

Vorlesungsmitschrift der Vorlesung „Mathematik
für Physiker“ gelesen von Prof. Pabel im SS2004
an der Bayrischen Julius-Maximilians-Universität
in Würzburg

Andreas Messer

18. Juli 2004

Inhaltsverzeichnis

1	Lineare Algebra	2
1.1	Vektorräume	3
1.2	Basis- und Dimensionssätze im endlich dimensionalen VR	11
1.3	Lineare Abbildungen	13
1.4	Matrizen	19
1.5	Lineare Gleichungssysteme	29
1.5.1	Homogene lineare Gleichungssysteme mit $b = 0 = \mathbb{K}^m$	30
1.5.2	Inhomogene lineare Gleichungssystem mit $b \in \mathbb{K}^m$ beliebig	31
1.6	Determinanten	34
1.7	Einführung in die Eigenwerttheorie	38
1.8	Euklidische und unitäre VR	43
1.8.1	Längenmessung	44
1.8.2	Orthogonalität	45
1.8.3	Winkelmessung (nur im Reellen)	47
1.8.4	Volumen- und Flächenberechnung (nur im \mathbb{R}^n)	48
1.8.5	Zum Vektor- oder Kreuzprodukt	48
2	Differential und Integralrechnung im \mathbb{R}^n	52
2.1	Konvergenz und Stetigkeit im \mathbb{R}^n	53
2.1.1	Topologische Grundbegriffe	53
2.1.2	Konvergenzbegriffe	53
2.1.3	Stetigkeit	54
2.2	Partielle und totale Differenzierbarkeit	55
2.3	Lokale und globale Extrema	63
2.4	Parameterabhängige Riemann-Integrale	69
2.5	Integralrechnung in mehreren Veränderlichen	70
2.5.1	Doppelintegrale über Rechtecke	70
2.5.2	Doppelintegrale über Normalbereiche	71
2.5.3	Der Transformationssatz für Doppelintegrale	72
2.5.4	Dreifachintegrale	74
2.5.5	Anwendung und Beispiele	75

Kapitel 1

Lineare Algebra

1.1 Vektorräume

Das Urbild der Vektorräume ist die anschauliche Vektorrechnung in der Ebene \mathbb{R}^2 oder im Raum \mathbb{R}^3 . Die Vektoraddition und die Vektorvervielfachung sind die Grundoperationen eines Vektors. Aus der (komponentenweisen) Vektoraddition $\vec{x} + \vec{y}$ ergeben sich folgende Eigenschaften:

- Assoziativität: $\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$
- Existenz eines Nullvektor: $\vec{0}$
- Existenz des negativen Vektors: $-\vec{x}$ mit $\vec{x} + (-\vec{x}) = \vec{0}$
- Kommutativität: $\vec{x} + \vec{y} = \vec{y} + \vec{x}$

Diese Eigenschaften bilden eine Abelsche Gruppe. Die Vektorvervielfachung $\lambda \cdot \vec{x}$ (Multiplikation mit einem Skalar aus dem Körper \mathbb{R}) ist mit der Vektoraddition und mit den Rechenregeln in \mathbb{R} („Körperaxiome“) verträglich.

$$\mu(\lambda \cdot \vec{x}) = (\mu \cdot \lambda) \vec{x}, \quad \lambda \cdot \vec{x} + \mu \cdot \vec{x} = (\lambda + \mu) \cdot \vec{x}$$

Diese Eigenschaften von Vektoren dienen als Vorbild für die Definition des Vektorraumes.

Definition: (Vektorraum) Ein Vektorraum oder linearer Raum über dem Körper \mathbb{K} mit Verknüpfungen $+$, \cdot besteht aus einer Menge V mit:

1. einer inneren Verknüpfung $+$: $V \times V \rightarrow V$ (x, y) $\mapsto x + y$, so dass $(V, +)$ eine additiv geschriebene abelsche Gruppe bildet und
2. einer äußeren Verknüpfung \cdot : $\mathbb{K} \times V \rightarrow V$ (λ, x) $\mapsto \lambda \cdot x$ so dass folgende Forderungen erfüllt sind:
 - $(\lambda \cdot \mu) x = \lambda(\mu \cdot x)$ [Assoziativgesetz]
 - $(\lambda + \mu) x = \lambda \cdot x + \mu \cdot x$ [1. Distributivgesetz]
 - $\lambda(x + y) = \lambda \cdot x + \lambda \cdot y$ [2. Distributivgesetz]
 - $1 \cdot x = x$ [Einsgesetz]

Die Elemente aus V heißen Vektoren, die Elemente aus \mathbb{K} Skalare. Die Gruppenoperationen $+$ in V nennt man die Vektoraddition, die äußere Verknüpfung \cdot die Vektorvervielfachung mit Skalaren. Beide zusammen werden als die linearen Operationen des Vektorraumes bezeichnet. Das Nullelement 0 in V heißt Nullvektor und ist verschieden von dem Skalar $0 \in K$.

Bemerkung:

- Die Vektorverknüpfung $+$, \cdot sind von den Körperoperationen $+$, \cdot zu unterscheiden. (obwohl gleich bezeichnet) Die Multiplikationszeichen werden meist weggelassen.
- Wir verwenden meist nur die skalaren Körper \mathbb{R} oder \mathbb{C} . Man spricht dann von einem reellen bzw. komplexen Vektorraum.

Beispiele:

- Die arithmetischen VRs \mathbb{K}^n ($= \mathbb{R}^n$ oder \mathbb{C}^n), bestehend aus n-tupeln

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n)^T$$

von Körperelementen. Die linearen Operationen sind komponentenweise erklärt:

$$x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \cdot x = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$$

und machen \mathbb{K}^n wirklich zu einem VR über \mathbb{K} .

- Die Abbildungsräume $\mathbb{K}^X = \text{Abb}(X, \mathbb{K}) = \{f \mid f : X \rightarrow \mathbb{K}\}$, bestehen aus allen Abbildungen einer (nichtleeren) Menge X in dem Körper \mathbb{K} (z. Bsp. den Funktionenraum $\text{Abb}(\mathbb{R}, \mathbb{R})$). Die linearen Operationen sind wertweise definiert:

$$\forall_{x \in X} (f + g)(x) := f(x) + g(x) \text{ für } f, g \in \text{Abb}(X, \mathbb{K})$$

$$\forall_{x \in X} (\lambda \cdot f)(x) := \lambda \cdot f(x) \text{ für } f \in \text{Abb}(X, \mathbb{K}), \lambda \in \mathbb{K}$$

und machen $\text{Abb}(X, \mathbb{K})$ zu einem VR über \mathbb{K} . Der Nullvektor 0 ist die Nullabbildung $0 : X \rightarrow \mathbb{K}, x \mapsto 0$.

Definition: Sei V ein VR über \mathbb{K} . Eine Teilung U von V heißt Unterraum oder linearer Teilraum von V ($U \subset V$), wenn U zusammen mit den auf U eingeschränkten linearen Operationen von V selbst ein VR über \mathbb{K} ist.

Satz 1.1.1 (Unterraumkriterium) *Eine nichtleere Teilmenge U eines VRs V ist genau dann ein UR von V , wenn sie abgeschlossen gegenüber den linearen Operationen von V ist, d.h. wenn gilt:*

1. $x, y \in U \Rightarrow x + y \in U$
2. $\lambda \in \mathbb{K}, x \in U \Rightarrow \lambda \cdot x \in U$
3. (gleichbedeutend mit 1. und 2.) $\lambda, \mu \in \mathbb{K}$ und $x, y \in U \Rightarrow \lambda \cdot x + \mu \cdot y \in U$

Bemerkung: Aus den Kriterien folgt, dass $+, \cdot$ auch Verknüpfungen auf U sind. Die VR-Gesetze sind dann automatisch erfüllt.

Beispiele:

- In einem VR V ist der Nullraum $\{0\}$ der kleinste UR und in jedem anderen UR U von V enthalten: $\{0\} \subset U \subset V$.

- Im Funktionenraum $Abb(\mathbb{R}, \mathbb{R})$ liegen die Unterräume

$$\begin{aligned} Pol(\mathbb{R}) &:= \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ Polynom} \} \\ \mathfrak{D}^1(\mathbb{R}) &:= \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ differenzierbar} \} \\ \ell^0(\mathbb{R}) &:= \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ stetig} \} \end{aligned}$$

mit der Eigenschaft $Pol(\mathbb{R}) \subset \mathfrak{D}^1(\mathbb{R}) \subset \ell^0(\mathbb{R}) \subset Abb(\mathbb{R}, \mathbb{R})$.

Seien jetzt U_1, U_2 zwei URs des VRs V . So gilt:

- $0 \in U_1, 0 \in U_2 \Rightarrow 0 \in U_1 \cap U_2$
- $x, y \in U_1 \cap U_2 \Rightarrow x, y \in U_1 \wedge x, y \in U_2 \Rightarrow x + y \in U_1 \wedge x + y \in U_2 \Rightarrow x + y \in U_1 \cap U_2$
- $\lambda \in \mathbb{K}, x \in U_1 \cap U_2 \Rightarrow \lambda \cdot x \in U_1 \cap U_2$

Dies lässt sich auf beliebige Systeme $\{U_i \subset V \mid i \in I\}$ von Unterräumen verallgemeinern und man erhält:

Satz 1.1.2 *Der Durchschnitt beliebig vieler Unterräume eines VRs V ist wieder ein Unterraum.*

$$\forall_{i \in I} U_i \subset V \Rightarrow \bigcap_{i \in I} U_i \subset V$$

Im VR \mathbb{R}^2 liegen die Unterräume

$$U_1 := \mathbb{R} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} := \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \text{ und}$$

$$U_2 := \mathbb{R} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} := \left\{ \lambda \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}.$$

Es ist

$$\begin{aligned} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\in U_1 \cup U_2 \text{ aber} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin U_1 \cup U_2. \end{aligned}$$

Daraus folgt, dass die Vereinigung von Unterräumen i.a. kein Unterraum ist. Vektorräume können sehr „groß“ sein. Manchmal genügt eine kleine Teilmenge von Vektoren, um alle Elemente eines VRs zu „erzeugen“.

Definition: Sei V ein VR über \mathbb{K}

1. Sind $X_1, \dots, X_n \mid m \in \mathbb{N}$ endlich viele Vektoren aus V , so heißt ein Vektor der Gestalt

$$X = \sum_{i=1}^n \lambda \cdot x_i = \lambda_1 x_1 + \dots + \lambda_n x_n \text{ mit } \lambda_1, \dots, \lambda_n \in \mathbb{K}$$

eine Linearkombination (LK) von x_1, \dots, x_n mit Koeffizienten $\lambda_1, \dots, \lambda_n$.

2. Ist $M \in V$ eine nichtleere Teilmenge von V , so heißt die Menge

$$\langle\langle M \rangle\rangle := \{\lambda_1 \cdot x_1 + \dots + \lambda_m \cdot x_m \mid m \in \mathbb{N}, x \in M, \lambda \in \mathbb{K}\}$$

aller Linearkombinationen von Elementen aus M die lineare Hülle von M oder der von M aufgespannte Unterraum.

3. Ist U ein UR von V und M eine Teilmenge mit $\langle\langle M \rangle\rangle = U$, so heißt M ein Erzeugendensystem von U .

Bemerkung:

- Offensichtlich ist $\langle\langle M \rangle\rangle$ ein UR von V , denn Summe und Vielfache von LK aus Elementen von M sind wieder solche $\langle\langle M \rangle\rangle$. $\langle\langle M \rangle\rangle$ ist der kleinste UR von V der M enthält.
- Bei einer endlichen Teilmenge $M = \{x_1, \dots, x_m\}$ ($\neq \emptyset$) schreibt man $\langle\langle x_1, \dots, x_m \rangle\rangle$ statt $\langle\langle\{x_1, \dots, x_m\}\rangle\rangle$ und nennt das Ergebnis die lineare Hülle der Vektoren x_1, \dots, x_m .
- Es gilt $U \subset V \Rightarrow \langle\langle U \rangle\rangle = U$ und $M' \subset M \subset V \Rightarrow \langle\langle M' \rangle\rangle \subset \langle\langle M \rangle\rangle \subset V$.
- Oft definiert man auch $\langle\langle \emptyset \rangle\rangle := \{0\}$.

Beispiele:

- Im \mathbb{K}^n erzeugen die n „Einheitsvektoren“

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

den ganzen \mathbb{K}^n , d.h. $\mathbb{K}^n = \langle\langle e_1, \dots, e_n \rangle\rangle$, denn jeder Vektor

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \cdot e_1 + x_2 \cdot e_2 + \dots + x_n \cdot e_n$$

ist eine LK der Einheitsvektoren

- Im Fkt. Raum $Abb(\mathbb{R}, \mathbb{R})$ erzeugen die abzählbar vielen Monome $x \in \mathbb{R} \mapsto x_m \in \mathbb{R}$ ($m \in \mathbb{N}_0$) den Unterraum $Pol(\mathbb{R})$ aller Polynome, denn jedes Polynom $x \mapsto Pol(x) = \sum_{k=0}^n a_k x^k$ is LK endlich vieler solcher Polynome.

Definition: Sind U_1, \dots, U_n endlich viele UR eines VRs V , so heißt die lineare Hülle der Vereinigung die Summe der Unterräume und man schreibt:

$$U_1 + \dots + U_n := \langle\langle U_1 \cup \dots \cup U_n \rangle\rangle$$

Bemerkung:

- Es gilt auch $U_1 + \dots + U_n = \{x_1 + \dots + x_n \mid x_1 \in U_1, \dots, x_n \in U_n\}$ denn jede LK $X = \sum_{i=1}^m \lambda_i y_i$ mit $y_1, \dots, y_m \in U_1 \cup \dots \cup U_m$ läßt sich nach geeignetem Umordnen und Zusammenfassen als Summe $x = x_1 + \dots + x_n \mid x_i \in U_i$ schreiben. Umgekehrt ist jede solche Summe eine spezielle LK aus $U_1 + \dots + U_n = \langle\langle U_1 \cup \dots \cup U_n \rangle\rangle$.
- Gilt $U_1 \cap U_2 = \{0\}$, so heißt die Summe $U_1 + U_2$ direkt und man schreibt $U_1 \oplus U_2$.

Bei der Bildung der linearen Hülle einer Menge $M \neq \emptyset$ braucht ein Vektor $x \in M$ nicht berücksichtigt zu werden, wenn er sich als LK $x = \sum_{i=1}^m \lambda_i x_i$ mit anderen Vektoren $x_1, \dots, x_m \in M$ darstellen lässt. Es gilt $\langle\langle M \setminus \{x\} \rangle\rangle = \langle\langle M \rangle\rangle$, denn jede LK von Elementen aus M , die x enthält, läßt sich mittels der LK in eine solche umformen, in der x nicht vorkommt. Wir suchen solche Erzeugendensysteme, die keine überflüssigen Elemente enthalten, also minimal sind.

Definition: Sei V ein VR über \mathbb{K}

1. (Endlich viele) Vektoren x_1, \dots, x_n ($n \in \mathbb{N}$) aus V heißen linear unabhängig, wenn sich der Nullvektor $0 \in V$ nur trivial als LK von x_1, \dots, x_n darstellen läßt.

$$\forall \lambda_1, \dots, \lambda_n \in \mathbb{K} \sum_{i=1}^n \lambda_i x_i = 0 \Rightarrow \lambda_1 = \dots = \lambda_n = 0$$

Sie heißen linear abhängig, wenn es eine nichttriviale Darstellung von $0 \in V$ als LK von x_1, \dots, x_n gibt.

$$\forall \lambda_1, \dots, \lambda_n \in \mathbb{K} \sum_{i=1}^n \lambda_i x_i = 0 \Rightarrow (\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$$

2. Eine nichtleere Teilmenge M von V heißt linear unabhängig, wenn je endlich viele Vektoren $x_1, \dots, x_n \in M$ linear unabhängig sind. Andernfalls ist sie linear abhängig. (zusätzlich sei auch die Leermenge $0 \in V$ linear unabhängig.

Beispiele:

- Im \mathbb{K}^n sind die Einheitsvektoren e_1, \dots, e_n linear unabhängig, denn es gilt

$$\lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_n e_n = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} = 0 \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_m = 0$$

- Im Funktionenraum $Abb(\mathbb{R}, \mathbb{R})$ ist die abzählbar unendliche Menge M aller Monome $x \mapsto x^n$ ($n \in \mathbb{N}_0$) linear unabhängig: Jede endliche LK von Elementen aus M ist ein Polynom $x \mapsto p(x) = \sum_{k=0}^n n_k x^k$ ($n \in \mathbb{N}_0$) und es gilt $P \equiv 0 \Leftrightarrow \forall x \in \mathbb{R} p(x) = 0 \Leftrightarrow \lambda_0 = \lambda_1 = \dots = \lambda_n = 0$ denn Polynome $p \neq 0$ besitzen nur endlich viele Nullstellen.
- Im \mathbb{R}^2 sind die Vektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ linear unabhängig, denn

$$\begin{aligned} \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} \lambda - \mu \\ \lambda + \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\Leftrightarrow \left\{ \begin{array}{l} \lambda - \mu = 0 \\ \lambda + \mu = 0 \end{array} \right\} \Leftrightarrow \lambda = 0 \wedge \mu = 0 \end{aligned}$$

Anders ausgedrückt sind die Vektoren x_1, \dots, x_n ($n \geq 2$) genau dann linear abhängig, wenn sich (mindestens) ein Vektor $x_j \mid 1 \leq j \leq n$ als LK der übrigen darstellen läßt.

Beweis:

- („ \Rightarrow “) x_1, \dots, x_n linear abhängig \Rightarrow es existieren $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ mit $\sum_{i=1}^n \lambda_i x_i = 0$ aber $(\lambda_1, \dots, \lambda_n) = (0, \dots, 0)$. Ist etwa $\lambda_j \neq 0$ so gilt $x_j = -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i x_i = \sum_{i \neq j} \left(-\frac{1}{\lambda_j} \lambda_i\right) x_i$.
- („ \Leftarrow “) Aus $x_i = \sum_{i \neq j} \mu_i x_i$ folgt $\mu_1 x_1 + \dots + (-1)x_j + \dots + \mu_n x_n = 0$

Definition: Eine Teilmenge B eines VRs V heißt Basis von V , wenn sie ein linear unabhängiges Erzeugendensystem für V ist.

$$B \text{ linear unabhängig} \wedge \langle\langle B \rangle\rangle = V$$

Beispiele:

- Im \mathbb{R}^n bildet die Menge $B = \{e_1, \dots, e_n\}$ der Einheitsvektoren eine (endliche) Basis, die sogenannte harmonische oder Standard Basis des \mathbb{R}^n .
- Im Funktionenraum $Abb(\mathbb{R}, \mathbb{R})$ bildet die Menge $B = \{x \mapsto x^n \mid n \in \mathbb{N}_0\}$ einen (abzählbar unendlichen) Raum des URs $Pol(\mathbb{R})$ aller Polynome
- Die Leermenge \emptyset wird als Basis des Nullraums $\{0\}$ aufgefaßt.

Satz 1.1.3

- Wenn ein VR eine unendliche Basis besitzt (z.Bsp. $Pol(\mathbb{R})$) sind alle seine Basen unendlich.
- Wenn ein VR eine endliche Basis besitzt (z.Bsp. \mathbb{K}^n), haben alle seine Basen die gleiche Anzahl von Vektoren (nämlich n)

(Ohne Beweis)

Definition:

1. Besitzt ein VR V eine endliche Basis aus n Vektoren ($n \in \mathbb{N}_0$) so heißt die allen Basis gemeinsame Basislänge die Dimension von V kurz $\dim V = n$.
($\dim \{0\} = 0$)
2. Besitzt ein VR eine unendliche Basis, so heißt er ∞ - dimensional.
($\dim V = \infty$)

Beispiele:

- $\dim \mathbb{R}^n = n$ (es gibt eine endliche Basis)
- $\dim \text{Pol}(\mathbb{R}) = \infty$ (es gibt eine abzählbar unendliche Basis)
- $\dim \text{Abb}(\mathbb{R}, \mathbb{R}) = ?$ (gibt es überhaupt eine Basis?)

Bemerkung:

1. Eine Basis B eines VRs V ist in folgendem Sinne maximal: Nimmt man aus V einen Vektor $x \in V \setminus B$ hinzu, so ist $B' = B \cup \{x\}$ nicht mehr linear unabhängig.
2. Sie ist im folgenden Sinne auch minimal: Nimmt man aus B nur einen Vektor $x \in B$ heraus, so ist $B' = B \setminus \{x\}$ nicht mehr erzeugend.

Satz 1.1.4 *Der VR V besitze eine Basis $B \neq \emptyset$. Dann läßt sich jeder Vektor aus V ($\neq 0$) eindeutig als LK von Basiselementen (ohne Nullkoeffizienten) darstellen.*

Beweis: Wegen $\langle\langle B \rangle\rangle = V$ lässt sich jedes $x \in V$ als LK von Basiselementen darstellen. Seien $x = \sum_{i=1}^m \lambda_i x_i$, $\sum_{k=1}^{\tilde{m}} \tilde{\lambda}_k \tilde{x}_k$ zwei solche Darstellungen mit $x_1, \dots, x_m, \tilde{x}_1, \dots, \tilde{x}_{\tilde{m}} \in B$. Wir können annehmen, dass $x = \sum_{i=1}^m \lambda_i x_i = \sum_{i=1}^m \mu_i x_i$ gilt. (andernfalls umbenennen und auffüllen mit Nullkoeffizienten) Dann folgt $\sum_{i=1}^m (\lambda_i - \mu_i) x_i = 0$. Wegen der linearen Unabhängigkeit von B gilt somit $\forall_{i=1}^m \lambda_i = \mu_i$. Die Darstellungen stimmen überein.

Spezialfall: Besitzt V eine endliche Basis $B = \{a_1, \dots, a_m\}$, so lässt sich jeder Vektor $x \in V$ in der Form $x = \lambda_1 a_1 + \dots + \lambda_m a_m$ mit eindeutig bestimmten Skalaren $\lambda_1, \dots, \lambda_m$ darstellen.

Beispiele:

- Im \mathbb{R}^2 ist

$$\{a_1, a_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

eine Basis. Für jeden Vektor $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ gilt

$$\lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda + \mu \\ \mu \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$
$$\Leftrightarrow \mu = y \wedge \lambda = x - y.$$

$\begin{pmatrix} x \\ y \end{pmatrix}$ besitzt also die Basisdarstellung

$$\begin{pmatrix} x \\ y \end{pmatrix} = (x - y) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

mit dem Koordinatenvektor

$$\begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} x - y \\ y \end{pmatrix}.$$

Satz 1.1.5 *Jeder VR besitzt eine Basis.*

Der Beweis ist für den allgemeinen Fall relativ aufwendig. Daher wird hier darauf verzichtet.

Folgerung 1.1.6 *Jeder endlich erzeugte VR V besitzt eine endliche Basis und damit eine Dimension $n = \dim V$, $n \in \mathbb{N}_0$. Diese gibt die Maximalanzahl linear unabhängiger Vektoren in V und die Minimalanzahl erzeugender Vektoren von V an.*

1.2 Basis- und Dimensionssätze im endlich dimensionalen VR

Satz 1.2.1 In einem VR V mit $\dim V = n < \infty$ bilden je n linear unabhängige Vektoren und auch je n erzeugenden Vektoren bereits eine Basis von V .

$$\begin{aligned} a_1, \dots, a_n \in V \text{ lin.unabh.} &\Leftrightarrow \langle\langle a_1, \dots, a_n \rangle\rangle = V \\ &\Leftrightarrow \{a_1, \dots, a_n\} \text{ Basis von } V \end{aligned}$$

Beweis:

- („ \Rightarrow “)

$B = \{a_1, \dots, a_n\}$ linear unabhängig $\Rightarrow \forall x \in V a_1, \dots, a_n, x$ lin. abh.

$$\Rightarrow \forall x \in V x \in \langle\langle a_1, \dots, a_n \rangle\rangle \Rightarrow \langle\langle a_1, \dots, a_n \rangle\rangle = V$$

- („ \Leftarrow “) Wäre $\langle\langle a_1, \dots, a_n \rangle\rangle = V$ und a_1, \dots, a_n linear abhängig, wäre mindestens ein a_j überflüssig und es gäbe weniger als n Vektoren, die V erzeugen. Widerspruch zu Satz 1.1.6.

In einer Basis $\{a_1, \dots, a_n\}$ kann man einen Vektor a_j gegen einen „schöneren“ Vektor $b \neq 0$ austauschen. Voraussetzung ist, dass in der Basisdarstellung $b = \sum_{i=1}^n \lambda_i a_i$ die j -te Koordinate λ_j nicht verschwindet. Denn dann gilt $a_j = \frac{1}{\lambda_j} \left(b - \sum_{i \neq j} \lambda_i a_i \right)$ und in jeder LK von a_1, \dots, a_n kann a_j durch b (und die übrigen a_i) ersetzt werden. Also ist $\langle\langle a_1, b, \dots, a_n \rangle\rangle = \langle\langle a_1, \dots, a_n \rangle\rangle = V$ und $\{a_1, b, \dots, a_n\}$ ebenfalls eine Basis von V . Dieses Verfahren kann auch auf mehrere linear unabhängige Vektoren angewandt werden.

Satz 1.2.2 (Austauschsatz von STEINITZ) Seien $\{a_1, \dots, a_n\}$ eine Basis des VR V und b_1, \dots, b_m linear unabhängige Vektoren mit $m < n$. Dann gibt es $n - m$ Vektoren $a_{i_{m+1}}, \dots, a_{i_n} \in \{a_1, \dots, a_n\}$, so dass auch $\{b_1, \dots, b_m, a_{i_{m+1}}, \dots, a_{i_n}\}$ eine Basis von V ist. Es können also m geeignete von den Basisvektoren a_1, \dots, a_n gegen die Vektoren b_1, \dots, b_m ausgetauscht werden.

Satz 1.2.3 (Basisergänzungssatz) In einem VR V mit $\dim V = n < \infty$ lassen sich je m linear unabhängige Vektoren b_1, \dots, b_m ($m < n$) zu einer Basis $\{b_1, \dots, b_m, b_{m+1}, \dots, b_n\}$ von V ergänzen.

Satz 1.2.4 In einem VR mit $\dim V < \infty$ gilt

1. $U \subset V \rightarrow \dim U \leq \dim V$
2. $U \subset V \rightarrow \dim U = \dim V \Rightarrow U = V$ (gilt nicht in ∞ -dimensionalen VRs)

Satz 1.2.5 (Dimensionsformel für Unterräume) Für endlich dimensionale UR U, U' eines VRs V gilt

$$\dim(U + U') = \dim U + \dim U' - \dim(U \cap U')$$

Im \mathbb{R}^3 sind die 1-dimensionalen UR Geraden durch den Nullpunkt und 2-dimensionale UR Ebenen durch den Nullpunkt.

Beispiele:

- Sind $E_1, E_2 \in \mathbb{R}^3$ zwei Ebenen mit der Schnittgeraden $G = E_1 \cap E_2$ so folgt

$$\dim(E_1 + E_2) = \dim E_1 + \dim E_2 - \dim G = 2 + 2 - 1 = 3.$$

Jeder Vektor $y \in \mathbb{R}^3$ läßt sich als Summe $y = y_1 + y_2$ mit $y_1 \in E_1, y_2 \in E_2$ darstellen. Diese Darstellung ist jedoch nicht eindeutig, denn mit einem beliebigen Vektor $u \in G = E_1 \cap E_2, u \neq 0$ gilt auch

$$y = (y_1 + u) + (y_2 - u) = \tilde{y}_1 + \tilde{y}_2 \text{ mit } \tilde{y}_1 \in E_1, \tilde{y}_2 \in E_2.$$

- Eine Gerade G schneide die Ebene E im Nullpunkt 0. D.h. $E \cap G = \{0\}$. Dann gilt

$$\dim(G + E) = \dim G + \dim E + \dim \{0\} = 1 + 2 - 0 = 3.$$

D.h. $G \oplus E = \mathbb{R}^3$. Jeder Vektor $y \in \mathbb{R}^3$ läßt sich also durch die Summe $y = y_1 + y_2$ mit $y_1 \in G, y_2 \in E$ eindeutig darstellen, da die Summe direkt ist.

Bemerkung: Geraden, Ebenen, usw. im \mathbb{R}^n , die nicht durch den Nullpunkt gehen sind keine (Vektor-)Unterräume, sondern sogenannte affine Unterräume und besitzen eine Darstellung $A = x_0 + U$ mit einem Verschiebungsvektor x_0 und einem (Vektor-)Unterraum U mit $0 \in U$.

1.3 Lineare Abbildungen

Definition: V, W seien VR über den selben Körper \mathbb{K} . Eine Abbildung $f : V \rightarrow W$ heißt linear (oder ein linearer Operator), wenn sie mit den linearen Operationen in V und U verträglich ist.

1. $\forall x, y \in V f(x + y) = f(x) + f(y)$
2. $\forall x \in V \forall \lambda \in \mathbb{K} f(\lambda \cdot x) = \lambda \cdot f(x)$

Diese Aussagen können auch zu

$$\forall x, y \in V \forall \lambda, \mu \in \mathbb{K} f(\lambda \cdot x + \mu \cdot y) = \lambda \cdot f(x) + \mu \cdot f(y)$$

zusammengefasst werden.

Beispiele:

- Die Nullabbildung $0 : V \rightarrow W, x \mapsto 0$ ist linear.
- Die Identität $id V : V \rightarrow V, x \mapsto x$ ist linear.
- Die Abbildung $f_\alpha : V \rightarrow V, x \mapsto f_\alpha(x) := \alpha x$ mit $\alpha \in \mathbb{K}$ ist linear. (Streckungen bzw. Homothetien)
- Die Abbildung $f_a : V \rightarrow V, x \mapsto f_a(x) := a + x$ mit $a \in V$ ist für $a \neq 0$ nicht linear. (Translationen)
- Die Abbildung $D : Pol(\mathbb{R}) \rightarrow Pol(\mathbb{R}), p \mapsto D(p) := p'$ ist linear. (siehe Summenregel, Konstanten bei Differentialrechnung)

Eine lineare Abbildung $f : V \rightarrow W$ wird auch als Homomorphismus bezeichnet. Ist diese zusätzlich bijektiv, so nennt man sie Isomorphismus. Eine lineare Selbstabbildung $f : V \rightarrow V$ ist ein Endomorphismus und eine bijektive Abbildung $f : V \rightarrow V$ ein Automorphismus.

Satz 1.3.1

1. $f : V \rightarrow W, g : W \rightarrow X$ linear $\Rightarrow g \circ f : V \rightarrow X$ linear
2. $f : V \rightarrow W$ Isomorphismus $\Rightarrow f^{-1} : W \rightarrow V$ Isomorphismus

Beweis:

1. $(g \circ f)(\lambda x + \mu y) = g(f(\lambda x + \mu y)) = g(\lambda f(x) + \mu f(y))$
 $\Rightarrow \lambda g(f(x)) + \mu g(f(y)) = \lambda (g \circ f)(x) + \mu (g \circ f)(y)$
2. $f(\lambda f^{-1}(x) + \mu f^{-1}(y)) = \lambda f(f^{-1}(x)) + \mu f(f^{-1}(y)) = \lambda x + \mu y$
 $\Rightarrow f^{-1}(\lambda x + \mu y) = f^{-1}(f(\lambda f^{-1}(x) + \mu f^{-1}(y))) = \lambda f^{-1}(x) + \mu f^{-1}(y)$

Satz 1.3.2 Ist B eine Basis des VR V und $f_0 : B \rightarrow W$ eine beliebige Abbildung in der VR W so gibt es genau eine lineare Abbildung $f : V \rightarrow W$ mit $\forall b \in B f(b) = f_0(b)$. Diese Abbildung wird lineare Fortsetzung von f_0 auf V genannt.

Beweis: (für den Fall, dass V eine endliche Basis $B = \{a_1, \dots, a_n\}$ besitzt) Zu zeigen ist, dass es zu beliebigen Vektoren $b_1, \dots, b_n \in W$ genau eine lineare Abbildung $f : V \rightarrow W$ mit $f(a_1) = b_1, \dots, f(a_n) = b_n$ gibt: Da jedes $x \in V$ eine eindeutige Basisdarstellung $x = \sum_{i=1}^n \lambda_i a_i$ besitzt, muss gelten:

$$f(x) = f\left(\sum_{i=1}^n \lambda_i a_i\right) = \sum_{i=1}^n \lambda_i f(a_i) = \sum_{i=1}^n \lambda_i b_i.$$

Die so definierte Abbildung $f : V \rightarrow W$ ist wirklich linear, wie man einfach nachrechnen kann.

Satz 1.3.3 Für eine lineare Abbildung $f : V \rightarrow W$ gilt:

1. $M \subset V \Rightarrow f[\langle\langle M \rangle\rangle] = \langle\langle f[M] \rangle\rangle$
2. $U \subset V \Rightarrow f[U] \subset W$
3. $Z \subset W \Rightarrow f^{-1}[Z] \subset V$
4. $M \subset V$ linear abhängig $\Rightarrow f[M]$ linear abhängig.
(aber $M \subset V$ linear unabhängig $\not\Rightarrow f[M]$ linear unabhängig)

Beweis: Sei $f : V \rightarrow W$

1. $f[\langle\langle M \rangle\rangle] = \langle\langle f[M] \rangle\rangle$

(a)

$$a = \sum_{i=1}^m \lambda_i x_i \in \langle\langle M \rangle\rangle \text{ mit } x_i \in M$$

$$\rightarrow f(x) = f\left(\sum_{i=1}^m \lambda_i x_i\right) = \sum_{i=1}^m \lambda_i f(x_i) \in \langle\langle f[M] \rangle\rangle$$

da $f(x_i) \in f[M]$. Also gilt $f[\langle\langle M \rangle\rangle] \subset \langle\langle f[M] \rangle\rangle$.

(b)

$$y = \sum_{i=1}^m \lambda_i y_i \in \langle\langle f[M] \rangle\rangle \text{ mit } y_i \in f[M].$$

Dann gibt es $x_1, \dots, x_m \in M$ mit $f(x_i) = y_i$ und es gilt

$$y = \sum_{i=1}^m \lambda_i f(x_i) = f\left(\sum_{i=1}^m \lambda_i x_i\right) \in f[\langle\langle M \rangle\rangle].$$

Also ist $\langle\langle f[M] \rangle\rangle \subset f[\langle\langle M \rangle\rangle]$.

2. $U \subset V \Rightarrow \langle\langle U \rangle\rangle = U \Rightarrow f[U] = f[\langle\langle U \rangle\rangle] = \langle\langle f[U] \rangle\rangle \Rightarrow f[U] \subset W$
3. $f^{-1}[Z] = \{x \in V \mid f(x) \in Z\} \subset V$, wenn $Z \subset W$. Beweis mit UR-Kriterium.

(a) $f^{-1}[Z] \neq \emptyset$ denn $0 \in f^{-1}[Z]$, da $f(0) = 0 \in Z$

(b) $x, y \in f^{-1}[Z] \Rightarrow f(x), f(y) \in Z$
 $\Rightarrow \lambda f(x) + \mu f(y) = f(\lambda x + \mu y) \in Z$
 $\Rightarrow \lambda x + \mu y \in f^{-1}[Z]$

4. Da M linear abhängig ist existieren $x_1, \dots, x_m \in M$ mit

$$\sum_{i=1}^m \lambda_i x_i = 0, \text{ aber } (\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0).$$

Für die Bilder $f(x_i) =: y_i \in f[M]$ gilt dann ebenfalls

$$\sum_{i=1}^m \lambda_i y_i = \sum_{i=1}^m \lambda_i f(x_i) = f\left(\sum_{i=1}^m \lambda_i x_i\right) = 0$$

aber $(\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$. Also ist auch $f[M]$ linear abhängig.

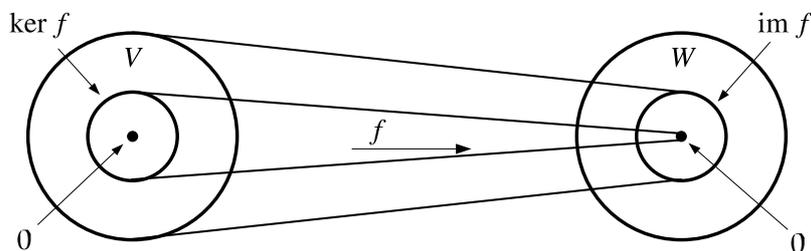


Abbildung 1.1: Kern(\ker) und Bild(im) einer linearen Abbildung(f)

Definition: (siehe Abbildung 1.1) Bei einer linearen Abbildung $f : V \rightarrow W$ heißt

1. Das Urbild des Nullraums $\{0\}$ unter f der Kern(\ker) von f und seine Dimension der Defekt(def) von f . $\ker f = f^{-1}[\{0\}] = \{x \in V \mid f(x) = 0\}$, $\text{def } f = \dim \ker f$
2. Das Bild von V selbst das Bild(im) von f und seine Dimension der Rang(rg) von f . $\text{im } f = f[V]$, $\text{rg } f = \dim f[V]$

Satz 1.3.4 Eine lineare Abbildung $f : V \Rightarrow W$ ist

1. genau dann injektiv, wenn $\ker f = \{0\}$.
2. genau dann surjektiv, wenn $\operatorname{im} f = W$.

Beweis:

1. („ \Rightarrow “): $x \in \ker f \Rightarrow f(x) = 0 = f(0) \Rightarrow x = 0$
 („ \Leftarrow “): $f(x) = f(y) \Rightarrow f(x) - f(y) = f(x - y) = 0$ da $\ker f = \{0\}$
 $\Rightarrow x - y = 0 \Rightarrow x = y$
2. gilt immer

Satz 1.3.5 Bei einer linearen Abbildung $f : V \Rightarrow W$ mit $\dim V < \infty$ gilt $\operatorname{def} f + \operatorname{rg} f = \dim V$.

Beweis: $0 \leq k := \operatorname{def} f \leq n := \dim V$ und eine Basis $\{b_1, \dots, b_k\}$ von $\ker f$ lässt sich zu einer Basis $\{b_1, \dots, b_k, b_{k+1}, \dots, b_n\}$ ergänzen. Wir zeigen, dass dann $B' = \{f(b_{k+1}), \dots, f(b_n)\}$ eine Basis von $\operatorname{im} f = f[V]$ ist:

1. $\langle\langle B' \rangle\rangle = f[V]$, denn $V = \langle\langle B \rangle\rangle = f[V] = f[\langle\langle B \rangle\rangle] = \langle\langle f[B] \rangle\rangle$
 $= \langle\langle f(b_1), \dots, f(b_k), f(b_{k+1}), \dots, f(b_n) \rangle\rangle$
 (mit $f(b_1), \dots, f(b_k) = 0$) $= \langle\langle f(b_{k+1}), \dots, f(b_n) \rangle\rangle = \langle\langle B' \rangle\rangle$
2. B' ist linear unabhängig, denn

$$\begin{aligned} \sum_{i=k+1}^n \lambda_i f(b_i) = f\left(\sum_{i=k+1}^n \lambda_i b_i\right) = 0 &\Rightarrow \sum_{i=k+1}^n \lambda_i b_i \in \ker f = \langle\langle b_1, \dots, b_k \rangle\rangle \\ &\Rightarrow \sum_{i=k+1}^n \lambda_i b_i = \sum_{i=1}^k (-\lambda_i b_i) \Rightarrow \sum_{i=1}^n \lambda_i b_i = 0 \\ &\Rightarrow (\lambda_1 = \dots = \lambda_k = 0) \lambda_{k+1} = \dots = \lambda_n = 0 \end{aligned}$$

Aus der Dimensionsformel und Satz 1.3.4 folgt

Satz 1.3.6

- Für eine lineare Abbildung $f : V \rightarrow W$ mit $\dim V, \dim W < \infty$ gilt
 - f injektiv $\Leftrightarrow \operatorname{rg} f = \dim V$
 - f surjektiv $\Leftrightarrow \operatorname{rg} f = \dim W$
 - f bijektiv (Isomorphismus) $\Leftrightarrow \dim V = \operatorname{rg} f = \dim W$
- Bei einer linearen Abbildung $f : V \rightarrow W$ mit $\dim V = \dim W < \infty$ gilt
 - f injektiv $\Leftrightarrow f$ surjektiv ($\Leftrightarrow f$ bijektiv)

Dies gilt insbesondere bei Endomorphismen.

Zum Abschluss (ohne Beweis) einige algebraische Eigenschaften von Mengen von linearen Abbildungen.

Satz 1.3.7 V, W seien VR über \mathbb{K}

1. Die Menge $L(V, W) := \operatorname{Hom}(V, W) := \{f : V \rightarrow W \mid f \text{ linear}\}$ aller linearen Abbildungen von V nach W bildet zusammen mit der wertweise definierten Addition $f + g$ und Vervielfachung $\lambda \cdot f$ einen Vektorraum über \mathbb{K} .
2. Zusätzlich bildet die Menge $\operatorname{End}(V) = L(V, V)$ aller Endomorphismen von V zusammen auf der Addition $f + g$ und der Komposition $g \circ f$ als Multiplikation einen Ring mit Einselement id_V , den sogenannten Endomorphismenring von V .
3. Die Menge $\operatorname{Gl}(V)$ aller Automorphismen bildet zusammen mit der Komposition \circ sogar eine (i.a.) nicht abelsche Gruppe, Automorphismengruppe oder allgemeine lineare Gruppe von V („General Linear Group“) genannt.

Bedeutung: Man kann mit linearen Abbildungen vernünftig rechnen.

Im endlich dimensionalen VR V ist es oft üblich, Basen $B = \{a_1, \dots, a_n\} \subset V$ als Basis n -Tupel $(a_1, \dots, a_n) \in V^n$ zu schreiben um die Reihenfolge der Basisvektoren festzulegen und man spricht dann von einer geordneten Basis. Sei V ein fester VR mit $n = \dim V < \infty$ über $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} . Ist $B = (a_1, \dots, a_n)$ eine geordnete Basis, so besitzt jeder Vektor $x \in V$ eine eindeutige Basisdarstellung $x = \sum_{i=1}^n \xi_i a_i$ mit Koeffizienten („Koordinaten“) $\xi_1, \dots, \xi_n \in \mathbb{K}$ deren Reihenfolge festgelegt ist. Dies liefert eine Abbildung

$$\Phi_B : V \rightarrow \mathbb{K}^n, x = \sum_{i=1}^n \xi_i a_i \mapsto \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n$$

von V in den Koordinatenraum \mathbb{K}^n , die bijektiv ist, da die Umkehrabbildung $\Phi_B^{-1} : \mathbb{K}^n \rightarrow V$, $\xi \mapsto x \sum_{i=1}^n \xi a_i$ linear ist. Wegen

$$x = \sum_{i=1}^n \xi_i a_i, \quad y = \sum_{i=1}^n \eta_i a_i$$

$$\Rightarrow \lambda x + \mu y = \sum_{i=1}^n (\lambda \xi_i + \mu \eta_i) a_i \Phi_B(\lambda x + \mu y) = \lambda \xi + \mu \eta = \lambda \Phi_B(x) + \mu \Phi_B(y)$$

ist auch Φ_B linear, also ein Isomorphismus.

Satz 1.3.8 *Jeder n -dim VR V über \mathbb{K} mit $n \in \mathbb{N}_0$ ist isomorph zu seinem Koordinatenraum \mathbb{K}^n . Jede geordnete Basis $B = (a_1, \dots, a_n)$ von V liefert einen Isomorphismus $\Phi_B : V \rightarrow \mathbb{K}^n$.*

Bedeutung: Bei endlich dimensionalen VR über K und linearen Abbildungen zwischen ihnen kann man sich auf \mathbb{K}^n beschränken und in Koordinaten rechnen.

1.4 Matrizen

Wir betrachten im folgenden lineare Abbildungen

$$f : \mathbb{K}^n \rightarrow \mathbb{K}^m, x \mapsto f(x) \quad (n, m \in \mathbb{N}).$$

Diese linearen Abbildungen sind eindeutig durch die Bilder $a_k := f(e_k) \in \mathbb{K}^m$ der harmonischen Basisvektoren $e_1, \dots, e_n \in \mathbb{K}^n$ definiert. Schreibt man diese Bilder als Spaltenvektor in der Form

$$a_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix}$$

so gilt für alle

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n x_i e_i \in \mathbb{K}^n; \quad y = f(x) = \sum_{k=1}^n x_k f(e_k) = \sum_{k=1}^n x_k a_k$$

d.h.

$$y_i = \sum_{k=1}^n a_{ik} x_k = a_{i1} x_1 + a_{i2} x_2 + \dots + a_{in} x_n \quad \text{mit } a_{11}, \dots, a_{mn}$$

Das Ergebnis kann man mit Hilfe der sogenannten Matrizenmultiplikation (hier Matrix x Vektor) berechnen:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad (\text{„Zeile mal Spalte“})$$

$$y = A \cdot x$$

(Kurzschreibweise für die Operation $x \in \mathbb{K}^n \mapsto y \in \mathbb{K}^m$) mit der (i.a. rechteckigen) $(m \times n)$ - Matrix

$$A = (a_1, \dots, a_n) = (a_{ik})_{\substack{i=1, \dots, m \\ k=1, \dots, n}} = (a_{ik})_{mn}$$

bestehend aus n Spalten (Vektoren) und m Zeilen. Vereinbarung bei den Matrixelementen a_{ik} : Erster Index ist der Zeilenindex und zweiter Index der Spaltenindex.

Beispiele:

- Von einer linearen Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ sei bekannt:

$$a_1 := f(e_1) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$a_2 := f(e_2) = f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

Dann gilt für alle $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ mit $A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$:

$$y = f(x) = A \cdot x = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$f\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 4 \cdot 1 \\ 2 \cdot 2 + 5 \cdot 1 \\ 3 \cdot 2 + 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ 12 \end{pmatrix}$$

Satz 1.4.1 Zu jeder linearen Abbildung $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ ($n, m \in \mathbb{N}$) gibt es genau eine $(m \times n)$ -Matrix $A = (a_1, \dots, a_n) = (a_{ik})_{mn}$ aus Körperelementen, so dass gilt:

$$\forall x \in \mathbb{K}^n \quad f(x) = A \cdot x$$

Diese wird als Darstellungsmatrix von f bezeichnet.

Ihre Spaltenvektoren $a_1, \dots, a_n \in \mathbb{K}^m$ sind die Bilder der kanonischen Basisvektoren $e_1, \dots, e_n \in \mathbb{K}^n$. Umgekehrt bestimmt jede $(m \times n)$ -Matrix A genau eine lineare Abbildung $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$.

Es gibt also eine (injektive) Zuordnung

$$\psi : L(\mathbb{K}^n, \mathbb{K}^m) \rightarrow M(m, n; \mathbb{K}), \quad f \mapsto A = (f(e_1), \dots, f(e_n))$$

zwischen der Menge aller linearen Abbildungen $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ und der Menge aller $(m \times n)$ -Matrizen über \mathbb{K} . Wir übertragen jetzt mit Hilfe dieser Bijektion die für lineare Abbildungen f, g erklärten Operationen $f + g, \lambda \cdot f, g \circ f$ auf die zugehörige Darstellungsmatrizen A, B und definierten damit Matrizenoperationen $A + B, \lambda \cdot A, B \circ A$.

1. **Matrizenaddition:** Sei $A = (a_1, \dots, a_n), B = (b_1, \dots, b_n) \in M(m, n; \mathbb{K})$

$$\left. \begin{array}{l} f(e_k) = a_k \\ f(e_k) = b_k \end{array} \right\} (f + g)(e_k) = f(e_k) + g(e_k) = a_k + b_k$$

$$A + B = (a_1 + b_1, \dots, a_n + b_n) = (a_{ik} + b_{ik})_{mn}$$

2. **Matrizenvervielfachung:** Sei $A = (a_1, \dots, a_n) \in M(m, n; \mathbb{K}), \lambda \in \mathbb{K}$

$$f(e_k) = a_k \Rightarrow (\lambda \cdot f)(e_k) = \lambda \cdot f(e_k) = \lambda \cdot a_k$$

$$\lambda \cdot A = (\lambda a_1, \dots, \lambda a_n) = (\lambda a_{ik})_{mn}$$

3. **Matrizenmultiplikation:** Sei

$$A = (a_{jk})_{mn} \in M(m, n; \mathbb{K}), \quad B = (b_{ij})_{mn} \in M(p, m; \mathbb{K})$$

und die zugehörigen linearen Abbildungen $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ und $g : \mathbb{K}^m \rightarrow \mathbb{K}^p$. Gesucht ist die Darstellungsmatrix $B \cdot A = C = (c_{ik})_{pm}$ von $g \circ f : \mathbb{K}^n \rightarrow \mathbb{K}^p$.

$$\left. \begin{aligned} y = f(x) = A \cdot x &\Leftrightarrow y_j = \sum_{k=1}^n a_{jk} x_k \quad (j = 1, \dots, m) \\ z = g(y) = B \cdot y &\Leftrightarrow z_i = \sum_{j=1}^m b_{ij} y_j \quad (i = 1, \dots, p) \end{aligned} \right\} \Rightarrow$$

$$z_i = \sum_{j=1}^m b_{ij} \sum_{k=1}^n a_{jk} x_k = \sum_{k=1}^n \left(\sum_{j=1}^m b_{ij} a_{jk} \right) x_k$$

Andererseits gilt

$$z = g \circ f(x) = C \cdot x \Leftrightarrow z_i = \sum_{k=1}^n c_{ik} x_k$$

Vergleich der beiden Gleichungen liefert

$$c_{ik} = \sum_{j=1}^m b_{ij} a_{jk} = b_{i1} a_{1k} + \dots + b_{im} a_{mk}$$

oder mit Hilfe der Matrizenmultiplikation („Zeile mal Spalte“)

$$\begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{p1} & \cdots & c_{pn} \end{pmatrix} = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{p1} & \cdots & b_{pm} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Das so definierte Matrizenprodukt $C = B \cdot A$ stellt also die Verkettung $g \circ f$ dar.

Satz 1.4.2 Die Darstellungsmatrix C einer Verkettung $g \circ f : \mathbb{K}^n \rightarrow \mathbb{K}^p$ linearer Abbildungen $F : \mathbb{K}^n \rightarrow \mathbb{K}^m$ und $q : \mathbb{K}^m \rightarrow \mathbb{K}^p$ mit Darstellungsmatrizen A und B ist das Matrizenprodukt $B \cdot A$ definiert durch

$$c_{ik} = \sum_{j=1}^m b_{ij} a_{jk}$$

Für diese gelten die selben Rechengesetze wie für die Verkettung linearer Abbildungen. Insbesondere ist es i.a. nicht kommutativ.

Beispiele:

•

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix}$$

[2 Zeilen] · [2 Spalten] = [2 × 2 - Matrix]

Spezialfälle:

1. Betrachtet man nur quadratische Matrizen $A, B \in M(n, n; \mathbb{K})$, also Darstellungsmatrizen von Endomorphismen $f, g: \mathbb{K}^n \rightarrow \mathbb{K}^n$, so definieren die Operationen $A + B$ und $B \cdot A$ eine Ringstruktur auf $M(n, n; \mathbb{K})$ "isomorph" zur Ringstruktur von $\text{End}(\mathbb{K}^n)$. Die Identität $\text{id}: \mathbb{K}^n \rightarrow \mathbb{K}^n$ wird durch die Einheitsmatrix $E = (e_1, \dots, e_n)$ mit $e \cdot x = x$ für alle x repräsentiert. Es gelten also die üblichen Ringsätze (z.Bsp. $A \cdot (B + C) = A \cdot B + A \cdot C$) für quadratische Matrizen.
2. Automorphismen $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$ für die eine Umkehrabbildung $f^{-1}: \mathbb{K}^n \rightarrow \mathbb{K}^n$ existiert, werden durch sogenannte reguläre Matrizen $A \in M(n, n; \mathbb{K})$, für die eine inverse Matrix $A^{-1} \in M(n, n; \mathbb{K})$ mit $A \cdot A^{-1} = A^{-1} \cdot A = E$ existiert, repräsentiert. ($f \circ f^{-1} = f^{-1} \circ f = \text{id}$)

Satz 1.4.3

1. Die Menge $M(m, n; \mathbb{K})$ aller $(m \times n)$ -Matrizen über \mathbb{K} bildet zusammen mit der elementweise definierten Addition $A + B$ und Vervielfachung $\lambda \cdot A$ einen Vektorraum über \mathbb{K} mit „Nullvektor“

$$0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \quad (= \text{Nullmatrix}).$$

2. Zusätzlich bildet die Menge $M(n, n; \mathbb{K})$ aller quadratischen $(n \times n)$ -Matrizen zusammen mit der Addition $A + B$ und der Multiplikation $B \cdot A$ einen Ring mit Einselement E (= Einheitsmatrix).
3. Die Menge $GL(n; \mathbb{K})$ aller regulären $(n \times n)$ -Matrizen (für die ein inverses A^{-1} existiert) bildet zusammen mit der Multiplikation sogar eine (i.a. nicht abelsche) Gruppe, die Allgemeine lineare Gruppe über \mathbb{K} .

Daher kann man mit Matrizen wie gewohnt rechnen. Aber im allgemeinen gilt $A \cdot B \neq B \cdot A$ und nicht jede Matrix A besitzt auch eine inverse Matrix A^{-1} .

Definition: Der Rang einer Matrix $A \in M(m, n; \mathbb{K})$ sei der Rang der untergeordneten linearen Abbildung $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $x \mapsto f(x) = A \cdot x$.

Satz 1.4.4

1. Der Rang $rg A$ einer Matrix $A \in M(m, n; \mathbb{K})$ ist gleich der Maximalanzahl linear unabhängiger Spaltenvektoren. (auch als Spaltenrang bezeichnet)
2. Eine Matrix $A \in M(n, n; \mathbb{K})$ ist genau dann regulär (also invertierbar), wenn $rg A = n$ maximal ist, d.h. wenn ihre Spaltenvektoren linear unabhängig sind.

Beispiele:

- $rg \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = 1$

Beweis:

1. Für $A = (a_1, \dots, a_n)$ gilt:

$$\begin{aligned} rg A = rg f &= \dim f[\langle e_1, \dots, e_n \rangle] = \dim \langle f(e_1), \dots, f(e_n) \rangle \\ &= \dim \langle a_1, \dots, a_n \rangle \end{aligned}$$

= maximale Anzahl linear unabhängiger Vektoren aus $\{a_1, \dots, a_n\}$

2. A regulär $\Leftrightarrow f$ Automorphismus $\Leftrightarrow rg A = rg f = \dim \langle a_1, \dots, a_n \rangle = n$
 $\Leftrightarrow a_1, \dots, a_n$ linear unabhängig

Satz 1.4.5 Der Rang einer Matrix bleibt bei Matrixmultiplikationen mit regulären Matrizen von links oder rechts unverändert.

$$A \text{ regulär} \Rightarrow \begin{cases} rg(B \cdot A) = rg B \\ rg(A \cdot B) = rg B \end{cases}$$

Definition: Die durch Vertauschen von Zeilen und Spalten einer $(m \times n)$ -Matrix $A = (a_{ik})_{mn}$ entstehende $(n \times m)$ -Matrix $A^T = (a_{ki}^T)_{nm}$ heißt die zu A transponierte Matrix.

Bemerkung: Bei diesem Vertauschen bleibt die Hauptdiagonale $\{a_{ii} \mid i = 1, \dots, \min\{m, n\}\}$ unverändert und A^T entsteht aus A durch Spiegeln an der Hauptdiagonalen.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Satz 1.4.6

1. $(A^T)^T = A$, $(A + B)^T = A^T + B^T$, $(\lambda \cdot A)^T = \lambda \cdot A^T$
2. $(B \cdot A)^T = A^T \cdot B^T$
3. $(A^{-1})^T = (A^T)^{-1}$, insbesondere gilt A regulär $\Leftrightarrow A^T$ regulär

Beweis:

1. klar
2. Für $C = B \cdot A$ gilt

$$c_{ij} = \sum_{k=1}^m b_{ik} a_{kj} \Rightarrow c_{ki}^T = \sum_{j=1}^m b_{ji}^T a_{kj}^T = \sum_{j=1}^m a_{kj}^T b_{ji}^T$$

$$\Rightarrow C^T = A^T \cdot B^T$$

$$3. A \cdot A^{-1} = E \Rightarrow (A^{-1})^T \cdot A^T = E^T = E \Rightarrow (A^T)^{-1} = (A^{-1})^T$$

Später können wir außerdem sagen, dass beim Transponieren auch der Rang erhalten bleibt:

$$rg(A^T) = rg A$$

Definition: Eine quadratische Matrix $A \in M(n, n; \mathbb{K})$ heißt symmetrisch, wenn $A = A^T$ und schiefsymmetrisch wenn $A = -A^T$

Beispiele:

- $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ ist symmetrisch $\begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$ ist schiefsymmetrisch
(Hauptdiagonale egal) (Hauptdiagonale immer 0)

Bei einer komplexen Matrix $A = (a_{ik})_{mn}$ kann man die konjugiert komplexe Matrix $\bar{A} = (\bar{a}_{ik})_{mn}$ definieren und nennt die Matrix hermitesch wenn $A = \bar{A}^T$ und schieferhermitesch wenn $A = -\bar{A}^T$

Sei $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ eine lineare Abbildung mit Darstellungsmatrix A . Wählt man in \mathbb{K}^n eine andere Basis $B = (b_1, \dots, b_n)$ und auch in \mathbb{K}^m eine andere Basis $B' = (b'_1, \dots, b'_m)$ statt der Einheitsbasen, so erhalten die Vektoren $x \in \mathbb{K}^n$ und $y = f(x) \in \mathbb{K}^m$ andere Koordinaten $\xi = \Phi_B(x)$ und $\eta = \Phi_{B'}(y)$. (siehe Satz ??)

$$\begin{array}{ccc} x \in \mathbb{K}^n & \xrightarrow[A]{} & y = f(x) \in \mathbb{K}^m \\ \Phi_B \downarrow T & & S \downarrow \Phi_{B'} \\ \xi \in \mathbb{K}^n & \xrightarrow{A'} & \eta \in \mathbb{K}^m \end{array}$$

Die lineare Abbildung $\Phi_B : \mathbb{K}^n \rightarrow \mathbb{K}^n$ bzw. $\Phi_{B'} : \mathbb{K}^m \rightarrow \mathbb{K}^m$ besitzen reguläre Darstellungsmatrizen T bzw. S , da sie bijektiv sind. Die Zusammensetzung

$$\xi \mapsto \Phi_B^{-1}(\xi) = x \mapsto y = f(x) \mapsto \eta = \Phi_{B'}(y)$$

besitzt dann die Darstellungsmatrix

$$A' = S \cdot A \cdot T^{-1}$$

und beschreibt die selbe lineare Abbildung f . (nur in anderen Koordinaten)

Definition: Zwei Matrizen $A, A' \in M$ heißen äquivalent, wenn reguläre Transformationsmatrizen $S \in M(m, m; \mathbb{K})$ und $T \in M(n, n; \mathbb{K})$ mit $A' = S \cdot A \cdot T^{-1}$ existieren.

Bemerkung:

- Dies ist wirklich eine mengenth. Äquivalenzrelation!
- Nach Satz 1.4.5 gilt $\text{rg } A = \text{rg } A'$ für äquivalente Matrizen.
- Wegen

$$\Phi_B^{-1}(\xi) = \sum_{i=1}^n \xi_i b_i \text{ also } \Phi_B^{-1}(e_i) = b_i \quad (i = 1, \dots, n)$$

und

$$\Phi_{B'}^{-1}(\eta) = \sum_{k=1}^m \eta_k b'_k \text{ also } \Phi_{B'}^{-1}(e_i) = b'_i \quad (i = 1, \dots, m)$$

gilt

$$T^{-1} = (b_1, \dots, b_n) \text{ und } S^{-1} = (b'_1, \dots, b'_m)$$

(Zur Berechnung von A^{-1} siehe später)

Man kann die Basen B und B' so geschickt wählen, das A' eine ganz einfache Gestalt annimmt.

Satz 1.4.7 Jede $(m \times n)$ - Matrix A mit $\text{rg } A = r (\leq \min\{n, m\})$ ist äquivalent zu einer Matrix der Form

$$D_r = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ & 0 & & 0 \end{pmatrix}$$

(Normalform der durch A (bzgl. $\text{rg } A$) bestimmte Äquivalenzklasse)

Bemerkung: Für die zur Normalenform D_r gehörenden Basen $B = (b_1, \dots, b_n)$ und $B' = (b'_1, \dots, b'_m)$ gilt (mit $f(x) = A \cdot x$)

$$\begin{aligned} f(b_i) &= A \cdot b_i = b'_i \quad (i = 1, \dots, r) \\ f(b_i) &= A \cdot b_i = 0 \quad (i = r + 1, \dots, n) \end{aligned} \quad \text{denn} \quad \begin{aligned} D_r \cdot e_i &= e_i \quad \text{für } i = 1, \dots, r \\ D_r \cdot e_i &= 0 \quad \text{für } i > r \end{aligned}$$

Daraus ergibt sich

$$\ker f = \langle \langle b_{r+1}, \dots, b_n \rangle \rangle$$

und

$$\operatorname{im} f = \langle \langle b'_1, \dots, b'_r \rangle \rangle$$

Folgerung 1.4.8 Für $(m \times n)$ - Matrizen gilt

1. A äquivalent $A' \Leftrightarrow \operatorname{rg} A = \operatorname{rg} A'$
2. $\operatorname{rg} A^T = \operatorname{rg} A$ (d.h. Spaltenrang = Zeilenrang)

Beweis:

1. („ \Leftarrow “): $\operatorname{rg} A = \operatorname{rg} A' = r \Rightarrow A \sim D_r \sim A' \Rightarrow A \sim A'$
2. $\operatorname{rg} A = r \Leftrightarrow D_r = SAT^{-1} \Leftrightarrow D_r^T = (T^{-1})^T A^T S^T = \tilde{S} A^T \tilde{T}^{-1}$
 $\Leftrightarrow D_r^T \sim A^T \Leftrightarrow \operatorname{rg} A^T = r$

Ergänzung: Bekannt sind Matrizen $A = (a_{ik})_{mn}$ und Matrizenoperationen $y = A \cdot x \Leftrightarrow y_i = \sum_{k=1}^n a_{ik} x_k$. Die Verallgemeinerung von Matrizen sind Tensoren $T = (t_{i_1, \dots, i_p})$ mit Tensoroperationen wie z.Bsp.

- $z_i = \sum_{j,k} t_{ijk} x_j y_k = t_{ijk} x_j y_k$
- $z_i = \sum_{j,k} t_{ijk} a_{jk} = t_{ijk} a_{jk}$.

Zum Beweis von Satz 1.4.4 muss man wissen, daß folgende Operationen den Rang einer Matrix $A \in M(m, n; \mathbb{K})$ nicht ändern.

1. Elementare Zeilentransformationen
 - (a) Vertausch zweier Zeilen $z_i \leftrightarrow z_j$.
 - (b) Multiplikation einer Zeile mit einem Vielfachen $\lambda \neq 0$ $z_i \rightarrow \lambda z_i$.
 - (c) Addition des Vielfachen einer Zeile zu einer anderen
 $z_i \rightarrow z_i + \lambda z_j$ $i \neq j$.
2. Elementare Spaltentransformationen
 - (a) Vertausch zweier Spalten $s_k \leftrightarrow s_l$.
 - (b) Multiplikation einer Spalte mit einem Vielfachen $\lambda \neq 0$ $s_k \rightarrow \lambda s_k$.
 - (c) Addition des Vielfachen einer Spalte zu einer anderen
 $s_k \rightarrow s_k + \lambda s_l$ $k \neq l$

Grund: Elementare Zeilen- und Spaltenoperationen lassen sich durch reguläre Elementarmatrizen realisieren, die von links (bei Zeilentransformationen) bzw. von rechts (bei Spaltentransformationen) an A heranzumultiplizieren sind.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} z_{11} & z_{12} & z_{13} \dots \\ z_{21} & z_{22} & z_{23} \dots \end{pmatrix} = \begin{pmatrix} z_{21} & z_{22} & z_{23} \dots \\ z_{11} & z_{12} & z_{13} \dots \end{pmatrix}$$

Man kann zeigen, Daß durch solche Transformationen - endlich oft durchgeführt - aus A eine Normalenform D_r hergestellt werden kann: Gaußscher Algorithmus zur Bestimmung der Normalenform D_r mit gleichzeitiger Berechnung der Transformationsmatrizen S und T und damit der neuen Basen (b_1, \dots, b_n) bzw. (b'_1, \dots, b'_m) .

Gaußscher Algorithmus

Ausgangsschema	$E_{(m)}$	$A_{(m,n)}$	$E_{(n)}$
Elementare	P_1	$P_1 A Q_1$	Q_1
Zeilen- und	$P_2 P_1$	$P_2 P_2 A Q_1 Q_2$	$Q_1 Q_2$
Spaltentransformationen			
\vdots	\vdots	\vdots	\vdots
bis A die Normalform D_r erreicht hat	$P_p \dots P_1$	$P_p \dots P_1 A Q_1 \dots Q_q$	$Q_1 \dots Q_q$
Ergebnis	$P = (S)$	$D_r = P A Q = (S A T^{-1})$	$Q = (T^{-1})$

Bemerkung: Ist man nur an $r = \text{rg } A$ interessiert, so genügt es, A in eine Dreiecksmatrix (Stufenmatrix) der Form

$$A' = \begin{pmatrix} \lambda_1 & & * & \\ & \ddots & & * \\ 0 & & \lambda_1 & \\ & 0 & & 0 \end{pmatrix} \quad \lambda_1 \dots \lambda_r \neq 0$$

zu bringen aus der der Rang einfach ablesbar ist.

Mit diesem Verfahren kann auch die Inverse A^{-1} einer regulären Matrix $A \in M(n, n; \mathbb{K})$ bestimmen. (hier ist $D_n = E$)

$$EA \xrightarrow{\text{elementare Zeilentransformationen}} PE$$

$$\Rightarrow (D_n = E) E = P \cdot A \Rightarrow A^{-1} = P$$

oder

$$AE \xrightarrow{\text{elementare Spaltentransformationen}} EQ$$

$$\Rightarrow (D_n = E) E = A \cdot Q \Rightarrow A^{-1} = Q$$

Beispiele:

- $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & \end{pmatrix}$. Gesucht sind A, S, T und D_r . Idee: Spalten werden nacheinander auf Normalform gebracht.

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$z_2 - 4z_1 \rightarrow z_2$
$\begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$(-\frac{1}{3})z_2 \rightarrow z_2$
$\begin{pmatrix} 1 & 0 \\ \frac{4}{3} & -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$z_1 - 2z_2 \rightarrow z_1$
$\begin{pmatrix} -\frac{5}{3} & \frac{2}{3} \\ \frac{4}{3} & -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$s_3 + s_1 - 2s_2 \rightarrow s_3$
$\begin{pmatrix} -\frac{5}{3} & \frac{2}{3} \\ \frac{4}{3} & -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$	
S	D_2	T^{-1}	

Berechnung von T

E	T^{-1}	
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$	$z_1 - z_3 \rightarrow z_1$
$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$	$z_2 + 2z_3 \rightarrow z_2$
$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	
T	E	

Berechnung von S^{-1}

$$S^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\frac{5}{9} - \frac{8}{9}} \begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} \\ -\frac{4}{3} & -\frac{5}{3} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$$

zugehörige Basen:

$$T^{-1} = (b_1, b_2, b_3) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

$$S^{-1} = (b'_1, b'_2) = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$$

$$A \cdot b_1 = b'_1, A \cdot b_2 = b'_2, A \cdot b_3 = 0$$

$$\ker A = \langle\langle b_3 \rangle\rangle = \left\langle\left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\rangle\right\rangle, \operatorname{im} A = \langle\langle b'_1, b'_2 \rangle\rangle = \mathbb{R}^2$$

1.5 Lineare Gleichungssysteme

Für ein lineares Gleichungssystem

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \vdots + \vdots + \dots + \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right\} \quad (=) \sum_{k=1}^n a_{ik}x_k = b_i \quad (i = 1, \dots, m) \quad (1.1)$$

mit gegebenen $a_{ik}, b_i \in \mathbb{K}$ und gesuchtem $x_k \in \mathbb{K}$ (m Gleichungen für n Unbekannte) suchen wir

- Angaben über die Existenz und Eindeutigkeit von Lösungen
- Angaben über die Struktur der Lösungsmenge
- Methoden zur praktischen Berechnung

Mit

$$A = (a_{ik})_{mn} \in M(m, n; \mathbb{K}), x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n, b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m$$

läßt sich 1.1 in der Form

$$A \cdot x = b$$

bzw. (mit $f : \mathbb{K}^n \rightarrow \mathbb{K}^m, x \mapsto f(x) = A \cdot x$)

$$f(x) = b$$

schreiben. (A heißt Koeffizientenmatrix, b die „rechte“ Seite des Gleichungssystems) Zu bestimmen ist also die Lösungsmenge der lineare Gleichung $f(x) = A \cdot x = b$. d.h. die Lösungsmenge

$$L(A, b) = f^{-1}[\{b\}] = \{x \in \mathbb{K}^n \mid f(x) = b\}$$

eines Vektors $b \in \mathbb{K}^m$ unter $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$.

1.5.1 Homogene lineare Gleichungssysteme mit $b = 0 = \mathbb{K}^m$

Satz 1.5.1

1. Ein homogenes lineares Gleichungssystem

$$f(x) = A \cdot x = 0 \text{ mit } A \in M(n, m; \mathbb{K})$$

besitzt stets die triviale Lösung

$$x = 0 \in \mathbb{K}^n.$$

2. Die Lösungsmenge $L(A, 0) = \ker f$ ist ein Untervektorraum des \mathbb{K}^n der Dimension $\dim f = n - \operatorname{rg} A$.
3. Die triviale Lösung ist genau dann die einzige Lösung, wenn f injektiv, also $\operatorname{rg} A = n$ ist.

Folgerung:

1. $m < n$ (weniger Gleichungen als Unbekannte)

$$\Rightarrow \operatorname{rg} A = \min\{m, n\} < n \Rightarrow A \cdot x = 0$$

ist nicht trivial lösbar.

2. $k = n - \operatorname{rg} A > 0 \Rightarrow$ es existieren k linear unabhängige Lösungsvektoren $x_1, \dots, x_k \in \mathbb{K}^n$ von $A \cdot x = 0$ und für die allgemeine Lösung gilt:

$$x = \lambda_1 x_1 + \dots + \lambda_k x_k \text{ mit } \lambda_1, \dots, \lambda_k \in \mathbb{K}$$

Eine solche Basis nennt man auch ein Fundamentalsystem von Lösungen.

1.5.2 Inhomogene lineare Gleichungssystem mit $b \in \mathbb{K}^m$ beliebig

Satz 1.5.2

1. Ein inhomogenes lineares Gleichungssystem

$$f(x) := A \cdot x = b \text{ mit } A = (a_1, \dots, a_n) \in M(m, n; \mathbb{K})$$

$b \in \mathbb{K}^m$ ist genau dann lösbar, wenn

- $b \in \text{im } f = f[\mathbb{K}^n] = \langle \langle a_1, \dots, a_n \rangle \rangle$ oder (äquivalent dazu)
- $\langle \langle a_1, \dots, a_n, b \rangle \rangle = \langle \langle a_1, \dots, a_n \rangle \rangle$
- $\text{rg}(A, b) = \text{rg } A$ mit der erweiterten Matrix

$$(A, b) = (a_1, \dots, a_n, b) \in M(m, n+1; \mathbb{K})$$

2. Ist $x_s \in \mathbb{K}^n$ eine spezielle Lösung von $A \cdot x = b$, so ist die Lösungsgesamtheit

$$L(A, b) = x_s + \ker f = \{x_s + x \mid A \cdot x = 0\}$$

ein affiner Unterraum des \mathbb{K}^n .

3. $A \cdot x = b$ ist genau dann eindeutig lösbar, wenn

$$\text{rg}(A, b) = \text{rg } A = n$$

Beweis:

1. klar
2. Ist x_s eine spezielle Lösung von $A \cdot x = b$, so gilt für jede andere Lösung (x) $A \cdot x = b = A \cdot x_s$
 $\Rightarrow A \cdot (x - x_s) = 0 \Rightarrow x - x_s \in \ker A \Rightarrow x \in x_s + \ker A$
3. klar

Folgerung:

1. $m < n \Rightarrow A \cdot x = b$ ist nicht eindeutig lösbar. (wenn überhaupt)
2. $\text{rg } A = m (\leq n) \Rightarrow A \cdot x = b$ ist für alle $b \in \mathbb{K}^m$ lösbar. (universell lösbar, denn $f[\mathbb{K}^n] = \mathbb{K}^m$)
3. Die allgemeine Lösung von $A \cdot x = b$ erhält man
 - aus einer speziellen Lösung von $A \cdot x = b$ und

- aus einem Fundamentalsystem x_1, \dots, x_k von Lösungen von $A \cdot x = 0$

$$x = x_s + \lambda_1 x_1 + \dots + \lambda_k x_k \text{ mit } \lambda_1, \dots, \lambda_k \in \mathbb{K}$$

Gaußscher Algorithmus zur Lösung linearer Gleichungssysteme

Ausgangsschema	$A_{(m,n)}$	$b_{(m)}$	$E_{(n)}$
	$P_1 A$	$P_1 b$	$P_1 E$
Elementare Zeilen- und Spaltentransformationen	$P_1 A Q_1$	$P_1 b$	Q_1
	\vdots	\vdots	\vdots
bis A Normalform annimmt	PAQ	Pb	Q
Ergebnis	D_r	b'	Q

Jetzt gilt:

1. x' Lösung von $D_r x' = b' \Leftrightarrow x := Q \cdot x'$ Lösung von $Ax = b$

Beweis: $D_r x' = P \cdot A \cdot (Qx') = P \cdot A \cdot x = P \cdot b = b' \Leftrightarrow A \cdot x = b$

Bemerkung: Die Matrix P braucht nicht berechnet zu werden. (Zeilentransformationen $\hat{=}$ Übergang zu äquivalenten Gleichungen)

2. Aus $D_r x' = n'$, d.h. $\left\{ \begin{array}{l} x'_1 = b'_1 \\ \vdots \\ x'_r = b'_r \\ 0 = b'_{r+1} \\ \vdots \\ 0 = b'_m \end{array} \right\}$, ist abzulesen:

- im Lösbarkeitsfalle ($b'_{r+1} = \dots = b'_m = 0$) ist

$$x'_s = (b'_1, \dots, b'_r, 0, \dots, 0)^T \in \mathbb{K}$$

eine spezielle Lösung von $D_r x' = b'$ und

-

$$x' = x'_s + \lambda_{r+1} e_{r+1} + \dots + \lambda_n e_n = \begin{pmatrix} b'_1 \\ \vdots \\ b'_r \\ \lambda_{r+1} \\ \vdots \\ \lambda_n \end{pmatrix}$$

die allgemeine Lösung. Somit ist die allgemeine Lösung von $Ax = b$

$$x = \underbrace{Qx'_s}_{x_s} + \underbrace{\lambda_{r+1} Qe_{r+1} + \dots + \lambda_n Qe_n}_{\text{Spalten von } Q}$$

Zusatz: Verzichtet man auf elementare Spaltentransformationen, so kann man (bis auf eine Ummummerierung der Variablen) eine Normalform

$$(A', b') = \begin{pmatrix} & & b'_{r+1} \\ E_r & \hat{A}' & \vdots \\ & & b'_r \\ & & b'_{r+1} \\ 0 & 0 & \vdots \\ & & b'_m \end{pmatrix}$$

erreichen. Dabei gilt

$$x \text{ Lösung von } A' \cdot x = b' \Leftrightarrow P \cdot Ax = Pb \Leftrightarrow x \text{ Lösung von } A \cdot x = b$$

(Jetzt muss eventuell noch rücknummeriert werden)

Beispiele:

A	b	
$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$	$\begin{pmatrix} 7 \\ 8 \end{pmatrix}$	$x_1 + 2x_2 + 3x_3 = 7$ $4x_1 + 5x_2 + 6x_3 = 8$
$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \end{pmatrix}$	$\begin{pmatrix} 7 \\ -20 \end{pmatrix}$	
$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 7 \\ \frac{20}{3} \end{pmatrix}$	
$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$	$\begin{pmatrix} -\frac{19}{3} \\ \frac{20}{3} \\ 3 \end{pmatrix}$	$x_1 - x_3 = -\frac{19}{3}$ $x_2 + 2x_3 = \frac{20}{3}$ $x_3 = x_3$

Die allgemeine Lösung ist also

$$x = \begin{pmatrix} -\frac{19}{3} \\ \frac{20}{3} \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

1.6 Determinanten

Sie liefern eine Möglichkeit, die Regularität einer quadratischen Matrix festzustellen. Zur Erinnerung: $S_n = \{\pi = \{1, \dots, n\} \Rightarrow \{1, \dots, n\}\}$ bezeichnen die Gruppe aller Permutationen der Zahlen $1, \dots, n$ (Anzahl $n!$). Das Vorzeichen oder signum $\text{sgn } \pi = (-1)^l$ gibt an, ob eine gerade oder ungerade Anzahl von Nachbervertauschungen nötig ist, um die Bilder $\pi(1), \dots, \pi(n)$ wieder in die natürliche Reihenfolge zu bringen.

Beispiele:

- $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix} \quad \text{sgn } \pi = (-1)^4 = 1$

Satz 1.6.1 Auf der Menge $M(n, n; \mathbb{K})$ aller $(n \times n)$ -Matrizen $A = (a_1, \dots, a_n)$ gibt es genau eine Determinantenfunktion

$$\det : M(n, n; \mathbb{K}) \rightarrow \mathbb{K}, A \mapsto \det A = (A)$$

mit der Eigenschaft

1. $\det(a_1, \dots, \lambda a_k + \mu a'_k, \dots, a_n) = \lambda \det(a_1, \dots, a_k, \dots, a_n) + \mu \det(a_1, \dots, a'_k, \dots, a_n)$ für alle $k = 1, \dots, n$ [Multilinearität]
2. a_1, \dots, a_n linear abhängig $\Rightarrow \det(a_1, \dots, a_n) = 0$ [Determinanteneigenschaft]
3. $\det E = \det(e_1, \dots, e_n) = 1$ [Normierung]

Beweis: Eindeutigkeit: Für eine Determinante muss gelten ($\sum_{i=1}^n a_{ik} e_k = a_k$)

$$\begin{aligned} \det A &= \det(a_1, \dots, a_n) = \det\left(\sum_{i_1=1}^n a_{i_1 1} e_{i_1}, \sum_{i_2=1}^n a_{i_2 2} e_{i_2}, \dots, \sum_{i_n=1}^n a_{i_n n} e_{i_n}\right) \\ &= \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n a_{i_1 1} \cdots a_{i_n n} \det(e_{i_1}, \dots, e_{i_n}) [n^n \text{ Summanden}] \end{aligned}$$

(Summanden nur dann $\neq 0$, wenn $(i_1, \dots, i_n) = (\sigma(1), \dots, \sigma(n))$)

$$\begin{aligned} &= \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) [n! \text{ Summanden}] \\ &= \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \text{sgn } \sigma \underbrace{\det(e_1, \dots, e_n)}_{=1} \\ &= \sum_{\sigma \in S_n} \text{sgn } \sigma a_{\sigma(1)1} \cdots a_{\sigma(n)n} \end{aligned}$$

mit $\sigma^{-1} = \pi$

$$= \sum_{\pi \in S_n} \operatorname{sgn} \pi a_{1\pi(1)} \cdots a_{n\pi(n)}$$

Beispiele:

- $n = 2$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \sum_{\pi \in S_2} \operatorname{sgn} \pi a_{1\pi(1)} \cdot a_{2\pi(2)} = a_{11}a_{22} - a_{12}a_{21}$$

Aus Satz 1.6.1 folgen weitere Eigenschaften:

1. $\det(a_1, \dots, a_i + \lambda a_k, \dots, a_k, \dots, a_n) = \det(a_1, \dots, a_n) + \lambda \underbrace{\det(a_1, \dots, a_k, \dots, a_k, \dots, a_n)}_{=0} = \det(a_1, \dots, a_n)$
2. $\det(a_1, \dots, a_k, \dots, a_i, \dots, a_n) = -\det(a_1, \dots, a_k, \dots, (-a_i), \dots, a_n) = -\det(a_1, \dots, a_i + a_k, \dots, (-a_i), \dots, a_n) = -\det(a_1, \dots, a_i + a_k, \dots, a_k, \dots, a_n) = -\det(a_1, \dots, a_i, \dots, a_k, \dots, a_n)$
(Die Determinantenfunktion ist also alternierend oder antisymmetrisch)
3. $\det(a_{\pi(1)}, \dots, a_{\pi(n)}) = (-1)^l \det(a_1, \dots, a_n) = \operatorname{sgn} \pi \det(a_1, \dots, a_n)$ für alle $\pi \in S_n$.
4. $\det A = |a_{ik}| = \sum_{\pi \in S_n} \operatorname{sgn} \pi a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}$ (Regel von SARRUS)

Beispiele:

- $n = 3$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum_{\pi \in S_3} \operatorname{sgn} \pi a_{1\pi(1)} a_{2\pi(2)} a_{3\pi(3)}$$

Mögliche Permutationen von 123:

$$\underbrace{123 \quad 231 \quad 321}_{\operatorname{sgn} \pi = +1} \quad \underbrace{132 \quad 213 \quad 321}_{\operatorname{sgn} \pi = -1}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Satz 1.6.2 (Eigenschaften der Determinanten)

1. A regulär $\Leftrightarrow \det A \neq 0$
2. $\det(B \cdot A) = \det B \cdot \det A$
3. $\det(A^{-1}) = (\det A)^{-1}$
4. $\det(A^T) = \det A$

Ein wichtiges Hilfsmittel zur Bestimmung der Determinante ist der Laplacesche Entwicklungssatz.

Definition: Bei einer Matrix $A \in M(n, n; \mathbb{K})$ heißt die Zahl

$$A_{ik} := (-1)^{i+k} \det S_{ik}$$

entstanden aus der Streichungsmatrix (durch Streichen der Spalten und Zeile ausgehend von a_{ik})

$$s_{ik} = \begin{pmatrix} a_{11} & | & a_{1n} \\ - & a_{ik} & - \\ a_{n1} & | & a_{nn} \end{pmatrix} \in M(n-1, n-1; \mathbb{K})$$

die Adjunkte oder das algebraische Komplement des Matrixelements a_{ik} .

Satz 1.6.3 (Entwicklungssatz für Determinanten) Für eine $(n \times n)$ -Matrix gilt

1. $\det A = a_{i1}A_{i1} + \dots + a_{in}A_{in}$ für jeden Zeilenindex i . (Entwicklung nach der i -ten Zeile)
2. $\det A = a_{1k}A_{1k} + \dots + a_{nk}A_{nk}$ für jeden Spaltenindex k . (Entwicklung nach der k -ten Spalte)

Beispiele:

•

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

•

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 4 & 5 & 6 \end{vmatrix} = -1 \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = 8 - 5 = 3$$

Mit Hilfe der Adjunkten kann man auch die Inverse einer Matrix explizit angeben.

Satz 1.6.4 Für die Inverse einer regulären Matrix $A = (a_{ik})$ mit der Adjunkten A_{ik} gilt

$$A^{-1} = \frac{1}{\det A} (A_{ik})^T$$

Beispiele:

•

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow A^{-1} = \frac{1}{\det A} A^T = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Satz 1.6.5 (CRAMERSche Regel) Ein lineares Gleichungssystem $Ax = b$ mit regulärer quadratischer Koeffizientenmatrix hat den Lösungsvektor $x = (x_1, \dots, x_n)^T = A^{-1} \cdot b$ mit

$$x_k = \frac{\det(a_1, \dots, b, \dots, a_n)}{\det(a_1, \dots, a_n)} \quad (\text{Ersetzen von } a_k \text{ durch } b)$$

Beispiele:

•

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot x = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

$$x = \frac{1}{-2} \left(\left| \begin{array}{c|c} 5 & 2 \\ 6 & 4 \end{array} \right|, \left| \begin{array}{c|c} 1 & 5 \\ 3 & 6 \end{array} \right| \right)^T = -\frac{1}{2} \begin{pmatrix} 8 \\ 9 \end{pmatrix} = \begin{pmatrix} -4 \\ -\frac{9}{2} \end{pmatrix}$$

Satz 1.6.6 Der Rang einer Matrix $A \in M(m, n; \mathbb{K})$ ist die größte Zahl r , für die eine von Null verschiedene Unterdeterminante existiert.

Beispiele:

•

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 2 \\ 3 & 1 & 4 \end{pmatrix} : \quad \left| \begin{array}{cc} 1 & 2 \\ 2 & 0 \end{array} \right| \neq 0 \Rightarrow r \geq 2 \wedge |A| = 0 \Rightarrow r = 2$$

1.7 Einführung in die Eigenwerttheorie

Sei $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ ein Endomorphismus mit einer quadratischen Darstellungsmatrix A . Benutzt man im \mathbb{K}^n eine andere (aber in Urbild und Zielraum gleiche) Basis $B = (b_1, \dots, b_n)$, so erhalten die Vektoren x und $y = f(x) = Ax$ andere Koordinaten $\xi = \Phi_B(x) = T \cdot x$, $\eta = \Phi_B(y) = T \cdot y$.

$$\begin{array}{ccc} x \in \mathbb{K}^n & \xrightarrow[A]{} & y \in \mathbb{K}^n \\ \Phi_B \downarrow T & & T \downarrow \Phi_B \\ \xi \in \mathbb{K}^n & \xrightarrow[A']{} & \eta \in \mathbb{K}^n \end{array}$$

Die Zusammensetzung

$$\xi \xrightarrow{T^{-1}} x \xrightarrow{A} y = f(x) \xrightarrow{T} \eta$$

besitzt dann die Darstellungsmatrix

$$A' = TAT^{-1}$$

und beschreibt den selben Endomorphismus. (nur in anderen Koordinaten)

Definition: Zwei quadratische Matrizen $A, A' \in M(n, n; \mathbb{K})$ heißen ähnlich, kurz $A \approx A'$, wenn eine reguläre Transformationsmatrix T mit

$$A' = TAT^{-1}$$

existiert.

Wieder werden wir versuchen die Basis B so geschickt zu wählen, daß A' eine einfache Form annimmt. (Eine Normalform D_r ist nicht mehr zu erreichen)

Definition: Eine Matrix $A \in M(n, n; \mathbb{K})$ (bzw. der dazugehörige Endomorphismus $f : \mathbb{K}^n \rightarrow \mathbb{K}^n, x \mapsto A \cdot x$) heißt diagonalisierbar, wenn sie ähnlich zu einer Diagonalmatrix

$$D = \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

ist. (also f bzw. eine Basis (b_1, \dots, b_n) des \mathbb{K}^n eine Diagonalmatrix D als Darstellungsmatrix besitzt.)

Satz 1.7.1 *Ein Endomorphismus $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ (bzw. seine Darstellungsmatrix A) ist genau dann diagonalisierbar, wenn im \mathbb{K}^n eine Basis (b_1, \dots, b_n) aus „Eigenvektoren“ mit der Eigenschaft*

$$\forall_{k=1}^n f(b_k) = A \cdot b_k = \lambda_k b_k$$

mit $\lambda_k \in \mathbb{K}$ existiert.

Definition: Sei V ein beliebiger VR über den Körper \mathbb{K} und $f : V \rightarrow V$ ein Endomorphismus. Ein Skalar $\lambda \in \mathbb{K}$ heißt Eigenwert (EW) von f , wenn ein Eigenvektor $x \in V \setminus \{0\}$ mit

$$f(x) = \lambda x \quad x \hat{=} \text{Eigenvektor}$$

existiert. $E_\lambda := \{x \in V \mid f(x) = \lambda x\} = \ker(f - \lambda \cdot \text{id}) \subset V$ heißt Eigenraum von f zu EW λ und seine Dimension $d := \dim(f - \lambda \cdot \text{id}) \geq 1$ die geometrische Vielfachheit des EWs λ .

Wie findet man die EW einer Matrix $A \in M(n, n; \mathbb{K})$? Es gilt

$$\begin{aligned} \lambda \text{ EW von } A &\Leftrightarrow E_\lambda \neq \{0\} \Leftrightarrow \dim \ker(A - \lambda E) > 0 \\ &\Leftrightarrow A - \lambda A \text{ nicht regulär} \Leftrightarrow \det(A - \lambda E) = 0 \end{aligned}$$

Es ist

$$\det(A - t \cdot E) = \begin{vmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{vmatrix} \quad (\text{mit } t \in \mathbb{K})$$

ein Polynom in t vom Grade n .

Satz 1.7.2 Die EW $\lambda \in \mathbb{K}$ einer Matrix $A \in M(n, n; \mathbb{K})$ sind genau die Nullstellen der charakteristischen Polynoms

$$t \mapsto \chi_A(t) = \det(A - t \cdot E)$$

vom Grade n . Den ER E_λ erhält man durch Lösen des linearen Gleichungssystems $(A - \lambda E) \cdot x = 0$

Eigenschaften:

1. Ähnliche Matrizen haben das gleiche charakteristische Polynom.

Beweis:

$$\begin{aligned} A \approx A' &\Rightarrow A' = TAT^{-1} \Rightarrow \chi_{A'}(t) = \det(A' - tE) \\ &= \det(T(A - tE)T^{-1}) = \det T \det(A - tE) \det T^{-1} \\ &= \det(A - tE) = \chi_A(t) \end{aligned}$$

2. Für die Koeffizienten k_m des charakteristischen Polynoms

$$\chi_A(t) = (-1)^n t^n + (-1)^{n-1} k_{n-1} t^{n-1} + \dots + (-1)^1 k_1 t + k_0$$

gilt stets

$$k_{n-1} = \text{spur } A = \sum_{i=1}^n a_{ii}, \quad k_0 = \det A$$

Folgerung: Ähnliche Matrizen haben gleiche Spur und Determinante.

Beispiele:

- Für eine (2×2) -Matrix gilt stets

$$\chi_A(t) = t^2 - \text{spur } A \cdot t + \det A$$

Satz 1.7.3 Eine Matrix $A \in (n, n; \mathbb{K})$ ist genau dann diagonalisierbar, wenn

1. ihr charakteristisches Polynom vollständig in Linearfunktionen zerfällt. d.h.

$$\chi_A(t) = \pm (t - \lambda_1)^{l_1} \dots (t - \lambda_m)^{l_m}$$

mit paarweise verschiedenen Nullstellen $\lambda_1, \dots, \lambda_m$ und algebraischen Vielfachheiten $l_1, \dots, l_m \in \mathbb{N}$.

2. für jeden EW λ_i ($1 \leq i \leq m$) von A gilt

geometrische Vielfachheit $d_i =$ algebraische Vielfachheit l_i

Bemerkung: Eine Zerlegung von χ_A wie in (1) läßt sich in \mathbb{C} immer durchführen. (\mathbb{C} ist algebraisch abgeschlossen) In \mathbb{R} sind auch quadratische Terme ohne (reelle) Nullstellen möglich (z.B. $t^2 + 1$)

Beweisskizze: Zu jedem Eigenwert λ_i findet man eine Basis des ERs E_{λ_i} aus $d_i = l_i$ Vektoren. Da EV zu verschiedenen EW linear unabhängig sind hat man insgesamt $l_1 + \dots + l_n = \text{grad}\chi_A = n$ linear unabhängige EV, also eine Basis aus EV. (siehe Satz 1.7.1)

Beispiele:

-

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} : \chi_A(t) = \begin{vmatrix} -t & 0 \\ 1 & -t \end{vmatrix} = t^2 \Rightarrow \lambda = 0, l = 2$$

Aber

$$(A - \lambda E) \cdot x = A \cdot x = \begin{pmatrix} 0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1 = 0$$

zeigt das $E_0 = \ker(A - \lambda E) = \left\langle \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle \right\rangle$ 1-dimensional ist, d.h. $d = 1 < 2 = l$. Also ist A zu keiner Diagonalmatrix ähnlich.

-

$$A = \begin{pmatrix} 3 & -2 \\ -2 & 0 \end{pmatrix} : \chi_A(t) = \begin{vmatrix} 3-t & -2 \\ -2 & -t \end{vmatrix} = t^2 - 3t - 4 = (t+1)(t-4) \\ \Rightarrow \lambda_1 = -1, \lambda_2 = 4, \underbrace{l_1 = l_2 = 1}_{\Rightarrow d_1 = d_2 = 1}$$

$$E_{-1} : (A - \lambda_1 E)x = 0 \Leftrightarrow \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \Leftrightarrow x_2 = 2x_1$$

$$\Leftrightarrow x \in \left\langle \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle \right\rangle = \langle\langle b_1 \rangle\rangle$$

$$E_4 : (A - \lambda_2 E)x = 0 \Leftrightarrow \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \Leftrightarrow x_1 = -2x_2$$

$$\Leftrightarrow x \in \left\langle \left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle \right\rangle = \langle\langle b_2 \rangle\rangle$$

Ergebnis:

$$A \approx D = \begin{pmatrix} -1 & \\ & 4 \end{pmatrix}$$

Transformationsmatrix:

$$S := T^{-1} = (b_1, b_2) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

Für sie gilt $D = TAT^{-1} = S^{-1}AS$.

Ergänzungen:

1. Ist die Bedingung (2) für einen Eigenwert verletzt ($D < l$) kann man wenigstens eine sogenannte Jordansche Normalform

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

mit den EW in der Hauptdiagonalen und höchstens Einsen in der (oberen) Nebendiagonale.

2. Bei einer diagonalisierbaren reellen Matrix treten echt komplexe EW immer als konjugiert komplexe Paare $\lambda, \bar{\lambda}$ auf. Auch die zugehörigen Eigenvektoren können als konjugiert komplexe Paare (z, \bar{z}) gewählt werden. (denn $Az = \lambda z \Leftrightarrow A\bar{z} = \lambda\bar{z}$) Ersetzt man das Paar (z, \bar{z}) durch $(y_1, y_2) := (\operatorname{Re} z, \operatorname{Im} z)$ so entsteht (mit $\lambda = \alpha + i\beta$) aus der Komplexen

Diagonalmatrix $D = \begin{pmatrix} \ddots & & & \\ & \lambda & & \\ & & \bar{\lambda} & \\ & & & \ddots \end{pmatrix}$ eine reelle Fast-Diagonalmatrix

$\tilde{D} = \begin{pmatrix} \ddots & & & \\ & \alpha & \beta & \\ & -\beta & \alpha & \\ & & & \ddots \end{pmatrix}$ mit (2×2) -Drehstreckungsblöcken

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = \begin{vmatrix} \operatorname{Re} \lambda & \operatorname{Im} \lambda \\ -\operatorname{Im} \lambda & \operatorname{Re} \lambda \end{vmatrix} = \begin{vmatrix} r \cos \phi & r \sin \phi \\ -r \sin \phi & r \cos \phi \end{vmatrix}$$

auf der Hauptdiagonalen.

Beispiele:

•

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \chi_A(t) = t^2 + 1 \Rightarrow \lambda = i, \bar{\lambda} = -i$$

EV zu $\lambda = i$ ist also etwa $\begin{pmatrix} 1 \\ i \end{pmatrix} \Rightarrow \bar{z} = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ ist EV zu $\bar{\lambda} = -i$.

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

komplexe Transformationsmatrix:

$$S = T^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$$

Reelle Normalenform:

$$\tilde{D} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = A$$

zugehörige reelle Basis:

$$y_1 = \operatorname{Re} z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, y_2 = \operatorname{Im} z = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Reelle Transformationsmatrix:

$$\tilde{S} = \tilde{T}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E$$

1.8 Euklidische und unitäre VR

Definition: Ein euklidischer VR ist ein reeller VR, ausgestattet mit einer positiv definierten symmetrischen Bilinearform

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

als Skalarprodukt (inneres Produkt) für das gilt

1. $\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle$ [Linear im 1. Argument]
 $\langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle$ [Linear im 2. Argument]
2. $\langle y, x \rangle = \langle x, y \rangle$ [Symmetrie]
3. $\langle x, x \rangle \geq 0 \wedge \langle x, x \rangle = 0 \Leftrightarrow x = 0$ [Positive Definiertheit]

Beispiele:

- Der \mathbb{R}^n mit dem Standard-Skalarprodukt

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^T \cdot y$$

- Der VR $\ell^0([a, b] \rightarrow \mathbb{R})$ aller stetigen Funktionen. $f : [a, b] \rightarrow \mathbb{R}$ mit dem SP

$$\langle f, g \rangle := \int_a^b f(x) g(x) dx$$

Mit der positiven Definiertheit des SP(3) gilt

$$\langle f, f \rangle = \int_a^b f^2(x) dx = 0 \Rightarrow f = 0$$

Diese Definition läßt sich nicht einfach auf den Komplexen VR übertragen. Ist etwa $\langle z, z \rangle > 0$, so folgt

$$\langle iz, iz \rangle = ii \langle z, z \rangle = -\langle z, z \rangle < 0 \text{ Widerspruch zu (3)}$$

Definition: Ein unitärer VR ist ein komplexer VR mit einer positiv definierten hermiteschen Semibilinearform $\langle \cdot, \cdot \rangle : V \times V \Rightarrow \mathbb{C}$ als Skalarprodukt für das gilt:

1. $\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle$ [Linear im 1. Argument]
2. $\langle y, x \rangle = \overline{\langle x, y \rangle}$ [Hermitesche Symmetrie]
 $(\Rightarrow \langle x, x \rangle = \overline{\langle x, x \rangle} \in \mathbb{R})$
3. $\langle x, x \rangle \geq 0 \wedge \langle x, x \rangle = 0 \Leftrightarrow x = 0$ [Positive Definiertheit]

Eigenschaften:

$$\begin{aligned} \langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle &= \overline{\langle \lambda_1 y_1 + \lambda_2 y_2, x \rangle} = \overline{\lambda_1 \langle y_1, x \rangle + \lambda_2 \langle y_2, x \rangle} \\ &= \overline{\lambda_1} \overline{\langle y_1, x \rangle} + \overline{\lambda_2} \overline{\langle y_2, x \rangle} \\ &= \overline{\lambda_1} \langle x, y_1 \rangle + \overline{\lambda_2} \langle x, y_2 \rangle \end{aligned}$$

daraus folgt speziell:

- $\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle$
- $\langle x, \lambda y \rangle = \bar{\lambda} \langle x, y \rangle$

Beispiele:

- Der \mathbb{C}^n mit dem Standard-SP

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i = x^T \cdot \bar{y}$$

- Der VR ℓ^0 ($[a, b] \rightarrow \mathbb{C}$) mit dem SP

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx \left(\Rightarrow \langle f, f \rangle = \int_a^b |f(x)|^2 dx \in \mathbb{R} \right)$$

1.8.1 Längenmessung

Satz 1.8.1 In einem euklidischen/unitären VR $(V; \langle, \rangle)$ gilt für die durch

$$|x| := \sqrt{\langle x, x \rangle}$$

definierte Längenfunktion die CAUCHY-SCHWARZsche Ungleichung

$$|\langle x, y \rangle| \leq |x| |y|$$

Beweis: Für $y \neq 0$ gilt (sonst trivial) $\left| |y|^2 x - \langle x, y \rangle y \right|^2$
 $= \langle |y|^2 x - \langle x, y \rangle y, |y|^2 x - \langle x, y \rangle y \rangle$
 $= |y|^4 \langle x, x \rangle - \langle x, y \rangle |y|^2 \langle y, x \rangle - |y|^2 \langle x, y \rangle \langle x, y \rangle + \langle x, y \rangle \langle x, y \rangle \langle y, y \rangle$
 $= |y|^2 \left(|x|^2 |y|^2 - |\langle x, y \rangle|^2 \right) \geq 0$

Satz 1.8.2 In einem euklidischen/unitären VR $(V; \langle, \rangle)$ gilt für die durch

$$x \mapsto |x|$$

die Normierungseigenschaft (1) bzw. (3) aus Kap. 1.1. Jeder euklidische/unitäre VR ist also ein normierter Raum.

Beweis:

1. [positive Definitheit] siehe oben.
2. [Homogenität] $|\lambda x|^2 = \langle \lambda x, \lambda x \rangle = \lambda \bar{\lambda} \langle x, x \rangle = |\lambda|^2 |x|^2$

$$\begin{aligned}
3. \text{ [\Delta - Ungleichung]} \quad & |x + y|^2 = |x|^2 + \langle x, y \rangle + \langle y, x \rangle |y|^2 \\
& |x|^2 + 2\operatorname{Re} \langle x, y \rangle + |y|^2 = |x|^2 + 2|\langle x, y \rangle| + |y|^2 \\
& \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2
\end{aligned}$$

Bemerkung: Nicht jede Norm wird von einem SP induziert. Z.B. die durch

$$|x|_\infty := \max_{i=1, \dots, n} |x_i|$$

im \mathbb{R}^n definierte Maximumsnorm. Notwendig und hinreichend dafür ist die Gültigkeit der Parallelogrammgleichung

$$|x + y|^2 + |x - y|^2 = 2|x|^2 + 2|y|^2$$

Die euklidische Norm $|x| = \sqrt{\langle x, x \rangle}$ erfüllt diese Gleichung!

Beispiele:

- Monome $p_n : [0, 1] \rightarrow \mathbb{R}, x \mapsto x^l$ haben bzgl. der vom SP $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$ induzierte Norm $\|f\| = \sqrt{\langle f, f \rangle}$ die Länge

$$\|p_k\|^2 = \int_0^1 (x^k)^2 dx = \frac{x^{2k+1}}{2k+1} \Big|_0^1 = \frac{1}{2k+1} \Rightarrow \|p_k\| = \frac{1}{\sqrt{2k+1}}$$

1.8.2 Orthogonalität

Definition: Sei V ein euklidischer/unitärer VR mit SP \langle, \rangle

1. Zwei Vektoren x, y heißen orthogonal, wenn

$$\langle x, y \rangle = 0$$

ist. (Schreibweise $x \perp y$)

2. Eine Teilmenge $M \subset V$ heißt Orthogonalsystem (OGS), wenn

$$0 \in M \wedge \forall_{x \neq y} x \perp y$$

3. Ein Orthonormalsystem (ONS) M ist ein OGS mit zusätzlich

$$\forall_{x \in M} |x| = 1$$

(Alle Vektoren sind also Einheitsvektoren)

4. Eine Basis von V , die gleichzeitig ein ONS bildet heißt Orthonormalbasis (ONB)

Jedes OGS und erst recht jede ONB ist linear unabhängig.

Beispiele:

- Im \mathbb{K}^n bilden die Einheitsvektoren eine ONB: Es gilt

$$\langle e_i, e_k \rangle = \delta_{ik} \begin{cases} 1 & \text{für } i = k \\ 0 & \text{für } i \neq k \end{cases} \quad (\text{Kronecker-Delta})$$

Je endlich viele, sogar je abzählbar viele linear unabhängige Vektoren können orthonormiert werden.

Satz 1.8.3 (SCHMIDT'sches Orthonormalisierungsverfahren)

Zu jedem m -tupel (b_1, \dots, b_m) bzw. zu jeder Folge $(b_n)_{n \in \mathbb{N}}$ linear unabhängiger Vektoren in einem euklidischen/unitären VR gibt es ein ONS (e_1, \dots, e_n) bzw. $(e_k)_{k \in \mathbb{N}}$ mit den Eigenschaften

$$\forall k \geq 1 \quad \langle \langle e_1, \dots, e_k \rangle \rangle = \langle \langle b_1, \dots, b_k \rangle \rangle$$

Beweis: Wir definieren die Vektoren e_1, e_2, \dots induktiv

$k = 1$: Mit $e_1 := \frac{b_1}{|b_1|}$ gilt $\langle \langle e_1 \rangle \rangle = \langle \langle b_1 \rangle \rangle$.

$k \rightarrow k+1$: Wir nehmen an, wir hätten schon ein ONS (e_1, \dots, e_k) mit $\forall_{i,j} \langle e_i, e_j \rangle = \delta_{ij}$ und $\langle \langle e_1, \dots, e_k \rangle \rangle = \langle \langle b_1, \dots, b_k \rangle \rangle$. Für den Vektor $\tilde{e}_{k+1} := b_{k+1} - \sum_{j=1}^k \langle b_{k+1}, e_j \rangle e_j$ gilt dann

1. $\tilde{e}_{k+1} \neq 0$ (sonst $b_{k+1} \in \langle \langle e_1, \dots, e_k \rangle \rangle = \langle \langle b_1, \dots, b_k \rangle \rangle$)
2. $\forall_{l=1}^k \langle \tilde{e}_{k+1}, e_l \rangle = \langle b_{k+1}, e_l \rangle - \sum_{j=1}^k \langle b_{k+1}, e_j \rangle \underbrace{\langle e_j, e_l \rangle}_{\delta_{jk}}$
 $= \langle b_{k+1}, e_l \rangle - \langle b_{k+1}, e_l \rangle = 0$
3. $\langle \langle e_1, \dots, e_k, \tilde{e}_{k+1} \rangle \rangle = \langle \langle e_1, \dots, e_k, b_{k+1} \rangle \rangle = \langle \langle b_1, \dots, b_{k+1} \rangle \rangle$

Die Normierung $e_{k+1} := \frac{\tilde{e}_{k+1}}{|\tilde{e}_{k+1}|}$ liefert dann ein ONS wie gewünscht.

Beispiele:

- $\mathbb{R} = \mathbb{K}, V = \ell^0([-1, +1]), \langle f, g \rangle = \int_{-1}^{+1} f(x)g(x)dx$
Orthonormalisierung der Folge der Monome (p_0, p_1, \dots) mit

$$P_k(x) = x^k$$

1. $\|p_0\| = \int_{-1}^{+1} 1dx = 2 \Rightarrow e_0(x) = \sqrt{\frac{1}{2}} \cdot 1$
2. $\tilde{e}_1 := p_1 - \langle p_1, e_0 \rangle e_0, \langle p_1, e_0 \rangle = \int_{-1}^{+1} x \sqrt{\frac{1}{2}} dx = 0$
 $\tilde{e}_1 = p_1, \|\tilde{e}_1\|^2 = \|p_1\|^2 = \int_{-1}^{+1} x^2 dx = \frac{2}{3} \Rightarrow e_1(x) = \sqrt{\frac{3}{2}} x$
3. $\tilde{e}_2 = p_2 - \langle p_2, e_0 \rangle e_0 - \langle p_2, e_1 \rangle e_1$
 $\langle p_2, e_0 \rangle = \int_{-1}^{+1} x^2 \sqrt{\frac{1}{2}} dx = \frac{\sqrt{2}}{3}$
 $\langle p_2, e_1 \rangle = \int_{-1}^{+1} x^2 \sqrt{\frac{3}{2}} x dx = 0$
 $\tilde{e}_2 = x^2 - \frac{\sqrt{2}}{3} \cdot \sqrt{\frac{1}{2}} \cdot 1 = x^2 - \frac{1}{3}$

$$\|e_2\|^2 = \int_{-1}^{+1} (x^2 - \frac{1}{3})^2 dx = \dots = \frac{2}{5} (\frac{2}{3})^2$$

$$e_2(x) = \sqrt{\frac{5}{2} \frac{3}{2}} (x^2 - \frac{1}{3})$$

Die Polynome

$$P_k(x) := \sqrt{\frac{2}{2k+1}} e_k(x)$$

heißen LEGENDRE-Polynome. $P_0(x) = 1$; $P_1(x) = x$; $P_2(x) = \frac{3}{2}(x^2 - \frac{1}{3})$ (sie bilden ein OGS) Bei diesem Verfahren wurde die Existenz einer Orthogonalprojektion benutzt.

Definition: Zu einem UR U eines euklidischen/unitären VRs V heißt der UR

$$U^\perp := \{x \in V \mid x \perp y \text{ für alle } y \in U\}$$

das orthogonale Komplement.

Satz 1.8.4 Auf einem endlich dimensionalen UR U eines euklidischen/unitären VRs V existiert eindeutig die Orthogonalprojektion

$$x \in V \mapsto x_p \in U$$

mit

$$\forall x \in V x - x_p \in U^\perp$$

Für eine beliebige ONB (e_1, \dots, e_m) von U gilt

$$x_p = \sum_{i=1}^m \langle x, e_i \rangle e_i$$

Insbesondere besitzt jeder Vektor $x \in U$ bezüglich (e_1, \dots, e_m) die Basisdarstellung

$$x = \sum_{i=1}^n \langle x, e_i \rangle e_i$$

1.8.3 Winkelmessung (nur im Reellen)

Aus der C-S-Ungleichung

$$-1 \leq \frac{\langle x, y \rangle}{|x| |y|} \leq +1$$

ergibt sich

Definition: In einem euklidischen VR heißt die durch

$$\cos \varphi = \frac{\langle x, y \rangle}{|x| |y|} = \left\langle \frac{x}{|x|}, \frac{y}{|y|} \right\rangle$$

eindeutig bestimmte Zahl $\varphi \in [0, \pi]$ der (nicht orientierte) Winkel zwischen $x, y \neq 0$.

Bemerkung: Die Gleichung $\langle x, y \rangle = |x| \cdot |y| \cdot \cos \varphi$ wird in der Schulgeometrie oft zur Definition des Skalarprodukts benutzt.

1.8.4 Volumen- und Flächenberechnung (nur im \mathbb{R}^n)

Im \mathbb{R}^n kann man das Volumen eines von Vektoren x_1, \dots, x_n aufgespannten n -dim Parallelotops mit Hilfe der Determinante bestimmen

$$V(x_1, \dots, x_n) = \det(x_1, \dots, x_n)$$

(für $n = 3$ Spatprodukt) Das Vorzeichen gibt die Orientierung an. Nachteil: Die Formel läßt sich nicht für niedriger-dimensionale Paralleleogramme, die von weniger als n Vektoren aufgespannt werden, benutzen. Aber es gilt auch

$$\begin{aligned} \det((\langle x_k, x_l \rangle)_{nn}) &= \det\left(\sum_{i=1}^n x_{ik} x_{il}\right) = \det\left(\sum_{i=1}^n x_{ki}^T x_{il}\right) \\ &= \det(x_{ki}^T) \cdot \det(x_{il}) = \det^2(x_{il}) = \det^2(x_1, \dots, x_n) \\ &= V^2(x_1, \dots, x_n) \end{aligned}$$

Diese Formel läßt sich verallgemeinern:

Satz 1.8.5 *Im euklidischen \mathbb{R}^n gilt für den p -dim Flächeninhalt ($1 \leq p \leq n$) eines von Vektoren x_1, \dots, x_p aufgespannten p -dim Parallelogramms:*

$$a_p(x_1, \dots, x_p) = \sqrt{\det(\langle x_i, x_k \rangle)_{pp}} \geq 0$$

mit der GRAMschen Determinante $\det(\langle x_i, x_k \rangle)_{pp}$ von x_1, \dots, x_p . (Für $p = n$ erhält man das nichtorientierte Volumen eines n -dim Parallelotops)

1.8.5 Zum Vektor- oder Kreuzprodukt

Im \mathbb{R}^3 kann man das Vektorprodukt $y = x_1 \times x_2$ durch folgende Eigenschaften charakterisieren:

1. $y \perp x_1 \wedge y \perp x_2$.
2. $|y| = a_2(x_1, x_2) = \sqrt{\begin{vmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle \end{vmatrix}} = \sqrt{|x_1|^2 |x_2|^2 - \langle x_1, x_2 \rangle^2}$
3. $\det(x_1, x_2, y) \geq 0$

Eine explizite Formel ist

$$y = x_1 \times x_2 = \sum_{i=1}^3 \det(x_1, x_2 e_i) e_i$$

oder

$$x_1 \times x_2 = \sum_{i=1}^3 \begin{vmatrix} x_{11} & x_{12} & \delta_{1i} \\ x_{21} & x_{22} & \delta_{2i} \\ x_{31} & x_{32} & \delta_{3i} \end{vmatrix} e_i = \begin{vmatrix} x_{11} & x_{12} & e_1 \\ x_{21} & x_{22} & e_2 \\ x_{31} & x_{32} & e_3 \end{vmatrix}$$

mit den Einheitsvektoren e_1, e_2, e_3 Zum Abschluss noch einige spezielle Endomorphismen

$$f : \mathbb{K}^n \rightarrow \mathbb{K}^n \text{ bzw. Matrizen } A \in M(n, n; \mathbb{K})$$

im euklidischen/unitären \mathbb{K}^n mit Standard-SP.

\mathbb{K}	Endomorph.	Eigenschaft	D.m. A bzgl. ONB	Eigen.
\mathbb{R}	selbstadjunkt	$\langle f(x), y \rangle = \langle x, f(y) \rangle$	symmetrisch	$A^T = A$
\mathbb{C}	selbstadjunkt	$\langle f(x), y \rangle = \langle x, f(y) \rangle$	hermitesch	$\bar{A}^T = A$
\mathbb{R}	orthogonal	$\langle f(x), f(y) \rangle = \langle x, y \rangle$	orthogonal	$A^T = A^{-1}$
\mathbb{C}	unitär	$\langle f(x), f(y) \rangle = \langle x, y \rangle$	unitär	$\bar{A}^T = A^{-1}$

Gestalt der Darstellungsmatrizen A

$$1. f \text{ selbstadjunkt} \Leftrightarrow \langle Ax, y \rangle = (Ax)^T \cdot \bar{y} = x^T A^T \bar{y} \\ = \langle x, Ay \rangle = x^T \bar{A} \bar{y} \text{ für alle } x, y \in \mathbb{K}^n \\ \Leftrightarrow A^T = \bar{A} \Leftrightarrow \bar{A}^T = A$$

$$2. f \text{ orthogonal/unitär} \Leftrightarrow \bar{A}^T = A^{-1}$$

Eigenschaften:

1. Alle diese Abb. f /Matrizen A sind (überall) stets diagonalisierbar, d.h. es existiert eine Basis aus (komplexen) EV.
2. Es existiert sogar eine ONB $S = (b_1, \dots, b_n)$ aus EV ($\Rightarrow S\bar{S}^T = E$ bzw. $\bar{S}^T = S^{-1}$)

Begründung: EV zu verschiedenen EW sind orthogonal und in den einzelnen Eigenräumen kann orthonormalisiert werden.

Sei x ein EV zum EW λ

1. A symmetrisch/hermitesch $\Rightarrow \langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda \cdot |x| \\ \stackrel{!}{=} \langle x, Ax \rangle = \langle x, \lambda x \rangle = \bar{\lambda} |x| \stackrel{x \neq 0}{\Rightarrow} \lambda = \bar{\lambda} \in \mathbb{R}$
2. A orthogonal/unitär $\Rightarrow \langle Ax, Ax \rangle = \langle \lambda x, \lambda x \rangle = \lambda \bar{\lambda} \langle x, x \rangle \\ |\lambda|^2 |x|^2 \stackrel{!}{=} \langle x, x \rangle = |x|^2 \\ \Rightarrow |\lambda| = 1$

Satz 1.8.6 Ein selbstadjunkter Endomorphismus $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ besitzt eine ONB aus EV, wobei die zugehörigen EW stets reel sind.

Folgerung: Jede symmetrische Matrix $A \in M(n, n; \mathbb{R})$ und jede hermitesche Matrix $A \in M(n, n; \mathbb{C})$ ist zu einer reellen Diagonalmatrix ähnlich. Die Transformationsmatrix S (mit $D = S^{-1}AS$) kann als orthogonale Matrix mit $S^{-1} = S^T$ bzw als unitäre Matrix mit $S^{-1} = \bar{S}^T$ gewählt werden. Schreibweise:

$$A \stackrel{\perp}{\approx} D$$

Satz 1.8.8

1. Die Menge aller unitären Matrizen $A \in M(n, n; \mathbb{C})$ bildet eine Untergruppe der linearen Gruppe $GL(n, \mathbb{C})$, die unitäre Gruppe $U(n, n; \mathbb{C})$.
2. Die Menge aller orthogonalen Matrizen $A \in M(n, n; \mathbb{R})$ bildet eine Untergruppe der linearen Gruppe $GL(n, \mathbb{R})$, die orthogonale Gruppe $O(n, \mathbb{R})$. In ihr liegt die spezielle orthogonale Gruppe $SO(n, \mathbb{R})$ aller eigentlich orthogonalen Matrizen A mit $\det A = +1$.

Bemerkung: Die uneigentlich orthogonalen Matrizen mit $\det A = -1$ bilden keine Gruppe. ($\det A = +1$ Drehung, $\det A = -1$ Spiegelung und Drehung)

Beispiele:

- $n = 2$: Jeder Drehmatrix $A \in SO(2, \mathbb{R})$ (mit $\det A = +1$) ist von der Gestalt $A = D_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ mit den Spezialfällen $D_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $D_\pi = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
- $n = 3$: Normalform einer Drehmatrix $A \in SO(3, \mathbb{R})$ (mit $\det A = +1$)

$$A \stackrel{\perp}{\approx} \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = D_\varphi$$

mit den Spezialfällen $D_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ und $D_\pi = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Einfache Bestimmung von D_φ und einer zugehörigen ONB (e_1, e_2, e_3) [für $\varphi \neq 0$]

1. e_3 ist ein normierter EV zum EW $+1$ und bestimmt die Drehachse $E_1 = \langle\langle e_3 \rangle\rangle$.
2. e_1 kann in der Drehebene $\langle\langle e_3 \rangle\rangle^\perp$ beliebig als Einheitsvektor gewählt werden und deshalb $e_2 := e_3 \times e_1$ zu einer positiv orientierten ONB ergänzt werden.
3. Für den Drehwinkel gilt:

$$\text{spur } A = \text{spur } D_\varphi = 2\cos \varphi + 1$$

und

$$\sin \varphi = \langle A \cdot e_1, e_2 \rangle$$

Kapitel 2

Differential und Integralrechnung im \mathbb{R}^n

2.1 Konvergenz und Stetigkeit im \mathbb{R}^n

Viele Definitionen und Sätze aus Kapitel ?? (Konvergenz und Stetigkeit im $\mathbb{K} = \mathbb{R}$ oder \mathbb{C}) können wörtlich übernommen werden, wenn die Betragsfunktion

$$x \in \mathbb{K} \mapsto |x| \in \mathbb{R}$$

durch die euklidische Norm

$$x \in \mathbb{R}^n \mapsto |x| = |x|_2 = \sqrt{\sum_{i=1}^n x_i^2} \in \mathbb{R}$$

oder eine andere Norm, z.Bsp. die Maximumsnorm

$$x \in \mathbb{R}^n \mapsto |x|_\infty = \max \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$$

ersetzt wird.

2.1.1 Topologische Grundbegriffe

- ϵ -Umgebung eines Punktes $\overset{\circ}{x} \in \mathbb{R}^n$:

$$U_\epsilon(\overset{\circ}{x}) = \{x \in \mathbb{R}^n \mid |x - \overset{\circ}{x}| < \epsilon\}$$

- Wenn um jeden Punkt $x \in U$ einer Teilmenge U eine ϵ -Umgebung, die ganz in U liegt (also $U_\epsilon(x) \subset U$) so ist U eine offene Teilmenge.
- Eine Menge heißt abgeschlossen, wenn ihr Komplement offen ist.
- Wenn eine abgeschlossene Menge auch noch beschränkt ist, so nennt man sie kompakt.

2.1.2 Konvergenzbegriffe

Eine Folge im $(x_k \in \mathbb{R}^n)_{k \in \mathbb{N}}$ konvergiert, wenn

$$\overset{\circ}{x} \in \mathbb{R}^n \Leftrightarrow \forall \epsilon > 0 \exists m \in \mathbb{N} \forall k \geq n \left| x_k - \overset{\circ}{x} \right| < \epsilon$$

gilt. Satz von BOLZANO-WEIERSTRASS; Konvergenz von Funktionen $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ Wenn für jede Folge

$$(x_k \in D \setminus \{\overset{\circ}{x}\})_{k \in \mathbb{N}} \text{ mit } \lim x_k = \overset{\circ}{x}$$

$$\lim_{k \rightarrow \infty} (f(x_k) \in \mathbb{R})_{k \in \mathbb{N}} = c$$

gilt, so hat f den Grenzwert

$$\lim_{x \rightarrow \overset{\circ}{x}} f(x) = c$$

2.1.3 Stetigkeit

$f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ stetig in $\overset{\circ}{x} \in D$:

$$\lim_{x \rightarrow \overset{\circ}{x}} f(x) = f(\overset{\circ}{x})$$

Folgenkriterium: Für jede Folge $(x_k \in D)_{k \in \mathbb{N}} \rightarrow \overset{\circ}{x}$ gilt:

$$(f(x_k) \in \mathbb{R})_{k \in \mathbb{N}} \rightarrow f(\overset{\circ}{x})$$

ϵ - δ -Kriterium

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in D \left(|x - \overset{\circ}{x}| < \delta \Rightarrow |f(x) - f(\overset{\circ}{x})| < \epsilon \right)$$

Beispiele:

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases} = \begin{cases} \sin \varphi \cos \varphi & r \neq 0 \\ 0 & r = 0 \end{cases}$$

f ist in $(0, 0)$ nicht stetig, denn $\lim_{x \rightarrow 0} f(x, x) = \lim_{x \rightarrow 0} \frac{1}{2} = \frac{1}{2}$ aber $\lim_{x \rightarrow 0} f(x, 0) = 0$.

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = \begin{cases} \frac{y^3}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

f ist in $(0, 0)$ stetig, denn

$$|f(x, y) - f(0, 0)| = \frac{|y^3|}{x^2+y^2} \leq |y| \leq \sqrt{x^2+y^2} < \epsilon$$

falls $|(x, y) - (0, 0)| = \sqrt{x^2+y^2} < \delta$.

Es gelten die üblichen Rechenregeln für stetige Funktionen wie z.Bsp

$$f, g \text{ stetig} \Rightarrow f + g; f \cdot g; \frac{f}{g}; g \circ f \text{ stetig}$$

und ebenso der Satz vom Maximum und Minimum: Eine stetige Funktion $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ auf einer kompakt Teilmenge $K \subset \mathbb{R}^n$ ist beschränkt und besitzt ein (globales) Maximum und Minimum.

Eine Abbildung

$$f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, x = (x_1, \dots, x_n) \mapsto f(x) = (f_1(x), \dots, f_m(x))$$

ist genau dann in $\overset{\circ}{x} \in D$ stetig, wenn jede Komponentenfunktion $f_i : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ in $\overset{\circ}{x}$ stetig ist.

Beispiele:

- $\phi : (r, t) \in \mathbb{R}^t \times \mathbb{R} \mapsto \phi(r, t) = (r \cos t, r \sin t) = (x, y) \in \mathbb{R}^2$

2.2 Partielle und totale Differenzierbarkeit

Der Definitionsbereich G der Funktion $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^{(m)}$ sei im Folgenden stets ein Gebiet, d.h. eine Teilmenge des \mathbb{R}^n , die offen und zusammenhängend ist. (Verallgemeinerung von offenen Intervallen im \mathbb{R}^n)

Definition:

1. Eine Funktion $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ heißt in $\overset{\circ}{x} \in G$ partiell differenzierbar, wenn die „partielle“ Funktionen

$$x_i \mapsto f\left(\overset{\circ}{x}_1, \dots, x_i, \dots, \overset{\circ}{x}_n\right)$$

(als Funktion einer Veränderlichen) in $\overset{\circ}{x}_i$ differenzierbar sind. Die partielle Ableitung von f nach der i -ten Komponente in $\overset{\circ}{x}$ wird durch

$$\frac{\partial f}{\partial x_i}\left(\overset{\circ}{x}\right) = \partial_i f\left(\overset{\circ}{x}\right)$$

dargestellt.

2. f heißt (auf G) partiell differenzierbar, wenn f in jedem $\overset{\circ}{x} \in G$ partiell differenzierbar ist.

Bemerkung:

1. Beim partiell Differenzieren interessiert man sich nur für das Verhalten von f auf achsenparallelen Geraden.
2. Bezeichnung bei Funktionen $(x, y, z, \dots) \mapsto f(x, y, z, \dots)$:

$$\partial_x f(x, y, z, \dots) = f_x(x, y, z, \dots)$$

$$\Delta = f_{xx} + f_{yy} + \dots$$

3. Rekursiv lassen sich auch höhere partielle Ableitungen definieren.

$$\frac{\partial^3 f}{\partial x_1 \partial x_k \partial x_i} = \partial_1 \partial_k \partial_i f, \quad \frac{\partial^3 f}{\partial x_k^2 \partial x_i} = \partial_k^2 \partial_i f$$

4. Abbildungen $f : G \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ heißen partiell differenzierbar, wenn jede Komponentenfunktion f_i ($i = 1, \dots, m$) partiell differenzierbar ist.

Beispiele:

•

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

Für $(x, y) \neq (0, 0)$ gilt

$$\partial_x f(x, y) = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2}, \quad \partial_y f(x, y) = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2}$$

Für $(x, y) = (0, 0)$ gilt

$$\begin{aligned} f(x, 0) = 0 &\Rightarrow \partial_x f(0, 0) = 0 \\ f(0, y) = 0 &\Rightarrow \partial_y f(0, 0) = 0 \end{aligned}$$

Also ist f überall partiell diffbar. (aber nicht überall stetig)

•

$$\begin{aligned} f(x, y) &= \begin{cases} \frac{xy^3}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases} \\ \partial_x f(x, y) &= \begin{cases} \frac{y^2(y^2-x^2)}{(x^2+y^2)^2} & \text{für } y \neq 0 \\ 0 & \text{für } y = 0 \end{cases} \\ \partial_y f(x, y) &= \begin{cases} \frac{x(3y^2(x^2-y^2)-2y^4)}{(x^2+y^2)^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases} \\ \partial_x f(0, y) &= \begin{cases} y & \text{für } y \neq 0 \\ 0 & \text{für } y = 0 \end{cases} = y \Rightarrow \partial_y \partial_x f(0, 0) = 1 \\ \partial_y f(x, 0) &= 0 \Rightarrow \partial_x \partial_y f(0, 0) = 0 \end{aligned}$$

Die Reihenfolge ist bei partiellen Ableitungen wesentlich.

Definition:

1. Eine Funktion $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ heißt in $\overset{\circ}{x} \in G$ (total) differenzierbar, wenn Funktionen $\Delta_1, \dots, \Delta_n : G \rightarrow \mathbb{R}$ existieren mit

(a)

$$\forall x \in G f(x) = f(\overset{\circ}{x}) + \sum_{i=1}^n \Delta_i(x) (x_i - \overset{\circ}{x}_i) = f(\overset{\circ}{x}) + \langle \Delta(x), x - \overset{\circ}{x} \rangle$$

(b)

$\Delta_1, \dots, \Delta_n$ sind in $\overset{\circ}{x}$ stetig

Der Vektor $\text{grad } f(\overset{\circ}{x}) = \nabla f(\overset{\circ}{x}) := (\Delta_1(\overset{\circ}{x}), \dots, \Delta_n(\overset{\circ}{x}))$ heißt der Gradient von f in $\overset{\circ}{x}$.

2. f heißt (auf G) (total) differenzierbar, wenn f in jedem $\overset{\circ}{x} \in G$ differenzierbar ist.

Geometrische Interpretation für $n = 2$

Für $(x, y) \neq (\overset{\circ}{x}, \overset{\circ}{y})$ gilt:

$$\begin{aligned} f(x, y) &= f(\overset{\circ}{x}, \overset{\circ}{y}) + \Delta_1(x, y) (x - \overset{\circ}{x}) + \Delta_2(x, y) (y - \overset{\circ}{y}) \\ &= \left\langle \begin{pmatrix} \Delta_1(x, y) \\ \Delta_2(x, y) \\ -1 \end{pmatrix}, \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix} - \begin{pmatrix} \overset{\circ}{x} \\ \overset{\circ}{y} \\ f(\overset{\circ}{x}, \overset{\circ}{y}) \end{pmatrix} \right\rangle = 0 \end{aligned}$$

Der Punkt $(x, y, f(x, y))^T$ liegt also in einer „Sekantenebene“ durch $(\overset{\circ}{x}, \overset{\circ}{y}, f(\overset{\circ}{x}, \overset{\circ}{y}))$ mit Normalenvektor

$$N(x, y) = \begin{pmatrix} \Delta_1(x, y) \\ \Delta_2(x, y) \\ -1 \end{pmatrix}$$

(Hessische Normalform)

$$N(\overset{\circ}{x}, \overset{\circ}{y}) = \begin{pmatrix} \nabla f(\overset{\circ}{x}, \overset{\circ}{y}) \\ -1 \end{pmatrix}$$

Bei Annäherung $(x, y) \rightarrow (\overset{\circ}{x}, \overset{\circ}{y})$ geht diese in eine Tangentialebene in $(\overset{\circ}{x}, \overset{\circ}{y}, f(\overset{\circ}{x}, \overset{\circ}{y}))^T$ mit Normalenvektor $N(\overset{\circ}{x}, \overset{\circ}{y}) = \begin{pmatrix} \nabla f(\overset{\circ}{x}, \overset{\circ}{y}) \\ -1 \end{pmatrix}$ über.

Ergebnis: Total differenzierbare Funktionen f besitzen in jedem Punkt $(\overset{\circ}{x}, f(\overset{\circ}{x}))$ ihres Graphen eine Tangentialhyperebene deren Normalenvektor

$$N(\overset{\circ}{x}) = \begin{pmatrix} \text{grad } f(\overset{\circ}{x}) \\ -1 \end{pmatrix}$$

durch den Gradienten bestimmt wird.

1. Äquivalent mit der Eigenschaft (1) und (2) ist die Darstellung

$$\begin{aligned} f(x) &= f(\overset{\circ}{x}) + \langle \Delta(x), x - \overset{\circ}{x} \rangle = f(\overset{\circ}{x}) + \langle \text{grad } f(\overset{\circ}{x}), x - \overset{\circ}{x} \rangle + \langle \Delta(x) - \Delta(\overset{\circ}{x}), x - \overset{\circ}{x} \rangle \\ &= f(\overset{\circ}{x}) + \langle \text{grad } f(\overset{\circ}{x}), x - \overset{\circ}{x} \rangle + R(x) \end{aligned}$$

mit einem Restglied $x \mapsto R(x)$ mit

$$\lim_{x \rightarrow \overset{\circ}{x}} \frac{R(x)}{|x - \overset{\circ}{x}|} = 0$$

denn

$$0 \leq \frac{|R(x)|}{|x - \overset{\circ}{x}|} < \left| \left\langle \Delta(x) - \Delta(\overset{\circ}{x}), \frac{x - \overset{\circ}{x}}{|x - \overset{\circ}{x}|} \right\rangle \right| \leq \underbrace{|\Delta(x) - \Delta(\overset{\circ}{x})|}_{0 \text{ für } x \rightarrow \overset{\circ}{x}} \cdot 1$$

2. Abbildungen $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißen (total) differenzierbar, wenn jede Komponente $f_j : G \rightarrow \mathbb{R}$ (total) differenzierbar ist. Die Matrix

$$Df(\overset{\circ}{x}) := \begin{pmatrix} \text{grad } f_1(\overset{\circ}{x}) \\ \vdots \\ \text{grad } f_m(\overset{\circ}{x}) \end{pmatrix} \in M(m, n; \mathbb{R})$$

heißt dann die Funktionalmatrix oder Jacobi-Matrix von f in $\overset{\circ}{x}$ und es gilt:

$$\begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} f_1(\overset{\circ}{x}) \\ \vdots \\ f_m(\overset{\circ}{x}) \end{pmatrix} + \begin{pmatrix} \text{grad } f_1(\overset{\circ}{x}) \\ \vdots \\ \text{grad } f_m(\overset{\circ}{x}) \end{pmatrix} \cdot \begin{pmatrix} x_1 - \overset{\circ}{x}_1 \\ \vdots \\ x_m - \overset{\circ}{x}_m \end{pmatrix} + \begin{pmatrix} R_1(x) \\ \vdots \\ R_m(x) \end{pmatrix}$$

$$\Leftrightarrow f(x) = f(\overset{\circ}{x}) + Df(\overset{\circ}{x}) \cdot (x - \overset{\circ}{x}) + R(x)$$

mit einem Restglied $x \mapsto R(x)$ mit

$$\lim_{x \rightarrow \overset{\circ}{x}} \frac{|R(x)|}{|x - \overset{\circ}{x}|} = 0$$

Satz 2.2.1 Eine in $\overset{\circ}{x} \in G$ total differenzierbare Funktion $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ist dort auch stetig sowie partiell differenzierbar und es gilt

$$\text{grad } f(\overset{\circ}{x}) = (\partial_1 f(\overset{\circ}{x}), \dots, \partial_n f(\overset{\circ}{x}))$$

Beweis:

1. f ist nach (1) und (2) als Verkettung in $\overset{\circ}{x}$ stetiger Funktionen in $\overset{\circ}{x}$ stetig.

2. Für die partielle Funktion $f_{(i)} : x_i \mapsto f(\overset{\circ}{x}_1, \dots, x_i, \dots, \overset{\circ}{x}_n)$ gilt

$$f_{(i)}(x_i) = f(\overset{\circ}{x}_i) + \Delta_i(\overset{\circ}{x}_1, \dots, x_i, \dots, \overset{\circ}{x}_n)(x_i - \overset{\circ}{x}_i)$$

Es existiert also

$$\begin{aligned} \partial_i f(\overset{\circ}{x}) &= \lim_{x_i \rightarrow \overset{\circ}{x}_i} \frac{f_{(i)}(x_i) - f_{(i)}(\overset{\circ}{x}_i)}{x_i - \overset{\circ}{x}_i} \\ &= \lim_{x_i \rightarrow \overset{\circ}{x}_i} \Delta_i(\overset{\circ}{x}_1, \dots, x_i, \dots, \overset{\circ}{x}_n) = \Delta_i(\overset{\circ}{x}) \end{aligned}$$

Folgerung: Für die Jacobi-Matrix einer differenzierbaren Abbildung $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ gilt

$$Df(\overset{\circ}{x}) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_m & \cdots & \partial_n f_m \end{pmatrix} (\overset{\circ}{x})$$

Satz 2.2.2 $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sei stetig partiell differenzierbar, d.h. es existieren die partiellen Ableitung $\partial_i f$ als stetige Funktionen $\partial_i f : G \rightarrow \mathbb{R}$. Dann ist f sogar total differenzierbar.

Beweis: Für $n = 2$ sei $(\overset{\circ}{x}, \overset{\circ}{y}) \in G$ beliebig. Für ein $(x, y) \neq (\overset{\circ}{x}, \overset{\circ}{y})$ gilt dann

$$f(x, y) - f(\overset{\circ}{x}, \overset{\circ}{y}) = \left(f(x, \overset{\circ}{y}) - f(\overset{\circ}{x}, \overset{\circ}{y}) \right) + \left(f(x, y) - f(x, \overset{\circ}{y}) \right)$$

Anwendung des MWS auf die differenzierbare Funktion $t \mapsto f(t, \overset{\circ}{y})$ und $t \mapsto$

$f(x, t)$ liefert Zahlen $\bar{x} \in \overline{\overset{\circ}{x}, x}, \bar{y} \in \overline{\overset{\circ}{x}, \overset{\circ}{y}}$ mit

$$f(x, y) - f(\overset{\circ}{x}, \overset{\circ}{y}) = \partial_x f(\bar{x}, \overset{\circ}{y})(x - \overset{\circ}{x}) + \partial_y f(x, \bar{y})(y - \overset{\circ}{y})$$

$$f(x, y) - f(\overset{\circ}{x}, \overset{\circ}{y}) = \Delta_1(x, y)(x - \overset{\circ}{x}) + \Delta_2(x, y)(y - \overset{\circ}{y})$$

Δ_1, Δ_2 sind also in $(\overset{\circ}{x}, \overset{\circ}{y})$ stetig fortsetzbar, d.h. f ist in $(\overset{\circ}{x}, \overset{\circ}{y})$ total differenzierbar.

Anwendung: Alle Funktionen denen man „ansieht“, daß sie partiell differenzierbar sind und die partiellen Ableitungen wieder stetig sind, sind total differenzierbar.

Beispiele:

- $f(x, y, z) = e^{\sin(x^2+y^2)} \cos(x \cdot y \cdot z)$
- $\ell^1(G) := \{f : G \subset \mathbb{R}^n \rightarrow \mathbb{R} \mid f \text{ stetig (partiell) differenzierbar}\}$
- $\ell^t(G) := \{f : G \subset \mathbb{R}^n \rightarrow \mathbb{R} \mid f \text{ } t\text{-mal stetig (partiell) differenzierbar}\}$

Satz 2.2.3 (Satz von SCHWARZ) $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sei in $\overset{\circ}{x} \in G$ 2-mal (total) differenzierbar. Dann gilt

$$\partial_l \partial_k f(\overset{\circ}{x}) = \partial_k \partial_l f(\overset{\circ}{x}) \text{ für alle } k, l = 1, \dots, n$$

(Analog für höhere Ableitungen)

Folgerung: Bei einer ℓ^2 -Funktion $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ist die HESSE-Matrix

$$H_f(x) = \begin{pmatrix} \partial_1 \partial_1 f & \cdots & \partial_n \partial_1 f \\ \vdots & & \vdots \\ \partial_1 \partial_n & \cdots & \partial_n \partial_n f \end{pmatrix}$$

für alle $x \in G$ symmetrisch mit reellen EW.

Für (total) differenzierbare Funktionen gelten die üblichen Rechenregeln. (z.Bsp. f, g diffbar $\Rightarrow f + g, f \cdot g, \frac{f}{g}$ diffbar)

Satz 2.2.4 (Kettenregel) $f : G \subset \mathbb{R}^n \rightarrow \tilde{G} \subset \mathbb{R}^m$ sei in $\overset{\circ}{x} \in G$ differenzierbar und ebenso $g : \tilde{G} \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$ in $f(\overset{\circ}{x}) \in \tilde{G}$. Dann ist auch $g \circ f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ in $\overset{\circ}{x}$ differenzierbar mit

$$\underbrace{D(g \circ f)}_{(p \times n)}(\overset{\circ}{x}) = \underbrace{(Dg)}_{(p \times m)}\left(f\left(\overset{\circ}{x}\right)\right) \cdot \underbrace{Df}_{(m \times n)}\left(\overset{\circ}{x}\right) \quad \text{Jacobi-Matrizen}$$

Beweis: $f(x) = f(\overset{\circ}{x}) + \Delta_1(x) \cdot (x - \overset{\circ}{x})$, $y = f(x)$ und

$$g(y) = g(\overset{\circ}{y}) + \Delta_2(y) \cdot (y - \overset{\circ}{y}), \overset{\circ}{y} = f(\overset{\circ}{x})$$

$$\Rightarrow g \circ f(x) = g \circ f(\overset{\circ}{x}) + \underbrace{\Delta_2(f(x)) - \Delta_1(x)}_{\text{Delta}(x)}(x - \overset{\circ}{x})$$

$$\text{mit } D(g \circ f)(\overset{\circ}{x}) = \Delta(\overset{\circ}{x}) = \Delta_2\left(f\left(\overset{\circ}{x}\right)\right) \Delta_1(\overset{\circ}{x}) = Dy\left(f\left(\overset{\circ}{x}\right)\right) Df\left(\overset{\circ}{x}\right).$$

Weitere Formen von Satz 2.2.4:

$$\partial_k (g \circ f)_i(\overset{\circ}{x}) = \sum_{j=1}^m \partial_j y_i\left(f\left(\overset{\circ}{x}\right)\right) \partial_k f_j\left(\overset{\circ}{x}\right)$$

oder

$$\frac{\partial}{\partial x_k} (g_i(f_1(x), \dots, f_n(x)))_{x=\overset{\circ}{x}} = \sum_{j=1}^m \frac{\partial g_i}{\partial y_j}\left(f\left(\overset{\circ}{x}\right)\right) \frac{\partial f_j}{\partial x_k}\left(\overset{\circ}{x}\right)$$

Beispiele:

- $(x, y) \mapsto g(x, y)$ sei differenzierbar und Φ durch $\Phi(r, t) = (r \cos t, r \sin t)$ definiert. Für die Funktion $(r, t) \mapsto F(r, t) := g \circ \Phi(r, t) = g(r \cos t, r \sin t)$ gilt dann
 $\partial_r F(r, t) = \partial_x g(r \cos t, r \sin t) \cos t + \partial_y g(\dots) \sin t$
 $\partial_t F(r, t) = \partial_x g(\dots) (-r \sin t) + \partial_y g(\dots) r \cos t$
 Überprüfung für $g(x, y) = x^2 + y^2$, also $F(r, t) = r^2$:
 $\partial_r F(r, t) = (2r \cos t) \cos t + (2r \sin t) \sin t = 2r$
 $\partial_t F(r, t) = (2r \cos t) (-r \sin t) + (2r \sin t) (r \cos t) = 0$
- $\partial_x (g(x^2, xy^5)) = \partial_1 g(x^2, xy^5) 2x + \partial_2 g(x^2, xy^5) y^5$
 $\partial_y (g(x^2, xy^5)) = 0 + \partial_2 g(x^2, xy^5) 5xy^4$

Satz 2.2.5 Sei $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar und $X \in \mathbb{R}^n$ ein Vektor. Dann gilt für die durch

$$d_X f(x) := \lim_{h \rightarrow 0} \frac{f(x + h \cdot X) - f(x)}{h}$$

definierte Richtungsableitung von f im Punkte x in Richtung X

$$d_X f(x) = \langle \text{grad } f(x), X \rangle$$

Beweis: Für die Funktion $h \mapsto g(h) := f(x + h \cdot X)$ gilt nach der Kettenregel $g'(0) = \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h} = \lim_{h \rightarrow 0} \frac{f(x + h \cdot X) - f(x)}{h} = d_X f(x)$

$\stackrel{!}{=} \sum_{k=1}^n \partial_k f(x) \cdot X_k = \langle \text{grad } f(x), X \rangle$
 Spezialfall: $d_{e_i} f(x) = \langle \text{grad } f(x), e_i \rangle = \partial_i f(x)$

Satz 2.2.6 Sei $f : G \subset \mathbb{R}^n \rightarrow \tilde{G} \subset \mathbb{R}^n$ ein ℓ^1 -Diffeomorphismus, also eine Bijektion dergestalt, daß f und f^{-1} ℓ^1 -Ableitungen sind. Dann ist für alle $x \in G$ die Jacobi-Matrix $Df(x)$ regulär und es gilt

$$(Df^{-1})(f(x)) = (Df(x))^{-1}$$

Beweis: $f^{-1} \circ f = id_G \Rightarrow Df^{-1}(f(x)) \cdot Df(x) = D(id_G)(x) = E$

Beispiele:

- Transformationen auf ebene/räumliche Polarkoordinaten

Satz 2.2.7 (Satz über lokale Umkehrbarkeiten) Sei $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und $\overset{\circ}{x}$ ein Punkt mit regulärer Jacobimatrix, also mit Funktionaldeterminante

$$\det Df\left(\overset{\circ}{x}\right) \neq 0$$

Dann gibt es offene Umgebungen U von $\overset{\circ}{x}$ und V von $f\left(\overset{\circ}{x}\right)$ so daß die Einschränkung $f_U : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ ein ℓ^1 -Diffeomorphismus ist. (f ist also ein lokaler Diffeomorphismus)

Gegeben sei ein differenzierbare Funktion $F : G \subset \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto F(x, y)$. Die Nullstellenmenge $N_F(G) = \{(x, y) \in G \mid F(x, y) = 0\}$ ist im allgemeinen eine Kurve in G . Damit $N_F(G)$ als Graph einer differenzierbaren Funktion $x \mapsto f(x) = y$ darstellbar ist, darf es keine Senkrechten Tangenten geben. (also $\frac{\partial F}{\partial y}(x, y) \neq 0$). Nun kann lokal nach y aufgelöst werden.

Beispiele:

- $F(x, y) = x^2 + y^2 - 1 = 0 \Leftrightarrow (x, y) = S^1$ hfill
 $\partial_y F(x, y) = 2y \neq 0 \Leftrightarrow y \neq 0$. Dann kann man lokal nach y auflösen:
 $y = f(x) = \pm\sqrt{1 - x^2}$.

Man kann sogar die Ableitung von $y = f(x)$ ausrechnen. Wegen $F(x, f(x)) = 0$ gilt nach der Kettenregel

$$\frac{d}{dx}(F(x, f(x))) = \partial_x F(x, f(x)) \cdot 1 + \underbrace{\partial_y F(x, f(x))}_{\neq 0} f'(x)$$

also

$$f'(x) = -\frac{\partial_x F(x, f(x))}{\partial_y F(x, f(x))}$$

Beispiele:

- $F(x, y) = e^{2x-3y} + 3x - 5y = 0$ kann nicht explizit nach x oder y aufgelöst werden. Es gilt aber $F(3, 2) = 0$. Wegen $\partial_y F(3, 2) = -8 \neq 0$; $\partial_x F(3, 2) = 5 \neq 0$ gibt es
 1. eine ℓ^1 -Funktion $x \in U(3) \mapsto y = f(x) \in V(2)$ mit $F(x, f(x)) = 0$ und $f'(3) = 5/8$.
 2. eine ℓ^1 -Funktion $y \in U(2) \mapsto x = g(y) \in V(3)$ mit $F(g(y), y) = 0$ und $g'(2) = 8/5$.

2.3 Lokale und globale Extrema

Definition: Lokales Minimum in $\overset{\circ}{x}$: es existiert eine Umgebung $U(\overset{\circ}{x})$ mit

$$\forall_{x \in U(\overset{\circ}{x})} f(x) \geq f(\overset{\circ}{x})$$

Bemerkung: Lokale Extrema können nur in Innenpunkten des Definitionsbereiches auftreten. Absolute Extrema (mit $\forall_{x \in D} f(x) \geq f(\overset{\circ}{x})$ bzw. $\leq f(\overset{\circ}{x})$) auch in Randpunkten eines nicht offenen Definitionsbereiches D auftreten.

Satz 2.3.1 (Notwendige Bedingung für ein lokales Extremum)

$f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sein (partiell) differenzierbar. Dann gilt

$$f \text{ besitzt in } \overset{\circ}{x} \in G \text{ ein lokales Extremum} \Rightarrow \text{grad } f(\overset{\circ}{x}) = 0$$

(Die Bedingung ist nicht hinreichend)

Beweis: f besitzt in $\overset{\circ}{x}$ ein lokales Extremum \Rightarrow Jede Partielle Funktion $x_i \mapsto f(\overset{\circ}{x}_1, \dots, x_i, \dots, \overset{\circ}{x}_n)$ besitzt auch in $\overset{\circ}{x}_i$ ein lokales Extremum \Rightarrow jede partielle Ableitung $\partial_i f(\overset{\circ}{x})$ muß 0 sein. $\Rightarrow \text{grad } f(\overset{\circ}{x}) = 0$.

Beispiele:

- $f(x, y) = x^2 - y^2$ mit $\text{grad } f(0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Aber es gilt:

$$\forall_{x \neq 0} f(x, 0) = x^2 > f(0, 0) = 0$$

$$\forall_{y \neq 0} f(0, y) = -y^2 < f(0, 0) = 0$$

Der Graph von f ist ein hyperbolisches Paraboloid. Der Punkt $(0, 0)$ ein Sattelpunkt.

Eine hinreichende Bedingung erhält man mit der HESSE-Matrix

$$H_f(x) = (\partial_k \partial_i f)(x)$$

Dazu benötigt man eine Taylorformel. Sei $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine ℓ^2 -Funktion und $\overset{\circ}{x} \in G$ beliebig zur Vereinfachung sei $\overset{\circ}{x} = 0$ (sonst Verschiebung des Koordinatensystem). Dann gibt es genau ein quadratisches Taylorpolynom

$$x = T(x) = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i,k=1}^n a_{ik} x_i x_k$$

mit $T(0) = f(0)$, $\partial_i T(0) = \partial_i f(0)$, $\partial_k \partial_i T(0) = \partial_k \partial_i f(0)$:

$$T(x) = f(0) + \sum_{i=1}^n \partial_i f(0) \cdot x_i + \frac{1}{2} \sum_{i,k=1}^n \partial_k \partial_i f(0) \cdot x_i x_k$$

$$= f(0) + \langle \text{grad } f(0), x \rangle + \frac{1}{2} x^T H_f(0) \cdot x$$

Um die Größe des Restgliedes $R(x) = f(x) - T(x)$ zu bestimmen betrachten wir für ein festes $x \in U_\delta(0) \subset G$ die Funktion einer Veränderlichen

$$t \in [0, 1] \mapsto g(t) := f(t \cdot x) \in \mathbb{R}$$

Nach der Kettenregel gilt:

$$y'(t) = \sum_{i=1}^n \partial_i f(tx) x_i, g''(t) = \sum_{i,k=1}^n \partial_k \partial_i f(tx) x_i x_k$$

Die 1-dimensionale Taylorformel (Lagrange'sche Form des Restgliedes) liefert

$$\forall t g(t) = g(0) + g'(0) t + \frac{1}{2} g''(\bar{t}) t^2 \text{ mit } \bar{t} \in [0, t]$$

also für $t = 1$

$$\begin{aligned} f(x) &= f(0) + \sum_{i=1}^n \partial_i f(0) x_i + \frac{1}{2} \sum_{i,k=1}^n \partial_k \partial_i f(\bar{x}) x_i x_k \\ &= f(0) + \langle \text{grad } f(0), x \rangle + \frac{1}{2} x^T H_f(0) x + R(x) \end{aligned}$$

mit $R(x) = \frac{1}{2} x^T (H_f(\bar{x}) - H_f(0)) x = \frac{1}{2} \sum_{i,k=1}^n (\partial_k \partial_i f(\bar{x}) - \partial_k \partial_i f(0)) x_i x_k$. Für dieses Restglied gilt

$$\lim_{x \rightarrow 0} \frac{R(x)}{|x|^2} = 0$$

denn

$$0 \leq \frac{R(x)}{|x|^2} \leq \frac{1}{2} \sum_{i,k=1}^n \underbrace{|\partial_k \partial_i f(\bar{x}) - \partial_k \partial_i f(0)|}_{\rightarrow 0} \underbrace{\frac{|x_i|}{|x|}}_{\leq 1} \underbrace{\frac{|x_k|}{|x|}}_{\leq 1}$$

Satz 2.3.2 Jede ℓ^2 -Funktion $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ besitzt um $0 \in G$ eine Taylorentwicklung

$$f(x) = f(0) + \langle \text{grad } f(0), x \rangle + \frac{1}{2} x^T H_f(0) \cdot x + R(x)$$

mit einem Restglied $x \mapsto R(x)$ mit

$$\lim_{x \rightarrow 0} \frac{R(x)}{|x|^2} = 0$$

Am Vorzeichen der HESSE-Matrix lässt sich erkennen, ob ein lokales Extremum vorliegt.

Definition: Sei $A \in M(n, n; \mathbb{R})$ eine symmetrische Matrix. Sie definiert eine quadratische Form

$$q_A : \mathbb{R}^n \rightarrow \mathbb{R}, y \mapsto q_A(y) = y^T \cdot A \cdot y = \langle y, A \cdot y \rangle = \sum_{i,k=1}^n a_{ik} y_i y_k$$

A bzw. q_A heißt

- positiv(negativ) definit, falls $y^T A y > 0 (< 0)$ für $y \neq 0$
- positiv(negativ) semidefinit, falls $y^T A y \geq 0 (\leq 0)$ für alle y
- indefinit falls $y, z \in \mathbb{R}^n$ existieren, mit $y^T A y < 0; z^T A z > 0$

Satz 2.3.3 $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sein ℓ^2 -differenzierbar und $\overset{\circ}{x} \in G$ ein kritischer Punkt von f mit $\text{grad } f(\overset{\circ}{x}) = 0$. Dann gilt:

1. $H_f(\overset{\circ}{x})$ positiv definit $\Rightarrow f$ hat in $\overset{\circ}{x}$ ein strenges lokales Minimum
2. $H_f(\overset{\circ}{x})$ negativ definit $\Rightarrow f$ hat in $\overset{\circ}{x}$ ein strenges lokales Maximum
3. $H_f(\overset{\circ}{x})$ indefinit $\Rightarrow f$ hat in $\overset{\circ}{x}$ kein lokales Extremum
4. $H_f(\overset{\circ}{x})$ echt semidefinit \Rightarrow es kann keine Aussage gemacht werden

Beweis: für $\overset{\circ}{x} = 0$

1. Es gilt $f(x) = f(0) + \frac{1}{2} x^T H_f(0) x + R(x)$ mit $\lim_{x \rightarrow 0} R(x) / |x|^2 = 0$. Da $H_f(0)$ positiv definit ist, nimmt die stetige Funktion $y \mapsto y^T H_f(0) y$ auf der kompakten Sphäre $\{y \in \mathbb{R}^n \mid |y| = 1\}$ ein positives Minimum $m > 0$ an. Wegen $\lim_{x \rightarrow 0} R(x) / |x|^2 = 0$ existiert eine δ -Umgebung um 0 mit $\forall_{x \in U \setminus \{0\}} \frac{R(x)}{|x|^2} < \epsilon := \frac{m}{2}$. Daraus folgt

$$\forall_{x \in U \setminus \{0\}} R(x) < \frac{1}{2} m |x|^2 \leq \frac{1}{2} \underbrace{\left(\left(\frac{x}{|x|} \right)^T H_f(0) \frac{x}{|x|} \right)}_{\geq m} \cdot |x|^2$$

$$= \frac{1}{2} x^T H_f(0) x \text{ d.h. für alle } x \neq 0 \text{ ist } f(x) > f(0)$$

2. analog
3. Für die Funktion $g_1 : t \mapsto f(ty)$ gilt in $U_{\text{delta}}(0)$:
 $g_1'(t) = \langle \text{grad } f(ty), y \rangle$, d.h. $g_1'(0) = 0$

$$g_1''(t) = y^T H_F(ty) y \text{ d.h. } g_1''(0) = y^T H_f(0) y < 0$$

Sie besitzt also in $t = 0$, d.h. $x = 0$, ein strenges lokales Maximum. Analog besitzt $g_2 : t \mapsto f(tz)$ wegen $g_2'(0) = 0, g_2''(0) = z^T H_f(0) z > 0$ in $x = 0$ ein strenges lokales Minimum. In $x = 0$ kann kein lokales Extremum von f vorliegen.

Beispiele:

- $f(x, y) = x^4 + y^4 \Rightarrow \text{grad } f(0, 0) = 0, H_f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow H_f$ echt semidefinit, aber $f(0, 0)$ (globales) Minimum
- $f(x, y) = x^4 - y^4 \Rightarrow \text{grad } f(0, 0) = 0, H_f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow H_f$ echt semidefinit, kein Extremum bei $(0, 0)$

Es gibt verschiedene Möglichkeiten Positive Definitheit zu erkennen

Satz 2.3.4 Eine symmetrische Matrix $A \in M(n, n; \mathbb{R})$ ist genau dann

- positiv(negativ) definit, wenn alle EW von A positiv(negativ) sind.
- positiv(negativ) semidefinit, wenn alle EW von $A \geq 0 (\leq 0)$ sind.
- indefinit, wenn positive und negative EW existieren.

Beweis: Es existieren ONB (e_1, \dots, e_n) zu EW $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Bezüglich der ONB-Darstellung $x = \sum_{i=1}^n \langle x, e_i \rangle e_i = \sum_{i=1}^n \xi_i e_i$ gilt dann

$$\begin{aligned} x^T A x &= \langle x, A x \rangle = \left\langle \sum_i \xi_i e_i, \sum_k \xi_k \underbrace{A e_k}_{\lambda_k e_k} \right\rangle \\ &= \sum_{i,k} \xi_i \xi_k \lambda_k \underbrace{\langle e_i, e_k \rangle}_{\delta_{ik}} = \sum_i \lambda_i \xi_i^2. \end{aligned}$$

Satz 2.3.5 (HURWITZ-Kriterium) Eine symmetrische Matrix $A \in M(n, n; \mathbb{R})$ ist genau dann positiv definit, wenn

$$\forall_{k=1}^n \det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0$$

und negativ definit, wenn $-A$ positiv definit ist.

Wiederholung: Spezialfall $n = 2$

$$A \sim D = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \Rightarrow \det A = \lambda_1 \lambda_2$$

$\det A < 0 \Leftrightarrow A$ indefiniert

$\det A = 0 \Leftrightarrow A$ echt semidefinit

$\det A > 0 \Leftrightarrow \begin{cases} a_{11} > 0 \Leftrightarrow A \text{ positiv definit} \\ a_{11} < 0 \Leftrightarrow A \text{ negativ definit} \end{cases}$

Gilt nur für $n = 2!$

Beispiele:

- $f(x, y) = x^3 + y^3 - 9xy + 27, G = \mathbb{R}^2$
 $\text{grad } f(x, y) = \begin{pmatrix} 3x^2 - 9y \\ 3y^2 - 9x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \end{pmatrix} \right\}$
 $H_f(x, y) = \begin{pmatrix} 6x & -9 \\ -9 & 6y \end{pmatrix}$
 $\Rightarrow H_f(0, 0) = \begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix}$ indefiniert ($\det < 0$) \Rightarrow kein lokales Extrema.
 $\Rightarrow H_f(3, 3) = \begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}$ positiv definit \Rightarrow lokales Minimum.

Bemerkung: Ist $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ auf dem kompakten Definitionsbereich D stetig, so existiert auf jeden Fall ein absolutes Minimum und Maximum. Diese Extremwerte können:

- im Inneren $\overset{\circ}{D}$ angenommen werden (und sind mit Hilfe der Differentialrechnung zu finden, falls f in $\overset{\circ}{D}$ ℓ^2 -differenzierbar) [lokale Extrema]
- oder auf dem Rand ∂D [Randextrema]

Beispiele:

- $f(x, y) = 2x^2 - y^2, D = \{(x, y) \mid x^2 + y^2 \leq 1\}$
 1. es existieren keine lokale Extrema in $\overset{\circ}{D} = \{(x, y) \mid x^2 + y^2 < 1\}$. $(0, 0)$ ist kritischer Punkt, aber Sattelpunkt.
 2. Randextrema auf $\partial D = \{(x, y) \mid x^2 + y^2 = 1\}$
 - Möglichkeit 1: Parametrisierung durch $y(x) = \pm\sqrt{1-x^2}$ und Untersuchen von $g(x) = f(x, y(x)) = 3x^2 - 1$ auf Extrema im Intervall $[-1, +1]$
Ergebnis: Absolutes Minimum in $x = 0$ ($y = \pm 1$) mit $f(0, \pm 1) = -1$ und absolutes Maximum in $x = \pm 1$ ($y = 0$) mit $f(\pm 1, 0) = 2$.
 - Möglichkeit 2: Parametrisierung durch $x(t) = \cos t, y(t) = \sin t$
Untersuchen von $h(t) := f(x(t), y(t)) = 2\cos^2 t - \sin^2 t$ auf absolute Extrema in \mathbb{R} . (es genügt $[-\pi, +\pi]$) Ist der Definitionsbereich nicht kompakt, brauchen keine absoluten Extrema zu existieren.

Manchmal sucht man Extrema einer Zielfunktion $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ unter der Nebenbedingung, daß sie Nullstellen einer zweiten Funktion $I : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sind.

Beispiele:

- $f(x, y) = x + y, G = \mathbb{R}^2, F(x, y) = x^2 + y^2 - 1 = 0 \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} \in S^1$

Sind es mehrere Nebenbedingungen kann man sie durch eine Abbildung $F : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ beschreiben.

Satz 2.3.6 (LAGRANGEsche Multiplikatoren) Seien $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ und $F : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($m < n$) ℓ^1 -Abbildungen und $\overset{\circ}{x} \in G$ ein Punkt mit $F(\overset{\circ}{x}) = 0$ und $\text{rg } DF(\overset{\circ}{x}) = m$. Ist dann f in $\overset{\circ}{x}$ lokal extremal unter der NB $F \equiv 0$, so existieren Lagrangesche Multiplikatoren $\lambda_1, \dots, \lambda_m$ mit

$$\text{grad } f(\overset{\circ}{x}) = \sum_{j=1}^m \lambda_j \text{grad } F_j(\overset{\circ}{x})$$

(Das ist nur eine notwendige Bedingung)

Möglicher Lösungsweg: Man löse im Bereich $G' = \{x \in G \mid \text{rg } DF(x) = m\}$ das Gleichungssystem

$$\left\{ \begin{array}{l} F(x) = 0 \\ \text{grad } f(x) = \sum_{j=1}^m \lambda_j \text{grad } F_j(x) \end{array} \right\} \begin{array}{l} \text{(m Gleichungen)} \\ \text{(n Gleichungen)} \end{array}$$

für die $n+m$ Variablen $x_1, \dots, x_n, \underbrace{\lambda_1, \dots, \lambda_m}_{\text{weniger wichtig}}$. Die Lösungen $x = (x_1, \dots, x_n)^T$

sind dann die Kandidaten für lokale Extrema von f unter der NB $F \equiv 0$ in G' .

Beispiele:

- $f(x, y, z) = x + y + z, F(x, y, z) = x^2 + y^2 + z^2 - 1 = 0 (\in S^2)$
 $\text{rg } DF(x, y, z) = \text{rg } (2x, 2y, 2z) = 1$ für $(x, y, z) \neq (0, 0, 0) \notin S^2$ zu lösen

ist

$$\left\{ \begin{array}{l} x^2 + y^2 + z^2 = 1 \\ 1 = \lambda \cdot 2x \\ 1 = \lambda \cdot 2y \\ 1 = \lambda \cdot 2z \end{array} \right\} \Rightarrow$$

$$x = y = z = \frac{1}{2\lambda}, x^2 + y^2 + z^2 = \frac{3}{4\lambda^2} = 1$$

$$x = y = z = \pm \frac{1}{\sqrt{3}} \text{ kritische Punkte}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = +\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} = -\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Da f auf S^2 ein absolutes Minimum und Maximum annimmt, besitzt f

wirklich in $\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ein absolutes Maximum und in $-\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ein abso-

lutes Minimum.

2.4 Parameterabhängige Riemann-Integrale

Manchmal hängt ein R-Integral $\int_a^b f(t, x) dt$ von einem oder mehreren Parametern $x = (x_1, \dots, x_n)$ ab und definiert eine Funktion $x \mapsto F(x) = \int_a^b f(t, x) dt$. Dann kann man nach diesen Parametern differenzieren.

Satz 2.4.1 Sei $f : [a, b] \times G \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$, $(t, x) \mapsto f(t, x)$ stetig und besitze die partiellen Ableitungen $(t, x) \mapsto \frac{\partial f}{\partial x_k}(t, x)$. Dann ist $x \in G \mapsto$

$F(x) := \int_a^b f(t, x) dt \in \mathbb{R}$ stetig (partiell) differenzierbar mit

$$\forall x \in G \quad \frac{\partial F}{\partial x_k}(x) = \frac{\partial}{\partial x_k} \int_a^b f(t, x) dt = \int_a^b \frac{\partial f}{\partial x_k}(t, x) dt$$

$$F(x) = \int_{a(x)}^{b(x)} f(t, x) dt$$

Voraussetzung: a, b, f stetig differenzierbar. Betrachte die Funktion

$$G(x, y, z) := \int_y^z f(t, x) dt$$

mit $\partial_x G(x, y, z) = \int_y^z \partial_x f(t, x) dt$, $\partial_z G(x, y, z) = f(z, x)$ und $\partial_y G(x, y, z) = -f(y, x)$. Für $F(x) = G(x, a(x), b(x))$ gilt dann nach der Kettenregel.
 $F'(x) = \partial_x G(x, a(x), b(x)) \cdot 1 + \partial_y G(x, a(x), b(x)) \cdot a'(x) + \partial_z G(x, a(x), b(x)) \cdot b'(x)$
 $= \int_{a(x)}^{b(x)} \partial_x f(t, x) dt + f(b(x), x) b'(x) - f(a(x), x) a'(x)$

2.5 Integralrechnung in mehreren Veränderlichen

2.5.1 Doppelintegrale über Rechtecke

Der Riemannsche Integralbegriff aus Kapitel ?? kann problemlos auf beschränkte Funktionen $f : Q \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ auf einem (abgeschl.) Rechteck $G : I_1 \times I_2 = [a, b] \times [c, d]$ übertragen werden. Gesucht ist jetzt das (vorzeichenbehaftete) Volumen der Ordinatenmenge $\{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in Q, z \in \overline{Of(x, y)}\}$.

Hilfsmittel für die Berechnung dieser Integrale:

Produktzerlegungen: $z = z_1 \times z_2$ von Q in Teilrechtecke $Q_{ik} = [x_{i-1}, x_i] \times [y_{k-1}, y_k]$ mit Flächeninhalt

$$|Q_{ik}| = \Delta x_i \cdot \Delta y_k$$

Riemansche Ober- und Untersummen:

$$\overline{R}_f(z) = \sum_{i,k} M_{ik} \cdot \Delta x_i \cdot \Delta y_k \text{ mit } M_{ik} = \sup \{f(x, y) \mid (x, y) \in Q_{ik}\}$$

$$\underline{R}_f(z) = \sum_{i,k} m_{ik} \cdot \Delta x_i \cdot \Delta y_k \text{ mit } m_{ik} = \inf \{f(x, y) \mid (x, y) \in Q_{ik}\}$$

Varianz: $V_f(z) = \overline{R}_f(z) - \underline{R}_f(z)$

f heißt R-integrierbar, wenn $\underline{R}_f = \sup_z \underline{R}_f(z) \stackrel{!}{=} \inf_z \overline{R}_f(z) = \overline{R}_f$ gilt bzw. wenn es zu jedem $\epsilon > 0$ eine Zerlegung z von Q gibt mit $V_f(z) = \overline{R}_f(z) - \underline{R}_f(z) < \epsilon$.

Der gemeinsame Wert vom Riemannschen Ober- und Unterintegral $\overline{R}_f = \underline{R}_f$ heißt R-Integral von f über Q bezeichnet mit

$$\int_Q f(x, y) d(x, y)$$

wieder kann man zeigen: Jede stetige Funktion $f : Q \rightarrow \mathbb{R}$ ist R-Integrierbar und es gelten die üblichen Rechenregeln wie z.B.

- f über Q R-integrierbar $\Rightarrow |f|$ über Q R-integrierbar mit $\left| \int_Q f(x, y) d(x, y) \right| \leq \int |f(x, y)| dy$.
- f, g über Q R-integrierbar $\Rightarrow f + g$ über Q R-integrierbar.

Satz 2.5.1 (Satz von FUBINI für Rechtecke) $f : Q \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $Q = [a, b] \times [c, d]$ sei R-integrierbar und für alle $y \in [c, d]$ existiert $F(y) = \int_a^b f(x, y) dx$. Dann ist auch F über $[c, d]$ R-integrierbar und es gilt

$$\int_Q f(x, y) d(x, y) = \int_c^d F(y) dy = \int_c^d \int_a^b f(x, y) dx dy$$

Folgerung: Für $Q = [a, b] \times [c, d]$ und eine stetige Funktion f ist

$$\int_Q f(x, y) d(x, y) = \int_c^d \int_a^b f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx$$

Beispiele:

- $f(x, y) = e^{x^2} \sin y, Q = [0, 1] \times [0, 2\pi]$
 $\int_Q f(x, y) d(x, y) = \int_0^{2\pi} \int_0^1 e^{x^2} \sin y dx dy = \int_0^1 \int_0^{2\pi} e^{x^2} \sin y dy dx$
 $= \int_0^1 e^{x^2} (\cos 0 - \cos 2\pi) dx = 0$

2.5.2 Doppelintegrale über Normalbereiche

Ziel ist die Erweiterung des Integralbegriffs auf beschränkte Funktionen f auf nicht notwendig rechteckige beschränkte Bereiche $B \subset \mathbb{R}^2$. Die lässt sich durch die Triviale Fortsetzung von f auf ein Rechteck $Q \supset B$ durch

$$\chi_B f(x, y) := \begin{cases} f(x, y) & \text{für } (x, y) \in B \\ 0 & \text{für } (x, y) \in Q \setminus B \end{cases}$$

Bemerkung: Die Funktion $\chi_B(x, y) = \begin{cases} 1 & \text{für } (x, y) \in B \\ 0 & \text{für } (x, y) \in Q \setminus B \end{cases}$ heißt charakteristische Funktion von B .

Definition: Eine beschränkte Funktion $f : B \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ auf einem beschränkten Bereich B heißt R-integrierbar, wenn ihre triviale Fortsetzung $\chi_B f : Q \rightarrow \mathbb{R}$ auf einem Rechteck $Q \supset B$ R-integrierbar ist und es sei

$$\int_B f(x, y) d(x, y) = \int_Q \chi_B f(x, y) d(x, y)$$

Bemerkung: Die Integrierbarkeit von f über B hängt damit von

- der Funktion f („Höhe“) und
- dem Bereich B („Grundfläche“)

ab. I.a. ist $\chi_B f$, auch wenn f stetig ist, nicht stetig.

Für $f \equiv 1$ erhält man nach der Regel Volumen = Grundfläche · Höhe

$$\int_B 1 d(x, y) = a(B) \cdot 1$$

Definition: Ein Bereich $B \subset \mathbb{R}^2$ der Form

$$B = \{(x, y) \mid a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\}$$

mit Funktion $\varphi, \psi : [a, b] \rightarrow \mathbb{R}$ heißt ein Normalbereich bzgl. y . Analog ist der Normalbereich bzgl. x mit

$$B' = \left\{ (x, y) \mid c \leq y \leq d, \tilde{\varphi}(x) \leq y \leq \tilde{\psi}(x) \right\}$$

definiert.

Satz 2.5.2 Sei $f : B \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ stetig auf dem Normalbereich $B = \{(x, y) \mid a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\}$ mit stetigen Funktionen $\varphi, \psi : [a, b] \rightarrow \mathbb{R}$. Dann ist f über B R -integrierbar mit

$$\int_Q f(x, y) d(x, y) = \int_a^b \int_{\varphi(x)}^{\psi(x)} f(x, y) dy dx$$

Für den Spezialfall mit $f \equiv 1$ ist der Flächeninhalt eines Normalbereiches

$$a(B) = \int_a^b (\psi(x) - \varphi(x)) dx = \int_a^b \int_{\varphi(x)}^{\psi(x)} 1 \cdot dy dx$$

Beispiele:

- Die Kreisscheibe $B = \{(x, y) \mid x^2 + y^2 \leq R^2\}$ ist ein Normalbereich, denn $B = \{(x, y) \mid -R \leq x \leq R, -\sqrt{R^2 - x^2} \leq y \leq +\sqrt{R^2 - x^2}\}$. also gilt $a(B) = \int_{-R}^R 2\sqrt{R^2 - x^2} dx = 2 \frac{1}{2} [R\sqrt{R^2 - x^2} + R^2 \text{Arcsin} \frac{x}{R}]_{-R}^{+R} = R^2 (\frac{\pi}{2} + \frac{\pi}{2}) = \pi R^2$
- Für $f(x, y) = \sqrt{R^2 - x^2 - y^2}$ liefert $\int_Q f(x, y) d(x, y)$, B wie oben, das Volumen der Halbkugel mit Radius R . Es gilt $V(H_R) = \int_{-R}^{+R} \int_{-\sqrt{R^2 - x^2}}^{+\sqrt{R^2 - x^2}} \sqrt{R^2 - x^2 - y^2} dy dx = \dots = \frac{2}{3} \pi R^3$

2.5.3 Der Transformationssatz für Doppelintegrale

Der Transformationssatz ist eine Verallgemeinerung der Substitutionsregel

$$\int_{\varphi(a)}^{\varphi(b)} f(x) dx = \int_a^b f(\varphi(t)) \varphi'(t) dt$$

auf Doppelintegrale.

Ist der Definitionsbereich B' einer Funktion f „kompliziert“, kann man versuchen, ihn als Bild $B' = \Phi[B]$ eines einfacheren Bereiches B unter einer Transformationsabbildung

$$\Phi : B \subset \mathbb{R}^2 \rightarrow B' \subset \mathbb{R}^2$$

darzustellen.

Satz 2.5.3 (Transformationsatz für Doppelintegrale)

Sei $\Phi : G \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2, (u, v) \mapsto (x, y) = \Phi(u, v) = (\Phi_1(u, v), \Phi_2(u, v))$ auf dem Gebiet G ℓ^1 -differenzierbar und auf dem Bereich $B \subset G$ injektiv und regulär, d.h. $\det D\Phi(u, v) \neq 0$ für alle $(u, v) \in B$. Weiter sei $f : B' = \Phi[B] \rightarrow \mathbb{R}$ stetig. Dann gilt

$$\int_{\Phi[B]} f(x, y) d(x, y) = \int_B f(\Phi(u, v)) |\det D\Phi(u, v)| d(u, v)$$

Bemerkung: Die Injektivität und Regularität von Φ kann auf dem Rande von B und Einzelpunkten (auf „Nullmengen“) verletzt sein.

Beispiele:

- **Ebene Polarkoordinaten:** Die Transformation $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (r, \varphi) \mapsto (x, y) = (r \cos \varphi, r \sin \varphi)$ ist ℓ^1 -differenzierbar mit Funktionaldeterminante $D\Phi(r, \varphi) = r$

- Für $r = 0$ und $\varphi = \pm\pi$ ist die Injektivität verletzt
- Für $r = 0$ ist Φ nicht regulär.

Dies sind jedoch nur Randpunkte. Auf $Q = [0, R] \times [-\pi, +\pi]$ (mit $\Phi[Q] = K_R = \{(x, y) \mid x^2 + y^2 \leq R^2\}$) ist Φ injektiv (bis auf $r = 0, \varphi = \pm\pi$) und regulär (bis auf $r = 0$). Also gilt für eine stetige Funktion $f : K_R \rightarrow \mathbb{R}$:

$$\begin{aligned} \int_{K_R} f(x, y) d(x, y) &= \int_Q f(\Phi(r, \varphi)) |\det D\Phi(r, \varphi)| d(r, \varphi) \\ &= \int_{-\pi}^{+\pi} \int_0^R f(r \cos \varphi, r \sin \varphi) r dr d\varphi \end{aligned}$$

$$- f \equiv 1: a(K_R) = \int_0^R \int_{-\pi}^{+\pi} 1 \cdot r d\varphi dr = 2\pi \cdot \frac{1}{2}R^2 = \pi R^2$$

$$- F(x, y) = \sqrt{R^2 - x^2 - y^2}:$$

$$\begin{aligned} V(H_R) &= \int_{K_R} f(x, y) d(x, y) = \int_{-\pi}^{+\pi} \int_0^R \sqrt{R^2 - r^2} r dr d\varphi \\ &= 2\pi - \frac{1}{2} \int_0^R \sqrt{R^2 - r^2} (-2r) dr = -\pi \left[\frac{2}{3} \sqrt{R^2 - r^2}^3 \right]_0^R \\ &= \pi \frac{2}{3} R^3 \end{aligned}$$

$$- f(x, y) = e^{-(x^2+y^2)}:$$

$$\begin{aligned} \int_{K_R} f(x, y) d(x, y) &= 2\pi \int_0^R e^{-r^2} \cdot r dr = -\pi \left[e^{-r^2} \right]_0^R \\ &= \pi \left(1 - e^{-R^2} \right) \end{aligned}$$

$$- \text{Für } Q_R = [-R, +R] \times [-R, +R] \text{ gilt } K_R < Q_R < K_{\sqrt{2}R}$$

$$\begin{aligned} \int_{Q_R} f(x, y) d(x, y) &= \int_{-R}^{+R} \int_{-R}^{+R} e^{-x^2} e^{-y^2} dx dy = \left(\int_{-R}^{+R} e^{-x^2} dx \right)^2 \\ \Rightarrow \pi \left(1 - e^{-R^2} \right) &\leq \left(\int_{-R}^{+R} e^{-x^2} dx \right)^2 \leq \pi \left(1 - e^{-2R^2} \right) \\ R \rightarrow \infty &:\Rightarrow \left(\int_{-R}^{+R} e^{-x^2} dx \right)^2 = \sqrt{\pi} \end{aligned}$$

2.5.4 Dreifachintegrale

Zunächst sind Dreifachintegrale analog zu Doppelintegralen für beschränkte Funktionen $f : Q \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ auf (abgeschlossenen) Quadern $Q = I_1 \times I_2 \times I_3 \subset \mathbb{R}^3 : \int_Q f(x, y, z) d(x, y, z)$ erklärt. (Hilfsmittel: Produktzerlegungen der Quader in Teilquader, Riemansche Ober- und Untersummen...). Der Satz von FUBINI für Quader $Q = [a, b] \times [c, d] \times [e, f]$ lautet dann

$$\int_Q f(x, y, z) dV = \int_e^f \int_c^d \int_a^b f(x, y, z) dx dy dz$$

(unabhängig von der Reihenfolge, falls f stetig; 4-dimensional)

Für eine Funktion $f : B \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ auf einem beschränkten Bereich $B \subset \mathbb{R}^3$ sei wieder

$$\int_B f(x, y, z) dV := \int_Q \chi_B f(x, y, z) dV$$

und die triviale Fortsetzung $\chi_B f$ von f auf einem Quader $Q \supset B$ $V(B) = \int_B 1 dV$ ist das (3-dimensionale) Volumen von B . Harmlose Bereiche sind wieder Normalbereiche der Form

$B = \{(x, y, z) \mid a \leq x \leq b, \varphi(x) \leq y \leq \psi(x), \alpha(x, y) \leq z \leq \beta(x, y)\}$. Für sie gilt

$$\int_B f(x, y, z) dV = \int_a^b \int_{\varphi(x)}^{\psi(x)} \int_{\alpha(x, y)}^{\beta(x, y)} f(x, y, z) dz dy dx$$

Beispiele:

- $B = \{(x, y, z) \mid 0 \leq x \leq z, 0 \leq y \leq x, 0 \leq zx + y + 1\}$
 $V(B) = \int_{x=0}^2 \int_{y=0}^x \int_{z=0}^{x+y+1} dz dx dy = \int_0^2 \int_0^x (x+y+1) dy dx$
 $= \int_0^2 \left[xy + \frac{y^2}{2} + y \right]_0^x dx = \left[\frac{1}{2}x^3 + \frac{1}{2}x^2 \right]_0^2 = 6$

Satz 2.5.4 (Prinzip von CAVALIERI) Sei $B \subset \mathbb{R}^3$ ein beschränkter Bereich mit $x \in [a, b]$ für alle $(x, y, z) \in B$ und für alle $x \in [a, b]$ existiert der Flächeninhalt $a(B_x)$ der Schnitte $B_x = \{(y, z) \in \mathbb{R}^2 \mid (x, y, z) \in B\}$. Dann gilt

$$V(B) = \int_a^b a(B_x) dx$$

Folgerung: Gilt $\forall_x a(B_x) = a(\tilde{B}_x)$, so ist auch $V(B) = V(\tilde{B})$.

Beispiele:

- Kugel $B: K_R = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq R^2\}$
 $B_x = \{(y, z) \mid y^2 + z^2 \leq R^2 - x^2\}$ (Kreis mit Radius $R_x = \sqrt{R^2 - x^2}$)
 $V(K_R) = \int_{-R}^{+R} R_x^2 \pi dx = \pi \int_{-R}^{+R} (R^2 - x^2) dx = \pi \left[R^2 x - \frac{1}{3} x^3 \right]_{-R}^{+R}$
 $= \frac{4}{3} \pi R^3$

Für Dreifachintegrale ist der Transformationssatz mit

$$\int_{\Phi[B]} f(x, y, z) d(x, y, z) = \int_B f(\Phi(u, v, w)) |\det D\Phi(u, v, w)| d(u, v, w)$$

wenn $\Phi : B \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3, (u, v, w) \mapsto (x, y, z) = \Phi(u, v, w)$ eine geeignete Transformation ist, gegeben.

Spezialfälle

1. **Zylinderkoordinaten:** $(r, \varphi, z) \mapsto \Phi(r, \varphi, z) = (r \cos \varphi, r \sin \varphi, z)$ mit $\det D\Phi(r, \varphi, z) = r$

Beispiele:

- $f \equiv 1$:

$$V(B) = \int_B 1 dV = \int_0^H \int_0^R \int_{-\pi}^{+\pi} 1 \cdot r d\varphi dr dz = \pi R^2 H$$

2. **Kugelkoordinaten:** (räumliche Polarkoordinaten)

$(r, \varphi, \vartheta) \mapsto \Phi(r, \varphi, \vartheta) = (r \cos \varphi \cos \vartheta, r \sin \varphi \cos \vartheta, r \sin \vartheta)$ mit $\det D\Phi(r, \varphi, \vartheta) = r^2 \cos \vartheta$

Beispiele:

- $f \equiv 1$:

$$V(K_R) = \int_{K_R} 1 dV = \int_0^R \int_{-\pi}^{+\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} 1 \cdot r^2 \cos \vartheta d\vartheta d\varphi dr = \frac{4}{3} \pi R^3$$

2.5.5 Anwendung und Beispiele

Rotationskörper

$z \in [a, b] \mapsto r(z) \geq 0$ erzeugt durch Rotation um die z -Achse einen Rotationskörper

$$B = \{(x, y, z) \mid a \leq z \leq b, x^2 + y^2 \leq r^2(z)\}$$

Nach CAVALIERI gilt

$$V(B) = \int_a^b a(B_z) dz = \pi \int_a^b r^2(z) dz$$

Beispiele:

- $r(z) = \frac{1}{z}, z \in [1, H]$
 $V(B) = \pi \int_1^H \frac{1}{z^2} dz = \pi \left[-\frac{1}{z} \right]_1^H = \pi \left(1 - \frac{1}{H} \right), \lim_{H \rightarrow \infty} \pi \left(1 - \frac{1}{H} \right) = \pi$

- Für $B = \{(x, y, z) \mid z \in [c, d], x^2 + 3y^2 \leq r^2(z)\}$ sind Scheiben B_z (Voll-) Ellipsen mit den Halbachsen $a = r(z)$ und $b = \frac{1}{\sqrt{3}}r(z)$, da $\left(\frac{x}{r(z)}\right)^2 + \left(\frac{y}{\frac{1}{\sqrt{3}}r(z)}\right)^2 \leq 1$. Es gilt also $a(B_z) = \pi ab = \pi \frac{1}{\sqrt{3}}r^2$ und

$$V(B) = \frac{\pi}{\sqrt{3}} \int_c^d r^2(z) dz$$

Schwerpunkte

Für die Koordinaten des Schwerpunktes eines Körpers konstanter Dichte gilt

$$s_x = \frac{1}{V(B)} \int_B x dV, \dots, s_z = \frac{1}{V(B)} \int_B z dV$$

Beispiele:

- **Holzseit(Halbkreis als Grundfläche):** In Zylinderkoordinaten mit $0 \leq r \leq R, 0 \leq \varphi \leq \pi, 0 \leq z \leq H$. Daher ist $s_z = \frac{1}{2}H$ und $s_x = 0$. Das Volumen ist mit $V(B) = \frac{\pi}{2}R^2H$ gegeben.

$$s_x = \frac{1}{V(B)} \int_0^R \int_0^\pi \int_0^H r^2 \cos \varphi dz d\varphi dr = 0$$

$$s_y = \frac{1}{V(B)} \int_0^R \int_0^\pi \int_0^H r^2 \sin \varphi dz d\varphi dr = \frac{1}{V(B)} \frac{R^3}{3} H [-\cos \varphi]_0^\pi = \frac{4}{3\pi} R$$

- **Halbkugel:** Aufgrund der Symmetrie ist $s_x = s_y = 0$.

$$s_z = \frac{1}{V(B)} \int_0^{\pi/2} \int_0^{2\pi} \int_0^R \underbrace{r \sin \vartheta}_z \underbrace{r^2 \cos \vartheta}_{\det D\Phi} dr d\varphi d\vartheta = \frac{1}{V(B)} \frac{\pi R^4}{2} \left[\frac{1}{2} \sin^2 \vartheta \right]_0^{\pi/2} = \frac{3}{8} R$$

- **Kreisegel:** $0 \leq \varphi \leq 2\pi, 0 \leq z \leq H$ und $0 \leq r \leq \frac{R}{H}z$. Das Volumen ist $V(b) = \frac{\pi}{3}R^2H$.

$$s_z = \frac{1}{V(B)} \int_0^{2\pi} \int_0^H \int_0^{z \cdot R/H} zr dr dz d\varphi = \frac{2\pi}{V(B)} \int_0^H \frac{R^2}{2H^2} z^3 dz = \frac{3}{4} H$$

Bilder von Rechtecken

Beispiele:

- Gegeben ist $B = \{(x^2 - y^2, 2xy) \mid x, y \in [0, 1]\}$. Gesucht ist der Flächeninhalt $a(B)$. B kann als das Bild $B = f[Q]$ von $Q = [0, 1] \times [0, 1]$ unter der Abbildung $f(x, y) = (x^2 - y^2, 2xy) = (u, v)$ aufgefasst werden. f ist injektiv und wegen $\det Df(x, y) = \begin{vmatrix} 2x & -2y \\ 2y & 2x \end{vmatrix} = 4(x^2 + y^2)$ regulär. (bis auf $(0, 0)$). Die Transformationsformel liefert $a(B) = \int_{f[Q]} 1 d(u, v) = \int_Q |\det Df(x, y)| d(x, y) = \int_0^1 \int_0^1 4(x^2 + y^2) dx dy = \frac{8}{3}$