Teil D

Lineare Algebra

Kapitel 34

Vektorräume

34.1 Motivation

- Im $\mathbb{R}^2$ und $\mathbb{R}^3$ kann man Vektoren addieren und mit einem Skalar multiplizieren.
- Wir wollen dieses Grundkonzept algebraisch formalisieren, um es auch auf andere Situationen anwenden zu können.
- Wichtig z.B. in der Robotik, der Codierungtheorie, in Computergrafik und Computer Vision.

34.2 Definition (K-Vektorraum)

Sei K ein Körper. Ein K-Vektorraum ist eine Menge V, auf der eine Verknüpfung $+ : V \times V \to V$ und eine skalare Multiplikation $\cdot : K \times V \to V$ definiert sind mit

(a) $(V, +)$ *ist eine kommutative Gruppe.*

(b) $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v \qquad \forall \lambda, \mu \in K, \forall v \in V$

(c) $1 \cdot v = v \qquad \forall v \in V$

(d) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w \qquad \forall \lambda \in K, \forall v, w \in V$

(e) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v \qquad \forall \lambda, \mu \in K, \forall v \in V$

Die Elemente von V heißen Vektoren, Elemente aus K heißen Skalare. Ist $K = \mathbb{R}$ oder $K = \mathbb{C}$, sprechen wir von einem reellen bzw. komplexen Vektorraum.

Bemerkung: Für Skalare verwenden wir meist kleine griechische Buchstabenb wie λ, μ, ν. Vektoren bezeichnen wir mit kleinen lateinischen Buchstaben wie u, v, w. Nur wenn Verwechslungsgefahr besteht, verwenden wir Pfeile auf Vektoren, z.B. um den Nullvektor $\vec{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ von der skalaren Null 0 zu unterscheiden.

34.3 Satz: (Rechenregeln für Vektorräume)

In einem K-Vektorraum V gilt:

(a) $\lambda \cdot \vec{0} = \vec{0} \qquad \forall \lambda \in K.$

(b) $0 \cdot v = \vec{0} \qquad \forall v \in V.$

(c) $(-1)v = -v \qquad \forall v \in V.$

Beweis:

(von a)

$$\lambda \cdot \vec{0} \overset{\vec{0}=\vec{0}+\vec{0}}{=} \lambda \cdot (\vec{0} + \vec{0}) \overset{34.2(d)}{=} \lambda \cdot \vec{0} + \lambda \cdot \vec{0}$$

Subtrahiert man auf beiden Seien $\lambda \cdot \vec{0}$, folgt $\vec{0} = \lambda \cdot \vec{0}$. $\square$.

34.4 Beispiele

(a) $\mathbb{R}^n$ ist ein $\mathbb{R}$-Vektorraum für alle $n \in \mathbb{N}$.

(b) $\mathbb{C}^n$ ist ein $\mathbb{C}$-Vektorraum für alle $n \in \mathbb{N}$.

(c) $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum.

(d) Für einen Körper K ist K^n für alle $n \in \mathbb{N}$ ein K-Vektorraum.

(e) Insbesondere ist $\mathbb{Z}_2^n$ ein Vektorraum über dem Körper $\mathbb{Z}_2$. Er besteht aus allen n-Tupeln von Nullen und Einsen. Beispielsweise lassen sich Integer-Werte in Binärdarstellung als Element des Vektorraums $\mathbb{Z}_2^{32}$ darstellen.
Die Vektorräume $\mathbb{Z}_2^n$ spielen eine große Rolle in der Codierungstheorie. Lineare Codes verwenden Teilmengen des $\mathbb{Z}_2^n$, bei denen beim Auftreten eines Übertragungsfehlers mit hoher Wahrscheinlichkeit ein Element außerhalb der Teilmenge entsteht (vgl. 11.3 auf Seite 91).

(f) Funktionsräume sind wichtige Vektorräume. Definiert man auf

$$\mathcal{F} \quad := \quad \{f \mid \mathbb{R} \to \mathbb{R}\}$$

eine Vektoraddition und eine skalare Multiplikation durch

$$\begin{aligned} (f+g)(x) &:= f(x) + g(x) \qquad f, g \in \mathcal{F} \\ (\lambda \cdot g)(x) &:= \lambda \cdot g(x) \qquad \lambda \in \mathbb{R}, \forall g \in \mathcal{F} \end{aligned}$$

so ist $\mathcal{F}$ ein $\mathbb{R}$-Vektorraum. Der Nullvektor $\vec{0}$ ist die Funktion $f(x) = 0 \;\; \forall x \in \mathbb{R}$.

(g) Die Polynome $K[x]$ über einem Körper K bilden einen K-Vektorraum, wenn man als skalare Multiplikation definiert:

$$\lambda \sum_{k=0}^{n} a_k x^k \quad := \quad \sum_{k=0}^{n} \lambda \cdot a_k x^k \qquad \forall \lambda \in K.$$

Zu Gruppen und Ringe haben wir in 8.5 auf Seite 74 und 9.5 auf Seite 80 Untergruppen und Unterringe definiert. Ähnliches ist auch für Vektorräume möglich.

34.5 Definition (Unterraum, Untervektorraum, Teilraum)

Sei V ein K-Vektorraum und $U \subset V$. Ist U mit den Verknüpfungen von V selbst wieder ein K-Vektorraum, so heißt U Unterraum (Untervektorraum, Teilraum) von V.

34.6 Satz: (Unterraumkriterium)

Sei V ein K-Vektorraum und $U \subset V$ eine nichtleere Teilmenge. Dann ist U genau dann ein Unterraum von V, wenn gilt:

(*a*) $u + v \in U \qquad \forall u, v \in U$

(*b*) $\lambda \cdot u \in U \qquad \forall u \in U, \forall \lambda \in K.$

Beweis:

„$\Rightarrow$" Offensichtlich, da U selbst ein Vektorraum ist.

„$\Leftarrow$" Nach 8.5 auf Seite 74 ist U Untergruppe von V, denn

$$\begin{array}{rcl} u+v \in U \quad \forall u,v & \in & U \\ -U \overset{34.3\ auf der vorherigen Seite (c)}{=} (-1)\cdot u & \overset{(b)}{\in} & U \end{array}$$

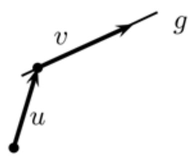
Da V kommutativ ist, ist U eine kommutative Gruppe, d.h. 34.2 (a) ist erfüllt.

Aus (b) folgt, dass $\cdot$ eine Abbildung von $K \times U$ nach $\underline{U}$ ist. Die Eigenschaften (b)-(e) der Vektorraumdefinition 34.2 übertragen sich von der Vektorraumeigenschaft von V.

34.7 Beispiele

(a) Ein K-Vektorraum hat die trivialen Unterräume $\{0\}$ und V.

(b) Lineare Codes sind Unterräume im $\mathbb{Z}_2^n$.



(c) Sind Geraden Unterräume des $\mathbb{R}^2$? Ein Unterraum muss stets den Nullvektor enthalten. Daher bilden nur die Ursprungsgeraden $\{\lambda v \mid \lambda \in \mathbb{R}\}$ Unterräume.

(d) Verallgemeinerung von (c):
In einem K-Vektorraum V bilden die Mengen $\{\lambda v \mid \lambda \in K\}$ Unterräume.

Das letzte Beispiel läßt sich noch weiter verallgemeinern:

34.8 Definition (Linearkombination,Erzeugnis)

Sei V ein K-Vektorraum. Ferner seien $u_1, \ldots, u_n \in V$ und $\lambda_1, \ldots, \lambda_n \in K$. Dann nennt man $\sum_{k=1}^{n} \lambda_k u_k$ eine Linearkombination von $u_1, \ldots, u_n$. Die Menge aller Linearkombinationen bildet das Erzeugnis $\operatorname{span}(u_1, \ldots, u_n)$, und man nennt $\{u_1, \ldots, u_n\}$ das Erzeugendensystem von $\operatorname{span}(u_1, \ldots, u_n)$.

34.9 Satz: (Erzeugnis als Unterraum)

Sei V ein K-Vektorraum und seien $u_1, \dots, u_n \in V$. Dann bildet $\operatorname{span}(u_1, \dots, u_n)$ einen Unterraum von V.

Beweis:

Wir wenden das Unterraumkriterium an:

(a) Seien $v, w \in \operatorname{span}(u_1, \dots, u_n) \Rightarrow \exists \lambda_1, \dots, \lambda_n$ und $\mu_1, \dots, \mu_n \in K$ mit

$$v = \sum_{i=1}^{n} \lambda_i u_i, \qquad w = \sum_{i=1}^{n} \mu_i u_i$$

$$\begin{aligned} \Rightarrow v + w &= \sum_{i=1}^{n} \lambda_i u_i + \sum_{i=1}^{n} \mu_i u_i \\ &\overset{34.2(e)}{=} \sum_{i=1}^{n} (\lambda_i + \mu_i) u_i \in \operatorname{span}(u_1, \dots, u_n). \end{aligned}$$

(b) Sei $\mu \in K$ und $u \in \operatorname{span}(u_1, \dots, u_n)$.

$$\begin{aligned} \Rightarrow \exists \lambda_1, \dots, \lambda_n \in K : u &= \sum_{i=1}^{n} \lambda_i u_i \\ \Rightarrow \mu u &= \mu \sum_{i=1}^{n} \underbrace{\lambda_i u_i}_{\in V} \\ &\overset{34.2(d)}{=} \sum_{i=1}^{n} \mu (\lambda_i u_i) \\ &\overset{34.2(b)}{=} \sum_{i=1}^{n} \underbrace{(\mu \lambda_i)}_{\in K} u_i \in \operatorname{span}(u_1, \dots, u_n). \end{aligned}$$

$\square$.

34.10 Beispiel:

Im $\mathbb{R}^3$ bildet die Menge aller Ebenen durch den Ursprung einen Unterraum.

34.11 Lineare Abhängigkeit

Offenbar ist $\operatorname{span}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right) = \mathbb{R}^3 = \operatorname{span}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right)$,
d.h. $\begin{pmatrix}1\\1\\0\end{pmatrix}$ läßt sich als Linearkombination von von $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$ darstellen.

Definition: *(linear abhängig,linear unabhängig)*

Ein Vektor u heißt linear abhängig von $u_1, \dots, u_n$, wenn es $\lambda_1, \dots, \lambda_n \in K$ gibt mit $\sum_{i=1}^{n} \lambda_i v_i$.

Eine Menge von Vektoren $u_1, \ldots, u_n$, bei denen sich keiner der Vektoren als Linearkombination der anderen ausdrücken lässt, heißt linear unabhängig.

Gibt es ein einfaches Kriterium, mit dem man lineare Unabhängigkeit nachweisen kann?

34.12 Satz: (Kriterium für lineare Unabhängigkeit)

Sei V ein K-Vektorraum. Die Vektoren $v_1, \ldots, v_n \in V$ sind genau dann linear unabhängig, wenn für jede Linearkombination mit $\sum_{i=1}^{n} \lambda_i v_i = \vec{0}$ gilt:

$$\lambda_1 = \ldots = \lambda_n = 0.$$

Beweis:

„$\Rightarrow$“ : Anm.: Es gibt eine Linearkombination, in der ein $\lambda_k \neq 0$ existiert und $\sum_{i=1}^{n} \lambda_i v_i = \vec{0}$ gilt.

$$\begin{aligned} \Rightarrow -\lambda_k v_k &= \sum_{\substack{i=1 \\ i\neq k}}^{n} \lambda_i v_i \\ \Rightarrow v_k &= \sum_{\substack{i=1 \\ i\neq k}}^{n} \underbrace{\frac{\lambda_i}{-\lambda_k}}_{\in K} v_i \end{aligned}$$

$$\Rightarrow v_k \text{ist linear abhängig von} \{v_1, \ldots, v_n\} \setminus \{v_k\}.$$

„$\Leftarrow$“ : Seien $v_1, \ldots, v_n$ linear abhängig, d.h. es gibt ein v_k mit $v_k = \sum_{\substack{i=1 \\ i\neq k}}^{n} \lambda_i v_i$.

Setze $\lambda_k := -1 \ \Rightarrow 0 = \sum_{i=1}^{n} \lambda_i v_i$. □.

Macht lineare Unabhängigkeit auch bei unendlichen vielen Vektoren Sinn?

34.13 Definition (Lineare Unabhängigkeit)

Ein unendliches System B von Vektoren heißt linear unabhängig, wenn jede endliche Auswahl von Vektoren aus B linear unabhängig ist.

34.14 Beispiel:

Betrachte $\mathbb{R}[x]$, d.h. die Menge aller Polynome mit Koeffizenten aus $\mathbb{R}$. Nach 34.4 (g) bildet $\mathbb{R}[x]$ einen $\mathbb{R}$-Vektorraum.

Wir zeigen, dass das unendliche System $B = \{1, x, x^2, \ldots\}$ linear unabhängig in $\mathbb{R}[x]$ ist.

Anm.: Dies trifft nicht zu. Dann existiert eine endliche Teilmenge $\{x^{m_1}, x^{m_2}, \ldots, x^{m_n}\} \subset B$, die linear abhängig ist.

$\Rightarrow$ Es gibt eine Linearkombination

$$\sum_{i=1}^{n} \lambda_i v^{m_i} = 0 \qquad (*)$$

mit einem $\lambda_k \neq 0$. Auf der linken Seite von $(*)$ steht ein Polynom, das nur endlich viele Nullstellen hat, rechts steht das Nullpolynom mit unendlich vielen Nullstellen. ↯ □.

Der Begriff der linearen Unabhängigkeit erlaubt uns ein minimales Erzeugensystem zu finden.

34.15 Definition (Basis)

Sei V ein Vektorraum. Eine Teilmenge $B \subset V$ heißt Basis von V, falls gilt:

(a) $\operatorname{span}(B) = V$

(b) *B ist linear unabhängig.*

34.16 Beispiele

(a) $\left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\}$ ist eine Basis des $\mathbb{R}^2$, denn:

(i) Sei $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$. Wir suchen λ, μ mit

$$\lambda \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Das Gleichungsystem

$$\begin{aligned} 2\lambda + 3\mu &= x \\ 3\lambda + 4\mu &= y \end{aligned}$$

hat die Lösung $\lambda = -4x + 3y, \mu = 3x - 2y$. $(*)$

$\Rightarrow \mathbb{R}^2 = \operatorname{span}\left(\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right)$

(ii) Sei

$$\lambda \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Mit $(*)$ folgt:

$$\begin{aligned} \lambda &= -4 \cdot 0 + 3 \cdot 0 = 0 \\ \mu &= 3 \cdot 0 - 2 \cdot 0 = 0 \end{aligned}$$

$\Rightarrow \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ sind linear unabhängig.

(b) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}$

bildet eine Basis des $\mathbb{R}^n$, die so genannte Standardbasis.

(c) Es gibt auch Vektorräume mit unendlichen Basen. Beispielsweise ist $\{1, x, x^2, \dots\}$ eine unendliche Basis von $\mathbb{R}[x]$:

(i) $\{1, x, x^2, \dots\}$ ist linear unabhängig nach 34.14 auf der vorherigen Seite.

(ii) Jedes Polynom aus $\mathbb{R}[x]$ lässt sich als Linearkombination von Elementen aus $\{1, x, x^2, \dots\}$ darstellen.

Hat jeder Vektorraum eine Basis? Man kann zeigen:

34.17 Satz: (Existenz einer Basis)

Jeder Vektorraum $V \neq \{\vec{0}\}$ hat eine Basis.
Offenbar sind Basen nicht eindeutig: So sind z.B.
$\left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\}$ und $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ Basen des $\mathbb{R}^2$.

Insbesondere kann man Basisvektoren austauschen:

34.18 Satz: (Austauschsatz)

Sei V ein Vektorraum mit einer endlichen Basis $B := \{b_1, \dots, b_n\}$. Ferner sei $v \in V$ und $v \neq \vec{0}$. Dann existiert ein b_k, so dass $\{b_1, \dots b_{k-1}, v, b_{k+1}, \dots, b_n\}$ eine Basis von V ist.

Beweis:

Da B Basis $\Rightarrow \exists \lambda_1, \dots, \lambda_n$ mit

$$V = \sum_{i=1}^{n} \lambda_i b_i. \qquad (*)$$

Da $v \neq 0$ ist mindestens ein $\lambda_i \neq 0$. Sei o.B.d.A. $\lambda_1 \neq 0$.

Wir zeigen, dass dann $\{v, b_2, b_3, \dots, b_n\}$ eine Basis von V ist.

(i) $\operatorname{span}(v, b_2, b_3, \dots, b_n) = V$, denn:
Sei $u \in V$. Da B Basis ist, existiert $\mu_1, \dots, \mu_n$ mit

$$u = \sum_{k=1}^{n} \mu_k b_k. \qquad (**)$$

Aus $(*)$ und $\lambda_1 \neq 0$ folgt:

$$b_1 \quad = \quad \frac{1}{\lambda_1} v_1 - \sum_{i=2}^{n} \frac{\lambda_i}{\lambda_1} b_i.$$

Einsetzen in $(**)$ zeigt, dass u als Linearkombination von $v, b_2, b_3, \dots, b_n$ geschrieben werden kann.

(a) $v, b_2, b_3, \dots, b_n$ ist linear unabhängig, denn:
Sei $\mu_1 \cdot v + \mu_2 \cdot b_2 + \mu_3 \cdot b_3 + \dots + \mu_n \cdot b_n = 0$.
Mit $(*)$:

$$\begin{aligned} 0 \quad &= \quad \mu_1 \sum_{i=1}^{n} \lambda_i b_i + \sum_{i=2}^{n} \mu_i b_i \\ &= \quad \mu_1 \lambda_1 b_1 + \sum_{i=2}^{n} (\mu_1 \lambda_i + \mu_i)\, b_i \end{aligned}$$

Da $\{b_1, \dots, b_n\}$ Basis, sind alle Koeffizienten 0:

$$\begin{aligned} \mu_1 \lambda_1 = 0 \quad &\overset{\lambda_1 \neq 0}{\Longrightarrow} \quad \mu_1 = 0 \\ \mu_1 \lambda_i + \mu_i = 0 \quad &\overset{\mu_1 \neq 0}{\Longrightarrow} \quad \mu_i = 0 \qquad (i = 2, \dots, n). \end{aligned}$$

□.Hat ein Vektorraum eine eindeutige Zahl an Basisvektoren?

34.19 Satz: (Eindeutigkeit der Zahl der Basisvektoren)

Hat ein Vektorraum V eine endliche Basis von n Vektoren, so hat jede Basis von V ebenfalls n Vektoren.

Beweis:

Seien $B = \{b_1, \dots, b_n\}$ und $C = \{c_1, \dots, c_n\}$ Basen.

(a) Anm.: $n > m$. Tausche m der Vektoren aus B durch C aus. Da C eine Basis ist, sind die nicht ausgetauschten Vektoren aus B als Linearkombination von $c_1, \dots, c_m$ darstellbar.
Dies widerspricht der Basiseigenschaft.

(b) Anm.: $n < m$. Widerspruch wird analog konstruiert.

$\square$.

Satz 34.19 motiviert:

34.20 Definition (Dimension)

Sei $B = \{b_1, \dots, b_n\}$ eine Basis des Vektoraums V. Dann heißt die Zahl $n := \dim V$ die Dimension von V. Der Nullraum $\{\vec{0}\}$ habe die Dimension 0

Mann kann zeigen:

34.21 Satz: (Basiskriterium linear unabhängiger Mengen)

In einem Vektorraum der Dimension n ist jede linear unabhängige Menge mit n Vektoren eine Basis.

Bemerkung: Insbesondere gibt es in einem n-dimensionalen Vektorraum keine Basen mit $n-1$ oder $n+1$ Elementen. Dies würde Satz 34.19 widersprechen.

Kapitel 35

Lineare Abbildungen

35.1 Motivation

- Nachdem wir wichtige Eigenschaften von Vektorräumen kennen, macht es Sinn zu untersuchen, wie Abbildungen zwischen Vektorräumen aussehen können. Lineare Abbildungen sind die wichtigsten Abbildungen zwischen Vektorräumen.
- Der Basisbegriff liefert ein wichtiges Werkzeug zur Beschreibung linearer Abbildungen.

35.2 Definition (lineare Abbildung, Vektorraumhomomorphismus)

Seien U, V K-Vektorräume. Eine Abbildung $f : U \to V$ heißt lineare Abbildung (Vektorraumhomomorphismus), falls gilt:

(*a*) $f(u + v) = f(u) + f(v) \qquad \forall u, v \in U$

(*b*) $f(\lambda \cdot u) = \lambda f(u) \qquad \forall \lambda \in K, \forall u \in U$

U und V heißen isomorph, wenn es eine bijektive lineare Abbildung $f : U \to V$ gibt. Wir schreiben hierfür $U \simeq V$.

Bemerkung: Ähnlich wie bei Gruppenhomomorphismen (vgl 8.1 auf Seite 71 (b)) überführt ein Vektorraumhomomorphismus die Verknüpfungen in U (skalare Multiplikation) in Verknüpfungen in V.

Die Bedingungen (a) und (b) fasst man oft in eine einzige Bedingung zusammen:

$$f(\lambda u + \mu v) \quad = \quad \lambda f(u) + \mu f(v) \qquad \forall \lambda, \mu \in K, \forall u, v \in U.$$

D.h. Linearkombinationen in U werden in Linearkombinationen in V überführt.

35.3 Beispiele

(a) $f : \mathbb{R}^3 \to \mathbb{R}^2$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 2x_1 + x_3 \\ -x_2 \end{pmatrix} \text{ ist linear:}$$

$$\begin{aligned}
f\left(\lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \mu \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) &= f\left(\begin{pmatrix} \lambda x_1 + \mu y_1 \\ \lambda x_2 + \mu y_2 \\ \lambda x_3 + \mu y_3 \end{pmatrix}\right) \\
&= \begin{pmatrix} 2 \cdot (\lambda x_1 + \mu y_1) + \lambda x_3 + \mu y_3 \\ -\lambda x_2 - \mu y_2 \end{pmatrix} \\
&= \lambda \begin{pmatrix} 2x_1 + x_3 \\ -x_2 \end{pmatrix} + \mu \begin{pmatrix} 2y_1 + y_3 \\ -y_2 \end{pmatrix} \\
&= \lambda f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \mu f \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\
&\qquad \forall \lambda, \mu \in \mathbb{R}, \forall \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3.
\end{aligned}$$

(b) $f : \mathbb{R} \to \mathbb{R}$ $\quad x \mapsto x + 1$ ist keine lineare Abbildung, denn
$1 = f(0 + 0) \neq f(0) + f(0) = 1 + 1$.

(c) $f : \mathbb{R}^2 \to \mathbb{R}^2 :$ $\quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_1 x_2 \end{pmatrix}$ ist ebenfalls nichtlinear:

$$f\left(2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 4 \end{pmatrix},$$
$$2 \cdot f \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

35.4 Bild, Kern, Monomorphismus,...

In Analogie zu 8.16 auf Seite 77 und 8.17 auf Seite 77 gibt es auch für Vektorräume Monomorphismen, Epimorphismen, Isomorphismen, Endomorphismen und Automorphismen. Ferner ist für $f : U \to V$:

$$\begin{aligned}
\mathrm{Im}(f) &:= \{f(u) \mid u \in U\} && \text{das Bild von } f. \\
\mathrm{Ker}(f) &:= \{u \in U \mid f(u) = 0\} && \text{den Kern von } f.
\end{aligned}$$

Für lineare Abbildungen lassen sich einige wichtige Eigenschaften zeigen:

35.5 Satz: (Eigenschaften linearer Abbildungen)

a) Ist $f : U \to V$ ein Isomorphismus, so ist auch die Umkehrabbildung f^{-1} linear.

b) Die lineare Abbildung $f : U \to V$ ist genau dann injektiv, wenn $\mathrm{Ker}(f) = \{0\}$ ist.

c) Ist $f : U \to V$ linear, so ist $\mathrm{Ker}(f)$ ein Unterraum von U und $\mathrm{Im}(f)$ ein Unterraum von V.

d) $\dim \mathrm{Ker}(f) + \dim \mathrm{Im}(f) = \dim U$.

35.6 Beispiel:

$f : \mathbb{R}^3 \to \mathbb{R}^2 : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 - x_3 \\ 0 \end{pmatrix}$

$\mathrm{Ker}(f) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid x_1 = x_3 \right\}$ ist eine Ursprungsebene im $\mathbb{R}^3$ (und damit ein 2-dimensionaler Unterraum des Unterraum des $\mathbb{R}^3$).

$\mathrm{Im}(f) = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathbb{R}^2 \mid x \in \mathbb{R} \right\}$ ist eine Ursprungsgerade im $\mathbb{R}^2$ (und damit ein 1-dimensionaler Unterraum des $\mathbb{R}^2$).

$\dim \mathrm{Ker}(f) + \dim \mathrm{Im}(f) = 2 + 1 = 3 \dim = (\mathbb{R}^3)$.

Welche Rolle spielen Basen bei der Beschreibung linearer Abbilungen?
Hierzu betrachten wir zunächst Basisdarstellungen von Vektoren.

35.7 Satz: (Eindeutigkeit der Darstellung in einer festen Basis)

Sei $B = \{b_1, \ldots, b_n\}$ eine Basis des K-Vektorraums V. Dann gibt es zu jedem Vektor $v \in V$ eindeutig(!) bestimmte Elemente $x_1, \ldots, x_n \in K$ mit $v = \sum_{i=1}^{n} x_i b_i$.

Bemerkung: Diese x_i heißen Koordinaten von v bzgl. B. Wir schreiben

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_B .$$

Beweis:

Die Existenz der Darstellung ist klar, da $V = \mathrm{span}(B)$. Zu zeigen ist also nur die Eindeutigkeit. Sei $\sum_{i=1}^{n} x_i b_i$ und $v = \sum_{i=1}^{n} y_i b_i$.

$$\Rightarrow \vec{0} = v - v = \sum_{i=1}^{n} (x_i - y_i) b_i$$

$$\Rightarrow x_i = y_i \qquad (i = 1, \ldots n), \text{da } b_i \text{ linear unabhängig.}$$

$\square$.

Selbstverständlich liefern unterschiedliche Basen auch unterschiedliche Koordinatendarstellungen eines Vektors. Wie kann man diese Darstellung in einander umrechnen?

35.8 Umrechnung von Koordinatendarstellungen

Beispiel: Im $\mathbb{R}^2$ sei eine Basis $B = \{b_1, b_2\}$ gegeben. Bzgl. B habe ein Vektor v die Darstellung $v = \begin{pmatrix} x \\ y \end{pmatrix}_B$.

Wir betrachten die neue Basis $C = \{c_1, c_2\}$ mit $c_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}_B$, $c_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B$. Wie lautet die Darstellung von v in C?

Ansatz: $v = \begin{pmatrix} \lambda \\ \mu \end{pmatrix}_C$

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix}_B &= v = \begin{pmatrix} \lambda \\ \mu \end{pmatrix}_C = \lambda c_1 + \mu c_2 \\ &= \lambda \begin{pmatrix} 2 \\ 2 \end{pmatrix}_B + \mu \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B \\ &= \begin{pmatrix} 2\lambda + \mu \\ 2\lambda + 2\mu \end{pmatrix}_B . \end{aligned}$$

Die Bestimmungsgleichungen

$$\begin{aligned} 2\lambda + \mu &= x \\ 2\lambda + 2\mu &= y \end{aligned}$$

haben die Lösung

$$\begin{aligned} \lambda &= x - \frac{1}{2}y \\ \mu &= -x + y \end{aligned}$$

Beispielsweise ist

$$\begin{pmatrix} 5 \\ 2 \end{pmatrix}_B = \begin{pmatrix} 5 - \frac{1}{2}2 \\ -5 + 2 \end{pmatrix}_C = \begin{pmatrix} 4 \\ -3 \end{pmatrix}_B .$$

Basen ermöglichen es, eine lineare Abbildung durch wenige Daten zu beschreiben: Es genügt zu wissen, was mit den Basisvektoren passiert.

35.9 Satz: (Charakterisierung einer linearen Abbildung durch ihre Wirkung auf die Basis)

Seien U, V K-Vekterräume und $b_1, \ldots, b_n$ sei eine Basis von U. Ferner seien $v_1, \ldots, v_n \in V$. Dann gibt es genau eine lineare Abbildung $f : U \to V$ mit $f(b_i) = v_i \quad (i = 1, \ldots, n)$.

Beweis:

Sei $u \in U$ und $u = \sum_{i=1}^{n} x_i b_i$. Setze $f(u) := \sum_{i=1}^{n} x_i v_i$.

Man prüft leicht nach, dass f linear ist und $f(b_i) = v_i$ für $i = 1, \ldots, n$.

Zum Nachweis der Eindeutigkeit nehmen wir an, dass g eine weitere lineare Abbildung mit $g(b_i) = v_i \ \ \forall i$ ist.

$$\begin{aligned} \Rightarrow g(u) &= g\left(\sum_{i=1}^{n} x_i b_i\right) \\ &\overset{lin.}{=} \sum_{i=1}^{n} x_i g(b_i) = \sum_{i=1}^{n} x_i v_i \\ &= \sum_{i=1}^{n} x_i f(b_i) \\ &\overset{lin.}{=} f\left(\sum_{i=1}^{n} x_i b_i\right) = f(u) \end{aligned}$$

Somit stimmen f und g überall überein. $\square$.Als weiteres wichtiges Resultat, das auf Basen beruht, kann man zeigen:

35.10 Satz: (Isomorphie endlichdimensionaler Vektorräume)

Seien U, V endlichdimensionale K-Vektorräume. Dann sind U und V genau dann isomorph, wenn sie die selbe Dimension haben.

$$U \simeq V \quad \Leftrightarrow \quad \dim U = \dim V.$$

Bemerkung: Dieser Satz besagt z.B., dass es im Wesentlichen nur einen einzigen n-dimensionalen Vektorraum über $\mathbb{R}$ gibt: den $\mathbb{R}^n$!

Beweisskizze:

Man zeigt:

a) Ein Isomorphismus zwischen endlichdimensionalen Vektorräumen bildet Basen auf Basen ab.

b) Eine lineare Abbildung zweier gleichdimensinaler Vektorräume, die eine Basis auf eine Basis abbildet, ist ein Isomorphismus.

Kapitel 36

Matrixschreibweise für lineare Abbildungen

36.1 Motivation

- Wir haben gesehen, dass lineare Abbildungen sich durch ihre Wirkung auf die Basisvektoren ausdrücken lassen.
- Mit Hilfe von Matrizen können wir dies kompakt aufschreiben und die Hintereinanderausführung linearer Abbildungen elegant berechnen.

36.2 Struktur linearer Abbildungen zwischen endlichdimensionalen Vektorräumen

Sei K ein Körper (meist $\mathbb{R}$ oder $\mathbb{C}$). Dann ist $f : K^n \to K^m$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n \\ \vdots & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n \end{pmatrix} \qquad (*)$$

eine lineare Abbildung. Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K^n$ und $\lambda, \mu \in K$ gilt:

$$\begin{aligned} f(\lambda x + \mu y) &= \begin{pmatrix} a_{11}(\lambda x_1 + \mu y_1) & + & \dots & + & a_{1n}(\lambda x_n + \mu y_n) \\ \vdots & & & & \vdots \\ a_{m1}(\lambda x_1 + \mu y_1) & + & \dots & + & a_{mn}(\lambda x_n + \mu y_n) \end{pmatrix} \\ &= \lambda \begin{pmatrix} a_{11}x_1 & + & \dots & + & a_{1n}x_n \\ \vdots & & & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n \end{pmatrix} \\ &+ \mu \begin{pmatrix} a_{11}y_1 & + & \dots & + & a_{1n}y_n \\ \vdots & & & & \vdots \\ a_{m1}y_1 & + & \dots & + & a_{mn}y_n \end{pmatrix} \\ &= \lambda f(x) + \mu f(y). \end{aligned}$$

Umgekehrt kann man sogar zeigen, dass jede lineare Abbildung von K^n nach K^m von diese Struktur besitzt: Nach 35.9 auf Seite 230 ist eine lineare Abbildung von K^n nach K^m durch die Bilder der Basisvektoren $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in K^n$ eindeutig bestimmt.

Sind diese Bilder durch $\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \in K^m$ gegeben, so ist $(*)$ die entsprechende Abbildung.

Die Struktur $(*)$ motiviert folgende Definition:

36.3 Definition (($m \times n$)-Matrix, Zeilenindex, Spaltenindex)

Sei K ein Körper. Das rechteckige Schema

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

mit $a_{ij} \in K$ für $i = 1, \dots, m$ und $j = 1, \dots, n$ heißt $(m \times n)$-Matrix. Die Menge aller $(m \times n)$-Matrizen über K bezeichnen wir mit $K^{m \times n}$. Die Vektoren $\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$ heißen Spaltenvektoren von A, und $(a_{11}, \dots, a_{1n}), \dots (a_{m1}, \dots, a_{mn})$ bilden die Zeilenvektoren. Die Matrix A wird auch als $A = (a_{ij})$ geschrieben. Dabei bezeichnet i den Zeilenindex (bleibt in jeder Zeile konstant) und j ist der Spaltenindex.

36.4 Matrix - Vektor - Produkt

Ist $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ und $c = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \in K^m$, so schreiben wir statt

$$\begin{pmatrix} a_{11}x_1 & + & \dots & + & a_{1n}x_n \\ \vdots & & & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$

jetzt

$$\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}}_{\text{Matrix } A} \cdot \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\text{Vektor } x} = \underbrace{\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}}_{\text{Vektor } c}.$$

also kurz: $Ax = c$

Eine solche Schreibweise ist sowohl für lineare Abbildungen als auch für lineare Gleichungssysteme (spätere Vorlesung) nützlich. Sie definiert das Produkt zwischen einer Matrix $A = (a_{ij}) \in K^{m \times n}$ und einem Vektor

$x \in K^n$ als $A \cdot x = c$ mit $c \in K^m$ und

$$c_i = \sum_{j=1}^{n} a_{ij} x_j \qquad (i = 1, \ldots, m)$$

Jede lineare Abbildung $f : K^n \to K^m$ kann geschrieben werden als $f(x) = Ax$ mit $A \in K^{m \times n}$. Die Spalten von A sind die Bilder der Basisvektoren.

36.5 Beispiele für lineare Abbildung in Matrixschreibweise

(a) **Streckung:**

Die Matrix $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \in \mathbb{R}^{2\times 2}$ bildet einen Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ auf

$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda \cdot x + 0 \cdot y \\ 0 \cdot x + \lambda \cdot y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}.$

(b) **Drehung im $\mathbb{R}^2$:**

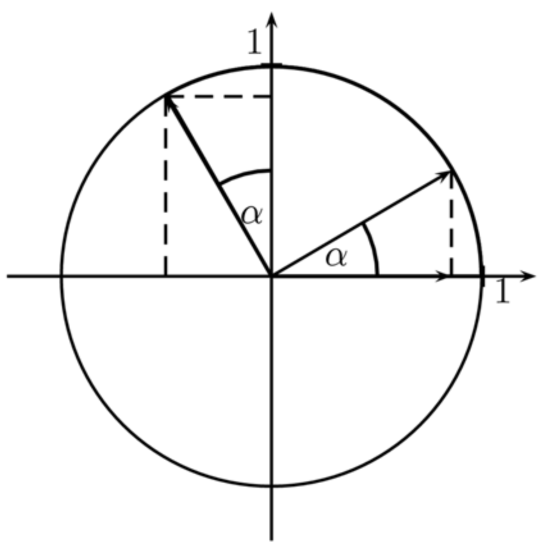
$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$$
$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$$

Die Matrix der Drehung um einen Winkel α lautet also

$$\Rightarrow A \quad = \quad \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Drehungen werden gegen den Uhrzeigersinn durchgeführt.

(c) **Drehung im $\mathbb{R}^3$:**



$$A \quad = \quad \begin{pmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{pmatrix}$$

beschreibt eine Drehung in der $x_1 - x_3$ - Ebene (entlang der x_2-Achse passiert nichts)

(d) **Translation:**

$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ sind **keine** linearen Abbildungen!

Welche Rechenoperationen lassen sich mit Matrizen durchführen?

36.6 Definition (Skalare Multiplikation, Matrixoperationen)

Sei K ein Körper.

(a) Für $A = (a_{ij}) \in K^{m \times n}$ und $\lambda \in K$ definiert man eine skalare Multiplikation komponentenweise:

$$\lambda A := (\lambda a_{ij}).$$

(b) Die Addition zweier Matrizen erfolgt ebenfalls komponentenweise:

$$A + B := (a_{ij} + b_{ij}) \qquad \forall\ A = a_{ij} \in K^{m \times n}, B = b_{ij} \in K^{m \times n}.$$

(c) Bei der Multiplikation von $A \in K^{l \times m}$ mit $B \in K^{m \times n}$ geht man nicht komponentenweise vor:

$$\begin{aligned} A \cdot B &= C \in K^{l \times n} \text{ mit} \\ c_{ij} &:= \sum_{k=1}^{m} a_{ik} b_{kj} \qquad (\text{„Zeile mal Spalte“}) \end{aligned}$$

36.7 Beispiele

(a) $3 \cdot \begin{pmatrix} 5 & 1 \\ 2 & -4 \end{pmatrix} = \begin{pmatrix} 15 & 3 \\ 6 & -12 \end{pmatrix}$

(b) $\begin{pmatrix} 2 & 3 & 1 \\ 4 & -1 & 7 \end{pmatrix} + \begin{pmatrix} 1 & 5 & -8 \\ 0 & 2 & 9 \end{pmatrix} = \begin{pmatrix} 3 & 8 & -7 \\ 4 & 1 & 16 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & 3 & 1 \\ 4 & -1 & 7 \end{pmatrix} \cdot \begin{pmatrix} 6 & 4 \\ 1 & 0 \\ 8 & 9 \end{pmatrix} = \begin{pmatrix} 2\cdot 6 + 3 \cdot 1 + 1 \cdot 8 & 2 \cdot 4 + 3 \cdot 0 + 1 \cdot 9 \\ 4 \cdot 6 - 1 \cdot 1 + 7 \cdot 8 & 4 \cdot 4 - 1 \cdot 0 + 7 \cdot 9 \end{pmatrix} = \begin{pmatrix} 23 & 17 \\ 79 & 79 \end{pmatrix}$

Welche Rechenregeln folgen aus Def 36.6?

36.8 Satz: (Eigenschaften der Matrixoperationen)

Sei K ein Körper.

a) $(K^{n \times n}, +, \cdot)$ ist ein Ring mit Eins, d.h. $(K^{n \times n}, +)$ ist eine kommutative Gruppe:

- *Assoziativgesetz: $(A + B) + C = A + (B + C)$.*
- *neutrales Element: $0 = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in K^{n \times n}$ (Nullmatrix).*
- *inverses Element zu A ist $-A = (-1)A$.*
- *Kommutativgesetz: $A + B = B + A$.*

$(K^{n\times n},\cdot)$ ist ein Monoid:

- *Assoziativgesetz:* $(A\cdot B)\cdot C = A\cdot(B\cdot C)$.
- *neutrales Element:*

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in K^{n\times n} \qquad \textit{Einheitsmatrix.}$$

Es gelten die Distributivgesetze:
$$\begin{aligned} A\cdot(B+C) &= A\cdot B + A\cdot C, \\ (A+B)\cdot C &= A\cdot C + B\cdot C. \end{aligned}$$

b) $K^{m\times n}$ ist ein K-Vektorraum:
$(K^{n\times n},+)$ ist eine kommutative Gruppe:

$$\begin{aligned} \lambda(\mu A) &= (\lambda\mu)A \\ 1\cdot A &= A \\ \lambda(A+B) &= \lambda A + \lambda B \\ (\lambda+\mu)A &= \lambda A + \mu A \end{aligned}$$

36.9 Bemerkung:

(a) Im Allgemeinen liegt <u>keine</u> Kommutativität bzgl. der Multiplikation vor

$$\begin{pmatrix} 2 & 1 \\ 7 & -4 \end{pmatrix}\cdot\begin{pmatrix} 5 & 3 \\ 4 & 6 \end{pmatrix} = \begin{pmatrix} 2\cdot5+1\cdot4 & 2\cdot3+1\cdot6 \\ 7\cdot5-4\cdot4 & 7\cdot3-4\cdot6 \end{pmatrix} = \begin{pmatrix} 14 & 12 \\ 19 & -3 \end{pmatrix}$$
$$\begin{pmatrix} 5 & 3 \\ 4 & 6 \end{pmatrix}\cdot\begin{pmatrix} 2 & 1 \\ 7 & -4 \end{pmatrix} = \begin{pmatrix} 5\cdot2+3\cdot7 & 5\cdot1-3\cdot4 \\ 4\cdot2+6\cdot7 & 4\cdot1-6\cdot4 \end{pmatrix} = \begin{pmatrix} 31 & -7 \\ 50 & -20 \end{pmatrix}$$

.

(b) Ebenso ist eine $(n\times n)$- Matrix im Allgemeinen <u>nicht</u> invertierbar bzgl. der Multiplikation.

(c) $(K^{n\times n},+,\cdot)$ ist <u>nicht</u> nullteilerfrei:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

.

(d) Der Vektorraum $K^{n\times m}$ ist isomorph zum Vektorraum $K^{n\cdot m}$.

(e) Spaltenvektoren aus K^m können als $(m\times 1)$-Matrizen aufgefasst werden. Vektoraddition und skalare Multiplikation sind Spezialfälle der entsprechenden Matrixoperationen.

(f) Die Hintereinanderausführung linearer Abbildungen entspricht der Multiplikation ihrer Matrizen.

$$K^l \xrightarrow[A\in K^{m\times l}]{g} K^m \xrightarrow[B\in K^{n\times m}]{f} K^n$$

$f\circ g : K^l \to K^n$ wird durch $B\cdot A \in K^{n\times l}$ repräsentiert.

36.10 Inverse Matrix

Im Allgemeinen hat $A \in K^{n\times n}$ kein multiplikatives Inverses A^{-1}. In vielen Fällen existiert jedoch eine inverse Matrix.

Beispiel: $A = \begin{pmatrix} 5 & 6 \\ 2 & 4 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{3}{4} \\ -\frac{1}{4} & \frac{5}{8} \end{pmatrix}$, denn

$$\begin{aligned} A\cdot A^{-1} &= \begin{pmatrix} 5 & 6 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & -\frac{3}{4} \\ -\frac{1}{4} & \frac{5}{8} \end{pmatrix} \\ &= \begin{pmatrix} 5 & 6 \\ 2 & 4 \end{pmatrix} \cdot \frac{1}{8} \cdot \begin{pmatrix} 4 & -6 \\ -2 & 5 \end{pmatrix} \\ &= \frac{1}{8} \begin{pmatrix} 5\cdot 4 - 2\cdot 6 & -5\cdot 6 + 6\cdot 5 \\ 2\cdot 4 - 4\cdot 2 & -2\cdot 6 + 4\cdot 5 \end{pmatrix} \\ &= \frac{1}{8} \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Ein Verfahren zur Berechnung der inversen Matrix werden wir in einer späteren Vorlesung kennenlernen.

36.11 Definition (Invertierbarkeit einer Matrix)

Eine Matrix $A \in K^{n\times n}$ heißt <u>invertierbar</u> (<u>umkehrbar</u>, <u>regulär</u>), falls eine Matrix $A^{-1} \in K^{n\times n}$ existiert mit $A \cdot A^{-1} = I$. Die Menge der invertierbaren Matrizen aus $K^{n\times n}$ wird mit $\underline{\mathrm{GL}(n,K)}$ bezeichnet.

36.12 Satz: (Gruppeneigenschaft von $\mathrm{GL}(n, K)$)

$\mathrm{GL}(n, K)$ *bildet mit der Matrizenmultiplikation eine multiplikative (nichtkommutative) Gruppe.*

36.13 Bemerkungen:

(a) $\mathrm{GL}(n, K)$ steht für <u>general linear group</u>.

(b) Sind $A, B \in \mathrm{GL}(n, K)$ so ist $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$, denn:
$A \cdot \underbrace{B \cdot B^{-1}}_{I} \cdot A^{-1} = AIA^{-1} = AA^{-1} = I.$

(c) Ist $n \neq m$, so nennt man $A \in K^{m\times n}$ eine <u>nichtquadratische</u> Matrix. Nichtquadratische Matrizen sind niemals invertierbar. Allerdings kann man eine so genannte Pseudoinverse angeben. ($\longrightarrow$ spätere Vorlesung)

(d) Die Invertierbarkeit einer Matrix entspricht der Bijektivität ihrer linearen Abbildung. Nach dem Beweis von Satz 35.10 auf Seite 231 ist eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen genau dann bijektiv, wenn Basen auf Basen abgebilet werden.

Folgerung: $A \in K^{n\times n}$ ist genau dann invertierbar, wenn die Spalten von A eine Basis des K^n bilden (d.h. wenn sie linear unabhängig sind).

Kapitel 37

Der Rang einer Matrix

37.1 Motivation

- Interpretiert man eine Matrix als lineare Abbildung, so haben die Spaltenvektoren eine besondere Bedeutung: Sie sind die Bilder der Basisvektoren.
- Wir wollen die lineare Abhängigkeit/Unabhängigkeit dieser Spaltenvektoren genauer untersuchen. Dies liefert u. A. ein wichtiges Kriterium für die Invertierbarkeit von Matrizen.

37.2 Definition (Spaltenrang)

Unter dem (Spalten-)Rang einer Matrix $A \in K^{m \times n}$ versteht man die maximale Anzahl linear unabhängiger Spaltenvektoren (Schreibweise: $\operatorname{rang} A$).

Welche Aussagen gelten für den Rang?

37.3 Satz: (Aussagen über den Rang einer Matrix)

Ist $f : K^n \to K^m$ eine lineare Abbildung mit zugehöriger Matrix $A \in K^{m \times n}$ und Spaltenvektoren $a_{*1}, a_{*2}, \ldots, a_{*n}$, so gilt:

(a) $\operatorname{Im}(f) = \operatorname{span}(a_{*1}, a_{*2}, \ldots, a_{*n})$.

(b) $\dim \operatorname{Im}(f) = \operatorname{rang} A$.

(c) $\dim \operatorname{Ker}(f) = n - \operatorname{rang} A$.

Beweis:

(a) „$\supset$“: Klar, da $a_{*1}, \ldots, a_{*n}$ die Bilder der Basisvektoren sind.

„$\subset$ “: $f(x) = Ax = \begin{pmatrix} a_{11}x_1 & + & \ldots & + & a_{1n}x_n \\ \vdots & & & & \vdots \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n \end{pmatrix} = x_1 a_{*1} + \ldots + x_n a_{*n}$.

(b) Folgt direkt aus (a).

(c) Folgt aus Satz 35.5 auf Seite 228.

□.

37.4 Beispiele

(a) Nach 36.13 (d) ist $A \in K^{n \times m}$ genau dann invertierbar, wenn $\operatorname{rang} A = n$ ist.

(b) $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ hat Rang 3:

$$\begin{aligned} a_{*2}, a_{*3} &\notin \operatorname{span}(a_{*1}) \\ a_{*3} &\notin \operatorname{span}(a_{*1}, a_{*2}) \end{aligned}$$

(c) $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix}$ hat Rang 2:

a_{*1}, a_{*2} sind linear unabhängig und
$a_{*3} = a_{*1} + 2 \cdot a_{*2}$.

37.5 Definition (Transponierte Matrix)

Vertauscht man bei einer Matrix $A \in K^{m \times n}$ die Rolle von Zeilen und Spalten, so entsteht die transponierte Matrix $A^\top \in K^{n \times m}$.

Der Rang von $A^\top$ beschreibt die maximale Zahl der linear unabhängigen Zeilen von A. Man nennt ihn daher auch den Zeilenrang von A.

37.6 Beispiel:

$$A = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 \\ 2 & 0 & 2 & 4 \end{pmatrix} \quad \Rightarrow \quad A^\top = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 2 & 2 \\ 2 & 1 & 4 \end{pmatrix} =: B$$

$$\begin{aligned} \operatorname{rang} A = 2 \quad &: \quad a_{*1}, a_{*2} \text{ linear unabhängig.} \\ & \quad\ \ a_{*3} = a_{*1} + 2 \cdot a_{*2} \\ & \quad\ \ a_{*4} = 2 \cdot a_{*1} + a_{*2} \\ \operatorname{rang} B = 2 \quad &: \quad b_{*1}, b_{*2} \text{ linear unabhängig.} \\ & \quad\ \ b_{*3} = 2 \cdot b_{*1}. \end{aligned}$$

In diesem Beispiel stimmen also Spaltenrang und Zeilenrang überein. Ist dies Zufall?

Man kann zeigen:

37.7 Satz: (Gleichheit von Spaltenrang und Zeilenrang)

Für $A \in K^{m \times n}$ sind Spalten- und Zeilenrang identisch.

Kapitel 38

Gauss'scher Algorithmus und lineare Gleichungssysteme

38.1 Motivation

- In Abschnitt 36.4 (Matrix-Vektor-Produkte) haben wir gesehen, dass Matrizen zur kompakten Notation linearer Gleichungssysteme benutzt werden können: Die Gleichheit $Ax = b$ (A Matrix, x, b Vektoren) verkörpert die Gleichungen
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$
Lineare Gleichungssysteme treten in einer Vielzahl von Anwendungen auf und müssen gelöst werden. Vorausgesetzt, A ist eine invertierbare Matrix (vgl. 36.10 - 36.13), dann erhalten wir durch Multiplikation der Gleichung $Ax = b$ mit A^{-1} auf der linken Seite den Vektor x, d.h. inverse Matrizen spielen eine Rolle im Zusammenhang mit der Lösung linearer Gleichungssysteme.
- Aus 37.4 kennen wir den Zusammenhang zwischen Rang und Invertierbarkeit
$$A(n \times n\text{-Matrix) ist invertierbar} \Leftrightarrow \operatorname{rang} A = n.$$
Aus diesen Gründen ist es daher von Interesse, ein Verfahren zur Hand zu haben, das
  - den Rang einer Matrix ermittelt
  - die Inverse (sofern vorhanden) berechnet
  - Lineare Gleichungssysteme löst.

  Dieses alles leistet der Gauss -Algorithmus.

38.2 Idee

In Beispiel 37.4 (b) $\left(A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\right)$ wurde eine obere Dreiecksmatrix betrachtet, die auf der Diagonalen nur von 0 verschiedene Einträge hatte. Allgemein hatte eine $(n \times n)$-Matrix mit dieser Eigenschaft stets den Rang n..

Die Idee des Gauß-Algorithmus besteht darin, eine beliebige Matrix in eine Dreiecksmatrix umzuwandeln, bzw. eine ähnliche Form, aus der der Rang leicht abzulesen ist, und zwar auf eine Weise, die sicher stellt, dass sich der Rangder Matrix dabei nicht ändert.

38.3 Definition (elementare Zeilenumformungen)

Die folgenden Operationen auf Matrizen heißen elementare Zeilenumformungen

(a) *Vertauschen zweier Zeilen,*

(b) *Addition des λ- fachen der Zeile a_{j*} zur Zeile a_{i*} $(\lambda \in \mathbb{R})$*

(c) *Multiplikation einer Zeile mit einer reellen Zahl $\lambda \neq 0$.*

38.4 Satz:

Bei elementaren Zeilenumformungen bleibt der Rang einer Matrix erhalten.

Beweis:

(a) offensichtlich, da $\operatorname{span}(a_{1*}, a_{2*}, \ldots, a_{m*})$ unverändert bleibt.

(b) Wir zeigen $\operatorname{span}(a_{1*}, a_{2*}, \ldots, a_{m*}) = \operatorname{span}(a_{1*}, \ldots, a_{i*} + \lambda a_{j*}, \ldots, a_{m*})$
„$\subset$ “: Es sei $v \in \operatorname{span}(a_{1*}, \ldots, a_{i*}, \ldots, a_{m*})$, also

$$\begin{aligned} v &= \lambda_1 a_{1*} + \ldots + \lambda_i a_{i*} + \ldots + \lambda_j a_{j*} + \ldots + \lambda_m a_{m*} \\ &= \lambda_1 a_{1*} + \ldots + \lambda_i (a_{i*} + \lambda a_{j*}) + \ldots + (\lambda_j - \lambda \lambda_i) a_{j*} + \ldots + \lambda_m a_{m*} \\ &\in \operatorname{span}(a_{1*}, \ldots, a_{i*} + \lambda a_{j*}, \ldots, a_{m*}) \end{aligned}$$

„$\supset$ “: wird analog gezeigt

(c) analog zu (b).

□.

38.5 Gauss'scher Algorithmus

Gegeben sei eine $(m \times n)$-Matrix $A = (a_{ij})$.

- Betrachte a_{11}
  - Ist $a_{11} \neq 0$, so subtrahiere von jeder Zeile $a_{i*}, i \geq 2$, das $\frac{a_{i1}}{a_{11}}$-fache der ersten Zeile. Danach ist a_{11} das einzige von Null verschiedene Element der ersten Spalte.
  - Ist $a_{11} = 0$, aber das erste Element a_{i1} einer anderen Zeilen $\neq 0$, so vertausche diese beiden Zeilen (a_{1*} und a_{i*}). Mit der neuen Matrix verfahre wie oben.
  - Sind alle $a_{i1}, i = 1, \ldots, m$ gleich 0, so beachte die erste Spalte nicht weiter und verfahre wie oben mit a_{12} (statt a_{11}), sofern auch die 2. Spalte nur Nullen enthält, gehe zur 3. Spalte usw.
    Im Ergebnis dieses 1. Schrittes erhalten wir eine Matrix

$$A' = \left(\begin{array}{c|ccc} a'_{11} & a'_{12} & \ldots & a'_{1n} \\ \hline & a'_{22} & \ldots & a'_{2n} \\ 0 & \vdots & & \vdots \\ & a'_{m2} & \ldots & a'_{mn} \end{array}\right)$$

(evtl. mit zusätzlichen Nullspalten links)

- Nun wird die erste Zeile und Spalte nicht mehr weiter betrachtet und auf die verbleibende Teilmatrix

$$\begin{pmatrix} a'_{22} & \dots & a'_{2n} \\ \vdots & \ddots & \vdots \\ a'_{m2} & \dots & a'_{mn} \end{pmatrix}$$

dasselbe Verfahren angewendet. Damit wird A' umgewandelt in

$$A'' \quad = \quad \left(\begin{array}{c|cccc} a'_{11} & a'_{12} & a'_{13} & \dots & a'_{1n} \\ \hline & a''_{22} & a''_{23} & \dots & a''_{2n} \\ \cline{3-5} 0 & & \multicolumn{1}{|c}{a''_{33}} & \dots & a''_{3n} \\ & 0 & \multicolumn{1}{|c}{\vdots} & & \vdots \\ & & \multicolumn{1}{|c}{a''_{m3}} & \dots & a''_{mn} \end{array}\right)$$

- Danach betrachtet man wiederum die um eine Zeile und Spalte verkleinerte Teilmatrix usw. bis die gesamte Matrix auf eine Form A^* wie die folgende gebracht ist.

$$A^* = \begin{pmatrix} a & * & * & & & & \dots & * \\ 0 & b & * & & & & \dots & * \\ 0 & 0 & 0 & c & * & & \dots & * \\ 0 & 0 & 0 & 0 & d & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & & & & \\ 0 & 0 & 0 & 0 & & 0 & \dots & 0 \end{pmatrix} \qquad (* \text{ beliebiger Wert})$$

In dieser Matrix heißt der erste von 0 verschiedene Eintrag einer Zeile (a, b, c, d) Leitkoeffizent. Im gesamten Bereich unterhalb und links eines Leitkoeffizienten enthält die Matrix nur Nullen.

$$\begin{pmatrix} a & * & * & & & \dots & * \\ 0 & b & * & & & \dots & * \\ 0 & 0 & 0 & c & & \dots & * \\ 0 & 0 & 0 & 0 & d & \dots & * \\ \vdots & & & \vdots & & \dots & * \\ 0 & 0 & 0 & 0 & & \dots & * \end{pmatrix}$$

Eine solche Matrix heißt Matrix in Zeilen-Stufen-Form.

Formulierung als rekursive Funktion

Gauss(i, j):
falls $i = m$ oder $j > n$
    Ende
falls $a_{ij} = 0$:
    suche $a_{kj} \neq 0, k > i$; wenn dies nicht existiert:
        Gauss$(i, j+1)$
        Ende
    vertausche Zeile i mit Zeile k
für alle $k > i$:
    subtrahiere $\frac{a_{kj}}{a_{ij}} \cdot$ Zeile i von Zeile k (*)
Gauss$(i+1, j+1)$
Ende.

Das Matrixelement a_{ij} das in der Zeile (*) als Nenner auftritt, heißt Pivotelement.

38.6 Beispiel:

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

Tausche 1. und 2. Zeile

$\Downarrow$

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

Addiere zur 3. Zeile das (-3)-fache der 1. Zeile und zur 4. Zeile das (-2)-fache der 1. Zeile.

$\Downarrow$

$$A' = \begin{pmatrix} 1 & 3 & 2 & 4 & 2 \\ 0 & 4 & 3 & 1 & 5 \\ 0 & -8 & -4 & -11 & -6 \\ 0 & -4 & -1 & -10 & -1 \end{pmatrix}$$

Addiere zur 3. Zeile das 2-fache der 2. Zeile und zur 4. Zeile das 1-fache der 2. Zeile.

$\Downarrow$

$$A'' = \begin{pmatrix} 1 & 3 & 2 & 4 & 2 \\ 0 & 4 & 3 & 1 & 5 \\ 0 & 0 & 2 & -9 & 4 \\ 0 & 0 & 2 & -9 & 4 \end{pmatrix}$$

Addiere zur 4. Zeile das (-1)-fache der 3. Zeile.

$\Downarrow$

$$A''' = \begin{pmatrix} 1 & 3 & 2 & 4 & 2 \\ 0 & 4 & 3 & 1 & 5 \\ 0 & 0 & 2 & -9 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A^*$$

38.7 Satz:

Der Rang einer Matrix in Zeilen-Stufen-Form ist gleich der Anzahl ihrer Leitkoeffizienten.

Beweis:

Die Anzahl der Leitkoeffizienten sei L.

Jede Zeile, die ein Leitkoeffizienten enthält, ist linear unabhängig von dem System der darunter stehenden Zeilen.
Nimmt man also von unten nach oben die linear unabhängigen Zeilen zur Basis hinzu, so folgt, dass $dim\ \mathrm{span}(a_{1*}, \ldots, a_{m*}) \geq L$.
Anderseits ist auch $dim\ \mathrm{span}(a_{1*}, \ldots, a_{m*}) \leq L$,

da es nur L Zeilenvektoren $\neq 0$ gibt. $\square$.

38.8 Satz:

Jede elementare Zeilenumformung für $m \times n$-Matrizen kann als Multiplikation von links mit einer geeigneten invertierbaren $m \times m$-Matrix D ausgedrückt werden:

(a) Vertauschung Zeile i und j.

$$
\begin{array}{cc}
\begin{pmatrix}
1 & & & & & & & & & \\
 & \ddots & & & \cdots & & & 0 & & \\
 & & 1 & & & & & & & \\
 & & & 0 & \cdots & & 1 & & & \\
 & & & & 1 & & & & & \\
 & \vdots & & \vdots & \ddots & & \vdots & \vdots & & \\
 & & & & & 1 & & & & \\
 & & & 1 & \cdots & & 0 & & & \\
 & & & & & & & 1 & & \\
 & 0 & & & \cdots & & & \ddots & & \\
 & & & & & & & & & 1
\end{pmatrix}
&
\begin{array}{c}
\\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\
\end{array}
\\
\begin{array}{cccccccccc}
 & & & \uparrow & & & \uparrow & & & \\
 & & & i & & & j & & &
\end{array}
&
\end{array}
$$

(b) Addition von $\lambda \cdot$ Zeile j zu Zeile i.

$$
\begin{array}{cc}
\begin{pmatrix}
1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \ddots & & & & & \vdots \\
\vdots & \ddots & 1 & \cdots & \boxed{\lambda} & & \\
 & & & \ddots & \vdots & & \vdots \\
\vdots & & & & 1 & & \\
\vdots & & & & & \ddots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{pmatrix}
&
\begin{array}{c}
\\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\
\end{array}
\end{array}
$$

(c) Multiplikation von Zeile i mit $\lambda \neq 0$:

$$
\begin{array}{cc}
\begin{pmatrix}
1 & 0 & \cdots & & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & & & & \vdots \\
\vdots & \ddots & 1 & & \ddots & & \vdots \\
 & & & \lambda & & & \\
\vdots & & \ddots & & 1 & \ddots & \vdots \\
\vdots & & & & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & & \cdots & 0 & 1
\end{pmatrix}
&
\begin{array}{c}
\\ \\ \\ \leftarrow i \\ \\ \\ \\
\end{array}
\end{array}
$$

(ohne Beweis)

38.9 Umformung einer invertierbaren Matrix zur Einheitsmatrix

- Es sei A eine invertierbare $n \times n$-Matrix. Wegen $\operatorname{rang} A = n$ hat sie die Zeilen-Stufen-Form

$$\begin{pmatrix} a_{11} & * & \dots & \dots & * \\ 0 & a_{22} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \dots & \dots & 0 & a_{nn} \end{pmatrix}$$

  mit $a_{ii} \neq 0, i = 1, \dots, n$.

- Multiplikation jeder Zeile (a_{i*}) mit $\dfrac{1}{a_{ii}}$ ergibt eine obere Dreiecksmatrix, deren Diagonalelemente sämtlich gleich 1 sind.

- Mittels weiterer elementarer Zeilenumformungen wird die Matrix in die Einheitsmatrix umgewandelt:
  - Subtrahiere für alle $i < n$ das a_{in}-fache der Zeile n von Zeile i. Danach ist $a_{nn} = 1$ das einzige Elemente $\neq 0$ in der letzten Spalte
  - Subtrahiere für alle $i < n - 1$ das $a_{i,n-1}$-fache der Zeile $n - 1$ von Zeile i, usw.

38.10 Satz: (Berechnung der inversen Matrix)

Wird eine invertierbare $n \times n$ -Matriz A durch elemetare Zeilenumformungen in die Einheitsmatrix E umgeformt, so erhält man die Inverse A^{-1}, indem man dieselben Umformungen auf die Einheitsmatrix anwendet.

Beweis:

Es seien $D_1, \dots D_k$ die Matrizen, die gemäß 38.8 die elementaren Zeilenumformungen darstellen, durch die A in E übergeht. Dann ist

$$E = D_k D_{k-1} \dots D_1 A = (D_k D_{k-1} \dots D_1)\, A$$

Damit ist

$$A^{-1} = D_k D_{k-1} \dots D_1 = D_k D_{k-1} \dots D_1 E.$$

$\square$.

38.11 Beispiel

$$\underset{A}{\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}} \qquad \underset{E}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}$$

Addiere zur 2 Zeile das (-2)-fache der 1. Zeile und
zur 3. Zeile das (-4)-fache der 1. Zeile

$$\Downarrow$$

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

Addiere zur 3. Zeile das 1-fache der 2. Zeile
zur 4. Zeile das 1-fache der 2. Zeile

$$\Downarrow$$

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

Multpliziere 2. und 3. Zeile mit -1

$$\Downarrow$$

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

Addiere zur 1. Zeile das (-2)-fache der 3. Zeile und
zur 1. Zeile das (-1)-fache der 3. Zeile

$$\Downarrow$$

$$\underset{E}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \qquad \underset{A^{-1}}{\begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}}$$

Probe:

$$\begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wie kann man den Gauß-Algorithmus zum Lösen linearer Gleichungssysteme verwenden?

38.12 Lineare Gleichungssysteme

Ein lineares Gleichungssystem mit m Gleichungen und n Unbekannten hat die Form $Ax = b$ mit $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$.

Falls $b \neq 0$, spricht man von einem <u>inhomogenen</u> linearen Gleichungsssystem. $Ax = 0$ (d.h. $b = 0$) heißt zugehörgiges <u>homogenes</u> Gleichungssystem.

Die Matrix $(A, b) \in \mathbb{R}^{m \times (n+1)}$, d.h. A mit rechts angefügter Spalte b, heißt <u>erweiterte Matrix</u> des Systems.

Interpretiert man A als lineare Abbildung, so sind Lösungen von $Ax = b$ gerade die Vektoren, die durch A und b abgebildet werden.

38.13 Satz: (Lösungsverhalten linearer Gleichungssysteme)

(*a*) *Die Lösungsmenge des homogenen Systems* $Ax = 0$ *ist* $\mathrm{Ker}A$ *und daher ein Unterraum von* $\mathbb{R}^n$.

(*b*) *Es sind äquivalent:*

- (*i*) $Ax = b$ *hat mindestens eine Lösung.*
- (*ii*) $b \in \mathrm{Im}A$.
- (*iii*) $\mathrm{rang}\ A = \mathrm{rang}\ (A, b)$.

(*c*) *Ist* w *eine* *Lösung von* $Ax = b$, *so ist die vollständige Lösungsmenge gleich*

$$w + \mathrm{Ker}A := \{w + x \mid x \in \mathrm{Ker}A\}$$

Beweis:

(a) Klar nach Def. von $\mathrm{Ker}A$.

(b) (i) $\Rightarrow$ (ii): Existiert eine Lösung w, so gilt $Aw = b$, d.h. $b \in \mathrm{Im}A$.
(ii) $\Rightarrow$ (i): Ist $b \in \mathrm{Im}A$, so ist jedes Urbild w von b Lösung von $Ax = b$.
(ii) $\Rightarrow$ (iii): Ist $b \in \mathrm{Im}A$, so ist b Linearkombination der Spalten von A. Damit ist $\mathrm{rang}\ (A, b) = \mathrm{rang}\ A$
(iii) $\Rightarrow$ (ii): Ist $\mathrm{rang}\ (A, b) = \mathrm{rang}\ A$, so ist b Linearkombination der Spalten von A. Somit ist $b \in \mathrm{Im}A$.

(c) „$\subset$“: Sei $y \in \mathrm{Ker}A$ und sei w Lösung von $Ax = b$. Dann gilt:

$$A(w + y) \quad = \quad Aw + Ay = b + 0 = b$$

d.h. $w + y$ ist Lösung von $Ax = b$.
Also liegt $w + \mathrm{Ker}A$ in der Lösungsmenge von $Ax = b$.
„$\supset$“: Sei v eine weitere Lösung von $Ax = b$. Dann gilt:

$$A(v - w) \quad = \quad Av - Aw = b - b = 0$$

d.h. $v - w \in \mathrm{Ker}A$.
$\Rightarrow v \in w + \mathrm{Ker}A$
Somit liegt die Lösungsmenge von $Ax = b$ in $w + \mathrm{Ker}A$.

38.14 Bemerkungen:

(a) Ist $\mathrm{rang}\ (A, b) > \mathrm{rang}\ A$, so hat $Ax = b$ keine Lösung. (folgt aus 38.13).

(b) Jedes homogene System hat mindestens eine Lösung: 0.

(c) Zur Lösung des inhomogenen Systems benötigt man

- die vollständige Lösungsmenge des homogenen Systems,
- eine spezielle Lösung des inhomogenen Systems.

(d) Aus (b) und (c) folgt: Das inhomogene System hat genau dann eine eindeutige Lösung, falls $\mathrm{Ker}A = 0$ ist.

Um mit Hilfe des Gauß- Algorithmus' das Lösungsverhalten von $Ax = b$ zu studieren , bezeichnen wir mit

$$\text{Lös}\ (A, b)$$

die Lösungsmenge von $Ax = b$. Wir benötigen

38.15 Lemma: *(Invarianz der Lösungsmenge unter Matrixmultiplikation)*

Ist $B \in \mathrm{GL}(m,\mathbb{R})$ und $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, so gilt:

$$\text{Lös}\,(A,b) = \text{Lös}\,(BA,Bb)$$

Beweis:

„$\Rightarrow$ ": Ist $x \in \text{Lös}\,(A,b)$, gilt $Ax = b$.
$\Rightarrow BAx = Bb \Rightarrow x \in \text{Lös}\,(BA,Bb)$.

„$\Leftarrow$ ": Sei $x \in \text{Lös}\,(BA,Bb) \Rightarrow BAx = Bb$.
Da $b \in \mathrm{GL}(m,\mathbb{R})$ existiert $B^{-1} \in \mathrm{GL}(m,\mathbb{R})$ nach Satz 36.12
$\Rightarrow B^{-1}BAx = B^{-1}Bb$
$\Rightarrow Ax = b \Rightarrow x \in \text{Lös}\,(A,b)$. □.Mit diesem Lemma gilt:

38.16 Satz: (Invarianz der Lösungsmenge unter elementaren Zeilenumformungen)

Ist die Matrix A' aus A durch elementare Zeilenumformungen entstanden, und der Vektor b' aus b durch die gleichen Zeilenumformungen, so gilt:

$$\text{Lös}\,(A',b') \quad = \quad \text{Lös}\,(A,b).$$

Beweisidee:

Elementare Zeilenumformungen entsprechen nach 38.8 der Multiplikation mit invertierbaren Matrizen $D_k, \ldots, D_1$. □.Nun zum eigentlichen Thema

38.17 Gauß-Algorithmus zum Lösen linearer Gleichungssysteme

Zur Lösung von $Ax = b$ geht man in vier Schritten vor:

Schritt 1: Bringe die Matrix (A,b) in Gauß-Jordan-Form (A',b').

Die Gauß-Jordan-Form ist eine spezielle Zeilen-Stufen-Form, bei der alle Leitkoeffizienten 1 sind und oberhalb der Leitkoeffizienten nur Nullen stehen.

Beispiel:

$$(A',b') \quad = \quad \begin{array}{c} \left(\begin{array}{cccccc|c} 1 & 0 & 3 & 0 & 0 & 8 & 2 \\ 0 & 1 & 2 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 & 0 & 5 & 6 \\ 0 & 0 & 0 & 0 & 1 & 4 & 0 \end{array}\right) \\ \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \end{array} \quad \text{ist in Gauß - Jordan Form.}$$

$\text{rang}\, A' = 4$ und $\text{rang}\,(A',b') = 4$
$Ax = b$ hat (mind.) eine Lösung.
$\Rightarrow$ Es existieren Lösungen, wir müssen weiter machen.

Schritt 2: Finde die Lösungsmenge U des homogenen Gleichungssystem $A'x = 0$.
Wähle hierzu die Unbekannten, die zu den Spalten ohne Leitkoeffizienten gehören, als freie Parameter.

Beispiel:
Im obigen Beispiel setzen wir $x_3 := \lambda, x_6 := \mu$.

Dann hat das homogene System die Lösungsmenge

$$\begin{aligned} x_1 &= -3\lambda - 8\mu \\ x_2 &= -2\lambda - \mu \\ x_3 &= \lambda \\ x_4 &= -5\mu \\ x_5 &= -4\mu \\ x_6 &= \mu \end{aligned}$$

Als Vektorgleichung:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} -3\lambda - 8\mu \\ -2\lambda - \mu \\ \lambda \\ -5\mu \\ -4\mu \\ \mu \end{pmatrix} = \lambda \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -8 \\ -1 \\ 0 \\ -5 \\ -4 \\ 1 \end{pmatrix}$$

$$U = \operatorname{span} \left(\begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -8 \\ -1 \\ 0 \\ -5 \\ -4 \\ 1 \end{pmatrix} \right)$$

Schritt 3: Suche eine spezielle Lösung w des inhomogenen Systems $A'x = b'$. Setze hierzu die Unbekannte, die zu den Spalten ohne Leitkoeffizienten gehören, gleich 0. (möglich, da dies freie Parameter sind).

Beispiel: Mit $x_3 = 0, x_6 = 0$ erlaubt die Gauß-Jordan-Form im obigen Beispiel die direkte Bestimmung von x_1, x_2, x_4, x_5:

$$x_1 = 2, x_2 = 4, x_4 = 6, x_5 = 0.$$

$$\Rightarrow w = \begin{pmatrix} 2 \\ 4 \\ 0 \\ 6 \\ 0 \\ 0 \end{pmatrix}$$

ist spezielle Lösung des inhomogenen Systems.

Schritt 4: Die Lösungsmenge von $Ax = b$ ist dann $w + U$.

Im Bespiel:

$$\text{Lös}\,(A, b) = \left\{ \begin{pmatrix} 2 \\ 4 \\ 0 \\ 6 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -8 \\ -1 \\ 0 \\ -5 \\ -4 \\ 1 \end{pmatrix} \middle| \lambda, \mu \in \mathbb{R} \right\}$$

38.18 Geometrische Interpretation liearer Gleichungssysteme

Sei $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$. Wir wissen: $\text{Lös}\,(A, 0) = \text{Ker} A$ ist Unterraum des $\mathbb{R}^n$.

(Jedoch ist $\text{Lös}\,(A, b)$ i.A. kein Unterraum)

Für die Dimension von Lös $(A,0)$ gilt:

$$\begin{aligned} \underbrace{\dim \mathrm{Ker} A}_{\dim \text{Lös } (A,0)} + \underbrace{\dim \mathrm{Im} A}_{\text{rang } A} &= \underbrace{\dim \mathbb{R}^n}_{n} \\ \Rightarrow \dim \text{Lös } (A,0) &= n - \text{rang } A. \end{aligned}$$

Beispiel:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}, b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$n = 3, \text{rang } A = 2$$

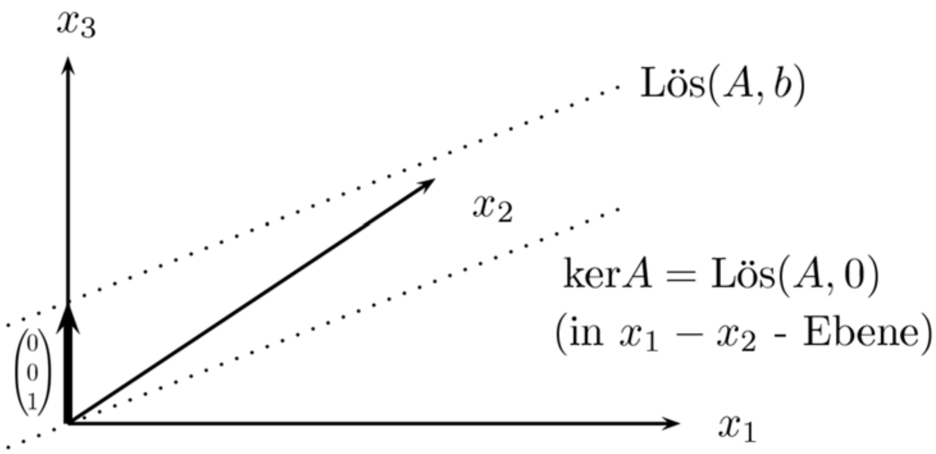
$$\Rightarrow \dim \text{Lös } (A,0) = 3 - 2 = 1.$$

$$\text{Lös } (A,0) = \left\{ \begin{array}{c} \lambda \\ \lambda \\ 0 \end{array} \middle| \lambda \in \mathbb{R} \right\} = \mathrm{Ker} A \qquad \text{Ursprungsgerade}$$

spezielle Lösung des inhomogen Systems: $w = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

allgemeine Lösung des inhomogen Systems:

$$\text{Lös } (A,b) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \middle| \lambda \in \mathbb{R} \right\} \qquad \text{Gerade.}$$



Kapitel 39

Iterative Verfahren für lineare Gleichungssysteme

39.1 Motivation

- Viele praktische Probleme führen auf sehr große lineare Gleichungssysteme, bei denen die Systemmatrix dünn besetzt ist, d.h. nur wenige Einträge von 0 verschieden sind.

  Beispiel:

  $$\begin{pmatrix} \ddots & \ddots & & & \ddots & & \mathbf{0} \\ \ddots & \ddots & \ddots & & & \ddots & \\ & \ddots & \ddots & \ddots & \mathbf{0} & & \ddots \\ & & \ddots & \ddots & \ddots & & \\ \ddots & & \mathbf{0} & \ddots & \ddots & \ddots & \\ & \ddots & & & \ddots & \ddots & \ddots \\ \mathbf{0} & & \ddots & & & \ddots & \ddots \end{pmatrix}$$

  Pentadiagonalmatrix bei 2-D Diffusionsfiltern in der Bildverarbeitung.

- Aus Speicherplatzgründen will man oft nur die von 0 verschiedenen Elemente abspeichern.
  Beispiel:
  Ein Grauwertbild mit 512x512 Pixel führt zu $512^2 = 262.144$ Gleichungen mit ebenso vielen Unbekannten. Bei 8 Byte / Eintrag und vollem Abspeichern benötigt die Pentadiagonalmatrix $8 \cdot 512^4 \approx 550$ GB, bei effizentem Abspeichern nur $5 \cdot 512^2 \approx$ 1.3 MB!

- Direkte Verfahren wie der Gauß-Algorithmus können die Nullen auffüllen und so zu einem enormen Speicherplatzbedarf führen.
  Zudem ist der Rechnenaufwand oft zu hoch: $\mathrm{O}(n^3)$ Operationen für ein $(n \times n)$-Gleichungssystem.

- Daher verwendet man oft iterative Näherungsverfahren, die kaum zusätzlichen Speicherplatz benötigen und nach wenigen Schritten eine brauchbare Approximation liefern.

39.2 Grundstruktur klassischer iterativerVerfahren

Geg.: $A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n$

Ges.: $x \in \mathbb{R}^n \text{ mit } Ax = b$

Falls $A = S - T$ mit einer einfach zu invertierenden Matrix S (z.B. Diagonalmatrix, Dreiecksmatrix), so kann man $Ax = b$ umformen in

$$Sx = Tx + b$$

und mit einem Startwert $x^{(0)}$ die Fixpunktiteration

$$Sx^{(k+1)} = Tx^{(k)} + b \qquad (k = 0, 1, 2, \ldots)$$

anwenden.

Wir wollen nun drei verschiedene Aufspaltungen $A = S - T$ untersuchen. Sei hierzu $A = D - L - R$ eine Aufspaltung in eine Diagonalmatrix D, eine strikte untere Dreiecksmatrix L und eine strikte obere Dreiecksmatrix R.

$$\underbrace{\begin{bmatrix} & & \\ & * & \\ & & \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} * & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & * \end{bmatrix}}_{D} - \underbrace{\begin{bmatrix} 0 & \ldots & 0 \\ & \ddots & \vdots \\ * & & 0 \end{bmatrix}}_{L} + \underbrace{\begin{bmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \ldots & 0 \end{bmatrix}}_{R} -$$

39.3 Das Jacobi-Verfahren (Gesamtschrittverfahren)

Hierzu wählt man S:= D und T:= L + R.

In jeder Iteration wird also nur das Diagonalsystem

$$Dx^{(k+1)} = (L + R)x^{(k)} + b$$

nach $x^{(k+1)}$ aufgelöst, d.h. nur durch die Diagonalelemente dividiert:

$$x^{(k+1)} = D^{-1}\Big((L + R)x^{(k)} + b\Big)$$

explizit:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(-\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)} + b_i \right) \qquad (i = 1, \ldots, n)$$

39.4 Beispiel

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

führt auf das Diagonalsystem

$$\left.\begin{aligned} 2x_1^{(k+1)} &= x_2^{(k)} + 3 \\ 2x_2^{(k+1)} &= x_1^{(k)} + 4 \end{aligned}\right\} \Rightarrow \begin{aligned} x_1^{(k+1)} &= \frac{1}{2}\Big(x_2^{(k)} + 3\Big) \\ x_2^{(k+1)} &= \frac{1}{2}\Big(x_1^{(k)} + 4\Big) \end{aligned} .$$

Bei vierstelliger Genauigkeit und Startvektor $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ erhält man

k	$x_1^{(k)}$	$x_2^{(k)}$
0	0	0
1	1,5	2
2	2,5	2,75
3	2,875	3,25
4	3,125	3,438
5	3,219	3,568
6	3,282	3,610
7	3,305	3,641
8	3,321	3,653
9	3,327	3,661
10	3,331	3,664
11	3,332	3,666
12	3,333	3,666

Exakte Lösung: $x_1 = 3\frac{1}{3}, x_2 = 3\frac{2}{3}$.

Das Jacobi-Verfahren ist gut geeignet für Parallelrechner, da $x_i^{(k+1)}$ nicht von $x_j^{(k+1)}$ abhängt.

39.5 Das Gauß-Seidel-Verfahren (Einzelschrittverfahren)

$S := D - L, T := R$

In jeder Iteration wird also das Dreieckssystem

$$(D-L)x^{(k+1)} = Rx^{(k)} + b$$

durch einfache Vorwärtssubstitution gelöst:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(-\sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} + b_i \right) \qquad (i = 1, \ldots, n)$$

d.h. neue Werte werden weiter verwendet, sobald sie berechnet wurden.

39.6 Beispiel:

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

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{2}\left(x_2^{(k)} + 3\right) \\ x_2^{(k+1)} &= \frac{1}{2}\left(x_1^{(k+1)} + 4\right) \end{aligned}$$

k	$x_1^{(k)}$	$x_2^{(k)}$
0	0	0
1	1,5	2,75
2	2,875	3,348
3	3,174	3,587
4	3,294	3,647
5	3,324	3,662
6	3,331	3,666

Konvergiert etwa doppelt so schnell wie Jacobi-Verfahren.

39.7 Das SOR-Verfahren (SOR = successive overrelaxation)

Beschleunigung des Gauß-Seidel-Verfahrens durch Extrapolation

$$x^{(k+1)} = x^{(k)} + \omega \left(\tilde{x}^{(k+1)} - x^{(k)} \right)$$

mit $\omega \in (1,2)$, wobei $\tilde{x}^{(k+1)}$ die Gauß-Seidel-Iteration zu $x^{(k)}$ ist.

Für große Gleichungssysteme kann damit die Iterationszahl (um eine bestimmte Genauigkeit zu erreichen) für ein geeignetes ω um eine Zehnerpotenz gesenkt werden.

39.8 Konvergenzresulate

(a) Die Jacobi-, Gauß-Seidel- und SOR-Verfahren können als Fixpunktiterationen aufgefasst werden. Nach 26.6 liegt Konvergenz vor, wenn die Abbildung kontrahierend ist. Dies ist <u>nicht</u> in allen Fällen erfüllt.

(b) Für wichtige, praxisrelevante Fälle existieren jedoch Konvergenzaussagen, z.B.: Ist die Systemmatrix $A = (A_{ij}) \in \mathbb{R}^{n \times n}$ <u>streng</u> <u>diagonaldominant</u> (d.h. $|a_{ij}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad \forall i$), so konvergieren das Jacobi- und das Gauß- Seidel-Verfahren.

39.9 Effizienz

(a) Die vorgestellten Verfahren sind die einfachsten, allerdings nicht die effizientesten iterativen Verfahren für lineare Gleichungssysteme.

(b) Effizientere, aber kompliziertere Verfahren :

- ADI - Verfahren (ADI = alternating directions implicit)
  (mehrdimensionale Probleme werden in eindimensionale Probleme umgewandelt)
- PCG - Verfahren (PCG = preconditioned conjugate gradients)
  (suchen in Unterräumen)
- Mehrgitterverfahren (multigrid)
  (erst grobe Berechnung, dann feiner)

Siehe numerische Spezialliteratur.

(c) Hocheffiziente Ansätze wie die Mehrgitterverfahren verwenden oft das Gauß-Seidel-Verfahren als Grundbaustein. Mit ihnen ist es z.T. möglich, lineare Gleichungssysteme in <u>optimaler</u> Komplexität (d.h. in $\mathrm{O}(n)$) zu lösen.

Kapitel 40

Determinaten

40.1 Motivation

- Gibt es eine möglichst „aussagekräftige" Abbildung, die eine Matrix $A \in K^{n \times n}$ auf eine Zahl in K reduziert.
- Die Determinate ist die „sinnvollste" Art, eine solche Abbildung zu definieren.
- Sie kann axiomatische fundierte werden und liefert nützliche Aussagen:
  - über die Invertierbarkeit einer Matrix
  - über die Lösung des linearen Gleichungssystem $Ax = b$
  - über das Volumen eines Parallelepipeds

Welche Forderung soll die Determinate erfüllen?

40.2 Definition (Determinantenfunktion, Determinante)

Sei K ein Körper: Eine Abbildung: $\det\ : K^{n \times n} \to K$ heißt Determinantenfunktion (Determinante), falls gilt:

(a) $\det$ *ist linear in jeder Zeile:*

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ \lambda z_i + \mu z_i' \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix} = \lambda \det \begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_i \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix} + \mu \det \begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_i' \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix}$$

für $i = 1, \ldots, n$ und $\lambda, \mu \in K$.

(Man sagt auch die Determinante ist eine Multilinearform)

(b) *Ist $\operatorname{rang} A < n$, so ist $\det A = 0$.*

(c) $\det I = 1$ *für die Einheitsmatrix $I \in K^{n \times n}$.*

Bemerkung:
Die Determinate ist nur für quadratische Matrizen definiert!

Gibt es sehr viele Determinantenfunktionen?
Nein! In einem aufwändigen Beweis (siehe z.B. Beutelspacher: Lineare Algebra) kann man zeigen:

40.3 Satz: (Eindeutigkeit der Determinante)

Zu jedem $n \in \mathbb{N}$ existiert genau eine Determinantenfunktion.

Mit anderen Worten: Die Forderungen (a) - (c) aus 40.2 liefern eine axiomatische Fundierung des Determinantenbegriffs.

Wie berechnet man Determinanten? Hierzu betrachten wir zunächst nur 2×2-Matrizen.

40.4 Satz: (Determinaten einer 2×2-Matrix)

Die Determinante einer Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2\times 2}$ *ist gegeben durch*

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{vmatrix} a & b \\ c & d \end{vmatrix} := a \cdot d - b \cdot c.$$

Beweis:

Wir zeigen, dass diese Abbildung Def. 40.2 erfüllt.

(a)

$$\begin{aligned} \det \begin{pmatrix} \lambda a_1 + \mu a_2 & \lambda b_1 + \mu b_2 \\ c & d \end{pmatrix} &= (\lambda a_1 + \mu a_2) \cdot d - (\lambda b_1 + \mu b_2) \cdot c \\ &= \lambda a_1 d + \mu a_2 d - \lambda b_1 c - \mu b_2 c \\ &= \lambda (a_1 d - b_1 c) + \mu (a_2 d - b_2 c) \\ &= \lambda \begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix} + \mu \begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix} \end{aligned}$$

Die Linearität in der 2. Zeile zeigt man analog.

(b) Wenn die Matrix nur aus Nullen besteht, ist ihrer Determinante Null. Hat die Matrix $\text{rang } 1$, so ist $\begin{pmatrix} a \\ c \end{pmatrix} = \lambda \begin{pmatrix} b \\ d \end{pmatrix}$, d.h.

$$\begin{aligned} \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \det \begin{pmatrix} \lambda b & b \\ \lambda d & d \end{pmatrix} \\ &= \lambda bd - \lambda bd = 0. \end{aligned}$$

(c) $\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot 1 - 0 \cdot 0 = 1.$

$\square$.

Bemerkung: Für eine 1×1-Matrix $a \in K^{1\times 1}$ ist $\det(a) = a$.

Determinanten von $(n \times n)$-Matrizen lassen sich rekursiv auf 2×2-Determinanten zurückführen. Hierzu benötigen wir:

40.5 Definition (Unterdeterminate)

Sei $A \in K^{n\times n}$. Die aus einer $n \times n$-Determinanten $D = \det A$ durch Streichung der i-ten Zeile und j-ten Spalte entstehende $(n-1) \times (n-1)$-Determinante D_{ij} nennen wir Unterdeterminate von D.

Der Ausdruck $A_{ij} := (-1)^{i+j} D_{ij}$ heißt algebraisches Komplement des Elementes a_{ij} in der Determinante D.

40.6 Beispiel

$$
\begin{aligned}
A &= \begin{pmatrix} 3 & 9 & 1 \\ 2 & 5 & 4 \\ -2 & 8 & 7 \end{pmatrix} \\
D_{23} &= \begin{vmatrix} 3 & 9 \\ -2 & 8 \end{vmatrix} = 3 \cdot 8 - (-2) \cdot 9 = 42 \\
A_{23} &= (-1)^{2+3} \cdot D_{23} = -42
\end{aligned}
$$

Kommen wir nur zur rekursiven Berechnung einer $n \times n$-Determinante.

40.7 Satz: (Laplace'scher Entwicklungssatz)

Man kann eine $n \times n$-Determinate berechnen, indem man die Elemente einer Zeile (oder Spalte) mit ihrem algebraischen Komplemente multipliziert und die Produkte addiert.

40.8 Beispiel:

(a) Entwicklung einer 3×3-Determinante nach der 2. Zeile:

$$
\begin{aligned}
\begin{vmatrix} 3 & 9 & 1 \\ 2 & 5 & 4 \\ -2 & 8 & 7 \end{vmatrix} &= -2 \begin{vmatrix} 9 & 1 \\ 8 & 7 \end{vmatrix} + 5 \begin{vmatrix} 3 & 1 \\ -2 & 7 \end{vmatrix} - 4 \begin{vmatrix} 3 & 9 \\ -2 & 8 \end{vmatrix} \\
&= -2(63-8) + 5(21+2) - 4(24+18) = \ldots = -163.
\end{aligned}
$$

(b) Entwicklung nach der 3. Spalte:

$$
\begin{aligned}
\begin{vmatrix} 3 & 9 & 1 \\ 2 & 5 & 4 \\ -2 & 8 & 7 \end{vmatrix} &= 1 \begin{vmatrix} 2 & 5 \\ -2 & 8 \end{vmatrix} - 4 \begin{vmatrix} 3 & 9 \\ 2 & 8 \end{vmatrix} + 7 \begin{vmatrix} 3 & 9 \\ -2 & 5 \end{vmatrix} \\
&= 1(16+10) - 4(24+18) + 7(15-18) = \ldots = -163.
\end{aligned}
$$

Wie rechnet man mit Determinanten?

40.9 Rechenregel für Determinanten

(a) Transponieren verändert den Wert einer Determinante nicht:

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

(folgt aus dem Laplace'schen Entwichklungssatz, indem man die Entwicklung nach Zeilen und Spalten vertauscht).

(b) Aus Def. 40.2 (b) folgt: Sind Spaltenvektoren oder Zeilenvektoren linear abhängig, so ist die Determinate Null.

Beispiel: $\begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 6 & 6 & 6 \end{vmatrix} = 0$

(c) Additiert man zu einer Zeile (oder Spalte) das vielfache einer anderen Zeile (Spalte), so bleibt die Determinante gleich.

Beispiel: $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 6 & 6 & 6 \end{vmatrix} \stackrel{(b)}{=} 0$

(1. Zeile von 2. und 3. abgezogen)

(d) Vertauscht man zweit Zeilen oder Spalten, so ändert die Determinante ihr Vorzeichen.

(e) Die Determinante von Dreiecksmatrizen ist das produkt der Diagonalelemente

(folgt durch rekursives Anwenden des Laplac'schen Entwicklungssatzes):

$\begin{vmatrix} 3 & 0 & 9 \\ 0 & -7 & 4 \\ 0 & 0 & 2 \end{vmatrix} = 3 \cdot (-7) \cdot 2 = -42$

Man kann also mit dem Gauß-Algorithmus die Matrix auf Dreiecksgestalt bringen (unter Beachtung von (d)), um dann ihre Determinante bequem zu berechnen. Für große n ist dies wesentlich effizenter als der Laplace'sche Entwicklungsatz.

(f) Für $A, B \in K^{n\times n}$ gilt: $\boxed{\det (A \cdot B) = \det A \cdot \det B}$

(g) Folgerung für eine invertierbare Matrix A gilt:

$1 = \det I = \det (A \cdot A^{-1}) = \det A \cdot (\det A^{-1})$

$\Rightarrow \boxed{\det (A^{-1}) = \dfrac{1}{\det A}}$

(h) Vorsicht: Für $A \in K^{n\times n}$, $\lambda \in K$ gilt:

$$\boxed{\det (\lambda A) = \lambda^n \det A}$$

(und nicht etwas $\det (\lambda A)) = \lambda \det A$, denn $\det$ ist linear in jeder Zeile).

Wozu sind Determinanten nützlich?

40.10 Bedeutung der Determinaten

(a) Mit ihnen kann man testen, ob eine Matrix $A \in K^{n\times n}$ invertierbar ist:

$$\boxed{A \in K^{n\times n} \text{ist invertierbar} \Leftrightarrow \det A \neq 0}$$

(b) Man kann mit ihnen lineare Gleichungssysteme lösen.
(für numerische Rechnungen ist dies jedoch ineffizient):

Cramer'sche Regel: Ist $A = (a_{*1}, \ldots, a_{*n}) \in K^{n\times n}$ invertierbar und $b \in K^n$, so läßt sich die Lösung des linearen Gleichungssystems $Ax = b$ angeben durch

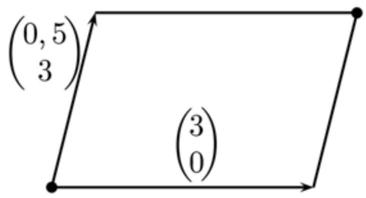
$$x_k = \frac{\det (a_{*1}, \ldots, a_{*k-1}, b, a_{*k+1}, \ldots, a_{*n})}{\det A} \qquad (k = 1, \ldots, n)$$

Beispiel:
$\begin{pmatrix} 2 & 5 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \end{pmatrix}$

$$x_1 = \frac{\begin{vmatrix} -2 & 5 \\ 6 & 4 \end{vmatrix}}{\begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix}} = \frac{-8-30}{8-5} = \frac{-38}{3}$$

$$x_2 = \frac{\begin{vmatrix} 2 & -2 \\ 1 & 6 \end{vmatrix}}{\begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix}} = \frac{12+2}{3} = \frac{14}{3}$$

(c) $|\det A|$ ist das Volumen des durch die Spaltenvektoren von A aufgespannten Parallelepids.
Beispiel: Parallelogrammfläche:



$$\left| \det \begin{pmatrix} 3 & 0.5 \\ 0 & 2 \end{pmatrix} \right| = |3 \cdot 2 - 0 \cdot 0,5| = |6|.$$

Kapitel 41

Euklidische Vektorräume

41.1 Motivation

Im $\mathbb{R}^2$ und $\mathbb{R}^3$ kann das Skalarprodukt zweier Vektoren gebildet werden. Mit seiner Hilfe lassen sich Längen von Vektoren bestimmen sowie feststellen, ob Vektoren senkrecht zueinander sind; allgemein können damit auch Winkel zwischen Vektoren berechnet werden.

Ziel: Wir wollen dieses Konzept auf andere Vektorräume ausdehnen und auch in diesen Skalarprodukt, Längen- und Winkelbestimmung, Orthogonalität bereitstellen.

41.2 Definition (Vektorengleichheit)

Es seien $u = (u_1, \dots, u_n)^\top \in \mathbb{R}^n, v = (v_1, \dots, v_n)^\top \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Die Vektoren u, v heißen gleich, falls $u_i = v_i$ $(i = 1, \dots, n)$. Die Summe von u und v ist definiert durch

$$u + v = (u_1 + v_1, \dots, u_n + v_n)^\top$$

sowie das skalare Vielfache αV durch

$$\alpha V = (\alpha v_1, \dots, \alpha v_n)^\top.$$

Bemerkung:

(a) Der Nullvektor im $\mathbb{R}^n$ ist gegeben durch $\quad 0 := (0, \dots, 0)^\top$.

(b) Das (additive) Inverse $-u$ des Vektors u lautet $\quad -u = (-u_1, \dots, -u_n)^\top$.

(c) Die Differenz zweier Vektoren lautet $\quad u - v := u + (-v) = (u_1 - v_1, \dots, u_n - v_n)^\top$.

Mit der in Def. 41.2 eingeführten Addition und Skalarmultiplikation wird der $\mathbb{R}^n$ zum Vektorraum. Es gilt nämlich

41.3 Satz: (Vektorraumeigenschaften des $\mathbb{R}^n$)

Es seien $u, v, w \in \mathbb{R}^n$ und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

(a) $(\mathbb{R}^n, +)$ *ist eine abelsche Gruppe:*

  i) *Kommutativgesetz:* $u + v = v + u$

  ii) *Assoziativgesetz:* $(u + v) + w = u + (v + w)$

  iii) *Neutrales Element:* $u + 0 = 0 + u = 0$

  iv) *Inverses Element:* $u + (-u) = 0$

(b) $\alpha(\beta u) = (\alpha\beta)u$

(c) $\alpha(u + v) = \alpha u + \alpha v$

(d) $(\alpha + \beta)u = \alpha u + \beta u$

(e) $1 \cdot v = v.$

Beweis:

Einfaches Anwenden von Def. 41.2.

41.4 Definition (Euklidisches Produkt)

Es seien $u = (u_1, \ldots, u_n)^\top, v = (v_1, \ldots, v_n)^\top$ Vektoren im $\mathbb{R}^n$. Das Euklidische Produkt $u \cdot v$ wird definiert durch

$$u \cdot v := \sum_{i=1}^{n} u_i v_i.$$

41.5 Beispiel:

$$u = \begin{pmatrix} -1 \\ 3 \\ 5 \\ 7 \end{pmatrix}, v = \begin{pmatrix} 5 \\ -4 \\ 7 \\ 0 \end{pmatrix} \Rightarrow u \cdot v = (-1) \cdot 5 + 3 \cdot (-4) + 5 \cdot 7 + 7 \cdot 0 = 18.$$

Bemerkung: Den Vektorraum $\mathbb{R}^n$ versehen mit dem euklidischen Produkt (Skalarprodukt) bezeichnet man als n-dimensionalen euklidischen Raum.

41.6 Satz: (Eigenschaften des euklidischen Produkts)

Es seien $u, v, w \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Dann gilt:

(a) $u \cdot v = v \cdot u$

(b) $(u + v) \cdot w = u \cdot w + v \cdot w$

(c) $(\alpha u) \cdot v = \alpha(u \cdot v)$

(d) $v \cdot v \geqq 0$,
$v \cdot v = 0 \Leftrightarrow v = 0.$

Beweis:

Wir zeigen nur (b) und (d).

(b)

$$\begin{aligned}(u+v)\cdot w &= (u_1+v_1,\dots,u_n+v_n)^\top \cdot (w_1,\dots,w_n)^\top \\ &= \sum_{i=1}^{n}(u_i+v_i)w_i = \sum_{i=1}^{n} u_i w_i + \sum_{i=1}^{n} v_i w_i \\ &= u\cdot w + v\cdot w\end{aligned}$$

(d) $v\cdot v = v_1^2+v_2^2+\ldots+v_n^2 \geqq 0$
Gleichheit gilt genau dann, wenn $v_1 = v_2 = \ldots = 0$ d.h. wenn $v = 0$.

□.

41.7 Definition (Euklidische Norm, Euklidischer Abstand)

Die euklidische Norm eines Vektors $v = (v_1,\dots,v_n)^\top \in \mathbb{R}^n$ ist definiert durch

$$\mid v \mid := \sqrt{v\cdot v} = \sqrt{v_1^2+v_2^2+\ldots+v_n^2}.$$

Der euklidische Abstand zweier Vektoren $u = (u_1,\dots,u_n)^\top$ und $v = (v_1,\dots,v_n)^\top$ ist definiert durch

$$d(u,v) := \mid u-v \mid = \sqrt{(u_1-v_1)^2+\ldots+(u_n-v_n)^2}.$$

41.8 Beispiel

$$u = \begin{pmatrix}1\\3\\-2\\7\end{pmatrix}, v = \begin{pmatrix}0\\7\\2\\2\end{pmatrix}$$

$\mid u \mid = \sqrt{1+9+4+49} = \sqrt{63} = 3\sqrt{7}$

$d(u,v) = \sqrt{(1-0)^2+(3-7)^2+(-2-2)^2+(7-2)^2} = \sqrt{1+16+16+25} = \sqrt{58}.$

41.9 Satz: (Cauchy-Schwarz'sche Ungleichung im $\mathbb{R}^n$)

Für $u,v\in\mathbb{R}^n$ gilt $|u\cdot v| \leqq |u|\cdot|v|$.

Beweis:

in der nächesten Vorlesung (in allgemeiner Form)

41.10 Satz: (Eigenschaften der euklidischen Norm)

Es seien $u, v \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Dann gilt:

(*a*) $|u| \geqq 0$

(*b*) $|u| = 0 \Leftrightarrow u = 0$

(*c*) $|\alpha u| = |\alpha|\ |u|$

(*d*) $|u+v| \leqq |u| + |v|$ *(Dreiecksungleichung).*

Beweis:

Wir zeigen nur (c) und (d):

(c)

$$\begin{aligned} |\ \alpha u\ | &= \sqrt{(\alpha u_1)^2 + \ldots + (\alpha u_n)^2} \\ &= \sqrt{\alpha^2 (u_1^2 + \ldots + u_n^2)} \\ &= \sqrt{\alpha^2}\sqrt{(u_1^2 + \ldots + u_n^2)} \\ &= |\alpha|\ |u| \end{aligned}$$

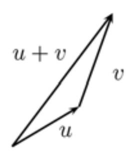
(d)

$$\begin{aligned} |\ u+v\ |^2 &\overset{\text{Def}}{=} (u+v)\cdot(u+v) \\ &\overset{41.6(b)}{=} u^2 + u\cdot v + v\cdot u + v\cdot v \\ &= |u|^2 + 2u\cdot v + |v|^2 \\ &\underbrace{\leq}_{Cauchy-Schwarz (41.9)} |u|^2 + 2\cdot|u|\ |v| + |v|^2 \\ &= (|u| + |v|)^2 \end{aligned}$$

$\Rightarrow |u+v| \leq |u| + |v|.$ $\square$.

Bedeutung der Dreiecksungleichung

Die Länge zweier Dreiecksseiten ist nie kleiner als die der dritten.



41.11 Satz: (Eigenschaften des euklidischen Abstandes)

Für $u, v, w \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$ gelten:

(a) $d(u, v) \geq 0$

(b) $d(u, v) = 0 \Leftrightarrow u = v$

(c) $d(u, v) = d(v, u)$

(d) $d(u, v) \leq d(u, w) + d(w, v)$.

Beweis:

Alle Eigenschaften ergeben sich als direkte Folgerung aus Satz 41.10

41.12 Definition (Orthogonalität)

Zwei Vektoren $u, v \in \mathbb{R}^n$ heißen orthogonal, falls $u \cdot v = 0$.

Beispiel:

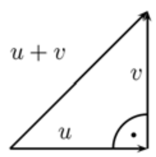
$$u = \begin{pmatrix} -2 \\ 3 \\ 1 \\ 4 \end{pmatrix}, v = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}$$

$u \cdot v = -2 + 6 + 0 - 4 = 0 \Rightarrow u$ und v sind orthogonal.

Für orthogonale Vektoren folgt aus der Dreiecksungleichung

41.13 Satz: (Satz des Pythagoras im $\mathbb{R}^n$)

Sind $u, v \in \mathbb{R}^n$ orthogonal, so gilt $\mid u + v \mid^2 = \mid u \mid^2 + \mid v \mid^2$.
Hypotenusenquadrat = Summe der Kathetenquadrate



Beweis:

$$|u + v|^2 = (u + v) \cdot (u + v) = |u|^2 + \underbrace{2u \cdot v}_{\substack{=0 \\ (Orthogonalität)}} + |v|^2 = |u|^2 + |v|^2. \qquad \square.$$

Beispiel:

$$u = \begin{pmatrix} -2 \\ 3 \\ 1 \\ 4 \end{pmatrix}, v = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}$$ sind orthogonal;

$$u + v = \begin{pmatrix} -1 \\ 5 \\ 1 \\ 3 \end{pmatrix}$$

$$\begin{aligned} |u|^2 &= 4 + 9 + 1 + 16 = 30 \\ |v|^2 &= 1 + 4 + 0 + 1 = 6 \\ u + v &= (-1, 5, 1, 3)^\top \\ |u + v|^2 &= 1 + 25 + 1 + 9 = 36 \end{aligned}$$

41.14 Interpretation des euklidischen Produktes als Matrixmultiplikation

Es seien $u, v \in \mathbb{R}^n$. Dann kann man das euklidische Produkt als Multiplikation der $1 \times n$-Matrix $u^\top$ mit der $n \times 1$-Matrix v auffassen.

$$u \cdot v = u^\top v$$

Beispiel:

$$u = \begin{pmatrix} 1 \\ -3 \\ 7 \\ 4 \end{pmatrix}, v = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 9 \end{pmatrix}$$

$$\begin{aligned} u^\top v &= \begin{pmatrix} 1 & -3 & 7 & 4 \end{pmatrix} \begin{pmatrix} 0, \\ 2, \\ 1, \\ 9, \end{pmatrix} \\ &= 1 \cdot 0 + (-3) \cdot 2 + 7 \cdot 1 + 4 \cdot 9 = 37 = u \cdot v. \end{aligned}$$

Beachte: Vektoren im $\mathbb{R}^n$ sind für uns stets Spaltenvektoren.

Kapitel 42

Funktionalanalytische Verallgemeinerungen

42.1 Motivation

Die Ideen des euklidischen Produktes, Norm und Abstandes sollen abstrahiert werden, um diese Konzepte auf andere Räume übertragen zu können. Dies ist auch für die Anwendungen wesentlich, z.B. in der Signal- und Bildverarbeitung (z.B. Fouriertransformation).

42.2 Definition (Prä-Hilbert-Raum)

Sei V ein reeller Vektorraum. Ein Skalarprodukt (inneres Produkt) ist eine Funktion $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ mit den folgenden Eigenschaften:

i)	*Symmetrie:*	$\langle u,v\rangle = \langle v,u\rangle$	$\forall u,v \in V$
ii)	*Additivität:*	$\langle u+v,v\rangle = \langle u,w\rangle + \langle v,w\rangle$	$\forall u,v,w \in V$
iii)	*Homogenität:*	$\langle \alpha u,v\rangle = \alpha\,\langle u,v\rangle$	$\forall\, u,v \in V, \alpha \in \mathbb{R}$
iv)	*Nichtnegativität:*	$\langle v,v\rangle\ \geq 0$	$\forall v \in V$
	Nichtdegeneriertheit:	$\langle v,v\rangle = 0 \Leftrightarrow v = 0.$	

$(V,\langle\cdot,\cdot\rangle)$ heißt Prä-Hilbert-Raum.

Bemerkung: Ist V zudem vollständig (d.h. jede Cauchy-Folge konvergiert), so heißt $(V,\langle\cdot,\cdot\rangle)$ Hilbertraum.

42.3 Beispiele

(a) Euklidische Räume
Der n-dimensionale euklidische Raum bildet einen Prä-Hilbert-Raum. Nach Satz 41.6 sind alle Eigenschaften von Def. 42.2 erfüllt.

(b) Gewichtete euklidische Räume
Seien $u = (u_1,u_2)^\top$ und $v = (v_1,v_2)^\top$ Vektoren im $\mathbb{R}^2$. Dann wird durch $\langle u,v\rangle := 3u_1v_1 + 5u_2v_2$ ein Skalarprodukt definiert.

Beweis:

$$\begin{array}{llll}
\text{i)} & \langle u,v\rangle & = & 3u_1v_1+5u_2v_2=\langle v,u\rangle \qquad \forall u,v\in\mathbb{R}^2\\
\text{ii)} & \langle u+v,w\rangle & = & 3(u_1+v_1)w_1+5(u_2+v_2)w_2\\
 & & = & (3u_1w_1+5u_2w_2)+(3v_1w_1+5v_2w_2)\\
 & & = & \langle u,w\rangle+\langle v,w\rangle \qquad \forall u,v,w\in\mathbb{R}^2\\
\text{iii)} & \langle \alpha u,v\rangle & = & 3\alpha u_1v_1+5\alpha u_2v_2=\alpha\,\langle u,v\rangle \qquad \forall \alpha\in\mathbb{R}\\
\text{iv)} & \langle v,v\rangle & = & \underbrace{3v_1^2}_{\geq 0}+\underbrace{5v_2^2}_{\geq 0}\geq 0
\end{array}$$

Klar: $< v,v >= 0 \Leftrightarrow v_1=v_2=0$.

(c) Polynomräume

Seien $p:=\sum\limits_{k=0}^{n}a_kx^k, q:=\sum\limits_{k=0}^{n}b_kx^k$ Polynome vom Grad $\leq n$.

Dann wird mittels $\langle p,g\rangle := \sum\limits_{k=0}^{n}a_kb_k$ der reelle Vektorraum der Polynome vom Grad $\leq n$ zum Prä-Hilbert-Raum.

(d) Funktionenraum $C[a,b]$

Sei $C[a,b]:=\{f:[a,b]\to\mathbb{R}\mid f$ stetig auf $[a,b]\}$ und seien $f,g\in C[a,b]$. Dann wird $C[a,b]$ mit

$$\langle f,g\rangle := \int\limits_a^b f(x)g(x)\,\mathrm{d}x$$

zum Prä-Hilbert-Raum.

Läßt sich die euklidische Norm verallgemeinern ?

42.4 Definition (Norm)

Sei V ein reeller Vektorraum. Unter einer Norm auf V versteht man eine Abbildung $\|\cdot\|:V\to\mathbb{R}$ mit folgenden Eigenschaften:

i) $\|v\|\geq 0 \qquad \forall v\in V$

ii) $\|v\|=0\Leftrightarrow v=0$

iii) $\|\alpha v\|=|\alpha|\,\|v\| \qquad \forall v\in V,\forall\alpha\in\mathbb{R}$

iv) $\|u+v\|\leq\|u\|+\|v\| \qquad \forall u,v\in V.$

$(V,\|\cdot\|)$ *heißt normierter Raum.*

Bemerkung: Ein vollständiger normierter Raum heißt auch Banachraum.

42.5 Satz: (Cauchy-Schwarz-Ungleichung im Prä-Hilbert-Raum)

Sei $(V, \langle\cdot,\cdot\rangle)$ Prä-Hilbert-Raum. Dann gilt:

$$\langle u, v\rangle^2 \;\le\; \langle u, u\rangle \cdot \langle v, v\rangle \qquad \forall u, v \in V.$$

Beweis:

Falls $u = 0$ ist, so sind beide Seiten der Ungleichung 0. Sei nun $u \neq 0$. Dann gilt für alle $x \in \mathbb{R}$ und $v \in V$:

$$0 \;\le\; \langle ux + v, ux + v\rangle \;= \underbrace{\langle u, u\rangle}_{=:a} x^2 + 2\underbrace{\langle u, v\rangle}_{=:b} x + \underbrace{\langle v, v\rangle}_{=:c}$$

Die Parabel $ax^2 + bx + c$ hat also höchstens eine reelle Nullstellen. Die Diskriminante erfüllt also $0 \ge b^2 - 4ac = 4\langle u, v\rangle^2 - 4\langle u, u\rangle \langle v, v\rangle$. Daraus folgt die Behauptung. □.

42.6 Satz: (Induzierte Norm von Prä-Hilbert-Räumen)

Jeder Prä-Hilbert-Raum $(V, \langle\cdot,\cdot\rangle)$ wird mit $\|v\| := \sqrt{\langle v, v\rangle}$ zum normierten Raum.

Beweis:

(i), (ii) folgen direkt aus Def 42.2 (iv).

zu iii)
$$\begin{aligned}
\|\alpha v\| &= \sqrt{\langle \alpha v, \alpha v\rangle} && \text{Def.}\\
&= \sqrt{\alpha \langle v, \alpha v\rangle} && \text{Homogenität}\\
&= \sqrt{\alpha \langle \alpha v, v\rangle} && \text{Symmetrie}\\
&= \sqrt{\alpha^2 \langle v, v\rangle} && \text{Homogenität}\\
&= |\alpha|\, \|v\| && \text{Def.}
\end{aligned}$$

zu iv)
$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v\rangle && \text{Def.}\\
&= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle && \text{Additivität und Symmetrie}\\
&= \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 && \text{Symmetrie, Def.}\\
&\le \|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2 && \text{Cauchy-Schwarz-Ungleichung}\\
&= (\|u\| + \|v\|)^2
\end{aligned}$$

□.

42.7 Beispiele

(a) Norm einer stetigen Funktion
$C[a, b]$ wird mit

$$\|f\| := \left(\int_a^b f^2(x)\,\mathrm{d}x\right)^{1/2} \qquad \forall f \in C[a, b]$$

zum normierten Raum. Beispielsweise hat $f(x) = \frac{1}{x}$ im Intervall $[1, 2]$ die Norm

$$\|f\| = \sqrt{\int_1^2 \frac{1}{x^2}\,\mathrm{d}x} = \sqrt{\left[-\frac{1}{x}\right]_1^2} = \sqrt{-\frac{1}{2} + 1} = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}.$$

(b) Gewichtete euklidische Norm
Der $\mathbb{R}^2$ versehen mit dem Skalarprodukt

$$\langle u, v\rangle = \frac{1}{9}u_1 v_1 + \frac{1}{4}u_2 v_2 \quad \forall u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^2, \quad \forall v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$$

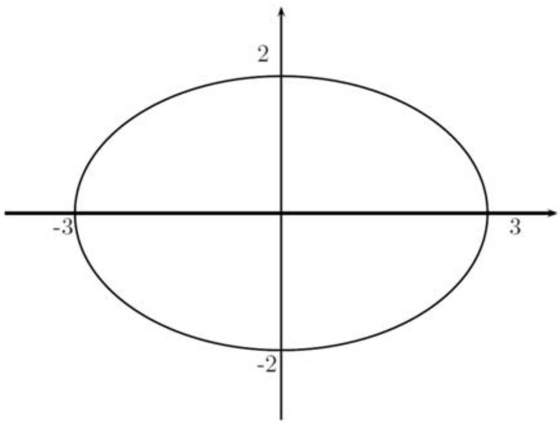
induziert die Norm

$$\|u\| := \sqrt{\frac{u_1^2}{9} + \frac{u_2^2}{4}}.$$

Der Einheitskreis bzgl. dieser Norm (d.h. die Menge aller $u \in \mathbb{R}^2$ mit $\|u\| = 1$) ist gegeben durch

$$\frac{u_1^2}{9} + \frac{u_2^2}{4} = 1.$$

Das ist eine Ellipsengleichung $\frac{u_1^2}{a^2} + \frac{u_2^2}{b^2}$ mit den Halbachsen $a = 3$ und $b = 2$.



Einheitskreise in solchen Normen sind nicht immer rund.

Kann man den Begriff des euklidischen Abstandes verallgemeinern?

42.8 Definition (Metrik, metrischer Raum)

Sei V ein Vektorraum über $\mathbb{R}$. Eine Abbildung $d : V \times V$ heißt Metrik, falls gilt:

i)	$d(u,v) \geq 0$	$\forall u, v \in V$
ii)	$d(u,v) = 0 \Leftrightarrow u = v$	
iii)	$d(u,v) = d(v,u)$	$\forall u, v \in V$
iv)	$d(u,v) \leq d(u,w) + d(w,v)$	$\forall u, v, w \in V$.

(V, d) heißt metrischer Raum.

Bemerkung: Für vollständige metrische Räume gibt es keinen besonderen Namen.

42.9 Satz: (Induzierte Metrik eines normierten Raumes)

Jeder normierte Raum $(V, \|\cdot\|)$ definiert mit

$$d(u,v) := \|u - v\| \ \forall u, v \in V$$

einen metrischen Raum (V, d).

Beweis:

einfache Folgerung aus Def. 42.4.

42.10 Beispiel: Metrik auf $C[a,b]$

Für $f, g \in C[a,b]$ kann man durch

$$d(f,g) := \left(\int_a^b (f(x) - g(x))^2 \,\mathrm{d}x \right)^{\frac{1}{2}}$$

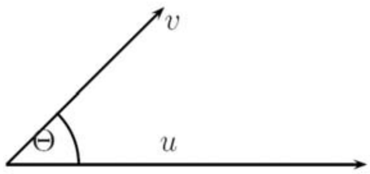
eine Metrik erklären. Ist z.B. $f(x) = 5x, g(x) = 2x - 1$, so haben f und g in dieser Metrik den Abstand:

$$d(f,g) := \left(\int_0^1 (3x-1)^2 \,\mathrm{d}x \right)^{\frac{1}{2}} = \sqrt{\int_0^1 (9x^2 + 6x + 1) \,\mathrm{d}x} = \sqrt{[3x^3 + 3x^2 + x]_0^1} = \sqrt{7}.$$

Kapitel 43

Orthogonalität

43.1 Motivation



Das euklidische Produkt zweier Vektoren $u, v \in \mathbb{R}^2$, die einen Winkel Θ einschließen, lautet:

$$u \cdot v = |u| \cdot |v| \cdot \cos \Theta.$$

Ist $\Theta = \frac{\pi}{2}$, so ist $u \cdot v = 0$ und u und v sind orthogonale Vektoren.
Wir wollen nun diese Begriffe des Winkels und der Orthogonalität in allgemeinen Prä-Hilbert-Räumen formulieren. Dies führt zu Darstellungen in Orthogonalbasen, die wichtige Anwendungen in der Informatik haben, z.B. in der geometrischen Datenverarbeitung, der Bildverarbeitung und im Information Retrieval.

43.2 Definition (Orthogonalität, Winkel)

Sei $(V, \langle\cdot,\cdot\rangle)$ ein Prä-Hilbert-Raum über $\mathbb{R}$ mit induzierter Norm $\|\cdot\|$. Für nicht verschwindende Vektoren $u, v \in V$ gibt es eine eindeutig bestimmte Zahl $\Theta \in [0, \pi)$ mit

$$\cos \Theta = \frac{\langle u, v\rangle}{\|u\| \cdot \|v\|},$$

die wir als Winkel zwischen u und v definieren. Wir nennen u und v orthogonal, falls $< u, v >= 0$ ist (und somit $\Theta = \frac{\pi}{2}$).

43.3 Beispiele

(a) Euklidischer Raum $\mathbb{R}^4$, $u = \begin{pmatrix} 4 \\ 3 \\ 1 \\ -2 \end{pmatrix}, v = \begin{pmatrix} -2 \\ 1 \\ 2 \\ 3 \end{pmatrix}$

$|u| = (16 + 9 + 1 + 4)^{\frac{1}{2}} = \sqrt{30}$
$|v| = (4 + 1 + 4 + 9)^{\frac{1}{2}} = \sqrt{18}$

$u \cdot v = -8 + 3 + 2 - 6 = -9$
$\cos\Theta = \dfrac{u \cdot v}{|u| \cdot |v|} = \dfrac{-9}{\sqrt{30} \cdot \sqrt{18}} \approx -0{,}3873$
$\Theta \approx 1{,}968 \quad (\stackrel{\triangle}{=} 112{,}8°)$

(b) $C[-1, 1]$ mit Skalarprodukt

$$\langle u, v \rangle \quad := \quad \int_{-1}^{1} u(x) \cdot v(x) \, \mathrm{d}x.$$

Mit $u(x) := x, v(x) := x^2$ ergibt sich:
$\langle u, v \rangle = \displaystyle\int_{-1}^{1} x^3 \, \mathrm{d}x = \left[\frac{x^4}{4}\right]_{-1}^{1} = \frac{1}{4} - \frac{1}{4} = 0.$
Die Funktionen $u(x) = x$ und $v(x) = x^2$ sind also orthogonal in $C[-1, 1]$.

Satz 41.13 verallgemeinern wir zu

43.4 Satz: (Satz des Pythagoras in Prä-Hilbert-Räumen)

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Prä-Hilbert-Raum über $\mathbb{R}$ und $\| \cdot \|$ die induzierte Norm. Dann gilt für orthogonale(!) Vektoren $u, v \in V$:

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Beweis:

$\|u + v\|^2 = \langle u + v, u + v \rangle = \|u\|^2 + 2 \underbrace{\langle u, v \rangle}_{0} + \|v\|^2.$ □.

43.5 Beispiel

$C[-1,1]$ mit Skalarprodukt $\langle u,v\rangle = \int\limits_{-1}^{1} u(x)v(x)\,\mathrm{d}x$. Nach 43.3(b) sind $u(x)=x, v(x)=x^2$ orthogonal.

$$\begin{aligned}
\|u+v\|^2 &= \int\limits_{-1}^{1}(x+x^2)^2\,\mathrm{d}x \\
&= \int\limits_{-1}^{1}(x^2+2x^3+x^4)\,\mathrm{d}x \\
&= \left[\frac{x^3}{3}+2\frac{x^4}{4}+\frac{x^5}{5}\right]_{-1}^{1} \\
&= \frac{1}{3}+\frac{1}{2}+\frac{1}{5}-(-\frac{1}{3}+\frac{1}{2}-\frac{1}{5}) \\
&= \frac{2}{3}+\frac{2}{5}=\frac{10+6}{15}=\frac{16}{15}. \\
\|u\|^2 &= \int\limits_{-1}^{1}x^2\,\mathrm{d}x = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{1}{3}-\left(-\frac{1}{3}\right)=\frac{2}{3} \\
\|v\|^2 &= \int\limits_{-1}^{1}x^4\,\mathrm{d}x = \left[\frac{x^5}{5}\right]_{-1}^{1} = \frac{1}{5}-\left(-\frac{1}{5}\right)=\frac{2}{5}
\end{aligned}$$

Wie erwartet gilt also $\|u\|^2+\|v\|^2=\frac{2}{3}+\frac{2}{5}=\frac{16}{15}=\|u+v\|^2$.

In Prä-Hilbert-Räumen ist es oft sinnvoll, Basen zu wählen, deren Elemente paarweise orthogonal sind.

43.6 Definition (Orthogonalmenge, Orthonormalmenge)

Eine Menge von Vektoren in einem Prä-Hilbert-Raum heißt orthogonale Menge, wenn ihre Elemente paarweise orthogonal sind. Haben sie außerdem die (induzierte) Norm 1, so heißt die Menge orthonormal. Bildet eine Basis eines Prä-Hilbert-Raums eine orthogonale (orthonormale) Menge, so spricht man von einer Orthogonalbasis (Orthonormalbasis).

43.7 Beispiel

$u_1=\begin{pmatrix}0\\1\\0\end{pmatrix}, u_2=\begin{pmatrix}1\\0\\1\end{pmatrix}, u_3=\begin{pmatrix}1\\0\\-1\end{pmatrix}$ bilden eine orthogonale Menge im euklidischen Raum $\mathbb{R}^3$, denn es gilt: $0=u_1\cdot u_2=u_1\cdot u_3=u_2\cdot u_3$.
Zwar ist $|u_1|=1$, aber wegen $|u_2|=\sqrt{2}=|u_3|$ ist $\{u_1,u_2,u_3\}$ keine orthonormale Menge. Um eine orthonormale Menge $\{v_1,v_2,v_3\}$ zu erhalten, muß man durch die euklidischen Normen dividieren:

$$v_1=\frac{u_1}{\|u_1\|}=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad v_2=\frac{u_2}{\|u_2\|}=\begin{pmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{pmatrix},\quad v_3=\frac{u_3}{\|u_3\|}=\begin{pmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{pmatrix}.$$

43.8 Satz: (Koordinatendarstellung in Orthonormalbasis)

Sei $S = \{v_1, \dots, v_n\}$ eine Orthonormalbasis eines endlich dimensionalen Prä-Hilbert-Raumes $(V, \langle \cdot, \cdot \rangle)$. Dann gilt für jeden Vektor $u \in V$:

$$\boxed{u = \langle u, v_1 \rangle \, v_1 + \ldots + \langle u, v_n \rangle \, v_n}$$

Beweis:

Da S Basis ist, existieren $\alpha_1, \dots, \alpha_n$ mit $u = \sum_{i=1}^{n} \alpha_i v_i$.

$\Rightarrow \langle u, v_k \rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_k \right\rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_k \rangle = \alpha_k \quad \forall\, k \in \{1, \dots, \}$.

Wegen $\langle v_i, v_k \rangle = \begin{cases} 1 & \text{falls } i = k \\ 0 & \text{sonst} \end{cases}$ gilt: $\langle u, v_k \rangle = \alpha_k$ für $k = 1, \dots, n$. □.

43.9 Beispiel

$v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} -\frac{4}{5} \\ 0 \\ \frac{3}{5} \end{pmatrix}, v_3 = \begin{pmatrix} \frac{3}{5} \\ 0 \\ \frac{4}{5} \end{pmatrix}$ bilden eine Orthonormalbasis des euklidischen Raums $\mathbb{R}^3$.

Man schreibe $u = \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix}$ als Linearkombination von v_1, v_2, v_3.

$$\begin{aligned}
u \cdot v_1 &= 2 \\
u \cdot v_2 &= -\frac{4}{5} + 7 \cdot \frac{3}{5} = \frac{17}{5} \\
u \cdot v_3 &= \frac{3}{5} + 7 \cdot \frac{4}{5} = \frac{31}{5} \\
\Rightarrow \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix} &= 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{17}{5} \begin{pmatrix} -\frac{4}{5} \\ 0 \\ \frac{3}{5} \end{pmatrix} + \frac{31}{5} \begin{pmatrix} \frac{3}{5} \\ 0 \\ \frac{4}{5} \end{pmatrix}.
\end{aligned}$$

43.10 Satz: (Koordinatendarstellung in Orthogonalbasis)

Sei $S = \{v_1, \dots, v_n\}$ eine Orthogonalbasis eines endlich-dimensionalen Prä-Hilbert-Raums $(V, \langle \cdot, \cdot \rangle)$ mit induzierter Norm $\| \cdot \|$. Dann gilt für jedes $u \in V$:

$$u = \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \ldots + \frac{\langle u, v_n \rangle}{\|v_n\|^2} v_n.$$

Beweis:

Aus der Orthogonalbasis S erhält man durch Normierung die Orthonormalbasis $S' = \left\{ \frac{v_1}{\|v_1\|}, \dots, \frac{v_n}{\|v_n\|} \right\}$.
Nach Satz 43.8 gilt:

$$\begin{aligned} u &= \left\langle u, \frac{v_1}{\|v_1\|} \right\rangle \frac{v_1}{\|v_1\|} + \ldots + \left\langle u, \frac{v_n}{\|v_n\|} \right\rangle \frac{v_n}{\|v_n\|} \\ &= \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \ldots + \frac{\langle u, v_n \rangle}{\|v_n\|^2} v_n. \end{aligned}$$

$\square$.

Bemerkung: Unter Zusatzvoraussetzungen gelten ähnliche Aussagen auch in unendlich dimensionalen Prä-Hilbert-Räumen.

43.11 Satz: (Lineare Unabhängigkeit orthogonaler Mengen)

Eine orthogonale Menge $S = \{v_1, \dots, v_n\}$ aus von 0 verschiedenen Elementen ist linear unabhängig.

Beweis:

Wir zeigen, dass aus $\sum_{i=1}^{n} \alpha_i v_i = 0$ stets $\alpha_1 = \ldots = \alpha_n = 0$ folgt.

Sei also $\sum_{i=1}^{n} \alpha_i v_i = 0$. Dann gilt für jedes $v_k,\ k = 1, \dots, n$:

$$\begin{aligned} 0 &= \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_k \right\rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_k \rangle \\ &= \alpha_k \underbrace{\|v_k\|^2}_{\neq 0}, \text{ da } \langle v_i, v_k \rangle = 0 \text{ für } i \neq k \\ &\Rightarrow \alpha_k = 0. \end{aligned}$$

$\square$.

43.12 Beispiel

Aus Beispiel 43.7 wissen wir, dass $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix} \right\}$ eine orthonormale Menge im euklidischen Raum $\mathbb{R}^3$ ist. Nach Satz 43.11 sind dies 3 linear unabhängige Vektoren. Sie bilden somit eine Orthonormalbasis des $\mathbb{R}^3$.

Nach Satz 43.11 wissen wir, dass orthogonale Mengen linear unabhängig sind. Kann man umgekehrt aus einer linear unabhängigen Menge eine orthogonale Menge konstruieren ?

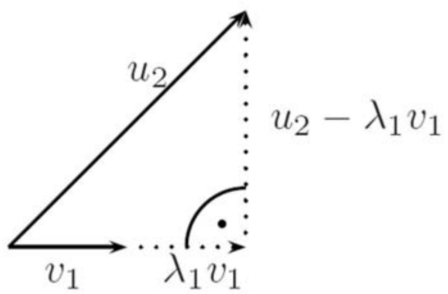
43.13 Orthogonalisierungsalgorithmus von Gram und Schmidt

Geg.: $(V, <\cdot,\cdot>)$ Prä-Hilbert-Raum mit endlich dimensionalem Unterraum W.
W habe die Basis $\{u_1, \dots, u_n\}$.

Ges.: Orthogonalbasis $\{v_1, \dots, v_n\}$ von W.

Schritt 1: $\boxed{v_1 := u_1}$

Schritt 2: $\lambda_1 v_1$ ist das Lot von u_2 auf $\operatorname{span}\{v_1\}$.



Ansatz: $v_2 := u_2 - \lambda_1 v_1$ mit Forderung $\langle v_2, v_1\rangle \overset{!}{=} 0$.
Dies erlaubt die Bestimmung von λ_1.

$$\begin{aligned} 0 = \langle v_2, v_1\rangle &= \langle u_2 - \lambda_1 v_1, v_1\rangle \\ &= \langle u_2, v_1\rangle - \lambda_1 \langle v_1, v_1\rangle \\ &\Rightarrow \lambda_1 = \frac{\langle u_2, v_1\rangle}{\|v_1\|^2} \\ &\Rightarrow \boxed{v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\|v_1\|^2} v_1} \end{aligned}$$

Schritt n: Seien $\{v_1, \dots, v_{n-1}\}$ orthogonal für $n \geq 2$.
Ansatz: $v_n := u_n - \sum_{i=1}^{n-1} \lambda_i v_i$ mit Forderungen
$\langle v_n, v_j\rangle = 0$ für $j = 1, \dots, n-1$ führt auf:

$$\begin{aligned} 0 = \langle v_n, v_j\rangle &= \left\langle u_n - \sum_{i=1}^{n-1} \lambda_i v_i, v_j \right\rangle \\ &= \langle u_n, v_j\rangle - \lambda_j \langle v_j, v_j\rangle \\ &\Rightarrow \lambda_j = \frac{\langle u_n, v_j\rangle}{\|v_j\|^2} \\ &\Rightarrow \boxed{v_n = u_n - \sum_{i=1}^{n-1} \frac{\langle u_n, v_i\rangle}{\|v_i\|^2} v_i} \end{aligned}$$

Das Gram-Schmidt-Orthogonalisierungsverfahren liefert somit einen konstruktiven Beweis von

43.14 Satz: (Existenz einer Orthogonalbasis)

Jeder endlich dimensionale Prä-Hilbert-Raum besitzt eine Orthogonalbasis.

Bemerkung: Ist der Raum vom Nullvektorraum verschieden, so erhält man eine Orthonormalbasis durch Normierung.

43.15 Beispiel

Konstruiere aus $u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ mit dem Gram-Schmidt-Algorithmus eine Orthogonalbasis des euklidischen Raumes $\mathbb{R}^3$. Konstruiere anschließend eine Orthonormalbasis.
Lösung:

$$\begin{aligned}
v_1 &= u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
v_2 &= u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix} \\
v_3 &= u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2 \\
&= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1/3}{2/3} \begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} - \begin{pmatrix} -1/3 \\ 1/6 \\ 1/6 \end{pmatrix} = \begin{pmatrix} 0 \\ -1/2 \\ 1/2 \end{pmatrix}
\end{aligned}$$

$\{v_1, v_2, v_3\}$ ist die gesuchte Orthogonalbasis im $\mathbb{R}^3$.

Die entsprechende Orthonormalbasis $\{q_1, q_2, q_3\}$ lautet

$$\begin{aligned}
q_1 &= \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
q_2 &= \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{6}/3} \begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix} = 1/\sqrt{6} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \\
q_3 &= \frac{v_3}{\|v_3\|} = \frac{1}{1/\sqrt{2}} \begin{pmatrix} 0 \\ -1/2 \\ 1/2 \end{pmatrix} = 1/\sqrt{2} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}
\end{aligned}$$

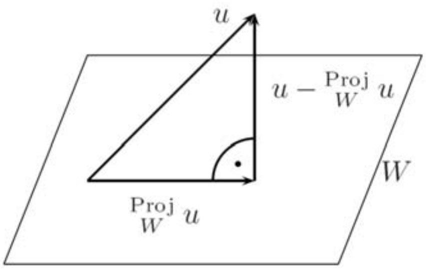
Ein weiteres „Abfallprodukt" des Gram-Schmidt-Verfahrens ist der

43.16 Satz: (Orthogonale Projektion auf Unterräume)

Sei $(V, \langle\cdot,\cdot\rangle)$ ein Prä-Hilbert-Raum, $u \in V$ und sei W ein endlich dimensionaler Unterraum mit Orthogonalbasis $\{v_1, \ldots v_n\}$. Dann beschreibt

$$\operatorname*{Proj}_W u := \sum_{i=1}^{n} \frac{\langle u, v_i \rangle}{\|v_i\|^2} v_i$$

eine orthogonale Projektion von u auf W; d.h.
$\operatorname*{Proj}_W u \in W$ *und* $\left\langle u - \operatorname*{Proj}_W u, w \right\rangle = 0 \quad \forall w \in W.$



Bemerkung: Ist $\{v_1, \ldots, v_n\}$ eine Ortho<u>normal</u>basis, gilt also $\underset{W}{\mathrm{Proj}}\, u = \sum_{i=1}^{n} \langle u, v_i \rangle\, v_i$.

Ist die Orthogonalprojektion eindeutig?

43.17 Definition (Orthogonales Komplement)

Sei W ein Unterraum eines Prä-Hilbert-Raums $(V, \langle \cdot, \cdot \rangle)$. Dann bezeichnet

$$W^\perp := \{v \in V \,|\, \langle v, w \rangle = 0 \quad \forall w \in W\}$$

das <u>orthogonale Komplement</u> von W.

43.18 Satz: (<u>Projektionssatz</u>)

Sei W ein endlich dimensionaler Unterraum eines Prä-Hilbert-Raums $(V, \langle \cdot, \cdot \rangle)$. Dann besitzt jedes $v \in V$ eine <u>eindeutige</u> Darstellung $v = w_1 + w_2$ mit $w_1 \in W$ und $w_2 \in W^\perp$.

(Man schreibt auch $V = W \oplus W^\perp$ und nennt $\oplus$ <u>direkte Summe</u>).

<u>Beweis:</u>

Es ist nur die Eindeutigkeit zu zeigen, die Existenz gilt wegen 43.16.
Seien also $w_1, w_1' \in W$ und $w_2, w_2' \in W^\perp$ mit

$$\begin{aligned} v &= w_1 + w_2 = w_1' + w_2' \\ &\Rightarrow w_1 - w_1' = w_2' - w_2 \end{aligned}$$

$w_1 - w_1' \in W$, da Unterräume abgeschlossen sind.
Andererseits gilt: $\langle w_2' - w_2, w \rangle = \underbrace{\langle w_2', w \rangle}_{=0 \text{ da } w_2' \in W^\perp} - \underbrace{\langle w_2, w \rangle}_{=0 \text{ da } w_2 \in W} = 0 \quad \forall w \in W.$

$\Rightarrow w_2' - w_2 \in W^\perp$.
Wegen $w_2' - w_2 = w_1 - w_1'$ gilt aber auch $w_2' - w_2 \in W$.

$\Rightarrow 0 = \left\langle \underbrace{w_2' - w_2}_{\in W}, \underbrace{w_2' - w_2}_{\in W^\perp} \right\rangle = \|w_2' - w_2\|^2$

$\Rightarrow 0 = w_2' - w_2 = w_1 - w_1'$
und daher $w_1 = w_1'$, $w_2 - w_2'$. $\square$.Gibt es weitere Anwendungen der Orthogonalprojektion?

43.19 Satz: (Approximationssatz)

Sei $(V, \langle\cdot,\cdot\rangle)$ ein Prä-Hilbert-Raum mit induzierter Norm $\|\cdot\|$ und W ein endlicher dimensionaler Unterraum. Zu $v \in V$ ist dann die $\underset{W}{\mathrm{Proj}}\, v$ die beste Approximation von v in W, d.h.

$$\|v - \underset{W}{\mathrm{Proj}}\, v\| < \|v - w\| \quad \forall w \in W \ \mathit{mit}\ W \neq \underset{W}{\mathrm{Proj}}\, v.$$

Beweis:

$$\begin{aligned}
\|v-w\|^2 &= \|\underbrace{v - \underset{W}{\mathrm{Proj}}\, v}_{\in W^\perp} + \underbrace{\underset{W}{\mathrm{Proj}}\, v - w}_{\in W}\|^2 \\
&\overset{\text{Pythagoras}}{=} \|v - \underset{W}{\mathrm{Proj}}\, v\|^2 + \|\underset{W}{\mathrm{Proj}}\, v - w\|^2 \\
&\geq \|v - \underset{W}{\mathrm{Proj}}\, v\|^2
\end{aligned}$$

Gleichheit geht nur, falls $w = \underset{W}{\mathrm{Proj}}\, v$. $\square$.

43.20 Beispiel

$V = C[0, \frac{\pi}{2}]$ mit Skalarprodukt $< u, v >:= \int\limits_0^{\pi/2} u(x)v(x)\,\mathrm{d}x$.

Bestimme Gerade, die die Funktion $u(x) = \sin x$ im Intervall $[0, \frac{\pi}{2}]$ optimal (bzgl. der induzierten Norm) approximiert.

Lösung: $W = \mathrm{span}\{\underbrace{1}_{v_1}, \underbrace{x}_{v_2}\}$ Unterraum aller Geraden.

Gesucht: $\underset{W}{\mathrm{Proj}}\, u = \sum\limits_{i=1}^{2} \lambda_i v_i$ mit $0 = \left\langle u - \sum\limits_{i=1}^{2} \lambda_i v_i, v_k \right\rangle \quad (k = 1, 2)$.

In unserem Fall:

$$\begin{aligned}
0 &= \langle \sin x - \lambda_1 - \lambda_2 x, 1\rangle \\
0 &= \langle \sin x - \lambda_1 - \lambda_2 x, x\rangle
\end{aligned}$$

gilt Gleichungssystem für Unbekannte λ_1, λ_2:

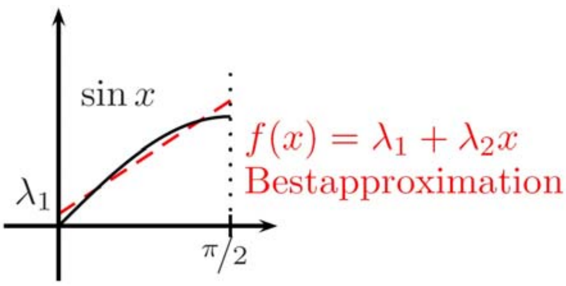
$$\begin{aligned}
\lambda_1 \langle 1, 1\rangle + \lambda_2 \langle x, 1\rangle &= \langle \sin x, 1\rangle \\
\lambda_1 \langle 1, x\rangle + \lambda_2 \langle x, x\rangle &= \langle \sin x, x\rangle
\end{aligned}$$

Mit

$$
\begin{aligned}
\langle 1,1\rangle &= \int_0^{\pi/2} \mathrm{d}x = \frac{\pi}{2} \\
\langle x,1\rangle &= <1,x>= \int_0^{\pi/2} x\ \mathrm{d}x = \frac{\pi^2}{8} \\
\langle \sin x,1\rangle &= \int_0^{\pi/2} \sin x\ \mathrm{d}x = [-\cos x]_0^{\pi/2} = 0-(-1) = 1 \\
\langle x,x\rangle &= \int_0^{\pi/2} x^2\ \mathrm{d}x = \left[\frac{x^3}{3}\right]_0^{\pi/2} = \frac{\pi^3}{24} \\
\langle x,\sin x\rangle &= \int_0^{\pi/2} x\sin x\ \mathrm{d}x = \Big[-x\cos x\Big]_0^{\pi/2} + \int_0^{\pi/2} \cos x\ \mathrm{d}x = -\frac{\pi}{2}\cdot 0 + 0\cdot 1 + \Big[\sin x\Big]_0^{\pi/2} = 1
\end{aligned}
$$

lautet das System $\begin{pmatrix} \pi/2 & \pi^2/8 \\ \pi^2/8 & \pi^3/24 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

$$
\begin{aligned}
\text{Es hat die Lösung } \lambda_1 &= 8\left(\frac{\pi-3}{\pi^2}\right) \approx 0,11 \\
\lambda_2 &= 24\left(\frac{4-\pi}{\pi^3}\right) \approx 0,66.
\end{aligned}
$$



Kapitel 44

Fourierreihen

44.1 Motivation

Ähnlich wie die Taylorreihen eine Funktion durch ein Polynom approximiert, wollen wir eine Funktion durch ein trigonometrisches Polynom annähern. Hierzu verwenden wir den Approximationssatz 43.19.

44.2 Herleitung der Fourierkoeffizienten

Gegeben:

- Vektorraum $V = C[0, 2\pi]$ mit Skalarprodukt

  $$\langle u, v \rangle := \int\limits_0^{2\pi} u(x)v(x)\ \mathrm{d}x$$

  und Norm

  $$\|u\| \quad := \quad \sqrt{\langle u, u \rangle}$$

- Funktion $u \in V$,
- endlich dimensionaler Unterraum:
  $W = \mathrm{span}\{\underbrace{1}_{=:v_0}, \underbrace{\cos x}_{v_1}, \underbrace{\cos(2x)}_{v_2}, \ldots, \underbrace{\cos(nx)}_{v_n}, \underbrace{\sin(x)}_{v_{n+1}}, \underbrace{\sin(2x)}_{v_{n+2}}, \ldots, \underbrace{\sin(vx)}_{v_{2n}}\}$
  trigonometrische Polynome vom Grad $\leq n$.

Gesucht: Koeffizienten $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n$ so dass das trigonometrische Polynom

$$\boxed{f(x) = \frac{a_0}{2} + \sum_{k=1}^{n} [a_k \cos(kx) + b_k \sin(kx)]}$$

die beste Approximation an $u(x)$ bzgl. der induzierten Norm $\| \cdot \|$ ist.
(„Approximation im quadratischen Mittel“). Diese Koeffizienten heißen Fourierkoeffizienten.

Lösung: Nach Satz 43.19 ist die Approximation im quadratischen Mittel durch die Orthogonalprojektion $f = \underset{W}{\text{Proj}}\, u$ gegeben. Man weist leicht nach, dass $\{1, \cos x, \ldots, \sin(nx)\}$ orthogonal sind:

$$\begin{aligned}
\int_0^{2\pi} \sin(kx)\sin(lx)\,\mathrm{d}x &= \begin{cases} 0 & (k \neq l) \\ \pi & (k = l) \end{cases} \quad (k, l \geq 1) \\
\int_0^{2\pi} \sin(kx)\cos(lx)\,\mathrm{d}x &= 0 \\
\int_0^{2\pi} \cos(kx)\cos(lx)\,\mathrm{d}x &= \begin{cases} 0 & (k \neq l) \\ \pi & (k = l \geq 1) \\ 2\pi & (k = l = 0) \end{cases}
\end{aligned}$$

Nach Satz 43.16 lautet damit die Orthogonalprojektion

$$\begin{aligned}
f &= \underset{W}{\text{Proj}}\, u \\
&= \sum_{k=0}^{2n} \frac{\langle u, v_k \rangle}{\|v_k\|^2} \cdot v_k \\
&= \frac{1}{2\pi}\langle u, 1 \rangle + \sum_{k=1}^{n} \frac{1}{\pi}[\langle u, \cos(kx)\rangle \cos(kx) + \langle u, \sin(kx)\rangle \sin(kx)]
\end{aligned}$$

Somit lauten die Fourierkoeffizienten

$$\boxed{\begin{aligned}
&a_0 = \frac{1}{\pi}\langle u, 1 \rangle = \frac{1}{\pi}\int_0^{2\pi} u(x)\,\mathrm{d}x \\
&\left.\begin{aligned}
a_k &= \frac{1}{\pi}\langle u, \cos(kx)\rangle = \frac{1}{\pi}\int_0^{2\pi} u(x)\cos(kx)\,\mathrm{d}x \\
b_k &= \frac{1}{\pi}\langle u, \sin(kx)\rangle = \frac{1}{\pi}\int_0^{2\pi} u(x)\sin(kx)\,\mathrm{d}x
\end{aligned}\right\} k = 1, \ldots, n
\end{aligned}}$$

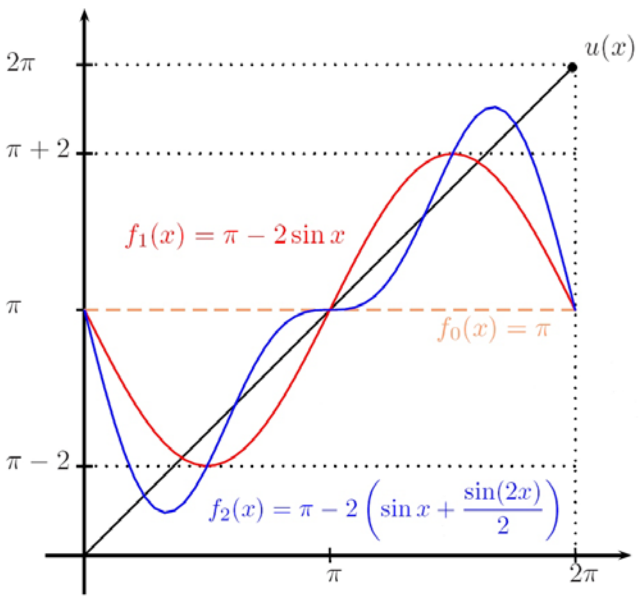
44.3 Beispiel

Die Funktion $u(x) = x$ soll auf $[0, 2\pi]$ im quadratischen Mittel mit einem trigonometrischen Polynom vom Grad $\leq n$ approximiert werden.

$$
\begin{aligned}
a_0 &= \frac{1}{\pi}\int_0^{2\pi} x \, dx = \frac{1}{\pi}\left[\frac{x^2}{2}\right]_0^{2\pi} = 2\pi \\
a_k &= \frac{1}{\pi}\int_0^{2\pi} x\cos(kx)\, dx = \frac{1}{\pi}\left[\underbrace{\left[x\frac{1}{k}\sin(kx)\right]_0^{2\pi}}_{0} - \int_0^{2\pi}\frac{1}{k}\sin(kx)\, dx\right] \\
&= \left[\frac{1}{\pi}\frac{1}{k^2}\cos(kx)\right]_0^{2\pi} = 0 \quad (k \geq 1) \\
b_k &= \frac{1}{\pi}\int_0^{2\pi} x\sin(kx)\, dx = \frac{1}{\pi}\left[\left[-x\frac{1}{k}\cos(kx)\right]_0^{2\pi} + \int_0^{2\pi}\frac{1}{k}\cos(kx)\, dx\right] \\
&= \frac{-2\pi}{k}\cdot\frac{1}{\pi} + \frac{1}{\pi}\underbrace{\left[\frac{1}{k^2}\sin(kx)\right]_0^{2\pi}}_{0} = -\frac{2}{k} \quad (k \geq 1)
\end{aligned}
$$

Das trigonometrische Approximationspolynom lautet also

$$
f_n(x) = \pi - 2\left(\sin x + \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} + \ldots + \frac{\sin(nx)}{n}\right)
$$



44.4 Definition (Fourierreihe)

Läßt man den Grad des trigonometrischen Approximationspolynoms gegen ∞ gehen, entsteht die Fourierreihe

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos(kx) + b_k \sin(kx)) \\ \text{mit} \quad a_k &= \frac{1}{\pi}\int_0^{2\pi} u(x)\cos(kx)\,\mathrm{d}x \qquad k = 0,1,\ldots \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} u(x)\sin(kx)\,\mathrm{d}x \qquad k = 1,2,\ldots \end{aligned}$$

Bem.: Falls u differenzierbar ist, kann man zeigen, dass die Fourierreihe punktweise gegen u konvergiert. An Sprungstellen zeigen die Fourierpolynome ein ausgeprägtes Über- und Unterschwingen (Gibbs-Phänomen).

44.5 Praktische Bedeutung

- Fourierreihen sind unentbehrlich in der Signalverarbeitung
- Die Fourierkoeffizienten eines Signals geben die einzelnen Frequenzanteile an:
  a_k, b_k mit kleinem k: niedrigen Frequenzen
  a_k, b_k mit hohem k: hohen Frequenzen
- Filterentwurf durch Spezifikation im Frequenzbereich:
  (a) Tiefpassfilter:
    – dämpfen hohe Frequenzen
    – zur Elimination von Rauschen (i.A. hochfrequent)
  (b) Hochpassfilter:
    – dämpfen tiefe Frequenzen (z.B. Brumm- und Rumpelgeräusche)
  (c) Bandpassfilter:
    – lassen nur vorgegebenen Frequenzbereich passieren
      (z.B. mittlere Frequenzen bei Stimmübertragung)
- ähnliche Bedeutung in der Bildverarbeitung:
  Grauwertbilder können als 2D-Signale aufgefasst werden.
- Signale und Bilder liegen meist diskret (gesampelt) vor.
  Dann verwendet man eine diskrete Fouriertransformation, die Integrale durch Summen ersetzt.
- Es existieren sehr schnelle Algorithen zur diskreten Fouriertransformation, die ein Signal mit N Werten mit einer Komplexität von $O(N \log N)$ in seine Frequenzanteile zerlegen. (FFT: Fast Fourier Transform)

44.6 Aktuelle Weiterentwicklung: Wavelets

- verwenden Basisfunktionen, die nicht nur in der Frequenz, sondern auch im Ort lokalisiert sind.
- effizientestesten Verfahren zur Signal- und Bildkompression (in zukünftigen jpeg- und mpeg-Standards):
  Viele der Waveletkoeffizienten sind sehr klein und können weggelassen werden, ohne dass es auffällt.
- hocheffiziente Algorithmen mit $O(N)$-Komplexität existieren.

Kapitel 45

Orthogonale Matrizen

45.1 Motivation

Im euklidischen Raum $\mathbb{R}^n$ haben wir gesehen, dass Orthonormalbasen zu besonders einfachen und schönen Beschreibungen führen.
Wir wollen nun das Konzept der Orthonormalität nicht mehr nur auf Vektoren beschränken, sondern auf Matrizen erweitern. Dies führt auf die wichtige Klasse der orthogonalen Matrizen, die eine Reihe von schönen Eigenschaften aufweisen. Mit ihnen lassen sich u.a. Drehungen und Spiegelungen beschreiben.

45.2 Definition (Orthogonale Matrix)

Hat eine Matrix $Q \in \mathbb{R}^{n\times n}$ orthonormale Spaltenvektoren $q_{*1}, \dots, q_{*n}$, so handelt es sich um eine orthogonale Matrix (Orthonormale Matrix wäre präziser, ist aber unüblich). Man definiert ferner $\mathrm{O}(N) := \{Q \in \mathbb{R}^{n\times n} \mid Q \text{ orthogonal }\}$.

Was sind nun die schönen Eigenschaften?

45.3 Satz: (Eigenschaften orthogonaler Matrizen)

Ist $Q \in \mathrm{O}(N)$, so gilt

(a) *Q ist invertierbar, und Q^{-1} hat eine sehr einfache Form:*

$$Q^{-1} = Q^\top.$$

(b) *Multiplikation mit Q erhält das euklidische Produkt zweier Vektoren:*

$$(Qu) \cdot (Qv) = u \cdot v \qquad \forall u, v \in \mathbb{R}^n.$$

(c) *Multiplikation mit Q erhält die euklidische Norm*

$$|Qv| = |v| \qquad \forall v \in \mathbb{R}^n.$$

Man nennt Q daher auch Isometrie.

Beweis:

(a) Sei $A = (a_{ij}) = Q^\top Q$. Dann gilt:

$$a_{ij} = \sum_{k=1}^{n} q_{ki} \cdot q_{kj} = q_{*i} \cdot q_{*j} = \begin{cases} 1 & (i = j) \\ 0 & (\text{sonst}) \end{cases}.$$

Also ist $Q^\top Q = I$. Ähnlich zeigt man $QQ^\top = I$. Somit ist Q invertierbar und $Q^{-1} = Q^\top$.

(b) $(Qu) \cdot (Qv) = (Qu)^\top (Qv) = u^\top \underbrace{Q^\top Q}_{I} v = u^\top v = u \cdot v$

(c) Folgt aus (b) mit $u = v$. □.

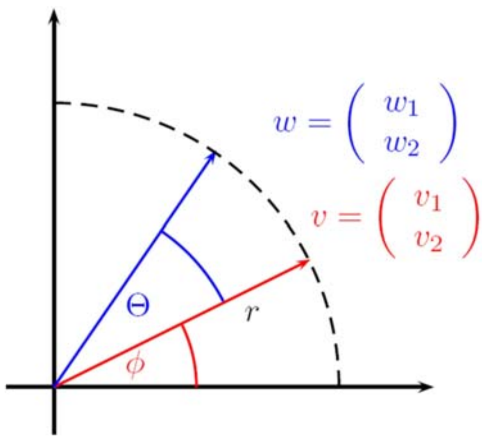
Bemerkung: Es gilt sogar: Q orthogonal $\Leftrightarrow Q^\top = Q^{-1}$.

45.4 Beispiele

(a) Rotationen können durch orthogonale Matrizen beschrieben werden.

$$Q = \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}$$

beschreibt Drehung um Winkel Θ, denn:



$$\begin{aligned}
v_1 &= r\cos\phi \\
v_2 &= r\sin\phi \\
w_1 &= r\cos(\phi + \Theta) \\
&= r(\cos\phi\cos\Theta - \sin\phi\sin\Theta) \\
&= v_1\cos\Theta - v_2\sin\Theta \\
w_2 &= r\sin(\phi + \Theta) \\
&= r(\sin\phi\cos\Theta + \cos\phi\sin\Theta) \\
&= v_2\cos\Theta + v_1\sin\Theta \\
\Rightarrow \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} &= \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.
\end{aligned}$$

Die inverse Matrix ist die Drehung um $-\Theta$:

$$\begin{aligned} Q^{-1} &= \begin{pmatrix} \cos(-\Theta) & -\sin(-\Theta) \\ \sin(-\Theta) & \cos(-\Theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{pmatrix} \\ &= Q^\top. \end{aligned}$$

Somit ist Q orthogonal.

Beachte: $\det Q = \cos^2\Theta + \sin^2\Theta = 1$.

(b) Es gibt auch orthogonale Matrizen, die keine Drehungen beschreiben, z.B.

$$Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Q beschreibt eine Spiegelung an der ersten Winkelhalbierenden, denn Q vertauscht x- und y-Komponente:

$$Q\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_2 \\ v_1 \end{pmatrix}.$$

Beachte: $\det Q = 0 - 1 = -1$.

Kann die Determinante von orthogonalen Matrizen auch andere Werte als ± 1 annehmen? Nein!

45.5 Satz: (Determinante orthogonaler Matrizen)

Ist $Q \in \mathrm{O}(n)$, so gilt $|\det Q| = 1$.

Beweis:

$1 = \det(I) = \det(QQ^\top) = \det Q \cdot \det Q^\top = (\det Q)^2.$ □.

Orthogonale Matrizen mit Determinante 1 sind noch einmal gesondert ausgezeichnet:

45.6 Definition:

$\mathrm{SO}(n) := \mathrm{O}^+(n) := \{Q \in \mathrm{O}(n) \mid \det Q = 1\}.$

45.7 Satz: (Gruppeneigenschaft von $\mathrm{O}(n)$ und $\mathrm{SO}(n)$)

$\mathrm{O}(n)$ und $\mathrm{SO}(n)$ sind Untergruppen der allgemeinen linearen Gruppe

$$\mathrm{GL}(n,\mathbb{R}) = \{A \in \mathbb{R}^{n\times n} \mid A \textit{ invertierbar}\}$$

bezüglich der Matrixmultiplikation. Man nennt $\mathrm{O}(n)$ die orthogonale Gruppe und $\mathrm{SO}(n)$ die spezielle orthogonale Gruppe.

Beweis:

Übungsaufgabe.

Wo treten Matrizen noch auf? Beim Wechsel von einer Orthonormalbasis in eine andere.

45.8 Wechsel zwischen Orthonormalbasen

Problem: Sei $\{v_1, \dots, v_n\}$ Orthonormalbasis (ONB) des euklidischen Raums $\mathbb{R}^n$. Dann existiert zu jedem Vektor $u \in \mathbb{R}^n$ eindeutig bestimmte Koeffizienten $a_1, \dots, a_n$ mit $u = \sum_{k=1}^{n} a_k v_k$.

$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ ist also der Koordinatenvektor von u bzgl. der ONB $\{v_1, \dots, v_n\}$.

Sei nun $\{w_1, \dots, w_n\}$ eine weitere ONB des $\mathbb{R}^n$ und u habe den Koordinatenvektor $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ bzgl. $\{w_1, \dots, w_n\}$.

Gibt es eine Übergangsmatrix Q mit $b = Qa$?

Lösung:

$$\begin{aligned} u &= \sum_{k=1}^{n} a_k v_k \\ &= \sum_{k=1}^{n} a_k \left(\sum_{i=1}^{n} (v_k^\top w_i) w_i \right) && v_k \text{ durch } \{w_1, \dots, w_n\} \text{ ausgedrückt} \\ &= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_k (v_k^\top w_i) \right) w_i && \text{Vertauschung der endlichen Summation} \\ &=: \sum_{i=1}^{n} b_j w_i \\ \text{mit } b_i &= \sum_{k=1}^{n} \underbrace{(v_k^\top w_i)}_{=: q_{ik}} a_k \end{aligned}$$

Für die gesuchte Übergangsmatrix $Q = (q_{ik})$ gilt also

$$q_{ik} = v_k^\top w_i = w_i^\top v_k$$

$$\Rightarrow \boxed{Q = \begin{pmatrix} w_1^\top \\ \hline \vdots \\ \hline w_n^\top \end{pmatrix} (v_1, \dots, v_k)}$$

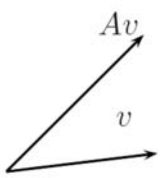
Q ist das Produkt zweier orthogonaler Matrizen (warum ist $\begin{pmatrix} w_1^\top \\ \hline \vdots \\ \hline w_n^\top \end{pmatrix}$ orthogonal?) und damit nach 45.7 selbst wieder orthogonal.

Kapitel 46

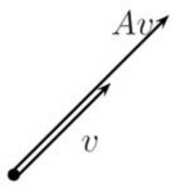
Eigenwerte und Eigenvektoren

46.1 Motivation

Sei $v \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n\times n}$. Dann sind v und Av normalerweise nicht parallel.



Gibt es ausgezeichnete Richtungen v, so dass Av ein skalares Vielfaches von v ist?



$$Av = \lambda v.$$

46.2 Definition (Eigenvektor, Eigenwert)

Sei $A \in \mathbb{R}^{n\times n}$. Ein von 0 verschiedener(!) Vektor $v \in \mathbb{R}^n$ heißt Eigenvektor von A, wenn es ein $\lambda \in \mathbb{R}$ gibt mit

$$\boxed{Av = \lambda v}$$

Der Skalar λ heißt dann Eigenwert von A.

46.3 Bedeutung von Eigenvektoren und Eigenwerten

Eigenvektor- und Eigenwertprobleme sind wichtig in der Statik, Maschinenbau, Elektrotechnik, Informatik, Biologie und Wirtschaftswissenschaften. Oft bezeichnen sie besondere Zustände von Systemen.

Beispiel: 1831 haben Soldaten eine Brücke zum Einsturz gebracht, indem sie mit einer Frequenz marschiert sind, die einen Eigenwert des Brückensystems getroffen hat. Es kam zur Resonanzkatastrophe. Seitdem geht man nicht mehr im Gleichschritt über Brücken.

46.4 Beispiel

$v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ist ein Eigenvektor von $A = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$, denn

$$Av = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 3v.$$

Der zugehörige Eigenwert ist 3.

Wie kann man Eigenwerte und Eigenvektoren bestimmen?

46.5 Bestimmung von Eigenwerten

Aus $Av = \lambda v$ folgt $(A - \lambda I)v = 0$.
$v = 0$ ist als Eigenvektor ausgenommen, da stets $A \cdot 0 = 0$ ist. Wir suchen also nichttriviale Lösungen von $(A - \lambda I)v = 0$. Sie existieren nur falls $\operatorname{rang}(A - \lambda I) < n$, d.h. für

$$\boxed{\det(A - \lambda I) = 0}$$

Für $A \in \mathbb{R}^{n \times n}$ ist dies ein Polynom n-ten Grades in λ (charakteristisches Polynom von A).

Seine Nullstellen sind die gesuchten Eigenwerte.

46.6 Beispiel

Für $A = \begin{pmatrix} 2 & 1 \\ 6 & 1 \end{pmatrix}$ erhalten wir:

$$\begin{aligned} 0 &= \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 1 \\ 6 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)(1-\lambda) - 6 \\ &= 2 - 2\lambda - \lambda + \lambda^2 - 6 = \lambda^2 - 3\lambda - 4 \\ \lambda_{1/2} &= \frac{3 \pm \sqrt{9+16}}{2} = \frac{3 \pm 5}{2} \\ \Rightarrow \lambda_1 &= 4 \\ \lambda_2 &= -1. \end{aligned}$$

46.7 Bemerkung

(a) Selbst wenn A nur reelle Einträge hat, kann das charakteristische Polynom komplexe Nullstellen besitzen. Komplexe Eigenwerte sind also nicht ungewöhnlich.

(b) Sucht man Eigenwerte einer $(n \times n)$-Matrix als Nullstellen des charakteristischen Polynoms, kann dies für $n \geq 3$ unangenehm werden. Für $n \geq 5$ ist dies i.A. nicht mehr analytisch möglich. Dann werden numerische Approximationen benötigt (ebenfalls nicht ganz einfach).

(c) Man kann zeigen, dass $\det A$ das Produkt der Eigenwerte ist und dass A genau dann invertierbar ist, wenn der Eigenwert 0 nicht auftritt.

Trotzdem ist das charakteristische Polynom in Spezialfällen sehr nützlich, z.B. bei Dreiecksmatrizen:

46.8 Definition (Dreiecksmatrix, Diagonalmatrix)

Sei $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. A heißt obere Dreiecksmatrix (untere Dreiecksmatrix), falls $a_{ij} = 0$ für $i > j$ ($i < j$) ist. Ist $a_{ij} = 0$ für alle $i \neq j$, so ist A eine Diagonalmatrix.

46.9 Beispiele

(a) $A = \begin{pmatrix} 3 & 1 & -5 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ ist obere Dreiecksmatrix.

(b) $B = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}$ ist Diagonalmatrix und damit auch obere/untere Dreiecksmatrix.

46.10 Satz: (Eigenwerte von Dreiecksmatrizen)

Ist $A \in \mathbb{R}^{n \times n}$ eine obere oder untere Dreiecksmatrix, so sind die Eigenwerte durch die Diagonaleinträge gegeben.

Beweis:

Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonaleinträge. Für $A = (a_{ij})$ folgt aus $0 = \det(A - \lambda I) = (a_{11} - \lambda) \cdot \ldots \cdot (a_{nn} - \lambda)$, dass die Eigenwerte durch $\lambda_1 = a_{11}, \ldots, \lambda_n = a_{nn}$ gegeben sind. $\square$.

46.11 Beispiel

$A = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ hat die Eigenwerte $\lambda_1 = 3$ und $\lambda_2 = -1$ (vgl. Beispiel 46.4).

46.12 Bestimmung der Eigenvektoren

Annahme: Ein Eigenwert λ der Matrix A sei bekannt.
Dann sind die Eigenvektoren zu α die nichttriviale Lösungen von

$$(A - \lambda I)v = 0 \qquad (*)$$

Eigenvektoren sind nicht eindeutig bestimmt: Mit v ist auch αv mit $\forall \alpha \in \mathbb{R} \setminus \{0\}$ Eigenvektor.
Der Lösungsraum von (*) heißt Eigenraum von A zum Eigenwert λ. Man sucht daher nach Basisvektoren im Eigenraum und gibt diese als Eigenvektoren an.

46.13 Beispiel

Bestimme die Basen der Eigenräume von $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$.

Lösung: $0 = \det(A - \lambda I) = \ldots = (\lambda - 1)(\lambda - 2)^2$.

i) Eigenraum zu $\lambda = 2$ ist Lösungsraum von

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

3 linear abhängige Gleichungen mit der Lösungsmenge

$$\left\{ \begin{pmatrix} s \\ t \\ -s \end{pmatrix} \middle| s, t \in \mathbb{R} \right\}.$$

Eine Basis dieses 2-dimensionalen Eigenraums ist z.B.

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

(a) Eigenraum zu $\lambda = 1$ ist Lösungsraum von

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

1. und 3. Gleichung sind linear abhängig.
Addition von Gleichung 1 und 2:

$$\begin{aligned} x_2 - x_3 &= 0 \\ x_2 &:= s \quad \Rightarrow x_3 = s \end{aligned}$$

in Gleichung 3: $x_1 = -2x_3 = -2s$.
Eindimensionaler Eigenraum $\left\{ \begin{pmatrix} -2s \\ s \\ s \end{pmatrix} \middle| s \in \mathbb{R} \right\}$ wird z.B. vom Basisvektor $\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ aufgespannt.

46.14 Satz: (Eigenwerte von Potenzen einer Matrix)

Sei $A \in \mathbb{R}^{n \times n}, k \in \mathbb{N}$ und λ sei Eigenwert von A mit Eigenvektor v. Dann ist λ^k Eigenwert von A^k mit zugehörigem Eigenvektor v.

Beweis:

$$\begin{aligned} A^k v &= A^{k-1}(Av) = A^{k-1}(\lambda v) = \lambda A^{k-1} v \\ &= \lambda A^{k-2}(Av) = \lambda^2 A^{k-2} v \\ &= \ldots = \lambda^k v. \end{aligned}$$

$\square$.

46.15 Beispiel

A^7 mit $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$ aus Beispiel 46.13 hat die Eigenwerte $\lambda_1 = 1^7$ und $\lambda_2 = 2^7 = 128$.

Kapitel 47

Eigenwerte und Eigenvektoren symmetrischer Matrizen

47.1 Motivation

Symmetrische Matrizen (d.h. $a_{ij} = a_{ji} \ \ \forall i,j$) kommen in der Praxis sehr häufig vor. Gibt es in diesem Fall besonders einfache Aussagen über Eigenwerte und Eigenvektoren?

47.2 Satz: (Eigenwerte, -vektoren symmetrischer Matrizen)

Für eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ gilt:

(*a*) *A hat nur reelle Eigenwerte.*

(*b*) *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beweis:

(a) Sei $z = x + \mathbf{i}y \in \mathbb{C}$ und $\overline{z} := x - \mathbf{i}y$ die komplex konjugierte Zahl. Dann ist $z\overline{z} = x^2 + y^2 \in \mathbb{R}$.
Für Vektoren und Matrizen definiert man die komplexe Konjugation komponentenweise. Sei nun λ Eigenwert von A zum Eigenvektor v.

$$\begin{aligned}
\Rightarrow \overline{\lambda}\overline{v}^\top v &= (\overline{\lambda v})^\top v \\
&= (\overline{Av})^\top v \\
&= \overline{v}^\top \overline{A}^\top v \\
&= \overline{v}^\top A v \qquad \text{da } A \text{ reell und symmetrisch} \\
&= \overline{v}^\top (\lambda v) \\
&= \lambda \overline{v}^\top v.
\end{aligned}$$

Da $\overline{v}^\top v \in \mathbb{R}$ und $\neq 0$ (Eigenvektoren sind $\neq 0$), ist $\overline{\lambda} = \lambda$, d.h. $\lambda \in \mathbb{R}$.

(b) Seien v_1, v_2 Eigenvektoren von A zu verschiedenen Eigenvektoren λ_1, λ_2.

$$\begin{aligned}
\Rightarrow \lambda_1 v_1^\top v_2 &= (Av_1)^\top v_2 \\
&= v_1^\top A^\top v_2 \\
&= v_1^\top (Av_2) && \text{da } A \text{ symmetrisch} \\
&= v_1^\top (\lambda_2 v_2) && \text{da } \lambda_2 \text{ Eigenwert von } A \text{ zum Eigenvektor } v_2 \\
&= \lambda_2 v_1^\top v_2 \\
\Rightarrow 0 &= \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} v_1^\top v_2
\end{aligned}$$

Also sind v_1 und v_2 orthogonal. □.

47.3 Beispiel

$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ symmetrisch.

Eigenwerte:

$$\begin{aligned}
0 &= \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) - 4 \\
&= 4 - 4\lambda + \lambda^2 - 4 = \lambda^2 - 5\lambda = \lambda(\lambda - 5)
\end{aligned}$$

$\Rightarrow$ zwei reelle Eigenwerte: $\lambda_1 = 0, \lambda_2 = 5$.

Eigenvektor zu $\lambda_1 = 0$:
$(A - \lambda_1 I)v = 0$
$\Rightarrow \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ 2 linear abhängige Gleichungen

$x_1 + 2x_2 = 0$
$x_2 := s \quad \Rightarrow x_1 = -2x_2 = -2s$

Eigenraum: $\left\{ \begin{pmatrix} -2s \\ s \end{pmatrix} \middle| s \in \mathbb{R} \right\}$
Eigenvektor: z.B. $v_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$

Eigenvektor zu $\lambda_2 = 5$:
$(A - \lambda_2 I)v = 0$
$\Rightarrow \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ 2 linear abhängige Gleichungen

$2x_1 - x_2 = 0$
$x_1 := s \quad \Rightarrow x_2 = 2s.$
Eigenraum $\left\{ \begin{pmatrix} s \\ 2s \end{pmatrix} \middle| s \in \mathbb{R} \right\}$
Eigenvektor: z.B. $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

v_1 und v_2 sind orthogonal.

Symmetrische Matrizen lassen sich mit Hilfe ihrer Eigenwerte und Eigenvektoren elegant zerlegen:

47.4 Satz: (Hauptachsentransformation, Spektraldarstellung)

Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Nach Satz 46.2 hat A ein Orthonormalsystem von Eigenvektoren $v_1, \dots, v_n$ mit zugehörigen (nicht notwendigerweise verschiedenen) Eigenwerten $\lambda_1, \dots, \lambda_n$. Dann ist

$$\boxed{A = Q\Lambda Q^\top}$$

mit der orthogonalen Matrix $Q = (v_1 \mid \dots \mid v_n)$ und der Diagonalmatrix

$$\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}.$$

Beweis:

Nach Satz 47.2 sind die Eigenräume zu verschiedenen Eigenwerten orthogonal. Verwendet man innerhalb jedes Eigenraums das Gram-Schmidt-Verfahren und normiert, entsteht ein Orthonormalsystem $\{v_1, \dots, v_n\}$ von Eigenvektoren von A. Somit ist $Q = (v_1 \mid \dots \mid v_n)$ eine orthogonale Matrix. Die k-te Spalte von $Q^\top AQ$ lautet

$$\begin{aligned} Q^\top \underbrace{Av_k}_{\lambda_k v_k} &= \lambda_k Q^\top v_k \\ &\overset{\text{Orthonormalität}}{=} \lambda_k e_k \\ \text{mit } e_k &= \begin{pmatrix} 0 \\ | \\ 0 \\ 1 \\ 0 \\ | \\ 0 \end{pmatrix} \leftarrow k\text{-te Stelle.} \end{aligned}$$

Somit ist $Q^\top AQ = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix} = 1.$ $\square$.

47.5 Bemerkungen

(a) Das bedeutet, dass A auf Diagonalgestalt transformiert werden kann:

$$\Lambda = Q^\top AQ$$

Durch den Übergang in das durch $Q = (v_1 \mid \dots \mid v_n)$ definierte Koordinatensystem hat A eine besonders einfache Gestalt.

(b) A lässt sich auch schreiben als

$$A = \lambda_1 v_1 v_1^\top + \dots + \lambda_n v_n v_n^\top$$

An dieser Schreibweise erkennt man sofort, dass $\lambda_1, \dots, \lambda_n$ Eigenwerte und $v_1, \dots, v_k$ Eigenvektoren von A sind, denn:

$$\begin{aligned} Av_k &= \sum_{i=1}^n \lambda_i v_i \underbrace{v_i^\top v_k}_{0 \text{ für } i\neq k} \\ &= \lambda_k v_k. \end{aligned}$$

47.6 Beispiel

Transformiere $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ auf Diagonalgestalt.

Lösung: Nach Beispiel 47.3 hat A die Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = 5$ mit zugehörigen Eigenvektoren $w_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ und $w_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Normierung der Eigenvektoren ergibt: $v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ und $v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Mit der orthogonalen Matrix $Q = (v_1 \mid v_2) = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$ ergibt sich:

$$
\begin{aligned}
Q^\top A Q &= \frac{1}{5} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 5 \\ 0 & 10 \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} = \operatorname{diag}(\lambda_1, \lambda_2)
\end{aligned}
$$

wie nach Satz 47.5 (a) zu erwarten war.

Kapitel 48

Quadratische Formen und Positive Definite Matrizen

48.1 Motivation

- Charakterisierung einer wichtigen Klasse symmetrischer Matrizen (Anwendungen in Computergrafik, Physik, u.a.)
- Verhalten quadratischer Funktionen in mehreren Variablen untersuchen
- wichtige Klassen geometrischer Kurven/Flächen kennen lernen

48.2 Definition (Quadratische Form, quadratisches Polynom, Quadrik)

Es sei $x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann heißt

$$x^\top Ax = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

quadratische Form.
Ferner seien $b \in \mathbb{R}^n$ und $c \in \mathbb{R}$. Dann heißt

$$q(x) = x^\top Ax + b^\top x + c$$

quadratisches Polynom in $x_1, \dots, x_n$.
Die Menge aller Punkte, die die quadratische Gleichung

$$q(x) = x^\top Ax + b^\top x + c = 0$$

erfüllen, heißt Quadrik.

48.3 Beispiel

(a)

$$\begin{aligned} 7x_1^2 + 6x_2^2 + 5x_3^2 - 4x_1x_2 + 2x_2x_3 &= 7x_1^2 + 6x_2^2 + 5x_3^2 - 2x_1x_2 - 2x_2x_1 + x_2x_3 + x_2x_3 \\ &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & 1 \\ 0 & 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

ist eine quadratische Form.

(b) $g(x) = 5x_1^2 - 3x_2^2 + 4x_1x_2 - 7x_2 + 3$ ist ein quadratisches Polynom.

(c) Die Ellipsengleichung

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - 1 = 0$$

beschreibt eine Quadrik mit $n = 2$.

Quadriken mit $n = 2$ können generell als Kegelschnitte (Ellipsen, Parabeln, Hyperbeln, ...) interpretiert werden.
Für $n = 3$ ergeben sich Ellipsoide, Hyperboloide, Paraboloide, ...

48.4 Definition (Definite, semidefinite, indefinite Matrizen)

Es sei $A \in \mathbb{R}^{n\times n}$ symmetrisch; $\lambda_1, \dots, \lambda_n$ seien die Eigenwerte von A. Dann heißt A

- *positiv definit, falls $\lambda_i > 0 \quad \forall i$*
- *positiv semidefinit, falls $\lambda_i \geqq 0 \quad \forall i$*
- *negativ definit, falls $\lambda_i < 0 \quad \forall i$*
- *negativ semidefinit, falls $\lambda_i \leqq 0 \quad \forall i$*
- *indefinit, falls λ_i, λ_j existieren mit $\lambda_i\lambda_j < 0$.*

Es besteht ein enger Zusammenhang zwischen positiver Definitheit, quadratischen Formen und Determinanten von Untermatrizen.

48.5 Satz: (Positiv definite Matrizen und quadratische Formen)

Es sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ symmetrisch. Dann ist A positiv definit genau dann, wenn

$$x^\top A x > 0 \qquad \forall x \in \mathbb{R}^n, \quad x \neq 0.$$

Beweis:

„$\Rightarrow$“: Es sei $\{v_1, \dots, v_n\}$ Orthonormalbasis von Eigenvektoren von A mit zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$.
Dann gilt für $x = \sum_{i=1}^{n} x_i v_i \quad (x \neq 0)$:

$$\begin{aligned}
x^\top A x &= \left(\sum_{i=1}^{n} x_i v_i\right)^\top A \left(\sum_{j=1}^{n} x_j v_j\right) \\
&= \left(\sum_{i=1}^{n} x_i v_i\right)^\top \left(\sum_{j=1}^{n} x_j \underbrace{A v_j}_{\lambda_j v_j}\right) \\
&= \sum_{i,j=1}^{n} \lambda_j x_i x_j \underbrace{(v_i^\top v_j)}_{=0 \text{ für } i\neq j} \\
&= \sum_{i=1}^{n} \lambda_i x_i^2 > 0.
\end{aligned}$$

„$\Leftarrow$“: Ist umgekehrt A nicht positiv definit, so existiert ein Eigenwert $\lambda \leqq 0$ von A mit zugehörigem Eigenvektor $v \neq 0$. Damit ist

$$v^\top Av = v^\top \lambda v = \underbrace{\lambda}_{\leq 0} \underbrace{(v^\top v)}_{>0} \leqq 0$$

im Widerspruch zu $x^\top Ax > 0 \quad \forall x \neq 0$. □.

Der Zusammenhang zwischen positiver Definitheit und Determinanten von Untermatrizen ist Gegenstand des folgenden Satzes. Er stellt ein wichtiges Kriterium zum Überprüfen der positiven Definitheit ohne Eigenwertberechnung dar.

48.6 Satz: (Hauptminorenkriterium)

Eine symmetrische Matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn ihre Hauptminoren

$$\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix}$$

für $k = 1, \dots, n$ positiv sind.

Bemerkung: Ein ähnliches Kriterium für Semidefinitheit anzugeben ist nicht so einfach, man muss dann alle quadratischen Untermatrizen (und nicht nur die Hauptminoren einbeziehen).

48.7 Beispiel

Geg.: $A = \begin{pmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{pmatrix}$. Ist A positiv definit?

Hauptminoren:

$$\begin{aligned}
\det(2) &= 2 > 0 \\
\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} &= 4 - 1 = 3 > 0 \\
\begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} &= 2\begin{vmatrix} 2 & 4 \\ 4 & 9 \end{vmatrix} - (-1)\begin{vmatrix} -1 & -3 \\ 4 & 9 \end{vmatrix} + (-3)\begin{vmatrix} -1 & -3 \\ 2 & 4 \end{vmatrix} \\
&= 2(18-16) + (-9+12) - 3(-4+6) = 4 + 3 - 6 = 1 > 0
\end{aligned}$$

$\Rightarrow A$ positiv definit.

Analog zu Quadratwurzeln aus nichtnegativen Zahlen lassen sich auch „Wurzeln“ aus einer positiv-semidefiniten Matrix definieren.

48.8 Satz: (Wurzel einer positiv semidefiniten Matrix)

Es sei $A \in \mathbb{R}^{n\times n}$ symmetrisch, positiv semidefinit. Dann existiert eine symmetrische, positiv semidefinite Matrix $B \in \mathbb{R}^{n\times n}$ mit $B^2 = A$.

Beweis:

Es seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A und $v_1, \dots v_n$ die zugehörigen normierten Eigenvektoren. Dann gilt mit $Q = (v_1|\dots|v_n)$ und $\Lambda := \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$, dass $A = Q\Lambda Q^\top$.

Wir setzen $\Lambda^{1/2} := \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}$ und $B := Q\Lambda^{1/2}Q^\top$.

Dann ist $B^2 = Q\Lambda^{1/2} \underbrace{Q^\top Q}_{I} \Lambda^{1/2} Q^\top = Q\Lambda^{1/2}\Lambda^{1/2}Q^\top = Q\Lambda Q^\top = A$. $\square$.

Positiv und negativ definite Matrizen spielen eine wichtige Rolle beim Nachweis von Minima und Maxima von Funktionen mehrerer Variablen ($\rightarrow$ Kap E).

Gibt es obere und untere Schranken für quadratische Formen?

48.9 Definition (Rayleigh-Quotient)

Es sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und $x \in \mathbb{R}^n$ mit $x \neq 0$. Dann nennt man

$$R_A(x) := R(x) := \frac{x^\top Ax}{x^\top x}$$

den Rayleigh-Quotienten.

Der Rayleigh-Quotient lässt sich durch die Eigenwerte von A abschätzen.

48.10 Satz: (Rayleigh-Prinzip)

Es sei $A \in \mathbb{R}^{n\times n}$ symmetrisch mit $\lambda_1 \geq \dots \geq \lambda_n$ und zugehörigen orthonormierten Eigenvektoren $v_1, \dots, v_n$. Dann gilt:

(a) $\lambda_n \leq R(x) \leq \lambda_1$

(b) *Diese Grenzen werden tatsächlich angenommen:*

$$\lambda_1 = \max_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} R(x), \quad \lambda_n = \min_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} R(x).$$

Beweis:

(a) Aus $x = \sum_{i=1}^{n} x_i v_i$ folgt $x^\top x = \sum_{i=1}^{n} x_i^2$; analog zum Beweis von Satz 48.5 ist außerdem

$$\begin{aligned} x^\top Ax &= \sum_{i=1}^{n} \lambda_i x_i^2 \\ \Rightarrow R(x) &= \frac{\sum_{i=1}^n \lambda_i x_i^2}{\sum_{i=1}^n x_i^2} \begin{cases} \leq \dfrac{\sum_{i=1}^n \lambda_1 x_i^2}{\sum_{i=1}^n x_i^2} = \lambda_1, \\[2ex] \geq \dfrac{\sum_{i=1}^n \lambda_n x_i^2}{\sum_{i=1}^n x_i^2} = \lambda_n. \end{cases} \end{aligned}$$

(b) Setzt man $x = v_k$, so folgt

$$R(v_k) = \frac{v_k^\top A v_k}{v_k^\top v_k} = \frac{v_k^\top \lambda_k v_k}{v_k^\top v_k} = \lambda_k$$

Insbesondere ist $R(v_1) = \lambda_1, R(v_n) = \lambda_n$. $\square$.

Kapitel 49

Quadriken

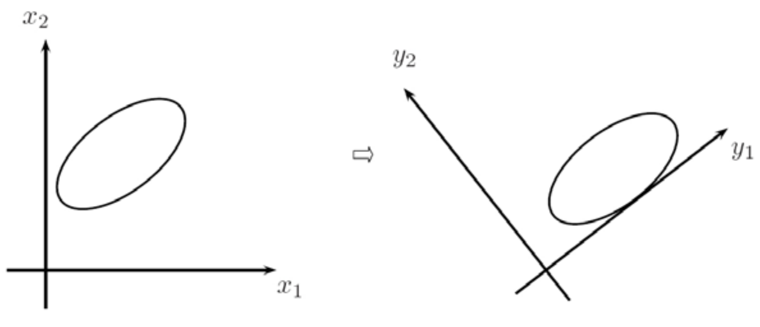
49.1 Motivation

Quadriken (→ Def. 48.2) stellen eine wichtige Klasse geometrischer Objekte dar, mit Anwendungen in Computergrafik, Physik, u.a.
Ziel: gegebene Quadrik auf einfache Form transformieren, sodass sich ihre geometrische Gestalt unmittelbar ablesen lässt.

49.2 Grundlegende Verfahrensweise

Gegeben: Quadrik $q(x) = x^\top Ax + b^\top x + c = 0$
$A \in \mathbb{R}^{n\times n}$ symmetrisch, $x, b \in \mathbb{R}^n, c \in \mathbb{R}$.

Schritt 1: Elimination der gemischten quadratischen Terme
Hierzu wird das Koordinatensystem so gedreht, dass A in eine Diagonalmatrix übergeht.



Berechne dazu die Eigenwerte λ_i von A und eine ONB $\{v_1, \dots v_n\}$ aus Eigenvektoren mit $\det(v_1 \mid \dots \mid v_n) = 1$.
(Falls $\det(v_1 \mid \dots \mid v_n) = -1$, ersetzt man v_1 durch $-v_1$)
Mit $Q := (v_1 \mid \dots \mid v_n)$ gilt dann

$$\Lambda = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix} = Q^\top AQ,$$

und aus $x^\top Ax + b^\top x + c = 0$ folgt

$$x^\top Q\Lambda Q^\top x + b^\top \underbrace{QQ^\top}_{I} x + c = 0.$$

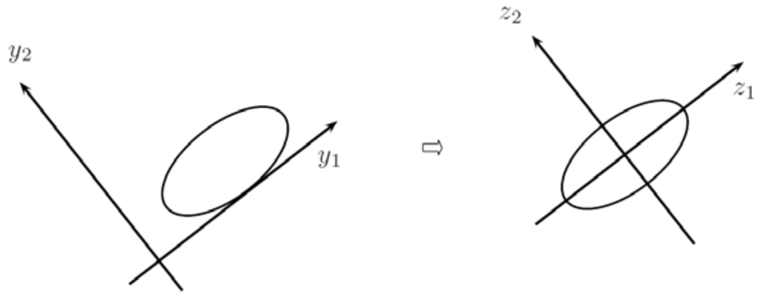
Mit $y := Q^\top x, \tilde{b} := Q^\top b$ ergibt sich daher $y^\top \Lambda y + \tilde{b}^\top y + c = 0$ bzw. ausgeschrieben:

$$\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 + \tilde{b}_1 y_1 + \ldots + \tilde{b}_n y_n + c = 0$$

(gemischte quadratische Terme weggefallen).

Schritt 2: Elimination linearer Terme (soweit möglich)

Durch Translation des Koordinatensystems kann erreicht werden, dass $\lambda_k y_k^2$ und $\tilde{b}_k y_k$ nicht zugleich vorkommen (für jedes k).



Es sei dazu o.B.d.A. $\lambda_i \neq 0$ für $i = 1, \ldots, r$ sowie $\lambda_{r+1} = \ldots = \lambda_n = 0$. Für $i = 1, \ldots, r$ wird der lineare Term $\tilde{b}_i y_i$ durch die quadratische Ergänzung eliminiert:

$$\begin{aligned} z_i &:= y_i + \frac{\tilde{b}_i}{2\lambda_i} \quad (i = 1, \ldots, r) \\ z_i &:= y_i \qquad\quad (i = r+1, \ldots, n) \end{aligned}$$

Damit erhält man

$$\lambda_1 z_1^2 + \ldots + \lambda_r z_r^2 + \tilde{b}_{r+1} z_{r+1} + \ldots + \tilde{b}_n z_n + \tilde{c} = 0$$

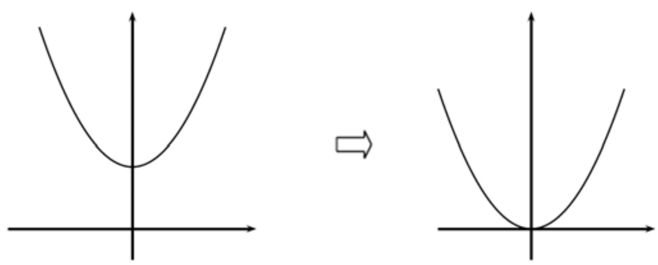
mit $\tilde{c} = c - \sum_{i=1}^{r} \frac{\tilde{b}_i^2}{4\lambda_i}$ und $r = \operatorname{rang} A$.

Schritt 3: Elimination der Konstanten (soweit möglich)

Ist einer der Koeffizienten $\tilde{b}_{r+1}, \ldots, \tilde{b}_n$ ungleich 0. (o.B.d.A. sei dies $\tilde{b}_n$), so kann $\tilde{c}$ eliminiert werden durch

$$z_n \mapsto z_n - \frac{\tilde{c}}{\tilde{b}_n}$$

ebenfalls Translation des Koordinatensystems, z.B.



Resultat: Normalformen der Quadrik

Darstellung in Koordinatensystem, in dem möglichst viele Koeffizienten verschwinden.

Für	$r := \operatorname{rang} A = n$:		$\lambda_1 z_1^2 + \ldots + \lambda_n z_n^2 + d = 0$
Für	$r < n$:	entweder	$\lambda_1 z_1^2 + \ldots + \lambda_r z_r^2 + e_{r+1} z_{r+1} + \ldots + e_n z_n = 0$
		oder	$\lambda_1 z_1^2 + \ldots + \lambda_r z_r^2 + d = 0.$

49.3 Beispiel

Die Quadrate

$$q(x) = 5x_1^2 - 4x_1x_2 + 8x_2^2 + \frac{20}{\sqrt{5}}x_1 - \frac{80}{\sqrt{5}}x_2 + 4 = 0$$

soll auf Normalform gebracht werden.

$$q(x) = x^\top Ax + b^\top x + c = 0$$

$$\text{mit } A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}, \quad b = \frac{1}{\sqrt{5}}\begin{pmatrix} 20 \\ -80 \end{pmatrix}, \quad c = 4$$

Schritt 1: (Hauptachsentransformation von A)

Eigenwerte $\lambda_1 = 9, \lambda_2 = 4$; $Q = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 & -2 \\ 2 & -1 \end{pmatrix}$ mit $\det Q = 1$.

Mit $\Lambda = Q^\top AQ = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix}$ und $\tilde{b} = Q^\top b = \begin{pmatrix} -36 \\ 8 \end{pmatrix}$ ergibt sich für

$$y = Q^\top x : \qquad 9y_1^2 + 4y_2^2 - 36y_1 + 8y_2 + 4 = 0.$$

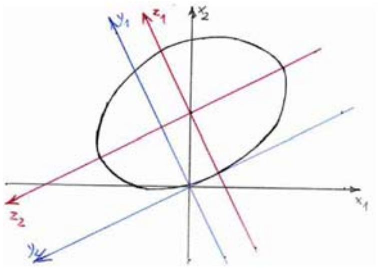
Schritt 2: (Elimination linearer Terme)

$9(y_1^2 - 4y_1 + 4) + 4(y_2^2 + 2y_2 + 1) = -4 + 36 + 4$
also mit $z_1 = y_1 - 2$ und $z_2 = y_2 + 1$:

$9z_1^2 + 4z_2^2 = 36$

$\Rightarrow \frac{z_1^2}{4} + \frac{z_2^2}{9} = 1$

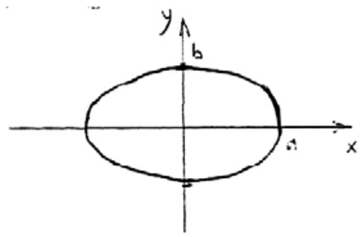
Ellipse mit Halbachsen 2 und 3.



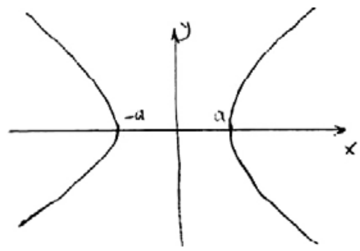
49.4 Normalformen der Quadriken im $\mathbb{R}^2$

(i) $\underline{\operatorname{rang} A = 2}$ (Alle Eigenwerte $\neq 0$)

a) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ Ellipse



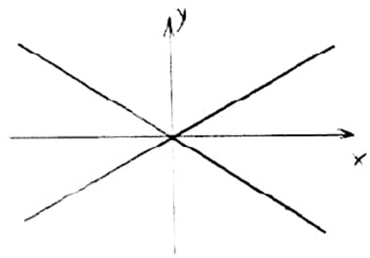
b) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$ Hyperbel



c) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$ leere Menge

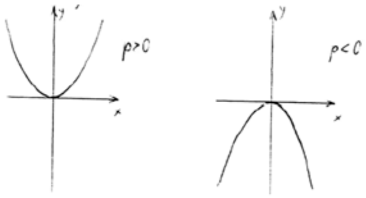
d) $x^2 + a^2y^2 = 0,\ a \neq 0$ Punkt $(0,0)$

e) $x^2 - a^2y^2 = 0,\ a \neq 0$ Geradenpaar $y = \pm\frac{1}{a}x$.

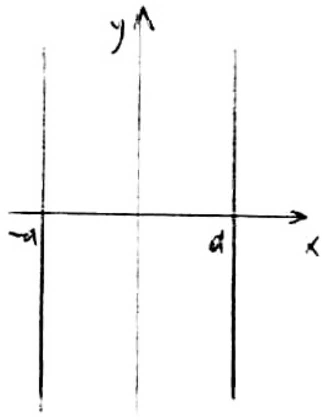


(ii) $\underline{\operatorname{rang} A = 1}$ (ein Eigenwert $= 0$)

a) $x^2 - 2py = 0$ <u>Parabel</u>



b) $x^2 - a^2 = 0$ <u>parallele Geraden</u> $x = \pm 0$



c) $x^2 + a^2 = 0$ <u>leere Menge</u>

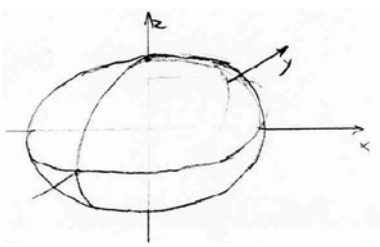
d) $x^2 = 0$ <u>„Doppelgerade"</u> $x = 0$ (y-Achse)

(iii) $\underline{\operatorname{rang} A = 0}$ (Alle Eigenwert $= 0$)
$b_1 x + b_2 y + c = 0$ Gerade

49.5 Normalformen der Quadriken im $\mathbb{R}^3$

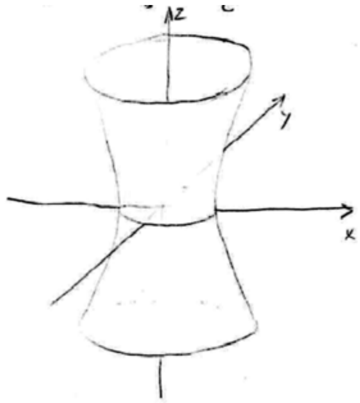
(i) $\underline{\operatorname{rang} A = 3}$ (Alle Eigenwerte $\neq 0$)

a) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} - 1 = 0$ <u>Ellipsoid</u>

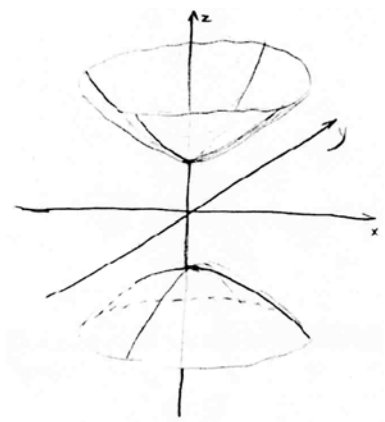


b) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} + 1 = 0$ <u>leere Menge</u>

c) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0$ <u>einschaliges Hyperboloid</u>
Ellipse in 1 Ebene ($x - y$-Ebene)
Hyperbel in 2 Ebenen ($x - z, y - z$)

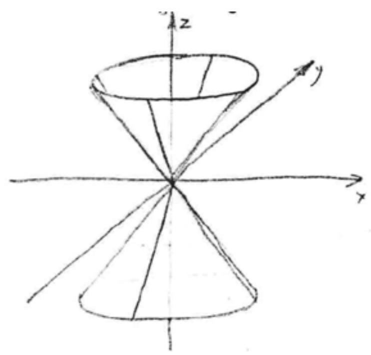


d) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} + 1 = 0$ <u>zweischaliges Hyperboloid</u>
Ellipse in 1 Ebene ($x - y$)
Hyperbeln in 2 Ebenen ($x - z,\ y - z$)



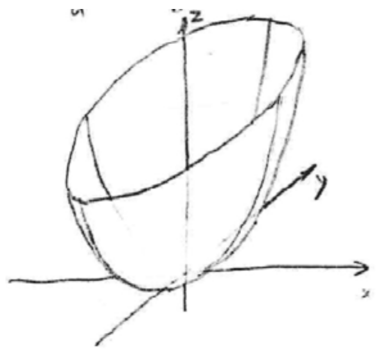
e) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$ <u>Punkt(0,0,0)</u>

f) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ <u>elliptischer Kegel</u>

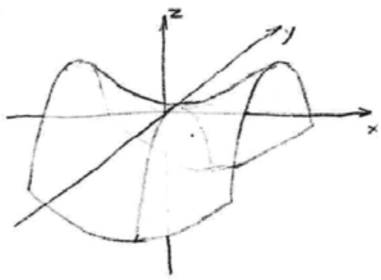


(ii) $\underline{\text{rang } A = 2}$ (Ein Eigenwert = 0)

a) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 2pz = 0$ elliptischer Paraboloid
in 1 Ebene Ellipse $(x - y)$
in 2 Ebenen Parabeln $(x - z, y - z)$

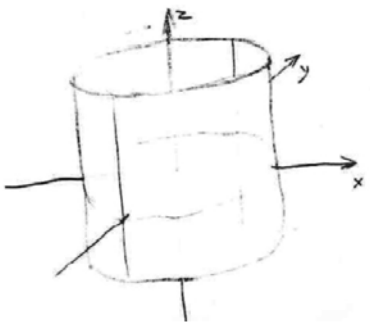


b) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 2pz = 0$ hyperbolisches Paraboloid (sattelartig)
in 1 Ebene Hyperbel $(x - y)$
in 2 Ebenen Parabeln $(x - z, y - z)$

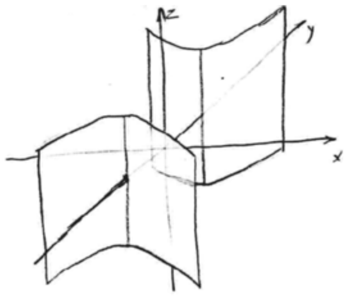


c) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$ leere Menge

d) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ elliptischer Zylinder

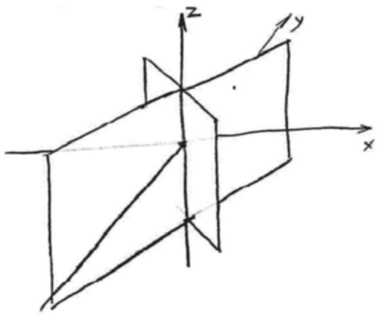


e) $\frac{x^2}{a^2} - \frac{y^2}{b^2} + 1 = 0$ hyperbolischer Zylinder



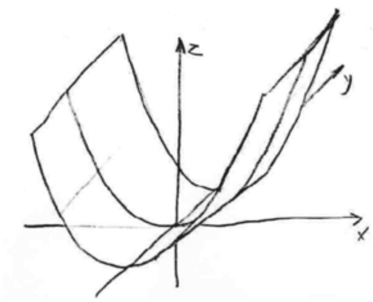
f) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$ Gerade (z-Achse)

g) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ Ebenenpaar mit Schnittgerade (z-Achse)



(iii) rang $A = 1$

a) $x^2 - 2pz = 0$ parabolischer Zylinder



b) $x^2 - a^2 = 0$ paralleles Ebenenpaar

c) $x^2 + a^2 = 0$ leere Menge

d) $x^2 = 0$ Ebene ($y - z$-Ebene)

(iv) rang $A = 0$

$b_1x + b_2y + b_3z + c = 0$ allgemeine Ebenengleichung

49.6 Satz:

Auf dem einschaligen Hyperboloid und auf dem hyperbolischen Paraboloid gibt es jeweils zwei Scharen von Geraden.

Beweis (für das einschalige Hyperboloid):

Gleichung des einschaligen Hyperboloiden:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}-1=0$$

Die Schnittkurve mit der $x-y$-Ebene ist die Ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

In jedem Punkt dieser Ellipse gibt es eine Berührungsebene an das Hyperboloiden und wir zeigen: Jede dieser Ebenen schneidet das Hyperboloid in zwei Geraden.

Fall 1: Berührungspunkt $(x_0,y_0,z_0)=(a,0,0)$ oder $(-a,0,0)$.
In diesem Fall verläuft die Berührungsebene parallel zur y- und zur z-Achse. Ihre Gleichung ist daher $x=x_0$, also $x=\pm a$.
Der Schnitt der Ebene mit dem Hyperboloid besteht aus allen Punkten, die die Gleichungen beider Flächen erhalten. Für jeden solchen Punkt muss also gelten:

$$\begin{aligned}
& x=\pm a \quad\wedge\quad \frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}-1=0\\
\overset{\substack{\text{Einsetzen in}\\ \text{Hyperboloidgl.}}}{\Leftrightarrow}\quad & x=\pm a \quad\wedge\quad 1+\frac{y^2}{b^2}-\frac{z^2}{c^2}-1=0\\
\Leftrightarrow\quad & x=\pm a \quad\wedge\quad \left(\frac{y}{b}+\frac{z}{c}\right)\left(\frac{y}{b}-\frac{z}{c}\right)=0\\
\Leftrightarrow\quad & \underbrace{\left(x=\pm a \quad\wedge\quad \frac{y}{b}+\frac{z}{c}=0\right)}_{\text{1. Schnittgerade}}\vee\underbrace{\left(x=\pm a \quad\wedge\quad \frac{y}{b}-\frac{z}{c}=0\right)}_{\text{2. Schnittgerade}}
\end{aligned}$$

Fall 2: Berührungspunkt (x_0,y_0,z_0) mit $\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}=1, y_0\neq 0$, und $z_0=0$.
Die Tangente an die Ellipse in der $x-y$-Ebene hat die Gleichung

$$\frac{x_0}{a^2}(x-x_0)+\frac{y_0}{b^2}(y-y_0)=0.$$

Dies ist auch die Gleichung der Berührungsebene an den Hyperboloid, da dieser parallel zur z-Achse ist.
Schnittpunkte der Ebene mit dem Hyperboloid:

$$\begin{aligned}
& \frac{x_0}{a^2}(x-x_0)+\frac{y_0}{b^2}(y-y_0)=0 \quad\wedge\quad \frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}-1=0\\
\overset{y_0\neq 0}{\Leftrightarrow}\quad & y=y_0-\frac{b^2x_0}{a^2y_0}(x-x_0) \quad\wedge\quad \frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}-1=0
\end{aligned}$$

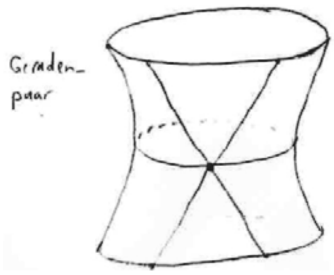
Einsetzen der ersten in die zweite Gleichung ergibt

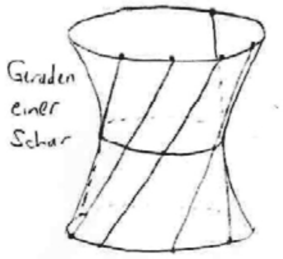
$$\begin{aligned}
0 &= \frac{x^2}{a^2}+\frac{y_0^2}{b^2}-2\frac{x_0}{a^2}(x-x_0)+\frac{b^2x_0^2}{a^2y_0^2}(x-x_0)^2-\frac{z^2}{c^2}-1\\
&= \frac{x^2-x_0^2}{a^2}-2\frac{x_0}{a^2}(x-x_0)+\frac{b^2x_0^2}{a^2y_0^2}(x-x_0)^2-\frac{z^2}{c^2}\\
&= \frac{1}{a^2}(x-x_0)(x+x_0-2x_0)+\frac{b^2}{a^2}\frac{x_0^2}{y_0^2}(x-x_0)^2-\frac{z^2}{c^2}\\
&= \frac{1}{a^2}\underbrace{\left(1+\frac{b^2x_0^2}{a^2y_0^2}\right)}_{=:e^2}(x-x_0)^2-\frac{z^2}{c^2}\\
&= \left(\frac{e}{a}(x-x_0)+\frac{z}{c}\right)\left(\frac{e}{a}(x-x_0)-\frac{z}{c}\right)
\end{aligned}$$

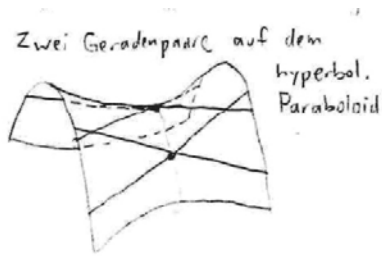
Schnittpunkte erfüllen also

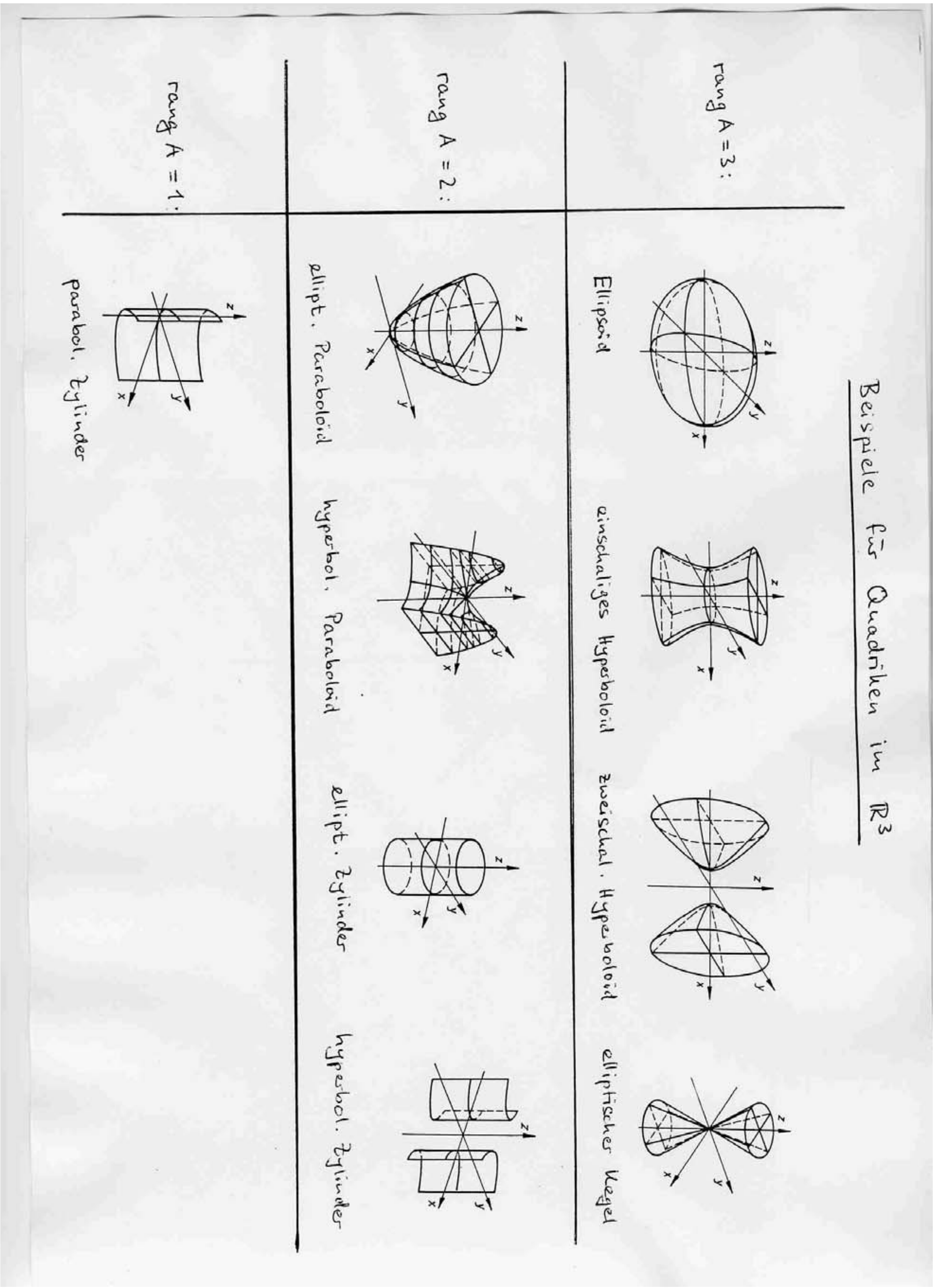
$$\begin{array}{rlcll} & y = y_0 - \dfrac{b^2 x_0}{a^2 y_0}(x - x_0) & \wedge & \dfrac{e}{a}(x - x_0) + \dfrac{z}{c} = 0 & \text{(1. Schnittgerade)} \\ \text{oder:} & y = y_0 - \dfrac{b^2 x_0}{a^2 y_0}(x - x_0) & \wedge & \dfrac{e}{a}(x - x_0) - \dfrac{z}{c} = 0 & \text{(2. Schnittgeraden)} \end{array}$$

$\square$.








Beispiele für Quadriken im ℝ³
rang A = 3:
Ellipsoid
einschaliges Hyperboloid
zweischal. Hyperboloid
elliptischer Kegel
rang A = 2:
ellipt. Paraboloid
hyperbol. Paraboloid
ellipt. Zylinder
hyperbol. Zylinder
rang A = 1:
parabol. Zylinder

Kapitel 50

Matrixnormen und Eigenwertabschätzungen

50.1 Motivation

Problem:

(a) Kann man die Eigenwerte einer Matrix mit geringrem Aufwand abschätzen?

(b) Dies spielt z.B. eine Rolle bei Konvergenzbetrachtungen von iterativen Algorithmen.

Ein wichtiges Hilfsmittel hierzu sind Matrixnormen.

50.2 Definition (Matrixnorm)

Unter einer Matrixnorm versteht man eine Funktion $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$ mit folgenden Eigenschaften:

(a) $\|A\| \geq 0 \quad \forall A \in \mathbb{R}^{n\times n}$
$\|A\| = 0 \Leftrightarrow A = 0$

(b) $\|\lambda A\| = |\lambda| \cdot \|A\| \quad \forall \lambda \in \mathbb{R}, A \in \mathbb{R}^{n\times n}$

(c) $\|A + B\| \leq \|A\| + \|B\|$

(d) $\|A \cdot B\| \leq \|A\| \cdot \|B\| \quad \forall A, B \in \mathbb{R}^{n\times n}$ *(Submultiplikativität)*

50.3 Beispiele

Sei $A \in \mathbb{R}^{n\times n}$.

(a) Gesamtnorm: $\|A\|_G := n \cdot \max_{i,k} |a_{ik}|$

(b) Zeilensummennorm: $\|A\|_Z := \max_i \sum_{k=1}^{n} |a_{ik}|$

(c) Spaltensummennorm: $\|A\|_S := \max_k \sum_{i=1}^{n} |a_{ik}|$

(d) Frobeniusnorm: $\|A\|_F := \left(\sum_{i,k=1}^{n} a_{ik}^2 \right)^{1/2}$

(e) Spektralnorm: $\|A\|_2 := \sqrt{\lambda_{\max}(A^\top A)}$
wobei $\lambda_{\max}(A^\top A)$ der größte Eigenwert von $A^\top A$ ist.
Falls A symmetrisch: $\|A\|_2 = \max_k \{|\lambda_k| \mid \lambda_k$ Eigenwert von $A\}$.

Da Matrizen und Vektoren oft gemeinsam auftreten, sollten Matrix- und Vektornormen verträglich sein.

50.4 Definition (Kompatibilität, Verträglichkeit)

Eine Matrixnorm $\|\cdot\|_M$ heißt kompatibel (verträglich) mit einer Vektornorm $\|\cdot\|_V$, falls gilt

$$\|Ax\|_V \leq \|A\|_M \cdot \|x\|_V \qquad \forall A \in \mathbb{R}^{n \times n}, \forall x \in \mathbb{R}^n.$$

50.5 Beispiele

Zu den p-Normen

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} & \text{für } 1 \leq p < \infty \\ \max_i \{|x_i|\} & \text{für } p = \infty \end{cases}.$$

bestehen folgende Verträglichkeiten:

(a) $\|A\|_G, \|A\|_S$ sind kompatibel zur Betragssummennorm $\|x\|_1$

(b) $\|A\|_G, \|A\|_F, \|A\|_2$ sind kompatibel zur euklidischen Norm $\|x\|_2 = |x|$

(c) $\|A\|_G, \|A\|_Z$ sind kompatibel zur Maximumsnorm $\|x\|_\infty$.

Beweis: Wir zeigen nur die Kompatibilität von $\|A\|_G$ und $\|x\|_\infty$:

$$\begin{aligned} \|Ax\|_\infty &= \max_i \left\{ \left| \sum_{k=1}^{n} a_{ik} x_k \right| \right\} \\ &\leq \max_i \left\{ \sum_{k=1}^{n} |a_{ik} x_k| \right\} && \text{Dreiecksungleichungen} \\ &\leq \max \left\{ \sum_{k=1}^{n} \max_{r,s} |a_{rs}| \max_l |x_l| \right\} \\ &= n \cdot \max_{r,s} |a_{rs}| \max_l |x_l| \\ &= \|A\|_G \cdot \|x\|_\infty . \end{aligned}$$

□.

Da zu einer gegebenen Vektornorm $\|\cdot\|_V$ oftmals viele kompatible Matrixnormen $\|\cdot\|_M$ existieren, verwendet man in der Praxis gerne diejenigen, für die die Abschätzung $\|Ax\|_V \leq \|A\|_M \cdot \|x\|_V$ am schärfsten ist:

50.6 Definition (Zugeordnete Matrixnorm)

Die zu einer gegebenen Vektornorm $\|x\|$ definierte Zahl

$$\|A\| := \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

heißt zugeordnete Matrixnorm.

Bemerkung: Man kann zeigen, dass die zugeordnete Matrixnorm alle Eigenschaften von Def. 50.2 erfüllt und die kleinste aller Matrixnormen ist, die zu einer vorgegebenen Vektornorm kompatibel sind.

50.7 Beispiele

Man kann zeigen:

Vektornorm		zugeordnete Matrixnorm	
Betragssummennorm	$\|x\|_1$	Spaltensummennorm	$\|A\|_S$
Euklidische Norm	$\|x\|_2$	Spektralnorm	$\|A\|_2$
Maximumsnorm	$\|x\|_\infty$	Zeilensummennorm	$\|A\|_Z$

Matrixnormen sind nützlich zur Abschätzung von Eigenwerten.

50.8 Satz: (Eigenwertabschätzung mit Matrixnormen)

Ist λ Eigenwert von $A \in \mathbb{R}^{n\times n}$ und $\|A\|$ eine beliebige, zu einer Vektornorm kompatible Matrixnorm, so gilt:

$$|\lambda| \leq \|A\|.$$

Beweis:

Sei v ein Eigenvektor zu λ: $\Rightarrow |\lambda|\,\|v\| = \|\lambda v\| = \|Av\| \leq \|A\| \cdot \|v\|$

Da $v \neq 0$, gilt $\|v\| \neq 0$. Daher ist $|\lambda| \leq \|A\|$. $\square$.

50.9 Beispiel

$$\begin{aligned}
A &= \begin{pmatrix} 1 & 0,1 & -0,1 \\ 0 & 2 & 0,4 \\ -0,2 & 0 & 3 \end{pmatrix} \\
\|A\|_G &= 3 \cdot \max_{i,k} |a_{ik}| = 3 \cdot 3 = 9 \\
\|A\|_Z &= \max\{1,2\,;2,4\,;3,2\} = 3,2 \\
\|A\|_S &= \max\{1,2\,;2,1\,;3,5\} = 3,5 \\
\|A\|_F &= \sqrt{1^2 + 0,1^2 + (-0,1)^2 + 2^2 + 0,4^2 + (-0,2)^2 + 3^2} = \sqrt{14,22} \approx 3,77
\end{aligned}$$

$\|A\|_Z$ liefert die schärfste Abschätzung: $|\lambda| \leq \|A\|_Z \leq 3,2$

Tatsächlich gilt:

$$\begin{aligned} \lambda_1 &\approx 3,0060 \\ \lambda_2 &\approx 2,0078 \\ \lambda_3 &\approx 0,9862. \end{aligned}$$

Gibt es auch Abschätzungen für alle Eigenwerte?

50.10 Satz: (Satz von Gerschgorin)

(*a*) *Die Vereinigung aller Kreisscheiben*

$$K_i := \left\{ \mu \in \mathbb{C} \,\middle|\, |\mu - a_{ii}| \le \sum_{\substack{k=1\\k\neq i}}^{n} |a_{ik}| \right\}$$

enthält alle Eigenwerte von $A = (a_{ij}) \in \mathbb{R}^{n\times n}$.

(*b*) *Jede Zusammenhangskomponente aus m solcher Kreise enthält genau m Eigenwerte (der Vielfachheit nach gezählt).*

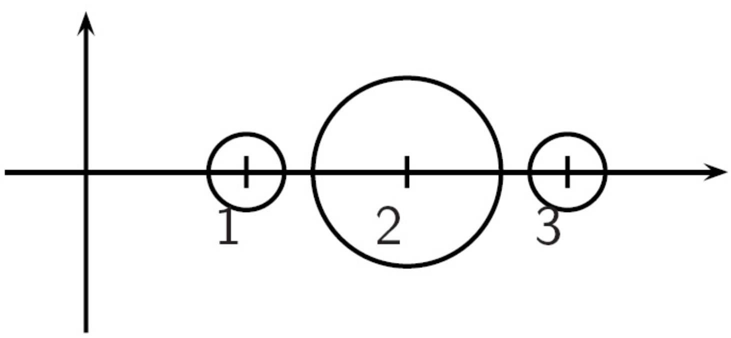
Beweis:

siehe Stoer / Burlisch: Einführung in die numerische Mathematik II, Springer, Berlin

50.11 Beispiel

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0,1 & -0,1 \\ 0 & 2 & 0,4 \\ -0,2 & 0 & 3 \end{pmatrix} \\ K_1 &= \{\mu \in \mathbb{C} \mid |\mu - 1| \le 0,2\} \\ K_2 &= \{\mu \in \mathbb{C} \mid |\mu - 2| \le 0,4\} \\ K_3 &= \{\mu \in \mathbb{C} \mid |\mu - 3| \le 0,2\} \end{aligned}$$

Sämtliche Eigenwerte liegen in $K_1 \cup K_2 \cup K_3$:



Da K_1, K_2, K_3 nicht überlappen, liegt nach (b) in jedem der Kreisscheiben genau ein Eigenwert. Ferner ist A invertierbar, da 0 außerhalb von $K_1 \cup K_2 \cup K_3$ liegt, also kein Eigenwert sein kann.

50.12 Korollar: *(Invertierbarkeit strikt diagonaldominanter Matrizen)*

*I*st $A \in \mathbb{R}^{n \times n}$ strikt diagonal dominant (d.h. $|a_{ii}| > \sum\limits_{\substack{k=1\\k\neq i}}^{n} |a_{ik}|$ für alle $i = 1, \ldots, n$), so ist A invertierbar.

Beweis:

Nach dem Satz von Gerschgorin liegt 0 außerhalb der Gerschgorinkreisscheiben, kann also kein Eigenwert sein. $\square$.

Kapitel 51

Numerische Berechnung von Eigenwerten und Eigenvektoren

51.1 Motivation

Die Berechnung der Eigenwerte von $A \in \mathbb{R}^{n \times n}$ über die Nullstellen von $\det(A - \lambda I)$ führt für $n \geq 5$ auf Polynome, für die keine analytischen Lösungsformeln existieren. Das macht numerische Verfahren notwendig.

51.2 Die einfache Vektoriteration (Potenzmethode, Von-Mises-Verfahren)

Idee: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und $u_0 \in \mathbb{R}^n$ ein beliebiger Startvektor. Lassen sich mit der Iteration

$$u_{k+1} = A \cdot u_k, \qquad k = 0, 1, \ldots$$

Aussagen über Eigenwerte und Eigenvektoren von A gewinnen?

Lösung: Seien $\lambda_1, \ldots, \lambda_n$ Eigenwerte und $v_1, \ldots, v_n$ die zugehörigen linear unabhängigen Eigenvektoren von A.
O.B.d.A. sei $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$.
Sei $u_0 = \sum_{i=1}^{n} \alpha_i v_i$.

$\Rightarrow u_1 := A \cdot u_0 = \sum_{i=1}^{n} \alpha_i A v_i = \sum_{i=1}^{n} \alpha_i \lambda_i v_i$

analog $u_k = \sum_{i=1}^{n} \alpha_i \lambda_i^k v_i = \lambda_1^k \cdot \left(\alpha_1 v_1 + \sum_{i=2}^{n} \frac{\lambda_i^k}{\lambda_1^k} \alpha_i v_i \right)$.

Falls $|\lambda_1| > |\lambda_2|$ (d.h. λ_1 ist dominanter Eigenwert), so konvergiert $\sum_{i=2}^{n} \frac{\lambda_i^k}{\lambda_1^k} \alpha_i v_i$ gegen 0 für $k \to +\infty$.

Falls der Startvektor u_0 „genügend allgemein" gewählt wurde (so dass $\alpha_1 \neq 0$), so konvergiert u_k mit geeigneter Normierung gegen den dominanten Eigenvektor v_1^* mit $\|v_1^*\| = 1, v_1^*$ parallel zu v_1. Die Konvergenz ist umso schneller, je kleiner die $\left|\frac{\lambda_i}{\lambda_1}\right|$ sind ($i = 2, \ldots, n$).

Der Rayleigh-Koeffizient $R_A(u_k) = \dfrac{< u_k, Au_k >}{< u_k, u_k >}$ konvergiert dann und ist gleich $\dfrac{< \alpha_1 v_1, \lambda_1 \alpha_1 v_1 >}{< \alpha_1 v_1, \alpha_1 v_1 >} = \lambda_1$.

Die Vektoriteration ist also ein einfaches Verfahren zur Approximation des dominanten Eigenwertes (und somit der Spektralnorm) und des dominanten Eigenvektors einer symmetrischen Matrix.

In der Praxis normiert man u_k nach jedem Iterationsschritt um zu vermeiden, dass u_k numerisch „explodiert“ (für $|\lambda_1| > 1$) bzw. gegen 0 geht (für $|\lambda_1| < 1$).

51.3 Beispiel

$$A = \begin{pmatrix} 1,04 & 0,72 \\ 0,72 & 1,46 \end{pmatrix}, \quad \text{Startvektor: } u_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = v_0 \text{ (bereits normiert)}$$

$$\begin{array}{lll}
u_1 = A \cdot v_0 = \begin{pmatrix} 1,04 \\ 0,72 \end{pmatrix} & v_1 = \dfrac{u_1}{|u_1|} \approx \begin{pmatrix} 0,8222 \\ 0,5682 \end{pmatrix} & R_A(v_0) = \dfrac{v_0^\top A v_0}{v_0^\top v_0} = v_0^\top \cdot u_1 \approx 1,04 \\
u_2 = A \cdot v_1 \approx \begin{pmatrix} 1,2649 \\ 1,4230 \end{pmatrix} & v_2 = \dfrac{u_2}{|u_2|} \approx \begin{pmatrix} 0,6645 \\ 0,7476 \end{pmatrix} & R_A(v_1) = v_1^\top \cdot u_2 \approx 1,8500 \\
u_3 = A \cdot v_2 \approx \begin{pmatrix} 1,2294 \\ 1,5699 \end{pmatrix} & v_3 = \dfrac{u_3}{|u_3|} \approx \begin{pmatrix} 0,6166 \\ 0,7873 \end{pmatrix} & R_A(v_2) = v_2^\top \cdot u_3 \approx 1,9906 \\
u_4 = A \cdot v_3 \approx \begin{pmatrix} 1,2081 \\ 1,5934 \end{pmatrix} & v_4 = \dfrac{u_4}{|u_4|} \approx \begin{pmatrix} 0,6042 \\ 0,7968 \end{pmatrix} & R_A(v_3) = v_3^\top \cdot u_4 \approx 1,9994 \\
u_5 = A \cdot v_4 \approx \begin{pmatrix} 1,2020 \\ 1,5984 \end{pmatrix} & v_5 = \dfrac{u_5}{|u_5|} \approx \begin{pmatrix} 0,6010 \\ 0,7992 \end{pmatrix} & R_A(v_4) = v_4^\top \cdot u_5 \approx 1,9999
\end{array}$$

Exakte Lösung für dominanten Eigenvektor bzw. Eigenwert: $v = \begin{pmatrix} 0,6 \\ 0,8 \end{pmatrix}, \lambda = 2$.

51.4 Das Jacobi-Verfahren

Ein einfaches und robustes Verfahren zur Bestimmung aller Eigenwerte und Eigenvektoren einer symmetrischen Matrix.

Grundidee:

(a) Wendet man auf eine symmetrische Matrix A eine orthogonale Transformation Q an, so haben $Q^\top AQ$ und A diesselben Eigenwerte.

(b) Mit Hilfe einer Sequenz $(Q_k)_{k=1,2...}$ von orthogonalen Matrizen transformiert man A auf Diagonalgestalt. Die Diagonalelemente geben die Eigenwerte an, und aus (Q_k) berechnet man die Eigenvektoren.

Beweis:

zu a.): A und $Q^\top AQ$ haben die selben charakteristischen Polynome, denn es gilt:

$$\begin{aligned}
P_{Q^\top AQ}(\lambda) &= \det\,(Q^\top AQ - \lambda I) \\
&= \det\,(Q^\top AQ - \lambda Q^\top Q)) \\
&= \det\,(Q^\top \cdot (A - \lambda I) \cdot Q) \\
&= \det\,(Q^\top) \cdot \det\,(A - \lambda I) \cdot \det\,(Q) \\
&= \det\,(\underbrace{Q^\top Q}_{I}) \cdot \det\,(A - \lambda I) \\
&= \det\,(A - \lambda I) \\
&= p_A(\lambda)
\end{aligned}$$

Also haben A und $Q^\top AQ$ diesselben Eigenwerte. □.

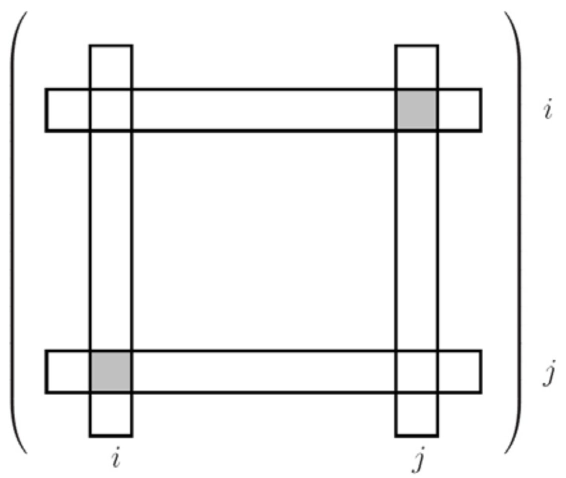
Einige Details zu b.)

Als orthogonale Transformationen verwendet man Rotationsmatrizen vom Typ

$$Q_k := Q(i,j,\varphi_k) \quad := \quad \begin{array}{cl}
\begin{pmatrix}
1 & & & & & & & & \\
& \ddots & & & \cdots & & & 0 & \\
& & 1 & & & & & & \\
& & & \cos\varphi_k & \cdots & \sin\varphi_k & & & \\
& & & & 1 & & & & \\
& \vdots & & \vdots & \ddots & \vdots & & \vdots & \\
& & & & 1 & & & & \\
& & & -\sin\varphi_k & \cdots & \cos\varphi_k & & & \\
& & & & & & 1 & & \\
& 0 & & & \cdots & & & \ddots & \\
& & & & & & & & 1
\end{pmatrix}
&
\begin{array}{l}
\leftarrow i\text{-te Spalte} \\[3em]
\leftarrow j\text{-te Spalte}
\end{array}
\\
\begin{array}{cc}
\uparrow & \uparrow \\
i\text{-te Spalte} & j\text{-te Spalte}
\end{array}
&
\end{array}$$

Man kann sich überlegen, dass $Q_k^\top AQ_k$ nur die i-ten und j-ten Zeilen und Spalten von A verändert:

$Q_k^T AQ_k =$



Den Rotationswinkel φ_k wählt man so, dass a_{ij} und a_{ji} zum Verschwinden gebracht werden. Führt man dies iterativ für alle Nichtdiagonalelemente von A durch, kann man zeigen, dass das Verfahren gegen eine Diagonalmatrix konvergiert:

$$Q_k^\top \cdot \ldots \cdot Q_2^\top Q_1^\top AQ_1Q_2 \cdot \ldots \cdot Q_k \to \ \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n) \text{ für } k \to +\infty.$$

Die m-te Spalte von $P := Q_1 \cdot Q_2 \cdot \ldots \cdot Q_k$ enthält somit eine Approximation an den Eigenvektor v_m zum Eigenwert λ_m. Da P orthogonal ist, erhält man insbesondere auch im Fall von mehrfachen Eigenwerten ein vollständiges ONS von Eigenvektoren. Gram-Schmidt-Orthonormierung ist daher nicht erforderlich.

Ein genauer Algorithmus zum Jacobi-Verfahren findet sich in H.R. Schwarz: „Numerische Mathematik“, Teubner, Stuttgart

Komplexität pro Zyklus (d.h. jedes Nichtdiagonalenelement wird einmal auf 0 transformiert) bei einer $n \times n$-Matrix $\approx 32n^3$ Multiplikationen. Typischerweise werden 6-8 Zyklen benötigt.

Fazit:
Eigenwertprobleme sind numerisch aufwändig! Es existieren wesentlich kompliziertere numerische Verfahren, insbesondere für nichtsymmetrische Matrizen.