Mathematik für Informatiker II
Teil D: Lineare Algebra

Prof. Dr. Joachim Weickert
LaTeXvon Sarah Diehl und Christian Spaniol

Sommersemester 2004

Inhaltsverzeichnis

40 Determinanten

40.1 Motivation

- Gibt es eine möglichst „aussagekräftige“ Abbildung, die jede Matrix $A \in K^{n \times n}$ auf eine Zahl aus K reduziert ?
- Die Determinante ist die „sinnvollste“ Art, eine solche Abbildung zu definieren.
- Sie kann axiomatisch fundiert werden und liefert wichtige Ausssagen:
  - über die Invertierbarkeit der Matrix A
  - über die Lösung des Gleichungssystems $Ax = b$
  - über das Volumen eines Parallelepipeds

Welche Forderungen soll die Determinante erfüllen?

40.2 Definition

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

a) det ist linear in jeder Zeile:

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

für $i = 1, \ldots, n$ und $\lambda, \mu \in K$.
(Man sagt auch: Die Determinante ist eine *Multilinearform*)

b) Ist rang $A < n$, so ist det $A = 0$

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

Beachte: Die Determinante ist nur für quadratische Matrizen definiert.

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

40.3 Satz (Eindeutigkeit der Determinante)

Zu jedem $n \in \mathbb{N}$ existiert genau eine Determinantenfunktion det $K^{n \times n} \to K$.

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

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

40.4 Satz (Determinante einer (2×2)-Matrix)

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

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

Beweis: Diese Abbildung erfüllt (a)-(c) der Definition 40.2:

a)

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

Die Linearität in der 2. Zeile zeigt man analog.

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

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

c)

$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot 1 - 0 \cdot 0 = 1 \qquad \square$$

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

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

40.5 Definition

Sei $A = (a_{ij}) \in K^{n\times n}$. Die aus einer $(n \times n)$-Determinante $D = \det A$ durch Streichung der i-ten Zeile und j-ten Spalte entstehende $(n-1)\times(n-1)$-Determinante D_{ij} nennen wir *Unterdeterminante* von D. Der Ausdruck $A_{ij} := (-1)^{i+j} D_{ij}$ heißt *algebraisches Komplement* des Elements a_{ij} in der Determinante D.

40.6 Beispiel

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

Kommen wir nun zur rekursiven Berechnung einer $(n \times n)$-Determinante:

40.7 Satz (Laplace'scher Entwicklungssatz)

Man kann eine $(n \times n)$-Determinante berechnen, indem man die Elemente einer Zeile (oder Spalte) mit ihren algebraischen Komplementen multipliziert und diese Produkte addiert.
Die Entwicklung nach der i-ten Zeile lautet also:

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij}$$

Entwickelt man nach der j-ten Spalte, ergibt sich:

$$\det A = \sum_{i=1}^{n} a_{ij} A_{ij}$$

40.8 Beispiel

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

$$\begin{aligned}\begin{vmatrix} 3 & 9 & 1 \\ 2 & 5 & 4 \\ -2 & 8 & 7 \end{vmatrix} &= -2 \cdot \begin{vmatrix} 9 & 1 \\ 8 & 7 \end{vmatrix} + 5 \cdot \begin{vmatrix} 3 & 1 \\ -2 & 7 \end{vmatrix} - 4 \cdot \begin{vmatrix} 3 & 9 \\ -2 & 8 \end{vmatrix} \\ &= -2(63 - 8) + 5(21 + 2) - 4(24 + 18) \\ &= -2 \cdot 55 + 5 \cdot 23 - 4 \cdot 42 = -163\end{aligned}$$

b) Entwicklung nach der 3. Spalte:

$$\begin{aligned}\begin{vmatrix} 3 & 9 & 1 \\ 2 & 5 & 4 \\ -2 & 8 & 7 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 2 & 5 \\ -2 & 8 \end{vmatrix} - 4 \cdot \begin{vmatrix} 3 & 9 \\ -2 & 8 \end{vmatrix} + 7 \cdot \begin{vmatrix} 3 & 9 \\ 2 & 5 \end{vmatrix} \\ &= 1(16 + 10) - 4(24 + 18) + 7(15 - 18) \\ &= 1 \cdot 26 - 4 \cdot 42 - 7 \cdot (-3) = -163\end{aligned}$$

Wie rechnet man mit Determinanten?

40.9 Rechenregeln für Determinanten

a) Transponieren verändert den Wert einer Determinante nicht: $\det A = \det A^\top$ (folgt aus dem Laplace'schen Entwicklungssatz, indem man die Entwicklung nach Zeilen und Spalten vertauscht).

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

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

c) Addiert man zu einer Zeile/Spalte das Vielfache einer anderen Zeile/Spalte, so bleibt die Determinante gleich:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 6 & 6 & 6 \end{vmatrix} \qquad \text{1. Zeile von 2. und 3. abgezogen}$$

d) Vertauscht man zwei Zeilen/Spalten, so ändert die Determinante ihre Vorzeichen.

e) Die Determinante von Dreiecksmatrizen ist das Produkt der Diagonalelemente:

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

(folgt durch rekursives Anwenden des Laplace'schen Entwicklungssatzes).

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 effizienter als der Laplace'sche Entwicklungssatz.

f) Für $A, B \in K^{n \times n}$ gilt:

$$\boxed{\det(A \cdot B) = \det\, A \cdot \det\, B}$$

g) Folgerung:

$$1 = \det\, I = \det(A \cdot A^{-1}) = \det A \cdot \det(A^{-1}) \Rightarrow \boxed{\det\, A^{-1} = \frac{1}{\det\, A}} \qquad \text{für } A \text{ invertierbar}$$

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

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

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

Wozu sind Determinanten nützlich?

40.10 Bedeutung der Determinanten

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

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

(vgl. Definition 40.2 (b))

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

Cramersche Regel: Ist $A = (a_{*1}, \ldots, a_{*n}) \in GL(n, K)$ und $b \in K^n$, so lässt sich die Lösung des linearen Gleichungssystems $Ax = b$ angeben durch

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

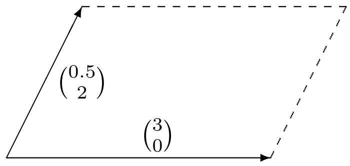
Beispiel:

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

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

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

c) $|\det A|$ ist das Volumen des durch die Spaltenvektoren von A aufgespannten Parallelepipeds:



$|\det \begin{pmatrix} 3 & 0,5 \\ 0 & 2 \end{pmatrix}| = |3 \cdot 2 - 0 \cdot 0,5| = |6|$ Parallelogrammfläche

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 ein Skalarprodukt, Längen- und Winkelbestimmung, Orthogonalität bereitstellen.

41.2 Definition

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
$0 := (0, \dots, 0)^\top$

b) Das (additive) *Inverse* $-v$ des Vektors v lautet
$-v := (-v_1, \dots, -v_n)^\top$

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

Mit der in Definition 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 $u, v, w \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$. 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: $u + 0 = 0 + u = u$
   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) $1v = v$

Beweis: Einfaches Anwenden von Definition 41.2

41.4 Definition

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

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

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) $v \cdot v \geq 0$
$v \cdot v = 0 \Leftrightarrow v = 0$

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) $v \cdot v = {v_1}^2 + {v_2}^2 + \dots + {v_n}^2 \geq 0$
Gleichheit gilt genau dann, wenn $v_1 = v_2 = \dots = v_n = 0$, d.h. wenn $v = 0$. □

41.7 Definition

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

$$|v| := \sqrt{v \cdot v} = \sqrt{{v_1}^2 + {v_2}^2 + \dots + {v_n}^2}$$

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

$$\mathrm{d}(u, v) := |u - v| = \sqrt{(u_1 - v_1)^2 + \dots + (u_n - v_n)^2}$$

41.8 Beispiel

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

$$\begin{aligned} |u| &= \sqrt{1+9+4+49} = \sqrt{63} = 3\sqrt{7} \\ \mathrm{d}(u,v) &= \sqrt{(1-0)^2+(3-7)^2+(-2-2)^2+(7-2)^2} \\ &= \sqrt{1+16+16+25} = \sqrt{58} \end{aligned}$$

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: siehe 42.5

41.10 Satz (Eigenschaften der euklidischen Norm)

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

a) $|v| \geq 0$

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

c) $|\alpha u| = |\alpha| \cdot |u|$

d) $|u+v| \leq |u| + |v|$ (Dreiecksungleichung)

Beweis: Wir zeigen nur (c) und (d)

c)

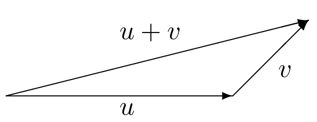
$$\begin{aligned} |\alpha u| &= \sqrt{(\alpha u_1)^2 + \cdots + (\alpha u_n)^2} = \sqrt{\alpha^2 \cdot ({u_1}^2 + \cdots + {u_n}^2)} \\ &= \sqrt{\alpha^2} \cdot \sqrt{{u_1}^2 + \cdots + {u_n}^2} = |\alpha| \cdot |u| \end{aligned}$$

d)

$$\begin{aligned} |u+v|^2 &= (u+v)\cdot(u+v) && \text{(Definition)} \\ &= u\cdot u + u\cdot v + v\cdot u + v\cdot v && \text{(41.6 b)} \\ &= |u^2| + 2u\cdot v + |v|^2 \\ &\leq |u|^2 + 2|u|\cdot|v| + |v|^2 && \text{(Cauchy-Schwarz'sche Ungleichung)} \\ &= (|u|+|v|)^2 \\ \Rightarrow |u+v| &\leq |u| + |v| \end{aligned}$$

□

Bedeutung der Dreiecksungleichung:
Die Länge zweier Dreieckseiten 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) $\mathrm{d}(u, v) \geq 0$

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

c) $\mathrm{d}(u, v) = \mathrm{d}(v, u)$

d) $\mathrm{d}(u, v) \leq \mathrm{d}(u, w) + \mathrm{d}(w, v)$

Beweis: Alle Eigenschaften ergeben sich als direkte Folgerung aus Satz 41.10

41.12 Definition

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

Beispiel:

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

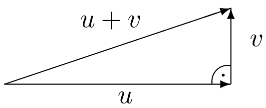
$$u \cdot v = -2 + 6 + 0 - 4 = 0$$

$$\Rightarrow u \text{ und } v \text{ sind orthogonal.}$$

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



Hypothenusenquadrat = Summe der Kathetenquadrate

Beweis:

$$\begin{aligned} |u + v|^2 &= (u + v) \cdot (u + v) \\ &= |u^2| + \underbrace{2u \cdot v}_{=0 \text{ (Orthogonalität)}} + |v|^2 \\ &= |u|^2 + |v|^2 \end{aligned}$$

$\square$

Beispiel:

$$u = \begin{pmatrix} -2 \\ 3 \\ 1 \\ 4 \end{pmatrix} \text{ und } v = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix} \text{ sind orthogonal; } u + v = \begin{pmatrix} -1 \\ 5 \\ 1 \\ 3 \end{pmatrix}$$

$$\left.\begin{aligned} |u|^2 &= 4 + 9 + 1 + 16 = 30 \\ |v|^2 &= 1 + 4 + 0 + 1 = 6 \\ |u + v|^2 &= 1 + 25 + 1 + 9 = 36 \end{aligned}\right\} \checkmark$$

41.14 Interpretation des euklidischen Produktes als Matrixmultiplikation

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

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

Beispiel:

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

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

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

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

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

i) Symmetrie: $\langle u, v\rangle = \langle v, u\rangle \quad \forall u, v \in V$

ii) Additivität: $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle \quad \forall u, v, w \in V$

iii) Homogenität: $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle \quad \forall u, v \in V, \forall \alpha \in \mathbb{R}$

iv) Nichtnegativität: $\langle v, v\rangle \geq 0 \quad \forall v \in V$
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 Definition 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 $\langle u, v\rangle := 3 \cdot u_1 \cdot v_1 + 5 \cdot u_2 \cdot v_2$ ein Skalarprodukt definiert.

Beweis:

i)

$$\begin{aligned}\langle u, v\rangle &= 3 \cdot u_1 \cdot v_1 + 5 \cdot u_2 \cdot v_2 \\ &= \langle v, u\rangle \qquad \forall u, v \in \mathbb{R}^2\end{aligned}$$

ii)

$$\begin{aligned}\langle u+v,w\rangle &= 3(u_1+v_1)w_1+5(u_2+v_2)w_2\\ &= 3(u_1\cdot w_1+5u_2\cdot w_2)+(3\cdot v_1\cdot w_1+5\cdot v_2\cdot w_2)\\ &= \langle u,w\rangle+\langle v,w\rangle \qquad \forall u,v,w,\in\mathbb{R}^2\end{aligned}$$

iii)

$$\begin{aligned}\langle \alpha u,v\rangle &= 3\alpha\cdot u_1\cdot v_1+5\alpha\cdot u_2\cdot v_2\\ &= \alpha\langle u,v\rangle \qquad \forall u,v\in\mathbb{R}^2,\forall\alpha\in\mathbb{R}\end{aligned}$$

iv)

$$\begin{aligned}\langle v,v\rangle &= \underbrace{3v_1^2}_{\geq 0}+\underbrace{5v_2^2}_{\geq 0}\geq 0 \qquad \forall v_1,v_2\in\mathbb{R}\\ klar: \langle v,v\rangle=0 &\Leftrightarrow v_1=v_2=0\end{aligned}$$

□

c) **Polynomräume**
Seien $p:=\sum_{k=0}^{n} a_k\cdot x^k,\quad q:=\sum_{k=0}^{n} b_k\cdot x^k$ Polynome vom Grad $\leq n$.
Dann wird mittels

$$\langle p,q\rangle := \sum_{k=0}^{n} a_k\cdot b_k$$

der reelle Vektorraum 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 $\langle f,g\rangle := \int_a^b f(x)\,g(x)\,dx$ zum Prä-Hilbert-Raum.

Lässt sich die euklidische Norm verallgemeinern?

42.4 Definition

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

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

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

Bemerkung: Ein vollständiger normierter Raum heißt auch *Banach-Raum.*

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

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

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

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

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

Die Parabel $ax^2 + bx + c$ hat also höchstens eine reelle Nullstelle. Die Diskriminante erfüllt also:

$$0 \geq b^2 - 4ac = 4\langle u, v \rangle^2 - 4\langle u, u \rangle \cdot \langle v, v \rangle$$

Daraus folgt die Behauptung.

$\square$

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 Definition 42.2 (iv)

(iii):

$$\begin{aligned} ||\alpha v|| &= \sqrt{\langle \alpha v, \alpha v \rangle} && \text{Definition} \\ &= \sqrt{\alpha \langle v, \alpha v \rangle} && \text{Homogenität} \\ &= \sqrt{\alpha \langle \alpha v, v \rangle} && \text{Symmetrie} \\ &= \sqrt{\alpha^2 \langle v, v \rangle} && \text{Homogenität} \\ &= |\alpha| \, ||v|| && \text{Definition} \end{aligned}$$

(iv):

$$\begin{aligned} ||u + v||^2 &= \langle u + v, u + v \rangle && \text{Definition} \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle && \text{Additivität (+ Symmetrie)} \\ &= ||u||^2 + 2\langle u, v \rangle + ||v||^2 && \text{Symmetrie, Definition} \\ &\leq ||u||^2 + 2||u|| \, ||v|| + ||v||^2 && \text{Cauchy-Schwarz-Ungleichung} \\ &= (||u|| + ||v||)^2 \end{aligned}$$

$\square$

42.7 Beispiele

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

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

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

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

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

$$\langle u, v\rangle = \frac{1}{9}u_1v_1 + \frac{1}{4}u_2v_2 \qquad \forall u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^2, \forall v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$$

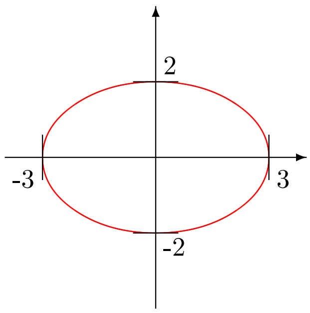
induziert die Norm

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

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 $\frac{u_1^2}{a^2} + \frac{u_2^2}{b^2} = 1$ 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

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

i) $d(u, v) \geq 0 \qquad \forall u, v \in V$

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

iii) $d(u, v) = d(v, u) \qquad \forall u, v \in V$

iv) $d(u, v) \leq d(u, w) + d(w, v) \qquad \forall u, v, w \in V$

(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 $d(u, v) := ||u-v|| \quad \forall u, v \in V$ einen metrischen Raum (V, d).

Beweis: einfache Folgerung aus Definition 42.4.

42.10 Beispiel:

Metrik auf $C[a, b]$

Für $f, g \in C[a, b]$ kann man durch

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

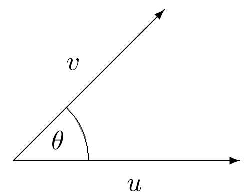
eine Metrik erklären. Ist z.B. $f(x) = 5x$, $g(x) = 2x - 1$, so lautet die Metrik auf $[0, 1]$:

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

43 Orthogonalität

43.1 Motivation

Das euklidische Produkt zweier Vektoren $u, v \in \mathbb{R}^2$, die einen Winkel θ einschließen, lässt sich auch schreiben als $u \cdot v = |u| \cdot |v| \cdot \cos\theta$.



Ist $\theta = \frac{\pi}{2}$, so ist $u \cdot v = 0$ und u und v nennt man 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

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 $\langle u, v\rangle = 0$ ist (und somit $\theta = \frac{\pi}{2}$).

43.3 Beispiele

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

$$\begin{aligned}
|u| &= (16+9+1+4)^{\frac{1}{2}} = \sqrt{30} \\
|v| &= (4+1+4+9)^{\frac{1}{2}} = \sqrt{18} \\
u \cdot v &= -8+3+2-6 = -9 \\
\cos\theta &= \frac{u \cdot v}{|u| \cdot |v|} = \frac{-9}{\sqrt{30}\sqrt{18}} \approx -0,3873 \\
\Rightarrow \theta &\approx 1,968 \; (\hat{=}\, 112,8\,^\circ)
\end{aligned}$$

b) $C[-1,1]$ mit Skalarprodukt $\langle u,v\rangle := \int_{-1}^{1} u(x)\, v(x)\mathrm{d}x$
Mit $u(x)=x$ und $v(x)=x^2$ ergibt sich
$\langle u,v\rangle = \int_{-1}^{1} x\cdot x^2\mathrm{d}x = \left[\frac{x^4}{4}\right]_{-1}^{1} = \frac{1}{4}-\frac{1}{4}=0.$
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 $\langle u,v\rangle = \int_{-1}^{1} u(x)\, v(x)\mathrm{d}x.$
Nach 43.3 (b) sind $u(x)=x$ und $v(x)=x^2$ orthogonal.

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

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

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:

$$u_1 \cdot u_2 = 0,\ u_1 \cdot u_3 = 0,\ u_2 \cdot u_3 = 0$$

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 euklidische Norm dividieren:

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

43.8 Satz (Koordinatendarstellung in Orthonormalbasis)

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

$$\boxed{u = \langle u, v_1\rangle v_1 + \langle u, v_2\rangle v_2 + \dots + \langle u, v_n\rangle v_n}$$

Beweis:

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

$$\Rightarrow \langle u, v_k\rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_k \right\rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_k\rangle = \alpha_k,$$

da $\langle v_i, v_k\rangle = \begin{cases} 1 & (i = k) \\ 0 & (sonst) \end{cases}$

Dies gilt für alle $k \in \{1, \dots, n\}$. □

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.

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

$$\Rightarrow \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{17}{5} \begin{pmatrix} -\frac{4}{5} \\ 0 \\ \frac{3}{5} \end{pmatrix} + \frac{31}{5} \begin{pmatrix} \frac{3}{5} \\ 0 \\ \frac{4}{5} \end{pmatrix}$$

Wie sieht Satz 43.8 aus, falls die Basis nicht normiert ist?

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 $u \in V$:

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

Beweis: Aus der Orthogonalbasis S erhält man durch Normierung die Orthonormalbasis S':

$$S' = \left\{\frac{v_1}{||v_1||}, \dots, \frac{v_n}{||v_n||}\right\}.$$

Nach Satz 43.8 gilt:

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

□

Bemerkung: Unter Zusatzvoraussetzungen gelten ähnliche Aussagen 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 $\alpha_1 = \dots = \alpha_n = 0$ folgt.

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

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

□

43.12 Beispiel

Aus Beispiel 43.7 wissen wir, dass

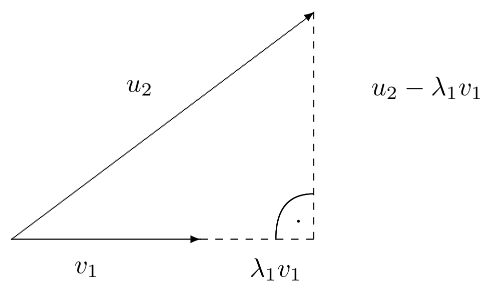
$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \right\}$$

eine orthonormale Menge im euklidischen Raum $\mathbb{R}^3$ ist. Nach Satz 43.11 sind dies drei 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.: $(V, \langle \cdot, \cdot \rangle)$ 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 $\text{span}\{v_1\}$.



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

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

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

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

Das Gram-Schmidt-Orthogonalisierungsverfahren liefert 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 Raums $\mathbb{R}^3$. Konstruiere anschließend eine Orthonormalbasis.
Lösung:

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

$\{v_1, v_2, v_3\}$ ist die gesuchte Orthogonalbasis des $\mathbb{R}^3$.
Die entsprechende Orthonormalbasis $\{q_1, q_2, q_3\}$ lautet

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

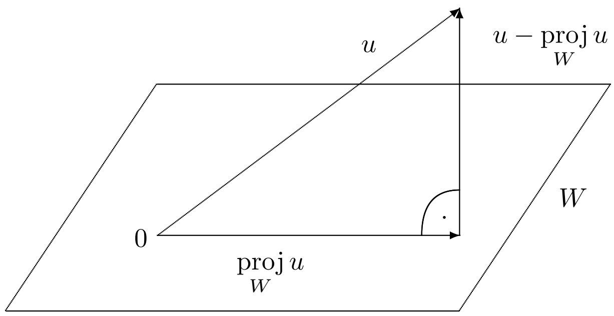
Ein weiteres „Abfallprodukt“ des Gram-Schmidt-Verfahrens ist der

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 $\langle u - \operatorname*{proj}_W u, w\rangle = 0 \quad \forall w \in W$.



Bemerkung: Ist $\{v_1, \ldots, v_n\}$ eine Ortho*normal*basis, gilt also $\operatorname*{proj}_W u = \sum_{i=1}^{n} \langle u, v_i\rangle v_i$.

Ist die Orthogonalprojektion eindeutig?

43.17 Definition

Sei W ein Unterraum eines Prä-Hilbert-Raums $(V, \langle\cdot,\cdot\rangle)$. Dann bezeichnet
$W^\perp := \{v \in V \mid \langle v, w\rangle = 0 \quad \forall w \in W\}$ das *orthogonale Komplement* von W.

43.18 Satz (Projektionssatz)

Sei W ein endlich dimensionaler Unterraum eines Prä-Hilbert-Raums $(V, \langle\cdot,\cdot\rangle)$. Dann besitzt jedes $v \in V$ eine *eindeutige* 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 $\oplus$ *direkte Summe.*)

Beweis: Es ist nur die Eindeutigkeit zu zeigen, die Existenz gilt wegen Satz 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

$$\begin{aligned} &\langle w_2' - w_2, w\rangle = \underbrace{\langle w_2', w\rangle}_{0,\ \text{da } w_2' \in W^\perp} - \underbrace{\langle w_2, w\rangle}_{0,\ \text{da } w_2 \in W^\perp} = 0 \quad \forall w \in W \\ &\Rightarrow w_2' - w_2 \in W^\perp. \end{aligned}$$

Wegen $w_2' - w_2 = w_1 - w_1'$ gilt auch $w_2' - w_2 \in W$.

$$\begin{aligned} &\Rightarrow \quad 0 = \langle \underbrace{w_2' - w_2}_{\in W}, \underbrace{w_2' - w_2}_{\in W^\perp}\rangle = ||w_2' - w_2||^2 \\ &\Rightarrow \quad 0 = w_2' - w_2 = w_1 - w_1', \text{ d.h. } w_1 = w_1',\ w_2 = w_2'. \end{aligned}$$

□

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 endlich dimensionaler Unterraum. Zu $v \in V$ ist dann $\operatorname*{proj}_W v$ die lokale Approximation von v in W. d.h.

$$||v - \operatorname*{proj}_W v|| < ||v - w|| \quad \forall w \in W \text{ mit } w \neq \operatorname*{proj}_W v.$$

Beweis:

$$\begin{aligned} ||v-w||^2 &= ||\underbrace{v - \operatorname*{proj}_W v}_{\in W^\perp} + \underbrace{\operatorname*{proj}_W v - w}_{\in W}||^2 \\ &\overset{\text{Pythagoras}}{=} ||v - \operatorname*{proj}_W v||^2 + ||\operatorname*{proj}_W v - w||^2 \\ &\geq ||v - \operatorname*{proj}_W v||^2 \end{aligned}$$

Gleichheit gilt nur, falls $w = \operatorname*{proj}_W v$. $\square$

43.20 Beispiel

$V = C[0, \frac{\pi}{2}]$ mit Skalarprodukt $\langle u, v\rangle := \int_1^{\frac{\pi}{2}} u(x)\, v(x) \mathrm{d}x$.
Bestimme Gerade, die die Funktion $u(x) = \sin x$ im Intervall $[0, \frac{\pi}{2}]$ optimal (bzgl. der induzierten Norm) approximiert.
Lösung: $W = \operatorname{span}\{1, x\}$ Unterraum aller Geraden.
Ges.: $\operatorname*{proj}_W u := \sum_{i=1}^{2} \lambda_i v_i$ mit $0 = \langle u - \sum_{i=1}^{2} \lambda_i v_i, v_k\rangle$ $(k = 1, 2)$.
In unserem Fall:

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

gibt Gleichungssystem für λ_1, λ_2:

$$\lambda_1\langle 1,1\rangle + \lambda_2\langle x, 1\rangle = \langle \sin x, 1\rangle$$

$$\lambda_1\langle 1,x\rangle + \lambda_2\langle x, x\rangle = \langle \sin x, x\rangle$$

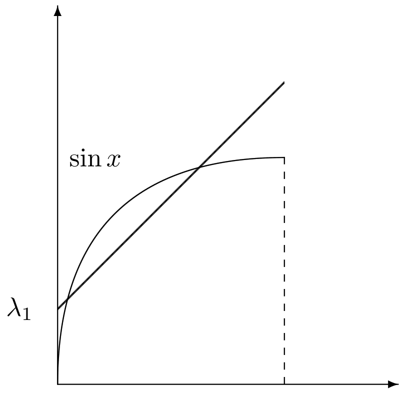
Mit

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

lautet das System

$$\begin{pmatrix} \frac{\pi}{2} & \frac{\pi^2}{8} \\ \frac{\pi^2}{8} & \frac{\pi^3}{24} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{aligned} \Rightarrow \lambda_1 &= 8\frac{\pi - 3}{\pi^2} \approx 0,11 \\ \lambda_2 &= 24\frac{4 - \pi}{\pi^3} \approx 0,66 \end{aligned}$$



Bestapproximation: $f(x) = \lambda_1 + \lambda_2 x$

44 Fourierreihen

44.1 Motivation

Ähnlich wie die Taylorreihe 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

Geg.:

- Vektorraum $V = C[0, 2\pi]$ mit Skalarprodukt $\langle u, v\rangle := \displaystyle\int_0^{2\pi} u(x)\, v(x) \mathrm{d}x$ und Norm $||u|| = \sqrt{\langle u, u\rangle}$
- Funktion $u \in V$
- endlich dimensionaler Unterraum $W = \operatorname{span}\{\underbrace{1}_{v_0}, \underbrace{\cos x}_{v_1}, \underbrace{\cos(2x)}_{v_2}, \dots, \underbrace{\cos(nx)}_{v_n}, \underbrace{\sin x}_{v_{n+1}}, \underbrace{\sin(2x)}_{v_{n+2}}, \dots, \underbrace{\sin(nx)}_{v_{2n}}\}$ trigonometrischen Polynome vom Grad $\leq n$.

Ges.: Koeffizienten $a_0, a_1, \dots, a_n, b_1, b_2, \dots, b_n$ 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 $f = \operatorname*{proj}_W u$ gegeben.
Man weist leicht nach, dass $\{1, \cos x, \dots, \sin(nx)\}$ orthogonal sind:

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

$$\int_0^{2\pi} \sin(kx)\cos(lx) \mathrm{d}x = 0$$

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

Nach Satz 43.16 lautet damit die Orthogonalprojektion

$$\begin{aligned} f &= \operatorname*{proj}_W u = \sum_{k=0}^{2n} \frac{\langle u, v_k \rangle}{||v_k||^2} v_k \\ &= \frac{1}{2\pi} \langle u, 1 \rangle + \sum_{k=1}^{n} \frac{1}{\pi} [\langle u, \cos(kx) \rangle \cos(kx) + \langle u, \sin(kx) \rangle \sin(kx)] \end{aligned}$$

Somit lauten die Fourierkoeffizienten

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

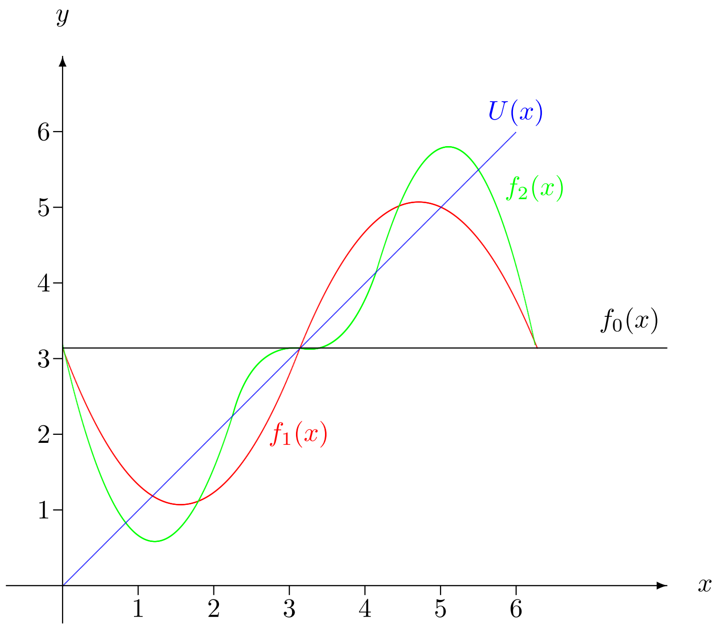
44.3 Beispiel

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

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

Das trigonometrische Approximationspolynom lautet also

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



$$
\begin{aligned}
u(x) &= x \\
f_0(x) &= \pi \\
f_1(x) &= \pi - 2\sin(x) \\
f_2(x) &= \pi - 2(\sin(x) + \frac{\sin(2x)}{2})
\end{aligned}
$$

44.4 Definition

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

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

mit

$$
\begin{aligned}
a_k &= \frac{1}{\pi}\int_0^{2\pi} u(x)\cos(kx)\mathrm{d}x \quad (k = 0, 1, \dots) \\
b_k &= \frac{1}{\pi}\int_0^{2\pi} u(x)\sin(kx)\mathrm{d}x \quad (k = 1, 2, \dots)
\end{aligned}
$$

Bemerkung: Falls u differenzierbar ist, kann man zeigen, dass die Fourierreihe punktweise gegen u konvergiert. An Sprungstellen zeigen die Fourierpolynome 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: niedrige Frequenzen
  a_k, b_k mit großem k: hohe 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 Stimmenübertragung)

- Ähnliche Bedeutung in der Bildverarbeitung:
  Grauwertbilder können als 2D-Signale aufgefasst werden.

- Signale und Bilder liegen meist diskret (gesampelt) vor. Dann verwendet man diskrete Fourierrepräsentationen, bei denen Integrale durch Summen ersetzt sind.

- Es existieren sehr schnelle Algorithmen 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.

- effizientesten 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.

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

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

$$O(n) := \{Q \in \mathbb{R}^{n\times n} \mid Q \text{ orthogonal}\}.$$

Was sind nun die schönen Eigenschaften?

45.3 Satz (Eigenschaften orthogonaler Matrizen)

Ist $Q \in O(n)$, 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:

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

c) Multiplikation mit Q erhält die euklidische Norm:

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

Man nennt Q daher auch *Isometrie*.

Beweis:

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

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

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

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

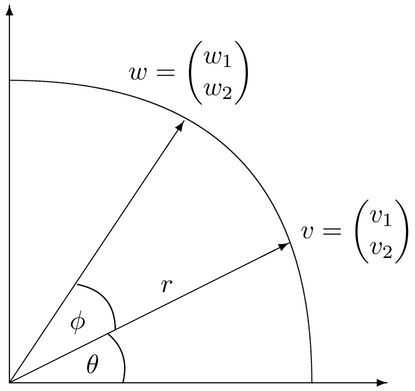
c) Folgt aus (b) mit $u = v$. $\square$

Bemerkung zu 45.3 (a): Es gilt sogar: Q orthogonal $\Leftrightarrow Q^{-1} = Q^\top$.

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 $-\theta$.

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

Somit ist Q orthogonal.

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

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

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

Q beschreibt eine *Spiegelung* an der 1. 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 $Q \in O(n)$, so gilt $|\det\, Q| = 1$.

Beweis:

$$1 = \det I = \det(QQ^\top) = \det Q \cdot \det(Q^\top) = (\det Q)^2 \qquad \square$$

Orthogonale Matrizen mit Determinante 1 sind noch einmal besonders ausgezeichnet.

45.6 Definition

$$SO(n) := O^+(n) := \{Q \in O(n) \,|\, \det Q = 1\}$$

45.7 Satz (Gruppeneigenschaft von $O(n)$ und $SO(n)$)

$O(n)$ und $SO(n)$ sind Untergruppen der allgemeinen linearen Gruppe

$$GL(n, \mathbb{R}) = \{A \in \mathbb{R}^{n\times n} \,|\, \text{A invertierbar}\}$$

bezüglich der Matrixmultiplikation. Man nennt $O(n)$ die *orthogonale Gruppe* und $SO(n)$ die *spezielle orthogonale Gruppe*.

Beweis: Übungsaufgabe

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

45.8 Wechsel zwischen Orthonormalbasen

Problem: Sei $\{v_1, \ldots, v_n\}$ Orthonormalbasis (ONB) des euklidischen Raums $\mathbb{R}^n$. Dann existiert zu jedem beliebigen Vektor $u \in \mathbb{R}^n$ eindeutig bestimmte Koeffizienten $a_1, \ldots, a_n$ 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 Orthonormalbasis $\{v_1, \ldots, v_n\}$.

Sei nun $\{w_1, \ldots, w_n\}$ eine weitere Orthonormalbasis des $\mathbb{R}^n$ und u habe den Koordinatenvektor

$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ bzgl. $\{w_1, \ldots, w_n\}$.

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

Lösung:

$$\begin{aligned}
u &= \sum_{k=1}^{n} a_k v_k \\
&= \sum_{k=1}^{n} a_k \Big(\sum_{i=1}^{n} ({v_k}^\top w_i) w_i \Big) && v_k \text{ durch } \{w_1, \dots, w_n\} \text{ ausgedrückt} \\
&= \sum_{i=1}^{n} \Big(\sum_{k=1}^{n} a_k ({v_k}^\top w_i) \Big) w_i && \text{Vertauschung der endlichen Summationen} \\
&=: \sum_{i=1}^{n} b_i w_i \text{ mit } b_i = \sum_{k=1}^{n} ({v_k}^\top w_i) a_k
\end{aligned}$$

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

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

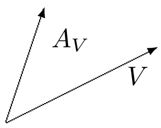
$$\Rightarrow \boxed{Q = \begin{pmatrix} {w_1}^\top \\ \hline \vdots \\ \hline {w_n}^\top \end{pmatrix} (v_1, \dots, v_n)}$$

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.

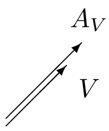
46 Eigenwerte und Eigenvektoren

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, für die v und Av parallel sind?



$$A_V = \lambda V, \quad \lambda \in \mathbb{R}$$

46.2 Definition

Sei $A \in \mathbb{R}^{n \times n}$. Ein von 0 verschiedener (!) Vektor $v \in \mathbb{R}^n$ heißt *Eigenvektor* von A, falls 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- bzw. Eigenwertprobleme sind wichtig in der Statik, Elektrotechnik, Maschinenbau, Biologie, Informatik und den Wirtschaftswissenschaften. Oft beschreiben 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

Die Matrix $A = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ hat einen Eigenvektor $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, denn

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

Der zugehörige Eigenwert ist 3.

Wie kann man Eigenwerte 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 nicht triviale Lösungen von $(A - \lambda I)v = 0$.
Sie existieren nur für $\operatorname{rang}(A - \lambda I) < n$, d.h. für

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

Für $A \in \mathbb{R}^n$ 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

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

46.7 Bemerkungen

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 eine $(n \times n)$-Matrix A als Nullstellen des charakteristischen Polynoms, kann dies für $n \geq 3$ unangenehm werden. Für $n \geq 5$ ist dies im Allgemeinen 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 *kein* Eigenwert 0 auftritt.

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

46.8 Definition

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 Beispiel

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

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 Diagonalelemente.
Für $A = (a_{ij})$ folgt aus

$$0 = \det(A - \lambda I) = (a_{11} - \lambda) \cdot ... \cdot (a_{nn} - \lambda)$$

dass die Eigenwerte durch $\lambda_1 = a_{11}, \dots, \lambda_n = a_{nn}$ 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).

Wie kann man Eigenvektoren berechnen?

46.12 Bestimmung der Eigenvektoren

Anmerkung: Ein Eigenwert λ der Matrix A sei bekannt.
Dann sind die Eigenvektoren zu λ die nichttrivialen Lösungen von

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

Eigenvektoren sind *nicht* eindeutig bestimmt:
Mit v ist auch αv für $\alpha \neq 0$ Eigenvektor ($\alpha \in \mathbb{R}$).
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: $0 = \det(A - \lambda I) = \dots = (\lambda - 1) \cdot (\lambda - 2)^2 \Rightarrow \lambda_1 = 2,\ \lambda_2 = 1$

i) Eigenraum zu $\lambda_1 = 2$ ist Lösungsraum von

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

drei 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 Eigenraum ist z.B. $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$.

ii) Eigenraum zu $\lambda_2 = 1$ ist die Lösungsmenge 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:

$$x_2 - x_3 = 0 \Rightarrow x_2 = x_3 = s$$

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 zugehörigem Eigenvektor v. Dann ist λ^k Eigenwert von A^k mit zugehörigem Eigenvektor v.

Beweis:

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

46.15 Beispiel

Sei $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$. Dann hat A die Eigenwerte 1 und 2 (vgl. Beispiel 46.13).
Somit hat A^7 die Eigenwerte $1^7 = 1$ und $2^7 = 128$.

47 Eigenwerte und Eigenvektoren symmetrischer Matrizen

47.1 Motivation

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

47.2 Satz (Eigenwerte und Eigenvektoren symmetrischer Matrizen)

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

a) Sämtliche Eigenwerte sind reell.

b) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.

Beweis:

a) Sei $z = x + iy \in \mathbb{C}$ und $\overline{z}$ die komplexe konjugierte Zahl $x - iy$.
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 Eigenwerten λ_1, λ_2.

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

$$\Rightarrow \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} {v_1}^\top v_2 = 0$$

Also sind v_1 und v_2 orthogonal $\square$

47.3 Beispiel

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

Eigenwerte:

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

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

Eigenvektor zu $\lambda_1 = 0$:
$(A - \lambda_1 I)v = 0$

$\Rightarrow \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ zwei linear abhängige Gleichungen

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

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 $\lambda_2 = s$:
$(A - \lambda_2 I)v = 0$

$\Rightarrow \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ zwei linear abhängige Gleichungen

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

Eigenraum: $\left\{ \begin{pmatrix} s \\ 2s \end{pmatrix} \middle| \, s \in \mathbb{R} \right\}$.

Eigenvektor: z.B. $v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

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 47.2 hat A ein Orthonormalsystem von Eigenvektoren $v_1, \ldots, v_n$ mit zugehörigen (nicht notwendigerweise verschiedenen) Eigenwerten $\lambda_1, \ldots, \lambda_n$. Dann ist

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

mit der orthogonalen Matrix $Q = (v_1 | \ldots | v_n)$ und der Diagonalmatrix

$$\Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

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 $Q = (v_1| \dots |v_n)$ eine orthogonale Matrix. Die k-te Spalte von $Q^\top AQ$ lautet:

$$Q^\top \underbrace{Av_k}_{\lambda_k v_k} = \lambda_k Q^\top v_k \overset{\text{Orthogonalität}}{=} \lambda_k e_k \text{ mit } e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow \text{ k-te Stelle} \\ \\ \\ \\ \end{matrix}$$

Somit ist

$$Q^\top AQ = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \Lambda. \qquad \square$$

47.5 Bemerkungen

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

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

(Durch den Übergang in das durch $Q = (v_1| \dots |v_n)$ 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 $v_1, \dots, v_n$ Eigenvektoren von A sind, denn:

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

47.6 Beispiel

Wir bestätigen Satz 47.5, indem wir $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ auf Diagonalgestalt transformieren.

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 $Q = (v_1|v_2) = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$ ergibt sich:

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

wie nach Satz 47.5 (a) zu erwarten war.

48 Quadratische Formen und positiv definite Matrizen

48.1 Motivation

- Charakterisierung einer wichtigen Klasse symmetrischer Matrizen (Anwendungen: Physik, Computergrafik)
- Verhalten quadratischer Funktionen in mehreren Variablen untersuchen
- Wichtige Klassen geometrischer Kurven/Flächen kennenlernen

48.2 Definition

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

$$x^\top A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

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

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

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

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

erfüllen, heißt *Quadrik.*

48.3 Beispiel

a) $7{x_1}^2 + 6{x_2}^2 + 5{x_3}^2 - 4x_1x_2 + 2x_2x_3$

$$\begin{aligned} &= 7{x_1}^2 + 6{x_2}^2 + 5{x_3}^2 - 2x_1x_2 - 2x_2x_1 + x_2x_3 + x_3x_2 \\ &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & 1 \\ 0 & 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

ist eine quadratische Form.

b) $q(x) = 5{x_1}^2 - 3{x_2}^2 + 4x_1x_2 - 7x_2 + 3$ 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 $n = 3$ entstehen Ellipsoide, Hyperboloide, Paraboloide, ...

48.4 Definition

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

- *positiv definit*, falls $\lambda_i > 0\ \forall i$
- *positiv semidefinit*, falls $\lambda_i \geqq 0\ \forall i$
- *negativ definit*, falls $\lambda_i < 0\ \forall i$
- *negativ semidefinit*, falls $\lambda_i \leqq 0\ \forall i$
- *indefinit*, falls $\exists i, j$ mit $\lambda_i \lambda_j < 0$

Es besteht ein enger Zusammenhang zwischen Definitheit, quadratischen Formen und Determinanten von Untermatrizen.

48.5 Satz (Charakterisierung positiv definiter Matrizen)

Es sei $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ symmetrisch. Dann ist A positiv definit genau dann, wenn

$$x^\top A x > 0 \quad \forall x \in \mathbb{R}^n \text{ mit } x \neq 0$$

Beweis: „⇒“: Es sei $\{v_1, \ldots, v_n\}$ Orthonormalbasis von Eigenvektoren von A mit Eigenwerten $\lambda_1, \ldots, \lambda_n$. Dann gilt für $x \in \mathbb{R}^n$, $x = \sum_{i=1}^n x_i v_i$ $(x \neq 0)$:

$$\begin{aligned}
x^\top A x &= \left(\sum_{i=1}^n x_i v_i\right)^\top A \left(\sum_{j=1}^n x_j v_j\right) \\
&= \left(\sum_{i=1}^n x_i v_i\right)^\top \left(\sum_{k=1}^n \lambda_k v_k v_k^\top\right)\left(\sum_{j=1}^n x_j v_j\right) \\
&= \sum_{i,j,k=1}^n \lambda_k x_i x_j \underbrace{(v_i^\top v_k)}_{0 \text{ für } i \neq k} \underbrace{(v_k^\top v_j)}_{0 \text{ für } k \neq j} = \sum_{k=1}^n \lambda_k {x_k}^2 > 0
\end{aligned}$$

„⇐“: Sei A nicht positiv definit. Dann existiert ein Eigenvektor $v \neq 0$ von A mit zugehörigem Eigenwert $\lambda \leq 0$. Damit ist

$$v^\top A v = v^\top \lambda v = \underbrace{\lambda}_{\leq 0} \underbrace{v^\top v}_{>0} \leq 0$$

im Widerspruch zu $x^\top A x > 0\ \forall x \neq 0$. □

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*

$$\begin{vmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{vmatrix}$$

für $k = 1, \ldots, n$ 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

$A = \begin{pmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{pmatrix}$ Ist A positiv definit?
Hauptminoren:

$$\begin{aligned} |2| &= 2 > 0 \\ \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} &= 4 - 1 = 3 > 0 \\ \begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} &= -3 \begin{vmatrix} -1 & -3 \\ 2 & 4 \end{vmatrix} - 4 \begin{vmatrix} 2 & -3 \\ -1 & 4 \end{vmatrix} + 9 \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} \\ &= -3(-4+6) - 4(8-3) + 9 \cdot 3 = -6 - 20 + 27 = 1 > 0 \end{aligned}$$

$\Rightarrow A$ positiv definit.

Ähnlich wie zu jeder nichtnegativen Zahl eine Wurzel existiert, gibt es auch die „Wurzel“ aus einer positiv semidefiniten Matrix.

48.8 Satz (Wurzel einer positiv semidefiniten Matrix)

Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und positiv semidefinit. Dann existiert eine 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 Orthonormalbasis der zugehörigen Eigenvektoren.

Dann gilt mit $Q = (v_1|v_2|\dots|v_n)$ und $\Lambda := \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{vmatrix}$, dass $A = Q\Lambda Q^\top$.

Wir setzen $\Lambda^{\frac{1}{2}} := \begin{vmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{\lambda_n} \end{vmatrix}$ und $B := Q\Lambda^{\frac{1}{2}}Q^\top$.

Dann ist $B^2 = Q\Lambda^{\frac{1}{2}} \underbrace{Q^\top Q}_{I} \Lambda^{\frac{1}{2}} Q^\top = Q\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}Q^\top = Q\Lambda Q^\top = A$ □

Positiv und negativ (semi-)definite Matrizen spielen einer wichtige Rolle beim Nachweis von Maxima/Minima von Funktionen mit mehreren Variablen (→ Kapitel E).

Gibt es obere und untere Schranken für Werte quadratischer Formen?

48.9 Definition (Rayleigh - Quotient)

Es sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und $x \in \mathbb{R}^n$, $x \neq 0$. 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 Eigenwerten $\lambda_1 \geq \ldots \geq \lambda_n$ und zugehörigen orthonormierten Eigenvektoren $v_1, \ldots, v_n$. Dann gilt:

a) $\lambda_n \leq R(x) \leq \lambda_1$

b) Diese Grenzen werden tatsächlich angenommen:

$$\lambda_1 = \max_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} R(x),\ \lambda_n = \min_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} R(x)$$

Beweis:

a) Aus $x = \sum_{i=1}^n x_i v_i$ folgt $x^\top x = \sum_{i=1}^n {x_i}^2$
sowie $Ax = \sum_{i=1}^n x_i A v_i = \sum_{i=1}^n x_i \lambda_i v_i$.

Damit gilt:

$$\begin{aligned} x^\top Ax &= \Big(\sum_{i=1}^n x_i v_i\Big)^\top \Big(\sum_{j=1}^n x_j \lambda_j v_j\Big) \\ &= \sum_{i,j=1}^n \lambda_j x_i x_j v_i^\top v_j \\ &= \sum_{i=1}^n \lambda_i {x_i}^2 \\ \mathrm{R}(x) &= \frac{\sum_{i=1}^n \lambda_i {x_i}^2}{\sum_{i=1}^n {x_i}^2} \begin{cases} \leq \lambda_1 \\ \geq \lambda_n \end{cases} \end{aligned}$$

b) Setzt man $x = v_k$, dann ist

$$\mathrm{R}(v_k) = \frac{v_k^\top A v_k}{v_k^\top v_k} = \frac{\lambda_k v_k^\top v_k}{v_k^\top v_k} = \lambda_k$$

Speziell: $\mathrm{R}(v_n) = \lambda_n$, $\mathrm{R}(v_1) = \lambda_1$ □

49 Quadriken

49.1 Motivation

Quadriken (→ Definition 48.2) stellen eine wichtige Klasse geometrischer Objekte dar, mit Anwendungen in Computergrafik, Physik u.a.
Ziel: gegebene Quadrik auf einfache Form transformieren, so dass sich ihre geometrische Gestalt unmittelbar ablesen lässt.
Geometrische Interpretation von Eigenwerten symmetrischer Matrizen.

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 Orthonormalbasis $\{v_1, \dots, v_n\}$ aus Eigenvektoren mit $\det\,(v_1|\dots|v_n) = 1$ (falls $\det\,(v_1|\dots|v_n) = -1$, ersetzt man v_1 durch $-v_1$).

Mit $Q = (v_1|\dots|v_n)$ gilt dann $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = Q^\top AQ$

und aus $x^\top Ax + b^\top x + c = 0$ folgt dann

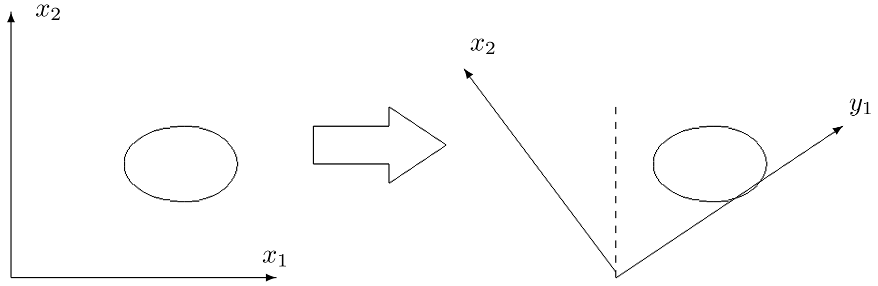
$$\underbrace{x^\top Q}_{(Q^\top x)^\top} \Lambda Q^\top x + \underbrace{b^\top Q}_{(Q^\top b)^\top} Q^\top x + c = 0$$

Mit $y := Q^\top x$, $\tilde{b} := Q^\top b$ ergibt sich daher

$$y^\top \Lambda y + \tilde{b}^\top y + c = 0$$

beziehungsweise ausgeschrieben (gemischte quadratische Terme weggefallen):

$$\lambda_1 {y_1}^2 + \dots + \lambda_n {y_n}^2 + \tilde{b}_1 y_1 + \dots + \tilde{b}_n y_n + c = 0$$



Schritt 2: Elimination linearer Terme (soweit möglich)

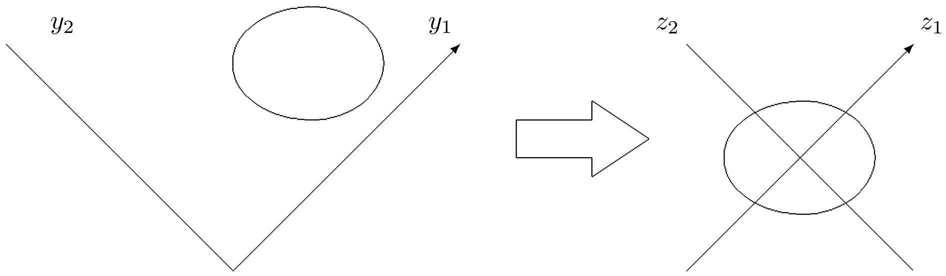
Durch Translation des Koordinatensystems kann erreicht werden, dass $\lambda_k {y_k}^2$ und $\tilde{b}_k y_k$ nicht zugleich vorkommen (für jedes k).
Es sei dazu o.B.d.A $\lambda_i \neq 0$ für $i = 1, \ldots, r$ sowie $\lambda_{r+1} = \cdots = \lambda_n = 0$. Dann wird für $i = 1, \ldots, r$ der lineare Term $\tilde{b}_i y_i$ durch die quadratische Ergänzung eliminert:

$$\begin{aligned} z_i &:= y_i + \frac{\tilde{b}_i}{2\lambda_i} \quad (i = 1, \ldots, r) \\ z_i &:= y_i \quad (i = r+1, \ldots, n) \end{aligned}$$

Damit erhält man:

$$\lambda_1 {z_1}^2 + \cdots + \lambda_r {z_r}^2 + \tilde{b}_{r+1} z_{r+1} + \cdots + \tilde{b}_n z_n + \tilde{c} = 0$$

mit $\tilde{c} = c - \sum_{i=1}^{r} \frac{{\tilde{b}_i}^2}{4\lambda_i}$ mit $r = \text{rang } A$.

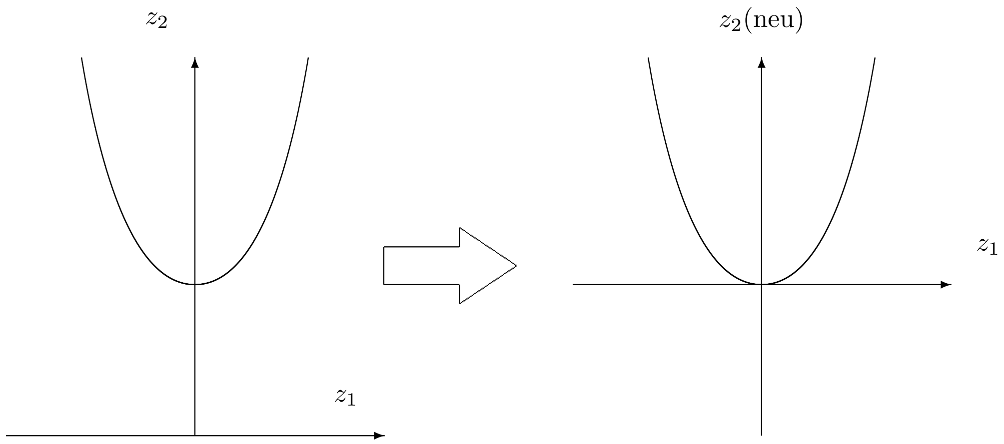


Schritt 3: Elimination der Konstanten (falls möglich)

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

$$z_n \mapsto z_n - \frac{\tilde{c}}{\tilde{b}_n}$$

(eine zusätzliche Translation des Koordinatensystems)



Resultat: Normalform der Quadrik

Darstellung in einem Koordinatensystem, in dem möglichst viele Koeffizienten verschwinden.

$$\begin{array}{lll} \text{Für } r := \text{rang } A = n: & & \lambda_1 {z_1}^2 + \cdots + \lambda_n {z_n}^2 + d = 0 \\ \text{Für } r < n: & \text{entweder} & \lambda_1 {z_1}^2 + \cdots + \lambda_r {z_r}^2 + e_{r+1} z_{r+1} + \cdots + e_n z_n = 0 \\ & \text{oder} & \lambda_1 {z_1}^2 + \cdots + \lambda_r {z_r}^2 + d = 0 \end{array}$$

49.3 Beispiel

Die Quadrik

$$q(x) = 5{x_1}^2 - 4x_1x_2 + 8{x_2}^2 + \frac{20}{\sqrt{5}}x_1 - \frac{80}{\sqrt{5}}x_2 + 4 = 0$$

soll auf Normalform gebracht werden.

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

$$A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}, b = \frac{1}{\sqrt{5}}\begin{pmatrix} 20 \\ -80 \end{pmatrix}, c = 4$$

1. Schritt (Hauptachsentransformation von A)
Eigenwerte $\lambda_1 = 9,\ \lambda_2 = 4$
Eigenvektoren $v_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $v_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ -1 \end{pmatrix}$

$$Q = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 & -2 \\ 2 & -1 \end{pmatrix} \quad \text{mit } \det Q = 1$$

$$\text{mit } \Lambda = Q^\top AQ = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix} \text{ und } \tilde{b} = Q^\top b = \begin{pmatrix} -36 \\ 8 \end{pmatrix}$$

ergibt sich für $y = Q^\top x$:

$$9{y_1}^2 + 4{y_2}^2 - 36y_1 + 8y_2 + 4 = 0$$

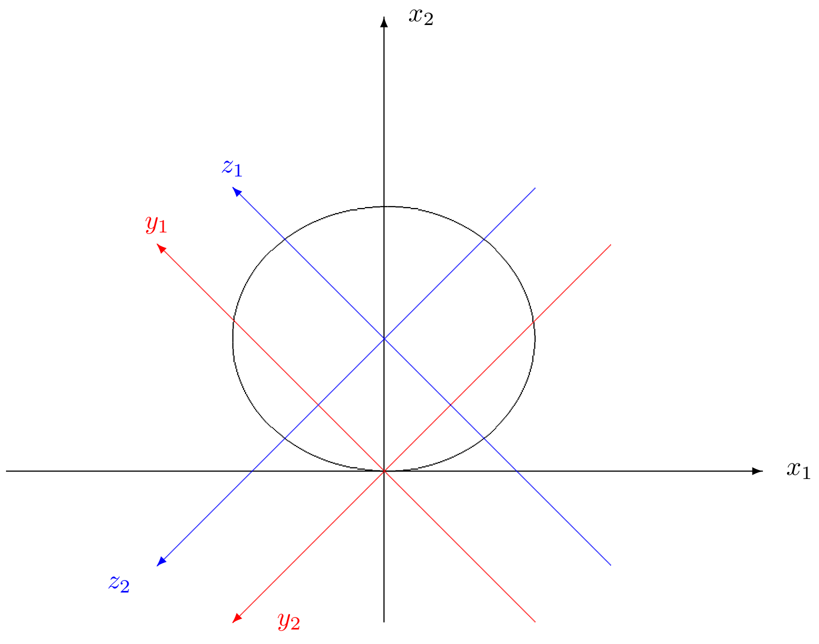
2. Schritt (Elimination linearer Terme)

$$(9{y_1}^2 - 36y_1 + 36) + (4{y_2}^2 + 8y_2 + 4) + (4 - 36 - 4) = 0$$

$$9({y_1}^2 - 4y_1 + 4) + 4({y_2}^2 + 2y_2 + 1) - 36 = 0$$

also mit $z_1 := y_1 - 2$ und $z_2 := y_2 + 1$:

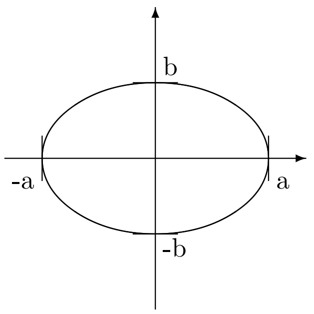
$$\begin{aligned} 9{z_1}^2 + 4{z_2}^2 &= 36 \\ \Rightarrow \frac{{z_1}^2}{4} + \frac{{z_2}^2}{9} &= 1 \qquad \text{Ellipse mit Halbachsen 2 und 3} \end{aligned}$$



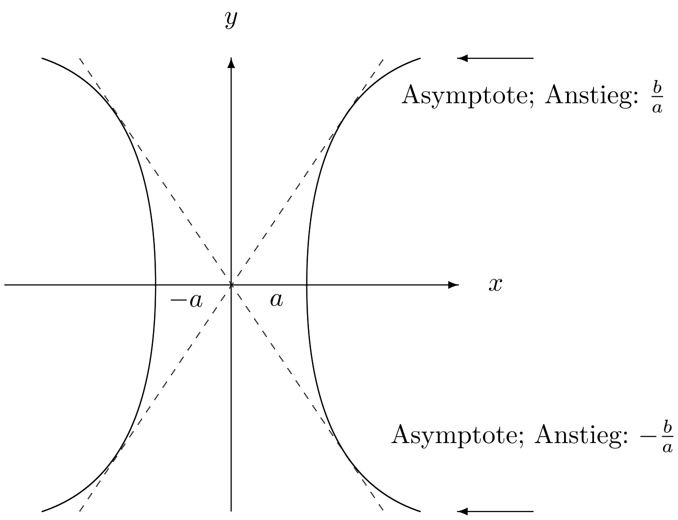
49.4 Normalform der Quadriken im $\mathbb{R}^2$ (Kegelschnitte)

(i) **rang $\boldsymbol{A} = 2$** (alle Eigenwerte $\neq 0$)

a) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ *Ellipse*



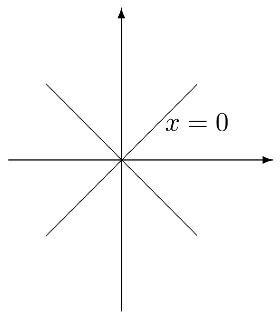
b) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$ *Hyperbel*



c) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$ *leere Menge*

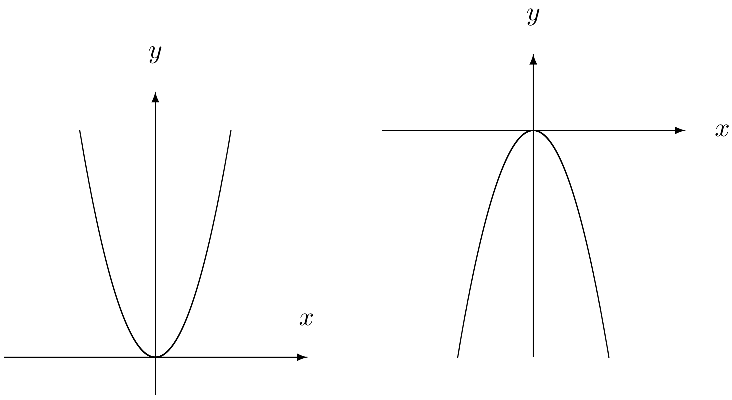
d) $x^2 + a^2y^2 = 0$, $a \neq 0$ *Punkt (0,0)*

e) $x^2 - a^2y^2 = 0$, $a \neq 0$ *Geradenpaar* $y = \pm\frac{1}{a}x$

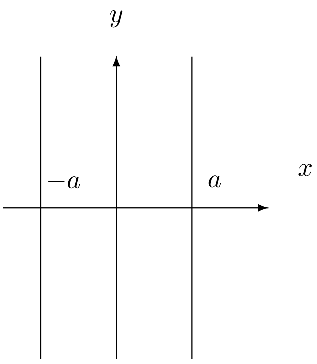


(ii) **rang $\boldsymbol{A} = 1$** (ein Eigenwert $= 0$)

a) $x^2 - 2py = 0$ *Parabel*

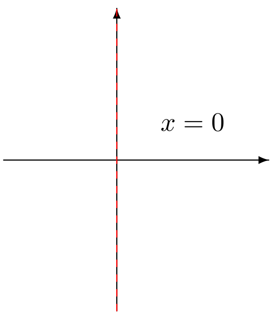


b) $x^2 - a^2 = 0$ *parallele Geraden* $x = \pm a$



c) $x^2 + a^2 = 0$ *leere Menge*

d) $x^2 = 0$ *„Doppelgerade“* $x = 0$

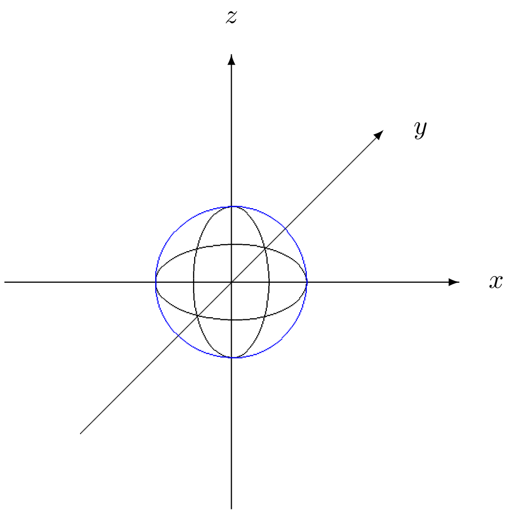


(iii) **rang $\boldsymbol{A} = 0$** (alle Eigenwerte $= 0$)
$b_1 x + b_2 y + c = 0$ *Gerade*

49.5 Normalform der Quadriken im $\mathbb{R}^3$

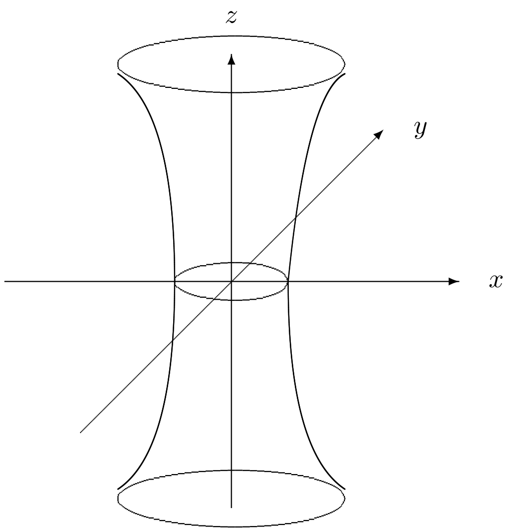
(i) **rang $\boldsymbol{A} = 3$** (alle Eigenwerte $\neq 0$)

a) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$ *Ellipsoid*



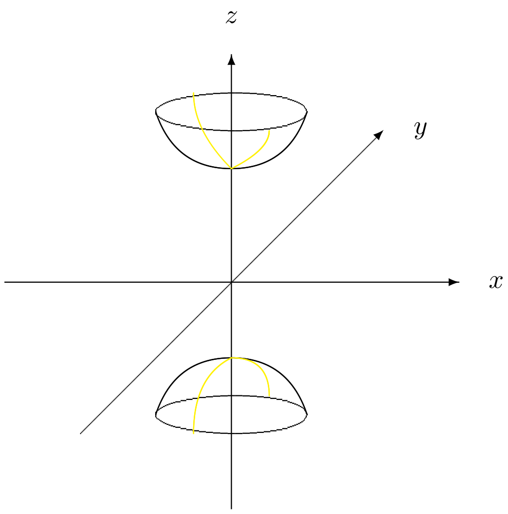
b) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} + 1 = 0$ *leere Menge*

c) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0$ *einschaliges Hyperboloid*



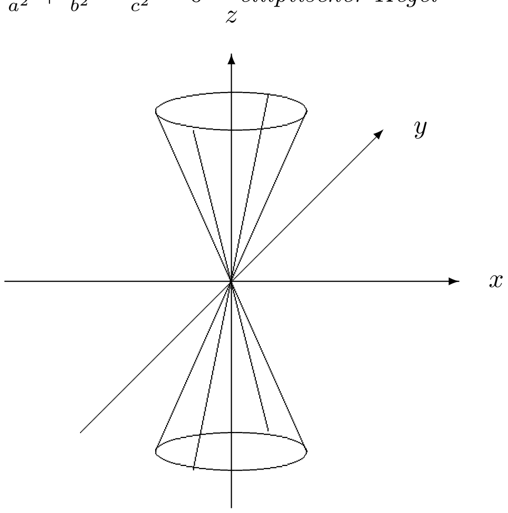
Ellipse in einer Ebene (parallel zur x-y-Ebene)
Hyperbel in zwei Ebenen $(x - z, y - z)$

d) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} + 1 = 0$ *zweischaliges Hyperboloid*



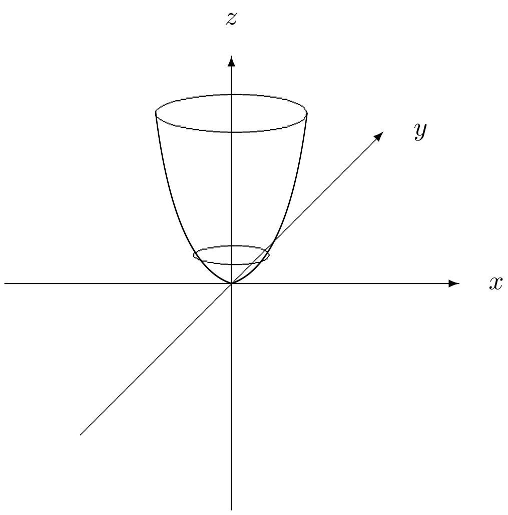
e) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$ *Punkt (0,0,0)*

f) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ *elliptischer Kegel*



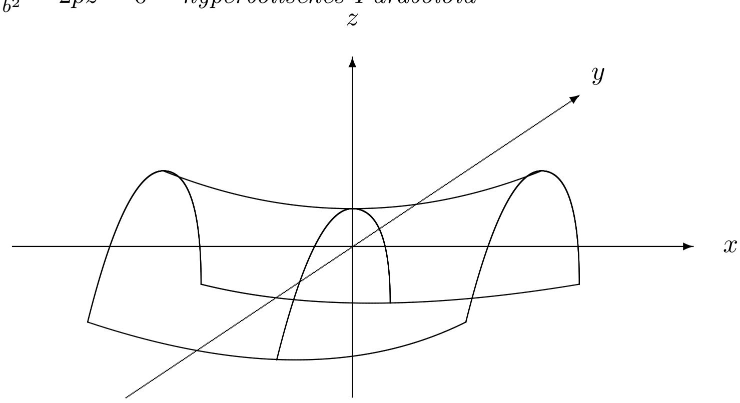
(ii) **rang $\boldsymbol{A} = \mathbf{2}$** (ein Eigenwert $= 0$)

a) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 2pz = 0$ *elliptisches Paraboloid*



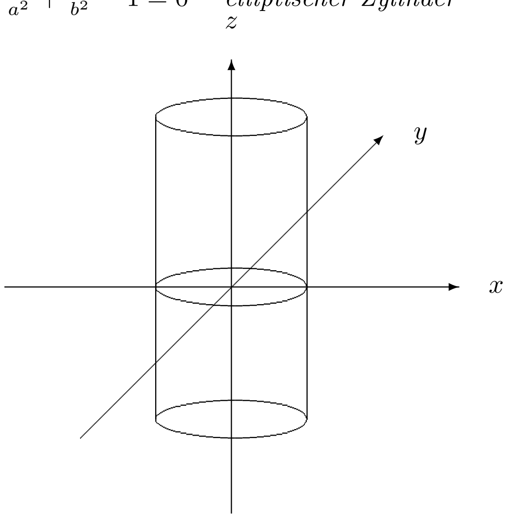
in einer Ebene (parallel zur x-y-Ebene) Ellipse
in zwei Ebenen Parabeln

b) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 2pz = 0$ *hyperbolisches Paraboloid*

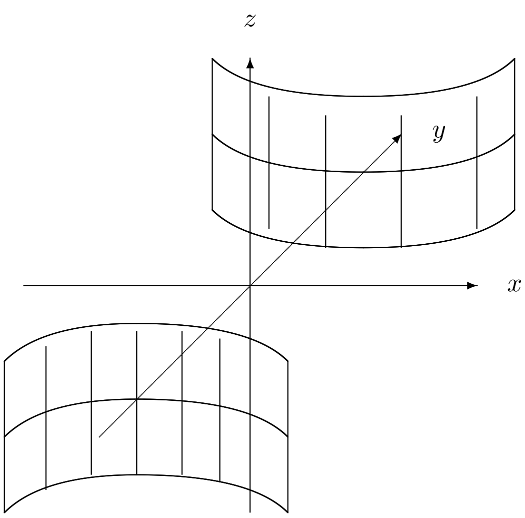


c) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$ *leere Menge*

d) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ *elliptischer Zylinder*

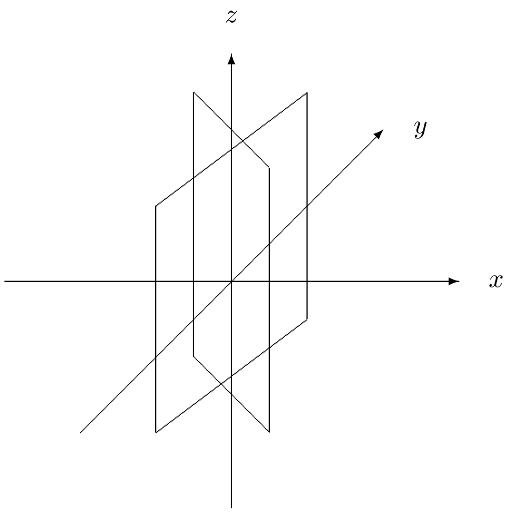


e) $\frac{x^2}{a^2} - \frac{y^2}{b^2} + 1 = 0$ *hyperbolischer Zylinder*



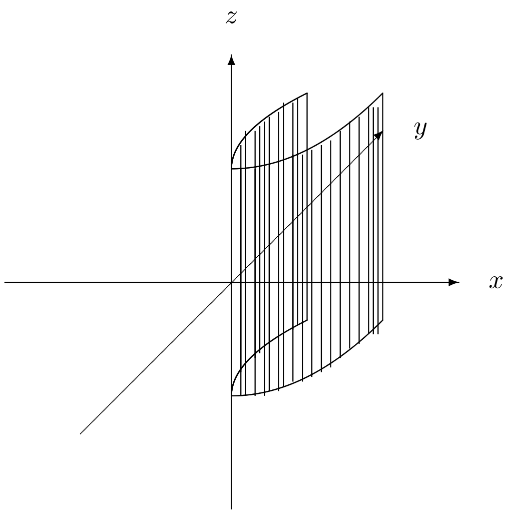
f) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$ *Gerade* (z-Achse)

g) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ *Ebenenpaar mit Schnittgerade* (z-Achse)



(iii) **rang $\boldsymbol{A} = 1$**

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 $\boldsymbol{A} = 0$**

$b_1x + b_2y + b_3z + c = 0$ *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)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0$$

Die x-y-Ebene schneidet das Hyperboloid in der Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

In jedem Punkt dieser Ellipse gibt es eine Berührungsebene an das Hyperboloid und wir zeigen: Jede dieser Ebenen schneidet das Hyperboloid in zwei Geraden.

Fall 1: Betrachte zunächst $x = \pm a$, $y = 0$, $z = 0$ und die zugehörigen Ebenen, die die Fläche berühren: $x = \pm a$. Schnitt zwischen Ebene und Hyperboloid besteht aus den Punkten, die beide Gleichungen erfüllen. Setzen daher $x = \pm a$ in die Hyperboloid-Gleichung ein:

$$0 = 1 + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = (\frac{y^2}{b^2} + \frac{z}{c})(\frac{y}{b} - \frac{z}{c})$$

$$\text{also: } \frac{y}{b} + \frac{z}{c} = 0 \text{ oder } \frac{y}{b} - \frac{z}{c} = 0$$

Jede dieser Gleichungen beschreibt zusammen mit $x = \pm a$ eine Gerade durch $(\pm a; 0; a)$.

Fall 2: Betrachte Punkt $(x_0, y_0, 0)$ mit $\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1$, $y_0 \neq 0$ und die zugehörige Berührungsebene:

$$(x - x_0)\frac{x_0}{a^2} + (y - y_0)\frac{y_0}{b^2} = 0 \quad (*)$$

Umformen dieser Ebenengleichung (beachte $y_0 \neq 0$) ergibt:

$$y = y_0 - \frac{b^2 x_0}{a^2 y_0}(x - x_0)$$

Einsetzen in Hyperboloidgleichung:

$$\begin{aligned}
0 &= \frac{x^2}{a^2} + (\frac{y_0}{b^2} - \frac{b}{a^2}\frac{x_0}{y_0}(x - x_0))^2 - \frac{z^2}{c^2} - 1 \\
0 &= \frac{x^2}{a^2} + \frac{y_0^2}{b^2} - 2\frac{x_0}{a^2}(x - x_0)\frac{b^2}{a^4}\frac{x_0^2}{y_0^2}(x - x_0)^2 - \frac{z^2}{c^2} - 1 \\
0 &= \frac{x^2}{a^2} - \frac{x_0^2}{a^2} - \frac{2}{a^2}x_0(x - x_0) + \frac{b^2}{a^4}\frac{x_0^2}{y_0^2}(x - x_0)^2 - \frac{z^2}{c^2} \\
0 &= \frac{x - x_0}{a^2}(x + x_0 - 2x_0) + \frac{b^2}{a^4}\frac{x_0^2}{y_0^2}(x - x_0)^2 - \frac{z^2}{c^2} \\
0 &= \frac{(x - x_0)^2}{a^2}\underbrace{(1 + \frac{b^2}{a^2}\frac{x_0^2}{y_0^2})}_{=:e^2} - \frac{z^2}{c^2} \\
0 &= (\frac{e}{a}(x - x_0) + \frac{z}{c})(\frac{e}{a}(x - x_0) - \frac{z}{c})
\end{aligned}$$

Jeder Faktor ergibt eine Ebenengleichung die zusammen mit $(*)$ eine Gerade beschreibt. □
Schnittpunkte erfüllen also:

$$\begin{aligned}
y &= y_0 - \frac{b^2 x_0}{a^2 y_0}(x - x_0) \quad \wedge \quad \frac{e}{a}(x - x_0) + \frac{z}{c} = 0 \text{ (1. Schnittgerade)} \\
\text{oder } y &= y_0 - \frac{b^2 x_0}{a^2 y_0}(x - x_0) \quad \wedge \quad \frac{e}{a}(x - x_0) - \frac{z}{c} = 0 \text{ (2. Schnittgerade)}
\end{aligned}$$

50 Matrixnormen und Eigenwertabschätzungen

50.1 Motivation

Problem:

- Kann man die Eigenwerte eine Matrix mit geringem Aufwand abschätzen?
- Dies spielt z.B. eine Rolle bei Konvergenzbetrachtungen von iterativen Algorithmen.

Ein wichtiges Hilfsmittel hierzu sind Matrixnormen.

50.2 Definition

Unter einer *Matrixnorm* versteht man eine Funktion $||\cdot|| : \mathbb{R}^{n\times n} \to \mathbb{R}$ mit folgenden Eigenschaften:

a) $||A|| \geq 0 \quad \forall A \in \mathbb{R}^{n\times n}$ und $(||A|| = 0 \Leftrightarrow A = 0)$.

b) $||\lambda A|| = |\lambda| \cdot ||A|| \quad \forall \lambda \in \mathbb{R},\ \forall A \in \mathbb{R}^{n\times n}$.

c) $||A + B|| \leq ||A|| + ||B|| \quad \forall A, B \in \mathbb{R}^{n\times n}$.

d) $||A \cdot B|| \leq ||A|| \cdot ||B|| \quad \forall A, B \in \mathbb{R}^{n\times n}$ (*Submultiplikativität*)

50.3 Beispiele

Sei $A = (a_{ik}) \in \mathbb{R}^{n\times n}$.

a) *Gesamtnorm*: $||A||_G := n \cdot \max\limits_{i,k} |a_{ik}|$

b) *Zeilensummennorm*: $||A||_Z := \max\limits_{i} \sum\limits_{k=1}^{n} |a_{ik}|$

c) *Spaltensummennorm*: $||A||_S := \max\limits_{k} \sum\limits_{i=1}^{n} |a_{ik}|$

d) *Frobeniusnorm*: $||A||_F := \left(\sum\limits_{i,k=1}^{n} {a_{ik}}^2 \right)^{\frac{1}{2}}$

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: $||A||_2 = \max\limits_{k} \{|\lambda_k| \mid \lambda_k \text{ Eigenwert von } A\}$

Da Matrizen und Vektoren oft gemeinsam auftreten, sollten Matrix- und Vektornormen verträglich sein:

50.4 Definition

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*

$$||x||_p := \begin{cases} \Big(\sum\limits_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} & \text{für } 1 \leq p < \infty \\ \max\limits_{i=1,\dots,n} \{|x_i|\} & \text{für } p = \infty \end{cases}$$

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* $||x||_2$.

c) $||A||_G$, $||A||_Z$ sind kompatibel zur *Maximumsnorm* $||x||_\infty$.

Beweis: Wir zeigen nur Kompatibilität von $||A||_G$ und $||x||_\infty$:

$$\begin{aligned} ||Ax||_\infty &= \max_i \Big\{\Big|\sum_{k=1}^{n} a_{ik}x_k\Big|\Big\} \\ &\leq \max_i \Big\{\sum_{k=1}^{n} |a_{ik}x_k|\Big\} \quad \text{Dreiecksungleichung} \\ &\leq \max_i \Big\{\sum_{k=1}^{n} \max_{r,s} |a_{rs}| \cdot \max_l |x_l|\Big\} \\ &= n \cdot \max_{r,s} |a_{rs}| \cdot \max_l |x_l| \\ &= ||A||_G \cdot ||x||_\infty. \end{aligned}$$

□

Da zu einer gegebenen Vektornorm $||\cdot||_V$ oftmals viele kompatible Matrixnormen $||\cdot||_M$ existieren, verwendet man in der Praxis gerne diejenige, für die die Abschätzung $||Ax||_V \leq ||A||_M \cdot ||x||_V$ am schärfsten ist.

50.6 Definition

Die zu einer gegebenen Vektornorm $||\cdot||$ 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 Definition 50.2 besitzt und die kleinste aller Matrixnormen ist, die zu einer gegebenen Vektornorm kompatibel sind.

50.7 Beispiele

Man erhält:

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 (Eigenwertabschätzungen mit Matrixnormen)

Ist λ ein Eigenwert von $A \in \mathbb{R}^{n\times n}$ und $||A||$ eine beliebige, zu einer Vektornorm verträgliche Matrixnorm, so ist $|\lambda| \le ||A||$.

Beweis: Sei v ein Eigenvektor zu λ.

$$\Rightarrow |\lambda|\cdot||v|| = ||\lambda v|| = ||Av|| \le ||A||\cdot||v||.$$

Da $v \neq 0$, gilt $||v|| > 0$. Daher ist $|\lambda| \le ||A||$. □

50.9 Beispiel

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

$$\begin{aligned}
||A||_G &= 3\cdot\max_{i,k}|a_{ik}| = 3\cdot 3 = 9 \\
||A||_Z &= \max\,\{1,2\ ;\ 2,4\ ;\ 3,2\} = 3,2 \\
||A||_S &= \max\,\{1,2\ ;\ 2,1\ ;\ 3,2\} = 3,5 \\
||A||_F &= \sqrt{1^2+0,1^2+(-0,1)^2+2^2+0,4^2+(-0,2)^2+3^2} = \sqrt{14,22} \approx 3,77
\end{aligned}$$

$||A||_Z$ liefert die schärfste Abschätzung: $|\lambda| \le ||A||_Z = 3,2$.
Tatsächlich gilt: $\lambda_1 \approx 3,0060$, $\lambda_2 \approx 2,0078$, $\lambda_3 \approx 0,9862$. □

Offenbar erlaubt Satz 50.8 nur die Abschätzung des betragsgrößten Eigenwerts. Gibt es auch Abschätzungen für alle Eigenwerte?

50.10 Satz (Satz von Gerschgorin)

a) Die Vereinigung aller Kreisscheiben

$$K_i := \{\mu \in \mathbb{C} \big|\ |\mu - a_{ii}| \le \sum_{\substack{k=1\\k\neq i}}^{n} |a_{ik}|\}$$

enthält alle Eigenwerte der Matrix $A = (a_{ik}) \in \mathbb{R}^{n\times n}$.

b) Jede Zusammenhangskomponente aus m solcher Kreisscheiben enthält genau m Eigenwerte (der Vielfachheit nach gezählt).

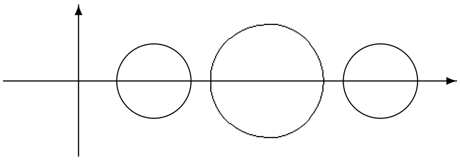
Beweis: siehe z.B. Stoer/Bulirsch: Einführung in die numerische Mathematik II, Springer, Berlin.

50.11 Beispiel

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

$$\begin{aligned} K_1 &= \{\mu \in \mathbb{C} \big| \, |\mu - 1| \leq 0,2\} \\ K_2 &= \{\mu \in \mathbb{C} \big| \, |\mu - 2| \leq 0,4\} \\ K_3 &= \{\mu \in \mathbb{C} \big| \, |\mu - 3| \leq 0,2\} \end{aligned}$$

Alle Eigenwerte liegen in $K_1 \cup K_2 \cup K_3$:



Da K_1, K_2, K_3 nicht überlappen, liegt nach 50.10 (b) in jeder der Kreisscheiben genau ein Eigenwert. Ferner ist A invertierbar, da $0 \notin K_1 \cup K_2 \cup K_3$, also 0 kein Eigenwert ist.

50.12 Korollar (Invertierbarkeit strikt diagonaldominanter Matrizen)

Ist $A \in \mathbb{R}^{n \times n}$ strikt diagonaldominant (d.h. $|a_{ii}| > \sum\limits_{\substack{k=1 \\ k \neq i}}^{n} |a_{ik}| \quad \forall i = 1, \ldots, n$) so ist A invertierbar.

Beweis: Nach dem Satz von Gerschgorin liegt 0 außerhalb der Gerschgorinkreisscheiben, kann also kein Eigenwert sein. $\square$

51 Numerische Berechnung von Eigenwerten und Eigenvektoren

51.1 Motivation

Die Berechnung der Eigenwerte von $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ (entspricht $\mathbb{R}^{n\times n}$) über die Nullstellen von $\det(A-\lambda I)$ führt für $n \geq 5$ auf Polynome, für die es keine analytischen Lösungsformeln gibt. Das macht numerische Verfahren notwendig.

51.2 Die einfache Vektoriteration

(Potenzmethode, Von-Mises-Verfahren)

Idee:
Sei $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ symmetrisch und $u_0 \in \mathbb{R}^n$ ein beliebiger Startvektor. Lassen sich mit der Iteration

$$u_{k+1} = Au_k,\ k = 0, 1, \dots$$

Aussagen über Eigenwerte und Eigenvektoren von A gewinnen?

Lösung:
Sei $\lambda_1, \dots, \lambda_n$ Eigenwerte und $v_1, \dots, v_n$ die zugehörigen linear unabhängigen Eigenvektoren von A.
O.B.d.A.: $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$
Sei $u_0 = \sum_{i=1}^{n} \alpha_i v_i$
$\Rightarrow u_1 = Au_0 = \sum_{i=1}^{n} \alpha_i A v_i = \sum_{i=1}^{n} \alpha_i \lambda_i v_i$
analog $u_k = \sum_{i=1}^{n} \alpha_i {\lambda_i}^k v_i = {\lambda_1}^k (\alpha_1 v_1 + \sum_{i=2}^{n} \frac{{\lambda_i}^k}{{\lambda_1}^k} \alpha_i v_i)$
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 den Nullvektor 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 v_1.
Die Konvergenz ist umso schneller, je kleiner die $\left|\frac{\lambda_i}{\lambda_1}\right|$ sind ($i = 2, \dots, n$).
Der *Rayleigh-Koeffzient*

$$R_A(u_k) = \frac{\langle u_k, Au_k\rangle}{\langle u_k, u_k\rangle}$$

konvergiert dann und ist gleich

$$\frac{\langle \alpha_1 v_1, \lambda_1 \alpha_1 v_1\rangle}{\alpha_1 v_1, \alpha_1 v_1\rangle} = \lambda_1.$$

Die Vektoriteration ist also ein einfaches Verfahren zur numerischen Approximation des dominanten Eigenwertes (und somit der Spektralnorm) und des dominanten Eigenvektors einer symmetrischen Matrix A.

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

$A = \begin{pmatrix} 1,04 & 0,72 \\ 0,72 & 1,46 \end{pmatrix}$, Startvektor $u_0 =: \begin{pmatrix} 1 \\ 0 \end{pmatrix} =: v_0$ (schon normiert)

$$\begin{aligned}
u_1 &= Av_0 = \begin{pmatrix} 1,04 \\ 0,72 \end{pmatrix}, \quad v_1 = \frac{u_1}{|u_1|} \approx \begin{pmatrix} 0,8222 \\ 0,5692 \end{pmatrix}, \; R_A(v_0) = \frac{v_0^\top A v_0}{v_0^\top v_0} = v_0^\top u_1 \approx 1,04 \\
u_2 &= Av_1 \approx \begin{pmatrix} 1,2649 \\ 1,4230 \end{pmatrix}, \; v_2 = \frac{u_2}{|u_2|} \approx \begin{pmatrix} 0,6645 \\ 0,7476 \end{pmatrix}, \; R_A(v_1) = v_1^\top u_2 \approx 1,8500 \\
u_3 &= Av_2 \approx \begin{pmatrix} 1,2294 \\ 1,5699 \end{pmatrix}, \; v_3 = \frac{u_3}{|u_3|} \approx \begin{pmatrix} 0,6166 \\ 0,7873 \end{pmatrix}, \; R_A(v_2) = v_2^\top u_3 \approx 1,9906 \\
u_4 &= Av_3 \approx \begin{pmatrix} 1,2081 \\ 1,5934 \end{pmatrix}, \; v_4 = \frac{u_4}{|u_4|} \approx \begin{pmatrix} 0,6042 \\ 0,7968 \end{pmatrix}, \; R_A(v_3) = v_3^\top u_4 \approx 1,9994 \\
u_5 &= Av_4 \approx \begin{pmatrix} 1,2020 \\ 1,5984 \end{pmatrix}, \; v_5 = \frac{u_5}{|u_5|} \approx \begin{pmatrix} 0,6010 \\ 0,7992 \end{pmatrix}, \; R_A(v_4) = v_4^\top u_5 \approx 1,9999
\end{aligned}$$

Exakte Lösung für dominanten Eigenvektor und Eigenwert:

$$v = \begin{pmatrix} 0,6 \\ 0,8 \end{pmatrix}, \; \lambda = 2$$

51.4 Das Jacobi-Verfahren

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 dieselben Eigenwerte.

b) Mit Hilfe einer Sequenz $(Q_k)_{k=1,\dots}$ von orthogonalen Matrizen transformiert man A auf Diagonalgestalt.
Die Diagonalelemente geben die Eigenwerte an und aus $(Q_k)_{k=1,\dots}$ berechnet man die Eigenvektoren.

Beweis:

a) A und $Q^\top AQ$ haben dieselben charakteristischen Polynome, denn

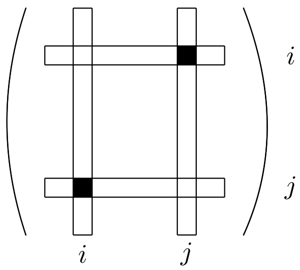
$$\begin{aligned}
p_{Q^\top AQ} &= \det\,(Q^\top AQ - \lambda I) = \det\,(Q^\top AQ - \lambda \underbrace{Q^\top Q}_{I}) \\
&= \det\,(Q^\top (A - \lambda I) Q) \\
&= \det\,(Q^\top) \cdot \det\,(A - \lambda I) \cdot \det\,(Q) \\
&= \det\,(\underbrace{Q^\top Q}_{I}) \cdot \det\,(A - \lambda I) = \det\,(A - \lambda I) = p_A(\lambda)
\end{aligned}$$

Also haben A und $Q^\top AQ$ dieselben Eigenwerte. □

b) Einige Details:
Als orthogonale Transformationen verwendet man Rotationsmatrizen vom Typ

$$Q_k := Q(i,j,\varphi_k) := \begin{array}{cl} \left(\begin{array}{ccccccccccc} 1 & & & & & & & & & & 0 \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & \cos\varphi_k & & \cdots & & \sin\varphi_k & & & \\ & & & & 1 & & & & & & \\ & & & \vdots & & \ddots & & \vdots & & & \\ & & & & & & 1 & & & & \\ & & & -\sin\varphi_k & & \cdots & & \cos\varphi_k & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ 0 & & & & & & & & & & 1 \end{array}\right) & \begin{array}{l} \\ \\ \\ \leftarrow \text{i-te Zeile} \\ \\ \\ \\ \leftarrow \text{j-te Zeile} \\ \\ \\ \\ \end{array} \\ \begin{array}{ccc} \uparrow & \qquad\qquad & \uparrow \\ \text{i-te Spalte} & & \text{j-te Spalte} \end{array} & \end{array}$$

Man kann zeigen, dass $Q_k^\top A Q_k$ nur die i-ten und j-ten Zeilen und Spalten von A verändert:

$$Q_k^\top A Q_k =$$ 

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 dieses Verfahren gegen eine Diagonalmatrix konvergiert:

$$Q_k^\top \cdot \ldots \cdot Q_2^\top \cdot Q_1^\top \cdot A \cdot Q_1 \cdot Q_2 \cdot \ldots \cdot Q_k \overset{k\to+\infty}{\longrightarrow} \operatorname{diag}\,(\lambda_1, \ldots, \lambda_n)$$

Die m-te Spalte von $P = Q_1 \cdot \ldots \cdot Q_k$ enthält 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 Orthonormalsystem von Eigenvektoren. Gram-Schmidt-Orthonormalisierung 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 Nichtdiagonalelement 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.