ii) = (iii) \end{cases}$$

$$\begin{cases} 2x_1 - x_2 = 1 & (i) \\ 0x_1 - 3x_2 = -9 & (i) + (iii) = (iv) \end{cases}$$

$$\begin{cases} 2x_1 - x_2 = 1 & (i) \\ x_2 = 3 & (iv) : (-3) \end{cases}$$

Einsetzen in (i)

$$\begin{aligned} 2x_1 - 3 &= 1 \\ 2x_1 &= 4x_1 = 2 \end{aligned}$$

Warum darf man diese Operationen durchführen? Geht das für $\mathbb{R}/\mathbb{Q}/\mathbb{Z}$?

→ In \mathbb{Z} darf man i.a. nicht dividieren.

Frage: Wie kann man eine »optimale« Eliminationsreihenfolge finden? Und was heißt optimal?

→ Numerik: »Strukturerkennung«, theoretische Informatik, konkrete Optimierung: »Komplexität«

Geometrische Interpretation: Was für Mengen sind

$$\begin{aligned} L_1 &= \{(x_1, x_2) : x_1, x_2 \in \mathbb{R} \wedge 2x_1 - x_2 = 1\} \\ L_2 &= \{(x_1, x_2) : x_1, x_2 \in \mathbb{R} \wedge x_1 + x_2 = 5\} \end{aligned}$$

Beachte:

$$\begin{aligned} 2x_1 - 1 - x_2 = 1 &\Leftrightarrow x_2 = -1 + 2x_1 \\ x_1 + x_2 = 5 &\Leftrightarrow x_2 = 5 - x_1 \end{aligned}$$

Wir können L_1 und L_2 als Graphen von linearen Abbildungen auffassen: $x_1 \mapsto -1 + 2x_1$; $x_1 \mapsto 5 - x_1$.

Lösung: Schnitt $L_1 \cap L_2$ der beiden zugehörigen Geraden.

$$(i), (ii) \xrightarrow{\text{Umformung}} \begin{cases} x_1 = 2 \\ x_2 = 3 \end{cases}$$

Beachte: Die Operationen verändern die Geraden, lassen aber die Lösungsmenge invariant. Die Auflösung $x_1 = 2, x_2 = 3$ entspricht den Geraden parallel zu den Achsen.

Diese Interpretation bezog sich auf die Zeilen des 2×2 -Gleichungssystems.

Eine weitere fundamentale Interpretation bezieht sich auf die Spalten. Dann schreiben wir das System in der Form

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} -1 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ sind Vektoren des \mathbb{R}^2 .

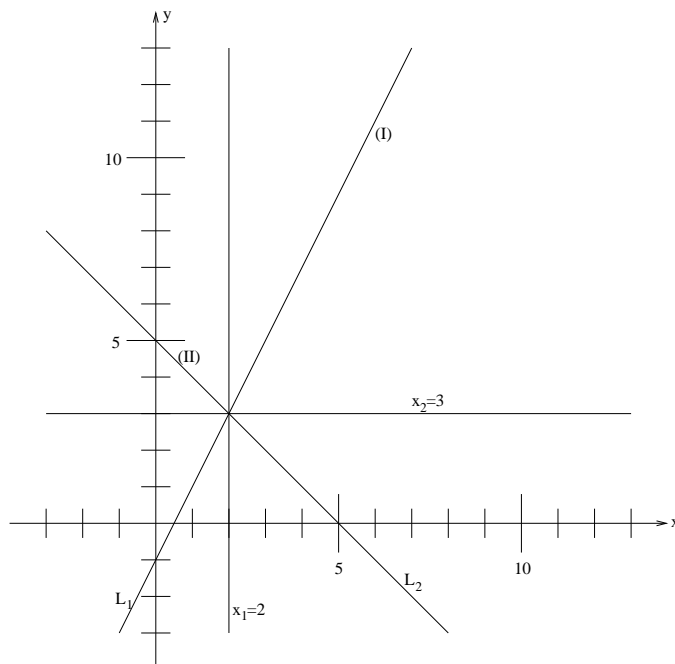


Abbildung 1.1: Geometrische Interpretation eines Gleichungssystems

Definition

Sei $n \in \mathbb{N}$. Dann sei $\mathbb{R}^n = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} ; v_1 \in \mathbb{R} \wedge \dots \wedge v_n \in \mathbb{R} \right\}$. Die Elemente von \mathbb{R}^n

werden als N-VEKTOREN (kurz Vektoren) bezeichnet. Die Einträge der Vektoren heißen KOMPONENTEN.

Durch $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$ ist für alle Vektoren $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ eine

Addition definiert. $\lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}$ definiert für alle Vektoren $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n, \lambda \in \mathbb{R}$

die Multiplikation eines Vektors mit einem Skalar.

Geometrische Interpretation

\mathbb{R}^1 Zahlengerade, \mathbb{R}^2 Ebene, \mathbb{R}^3 »Anschauungsraum«, analog \mathbb{R}^n . Die Vektoren $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ können als Punkte im \mathbb{R}^n interpretiert werden, wenn man ihre Komponenten als kartesische Koordinaten auffasst (s. Abbildung 1.2).

Lösen des Gleichungssystems heißt jetzt: Finde Skalare x_1, x_2 , so dass die Summe des um den Faktor x_1 verlängerten Vektors $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und des um den Faktor x_2 verlängerten Vektors $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ gerade den Vektor $\begin{pmatrix} 1 \\ 5 \end{pmatrix}$ ergibt.

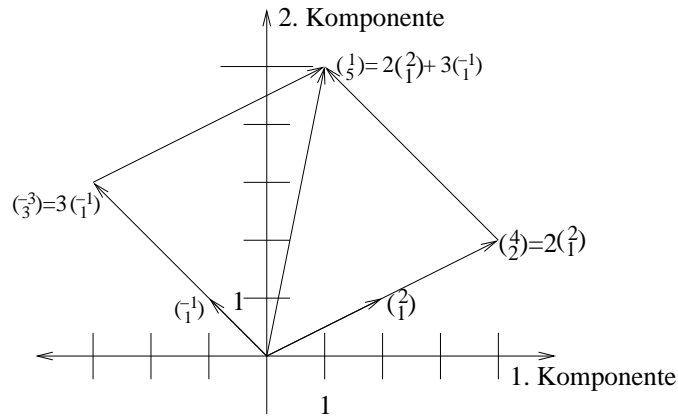


Abbildung 1.2: Interpretation eines Gleichungssystems durch Vektoren

Beispiel:

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ x_1 + x_2 &= 1 \end{aligned}$$

Geometrische Interpretation: s. Abbildung 1.3

Auflösung:

$$\begin{cases} 2x_1 + x_2 + x_3 = 4 & (i) \\ x_1 + x_2 = 1 & (ii) \end{cases}$$

$$\begin{cases} 2x_1 + x_2 + x_3 = 4 & (i) \\ -x_2 + x_3 = 2 & (-2)(ii) - (i) = (iii) \end{cases}$$

$$\begin{cases} 2x_1 + x_2 + x_3 = 4 & (i) \\ x_2 - x_3 = -2 & (-1)(iii) = (iv) \end{cases}$$

»Die« Lösung ist nicht eindeutig bestimmt, sondern eine Gerade im \mathbb{R}^3 . Sie ist durch die Gleichungen (i) und (iv) bestimmt - ebenso wie durch die ursprünglichen Gleichungen (i) und (ii).

Nach der Umformung ist es aber einfach, eine »parametrische« Darstellung als Funktion von \mathbb{R} nach \mathbb{R}^3 anzugeben. Als Parameter wählen wir den Wert der letzten Koordinate, nämlich x_3 und setzen ihn gleich t .

Dann ist $x_3 = t$, aus (iv) ergibt sich $x_2 = -2 + t$, und durch Einsetzen in (i) erhalten wir $x_1 = \frac{1}{2}(4 - x_2 - x_3) = \frac{1}{2}(4 + 2 - t - t) = \frac{1}{2}(6 - 2t) = 3 - t$. Der Lösungsraum

ist also gegeben als $\left\{ \begin{pmatrix} 3-t \\ t-2 \\ t \end{pmatrix} : t \in \mathbb{R} \right\}$ (Darstellung der Lösungsgeraden als »parametrische Kurve« im \mathbb{R}^3).

Natürlich können wir diese Gerade auch in der Form $\begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} t, t \in \mathbb{R}$

schreiben. Wir stellen den Lösungsraum dann dar mittels einer speziellen Lösung und einer Geraden durch den Nullpunkt (s. Abbildung 1.4).

$$\mathbb{R} \rightarrow \mathbb{R}^3; t \mapsto \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} t \quad \text{Parametrisierung der Lösungsgeraden}$$

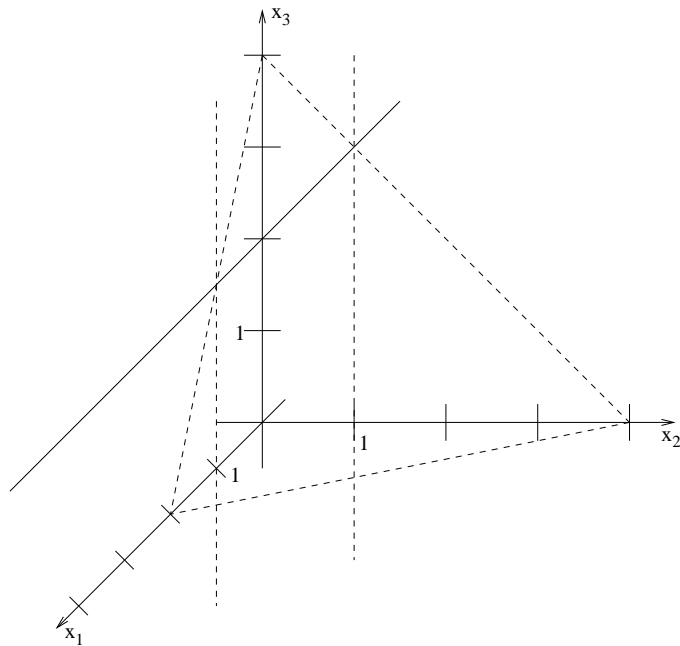


Abbildung 1.3: Lösungsgerade, beschrieben durch zwei Ebenen

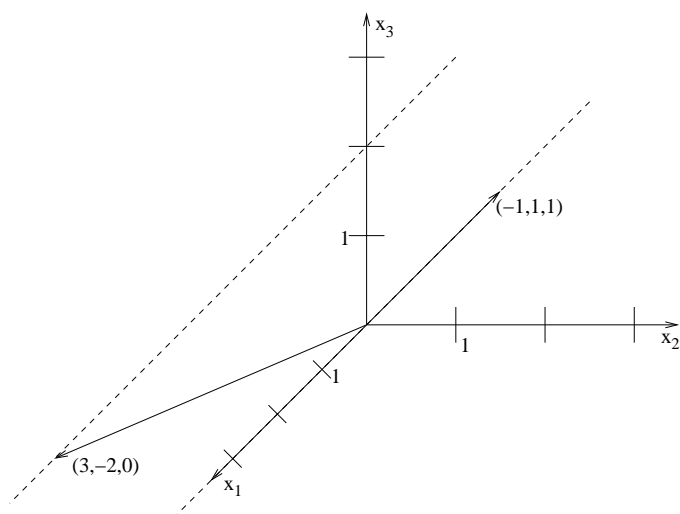


Abbildung 1.4: Darstellung des Lösungsraum durch eine spezielle Lösung und eine Gerade durch den Nullpunkt

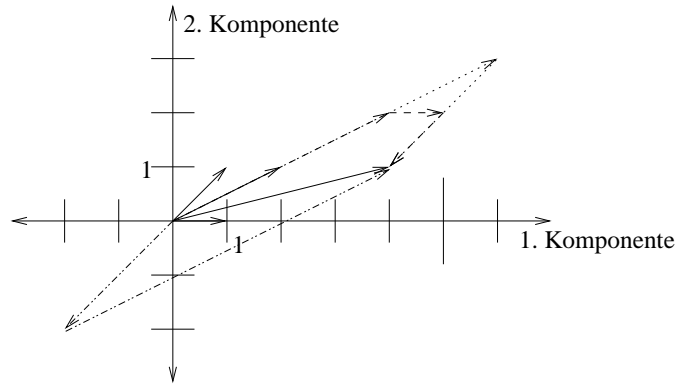


Abbildung 1.5: Interpretation eines Gleichungssystems mittels seiner Spalten

Interpretation des Gleichungssystems mittels seiner Spalten

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} x_2 + \begin{pmatrix} 1 \\ 0 \end{pmatrix} x_3 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

Die Darstellung kann auch als Abbildung aufgefasst werden:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} x_2 + \begin{pmatrix} 1 \\ 0 \end{pmatrix} x_3$$

Interpretiert man dieses Beispiel als Fabrikation, die sich als Kombination von drei Produktionsprozessen ergibt, so beschreibt diese Abbildung die Paare von Quantitäten von zwei Gütern, die auf diese Weise erzeugt werden.

Das Gleichungssystem fragt dann nach dem Urbild von $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$.

Frage: »Ökonomie«: Welche Möglichkeiten (»Lösungsgerade«), den Bedarf $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$ zu erfüllen, soll man nehmen?

Man benötigt zur Beantwortung dieser Frage eine »Bewertung« der verschiedenen Lösungen (»Zielfunktion«) \rightarrow math. Optimierung

Offenbar ist alles Wesentliche dieser Abbildung in der »Tabelle«

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

zusammengefasst.

Definition:

Seien $m, n \in \mathbb{N}$ und für $i = 1, \dots, m$ und $j = 1, \dots, n$ seien $a_{ij} \in \mathbb{R}$. Dann heißt

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1,n-1} & a_{1,n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2,n-1} & a_{2,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{m-1,1} & a_{m-1,2} & a_{m-1,3} & \dots & a_{m-1,n-1} & a_{m-1,n} \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{m,n-1} & a_{m,n} \end{pmatrix} = A$$

$m \times n$ - Matrix. Die Einträge a_{ij} heißen KOMPONENTEN (KOEFFIZIENTEN, EINTRÄGE, ELEMENTE, ...) der Matrix. Die n - Tupel $(a_{i1}, a_{i2}, \dots, a_{in}), i = 1, \dots, m$

heißen ZEILEN, die Vektoren $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, j = 1, \dots, n$ heißen SPALTEN.

$$\text{Ist } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ so ist } Ax := \begin{pmatrix} a_{11}x_1 & \dots & a_{1n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 & \dots & a_{mn}x_n \end{pmatrix}.$$

Beispiel:

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

Bemerkung

Fasst man die Matrix A als Abbildung $x \mapsto Ax$ auf, so kann man (mit Hilfe der entsprechenden Regeln für die Abbildungen) leicht Addition, Multiplikation und andere Operationen auf Matrizen erklären. (später detaillierter)

1.2 Differenzengleichungen

Beispiel: Durchseuchung einer Population

Jeden Monat infizieren sich 10% des vormals gesunden Teils einer Bevölkerung, 20% des vormals infizierten Teils wird geheilt. Erneute Infektion ist möglich, Geburten und Todesfälle werden (zunächst) nicht berücksichtigt. Sei k_0 die Anzahl der zu Beginn der Untersuchung infizierten, g_0 die der zu Beginn gesunden Personen. Für $i \in \mathbb{N}$ bezeichnen k_i, g_i die entsprechenden Zahlen nach dem i -ten Monat. Dann gilt:

$$\begin{aligned} k_1 &= 0,8k_0 + 0,1g_0 \\ g_1 &= 0,2k_0 + 0,9g_0 \end{aligned}$$

Mit $A = \begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,9 \end{pmatrix}$ gilt $\begin{pmatrix} k_1 \\ g_1 \end{pmatrix} = A \begin{pmatrix} k_0 \\ g_0 \end{pmatrix}$. Analog gilt $\begin{pmatrix} k_i \\ g_i \end{pmatrix} = A \begin{pmatrix} k_{i-1} \\ g_{i-1} \end{pmatrix} \forall i \in \mathbb{N}$. Setzt man diese Beziehung sukzessive ein, so folgt $\begin{pmatrix} k_i \\ g_i \end{pmatrix} = \underbrace{A \cdot A \cdot \dots \cdot A}_{i \text{ Mal}} \begin{pmatrix} k_0 \\ g_0 \end{pmatrix}$. Man benutzt

oft die Abkürzung A^i .

Berechne z.B. A^2 :

$$k_2 = 0,8k_1 + 0,1g_1 = 0,8(0,8k_0 + 0,1g_0) + 0,1(0,2k_0 + 0,9g_0) = 0,66k_0 + 0,17g_0$$

g_2 analog:

$$g_2 = 0,34k_0 + 0,83g_0$$

Wir erhalten also $A^2 = \begin{pmatrix} 0,8 \cdot 0,8 + 0,1 \cdot 0,2 & 0,8 \cdot 0,1 + 0,1 \cdot 0,9 \\ 0,2 \cdot 0,8 + 0,9 \cdot 0,2 & 0,2 \cdot 0,1 + 0,9 \cdot 0,9 \end{pmatrix}$.

Aus diesem Grunde werden wir später eine Matrizenmultiplikation wie folgt definieren:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Um die Durchseuchung der Bevölkerung nach i Monaten zu berechnen, muss A^i bestimmt werden. Berechnet man A^i direkt, so ist die Rechnung für großes i aufwändig und (möglicherweise) »numerisch instabil«.

Frage: Kann man durch geeignete Koordinatentransformation A auf eine für die Rechnung besonders einfache Gestalt bringen?

Beispiel:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0,7 \end{pmatrix} \quad D^i = \begin{pmatrix} 1^i & 0 \\ 0 & (0,7)^i \end{pmatrix}$$

Lineare Algebra: Normalformen, Eigenwerte, Eigenvektoren

Numerik: Effiziente, »stabile« Berechnung

In unserem Beispiel gibt es einen (nicht trivialen) Gleichgewichtszustand $k_0 = \frac{1}{3}, g_0 = \frac{2}{3}$.

$$\begin{aligned} k_1 &= \frac{8}{10} \cdot \frac{1}{3} + \frac{1}{10} \cdot \frac{2}{3} = \frac{1}{3} \\ g_1 &= \frac{2}{10} \cdot \frac{1}{3} + \frac{9}{10} \cdot \frac{2}{3} = \frac{2}{3} \\ \begin{pmatrix} k_1 \\ g_1 \end{pmatrix} &= \begin{pmatrix} \frac{8}{10} & \frac{1}{10} \\ \frac{2}{10} & \frac{9}{10} \end{pmatrix} \cdot \begin{pmatrix} k_0 \\ g_0 \end{pmatrix} \end{aligned}$$

Interpretiert man die Prozentzahlen als Wahrscheinlichkeiten, so liegt hier ein spezieller stochastischer Prozess vor. Solche Prozesse werden in der Stochastik und der stochastischen Optimierung untersucht (z.B. Warteschlangen, Internet).

Fazit: Die lineare Algebra bildet zusammen mit der Analysis Grundlage für die gesamte Mathematik.

Kapitel 2

Das Eliminationsverfahren von Gauß

2.1 Elementare Umformungen & Zeilenstufenform

Allgemeines lineares $m \times n$ - Gleichungssystem mit reellen Koeffizienten: $m, n \in \mathbb{N}, a_{ij} \in \mathbb{R} (i = 1, \dots, m; j = 1, \dots, n), b_i \in \mathbb{R} (i = 1, \dots, m)$:

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

In Summationsschreibweise:

$$\begin{aligned} \sum_{j=1}^n a_{1j}x_j &= b_1 \\ &\vdots \\ \sum_{j=1}^n a_{mj}x_j &= b_m \end{aligned}$$

Kurzform:

$$\sum_{j=1}^n a_{ij}x_j = b_i \quad (i = 1, \dots, m)$$

Matrixschreibweise:

$$Ax = b$$

mit:

- A ... Koeffizientenmatrix

- $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$... »rechte Seite«

$$(A, b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \quad \text{»erweiterte Koeffizientenmatrix«}$$

Kurzform:

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}, b = (b_i)_{i=1,\dots,m}$$

Bezeichnung

$\text{Lös}(A, b)$ Lösungsmenge (zulässiger Bereich)

$$\text{Lös}(A, b) := \{x : x \in \mathbb{R}^n \wedge Ax = b\}$$

Jedes Element von $\text{Lös}(A, b)$ heißt Lösung. $Ax = b$ heißt lösbar $\Leftrightarrow \text{Lös}(A, b) \neq \emptyset$.

Ziel

Forme das Gleichungssystem äquivalent um, so dass A Zeilenstufen- oder Treppenform erhält.

Definition

Sei $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$. A heißt in ZEILENSTUFENFORM, wenn folgende Bedingungen erfüllt sind:

- (i) Es gibt ein r mit $r \in \mathbb{N}_0$ und $0 \leq r \leq m$, so dass *jede* der ersten r Zeilen ein von 0 verschiedenes Element enthält und alle Elemente der letzten $m - r$ Zeilen 0 sind.
- (ii) Für $i = 1, \dots, r$ sei $j_i := \{\min j : a_{ij} \neq 0\}$. Dann gilt $j_1 < j_2 < \dots < j_r$ (Stufenbedingung).

Bemerkung

Der Fall $r = 0$ ist zugelassen; dieser entspricht der Nullmatrix $A = 0$.

Bezeichnung

Die Elemente $a_{1j_1}, \dots, a_{rj_r}$ heißen PIVOTELEMENTE.

Beispiele

(a)

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} r = 2, j_1 = 1 < j_2 = 2$$

(b)

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 7 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} m = 3, n = 5, r = 2, j_1 = 2 < j_2 = 4$$

2.1.1 Bemerkung

Durch Umordnung der Spalten, d.h. durch Permutation der Variablen des Gleichungssystems, kann man eine Matrix in Zeilenstufenform stets in eine Gestalt bringen, so dass $j_1 = 1, j_2 = 2, \dots, j_r = r$ gilt (KANONISCHE ZEILENSTUFENFORM).

2.2 Lösen von Gleichungssystemen in Zeilenstufenform

Nach (2.1.1) können wir o.B.d.A. annehmen, dass die erweiterte Koeffizientenmatrix (A, b) die Gestalt

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ 0 & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \dots & a_{rr} & \dots & a_{rn} & b_r \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & b_m \end{pmatrix} \quad (2.1)$$

hat mit $a_{ii} \neq 0$ für $i = 1, \dots, r$.

2.2.1 Satz

(A, b) sei in der Form von (2.1). $\text{Lös}(A, b) \neq \emptyset$ gilt *genau* dann, wenn $b_{r+1} = \dots = b_m = 0$.

Beweis

\Rightarrow Sei $x^* \in \text{Lös}(A, b)$ und $i \in \{r+1, \dots, m\}$. Dann erfüllt x^* insbesondere die i -te Gleichung, und es folgt $b_i = 0x_1^* + 0x_2^* + \dots + 0x_n^* = 0$

\Leftarrow Es gelte $b_i = 0$ für $i = r+1, \dots, m$. Wir fassen x_{r+1}, \dots, x_n als *freie* Variablen auf und bestimmen dann x_1, \dots, x_r als Funktionen der Variablen x_{r+1}, \dots, x_n aus den Gleichungen

$$a_{jj}x_j + \dots + a_{jr}x_r + a_{j,r+1}x_{r+1} + \dots + a_{jn}x_n = b_j$$

für $j = r, r-1, \dots, 1$. Hierzu berechnen wir sukzessive x_{r-l}^* für $l = 0, \dots, r-1$:

$$x_{r-l}^* := \frac{1}{a_{r-l,r-l}} \left(b_{r-l} - \sum_{i=r-l+1}^r a_{r-l,i}x_i^* - \sum_{j=r+1}^n a_{r-l,j}x_j \right)$$

Offenbar liefert jede Setzung von x_{r+1}, \dots, x_n eindeutige Werte x_1^*, \dots, x_r^* für die anderen Variablen x_1, \dots, x_r . Man erhält so eine Lösung.

Bezeichnung

Die Abbildung $\Phi : \mathbb{R}^{n-r} \rightarrow \mathbb{R}^n$, definiert durch

$$\Phi \begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1^* \\ \vdots \\ x_r^* \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix}$$

heißt **PARAMETRISIERUNG** von $\text{Lös}(A, b)$.

Beachte

$\mathbb{R}^0 := \{0\}$, d.h. für $r = n$ gibt es genau eine Lösung, falls Bedingung (2.2.1) erfüllt ist.

Beispiel

$$(A, b) = \begin{pmatrix} 1 & 3 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Es ist $m = 2, n = 4, r = 2$. Es gilt

$$\begin{aligned} x_2^* &= 1 - x_4 \\ x_1^* &= 2 - 3x_2^* - 2x_3 - x_4 \\ &= 2 - 3 + 3x_4 - 2x_3 - x_4 &= -1 + 2x_4 - 2x_3 \end{aligned}$$

Also $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^4$

$$\Phi \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_4 - 2x_3 - 1 \\ 1 - x_4 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix} x_4$$

2.3 Überführen eines Gleichungssystems in Zeilenstufenform**Bezeichnung**

Die folgenden drei Operationen auf einer beliebigen erweiterten Koeffizientenmatrix (A, b) heißt **ELEMENTARE ZEILENUMFORMUNGEN**:

- (I) Vertauschung von zwei Zeilen
- (II) Multiplikation einer Zeile (d.h. jedes Elements der Zeile) mit einer reellen Zahl λ mit $\lambda \neq 0$.
- (III) Addition von zwei Zeilen (komponentenweise)

2.3.1 Satz

Die erweiterte Koeffizientenmatrix (\hat{A}, \hat{b}) gehe aus (A, b) durch endlich viele elementare Zeilenumformungen hervor. Dann gilt $\text{Lös}(\hat{A}, \hat{b}) = \text{Lös}(A, b)$.

Beweis

Es reicht zu zeigen, dass jede einzelne der drei elementaren Umformungen die Lösungsmenge unverändert lässt. Dann gilt das auch für eine beliebige - endliche - Kombination. (Zur Vereinfachung der Bezeichnungen: Ausgangssystem (A, b) , entstehendes (\hat{A}, \hat{b}) .)

(I) Da alle Bedingungen erfüllt sein müssen, ist die Aussage trivial.

(II)

$$\begin{aligned} \sum_{j=1}^n a_{ij} x_j = b_i &\Leftrightarrow_{\lambda \neq 0} \lambda \left(\sum_{j=1}^n a_{ij} x_j \right) = \lambda b_i \\ &\Leftrightarrow \sum_{j=1}^n \underbrace{\lambda a_{ij}}_{\hat{a}_{ij}} x_j = \underbrace{\lambda b_i}_{\hat{b}_i} \end{aligned}$$

(III)

$$\begin{aligned}
& \sum_{j=1}^n a_{ij}x_j = b_i \wedge \sum_{j=1}^n a_{kj}x_j = b_k \\
& \Rightarrow \sum_{j=1}^n a_{ij}x_j = b_i \wedge \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n a_{kj}x_j = b_i + b_k \\
& \Rightarrow \sum_{j=1}^n a_{ij}x_j = b_i \wedge \sum_{j=1}^n \underbrace{(a_{ij} + a_{kj})}_{\hat{a}_{kj}}x_j = \underbrace{b_i + b_k}_{\hat{b}_k} \\
& \stackrel{(II)}{\Rightarrow} \sum_{j=1}^n (-a_{ij})x_j = -b_i \wedge \sum_{j=1}^n (a_{ij} + a_{kj})x_j = b_i + b_k \\
& \Rightarrow \sum_{j=1}^n (-a_{ij})x_j = -b_i \wedge \sum_{j=1}^n ((a_{ij} + a_{kj}) - a_{ij})x_j = (b_i + b_k) - b_i \\
& \stackrel{(II)}{\Rightarrow} \sum_{j=1}^n a_{kj}x_j = b_k \wedge \sum_{j=1}^n a_{kj}x_j = b_k
\end{aligned}$$

2.3.2 Satz

Jedes lineare Gleichungssystem kann durch elementare Zeilenumformungen in Zeilenstufenform gebracht werden.

Beweis

Wir geben einen entsprechenden Algorithmus an.

Ist $A = 0$, so gilt die Behauptung. Sei also im Folgenden $A \neq 0$.

Sei $j_i = \min \{j : j \in \{1, \dots, n\} \wedge \exists i : i \in \{1, \dots, m\} \wedge a_{ij} \neq 0\}$. Durch Anwendung von Operation (I) - falls erforderlich - bringt man A auf die Gestalt

$$\begin{pmatrix} 0 & \dots & 0 & \hat{a}_{1j} & & \\ \vdots & & \vdots & \vdots & & * \\ 0 & \dots & 0 & \hat{a}_{mj} & & \end{pmatrix}$$

mit $\hat{a}_{1j} \neq 0$.

Für jedes $i \in \{2, \dots, m\}$ addieren wir das $-\frac{\hat{a}_{ij}}{\hat{a}_{1j}}$ -fache der ersten Zeile zur i -ten Zeile (Operationen (II) und (III)). Dann erhalten wir ein System der Gestalt

$$\begin{pmatrix} 0 & \dots & 0 & \hat{a}_{1j} & * & \dots & * \\ \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & \vdots & & A2 & \\ 0 & \dots & 0 & 0 & & & \end{pmatrix}$$

Das Verfahren wird nun auf $A2$ fortgesetzt. Sukzessive folgt die Behauptung.

Beispiel

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 2 & 2 & 0 & 5 \\ 0 & 4 & 6 & 0 & 8 \end{pmatrix} \xrightarrow{(II),(III)} \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 & 4 \end{pmatrix}$$

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

Eliminationsverfahren

- Bringe (A, b) mittels elementarer Zeilenoperationen auf Zeilenstufenform (\hat{A}, \hat{b}) .
- Löse dann das System $\hat{A}x = \hat{b}$.

Fragen

- (1) Wie unterscheiden sich verschiedene Parametrisierungen von $\text{Lös}(A, b)$?

$$\Phi : \mathbb{R}^{n-r} \rightarrow \text{Lös}(A, b) \subseteq \mathbb{R}^n$$

- (2) Welche »Struktur« hat die Lösungsmenge?

Bemerkungen zu (1)

- (1) r ist *nicht* Eigenschaft des Eliminationsverfahrens, sondern der Matrix. (r ... Zeilenrang von A)
- (2) Φ ist bijektiv

2.3.3 Bemerkung

Sei (A, b) die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems. Dann kann (A, b) durch elementare Zeilenumformungen auf eine Gestalt gebracht werden, so dass alle Pivotelemente 1 sind und in den Spalten der Pivotelemente, bis auf die Pivotelemente selbst, nur Nullen stehen.

Bemerkung

Die zugehörige kanonische Zeilenstufenform hat die Gestalt

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

Beispiel

$$\begin{pmatrix} 2 & 1 & 1 & 4 \\ 1 & 1 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 1 & 4 \\ 0 & \frac{1}{2} & -\frac{1}{2} & -1 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 2 & 1 & 1 & 4 \\ 0 & 1 & -1 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & 2 & 6 \\ 0 & 1 & -1 & -2 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -2 \end{pmatrix}$$

In dieser Form kann man die Lösungen direkt ablesen:

$$\begin{aligned}x_1 &= 3 - x_3 \\x_2 &= x_3 - 2 \\x_3 &\text{ frei}\end{aligned}$$

2.4 Praktische Aspekte des Eliminationsverfahrens

2.4.1 Bemerkung

Ein lineares Gleichungssystem kann mittels $O(n^2m)$ Rechenoperationen gelöst werden.

Beachte: $O(n^2m)$ asymptotisch, d.h. für $m, n \rightarrow \infty$ wächst der Rechenaufwand wie $\text{const} \cdot m \cdot n^2$.

Im Folgenden beschäftigen wir uns exemplarisch mit dem Effekt von Rundungsfehlern. (Hierzu gibt es umfangreiche Theorien in der Numerik.)

Beispiel

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 + 10^{-k} \end{pmatrix} \quad \text{für ein festes } k \in \mathbb{N}$$

Zeilenstufenform:

$$\hat{A} = \begin{pmatrix} 1 & 1 \\ 0 & 10^{-k} \end{pmatrix}$$

Man sieht, dass \hat{A} »fast« (aber *nur* fast) eine Nullzeile besitzt. Bei exakter Rechnung hat das Gleichungssystem $Ax = b$ für jeden passenden Vektor $b \in \mathbb{R}^2$ eine eindeutige Lösung. Bei Rundung auf weniger als k Nachkommastellen gibt es (nach (2.2.1)) rechte Seiten, so dass $\text{Lös}(A, b) = \emptyset$ ist - z.B. $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ -, aber auch andere, für die $\text{Lös}(A, b)$ eine Gerade ist - z.B. $b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Selbst bei exakter Rechnung hängt die Lösung sehr sensitiv von der rechten Seite ab; für $b = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ hat man die Lösung $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, für $b = \begin{pmatrix} 2 \\ 2 + 10^{-k} \end{pmatrix}$ die Lösung $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Eine Änderung in der k -ten Stelle führt also zu Änderungen in der ersten Vorkommastelle (» A ist schlecht konditioniert«).

Warum ist das so?

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix} \text{ wird gebildet als } 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 1 + 10^{-k} \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 2 + 10^{-k} \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 + 10^{-k} \end{pmatrix}$$

Diese Probleme liegen offenbar an der Koeffizientenmatrix A . Selbst, wenn A »gut konditioniert« ist, kann das Eliminationsverfahren durch »Fehlerfortpflanzung« in große numerische Probleme laufen.

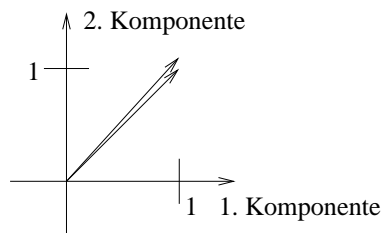


Abbildung 2.1: »fast gleiche Spalten«

Beispiel

$$A = \begin{pmatrix} 10^{-k} & 1 \\ 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

(10^{-k} Pivotelement)

Mit a_{11} als erstem Pivotelement ergibt sich

$$\begin{pmatrix} 10^{-k} & 1 & 1 \\ 0 & 1 - 10^k & 2 - 10^k \end{pmatrix}$$

also

$$x_2 = \frac{2 - 10^k}{1 - 10^k} = 1 - \frac{1}{10^k - 1}$$

Wir benutzen die Gleitkommadarstellung mit $k - 1$ Stellen Genauigkeit, d.h. die Zahlen werden gerundet dargestellt mit einer Vorkomma- und $k - 2$ Nachkommastellen sowie der passenden Zehnerpotenz. Dann sind

$$\begin{aligned} 10^{-k} &\rightarrow 1 \cdot 10^{-k} \\ 1 - 10^k &= \underbrace{-99 \dots 9}_k \rightarrow -1 \cdot 10^k \\ 2 - 10^k &= \underbrace{-99 \dots 98}_{k-1} \rightarrow -1 \cdot 10^k \end{aligned}$$

Die zweite Gleichung ist dann

$$-1 \cdot 10^k x_2 = -1 \cdot 10^k$$

und man erhält $x_2 = 1$.

Verglichen mit dem exakten Wert: kein wesentlicher Rundungsfehler. Aber: Einsetzen zur Bestimmung von x_1 :

- ohne Rundung:

$$\begin{aligned} 10^{-k} x_1 + \left(1 - \frac{1}{10^k - 1}\right) &= 1, \text{ also} \\ x_1 &= \frac{10^k}{10^k - 1} = 1 + \frac{1}{10^k - 1} \approx 1 \end{aligned}$$

- mit Rundung:

$$10^{-k} x_1 + 1 = 1, \text{ also } x_1 = 0$$

Die Rechnung ist also »instabil«.

Fragen

- Welche Pivotwahl ist bezüglich Anfälligkeit für Rundungsfehler geeignet?
(Eine Möglichkeit: Wähle jeweils den größten zur Verfügung stehenden Eintrag: »vollständige Pivotwahl«. Problem: aufwändig)
- Welche Abschätzungen gelten für die Fehlerfortpflanzung?
- Gibt es numerisch bessere Verfahren zur Lösung linearer Gleichungssysteme?
- Kann man auch Ungleichungen berücksichtigen?

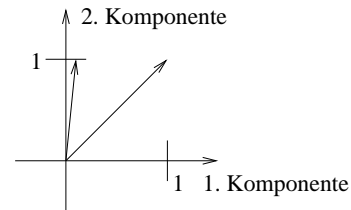


Abbildung 2.2: Spalten von A
»gutartig«

- ➔ Numerik: Rundungsfehler, iterative Verfahren
- ➔ Geometrie: Polyeder
- ➔ Lineare Optimierung: Algorithmen

Kapitel 3

Grundbegriffe der Algebra

Wir betrachten Strukturen im Zusammenhang mit folgenden (und weiteren) Fragen:

- (1) Welche Eigenschaften müssen Zahlensysteme haben, damit über ihnen das Eliminationsverfahren funktioniert?
- (2) Welche Struktur haben Lösungsmengen von linearen Gleichungssystemen?

3.1 Gruppen

Im Folgenden betrachten wir (nichtleere) Mengen G zusammen mit einer Verknüpfung $*$: $G \times G \rightarrow G$. Oft: $g_1 * g_2$ statt $*(g_1, g_2)$.

Beispiele

(a) $(G, *) = (\mathbb{R}, +), (\mathbb{R} \setminus \{0\}, \cdot), (\mathbb{Z}, +), (\mathbb{Z} \setminus \{0\}, \cdot)$

(b) $G = \{0, 1\}$, $*$ gemäß Tabelle:

| | | |
|---|---|---|
| * | 0 | 1 |
| 0 | 0 | 1 |
| 1 | 1 | 0 |

(Addition modulo 2, $(\mathbb{Z}_2, +)$)

(c) $p \in \mathbb{N}$, p Primzahl; $G = \{1, \dots, p-1\}$

$g_1 * g_2 := \min \{r; r \in \mathbb{N}_0 \wedge \exists b : b \in \mathbb{N}_0 \wedge g_1 \cdot g_2 = pb + r\}$ für $g_1, g_2 \in G$

$g_1 * g_2 \dots$ Rest bei Teilung von $g_1 \cdot g_2$ durch p (Multiplikation modulo p):

$(\mathbb{Z}_p \setminus \{0\}, \cdot)$

Beachte: Ist p keine Primzahl, so liegt $g_1 * g_2$ i.a. nicht in G .

(d) $G = \mathbb{N}, g_1 * g_2 = g_1^{g_2}$

(e) Sei X eine (nichtleere) Menge. $G = \text{Abb}(X, X) := \{f; f : X \rightarrow X\}, f * g := f \circ g$

Definition

Sei $(G, *)$ eine Menge mit einer Verknüpfung. $(G, *)$ heißt GRUPPE, wenn gilt:

(G1)

$$\exists e : e \in G \wedge (g \in G \Rightarrow e * g = g)$$

(Existenz (mindestens) eines (links-)NEUTRALEN ELEMENTS)

(G2) Es existiert ein (links-)neutrales Element e mit folgender Eigenschaft:

$$g \in G \Rightarrow \exists \hat{g} : \hat{g} \in G \wedge \hat{g} * g = e$$

(Existenz (links-)INVERSER ELEMENTE)

(G3)

$$g_1, g_2, g_3 \in G \Rightarrow (g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$$

(Assoziativgesetz)

Definition

Sei $(G, *)$ eine Gruppe. Zwei Elemente $g_1, g_2 \in G$ KOMMUTIEREN, wenn $g_1 * g_2 = g_2 * g_1$.

$(G, *)$ heißt kommutativ (oder ABELSCH), wenn je zwei Elemente von G kommutieren, d.h.

$$g_1, g_2 \in G \Rightarrow g_1 * g_2 = g_2 * g_1 \quad (\text{Kommutativgesetz})$$

Beispiele

- (a) $(\mathbb{R}, +)$, $(\mathbb{R} \setminus \{0\}, \cdot)$, $(\mathbb{Z}, +)$ sind kommutative Gruppen.
 $(\mathbb{Z} \setminus \{0\}, \cdot)$: 1 ist einziges neutrales Element, zu 2 gibt es aber *kein* Inverses.
 Assoziativ- und Kommutativgesetz gelten.
- (b) $(\mathbb{Z}_2, +)$: 0 ist neutrales Element, 0 ist invers zu 0 und 1 ist invers zu 1. Ferner gelten Kommutativ- und Assoziativgesetz, wie aus der Tabelle ersichtlich wird:

| g_1 | g_2 | g_3 | $g_1 * g_2$ | $g_2 * g_1$ | $(g_1 * g_2) * g_3$ | $g_2 * g_3$ | $g_1 * (g_2 * g_3)$ |
|-------|-------|-------|-------------|-------------|---------------------|-------------|---------------------|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |

- (c) $(\mathbb{Z}_p \setminus \{0\}, \cdot)$: 1 ist neutrales Element, Assoziativ- und Kommutativgesetze gelten. Der Beweis der Existenz inverser Elemente erfordert elementare Voraussetzungen der Zahlentheorie.

Sei $g \in \{1, \dots, p-1\} =: G$

Es reicht zu zeigen, dass die Abbildung $\varphi_g : G \rightarrow G$, definiert durch $\varphi_g(a) = a \cdot g$ surjektiv ist. Da G endlich ist, reicht es zu zeigen, dass φ_g injektiv ist.

Seien $a_1, a_2 \in G$ und $\varphi_g(a_1) = \varphi_g(a_2) =: r$. Nach Definition gibt es $q_1, q_2 \in \mathbb{N}_0$ mit $a_1 g = q_1 p + r$ und $a_2 g = q_2 p + r$.

Es folgt $(a_1 - a_2)g = (q_1 - q_2)p$. Somit teilt das Produkt $(a_1 - a_2)g$, also $a_1 - a_2$ oder g . Wegen $|a_1 - a_2| < p$ und $1 \leq g < p$ folgt $a_1 - a_2 = 0$.

Wir haben also gezeigt: $\varphi_g(a_1) = \varphi_g(a_2) \Rightarrow a_1 = a_2$, d.h. φ_g ist injektiv.

Folglich ist $(\mathbb{Z}_p \setminus \{0\}, \cdot)$ eine kommutative Gruppe.

- (d) $g_1, g_2 \in \mathbb{N}, g_1 * g_2 = g_1^{g_2}$
 Annahme: $e \in \mathbb{N}$ neutral. Dann gilt insbesondere $e^2 = 2$. Wegen $1^2 = 1$ und $e^2 > 2$ für $e \in \mathbb{N} \setminus \{1\}$ Widerspruch.

- (e) $(\text{Abb}(X, X), \circ)$: $\text{id} := \text{id}_x : X \rightarrow X, \text{id}(x) := x$ ist neutrales Element.
 Inverses Element: Sei $f \in \text{Abb}(X, X)$. Gesucht: $\hat{f} \in \text{Abb}(X, X)$ mit $\hat{f} \circ f = \text{id}$.
 Ist f nicht injektiv, so kann \hat{f} nicht existieren.
 Wegen $(f \circ g) \circ h(x) = (f \circ g)(h(x)) = f(g(h(x)))$ und $f \circ (g \circ h)(x) = f((g \circ h)(x)) = f(g(h(x)))$ gilt das Assoziativgesetz.
 Kommutativ ist \circ i.a. nicht.
Spezialfall: $S(X) := \{f; f \in \text{Abb}(X, X) \wedge f \text{ bijektiv}\}$. Sei $f \in S(X)$. Definiere $\hat{f} : X \rightarrow X$ durch $f(x) \mapsto x$ (wohldefiniert). Es gilt $\hat{f} \circ f = \text{id}$. Also ist $(S(X), \circ)$ eine Gruppe: die SYMMETRISCHE GRUPPE auf X .
 Ist speziell X endlich, so ist $(S(X), \circ)$ die GRUPPE DER PERMUTATIONEN auf X .

Wir ziehen nun Folgerungen aus den Gruppenaxiomen.

3.1.1 Lemma

Seien $(G, *)$ eine Gruppe, e ein (links-)neutrales Element, für das zu jedem Element aus G ein (links-)inverses Element existiert.
 Sind $g, \hat{g} \in G$ mit $\hat{g} * g = e$, so gilt $g * \hat{g} = e$.

Beweis

Sei \hat{g} ein (links-)inverses Element zu g . Dann gilt

$$\begin{aligned} e &= \hat{g} * g = \hat{g} * (e * g) = \hat{g} * ((\hat{g} * g) * g) \\ &= (\hat{g} * \hat{g}) * (g * g) = e * (g * g) = g * g \end{aligned}$$

3.1.2 Lemma

Es gelten die Voraussetzungen von (3.1.1).
 Ist $g \in G$, so gilt $g * e = g$.

Beweis

Sei $\hat{g} \in G$ mit $\hat{g} * g = e$. Nach (3.1.1) gilt $g * \hat{g} = e$, und es folgt

$$g * e = g * (\hat{g} * g) = (g * \hat{g}) * g = e * g = g$$

3.1.3 Lemma

Es gelten die Voraussetzungen von (3.1.1).

- (i) $g_1, g_2, x \in G \wedge x * g_1 = x * g_2 \Rightarrow g_1 = g_2$ (LINKSKÜRZUNGSREGEL)
 (ii) $g_1, g_2, x \in G \wedge g_1 * x = g_2 * x \Rightarrow g_1 = g_2$ (RECHTSKÜRZUNGSREGEL)

Beweise

Sei $\hat{x} \in G$ (links-)invers zu x . Dann gilt

$$\begin{aligned} x * g_1 = x * g_2 &\Rightarrow \hat{x} * (x * g_1) = \hat{x} * (x * g_2) \\ &\Rightarrow (\hat{x} * x) * g_1 = (\hat{x} * x) * g_2 \Rightarrow e * g_1 = e * g_2 \\ &\Rightarrow g_1 = g_2 \end{aligned}$$

Mit Hilfe von (3.1.1) folgt analog

$$g_1 * x = g_2 * x \Rightarrow g_1 * (x * \hat{x}) = g_2 * (x * \hat{x}) \Rightarrow g_1 = g_2$$

3.1.4 Satz

Sei $(G, *)$ eine Gruppe. Dann gilt:

- (i) Es existiert jeweils genau ein Element $e \in G$ bzw. $\bar{e} \in G$ mit $g \in G \Rightarrow e * g = g$ bzw. $g \in G \Rightarrow g * \bar{e} = g$ und es gilt $e = \bar{e}$. (e heißt DAS NEUTRALE ELEMENT.)
- (ii) Zu jedem $g \in G$ existiert jeweils genau ein $\hat{g} \in G$ bzw. $\bar{g} \in G$ mit $\hat{g} * g = e$ bzw. $g * \bar{g} = e$ und es gilt $\hat{g} = \bar{g}$. (\hat{g} ist DAS ZU g INVERSE ELEMENT.)

Beweis

- (i) Seien $e_1, e_2 \in G$ mit $e_1 * g = e_2 * g = g$ für alle g . Aus (3.1.3(ii)) folgt $e_1 = e_2$; wer nennen dieses das eindeutig bestimmte (links-)neutrale Element zu e .
Nach (3.1.2) gilt $g * e = g$ für alle g , und nach (3.1.3(i)) gibt es nur ein solches Element.
- (ii) Analog zu (i), unter Verwendung von (3.1.1) statt (3.1.2).

Bezeichnungen

Oft wird das Zeichen $\gg * \ll$ für die Gruppenoperation einfacher als $\gg \cdot \ll$ geschrieben oder weggelassen. Das neutrale Element wird dann mit $\gg 1 \ll$ bezeichnet (Einselement), das zu einem g gehörige Inverse mit g^{-1} .
Ist $(G, *)$ kommutativ, so wird statt $\gg * \ll$ oft auch $\gg + \ll$ geschrieben, $\gg 0 \ll$ bezeichnet das neutrale Element (Nullelement) und $-g$ das zu g Inverse.

3.1.5 Lemma

Sei (G, \cdot) eine Gruppe, $g \in G$. Dann gilt $(g^{-1})^{-1} = g$.

Beweis

$$gg^{-1} = e = (g^{-1})^{-1}g^{-1}$$

Mit (3.1.3) folgt die Behauptung.

3.1.6 Lemma

Seien (G, \cdot) eine Gruppe, $k \in \mathbb{N}$ und $g_1, \dots, g_k \in G$. Dann haben alle Produkte, die mit g_1, \dots, g_k in dieser Reihenfolge, aber mit beliebiger Klammerung gebildet werden können, den gleichen Wert. Bezeichnung: $g_1 \cdot g_2 \cdot \dots \cdot g_k$.

3.1.7 Lemma

Seien (G, \cdot) eine Gruppe, $k \in \mathbb{N}$ und $g_1, \dots, g_k \in G$. Dann gilt $(g_1 \cdot \dots \cdot g_k)^{-1} = g_k^{-1} \cdot \dots \cdot g_1^{-1}$.

Beweis

Mit Hilfe von (3.1.6) folgt die Behauptung wegen $(g_k^{-1} \cdot \dots \cdot g_1^{-1})(g_1 \cdot \dots \cdot g_k) = (g_k^{-1} \cdot \dots \cdot g_2^{-1})(g_1^{-1}g_1)(g_2 \cdot \dots \cdot g_k) = e$ aus (3.1.4).

Bezeichnung

Seien (G, \cdot) eine Gruppe, $k \in \mathbb{N}$ und $g \in G$. Dann sei $g^k := \underbrace{(g \cdot \dots \cdot g)}_{k \text{ Mal}}$. (\gg wohldefiniert \ll wegen (3.1.6)).

Bemerkung

Nach (3.1.7) gilt $(g^k)^{-1} = (g^{-1})^k$.

Bezeichnung

Seien (G, \cdot) eine Gruppe, $k \in \mathbb{N}$ und $g \in G$. Dann sei $g^{-k} := (g^{-1})^k$. Ferner sei $g^0 := e$.

3.1.8 Lemma

Seien (G, \cdot) eine Gruppe, $k, l \in \mathbb{Z}$ und $g \in G$. Dann gilt:

- (i) $g^k \cdot g^l = g^{k+l}$
- (ii) $(g^k)^l = g^{kl}$

Definition

Seien (G, \cdot) eine Gruppe und $U \subseteq G$. Ist (U, \cdot) eine Gruppe (mit der selben Verknüpfung wie in G), so heißt (U, \cdot) **UNTERGRUPPE** von (G, \cdot) .

3.1.9 Lemma

Seien (G, \cdot) eine Gruppe und $U \subset G$. Dann ist (U, \cdot) eine Untergruppe von (G, \cdot) , falls gilt:

- (i) $U \neq \emptyset$
- (ii) $u, v \in U \Rightarrow uv \in U$
- (iii) $u \in U \Rightarrow u^{-1} \in U$

Beweis

trivial

Definition

Seien $(G, \cdot), (H, *)$ Gruppen, $\varphi : G \rightarrow H$ eine Abbildung und es gelte

$$g_1, g_2 \in G \Rightarrow \varphi(g_1 \cdot g_2) = \varphi(g_1) * \varphi(g_2)$$

Dann heißt φ (Gruppen-)HOMOMORPHISMUS. Ist φ ein Homomorphismus und ist φ bijektiv, so heißt φ (Gruppen-)ISOMORPHISMUS.

Beispiel

$(\mathbb{Z}, +)$ ist Untergruppe von $(\mathbb{R}, +)$. Die Abbildung $\varphi : \mathbb{Z} \rightarrow \mathbb{R}; x \mapsto x$ ist ein Homomorphismus (aber kein Isomorphismus).

Bemerkung

Gruppen und Gruppenhomomorphismen werden detailliert in der Gruppentheorie studiert. Im Sinne der Gruppentheorie sind isomorphe Gruppen nicht unterscheidbar, so dass man nach geeigneten »Darstellungen« (d.h. isomorphen Bildern) suchen kann.

3.2 Ringe und Körper

$(\mathbb{R}, +)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ sind Gruppen; beide Operationen treten in Gleichungssystemen auf. Wir betrachten im Folgenden daher Mengen R mit zwei Operationen; wir bezeichnen sie mit $\gg+\ll$ und $\gg\cdot\ll$ und nennen sie »Addition« und »Multiplikation« (obwohl sie mit den gleichnamigen Operationen auf \mathbb{R} i.a. nichts zu tun haben).

Definition

Sei R eine Menge mit zwei Operationen

$$+ : R \times R \rightarrow R \quad \cdot : R \times R \rightarrow R$$

$(R, +, \cdot)$ heißt RING, falls folgende Eigenschaften erfüllt sind:

- (R1) $(R, +)$ ist eine kommutative Gruppe
- (R2) \cdot erfüllt das Assoziativgesetz
- (R3) Es gelten die Distributivgesetze

$$a, b, c \in R \Rightarrow a \cdot (b + c) = a \cdot b + a \cdot c$$

$$a, b, c \in R \Rightarrow (a + b) \cdot c = a \cdot c + b \cdot c$$

Sei $(R, +, \cdot)$ ein Ring. Ein Element $1 \in R$ heißt EINSELEMENT, wenn gilt

$$a \in R \Rightarrow 1 \cdot a = a \cdot 1 = a$$

$(R, +, \cdot)$ heißt KOMMUTATIV, wenn in (R, \cdot) das Kommutativgesetz gilt.

Beispiel

$$(\mathbb{Z}, +, \cdot)$$

3.2.1 Lemma

Seien $(R, +, \cdot)$ ein Ring und 0 das neutrale Element der Addition. Dann gilt für alle $a \in R$:

$$0 \cdot a = a \cdot 0 = 0$$

Beweis

$$0 \cdot a = (0 + 0)a = 0a + 0a \Rightarrow 0 = 0a + (-0a) = 0a + 0a + (-0a) = 0a + 0 = 0a$$

Analog folgt $a \cdot 0 = 0$.

Beispiel

Sei $R = \text{Abb}([0, 1], \mathbb{R}) := \{f; f : [0, 1] \rightarrow \mathbb{R}\}$. Definiere $+$: $R \times R \rightarrow R$, \cdot : $R \times R \rightarrow R$ durch $(f + g)(x) := f(x) + g(x)$, $(f \cdot g)(x) = f(x) \cdot g(x)$. Ferner sei $h \in R$ mit $h(x) = 1 \forall x \in [0, 1]$. Dann gilt für jedes $f \in R$:

$$(fh)(x) = f(x)h(x) = f(x) = h(x)f(x) = (hf)(x)$$

$(R, +, \cdot)$ ist kommutativer Ring mit Einselement.

Definition

Sei $(R, +, \cdot)$ ein Ring. $(R, +, \cdot)$ heißt NULLTEILERFREI, wenn gilt

$$a, b \in R \wedge ab = 0 \Rightarrow a = 0 \vee b = 0$$

Beispiele

- (a) Der oben definierte Ring $(\text{Abb}([0, 1], \mathbb{R}), +, \cdot)$ ist nicht nullteilerfrei. $f, g \in R$, $f, g \neq 0$ (f, g sind nicht die Nullfunktion), aber $fg = 0$ (Nullfunktion), z.B.

$$f = \begin{cases} 0 & \text{für } 0 \leq x \leq \frac{1}{2} \\ 1 & \text{sonst} \end{cases}$$

$$g = \begin{cases} 1 & \text{für } \frac{1}{2} < x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

- (b) $(\mathbb{Z}, +, \cdot)$ ist nullteilerfrei.

Bemerkung

Ist $(R, +, \cdot)$ ein nullteilerfreier Ring, so ist die Multiplikation eine assoziative Verknüpfung auf $\mathbb{R} \setminus \{0\}$.

Beispiel

Sei $2\mathbb{Z} := \{2z; z \in \mathbb{Z}\}$, $(2\mathbb{Z}, +, \cdot)$ ist ein nullteilerfreier kommutativer Ring. Er enthält *kein* Einselement und zu keinem Element ein multiplikatives Inverses.

$(\mathbb{R}, +, \cdot)$ ist ein nullteilerfreier, kommutativer Ring mit Einselement. $(\mathbb{R} \setminus \{0\}, \cdot)$ ist eine Gruppe.

Definition

Sei K eine Menge mit zwei Operationen $+, \cdot$. $(K, +, \cdot)$ heißt KÖRPER, wenn die folgenden Eigenschaften erfüllt sind:

- (K1) $(K, +)$ ist eine kommutative Gruppe
 (K2) $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe
 (K3) Es gilt das Distributivgesetz:

$$a, b, c, \in K \Rightarrow a(b + c) = ab + ac$$

Bezeichnung

$$K^* := K \setminus \{0\}$$

Bemerkungen

- (a) (K2) schließt die Bedingung $\cdot : K^* \times K^* \rightarrow K^*$, d.h. die Nullteilerfreiheit, ein.
- (b) Wegen der Kommutativität von (K^*, \cdot) gilt auch das »zweite« Distributivgesetz.
- (c) Ist $(K, +, \cdot)$ ein Körper, so ist $(K, +, \cdot)$ ein Ring. (Warum?) (Aus (K2) folgt (R2) nur für Elemente aus K^* !)

3.2.2 Lemma

Sei $(K, +, \cdot)$ ein Körper. Dann gilt:

- (i) $0 \neq 1$
- (ii) $a, b \in K \Rightarrow a(-b) = -(ab)$
- (iii) $a, b \in K \Rightarrow (-a)(-b) = ab$
- (iv) $a, b, x \in K \wedge ax = bx \wedge x \neq 0 \Rightarrow a = b$

Beweis

- (i) $0 \notin K^*, 1 \in K^*$
- (ii) $ab + a(-b) = a(b - b) = a \cdot 0 \stackrel{(3.2.1)}{=} 0 \stackrel{(3.1.4)}{\Rightarrow} a(-b) = -(ab)$
- (iii) $(-a)(-b) \stackrel{(ii)}{=} -((-a)b) \stackrel{(ii)}{=} -(-ab) \stackrel{(3.1.5)}{=} ab$
- (iv) Die Kürzungsregeln (3.1.3) gelten in K^* . Also folgt die Behauptung für $a, b \neq 0$. Gilt $a = 0$, so folgt aus (3.2.1), dass $ax = 0$, also wegen $ax = bx$ auch $bx = 0$. Wegen $x \neq 0$ folgt aus der Nullteilerfreiheit $b = 0$, also $a = b$. Vertauschung der Rollen von a und b in diesem Argument liefert die Behauptung.

Beispiele

- (a) $(\mathbb{R}, +, \cdot), (\mathbb{Q}, +, \cdot)$ sind Körper, $(\mathbb{Z}, +, \cdot)$ ist keiner. $(\mathbb{C}, +, \cdot)$ ist ein Körper.
- (b) $(\mathbb{Z}_2, +, \cdot)$ ist Körper; wegen (3.2.2(ii)) gibt es keinen mit weniger Elementen.

| | | | | | |
|---|---|---|---|---|---|
| + | 0 | 1 | · | 0 | 1 |
| 0 | 0 | 1 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 0 | 1 |

- (c) $(\mathbb{Z}_2, +, \cdot)$ ist Spezialfall von $(\mathbb{Z}_p, +, \cdot)$ für eine Primzahl p . Dass $(\mathbb{Z}_p^*, +, \cdot)$ eine kommutative Gruppe ist, haben wir bereits im Abschnitt über Gruppen (Seite 17ff) nachgewiesen. Addition modulo p macht $(\mathbb{Z}_p, +)$ zu einer kommutativen Gruppe. Ferner gilt das Distributivgesetz.

Bemerkung

Sei $(K, +, \cdot)$ ein Körper. Wegen (K2) kann die Multiplikation nicht aus K^* hinausführen. Wie (b) und (c) zeigen, kann das bezüglich der Addition passieren, denn

$$\begin{aligned} \mathbb{Z}_2^* : 1 \in \mathbb{Z}_2^*, 0 \notin \mathbb{Z}_2^*, \text{ aber } 1 + 1 = 0 \\ \mathbb{Z}_p^* : 1 \in \mathbb{Z}_p^*, 0 \notin \mathbb{Z}_p^*, \underbrace{1 + \dots + 1}_{p \text{ Mal}} = 0 \end{aligned}$$

Beachte, dass $p \notin \{1, \dots, p-1\}$; tatsächlich liegt p in der selben Äquivalenzklasse (genannt RESTKLASSE) wie 0. Schreibweise: $p = 0 \pmod{p}$. $p \cdot 1 = 0$ steht also nicht im Widerspruch zur Nullteilerfreiheit.

In \mathbb{Z} kann das nicht passieren.

Definition

Sei $(R, +, \cdot)$ ein Ring mit Einselement.

$$\text{char}(R, +, \cdot) := \begin{cases} 0 & \text{falls } n \in \mathbb{N} \Rightarrow n \cdot 1 \neq 0 \\ \min \{n : n \in \mathbb{N}; n \cdot 1 = 0\} & \text{sonst} \end{cases}$$

heißt CHARAKTERISTIK von $(R, +, \cdot)$.

Bezeichnung

Wenn für Gruppen, Ringe oder Körper klar ist, welche Verknüpfungen gemeint sind, so lässt man diese meistens weg und schreibt nur G, R, K . Insbesondere wird dann auch $\text{char}(R)$ geschrieben.

Beispiel

$$\begin{aligned} \text{char}(\mathbb{Z}) &= 0 \\ \text{char}(\mathbb{Z}_p) &= p \end{aligned}$$

3.2.3 Lemma

Sei K ein Körper. Dann ist $\text{char}(K)$ eine Primzahl, oder es gilt $\text{char}(K) = 0$.

Beweis

Sei $\text{char}(K) = m \cdot n \neq 0$. Dann gilt $0 = (m \cdot n) \cdot 1 = (m \cdot 1)(n \cdot 1)$, $m \cdot 1 = \underbrace{1 + \dots + 1}_{m \text{ Mal}}$, $n \cdot 1 = \underbrace{1 + \dots + 1}_{n \text{ Mal}}$ liegen wegen der Minimalität von (mn) in K^* , falls $m, n < mn$ gilt. Aus der Nullteilerfreiheit von K^* folgt, dass $m = 1$ oder $n = 1$ ist. Also kann $\text{char}(K)$ keine echten Teiler besitzen.

Definition

- (a) Seien $(R, +, \cdot)$ ein Ring und $U \subseteq R$. Ist $(U, +, \cdot)$ ein Ring, so heißt $(U, +, \cdot)$ **UNTERRING** von $(R, +, \cdot)$.
Sei $(K, +, \cdot)$ ein Körper und $U \subseteq K$ und ist $(U, +, \cdot)$ ein Körper, so heißt $(U, +, \cdot)$ **UNTERKÖRPER** von $(K, +, \cdot)$.
- (b) Seien $(R, +, \cdot)$ und (H, \oplus, \odot) Ringe (bzw. Körper). Eine Abbildung $\varphi : R \rightarrow H$ heißt **(Ring- bzw. Körper-)HOMOMORPHISMUS** (falls φ bijektiv ist, **ISOMORPHISMUS**), falls gilt:

$$\begin{aligned} a, b \in R &\Rightarrow \varphi(a + b) = \varphi(a) \oplus \varphi(b) \\ a, b \in R &\Rightarrow \varphi(a \cdot b) = \varphi(a) \odot \varphi(b) \end{aligned}$$

3.3 Polynome

In der Analysis treten Polynome als spezielle Funktionen auf. Wir führen im Folgenden Polynome als algebraische Struktur ein. Später wird der Zusammenhang mit den entsprechenden Funktionen hergestellt. Im Folgenden sei $K = (K, +, \cdot)$ ein beliebiger Körper.

Bezeichnung

Sei $\sigma : \mathbb{N}_0 \rightarrow K$. Dann heißt σ SEQUENZ über K . Wir schreiben σ auch in der Form $(\sigma(i))_{i \in \mathbb{N}_0}, (\sigma(0), \sigma(1), \dots)$ oder mit $\sigma(i) := a_i$ für $i \in \mathbb{N}_0$ einfacher $(a_i)_{i \in \mathbb{N}_0}$. $0 := (0)_{i \in \mathbb{N}_0}$.

Definition

Sei $\wp := \{(a_i)_{i \in \mathbb{N}_0}; a_i = 0 \text{ für alle bis auf endlich viele Indizes } i \in \mathbb{N}_0\}$. Jedes Element f von \wp heißt POLYNOM über K . Ist $f = (a_i)_{i \in \mathbb{N}_0} \in \wp$, so heißt $\deg(f)$ GRAD (degree) von f .

$$\deg(f) := \begin{cases} \max \{i : i \in \mathbb{N}_0 \wedge a_i \neq 0\} & \text{falls } f \neq 0 \\ -\infty & \text{falls } f = 0 \end{cases}$$

Für $i \in \mathbb{N}_0$ heißt a_i I-TER KOEFFIZIENT von f ; mit $n = \deg(f) \geq 0$ heißt a_n FÜHRENDER KOEFFIZIENT von f . f heißt NORMIERT, wenn $a_n = 1$ ist. 0 heißt NULLPOLYNOM. Addition (+) und Multiplikation (\cdot) von Polynomen sind wie folgt definiert: Seien $f = (a_i)_{i \in \mathbb{N}_0}, g = (b_i)_{i \in \mathbb{N}_0}, f, g \in \wp$. Dann sei

$$f + g = (a_i + b_i)_{i \in \mathbb{N}_0}$$

$$f \cdot g = (c_i)_{i \in \mathbb{N}_0} \text{ mit } c_i = \sum_{j=0}^i a_j b_{i-j}$$

3.3.1 Satz

$(\wp, +, \cdot)$ ist ein nullteilerfreier, kommutativer Ring mit Einselement (\gg Polynomring über $K \ll$).

Beweis

(R1), (R2), (R3) rechnet man leicht nach, ebenso die Kommutativität der Multiplikation. Ferner ist klar, dass $(1, 0, \dots)$ Einselement ist. Es bleibt also nur die Nullteilerfreiheit zu zeigen.

Seien $f = (a_i)_{i \in \mathbb{N}_0}, g = (b_i)_{i \in \mathbb{N}_0}$ Polynome, $f \neq 0, g \neq 0$. Sei $n := \deg(f), m := \deg(g)$. Dann gilt $f \cdot g = (c_i)_{i \in \mathbb{N}_0}$ mit $\sum_{j=0}^i a_j b_{i-j}$, also ist insbesondere $c_{n+m} =$

$$\sum_{j=0}^{n+m} a_j b_{n+m-j} = \underbrace{\sum_{j=0}^{n-1} a_j b_{n+m-j}}_{=0} + \underbrace{a_n b_m}_{\neq 0} + \underbrace{\sum_{j=n+1}^{n+m} a_j b_{n+m-j}}_{=0} = a_n b_m \neq 0, \text{ d.h.}$$

$f \cdot g \neq 0$.

Formale Setzung: Seien $n, m \in \mathbb{N}_0$. Dann sei $n + (-\infty) = (-\infty) + m = (-\infty) + (-\infty) = (-\infty)$.

3.3.2 Korollar

Seien $f, g \in \wp$. Dann gilt $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Beweis

Folgt mit obiger formaler Setzung wie im Beweis der Nullteilerfreiheit in (3.3.1). Genau so wie über \mathbb{Z} können wir auch im Polynomring mit Rest dividieren.

3.3.3 Satz

Seien $f, g \in \wp, g \neq 0$. Dann gibt es eindeutig bestimmte Polynome $q, r \in \wp$ mit $\deg(r) \leq \deg(g) - 1$, so dass $f = q \cdot g + r$.

Beweis

Eindeutigkeit: Seien $q, \bar{q}, r, \bar{r} \in \wp$ mit $\deg(r), \deg(\bar{r}) \leq \deg(g) - 1$ und $f = q \cdot g + r$, $f = \bar{q} \cdot g + \bar{r}$. Dann folgt $0 = (q - \bar{q}) \cdot g + (r - \bar{r})$, also $(q - \bar{q}) \cdot g = \bar{r} - r$. Mit (3.3.2) folgt, dass $\deg(\bar{r} - r) = \deg((q - \bar{q}) \cdot g)$ sein muss: $\deg(\bar{r} - r) = \deg(q - \bar{q}) + \deg(g)$. Da $g \neq 0$ vorausgesetzt war und außerdem $\deg(\bar{r} - r) \leq \max\{\deg(\bar{r}), \deg(r)\} \leq \deg(g) - 1$ ist, folgt $\deg(q - \bar{q}) = -\infty$, also $q = \bar{q}$. Wegen $q - \bar{q} = \bar{r} - r$ folgt auch $r = \bar{r}$.

Existenz: Ist $\deg(f) < \deg(g)$, so kann man $q := 0$ und $r := f$ wählen. Sei also $n := \deg(f)$, $n \geq \deg(g) := m$ mit $f := (a_i)_{i \in \mathbb{N}_0}, g := (b_i)_{i \in \mathbb{N}_0}$. Setze $q_i = (0, \dots, 0, \frac{a_n}{b_n}, 0, \dots, 0)$ ($\frac{a_n}{b_n}$ steht an der Stelle $n - m$ im Polynom), $f_1 = f - q_1 \cdot g$. Es ist $\deg(f_1) < \deg(f)$. Ist bereits $\deg(f_1) < m$, so setzen wir $q := q_1, r := f_1$. Andernfalls wiederholen wir das Verfahren mit f_1 und g . Sukzessive erhalten wir so Polynome f_1, \dots, f_k mit $\deg(f) > \deg(f_1) > \dots > \deg(f_k)$, d.h. das Verfahren liefert nach höchstens $n - m$ Schritten ein f_k mit $\deg(f_k) < m$. Man erhält

$$f = q_1 g + f_1 = q_1 g + (q_2 g + f_2) = (q_1 + q_2)g + f_2 = \dots = (q_1 + \dots + q_k)g + f_k$$

und es folgt die Behauptung.

Wir betrachten im Folgenden die Struktur von $(\wp, +, \cdot)$ genauer.

3.3.4 Bemerkung

Sei $(U, +, \cdot)$ der Unterring von $(\wp, +, \cdot)$ der Polynome vom Grad höchstens 0. Dann ist $(U, +, \cdot)$ ein Körper. Die Abbildung $\varphi : K \rightarrow U$, definiert durch $\varphi(a) = (a, 0, \dots)$ ist ein Körperisomorphismus.

Beweis: trivial

Hinweis

Man kann also die Skalare $a \in K$ mit den Polynomen $(a, 0, \dots) \in \wp$ identifizieren. Multiplikation in \wp geht auf das Bilden von Potenzen zurück.

3.3.5 Bemerkung

Sei $X := (0, 1, 0, \dots)$, und setze $X^0 := (1, 0, \dots)$. Sei $f = (a_i)_{i \in \mathbb{N}_0} \in \wp$ mit $\deg(f) \leq n$. Dann gilt $f = \sum_{i=0}^n a_i X^i$.

Beweis

Es reicht zu zeigen, dass $X^i = (0, \dots, 0, \underset{i}{1}, 0, \dots)$ gilt für $i \in \mathbb{N}_0$. Für $i = 0, 1$ folgt das aus der Definition, der Rest mittels vollständiger Induktion, denn $X \cdot X^i = (0, 1, 0, \dots) \cdot (0, \dots, 0, \underset{i}{1}, 0, \dots) = (0, \dots, 0, \underset{i+1}{1}, 0, \dots)$.

Bezeichnung

X wird oft UNBESTIMMTE genannt. $(\wp, +, \cdot)$ wird dann mit $K[X]$ bezeichnet und POLYNOMRING IN EINER UNBESTIMMTEN genannt. Für f schreibt man auch $f(X)$. Die Verbindung zu den Polynomen in der Analysis erhält man, wenn man »für X Werte x aus K einsetzt«.

Definition

Sei K ein Körper. Eine Funktion $p : K \rightarrow K$ heißt POLYNOMFUNKTION auf K , wenn es ein $n \in \mathbb{N}_0$ und $a_0, \dots, a_n \in K$ gibt mit $p(x) = \sum_{i=0}^n a_i x^i$ für alle $x \in K$. Ist p eine Polynomfunktion und $p(x) = \sum_{i=0}^n a_i x^i$ für alle $x \in K$, so heißt $\sum_{i=0}^n a_i x^i$ (POLYNOM-)DARSTELLUNG von p .

Die Rechenregeln für Darstellungen $\sum_{i=0}^n a_i x^i, \sum_{i=0}^m b_i x^i$ sind definiert wie für die zugehörigen Polynome $(a_0, a_1, \dots, a_n, 0, \dots), (b_0, b_1, \dots, b_m, 0, \dots)$.

Bemerkung

Polynomfunktionen sind gleich, wenn ihre *Werte* übereinstimmen. Darstellungen sind gleich, wenn ihre *Koeffizienten* übereinstimmen.

3.3.6 Bemerkung

Ist K endlich, so sind die Darstellungen von Polynomfunktionen *nicht* eindeutig.

Beweis

Es reicht zu zeigen, dass die Nullfunktion eine nichttriviale Darstellung besitzt (da $f = f+0$). Sei $K = \{k_1, \dots, k_n\}$. Definiere $p(x) = (x-k_1) \cdot \dots \cdot (x-k_n)$ für alle $x \in K$. Dann ist p (als Funktion) identisch 0 (*Schreibweise*: $p \equiv 0$), aber die Koeffizienten der zugehörigen Darstellung sind *nicht* alle 0; speziell ist der Koeffizient von $x^n = 1$.

Bezeichnung

\mathcal{F} sei die Menge der Polynomfunktionen, \mathcal{D} die Menge der Darstellungen.

3.3.7 Bemerkung

$(\mathcal{F}, +, \cdot)$ und $(\mathcal{D}, +, \cdot)$ sind kommutative Ringe mit Einselement. Die Abbildung $\varphi : \mathcal{D} \rightarrow \mathcal{F}$, definiert durch $\varphi(\sum_{i=0}^n a_i x^i) := p$ mit $p(x) = \sum_{i=0}^n a_i x^i$ ist ein surjektiver (Ring-)Homomorphismus (kanonischer EPIMORPHISMUS). Die Abbildung $\psi : K[X] \rightarrow \mathcal{D}$, definiert durch $\psi(\sum_{i=0}^n a_i X^i) := \sum_{i=0}^n a_i x^i$ ist ein ISOMORPHISMUS.

Hinweis

$K[X]$ und $(\mathcal{D}, +, \cdot)$ können also identifiziert werden. $(\mathcal{F}, +, \cdot)$ ist dagegen für endliche Körper nicht einmal nullteilerfrei.

Bezeichnung

$(\mathcal{D}, +, \cdot)$ wird auch mit $K[x]$ bezeichnet; die Elemente von $K[x]$ werden auch POLYNOME genannt.

Bemerkung

Wegen (3.3.7) übertragen sich die Aussagen über $K[X]$ auf $K[x]$, ebenso Begriffe wie Grad.

Definition

Seien $p \in K[x]$ und $x_0 \in K$ mit $p(x_0) = 0$. Dann heißt x_0 NULLSTELLE von p .

Beispiel

Sei $K = \{k_1, \dots, k_m\}$, $p_j(x) = k_j + \prod_{i=1}^m (x - k_i)$ für $j = 1, \dots, m$. Dann gilt $p_j(x) \equiv k_j$, besitzt also keine Nullstelle, wenn $k_j \neq 0$ und m Nullstellen, falls $k_j = 0$ ist.

Die folgende Aussage zeigt, dass den Nullstellen x_0 von Polynomen sog. LINEARFAKTOREN $X - x_0$ bzw. $x - x_0$ entsprechen.

3.3.8 Satz

Seien $p \in K[x]$ und x_0 Nullstelle von p . Ferner sei $g \in K[x]$ definiert durch $g(x) = x - x_0$. Dann gibt es genau ein $q \in K[x]$ mit $gq = p$ und $\deg(q) = \deg(p) - 1$.

Beweis

Nach (3.3.3) gibt es eindeutig bestimmte $q, r \in K[x]$ mit $p = gq + r$ und $\deg(r) \leq \deg(g) - 1 = 0$. Also ist $r \equiv r_0 \in K$ eine Konstante und wegen $0 = p(x_0) = 0 \cdot q(x_0) + r_0$ gilt $r \equiv 0$. Aus (3.3.2) folgt $\deg(p) = \deg(g) + \deg(q) = 1 + \deg(q)$.

Bemerkung

Ist f eine Polynomfunktion und ist $x_0 \in K$ mit $f(x_0) = 0$, so gibt es ein $q \in \mathcal{F}$ mit $f(x) = (x - x_0)q(x)$; die Eindeutigkeit geht i.a. aber verloren.

Beispiel

$K = \mathbb{Z}_3$; $f(x) = p(x) = x(x-1)(x-2)$ für $x \in \{0, 1, 2\}$. Die Polynomfunktion $f \in \mathcal{F}$ ist also die Nullfunktion; $p \in K[x]$ ist *nicht* das Nullpolynom. Es gilt $p(x) = xq(x)$ mit $q(x) = (x-1)(x-2)$ und $q \in K[x]$ ist eindeutig bestimmt (nach (3.3.8)).

Seien $g_1, g_2 \in \mathcal{F}$ mit $g_1(x) = (x-1)(x-2)$, $g_2(x) = 2(x-1)(x-2)$. Dann gilt $g_1(0) = 2$, $g_2(0) = 1$, d.h. $g_1 \neq g_2$ auch in \mathcal{F} . Andererseits gilt aber $f(x) = xg_1(x)$ und $f(x) = xg_2(x)$ für alle $x \in \{0, 1, 2\}$.

3.3.9 Korollar

Sei $p \in K[x]$ nicht das Nullpolynom. Dann hat p höchstens $\deg(p)$ viele Nullstellen.

Beweis

(durch vollständige Induktion nach $n := \deg(p)$)

$n = 0$: $p \equiv a_0 \in K \setminus \{0\}$

$n \geq 0$: Falls p keine Nullstelle besitzt, folgt die Behauptung. Andernfalls sei x_0 Nullstelle. Nach (3.3.8) gibt es $q \in K[x]$ mit $p(x) = (x - x_0)q(x)$ für alle $x \in K$ und $\deg(q) = \deg(p) - 1$. Aus der Induktionsannahme folgt die Behauptung.

3.3.10 Korollar

Ist K unendlich, so ist die Abbildung $\varphi : K[x] \rightarrow \mathcal{F}$, definiert durch

$$\varphi \left(\sum_{i=0}^n a_i x^i \right) = f$$

mit $f(x) = \sum_{i=0}^n a_i x^i$ für $x \in K$ ein (Ring-)Isomorphismus.

Beweis

Nach (3.3.7) reicht es zu zeigen, dass φ injektiv ist. Seien $p_1, p_2 \in K[x]$ mit $\varphi(p_1) = \varphi(p_2)$. Dann gilt $\varphi(p_1 - p_2) = 0$, d.h. die Polynomfunktion $\varphi(p_1 - p_2)$ hat unendlich viele Nullstellen. Sei $\sum_{i=1}^n a_i x^i$ eine Darstellung von $\varphi(p_1 - p_2)$. Dann hat auch $\sum_{i=0}^n a_i x^i$ unendlich viele Nullstellen. Aus (3.3.9) folgt $a_0 = \dots = a_n = 0$. Also gilt $p_1 = p_2$ (in $K[x]$).

Definition

Seien $p \in K[x] \setminus \{0\}$, $x_0 \in K$. Dann heißt

$$\mu(p, x_0) := \max \{m : m \in \mathbb{N}_0 : \exists q \in K[x] \wedge x \in K \Rightarrow p(x) = (x - x_0)^m q(x)\}$$

die VIELFACHHEIT DER NULLSTELLE x_0 von p .

Hinweis

$\mu(p, x_0) = \max \{m : m \in \mathbb{N}_0 \wedge f(x_0) = f'(x_0) = \dots = f^{(m-1)}(x_0) = 0\}$ für $K = \mathbb{R}$, d.h. genau die 0-te, 1-te, \dots , $(\mu(p, x_0) - 1)$ -te Ableitung in x_0 verschwindet.

3.3.11 Satz (Fundamentalsatz der Algebra)

Sei $f \in \mathbb{C}[x]$. Dann gibt es $z_1, \dots, z_{\deg(f)} \in \mathbb{C}$ und $a \in \mathbb{C}$ mit $f(x) = a(x - z_1) \cdot \dots \cdot (x - z_{\deg(f)})$.

Beweis

siehe z.B. G. Fischer, Einführung in die Algebra, 1974

Bezeichnung

Seien $z \in \mathbb{C}$, $a, b \in \mathbb{R}$ und $z = a + bi$. Dann heißt $\bar{z} = a - bi$ zu z KONJUGIERT KOMPLEXE ZAHL.

3.3.12 Lemma

Seien $p \in \mathbb{R}[x]$, $z \in \mathbb{C}$ eine Nullstelle von p . Dann gilt $\mu(p, z) = \mu(p, \bar{z})$.

Hinweis

Die (komplexen) Nullstellen von Polynomen mit *reellen* Koeffizienten liegen *symmetrisch* zur reellen Achse.

3.3.13 Satz

Sei $p \in \mathbb{R}[x]$ mit $n := \deg(p) \geq 1$. Dann gibt es ein $a \in \mathbb{R} \setminus \{0\}$, $m \in \mathbb{N}_0$, $x_1, \dots, x_m \in \mathbb{R}$, $s \in \mathbb{N}_0$, $q_1, \dots, q_s \in \mathbb{R}[x]$ normiert vom Grad 2, ohne reelle Nullstelle, so dass für alle $x \in \mathbb{R}$ gilt:

$$p(x) = a(x - x_1) \cdot \dots \cdot (x - x_m) \cdot q_1(x) \cdot \dots \cdot q_s(x)$$

Ferner gilt $m + 2s = n$.

Beweis

Die letzte Identität folgt aus (3.3.2).

Die Existenz der behaupteten Faktorisierung folgt aus dem Fundamentalsatz der Algebra, denn:

Nach (3.3.11) gibt es $z_1, \dots, z_n \in \mathbb{C}$ und $a \in \mathbb{C}$ mit $p(x) = a(x - z_1) \cdot \dots \cdot (x - z_n)$ für alle $x \in \mathbb{R}$. Da $p \in \mathbb{R}[x]$ ist und a der führende Koeffizient ist, folgt $a \in \mathbb{R}$. Sei $m \in \mathbb{N}_0$ mit $z_1, \dots, z_m \in \mathbb{R}$, $z_{m+1}, \dots, z_n \in \mathbb{C} \setminus \mathbb{R}$. Nach (3.3.12) ist $n - m$ gerade und durch Ummummerierung kann erreicht werden, dass z_i und \bar{z}_i für $i = (m + 1) + 2j$, $j = 0, \dots, \frac{1}{2}(n - m) - 1$ konjugiert komplex sind. Nun gilt für $z = a + bi$ ($a, b \in \mathbb{R}$):

$$(x - z)(x - \bar{z}) = (x - a - bi)(x - a + bi) = x^2 - 2ax + (a^2 + b^2) \in \mathbb{R}[x]$$

Das Polynom $x^2 + 2ax + a^2 + b^2$ besitzt *keine* reelle Nullstelle (denn $a^2 - (a^2 + b^2) = -b^2 \leq 0$), falls $b \neq 0$. Insgesamt folgt die Behauptung.

Hinweis

Es gibt verschiedene strukturelle Aussagen über reelle Nullstellen und numerische Algorithmen zu ihrer Bestimmung.

Kapitel 4

Vektorräume

In (1.1): Vektoren des $\mathbb{R}^n, \mathbb{R}^m$ sind wichtig zur Beschreibung linearer Gleichungssysteme über \mathbb{R} .

Jetzt: allgemeiner Standpunkt.

4.1 (Unter-)Vektorräume und lineare Hülle

Im Folgenden sei K ein beliebiger Körper.

Beispiel

$$K^n = \left\{ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} : v_1 \in K \wedge \dots \wedge v_n \in K \right\}$$

Definiere für $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in K^n, \lambda \in K$:

$$v + u = \begin{pmatrix} v_1 + u_1 \\ \vdots \\ v_n + u_n \end{pmatrix}$$
$$\lambda v = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}$$

Beachte: $+$: $K^n \times K^n \rightarrow K^n$, aber \cdot : $K \times K^n \rightarrow K^n$. $+$ ist daher eine INNERE VERKNÜPFUNG, \cdot ist eine ÄUSSERE VERKNÜPFUNG, die Multiplikation mit einem Skalar.

Hinweis

Diese beiden Verknüpfungen traten bereits bei der Formulierung $x_1 v_1 + \dots + x_n v_n = b$ mit $b, v_1, \dots, v_n \in \mathbb{R}^n, x_1, \dots, x_n \in \mathbb{R}$ von Gleichungssystemen mittels der Spalten ihrer Koeffizientenmatrix auf, sowie bei der Parametrisierung ihrer Lösungsmengen. (Beachte: Wir haben in (1.1) auch $v_1 x_1 + \dots + v_n x_n$ für $x_1 v_1 + \dots + x_n v_n$ geschrieben. - In Analogie zu einer Interpretation als Abbildung (vgl Kapitel 5) - obwohl \cdot formal nur auf Paaren $K \times K^n$ wirkt.)

Definition

Seien V eine Menge, K ein Körper und $\oplus : V \times V \rightarrow V, \odot : K \times V \rightarrow V$ Verknüpfungen. (V, \oplus, \odot) heißt K -VEKTORRAUM, wenn gilt:

(V1) (V, \oplus) ist eine kommutative Gruppe

(V2) (a)

$$\lambda, \mu \in K, v \in V \Rightarrow \lambda \odot (\mu \odot v) = (\lambda\mu) \odot v$$

(ASSOZIATIVITÄT)

(b)

$$\begin{aligned} \lambda, \mu \in K \wedge v \in V &\Rightarrow (\lambda + \mu) \odot v = (\lambda \odot v) \oplus (\mu \odot v) \\ \lambda \in K \wedge v, w \in V &\Rightarrow \lambda \odot (v \oplus w) = (\lambda \odot v) \oplus (\lambda \odot w) \end{aligned}$$

(DISTRIBUTIVITÄT)

(c) $v \in V \Rightarrow 1 \odot v = v$

Konvention

\odot bindet stärker als \oplus .

Bemerkung

(V2) überträgt sich auf endliche Summen bzw. Produkte

Bezeichnung

Die Operationen in K und die Verknüpfungen in V werden im Folgenden beide mit $+$ und \cdot bezeichnet.

Der NULLVEKTOR, d.h. das Nullelement der Addition in V wird mit 0 bezeichnet; die Elemente von V heißen VEKTOREN.

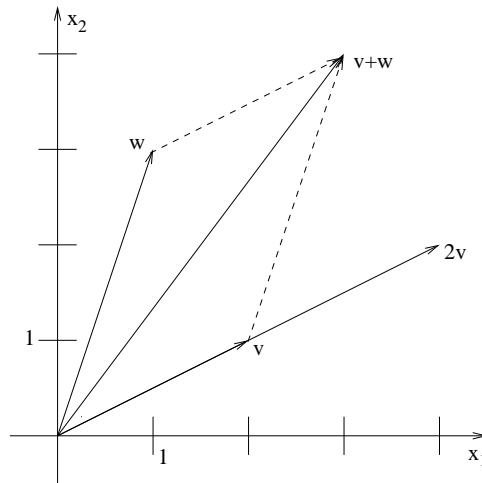
Beispiele

- (a) $(\mathbb{R}^n, +, \cdot)$ ist ein \mathbb{R} -Vektorraum. Die Vektoren kann man als »Ortsvektoren« interpretieren; die Verknüpfungen lassen sich geometrisch darstellen. (vgl. 1.1 und Abbildung 4.1)
- (b) Allgemeiner ist $(K^n, +, \cdot)$ für einen beliebigen Körper K ein K -Vektorraum.
- (c) Der Polynomring $K[x]$ ist auch ein K -Vektorraum (Die Multiplikation $\cdot : K \times K[x] \rightarrow K[x]$ wird mittels (3.3.4) als Spezialfall der Multiplikation von Polynomen erklärt, wie bereits - implizit - in (3.3.5)).
- (d) Ein weiteres Beispiel bilden die $m \times n$ -Matrizen (vgl. 1.1).

Definition

Seien K ein Körper und $m, n \in \mathbb{N}$. A heißt $m \times n$ -MATRIX (mit Koeffizienten in K), wenn für $i = 1, \dots, m$ und $j = 1, \dots, n$ Elemente $a_{ij} \in K$ existieren, so dass

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Abbildung 4.1: Geometrische Darstellung der Verknüpfungen im \mathbb{R}^2

Die Begriffe Koeffizient etc. werden analog zur Definition in (1.1) bzw. (2.1) (für $K = \mathbb{R}$) benutzt. Die Menge aller $m \times n$ -Matrizen mit Koeffizienten in K wird mit $K^{m \times n}$ bezeichnet. Die Operationen $+$: $K^{m \times n} \times K^{m \times n} \rightarrow K^{m \times n}$ und \cdot : $K \times K^{m \times n} \rightarrow K^{m \times n}$ seien definiert durch:

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}, B = (b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}, \lambda \in K \Rightarrow \\ A + B := (a_{ij} + b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \\ \lambda A := (\lambda a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

4.1.1 Bemerkung

Seien K ein Körper und $m, n \in \mathbb{N}$. Dann ist $(K^{m \times n}, +, \cdot)$ ein K -Vektorraum. Beweis: trivial.

Hinweis

$(K^{m \times n}, +, \cdot)$ ist gerade $(K^{mn}, +, \cdot)$ (bis auf die andere Art, Vektoren zu schreiben). (Später: Matrizenmultiplikation; Interpretation von Matrizen als Abbildungen)

Bezeichnung

Im Folgenden sei K stets ein Körper. Ein K -Vektorraum $(V, +, \cdot)$ wird einfacher mit V bezeichnet. Wenn Verwechslungen ausgeschlossen sind, sprechen wir auch einfacher nur von einem Vektorraum.

4.1.2 Lemma

Sei V ein K -Vektorraum. Dann gilt:

- (i) $v \in V \Rightarrow 0 \cdot v = 0$
- (ii) $\lambda \in K \Rightarrow \lambda \cdot 0 = 0$
- (iii) $\lambda \in K \wedge v \in V \wedge \lambda v = 0 \Rightarrow \lambda = 0 \vee v = 0$
- (iv) $v \in V \Rightarrow (-1) \cdot v = -v$

Beispiel

Sei $A \in \mathbb{R}^{m \times n}$. Dann ist $\text{Lös}(A, 0)$ ein Vektorraum. Andererseits ist $\text{Lös}(A, 0)$ eine Teilmenge des Vektorraums \mathbb{R}^n , die ihre Verknüpfungen vom \mathbb{R}^n »erbt«.

Definition

Seien $(V, +, \cdot)$ ein K -Vektorraum, $U \subseteq V$. Ist $(U, +, \cdot)$ ein K -Vektorraum, so heißt $(U, +, \cdot)$ UNTERVEKTORRAUM von $(V, +, \cdot)$.

4.1.3 Lemma

Seien V ein K -Vektorraum und $U \subset V$. U ist genau dann ein Untervektorraum von V , wenn gilt:

- $U \neq \emptyset$
- $u_1, u_2 \in U \Rightarrow u_1 + u_2 \in U$
- $\lambda \in K \wedge u \in U \Rightarrow \lambda u \in U$

Beweis

Aus (3.1.9) und (4.1.2(iv)) folgt, dass $(U, +)$ Untergruppe von $(V, +)$ ist. $(U, +)$ ist kommutativ, da $(V, +)$ kommutativ ist. Also gilt (V1). Wegen $U \subset V$ folgt auch (V2).

Beispiele

- (a) Wegen $0 + 0 = 0, \lambda 0 = 0$ für $\lambda \in K$ (4.1.2(ii)), ist $\{0\}$ (mit den »geerbten« Verknüpfungen) Untervektorraum jedes Vektorraums.

- (b)

$$V = \mathbb{R}^2, U = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1, x_2 \in \mathbb{R} \wedge x_1 \geq 0 \wedge x_1 - x_2 \leq 0 \right\}$$

U ist kein Untervektorraum. Es gelten zwar (4.1.3(i)) und (4.1.3(ii)), aber (4.1.3(iii)) gilt nur für $\lambda \geq 0$. Werte $\lambda < 0$ führen zu einer Spiegelung am Ursprung. U ist ein KEGEL.

- (c) Sei $K[x]_n$ die Menge aller Polynome vom Grad höchstens n . $K[x]_n$ ist *kein* Unterring von $K[x]$, da die Multiplikation von Polynomen den Grad i.a. erhöht. Aber $K[x]_n$ ist ein Untervektorraum des K -Vektorraums $K[x]$. ($K[x]_0 \subset_{\text{UVR}} K[x]_1 \subset_{\text{UVR}} K[x]_2 \subset_{\text{UVR}} \dots \subset_{\text{UVR}} K[x]_n \subset_{\text{UVR}} \dots$)

4.1.4 Lemma

Seien V ein K -Vektorraum, I eine beliebige Indexmenge und für jedes $\alpha \in I$ sei U_α ein Untervektorraum von V . Dann ist $U := \bigcap_{\alpha \in I} U_\alpha$ ein Untervektorraum von V .

(Untervektorraumbegriff ist abgeschlossen unter beliebiger Durchschnittsbildung.)

Beweis

Es gilt $0 \in U_\alpha \forall \alpha \in I$, also $0 \in U \neq \emptyset$. Seien $u_1, u_2 \in U$, so gilt $u_1, u_2 \in U_\alpha$ für jedes $\alpha \in I$, also $u_1 + u_2 \in U_\alpha$ für alle $\alpha \in I$ (denn U_α Untervektorraum) und somit $u_1 + u_2 \in U$. Sind $\lambda \in K, u \in U$, so ist $u \in U_\alpha$ für $\alpha \in I$, also $\lambda u \in U_\alpha$ für $\alpha \in I$, somit $\lambda u \in U$.

Beachte

Die Vereinigung von Untervektorräumen ist i.a. *kein* Untervektorraum.

$$U_1 = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix} := \left\{ \begin{pmatrix} x \\ x \end{pmatrix} \in \mathbb{R} \right\}$$

$$U_2 = \mathbb{R} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$U_1 \cap U_2$ ist nach (4.1.4) Untervektorraum, $U_1 \cup U_2$ erfüllt Bedingung (4.1.3(ii)) *nicht*. (4.1.4) erlaubt es, einen HÜLLENOPERATOR zu definieren (analog zur abgeschlossenen Hülle der Analysis).

Definition

Seien V ein K -Vektorraum und $M \subset V$. Die LINEARE HÜLLE $\text{lin}(M)$ ist der Durchschnitt aller Untervektorräume von V , die M enthalten.

Bemerkung

$\text{lin}(M)$ ist Untervektorraum von V .

Beispiel

Sei $U_1 = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, U_2 = \mathbb{R} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Dann ist $\text{lin}(U_1 \cup U_2) = \mathbb{R}^2$.

In der Analysis erhält man den Abschluss einer Menge durch Hinzunahme ihrer Häufungspunkte. Im Folgenden soll untersucht werden, wie $\text{lin}(M)$ »konstruktiv« aus den Elementen von M aufgebaut werden kann.

Definition

Sei V ein K -Vektorraum. Sind $r \in \mathbb{N}, v_1, \dots, v_r \in V, \lambda_1, \dots, \lambda_r \in K$, so heißt $\lambda_1 v_1 + \dots + \lambda_r v_r$ LINEARKOMBINATION von v_1, \dots, v_r . Sei $M \subset V$. Das LINEARE ERZEUGNIS $\text{span}(M)$ von M sei definiert durch $\text{span}(M) = \{0\}$ für $M = \emptyset$ und für $M \neq \emptyset$ durch

$$\text{span}(M) = \left\{ v : \exists r \in \mathbb{N} \exists v_1, \dots, v_r \in M \wedge \exists \lambda_1, \dots, \lambda_r \in K : v = \lambda_1 v_1 + \dots + \lambda_r v_r \right\}$$

Bemerkung

- (a) Für Linearkombinationen werden nur *endliche* (aber beliebig große) Summen $\sum_{i=1}^r \lambda_i v_i$ betrachtet.
- (b) $0 \in \text{span}(M)$. Für $M \neq \emptyset$ ist $\text{span}(M)$ die Menge aller Linearkombinationen von Elementen in M .

Beispiel

Seien $U_1 = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, U_2 = \mathbb{R} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ($M = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$)

$r = 1$ $\lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \lambda \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rightarrow U_1 \cup U_2$

$r = 2$ Da U_1, U_2 Untervektorräume von \mathbb{R}^2 sind, gilt

$$\lambda_1, \lambda_2 \in K \wedge (u_1, u_2 \in U_1 \vee u_1, u_2 \in U_2) \Rightarrow \lambda_1 u_1 + \lambda_2 u_2 \in U_1 \vee U_2$$

Neue Vektoren können also nur von der Form $\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ($\lambda_1, \lambda_2 \in K$) sein. Welche Vektoren $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ werden so erzeugt? Löse Gleichungssystem $\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, also $\lambda_1 - \lambda_2 = v_1, \lambda_1 + \lambda_2 = v_2$. Lösung: $\lambda_1 = \frac{v_1+v_2}{2}, \lambda_2 = \frac{v_2-v_1}{2}$. Somit lässt sich jeder Vektor des \mathbb{R}^2 als Linearkombination von $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ darstellen. Es folgt $\text{span}(M) = \mathbb{R}^2$.

$r \geq 3$ Da \mathbb{R}^2 ein Vektorraum ist, kommen keine weiteren Vektoren hinzu.

Insgesamt folgt: $\text{span}(U_1 \cup U_2) = \text{span}(\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}\}) = \mathbb{R}^2$.

Beobachtungen

(a) $\text{lin}(U_1 \cup U_2) = \text{span}(U_1 \cup U_2)$

(b) Zur Erzeugung von $\text{span}(U_1 \cup U_2)$ »reichen« zwei Vektoren.

Die Beobachtung (b) führt auf den zentralen Begriff der BASIS eines Vektorraums, Beobachtung (a) gilt allgemein.

4.1.5 Satz

Seien V ein Vektorraum und $M \subset V$. Dann gilt $\text{lin}(M) = \text{span}(M)$.

Beweis

Da $\text{lin}(M)$ ein Vektorraum ist, liegen insbesondere alle Linearkombinationen von Elementen aus M in $\text{lin}(M)$, d.h. es gilt $\text{lin}(M) \supset \text{span}(M)$.

Nach Definition von $\text{lin}(M)$ reicht es zu zeigen, dass $\text{span}(M)$ ein Untervektorraum von V ist. Für $M = \emptyset$ ist das klar, für $M \neq \emptyset$ folgt das unmittelbar aus (4.1.3).

4.2 Lineare (Un-)Abhängigkeit, Basis, Dimension

Frage: Wie kann man eine »möglichst kleine« Menge von Vektoren angeben, die V linear erzeugen?

Beispiele

(a) $\mathbb{R}^2 = \text{span}(\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}\})$ (ein Vektor reicht nicht)

(b) $\mathbb{R}^2 = \text{span}(\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\})$

(c) Für $i = 1, \dots, n$ sei $e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^n$ (die 1 jeweils an der i-ten Stelle).

e_i heißt i-ter KOORDINATENVEKTOR des K^n . Es gilt

$$K^n = \text{span}(\{e_1, \dots, e_n\})$$

(d) Für $i \in \mathbb{N}_0$ sei $v_i \in K[x]$ mit $v_i(x) = x^i$. v_i heißt i-tes MONOM. Es gilt $K[x] = \text{span}(\{v_i : i \in \mathbb{N}_0\})$. Endlich viele Polynome erzeugen $K[x]$ *nicht*.

Definition

Sei V ein K -Vektorraum. Eine Teilmenge E von V heißt ERZEUGENDENSYSTEM, wenn $\text{span}(E) = V$ gilt. Ist E ein Erzeugendensystem, so heißt E MINIMAL, wenn gilt: $v \in E \Rightarrow \text{span}(E \setminus \{v\}) \neq V$.

Existiert zu V ein endliches Erzeugendensystem, so heißt V ENDLICH ERZEUGT.

Beispiele

(i) Die obigen Beispiele (a), ..., (d) sind minimale Erzeugendensysteme.

(ii) Es gilt $\mathbb{R}^2 = \text{span}\left(\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}\right)$. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ sind *kein* minimales Erzeugendensystem, denn es gilt $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} -1 \\ -1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, d.h. $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Rightarrow \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = (\lambda_2 - \frac{1}{2}\lambda_1)\begin{pmatrix} -1 \\ -1 \end{pmatrix} + (\lambda_3 - \frac{1}{2}\lambda_1)\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Eine »symmetrischere« Art, obige »Abhängigkeit« der drei Vektoren darzustellen, ist:

$$1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$$

Natürlich gilt auch:

$$0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$$

d.h. das Erzeugendensystem ist nicht minimal, weil es eine nichttriviale Linearkombination gibt, die den Nullvektor erzeugt.

Definition

Seien V ein K -Vektorraum und $v_1, \dots, v_n \in V$. v_1, \dots, v_n heißen LINEAR UNABHÄNGIG, falls gilt:

$$\lambda_1, \dots, \lambda_n \in K \wedge \sum_{k=1}^n \lambda_k v_k = 0 \Rightarrow \lambda_1 = \dots = \lambda_n = 0$$

Andernfalls heißen v_1, \dots, v_n LINEAR ABHÄNGIG.

Beispiele

(a) $\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ sind linear abhängig in \mathbb{R}^2 .

(b) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ sind linear unabhängig in \mathbb{R}^2 .

(c) v, v sind linear abhängig in V , ebenso 0.

(d) Die Koordinatenvektoren e_1, \dots, e_n aus K^n sind linear unabhängig in K^n .

(e) Die Monome aus $K[x]_n$ sind linear unabhängig in $K[x]$.

(f) $1, \sqrt{2}$ sind linear abhängig im \mathbb{R} -Vektorraum \mathbb{R} , aber linear unabhängig im \mathbb{Q} -Vektorraum \mathbb{R} , denn $\lambda_1, \lambda_2 \in \mathbb{Q} \wedge \lambda_1 + \lambda_2 \sqrt{2} = 0 \Rightarrow (\lambda_1 = \lambda_2 = 0) \vee (-\frac{\lambda_1}{\lambda_2} = \sqrt{2}) \Rightarrow \lambda_1 = \lambda_2 = 0$.

Das Beispiel $K[x] = \text{span}(\{v_i; i \in \mathbb{N}_0\})$ zeigt, dass es nicht reicht, die lineare (Un-)Abhängigkeit nur für *endlich* viele Vektoren zu definieren.

Definition

Seien V ein K -Vektorraum, $M \subseteq V$. Dann heißt M LINEAR ABHÄNGIG, wenn es ein $r \in \mathbb{N}$ und $v_1, \dots, v_r \in M$ gibt, so dass v_1, \dots, v_r linear abhängig sind. Andernfalls heißt M LINEAR UNABHÄNGIG.

Bemerkung

Gilt $S \subset M \subset V$ und ist M linear unabhängig, so ist auch S linear unabhängig.

4.2.1 Lemma

Sei V ein K -Vektorraum, $M \subset V$. M ist genau dann linear unabhängig, wenn $M \neq \emptyset$ oder wenn sich jeder Vektor v aus $\text{span}(M)$ auf *genau* eine Weise als Linearkombination von Vektoren aus M darstellen lässt.

Beweis

Der Fall $M = \emptyset$ ist trivial. Sei also $M \neq \emptyset$ linear unabhängig. Ferner seien $v \in \text{span}(M)$, $r, s \in \mathbb{N}$, $v_1, \dots, v_r, u_1, \dots, u_s \in V$, $\lambda_1, \dots, \lambda_r, \mu_1, \dots, \mu_s \in K$ mit $v = \sum_{i=1}^r \lambda_i v_i = \sum_{i=1}^s \mu_i u_i$. Also gilt $\sum_{i=1}^r \lambda_i v_i + \sum_{i=1}^s (-\mu_i) u_i = 0$. O. B. d. A. sei $t \in \mathbb{N}_0$ mit $\{v_1, \dots, v_r\} \cap \{u_1, \dots, u_s\} = \{v_1, \dots, v_t\}$. Also ist $\sum_{i=0}^t (\lambda_i - \mu_i) v_i + \sum_{i=t+1}^r \lambda_i v_i + \sum_{i=t+1}^s \mu_i u_i = 0$. Da $\{v_i, \dots, v_r\} \cup \{u_1, \dots, u_s\} \subset M$ und M linear unabhängig folgt $\lambda_1 = \mu_1, \dots, \lambda_t = \mu_t, \lambda_{t+1} = \dots = \lambda_r = \mu_{t+1} = \dots = \mu_s = 0$. Ist andererseits M linear abhängig, so besitzt 0 eine nichttriviale Darstellung, d.h. 0 ist nicht eindeutig darstellbar.

Beispiel

Seien $b \in K^m$, $A \in K^{m \times n}$, \hat{A}, \hat{b} eine zugehörige Matrix in Zeilenstufenform bzw. die zugehörige rechte Seite.

Zeilenvektoren: Die ersten r Zeilenvektoren von \hat{A} sind linear unabhängig. Alle Zeilenvektoren von A werden von den zu diesen r gehörigen linear erzeugt. (\gg Rückgängigmachen \ll der elementaren (linearen) Zeilenoperationen). Der Zeilenrang ist also das Maximum der Kardinalität linear unabhängiger Teilmengen der Zeilenvektoren von A .

Spaltenvektoren: Die spalten von \hat{A} mit Indizes aus $\{j_1, \dots, j_r\}$ sind linear unabhängig. \hat{b} ist *eindeutig* als Linearkombination dieser Vektoren darstellbar, wenn das Gleichungssystem lösbar ist (vgl. Beweis (2.2.1) auf Seite 11).

Bemerkung

Die Ergebnisse von Kapitel 2 gelten (mit den selben Beweisen) für Gleichungssysteme über beliebigen Körpern.

Definition

Seien V ein K -Vektorraum und $B \subset V$. B heißt BASIS des Vektorraums V , wenn gilt:

- (i) $\text{span}(B) = V$
- (ii) B ist linear unabhängig

Ist B endlich, so heißt die Anzahl $|B|$ ($\text{card}(B)$) der Elementen von B LÄNGE DER BASIS.

Wir beweisen zunächst eine fundamentale Charakterisierung von Basen.

4.2.2 Satz

Seien V ein K -Vektorraum und $B \subset V$. Dann sind die folgenden drei Aussagen äquivalent:

- (i) B ist eine Basis von V .
- (ii) B ist ein minimales Erzeugendensystem für V .
- (iii) B ist eine (inklusions-)maximale linear unabhängige Teilmenge von V , d.h. $B \cup \{v\}$ ist linear abhängig für alle $v \in V \setminus B$.

Beweis

Ist $V = \{0\}$, so ist B in allen drei Fällen \emptyset . Wir setzen daher im Folgenden $V \neq \{0\}$ voraus.

$(i) \Rightarrow (ii)$: Sei B eine Basis von V . Ist das Erzeugendensystem B nicht minimal, so existiert ein $b_0 \in B$ mit $V = \text{span}(B \setminus \{b_0\})$, d.h. es existiert $r \in \mathbb{N}, b_1, \dots, b_r \in B \setminus \{b_0\}, \lambda_1, \dots, \lambda_r \in K$ mit $b_0 = \sum_{i=1}^r \lambda_i b_i$. Somit wäre B linear abhängig. Widerspruch.

$(ii) \Rightarrow (iii)$: Sei B ein minimales Erzeugendensystem. Wäre B linear abhängig, so gäbe es $r \in \mathbb{N}, b_1, \dots, b_r \in B, \lambda_1, \dots, \lambda_r \in K \setminus \{0\}$ mit $\sum_{i=1}^r \lambda_i b_i = 0$. Es folgt $b_1 = -\frac{1}{\lambda_1} \sum_{i=2}^r \lambda_i b_i$. Somit gilt $\text{span}(B \setminus \{b_1\}) = V$. Widerspruch zur Minimalität von B . Ist $v \in V \setminus B$, so folgt $v \in \text{span}(B)$, also ist $B \cup \{v\}$ linear abhängig.

$(iii) \Rightarrow (i)$: Sei B eine maximale linear unabhängige Teilmenge von V . Ist $b_0 \in V \setminus B$, so ist $B \cup \{b_0\}$ linear abhängig, d.h. es gibt $r \in \mathbb{N}, b_1, \dots, b_r \in B, \lambda_0, \dots, \lambda_r \in K$, nicht alle 0, mit der Eigenschaft $\sum_{i=0}^r \lambda_i b_i = 0$. Da B linear unabhängig ist, folgt $\lambda_0 \neq 0$ und damit $b_0 = -\frac{1}{\lambda_0} \sum_{i=1}^r \lambda_i b_i$. Da $b_0 \in V \setminus B$ beliebig gewählt war, folgt $V = \text{span}(B)$.

4.2.3 Korollar

Seien V ein endlich erzeugter Vektorraum und E ein endliches Erzeugendensystem. Dann besitzt V eine Basis B mit $B \subset E$.

Beweis

Die Aussage folgt aus der Äquivalenz von (i), (ii) in (4.2.2) durch sukzessives Entfernen von abhängigen Elementen.

Frage

Besitzt jeder Vektorraum (auch solche, die nicht endlich erzeugt sind) eine Basis?

Definition

Seien M eine Menge. Eine Menge von Teilmengen von M heißt MENGENSYSTEM (über M). Seien \mathcal{S} ein nichtleeres Mengensystem, K eine nichtleere Teilmenge von \mathcal{S} . K heißt KETTE, wenn gilt

$$S_1, S_2 \in K \Rightarrow S_1 \subset S_2 \vee S_1 \supset S_2$$

Sei $T \in \mathcal{S}$. T heißt MAXIMALES ELEMENT, wenn gilt

$$S \in \mathcal{S} \wedge T \subset S \Rightarrow T = S$$

4.2.4 Satz (Lemma von Zorn)

Sei \mathcal{S} ein nichtleeres Mengensystem. Gilt für jede Kette K von \mathcal{S} , dass $\bigcup_{S \in K} S \in \mathcal{S}$, so gibt es in \mathcal{S} ein maximales Element.

Beweis

siehe z.B. D. Klaua, Allgemeine Mengenlehre, 1964

4.2.5 Satz

Seien V ein K -Vektorraum und A eine linear unabhängige Teilmenge von V . Dann gibt es eine Basis B von V mit $A \subset B$.

Beweis

Sei \mathcal{S} das Mengensystem aller linear unabhängigen Teilmengen von V . Wegen $A \in \mathcal{S}$ ist $\mathcal{S} \neq \emptyset$. Ferner seien K eine beliebige Kette von \mathcal{S} und $T := \bigcup_{S \in K} S$. Wir zeigen zunächst, dass T wieder unabhängig ist. Wäre T nicht linear unabhängig, so gäbe es ein $r \in \mathbb{N}$, $\{t_1, \dots, t_r\} \subset T$, so dass t_1, \dots, t_r linear abhängig sind. Für $i = 1, \dots, r$ sei $T_i \in K$ mit $t_i \in T_i$. Da K eine Kette ist, existiert ein i_0 , so dass $T_i \subset T_{i_0}$ für alle $i = 1, \dots, r$. (Warum?) T_{i_0} wäre somit linear abhängig; Widerspruch zu $T_{i_0} \in \mathcal{S}$. Also ist T linear unabhängig, d.h. $T \in \mathcal{S}$. Nach (4.2.4) existiert in \mathcal{S} ein maximales Element B . Aus (4.2.2) folgt die Behauptung.

4.2.6 Lemma

Seien V ein K -Vektorraum und $B = \{v_1, \dots, v_n\}$ eine Basis. Seien $\eta_1, \dots, \eta_n \in K$, $I = \{i; i \in \{1, \dots, n\} \wedge \eta_i \neq 0\}$, und $v_0 = \sum_{j=1}^n \eta_j v_j$. Dann ist für jedes $i \in I$ $(B \setminus \{v_i\}) \cup \{v_0\}$ eine Basis von V .

Beweis

Sei $i \in I$ und $\bar{B} = (B \setminus \{v_i\}) \cup \{v_0\}$. Seien $v \in V$ und $\mu_1, \dots, \mu_n \in K$ mit $v = \sum_{j=1}^n \mu_j v_j$. Dann gilt wegen $v_i = \frac{1}{\eta_i} v_0 + \sum_{\substack{j=1 \\ j \neq i}}^n (-\frac{\eta_j}{\eta_i}) v_j$

$$v = \frac{\mu_i}{\eta_i} v_0 + \sum_{\substack{j=1 \\ j \neq i}}^n (\mu_j - \frac{\eta_j}{\eta_i} \mu_i) v_j$$

Also ist \bar{B} ein Erzeugendensystem.

Sei $\lambda_0 v_0 + \sum_{\substack{j=1 \\ j \neq i}}^n \lambda_j v_j = 0$. Es folgt $0 = \lambda_0 \sum_{j=1}^n \eta_j v_j + \sum_{\substack{j=1 \\ j \neq i}}^n \lambda_j v_j = \lambda_0 \eta_i v_i +$

$\sum_{\substack{j=1 \\ j \neq i}}^n (\lambda_0 \eta_j + \lambda_j) v_j$ und somit für $j = 1, \dots, n, j \neq i$ $\lambda_0 \eta_i = (\lambda_0 \eta_j + \lambda_j) = 0$.

Wegen $i \in I$ ist $\eta_i \neq 0$, also gilt $\lambda_0 = 0$ und somit $\lambda_j = 0$ für alle $j = 1, \dots, n, j \neq i$.

4.2.7 Satz (Austauschsatz von Steinitz)

Seien V ein K -Vektorraum, $B = \{v_1, \dots, v_n\}$ eine Basis und $w_1, \dots, w_r \in V$ linear unabhängig. Dann gilt $n \geq r$ und es gibt $i_1, \dots, i_r \in \{1, \dots, n\}$, so dass $\bar{B} = \{v_i; i \notin \{i_1, \dots, i_r\}\} \cup \{w_1, \dots, w_r\}$ eine Basis von V ist.

Beweis

Induktion nach r : Für $r = 0$ ist nichts zu zeigen. Sei $r \geq 1$. Da insbesondere w_1, \dots, w_{r-1} linear unabhängig sind, gibt es nach Induktionsannahme $i_1, \dots, i_r \in \{1, \dots, n\}$, so dass $\{v_i; i \notin \{i_1, \dots, i_{r-1}\}\} \cup \{w_1, \dots, w_{r-1}\}$ eine Basis von V ist. Ferner gilt nach Induktionsannahme $n \geq r - 1$. Zum Beweis von $n \geq r$ nehmen wir an, dass $n = r - 1$ ist. Dann ist $\{w_1, \dots, w_{r-1}\}$ eine Basis von V . Andererseits sind w_1, \dots, w_r linear unabhängig. Widerspruch zu (4.2.2). Also gilt $n \geq r$. O.B.d.A. können wir annehmen, dass $\{i_1, \dots, i_{r-1}\} = \{1, \dots, r - 1\}$ gilt. Da $\{v_i : i = r, \dots, n\} \cup \{w_1, \dots, w_{r-1}\}$ eine Basis ist, gibt es $\lambda_1, \dots, \lambda_n \in K$ mit $w_r = \sum_{i=1}^{r-1} \lambda_i w_i + \sum_{i=r}^n \lambda_i v_i$. Da w_1, \dots, w_r linear unabhängig sind, gibt es ein $i \in \{r, \dots, n\}$ mit $\lambda_i \neq 0$. Die Behauptung folgt nun aus (4.2.6).

4.2.8 Satz

Sei V ein endlich erzeugter K -Vektorraum. Dann gibt es ein $n \in \mathbb{N}_0$, so dass *jede* Basis von V endlich ist und Länge n hat.

Beweis

Da V endlich erzeugt ist, besitzt V nach (4.2.3) eine endliche Basis B . Sei \bar{B} eine zweite Basis, so ist \bar{B} linear unabhängig und aus (4.2.7) folgt für jede endliche Teilmenge T von \bar{B} $|T| \leq |B|$. Somit ist \bar{B} endlich, und es gilt $|\bar{B}| \leq |B|$. Vertauschung der Rollen von B und \bar{B} in (4.2.7) liefert $|B| \leq |\bar{B}|$.

Korollar

(4.2.8) erlaubt es, die Dimension eines endlich erzeugten Vektorraumes zu definieren.

Definition

Sei V ein K -Vektorraum, B eine Basis von V . Dann heißt

$$\dim V := \begin{cases} |B| & \text{falls } B \text{ endlich} \\ \infty & \text{sonst} \end{cases}$$

die DIMENSION von V .

Soll der Körper K besonders betont werden, so schreibt man auch $\dim_K V$.

Beispiele

- (a) $\dim K^n = n$
- (b) $\dim K[x] = \infty$
- (c) $\dim_{\mathbb{R}} \mathbb{R} = 1$; $\dim_{\mathbb{Q}} \mathbb{R} = \infty$, da \mathbb{Q} abzählbar ist, \mathbb{R} aber nicht.

4.2.9 Lemma

Seien V ein endlich erzeugter K -Vektorraum und U ein Untervektorraum von V . Dann ist auch U endlich erzeugt, und es gilt $\dim U \leq \dim V$. Gilt $\dim U = \dim V$, so folgt $U = V$.

Beweis

Die erste Aussage folgt aus (4.2.2) und (4.2.5); die zweite aus (4.2.7) und die dritte mit (4.2.2) ebenfalls aus (4.2.7).

4.3 Summen von Vektorräumen

Beispiel

$$U_1 = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad U_2 = \mathbb{R} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbb{R}^2 = (U_1 \cup U_2) = \{x : \exists u_1 \in U_1 \wedge \exists u_2 \in U_2 : x = u_1 + u_2\}$$

Definition

Sei V ein K -Vektorraum, I eine beliebige Indexmenge und für jedes $\alpha \in I$ sei U_α Untervektorraum von V . Die SUMME $\sum_{\alpha \in I} U_\alpha$ der Untervektorräume U_α ist definiert durch

$$\sum_{\alpha \in I} U_\alpha = \left\{ \sum_{\alpha \in I} u_\alpha : (\alpha \in I \Rightarrow u_\alpha \in U_\alpha) \wedge u_\alpha \neq 0 \text{ für höchstens endlich viele } \alpha \right\}$$

$\sum_{\alpha \in I} U_\alpha$ heißt DIREKT, wenn für jedes $\alpha \in I$ gilt:

$$U_\alpha \neq \{0\} \quad \text{und} \quad U_\alpha \cap \sum_{\beta \in I \setminus \{\alpha\}} U_\beta = \{0\}$$

Bezeichnung:

$$\bigoplus_{\alpha \in I} U_\alpha$$

Bezeichnung

Ist I endlich, etwa $I = \{1, \dots, r\}$, so werden die Bezeichnungen $U_1 + \dots + U_r$ bzw. $U_1 \oplus \dots \oplus U_r$ verwendet.

Bemerkung

$$\sum_{\alpha \in I} U_\alpha = \text{span} \left\{ \bigcup_{\alpha \in I} U_\alpha \right\}$$

4.3.1 Lemma

Seien V ein K -Vektorraum und U_1, U_2 endlich - dimensionale Untervektorräume von V . Dann gilt:

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

Beweis

Seien $B_{1,2}$ eine Basis von $U_1 \cap U_2$, B_1, B_2 Basen von U_1 bzw. U_2 mit $B_{1,2} \subset B_1, B_2$. (Solche Basen existieren nach (4.2.5)). Natürlich erzeugen $B_1 \setminus B_{1,2}$, $B_2 \setminus B_{1,2}$ und $B_{1,2}$ zusammen $U_1 + U_2$. Die Behauptung folgt, wenn gezeigt ist, dass diese Vektoren linear unabhängig sind. Sei

$$\sum_{a \in B_1 \setminus B_{1,2}} \lambda_a a + \sum_{b \in B_2 \setminus B_{1,2}} \lambda_b b + \sum_{c \in B_{1,2}} \lambda_c c = 0$$

Es gilt

$$v := \sum_{a \in B_1 \setminus B_{1,2}} \lambda_a a + \sum_{b \in B_{1,2}} \lambda_b b = - \sum_{c \in B_2 \setminus B_{1,2}} \lambda_c c = 0$$

also $v \in U_1$, $-v \in U_2$, somit $v \in U_1 \cap U_2$. Es gibt also für jedes $b \in B_{1,2}$ ein $\mu_b \in K$ mit $v = \sum_{b \in B_{1,2}} \mu_b b$. Wegen (4.2.1) folgt $\lambda_a = 0$ für alle $a \in B_1 \setminus B_{1,2}$, sowie $\lambda_c = 0$ für $c \in B_2 \setminus B_{1,2}$. Somit folgt auch $\lambda_b = 0$ für $b \in B_{1,2}$.

4.3.2 Satz

Seien V ein K -Vektorraum, I eine Indexmenge und für jedes $\alpha \in I$ U_α ein Untervektorraum von V mit $U_\alpha \neq \{0\}$. $\sum_{\alpha \in I} U_\alpha$ ist genau dann direkt, wenn sich jeder von 0 verschiedene Vektor von $\sum_{\alpha \in I} U_\alpha$ auf genau eine Weise in der Form $u_{\alpha_1} + \dots + u_{\alpha_r}$ mit paarweise verschiedenen Indizes $\alpha_1, \dots, \alpha_r \in I$ und Vektoren $u_{\alpha_i} \in U_{\alpha_i} \setminus \{0\}$ ($i = 1, \dots, r$) schreiben lässt.

Beweis

Sei $\sum_{\alpha \in I} U_\alpha$ direkt und $\sum_{i=1}^r u_{\alpha_i} = \sum_{j=1}^s v_{\beta_j}$ zwei entsprechende Darstellungen. Gilt $\alpha_i \notin \{\beta_1, \dots, \beta_s\}$ für ein $i \in \{1, \dots, r\}$, so folgt

$$0 \neq u_{\alpha_i} = \sum_{j=1}^s v_{\beta_j} - \sum_{\substack{j=1 \\ j \neq i}}^r u_{\alpha_j} \in U_{\alpha_i} \cap \sum_{\substack{\alpha \in I \\ \alpha \neq \alpha_i}} U_\alpha$$

im Widerspruch zur Direktheit der Summe. Also folgt $r = s$ und bei geeigneter Numerierung $\alpha_i = \beta_i$ für $i = 1, \dots, r$. Daher gilt für $i = 1, \dots, r$

$$u_{\alpha_i} - v_{\alpha_i} = \sum_{\substack{j=1 \\ j \neq i}}^r (v_{\alpha_j} - u_{\alpha_j}) \in U_{\alpha_i} \cap \sum_{\substack{\alpha \in I \\ \alpha \neq \alpha_i}} U_\alpha$$

also gilt $u_{\alpha_i} = v_{\alpha_i}$.

Ist andererseits $\sum_{\alpha \in I} U_\alpha$ nicht direkt, dann gibt es ein $\alpha \in I$ mit

$$U_\alpha \cap \sum_{\beta \in I \setminus \{\alpha\}} U_\beta \ni u_\alpha \neq 0$$

Also gibt es $r \in \mathbb{N}$, $\beta_1, \dots, \beta_r \in I \setminus \{\alpha\}$, $U_{\beta_i} \in U_{\beta_i} \setminus \{0\}$ für $i = 1, \dots, r$ mit $u_\alpha = \sum_{i=1}^r u_{\beta_i}$; das sind zwei verschiedene Darstellungen.

4.3.3 Bemerkung

Sei V ein K -Vektorraum und $M \subset V$ mit $M \neq \emptyset$ und $0 \notin M$. M ist genau dann linear unabhängig, wenn $\sum_{v \in M} K v$ direkt ist. M ist genau dann eine Basis von V , wenn gilt $V = \bigoplus_{v \in M} K v$.

Beweis

Die Behauptungen folgen aus (4.2.1) und (4.3.2).

Definition

Seien V ein K -Vektorraum und U, W Untervektorräume von V . W heißt KOMPLEMENT von U , falls $V = U + W$ und $U \cap W = \{0\}$.

Beispiel

(a) $V = K^2$, $U = \left\{ \begin{pmatrix} a \\ 0 \end{pmatrix} : a \in K \right\}$, $W_1 = \left\{ \begin{pmatrix} 0 \\ b \end{pmatrix} : b \in K \right\}$, $W_2 = \left\{ \begin{pmatrix} c \\ c \end{pmatrix} : c \in K \right\}$. W_1 und W_2 sind Komplemente von U .

(b) $\{0\}$ ist Komplement von V in V .

4.3.4 Satz

Seien V ein K -Vektorraum und U ein Untervektorraum. Dann besitzt U ein Komplement in V .

Beweis

Folgt aus (4.2.5) und (4.3.3).

Kapitel 5

Lineare Abbildungen und Matrizen

Bereits früher haben wir das Lösen eines linearen Gleichungssystems $Ax = b$ implizit als Frage aufgefasst, ob der Vektor $b \in K^m$ als Bild Ax eines Vektors $x \in K^n$ dargestellt werden kann, d.h. wir haben implizit die Abbildung $x \mapsto Ax$ betrachtet. Eine geometrische Interpretation dieser Abbildung für spezielle Matrizen wurde in H15, H20 gegeben (Drehung, Spiegelung). Es zeigt sich, dass diese Abbildungen gerade den »Vektorraum-Homomorphismen« zwischen endlich - dimensionalen K -Vektorräumen entsprechen.

5.1 Lineare Abbildungen

Definition

Seien K ein Körper und V, W K -Vektorräume. Eine Abbildung $F : V \rightarrow W$ heißt LINEAR (oder (Vektorraum-)HOMOMORPHISMUS), wenn gilt:

$$(L1) \quad u, v \in V \Rightarrow F(u + v) = F(u) + F(v)$$

$$(L2) \quad v \in V \wedge \lambda \in K \Rightarrow F(\lambda v) = \lambda F(v)$$

5.1.1 Bemerkung

Die Abbildung $F : V \rightarrow W$ ist genau dann linear, wenn gilt:

$$u, v \in W \wedge \lambda, \mu \in K \Rightarrow F(\lambda u + \mu v) = \lambda F(u) + \mu F(v)$$

Beweis trivial.

Beispiele

(a) Sei $A \in K^{m \times n}$. Die Abbildung $F_A : K^n \rightarrow K^m$, definiert durch $F_A(x) = Ax$ ist linear.

(b) Die Abbildung $F : K^{m \times n} \rightarrow K^{n \times m}$ sei definiert durch

$$(a_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} \mapsto (a_{ji})_{\substack{j=1, \dots, n \\ i=1, \dots, m}}$$

Dann ist F linear.

- (c) Seien V der \mathbb{R} -Vektorraum der einmal stetig differenzierbaren Funktionen auf einem (offenen) reellen Intervall I und W der \mathbb{R} -Vektorraum der auf I stetigen Funktionen. Dann ist die Abbildung $F : V \rightarrow W$, definiert durch $F(f) = f' (= \frac{d}{dx}f)$ linear.
- (d) Analog zum obigen »Differentialoperator« ist (auf geeigneten Räumen) auch der »Integraloperator« linear, denn

$$\int_M \lambda f(x) + \mu g(x) dx = \lambda \int_M f(x) dx + \mu \int_M g(x) dx$$

Bezeichnung

Sei $A \in K^{m \times n}$. Dann wird mit F aus Beispiel (b) $F(A)$ als A^T bezeichnet und die TRANSPONIERTE von A genannt.

Definition

Sei $F : V \rightarrow W$ linear. F heißt (Vektorraum-)ISOMORPHISMUS, wenn F bijektiv ist. Ist $V = W$, so heißt F ENDOMORPHISMUS. Gilt $V = W$ und ist F bijektiv, so heißt F AUTOMORPHISMUS.

Beispiel

Die Transformation $F_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, definiert durch $x \mapsto Ax$ mit $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ ist linear. Ferner ist F surjektiv, denn der Zeilenrang von A ist 2 (vgl. (2.2.1) und (4.2)). F ist injektiv, denn $Ax = 0$ hat nur die triviale Lösung $x = 0$. Also ist F ein Automorphismus des \mathbb{R}^2 . (vgl. (5.2.10)). Natürlich gilt $Ax = x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 + x_2 \end{pmatrix} = (x_1 - x_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (x_1 + x_2) \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, d.h. die Koordinaten x eines Vektors bezüglich der Basis $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ werden in die Koordinaten des selben Vektors bezüglich der kanonischen Basis $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ überführt. F_A beschreibt also eine KOORDINATENTRANSFORMATION.

5.1.2 Lemma

Sei $F : V \rightarrow W$ linear. Dann gilt:

- (i) $n \in \mathbb{N} \wedge v_1, \dots, v_n \in V \wedge \lambda_1, \dots, \lambda_n \in K \Rightarrow F(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i F(v_i)$
- (ii) Sei I eine Indexmenge, $v_\alpha \in V$ für $\alpha \in I$ und $(v_\alpha)_{\alpha \in I}$ linear abhängig. Dann ist $(F(v_\alpha))_{\alpha \in I}$ linear abhängig in W .
- (iii) Seien \bar{V} Untervektorraum von V und \bar{W} Untervektorraum von W . Dann sind $F(\bar{V})$ ein Untervektorraum von W und das Urbild

$$F^{-1}(\bar{W}) = \{v : v \in V \wedge F(v) \in \bar{W}\}$$

ein Untervektorraum von V .

- (iv) $\dim(F(V)) \leq \dim V$

- (v) Ist F bijektiv, so ist die Umkehrabbildung $F^{-1} : W \rightarrow V$ linear.

Beweis

(i) Folgt aus (5.1.1) durch vollständige Induktion.

$$(ii) \sum_{i=1}^n \lambda_i v_i = 0 \stackrel{(L2)}{\Rightarrow} 0 = F(0) = F\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i F(v_i)$$

(iii) Wir überprüfen die Bedingungen von (4.1.3).

Wegen $0 \in \bar{V}$ gilt $0 = F(0) \in F(\bar{V}) \neq \emptyset$.

Seien $w_1, w_2 \in F(\bar{v})$. Ferner seien $v_1, v_2 \in \bar{V}$ mit $w_i = F(v_i)$ ($i = 1, 2$). Dann gilt $w_1 + w_2 = F(v_1) + F(v_2) = F(v_1 + v_2) \in F(\bar{V})$.

Seien $\lambda \in K, w \in F(\bar{V})$ und $v \in \bar{V}$ mit $w = F(v)$. Dann gilt $\lambda w = \lambda F(v) = F(\lambda v) \in F(\bar{V})$.

Somit ist $F(\bar{V})$ Untervektorraum von W . Die zweite Aussage folgt analog.

(iv) folgt aus (ii)

(v) Seien $\lambda_1, \lambda_2 \in K, w_1, w_2 \in W$ und $v_1, v_2 \in V$ mit $w_i = F(v_i)$ ($i = 1, 2$). Dann gilt $v_i = F^{-1}(w_i)$ ($i = 1, 2$) und es folgt aus $F(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 F(v_1) + \lambda_2 F(v_2) = \lambda_1 w_1 + \lambda_2 w_2$ durch Anwendung von F^{-1} . $F^{-1}(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 F^{-1}(w_1) + \lambda_2 F^{-1}(w_2)$.

5.1.3 Lemma

Seien U, V, W K -Vektorräume und $F : U \rightarrow V, G : V \rightarrow W$ linear. Dann ist auch $G \circ F : U \rightarrow W$ linear.

Beweis

$$\begin{aligned} (G \circ F)(\lambda_1 u_1 + \lambda_2 u_2) &= G(F(\lambda_1 u_1 + \lambda_2 u_2)) = \\ G(\lambda_1 F(u_1) + \lambda_2 F(u_2)) &= \lambda_1 G(F(u_1)) + \lambda_2 G(F(u_2)) = \\ \lambda_1 (G \circ F)(u_1) + \lambda_2 (G \circ F)(u_2) \end{aligned}$$

5.1.4 Lemma

Seien V, W K -Vektorräume, $F, G : V \rightarrow W$ linear und $\lambda \in K$. Dann sind auch $F + G$ und λF linear.

Beweis trivial

Bezeichnung

$$\text{Hom}_K(V, W) := \{F : V \rightarrow W : F \text{ ist linear}\}$$

Nullabbildung: 0

5.1.5 Korollar

$\text{Hom}_K(V, W)$ ist Untervektorraum von $\text{Abb}(V, W)$.

Bezeichnung

$\text{End}_K(V) := \text{Hom}_K(V, V)$ bezeichnet die Menge der Endomorphismen von V .

Bemerkung

Nach (5.1.5) ist $\text{End}_K(V)$ ein Untervektorraum von $\text{Abb}(V, V)$; nach (5.1.4) können wir auch noch die Komposition als Operation (Multiplikation) hinzunehmen, denn $F \cdot G = F \circ G \in \text{End}_K(V)$, wenn $F, G \in \text{End}_K(V)$ sind.

5.1.6 Bemerkung

Sei V ein K -Vektorraum. Dann ist $(\text{End}_K(V), +, \cdot)$ ein Ring.
Beweis trivial

5.2 Die Geometrie linearer Abbildungen und Gleichungssysteme

Definition

Sei $F : V \rightarrow W$ linear. Sei $\text{bild}(F) := F(V)$ (das BILD von F), für $w \in W$ $F^{-1}(w) := \{v : v \in V \wedge F(v) = w\}$ die FASER von w und $\text{kern}(F) := F^{-1}(0)$ der KERN von F .

Hinweis

Nach (5.1.2(iii)) sind $\text{bild}(F)$ und $\text{kern}(F)$ Untervektorräume von W bzw. V .

5.2.1 Bemerkung

Sei $F : V \rightarrow W$ linear.

- (i) F ist genau dann injektiv, wenn $\text{kern}(F) = \{0\}$.
- (ii) Seien F injektiv und $v_1, \dots, v_r \in V$ linear unabhängig. Dann sind auch $F(v_1), \dots, F(v_r)$ linear unabhängig.

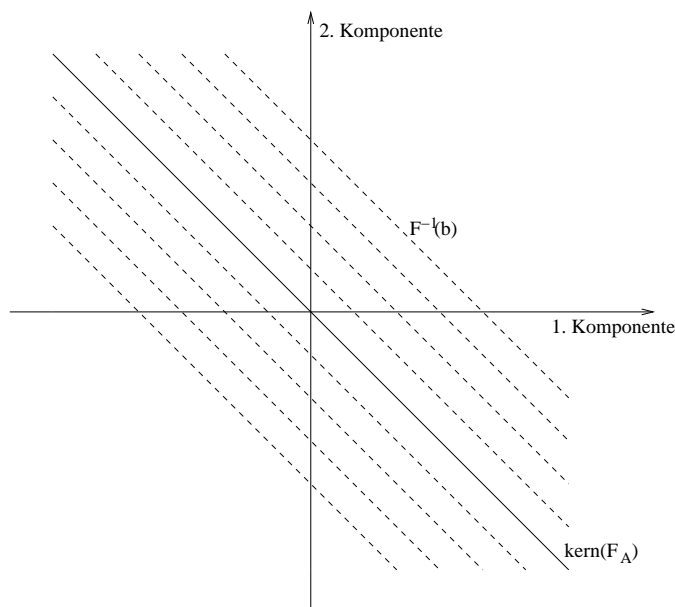
Beweis

- (i) \Rightarrow : trivial
 \Leftarrow : Seien $v_1, v_2 \in V$ und $F(v_1) = F(v_2)$. Dann gilt $F(v_1 - v_2) = 0$, d.h. $v_1 - v_2 \in \text{kern}(F)$.
- (ii) $0 = \sum_{i=1}^n \lambda_i F(v_i) = F(\sum_{i=1}^n \lambda_i v_i) \stackrel{(i)}{\Rightarrow} \sum_{i=1}^n \lambda_i v_i = 0 \Rightarrow \lambda_1 = \dots = \lambda_n = 0$

Beispiele

$A = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, F_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, F_B : \mathbb{R}^3 \rightarrow \mathbb{R}^2$. Es ist:

- $\text{bild}(F_A) = \mathbb{R} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$
- $\text{bild}(F_B) = \mathbb{R}^2$, da B Zeilenrang 2 hat.
- $\text{kern}(F_A) = \mathbb{R} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$
- $\text{kern}(F_B) = \mathbb{R} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$
- Sei $b \in \text{bild}(F_A)$. Dann gibt es ein $\lambda \in \mathbb{R}$ mit $b = \begin{pmatrix} 2\lambda \\ \lambda \end{pmatrix}$ und es gilt $F^{-1}(b) = F^{-1} \begin{pmatrix} 2\lambda \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda \end{pmatrix} + \text{kern}(F_A)$, d.h. die Fasern von b unter F_A sind Geraden parallel zum Kern von F_A (s. Abbildung 5.1).
- Analog sind auch für F_B die Fasern parallel zu $\text{kern}(F_B)$.

Abbildung 5.1: Geometrische Interpretation von Kern und Fasern von F_A **Beobachtung**

$$\dim(\text{kern } F_A) + \dim(\text{bild } F_A) = \dim(\mathbb{R}^2)$$

$$\dim(\text{kern } F_B) + \dim(\text{bild } F_B) = \dim(\mathbb{R}^3)$$

In (5.2.7) werden wir sehen, dass das Spezialfälle einer allgemeinen Dimensionsformel sind.

5.2.2 Hinweis

Sei $F : V \rightarrow W$ linear. Dann ist

$$\bigcup_{w \in W} F^{-1}(w) = \bigcup_{w \in \text{bild}(F)} F^{-1}(w)$$

eine (disjunkte) Zerlegung von V in Mengen der Form $u + \text{kern}(F)$.

Beweis

Seien $w \in \text{bild}(F)$ und $u, v \in F^{-1}(w)$. Dann gilt $F(u) = F(v) = w$, also $v - u \in \text{kern}(F)$; somit gilt $F^{-1}(w) \subset u + \text{kern}(F)$. Trivialerweise gilt auch \supseteq , und es folgt die Behauptung.

5.2.3 Korollar

Seien $A \in K^{m \times n}$, $b \in K^m$, $y \in K^n$ mit $Ay = b$. Dann gilt $\text{Lös}(A, b) = y + \text{Lös}(A, 0)$.

Beweis

Sei $F : K^n \rightarrow K^m$ definiert durch $F(x) = Ax$. Dann gilt $\text{Lös}(A, b) = F^{-1}(b)$, $\text{Lös}(A, 0) = \text{kern}(F)$.

Sprechweise

Die ALLGEMEINE Lösung des (INHOMOGENEN) Gleichungssystems $Ax = b$ ergibt sich als Summe einer (PARTIKULÄREN) Lösung und der allgemeinen Lösung des HOMOGENEN Systems $Ax = 0$.

Aus (4.3.3) ergibt sich sofort der folgende Struktursatz:

5.2.4 Korollar

Unter den Voraussetzungen von (5.2.3) sei $\{w_1, \dots, w_s\}$ eine Basis von $\text{Lös}(A, 0)$. Dann gilt $\text{Lös}(A, b) = y + Kw_1 + \dots + Kw_s$.

Offenbar spielen »Parallelverschiebungen« von Untervektorräumen als Fasern, d.h. für die Lösung linearer Gleichungssysteme eine besondere Rolle:

Definition

Seien V ein K -Vektorraum und $X \subset V$. X heißt AFFINER UNTERRAUM, falls es einen Vektor $v \in V$ und einen Untervektorraum U von V gibt mit $X = v + U$.

5.2.5 Lemma

Seien V ein K -Vektorraum, U, \bar{U} Untervektorräume und $v, \bar{v} \in V$. Dann gilt $v + U = \bar{v} + \bar{U} \Leftrightarrow U = \bar{U} \wedge v - \bar{v} \in U$.

Beweis

» \Leftarrow «: Sei $u \in U$ mit $v - \bar{v} = u$. Dann gilt $v + U = \bar{v} + u + U$, da U Untervektorraum ist.

» \Rightarrow «: $U = \bar{v} - v + \bar{U}$. Wegen $0 \in U$ ist $v - \bar{v} \in \bar{U}$. Damit folgt $U = \bar{U}$.

Zu einem gegebenen affinen Unterraum ist der zugehörige Untervektorraum U (LINEARER Teilraum) eindeutig bestimmt, »der« TRANSLATIONSVEKTOR v aber nur bis auf Summanden aus U .

Definition

Seien V ein K -Vektorraum, $v \in V$, U Untervektorraum und $X = v + U$. Dann sei $\dim X := \dim U$ die DIMENSION von X .

Affine Teilräume der Dimension 0, 1, 2 bzw. $(\dim V) - 1$ (falls V endlich - dimensional ist) heißen PUNKTE, GERADEN, EBENEN bzw. HYPEREBENEN.

5.2.6 Satz

Seien $F : V \rightarrow W$ linear und $\dim V < \infty$, v_1, \dots, v_s eine Basis von $\text{kern}(F)$, w_1, \dots, w_r eine Basis von $\text{bild}(F)$ und für $i = 1, \dots, r$ $u_i \in F^{-1}(w_i)$. Dann bilden $v_1, \dots, v_s, u_1, \dots, u_r$ eine Basis von V .

Beweis

Sei $v \in V$. Seien $\mu_1, \dots, \mu_r \in K$ mit $F(v) = \sum_{i=1}^r \mu_i w_i$. Ferner sei $\bar{v} = \sum_{i=1}^r \mu_i u_i$. Dann gilt $\bar{v} \in V$, $F(\bar{v}) = \sum_{i=1}^r \mu_i F(u_i) = F(v)$, also $v - \bar{v} \in \text{kern}(F)$. Es gibt also $\lambda_1, \dots, \lambda_s \in K$ mit $v - \bar{v} = \sum_{i=1}^s \lambda_i v_i$. Somit $v = \sum_{i=1}^r \mu_i u_i + \sum_{i=1}^s \lambda_i v_i$, d.h. $V = \text{span}\{v_1, \dots, v_s, u_1, \dots, u_r\}$.

Gilt andererseits $\sum_{i=1}^s \lambda_i v_i + \sum_{i=1}^r \mu_i u_i = 0$, so folgt $F(\sum_{i=1}^s \lambda_i v_i + \sum_{i=1}^r \mu_i u_i) = \sum_{i=1}^r \mu_i F(u_i) = \sum_{i=1}^r \mu_i w_i = 0$, also $\mu_1 = \dots = \mu_r = 0$. Es folgt $\sum_{i=1}^s \lambda_i v_i = 0$, somit auch $\lambda_1 = \dots = \lambda_s = 0$.

5.2.7 Korollar

Seien $F : V \rightarrow W$ linear, $\dim V < \infty$. Dann gilt

$$\dim V = \dim(\text{bild } F) + \dim(\text{kern } F)$$

Bezeichnung

$\text{rang } F := \dim(\text{bild } F)$ ist der RANG von F .

5.2.8 Korollar

Seien $F : V \rightarrow W$ linear, $\dim V < \infty$ und $w \in \text{bild}(F)$. Dann gilt

$$\dim(F^{-1}(w)) = \dim V - \text{rang } F$$

Beweis

Die Aussage folgt aus (5.2.2) und (5.2.6).

5.2.9 Korollar

Seien V, W endlich - dimensionale K -Vektorräume. Dann existiert genau dann ein Isomorphismus $F : V \rightarrow W$, wenn $\dim V = \dim W$.

Beweis

Sei F ein Isomorphismus. Dann gilt $W = \text{bild } F$, $\text{kern } F = \{0\}$ und die Behauptung folgt aus (5.2.7).

Gilt umgekehrt $\dim V = \dim W$ und sind v_1, \dots, v_n bzw. w_1, \dots, w_n Basen von V bzw. W , so ist durch $F(\sum_{i=1}^n \lambda_i v_i) := \sum_{i=1}^n \lambda_i w_i$ ein Isomorphismus gegeben.

5.2.10 Korollar

Seien V, W endlich - dimensionale K -Vektorräume, $\dim V = \dim W$ und $F : V \rightarrow W$ linear. Dann sind die folgenden Aussagen äquivalent:

- (i) F ist surjektiv.
- (ii) F ist injektiv.
- (iii) F ist bijektiv.

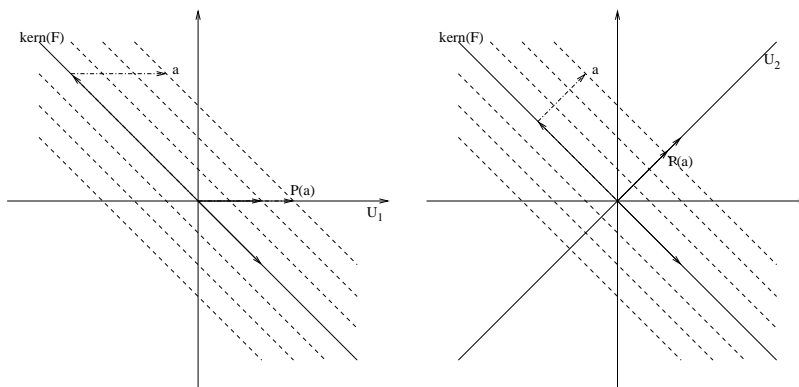
Beweis

Folgt unmittelbar aus (5.2.7) und (5.1.2).

5.2.11 Satz (Faktorisierung)

Seien V, W K -Vektorräume, $u_1, \dots, u_r, v_1, \dots, v_s$ eine Basis von V mit $\text{kern } F = \text{span}\{v_1, \dots, v_s\}$ und setze $U := \text{span}\{u_1, \dots, u_r\}$. Dann gilt:

- (i) $V = U + \text{kern}(F)$; U und $\text{kern}(F)$ sind Komplemente.
- (ii) Die Einschränkung $F|_U : U \rightarrow \text{bild}(F)$ ist ein Isomorphismus.
- (iii) Sei $P : V \rightarrow U$ die PROJEKTION $P(u + v) := u$ für $u \in U, v \in \text{kern}(F)$, so gilt $F = F|_U \circ P$.



Beispiel

Seien

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}, F(x) = Ax$$

Dann sind

$$\begin{aligned} \text{bild}(F) &= \mathbb{R} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ \text{kern}(F) &= \mathbb{R} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

Wähle

$$U_1 = \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad U_2 = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Es gilt dann

$$V = \mathbb{R}^2 = U_1 \oplus \text{kern}(F) = U_2 \oplus \text{kern}(F)$$

Weiter ist

$$\begin{aligned} F|_{U_1} \begin{pmatrix} \lambda \\ 0 \end{pmatrix} &= \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ F|_{U_1} : U_1 &\rightarrow \text{bild}(F) \text{ ist Isomorphismus} \end{aligned}$$

Ebenso ist auch $F|_{U_2} \begin{pmatrix} \lambda \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ ein Isomorphismus.

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \Rightarrow a = (a_1 + a_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (-a_2) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in U_1 \oplus \text{kern } F$$

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \Rightarrow a = \frac{1}{2}(a_1 + a_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2}(a_1 - a_2) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in U_2 \oplus \text{kern } F$$

Also:

$$\begin{aligned} P_1(a) &= (a_1 + a_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ P_2(a) &= \frac{1}{2}(a_1 + a_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

$$F|_{U_1} \circ P_1(a) = F_{U_1} \begin{pmatrix} a_1 + a_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2(a_1 + a_2) \\ a_1 + a_2 \end{pmatrix} = F(a)$$

$$F|_{U_2} \circ P_2(a) = F(a)$$

Also:

$$F^{-1}(F(a)) \cap U_1 = (a_1 + a_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = P_1(a)$$

$$F^{-1}(F(a)) \cap U_2 = P_2(a)$$

Bezeichnung

Die Umkehrabbildung $(F|_U)^{-1} : \text{bild}(F) \rightarrow U$ heißt **SCHNITT**.

Wir betrachten jetzt noch einmal die in (2.2) hergeleiteten (aber für beliebige K korrekten) Parametrisierungen Φ vor Lösungsmengen linearer Gleichungssysteme.

Erinnerung

Aus (2.2.1): $r = \text{Zeilenrang von } A = \text{Zeilenrang von } (A, b)$.

$$\Phi_{A,b} : K^{n-r} \rightarrow K^n$$

$$\Phi \begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1^* \\ \vdots \\ x_r^* \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix}$$

mit

$$x_{r-l}^* = \frac{1}{a_{r-l,r-l}} \left(b_{r-l} - \sum_{j=r-l+1}^r a_{r-l,j} x_j^* - \sum_{j=r+1}^n a_{r-l,j} x_j \right)$$

für $l = 0, \dots, r-1$. (Beachte: Die Pivotelemente stehen o.B.d.A. in den ersten r Spalten.)

5.2.12 Hinweis

$\Phi_{A,0} : K^{n-r} \rightarrow \text{Lös}(A, 0)$ ist ein Vektorraum - Isomorphismus.

Beweis

Die Linearität folgt aus der expliziten Darstellung von x_1^*, \dots, x_r^* , die Bijektivität aus den Ergebnissen von Kapitel 2.

Definition

Seien V, W K -Vektorräume und $\varphi : V \rightarrow W$. φ heißt **AFFIN**, wenn es eine lineare Abbildung $F : V \rightarrow W$ und einen Vektor $a \in W$ gibt, so dass $\varphi(v) = a + F(v)$ für alle $v \in V$. φ heißt außerdem **AFFINE TRANSFORMATION**, wenn φ affin und F ein Isomorphismus ist.

5.2.13 Bemerkung

$\Phi_{A,b} : K^{n-r} \rightarrow \text{Lös}(A, b)$ ist eine affine Transformation.

Nach Kapitel 2 ist die Dimension von $\text{Lös}(A, b)$ genau $n - r$. Nach (5.2.7) gilt $n = \dim V = \text{rang } F_A + \dim(\text{Lös}(A, 0)) = \text{rang } F_A + n - r$ und es ist $\text{rang } F_A = r$ (vgl. (4.2)).

Ist e_1, \dots, e_n die kanonische Basis von K^n und bezeichnen a_1, \dots, a_n die Spaltenvektoren von A , so gilt

$$A \sum_{i=1}^n \lambda_i e_i = \sum_{i=1}^n \lambda_i A e_i = \sum_{i=1}^n \lambda_i a_i$$

d.h. die Spaltenvektoren erzeugen bild F_A . Nach (4.2.3) gibt es somit eine Basis von bild F_A , die aus Spaltenvektoren von A besteht. Nach (4.2.2) ist daher der SPALTENRANG von A , d.h. die maximale Anzahl linear unabhängiger Spaltenvektoren gleich $\text{rang } F_A$. Wir haben somit bewiesen:

5.2.14 Satz

Seien $A \in K^{m \times n}$ und $F : K^n \rightarrow K^m$, definiert durch $F(x) = Ax$. Dann gilt:

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A) = \text{rang } F_A$$

Bezeichnung

Im Folgenden bezeichnen wir F_A (oft) einfacher mit A , d.h. wir *identifizieren* die Matrix A mit der durch sie induzierten Abbildung.

5.2.15 Korollar

Seien $A \in K^{m \times n}, b \in K^m$. Dann ist $\text{Lös}(A, b) \neq \emptyset$ genau dann, wenn $\text{rang}(A) = \text{rang}(A, b)$ gilt.

Beweis

Folgt mit (5.2.14) aus (2.2.1).

5.2.16 Bemerkung

Seien $A \in K^{m \times n}, b \in K^m$. Das Gleichungssystem $Ax = b$ ist genau dann *eindeutig* lösbar, wenn $\text{rang}(A) = \text{rang}(A, b) = n$ gilt.

Beweis

Folgt aus (5.2.13), (5.2.14) und (5.2.15).

5.2.17 Korollar

Seien $A \in K^{n \times n}, b \in K^n$ und gelte $\text{rang}(A) = n$. Dann existiert die Umkehrabbildung F_A^{-1} von F_A und $F_A^{-1}(b)$ ist die eindeutig bestimmte Lösung von $Ax = b$.

Beweis

Folgt aus (5.2.16) und (5.2.10).

5.3 Lineare Abbildungen und Matrizen

Die Abbildung F_A^{-1} aus (5.2.17) ist nach (5.1.2) linear. Wir zeigen, dass zu dieser Abbildung selbst wieder eine Matrix gehört. Genauer kann *jede* lineare Abbildung zwischen endlich - dimensionalen Vektorräumen durch Matrizen dargestellt werden.

5.3.1 Satz

Seien V, W endlich - dimensionale K -Vektorräume, $r \in \mathbb{N}$ und $v_1, \dots, v_r \in V$, $w_1, \dots, w_r \in W$.

- (i) Sind v_1, \dots, v_r linear unabhängig, so existiert eine lineare Abbildung $F : V \rightarrow W$ mit $F(v_i) = w_i$ für $i = 1, \dots, r$.
- (ii) Ist $\{v_1, \dots, v_r\}$ eine Basis von V , so gibt es *genau* eine lineare Abbildung $F : V \rightarrow W$ gemäß (i). Es gilt: $\text{bild}(F) = \text{span}\{w_1, \dots, w_r\}$. F ist genau dann injektiv, wenn w_1, \dots, w_r linear unabhängig sind.

Beweis

- (ii) Analog zur Konstruktion im Beweis von (5.2.9) setzen wir

$$F\left(\sum_{i=1}^r \lambda_i v_i\right) := \sum_{i=1}^r \lambda_i w_i$$

Dann hat F die gewünschten Eigenschaften. Sei \hat{F} eine weitere linear Abbildung mit $\hat{F}(v_i) = w_i$ für $i = 1, \dots, r$. Dann folgt ebenfalls

$$\hat{F}\left(\sum_{i=1}^r \lambda_i v_i\right) = \sum_{i=1}^r \lambda_i \hat{F}(v_i) = \sum_{i=1}^r \lambda_i w_i$$

d.h. $F = \hat{F}$.

Natürlich gilt $\text{bild } F \subset \text{span}\{w_1, \dots, w_r\}$. Ist andererseits $w = \sum_{i=1}^r \mu_i w_i$, so folgt

$$F\left(\sum_{i=1}^r \mu_i v_i\right) = \sum_{i=1}^r \mu_i F(v_i) = \sum_{i=1}^r \mu_i w_i = w$$

d.h. es gilt auch die umgekehrte Inklusion.

F ist genau dann injektiv, wenn $\text{kern } F = \{0\}$ ist (5.2.1). Aus (5.2.7) folgt $\dim V = \dim(\text{span}\{w_1, \dots, w_r\})$ genau dann, wenn F injektiv ist und es gilt $\dim(\text{span}\{w_1, \dots, w_r\}) = r$ genau dann, wenn w_1, \dots, w_r linear unabhängig sind.

- (i) Wir ergänzen v_1, \dots, v_r gemäß (4.2.5) durch Hinzunahme von entsprechenden Vektoren v_{r+1}, \dots, v_n zu einer Basis von V und wählen w_{r+1}, \dots, w_n beliebig. Die Aussage folgt aus (ii).

5.3.2 Korollar

Seien V ein K -Vektorraum und $B = \{v_1, \dots, v_r\}$ eine Basis. Dann gibt es genau einen Isomorphismus $\Phi_B : K^n \rightarrow V$ mit $\Phi_B(e_i) = v_i$ für $i = 1, \dots, n$ (e_i ist der i -te Koordinatenvektor).

Bezeichnung

Φ_B heißt KOORDINATENSYSTEM.

Beispiele

$$(a) V = \text{span} \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\}$$

$$\Phi_B \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \Phi_B \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

$$\text{Dann gilt } \Phi_B \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = x_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \text{ d.h. mit } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ gilt } \Phi_B(x) = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ (vgl. Beispiel vor (5.1.2)).}$$

(b) $F : K^n \rightarrow K^m$, setze $F(e_j) = a_j$, $A = (a_1, \dots, a_n)$. Dann ist der Koordinatenvektor Ax das Bild des Koordinatenvektors x unter F .

Da die linearen Abbildungen durch ihr Wirken auf eine Basis des Urbildraumes bestimmt sind, können sie durch Matrizen beschrieben werden.

5.3.3 Satz

Seien V, W K -Vektorräume, $B_V = \{v_1, \dots, v_n\}$ Basis von V , $B_W = \{w_1, \dots, w_m\}$ Basis von W und sei $F : V \rightarrow W$ linear. Dann gibt es genau eine Matrix $A = M_{B_W}^{B_V}(F) = (a_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} \in K^{m \times n}$ mit folgender Eigenschaft:

Seien $v \in V, x_1, \dots, x_n, y_1, \dots, y_m \in K$ mit $v = \sum_{j=1}^n x_j v_j, F(v) = \sum_{i=1}^m y_i w_i$ und

$$\text{setze } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}. \text{ Dann gilt } Ax = y.$$

Beweis

Für $i = 1, \dots, m, j = 1, \dots, n$ sei $a_{ij} \in K$, so dass

$$F(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{für } j = 1, \dots, n$$

Es folgt $F(v) = \sum_{j=1}^n x_j F(v_j) = \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} w_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) w_i$, also $Ax = y$.

Der Koordinatenvektor x von v in der Basis B_V wird durch A in den Koordinatenvektor von $F(v)$ in der Basis B_W überführt. Nach (4.2.1) ist für jedes x Ax eindeutig. Wählt man die kanonischen Basisvektoren in K^n und K^m , so folgt die Behauptung.

5.3.4 Satz

Unter den Voraussetzungen und mit den Bedingungen von (5.3.3) gilt: Die Abbildung $\Phi = \Phi_{B_V, B_W} : \text{Hom}_K(V, W) \rightarrow K^{m \times n}$, definiert durch $F \mapsto A_{B_V}^{B_W}(F)$ ist ein Vektorraum - Isomorphismus.

Beweis

Seien $F, G \in \text{Hom}_K(V, W)$, (a_{ij}) bzw. (b_{ij}) die nach (5.3.3) jeweils zugehörige Matrix und $\lambda, \mu \in K$. Dann gilt $(\lambda F + \mu G)(v_j) = \lambda F(v_j) + \mu G(v_j) = \lambda \sum_{i=1}^m a_{ij} w_i + \mu \sum_{i=1}^m b_{ij} w_i = \sum_{i=1}^m (\lambda a_{ij} + \mu b_{ij}) w_i$, also ist $(\lambda a_{ij} + \mu b_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$ die $(\lambda F + \mu G)$ zugeordnete Matrix. Die Bijektivität folgt aus (5.3.1).

Bemerkung

(5.3.4) zeigt, dass nach Festlegung von Basen B_V, B_W lineare Abbildungen und Matrizen identifiziert werden können. Unterschiedliche Basen führen aber bei gleicher Abbildung i.a. zu verschiedenen Matrizen und umgekehrt.

Fragen

- (1) Wie ändert sich $M_{B_W}^{B_V}(F)$ bei Übergang zu anderen Basen?
- (2) Kann man Basen B_V, B_W so wählen, dass die Matrix $M_{B_W}^{B_V}(F)$ besonders »einfach« wird?

Beispiel

Wie sieht eigentlich $M_{B_W}^{B_V}(F)$ aus, wenn die Basen aus (5.2.6) gewählt werden? Seien v_1, \dots, v_{n-r} Basis von kern F , w_1, \dots, w_r Basis von bild F und für $i = 1, \dots, r$ sei $u_i \in F^{-1}(w_i)$. Es ist $B_V = \{u_1, \dots, u_r, v_1, \dots, v_{n-r}\}$. Es gilt also $F(u_j) = w_j$ für $j = 1, \dots, r$, $F(v_j) = 0$ für alle $j = 1, \dots, n-r$. Wir ergänzen w_1, \dots, w_r zu einer Basis B_W von W . Dann hat $M_{B_W}^{B_V}$ die Gestalt

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & 0 & \vdots & & \vdots \\ 0 & \dots & 0 & 1 & 0 & & \vdots \\ \vdots & & & & \vdots & & \vdots \\ 0 & \dots & \dots & \dots & 0 & \dots & 0 \end{pmatrix}$$

ist also von der Form $\begin{pmatrix} I_R & 0 \\ 0 & 0 \end{pmatrix}$, wobei $I_R = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ die $r \times r$ -Einheitsmatrix

und 0 Nullmatrizen bezeichnet. Diese Darstellung spiegelt auch die Strukturaussagen von (5.2.11) wieder.

Wenn wir also die volle Freiheit bei der Wahl der Basen haben, so kann man die zu einer linearen Abbildung gehörende Matrix sehr einfach machen. Betrachtet man Endomorphismen $F : V \rightarrow V$ und verlangt, dass die im Urbild- und Bildraum gewählten Basen übereinstimmen, so kann man i.a. keine so einfache Gestalt erwarten. Andererseits enthält die Matrix »mehr Informationen« über F .

Beispiel

$V = \mathbb{R}^2 = \text{span}\{e_1, e_2\}$, $F(e_1) = 4e_1 + e_2$, $F(e_2) = -e_1 + 2e_2$.

Zugehörige Matrix: $\begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$.

Seien $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $B_V = \{v_1, v_2\}$. Dann gilt $F(v_1) = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, $F(v_2) = \begin{pmatrix} -1 \\ 2 \end{pmatrix} = -v_1 + 3v_2$.

Die bezüglich B_V (im Urbild- und Bildraum) zugehörige Matrix ist $\begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix}$.

Wieso ist der Koeffizient a_{21} gleich 0?

5.3.5 Bemerkung

Seien V ein K -Vektorraum, v_1, \dots, v_r eine Basis, $F : V \rightarrow V$ linear und $A = (a_{ij})_{i,j=1,\dots,n}$ die zugehörige Matrix. Dann gilt $a_{21} = \dots = a_{n1} = 0$ genau dann, wenn $F(v_1) = a_{11}v_1$ ist.

Beweis

Gilt $F(v_1) = a_{11}v_1$, so ist der erste Spaltenvektor von A nach (5.3.3) gerade

$$a_{11} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \text{ Hat andererseits der erste Spaltenvektor diese Gestalt, so folgt } F(v_1) = \sum_{i=1}^n a_{i1}v_i = a_{11}v_1.$$

Definition

Seien V ein K -Vektorraum, $\lambda \in K$, $F \in \text{End}_K(V)$. λ heißt EIGENWERT von F , wenn es ein $v \in V \setminus \{0\}$ gibt mit $F(v) = \lambda v$. Jeder solche Vektor v heißt EIGENVEKTOR von F (zum Eigenwert λ).

Beachte

λ ist beliebig, aber $v \neq 0$, da $F(0) = 0$ keine »neue Strukturaussage« ist.

Hinweis

Da für Eigenvektoren $v \neq 0$ vorausgesetzt ist, gibt es eine Basis von V , die einen gegebenen Eigenvektor enthält; wir können also die Situation von (5.3.5) erreichen.

Bemerkung

Ist v ein Eigenvektor und der zugehörige Eigenwert $\lambda \neq 0$, so gilt $F(\text{span}\{v\}) = \text{span}\{v\}$, d.h. F lässt die Gerade Kv *invariant*. Auf Kv wirkt F lediglich als Streckung oder Stauchung bzw. Spiegelung.

Fragen

- (1) Wann besitzt ein Endomorphismus Eigenwerte?
- (2) Welche Endomorphismen sind »diagonalisierbar«, können also in die Form

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ gebracht werden?}$$

Das gilt nach (5.3.3) offenbar genau dann, wenn es eine Basis aus Eigenvektoren gibt.

- (3) Wie kann man Eigenwerte (und Eigenvektoren) bestimmen? ($Av = \lambda v$: Gleichungssystem in v, λ ; nicht linear)

Diese Fragen werden in Kapitel 7 diskutiert, in Kapitel 6 werden wesentliche Hilfsmittel zur Verfügung gestellt.

Im Folgenden: Was geschieht bei Komposition linearer Abbildungen mit den zugehörigen Matrizen und was bei Bildung der Umkehrfunktion?

5.3.6 Satz

Seien U, V, W K -Vektorräume mit Basen $B_U = \{u_1, \dots, u_r\}$, $B_V = \{v_1, \dots, v_r\}$, $B_W = \{w_1, \dots, w_t\}$, $F : U \rightarrow V, G : V \rightarrow W$ linear, $A = (a_{kj})_{\substack{k=1, \dots, s \\ j=1, \dots, r}} = M_{B_V}^{B_U}(F)$, $B = (b_{ik})_{\substack{i=1, \dots, t \\ k=1, \dots, s}} = M_{B_W}^{B_V}(G)$ und $C = M_{B_W}^{B_U}(G \circ F)$. Dann ist $C = (c_{ij}) \in K^{t \times r}$ und es gilt

$$c_{ij} = \sum_{k=1}^s b_{ik} a_{kj}$$

Beweis

Nach (5.3.3) sind die Koeffizienten c_{1j}, \dots, c_{tj} die Koordinaten von $(G \circ F)(u_j)$ für $j = 1, \dots, r$, d.h. $C \in K^{t \times r}$.

Nach Voraussetzung gilt:

$$F(u_j) = \sum_{k=1}^s a_{kj} v_k \quad (j = 1, \dots, r)$$

$$G(v_k) = \sum_{i=1}^t b_{ik} w_i \quad (k = 1, \dots, s)$$

Somit gilt für $j = 1, \dots, r$:

$$(G \circ F)(u_j) = \sum_{k=1}^s a_{kj} \sum_{i=1}^t b_{ik} w_i = \sum_{i=1}^t \underbrace{\left(\sum_{k=1}^s b_{ik} a_{kj} \right)}_{c_{ij}} w_i$$

und es folgt die Behauptung.

Beachte

Dieses entspricht für 2×2 -Matrizen genau der Matrix-Multiplikation aus (1.2).

Definition

Seien $r, s, t \in \mathbb{N}$. Dann sei $\cdot : K^{t \times s} \times K^{s \times r} \rightarrow K^{t \times r}$ definiert durch $B \cdot A = C$, wobei $A = (a_{kj}) \in K^{s \times r}$, $B = (b_{ik}) \in K^{t \times s}$ und $C = (c_{ij}) \in K^{t \times r}$ mit $c_{ij} = \sum_{k=1}^s b_{ik} a_{kj}$ für $i = 1, \dots, t, j = 1, \dots, r$.

Beachte

Bei der Matrix-Multiplikation $B \cdot A$ muss die Spaltenanzahl von B mit der Zeilenanzahl von A übereinstimmen.

Beispiel

$$\begin{pmatrix} 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 4 & 0 \\ 3 & 0 & 0 \\ 1 & 0 & 7 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 1 & 15 \\ 9 & 4 & 7 \end{pmatrix}$$

Bemerkung

Die Multiplikation $\gg 2 \times 5$ - Matrix \ll mal $\gg 5 \times 3$ - Matrix \ll ist definiert nicht aber die Multiplikation einer 5×3 - Matrix mit einer 2×5 - Matrix.

Beispiele

- Für quadratische Matrizen $A, B \in K^{n \times n}$ sind $A \cdot B$ und $B \cdot A$ definiert, stimmen aber i.a. nicht überein.

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

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

Selbst für quadratische Matrizen ist die Multiplikation i.a. *nicht* kommutativ.

- Die Matrix-Multiplikation ist i.a. nicht nullteilerfrei.

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

$$(1 \ 1) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$$

Bemerkung

Gilt $A \in K^{m \times n}$, $x \in K^{n \times 1} = K^n$, dann stimmt die Definition von Ax aus (1.1) mit der oben gegebenen überein. Man kann $n \times 1$ - Matrizen also mit (Spalten-) Vektoren und $1 \times n$ - Matrizen mit Zeilenvektoren identifizieren.

Beispiele

- *Vektorprodukte:* Seien $A \in K^{1 \times n}$, $B \in K^{n \times 1}$. Dann ist $A \cdot B \in K^{1 \times 1} = K$ das SKALARPRODUKT und $B \cdot A \in K^{n \times n}$.

$$(1 \ 2 \ 0 \ 1) \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = 2$$

$$\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \cdot (1 \ 2 \ 0 \ 1) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ -1 & -2 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Kettenregel für $f \circ g$ mit $f : \mathbb{R}^s \rightarrow \mathbb{R}^t$, $g : \mathbb{R}^r \rightarrow \mathbb{R}^s$ differenzierbar (Jakobi - Matrizen).

5.3.7 Lemma: Rechenregeln für Matrizen

(i) Distributivgesetz

$$A \in K^{m \times n} \wedge B_1, B_2 \in K^{n \times r} \Rightarrow A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2$$

$$A_1, A_2 \in K^{m \times n} \wedge B \in K^{n \times r} \Rightarrow (A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B$$

(ii) $A \in K^{m \times n} \wedge B \in K^{n \times r} \wedge \lambda \in K \Rightarrow A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B)$

(iii) Assoziativgesetz

$$A \in K^{m \times n} \wedge B \in K^{n \times r} \wedge C \in K^{r \times s} \Rightarrow (A \cdot B) \cdot C = A \cdot (B \cdot C)$$

(iv) $A \in K^{m \times n} \wedge B \in K^{n \times r} \Rightarrow (A \cdot B)^T = B^T \cdot A^T$

(v) Neutralität der passenden Einheitsmatrix

$$A \in K^{m \times n} \Rightarrow I_m \cdot A = A \cdot I_n = A$$

Beweis

Durch Nachrechnen oder einfacher mit Hilfe von (5.3.4) und (5.3.6).

Bezeichnung

$A \cdot B$ wird kürzer AB geschrieben, ebenso λAB statt $\lambda(AB)$ und ABC statt $A(BC)$ bzw. $(AB)C$.

5.3.8 Korollar

$(K^{n \times n}, +, \cdot)$ ist ein Ring.

Beweis

Folgt aus (4.1.1) und (5.3.7).

5.3.9 Bemerkung

Seien V, W K -Vektorräume, $B_V = \{v_1, \dots, v_n\}$, $B_W = \{w_1, \dots, w_n\}$ Basen von V bzw. W , $F: V \rightarrow W$, $G: W \rightarrow V$ Isomorphismen, $A = M_{B_V}^{B_V}(F)$, $B = M_{B_V}^{B_W}(G)$. Dann gilt $G \circ F = \text{id}_V$ genau dann, wenn $BA = I_n$ ist.

Beweis

Nach (5.3.6) ist $BA = M_{B_V}^{B_V}(G \circ F)$. Nach (5.3.3) ist die j -te Spalte von BA der Koordinatenvektor von $G \circ F(v_j)$ in der Basis B_V . Hieraus folgt die Behauptung.

Bemerkung

Natürlich gilt $F^{-1} \circ F = \text{id}_V$, $F \circ F^{-1} = \text{id}_W$ (vgl. (3.1.4)). Also kommutieren die zugehörigen Matrizen bei Endomorphismen.

Definition

Es sei $A \in K^{n \times n}$. A heißt INVERTIERBAR, wenn es ein $\hat{A} \in K^{n \times n}$ gibt mit $A\hat{A} = \hat{A}A = I_n$.

Sei $\text{GL}(n, K) := \{A : A \in K^{n \times n} \wedge A \text{ ist invertierbar}\}$. Dann heißt $(\text{GL}(n, K), \cdot)$ ALLGEMEINE LINEARE GRUPPE.

5.3.10 Hinweis

$(\text{GL}(n, K), \cdot)$ ist eine Gruppe.

Beweis

Wir zeigen, dass $A, B \in \text{GL}(n, K)$ auch $AB \in \text{GL}(n, K)$ impliziert. Der Rest der Behauptung ist trivial.

Seien $A, B \in \text{GL}(n, K)$ und $\hat{A}, \hat{B} \in \text{GL}(n, K)$ mit $\hat{A}A = \hat{B}B = I_n$. Dann gilt $(\hat{B}\hat{A})(AB) = \hat{B}(\hat{A}A)B = \hat{B}B = I_n$.

Bezeichnung

Es sei $A \in \text{GL}(n, K)$. A^{-1} bezeichnet die zu A inverse Matrix, d.h. $AA^{-1} = A^{-1}A = I_n$.

Hinweis

Rechenregeln für inverse Matrizen gelten gemäß (3.1.5), (3.1.7) und (3.1.8).

5.3.11 Hinweis

Sei $A \in K^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:

- (i) A ist invertierbar.
- (ii) A^T ist invertierbar.
- (iii) $\text{rang}(A) = n$

Ferner gilt $(A^{-1})^T = (A^T)^{-1}$.

Berechnung von A^{-1}

Sei $A \in K^{n \times n}$ mit $\text{rang}(A) = n$. Dann existiert die inverse Matrix A^{-1} . Wir setzen

$$A^{-1} = X = (x_{ij})_{i,j=1,\dots,n}, \quad x_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix} \quad \text{für } j = 1, \dots, n. \quad \text{Dann gilt } AX = I_n,$$

also für $j = 1, \dots, n$ $Ax_j = e_j$. Zur Bestimmung von A^{-1} sind somit diese n Gleichungssysteme zu lösen. Das erfolgt am Einfachsten simultan und vollständig (vgl. (2.3.3)), d.h. (A, I_n) wird durch elementare Zeilenoperationen in (I_n, A^{-1}) überführt.

Beispiel

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array} \right) &\rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{array} \right) \rightarrow \\ \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & 3 & -1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & -1 \end{array} \right) &\rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{3} & -\frac{1}{3} & \frac{5}{3} \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right) \end{aligned}$$

d.h. die Inverse zu $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{pmatrix}$ ist $\frac{1}{3} \begin{pmatrix} -2 & 1 & 5 \\ 3 & 0 & -3 \\ -1 & 1 & 1 \end{pmatrix}$.

Bemerkung

Elementare Zeilenoperationen lassen sich mit Hilfe von Matrizenmultiplikation beschreiben. Seien $A \in K^{n \times n}$ und a_1, \dots, a_n seien die Zeilenvektoren von A . Ferner sei $\lambda \in K$. Dann gilt

$$\begin{pmatrix} 1 & & & \\ \lambda & 1 & 0 & \\ & 0 & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ \lambda a_1 + a_2 \\ \vdots \\ a_n \end{pmatrix}$$

d.h. die Multiplikation von A mit der angegebenen »Elementarmatrix« führt zur Addition des λ -fachen der ersten Zeile zur zweiten.

Kapitel 6

Determinanten

Die Determinantenabbildung \det ist eine Funktion $\det : K^{n \times n} \rightarrow K$ (bzw. $\det : V \times \dots \times V \rightarrow K$), ordnet also einer $n \times n$ -Matrix ein Körperelement zu. Sie tritt beim Lösen von Gleichungssystemen auf (»Cramersche Regel«), bei der Volumenberechnung (»Gramsche Determinante«) und ist zentral für die Bestimmung von Eigenwerten. (»charakteristisches Polynom«).

6.1 Einführung

Fragen

- (1) Seien $A \in GL(n, K)$, $b \in K^n$. Kann man die Lösung von $Ax = b$ explizit als Funktion (der Einträge) von A und b angeben? Da $A^{-1}b = x$ ist, heißt das auch:
- (2) Kann man A^{-1} explizit als Funktion von A darstellen?

Wir betrachten als Beispiel das allgemeine 2×2 -System.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 & (I) \\ a_{21}x_1 + a_{22}x_2 = b_2 & (II) \end{cases}$$
$$\begin{cases} (a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2 & a_{22}(I) - a_{12}(II) \\ (a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1 & a_{11}(II) - a_{21}(I) \end{cases}$$

Man beachte, dass der Übergang von ersten zum zweiten System beschrieben wird durch

$$\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

d.h. $\text{Lös}(A, b)$ wird genau dann nicht verändert, wenn $\text{rang} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = 2$ ist und das ist äquivalent zu $\text{rang}(A) = 2$. Die Rangbedingung ist genau dann erfüllt, wenn $a_{11}a_{22} - a_{12}a_{21} \neq 0$ ist. In diesem Fall können wir nach x_1, x_2 auflösen.

Bezeichnung

Sei $A = (a_{ij})_{i,j=1,2} \in K^{2 \times 2}$. Dann heißt $\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{12}a_{21}$ DETERMINANTE von A .

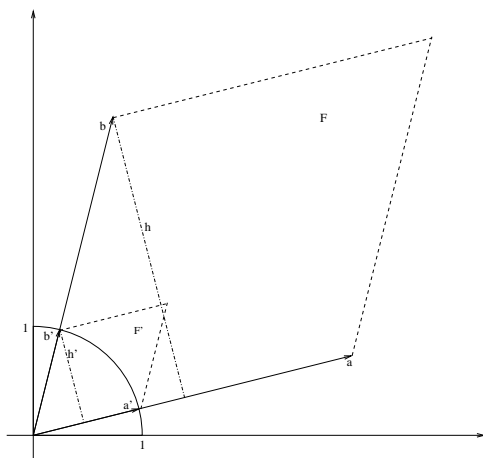


Abbildung 6.1: zur Fläche eines Parallelogramms

6.1.1 Hinweis

Seien $A \in K^{2 \times 2}$, $b \in K^2$. Das Gleichungssystem $Ax = b$ ist genau dann eindeutig lösbar, wenn $\det(A) \neq 0$ gilt. In diesem Fall gilt

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

(Spezialfall $n = 2$ der Cramerschen Regel).

Übung: Wie sieht in dieser Schreibweise A^{-1} aus?

6.1.2 Hinweis

Sei $A \in K^{2 \times 2}$. $\text{rang}(A)$ ist genau dann kleiner 2, wenn $\det(A) = 0$ ist.

6.1.3 Korollar

Seien $A \in K^{2 \times 2}$, $\lambda \in K$. λ ist genau dann Eigenwert von A , wenn $\det(A - \lambda I_2) = 0$ gilt.

Beachte

$$\begin{aligned} \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \\ &= a_{11}a_{22} - a_{12}a_{21} - \lambda(a_{11} + a_{22}) + \lambda^2 \end{aligned}$$

Mit der Bezeichnung $\text{spur}(A) = a_{11} + a_{22}$ ist ein Eigenwert λ also Nullstelle des CHARAKTERISTISCHEN POLYNOMS $x^2 - \text{spur}(A)x + \det(A)$. Eigenwerte zu berechnen bedeutet Nullstellen des charakteristischen Polynoms zu bestimmen.

Beispiel

Flächeninhalt eines Parallelogramms (s. Abbildung 6.1)

Seien α der Winkel zwischen der horizontalen Achse und dem Vektor a , β der

Winkel zwischen der horizontalen Achse und dem Vektor b , h die Höhe und nach Voraussetzung $0 \leq \beta - \alpha \leq \pi$.

Weiter seien $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \varrho \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \sigma \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix}$. Dann ist $F = h\varrho$ (\rightarrow Cavalieri, Fubini).

$$h' = \sin(\beta - \alpha)$$

$$a' = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$$

$$b' = \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix}$$

Es gilt $F = \varrho\sigma F' = \varrho\sigma \sin(\beta - \alpha)$. Wegen

$$\sin(\beta - \alpha) = \sin \beta \cos \alpha - \sin \alpha \cos \beta = \begin{vmatrix} \cos \alpha & \cos \beta \\ \sin \alpha & \sin \beta \end{vmatrix}$$

gilt

$$F = \varrho\sigma \begin{vmatrix} \cos \alpha & \cos \beta \\ \sin \alpha & \sin \beta \end{vmatrix} = \begin{vmatrix} \varrho \cos \alpha & \sigma \cos \beta \\ \varrho \sin \alpha & \sigma \sin \beta \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

6.2 Permutation

Erinnerung und Bezeichnung

Sei $X = \{1, \dots, n\}$, dann sei $S_n := S(X)$, d.h. S_n bezeichne die Menge aller bijektiven Abbildungen $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$. Die Elemente von S_n heißen PERMUTATIONEN. (S_n, \circ) ist die SYMMETRISCHE GRUPPE. Die ORDNUNG $|S_n|$ von S_n , d.h. die Anzahl der Elemente von S_n ist $n!$.

Schreibweisen: $\begin{pmatrix} 1 & \dots & n \\ \pi(1) & \dots & \pi(n) \end{pmatrix}$ oder kürzer $(\pi(1), \dots, \pi(n))$.

Seien $\pi \in S_n$ und $i \in X$. Dann heißt i genau dann FIXPUNKT von π , wenn $\pi(i) = i$ gilt. Verkürzte Schreibweise: Sei $I \subset X$ und alle $i \in X \setminus I$ seien Fixpunkte von π . Dann schreibt man auch $(\pi(i) : i \in I)$ statt $(\pi(i) : i \in X)$.

Beispiel

$(2, 1) \in S_3$ bedeutet $(2, 1, 3)$.

$(4, 2) \in S_5$ bedeutet $(1, 4, 3, 2, 5)$.

Definition

Seien $n \geq 2$ und $\pi \in S_n$. π heißt TRANSPOSITION, wenn $i, j \in \{1, \dots, n\}$ mit $i < j$ und $\pi = (j, i)$.

6.2.1 Satz

Seien $n \geq 2$ und $\pi \in S_n$. Dann ist π Produkt endlich vieler Transpositionen.

Beweis

(vollständige Induktion über n)

$S_2 = \{\text{id}, (2, 1)\}$. Es gilt $(2, 1)(2, 1) = \text{id}$, d.h. S_2 wird durch die Transposition $(2, 1)$

erzeugt.

Für $i = 1, \dots, n-1$ sei $\pi_i = (n, i)$. Dann gilt

$$S_n = S_{n-1} \cup \bigcup_{i=1}^{n-1} \pi_i S_{n-1}$$

wobei S_{n-1} als Teilmenge von S_n aufgefasst wird, deren Elemente n fix lassen. Die Behauptung folgt aus der Induktionsannahme.

Beispiel

$\pi = (3, 1, 2) \in S_3$. Dann gilt $\pi = (2, 1)(3, 1) = (3, 1)(3, 2)(5, 4)(5, 4)$, d.h. die Darstellungen gemäß (6.2.1) sind nicht eindeutig. Die Anzahl der benötigten Permutationen ist aber in beiden Fällen gerade.

Definition

Seien $\pi \in S_n$, $i, j \in \{1, \dots, n\}$. Das Paar (i, j) heißt FEHLSTAND von π , wenn $i \leq j$, aber $\pi(i) > \pi(j)$.

Das SIGNUM $\text{sign}(\pi)$ ist definiert durch $\text{sign}(\pi) = (-1)^{\alpha(\pi)}$, wobei $\alpha(\pi)$ die Anzahl der Fehlstände von π ist. π heißt GERADE, wenn $\text{sign}(\pi) = 1$ ist und UNGERADE, wenn $\text{sign}(\pi) = -1$ ist.

Beispiel

Sei $\pi = (3, 1, 2) \in S_3$.

$$\begin{array}{ccc} 1 & 2 & \rightarrow 3 & 1 & \rightarrow \text{Fehlstand} \\ 1 & & 3 & \rightarrow 3 & 2 & \rightarrow \text{Fehlstand} \\ & 2 & 3 & \rightarrow & 1 & 2 \end{array}$$

Also: $\text{sign}(\pi) = (-1)^2 = 1$; π ist eine gerade Permutation.

6.2.2 Lemma

Seien $n \geq 2$ und $\pi \in S_n$. Dann gilt

$$\text{sign}(\pi) = \prod_{\substack{i, j=1 \\ i < j}}^n \frac{\pi(j) - \pi(i)}{j - i}$$

Beachte

Da $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ bijektiv ist, stehen im Zähler und im Nenner die selben Differenzen - bis auf die Vorzeichen.

Beweis

Sei α die Anzahl der Fehlstände von π . Dann gilt

$$\begin{aligned} \prod_{i < j} (\pi(j) - \pi(i)) &= \prod_{\substack{i < j \\ \pi(i) < \pi(j)}} (\pi(j) - \pi(i)) \cdot (-1)^\alpha \prod_{\substack{i < j \\ \pi(i) > \pi(j)}} |\pi(j) - \pi(i)| = \\ &(-1)^\alpha \prod_{i < j} |\pi(j) - \pi(i)| = (-1)^\alpha \prod_{i < j} (j - i) \end{aligned}$$

6.2.3 Satz

Seien $\pi, \sigma \in S_n$. Dann gilt

- (i) $\text{sign}(\pi^{-1}) = \text{sign}(\pi)$
- (ii) $\text{sign}(\pi\sigma) = \text{sign}(\pi) \text{sign}(\sigma)$

Beweis

(i) folgt aus (ii)

(ii) Nach (6.2.2) gilt

$$\begin{aligned} \text{sign}(\pi\sigma) &= \prod_{i < j} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{j - i} = \\ &= \prod_{i < j} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(j) - \sigma(i)} \cdot \frac{\sigma(j) - \sigma(i)}{j - i} = \\ &= \text{sign}(\sigma) \prod_{i < j} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(j) - \sigma(i)} = \\ &= \text{sign}(\sigma) \prod_{\substack{i < j \\ \sigma(i) < \sigma(j)}} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(j) - \sigma(i)} \prod_{\substack{i < j \\ \sigma(i) > \sigma(j)}} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(j) - \sigma(i)} = \\ &= \text{sign}(\sigma) \prod_{\substack{i < j \\ \sigma(i) < \sigma(j)}} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(j) - \sigma(i)} \prod_{\substack{i > j \\ \sigma(i) < \sigma(j)}} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(j) - \sigma(i)} = \\ &= \text{sign}(\sigma) \prod_{\sigma(i) < \sigma(j)} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(j) - \sigma(i)} = \text{sign}(\sigma) \text{sign}(\pi) \end{aligned}$$

6.2.4 Korollar

Sei $\pi \in S_n$ Produkt von k Transpositionen. Dann gilt $\text{sign}(\pi) = (-1)^k$.

Beweis

Es reicht zu zeigen, dass für eine Transposition $\tau \in S_n$ $\text{sign}(\tau) = -1$ gilt. Der Rest folgt aus (6.2.3(ii)).

Seien $i, j \in \{1, \dots, n\}$, $i < j$, $\tau = (j, i)$. Dann gibt es genau die folgenden Fehlstände:

- $\{j, k\}$ mit $k \in \{i + 1, \dots, j - 1\}$
- $\{j, i\}$
- $\{k, i\}$ mit $k \in \{i + 1, \dots, j - 1\}$

Die Anzahl der Fehlstände ist also ungerade.

Bezeichnung

$$A_n := \{\pi \in S_n : \text{sign}(\pi) = 1\}$$

6.2.5 Hinweis

A_n ist eine Untergruppe von S_n : die ALTERNIERENDE GRUPPE.

6.2.6 Bemerkung

Sei $n \geq 2$ und $\tau \in S_n$ mit $\text{sign}(\tau) = -1$. Dann gilt $S_n = A_n \cup \tau A_n$, $A_n \cap \tau A_n = \emptyset$ und ferner $|a_n| = \frac{1}{2}n!$.

Beweis

Sei $\pi \in S_n$ mit $\text{sign}(\pi) = -1$. Dann gilt nach (6.2.3) $\text{sign}(\tau^{-1}\pi) = 1$, d.h. $\pi \in \tau A_n$. Da A_n die geraden und τA_n die ungeraden Permutationen enthält, gilt natürlich $A_n \cap \tau A_n = \emptyset$. Wegen $|S_n| = n!$ folgt die Behauptung.

6.3 Multilineare Abbildungen und Determinantenformen

Definition

Seien V_1, \dots, V_r, W K-Vektorräume und $F : V_1 \times \dots \times V_r \rightarrow W$. F heißt R-FACH LINEAR oder R-MULTILINEAR, wenn für jedes $i \in \{1, \dots, r\}$ und jede Wahl von $v_j \in V_j$ für $j \in \{1, \dots, r\} \setminus \{i\}$ die Abbildung $F_i : V_i \rightarrow W$, definiert durch $F_i(v) := F(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_r)$ für $v \in V$ linear ist. Gilt $V_1 = \dots = V_r$ und $W = K$, so heißt eine multilineare Abbildung auch R-FACHE LINEARFORM oder R-MULTILINEARFORM, für $r = 2$ BILINEARFORM.

Hinweis

$F : K^2 \times K^2 \rightarrow K$, definiert durch $F\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) := \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, ist Bilinearform.

Definition

Seien V ein K-Vektorraum und $F : V^r \rightarrow K$ eine r-Multilinearform. F heißt ALTERNIEREND, wenn für jedes $i \in \{1, \dots, r-1\}$ und jede Wahl von $a_1, \dots, a_{i-1}, a_{i+2}, \dots, a_r, a, b \in V$ gilt:

$$F(a_1, \dots, a_{i-1}, a, b, a_{i+1}, \dots, a_r) = -F(a_1, \dots, a_{i-1}, b, a, a_{i+1}, \dots, a_r)$$

Hinweis

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = -\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix}$$

6.3.1 Lemma

Sei $F : V^r \rightarrow K$ multilinear. F ist genau dann alternierend, wenn gilt

$$a_1, \dots, a_r \text{ linear abhängig} \Rightarrow F(a_1, \dots, a_r) = 0$$

Beweis

»⇐«: Seien $a, b, a_1, \dots, a_{i-1}, a_{i+2}, \dots, a_r \in V$. Dann gilt

$$\begin{aligned} 0 &= F(a_1, \dots, a_{i-1}, a + b, a + b, a_{i+2}, \dots, a_r) = \\ &F(\dots, a, a, \dots) + F(\dots, a, b, \dots) + F(\dots, b, a, \dots) + F(\dots, b, b, \dots) = \\ &F(\dots, a, b, \dots) + F(\dots, b, a, \dots) \end{aligned}$$

»⇒«: Seien a_1, \dots, a_r linear abhängig. Dann gibt es $\lambda_1, \dots, \lambda_r$, nicht alle 0, mit $\sum_{i=1}^r \lambda_i a_i = 0$. Sei $i_0 \in \{1, \dots, r\}$ mit $\lambda_{i_0} \neq 0$. Dann gilt $a_{i_0} = \sum_{\substack{i=1 \\ i \neq i_0}}^r \mu_i a_i$ mit $\mu_i = -\frac{\lambda_i}{\lambda_{i_0}}$. Aus der Multilinearität von F folgt

$$F(a_1, \dots, a_r) = \sum_{\substack{i=1 \\ i \neq i_0}}^r \mu_i F(a_1, \dots, a_{i_0-1}, a_i, a_{i_0+1}, \dots, a_r)$$

Es reicht also zu zeigen, dass $F(\dots, a, \dots, a, \dots) = 0$ ist. Durch sukzessives Vertauschen des linken a mit seinem rechten Nachbarn ergibt sich, dass es reicht, $F(\dots, a, a, \dots) = 0$ zu zeigen. Nach Definition ist aber $F(\dots, a, a, \dots) = -F(\dots, a, a, \dots)$ und es folgt die Behauptung.

6.3.2 Lemma

Seien $F : V^r \rightarrow K$ multilinear und alternierend und $\pi \in S_n$. Dann gilt für alle $a_1, \dots, a_r \in V$:

$$F(a_{\pi(1)}, \dots, a_{\pi(r)}) = \text{sign}(\pi) F(a_1, \dots, a_r)$$

Beweis

Für $r = 1$ ist die Behauptung trivial. Sei also $r \geq 2$. Nach (6.2.1) und (6.2.3(ii)) reicht es, die Behauptung für Transpositionen zu zeigen. Seien also $i, j \in \{1, \dots, r\}$ mit $i < j$ und $\tau = (j, i)$. Dann gilt

$$\begin{aligned} 0 &= F(\dots, (a_i + a_j), \dots, (a_i + a_j), \dots) = \\ &F(\dots, a_i, \dots, a_j, \dots) + f(\dots, a_j, \dots, a_i, \dots) \end{aligned}$$

also $F(a_{\tau(1)}, \dots, a_{\tau(i)}, \dots, a_{\tau(j)}, \dots, a_{\tau(r)}) = \text{sign}(\tau) F(a_1, \dots, a_r)$.

Definition

Seien V ein K -Vektorraum, $\dim V = n$, $F : V^n \rightarrow K$. F heißt DETERMINANTENFORM, wenn F n -multilinear und alternierend, aber von der Nullform verschieden ist, d.h. $F \not\equiv 0$.

6.3.3 Lemma

Seien V ein K -Vektorraum, $\dim V = n$ und $D : V^n \rightarrow K$ Determinantenform. Dann gilt

$$a_1, \dots, a_n \text{ linear abhängig} \Leftrightarrow D(a_1, \dots, a_n) = 0$$

Beweis

\Rightarrow : (6.3.1)

\Leftarrow : Seien a_1, \dots, a_n linear unabhängig, d.h. eine Basis von V . Da $D \neq 0$ ist, gibt es Vektoren $v_1, \dots, v_n \in V$ mit $D(v_1, \dots, v_n) \neq 0$. Die Basisdarstellungen seien $v_j = \sum_{i=1}^n c_{ij} a_i$. Es folgt

$$\begin{aligned} 0 \neq D(v_1, \dots, v_n) &= \\ &= \sum_{i=1}^n c_{i1} D(a_1, v_2, \dots, v_n) = \\ &= \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n c_{i_1,1} \cdots c_{i_n,n} D(a_{i_1}, \dots, a_{i_n}) = \\ &= \sum_{\pi \in S_n} c_{\pi(1),1} \cdots c_{\pi(n),n} D(a_{\pi(1)}, \dots, a_{\pi(n)}) \end{aligned}$$

Nach (6.3.2) gilt somit $0 \neq D(a_1, \dots, a_n) \sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(n),n}$. Also $D(a_1, \dots, a_n) \neq 0$.

6.3.4 Hinweis

(i)

$$\begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = \sum_{\pi \in S_2} \text{sign}(\pi) c_{\pi(1),1} c_{\pi(2),2}$$

(ii) Seien

$$\begin{aligned} A &= (a_1, \dots, a_n) \in K^{n \times n} \\ M &= (v_1, \dots, v_n) \in K^{n \times n} \\ C &= \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} \end{aligned}$$

Dann gilt $M = A \cdot C$.

Fassen wir D als Abbildung $K^{n \times n} \rightarrow K$ auf (die Argumente sind die Spaltenvektoren), so ergibt sich die Rechnung aus dem Beweis von (6.3.3): $D(M) = D(A) \sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(n),n}$.

6.3.5 Lemma

Seien V ein K -Vektorraum, $\{a_1, \dots, a_n\}$ Basis von V , $\gamma \in K \setminus \{0\}$. Ferner sei $D : V^n \rightarrow K$ definiert durch $D(v_1, \dots, v_n) := \gamma \sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(n),n}$ für $v_1, \dots, v_n \in V$, $v_j = \sum_{i=1}^n c_{ij} a_i$. Dann ist D eine Determinantenform.

Beweis

Offenbar ist D multilinear. Für $v_1 = a_1, \dots, v_n = a_n$ gilt $c_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst} \end{cases}$, also $\sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(n),n} = 1$. Somit gilt $D(a_1, \dots, a_n) = \gamma \neq 0$.

Sei nun $i_0 \in \{1, \dots, r-1\}$. Dann gilt mit $\tau = (i_0 + 1, i_0) \in S_n$

$$\begin{aligned} D(v_1, \dots, v_{i_0-1}, v_{i_0+1}, v_{i_0}, v_{i_0+2}, \dots, v_r) &= \\ \sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(i_0-1),i_0-1} \cdot c_{\pi(i_0),i_0+1} \cdot c_{\pi(i_0+1),i_0} \cdots c_{\pi(n),n} &= \\ \sum_{\pi \in \tau^{-1}S_n} \text{sign}(\pi\tau) c_{\pi\tau(1),1} \cdots c_{\pi\tau(n),n} &= \\ \text{sign}(\tau) \sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(n),n} & \end{aligned}$$

und aus $\text{sign}(\tau) = -1$ folgt die Behauptung.

6.3.6 Satz

Seien V ein K -Vektorraum, $\dim V = n$ und $D_1, D_2 : V^n \rightarrow K$ Determinantenformen. Dann existiert $\gamma \in K \setminus \{0\}$ mit $D_1 = \gamma D_2$.

Beweis

Sei a_1, \dots, a_n eine Basis von V . Nach (6.3.3) gilt

$$\begin{aligned} D_1(a_1, \dots, a_n) &\neq 0 \\ D_2(a_1, \dots, a_n) &\neq 0 \end{aligned}$$

Setze

$$\gamma = \frac{D_1(a_1, \dots, a_n)}{D_2(a_1, \dots, a_n)}$$

Mit (6.3.4)(ii) folgt die Behauptung.

6.3.7 Korollar

Sei V ein K -Vektorraum, $\dim V = n$. Die Menge der alternierenden n -Multilinearformen bildet einen eindimensionalen Vektorraum.

6.4 Determinanten von Endomorphismen und Matrizen

6.4.1 Satz

Es seien V ein K -Vektorraum, $\{a_1, \dots, a_n\}, \{b_1, \dots, b_n\}$ Basen von V , $D : V^n \rightarrow K$ eine Determinantenform und $F \in \text{End}_K(V)$. Dann gilt

$$\frac{D(F(a_1), \dots, F(a_n))}{D(a_1, \dots, a_n)} = \frac{D(F(b_1), \dots, F(b_n))}{D(b_1, \dots, b_n)}$$

Beweis

Wir nehmen zunächst an, dass F nicht bijektiv ist. Dann ist $\text{rang}(F) < n$ und aus (6.3.3) folgt $D(F(a_1), \dots, F(a_n)) = D(F(b_1), \dots, F(b_n)) = 0$.

Sei also $F \in \text{Aut}_K(V)$, d.h. ein Automorphismus. Definiere $D_F : V^n \rightarrow K$ durch $D_F(v_1, \dots, v_n) := D(F(v_1), \dots, F(v_n))$. Dann ist D_F wegen (6.3.1) und (6.3.3) eine Determinantenform, und nach (6.3.6) gibt es eine Konstante $\gamma \in K \setminus \{0\}$ mit $D_F = \gamma D$. Es folgt somit insbesondere

$$\frac{D_F(b_1, \dots, b_n)}{D(b_1, \dots, b_n)} = \frac{D_F(a_1, \dots, a_n)}{D(a_1, \dots, a_n)}$$

Hinweis

Natürlich hängt der Quotient $\frac{D(F(a_1), \dots, F(a_n))}{D(a_1, \dots, a_n)}$ nach (6.3.6) auch nicht von der gewählten Determinantenform ab.

Aufgrund dieser Invarianzen ist folgende Definition gerechtfertigt:

Definition

Seien V ein K -Vektorraum, $\{a_1, \dots, a_n\}$ eine Basis von V , $D : V^n \rightarrow K$ eine Determinantenform und $\det : \text{End}_K(V) \rightarrow K$ sei definiert durch

$$\det(F) = \frac{D(F(a_1), \dots, F(a_n))}{D(a_1, \dots, a_n)}$$

für alle $F \in \text{End}_K(V)$. Dann heißt \det DETERMINANTENFUNKTION oder kürzer DETERMINANTE, und für $F \in \text{End}_K(V)$ heißt $\det(F)$ DETERMINANTE VON F .

Hinweis

$\det(F)$ ist ein »Verzerrungsmaß« des Endomorphismus F (vgl. Substitutionsregel für mehrdimensionale Integrale, vgl. (6.1) (Flächeninhalt eines Parallelogramms), s. auch (6.5)).

6.4.2 Korollar

Seien V ein n -dimensionaler K -Vektorraum und $F, G \in \text{End}_K(V)$. Dann gilt:

- (i) $F \in \text{Aut}_K(V) \Leftrightarrow \det(F) \neq 0$
- (ii) $\det(F \circ G) = \det(F) \det(G)$
- (iii) $\det(\text{id}_V) = 1$
- (iv) $F \in \text{Aut}_K(V) \Rightarrow \det(F^{-1}) = \frac{1}{\det(F)}$

Beweis

- (i) folgt aus (6.3.3)
- (ii) Nach (i) können wir o.B.d.A. $F, G \in \text{Aut}_K(V)$ voraussetzen. Ist a_1, \dots, a_n eine Basis von V , so ist auch $G(a_1), \dots, G(a_n)$ eine Basis von V . Mit (6.4.1) und (6.4.2) folgt somit

$$\begin{aligned} \det(F \circ G) &= \frac{D(FG(a_1), \dots, FG(a_n))}{D(a_1, \dots, a_n)} \\ &= \frac{D(FG(a_1), \dots, FG(a_n))}{D(G(a_1), \dots, G(a_n))} \cdot \frac{D(G(a_1), \dots, G(a_n))}{D(a_1, \dots, a_n)} \\ &= \det(F) \cdot \det(G) \end{aligned}$$

- (iii) Sei a_1, \dots, a_n Basis von V . Dann gilt

$$\det(\text{id}) = \frac{D(\text{id}(a_1), \dots, \text{id}(a_n))}{D(a_1, \dots, a_n)} = 1$$

- (iv) $1 = \det(\text{id}) = \det(F^{-1} \circ F) = \det(F^{-1}) \cdot \det(F)$

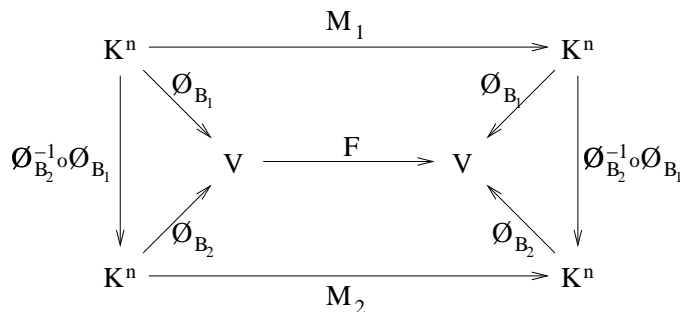


Abbildung 6.2: Schema zur Invarianz der Determinante (1)

6.4.3 Kommentar

Seien V ein n -dimensionaler K -Vektorraum, a_1, \dots, a_n eine Basis von V und $F \in \text{End}_K(V)$. Ferner sei $C = (c_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$ mit $F(a_j) = \sum_{i=1}^n c_{ij} a_i$. Sei D eine Determinantenform auf V . Nach (6.3.5) und (6.4.1) gilt dann

$$\det(F) = \frac{D(F(a_1), \dots, F(a_n))}{D(a_1, \dots, a_n)} = \sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(n),n}$$

Nun ist mit $B = \{a_1, \dots, a_n\}$ $C = M_B^B(F)$.
Im Sinne der folgenden Definition gilt

$$\det(F) = \det(M_B^B(F))$$

Definition

Für $C = (c_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$ sei

$$\det(C) := \begin{vmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{vmatrix} = \sum_{\pi \in S_n} \text{sign}(\pi) c_{\pi(1),1} \cdots c_{\pi(n),n}$$

Beachte

Hiermit ist die Definition aus 6.1 auf $n \times n$ -Matrizen verallgemeinert.

Nach (5.3.4) sind nach Wahl der Basis $K^{n \times n}$ und $\text{End}_K(V)$ isomorph. Nach (6.4.1) ist aber $\det(M_B^B(F))$ sogar unabhängig von der Wahl von B . Wieso ist das so?

Seien $F \in \text{End}_K(V)$, $B_1 = \{a_1, \dots, a_n\}$, $B_2 = \{b_1, \dots, b_n\}$ Basen von V und $M_1 = M_{B_1}^{B_1}(F)$, $M_2 = M_{B_2}^{B_2}(F)$. Ferner seine Φ_{B_1} , Φ_{B_2} die zu B_1 bzw. B_2 gehörigen Koordinatensysteme. Gemäß (5.3) liegt die schematisch in Abb. 6.2 dargestellte Situation vor.

Sei S die zu $\Phi_{B_2}^{-1} \Phi_{B_1}$ gehörige Matrix. Dann folgt (vgl. (5.3.6)) $M_2 = S M_1 S^{-1}$. Nach (6.4.2) gilt $\det(M_2) = \det(S) \det(M_1) \det(S^{-1}) = \det(M_1)$.

Definition

Seien $M_1, M_2 \in K^{n \times n}$. M_1 und M_2 heißen **ÄHNLICH**, wenn es $S \in \text{GL}(n, K)$ gibt mit $M_2 = S M_1 S^{-1}$.

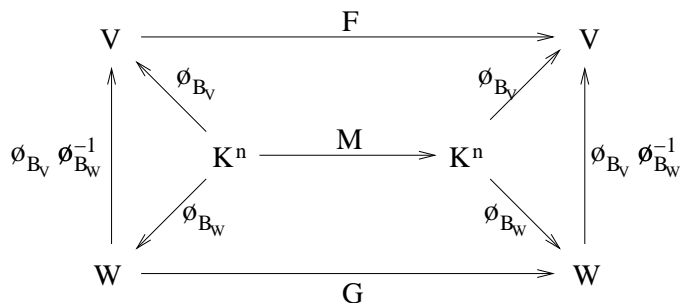


Abbildung 6.3: Schema zur Invarianz der Determinante (2)

6.4.4 Korollar

Ähnliche Matrizen besitzen die selbe Determinante. Gemäß (6.4.3) liegt noch eine zweite Invarianz vor: Beschreibt eine Matrix zwei verschiedene Endomorphismen, so ist deren Determinante gleich (s. Abb. 6.3).

6.5 Berechnung von Determinanten

6.5.1 Satz

Seien $A, B \in K^{n \times n}$.

- (i) $\det(A) = \det(A^T)$
- (ii) B geht aus A durch Vertauschung zweier Zeilen (oder Spalten) hervor. Dann gilt $\det(B) = -\det(A)$.
- (iii) B entstehe dadurch, dass zu einer Zeile (oder Spalte) von A eine Linearkombination der übrigen Zeilen (Spalten) addiert wird. Dann gilt $\det(B) = \det(A)$.
- (iv) B entstehe aus A durch Multiplikation einer Zeile (Spalte) mit einem Skalar λ . Dann gilt $\det(B) = \lambda \det(A)$.
- (v) Für $\lambda \in K$ gilt $\det(\lambda A) = \lambda^n \det(A)$.
- (vi) Sei $A \in \text{GL}(n, K)$. Dann gilt $\det(A^{-1}) = \frac{1}{\det(A)}$.
- (vii) $\det(AB) = \det(A) \det(B)$
- (viii) $\det(I_n) = 1$

Beweis

Es reicht, (i) zu zeigen; die übrigen Aussagen folgen direkt aus den Ergebnissen von (6.3) und (6.4).

Sei $A = (a_{ij})_{i,j=1,\dots,n}$. Dann gilt

$$\begin{aligned}
 \det(A) &= \sum_{\pi \in S_n} \text{sign}(\pi) a_{\pi(1),1} \cdot \dots \cdot a_{\pi(n),n} \\
 &= \sum_{\pi \in S_n} \text{sign}(\pi) a_{1,\pi^{-1}(1)} \cdot \dots \cdot a_{n,\pi^{-1}(n)} \\
 &= \sum_{\pi^{-1} \in S_n} \text{sign}(\pi) a_{1,\pi^{-1}(1)} \cdot \dots \cdot a_{n,\pi^{-1}(n)} \\
 &= \det(A^T)
 \end{aligned}$$

6.5.2 Hinweis

Sei $A \in K^{n \times n}$. Dann gilt $\text{rang}(A) = n \Leftrightarrow \det(A) \neq 0$.

Beweis

(6.4.2) und (6.4.3)

Beispiel

$$S_3 = \left\{ \underbrace{(1, 2, 3), (2, 3, 1), (3, 1, 2)}_{\text{gerade}}, \underbrace{(3, 2, 1), (2, 1, 3), (1, 3, 2)}_{\text{ungerade}} \right\}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}$$

Für größeres n ist es wesentlich effizienter, die Invarianz (6.5.1)(iii) sowie (6.5.1)(ii) auszunutzen und (2.3.2) anzuwenden. Mit Hilfe des Gauß - Algorithmus überführe man

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \quad \text{in} \quad \begin{vmatrix} b_{11} & \dots & b_{1n} \\ & \ddots & \vdots \\ 0 & & b_{nn} \end{vmatrix}$$

Es gilt $\det(A) = (-1)^k \det(B)$, wobei k die Anzahl der vorgenommenen Zeilenvertauschungen ist. (Im Unterschied zur Lösung von linearen Gleichungssystemen kann man hier auch Spaltenoperationen oder beides anwenden).

6.5.3 Hinweis

$$\begin{vmatrix} b_{11} & \dots & b_{1n} \\ & \ddots & \vdots \\ 0 & & b_{nn} \end{vmatrix} = \prod_{k=1}^n b_{kk}$$

Beweis

Es gilt $b_{ij} = 0$ für $i > j$. Sei $\pi \in S_n \setminus \{\text{id}\}$. Dann existiert ein j mit $\pi(j) > j$ und es folgt die Behauptung.

Beispiel

$$\begin{vmatrix} 1 & 2 & 2 & 4 \\ 1 & 3 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 2 & 4 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & -3 \end{vmatrix} = -6$$

Die folgende Formel führt die Determinantenberechnung rekursiv auf die Berechnung der Determinanten von Matrizen kleinerer Ordnung zurück.

Definition

Seien $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$, $i_0, j_0 \in \{1, \dots, n\}$; $M_{i_0 j_0} \in K^{(n-1) \times (n-1)}$ entstehe aus A durch Streichen der i_0 -ten Zeile und der j_0 -ten Spalte. $M_{i_0 j_0}$ heißt UNTERMATRIX von A bezüglich $a_{i_0 j_0}$, $\det(M_{i_0 j_0})$ heißt UNTERDETERMINANTE von $\det(A)$ bezüglich $a_{i_0 j_0}$ und $A_{i_0 j_0} = (-1)^{i_0+j_0} \det(M_{i_0 j_0})$ heißt ADJUNKTE zu $a_{i_0 j_0}$.

6.5.4 Entwicklungssatz von Laplace

Sei $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$. Dann gilt für $i_0, j_0 \in \{1, \dots, n\}$:

- (i) $\det(A) = \sum_{j=0}^n a_{i_0 j} A_{i_0 j}$ (Entwicklung nach i_0 -ter Zeile)
- (ii) $\det(A) = \sum_{i=0}^n a_{i j_0} A_{i j_0}$ (Entwicklung nach j_0 -ter Spalte)

Beweis

- (i) Für $i, j \in \{1, \dots, n\}$ seien $S_n^{i,j} = \{\pi \in S_n : \pi(j) = i\}$. Dann gilt für jedes $i_0 \in \{1, \dots, n\}$

$$\det(A) = \sum_{\pi \in S_n} \text{sign}(\pi) a_{\pi(1),1} \cdot \dots \cdot a_{\pi(n),n} = \sum_{j=1}^n a_{i_0 j} \hat{A}_{i_0 j}$$

mit $\hat{A}_{i_0 j} = \sum_{\pi \in S_n^{i_0, j}} \text{sign}(\pi) \prod_{\substack{k=1 \\ k \neq j}}^n a_{\pi(k),k}$. (D.h. die relevanten Untermatrizen sind korrekt und wir brauchen uns nur noch um die Vorzeichen zu kümmern.) Für $i_0 = j = n$ gilt $\hat{A}_{nn} = \sum_{\pi \in S_{n-1}} \text{sign}(\pi) \prod_{k=1}^{n-1} a_{\pi(k),k} = \det(M_{nn}) = (-1)^{n+n} \det(M_{nn}) = A_{nn}$. Im allgemeinen Fall von $\hat{A}_{i_0 j}$ wende (6.5.1)(ii) an und mache die i_0 -te Zeile von A durch sukzessives Vertauschen mit der höher indizierten Nachbarzeile zur n -ten Zeile. Danach verfähre man analog mit der j -ten Spalte. $|\det(M_{i_0 j})|$ ändert sich nicht, aber das Vorzeichen um $(-1)^{n-i_0+n-j} = (-1)^{i_0+j}$. Somit gilt $\hat{A}_{i_0 j} = (-1)^{i_0 j} \det(M_{i_0 j}) = A_{i_0 j}$.

- (ii) folgt aus (i) wegen (6.5.1)(i)

Beispiel

$$\begin{vmatrix} 1 & 2 & 2 & 4 \\ 1 & 3 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{vmatrix} = (-1)^{4+3} \cdot 2 \cdot \begin{vmatrix} 1 & 2 & 4 \\ 1 & 3 & 1 \\ 0 & 1 & 1 \end{vmatrix} + (-1)^{4+4} \cdot 1 \cdot \begin{vmatrix} 1 & 2 & 2 \\ 1 & 3 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -2 \cdot 4 + 2 = -6$$

6.6 Anwendungen

(vgl. (6.1) lineare Gleichungssysteme, Eigenwerte, Volumenberechnung)

Definition

Seien $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$ und für $i, j \in \{1, \dots, n\}$ A_{ij} die Adjunkte zu a_{ij} . Dann heißt die durch $\text{adj}(A) = (A_{ij})_{i,j=1,\dots,n}^T$ definierte Matrix die ADJUNKTE VON A .

6.6.1 Satz

Sei $A \in K^{n \times n}$. Dann gilt

- (i) $A \cdot \text{adj}(A) = \text{adj}(A) \cdot A = \det(A) \cdot I_n$
- (ii) A invertierbar $\Rightarrow A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$

Beweis

- (i) Sei $A = (a_{ik})_{i,k=1,\dots,n}$. Dann gilt

$$A \cdot \text{adj}(A) = \left(\sum_{k=1}^n a_{ik} A_{jk} \right)_{i,j=1,\dots,n}$$

Nach (6.5.4) gilt für $i = j$ stets $\sum_{k=1}^n a_{ik} A_{jk} = \det(A)$. Ferner ist für $i \neq j$ nach (6.5.4) $\sum_{k=1}^n a_{ik} A_{jk}$ die Determinante der Matrix, die aus A durch Ersetzen der j -ten Zeile durch die i -te entsteht. Da diese zwei gleiche Zeilen enthält, ist ihre Determinante nach (6.5.2) gleich 0. Es folgt somit

$$\sum_{k=1}^n a_{ik} A_{jk} = \begin{cases} \det(A) & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

Die zweite Gleichung folgt analog.

- (ii) folgt aus (i) durch Multiplikation mit A^{-1} .

Beispiel

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{pmatrix} \quad \det(A) = 3$$

$$A_{11} = -2A_{12} = 3 \quad A_{13} = -1$$

$$A_{21} = -1A_{22} = 0 \quad A_{23} = 1$$

$$A_{31} = 5 \quad A_{32} = -3A_{33} = 1$$

also:

$$A^{-1} = \frac{1}{3} \begin{pmatrix} -2 & -1 & 5 \\ 3 & 0 & -3 \\ -1 & 1 & 1 \end{pmatrix}$$

6.6.2 Korollar: Cramersche Regel

Seien $A \in K^{n \times n}$, $b \in K^n$ und es gelte $\det(A) \neq 0$. Dann ist das lineare Gleichungssystem $Ax = b$ eindeutig lösbar und für den Lösungsvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ gilt für

$$j = 1, \dots, n \quad x_j = \frac{1}{\det(A)} \sum_{i=1}^n b_i A_{ij}.$$

Beweis

Der erste Teil der Behauptung folgt aus (5.2.16) und (6.5.2). Nach (6.6.1) gilt für den Lösungsvektor $x = A^{-1}b = \frac{1}{\det(A)} \text{adj}(A)b = \frac{1}{\det(A)} \left(\sum_{i=1}^n b_i A_{ij} \right)_{j=1,\dots,n}$.

6.6.3 Bemerkung

Unter den Vorbedingungen von (6.6.2) gilt

$$x_j = \frac{\begin{pmatrix} a_{11} & \dots & a_{1,j-1} & b_1 & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & b_n & a_{n,j+1} & \dots & a_{nn} \end{pmatrix}}{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}}$$

Beweis

Folgt aus (6.6.2) und (6.5.4) durch Entwicklung nach der j -ten Spalte.

Beispiel

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ x_2 + 3x_3 &= 3 \\ x_1 + x_2 + x_3 &= 1 \end{aligned}$$

$$\begin{aligned} x_1 &= \frac{1}{3} (b_1 (-2) + b_2 (-1) + b_3 (5)) = 0 \\ x_2 &= \frac{1}{3} (b_1 (3) + b_2 (0) + b_3 (-3)) = 0 \\ x_3 &= \frac{1}{3} (b_1 (-1) + b_2 (1) + b_3 (1)) = 1 \end{aligned}$$

Hinweis

Damit sind für Gleichungssysteme die zu Beginn von (6.1) gestellten Fragen beantwortet. (6.6.1) und (6.6.2) sind als explizite Formeln von großer theoretischer Bedeutung (vgl. z.B. (2.4), »Stabilität«). Für praktische Zwecke sind sie i.a. nicht »effizient« genug.

Wir kommen kurz zur Verallgemeinerung von (6.1.2); Kapitel 7 wird sich ausführlich mit Eigenwerten etc. befassen.

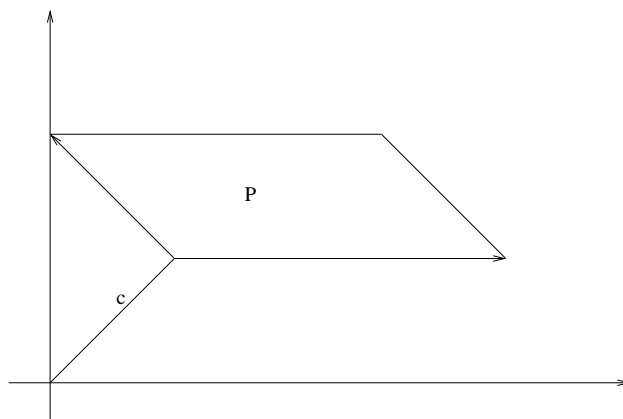
6.6.4 Bemerkung

Seien V ein endlich-dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$ und $\lambda \in K$. λ ist genau dann Eigenwert von F , wenn $\det(F - \lambda \text{id}_V) = 0$ gilt.

Beweis

Nach Definition ist λ genau dann Eigenwert von F , wenn es $v \in V \setminus \{0\}$ gibt mit $F(v) = \lambda v$, d.h. wenn $\text{kern}(F - \lambda \text{id}) \neq \{0\}$. Die Behauptung folgt aus (6.4.2).

Wir betrachten zum Abschluss dieses Kapitels noch Anwendungen auf die Volumenberechnung von »Parallelotopen« des \mathbb{R}^n .

Abbildung 6.4: Parallelotop im \mathbb{R}^2

Definition

Sei $P \subset \mathbb{R}^n$. P heißt PARALLELOTOP, wenn es $c \in \mathbb{R}^n$ und linear unabhängige Vektoren a_1, \dots, a_n des \mathbb{R}^n gibt mit $P = c + \sum_{i=1}^n [0, 1] a_i$.

Beispiel

$$n = 2, c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, a_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

$$[0, 1] a_1 + [0, 1] a_2 = \left\{ \begin{pmatrix} 2\lambda - \mu \\ \mu \end{pmatrix} : \lambda, \mu \in [0, 1] \right\}$$

Sei $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definiert durch $F \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = c + x_1 a_1 + x_2 a_2$. Dann ist $P = c + [0, 1] a_1 + [0, 1] a_2 = F \left([0, 1]^2 \right)$, d.h. das Parallelogramm P ist das Bild des Einheitswürfels $[0, 1]^2$ unter der affinen Abbildung F .

6.6.5 Hinweis

Parallelotope des \mathbb{R}^n sind affine Bilder des EINHEITSWÜRFELS $[0, 1]^n$.

Bemerkung

Der Würfel $[0, 1]^n$ hat 2^n Ecken und $2n$ Facetten (d.h. $(n-1)$ -dimensionale Seiten)

Beweis

Seien $c, a_1, \dots, a_n \in \mathbb{R}^n$, a_1, \dots, a_n linear unabhängig, $P = c + \sum_{i=1}^n [0, 1] a_i$ und $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definiert durch $F(x) = c + Ax$, wobei $A = (a_1, \dots, a_n)$ ist. Dann gilt $F([0, 1]^n) = c + A[0, 1]^n = c + \sum_{i=1}^n [0, 1] a_i = P$.

Mit Hilfe der Transformationsformel für mehrdimensionale Integrale (vgl. z.B. Königsberger, Analysis 2, S. 295ff) berechnen wir nun das Volumen eines Parallelotops.

$P = c + A[0, 1]^n$. Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \equiv 1$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $F(x) = c + Ax$. Dann gilt

$$\text{Vol}(P) = \int_P dx = \int_P f(x) dx = \int_{[0,1]^n} f(F(y)) |\det(F(y))| dy$$

Es gilt $\frac{\partial F(y)}{\partial y_i} = a_i$, also $F'(y) = A$. Es folgt

$$\text{Vol}(P) = \int_{[0,1]^n} |\det(A)| dy = |\det(A)| \int_{[0,1]^n} dy$$

also

$$\text{Vol}(P) = |\det(A)| \text{Vol}([0, 1]^n) = |\det(A)|$$

Das Vorzeichen der Determinante von $A \in \text{GL}(n, \mathbb{R})$ gibt die ORIENTIERUNG der aus den Spaltenvektoren bestehenden Basis an.

Bezeichnung

Die Relation \sim auf $\text{GL}(n, \mathbb{R})$ sei definiert durch $A \sim B \Leftrightarrow \det(A) \det(B) > 0$ für alle $A, B \in \text{GL}(n, \mathbb{R})$.

Hinweis

\sim ist eine Äquivalenzrelation, die Äquivalenzklassen sind:

$$G_+ := \{A : A \in \text{GL}(n, \mathbb{R}) \wedge \det(A) > 0\}$$

$$G_- := \{A : A \in \text{GL}(n, \mathbb{R}) \wedge \det(A) < 0\}$$

Definition

Seien a_1, \dots, a_n und b_1, \dots, b_n Basen des \mathbb{R}^n und die Matrizen $A = (a_1, \dots, a_n)$, $B = (b_1, \dots, b_n)$. Die Basen heißen GLEICH ORIENTIERT, wenn $A \sim B$ gilt.

Beispiel

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$|A| = 1, |B| = -1 \Rightarrow a_1, a_2$ und b_1, b_2 sind nicht gleich orientiert.

Definition

Seien $A, B \in \text{GL}(n, \mathbb{R})$. A und B heißen STETIG INEINANDER VERFORMBAR (HOMOTOP), wenn es eine stetige Abbildung $\varphi : [0, 1] \rightarrow \text{GL}(n, \mathbb{R})$ gibt mit $\varphi(0) = A$ und $\varphi(1) = B$.

Beispiel

$n = 1, \text{GL}(1, \mathbb{R}) = \mathbb{R} \setminus \{0\}, A, B \in \text{GL}(1, \mathbb{R})$

Für $A, B > 0$ oder $A, B < 0$ ist $\varphi(t) = (1-t)A + tB$ eine gesuchte Abbildung (s. Abb. 6.5).

Haben A und B verschiedene Vorzeichen, so kann es nach dem Zwischenwertsatz keine solche Abbildung geben.

6.6.6 Satz

Zwei Basen $A = (a_1, \dots, a_n)$ und $B = (b_1, \dots, b_n)$ des \mathbb{R}^n sind genau dann gleichorientiert, wenn A und B stetig ineinander verformbar sind.

Beweis

s. z.B. Fischer, S. 215ff

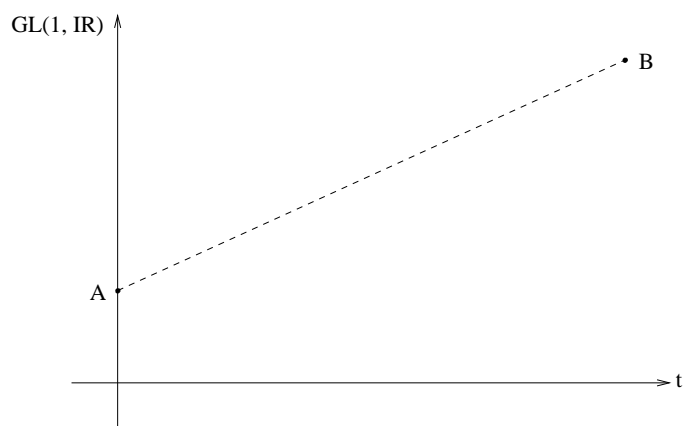


Abbildung 6.5: Stetig ineinander verformbare Matrizen

Kapitel 7

Eigenwerte und Normalformen

Wir haben bereits gesehen, dass Eigenwerte und Eigenvektoren für die »Diagonalisierung« von Matrizen, für die Berechnung von Potenzen von Matrizen und für Differenzgleichungen wichtig sind. All diese Aspekte werden wir jetzt gründlicher studieren.

Zur Einführung beginnen wir mit einem weiteren Anwendungsbeispiel: lineare Differentialgleichungen mit konstanten Koeffizienten.

7.1 Einführung

Wir betrachten folgendes Anfangswertproblem: $u_i : \mathbb{R} \rightarrow \mathbb{R}$

$$\begin{aligned} u_1'(t) &= a_{11}u_1(t) + \cdots + a_{1n}u_n(t) & u_1(0) &= u_1^* \\ &\vdots \\ u_n'(t) &= a_{n1}u_1(t) + \cdots + a_{nn}u_n(t) & u_n(0) &= u_n^* \end{aligned}$$

Mit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, u_0 = \begin{pmatrix} u_1^* \\ \vdots \\ u_n^* \end{pmatrix}, u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} : \mathbb{R} \rightarrow \mathbb{R}$$

lässt sich das Anfangswertproblem in der Form

$$u' = Au, u(0) = u_0$$

schreiben.

Wie kann man ein solches Problem lösen?

Aufgabe: Finde $u : \mathbb{R} \rightarrow \mathbb{R}^n$ mit $Au = u'$ und $u(0) = u_0$.

Betrachten wir den eindimensionalen Fall:

$$v : \mathbb{R} \rightarrow \mathbb{R}, a \in \mathbb{R}, v_0 \in \mathbb{R}, v' = av, v(0) = v_0$$

Lösung ist $v(t) = v_0 e^{at}$.

Man beachte, dass die Lösung für $t \rightarrow \infty$ *instabil* ist, falls $a > 0$ ist, für $a < 0$ aber gegen 0 strebt.

Für unser n -dimensionales Problem machen wir nun den *Ansatz* $u(t) = xe^{\lambda t}$ ($\lambda \in \mathbb{R}, x \in \mathbb{R}^n$). Durch Einsetzen in die Bedingung erhalten wir

$$\lambda x e^{\lambda t} = A(x e^{\lambda t}) = A x e^{\lambda t}$$

und das ist äquivalent zum *Eigenwertproblem* $\lambda x = Ax$.

Beispiel

$$\begin{aligned} u_1'(t) &= 4u_1(t) - 5u_2(t) & u_1(0) &= 8 \\ u_2'(t) &= 2u_1(t) - 3u_3(t) & u_2(0) &= 5 \end{aligned}$$

$$\begin{vmatrix} 4 - \lambda & -5 \\ 2 & -3 - \lambda \end{vmatrix} = (4 - \lambda)(-3 - \lambda) + 10 = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2)$$

Eigenwerte: $\lambda_1 = -1$, $\lambda_2 = 2$

Zugehörige Eigenvektoren:

$$\begin{aligned} \begin{pmatrix} 5 & -5 \\ 2 & -2 \end{pmatrix} v_1 = 0 &\Rightarrow v_1 = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \alpha \in \mathbb{R} \\ \begin{pmatrix} 2 & -5 \\ 2 & -5 \end{pmatrix} v_2 = 0 &\Rightarrow v_2 = \beta \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \beta \in \mathbb{R} \end{aligned}$$

Fundamentallösungen der Differentialgleichung:

$$w_1(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t} \quad w_2(t) = \begin{pmatrix} 5 \\ 2 \end{pmatrix} e^{2t}$$

Jede Linearkombination $c_1 w_1(t) + c_2 w_2(t)$ mit $c_1, c_2 \in \mathbb{R}$ ist ebenfalls Lösung des Differentialgleichungssystems. Um die Anfangsbedingungen zu erfüllen, sind c_1, c_2 so zu bestimmen, dass $c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \end{pmatrix}$ gilt. Da $\begin{pmatrix} 1 & 5 \\ 1 & 2 \end{pmatrix}$ vollen Rang hat, gibt es genau eine Lösung, nämlich $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Das Anfangswertproblem hat also die Lösung $3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t} + \begin{pmatrix} 5 \\ 2 \end{pmatrix} e^{2t}$.

Wegen $\lambda = 2 > 0$ ist die Lösung für $t \rightarrow \infty$ unbeschränkt. Da die Eigenvektoren der Koeffizientenmatrix den \mathbb{R}^2 aufspannen, ist das Differentialgleichungssystem für *jeden* Anfangswert lösbar. Nach (5.3) und (6.4) ist $\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix}$ zur Diagonalmatrix $\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$ ähnlich. Genauer gilt

$$\underbrace{\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix}}_M = \underbrace{\begin{pmatrix} 1 & 5 \\ 1 & 2 \end{pmatrix}}_S \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}}_D \underbrace{\begin{pmatrix} -\frac{2}{3} & \frac{5}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}}_{S^{-1}}$$

Man beachte, dass die Spalten von S gerade die gefundenen Einheitsvektoren von M sind.

Literatur

W. Walter, Gewöhnliche Differentialgleichungen, Springer-Verlag

7.2 Diagonalisierbarkeit**Definition**

Sei $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$. A heißt DIAGONALMATRIX, wenn a_{ij} für alle $i, j \in \{1, \dots, n\}$, $i \neq j$ gilt.

A heißt DIAGONALISIERBAR, wenn es eine Diagonalmatrix D und $S \in \text{GL}(n, K)$ gibt mit $A = SDS^{-1}$.

Seien V ein endlich-dimensionaler K -Vektorraum und $F \in \text{End}_K(V)$. F heißt DIAGONALISIERBAR, wenn es eine Basis B von V gibt, so dass $M_B^B(F)$ eine Diagonalmatrix ist.

7.2.1 Bemerkung

Seien V ein n -dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$, B Basis von V , $A = M_B^B(F)$, $\lambda \in K$.

- (i) F ist genau dann diagonalisierbar, wenn F eine Basis aus Eigenvektoren von F besitzt.
- (ii) F ist genau dann diagonalisierbar, wenn A diagonalisierbar ist.
- (iii) λ ist genau dann Eigenwert von F , wenn λ Eigenwert von A ist.
- (iv) Ähnliche Matrizen haben die selben Eigenwerte.

Beweis

- (i) nach (5.3) trivial.
- (ii) Basiswechsel in V überführt A in eine zu A ähnliche Matrix und umgekehrt (vgl. (6.4)).
- (iii) trivial
- (iv) folgt aus (iii)

7.2.2 Lemma

Seien V ein K -Vektorraum, $F \in \text{End}_K(V)$, $\lambda_1, \dots, \lambda_m$ paarweise verschiedene Eigenwerte von F und für $i = 1, \dots, m$ sei v_i zu λ_i gehöriger Eigenvektor. Dann sind v_1, \dots, v_m linear unabhängig.

Beweis

(vollständige Induktion nach m)

Für $m_0 = 1$: klar wegen $v_1 \neq 0$.

Die Aussage sei für $m \leq m_0 - 1$ bereits bewiesen. Seien $\mu_1, \dots, \mu_{m_0} \in K$ und $\sum_{i=1}^{m_0} \mu_i v_i = 0$. Dann gilt

$$0 = F(0) = F\left(\sum_{i=1}^{m_0} \mu_i v_i\right) = \sum_{i=1}^{m_0} \mu_i F(v_i) = \sum_{i=1}^{m_0} \mu_i \lambda_i v_i$$

Es folgt

$$0 = \left(\sum_{i=1}^{m_0} \mu_i \lambda_i v_i\right) - \lambda_{m_0} \left(\sum_{i=1}^{m_0} \mu_i v_i\right) = \sum_{i=1}^{m_0-1} \mu_i (\lambda_i - \lambda_{m_0}) v_i$$

Nach Induktionsvoraussetzung gilt $\mu_i (\lambda_i - \lambda_{m_0}) = 0$ für $i = 1, \dots, m_0 - 1$. Da die Eigenwerte paarweise verschieden sind, folgt $\mu_1 = \dots = \mu_{m_0-1} = 0$. Also gilt $\mu_{m_0} v_{m_0} = 0$. Da aber $v_{m_0} \neq 0$ ist, folgt auch $\mu_{m_0} = 0$.

7.2.3 Korollar

Seien V ein n -dimensionaler K -Vektorraum und $F \in \text{End}_K(V)$. Dann besitzt F höchstens n verschiedene Eigenwerte.

Definition

Seien V ein K -Vektorraum, $F \in \text{End}_K(V)$ und $\lambda \in K$. Dann heißt $\text{Eig}(F, \lambda) := \{v : v \in V \wedge F(v) = \lambda v\}$ EIGENRAUM von F bezüglich λ .

7.2.4 Hinweis

Seien V ein K -Vektorraum, $F \in \text{End}_K(V)$ und $\lambda \in K$. Dann gilt:

- (i) $\text{Eig}(F, \lambda)$ ist Untervektorraum von V .
- (ii) λ ist Eigenwert von $F \Leftrightarrow \text{Eig}(F, \lambda) \neq \{0\}$
- (iii) $\text{Eig}(F, \lambda) = \text{kern}(F - \lambda \text{id}_V)$
- (iv) $\lambda_1, \lambda_2 \in K \wedge \lambda_1 \neq \lambda_2 \Rightarrow \text{Eig}(F, \lambda_1) \cap \text{Eig}(F, \lambda_2) = \{0\}$

Beweis

- (i) folgt nach (4.1.3)
- (ii) folgt aus der entsprechenden Definition
- (iii) folgt aus der entsprechenden Definition
- (iv) folgt aus (7.2.2)

Nach (6.4.4) ist λ genau dann Eigenwert von F , wenn λ Nullstelle der durch $f(x) = \det(F - x \text{id}_V)$ definierten Funktion $f : K \rightarrow K$ ist. Es ist daher naheliegend, Determinanten auch mit Einträgen aus $K[X]$ zu betrachten. Da aber $K[X]$ kein Körper, sondern nur ein (kommutativer) Ring (mit Einselement) ist, sind *nicht* alle Aussagen aus Kapitel 6 verallgemeinerbar. Zur multilinearen Algebra über Modulen (\gg VR \ll über kommutativen Ringen mit Einselement) siehe etwa Kowalsky, Michler: Lineare Algebra, 1995, Kapitel 10.

Im konkreten Fall von Determinanten mit Einträgen aus $K[X]$ bzw. $K[x]$ kann man zum Körper $K(X)$ bzw. $K(x)$ übergehen.

Wir wählen im Folgenden den direkten Weg zur Definition von

$$\begin{vmatrix} a_{11} - X & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} - X \end{vmatrix}$$

über (6.3.5).

Definition

Seien V ein K -Vektorraum, $\dim V = n$, $F \in \text{End}_K(V)$, $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$ und X Unbestimmte über K . Die Funktion $f : K \rightarrow K$, definiert durch $f(x) = \det(F - x \text{id}_V)$ für alle $x \in K$ heißt CHARAKTERISTISCHE FUNKTION von F .

Das Polynom $p_A \in K[X]$ sei definiert durch

$$p_A = \sum_{\pi \in S_n} \text{sign}(\pi) \prod_{i=1}^n b_{\pi(i), i}$$

wobei

$$b_{ij} = \begin{cases} a_{ij} - X & \text{für } i = j \\ a_{ij} & \text{für } i \neq j \end{cases}$$

ist. p_A heißt CHARAKTERISTISCHES POLYNOM von A .

Hinweis

Wegen der Isomorphie von $K[X]$ und $K[x]$ (vgl. (3.3.7)) benutzen wir die selbe Notation p_A bzw. $p_A(x)$ auch für die Polynomdarstellung, die sich durch Ersetzen von X durch die Variable x ergibt.

In diesem Sinn ist $p_A \in K[x]$ Polynomdarstellung der charakteristischen Funktion f von F , falls B eine Basis von V ist mit $A = M_B^B(F)$ (vgl. (3.3)).

Bezeichnung

Mit den Bezeichnung der vorherigen Definition setzen wir

$$p_A := \begin{vmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ a_{12} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \dots & a_{n,n-1} & a_{nn} - X \end{vmatrix}$$

7.2.5 Bemerkung

- (1) Sei K unendlich und $M_1, M_2 \in K^{n \times n}$ ähnlich. Dann gilt $p_{M_1} = p_{M_2}$.
- (2) Sei K endlich und $M_1, M_2 \in K^{n \times n}$ ähnlich. Dann gilt $p_{M_1} = p_{M_2}$.

Beweis

- (1) Da K unendlich ist, folgt die Aussage wegen (3.3.10) aus (6.4.4).
- (2) Ist K ein beliebiger Körper, so folgt aus den bisherigen Ergebnissen nur, dass p_{M_1} und p_{M_2} Darstellungen der selben Polynomfunktion sind. Es gilt aber sogar $p_{M_1} = p_{M_2}$.

Beweisskizze:

$$\begin{aligned} |M_1 - XI| &= |SM_2S^{-1} - XI| = |S(M_2 - XI)S^{-1}| \\ &= |S||M_2 - XI||S^{-1}| = |M_2 - XI| \end{aligned}$$

ist »formal« zu verifizieren (vgl. Fischer, 3.2.8, 4.1.3; Kowalsky, Michler, Kap. 10)

Bezeichnung

Seien V ein endlich-dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$, B eine Basis von V , $A = M_B^B(F)$ und p_A charakteristisches Polynom von A . Dann heißt p_A CHARAKTERISTISCHES POLYNOM VON F und wird auch mit p_F bezeichnet.

7.2.6 Hinweis

Seien V ein endlich-dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$ und $\lambda \in K$. λ ist genau dann Eigenwert von F , wenn λ Nullstelle von p_F ist.

Hiermit ist die Verallgemeinerung von (6.1.3) - das geometrische Problem invarianter Unterräume $\text{Eig}(F, \lambda)$ - auf die algebraischen Probleme der Berechnung der Nullstellen von p_F und der Bestimmung von $\text{kern}(F - \lambda \text{id}_V)$ zurückgeführt.

Beispiel

Für $\alpha \in [0, 2\pi[$ sei

$$A = A(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

Dann ist

$$p_A(x) = \begin{vmatrix} \cos(\alpha) - x & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) - x \end{vmatrix} = x^2 - 2\cos(\alpha)x + 1$$

p_A besitzt genau dann reelle Nullstellen, wenn $\cos^2(\alpha) \geq 1$, d.h. $\alpha \in \{0, \pi\}$ ist. Es gilt

$$A(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -A(\pi)$$

In (6.1) wurde bereits gezeigt, welche Gestalt p_A für 2×2 -Matrizen hat, nämlich $p_A(x) = x^2 - \text{spur}(A)x + \det(A)$. Das soll jetzt verallgemeinert werden.

Definition

Die Abbildung $\text{spur} : K^{n \times n} \rightarrow K$ sei definiert durch

$$\text{spur} \left((a_{ij})_{i,j=1,\dots,n} \right) = \sum_{i=1}^n a_{ii}$$

Für jedes $A \in K^{n \times n}$ heißt $\text{spur}(A)$ die SPUR von A (alternative Bezeichnungen: $\text{trace}(A)$, $\text{tr}(x)$).

7.2.7 Hinweis

Ähnliche Matrizen besitzen die selbe Spur.

7.2.8 Bemerkung

Sei $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$. Seien $q_1, \dots, q_n \in K$ mit $p_A = \sum_{i=0}^n q_i X^i$. Dann gilt $q_0 = \det(A)$, $q_{n-1} = (-1)^{n-1} \text{spur}(A)$ und $q_n = (-1)^n$.

Beweis

Nach Definition gilt

$$p_A = \sum_{\pi \in S_n} \text{sign}(\pi) \prod_{i=1}^n b_{\pi(i),i}$$

mit

$$b_{ij} = \begin{cases} a_{ij} - X & i = j \\ a_{ij} & i \neq j \end{cases}$$

Für $\pi_0 = \text{id} \in S_n$ ist somit $\prod_{i=1}^n b_{\pi_0(i),i} = \prod_{i=1}^n (a_{ii} - X)$, d.h. ein Polynom vom Grad n mit führendem Koeffizienten $(-1)^n$. Für alle anderen Permutationen π gilt $\pi(i) \neq i$ für mindestens zwei Indizes $i \in \{1, \dots, n\}$. Also ist $\sum_{\pi \in S_n \setminus \{\pi_0\}} \text{sign}(\pi) \prod_{i=1}^n b_{\pi(i),i}$ ein Polynom vom Grad höchstens $n-2$ und es folgt $q_n = (-1)^n$ und $q_{n-1} = (-1)^{n-1} \sum_{i=1}^n a_{ii} = (-1)^{n-1} \text{spur}(A)$. Ferner gilt $q_0 = \sum_{\pi \in S_n} \text{sign}(\pi) \prod_{i=1}^n a_{\pi(i),i} = \det(A)$.

7.2.9 Hinweis

Seien V ein K -Vektorraum, $\dim V = n$ und $F \in \text{End}_K(V)$.

- (i) Ist F diagonalisierbar, so zerfällt p_F in Linearfaktoren, d.h. p_F ist von der Form $\pm(x - \lambda_1) \dots (x - \lambda_n)$.
- (ii) Ist $p_F = \pm(x - \lambda_1) \dots (x - \lambda_n)$ mit paarweise verschiedenen $\lambda_1, \dots, \lambda_n \in K$, so ist F diagonalisierbar.

Beweis

(i) trivial

(ii) folgt aus (7.2.2)

Es bleibt noch zu klären, was im Fall von mehrfachen Nullstellen von p_F passiert. Wegen (7.2.6) können wir von der *Vielfachheit eines Eigenwerts* sprechen und mit $r_i = \mu(p_F, \lambda_i)$ (vgl. (3.3)) ist $p_F(x) = \pm(x - \lambda_1)^{r_1} \dots (x - \lambda_k)^{r_k}$, wenn $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Nullstellen sind. Wegen (7.2.9) gilt für diagonalisierbares F natürlich $\sum_{i=1}^k r_k = n$.

7.2.10 Lemma

Seien V ein n -dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$ und $\lambda \in K$ Eigenwert von F . Dann gilt

$$\dim(\text{Eig}(F, \lambda)) \leq \mu(p_F, \lambda)$$

Beweis

Seien $s \in \mathbb{N}$, $B = \{v_1, \dots, v_n\}$ Basis von V , so dass $\text{Eig}(F, \lambda) = \text{span}\{v_1, \dots, v_s\}$ (vgl. (4.2.5)). Dann hat $M_B^B(F)$ die Gestalt

$$\begin{pmatrix} \lambda & & 0 & & \\ & \ddots & & & * \\ 0 & & \lambda & & \\ & & 0 & & \hat{A} \end{pmatrix}$$

Mit (7.2.5) folgt $p_F = (\lambda - x)^s p_{\hat{A}}$ und es folgt die Behauptung.

Beispiel

Es sei

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 & 1 \\ 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}$$

Dann gilt $p_A(x) = (1 - x)^n$, d.h. $\mu(p_A, 1) = n$. Ferner ist $\text{rang}(A - I_n) = 1$, d.h. $\dim \text{Eig}(A, 1) = n - 1$, d.h. A ist nicht diagonalisierbar.

7.2.11 Satz

Seien V ein endlich-dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:

- (i) F ist diagonalisierbar.
- (ii) p_F zerfällt in Linearfaktoren und $\dim \text{Eig}(F, \lambda) = \mu(p_F, \lambda)$ gilt für jeden Eigenwert λ von F .
- (iii) Bezeichnen $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von F , so ist $V = \text{Eig}(F, \lambda_1) \oplus \dots \oplus \text{Eig}(F, \lambda_k)$.

Beweis

(i) \Rightarrow (ii) folgt aus (7.2.1) und (7.2.9).

(ii) \Rightarrow (iii) folgt aus (7.2.4)(iv) und $\sum_{i=1}^k \mu(p_F, \lambda_i) = n$.

(iii) \Rightarrow (i) trivial

7.3 Die Jordansche Normalform

Nach (7.2.11) gibt es zwei Bedingungen für die Diagonalisierbarkeit eines Endomorphismus: Das charakteristische Polynom zerfällt in Linearfaktoren und die Eigenräume haben maximale Dimension. Die erste Bedingung ist wegen (3.3.11) über \mathbb{C} stets erfüllt. Wir untersuchen daher, ob ein Endomorphismus über \mathbb{C} wenigstens »nahezu« diagonalisierbar ist. Wir führen zunächst die zugehörigen Matrizen ein.

Definition

Seien $J \in K^{n \times n}$. J heißt JORDAN-MATRIX oder JORDANSCHER NORMALFORM, wenn es $\lambda_1, \dots, \lambda_s \in K$ und $n_1, \dots, n_s \in \mathbb{N}$ mit $n_1 + \dots + n_s = n$ gibt, so dass folgendes gilt: Für $i = 1, \dots, s$ sei

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \vdots & \vdots & 0 & \lambda_i \end{pmatrix} \in K^{n_i \times n_i}$$

Dann hat J die Gestalt

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_s \end{pmatrix}$$

Für $i = 1, \dots, s$ heißt J_i JORDAN-BLOCK (ZUM EIGENWERT λ_i).

Bemerkung

Hauptziel von (7.3) ist es, zu zeigen, dass für $K = \mathbb{C}$ jede Matrix A einer Jordan-Matrix J ähnlich ist. Äquivalent bedeutet das, dass zu jedem Endomorphismus F eine Basis B existiert, so dass $M_B^B(F)$ Jordan-Matrix ist.

Beispiel

$F \in \text{End}_{\mathbb{C}}(\mathbb{C}^3)$, $B = \{b_1, b_2, b_3\}$ Basis von \mathbb{C}^3 , $M_B^B(F) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$. Es folgt

$$(I) \quad F(b_1) = \lambda b_1,$$

$$(II) \quad F(b_2) = \lambda b_2 + b_1 \text{ und}$$

$$(III) \quad F(b_3) = \lambda b_3 + b_2$$

(I) besagt, dass b_1 Eigenvektor zum Eigenwert λ ist, bzw. $b_1 \in \text{kern}(F - \lambda \text{id}) = \text{Eig}(F, \lambda)$. Da $\text{rang}(F - \lambda \text{id}) = 2$ ist, gilt wegen (5.2.7) $\dim \text{Eig}(F, \lambda) = 1$; es gibt also keinen weiteren von b_1 linear unabhängigen Eigenvektor.

Gleichung (II) bedeutet $b_1 = F(b_2) - \lambda b_2 \in \text{bild}(F - \lambda \text{id})$, also $b_1 \in \text{kern}(F - \lambda \text{id}) \cap \text{bild}(F - \lambda \text{id})$ und es gilt $b_2 \in \text{kern}\left((F - \lambda \text{id})^2\right)$.

Analog liefert (III) $b_2 = F(b_3) - \lambda b_3 \in \text{bild}(F - \lambda \text{id}) \cap \text{kern}(F - \lambda \text{id})^2$, also $b_3 \in \text{kern}(F - \lambda \text{id})^3$.

Natürlich gilt $\text{Eig}(F, \lambda) = \text{kern}(F - \lambda \text{id}) \subset \text{kern}(F - \lambda \text{id})^2 \subset \text{kern}(F - \lambda \text{id})^3$ und die Inklusionen sind sämtlich strikt. $\text{kern}(F - \lambda \text{id})^3$ ist dreidimensional, d.h. ganz \mathbb{C}^3 .

Definition

Seien V ein K -Vektorraum, U ein Untervektorraum von V und $F \in \text{End}_K(V)$. U heißt F -INVARIANT, wenn $F(U) \subset U$ ist.

7.3.1 Hinweis

Seien V ein K -Vektorraum, U ein Untervektorraum und $F \in \text{End}_K(V)$. Dann gilt:

- (i) $U \subset \text{kern}(F) \Rightarrow U$ ist F -invariant
- (ii) $\text{bild}(F) \subset U \Rightarrow U$ ist F -invariant

Beweis

trivial

7.3.2 Lemma

Seien V ein n -dimensionaler K -Vektorraum mit $n \geq 2$, $F \in \text{End}_K(V)$ und es gelte $\text{kern}(F) \not\subset \text{bild}(F)$. Dann gibt es F -invariante Unterräume U_1, U_2 von V mit $V = U_1 \oplus U_2$.

Beweis

Für $\text{kern}(F) = V$ ist die Aussage nach (7.3.1)(i) klar. Es gelte also $\text{kern}(F) \neq V$. Sei $W = \text{kern}(F) \cap \text{bild}(F)$. Nach (4.3.4) gibt es in $\text{kern}(F)$ ein Komplement U_1 von W . Nach Voraussetzung ist $\text{kern}(F) \setminus W \neq \emptyset$, also $1 \leq \dim(U_1) \leq n-1$. Ferner sei U_2 komplementär zu U_1 so, dass $\text{bild}(F) \subset U_2$ ist. Nach (7.3.1) sind U_1 und U_2 F -invariant und es folgt die Behauptung.

7.3.3 Lemma

Seien V ein endlich-dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$ und W_1, W_2 F -invariante Unterräume mit $\text{bild}(F) = W_1 \oplus W_2$. Dann gibt es Untervektorräume U_1, U_2 mit folgenden Eigenschaften:

- (i) U_1, U_2 sind F -invariant.
- (ii) $W_1 \subset U_1 \wedge W_2 \subset U_2$
- (iii) $V = U_1 \oplus U_2$

Beweis

Nach (5.1.2)(iii) sind $F^{-1}(W_1), F^{-1}(W_2)$ Untervektorräume von V , und wegen $\text{bild}(F) = W_1 \oplus W_2$ gilt $V = F^{-1}(W_1) + F^{-1}(W_2)$. Da W_1, W_2 F -invariant sind, gilt $W_1 \subset F^{-1}(W_1)$ und $W_2 \subset F^{-1}(W_2)$. Wir können daher die Untervektorräume U_1, U_2 so wählen, dass gilt:

$W_1 \subset U_1 \subset F^{-1}(W_1)$; U_1 ist komplementär in $F^{-1}(W_1)$ zu $(F^{-1}(W_1) \cap W_2)$
 $W_2 \subset U_2 \subset F^{-1}(W_2)$; U_2 ist komplementär in $F^{-1}(W_2)$ zu $(F^{-1}(W_2) \cap U_1)$
 Dann sind (i) und (ii) erfüllt und es gilt $U_1 \cap U_2 = \{0\}$. Ferner gilt

$$\begin{aligned} F^{-1}(W_1) &\subset U_1 + (F^{-1}(W_1) \cap W_2) \subset U_1 + W_2 \subset U_1 + U_2 \\ F^{-1}(W_2) &\subset (F^{-1}(W_2) \cap U_1) + U_2 \subset U_1 + U_2 \end{aligned}$$

also $V = F^{-1}(W_1) + F^{-1}(W_2) \subset U_1 + U_2$.

Definition

Seien V ein n -dimensionaler K -Vektorraum, $B = \{b_1, \dots, b_n\}$ Basis von V und $F \in \text{End}_K(V)$. B heißt **JORDANBASIS** von V bezüglich F , wenn es $\delta_2, \dots, \delta_n \in \{0, 1\} \subset K$ und $\lambda_1, \dots, \lambda_n \in K$ gibt mit $i \in \{2, \dots, n\} \wedge \lambda_{i-1} \neq \lambda_i \Rightarrow \delta_i = 0$, so dass $F(b_1) = \lambda_1 b_1$ und $F(b_i) = \delta_i b_{i-1} + \lambda_i b_i$ für $i = 2, \dots, n$ ist.

7.3.4 Hinweis

Seien V ein endlich-dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$ und U_1, U_2 F -invariante Untervektorräume mit $V = U_1 \oplus U_2$. Besitzt U_1 eine Jordanbasis bezüglich $F|_{U_1}$ und U_2 eine Jordanbasis bezüglich $F|_{U_2}$, so besitzt V eine Jordanbasis bezüglich F .

Beweis

Die Vereinigung der Jordanbasen von U_1 und U_2 bildet (bei geeigneter Nummerierung) eine Jordanbasis von V .

7.3.5 Hinweis

Seien V ein K -Vektorraum, U ein Untervektorraum von V , $F \in \text{End}_K(V)$ und $\lambda \in K$. Dann ist U genau dann F -invariant, wenn U $(F - \lambda \text{id}_V)$ -invariant ist.

Beweis

\Rightarrow : $F(U) \subset U \Rightarrow (F - \lambda \text{id}_V)U \subset F(U) + U \subset U$

\Leftarrow : Gilt $(F - \lambda \text{id}_V)(u) \in U$ für jedes $u \in U$, so folgt $F(u) \in \lambda u + U = U$ für alle $u \in U$.

7.3.6 Hinweis

Seien V ein endlich-dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$ und U ein F -invarianter Untervektorraum. Zerfällt p_F in Linearfaktoren, so auch $p_{F|_U}$.

Beweis

$p_{F|_U}$ ist Teiler von p_F .

7.3.7 Satz

Seien V ein n -dimensionaler K -Vektorraum, $F \in \text{End}_K(V)$ und p_F zerfalle (über K) in Linearfaktoren. Dann besitzt V eine Jordanbasis bezüglich F .

Beweis

(vollständige Induktion nach n)

Für $n = 1$ ist die Aussage trivial.

Sei also $n \geq 2$ und die Behauptung gelte für unterdimensionale Vektorräume. Da p_F in Linearfaktoren zerfällt, besitzt F nach (7.2.6) einen Eigenwert. Sei also λ Eigenwert von F . Dann gilt $\text{rang}(F - \lambda \text{id}) < n$.

Gilt $\text{kern}(F - \lambda \text{id}) \not\subset \text{bild}(F - \lambda \text{id})$, so gibt es nach (7.3.2) $(F - \lambda \text{id})$ -invariante Untervektorräume U_1, U_2 mit $V = U_1 \oplus U_2$. Nach (7.3.5) sind U_1, U_2 F -invariant. Wegen (7.3.6) besitzen U_1, U_2 nach Induktionsvoraussetzung Jordanbasen bezüglich $F|_{U_1}$ bzw. $F|_{U_2}$ und die Behauptung folgt aus (7.3.4).

Im Folgenden gelte $\text{kern}(F - \lambda \text{id}) \subset \text{bild}(F - \lambda \text{id})$. Nach (7.3.1) und (7.3.5) ist $\text{bild}(F - \lambda \text{id})$ ein F -invarianter Untervektorraum von V . Wegen (7.3.6) besitzt $\text{bild}(F - \lambda \text{id})$ eine Jordanbasis $\{b_1, \dots, b_k\}$ bezüglich $F|_{\text{bild}(F - \lambda \text{id})}$, d.h. es gibt $\delta_2, \dots, \delta_k \in \{0, 1\}$ und $\lambda_1, \dots, \lambda_k \in K$ mit $i \in \{2, \dots, k\} \wedge \lambda_{i-1} \neq \lambda_i \Rightarrow \delta_i = 0$, so dass $F(b_1) = \lambda_1 b_1$ und $F(b_i) = \delta_i b_{i-1} + \lambda_i b_i$ für $i = 2, \dots, k$ ist.

Wir nehmen zunächst an, dass es einen Index $i_0 \in \{2, \dots, k\}$ gibt mit $\delta_{i_0} = 0$. Seien $W_1 = \text{span}\{b_1, \dots, b_{i_0-1}\}$ und $W_2 = \text{span}\{b_{i_0}, \dots, b_k\}$. Dann sind W_1, W_2 F -invariant, nach (7.3.5) also auch $(F - \lambda \text{id})$ -invariant und es gilt $\text{bild}(F - \lambda \text{id}) = W_1 \oplus W_2$. Nach (7.3.3) und (7.3.5) ist V direkte Summe von zwei $(F - \lambda \text{id})$ -invarianten Untervektorräumen und nach (7.3.4) und (7.3.6) folgt die Behauptung aus der Induktionsvoraussetzung.

Wir können also im Folgenden annehmen, dass $\delta_2 = \dots = \delta_k = 1$ gilt. Insbesondere folgt $\lambda_1 = \dots = \lambda_k$. Da $\{0\} \neq \text{kern}(F - \lambda \text{id}) \subset \text{bild}(F - \lambda \text{id})$ ist, gilt $\lambda = \lambda_i$ und $\dim \text{kern}(F - \lambda \text{id}) = 1$. Somit ist $k = \text{rang}(F - \lambda \text{id}) = n - 1$. Sei nun $b \in V \setminus \text{bild}(F - \lambda \text{id})$. Dann ist $\bar{B} = \{b_1, \dots, b_{n-1}, b\}$ eine Basis von V .

Seien $\mu_1, \dots, \mu_{n-1}, \mu \in K$ mit $F(b) = \left(\sum_{i=1}^{n-1} \mu_i b_i\right) + \mu b$. Wegen $(F - \lambda \text{id}) \in \text{bild}(F - \lambda \text{id}) = \text{span}\{b_1, \dots, b_{n-1}\}$ folgt $\mu = \lambda$ und es gilt

$$M_{\bar{B}}^{\bar{B}} = \begin{pmatrix} \lambda & 1 & & 0 & \mu_1 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & 1 & \vdots \\ 0 & & & \lambda & \mu_{n-1} \\ & & & & \lambda \end{pmatrix}$$

Da $\text{kern}(F - \lambda \text{id})$ eindimensional ist, folgt $\mu_{n-1} \neq 0$. Gesucht ist nun ein Vektor b_n , so dass $\{b_1, \dots, b_n\}$ Jordanbasis von V bezüglich F ist, d.h. b_1, \dots, b_n linear unabhängig und $F(b_n) = \delta_n b_{n-1} + \lambda_n b_n$ mit $\delta_n \in \{0, 1\}$, $\lambda \in K$ und $\lambda_{n-1} \neq$

$\lambda_n \Rightarrow \delta_n = 0$. Mit $b_n = \left(\sum_{i=1}^{n-1} \alpha_i b_i\right) + \alpha b$ führt dieses auf das folgende lineare Gleichungssystem:

$$\begin{pmatrix} \lambda & 1 & & 0 & \mu_1 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & 1 & \vdots \\ 0 & & & \lambda & \mu_{n-1} \\ & & & & \lambda \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{n-2} \\ \alpha_{n-1} \\ \alpha \end{pmatrix} = \begin{pmatrix} \lambda_n \alpha_1 \\ \vdots \\ \lambda_n \alpha_{n-2} \\ \lambda_n \alpha_{n-1} + \delta_{n-1} \\ \lambda_n \alpha \end{pmatrix}$$

Da b_1, \dots, b_n linear unabhängig sein sollen, muss $\alpha \neq 0$ gelten, und es folgt $\lambda = \lambda_n$ und wegen $\mu_{n-1} \neq 0$ auch $0 \neq \delta_n = 1$. Somit verbleibt das lineare Gleichungssystem

$$\begin{pmatrix} 0 & 1 & & 0 & \mu_1 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \dots & 0 & 1 & \mu_{n-1} \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{n-1} \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

also

$$\begin{pmatrix} 1 & 0 & \dots & 0 & \mu_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & \mu_{n-1} \end{pmatrix} \begin{pmatrix} \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Nach (5.2.15) ist das System für $\alpha_2, \dots, \alpha_{n-1}, \alpha$ eindeutig lösbar; α_1 kann beliebig gewählt werden. (Woher kommt der Freiheitsgrad?)

Insgesamt folgt hiermit die Behauptung.

Beispiel

Im Folgenden wollen wir die Jordan-Normalform mit Hilfe von Rangbetrachtungen bestimmen. Ist J bekannt, so erhält man aus $A = SJS^{-1}$ das Gleichungssystem $AS - SJ = 0$ in den Einträgen von S (das von jeder Jordanbasis erfüllt wird).

Seien $V = \mathbb{Q}^4$ und

$$A = \begin{pmatrix} 1 & -\frac{4}{3} & -\frac{2}{3} & \frac{10}{3} \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Es gilt

$$\begin{aligned} p_F(x) &= \det(A - xI) \\ &= (1-x) \begin{vmatrix} 2-x & 0 & -1 \\ 0 & 1-x & 0 \\ 1 & 0 & -x \end{vmatrix} \\ &= (1-x)^4 \end{aligned}$$

Für die JNF J kommen hier (bis auf Permutation der Jordanblöcke) nur die folgenden Matrizen in Frage:

$$J_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad J_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad J_3 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$J_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad J_5 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Nun gilt $\text{rang}(F - \lambda \text{id}) = 2$ und ferner $\text{rang}(J_1 - I) = 0$, $\text{rang}(J_2 - I) = 1$, $\text{rang}(J_3 - I) = 2$, $\text{rang}(J_4 - I) = 2$ und $\text{rang}(J_5 - I) = 3$. J muss also (bis auf Permutation der Blöcke) gleich J_3 oder J_4 sein.

Nun gilt

$$(J_3 - I)^2 = 0$$

$$(J_4 - I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

also $\text{rang}(J_3 - I)^2 = 0$ und $\text{rang}(J_4 - I)^2 = 1$. Ferner gilt

$$(A - I)^2 = \begin{pmatrix} 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

d.h. $\text{rang}(A - I)^2 = 1$. Da die Ränge bei Koordinatentransformation invariant sind, muss J_4 die JNF von A sein. Das zugehörige lineare Gleichungssystem lautet mit $S = (s_1, s_2, s_3, s_4)$:

$$AS - S - (0, 0, s_2, s_3) = (A - I)S - (0, 0, s_2, s_3)$$

$$\begin{pmatrix} A - I & 0 & 0 & 0 \\ 0 & A - I & 0 & 0 \\ 0 & -I & A - I & 0 \\ 0 & 0 & -I & A - I \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = 0$$

Man beachte, dass S vollen Rang haben muss.

Bezeichnung

Sei $J \in K^{n \times n}$ eine Jordanmatrix. Seien $t \in \mathbb{N}$, $\lambda_1, \dots, \lambda_t \in K$ paarweise verschieden und $r_1, \dots, r_t \in \mathbb{N}$, so dass

$$p_J = \pm \prod_{i=1}^t (x - \lambda_i)^{r_i}$$

Ferner sei für $i = 1, \dots, t$ und $j = 1, \dots, n$ k_{ij} die Anzahl der Jordanblöcke von J zum Eigenwert λ_i der Größe j . Dann heißen $(n, t, \lambda_i, r_i, k_{ij})$ PARAMETER von J .

7.3.8 Hinweis

Jede Jordanmatrix J ist bis auf Permutation ihrer Blöcke eindeutig durch ihre Parameter bestimmt. Ferner gilt $\sum_{i=1}^t r_i = n$ und für jedes $i = 1, \dots, t$ gilt $\sum_{j=1}^n j k_{ij} = r_i$, $\sum_{i=1}^t \sum_{j=1}^n k_{ij}$ ist die Anzahl s der Jordanblöcke von J .

7.3.9 Lemma

Es sei J eine Jordanmatrix mit den Parametern $(n, t, \lambda_i, r_i, k_{ij})$. Dann gilt für $i = 1, \dots, t$:

$$(i) \sum_{j=1}^n k_{ij} = n - \text{rang}(J - \lambda_i I)$$

$$(ii) \text{ und für } l = 2, \dots, n: \sum_{j=l}^n k_{ij} = \text{rang}(J - \lambda_i I)^{l-1} - \text{rang}(J - \lambda_i I)^l$$

Beweis

Seien $J_1 \in K^{n_1 \times n_1}, \dots, J_s \in K^{n_s \times n_s}$ die Jordanblöcke von J , $p \in \{1, \dots, s\}$, $i, q \in \{1, \dots, t\}$ und sei λ_q Eigenwert von J_p . Dann gilt für $l = 1, \dots, n$ mit der Setzung $M^0 = I$

$$\begin{aligned} (J_p - \lambda_i I)^l &= \begin{pmatrix} \lambda_q - \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda_q - \lambda_i \end{pmatrix}^l \\ &= \begin{pmatrix} & \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \\ (\lambda_q - \lambda_i) I + & \end{pmatrix}^l \\ &= \sum_{k=0}^l \binom{l}{k} (\lambda_q - \lambda_i)^k \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}^{l-k} \end{aligned}$$

also

$$\text{rang}(J_p - \lambda_i I)^l = \begin{cases} n_p & \text{für } q \neq i \\ n_p - l & \text{für } q = i \wedge n_p \geq l + 1 \\ 0 & \text{für } q = i \wedge n_p \leq l \end{cases}$$

Somit gilt

$$\begin{aligned} \text{rang}(J - \lambda_i I)^{l-1} - \text{rang}(J - \lambda_i I)^l &= \sum_{p=1}^s (\text{rang}(J_p - \lambda_i I)^{l-1} - \text{rang}(J_p - \lambda_i I)^l) \\ &= \sum_{j=l}^n k_{ij} \end{aligned}$$

und somit die Behauptung.

Bemerkung

Aus (7.3.9) lassen sich alle k_{ij} bestimmen; ferner sind alle Parameter invariant gegenüber Ähnlichkeitstransformationen (Basiswechsel).

7.3.10 Satz

Seien $A, B \in K^{n \times n}$ und J_A, J_B zu A bzw. B ähnliche Jordanmatrizen. Dann sind A und B ähnlich genau dann, wenn J_A bis auf Permutation der Jordanblöcke mit J_B übereinstimmt.

Beweis

$\gg\leftarrow\leftarrow$: trivial

$\gg\Rightarrow\leftarrow$: Seien $S \in \text{GL}(n, K)$ mit $A = SBS^{-1}$ und $l \in \mathbb{N}$. Dann gilt $A^l = (SBS^{-1})^l = SB^lS^{-1}$. Da ähnliche Matrizen den selben Rang besitzen, haben J_A und J_B nach (7.3.9) die selben Parameter und die Behauptung folgt aus (7.3.8).

Beispiel

(s. oben)

$$\begin{aligned} k_1 + k_2 + k_3 + k_4 &= 4 - 2 = 2 \\ k_2 + k_3 + k_4 &= 2 - 1 = 1 \\ k_3 + k_4 &= 1 - 0 = 1 \\ k_4 &= 0 \end{aligned}$$

Es folgt $k_1 = k_3 = 1$ und $k_2 = k_4 = 0$.

Bemerkung

Aus (7.3.10) folgt insbesondere, dass zwei Jordanmatrizen genau dann ähnlich sind, wenn sie bis auf Permutation der Jordanblöcke übereinstimmen.

Natürlich kann man wegen der speziellen Beziehung von \mathbb{R} und \mathbb{C} mittels (7.3.7) auch reelle Normalformen herleiten (falls p_F über \mathbb{R} nicht in Linearfaktoren zerfällt); vgl. Biskorn, Lineare Algebra und analytische Geometrie II, S. 165ff oder Kowalsky, Michler, Lineare Algebra, S. 338f.

7.4 Anwendungen auf Differenzen- und Differentialgleichungen

Wir betrachten zunächst noch einmal das Beispiel aus (1.2).

$$a_i = Ax_{i-1}, \quad A = \begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,9 \end{pmatrix}, \quad x_0, \dots \in \mathbb{R}^2$$

Es gilt

$$p_A(x) = (0,8 - x)(0,9 - x) - 0,02 = (x - 1)(x - 0,7), \quad x \in \mathbb{R}$$

Nach (7.2.2) bzw. (7.2.11) ist A zur Diagonalmatrix

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0,7 \end{pmatrix}$$

ähnlich. Es gilt

$$\begin{aligned} \text{kern}(A - I) &= \{x \in \mathbb{R}^2 : (-0,2, 0,1)x = 0\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\} \\ \text{kern}(A - 0,7I) &= \{x \in \mathbb{R}^2 : (0,1, 0,1)x = 0\} = \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \end{aligned}$$

und mit

$$S = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}, \quad S^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$$

folgt $A = SDS^{-1}$. Ferner ist

$$\begin{aligned} x_i &= A^i x_0 = (SD^i S^{-1})x_0 \\ &= \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0,7^i \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} x_0 \\ &= \frac{1}{3} \begin{pmatrix} 1+2 \cdot 0,7^i & 1-0,7^i \\ 2(1-0,7^i) & 2+0,7^i \end{pmatrix} x_0 \\ &\xrightarrow{i \rightarrow \infty} \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} x_0 \\ &= \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} (x_{01} + x_{02}) \end{aligned}$$

wobei $x_0 = \begin{pmatrix} x_{01} \\ x_{02} \end{pmatrix}$.

Es findet also Konvergenz statt gegen den Gleichgewichtszustand $\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} p$, wobei p die Gesamtpopulation ist.

Beispiel

($n = 3$)

Führt die Matrix einer 3×3 -Differenzgleichung etwa auf

$$J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

so ist wegen $A^i = SJ^i S^{-1}$ die Potenz J^i zu berechnen. Es gilt $J = \lambda I + C$ mit

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

Wegen

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

und $C^i = 0$ für $i \geq 3$ folgt für $i \geq 2$

$$\begin{aligned} J^i &= \sum_{l=0}^i \binom{i}{l} \lambda^l I^l C^{i-l} \\ &= \binom{i}{i} \lambda^i I C^0 + \binom{i}{i-1} \lambda^{i-1} I C + \binom{i}{i-2} \lambda^{i-2} I C^2 \\ &= \lambda^i I + i \lambda^{i-1} I C + \frac{1}{2} i(i-1) \lambda^{i-2} I C^2 \\ &= \begin{pmatrix} \lambda^i & i \lambda^{i-1} & \frac{1}{2} i(i-1) \lambda^{i-2} \\ 0 & \lambda^i & i \lambda^{i-1} \\ 0 & 0 & \lambda^i \end{pmatrix} \end{aligned}$$

In (7.1) haben wir gesehen, wie man im Falle der Diagonalisierbarkeit der Koeffizientenmatrix lineare Anfangswertprobleme mit konstanten Koeffizienten behandeln kann:

$$\begin{aligned} u' = Au &\rightarrow u' = (SDS^{-1})u \rightarrow (S^{-1}u)' = D(S^{-1}u) \\ y := S^{-1}u &\rightarrow y' = Dy \end{aligned}$$

Fundamentalsystem:

$$e^{\lambda_1 t} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + e^{\lambda_n t} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Beispiel

Sei

$$J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

Mit

$$y_1 := \begin{pmatrix} e^{\lambda t} \\ 0 \\ 0 \end{pmatrix} \quad y_2 := \begin{pmatrix} te^{\lambda t} \\ e^{\lambda t} \\ 0 \end{pmatrix} \quad y_3 := \begin{pmatrix} \frac{1}{2}t^2 e^{\lambda t} \\ te^{\lambda t} \\ e^{\lambda t} \end{pmatrix}$$

gilt dann

$$\begin{aligned} y_1' &= Jy_1 \\ y_2' &= \begin{pmatrix} e^{\lambda t} + \lambda te^{\lambda t} \\ \lambda e^{\lambda t} \\ 0 \end{pmatrix} = Jy_2 \\ y_3' &= \begin{pmatrix} te^{\lambda t} + \frac{1}{2}\lambda t^2 e^{\lambda t} \\ e^{\lambda t} + \lambda te^{\lambda t} \\ \lambda e^{\lambda t} \end{pmatrix} = Jy_3 \end{aligned}$$

d.h. man erhält ein Fundamentalsystem.

Der Zusammenhang wird noch deutlicher, wenn man die Fundamentallösungen in einer Matrix zusammenfasst:

- diagonalisierbar: $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \rightarrow \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix}$
- Jordanmatrix: $J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_s \end{pmatrix} \rightarrow \begin{pmatrix} L_1 & & 0 \\ & \ddots & \\ 0 & & L_s \end{pmatrix}$

mit

$$L_i = \begin{pmatrix} e^{\lambda_i t} & te^{\lambda_i t} & \frac{1}{2!}t^2 e^{\lambda_i t} & \dots & \frac{1}{(n_i-1)!}t^{n_i-1} e^{\lambda_i t} \\ 0 & e^{\lambda_i t} & te^{\lambda_i t} & \ddots & \vdots \\ \vdots & \ddots & e^{\lambda_i t} & \ddots & \frac{1}{2!}t^2 e^{\lambda_i t} \\ \vdots & & \ddots & \ddots & te^{\lambda_i t} \\ 0 & \dots & \dots & 0 & e^{\lambda_i t} \end{pmatrix}$$

Kapitel 8

Euklidische und unitäre Räume

Wir haben bereits gesehen, dass Determinanten zur Volumenberechnung von Parallelotopen führen. Wir befassen uns jetzt auch mit Winkeln und Längen.

8.1 Einführung: Kanonische Skalarprodukte im \mathbb{R}^n und \mathbb{C}^n

Bezeichnung

Die Abbildung

$$\begin{aligned} \langle, \rangle_R : \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \langle x, y \rangle_R := \sum_{i=1}^n x_i y_i \end{aligned}$$

für alle $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ heißt (KANONISCHES) SKALARPRODUKT DES \mathbb{R}^n .

8.1.1 Hinweis

Seien $x, x_1, x_2, y, y_1, y_2 \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. Dann gilt:

$$\begin{aligned} (i) \quad &\langle x, y \rangle_R = x^T y = y^T x \\ (B1) \quad &\langle x_1 + x_2, y \rangle_R = \langle x_1, y \rangle_R + \langle x_2, y \rangle_R \\ &\langle x, y_1 + y_2 \rangle_R = \langle x, y_1 \rangle_R + \langle x, y_2 \rangle_R \\ (B2) \quad &\langle \lambda x, y \rangle_R = \lambda \langle x, y \rangle_R \\ &\langle x, \lambda y \rangle_R = \lambda \langle x, y \rangle_R \\ (S) \quad &\langle x, y \rangle_R = \langle y, x \rangle_R \\ (PD) \quad &\langle x, x \rangle_R \geq 0 \\ &\langle x, x \rangle_R = 0 \Leftrightarrow x = 0 \end{aligned}$$

Beweis

trivial

Bemerkung

(i), (B1), (B2), (S) kann man über beliebigen Vektorräumen definieren. (PD) setzt eine *Ordnungsrelation* \geq in K voraus.

Bezeichnung

Die Abbildung $\|\cdot\|_{(2)} : \mathbb{R}^n \rightarrow \mathbb{R}$, definiert durch $\|x\|_{(2)} = \sqrt{\langle x, x \rangle_R}$ für $x \in \mathbb{R}^n$ heißt EUKLIDISCHE NORM des \mathbb{R}^n .

Beispiel

$$x \in \mathbb{R}^1 \Rightarrow \|x\|_{(2)} = \sqrt{x^2} = |x|$$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \Rightarrow \|x\|_{(2)} = \sqrt{x_1^2 + x_2^2}$$

Nach dem Satz des Phytagoras ist $\|x\|_{(2)}$ die Länge des Vektors x . Induktiv zeigt man, dass allgemeiner für $x \in \mathbb{R}^n$ $\|x\|_{(2)}$ der Abstand des Punktes x zum Nullpunkt ist.

Bezeichnung

Die Abbildung $d_{(2)} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, definiert durch $d_{(2)}(x, y) = \|x - y\|_{(2)}$ für alle $x, y \in \mathbb{R}^n$ heißt EUKLIDISCHER ABSTAND (EUKLIDISCHE DISTANZ, EUKLIDISCHE METRIK).

8.1.2 Hinweis

Seien $x, y, z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. Dann gilt:

$$(N1) \quad \|x\|_{(2)} \geq 0 \wedge \|x\|_{(2)} = 0 \Leftrightarrow x = 0$$

$$(N2) \quad \|\lambda x\|_{(2)} = |\lambda| \|x\|_{(2)}$$

$$(N3) \quad \|x - y\|_{(2)} \leq \|x\|_{(2)} + \|y\|_{(2)}$$

$$(D1) \quad d_{(2)}(x, y) \geq 0 \wedge d_{(2)}(x, y) = 0 \Leftrightarrow x = y$$

$$(D2) \quad d_{(2)}(x, y) = d_{(2)}(y, x)$$

$$(D3) \quad d_{(2)}(x, z) \leq d_{(2)}(x, y) + d_{(2)}(y, z)$$

Beweis

(N1), (N2), (D1), (D2) sind trivial. (D3) folgt aus (N3). $(d_{(2)}(x, z) = \|x - z\|_{(2)} = \|x - y + y - z\|_{(2)} \leq \|x - y\|_{(2)} + \|y - z\|_{(2)})$.

Wegen $\|x + y\|_{(2)}^2 = \langle x + y, x + y \rangle_R = \langle x, x \rangle_R + 2\langle x, y \rangle_R + \langle y, y \rangle_R = \|x\|_{(2)}^2 + 2\langle x, y \rangle_R + \|y\|_{(2)}^2 = (\|x\|_{(2)} + \|y\|_{(2)})^2 + 2(\langle x, y \rangle_R - \|x\|_{(2)}\|y\|_{(2)})$ folgt (N3) aus folgender Ungleichung.

8.1.3 Lemma (Cauchy-Schwarzsche Ungleichung)

Seien $x, y \in \mathbb{R}^n$. Dann gilt $|\langle x, y \rangle_R| \leq \|x\|_{(2)}\|y\|_{(2)}$ mit Gleichheit genau dann wenn x und y linear abhängig sind.

Beweis

Für $n = 1$ ist die Aussage trivial. Sei also $n \geq 2$. Es gilt $\|x\|_{(2)}^2 \|y\|_{(2)} - \langle x, y \rangle_R^2 =$

$$(x^T x)(y^T y) - (x^T y)(y^T x) = \begin{vmatrix} x^T x & x^T y \\ y^T x & y^T y \end{vmatrix} = |A^T A| \text{ mit } A = (x, y) \in \mathbb{R}^{n \times 2}.$$

Mittels vollständiger Induktion nach n folgt

$$\det(A^T A) = \sum_{\substack{i, j=1, \dots, n \\ i < j}} \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix}^2 \left(x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right)$$

Also gilt $\|x\|_{(2)}^2 \|y\|_{(2)}^2 - \langle x, y \rangle_R^2 = \sum_{\substack{i, j=1, \dots, n \\ i < j}} \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix}^2 \geq 0$ mit Gleichheit genau

dann, wenn $\begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} = 0$ für alle $i, j = 1, \dots, n$ und $i < j$ und es folgt die Behauptung.

8.1.4 Korollar

Seien $x, y \in \mathbb{R}^n \setminus \{0\}$. Dann gilt

$$-1 \leq \frac{\langle x, y \rangle_R}{\|x\|_{(2)} \|y\|_{(2)}} \leq 1$$

Beweis

folgt aus (8.1.3)

Interpretation

$x, y \in \mathbb{R}^n \setminus \{0\}$

$$\begin{aligned} \frac{\langle x, y \rangle_R}{\|x\|_{(2)} \|y\|_{(2)}} &= \left\langle \frac{x}{\|x\|_{(2)}}, \frac{y}{\|y\|_{(2)}} \right\rangle = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}^T \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix} \\ &= \cos \alpha \cos \beta + \sin \alpha \sin \beta = \cos(\beta - \alpha) \end{aligned}$$

Insbesondere ist also $\langle x, y \rangle_R \geq 0$ genau dann, wenn x und y aufeinander *senkrecht* stehen.

Bezeichnet $\angle(x, y)$ den (zwischen 0 und π liegenden unorientierten) Winkel zwischen x und y , so ist also $\langle x, y \rangle_R = \|x\|_{(2)} \|y\|_{(2)} \cos(\angle(x, y))$, d.h. Produkt von $\|x\|_{(2)}$ und der Länge der senkrechten (*orthogonalen*) Projektion von y auf $\text{span}(\{x\})$.

Wir betrachten jetzt den komplexen Fall. Da \mathbb{C} nicht angeordnet ist, kann man (PD) nur erwarten, wenn $\langle z, z \rangle_R \in \mathbb{R}$ für alle $z \in \mathbb{C}$ ist. (B2) ergäbe sich dann aber zu $\langle iz, iz \rangle_R = i^2 \langle z, z \rangle_R = -\langle z, z \rangle_R$, d.h. $\langle z, z \rangle_R \geq 0$ und $\langle iz, iz \rangle_R = -\langle z, z \rangle_R \geq 0$, also $\langle z, z \rangle_R = 0$.

Wenn man also (B1), (B2), (S) und (PD) fordert, so erhält man nur die triviale Funktion 0. Für $z = a + ib$ mit $a, b \in \mathbb{R}$ gilt $|z|^2 = a^2 + b^2 = z\bar{z}$. Wir definieren daher:

Bezeichnung

Die Abbildung $\langle \cdot, \cdot \rangle_C : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$, definiert durch $\langle z, w \rangle_C = \sum_{i=1}^n z_i \bar{w}_i$ für alle

$$z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \text{ und } w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \text{ heißt (KANONISCHES) SKALARPRODUKT des } \mathbb{C}^n.$$

8.1.5 Hinweis

Seien $z, z_1, z_2, w, w_1, w_2 \in \mathbb{C}^n$ und $\lambda \in \mathbb{C}$. Dann gilt:

$$\begin{aligned} (B1) \quad & \langle z_1 + z_2, w \rangle_C = \langle z_1, w \rangle_C + \langle z_2, w \rangle_C \\ & \langle z, w_1 + w_2 \rangle_C = \langle z, w_1 \rangle_C + \langle z, w_2 \rangle_C \\ (\overline{B2}) \quad & \langle \lambda z, w \rangle_C = \lambda \langle z, w \rangle_C \\ & \langle z, \lambda w \rangle_C = \overline{\lambda} \langle z, w \rangle_C \\ (\overline{S}) \quad & \langle z, w \rangle_C = \overline{\langle w, z \rangle_C} \\ (PD) \quad & \langle z, z \rangle_C \in \mathbb{R}, \langle z, z \rangle_C \geq 0 \\ & \langle z, z \rangle_C = 0 \Rightarrow z = 0 \end{aligned}$$

Beweis

Wegen (PD) kann man mittels $\|z\| = \sqrt{\langle z, z \rangle_C}$ und durch $d(z, w) = \|z - w\|$ NORM und METRIK definieren.

8.2 Skalarprodukte und hermitesche Formen

Definition

Seien V ein \mathbb{R} -Vektorraum und $F : V \times V \rightarrow \mathbb{R}$ eine Bilinearform. F heißt SKALAR-PRODUKT, wenn gilt:

$$\begin{aligned} (S) \quad & x, y \in V \Rightarrow F(x, y) = F(y, x) \quad (\text{symmetrisch}) \\ (PD) \quad & x \in V \setminus \{0\} \Rightarrow F(x, x) > 0 \quad (\text{positiv definit}) \end{aligned}$$

Hinweis

Ist F ein Skalarprodukt, so gilt natürlich $F(0, 0) = 0$, also $x \in V \Rightarrow F(x, x) \geq 0$.

Beispiele

(a) Seien

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

und $F : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch $F(x, y) = x^T A y$. Dann ist F bilinear. Ferner gilt

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

d.h. F ist symmetrisch. Mit $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ folgt ferner

$$F(x, x) = (x_1, x_2) \begin{pmatrix} 4 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (2x_1 - x_2)^2 + 2x_2^2 \geq 0$$

Aus $F(x, x) = 0$ folgt $2x_1 - x_2 = x_2 = 0$, d.h. $x = 0$. F ist somit positiv definit.

(b) Seien V der Vektorraum der auf $[-\pi, \pi]$ stetigen reellen Funktionen, $h \in V$ mit $h(t) > 0 \forall t \in [-\pi, \pi]$ und $F : V \times V \rightarrow \mathbb{R}$ sei definiert durch $F(f, g) = \int_{-\pi}^{\pi} h(t) f(t) g(t) dt$ für $f, g \in V$. F ist eine symmetrische Bilinearform. Ferner gilt $F(f, f) = \int_{-\pi}^{\pi} h(t) f^2(t) dt \geq 0$; wegen der Stetigkeit von f gilt $F(f, f) = 0$ nur für $f = 0$.

Definition

Seien V ein \mathbb{C} -Vektorraum, $F : V \times V \rightarrow \mathbb{C}$. F heißt SESQUILINEAR, wenn gilt:

$$(B1) \quad z, z_1, z_2, w, w_1, w_2 \in V \Rightarrow F(z_1 + z_2, w) = F(z_1, w) + F(z_2, w) \wedge \\ \wedge F(z, w_1 + w_2) = F(z, w_1) + F(z, w_2)$$

$$(\overline{B2}) \quad z, w \in V \wedge \lambda \in \mathbb{C} \Rightarrow F(\lambda z, w) = \lambda F(z, w) \wedge F(z, \lambda w) = \overline{\lambda} F(z, w)$$

F heißt HERMITESCH oder HERMITESISCHE FORM, wenn zusätzlich gilt:

$$(\overline{S}) \quad z, w \in V \Rightarrow F(z, w) = \overline{F(w, z)}$$

Gilt ferner

$$(PD) \quad z \in V \setminus \{0\} \Rightarrow F(z, z) > 0$$

so heißt F SKALARPRODUKT.

Definition

Seien $K \in \{\mathbb{R}, \mathbb{C}\}$, V ein K -Vektorraum und F ein Skalarprodukt. Dann heißt (V, F) (oder kürzer V) EUKLIDISCHER RAUM, falls $K = \mathbb{R}$ und UNITÄRER RAUM falls $K = \mathbb{C}$.

Bezeichnung

Skalarprodukte werden im Folgenden (meistens) mit $\langle \cdot, \cdot \rangle$ bezeichnet.

Wir zeigen nun, in welcher Beziehung euklidische und unitäre Räume stehen.

Definition

Seien V ein \mathbb{R} -Vektorraum, $Z = Z(V) := \{(x, y) : x, y \in V\}$. $+$: $Z \times Z \rightarrow Z$ und \cdot : $\mathbb{C} \times Z \rightarrow Z$ seien definiert durch

$$z_1 = (x_1, y_1) \in Z \wedge z_2 = (x_2, y_2) \in Z \Rightarrow z_1 + z_2 := (x_1 + x_2, y_1 + y_2)$$

$$\lambda = \lambda_1 + \lambda_2 i \wedge \lambda_1, \lambda_2 \in \mathbb{R} \wedge z = (x, y) \in Z \Rightarrow \lambda z := (\lambda_1 x - \lambda_2 y, \lambda_1 y + \lambda_2 x)$$

Dann heißt $(Z, +, \cdot)$ (oder kürzer Z) KOMPLEXE ERWEITERUNG von V .

8.2.1 Hinweis

Sei V ein \mathbb{R} -Vektorraum. Dann ist $Z(V)$ ein \mathbb{C} -Vektorraum.

Beweis

nachrechnen!

V kann in $Z(V)$ in folgendem Sinn eingebettet werden:

Bezeichnung

Seien V ein \mathbb{R} -Vektorraum und $\psi : V \rightarrow Z(V)$ definiert durch $\psi(v) = (v, 0)$ für $v \in V$. Dann heißt ψ KANONISCHE EINBETTUNG von V in $Z(V)$.

8.2.2 Hinweis

Seien V ein \mathbb{R} -Vektorraum, Z seine komplexe Erweiterung und ψ die kanonische Einbettung. Dann ist ψ injektiv und es gilt für alle $x, y \in V$ und $\lambda, \mu \in \mathbb{R}$: $\psi(\lambda x + \mu y) = \lambda \psi(x) + \mu \psi(y)$.

Beweis

trivial

Bemerkung

Wir identifizieren ein Paar $(x, 0) \in Z$ mit dem zugehörigen Vektor $x \in V$. Dann gilt wegen $i(y, 0) = (0, y)$ stets $(x, y) = x + iy$.

8.2.3 Satz

Seien V_1, V_2 \mathbb{R} -Vektorräume, $F : V_1 \rightarrow V_2$ linear und $Z_1 = Z(V_1)$, $Z_2 = Z(V_2)$. Dann gibt es *genau eine* lineare Abbildung $G = G(F) : Z_1 \rightarrow Z_2$ mit $G(x) = F(x) \forall x \in V_1$.

Sprechweise

F kann auf genau eine Weise zu einer linearen Abbildung $G : Z_1 \rightarrow Z_2$ fortgesetzt werden.

Beweis

Sei zunächst $G : Z_1 \rightarrow Z_2$ eine solche Fortsetzung. Dann gilt für jeden Vektor $z = x + iy$ mit $x, y \in V$ $G(z) = G(x + iy) = G(x) + iG(y) = F(x) + iF(y)$, d.h. G ist eindeutig durch F bestimmt.

Andererseits wird durch $G(z) = F(x) + iF(y)$ für $z = x + iy$ eine solche Fortsetzung definiert.

8.2.4 Satz

Sei (V, F) ein euklidischer Raum. Dann gibt es *genau ein* Skalarprodukt G auf $Z(V)$ mit $G(x, y) = F(x, y)$ für alle $x, y \in V$ (\gg FORTSETZUNG \ll).

Beweis

Sei G eine solche Fortsetzung. Mit $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$, $x_1, x_2, y_1, y_2 \in V$ gilt dann

$$\begin{aligned} G(z_1, z_2) &= G(x_1 + iy_1, x_2 + iy_2) \\ &= G(x_1, x_2) + G(y_1, y_2) + i(G(y_1, x_2) - G(x_1, y_2)) \\ &= F(x_1, x_2) + F(y_1, y_2) + i(F(y_1, x_2) - F(x_1, y_2)) \end{aligned}$$

d.h. G ist durch F eindeutig bestimmt.

Andererseits definiert diese Beziehung ein Skalarprodukt in $Z(V)$.

Bemerkung

Satz (8.2.4) besagt, dass sich jeder euklidische Raum in einen unitären Raum einbetten lässt. Sätze über Skalarprodukte brauchen daher im Allgemeinen nur für unitäre Räume bewiesen werden und lassen sich dann auf den reellen Fall übertragen.

8.2.5 Bemerkung

Sei V ein endlich-dimensionaler \mathbb{C} -Vektorraum, $\mathcal{B} = \{b_1, \dots, b_n\}$ eine Basis von V und F eine Sesquilinearform. Ferner sei $A := M_{\mathcal{B}}(F) = (F(b_i, b_j))_{i,j=1,\dots,n}$. Dann gilt für alle $v, w \in V$ mit $x = \Phi^{-1}(v)$, $y = \Phi^{-1}(w)$

$$F(v, w) = x^T A \bar{y}$$

Beweis

Seien $x = (x_1, \dots, x_n)^T$, $y = (y_1, \dots, y_n)^T$. Dann gilt

$$\begin{aligned} F(v, w) &= F\left(\sum_{i=1}^n x_i b_i, \sum_{j=1}^n y_j b_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i \overline{y_j} F(b_i, b_j) \\ &= x^T A \overline{y} \end{aligned}$$

Hinweis

Reelle Matrizen und Bilinearformen bzw. komplexe Matrizen und Sesquilinearformen entsprechen einander. Es gilt $\overline{F(w, v)} = \overline{y^T A x} = (\overline{y^T A x})^T = x^T \overline{A}^T \overline{y}$. Ist F also hermitesch, so folgt $x^T A \overline{y} = x^T \overline{A}^T \overline{y}$ für alle $x, y \in C^n$, also $A = \overline{A}^T$.

Definition

Seien $K \in \{\mathbb{R}, \mathbb{C}\}$ und $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$. Für $K = \mathbb{R}$ ($K = \mathbb{C}$) heißt A SYMMETRISCH (HERMITESCH), wenn $A = A^T$ ($A = \overline{A^T}$) und POSITIV DEFINIT, wenn $x^T A x > 0$ ($x^T A \overline{x} > 0$) für alle $x \in K^n \setminus \{0\}$. $\overline{A} = (\overline{a_{ij}})_{i,j=1,\dots,n}$ heißt die zu A KONJUGIERT KOMPLEXE MATRIX.

8.2.6 Hinweis (Basiswechsel)

Seien $K \in \{\mathbb{R}, \mathbb{C}\}$, V ein n -dimensionaler K -Vektorraum und F eine Sesquilinearform. Sind \mathcal{B}_1 und \mathcal{B}_2 Basen von V , $M_1 = M_{\mathcal{B}_1}(F)$, $M_2 = M_{\mathcal{B}_2}(F)$ und S die zu $\Phi_{\mathcal{B}_1}^{-1} \Phi_{\mathcal{B}_2}$ gehörige Matrix, so gilt $M_2 = S^T M_1 S$ für $K = \mathbb{R}$ und $M_2 = S^T M_1 \overline{S}$ für $K = \mathbb{C}$.

Beweis

Seien $v, w \in V$ und für $i = 1, 2$ seien $x_i = \Phi_{\mathcal{B}_1}^{-1}(v)$ und $y_i = \Phi_{\mathcal{B}_1}^{-1}(w)$. Nach (8.2.5) gilt dann $F(v, w) = x_1^T M_1 \overline{y_1} = x_2^T M_2 \overline{y_2}$.

Nun gilt andererseits $x_1 = \Phi_{\mathcal{B}_1}^{-1} \Phi_{\mathcal{B}_2}(x_2)$ und $y_1 = \Phi_{\mathcal{B}_1}^{-1} \Phi_{\mathcal{B}_2}(y_2)$. Somit folgt

$$\begin{aligned} x_2^T M_2 \overline{y_2} &= x_1^T M_1 \overline{y_1} \\ &= (S x_2)^T M_1 \overline{(S y_2)} \\ &= x_2^T (S^T M_1 \overline{S}) \overline{y_2} \end{aligned}$$

Da dieses für alle $v, w \in V$ gilt, folgt $M_2 = S^T M_1 \overline{S}$.

8.3 Orthogonalität und Orthonormalität**Definition**

Sei V ein euklidischer oder unitärer Vektorraum. Dann heißt $\|\cdot\| : V \rightarrow \mathbb{R}$, definiert durch $\|v\| := \sqrt{\langle v, v \rangle}$ für $v \in V$ die (\langle, \rangle) zugeordnete) NORM.

Die Abbildung $d : V \times V \rightarrow \mathbb{R}$, definiert durch $d(v, w) = \|v - w\|$ für $v, w \in V$ heißt (von $\|\cdot\|$) INDUZIERT METRIK.

8.3.1 Satz (Cauchy-Schwarzsche Ungleichung)

Sei V ein euklidischer oder unitärer Vektorraum. Dann gilt für alle $v, w \in V$: $|\langle v, w \rangle| \leq \|v\| \|w\|$ mit Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis

Für $w = 0$ gilt $\langle v, w \rangle = 0 = \|w\|$; wir können also im Folgenden $w \neq 0$ voraussetzen. Dann gilt auch $\|w\| > 0$ und wir können $\lambda = \frac{\langle v, w \rangle}{\langle w, w \rangle}$ setzen. Es folgt

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w \rangle \\ &= \langle v, v \rangle - \lambda \overline{\langle v, w \rangle} - \bar{\lambda} \langle v, w \rangle + \lambda \bar{\lambda} \langle w, w \rangle \\ &= \langle v, v \rangle - \frac{\langle v, w \rangle \overline{\langle v, w \rangle}}{\langle w, w \rangle} - \frac{\langle v, w \rangle \langle v, w \rangle}{\langle w, w \rangle} + \frac{\langle v, w \rangle \overline{\langle v, w \rangle}}{\langle w, w \rangle} \end{aligned}$$

also nach Multiplikation mit $\langle w, w \rangle$:

$$0 \leq \langle v, v \rangle \langle w, w \rangle - |\langle v, w \rangle|^2 = \|v\|^2 \|w\|^2 - |\langle v, w \rangle|^2$$

Da \langle, \rangle positiv definit ist, gilt das Gleichheitszeichen genau dann, wenn $v - \lambda w = 0$ gilt, und es folgt die Behauptung.

8.3.2 Korollar

Sei V ein euklidischer oder unitärer Vektorraum. Dann gelten die Eigenschaften (N1) - (N3) und (D1) - (D3) aus (8.1.2), d.h. für $u, v, w \in V$ und $\lambda \in K$:

- $\|v\| \geq 0$, $\|v\| = 0 \Leftrightarrow v = 0$
- $\|\lambda v\| = |\lambda| \|v\|$
- $\|v + w\| \leq \|v\| + \|w\|$
- $d(v, w) \geq 0$, $d(v, w) = 0 \Leftrightarrow v = w$
- $d(v, w) = d(w, v)$
- $d(u, w) \leq d(u, v) + d(v, w)$

Hinweis

Mit Hilfe der Eigenschaften aus (8.3.2) kann man allgemein Normen und Metriken definieren (\rightarrow Funktionalanalysis). In euklidischen Räumen kann man wie in (8.1) mittels (8.3.1) Winkel erklären.

Wir betrachten im Folgenden den Spezialfall des »senkrecht stehens«.

Definition

Sei V ein euklidischer oder unitärer Vektorraum.

- (i) Sei $v \in V \setminus \{0\}$. v heißt **NORMIERT**, wenn $\|v\| = 1$ ist.
- (ii) Seien $v, w \in V$. v und w heißen **ORTHOGONAL**, wenn $\langle v, w \rangle = 0$ gilt.
Schreibweise: $v \perp w$.
- (iii) Seien U, W Untervektorräume von V . U und W heißen **ORTHOGONAL**, wenn gilt: $u \in U \wedge v \in W \Rightarrow u \perp v$.
Schreibweise: $U \perp W$.
Sprechweise: \perp : »ist orthogonal zu«, »steht senkrecht auf«.

- (iv) Sei $M \subset V$ nicht leer. Dann heißt $M^\perp := \{v : v \in V \wedge x \in M \Rightarrow v \perp x\}$ ORTHOGONALES KOMPLEMENT von M .
- (v) Sei $M \subset V$. M heißt ORTHOGONALSYSTEM, wenn gilt $0 \notin M \neq \emptyset$ und $x, y \in M \Rightarrow x \perp y$.
 M heißt ORTHONORMALSYSTEM, wenn M ein Orthogonalsystem ist und jeder Vektor aus M normiert ist.
 M heißt ORTHONORMALBASIS, wenn M Basis von V und Orthonormalsystem ist.
- (vi) Seien U_1, \dots, U_k Untervektorräume von V mit $V = U_1 \oplus \dots \oplus U_k$. Gilt $U_i \perp U_j$ für $i, j = 1, \dots, k, i \neq j$, so heißt die direkte Summe ORTHOGONAL. Schreibweise¹: $V = U_1 \boxplus \dots \boxplus U_k$

Bemerkung

- (1) Ist B Orthonormalbasis von V , so gilt für $x, y \in B$:

$$\langle x, y \rangle = \begin{cases} 0 & \text{für } x \neq y \\ 1 & \text{für } x = y \end{cases}$$

- (2) Ist $\emptyset \neq M \subset V$, so ist M^\perp ein Untervektorraum von V .

- (3) Ist M ein Orthogonalsystem, so ist $\hat{M} = \left\{ \frac{x}{\|x\|} : x \in M \right\}$ ein Orthonormalsystem.

8.3.3 Bemerkung

Seien V ein euklidischer oder unitärer Raum und M ein Orthogonalsystem. Dann ist M linear unabhängig.

Beweis

Seien $v_1, \dots, v_k \in M$, $\lambda_1, \dots, \lambda_k \in K$ mit $\sum_{i=1}^k \lambda_i v_i = 0$. Dann gilt für $j \in \{1, \dots, k\}$:

$$0 = \left\langle \sum_{i=1}^n \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^n \lambda_i \langle v_i, v_j \rangle = \lambda_j \|v_j\|^2$$

also $\lambda_j = 0$. Da j beliebig gewählt war, folgt die Behauptung.

8.3.4 Bemerkung

Seien V ein euklidischer oder unitärer Vektorraum, $\{b_1, \dots, b_n\}$ eine Orthogonalbasis und $v, w \in V$, $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in K$ mit $v = \sum_{i=1}^n \lambda_i b_i$ und $w = \sum_{i=1}^n \mu_i b_i$. Dann gilt $\langle v, w \rangle = \sum_{i=1}^n \lambda_i \mu_i$ und speziell $\lambda_i = \langle v, b_i \rangle$ für $i = 1, \dots, n$.

¹Das Symbol \boxplus , das in diesem Script für die direkte orthogonale Summe verwendet wird, entspricht nicht dem in der Vorlesung verwendeten (ein eingekreistes \perp). Dieses Symbol scheint es in \LaTeX nicht zu geben.

Beweis

$$\begin{aligned}
\langle v, w \rangle &= \left\langle \sum_{i=1}^n \lambda_i b_i, \sum_{j=1}^n \mu_j b_j \right\rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \overline{\mu_j} \langle b_i, b_j \rangle \\
&= \sum_{i=1}^n \lambda_i \overline{\mu_i}
\end{aligned}$$

8.3.5 Hinweis

Seien $K \in \{\mathbb{R}, \mathbb{C}\}$, V ein K -Vektorraum, $B = \{b_1, \dots, b_n\}$ eine Basis. Dann ist durch

$$\langle v, w \rangle = (\Phi_B^{-1}(v))^T \overline{\Phi_B^{-1}(w)}$$

ein Skalarprodukt definiert, bezüglich dessen B eine Orthonormalbasis ist.

Beispiel

Sei V der Vektorraum aller in $[-\pi, \pi]$ stetigen reellen Funktionen und \langle, \rangle das durch $\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)g(t) dt$ für $f, g \in V$ definierte Skalarprodukt. Sei M die Menge der Funktionen $\frac{1}{\sqrt{2}}, \cos(nt), \sin(nt), n \in \mathbb{N}$. Dann ist M ein Orthonormalsystem.

Im Folgenden geben wir ein Verfahren an, wie man aus einer gegebenen Basis eine Orthonormalbasis konstruieren kann.

Definition

Eine Menge heißt ABZÄHLBAR UNENDLICH, wenn es eine bijektive Abbildung $\varphi : M \rightarrow \mathbb{N}$ gibt. Ein K -Vektorraum heißt VON ABZÄHLBAR UNENDLICHER DIMENSION, wenn er eine abzählbar unendliche Basis besitzt.

Beispiel

$K[X]$ hat abzählbar unendliche Dimension, $\text{Abb}(\mathbb{R}, \mathbb{R})$ besitzt kein abzählbares Erzeugendensystem.

8.3.6 Satz (Gram-Schmidtsches Orthogonalisierungsverfahren)

Seien V ein euklidischer oder unitärer Vektorraum, $N \subset \mathbb{N}$ endlich oder $N = \mathbb{N}$ und $\mathcal{A} = \{a_i : i \in N\}$ eine linear unabhängige Teilmenge von V . Dann gibt es genau ein Orthonormalsystem $\mathcal{B} = \{b_i : i \in N\}$, so dass mit $A_k = \{a_i : i = 1, \dots, k\}$ und $B_k = \{b_i : i = 1, \dots, k\}$ für $k \in N$ gilt:

- (i) $\text{span}(A_k) = \text{span}(B_k)$
- (ii) Sei S_k die Transformationsmatrix des Basiswechsels $\Phi_{A_k}^{-1} \circ \Phi_{B_k}$. Dann ist $\det(S_k) \in]0, \infty[$.

Beweis

Die Vektoren werden induktiv definiert.

Da \mathcal{A} linear unabhängig ist, gilt $a_1 \neq 0$ und wir setzen $b_1 = \frac{1}{\|a_1\|} a_1$. Natürlich gilt $\text{span}(A_1) = \text{span}(B_1)$ und es gilt $\det(S_1) = \frac{1}{\|a_1\|} \in]0, \infty[$. Ist andererseits \hat{b}_1 ein anderer Vektor mit (i) und (ii), so folgt wegen (i) $\hat{b}_1 = \lambda a_1$ mit einem $\lambda \in K$; wegen $\|\hat{b}_1\| = 1$ gilt $|\lambda| = \frac{1}{\|a_1\|}$. Wegen $\lambda = \det(\hat{S}_1) \in]0, \infty[$ folgt $\lambda = \frac{1}{\|a_1\|}$, also gilt $\hat{b}_1 = b_1$.

Wir nehmen nun an, dass b_1, \dots, b_n bereits konstruiert sind, so dass (i) und (ii) gelten. Sei $c_{n+1} = a_{n+1} - \sum_{i=1}^n \langle a_{n+1}, b_i \rangle b_i$. Dann folgt mit Hilfe der Induktionsvoraussetzung $\text{span}(A_{n+1}) = \text{span}(B_n \cup \{a_{n+1}\}) = \text{span}(B_n \cup \{c_{n+1}\})$. Insbesondere ist $B_n \cup \{c_{n+1}\}$ linear unabhängig und wir können $b_{n+1} = \frac{1}{\|c_{n+1}\|} c_{n+1}$ setzen. Es gilt für $i = 1, \dots, n$

$$\begin{aligned} \langle b_{n+1}, b_j \rangle &= \frac{1}{\|c_{n+1}\|} \left\langle a_{n+1} - \sum_{i=1}^n \langle a_{n+1}, b_i \rangle b_i, b_j \right\rangle \\ &= \frac{1}{\|c_{n+1}\|} (\langle a_{n+1}, b_j \rangle - \langle a_{n+1}, b_j \rangle) = 0 \end{aligned}$$

B_{n+1} erfüllt also (i).

$S_{n+1} = (s_{ij})_{i,j=1,\dots,n+1}$ ist die durch folgende Gleichungen bestimmte Dreiecksmatrix:

$$\begin{aligned} b_1 &= s_{11} a_1 \\ b_2 &= s_{12} a_1 + s_{22} a_2 \\ &\vdots \\ b_n &= s_{1n} a_1 + \dots + s_{nn} a_n \\ b_{n+1} &= s_{1,n+1} a_1 + \dots + s_{n,n+1} a_n + \frac{1}{\|c_{n+1}\|} a_{n+1} \end{aligned}$$

Es folgt $\det(S_{n+1}) = \det(S_n) \frac{1}{\|c_{n+1}\|} \in]0, \infty[$. Ist andererseits \hat{b}_{n+1} ein anderer Vektor, für den (i) und (ii) erfüllt sind, so folgt $\hat{b}_{n+1} = \sum_{i=1}^n \lambda_i b_i + \lambda_{n+1} b_{n+1}$ mit $\lambda_{n+1} \in]0, \infty[$. Wegen $\langle \hat{b}_{n+1}, b_i \rangle = 0$ für $i = 1, \dots, n$ folgt $\hat{b}_{n+1} = \lambda_{n+1} b_{n+1}$ und $\|\hat{b}_{n+1}\| = 1 = |\lambda_{n+1}|$.

8.3.7 Korollar

Seien V ein euklidischer oder unitärer Vektorraum endlicher oder abzählbar unendlicher Dimension und $\{b_1, \dots, b_n\} \subset V$ Orthogonalsystem. Dann gibt es eine Orthonormalbasis von V , die $\{b_1, \dots, b_n\}$ enthält.

Beweis

Folgt aus (4.2.5) und (8.3.6).

Beispiele

(a) $(\mathbb{R}^4, \langle \cdot, \cdot \rangle_R)$

$$a_1 = \begin{pmatrix} 4 \\ 2 \\ -2 \\ -1 \end{pmatrix} \quad a_2 = \begin{pmatrix} 2 \\ 2 \\ -4 \\ -5 \end{pmatrix} \quad a_3 = \begin{pmatrix} 0 \\ 8 \\ -2 \\ -5 \end{pmatrix}$$

$$\begin{aligned}
b_1 &= \frac{1}{\|a_1\|} a_1 = \frac{1}{5} \begin{pmatrix} 4 \\ 2 \\ -2 \\ -1 \end{pmatrix} \\
c_2 &= a_2 - \langle a_2, b_1 \rangle_R b_1 = \begin{pmatrix} 2 \\ 2 \\ -4 \\ -5 \end{pmatrix} - 25 \cdot \frac{1}{5} \cdot \frac{1}{5} \begin{pmatrix} 4 \\ 2 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ -2 \\ -4 \end{pmatrix} \\
b_2 &= \frac{1}{\|c_2\|} c_2 = \frac{1}{\sqrt{24}} c_2 \\
c_3 &= a_3 - \langle a_3, b_1 \rangle_R b_1 - \langle a_3, b_2 \rangle_R b_2 \\
&= \begin{pmatrix} 0 \\ 8 \\ -2 \\ -5 \end{pmatrix} - 25 \cdot \frac{1}{5} \cdot \frac{1}{5} \begin{pmatrix} 4 \\ 2 \\ -2 \\ -1 \end{pmatrix} - 24 \cdot \frac{1}{i\sqrt{24}} \cdot \frac{1}{\sqrt{24}} \begin{pmatrix} -2 \\ 0 \\ -2 \\ -4 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \\ 2 \\ 0 \end{pmatrix} \\
b_3 &= \frac{1}{\|c_3\|} c_3 = \frac{1}{\sqrt{44}} \begin{pmatrix} -2 \\ 6 \\ 2 \\ 0 \end{pmatrix}
\end{aligned}$$

- (b) $K[X]_2$, $\langle p, q \rangle := \int_0^1 p(x)q(x) dx$ für $p, q \in K[X]_2$ und für $i = 1, 2, 3$ seien $a_i \in K[X]_2$ mit $a_i(x) = x^{i-1}$.

$$\begin{aligned}
\|a_1\|^2 &= \int_0^1 1 \cdot 1 dx = 1 \Rightarrow b_1 = a_1 \\
c_2 &= a_2 - \langle a_2, b_1 \rangle b_1, \text{ d.h. } c_2(x) = x^1 - \left(\int_0^1 x \cdot 1 dx \right) x^0 = x - \frac{1}{2} \\
\|c_2\|^2 &= \int_0^1 \left(x - \frac{1}{2} \right)^2 dx = \frac{1}{12} \\
b_2(x) &= 2\sqrt{3} \left(x - \frac{1}{2} \right) \\
c_3 &= a_3 - \langle a_3, b_1 \rangle b_1 - \langle a_3, b_2 \rangle b_2 = x^2 - x^1 + \frac{1}{6} \\
\|c_3\|^2 &= \frac{1}{180} \\
b_3(x) &= 6\sqrt{5} \left(x^2 - x + \frac{1}{6} \right)
\end{aligned}$$

8.3.8 Korollar

Seien V ein n -dimensionaler euklidischer oder unitärer Vektorraum und U ein Untervektorraum. Dann gilt:

- (i) $\dim U^\perp = n - \dim U$
- (ii) $(U^\perp)^\perp = U$
- (iii) U^\perp ist Komplement von U und für $\{0\} \neq U \neq V$ gilt $V = U \boxplus U^\perp$

Beweis

Nach (8.3.6) besitzt U eine Orthonormalbasis $B = \{v_1, \dots, v_r\}$. B lässt sich nach (8.3.7) zu einer Orthonormalbasis von V ergänzen. Es folgt $\text{span}\{v_{r+1}, \dots, v_n\} \subset$

U^\perp , d.h. $\dim U^\perp \geq n - r$. Ist aber $u \in (U \cap U^*)$, so folgt $\langle u, u \rangle = 0$, also $u = 0$. Nach (4.3.1) gilt $\dim U^* \leq n - \dim U$ und es folgt (i) und (iii).

Zu (ii): Für alle $u \in U$ und $v \in U^*$ gilt $\langle u, v \rangle = 0$, also $U \subset (U^\perp)^\perp$. Nach (i) gilt $\dim (U^\perp)^\perp = n - \dim U^\perp = n - (n - \dim U) = \dim U$ und es folgt die Behauptung.

Bemerkung

(8.3.8) lässt sich *nicht* auf unendlich-dimensionale Vektorräume verallgemeinern.

Bezeichnung

Seien V ein euklidischer Vektorraum, $m \leq \dim V < \infty$ und $v_1, \dots, v_m \in V$.

$$G(v_1, \dots, v_m) := \det \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_m \rangle \\ \vdots & & \vdots \\ \langle v_m, v_1 \rangle & \dots & \langle v_m, v_m \rangle \end{pmatrix}$$

heißt GRAMSCHE DETERMINANTE von v_1, \dots, v_m .

8.3.9 Lemma

Seien V ein n -dimensionaler euklidischer Vektorraum, $m \leq n$ und $v_1, \dots, v_m \in V$. Dann gilt $G(v_1, \dots, v_m) \geq 0$ und $G(v_1, \dots, v_m) = 0$ genau dann, wenn v_1, \dots, v_m linear abhängig sind.

Beweis

Seien $U = \text{span}\{v_1, \dots, v_m\}$, $A = (v_1, \dots, v_m)$.

Sei zunächst $m \leq n$. Nach (8.3.6) und (8.3.7) gibt es ein Orthonormalsystem $\{v_{m+1}, \dots, v_n\}$, so dass $\text{span}\{v_{m+1}, \dots, v_n\} \perp U$ ist. Dann gilt

$$\begin{aligned} G(v_1, \dots, v_n) &= \begin{vmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_m \rangle & \langle v_1, v_{m+1} \rangle & \dots & \langle v_1, v_n \rangle \\ \vdots & & \vdots & \vdots & & \vdots \\ \langle v_m, v_1 \rangle & \dots & \langle v_m, v_m \rangle & \langle v_m, v_{m+1} \rangle & \dots & \langle v_m, v_n \rangle \\ \langle v_{m+1}, v_1 \rangle & \dots & \langle v_{m+1}, v_m \rangle & \langle v_{m+1}, v_{m+1} \rangle & \dots & \langle v_{m+1}, v_n \rangle \\ \vdots & & \vdots & \vdots & & \vdots \\ \langle v_n, v_1 \rangle & \dots & \langle v_n, v_m \rangle & \langle v_n, v_{m+1} \rangle & \dots & \langle v_n, v_n \rangle \end{vmatrix} \\ &= \begin{vmatrix} A^T A & 0 \\ 0 & I \end{vmatrix} \\ &= G(v_1, \dots, v_m) \end{aligned}$$

Andererseits gilt mit $B = (v_1, \dots, v_n)$ $G(v_1, \dots, v_n) = \det(B^T B) = (\det B)^2 \geq 0$ und es folgt die erste Behauptung für alle $m \leq n$.

Ferner ist $G(v_1, \dots, v_n)$ genau dann 0, wenn v_1, \dots, v_n linear abhängig sind. Nach Konstruktion gilt das genau dann, wenn v_1, \dots, v_m linear abhängig sind.

Definition

Seien V ein n -dimensionaler Vektorraum, $m \leq n$ und $v_1, \dots, v_m \in V$. Ferner sei $P = \sum_{i=1}^m [0, 1]v_i$. Dann heißt

$$\text{Vol}_m(P) = \sqrt{G(v_1, \dots, v_m)}$$

das m -dimensionale Volumen von P .

Hinweis

In $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_R)$ stimmt diese Definition genau mit der geometrischen Berechnung und für $m = n$ auch mit dem entsprechenden Ergebnis aus (6.1) überein.

Die folgende Ungleichung behandelt das Optimierungsproblem, das Volumen eines von m Vektoren aufgespannten Parallelotops im \mathbb{R}^n zu maximieren.

8.3.10 Satz (Hadamonsche Ungleichung)

Seien V ein n -dimensionaler euklidischer Vektorraum, $m \leq n$, $v_1, \dots, v_m \in V$ und $P = \sum_{i=1}^m [0, 1]v_i$. Dann gilt

$$\text{Vol}_m(P) \leq \|v_1\| \cdots \|v_m\|$$

mit Gleichheit genau dann, wenn $\{v_1, \dots, v_m\}$ ein Orthogonalsystem bilden.

Beweis

Übung

8.4 Adjungierte Abbildungen und normale Endomorphismen**Definition**

Seien V, W euklidische (oder unitäre) Vektorräume, $F : V \rightarrow W$ linear und $G : W \rightarrow V$ linear. G heißt zu F ADJUNGIERTE ABBILDUNG, wenn gilt:

$$x \in V \wedge y \in W \Rightarrow \langle F(x), y \rangle = \langle x, G(y) \rangle$$

8.4.1 Hinweis

Seien V, W euklidische oder unitäre Vektorräume., $F : V \rightarrow W$ linear und $G : W \rightarrow V$ zu F adjungiert. Dann ist G eindeutig bestimmt.

Beweis

Sei $H : W \rightarrow V$ zu F adjungiert. Dann gilt für alle $x \in V$ und $y \in W$

$$\langle x, G(y) \rangle = \langle F(x), y \rangle = \langle x, H(y) \rangle$$

also $\langle x, G(y) - H(y) \rangle = 0$. Setzt man für x speziell $G(y) - H(y)$ ein, so folgt $G(y) = H(y)$ für alle $y \in W$.

Bezeichnung

Die Adjungierte zu F wird - falls sie existiert - mit F^{ad} bezeichnet.

8.4.2 Lemma

Seien V, W euklidische oder unitäre Vektorräume, $F : V \rightarrow W$ linear und es gelte $\dim V < \infty$. Dann existiert die zu F adjungierte Abbildung F^{ad} . Ist ferner $\{b_1, \dots, b_n\}$ eine Orthonormalbasis von V , so gilt $F^{\text{ad}}(y) = \sum_{i=1}^n \langle y, F(b_i) \rangle b_i$.

Hinweis

Ist V unendlich-dimensional, so existiert F^{ad} im Allgemeinen nicht.

Beweis

Nach (8.3.7) besitzt V eine Orthonormalbasis $\{b_1, \dots, b_n\}$ und nach (8.3.4) gilt für jedes $x \in V$ $x = \sum_{i=1}^n b_i \langle x, b_i \rangle$. Also gilt für $y \in W$

$$\begin{aligned} \langle F(x), y \rangle &= \sum_{i=1}^n \langle x, b_i \rangle \langle F(b_i), y \rangle \\ &= \sum_{i=1}^n \left\langle x, \overline{\langle F(b_i), y \rangle} b_i \right\rangle \\ &= \sum_{i=1}^n \langle x, \langle y, F(b_i) \rangle b_i \rangle \\ &= \left\langle x, \sum_{i=1}^n \langle y, F(b_i) \rangle b_i \right\rangle \end{aligned}$$

$G(y) := \sum_{i=1}^n \langle y, F(b_i) \rangle b_i$ ist linear und zu F adjungiert. Aus (8.4.1) folgt $F^{\text{ad}} = G$.

Wir zeigen nun, warum in (8.2) \overline{A}^T die zu A adjungierte Matrix genannt wurde.

8.4.3 Satz

Seien V, W endlich-dimensionale euklidische (bzw. unitäre) Vektorräume, $B_V = \{v_1, \dots, v_n\}$ Orthonormalbasis von V , $B_W = \{w_1, \dots, w_r\}$ Orthonormalbasis von W und $F : V \rightarrow W$ linear. Dann gilt

$$M_{B_V}^{B_W}(F^{\text{ad}}) = \overline{M_{B_W}^{B_V}(F)}^T$$

Beweis

Sei $M_{B_V}^{B_W}(F) = (a_{ij})_{\substack{i=1, \dots, r \\ j=1, \dots, n}}$. Dann gilt $F(v_j) = \sum_{i=1}^r a_{ij} w_i$. Da B_W Orthonormalbasis ist, folgt $\langle F(v_j), w_i \rangle = a_{ij}$ für $i = 1, \dots, r$, $j = 1, \dots, n$.

Ist andererseits $M_{B_V}^{B_W}(F^{\text{ad}}) = (b_{ji})_{\substack{i=1, \dots, r \\ j=1, \dots, n}}$, so folgt analog $b_{ji} = \langle F^{\text{ad}}(w_i), v_j \rangle$, also $b_{ji} = \overline{a_{ij}}$ und damit die Behauptung.

8.4.4 Lemma

Seien U, V, W euklidische (unitäre) Vektorräume, $F : U \rightarrow V$, $G : U \rightarrow V$ und $H : V \rightarrow W$ linear und es existieren die zugehörigen adjungierten Abbildungen F^{ad} , G^{ad} , H^{ad} . Dann existieren die im Folgenden jeweils links stehenden adjungierten Abbildungen und es gelten die folgenden Identitäten:

- (i) $(F^{\text{ad}})^{\text{ad}} = F$
- (ii) $(F + G)^{\text{ad}} = F^{\text{ad}} + G^{\text{ad}}$
- (iii) $(cF)^{\text{ad}} = \overline{c} F^{\text{ad}} \quad \forall c \in K$
- (iv) $(H \circ G)^{\text{ad}} = G^{\text{ad}} \circ H^{\text{ad}}$

Beweis

Die Aussagen ergeben sich mit (8.4.1) wie folgt:

$$(i) \quad \langle F^{\text{ad}}(y), x \rangle = \overline{\langle x, F^{\text{ad}}(y) \rangle} = \overline{\langle F(x), y \rangle} = \langle y, F(x) \rangle$$

$$(ii) \quad \begin{aligned} \langle (F + G)(x), y \rangle &= \langle F(x), y \rangle + \langle G(x), y \rangle \\ &= \langle x, F^{\text{ad}}(y) \rangle + \langle x, G^{\text{ad}}(y) \rangle \\ &= \langle x, (F^{\text{ad}} + G^{\text{ad}})(y) \rangle \end{aligned}$$

$$(iii) \quad \langle (cF)(x), y \rangle = c \langle F(x), y \rangle = c \langle x, F^{\text{ad}}(y) \rangle = \langle x, (\bar{c}F^{\text{ad}})(y) \rangle$$

$$(iv) \quad \langle (H \circ G)(x), y \rangle = \langle G(x), H^{\text{ad}}(y) \rangle = \langle x, (G^{\text{ad}} \circ H^{\text{ad}})(y) \rangle$$

8.4.5 Hinweis

Seien V ein euklidischer oder unitärer Vektorraum, $\dim V < \infty$ und $F : V \rightarrow V$ linear. Dann gilt $\det(F^{\text{ad}}) = \overline{\det F}$.

Beweis

folgt aus (8.4.3).

Definition

Seien V ein euklidischer oder unitärer Vektorraum und F ein Endomorphismus. F heißt **NORMAL**, wenn F^{ad} existiert und $F \circ F^{\text{ad}} = F^{\text{ad}} \circ F$ gilt.

Hinweis

Jeder selbstadjungierte Endomorphismus ist normal.

8.4.6 Lemma

Seien V ein euklidischer oder unitärer Vektorraum und F ein Endomorphismus. F ist genau dann normal, wenn F^{ad} existiert und für alle $x, y \in V$ $\langle F(x), F(y) \rangle = \langle F^{\text{ad}}(x), F^{\text{ad}}(y) \rangle$ gilt.

Beweis

\Rightarrow :

$$\begin{aligned} \langle F(x), F(y) \rangle &= \langle x, F^{\text{ad}}F(y) \rangle \\ &= \langle x, FF^{\text{ad}}(y) \rangle \\ &= \overline{\langle FF^{\text{ad}}(x), y \rangle} \\ &= \overline{\langle F^{\text{ad}}(y), F^{\text{ad}}(x) \rangle} \\ &= \langle F^{\text{ad}}(x), F^{\text{ad}}(y) \rangle \end{aligned}$$

⇐:

$$\begin{aligned}\langle FF^{\text{ad}}(x), y \rangle &= \langle F^{\text{ad}}(x), F^{\text{ad}}(y) \rangle \\ &= \langle F(x), F(y) \rangle \\ &= \langle F^{\text{ad}}F(x), y \rangle\end{aligned}$$

Somit gilt für alle $x, y \in V$: $\langle FF^{\text{ad}}(x) - F^{\text{ad}}F(x), y \rangle = 0$. Für $y = FF^{\text{ad}}(x) - F^{\text{ad}}F(x)$ folgt aus der Positivdefinitheit $FF^{\text{ad}}(x) = F^{\text{ad}}F(x)$ für alle $x \in V$, also $FF^{\text{ad}} = F^{\text{ad}}F$.

8.4.7 Korollar

Seien V ein euklidischer oder unitärer Vektorraum und F ein normaler Endomorphismus. Dann gilt $\text{kern}(F) = \text{kern}(F^{\text{ad}})$.

Beweis

$$\|F(x)\|^2 = \langle F(x), F(x) \rangle = \langle F^{\text{ad}}(x), F^{\text{ad}}(x) \rangle = \|F^{\text{ad}}(x)\|^2$$

8.4.8 Korollar

Seien V ein euklidischer oder unitärer Vektorraum und F ein normaler Endomorphismus.

- (i) F und F^{ad} besitzen die selben Eigenwerte.
- (ii) Seien $v \in V$ und $\lambda \in K$. Dann gilt $F(v) = \lambda v \Leftrightarrow F^{\text{ad}}(v) = \bar{\lambda}v$.

Beweis

Nach (8.4.4) gilt $(F - \lambda I)^{\text{ad}} = F^{\text{ad}} - \bar{\lambda}I$ und aus (8.4.7), angewandt auf $F - \lambda I$, folgt die Behauptung.

Der folgende Satz konkretisiert (7.2.11).

8.4.9 Satz

Seien V ein endlich-dimensionaler unitärer Vektorraum und $F \in \text{End}_{\mathbb{C}}(V)$. F ist genau dann normal, wenn es eine Orthonormalbasis von V gibt, die aus lauter Eigenvektoren von F besteht.

Beweis

⇒: Nach (3.3.11) zerfällt das charakteristische Polynom in Linearfaktoren. Nach (7.2.6) und (7.2.4) besitzt F also einen Eigenwert λ und einen zugehörigen Eigenvektor v . O.B.d.A. kann $\|v\| = 1$ vorausgesetzt werden. Die Behauptung gilt also für $n := \dim V = 1$. Wir setzen nun induktiv voraus, dass die Behauptung für $n - 1$ bewiesen sei. Sei $U = (\text{span}\{v\})^{\perp}$. Für $x \in U$ gilt nach (8.4.8) $\langle F(x), v \rangle = \langle x, F^{\text{ad}}(v) \rangle = \langle x, \bar{\lambda}v \rangle = \lambda \langle x, v \rangle = 0$, also $F(U) \subset U$. U ist somit F -invariant, nach (8.3.8) gilt $\dim U = n - 1$ und da $F|_U$ normal ist, folgt die Behauptung.

⇐: Seien $\{v_1, \dots, v_n\}$ eine Orthonormalbasis von V , $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ und $F(v_i) =$

$\lambda_i v_i$ für $i = 1, \dots, n$. Durch $G(v_i) = \overline{\lambda_i} v_i$ für $i = 1, \dots, n$ wird ein Endomorphismus $G: V \rightarrow V$ definiert. Für $i, j = 1, \dots, n$ gilt

$$\langle F(v_i), v_j \rangle = \lambda_i \langle v_i, v_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Für $i = j$ gilt $\langle F(v_i), v_i \rangle = \langle v_i, \overline{\lambda_i} v_i \rangle = \langle v_i, G(v_i) \rangle$ und hieraus folgt $\langle F(x), y \rangle = \langle x, G(y) \rangle$ für alle $x, y \in V$. G ist also der zu F adjungierte Endomorphismus F^{ad} . Ferner gilt $F^{\text{ad}}(F(v_i)) = F^{\text{ad}}(\lambda_i v_i) = \lambda_i \overline{\lambda_i} v_i = F(\overline{\lambda_i} v_i) = F(F^{\text{ad}}(v_i))$, also $F^{\text{ad}} \circ F = F \circ F^{\text{ad}}$, d.h. F ist normal.

8.4.10 Korollar

Seien V ein endlich-dimensionaler euklidischer Vektorraum und $F \in \text{End}_{\mathbb{R}}(V)$. Es existiert genau dann eine Orthonormalbasis aus Eigenvektoren von F , wenn F normal ist und lauter reelle Eigenwerte besitzt.

Beweis

folgt mit (8.2.4) aus (8.4.9).

8.4.11 Satz (Hauptachsentheorem für selbstadjungierte Endomorphismen)

Seien V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und F ein selbstadjungierter Endomorphismus. Dann gilt:

- (i) Alle Eigenwerte von F sind reell.
- (ii) V besitzt eine Orthonormalbasis aus Eigenvektoren von F .
- (iii) Seien λ_1, λ_2 Eigenwerte, $\lambda_1 \neq \lambda_2$, v_1 Eigenvektor zu λ_1 und v_2 Eigenvektor zu λ_2 . Dann gilt $v_1 \perp v_2$.

Beweis

- (i) Seien λ Eigenwert und v ein zugehöriger Eigenvektor. Dann gilt nach (8.4.8) wegen $F = F^{\text{ad}}$ $\lambda v = F(v) = F^{\text{ad}}(v) = \overline{\lambda} v$, d.h. $\lambda = \overline{\lambda}$.
- (ii) folgt aus (8.4.9) und (8.4.10).
- (iii)

$$\begin{aligned} \lambda_1 \langle v_1, v_2 \rangle &= \langle \lambda_1 v_1, v_2 \rangle = \langle F(v_1), v_2 \rangle \\ &= \langle v_1, F^{\text{ad}}(v_2) \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle \end{aligned}$$

, also $(\lambda_1 - \lambda_2) \langle v_1, v_2 \rangle = 0$. Wegen $\lambda_1 \neq \lambda_2$ folgt $v_1 \perp v_2$.

8.4.12 Korollar

Seien V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und F ein selbstadjungierter Endomorphismus. Ferner seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von F . Dann gilt $V = \text{Eig}(F, \lambda_1) \boxplus \dots \boxplus \text{Eig}(F, \lambda_k)$.

Beweis

folgt aus (8.4.11).

Wir wollen nun eine zu (8.4.9) analoge Aussage im euklidischen Fall herleiten. Dazu gehen wir von einem (endlich-dimensionalen) euklidischen Vektorraum V aus und konstruieren den gemäß (8.2.1) und (8.2.4) zugehörigen unitären Raum Z . Ein Endomorphismus F wird gemäß (8.2.3) zu einem Endomorphismus G von Z fortgesetzt.

8.4.13 Hinweis

Seien V ein euklidischer Vektorraum, $F \in \text{End}_{\mathbb{R}}(V)$, Z der zugehörige unitäre Raum und G der zugehörige Endomorphismus von Z . Dann gilt: F ist genau dann normal, wenn G normal ist.

Beweis

\Rightarrow : Seien $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, $x_1, x_2, y_1, y_2 \in V$. Dann gilt gemäß (8.2.3) und (8.2.4):

$$\begin{aligned} \langle G(z_1), z_2 \rangle_Z &= \langle G(x_1) + iG(y_1), x_2 + iy_2 \rangle_Z \\ &= \langle G(x_1), x_2 \rangle_Z + \langle G(y_1), y_2 \rangle_Z + i(\langle G(y_1), x_2 \rangle_Z - \langle G(x_1), y_2 \rangle_Z) \\ &= \langle F(x_1), x_2 \rangle_V + \langle F(y_1), y_2 \rangle_V + i(\langle F(y_1), x_2 \rangle_V - \langle F(x_1), y_2 \rangle_V) \\ &= \langle x_1, F^{\text{ad}}(x_2) \rangle_V + \langle y_1, F^{\text{ad}}(y_2) \rangle_V + \\ &\quad i(\langle y_1, F^{\text{ad}}(x_2) \rangle_V - \langle x_1, F^{\text{ad}}(y_2) \rangle_V) \\ &= \langle x_1 + iy_1, F^{\text{ad}}(x_2) + iF^{\text{ad}}(y_2) \rangle_Z \end{aligned}$$

Also existiert G^{ad} und es gilt $G^{\text{ad}}(z_2) = F^{\text{ad}}(x_2) + iF^{\text{ad}}(y_2)$. Hieraus folgt durch Nachrechnen, dass G normal ist.

\Leftarrow : Übung.

8.4.14 Lemma

Seien V ein euklidischer Vektorraum und $F \in \text{End}_{\mathbb{R}}(V)$ normal. Ferner sei Z der zugehörige unitäre Raum und G der zugehörige Endomorphismus. Seien $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_2 \neq 0$, $\lambda = \lambda_1 + i\lambda_2$, $x, y \in V$, $z = x + iy$, $\|z\| = 1$, $G(z) = \lambda z$. Dann ist $\hat{z} = x - iy$ normierter Eigenvektor von G mit Eigenwert $\bar{\lambda}$. Ferner sind z und \hat{z} orthogonal.

Beweis

Es gilt

$$\begin{aligned} \langle \hat{z}, \hat{z} \rangle_Z &= \langle x, x \rangle_V + \langle y, y \rangle_V + i(\langle x, y \rangle_V - \langle y, x \rangle_V) \\ &= \langle x, x \rangle_V + \langle y, y \rangle_V \\ &= \langle z, z \rangle_Z \end{aligned}$$

Ferner gilt

$$\begin{aligned} F(x) + iF(y) &= G(z) = \lambda z \\ &= \lambda_1 x - \lambda_2 y + i(\lambda_1 y + \lambda_2 x) \end{aligned}$$

also

$$\begin{aligned} F(x) &= \lambda_1 x - \lambda_2 y \\ F(y) &= \lambda_1 y + \lambda_2 x \end{aligned}$$

und somit

$$\begin{aligned} G(\hat{z}) &= F(x) - iF(y) \\ &= \lambda_1 x - \lambda_2 y - i(\lambda_1 y + \lambda_2 x) \\ &= (\lambda_1 - i\lambda_2)(x - iy) = \bar{\lambda}\hat{z} \end{aligned}$$

Nach (8.4.13) ist G normal und nach (8.4.8) ist \hat{z} Eigenwert von G^{ad} zum Eigenwert $\lambda = \bar{\lambda}$. Es folgt

$$\lambda \langle z, \hat{z} \rangle_Z = \langle \lambda z, \hat{z} \rangle_Z = \langle G(z), \hat{z} \rangle_Z = \langle z, G^{\text{ad}}(\hat{z}) \rangle_Z = \bar{\lambda} \langle z, \hat{z} \rangle_Z$$

also $(\lambda - \bar{\lambda}) \langle z, \hat{z} \rangle_Z = 0$. Wegen $\lambda_2 \neq 0$ folgt $z \perp \hat{z}$.

8.4.15 Satz

Seien V ein endlich-dimensionaler euklidischer Vektorraum und $F \in \text{End}_{\mathbb{R}}(V)$. F ist genau dann normal, wenn es eine Orthonormalbasis \mathcal{B} von V gibt, so dass $M_{\mathcal{B}}^{\mathcal{B}}(F)$ von der Form

$$\begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda_k & \ddots & \vdots \\ \vdots & & \ddots & N_1 & \ddots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & N_r \end{pmatrix}$$

ist, wobei $\lambda_1, \dots, \lambda_k$ die reellen Eigenwerte von F sind und jede Matrix $N_j \in \mathbb{R}^{2 \times 2}$ von der Form

$$\begin{pmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{pmatrix}$$

ist. Jedem solchen Zweierkästchen entspricht dabei ein Paar $\mu_j, \bar{\mu}_j$ konjugiert komplexer Eigenwerte der komplexen Fortsetzung $G \in \text{End}_{\mathbb{C}}(Z)$ von F und es gilt $\mu_j = \alpha_j + i\beta_j$ für $j = 1, \dots, r$.

Beweis

\Rightarrow : (vollständige Induktion nach $n := \dim V$) Für $n = 1$ ist die Behauptung trivial. Sei $n > 1$ und wir nehmen an, dass die Aussage für kleinere Dimensionen bereits bewiesen sei. Besitzt F einen reellen Eigenwert, so folgt die Induktionsbehauptung analog zum Beweis von (8.4.9). Anderenfalls betrachten wir den zu V gehörenden unitären Raum Z und den zu F gehörenden Endomorphismus G von Z . Seien μ ein Eigenwert von G und $z \in Z$ ein zugehöriger normierter Eigenvektor, und es gelte $z = x + iy$ mit $x, y \in V$. Nach (8.4.14) ist auch $\hat{z} = x - iy$ ein normierter und z orthogonaler Eigenvektor von G zum Eigenwert $\bar{\mu}$. Wir setzen:

$$\begin{aligned} v &= (z + \hat{z}) \frac{1}{\sqrt{2}} = \sqrt{2}x \\ w &= (z - \hat{z}) \frac{1}{i\sqrt{2}} = \sqrt{2}y \end{aligned}$$

Es gilt $v, w \in V$.

$$\begin{aligned}\langle v, v \rangle_V &= \frac{1}{2} \langle z + \hat{z}, z + \hat{z} \rangle_Z \\ &= \frac{1}{2} (\langle z, z \rangle_Z + \langle \hat{z}, \hat{z} \rangle_Z) = 1 \\ &= \frac{1}{2} \langle z - \hat{z}, z - \hat{z} \rangle_Z = \langle w, w \rangle_V\end{aligned}$$

und

$$\langle v, w \rangle_V = \frac{-1}{2i} \langle z + \hat{z}, z - \hat{z} \rangle_Z = -\frac{1}{2i} (\langle z, z \rangle_Z - \langle \hat{z}, \hat{z} \rangle_Z) = 0$$

Ferner ist mit $\mu = \alpha + i\beta$

$$\begin{aligned}F(v) + iF(w) &= \sqrt{2}G(z) = \sqrt{2}\mu z = \sqrt{2}(\alpha x - \beta y + i(\alpha y + \beta x)) \\ &= (\alpha v - \beta w) + i(\alpha w + \beta v)\end{aligned}$$

d.h. $F(v) = \alpha v - \beta w$, $F(w) = \beta v + \alpha w$. Bezüglich $W = \text{span}\{v, w\}$ entspricht F also einem Zweierkästchen der behaupteten Art. Wir setzen wie in (8.4.9) $U = W^\perp$. Für $a \in U$ gilt dann

$$\begin{aligned}\langle F(a), x \rangle_V - i \langle F(a), y \rangle_V &= \langle F(a), x + iy \rangle_Z \\ &= \langle G(a), z \rangle_Z \\ &= \langle a, G^{\text{ad}}(z) \rangle_Z \\ &= \langle a, \bar{\mu}z \rangle_Z \\ &= \mu \langle a, z \rangle_Z \\ &= \mu \langle a, x + iy \rangle_Z \\ &= \mu \langle a, x \rangle_V - i\mu \langle a, y \rangle_V \\ &= 0\end{aligned}$$

d.h. $F(a) \perp W$, also $F(U) \subset U$. U ist somit F -invariant und die Behauptung folgt durch Anwendung der Induktionsvoraussetzung auf $F|_U$.

⇐: Analog zu (8.4.9); beachte $M_B^B(F^{\text{ad}}) = M_B^B(F)^T$.

8.5 Orthogonale und unitäre Abbildungen

Wir kommen jetzt zu den wichtigsten normalen Endomorphismen zwischen euklidischen bzw. unitären Vektorräumen, nämlich solchen, die Skalarprodukte erhalten.

Definition

Seien V, W euklidische (unitäre) Vektorräume und $F : V \rightarrow W$ linear. F heißt ORTHOGONAL (UNITÄR), wenn gilt:

$$x, y \in V \Rightarrow \langle x, y \rangle = \langle F(x), F(y) \rangle$$

8.5.1 Lemma

Seien V, W euklidische (unitäre) Vektorräume und $F : V \rightarrow W$ linear. Dann sind die folgenden Aussagen äquivalent:

- (i) F ist orthogonal (unitär).

$$(ii) \quad x \in V \wedge \|x\| = 1 \Rightarrow \|F(x)\| = 1$$

$$(iii) \quad x \in V \Rightarrow \|x\| = \|F(x)\|$$

$$(iv) \quad \{b_1, \dots, b_n\} \text{ ist Orthonormalsystem von } V \Rightarrow \{F(b_1), \dots, F(b_n)\} \text{ ist Orthonormalsystem von } W$$

Beweis

(i) \Rightarrow (ii): trivial

(ii) \Rightarrow (iii): trivial

(iii) \Rightarrow (iv): Seien $j, k \in \{1, \dots, n\}$, $j \neq k$. Dann gilt

$$\begin{aligned} \|b_j\|^2 + \|b_k\|^2 &= \|b_j + b_k\|^2 \\ &= \|F(b_j) + F(b_k)\|^2 \\ &= \|F(b_j)\|^2 + \|F(b_k)\|^2 + \langle F(b_j), F(b_k) \rangle + \langle F(b_k), F(b_j) \rangle \\ &= \|b_j\|^2 + \|b_k\|^2 + \langle F(b_j), F(b_k) \rangle + \overline{\langle F(b_j), F(b_k) \rangle} \\ &= \|b_j\|^2 + \|b_k\|^2 \end{aligned}$$

und ebenso

$$\begin{aligned} \|b_j\|^2 + i\|b_k\|^2 &= \|b_j + ib_k\|^2 \\ &= \|b_j\|^2 + \|b_k\|^2 - i\langle F(b_j), F(b_k) \rangle + i\overline{\langle F(b_j), F(b_k) \rangle} \\ &= \|b_j\|^2 + \|b_k\|^2 \end{aligned}$$

Also:

$$\langle F(b_j), F(b_k) \rangle = -\overline{\langle F(b_j), F(b_k) \rangle} = \overline{\langle F(b_j), F(b_k) \rangle}$$

und es folgt (iv).

(iv) \Rightarrow (i): Seien $x, y \in V$. Sind x und y linear abhängig, so gibt es ein $c \in K$ mit $y = cx$. Ist $x = 0$, so folgt $\langle F(x), F(y) \rangle = 0 = \langle x, y \rangle$. Für $x \neq 0$ ist aber $\left\{ \frac{1}{\|x\|} x \right\}$ ein Orthonormalsystem, also nach Voraussetzung auch $\left\{ \frac{1}{\|x\|} F(x) \right\}$. Es folgt

$$\langle F(x), F(y) \rangle = \bar{c}\|x\|^2 \left\langle \frac{1}{\|x\|} F(x), \frac{1}{\|x\|} F(x) \right\rangle = \bar{c}\|x\|^2 = \bar{c}\langle x, x \rangle = \langle x, y \rangle$$

Seien nun x, y linear unabhängig. Sei $\{b_1, b_2\}$ eine Orthonormalbasis von $\text{span}\{x, y\}$ und $c_1, c_2, d_1, d_2 \in K$ mit $x = c_1b_1 + c_2b_2$ und $y = d_1b_1 + d_2b_2$. Da nach Voraussetzung $\{F(b_1), F(b_2)\}$ Orthonormalbasis ist, folgt

$$\begin{aligned} \langle F(x), F(y) \rangle &= \langle c_1F(b_1) + c_2F(b_2), d_1F(b_1) + d_2F(b_2) \rangle \\ &= c_1\bar{d}_1 + c_2\bar{d}_2 \\ &= \langle c_1b_1 + c_2b_2, d_1b_1 + d_2b_2 \rangle \\ &= \langle x, y \rangle \end{aligned}$$

8.5.2 Korollar

Seien V, W euklidische Vektorräume und $F : V \rightarrow W$ orthogonal. Sei G die komplexe Fortsetzung von F . Dann ist G unitär.

Beweis

trivial

8.5.3 Korollar

Seien V, W euklidische (unitäre) Vektorräume und $F : V \rightarrow W$ orthogonal (unitär). Dann ist F injektiv.

Beweis

trivial

Beispiele

- (i) $V = W = (\mathbb{R}^n, \langle \cdot, \cdot \rangle_R)$, $A \in \mathbb{R}^{n \times n}$, $F = F_A$. Dann ist $\langle F(x), F(y) \rangle_R = \langle Ax, Ay \rangle_R = x^T A^T A y$. F ist genau dann orthogonal, wenn $x^T A^T A y = x^T y$ ist, d.h. $x^T (A^T A - I)y = 0 \forall x, y \in \mathbb{R}^n$, d.h. $A^T A = I_n$.
- (ii) $V = W = (\mathbb{R}^2, \langle \cdot, \cdot \rangle_R)$, $A = (a_{ij})_{i,j=1,2} \in \mathbb{R}^{2 \times 2}$. F_A ist orthogonal genau dann, wenn $A^T A = I_2$ ist, d.h.

$$\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Leftrightarrow a_{11}^2 + a_{21}^2 = 1 \wedge a_{11}a_{12} + a_{21}a_{22} = 0 \wedge a_{12}a_{11} + a_{22}a_{21} = 0 \wedge a_{12}^2 + a_{22}^2 = 1$$

$$\Leftrightarrow \exists \alpha, \beta : \alpha, \beta \in [0, 2\pi[\wedge a_{11} = \cos \alpha \wedge a_{21} = \sin \alpha \wedge a_{12} = \sin \beta$$

$$\wedge a_{22} = \cos \beta \wedge \cos \alpha \sin \beta + \sin \alpha \cos \beta (= \sin(\alpha + \beta)) = 0$$

Die letzte Gleichung gilt genau für $\alpha + \beta \in \{0, \pi, 2\pi, 3\pi\}$. F_A ist also genau dann orthogonal, wenn A von der Form

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}$$

ist. (\gg Drehung \ll (Orientierung bleibt erhalten), \gg Spiegelung an Gerade \ll (Orientierung bleibt nicht erhalten))

- (iii) $V = W = (\mathbb{C}^n, \langle \cdot, \cdot \rangle_C)$, $A \in \mathbb{C}^{n \times n}$, $F = F_A$. $\langle F(x), F(y) \rangle = \langle Ax, Ay \rangle = x^T A^T \overline{Ay}$. Also ist F genau dann unitär, wenn $A^T \overline{A} = I$ ist.

Definition

Seien $K \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \text{GL}(n, K)$. A heißt ORTHOGONAL (UNITÄR) wenn für $K = \mathbb{R}$ ($K = \mathbb{C}$) $A^{-1} = A^T$ ($A^{-1} = \overline{A}^T$) gilt.

Hinweis

Ist A orthogonal (unitär), so gilt $|\det A| = 1$. Ist $A \in \mathbb{R}^{n \times n}$, so entspricht $\det A = 1$ nach (6.6) der Orientierungserhaltung.

8.5.4 Hinweis

Seien $K \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in K^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:

- (i) A ist orthogonal (unitär).
- (ii) Die Zeilen von A bilden ein Orthonormalsystem.
- (iii) Die Spalten von A bilden ein Orthonormalsystem.

Beweis

trivial

8.5.5 Bemerkung

Seien V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und F ein Endomorphismus. F ist genau dann orthogonal (unitär), wenn F invertierbar ist und $F^{-1} = F^{\text{ad}}$ gilt.

Beweis

\Rightarrow : Nach (8.5.3) und (5.2.10) ist F invertierbar. Nach (8.4.2) existiert F^{ad} und es gilt für alle $x, y \in V$:

$$\begin{aligned} \langle x, F^{\text{ad}}(y) - F^{-1}(y) \rangle &= \langle x, F^{\text{ad}}(y) \rangle - \langle x, F^{-1}(y) \rangle \\ &= \langle F(x), y \rangle - \langle x, F^{-1}(y) \rangle \\ &= \langle F(x), y \rangle - \langle F(x), y \rangle \\ &= 0 \end{aligned}$$

also $F^{\text{ad}} = F^{-1}$.

\Leftarrow : Es gilt für alle $x, z \in V$ mit $y = F^{-1}(z)$:

$$\langle x, y \rangle = \langle x, F^{-1}(z) \rangle = \langle x, F^{\text{ad}}(z) \rangle = \langle F(x), z \rangle = \langle F(x), F(y) \rangle$$

Da F bijektiv ist, folgt die Behauptung.

8.5.6 Korollar

Seien V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und $F : V \rightarrow V$ orthogonal (unitär). Dann ist F normal.

Beweis

Nach (8.5.5) gilt $FF^{\text{ad}} = FF^{-1} = I = F^{-1}F = F^{\text{ad}}F$.

Bemerkung

Damit sind insbesondere (8.4.9) und (8.4.15) anwendbar.

8.5.7 Bemerkung

Seien V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum, $\mathcal{B} = \{b_1, \dots, b_n\}$ Orthonormalbasis von V und F ein Endomorphismus. F ist genau dann orthogonal (unitär), wenn $M_{\mathcal{B}}^{\mathcal{B}}(F)$ orthogonal (unitär) ist.

Beweis

folgt aus (8.4.3) und (8.5.5)

Bezeichnung

$$\begin{aligned} O(n) &:= \{A : A \in \text{GL}(n, \mathbb{R}) \wedge A^{-1} = A^T\} \\ SO(n) &:= \{A : A \in O(n) \wedge \det A = 1\} \\ U(n) &:= \{A : A \in \text{GL}(n, \mathbb{C}) \wedge A^{-1} = \overline{A}^T\} \end{aligned}$$

8.5.8 Bemerkung

- (i) $O(n)$ ist Untergruppe von $GL(n, \mathbb{R})$.
- (ii) $SO(n)$ ist Untergruppe von $O(n)$.
- (iii) $U(n)$ ist Untergruppe von $GL(n, \mathbb{C})$.

Beweis

trivial

Bezeichnung

$O(n)$ heißt ORTHOGONALE GRUPPE, $SO(n)$ SPEZIELLE ORTHOGONALE GRUPPE und $U(n)$ UNITÄRE GRUPPE.

8.5.9 Satz (Hauptachsentheorem)

Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch (oder $A \in \mathbb{C}^{n \times n}$ hermitesch). Dann gibt es eine orthogonale (unitäre) Matrix S , so dass $S^{-1}AS$ eine reelle Diagonalmatrix ist.

Beweis

folgt aus (8.4.3), (8.4.11) und (8.5.4).

8.5.10 Satz

Sei $A \in \mathbb{R}^{n \times n}$ orthogonal. Dann gibt es eine orthogonale Matrix S , so dass S^TAS von der Form

$$\begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & & & & \vdots \\ \vdots & & \ddots & -1 & \ddots & & & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & & & \ddots & -1 & \ddots & & & \vdots \\ \vdots & & & & & \ddots & M_1 & \ddots & & \vdots \\ \vdots & & & & & & \ddots & \ddots & & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & M_r \end{pmatrix}$$

ist, wobei für $j = 1, \dots, r$

$$M_j = \begin{pmatrix} \cos \varphi_j & -\sin \varphi_j \\ \sin \varphi_j & \cos \varphi_j \end{pmatrix} \in SO(2)$$

mit $\varphi_j \in]0, \pi[\cup]\pi, 2\pi[$ ist.

Beweis

Wegen (8.5.6) ist (8.4.15) anwendbar und es existiert eine orthogonale Matrix S , so dass

$$S^T A S = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda_k & \ddots & \vdots \\ \vdots & & \ddots & N_1 & \ddots \\ \vdots & & & \ddots & \ddots \\ 0 & \dots & \dots & 0 & N_r \end{pmatrix} \quad \text{ist mit} \quad N_j = \begin{pmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{pmatrix}$$

Wegen (8.5.1) und (8.5.2) gilt $|\lambda_1| = \dots = |\lambda_k| = 1$ und $|\alpha_1 + i\beta_1| = \dots = |\alpha_r + i\beta_r| = 1$. Also gibt es für jedes $j = 1, \dots, r$ ein $\varphi_j \in [0, 2\pi[$ mit $\alpha_j = \cos \varphi_j$ und $\beta_j = \sin \varphi_j$, d.h.

$$N_j = \begin{pmatrix} \cos \varphi_j & -\sin \varphi_j \\ \sin \varphi_j & \cos \varphi_j \end{pmatrix}$$

Da die Zweierkästchen zu den nicht-reellen Eigenwerten gehören, gilt $\beta_j \neq 0$, also $\varphi_j \notin \{0, \pi\}$.

Beispiel

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix} \in SO(2) \quad \gg \text{Spiegelung} \ll$$

Es ist

$$p_A = ((\cos \alpha) - x)(-(\cos \alpha) - x) - \sin^2 \alpha = -1 + x^2 = (x+1)(x-1)$$

Normalform:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Wir beweisen nun noch eine Erweiterung von (8.2.5) und (8.3.5).

8.5.11 Satz

Seien $K \in \{\mathbb{R}, \mathbb{C}\}$, V ein n -dimensionaler K -Vektorraum, $\mathcal{B} = \{v_1, \dots, v_n\}$ eine (beliebige) Basis von V , $A \in K^{n \times n}$ und $F: V \times V \rightarrow K$ sei definiert durch

$$F(v, w) = \Phi_{\mathcal{B}}^{-1}(v)^T \overline{\Phi_{\mathcal{B}}^{-1}(w)}$$

F ist genau dann ein Skalarprodukt, wenn A für $K = \mathbb{R}$ symmetrisch bzw. für $K = \mathbb{C}$ hermitesch ist und alle Eigenwerte positiv sind.

Beweis

\Rightarrow : Sei $A = (a_{ij})_{i,j=1,\dots,n}$. Es gilt

$$a_{ij} = F(v_i, v_j) = \overline{F(v_j, v_i)} = \overline{a_{ji}}$$

also $A = \overline{A}^T$. Sei λ ein Eigenwert und x ein zugehöriger Eigenvektor. Nach (8.5.9) ist λ reell und es folgt

$$0 < F(\Phi_{\mathcal{B}}(x), \Phi_{\mathcal{B}}(x)) = x^T A \overline{x} = (\lambda x)^T \overline{x} = \lambda x^T \overline{x} = \lambda \|x\|^2$$

also $\lambda > 0$.

⇐: Wir zeigen, dass F positiv definit ist; die anderen Eigenschaften sind klar. Nach (8.5.9) gibt es eine unitäre Matrix S , so dass

$$D = \bar{S}^T A S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

gilt, wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A sind. Es folgt mit $x = \Phi_B^{-1}(v)$:

$$F(v, v) = \Phi_B^{-1}(v)^T A \overline{\Phi_B^{-1}(v)} = x^T A \bar{x} = x^T S D \bar{S}^T \bar{x}$$

Mit $y = S^T x$ gilt somit $F(v, v) = y^T D \bar{y}$ und mit $y = (y_1, \dots, y_n)^T$ folgt

$$F(v, v) = \sum_{i=1}^n \lambda_i y_i \bar{y}_i = \sum_{i=1}^n \lambda_i |y_i|^2$$

Für $v \neq 0$ gilt $x \neq 0$ und damit $y \neq 0$. Da nach Voraussetzung $\lambda_1, \dots, \lambda_n > 0$ gilt, folgt die Behauptung.

Zum Abschluss dieses Paragraphen beweisen wir einen »Trägheitssatz«, der insbesondere zur Charakterisierung quadratischer Formen benutzt wird.

Bezeichnung

Seien $A \in \mathbb{R}^{n \times n}$ symmetrisch bzw. $A \in \mathbb{C}^{n \times n}$ hermitesch und $t(A)$ die Anzahl der positiven Eigenwerte von A . $t(A)$ heißt TRÄGHEITSSINDEX von A .

8.5.12 Satz

Seien $A \in \mathbb{R}^{n \times n}$ symmetrisch (oder $A \in \mathbb{C}^{n \times n}$ hermitesch), $Q \in \text{GL}(n, \mathbb{R})$ ($Q \in \text{GL}(n, \mathbb{C})$) und $B = Q^T A Q$ ($B = \bar{Q}^T A Q$). Dann ist B symmetrisch (hermitesch) und es gilt $t(A) = t(B)$.

Beweis

(komplexer Fall)

$$\bar{B}^T = (\bar{Q}^T A Q)^T = \bar{Q}^T \bar{A}^T Q = B$$

Seien $\lambda_1, \dots, \lambda_n$ bzw. μ_1, \dots, μ_n die Eigenwerte von A bzw. B . Da sie reell sind und $r = \text{rang}(A) = \text{rang}(B)$ gilt, können wir annehmen, dass mit geeigneten t und s

$$\lambda_j \begin{cases} > 0 & \text{für } j = 1, \dots, t \\ < 0 & \text{für } j = t+1, \dots, r \\ = 0 & \text{für } j = r+1, \dots, n \end{cases} \quad \text{und} \quad \mu_j \begin{cases} > 0 & \text{für } j = 1, \dots, s \\ < 0 & \text{für } j = s+1, \dots, r \\ = 0 & \text{für } j = r+1, \dots, n \end{cases} \quad \text{gilt.}$$

Seien $S, T \in U(n)$ mit

$$D_A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \bar{S}^T A S$$

und

$$D_B = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix} = \bar{T}^T B T$$

Für $x = (x_1, \dots, x_n)^T \in \mathbb{C}^n$ gilt dann

$$\bar{x}^T D_A x = \sum_{i=1}^t \lambda_j |x_j|^2 - \sum_{j=t+1}^r |\lambda_j| |x_j|^2$$

und wegen

$$\begin{aligned} D_A &= \bar{S}^T A S = \bar{S}^T (\bar{Q}^T)^{-1} B Q^{-1} S = \bar{S}^T (\bar{Q}^{-1})^T (\bar{T}^{-1})^T D_B T^{-1} Q^{-1} S \\ &= \bar{S}^T (\bar{Q}^{-1})^T T D_B \bar{T}^T Q^{-1} S = \bar{C}^T D_B C \end{aligned}$$

mit $C = \bar{T}^T Q^{-1} S = (c_{jk})_{j,k=1,\dots,n}$ und $y = (y_1, \dots, y_n)^T = Cx$ auch

$$\bar{x}^T D_A x = \bar{x}^T \bar{C}^T D_B C x = \bar{y}^T D_B y = \sum_{j=1}^s \mu_j |y_j|^2 - \sum_{j=s+1}^r |\mu_j| |y_j|^2$$

Wäre $t < s$, so hätte das lineare Gleichungssystem

$$\begin{aligned} x_j &= 0 & j &= 1, \dots, t \\ y_j &= \sum_{k=1}^n c_{jk} x_k = 0 & j &= s+1, \dots, n \end{aligned}$$

eine nichttriviale Lösung $z = (z_1, \dots, z_n)^T \in \mathbb{C}^n$ und es wäre

$$\begin{aligned} \bar{z}^T D_A z &= - \sum_{j=t+1}^r |\lambda_j| |z_j|^2 \leq 0 \\ \bar{z}^T D_A z &= \sum_{j=t+1}^s \mu_j |z_j|^2 > 0 \end{aligned}$$

Somit folgt $t \geq s$. Analog folgt $t \leq s$, also $t(A) = t = s = t(B)$.

8.6 Anwendungen

Wir geben nun verschiedene Anwendungen der entwickelten Theorie an. Wir beginnen mit linearen Gleichungssystemen, der Einfachheit halber über \mathbb{R} . Wie wir gesehen haben, sind diesen vier fundamentale Teilräume zugeordnet.

Definition

Sei $A \in \mathbb{R}^{m \times n}$, $a_1, \dots, a_n \in \mathbb{R}^m$ die Spaltenvektoren und $b_1^T, \dots, b_m^T \in \mathbb{R}^n$ die Zeilenvektoren von A . Seien

$$\begin{aligned} \mathcal{R}(A) &:= \text{bild}(F_A) = \text{span}\{a_1, \dots, a_n\} \\ \mathcal{R}(A^T) &:= \text{bild}(F_{A^T}) = \text{span}\{b_1, \dots, b_m\} \\ \mathcal{N}(A) &:= \text{kern}(F_A) = \{x \in \mathbb{R}^n : Ax = 0\} \\ \mathcal{N}(A^T) &:= \text{kern}(F_{A^T}) = \{x \in \mathbb{R}^m : x^T A = 0\} \end{aligned}$$

$\mathcal{R}(A)$ heißt auch SPALTENRAUM von A , $\mathcal{R}(A^T)$ ZEILENRAUM, $\mathcal{N}(A)$ NULLRAUM und $\mathcal{N}(A^T)$ LINKER NULLRAUM von A .

8.6.1 Satz

Sei $A \in \mathbb{R}^{m \times n}$. Dann gilt:

$$(i) \mathcal{R}(A)^\perp = \mathcal{N}(A^T)$$

$$(ii) \mathcal{R}(A^T)^\perp = \mathcal{N}(A)$$

Beweis

(i) Seien $y = Ax \in \mathcal{R}(A)$ und $z \in \mathcal{N}(A^T)$. Dann gilt

$$\langle z, y \rangle_R = z^T y = z^T Ax = 0$$

Ferner folgt aus (5.2.7), (5.2.14) und (8.3.8)

$$\begin{aligned} \dim(\mathcal{R}(A)^\perp) &= m - \dim(\mathcal{R}(A)) = m - \dim(\mathcal{R}(A^T)) \\ &= m - (m - \dim \mathcal{N}(A^T)) \end{aligned}$$

und insgesamt folgt $\mathcal{R}(A)^\perp = \mathcal{N}(A^T)$.

(ii) Folgt analog oder durch Anwendung von (i) auf A^T .

8.6.2 Korollar (Alternativsatz)

Seien $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Das lineare Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn gilt

$$y \in \mathbb{R}^m \wedge A^T y = 0 \Rightarrow b^T y = 0$$

Beweis

$Ax = b$ ist genau dann lösbar, wenn $b \in \mathcal{R}(A)$ gilt. Nach (8.6.1) ist $\mathcal{R}(A) = \mathcal{N}(A^T)^\perp$, d.h. $Ax = b$ ist genau dann lösbar, wenn $b \perp \mathcal{N}(A^T)$ ist.

8.6.3 Korollar (Friedhelms Alternative)

Seien $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Dann hat genau eines der folgenden zwei Systeme eine Lösung:

$$\begin{aligned} Ax &= b \\ A^T y &= 0 \wedge b^T y \neq 0 \end{aligned}$$

Beweis

Folgt direkt aus (8.6.2).

Wir betrachten nun das Problem, eine *kleinste-Quadrate-Approximation* für ein unzulässiges Gleichungssystem zu finden; dieses führt unmittelbar zur Berechnung der orthogonalen Projektion von Punkten auf einen Teilraum.

Definition

Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und es gelte $\text{Lös}(A, b) = \emptyset$. Sei $x^* \in \mathbb{R}^n$ mit

$$\|Ax^* - b\|_{(2)}^2 = \min_{x \in \mathbb{R}^n} \|Ax - b\|_{(2)}^2$$

Dann heißt x^* KLEINSTE-QUADRATE-(APPROXIMATIONS-)LÖSUNG von $Ax = b$.

8.6.4 Lemma

Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ mit $\text{Lös}(A, b) = \emptyset$. x^* ist genau dann kleinste-Quadrate-Lösung, wenn gilt

$$(Ax^* - b) \perp \mathcal{R}(A)$$

Beweis

Nach (8.6.1) gibt es eindeutig bestimmte Vektoren $y_0 \in \mathcal{R}(A)$ und $z_0 \in \mathcal{N}(A^T)$ und $b = y_0 + z_0$. Ferner gilt für jedes $x \in \mathbb{R}^n$

$$\|Ax - b\|_{(2)}^2 = \|Ax - y_0 + y_0 - b\|_{(2)}^2 = \|Ax - y_0\|_{(2)}^2 + \|z_0\|_{(2)}^2$$

Es folgt

$$\|Ax - b\|_{(2)}^2 \geq \|z_0\|_{(2)}^2$$

Mit $y_0 = Ax_0$ gilt andererseits

$$\|Ax_0 - b\|_{(2)}^2 = \|y_0 - b\|_{(2)}^2 = \|z_0\|_{(2)}^2$$

also

$$\|z_0\|_{(2)}^2 = \min_{x \in \mathbb{R}^n} \|Ax - b\|_{(2)}^2$$

x^* ist also genau dann kleinste-Quadrate-Lösung, wenn $Ax^* = y_0$ ist.

8.6.5 Korollar

Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ mit $\text{Lös}(A, b) = \emptyset$. $x^* \in \mathbb{R}^n$ ist genau dann kleinste-Quadrate-Lösung von $Ax = b$, wenn $A^T(Ax^* - b) = 0$, also $(Ax^* - b) \in \mathcal{N}(A^T)$ gilt.

Beweis

Folgt aus (8.6.4) und (8.4.1).

Hinweis

Die gleiche Bedingung erhält man auch aus der Analysis. Sei $f(x) = \|Ax - b\|_{(2)}^2$. Dann gilt

$$\begin{aligned} f(x) &= (x^T A^T - b^T)(Ax - b) \\ &= x^T A^T Ax - x^T A^T b - b^T Ax + b^T b \\ &= x^T A^T Ax - 2x^T A^T b + b^T b \end{aligned}$$

also

$$f'(x) = 2A^T Ax - 2A^T b \stackrel{!}{=} 0$$

ist notwendige Bedingung. $f'(x) = 0$ ist äquivalent zu $A^T(Ax - b) = 0$.

Bezeichnung

$A^T Ax = A^T b$ heißt NORMALGLEICHUNG.

8.6.6 Satz

Seien $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$ und es gelte $\text{rang}(A) = n$. Dann ist $A^T A$ invertierbar, die kleinste-Quadrate-Lösung x^* von $Ax = b$ ist eindeutig bestimmt und es gilt

$$x^* = (A^T A)^{-1} A^T b$$

Für die orthogonale Projektion y_0 von b auf $\mathcal{R}(A)$ gilt

$$y_0 = Ax^* = A(A^T A)^{-1} A^T b$$

Beweis

Die Invertierbarkeit von $A^T A$ folgt aus (8.3.9); der Rest folgt aus (8.6.4) und (8.6.5), falls $\text{Lös}(A, b) = \emptyset$ ist.

Ist andererseits $b \in \mathcal{R}(A)$, dann gibt es ein $\hat{x} \in \mathbb{R}^n$ mit $b = A\hat{x}$ und es folgt

$$\hat{x} = (A^T A)^{-1} (A^T A) \hat{x} = (A^T A)^{-1} A^T b$$

sowie

$$b = A\hat{x} = A(A^T A)^{-1} A^T b = y_0$$

Bezeichnung

Seien $A \in \mathbb{R}^{m \times n}$ mit $\text{rang}(A) = n$ und $P = A(A^T A)^{-1} A^T$. Dann heißt P die zu A gehörige PROJEKTIONSMATRIX.

8.6.7 Satz

Seien $A \in \mathbb{R}^{m \times n}$ mit $\text{rang}(A) = n$ und P die zugehörige Projektionsmatrix. Dann ist P symmetrisch und es gilt $P^2 = P$. Umgekehrt ist jede symmetrische Matrix P mit $P^2 = P$ eine Projektionsmatrix.

Beweis

\Rightarrow : nachrechnen

\Leftarrow : Es reicht zu zeigen, dass $(b - Pb) \perp \mathcal{R}(P)$ ist. Sei also $c \in \mathbb{R}^m$. Dann gilt

$$\begin{aligned} \langle b - Pb, Pc \rangle_R &= (b - Pb)^T Pc \\ &= (b^T - b^T P^T) Pc \\ &= b^T Pc - b^T P^T Pc \\ &= b^T Pc - b^T P^2 c \\ &= 0 \end{aligned}$$

Also ist Pb die orthogonale Projektion auf $\mathcal{R}(P)$.

Als nächstes betrachten wir *Dualräume* und zeigen, wie lineare Funktionale $F : V \rightarrow \mathbb{R}$ mit \langle, \rangle zusammenhängen.

Definition

Seien V ein K -Vektorraum und

$$V^* = \text{Hom}_K(V, K) = \{F : V \rightarrow K; F \text{ linear}\}$$

Dann heißt V^* DUALRAUM von V . Die Elemente von V^* heißen LINEARFORMEN AUF V .

Beispiel

Sei V der Vektorraum der auf $[0, 1]$ differenzierbaren reellen Funktionen. Dann sind

$$F_1 : V \rightarrow \mathbb{R}, f \mapsto F_1(f) := \int_0^1 f(x) dx$$

$$F_2 : V \rightarrow \mathbb{R}, f \mapsto F_2(f) := f' \left(\frac{1}{2} \right)$$

Linearformen von V .

Bemerkung

Die Funktionalanalysis untersucht insbesondere Linearformen auf unendlich-dimensionalen Vektorräumen.

8.6.8 Bemerkung

Seien V ein n -dimensionaler K -Vektorraum, $\mathcal{B} = \{b_1, \dots, b_n\}$ eine Basis von V und für jeden Index $j = 1, \dots, n$ sei $b_j^* \in V^*$ definiert durch

$$b_j^*(b_k) = \delta_{jk} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

Dann ist $\mathcal{B}^* = \{b_1^*, \dots, b_n^*\}$ eine Basis von V^* .

Beweis

Nach (5.3.1) sind b_1^*, \dots, b_n^* eindeutig bestimmt. Seien $\lambda_1, \dots, \lambda_n \in K$ mit

$$\sum_{j=1}^n \lambda_j b_j^* = 0$$

Es folgt

$$0 = 0(b_k) = \left(\sum_{j=1}^n \lambda_j b_j^* \right) (b_k) = \sum_{i=1}^n \lambda_i b_i^*(b_k) = \lambda_k$$

Ferner gilt für $F \in V^*$ $F = \sum_{j=1}^n F(b_j) b_j^*$ und es folgt die Behauptung.

Bezeichnung

\mathcal{B}^* heißt die zu \mathcal{B} DUALE BASIS von V^* .

8.6.9 Korollar

Seien V ein n -dimensionaler K -Vektorraum und $\mathcal{B} = \{b_1, \dots, b_n\}$ eine Basis. Dann gibt es einen Isomorphismus $\psi_{\mathcal{B}} : V \rightarrow V^*$ mit $\psi_{\mathcal{B}}(b_j) = b_j^*$ für $j = 1, \dots, n$.

Beweis

Folgt aus (8.6.8).

8.6.10 Lemma (kleines Lemma von Riesz)

Seien V ein n -dimensionaler euklidischer Vektorraum und $F \in V^*$. Dann existiert genau ein $a \in V$, so dass $F(v) = \langle a, v \rangle$ ist für alle $v \in V$.

Beweis

Sei b_1, \dots, b_n eine Basis von V und $G : V \rightarrow \mathbb{R}^n$ sei definiert durch

$$G(v) = \begin{pmatrix} \langle v, b_1 \rangle \\ \vdots \\ \langle v, b_n \rangle \end{pmatrix}$$

G ist linear und injektiv, denn aus $G(v) = 0$ folgt

$$\left\langle v, \sum_{j=1}^n \lambda_j b_j \right\rangle = \sum_{j=1}^n \lambda_j \langle v, b_j \rangle = 0$$

für beliebige $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, d.h. $\langle v, w \rangle = 0$ für alle $w \in V$, also $v = 0$, d.h. $G = 0$.

Nach (5.2.10) ist G auch surjektiv, d.h. zu $\begin{pmatrix} F(b_1) \\ \vdots \\ F(b_n) \end{pmatrix}$ existiert genau ein Urbild $a \in V$. Seien nun $v \in V$, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit $v = \sum_{j=1}^n \lambda_j b_j$. Dann gilt

$$\langle a, v \rangle = \sum_{j=1}^n \lambda_j \langle a, b_j \rangle = \sum_{j=1}^n \lambda_j F(b_j) = F \left(\sum_{j=1}^n \lambda_j b_j \right) = F(v)$$

Geometrische Interpretation

Sei $F \in V^*$ und $a \in V \setminus \{0\}$ mit $F(v) = \langle v, a \rangle$ für alle $v \in V$. Nach (5.2.2) ist $\bigcup_{\lambda \in \mathbb{R}} F^{-1}(\lambda)$ eine disjunkte Zerlegung von V in Mengen der Form $u + \ker(F)$. Nach (5.2.7) gilt $\dim(\ker(F)) = n - 1$ ($n := \dim V$), d.h. die Fasern sind Hyperebenen. Ferner folgt mit $H := \ker(F)$ für $x \in H$ $0 = F(x) = \langle a, x \rangle$, d.h. a ist orthogonal zu H . Es gilt $V = H \boxplus \text{span}\{a\}$. F ist also bereits durch $F(a)$ eindeutig bestimmt.

Wir zeigen nun, wie man Eigenwerte symmetrischer Matrizen durch ein Optimierungsprinzip charakterisieren kann.

Definition

Seien $A \in \mathbb{R}^{n \times n}$ und $q : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ definiert durch

$$q(x) = \frac{\langle x, Ax \rangle_{\mathbb{R}}}{\langle x, x \rangle_{\mathbb{R}}}$$

für alle $x \in \mathbb{R}^n \setminus \{0\}$. q heißt RAYLEIGH-QUOTIENT.

Bezeichnung

Für $k = 0, \dots, n$ sei \mathcal{U}^k die Menge der k -dimensionalen Untervektorräume von \mathbb{R}^n .

8.6.11 Satz (Rayleigh-Prinzip)

Seien $A \in \mathbb{R}^{n \times n}$ symmetrisch, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ die Eigenwerte von A und sei $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Dann gilt für $k = 1, \dots, n$:

$$\lambda_k = \max_{U \in \mathcal{U}^{n-k+1}} \min_{x \in U} q(x)$$

$$\lambda_k = \min_{U \in \mathcal{U}^k} \max_{x \in U} q(x)$$

Beweis

Nach (8.5.9) gibt es eine orthogonale Matrix S mit

$$S^T A S = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Die Spalten s_1, \dots, s_n von S sind zu $\lambda_1, \dots, \lambda_n$ gehörige Eigenvektoren. Es folgt für $x \in \mathbb{R}^n \setminus \{0\}$ mit $x = S y$, $y = (y_1, \dots, y_n)^T$:

$$q(x) = \frac{x^T A x}{x^T x} = \frac{(S y)^T A S y}{(S y)^T S y} = \frac{y^T S^T A S y}{y^T S^T S y} = \frac{y^T D y}{y^T y}$$

Für $k = 1, \dots, n$ sei $U_k = \text{span}\{s_1, \dots, s_k\}$. Dann gilt

$$\begin{aligned} \max_{U \in \mathcal{U}^{n-k+1}} \min_{x \in U} q(x) &\geq \min_{x \in U_{k-1}^\perp} q(x) \\ &= \min_{y \in \mathbb{R}^n \setminus \{0\}} \frac{\sum_{j=k}^n \lambda_j y_j^2}{\sum_{j=k}^n y_j^2} \\ &\geq \min_{y \in \mathbb{R}^n \setminus \{0\}} \frac{\lambda_k \sum_{j=k}^n y_j^2}{\sum_{j=k}^n y_j^2} \\ &= \lambda_k \end{aligned}$$

Sei andererseits $U \in \mathcal{U}^{n-k+1}$ beliebig. Dann gilt nach (4.3.1) $\dim(U \cap U_k) \geq 1$. Seien $x \in (U \cap U_k) \setminus \{0\}$ und $x = S y$. Dann ist

$$q(x) = \frac{\sum_{j=1}^k \lambda_j y_j^2}{\sum_{j=1}^k y_j^2} \leq \lambda_k$$

und es folgt die erste Behauptung. Die zweite Behauptung ergibt sich analog.

Zum Abschluss betrachten wir *Nullstellenmengen quadratischer Formen*.

Definition

Sei $Q : \mathbb{R}^n \rightarrow \mathbb{R}$. Q heißt QUADRATISCHE FORM, wenn es $A \in \mathbb{R}^{n \times n}$ mit $0 \neq A = A^T$, $b \in \mathbb{R}^n$ und $c \in \mathbb{R}$ gibt, so dass

$$Q(x) = x^T A x + 2b^T x + c$$

für alle $x \in \mathbb{R}^n$ gilt. M heißt QUADRIK (oder HYPERFLÄCHE ZWEITER ORDNUNG), wenn es eine quadratische Form Q gibt mit

$$M = \{x : x \in \mathbb{R}^n \wedge Q(x) = 0\}$$

Bezeichnung

Seien $A \in \mathbb{R}^{n \times n}$ mit $0 \neq A = A^T$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ und

$$\hat{A} = \begin{pmatrix} c & b^T \\ b & A \end{pmatrix}$$

Dann heißt \hat{A} die zu (A, b, c) gehörige ERWEITERTE MATRIX.

8.6.12 Hinweis

Sei $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ die quadratische Form $Q(x) = x^T A x + 2b^T x + c$. Dann gilt

$$Q(x) = (1, x^T) \hat{A} \begin{pmatrix} 1 \\ x \end{pmatrix}$$

Ist $\hat{Q} : \mathbb{R}^n \rightarrow \mathbb{R}$ die durch

$$\hat{Q} \begin{pmatrix} x_1 \\ x \end{pmatrix} = (x_1, x^T) \hat{A} \begin{pmatrix} x_1 \\ x \end{pmatrix}$$

definierte quadratische Form des \mathbb{R}^{n+1} , so gilt für die zu Q gehörige Quadrik

$$\begin{aligned} M &= \{x : x \in \mathbb{R}^n \wedge Q(x) = 0\} \\ &= \{y : y \in \mathbb{R}^{n+1} \wedge \hat{Q}(y) = 0\} \cap \{y : y \in \mathbb{R}^{n+1} \wedge e_1^T y = 1\} \end{aligned}$$

Beweis

trivial

Beispiel

$$Q(x_1, x_2) = x_1^2 + x_2^2 - 1 = (x_1, x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 1$$

Es gilt

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \hat{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \hat{Q} = -y_1^2 + y_2^2 + y_3^2$$

$\{y : y \in \mathbb{R}^3 \wedge \hat{Q}(y) = 0\}$ definiert einen Kegel im \mathbb{R}^3 .

$M = \{y : y \in \mathbb{R}^3 \wedge \hat{Q}(y) = 0\} \cap \{y : y = (y_1, y_2, y_3) \in \mathbb{R}^3 \wedge y_1 = 1\}$ bedeutet einen Schnitt des Kegels mit der Ebene $y_1 = 1$ (vgl. Abb. 8.1).

Definition

Seien $Q_1, Q_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ quadratische Formen. Q_1 und Q_2 heißen **AFFIN ÄQUIVALENT**, wenn es eine affine Transformation φ gibt mit $Q_2(x) = Q_1(\varphi(x))$ für alle $x \in \mathbb{R}^n$.

Q_1 und Q_2 heißen **KONGRUENT**, wenn es eine affine Transformation $\varphi = a + F$ gibt, wobei die zugehörige lineare Abbildung F orthogonal ist, so dass $Q_2(x) = Q_1(\varphi(x))$ für alle $x \in \mathbb{R}^n$.

Ziel ist es, Normalformen bezüglich affiner Äquivalenz bzw. Kongruenz zu finden.

Definition

Sei $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ die quadratische Form $Q(x) = x^T A x + 2b^T x + c$ und es seien $r(Q) := \text{rang}(A)$, $\hat{r}(Q) := \text{rang}(\hat{A})$ und $t(Q) := t(A)$. $r(Q)$, $\hat{r}(Q)$, $t(Q)$ heißen **RANG**, **ERWEITERTER RANG** bzw. **TRÄGHEITSINDEX** von Q .

Beispiel

Seien $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $c = 0$. Dann ist

$$\text{rang}(A) = 1$$

$$\text{rang}(\hat{A}) = \text{rang} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 3$$

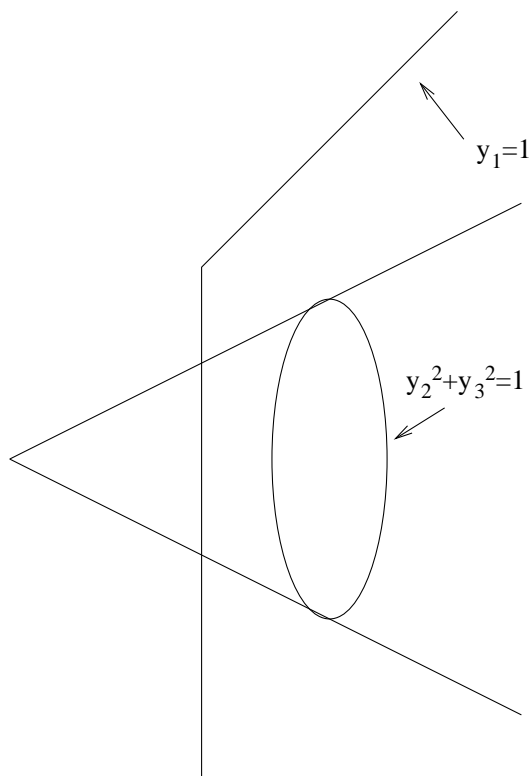


Abbildung 8.1: Kegelschnitt

d.h. mit $Q(x) = x^T Ax + 2b^T x + c$ gilt $r(Q) = 1$, $\hat{r}(Q) = 3$, $t(Q) = 1$.

Das folgende Ergebnis ist eine einfache Folgerung aus (8.5.12).

8.6.13 Satz (Trägheitssatz von Sylvester)

Affin äquivalente quadratische Formen haben den selben Trägheitsindex.

8.6.14 Satz

Sei $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ eine quadratische Form. Q ist genau zu einer der nachfolgenden quadratischen Formen kongruent:

(1) $\hat{r}(Q) = r(Q)$:

$$\sum_{j=1}^{t(Q)} a_j x_j^2 - \sum_{j=1}^{t(Q)} a_j x_j^2$$

(2a) $\hat{r}(Q) = r(Q) + 1$:

$$\sum_{j=1}^{t(Q)} a_j x_j^2 - \sum_{j=1}^{t(Q)} a_j x_j^2 - d$$

(2b) $\hat{r}(Q) = r(Q) + 1$:

$$\sum_{j=1}^{t(Q)} a_j x_j^2 - \sum_{j=1}^{t(Q)} a_j x_j^2 + d$$

(3) $\hat{r}(Q) = r(Q) + 2$:

$$\sum_{j=1}^{t(Q)} a_j x_j^2 - \sum_{j=1}^{t(Q)} a_j x_j^2 + dx_{r(Q)+1}$$

Die auftretenden Koeffizienten sind alle positiv, und es gelten die Ungleichungen

$$a_1 \geq \cdots \geq a_{t(Q)}$$

und

$$a_{t(Q)+1} \geq t_{t(Q)}$$

Die Aussage gilt ebenso für »affin-äquivalent« statt »kongruent«, wenn nur Koeffizienten $a_j = 1$ zugelassen sind.

Mit Hilfe von (8.6.14) kann man auch alle Quadriken klassifizieren; für $n = 3$ erhält man die folgenden Typen:

- Punkt
- Kegel
- Gerade
- Paar sich schneidender Ebenen
- Ebene
- Ellipsoid
- ein- / zweisechaliges Hyperboloid
- \emptyset
- elliptischer / hyperbolischer Zylinder
- elliptischer / hyperbolischer Paraboloid
- parabolischer Zylinder