

Mathematik III Mitschrift

Thilo Fester
Till Helge Helwig

30. Januar 2007

Vorlesung wurde gehalten von

Dr. Michael Huber

im Wintersemester 2006/2007.

*Die Mitschrift entspricht nicht 1:1 der Vorlesung,
da wir an einigen Stellen die Formulierungen geändert haben.*

Inhaltsverzeichnis

1	Lineare Algebra (Fortsetzung)	3
1.1	Vektorräume	3
1.2	Lineare Abbildungen und Matrizen	3
1.3	Lineare Gleichungssysteme	3
1.4	Determinanten	8
1.5	Theorie der Eigenwerte	11
1.6	Euklidische Vektorräume	17
2	Mehrdimensionale Analysis (Funktionen mehrerer Variablen)	20
2.1	Metrische Räume	20
2.2	Grenzwerte und Stetigkeit	25
2.3	Mehrdimensionale Analysis	28
3	Anwendungen	34
3.1	Lineare Algebra und Google	34
3.2	Codierungstheorie	38

1 Lineare Algebra (Fortsetzung)

1.1 Vektorräume

1.2 Lineare Abbildungen und Matrizen

1.3 Lineare Gleichungssysteme

Motivation:

- Lösen (linearer) Gleichungssysteme ist von fundamentaler Bedeutung bei vielen Anwendungen und Beschreibungen mittels Matrizen.
- Es gibt ein algorithmisches Verfahren, das lineare Gleichungssysteme löst und (in etwas modifizierter Form) die Invertierbarkeit von Matrizen überprüft und die inversen Matrizen berechnet (falls sie existieren).
 ⇒ Sogenannter Gauss'scher Eliminationsalgorithmus (C. F. Gauss, 1777-1855)

Lineare Gleichungssysteme, Homogenität (6.37)

Def. 6.37

Sei $A \in \mathcal{M}_{m,n}(K)$ und $b \in K^m$, dann heißt das Paar (A, b) lineares Gleichungssystem (LGS). Ein Vektor $x \in K^n$ heißt eine Lösung des LGS, falls $A \cdot x = b$. Die Menge $\{x \in K^n \mid Ax = b\}$ heißt Lösungsmenge des LGS. Das LGS (A, b) heißt homogen, falls $b = 0$, sonst inhomogen. Jedes homogene LGS hat die triviale Lösung $x = 0$. Mit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

ist die Vektorgleichung $Ax = b$ äquivalent zum LGS der m Gleichungen mit n Variablen:

$$\sum_{j=1}^n a_{ij}x_j = b_i \text{ für } 1 \leq i \leq m$$

Beispiel:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Dann ist $Ax = b$ gleichwertig mit:

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

Affiner Unterraum (6.38)

Def. 6.38

Sei U ein Unterraum des Vektorraums V , dann heißt die Menge $U + a := \{u + a | u \in U\}$ für ein $a \in V$ ein affiner Unterraum von V mit $\dim(U + a) := \dim U$. Insbesondere ist jeder Unterraum ein affiner Unterraum ("Parallelverschiebungen").

Strukturelle Lösbarkeit eines LGS (6.39)

Satz 6.39

Sei $A \in \mathcal{M}_{m,n}(K)$ und $b \in K^m$, dann tritt genau einer der beiden Fälle auf:

1. Das LGS (A, b) ist unlösbar (d.h. es existiert kein $x \in K^n$ mit $Ax = b$)
2. Es gibt eine Lösung $a \in K^n$. Zum LGS (A, b) ist die Lösungsmenge genau der affine Unterraum $U + a$ mit $U := \{x \in K^n | Ax = 0\}$. (U ist die Lösungsmenge des zugehörigen homogenen LGS $(A, 0)$.)
Es gilt: $\dim(U + a) = n - \text{rg}(A)$.
[Dimensionssatz 6.26: $\dim \text{Im} f = \text{rg}(f) = \text{rg}(A) = \text{maximale Anzahl linear unabhängiger Spalten von } A$]

Beweis:

Sei $a \in K^n$ eine Lösung von (A, b) , d.h. $Aa = b$, und $f : K^n \rightarrow K^m, x \mapsto f(x) = Ax$ (bzgl. der Standardbasis, Definition 6.32) linear, dann ist $U = \ker f = f^{-1}(\{0\})$ und das LGS (A, b) hat die Lösungsmenge $f^{-1}(\{b\}) = U + a$. [$x \in U + a \Leftrightarrow x - a \in U \Leftrightarrow f(x - a) = 0 \Leftrightarrow f(x) - f(a) \Leftrightarrow f(x) = b \Leftrightarrow x \in f^{-1}(\{b\})$] mit $\dim(U + a) = \dim U = \dim \ker f = n - \text{rg}(f)$. ■

Bemerkung:

Die Lösungen eines inhomogenen LGS (A, b) erhält man durch Addition einer speziellen Lösung a von (A, b) zu allen Lösungen des homogenen LGS $(A, 0)$.

Korollar (6.40)

Kor. 6.40

Sei $n > m$, dann hat jedes homogene LGS $(A, 0)$ mit n Gleichungen und m Variablen stets eine nicht-triviale Lösung.

Beweis:

$a = 0$ ist stets eine Lösung von $(A, 0)$, dann gilt $\text{rg}(A) = \text{rg}(f) \leq m$ und mit 6.39 folgt $\dim U = n - \text{rg}(A) \geq n - m > 0$. Daher existiert eine nicht-triviale Lösung. □

Korollar (6.41)

Kor. 6.41

Sei $A \in \mathcal{M}_{m,n}(K)$ und $f : K^n \rightarrow K^m, x \mapsto Ax$ (mit Standardbasis), dann:

1. Das LGS (A, b) ist für jedes $b \in K^m$ lösbar $\Leftrightarrow f$ ist surjektiv $\Leftrightarrow \text{rg}(A) = m$.
2. Das LGS (A, b) hat für ein $b \in K^m$ höchstens eine Lösung $\Leftrightarrow f$ ist injektiv $\Leftrightarrow \text{rg}(A) = n$.

Korollar (6.42)

Kor. 6.42

Sei $A \in \mathcal{M}_{n,m}(K)$ quadratisch, dann sind äquivalent:

1. Das LGS (A, b) ist für alle $b \in K^n$ lösbar.
2. Das LGS (A, b) ist für alle $b \in K^n$ eindeutig lösbar.
3. Das LGS (A, b) ist für ein $b \in K^n$ eindeutig lösbar.
4. Das zugehörige homogene LGS $(A, 0)$ ist nur trivial lösbar.
5. $\text{rg}(A) = n$
6. A ist invertierbar.

Beweis folgt sofort aus 6.41 und 6.34.

24.10.2006

Wir betrachten im Folgenden folgende algorithmische Probleme der linearen Algebra:

- Ist (A, b) lösbar?
- Finde eine oder alle Lösungen von (A, b) .
- Finde nicht-triviale Lösungen des zugehörigen homogenen Systems $(A, 0)$. (Diese existieren genau dann, wenn Spalten von A linear abhängig sind.)
- Bestimme den Rang einer Matrix $\text{rg}(A)$ und ihre Inverse A^{-1} (falls existent).

Elementare Zeilenumformungen (6.43)

Def. 6.43

Sei $A \in \mathcal{M}_{m,n}(K)$. Als elementare Zeilenumformungen werden folgende drei Typen bezeichnet:

- (i) Vertauschung zweier Zeilen von A .
- (ii) Zu einer Zeile von A wird das k -fache einer anderen Zeile hinzuaddiert ($k \in K$).
- (iii) Eine Zeile wird mit $k \neq 0$ ($k \in K^* = K \setminus \{0\}$) multipliziert.

Bemerkung:

Diese elementaren Zeilenumformungen lassen sich durch geeignete Matrizenmultiplikationen realisieren, z.B.:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \stackrel{(i)}{=} \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \stackrel{(ii)}{=} \begin{pmatrix} 9 & 12 & 15 \\ 4 & 5 & 6 \end{pmatrix}$$

Stufenform und Gauss-Algorithmus (6.44)

Def. 6.44

Eine Matrix $A \in \mathcal{M}_{m,n}(K)$ hat Stufenform, falls sie von der Form

$$A = \begin{pmatrix} 0 & \dots & 0 & 1 & * & * & * \\ 0 & \dots & \dots & 0 & 1 & * & * \\ 0 & \dots & \dots & \dots & 0 & 1 & * \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix} \begin{array}{l} r \text{ Zeilen} \\ m - r \text{ Zeilen} \end{array}$$

ist mit $r \in \{0, 1, \dots, m\}$. Der folgende Gauss-Algorithmus transformiert eine beliebige Matrix A durch elementare Zeilenumformungen auf Stufenform:

- Falls $A = 0$, dann fertig ($r = 0$).
- Falls $A \neq 0$, suche von links die erste Spalte von A , die nicht nur Nullen enthält; diese erhält den Index j_1 . Durch Umformung vom Typ (i) erreicht man $a_{1,j_1} \neq 0$ und sogar $a_{1,j_1} = 1$ durch Umformung vom Typ (iii).
- Nun addiert man das $-(a_{i,j_1})$ -fache der ersten Zeile zur i -ten Zeile ($i \in \{2, 3, \dots, m\}$). Somit entstehen unterhalb von $a_{1,j_1} = 1$ lauter Nullen:

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

mit einer $(m-1) \times (n-j_1)$ -Matrix A' .

- Wiederholung des Verfahrens mit A' bis die Stufenform erreicht ist.

Beispiel:

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 2 & 0 & 1 & 3 \\ 1 & 2 & 3 & 0 \end{pmatrix} \xrightarrow{(i)} \begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 0 & 1 & 3 \\ 0 & 1 & 1 & 1 \end{pmatrix} \xrightarrow{(ii)} \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -4 & -5 & 3 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow{(i)} \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -4 & -5 & 3 \end{pmatrix} \xrightarrow{(ii)} \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 7 \end{pmatrix} \xrightarrow{(iii)} \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -7 \end{pmatrix}$$

Bemerkung:

Das Element $a_{i,j_1} \neq 0$ heißt *Pivotelement*; man kann jedes von Null verschiedene Element der j_1 -ten Spalte als Pivotelement wählen.

Erweiterte Matrix (6.45)

Def. 6.45

Gegeben sei ein LGS (A, b) mit $A = (a_{ij}) \in \mathcal{M}_{m,n+1}(K)$, $b \in K^m$. Die *erweiterte Matrix* $(A | b) \in \mathcal{M}_{m,n+1}(K)$ ist die Matrix

$$(A | b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

Entsteht $(A' | b')$ aus $(A | b)$ durch elementare Zeilenumformungen, so haben die LGS (A, b) und (A', b') den *gleichen Lösungsraum*. (Da $(A' | b') = M \cdot (A | b)$ mit einer Matrix $M \in \mathcal{M}_{m,m}(K)$, die invertierbar ist. Also gilt $A' = M \cdot A$ und $b' = M \cdot b$ und somit für alle $x \in K^n$: $A \cdot x = b \Leftrightarrow M \cdot A \cdot x = M \cdot b \Leftrightarrow A' \cdot x = b'$.) Insbesondere gilt $\text{rg}(A) = \text{rg}(A')$. Daher genügt es den Spezialfall zu betrachten, dass A Stufenform hat.

25.10.2006

Es ergibt sich:

Praktische Lösbarkeit eines LGS (6.46)

Satz 6.46

Sei $A \in \mathcal{M}_{m,n}(K)$ in Stufenform, $b \in K^m$, d.h.

$$(A|b) = \begin{pmatrix} 0 & \dots & 0 & 1 & * & * & b_1 \\ 0 & \dots & \dots & 0 & 1 & * & \vdots \\ 0 & \dots & \dots & \dots & 0 & 1 & b_r \\ 0 & \dots & \dots & \dots & \dots & 0 & b_{r+1} \\ 0 & \dots & \dots & \dots & \dots & 0 & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 & b_m \end{pmatrix} \quad r \text{ Zeilen}$$

Dann gilt:

1. $\text{rg}(A) = r$
2. Das LGS (A, b) ist lösbar. $\Leftrightarrow b_{r+1} = \dots = b_m = 0$. In diesem Fall ergeben sich alle Lösungen $x \in K^n$ durch sogenannte *Rücksstitution*.

Beispiel:

$$\begin{pmatrix} 1 & 1 & 4 & 3 & b_1 \\ 0 & 0 & 1 & 1 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \\ 0 & 0 & 0 & 0 & b_4 \end{pmatrix}$$

$r = 3$ und (A, b) ist lösbar. $\Leftrightarrow b_4 = 0$. In diesem Fall ergibt sich:

$$\begin{aligned} x_1 + x_2 + 4x_3 + 3x_4 &= b_1 \\ x_3 + x_4 &= b_2 \\ x_4 &= b_3 \end{aligned}$$

Man erhält:

$$\begin{aligned} x_4 &= b_3 \\ x_3 &= b_2 - x_4 = b_2 - b_3 \\ x_2 &\dots \text{ freie Variable } (x_2 := t \in K) \\ x_1 &= b_1 - x_2 - 4x_3 - 3x_4 = b_1 - x_2 - 4b_2 + b_3 \end{aligned}$$

Daraus ergibt sich der Lösungsraum: $\left\{ (b_1 - x_2 - 4b_2 + b_3, x_2, b_2 - b_3, b_3)^t \mid x_2 \in K \right\}$

Praktisches Berechnen der inversen Matrix (6.47)

Kor. 6.47

Nach 6.42 ist $A \in \mathcal{M}_{n,n}(K)$ genau dann invertierbar, wenn $\text{rg}(A) = n$. 6.46 liefert den Test auf Invertierbarkeit. Daraus ergibt sich folgende Berechnung der inversen Matrix:

Wende elementare Zeilenumformungen auf das Paar (A, E_n) an, bis aus A die Einheitsmatrix E_n wird. Dann steht rechts die inverse Matrix A^{-1} .

$[(A, E_n)$ wird ersetzt durch (MA, M) mit $MA = E_n$, also $M = A^{-1}$.]

Beispiel:

$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$ ist gegeben.

$$\begin{aligned} (A, E_3) &= \begin{pmatrix} 1 & 2 & 1 & \vdots & 1 & 0 & 0 \\ 1 & 1 & 1 & \vdots & 0 & 1 & 0 \\ 1 & 1 & -1 & \vdots & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & \vdots & 1 & 0 & 0 \\ 1 & -1 & 0 & \vdots & -1 & 1 & 0 \\ 0 & 1 & -2 & \vdots & -1 & 0 & 1 \end{pmatrix} \\ &\rightarrow \begin{pmatrix} 1 & 2 & 1 & \vdots & -1 & 2 & 0 \\ 0 & -1 & 0 & \vdots & 1 & -1 & 0 \\ 0 & 0 & -2 & \vdots & 0 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & \vdots & -1 & \frac{3}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \vdots & 1 & -1 & 0 \\ 0 & 0 & 1 & \vdots & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \end{aligned}$$

1.4 Determinanten

Motivation:

- "Linearform": $f: \underbrace{K^n}_V \rightarrow K$ linear, eindimensionaler Vektorraum über sich selbst
"Multilinearform": $\mathcal{M}_{n,n}(K) \rightarrow K$
- Determinanten sind wichtig für die Eigenwerttheorie (6.5).

Determinantenfunktion \det (6.48)

Def. 6.48

Sei K ein beliebiger Körper, $n \in \mathbb{N}$. Eine Determinantenfunktion ist eine Abbildung $\det: \mathcal{M}_{n,n}(K) \rightarrow K$ ("Multilinearform") mit folgenden Eigenschaften:

1. \det ist linear in jeder Zeile, d.h.:

$$\det \begin{pmatrix} z_1 \\ \vdots \\ k_1 \cdot z_i + k_2 \cdot z'_i \\ \vdots \\ z_n \end{pmatrix} = k_1 \cdot \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix} + k_2 \cdot \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix} \quad \text{für jede Zeile } i \in \{1, \dots, n\}$$

und $k_1, k_2 \in K$

2. $\det A = 0$, falls $\text{rg}(A) < n$ (mit anderen Worten: falls z Zeilen von A übereinstimmen, "det ist alternierend")
3. $\det E_n = 1$

Alternative Schreibweise: $\det(A) = |A|$

Man kann zeigen:

Rechenregeln für Determinanten (6.49)

Satz 6.49

Für jede Determinantenfunktion $\det : \mathcal{M}_{n,n}(K) \rightarrow K$ gilt:

- Entsteht B aus A durch Vertauschung zweier Zeilen, so gilt: $\det(B) = -\det(A)$
- Entsteht B aus A durch Addition eines Vielfachen einer Zeile zu einer anderen Zeile, so gilt: $\det(B) = \det(A)$
- Entsteht B aus A durch Multiplikation einer Zeile mit $k \in K$, so gilt: $\det(B) = k \cdot \det(A)$
- Die Determinante einer oberen (oder unteren) Dreiecksmatrix ist das Produkt der Hauptdiagonalelemente:

$$\det \begin{pmatrix} a_{11} & * & * \\ 0 & \ddots & * \\ 0 & 0 & a_{nn} \end{pmatrix} = \prod_{i=1}^n a_{ii}$$

Damit lässt sich zu jeder Matrix "die" Determinante berechnen:

Beispiel (Vandermonde'sche Matrix, 1771):

$$\det \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix} = ?$$

25.10.2006

$$\begin{aligned} &= \det \begin{pmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{pmatrix} = (b-a)(c-a) \cdot \det \begin{pmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 1 & c+a \end{pmatrix} \\ &= (b-a)(c-a) \cdot \det \begin{pmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 0 & c-b \end{pmatrix} = (b-a)(c-a)(c-b) \end{aligned}$$

Existiert \det überhaupt? Falls ja, wieviele? Es lässt sich nur sehr aufwendig zeigen:

Existenz und Eindeutigkeit der Determinantenfunktion \det (6.50)

Satz 6.50

Sei K ein Körper, $n \in \mathbb{N}$. Dann existiert genau eine Determinantenfunktion $\det : \mathcal{M}_{n,n}(K) \rightarrow K$.

Beweis: Siehe HWK, Seite 338.

Beispiele:

- $\det(a) = a$
- $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$
- $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} - a_{13} \cdot a_{23} \cdot a_{31} - a_{11} \cdot a_{23} \cdot a_{32} - a_{12} \cdot a_{21} \cdot a_{33}$
(Regel von Sarrus, Summe der Hauptdiagonalen minus Summe der Nebendiagonalen)

Korollar (6.51)

Kor. 6.51

Für alle $A \in \mathcal{M}_{n,n}(K)$ gilt: $\det(A) = \det(A^t)$.

(Wiederholung: A^t ist die zu A transponierte Matrix $A^t = (a_{ij})^t = (a_{ji})$ («Aus Spalten werden Zeilen»))

Zusammenhang Determinante und Invertierbarkeit (6.52)

Satz 6.52

Eine Matrix $A \in \mathcal{M}_{n,n}(K)$ ist invertierbar (d.h. $\text{rg}(A) = n$) $\Leftrightarrow \det(A) \neq 0$.

Beweis:

“ \Leftarrow “: Via Kontraposition. Sei $\text{rg}(A) < n$. Durch elementare Zeilenumformungen bringe A auf Stufenform. Dann ist mindestens ein Hauptdiagonalelement 0. Mit Satz 6.49d folgt $\det(A) = 0$.

“ \Rightarrow “: Sei $\text{rg}(A) = n$. Elementare Zeilenumformungen liefern Stufenform mit Einsen auf der Hauptdiagonalen und elementare Zeilenumformungen ändern die Determinante höchstens um einen Faktor $k \in K \setminus \{0\} \Rightarrow \det(A) \neq 0$ \square

Die Determinantenfunktion \det ist multiplikativ (6.53)

Satz 6.53

Für alle $A, B \in \mathcal{M}_{n,n}(K)$ gilt: $\det(A \cdot B) = \det(A) \cdot \det(B)$. Ist A invertierbar, dann $\det(A^{-1}) = \det(A)^{-1}$.

Beweis:

Sei $\text{rg}(B) < n$, dann $\text{rg}(A \cdot B) < n$ (dann $\text{Im}(AB) \leq \text{Im}(B)$). Nach Satz 6.52 folgt: $\det(A \cdot B) = 0 = \det(A) \cdot \underbrace{\det(B)}_{=0}$. Sei $\text{rg}(B) = n$, also $\det(B) \neq 0$. Wir definieren $\widetilde{\det} : \mathcal{M}_{n,n}(K) \rightarrow K$,

$\widetilde{\det} := \det(X \cdot B) \cdot \det(B)^{-1}$ für alle $x \in \mathcal{M}_{n,n}(K)$. Es ist leicht zu zeigen: $\widetilde{\det}$ ist die Determinantenfunktion (Definition 6.48). Wegen Satz 6.52 gilt daher $\det = \widetilde{\det}$. Also: $\det(A) = \widetilde{\det}(A) \stackrel{(*)}{=} \det(A \cdot B) = \widetilde{\det}(B)^{-1} \det(A \cdot B)$ und $\det(AB) = \det(A) \cdot \det(B)$. Ist A invertierbar, so gilt $A \cdot A^{-1} = E_n$, also $\det(A) \cdot \det(A^{-1}) \stackrel{\det \text{ multipl.}}{=} \det(A \cdot A^{-1}) = \det(E_n) = 1$. \blacksquare

Bemerkung:

Wie wir wissen, bilden die invertierbaren $(n \times n)$ -Matrizen die Gruppe $GL(K^n)$. Nach Satz 6.53 ist $\det : (GL(K^n), \cdot) \rightarrow (K \setminus \{0\}, \cdot)$ ein Gruppenhomomorphismus.

Zur Berechnung größerer Determinanten lässt sich zeigen:

Laplace'sche Entwicklung (6.54)

Satz 6.54

Sei $A \in \mathcal{M}_{n,n}(K)$ und $i \in \{1, \dots, n\}$, dann gilt:

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij})$$

Dabei bezeichnet A_{ij} die aus Streichen der i -ten Zeile und j -ten Spalte entstehende $(n-1) \times (n-1)$ -Matrix. («Laplace Entwicklung der i -ten Zeile».)

Bemerkung:

Entsprechend lässt sich auch nach der j -ten Spalte entwickeln.

Beispiel:

$$\begin{aligned}
\det \begin{pmatrix} 2 & 3 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 1 & 0 & 3 & 4 \\ 1 & 0 & 5 & 6 \end{pmatrix} & \stackrel{(*)}{=} 2 \cdot \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 5 & 6 \end{pmatrix} - 3 \cdot \det \begin{pmatrix} 0 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 5 & 6 \end{pmatrix} \\
& = 2 \cdot \det \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} - 3 \cdot \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} - 3 \cdot \det \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \\
& = -4 - 9 + 3 = -10
\end{aligned}$$

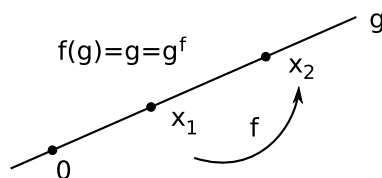
(*) ... 1. Zeile und 1. Spalte werden gestrichen.

07.11.2006

1.5 Theorie der Eigenwerte

Motivation:

- Die Bestimmung von sogenannten Eigenvektoren ist die (geometrische) Frage nach der Bestimmung aller Ursprungsgeraden, die durch eine gegebene lineare Abbildung als Ganzes fest gelassen wird.



- Diagonalisierbarkeit, Trigonalisierbarkeit von Matrizen.
- Zahlreiche Anwendungen: Spektraltheorie, Physik, (diskrete) dynamische Systeme.

Eigenwerte (6.55)

Def. 6.55

Sei V ein K -Vektorraum und $f : V \rightarrow V$ eine lineare Abbildung, dann:

- (i) heißt $k \in K$ Eigenwert von f , falls ein Vektor $v \in V \setminus \{0\}$ existiert mit $f(v) = k \cdot v$, d.h. $\ker(f - k \cdot \text{id}_V) \neq \{0\}$. Jeder solche Vektor $v \neq 0$ heißt *Eigenvektor*.
- (ii) heißt der Unterraum $\{v \in V \mid f(v) = k \cdot v\} = \ker(f - k \cdot \text{id}_V)$ *Eigenraum* von f zum Eigenwert k , falls k ein Eigenwert von f ist.

Bemerkungen:

- zu (i) Die Eigenvektoren zu einem Eigenwert von f sind genau die Lösungen $\neq 0$ des homogenen LGS $f(v) - k \cdot v = 0$.
- zu (ii) Der Eigenraum zum Eigenwert k besteht gerade aus allen Eigenvektoren zum Eigenwert k und dem Nullvektor.

Diagonalisierbare lineare Abbildung (6.56)

Def. 6.56

Sei V ein K -Vektorraum und $f : V \rightarrow V$ linear, dann heißt f *diagonalisierbar*, falls eine Basis B von V existiert, welche aus Eigenvektoren von f besteht. Ist $\dim V = n (< \infty)$ so bedeutet dies, dass

$$A_f^{B,B} = \begin{pmatrix} k_1 & & 0 \\ & \ddots & \\ 0 & & k_n \end{pmatrix}$$

eine Diagonalmatrix ist mit den Eigenwerten k_i von f .

Beispiele:

- $f : V \rightarrow V$ hat genau dann den Eigenwert 0, wenn $\ker f \neq \{0\}$ (d.h. f ist nicht injektiv).
- $\text{id}_V : V \rightarrow V$ hat lediglich den Eigenwert 1. Der zugehörige Eigenraum ist ganz V .

Bemerkung:

f ist diagonalisierbar genau dann, wenn $A_f^{B,B}$ ähnlich zu einer Diagonalmatrix bezüglich einer beliebigen Basis B ist.

Bestimmung von Eigenwerten (6.57)

Satz 6.57

Sei V ein n -dimensionaler K -Vektorraum, $f : V \rightarrow V$ linear mit $A = A_f^{B,B} \in \mathcal{M}_{n,n}(K)$ als zugehöriger Darstellungsmatrix bezüglich einer Basis B von V , dann

- (i) ist $k \in K$ genau dann ein Eigenwert, wenn $\det(A - k \cdot E_n) = 0$.
- (ii) sind die Hauptdiagonalelemente a_{11}, \dots, a_{nn} genau die Eigenwerte von f , wenn A eine Dreiecksmatrix ist.

Beweis:

- ad (i) k ist Eigenwert von $f \Leftrightarrow \ker(f - k \cdot \text{id}_V) \neq \{0\}$
 $\Leftrightarrow f - k \cdot \text{id}_V$ nicht injektiv
 $\Leftrightarrow f - k \cdot \text{id}_V$ nicht surjektiv
 $\Leftrightarrow \text{rg}(A - k \cdot E_n) < n$
 $\Leftrightarrow \det(A - k \cdot E_n) = 0$ (6.52)

ad (ii) Nach 6.49d gilt

$$\det(A - k \cdot E_n) = \prod_{i=1}^n (a_{ij} - k)$$

Sukzessives Anwenden von (i). ■

Charakteristisches Polynom (6.58)

Def. 6.58

Ist $A \in \mathcal{M}_{n,n}(K)$ quadratisch, so nennt man $\chi_A := \det(x \cdot E_n - A)$ das *charakteristische Polynom* von A . (Die Einträge der Matrix $x \cdot E_n - A$ sind Polynome.) $\chi_A \in K[x]$ ist ein Polynom vom Grad n mit einem Leitkoeffizient (Koeffizient vor x^n) von 1.

08.11.2006

Bemerkung:

Entsprechend für einen n -dimensionalen K -Vektorraum V , $f : V \rightarrow V$ linear: $\chi_f := \det(x \cdot E_n - A_f^{B,B})$ "charakteristisches Polynom" von f (mit einer beliebigen Basis B von V).

Korollar (6.59)

Kor. 6.59

Es gilt: Die Eigenwerte von f sind genau die Nullstellen des charakteristischen Polynoms χ_f .

Beweis folgt aus Satz 6.57 und $\det(x \cdot E_n - A) = (-1)^n \cdot \det(A - x \cdot E_n)$.

Beispiele:

1. $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\chi_A = \det \begin{pmatrix} x & -1 \\ 1 & x \end{pmatrix} = x^2 + 1$

Über $K = \mathbb{R}$ hat χ_A keine Nullstellen, d.h. A hat keine reellen Eigenwerte. Über $K = \mathbb{C}$ hat χ_A die beiden Nullstellen $\pm i$. Seien v_1, v_2 die zu $\pm i$ gehörigen Eigenvektoren, dann bilden $\{v_1, v_2\}$ eine Basis von \mathbb{C}^2 . Somit hat bezüglich dieser Basis die lineare Abbildung

$$f : \mathbb{C}^2 \rightarrow \mathbb{C}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \end{pmatrix}$$

die Darstellungsmatrix

$$A_f^{B,B} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

Folglich ist f (komplexwertig) diagonalisierbar.

2. $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \chi_B = \det \begin{pmatrix} x & -1 \\ 0 & x \end{pmatrix} = x^2$

hat genau den Eigenwert 0. Eigenraum zum Eigenwert 0:

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in K^2 \mid B \begin{pmatrix} x \\ y \end{pmatrix} = 0 \right\} = \left\{ \begin{pmatrix} c \\ 0 \end{pmatrix} \mid c \in K \right\}$$

ist eindimensional.

⇒ Es existiert keine Basis aus Eigenvektoren. Daher ist B nicht diagonalisierbar (über jedem Körper K).

3. $C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \chi_C = x^2$

C ist Diagonalmatrix.

(Im Allgemeinen lässt sich Diagonalisierbarkeit auch nicht alleine am charakteristischen Polynom ablesen.)

Es lässt sich zeigen:

Trigonalisierbarkeit (6.60)

Satz 6.60

Sei V ein n -dimensionaler K -Vektorraum, $f : V \rightarrow V$ linear, $A_f^{B,B}$ bezüglich einer Basis B von V und $\chi = \chi_A = \chi_f$ das zugehörige charakteristische Polynom, dann gilt:

χ zerfällt über K in Linearfaktoren (d.h. $\chi = \prod_{i=1}^n (x - a_i)$ (*)) $\Leftrightarrow A$ ist ähnlich zu einer oberen Dreiecksmatrix (d.h. f ist trigonalisierbar).

Bemerkung:

(*) lässt sich durch Vergrößern von K immer erreichen. Insbesondere ist jede quadratische Matrix über \mathbb{C} trigonalisierbar.

$$(\det(x \cdot E_n - A) \in K[x])$$

Was passiert beim Einsetzen von A in χ_A ?

Es gilt der tiefliegende Sachverhalt:

Cayley-Hamilton (6.61)

Satz 6.61

Für jede quadratische Matrix $A \in \mathcal{M}_{n,n}(K)$ gilt: $\chi_A(A) = 0$ (Nullmatrix!)

Beweis ist nachzulesen in "Lineare Algebra" (Fischer).

Minimalpolynom (6.62)

Def. 6.62

Ein Polynom

$$p = \sum_{i=0}^t \dots \in K[x]$$

heißt *normiert*, falls $a_t = 1$.

Sei $A \in \mathcal{M}_{n,n}(K)$, dann existieren nach 6.61 normierte Polynome $p \in K[x]$ mit $P(A) = 0$.

Das normierte Polynom $p \in K[x]$ mit minimalem Grad t und $p(A) = 0$ ist eindeutig bestimmt [denn für jedes weitere Polynom q gilt: $(p - q)(A) = p(A) - q(A) = 0$ und der Grad von $p - q$ ist kleiner als t , also $\tilde{t} = 0$ wegen der Minimalität von t .]

Man nennt dieses Polynom das *Minimalpolynom* von A (über K).

Schreibweise: μ_A, m_A

Entsprechend: m_f für eine lineare Abbildung $f : V \rightarrow V$

14.11.2006

Rekapitulation:

Sei $A \in \mathcal{M}_{n,n}(K)$. Das charakteristische Polynom ist $\chi_A := \det(x \cdot E_n - A)$. Das Minimalpolynom m_A : normiertes Polynom $p \in K[x]$ mit minimalem Grad, so dass $p(A) = 0$.

Zusammenhang zwischen charakteristischem und Minimalpolynom (6.63)

Lem. 6.63

Sei $A \in \mathcal{M}_{n,n}(K), p \in K[x]$, dann gilt: $p(A) = 0 \Leftrightarrow m_A \mid p$, sowie insbesondere $m_A \mid \chi_A$ und m_A hat die gleichen Nullstellen in K wie χ_A .

Beweis:

« \Leftarrow » $m_A \mid p \stackrel{MI}{\Leftrightarrow} p = q \cdot m_A$ für ein $q \in K[x]$. $\Rightarrow p(A) = q(A) \cdot \underbrace{m_A(A)}_0 = 0$.

« \Rightarrow » Sei $p(A) = 0$, dann existieren $q, r \in K[x]$ mit $p = q \cdot m_A + r$ und $\deg(r) < \deg(m_A)$ (MI). Somit gilt: $0 = p(A) = q(A) \cdot \underbrace{m_A(A)}_0 + r(A) = r(A)$. $\Rightarrow r = 0$. Ergo $m_A \mid p$.

«Insb.» Wegen Satz 6.61 folgt $m_A \mid \chi_A$, daher ist jede Nullstelle von m_A in K auch Nullstelle von χ_A . Umgekehrt sei $k \in K$ und $\chi_A(k) = 0$. Dann ist k ein Eigenwert der Matrix von A , also $A \cdot v = k \cdot v$ für $v \in K^n \setminus \{0\}$. Sei $m_A = \sum_{i=0}^t a_i \cdot x^i$ mit $a_i \in K$. Dann gilt

$$0 = \underbrace{m_A(A)}_0 \cdot v = \left(\sum_{i=0}^t a_i \cdot A^i \right) \cdot v = \sum_i a_i A^i \stackrel{\text{bungsaufgabe}}{=} \sum_i a_i k^i v = \left(\sum_i a_i k^i \right) v = m_A(k) v \Rightarrow m_A(k) = 0 \quad \blacksquare$$

Beispiele: Man findet m_A unter den (endlich vielen) Teilern von χ_A .

- $A = E_n : \chi_A = (x - 1)^n, m_A = x - 1$
- $A = 0$ (Nullmatrix, Dimension n) : $\chi_A = x^n, m_A = x$.
- $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} : \chi_A = x^2, m_A = x^2$ (beachte $m_A \neq 1$ und $m_A \neq x$)

Es lässt sich zeigen:

Charakterisierung der Diagonalisierbarkeit (6.64)

Satz 6.64

Eine Matrix $A \in \mathcal{M}_{n,n}(K)$ ist genau dann diagonalisierbar (über K), wenn $m_A = \prod_{i=1}^t (x - a_i)$ mit paarweise verschiedenen $a_i \in K$ (mit anderen Worten: wenn m_A in paarweise verschiedene Linearfaktoren zerfällt).

15.11.2006

Beispiel:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, m_A = x^2$$
$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, m_B = (x - 1)^2$$

Nach 6.64 sind A und B offensichtlich nicht diagonalisierbar (über jedem Körper).

Korollar (6.65)

Kor. 6.65

Hat die Matrix $A \in \mathcal{M}_{n,n}(K)$ genau n verschiedene Eigenwerte, so ist sie nicht diagonalisierbar.

Beweis:

Gegeben sind n verschiedene Eigenwerte von A .
 $\Rightarrow \chi_A$ ist das Produkt von n verschiedenen Linearfaktoren.
Nach 6.63 gilt: $\chi_A = m_A$. Mit 6.64 folgt, dass A diagonalisierbar ist. ■

Jordan-Kästchen (6.66)

Def. 6.66

Sei $n \in \mathbb{N}$, $k \in K$, dann heißt die $(n \times n)$ -Matrix

$$J := J(n, k) = \begin{pmatrix} k & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & k \end{pmatrix}$$

das $(n \times n)$ -te Jordan-Kästchen zum Eigenwert k .

Es gilt: $\chi_J = m_J = (x - k)^n$ und der Eigenraum $\ker(J - k \cdot E_n)$ ist eindimensional.

Jordan'sche Normalform (6.67)

Satz 6.67

Sei $A \in \mathcal{M}_{n,n}(K)$ und χ_A zerfällt über K in Linearfaktoren (stets zu erreichen durch Vergrößerung von K), dann ist A ähnlich zu einer "Blockmatrix"

$$\begin{pmatrix} J(n_1, k_1) & 0 & \dots & 0 \\ 0 & J(n_2, k_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J(n_r, k_r) \end{pmatrix}$$

von Jordan-Kästchen $J(n_i, k_i)$ mit $1 \leq i \leq r$. Diese Jordan-Kästchen sind durch A eindeutig bestimmt (bis auf Umnummerierung). Man nennt eine solche "Blockmatrix" auch *Jordan'sche Normalform*.

Es gilt: A ist diagonalisierbar $\Leftrightarrow n_i = 1$ für alle i .

Beweis nachzulesen in "Lineare Algebra" (Fischer).

Ferner gilt:

- $\chi_A = \prod_{i=1}^r (x - k_i)^{n_i}$ mit $n = \sum_{i=0}^r n_i$
- Sind a_1, \dots, a_t ($t \leq r$) die verschiedenen Nullstellen von χ_A , so gilt: $m_A = \prod_{i=1}^t (x - a_i)^{l_i}$ mit l_i als der Größe eines größten Jordan-Kästchens zum Eigenwert a_i .
- Die Anzahl der Jordan-Kästchen zu Eigenwerten $k_i = n - \text{rg}(A - k_i \cdot E_n)$

Bemerkung:

Für $K = \mathbb{C}$ sind zwei Matrizen genau dann ähnlich zueinander, wenn sie die gleiche Jordan'sche Normalform haben (bis auf Umnummerierung).

1.6 Euklidische Vektorräume

Motivation:

- Modell z.B. der realen Welt, für das ein "Abstandsbegriff" benötigt wird.
- Beschreibung von Winkeln und der Orthogonalität.

Bilinearform (6.68)

Def. 6.68

Sei V ein K -Vektorraum. Eine Abbildung $\beta : V \times V \rightarrow K$ heißt *Bilinearform*, falls für jeden Vektor $v \in V$ die beiden Abbildungen

$$\beta(v, \cdot) : V \rightarrow K, x \mapsto \beta(v, x) \quad \text{und} \quad \beta(\cdot, v) : V \rightarrow K, x \mapsto \beta(x, v)$$

linear sind.

21.11.2006

Ferner gilt:

- β ist symmetrisch $\Leftrightarrow \beta(x, y) = \beta(y, x) \forall x, y \in V$ [Notation in HWK u.a.: $\langle x, y \rangle, \langle x | y \rangle, (x, y)$ statt $\beta(x, y)$]
- $x, y \in V$ sind *orthogonal* (bezügl. β), falls $\beta(x, y) = 0$

Beispiel:

- a) VR: $\mathcal{M}_{n,1}(K)$, $\beta(x, y) = x^t \cdot y = \sum_{i=1}^n x_i \cdot y_i$ ist eine symmetrische Bilinearform (BLF) («Standard-Skalarprodukt»). Allgemeiner: Ist $A \in \mathcal{M}_{n,n}(K)$, so ist $\beta(x, y) = x^t \cdot A \cdot y = \sum_{i=1}^n \sum_{j=1}^n x_i \cdot a_{ij} \cdot y_j$ eine Bilinearform (β ist symmetrisch $\Leftrightarrow A = A^t$, d.h. A symmetrisch)
- b) VR: $C[0, 1] = \{f \mid f : [0, 1] \rightarrow \mathbb{R} \text{ stetig}\}$ ¹. Dann gilt: $\beta(f, g) = \int_0^1 f(x) \cdot g(x) dx$ ist eine symmetrische Bilinearform (leichte Übung: nachrechnen)

Cauchy-Schwarz'sche Ungleichung (6.69)

Satz 6.69

Sei V ein \mathbb{R} -VR und $\beta : V \times V \rightarrow K$ eine symmetrische Bilinearform mit der zusätzlichen Eigenschaft $\beta(x, x) > 0$ für $x \in V \setminus \{0\}$ (d.h. β ist *positiv definiert*). Nun gilt für alle $x, y \in V$: $\beta(x, y)^2 \leq \beta(x, x) \cdot \beta(y, y)$. (« \Rightarrow » Gleichheit gilt genau dann, wenn x, y linear abhängig sind).

Beweis:

Seien x, y linear unabhängig, also gilt insbesondere $y \neq 0$. Definiere $r := \frac{\beta(x, y)}{\beta(y, y)}$, dann ist $x + y \cdot r \neq 0$. Somit

$$0 < \beta(x + y \cdot r, x + y \cdot r) \stackrel{BLF}{=} \beta(x, x) + 2 \cdot r \beta(x, y) + r^2 \cdot \beta(y, y) = \beta(x, x) - 2 \cdot \frac{\beta(x, y)^2}{\beta(y, y)} + \frac{\beta(x, y)^2}{\beta(y, y)} = \beta(x, x) - \frac{\beta(x, y)^2}{\beta(y, y)}$$

also $\beta(x, y)^2 < \beta(x, x) \cdot \beta(y, y)$ (unter der Voraussetzung, dass x und y linear unabhängig sind). Aufgrund der positiven Definiertheit gilt

$$\beta(x + y \cdot r, x + y \cdot r) = 0 \Leftrightarrow x + y \cdot r = 0 \Leftrightarrow x, y$$

linear abhängig. ■

Beispiele:

- a) $\left(\sum_{i=1}^n x_i \cdot y_i \right)^2 \leq \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i^2 \right)$ (vergleiche mit Beispiel a) «Standard-Skalarprodukt»)
- b) $\left(\int_0^1 f(x) \cdot g(x) dx \right)^2 \leq \left(\int_0^1 f(x)^2 dx \right) \cdot \left(\int_0^1 g(x)^2 dx \right)$ mit $f, g \in C[0, 1]$

Euklidischer Vektorraum (6.70)

Def. 6.70

Sei V ein endlich dimensionaler \mathbb{R} -VR und $\beta : V \times V \rightarrow \mathbb{R}$ eine positiv definierte, symmetrische Bilinearform, dann heißt das Paar (V, β) ein *euklidischer Vektorraum*. Man definiert ferner die (euklidische) *Norm* von $x \in V$ via $\|x\| := \sqrt{\beta(x, x)}$ und den (euklidischen) *Abstand* von $x, y \in V$ via $d(x, y) := \|x - y\|$.

Es gilt:

¹ C steht für stetig! «Continues»

1. die sogenannte Dreiecksungleichung $\|x + y\| \leq \|x\| + \|y\|$ für $x, y \in V$

$$\begin{aligned} \|x + y\|^2 &= \beta(x+y, x+y) \stackrel{BLF}{=} \beta(x, x) + 2 \cdot \beta(x, y) + \beta(y, y) \stackrel{\text{Satz 6.69}}{\leq} \|x\|^2 + 2 \cdot \|x\| \cdot \|y\| + \|y\|^2 = \\ &= \|x + y\|^2 \text{] und } d(x, z) \leq d(x, y) + d(y, z) \text{ für } x, y, z \in \\ \|x - z\| &= \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| \end{aligned}$$

Dreiecksungleichung

2. Satz des Pythagoras: $\|x + y\|^2 - \|x\|^2 - \|y\|^2 = 2 \cdot \beta(x, y)$ für $x, y \in V$. Daher sind $x, y \in V$ genau dann orthogonal, wenn $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

3. Wir können Winkel in (V, β) einführen: für $x, y \in V \setminus \{0\}$ gilt $-1 \leq \frac{\beta(x, y)}{\|x\| \cdot \|y\|} \leq +1$ (Satz 6.69!). Daher existiert genau eine reelle Zahl α mit $0 \leq \alpha \leq \pi$ und $\cos \alpha = \frac{\beta(x, y)}{\|x\| \cdot \|y\|}$. Wir nennen α den Winkel zwischen x und y . (Es gilt: $\alpha = \frac{\pi}{2} \Leftrightarrow \beta(x, y) = 0 \Leftrightarrow x$ und y orthogonal)

28.11.2006

Orthonormalbasis (6.71)

Def. 6.71

Sei (V, β) ein euklidischer Raum. Eine Basis B heißt *Orthonormalbasis* (in Zukunft auch ONB), falls für Vektoren $b, b' \in B$ gilt:

$$\beta(b, b') = \begin{cases} 1, & b = b' \quad (\text{'normal' (Norm=1)}) \\ 0, & b \neq b' \quad (\text{'orthogonal'}) \end{cases}$$

Man kann leicht zeigen, dass $\forall v \in V$ gilt:

- $v = \sum_{b \in B} \beta(v, b) \cdot b$
- $\beta(v, v) = \sum_{b \in B} \beta(v, b)^2$ («einfach mal nachrechnen»)

Orthonormalisierungsverfahren nach Gram-Schmidt (6.72)

Satz 6.72

Sei (V, β) ein euklidischer Raum und U ein Unterraum von V sowie $\{b_1, b_2, \dots, b_t\}$ sei eine Orthonormalbasis von U . Sei $v \in V \setminus U$ beliebig und

$$v' := v - \sum_{i=1}^t \beta(v, b_i) \cdot b_i$$

, dann ist die Menge $\{b_1, b_2, \dots, b_t, \frac{v'}{\|v'\|}\}$ eine Orthonormalbasis von $\langle U, v \rangle$. Somit lässt sich jede Orthonormalbasis von U zu einer Orthonormalbasis von V ergänzen. Insbesondere existiert eine Orthonormalbasis von V .

Beweis:

Es gilt $v' \neq 0$ (da $v \in V \setminus U$). Ferner gilt, dass $\beta(v', b_i) = \beta(v, b_i) - \beta(\sum_{j=1}^t \beta(v, b_j) b_j, b_i) = \beta(v, b_i) -$

$$\sum_{j=1}^t \beta(v, b_j) \cdot \underbrace{\beta(b_j, b_i)}_{\delta_{ij}} = \beta(v, b_i) - \beta(v, b_i) \cdot 1 = 0 \text{ für } 1 \leq i \leq t.$$

² Kroneckersymbol: $\delta_{ij} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

$$\beta\left(\frac{v'}{\|v'\|}, \frac{v'}{\|v'\|}\right) = \frac{1}{\|v'\|^2} \cdot \beta(v', v') = \frac{\|v'\|^2}{\|v'\|^2} = 1. \Rightarrow \{b_1, b_2, \dots, b_t, \frac{v'}{\|v'\|}\} \text{ ist eine Orthonormalbasis von } \langle U, v \rangle = \langle U, v' \rangle.$$

Durch Wiederholung des Verfahrens erhält man eine Orthonormalbasis von V . Mit $U = \emptyset$ erhält man die Existenz einer Orthonormalbasis von V . ■

Es gilt:

Satz über die Hauptachsentransformation (6.73)

Satz 6.73

Sei $A \in \mathcal{M}_{n,n}(\mathbb{R})$ symmetrisch, dann hat der euklidische Raum \mathbb{R}^n eine Orthonormalbasis, die aus den Eigenvektoren der Matrix A besteht. Insbesondere ist A diagonalisierbar (über \mathbb{R}).

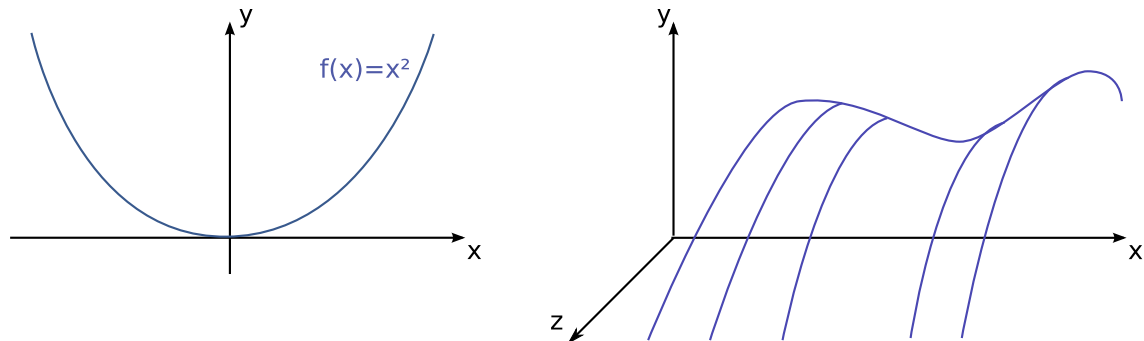
2 Mehrdimensionale Analysis (Funktionen mehrerer Variablen)

Literatur:

- Forster, Analysis 2, Vieweg
- Königsberger, Analysis 2, Springer

Motivation:

- Oft werden (in der Bio-/Informatik) Funktionen mit mehreren Veränderlichen betrachtet.



- Grundlagen der *Topologie*

2.1 Metrische Räume

Metrischer Raum (7.1)

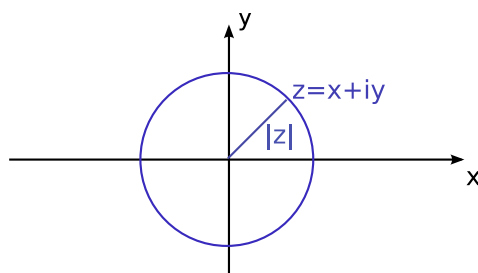
Def. 7.1

Sei X eine Menge und wir sagen eine Abbildung $d : X \times X \rightarrow \mathbb{R}$ heißt eine *Metrik* auf X , falls für $x, y, z \in X$ gilt:

- (i) $d(x, y) \geq 0$ und $d(x, y) = 0 \Leftrightarrow x = y$ (positive Definitheit)
- (ii) $d(x, y) = d(y, x)$ (Symmetrie)
- (iii) $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecksungleichung)

Beispiele:

- a) $(\mathbb{R}, |\cdot|)$ ist ein metrischer Raum mit $d(x, y) := |x - y|$
- b) $(\mathbb{C}, |\cdot|)$ ist ein metrischer Raum mit $d(z, w) := |z - w|$
 $\mathbb{C} \ni \underbrace{z}_{x+iy} : |z| = \sqrt{x^2 + y^2} \in \mathbb{R}$



29.11.2006

Normierter Raum (7.2)

Def. 7.2

Sei V ein \mathbb{R} -Vektorraum (bzw. \mathbb{C} -Vektorraum). Eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ heißt *Norm* auf V , und dann $(V, \|\cdot\|)$ *normierter reeller (bzw. komplexer) Vektorraum*, falls gilt:

- (i) $\|x\| \geq 0$ und $\|x\| = 0 \Leftrightarrow x = 0$ («positive Definitheit»)
- (ii) $\|c \cdot x\| = |c| \cdot \|x\|$ ($c \in \mathbb{R}$ bzw. \mathbb{C}) («Symmetrie»)
- (iii) $\|x - y\| \leq \|x - z\| + \|z - y\|$ («Dreiecksungleichung»)

Dann ist V ein metrischer Raum via der induzierten Metrik: $(x, y) \mapsto \|x - y\|$.

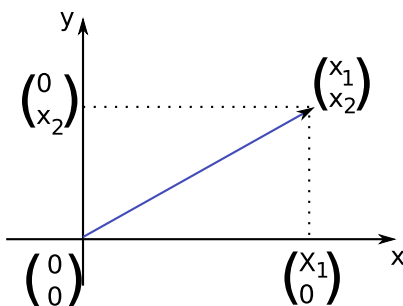
Beispiele:

- a) \mathbb{R}^n ist ein normierter Raum Vektorraum bezüglich folgender Normen:

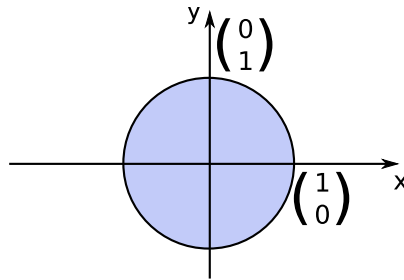
- (i) *Euklidische Norm (L^2 -Norm).*

$$\mathbb{R}^n \ni x : \|x\|_2 := \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}$$

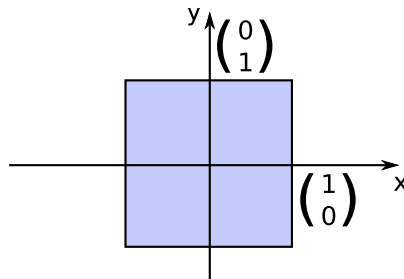
$n = 2:$



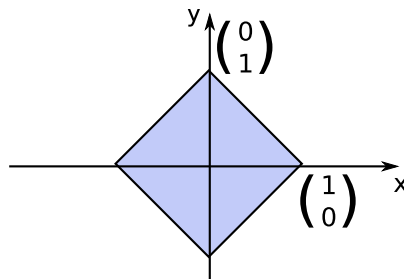
Offene 1-Kugel um $\mathcal{O} := \{x \in \mathbb{R}^n \mid \|x\|_2 < 1\}$



- (ii) *Supremums-Norm: (L^∞ -Norm).*
 $\|x\|_\infty := \sup_{1 \leq i \leq n} |x_i| = \max_{1 \leq i \leq n} |x_i|$
 Offene 1-Kugel um $\mathcal{O} := \{x \in \mathbb{R}^n \mid \|x\|_\infty < 1\}$



- (iii) *Summen-Norm: (L^1 -Norm):*
 $\|x\|_1 := \sum_{i=1}^n |x_i|$
 Offene 1-Kugel um $\mathcal{O} := \{x \in \mathbb{R}^n \mid \|x\|_1 < 1\}$



$\mathbb{R}^n, x \in \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Es gilt $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \cdot \|x\|_\infty$. (d.h. alle Normen sind äquivalent (mit anderen Worten: «alle gleich gut»))

b) $\underbrace{C[0,1]}_{\{f|f:[0,1] \rightarrow \mathbb{R} \text{ ist stetig}\}}$ ist ein normierter (reeller) Vektorraum bezüglich:

- (i) Euklidischer Norm (L^2 -Norm):
 $C[0,1] \ni f : \|f\|_2 := \left(\int_0^1 |f(x)|^2 dx \right)^{\frac{1}{2}}$

- (ii) Supremums-Norm (L^∞ -Norm):

$$\|f\|_\infty := \sup_{0 \leq x \leq 1} |f(x)| = \max_{0 \leq x \leq 1} |f(x)|$$
- (iii) L^1 -Norm:

$$\|f\|_1 := \int_0^1 |f(x)| dx$$

Bemerkung:

Die L^2 -Normen resultieren von positiv definiten, symmetrischen Bilinearformen über \mathbb{R} (Kapitel 6.6).

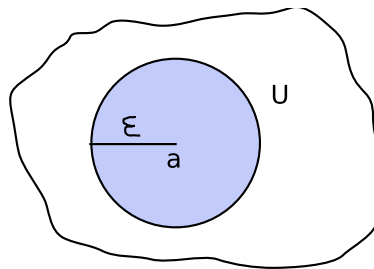
$$\|x\| := \sqrt{\beta(x, x)} = (x^t \cdot x)^{\frac{1}{2}} = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}} \text{ Standardskalarprodukt bezüglich der Standardbasis}$$

Offene Menge (7.3)

Def. 7.3

Sei (X, d) ein metrischer Raum, $a \in X$ sowie $U \subseteq X$ ein Unterraum; $0 \leq r \in \mathbb{R}$.

- a) $U_r(a) := \{x \in X \mid d(a, x) < r\}$ ist die «offene Kugel» um a mit dem Radius r (in X).
- b) U heißt *Umgebung* von a : $\Leftrightarrow \exists \epsilon > 0 : U_\epsilon(a) \subseteq U$.



- c) U heißt *offen* in X ($U \underset{off}{\subseteq} X$) : $\Leftrightarrow U$ ist Umgebung jeder seiner Punkte $\Leftrightarrow \forall x \in U \exists \epsilon = \epsilon(x) > 0 : U_\epsilon(x) \subseteq U \Leftrightarrow \forall x \in U \exists \epsilon = \epsilon(x) > 0 : U_\epsilon(x) \subseteq U$

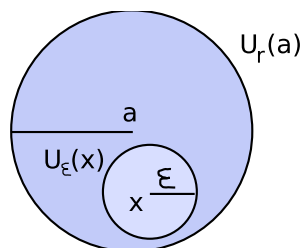
Lemma 7.4

Lemma 7.4

- (i) $U_r(a) \underset{off}{\subseteq} X$ (insbesondere sind offene Intervalle in $(\mathbb{R}, |\cdot|)$ offen.)
- (ii) $\emptyset, X \underset{off}{\subseteq} X$
- (iii) Sind $U_i \underset{off}{\subseteq} X$ ($i \in I$), so ist $\bigcup_{i \in I} U_i \underset{off}{\subseteq} X$
- (iv) Sind $U_1, \dots, U_n \underset{off}{\subseteq} X$, so $\bigcap_{i=1}^n U_i \underset{off}{\subseteq} X$.

Beweis:

nur (i): Sei $x \in U_r(a)$. z.Z. $\exists \epsilon > 0 : U_\epsilon(x) \subseteq U_r(a)$



. Setze $\epsilon := r - d(a, x) > 0$. Es bleibt zu zeigen, dass $U_\epsilon(x) \subseteq U_r(a)$. D.h. für jedes $y \in U_\epsilon(x)$ gilt $d(a, y) < r$. Das ist auch so, denn es gilt die Dreiecksungleichung:
 $d(a, y) \leq d(a, x) + d(x, y) < d(a, x) + \epsilon = r$ ■

05.12.2006

Abgeschlossene Mengen (7.5)

Def. 7.5

Sei (X, d) ein metrischer Raum und $a \in X$ sowie $Y \subseteq X$ eine Teilmenge. Dann:

- (i) Y abgeschlossen in X ($Y \subseteq_{abg.} X$) $\Leftrightarrow X \setminus Y := \{x \in X \mid x \notin Y\} \subseteq_{off} X$.
- (ii) a heißt *Berührungspunkt* von Y (in X) $\Leftrightarrow \forall \epsilon > 0 : U_\epsilon(a) \cap Y \neq \emptyset$

Bezeichnung:

\bar{Y} := Menge aller Berührungspunkte Y (in X).

Klar ist, dass $\bar{\bar{Y}} \supseteq Y$ und $\bar{\bar{\bar{Y}}} = \bar{Y}$.

Lemma 7.6:

Lemma 7.6

Sei (X, d) ein metrischer Raum und $Y \subseteq X$, dann gilt $Y \subseteq_{abg.} X \Leftrightarrow \bar{Y} = Y$.

Beweis:

« \Rightarrow » Nach Voraussetzung $X \setminus Y \subseteq_{off} X$. Wir nehmen jetzt an, dass die Menge der Berührungspunkte \bar{Y} echt größer ist, als Y : $\bar{Y} \supset Y$. In diesem Fall $\exists x \in \bar{Y} \setminus Y \subseteq X \setminus Y \subseteq_{off} X$ und daher $\exists \epsilon > 0 : U_\epsilon(x) \subseteq X \setminus Y$. Folglich ist $U_\epsilon(x) \cap Y = \emptyset$. Widerspruch, denn x ist ein Berührungspunkt ($x \in \bar{Y}$)!

« \Leftarrow » Sei $x \in X \setminus Y = X \setminus \bar{Y}$. Zu zeigen ist: $\exists \epsilon > 0 : U_\epsilon(x) \subseteq X \setminus Y$ ($\Rightarrow Y \subseteq_{abg.} X$)
 Da x kein Berührungspunkt von Y (in X) ist, $\exists \epsilon > 0 : U_\epsilon(x) \cap Y = \emptyset$, also $U_\epsilon(x) \subseteq X \setminus Y$ ■

2.2 Grenzwerte und Stetigkeit

Grenzwert und Häufungswert (7.7)

Def. 7.7

Sei (X, d) ein metrischer Raum und $(x_n)_{n \in \mathbb{N}}$ eine Folge in X sowie $a \in X$.

- (i) (x_n) konvergiert gegen a $\Leftrightarrow \forall \epsilon > 0 \exists N \in \mathbb{N} : \underbrace{d(a, x_n)}_{x_n \in U_\epsilon(a)} < \epsilon$ für alle $n \geq N$. Dann ist $a = \lim_{x \rightarrow \infty} x_n$ eindeutig bestimmt und heißt «Grenzwert». (vgl. Kapitel 5.1 (Mathematik für Informatiker und Bioinformatiker II))
- (ii) a ist der Häufungswert von (x_n) \Leftrightarrow In jeder $U_\epsilon(a)$ liegen unendlich viele Folgenglieder von (x_n) .
 $[a^0 \leftarrow (x_{n_i})_i \subset (x_n)]$

Stetigkeit (7.8)

Def. & Satz 7.8

Seien (X, d) und (X', d') zwei metrische Räume, $f : X \rightarrow X'$ und $a \in X$.

- (i) f stetig in a $\Leftrightarrow \forall \epsilon > 0 \exists \delta > 0 : d'(f(x), f(a)) < \epsilon$ für $d(a, x) < \delta \Leftrightarrow \forall \epsilon > 0 \exists \delta > 0 : f(U_\delta(a)) \subseteq U'_\epsilon(f(a)) \Leftrightarrow$ Ist U' eine Umgebung des Bildes $f(a)$ in X' ist, so ist das Urbild $U := \underbrace{f^{-1}(U')}_{\{x \in X \mid f(x) \in U'\}}$ Umgebung von a in X .
Schreibweise: $\lim_{x \rightarrow a} f(x) = f(a)$.
- (ii) f ist stetig (auf X , das ist eine globale Eigenschaft) $\Leftrightarrow f$ stetig in allen Punkten von X
 \Leftrightarrow das Urbild $f^{-1}(U')$ jeder offenen Menge $U' \underset{\text{off}}{\subseteq} X'$ ist offen in X .

Beweis (!):

« \Rightarrow » Sei $U' \underset{\text{off}}{\subseteq} X'$. Zu zeigen ist, dass das Urbild von U' offen ist: $U = f^{-1}(U') \underset{\text{off}}{\subseteq} X$.
Wir haben ein U' als Umgebung jedes Punktes von U' in X' . Also ist doch dann U Umgebung jeder seiner Punkte in X , das heißt, dass $U \underset{\text{off}}{\subseteq} X$.

« \Leftarrow » analog... ■

06.12.2006

Folgenstetigkeit (7.9)

Def. 7.9

$f : X \rightarrow X'$ (X, X' seien metrische Räume) ist genau dann stetig in $a \in X$, falls für jede Folge (x_n) in X mit $\lim_n x_n = a$ gilt: $\lim_n f(x_n) = f(a)$.

Bemerkung:

Kompositionen stetiger Funktionen sind stetig. Ebenso gilt, wenn $f, g : X \rightarrow \mathbb{R}$ stetig sind, so sind auch $f + g, f \cdot g$ und $\frac{f}{g}$ mit $(g \neq 0)$.

Lemma (7.10)

Lemma 7.10

Sei (X, d) ein metrischer Raum $Y \subseteq X$, dann ist $Y \stackrel{abg}{\subseteq} X \Leftrightarrow$ (Ist (y_n) eine Folge in Y , die in X konvergiert, so ist $\lim_n y_n \in Y$.)

Beispiele: Sei $(\mathbb{R}^n, \|\cdot\|_2)$, $n \in \mathbb{N}$.

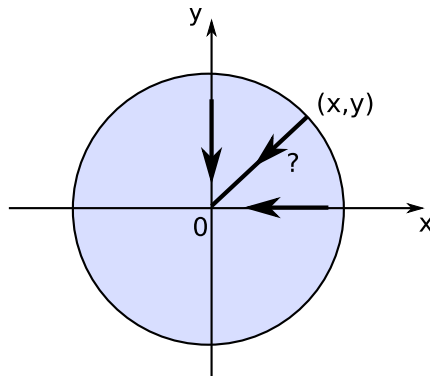
a) $n = 2 : f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = \begin{cases} x \cdot y \cdot \frac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & \text{sonst} \end{cases}$

Behauptung: f ist stetig.

Beweis: Bleibt nur die Steigkeits von f in $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ zu zeigen: Sei $\epsilon > 0!$ Wähle $\delta = \sqrt{2} \cdot \epsilon$

Ist $\|(x, y) - \sigma\|_2 = \sqrt{x^2 + y^2} < \delta$, so gilt $|f(x, y) - 0| = \left| x \cdot y \cdot \frac{x^2 - y^2}{x^2 + y^2} \right| \leq |x \cdot y| \leq \frac{1}{2}(x^2 + y^2) < \epsilon \square$

b) $n = 2 : g : \mathbb{R}^2 \rightarrow \mathbb{R}, g(x, y) = \begin{cases} \frac{x \cdot y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & \text{sonst} \end{cases}$



Behauptung: g ist nicht stetig in $\sigma = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Beweis: Es gilt $g(x, 0) = 0 = g(0, y)$ für $x \neq 0, y \neq 0$. Andererseits: $g(x, x) = \frac{1}{2}$ für $x \neq 0$ zu jedem (noch so kleinen) $\delta > 0 \exists (x, y) \in U_\delta(0)$ mit $|g(x, y)| \geq \frac{1}{2}$.

c) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine \mathbb{R} -lineare Abbildung $(\mathbb{R}^m = (\mathbb{R}^m, \|\cdot\|_2))$ mit der dazugehörigen Darstellungsmatrix $A = A_f$ (bezüglich der Standardbasen). Setze $\|A\|_2 := \left(\sum_{i,j} a_{ij}^2 \right)^{\frac{1}{2}}$

$(L^2$ -Norm auf $M_{n,m}(\mathbb{R})$) entsprechend für L^∞ - und L^1 -Norm. Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ gilt

mit Cauchy-Schwarz: $\|A\|_2^2 = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right)^2 = \left(\sum_{i=1}^m \beta \left(\begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix}^t, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) \right)^2 \stackrel{CS}{\leq}$

$\sum_{i=1}^m \left(\left(\sum_{j=1}^m a_{ij}^2 \right), \left(\sum_{j=1}^m x_j^2 \right) \right) = \left(\sum_{i,j} a_{ij}^2 \right) \|x\|_2^2 = \|A\|_2^2 \cdot \|x\|_2^2$. Folglich gilt: $\|Ax\|_{(2)} \leq$

$$\|A\|_{(2)} \cdot \|x\|_{(2)} \quad (x \in \mathbb{R}^n). \text{ Daher ist } A \text{ stetig (sogar stetig nach Lipschitz): } \|Ax - Ay\|_2 = \|A(x - y)\|_2 \leq \underbrace{\|A\|}_{\text{Konstante}} \cdot \|x - y\|_2 \text{ f\"ur } x, y \in \mathbb{R}^n$$

A ist linear

Vollstandigkeit (7.11)

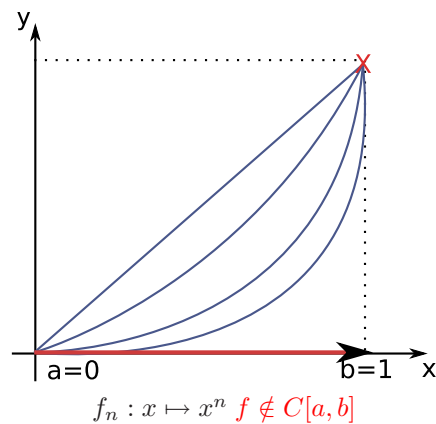
Def. 7.11

Sei (X, d) ein metrischer Raum und (x_n) eine Folge in X .

- (i) (x_n) ist eine *Cauchyfolge* in X $\Leftrightarrow \forall \epsilon > 0 \exists N \in \mathbb{N} : d(x_n, x_m) < \epsilon$ fur alle $n, m \geq N$ ($n, m \in \mathbb{N}$).
- (ii) X heit *vollstandig metrischer Raum* \Leftrightarrow jede Cauchyfolge in X konvergiert in X .
- (iii) *Banachraum* := vollstandiger, normierter Vektorraum
- (iv) *Hilbertraum* := vollstandiger, euklidischer Vektorraum («vollstandiger Raum mit Skalarprodukt»)

Beispiele:

- a) \mathbb{R}, \mathbb{C} mit dem Betrag sind Banachraume
- b) $(C[a, b], \|\cdot\|_\infty)$ ist ein Banachraum. ($a, b \in \mathbb{R}, a < b$)
- c) $(C[a, b], \|\cdot\|_1)$ ist kein Banachraum.



12.12.2006

Es gilt der folgende wichtige Satz:

Satz (7.12)

Satz 7.12

$\mathbb{R}^n = (\mathbb{R}^n, \|\cdot\|_\infty)$ mit $n \in \mathbb{N}$, dann gilt:

- a) \mathbb{R}^n ist ein Banachraum
- b) Jede beschrankte Folge in \mathbb{R}^n hat einen Haufungswert (Bolzano-Weierstrass)
- c) Je 2 Normen auf dem \mathbb{R}^n sind aquivalent

2.3 Mehrdimensionale Analysis

A. Kurven im \mathbb{R}^n

Sei $\mathbb{R}^n = (\mathbb{R}^n, \beta)$ euklidischer Vektorraum der Dimension $n \in \mathbb{N}$. L^2 -Norm $\|x\|_2 = \beta(x, x)^{\frac{1}{2}} = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$ für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$

$D = [a, b]$ kompaktes (d.h. beschränkt und abgeschlossen) Intervall, $U \subseteq \mathbb{R}^n$.

Kurve (7.13)

Def. 7.13

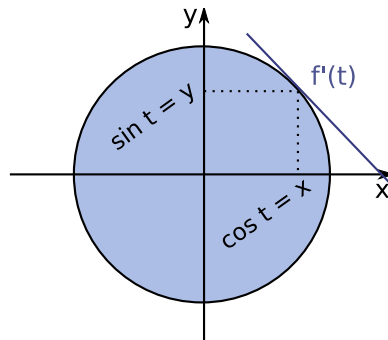
Sei $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : D \rightarrow U$ mit Koordinatenfunktionen $f_i : D \rightarrow \mathbb{R}$. f heißt *Kurve* (Weg) $U \Leftrightarrow f$ stetig (d.h. f_i stetig für $i = 1, \dots, n$)

i) f differenzierbar $\Leftrightarrow f_i$ differenzierbar für $i = 1, \dots, n$ und dann $f'(t) = \begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix} \in \mathbb{R}^n = \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h}$ «Tangentialvektor von f bei t » («Geschwindigkeitsvektor»). f ist eine *glatte (reguläre) Kurve* $\Leftrightarrow f$ stetig differenzierbar (f C^1 -Kurven) und $f'(t) \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ für alle $t \in D$.

ii) f ist integrierbar $\Leftrightarrow f_i$ (Regel-/Riemann-)integrierbar für $i = 1, \dots, n$ und dann $\int_a^b f := \begin{pmatrix} \int_a^b f_1 \\ \vdots \\ \int_a^b f_n \end{pmatrix} \in \mathbb{R}^n$.

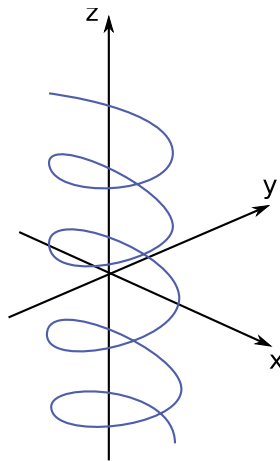
Beispiele:

a) *1-Sphäre*: $n = 2$: $S^1 := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \right\}$ Kurve («Parameterdarstellung»)
 $f : [0, 2\pi] \rightarrow \mathbb{R}^2$, $f(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$
 $f([0, 2\pi]) = S^1$. $f'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. f ist glatt. $\beta \left(\begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}, \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \right) = \begin{pmatrix} -\sin(t) \cdot \cos(t) \\ \cos(t) \cdot \sin(t) \end{pmatrix} = 0$



b) Schraubenlinie: $n = 3$

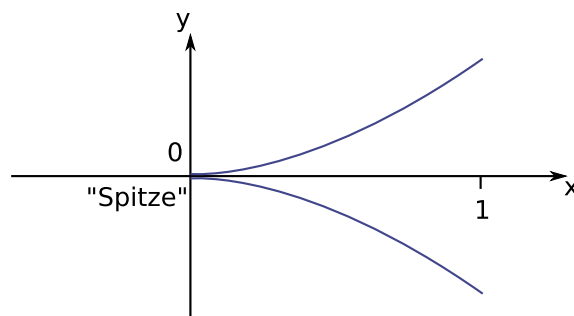
$$f : [0, 2\pi] \rightarrow \mathbb{R}^3, f(t) = \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}, f \text{ glatt}, f'(t) = \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix}$$



c) Neil'sche Parabel: $n = 2$

$$N := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y^2 = x^3, 0 \leq x \leq 1 \right\}$$

$$f : [-1, 1] \rightarrow \mathbb{R}^2, f(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}, f'(t) = \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ nur für } t = 0 \Rightarrow f \text{ nicht glatt}$$



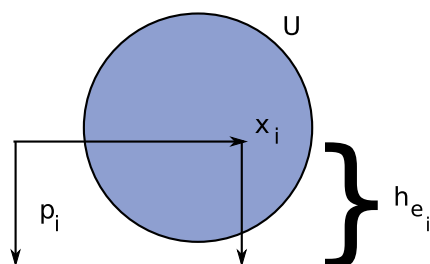
B. Partielle Ableitungen

Bezeichnungen wie in A.

Partielle Ableitung (Def. 7.14)

Def. 7.14

f ist in U in Richtung x_i (besser: e_i), $1 \leq i \leq n$, *partiell differenzierbar*, falls existiert: $D_i f(x) = \frac{\partial f}{\partial x_i}(x) := \lim_{h \rightarrow 0} \frac{f(x+h \cdot e_i) - f(x)}{h}$ ($f(x+h \cdot e_i) = f(x_1, \dots, x_{i-1}, x_i+h, x_{i+1}, \dots, x_n)$)



(Fasse f als Funktion der i -ten Koordinate auf.)

f ist in $x \in U$ *partiell differenzierbar* $\Leftrightarrow D_i f(x)$ existiert für alle $i = 1, \dots, n$ und der *Gradient* $\text{grad} f(x) := (D_1 f(x), \dots, D_n f(x))$.

f *partiell differenzierbar* (auf U) $\Leftrightarrow \text{grad} f : U \rightarrow \mathbb{R}^n$ existiert ($\cong \text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$)

Beispiele:

a)

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = 3x^3 y^2 \quad (x, y \in \mathbb{R})$$

$$D_1 f(x, y) = 9x^2 y^2$$

$$D_2 f(x, y) = 6x^3 y$$

$$\underbrace{D_1 D_1 f(x, y)}_{:= D_1^2} = 18xy^2$$

$$D_2^2 f(x, y) = 6x^3$$

$$D_2 D_1 f(x, y) = 18x^2 y$$

$$D_1 D_2 f(x, y) = 18x^2 y$$

19.12.2006

b)

$$f : \underbrace{\mathbb{R}^3 \setminus \{0\}}_{\substack{U \subseteq \mathbb{R}^3 \\ \text{off.}}} \rightarrow \mathbb{R} \quad (n = 3):$$

$$f(x_1, x_2, x_3) = \frac{1}{\|x\|_2} = (x_1^2 + x_2^2 + x_3^2)^{-\frac{1}{2}}$$

$$D_1 f(x_1, x_2, x_3) = \underbrace{-x_1}_f \cdot \underbrace{(x_1^2 + x_2^2 + x_3^2)^{-\frac{3}{2}}}_g$$

$$D_1^2 f(x_1, x_2, x_3) = - \underbrace{(x_1^2 + x_2^2 + x_3^2)^{-\frac{3}{2}}}_{(2x_1^2 + x_2^2 + x_3^2)} + x_1 \cdot \frac{3}{2} \cdot 2x_1 \cdot (x_1^2 + x_2^2 + x_3^2)^{-\frac{5}{2}} = \dots = (x_1^2 + x_2^2 + x_3^2)^{-\frac{5}{2}}.$$

entsprechend:

$$\begin{aligned}
D_2^2 f(x_1, x_2, x_3) &= (x_1^2 + x_2^2 + x_3^2)^{-\frac{5}{2}} \cdot (-x_1^2 + 2x_2^2 + x_3^2) \\
D_3^2 f(x_1, x_2, x_3) &= (x_1^2 + x_2^2 + x_3^2)^{-\frac{5}{2}} \cdot (-x_1^2 - x_2^2 + 2x_3^2) \\
\Delta f(x_1, x_2, x_3) &:= D_1^2 f(x_1, x_2, x_3) + D_2^2 f(x_1, x_2, x_3) + D_3^2 f(x_1, x_2, x_3) = 0 \text{ ("Laplace-Operator", } f \text{ ist "harmonische Funktion")}
\end{aligned}$$

Bemerkung:

$\text{grad} f(x)$ zeigt in Richtung maximaler Zunahme von f bei x .

Es gilt:

Satz von Schwarz (7.15)

Satz 7.15

Sei $f : U \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar (d.h. $D_i D_j f$ existiert und sind stetig $\forall i, j$), dann gilt: $D_i D_j f = D_j D_i f \quad \forall i, j$.

C. Differenzierbare Abbildungen

$U \underset{\text{off}}{\subseteq} \mathbb{R}^n, f = \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} : U \rightarrow \mathbb{R}^k, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U$ und $\mathbb{R}^n, \mathbb{R}^k, \text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^k) = \mathcal{M}_{n,k}(\mathbb{R})$ werden als euklidische Vektorräume aufgefasst.

Jacobi-Matrix (7.15)

Def. 7.15

Existieren für $f = \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} : U \rightarrow \mathbb{R}^k$ alle partiellen Ableitungen $D_j f_i(x)$ ($j = 1, \dots, n, i = 1, \dots, k$), so definiere die *Jacobi-Matrix*:

$$J_f(x) := \begin{pmatrix} \text{grad} f_1(x) \\ \text{grad} f_2(x) \\ \vdots \\ \text{grad} f_k(x) \end{pmatrix} = \begin{pmatrix} D_1 f_1(x) & \dots & D_n f_1(x) \\ \vdots & & \vdots \\ \vdots & & \vdots \\ D_1 f_k(x) & \dots & D_n f_k(x) \end{pmatrix} \in \mathcal{M}_{k,n}(\mathbb{R})$$

(Totale) Ableitung (7.17)

Def. & Satz 7.17

- f (total) differenzierbar in $x \in U$, falls eine lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^k$ ($A = (a_{ij}) \in \mathcal{M}_{k,n}(\mathbb{R})$) existiert mit: $f(x+k) = f(x) + A \cdot h + o(h)$, wobei $o(h)$ einer Funktion $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ entspricht mit

(i) $\varphi(0) = 0$

(ii) $\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|_2} = 0$ für kurze $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$, d.h. $\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - A \cdot h}{\|h\|_2} = \underbrace{0}_{\mathbb{R}^k}$

dann:

- a) f ist stetig in x .
 b) Es existierten alle partiellen Ableitungen $D_j f_i(x)$ ($i = 1, \dots, k, j = 1, \dots, n$)

Und es gilt eindeutig: $D_j f_i(x) = a_{ij}$ ($i = 1, \dots, k, j = 1, \dots, n$). Somit identifiziere: $A = J_f(x) =: f'(x)$ "(totale) Ableitung von f in x " (eindeutig bestimmt durch f und x).

- f (total) differenzierbar auf U , falls $f'(x)$ existiert $\forall x \in U$ und dann $f' : U \rightarrow \mathcal{M}_{k,n}(\mathbb{R})$.

Bemerkungen:

- Für $n = k = 1$ erhalte die gewohnte Differentiation für Funktionen einer Veränderlichen (vgl. Kap. 5).
- Partielle Ableitungen müssen nicht notwendig stetig sein. Es gilt: $f : U \rightarrow \mathbb{R}^k$ ist stetig differenzierbar (C^1 -Funktion) \Leftrightarrow Alle partiellen Ableitungen $D_j f_i : U \rightarrow \mathbb{R}$ existieren und sind stetig $\forall i, j$.

Kettenregel (7.18)

Satz 7.18

$f : U \rightarrow \mathbb{R}^k, x \in U, f(x) = y \in V \underset{\text{off}}{\subseteq} \mathbb{R}^k$ und $g : V \rightarrow \mathbb{R}^l$. Ist f differenzierbar in x und g differenzierbar in $y = f(x)$, so $g \circ f$ differenzierbar in x und $(g \circ f)'(x) = g'(f(x)) \circ f'(x)$.

Beweis: analog wie in Kap. 5.

Beispiel:

a) $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & , (x, y) \neq (0, 0) \\ 0 & , \text{sonst} \end{cases}$

Gezeigt in 7.9, Beispiel b: f ist nicht stetig in $(0, 0)$. Daher auch nicht (total) differenzierbar in $(0, 0)$ [Satz 7.16]. Aber:

$$D_1 f(x, y) = \begin{cases} \frac{y(x^2+y^2) - 2x^3y}{(x^2+y^2)^2} & , (x, y) \neq (0, 0) \\ 0 & , \text{sonst} \end{cases}$$

Ebenso existiert $D_2 f(x, y)$ überall. Somit: $\text{grad} f(0, 0) = J_f(0, 0) = (0, 0)$ existiert, aber $f'(0, 0)$ existiert nicht.

b) $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, f(x, y) = \begin{pmatrix} e^x \cdot \cos y \\ e^x \cdot \sin y \end{pmatrix}$

Es gilt: $J_f(x, y) = \begin{pmatrix} e^x \cdot \cos y & -e^x \cdot \sin y \\ e^x \cdot \sin y & e^x \cdot \cos y \end{pmatrix} = f'(x, y)$, denn $(x, y) \mapsto e^x \cdot \cos y, \pm e^x \cdot \sin y$ sind stetig. [Wende Bemerkung zu 7.16 an.]

Hesse-Matrix (7.19)

Def. 7.19

Sei $f : U \underset{\text{off}}{\subseteq} \mathbb{R}^n \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann heißt die Matrix der zweiten partiellen Ableitungen

$$H_f = \begin{pmatrix} D_1^2 f & \dots & D_n D_1 f \\ \vdots & \ddots & \vdots \\ D_1 D_n f & \dots & D_n^2 f \end{pmatrix}$$

die Hesse-Matrix von f .

Bemerkung:

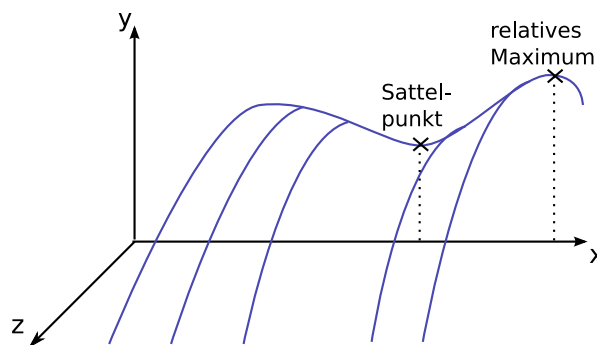
Nach dem Satz von Schwarz (7.15) ist H_f symmetrisch.

Extremwertberechnung: $n = 2$ (7.20)

Anwendung
7.20

Sei $f : U \underset{\text{off}}{\subseteq} \mathbb{R}^2 \rightarrow \mathbb{R}$ zweimal stetig differenzierbar.

- a) *Bestimmung der kritischen Punkte:*
Man sucht alle Zahlenpaare $(x_0, y_0) \in U$, so dass $D_1 f(x, y) = 0$, $D_2 f(x, y) = 0$ simultan erfüllt sind.
- b) *Test der zweiten Ableitung:*
Ist (x_0, y_0) ein kritischer Punkt von f und gilt:
- (i) $\det(H_f(x_0, y_0)) < 0$, so ist (x_0, y_0) *Sattelpunkt*.
 - (ii) $\det(H_f(x_0, y_0)) > 0$ und $D_1^2 f(x_0, y_0) > 0$, so ist (x_0, y_0) ein *lokales Minimum*.
 - (iii) $\det(H_f(x_0, y_0)) > 0$ und $D_1^2 f(x_0, y_0) < 0$, so ist (x_0, y_0) ein *lokales Maximum*.

**Beispiel:**

$$f(x, y) = x + y + \frac{8}{xy}, \quad x, y > 0$$

$$D_1 f(x, y) = 1 - \frac{8}{x^2 y} = 0 \Rightarrow y = \frac{8}{x^2}$$

$$D_2 f(x, y) = 1 - \frac{8}{xy^2} = 0 \Rightarrow 1 - \frac{8 \cdot x^4}{x \cdot 64} = 0 \Rightarrow x^3 = 8, \quad x = 2, \quad y = 2 \text{ (kritischer Punkt)}$$

$$D_1^2 f(x, y) = \frac{16}{x^3 y}$$

$$D_2^2 f(x, y) = \frac{16}{xy^3}, \quad D_1 D_2 f(x, y) = \frac{8}{x^2 y^2}$$

$$\det(H_f(2, 2)) = 1 \cdot 1 - \left(\frac{1}{2}\right)^2 > 0 \text{ und } D_1^2 f(2, 2) = 1 > 0 \Rightarrow (x_0, y_0) = (2, 2) \text{ ist relatives Minimum.}$$

09.01.2007

DER FOLGENDE TEIL DER VORLESUNG GEHÖRT NICHT MEHR ZUM OFFIZIELLEN STOFFUMFANG.

3 Anwendungen

3.1 Lineare Algebra und Google

K. Bryan, T. Leise: The \$ 25.000.000.000 Eigenvector. The Linear Algebra behind Google. SIAM Reviews 48 (3), 2006, S. 569-581.

Suchmaschinen: "wichtige" Webseiten zuerst auflisten.

Besonderheit bei Google: Seitenrang-Algorithmus ("Page Rank Algorithm")

1. Überblick über das Web:
Lokalisiere alle (öffentlich zugänglichen) Webseiten
2. Indiziere die Daten aus 1 (z.B. nach Stichwörtern, Phrasen, ...)
3. Ordne jeder Webseite einen Wert zu, wie "wichtig" diese Seite ist ("Importance score")

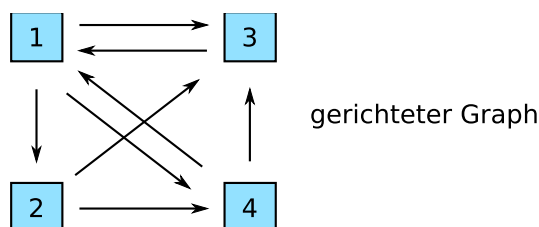
Frage: Wie definiert man "wichtig"?

Hängt sicherlich davon ab, wieviele Links von anderen Seiten auf eine gegebene Seite gemacht werden. Einen Link auf eine Seite nennen wir einen "Backlink". ("Demokratie des Webs": einzelne Seiten geben ihre Stimme ab über Bedeutung einer anderen)

Mathematisierung des Problems:

Annahme: Web hat n Seiten, jede Seite hat ein Index k ($k \in \mathbb{N}, 1 \leq k \leq n$).

x_k gibt an, wie wichtig die Seite k im Web ist ($x_k \geq 0$ und $x_j > x_i$ bedeutet, dass Seite j wichtiger ist als Seite i).



Beispiel:

Einfachste Überlegung: x_k ... Anzahl der Backlinks der Seite k
 $x_1 = 2, x_2 = 1, \underbrace{x_3 = 3}_{\text{wichtigste!}}, x_4 = 2$

Aber:

Ein Backlink von einer vib euber bedeutenden Seite soll mehr zählen soll mehr zählen als der von einer unbedeutenden.

- a) Summation über Backlinks
z.B. $x_1 = x_3 + x_4$
- b) Wert einer Seite soll sich nicht erhöhen, dass viele Links von dieser Seite ausgehen ("Demokratie")

Page's importance score (8.1)

Def. 8.1

Sei $L_k \subseteq \{1, 2, \dots, n\}$ die Menge aller Backlinks der Seite k . Für jedes k definiere $x_k := \sum_{j \in L_k} \frac{x_j}{n_j}$, wobei n_j die Anzahl der von der Seite j ausgehenden Links angibt.

Beispiel:

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

Mit $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ erhalte $Ax = x$ mit $A = \begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix}$ ("Link-Matrix")

⇒ Problem der linearen Algebra: Finde Eigenvektoren zum Eigenwert 1 von A.

Eigenvektoren: $\left\{ t \cdot \begin{pmatrix} 12 \\ 4 \\ 9 \\ 6 \end{pmatrix} \mid t \in \mathbb{Z} \setminus \{0\} \right\}$

Skalierung: $x_1 = \frac{12}{31} \approx 0,387 \Rightarrow$ Seite 1 wichtigste!

$$\begin{aligned}x_2 &= \frac{4}{31} \approx 0,129 \\x_3 &= \frac{9}{31} \approx 0,290 \\x_4 &= \frac{6}{31} \approx 0,194 \\ \Rightarrow \sum &= 1\end{aligned}$$

Ergebnis: Seite 3 ist nicht die wichtigste (obwohl sie Backlinks von allen anderen hat). Seite 3 macht nur einen Link und zwar auf Seite 1 ("wirft ganzes Gewicht auf Seite 1").

Allgemeiner gilt: Die Link-Matrix A für ein beliebiges Web hat stets den Eigenwert 1, falls das zugrundeliegende Web keine Seiten enthält, von denen keine Links ausgehen ("dangling nodes").

Stochastische Matrizen (8.2)

Def. 8.2

Eine quadratische $(n \times n)$ -Matrix heißt (spalten-)stochastisch, falls alle Einträge $a_{ij} \geq 0 \forall i, j$ (nicht-negativ) und $\sum_{j=1}^n a_{ij} = 1 \forall i$ (Einträge jeder Spalte summieren sich zu Eins).

10.01.2007

Bemerkung:

Die Link-Matrix eines Webs ohne "dangling nodes" ist spalten-stochastisch.

Satz

Satz

Jede spalten-stochastische Matrix hat den Eigenwert 1.

Beweis:

Sei A eine spalten-stochastische $(n \times n)$ -Matrix und $v = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. Dann $A^t v = v$, somit ist 1 stets ein Eigenwert von A^t und daher auch von A . ■

Bezeichnung:

$V_1(A)$: Eigenraum zum Eigenwert 1 einer spalten-stochastischen Matrix A .

Bemerkungen:

- (i) Eindeutiges Ranking, wenn $\dim V_1(A) = 1$ (d.h. es existiert ein eindeutig bestimmter Eigenvektor x mit $\sum_{i=1}^n x_i = 1$). Dies gilt stets für *stark zusammenhängende* Web-Graphen (von jeder Seite aus kommt man auf jede andere durch endlich viele Klicks). Für Webs mit r nicht-zusammenhängenden Teilwebs W_1, \dots, W_r gilt $\dim V_1(A) \geq r$.
- (ii) Webs mit "dangling nodes": Link-Matrix enthält Spalten mit allen Einträgen Null. Modifizierte Verfahren: spalten-substochastische Matrizen ($\sum_{i=1}^n x_i \leq 1$) \rightarrow Eigenwert ≤ 1 , aber 1 muss nicht notwendig Eigenwert sein.

Nähere Betrachtung zu (i):

Fall $\dim V_1(A) > 1$ (liefert einen Spezialfall des Satzes von Perron und Frobenius (B. Huppert: Lineare Algebra)): Sei S eine $(n \times n)$ -Matrix mit Einträgen stets $\frac{1}{n}$, dann ist S spalten-stochastisch und $\dim V_1(S) = 1$ (leichte Überlegung). Ersetze A durch die Matrix $M = (1 - m) \cdot A + m \cdot S$ ("gewichteter Durchschnitt von A und S "), wobei m konstant und $0 \leq m \leq 1$. (Google: $m = 0,15$)

Es gilt: Für $0 \leq m \leq 1$ ist M spalten-stochastisch. Es bleibt zu zeigen: $V_1(M)$ ist eindimensional für alle $0 < m \leq 1$. \Rightarrow Verwende M zur Berechnung des "page's importance score".

[Für $m = 0$ erhalte das ursprüngliche Problem mit $M = A$. Falls $m = 1$, so $M = S$. ("alle Seiten sind gleichwertig": $x_i = \frac{1}{n} \forall i$.)]

Beispiel:

Für $m = 0,15$ erhält man

$$M = \begin{pmatrix} 0,0375 & 0,0375 & 0,8875 & 0,4625 \\ 0,32083 & 0,0375 & 0,0375 & 0,0375 \\ 0,32083 & 0,4625 & 0,0375 & 0,4625 \\ 0,32083 & 0,4625 & 0,0375 & 0,0375 \end{pmatrix}$$

und somit $x_1 \approx 0,368$, $x_2 \approx 0,142$, $x_3 \approx 0,288$, $x_4 \approx 0,202$.

\Rightarrow selbes Ranking, leicht veränderte Scores.

Satz (8.4)

Satz 8.4

Falls M spalten-stochastisch und $m_{ij} > 0 \forall i, j$ (d.h. M ist positiv) ist, dann ist jeder Eigenvektor in $V_1(M)$ stets lauter positive oder lauter negative Komponenten.

Beweis:

Angenommen $x \in V_1(M)$ enthalte Komponenten mit unterschiedlichen Vorzeichen. Aus $Mx = x$ folgt $x_i = \sum_{j=1}^n m_{ij} x_j$ und alle Summanden $\underbrace{m_{ij}}_{>0 \forall i,j} x_j$ haben unterschiedliche Vorzeichen. Somit

$$|x_i| = \left| \sum_{j=1}^n m_{ij} x_j \right| < \sum_{j=1}^n m_{ij} \cdot |x_j| \text{ und daher } \sum_{i=1}^n |x_i| < \sum_{i=1}^n \sum_{j=1}^n m_{ij} \cdot |x_j| = \sum_{j=1}^n \underbrace{\left(\sum_{i=1}^n m_{ij} \right)}_{=1 \forall j} \cdot |x_j| = \sum_{j=1}^n |x_j|. \quad \blacksquare$$

Satz (8.5)

Satz 8.5

Seien $v, w \in \mathbb{R}^n, n \geq 2$, linear unabhängig, dann hat $x = s \cdot v + t \cdot w$ ($s, t \in \mathbb{R}$ geeignet) Komponenten mit unterschiedlichen Vorzeichen.

Beweis:

Klar: $v \neq 0, w \neq 0$ (linear unabhängig). Sei $d := \sum_{i=1}^n v_i$. Für $d = 0$ muss v Komponenten mit unterschiedlichen Vorzeichen haben, $s = 1$ und $t = 0$ liefert Widerspruch. Für $d \neq 0$ setze $s := \frac{-\sum_i w_i}{d}$ und $t = 1$. Da v, w linear unabhängig, ist $0 \neq x = s \cdot v + w$. Wegen $d = \sum_{i=1}^n v_i$ folgt $\sum_{i=1}^n x_i = 0$. Somit hat x Komponenten mit unterschiedlichen Vorzeichen. \blacksquare

Sätze 8.4 und 8.5 liefern:

Satz (8.6)

Satz 8.6

Falls M spalten-stochastisch und positiv ist, dann gilt $\dim V_1(M) = 1$.

Beweis:

Angenommen $v, w \in V_1(M)$ seien linear unabhängig. Für beliebiges $s \in \mathbb{R}$ gilt $x = v + s \cdot w \in V_1(M)$ und nach 8.4 hat dann x nur Komponenten, die entweder alle positiv oder alle negativ sind. Andererseits liefert Satz 8.5, dass x Komponenten hat mit unterschiedlichen Vorzeichen. \blacksquare

Somit haben wir stets einen eindeutig bestimmten Eigenvektor $x \in V_1(M)$ mit $\sum_{i=1}^n x_i = 1$ (wähle alle Komponenten positiv).

16.01.2007

Wie berechnet man den "pages's score"-Eigenvektor?

Iteratives Verfahren (Potenzmethode):

1. Wähle einen geeigneten Startvektor x_0
2. Bilde $x_k = Mx_{k-1}$ (somit $x_k = M^k \cdot x_0$)

3. $k \rightarrow \infty$

Dann ist x_k approximativ einen Eigenvektor für den größten Eigenwert von M . (Falls x_k gegen ∞ wächst oder gegen 0 geht, dann reskalisieren bei jeder Iteration via $x_k = \frac{Mx_{k-1}}{\|Mx_{k-1}\|}$.) Um sicherzustellen, dass die Potenzmethode konvergiert in vernünftig vielen Schritten, fordert man häufig, dass der größte Eigenwert ($k = 1$) für M nur einfach vorkommt, d.h. $z_M(k) = (k-1) \cdot r(k)$ mit $r(k)$ ein Polynom vom Grad $n - 1$ und $k - 1 \nmid r(k)$.

Etwas technisch erhält man:

Lemma (8.7)

Lemma 8,7

Sei M eine positive spalten-stochastische $(n \times n)$ -Matrix und $V \subseteq \mathbb{R}^n$ bezeichne den Unterraum, der die Vektoren v enthält mit $\sum_{i=1}^n v_i = 0$, dann gilt $M \cdot v \in V$ und für jedes $v \in V$ gilt: $\|Mv\|_1 \leq c \cdot \|v\|_1$ mit $0 < c < 1$ geeignet ($c := \max_{1 \leq j \leq n} \left| 1 - 2 \cdot \min_{1 \leq i \leq n} m_{ij} \right|$). Damit erhalten wir:

Satz (8.8)

Satz 8.8

Jede positive spalten-stochastische Matrix M hat einen eindeutig bestimmten Vektor q mit positiven Komponenten, so dass $Mq = q$ und $\|q\|_1 (= \sum_{i=1}^n |q_i|) = 1$. Der Vektor q lässt sich iterativ berechnen via $q = \lim_{k \rightarrow \infty} M^k \cdot x_0$ für einen beliebigen Startwert x_0 mit positiven Komponenten, so dass $\|x_0\|_1 = 1$.

Beweis:

Es genügt die Iteration nachzuweisen. Sei $x_0 \in \mathbb{R}^n$ beliebig mit positiven Komponenten, so dass $\|x_0\|_1 = 1$, dann schreibe $x_0 = q + v$ mit $v \in V$ (und V wie in 8.7). Daher $M^k x_0 = M^k \cdot q + M^k \cdot v = q + M^k \cdot v$, somit $M^k x_0 - q = M^k \cdot v$. Induktiv erhält man mit 8.7: $\|M^k v\|_1 \leq c^k \cdot \|v\|_1$ mit $0 < c < 1$ wie in 8.7. Folglich $\lim_{k \rightarrow \infty} \|M^k \cdot v\|_1 = 0$ und daher: $q = \lim_{k \rightarrow \infty} M^k x_0$. ■

Bemerkung:

M positiv (also keine Nulleinträge). Multiplikation $M \cdot v$ benötigt $\mathcal{O}(u^2)$ Additionen und Multiplikationen ($\mathcal{O}(u^3)$ bei Matrizenprodukt). Web $n \approx 8 \cdot 10^3$ (Stand: Januar 2005) geht zwar, aber besser: Verwende $M = (1 - m)Ax + m \cdot Sx$ (A mit sehr vielen Nulleinträgen) \Rightarrow wesentlich effizienter (die weiteren Webseiten haben in der Tat nur wenige Links mit anderen Seiten).

3.2 Codierungstheorie

Wörterbuch der Codierung (8.9)

Def. 8.9

Ein Code ist eine Menge C (von Zeichen, die Informationen speichern und übermitteln sollen). Eine Codierung (Verschlüsselung, Chiffrierung) ist eine Abbildung $f : C_1 \rightarrow C_2$ eines Codes C_1 in einen anderen Code C_2 . Eine Decodierung (Entschlüsselung, Dechiffrierung) einer Codierung $f : C_1 \rightarrow C_2$ ist eine Abbildung $g : C_2 \rightarrow C_1$ mit $g \circ f = \text{id}$. Die Quelle einer Codierung $f : C_1 \rightarrow C_2$ heißt Klartext, ein Element der Quelle heißt Nachrichtenwort. Ein Element des Bildes einer Codierung heißt Codewort. Eine Codierung heißt lineare Codierung, wenn C_1, C_2 Vektorräume sind und $f : C_1 \rightarrow C_2$ linear. Ferner sollte sinnvollerweise f injektiv sein.

Beispiele:

- Morsealphabet: \cdot und $-$
- Zahlensystem: $0, 1, \dots, 9$ und Zeichen \cdot und $-$
- Barcodes zur Bezeichnung von Waren im Supermarkt
- ISBN-Code

Hamming-Metrik, Hamming-Abstand (8.10)

Def. 8.10

Sei K^n ein linearer Code. Die *Hamming-Metrik* auf K^n ist die Abbildung $d : K^n \times K^n \rightarrow \mathbb{N}_0$ mit $d(x, y) := \#$ der $i \in \{1, \dots, n\}$ mit $\xi_i \neq \eta_i$. ("Anzahl der Koeffizienten von x , die in y anders (d.h. falsch) übertragen werden".) Dabei heißt $d(x, y)$ der *Hamming-Abstand* von x und y . Die *Hamming-Gewichtsfunktion* ist die Abbildung $\|\cdot\| : K^n \rightarrow \mathbb{N}_0$ mit $\|x\| := d(x, 0)$. ("Anzahl der von Null verschiedenen Komponenten von x ".)

Lemma (8.11)

Lemma 8.11

Die Hamming-Metrix ist eine Metrix auf K^n .

Beweis folgt direkt aus 7.1.

Bemerkung:

d ist durch das Hamming-Gewicht vollständig bestimmt via $d(x, y) = \|x - y\|$.

Lemma (8.12)

Lemma 8.12

Für das Hamming-Gewicht $\|\cdot\| : K^n \rightarrow \mathbb{N}_0$ gilt:

- (i) $\|x\| = 0 \Leftrightarrow x = 0$
- (ii) $\|c \cdot x\| = c \cdot \|x\| \forall c \neq 0$
- (iii) $\|x + y\| \leq \|x\| + \|y\| \forall x, y$

Beweis folgt unmittelbar aus 7.2 (oder mittels $\|x\| := d(x, 0)$).

Definition (8.13)

Def. 8.13

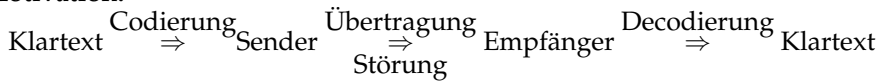
Ein Monomorphismus $f : K^k \rightarrow K^n$ heißt eine (k, n) -*lineare Codierung*. Die Darstellungsmatrix von f heißt *erzeugende Matrix*. Eine lineare Abbildung $h : K^n \rightarrow K^{n-k}$ heißt *Kontrollabbildung*, wenn $\ker(h) = \text{Im}(f)$. Die Darstellungsmatrix von h heißt *Kontrollmatrix*. Die *Hamming-Norm* oder der *Hamming-Abstand* einer (k, n) -linearen Codierung f ist definiert als

$$\|f\| := \min \{ \|f(x)\| \mid x \in K^n, x \neq 0 \}$$

Bemerkung:

- $x \in \text{Im}(f) \Leftrightarrow h(x) = 0$
- die Kontrollabbildung h existiert immer
- ferner gilt: $\|f\| = \min \{d(f(x), f(y)) \mid x, y \in K^k, x \neq y\}$
 $[d(f(x), f(y)) = d(f(x) - f(y), 0) = d(f(x - y), 0) = \|f(x - y)\|]$

Motivation:



Finde Methoden, die Fehler bei der Übertragung erkennen und möglichst auch korrigieren. Mit anderen Worten: Wenn x ein codiertes, ausgesuchtes Wort und y das empfangene Wort ist, dann soll festgestellt werden, ob es tatsächlich durch die Codierung entstanden ist oder verändert wurde und ob man daraus x rekonstruieren kann.

Wenn bei einer Codierung $f : K^k \rightarrow K^n$ Fehler an höchstens r Stellen immer erkannt werden, dann ist die Codierung r -fehlererkennend. Wenn die Fehler an höchstens s Stellen durch die restliche Information im übertragenden Wort korrigiert werden können, so heißt die Codierung s -fehlerkorrigierend.

Satz (8.14)

Satz 8.14

Sei $f : K^k \rightarrow K^n$ eine (k, n) -lineare Codierung und sei $C := \text{Im}(f)$ und $y \in K^n$, dann gilt:

- Fehlererkennung: Wenn $x \in C$ existiert mit $x \neq y$, so dass $d(x, y) < \|f\|$, dann ist $y \notin C$ (d.h. y ist kein Codewort, ist also falsch).
- Feherkorrektur: Wenn ein $x \in C$ existiert mit $d(x, y) < \frac{1}{2} \|f\|$, dann gilt für $z \in C$ mit $z \neq x$: $d(x, y) < d(z, y)$ (d.h. x ist das einzige Element von C mit dem gegebenen Abstand $d(x, y)$ und daher durch y eindeutig bestimmt).

Beweis:

- Angenommen $y \in C$, so wäre $d(x, y) \geq \|f\|$ oder $x = y$ nach Definition von $\|f\|$.
- Für $z \in C$ beliebig mit $x \neq z$ gilt nach Voraussetzung $d(x, y) < \frac{1}{2} \|f\| \leq d(x, z) \leq d(x, y) + d(y, z)$. Also $d(x, y) < d(y, z)$. ■

Bemerkung:

Wenn bei der Übertragung von $f(a)$ weniger als $\frac{1}{2} \|f\|$ Fehler auftreten und $y \in K^n$ empfangen, dann ist $x \in C$ mit $d(x, y) < \frac{1}{2} \|f\|$ das übertragende Element. Somit ist eine (k, n) -lineare Codierung stets $(\|f\| - 1)$ -fehlererkennend und $[\frac{1}{2}(\|f\| - 1)]$ -fehlerkorrigierend³. \Rightarrow Finde Codierungen mit $\|f\|$ möglichst groß.

24.01.2007, Wiederholung (Oliver Burger)

³ $[x] :=$ größte ganze Zahl $n < x$

Code:

Menge C von Zeichen zur Übermittlung von Nachrichten. Die Zeichen aus C stammen aus einem *Alphabet*.

Codierung:

Abbildung $f : C_1 \rightarrow C_2$

Decodierung:

Abbildung $g : C_2 \rightarrow C_1$ mit $g \circ f = \text{id}$

f sollte injektiv sein, da sonst keine eindeutige Decodierung möglich ist.

Urbild von f heißt *Klartext*. Elemente des Klartexts heißen *Nachrichtenwort*. Elemente des Bildes heißen *Codewort*.

Beispiele für Alphabete:

- lateinisches Alphabet: $\{a, b, c, \dots, z, A, B, C, \dots, Z\}$
- Morse-Alphabet: $\{-, \cdot\}$
- $\{0, 1\} = GF_2 = \mathbb{F}_2$
- $\{0, 1, \dots, 9\}$

Lineare Codes:

Eine Codierung $f : R \rightarrow C$ heißt *linear*, wenn K, C Vektorräume sind und f *linear* ist. Die zugehörigen Codes heißen *lineare Codes*. Das Alphabet ist hier ein (endlicher) Körper.

Hamming-Metrik:

Sei K^n ein linearer Code, $d : K^n \times K^n \rightarrow \mathbb{N}_0$ mit $d(x, y) = \#\{i \in \{1, \dots, n\} \mid \xi_i = \eta_i\}$.

Hamming-Gewicht:

$\|x\| : K^n \rightarrow \mathbb{N}, \|x\| = d(x, 0)$

Es gilt $d(x, y) = \|x - y\|$.

- (i) $\|x\| = 0 \Leftrightarrow x = 0$
- (ii) $\|c \cdot x\| = c \cdot \|x\| \forall c \in K$
- (iii) $\|x + y\| \leq \|x\| + \|y\|$

Eine (k, n) -lineare Codierung ist ein Monomorphismus $f : K^k \rightarrow K^n$. Die Darstellungsmatrix von f heißt *erzeugende Matrix*. Eine lineare Abbildung $h : K^n \rightarrow K^{n-k}$ heißt *Kontrollabbildung* genau dann, wenn $\ker(h) = \text{Im}(f)$. Die Darstellungsmatrix heißt *Kontrollmatrix*.

$\|f\| := \min \{\|f(x)\| : x \in K^k\}$

$d(f) := \min \{d(x, y) : x, y \in K^n\}$

Es gilt $d(f) = \|f\|$.

Eine Codierung, die Fehler an höchstens r Stellen erkennt, heißt *r-fehlererkennend*. Eine Codierung, die Fehler an höchstens s Stellen korrigiert, heißt *s-fehlerkorrigierend*.

Sei $f : K^k \rightarrow K^n, C = \text{Im}(f), y \in K^n$

- a) Sei $x \in C$ mit $x \neq y, d(x, y) \leq \|f\|$
 $\Rightarrow y \notin C$
- b) Sei $x \in C$ mit $d(x, y) < \frac{1}{2} \|f\|$
 $\Rightarrow z \in C$ mit $x \neq z$ ist $d(x, y) < d(z, y)$
 $\Rightarrow x$ durch y endlich bestimmt, man kann y korrigieren

Beispiel:

Der einfachste lineare Code ist der Wiederholungscode $f : K \rightarrow K^3, x \mapsto xxx$.

$$f(0) = 000$$

$$f(1) = 111$$

$$F = (111)$$

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

$$\|f\| = 3$$

30.01.2007 (Oliver Burger)

Lineare Codes:

Als Alphabet ist ein (meist endlicher) Körper notwendig, hier meist $GF(2) = \mathbb{F}_2$.

Beispiel:

$$\text{Sei } f : GF(2)^4 \rightarrow GF(2)^8, (x_1 x_2 x_3 x_4) \mapsto (x_1, x_2, x_3, x_4, x_1 + x_2, x_3 + x_4, x_2 + x_3, x_2 + x_4)$$

Zuerst berechnen wir F mit Hilfe der kanonischen Basis von $GF(2)^4$

$$f(1000) = 10001010$$

$$f(0100) = 01001001$$

$$f(0010) = 00100110$$

$$f(0001) = 00010101$$

$$F = \begin{pmatrix} 10001010 \\ 01001001 \\ 00100110 \\ 00010101 \end{pmatrix}$$

Es ist bekannt, dass die Kontrollmatrix $Hx^t = 0 \Leftrightarrow x \in C$.

Für jedes Element $y \in K = GF(2)^4$ gilt also $H(yF)^t = 0 \Rightarrow F \cdot H^t = 0$. H besteht aus den Lösungen der Gleichung $F \cdot z = 0$.

$$\begin{pmatrix} 10001010 \\ 01001001 \\ 00100110 \\ 00010101 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_j \end{pmatrix} = 0$$

- | | |
|---|---------------------------------|
| $\left. \begin{array}{l} 1) \quad x_5 = 1, x_6 = x_7 = x_8 = 0 \\ 2) \quad x_6 = 1, x_5 = x_7 = x_8 = 0 \\ 3) \quad x_7 = 1, x_5 = x_6 = x_8 = 0 \\ 4) \quad x_8 = 1, x_5 = x_6 = x_7 = 0 \end{array} \right\}$ | Definition der freien Variablen |
|---|---------------------------------|

Daraus erhalten wir:

$$1) \quad x_4 = 0, x_3 = 0, x_2 = 1, x_1 = 1$$

$$2) \quad x_4 = 1, x_3 = 1, x_2 = 0, x_1 = 0$$

$$3) \quad x_4 = 0, x_3 = 1, x_2 = 0, x_1 = 1$$

$$4) \quad x_4 = 1, x_3 = 0, x_2 = 1, x_1 = 0$$

$$H = \begin{pmatrix} 11001000 \\ 00110100 \\ 10100010 \\ 01010001 \end{pmatrix}$$

Da die Fehlerkorrekturfähigkeit eines Codes dem Hamming-Gewicht folgt, suchen wir eine Abschätzung für dieses.

Satz (8.15)

Satz 8.15

Sei $f : K^k \rightarrow K^n$ eine (k, n) -lineare Kodierung, sei $C = \text{Im}(f) \neq 0, K^n$. Sei H die Kontrollmatrix von f , dann gilt $d(f) = \|f\| = \min \{r \in \mathbb{N} : \}$

Dann gilt

$$\begin{aligned} d(f) &= \|f\| = \min \{r \in \mathbb{N} | H \text{ besitzt } r \text{ linear abhängige Spalten}\} \\ &= \max \{r \in \mathbb{N} | \text{je } r - 1 \text{ Spalten aus } H \text{ sind linear unabhängig}\} \end{aligned}$$

Beweis:

Seien s_1, \dots, s_n die Spalten von H , da $C \neq 0$ sind diese linear abhängig in K^{n-k} . Sei w minimal, so dass es w linear abhängige Spalten gibt, seien diese $s_{i_1}, \dots, s_{i_w} \Rightarrow \exists c_j \in K$ mit $\sum_{j=1}^w c_j s_j = 0$ mit

$c_j \neq 0 \Leftrightarrow j \in \{i_1, \dots, i_w\}$. Sei $c = (c_1, \dots, c_n) \Rightarrow H^t c^t = 0 \Rightarrow c \in C$ mit $\|c\| = w \Rightarrow \|f\| \leq w$.

Annahme: $\exists 0 \neq \bar{c} \in C$ mit $\|\bar{c}\| < w$

$\Rightarrow H \bar{c}^t = 0 \Rightarrow \exists \bar{w}$ linear abhängige Spalten von H . Widerspruch zu w minimal

$\Rightarrow \|f\| = w$ ■

Singleton-Schranke (8.16)

Kor. 8.16

Sei f eine (k, n) -lineare Codierung und $d(f) = d$, so gilt $d \leq n - k + 1$.

Beweis:

$\text{rg}(H) \leq n - k \Rightarrow$ es können maximal $n - k$ linear unabhängige Spalten vorhanden sein.

Durch elementare Umformung folgt $n \geq d + k - 1$.

Beispiel:

Sei f eine $(3, 7)$ -lineare Codierung über $GF(2)$ mit $F = \begin{pmatrix} 0011011 \\ 1101101 \\ 0101001 \end{pmatrix}$. Um den Minimalabstand der Codierung zu erhalten, berechnen wir H . Löse $Fz = 0, z = (z_1, \dots, z_7)^t$. Durch Zeilenvertauschungen erhalten wir: $\begin{pmatrix} 1101101 \\ 0101001 \\ 0011011 \end{pmatrix}$

Wir setzen:

1. $x_4 = 1, x_5 = x_6 = x_7 = 0$
2. $x_5 = 1, x_4 = x_6 = x_7 = 0$
3. $x_6 = 1, x_4 = x_5 = x_7 = 0$
4. $x_7 = 1, x_4 = x_5 = x_6 = 0$

und erhalten:

1. $x_3 = 1, x_2 = 1, x_1 = 0$
2. $x_3 = 0, x_2 = 0, x_1 = 1$

3. $x_3 = 1, x_2 = 0, x_1 = 0$

4. $x_3 = 1, x_2 = 1, x_1 = 0$

Also $H = \begin{pmatrix} 0111000 \\ 1000100 \\ 0110001 \end{pmatrix}$ mit 1. und 5. Spalte linear abhängig.

$\Rightarrow d(f) = 2$

Singleton: $d \leq n - k + 1 \leq 7 - 3 + 1 = 5$

Index

- Ableitung, partielle, 30
- Ableitung, totale, 31
- Affiner Unterraum, 4

- Banachraum, 27
- Bilinearform, 17
- Bolzano-Weierstrass, 27

- Cauchy-Schwarz, 18
- Cayley-Hamilton, 14
- Charakteristisches Polynom, 13, 15

- Determinante, 9, 10
- Determinantenfunktion, 9
- Determinantenfunktion, 8, 10
- Diagonalisierbar, 12
- Diagonalisierbarkeit, 16

- Eigenraum, 12
- Eigenvektor, 12
- Eigenwert, 12
- Eindeutigkeit, 9
- Elementare Zeilenumformungen, 5
- Erweiterte Matrix, 6
- Euklidischer Vektorraum, 18
- Existenz, 9
- Extremwertberechnung, 33

- Gauss-Algorithmus, 6
- Gram-Schmidt, 19
- Grenzwert, 25
- Gruppenhomomorphismus, 11

- Häufungswert, 25
- Hamilton, 14
- Hesse-Matrix, 32
- homogen, 3
- Homogenität, 3

- inhomogen, 3
- inverse Matrix, 8
- Invertierbarkeit, 10

- Jacobi-Matrix, 31
- Jordan'sche Normalform, 16
- Jordan-Kästchen, 16

- Kettenregel, 32
- Kurve, 28

- Lösbarkeit, praktische, 7
- Lösbarkeit, strukturelle, 4

- Laplace, 11
- Lineare Gleichungssysteme, 3

- Menge, abgeschlossene, 24
- Menge, offene, 23
- Minimalpolynom, 14, 15
- multiplikativ, 10

- Nullstellen, 13

- Orthonormalbasis, 19
- Orthonormalisierungsverfahren, 19

- Pivotelement, 6

- Rücksstitution, 7
- Raum, metrischer, 20, 24
- Raum, normierter, 21
- Rechenregeln, 9

- Schwarz, Satz von, 31
- Stetigkeit, 25
- Stetigkeit von Folgen, 25
- Stufenform, 6

- Trigonalisierbarkeit, 14

- Vektorraum, 18
- Vollständigkeit, 27