

Angewandte Mathematik: Numerik

Prof. Dr. Reinhard Klein

Übungen: Dipl.-Inform. Alexander Schier

Wintersemester 2022/23

Dieses Skript basiert auf und folgt in weiten Teilen den folgenden Lehrbüchern:

- Numerical Linear Algebra

Lloyd N. Trefethen, David Bau, III

ISBN: 978-0-898713-61-9 [?]

- Analysis 2

2. erweiterte Auflage

Konrad Königsberger

ISBN: 978-3540203896 [?]

Zum Lernen und zur weiteren Vertiefung empfehlen wir, diese Bücher zusätzlich zum Skript zu verwenden.

Hinweis:

Fehler und Anmerkungen für Ergänzungen zum Skript können an schier@uni-bonn.de geschickt werden.

Kapitel 1: Grundlagen

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Matrix-Vektor-Multiplikation

Seien x ein n -dimensionaler Spaltenvektor und A eine $(m \times n)$ -Matrix (m Zeilen, n Spalten) jeweils über \mathbb{C} . $b = Ax$ ist der m -dimensionale Spaltenvektor definiert durch

$$b_i = \sum_{j=1}^n a_{ij}x_j, \quad i = 1, \dots, m \tag{1.1}$$

wobei b_i der i -te Eintrag von b und (a_{ij}) der Eintrag an ij -ter Stelle (i -te Zeile, j -te Spalte) in A ist.

Matrix-Vektor-Multiplikation

Interpretation:

Sei a_j die j -te Spalte von A (d.h. a_j ist ein Vektor mit m Einträgen). Dann können wir das Vektor-Matrixprodukt in Gleichung (1.1) schreiben als

$$b = Ax = \sum_{j=1}^n x_j a_j, \quad (1.2)$$

d.h. b ist eine *Linearkombination der Spaltenvektoren* von A . Schematisch:

$$\begin{pmatrix} b \\ \end{pmatrix} = \left(\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_n \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \underbrace{x_1 \begin{pmatrix} a_1 \\ \end{pmatrix} + x_2 \begin{pmatrix} a_2 \\ \end{pmatrix} + \dots + x_n \begin{pmatrix} a_n \\ \end{pmatrix}}_{\text{Linearkombination}}$$

Matrix-Vektor-Multiplikation

Beispiel (Matrix-Vektor-Multiplikation)

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} = 0.5 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0.5 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Bemerkung 1.1 (Bilder der kanonischen Basisvektoren)

Die Spaltenvektoren $a_i, i = 1, \dots, n$, von A geben an, wohin die kanonischen Einheitsvektoren e_i , d.h. die Vektoren mit 1 an der i -ten Stelle und 0 sonst, abgebildet werden.

Matrix-Matrix-Multiplikation

Seien A eine $(l \times m)$ -Matrix und C eine $(m \times n)$ -Matrix definiert durch

$$b_{ij} = \sum_{k=1, \dots, m} a_{ik} c_{kj}, \quad i = 1, \dots, l, j = 1, \dots, n. \quad (1.3)$$

Schematisch:

$$\left(\begin{array}{c|c|c|c} b_1 & b_2 & \dots & b_n \end{array} \right) = \left(\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_m \end{array} \right) \left(\begin{array}{c|c|c|c} c_1 & c_2 & \dots & c_n \end{array} \right)$$

d.h. $b_j = Ac_j = \sum_{k=1, \dots, m} c_{kj} a_k$, also ist b_j eine *Linearkombination* der Spalten a_k mit den Koeffizienten c_{kj} .

Äußeres Produkt

Das Produkt eines Spaltenvektors u ($l \times 1$ -Matrix) mit einem Zeilenvektor v ($1 \times n$ -Matrix) heißt *Äußeres Produkt*:

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_l \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} = \left(\begin{array}{c|c|c|c} & v_1 u & v_2 u & \cdots & v_n u \end{array} \right) = \begin{pmatrix} u_1 v \\ u_2 v \\ \vdots \\ u_l v \end{pmatrix}$$

$$= \begin{pmatrix} u_1 v_1 & \cdots & u_1 v_n \\ \vdots & & \vdots \\ u_l v_1 & \cdots & u_l v_n \end{pmatrix}$$

Alle Spalten sind Vielfache des Vektors u und alle Zeilen sind Vielfache von v .

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Bild- und Nullraum

Definition 1.1 (Bild einer Matrix)

Das Bild einer Matrix $A \in \mathbb{C}^{m \times n}$, geschrieben als $\text{Bild}(A)$, ist die Menge der Vektoren $v \in \mathbb{C}^m$, für die es ein $x \in \mathbb{C}^n$ gibt, so dass gilt $v = Ax$, d.h.

$$\text{Bild}(A) = \{Ax \mid x \in \mathbb{C}^n\}$$

Definition 1.2 (Nullraum (Kern))

Der Nullraum (auch als Kern bezeichnet) einer Matrix $A \in \mathbb{C}^{m \times n}$, geschrieben als $\text{Null}(A)$, ist die Menge der Vektoren $x \in \mathbb{C}^n$, für die gilt $Ax = 0$, wobei 0 der 0 -Vektor im \mathbb{C}^m ist, d.h.

$$\text{Null}(A) = \{x \in \mathbb{C}^n \mid Ax = 0\}$$

Bild- und Nullraum

Theorem 1.1 (Bildraum)

Bild(A) ist der lineare Vektorraum, der von den Spalten von A aufgespannt wird.

Beweis:

Sei $y = Ax$. Dann ist $y = \sum_{1 \leq j \leq n} x_j a_j$ Linearkombination von

Spaltenvektoren, d.h. $y \in L\{a_1, \dots, a_n\}$. Sei umgekehrt y in der linearen Hülle der Spalten von A , so gilt $y = \sum_{1 \leq j \leq n} x_j a_j$ für Koeffizienten

$x_j, j = 1, \dots, n$. Wählt man $x = (x_1, \dots, x_n)^T$, so ist $y = Ax$.

□

Rang und Inverse einer Matrix

Definition 1.3 (Rang einer Matrix)

Der Spaltenrang einer Matrix ist die Dimension des Spaltenraums. Der Zeilenrang einer Matrix ist die Dimension des Zeilenraums. Eine $(m \times n)$ -Matrix ist von *vollem Rang*, falls sie maximal möglichen Rang besitzt, d.h. eine $(m \times n)$ -Matrix mit $m \geq n$ mit vollem Rang hat n linear unabhängige Spalten.

Definition 1.4 (invertierbar)

Eine invertierbare (nicht singuläre) Matrix A ist eine Matrix von vollem Rang, d.h. die Spalten einer invertierbaren $(m \times m)$ -Matrix bilden eine Basis des \mathbb{C}^m .

Inverse einer Matrix

Für alle Einheitsvektoren e_j gibt es einen Vektor z_j , sodass:

$$e_j = Az_j = \sum_{i=1}^m z_{ij} a_i, \quad j = 1, \dots, m, \quad A \in \mathbb{C}^{m \times m} \quad (1.4)$$

Sei Z die Matrix mit den Spaltenvektoren z_j , dann gilt:

$$\left(\begin{array}{c|c|c|c} e_1 & e_2 & \dots & e_n \end{array} \right) = \left(\begin{array}{cccc} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{array} \right) = I = AZ$$

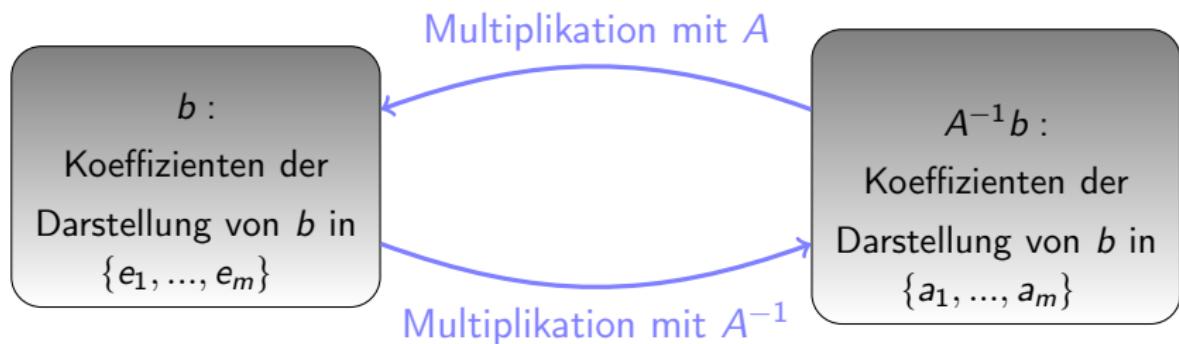
I ist hierbei die Einheitsmatrix und Z ist die *Inversen* von A .

Es gilt $I = AA^{-1} = A^{-1}A$.

Basiswechsel

Produkt von Inverser Matrix und Vektor:

Sei $x = A^{-1}b$. Dann gilt $Ax = AA^{-1}b = Ib = b$. x ist also der eindeutige Vektor, der die Gleichung $Ax = b$ erfüllt. D.h. die x_i sind die Koeffizienten der Darstellung von b als Linearkombination der Spaltenvektoren von A . Die Multiplikation mit A^{-1} ist ein Basiswechsel:



Adjungierte Matrix

Definition 1.5 (Adjungierte Matrix)

Die hermitesch adjungierte Matrix A^* (oder nur Adjungierte) einer Matrix $A \in \mathbb{C}^{m \times n}$ ist die $(n \times m)$ -Matrix, deren (i, j) -ter Eintrag das komplexe Konjugierte des (j, i) -ten Eintrags ist, d.h. $A^* = \overline{A}^T$.

Es gilt:

- Falls $A = A^*$ gilt, so heißt A *hermitesch*.
- Für reelle A ist $A^* = A^T$.
- Falls $A = A^T$ gilt, so heißt A *symmetrisch*.

Skalarprodukt

Definition 1.6 (Skalarprodukt)

Für zwei Spaltenvektoren $x, y \in \mathbb{C}^m$ definieren wir das Skalarprodukt als

$$x^*y = (\bar{x}_1 \quad \bar{x}_2 \quad \cdots \quad \bar{x}_m) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \sum_{i=1}^m \bar{x}_i y_i$$

Für $x, y \in \mathbb{R}^m$ entfällt die komplexe Konjugation und man erhält
 $x^*y = \sum_{i=1}^m x_i y_i$.

Bemerkung 1.2 (Alternative Schreibweisen für das Skalarprodukt)

Statt x^*y schreiben wir auch $\langle x, y \rangle$ oder $\langle x | y \rangle$.

Orthogonalität

Definition 1.7 (Orthogonale Vektoren)

Ein Paar von Vektoren x, y heißt orthogonal falls $x^*y = 0$. D.h. falls $x, y \in \mathbb{R}^m$, so bedeutet das, dass $x \perp y$ ist.

Eine Menge von Null verschiedener Vektoren S ist *orthogonal*, falls ihre Elemente paarweise orthogonal sind.

Eine Menge von Vektoren S heißt *orthonormal*, falls sie orthogonal ist und $\forall x \in S$ gilt $\|x\| = 1$.

Beispiel (Orthogonale Vektoren)

Die Vektoren $a = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ und

$b = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$ stehen orthogonal

zueinander, denn:

$$\begin{aligned} a^* b &= a_1 * b_1 + a_2 * b_2 \\ &= 3 * (-1) + 2 * 1.5 = 0 \end{aligned}$$

$a' := \frac{a}{\|a\|}$ und $b' := \frac{b}{\|b\|}$ sind
orthonormale Vektoren:

$$\|a\| = \sqrt{a_1^2 + a_2^2} = \sqrt{3^2 + 2^2} = \sqrt{13}$$

$$\|a'\| = \left\| \frac{a}{\|a\|} \right\| = \left\| \frac{1}{\sqrt{13}} \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\| = \sqrt{\left(\frac{3}{\sqrt{13}} \right)^2 + \left(\frac{2}{\sqrt{13}} \right)^2} = \sqrt{\frac{\sqrt{13}}{\sqrt{13}}} = 1$$

$\|b'\|$ analog

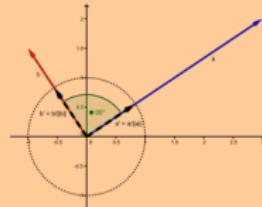


Abbildung 1.1: Orthogonale und orthonormale Vektoren.

Orthogonalität

Theorem 1.2 (Vektoren in einer orthogonalen Menge)

Die Vektoren in einer orthogonalen Menge $S := \{v_1, \dots, v_n\} \subset \mathbb{C}^m \setminus \{0\}$ sind linear unabhängig.

Beweis: (durch Widerspruch)

Angenommen, die Vektoren einer orthogonalen Menge sind nicht linear unabhängig. Dann kann der Vektor $v_k \in S$ als Linearkombination der übrigen Vektoren $v \in S \setminus \{v_k\}$ geschrieben werden, d.h. $v_k = \sum_{i=1, i \neq k}^n \lambda_i v_i$ ($\lambda_1, \dots, \lambda_n \in \mathbb{K}$). Da $v_k \neq 0$ folgt $v_k^* v_k = \|v_k\|^2 > 0$:

$$0 < v_k^* v_k = \sum_{i=1, i \neq k}^n \lambda_i v_k^* v_i = 0.$$

Wegen $v_k^* v_i = 0$ folgt der Widerspruch.

Komponenten eines Vektors

Das Skalarprodukt kann verwendet werden, um einen beliebigen Vektor in orthogonale Komponenten zu zerlegen.

Sei $\{q_1, \dots, q_n\}$ eine orthonormale Menge von Vektoren und v ein beliebiger Vektor. Dann ist der Vektor

$$r = v - (q_1^* v)q_1 - (q_2^* v)q_2 - \dots - (q_n^* v)q_n$$

orthogonal zu $\{q_1, \dots, q_n\}$:

$$q_i^* r = q_i^* v - \sum_{j=1}^n (q_j^* v) \underbrace{(q_i^* q_j)}_{=0, \text{ wenn } i \neq j} = q_i^* v - (q_i^* v) \underbrace{(q_i^* q_i)}_{=1} = 0$$

Es ist also möglich, v als Linearkombination von $n + 1$ orthogonalen Vektoren $\{q_1, \dots, q_n, r\}$ darzustellen:

$$v = r + \sum_{j=1}^n (q_j^* v)q_j.$$

Unitäre Matrizen

Definition 1.8 (Unitäre Matrizen)

Eine quadratische Matrix Q ist unitär (im reellen Fall orthogonal), falls $Q^{-1} = Q^*$ ist, und daher gilt: $Q^*Q = I$.

$$\begin{pmatrix} q_1^* \\ \hline q_2^* \\ \vdots \\ \hline q_m^* \end{pmatrix} \begin{pmatrix} q_1 & | & q_2 & | & \dots & | & q_m \end{pmatrix} = \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \end{pmatrix}$$

Unitäre Matrizen

Es gilt also $q_i^* q_j = \delta_{ij}$. Die Spalten einer unitären Matrix bilden somit eine orthonormale Basis des \mathbb{C}^m .

Satz 1.3 (Invarianz des Skalarprodukts)

Die Multiplikation mit einer unitären Matrix erhält das Skalarprodukt.

Beweis:

$$(Qx)^*(Qy) = x^* Q^* Qy = x^* Iy = x^* y$$



Unitäre Matrizen erhalten daher Winkel und Längen. Im reellen Fall entspricht die Multiplikation mit einer unitären Matrix einer Rotation ($\det Q = 1$) oder einer Spiegelung ($\det Q = -1$). In diesem Fall spricht man auch von *orthogonalen Matrizen*.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Norm

Definition 1.9 (Norm)

Eine Norm ist eine Abbildung $\|\cdot\| : \mathbb{C}^n \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

- 1 $\|x\| \geq 0$, $\|x\| = 0$ gdw. $x = 0$
- 2 $\|\alpha x\| = |\alpha| \cdot \|x\|$
- 3 $\|x + y\| \leq \|x\| + \|y\|$

Norm

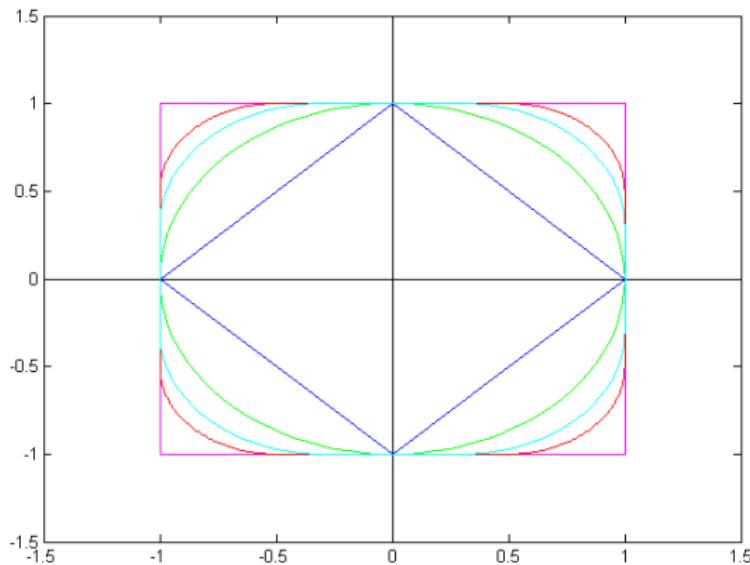


Abbildung 1.2: Beispiel für p -Normen: von innen (1-Norm) nach außen (∞ -Norm).

p -Normen

Definition 1.10 (p -Norm (Vektoren))

Für eine reelle Zahl $p \geq 1$ definiert man für einen Vektor $x \in \mathbb{C}^n$ die p -Norm durch

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

Beispiel (Euklidsche Norm)

Die p -Norm für $p = 2$ wird *euklidische Norm* genannt. Betrachte den Vektor $x = (1, 2, 3)$:

$$\|x\|_2 = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$

Matrixnorm

Eine $(m \times n)$ -Matrix kann als Vektor im $\mathbb{C}^{m \cdot n}$ aufgefasst werden. Damit kann jede Vektornorm auf dem $\mathbb{C}^{m \cdot n}$ auch als Matrixnorm verwendet werden. Es gibt jedoch andere Matrixnormen, die in vielen Fällen nützlicher sind, die *induzierten Matrixnormen*:

Definition 1.11 (Induzierte Matrixnorm)

Seien $\|\cdot\|_n$ und $\|\cdot\|_m$ Vektornormen des Definitions- bzw. Bildbereichs von $A \in \mathbb{C}^{m \times n}$. Dann ist $\|A\|_{(m,n)}$ definiert als die kleinste Zahl C , für die $\forall x \in \mathbb{C}^n$ gilt

$$\|Ax\|_m \leq C \cdot \|x\|_n,$$

d.h.

$$\|A\|_{(m,n)} = \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|_m}{\|x\|_n}.$$

Eigenschaften induzierter Matrixnormen

Satz 1.4 (Eigenschaften induzierter Matrixnormen)

Induzierte Matrixnormen sind Normen und es gilt für alle $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times l}$ und $x \in \mathbb{C}^n$:

- $\|A \cdot x\|_m \leq \|A\|_{(m,n)} \cdot \|x\|_n$
- $\|A \cdot B\|_{(m,l)} \leq \|A\|_{(m,n)} \cdot \|B\|_{(n,l)}$

Beweis: Siehe Skript.



2-Norm einer Diagonalmatrix

Lemma 1.5 (2-Norm einer Diagonalmatrix)

Sei D eine Diagonalmatrix mit $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$.

Dann gilt:

$$\|D\|_2 = \sup_{\|x\|=1} \|Dx\|_2 = \max_{1 \leq i \leq m} |d_i|.$$

Beweis: Die Matrix D skaliert die Einheitskugel entlang der Koordinatenachsen. Betrachte dazu die Multiplikation von D mit den kanonischen Einheitsvektoren e_i . Jeder Vektor e_i wird um den Faktor d_i gestreckt. Dabei entspricht die maximale Streckung dem betragsmäßig größtem Diagonaleintrag von D .



1-Norm einer Matrix

Definition 1.12 (1-Norm einer Matrix (maximale Spaltensumme))

Sei A eine $(m \times n)$ -Matrix. Dann entspricht $\|A\|_1$ der maximalen Spaltensumme:

$$\|A\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1.$$

1-Norm einer Matrix

Theorem 1.6 (1-Norm einer Matrix)

Die 1-Norm $\|A\|_1$ einer Matrix A mit $A \in \mathbb{C}^{m \times n}$ ist eine induzierte Matrixnorm.

Beweis: Um dies zu zeigen, betrachten wir die Einheitskugel $\{x \in \mathbb{C}^m \mid \sum_{j=1}^n |x_j| \leq 1\}$ unter der $\|\cdot\|_1$ Norm. Für jeden Vektor Ax im Bild dieser Menge gilt

$$\|Ax\|_1 = \left\| \sum_{j=1}^n x_j a_j \right\|_1 \stackrel{\text{Def. 1.9}}{\leq} \sum_{j=1}^n |x_j| \|a_j\|_1 \leq \max_{1 \leq j \leq n} \|a_j\|_1.$$

Es gilt also

$$\|Ax\|_1 \leq \max_{1 \leq j \leq n} \|a_j\|_1.$$

Wählt man $x = e_j$, wobei j so gewählt wird, dass $\|a_j\|_1$ maximal ist, d.h. $j = \arg \max \|a_j\|_1$, so gilt $\|A\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1$.

∞ -Norm einer Matrix

Definition 1.13 (**Unendlich-Norm einer Matrix (max. Zeilensumme)**)

Sei A eine $(m \times n)$ -Matrix. Dann entspricht $\|A\|_\infty$ der maximalen Zeilensumme:

$$\|A\|_\infty = \max_{1 \leq i \leq m} \|a_i^*\|_1$$

Theorem 1.7 (**Unendlich-Norm einer Matrix**)

Die ∞ -Norm $\|A\|_\infty$ einer Matrix A mit $A \in \mathbb{C}^{m \times n}$ ist eine induzierte Matrixnorm.

Hölder-Ungleichung

Lemma 1.8 (Hölder-Ungleichung)

Seien $p, q \in \mathbb{R}$ mit $1 < p, q < \infty$ so gewählt, dass $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für alle Vektoren $x, y \in \mathbb{C}^n$:

$$\left| \sum_{i=1}^n \overline{x_i} y_i \right| = |x^* y| \leq \|x\|_p \|y\|_q.$$

Beweis: Der klassische Beweis greift auf die Youngsche-Ungleichung zurück. □

Cauchy-Schwarz-Ungleichung

Spezialfall der Hölder-Ungleichung für $p = q = 2$:

Lemma 1.9 (Cauchy-Schwarz-Ungleichung)

Es gilt für alle Vektoren $x, y \in \mathbb{C}^n$:

$$|x^*y| \leq \|x\|_2 \|y\|_2.$$

Matrixnorm

Beispiel (1-Norm)

$$A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$

$$\begin{aligned}\|A\|_1 &= \sup_{\|x\|=1} \|Ax\|_1 \\ &= 2 + 2 \\ &= 4.\end{aligned}$$

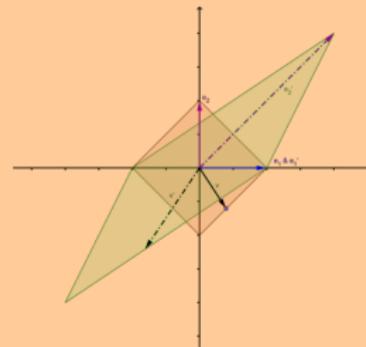


Abbildung 1.3: Anwendung der 1-Norm.

2-Norm eines Zeilenvektors

Enthalte A nur eine Zeile, d.h. $A = a^*$, wobei a ein Spaltenvektor ist:

$$\|Ax\|_2 = |a^*x| \leq \|a\|_2 \|x\|_2.$$

Für $x = a$ gilt:

$$\|Aa\|_2 = \|a\|_2^2 \Rightarrow \|A\|_2 = \sup_{x \neq 0} \left\{ \frac{\|Ax\|_2}{\|x\|_2} \right\} = \|a\|_2. \quad (1.5)$$

Begrenzung von $\|AB\|$ in induzierter Matrixnorm:

$$\begin{aligned} \|ABx\|_I &\leq \|A\|_{(I,m)} \|Bx\|_m \leq \|A\|_{(I,m)} \|B\|_{(m,n)} \|x\|_n \\ \Rightarrow \|AB\|_{(I,n)} &\leq \|A\|_{(I,m)} \|B\|_{(m,n)} \quad (\text{wird im Allgemeinen nicht angenommen}) \end{aligned}$$

Allgemeine Matrixnormen

Definition 1.14 (Allgemeine Matrixnormen)

Eine Matrixnorm kann analog zu einer Vektornorm definiert werden. Sie muss die folgenden Bedingungen erfüllen:

- 1 $\|A\| \geq 0$ und $\|A\| = 0 \Leftrightarrow A = 0$
- 2 $\|\alpha A\| = |\alpha| \|A\|$
- 3 $\|A + B\| \leq \|A\| + \|B\|$

Kapitel 2:

Lineare Ausgleichsrechnung

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

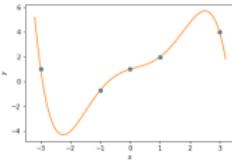
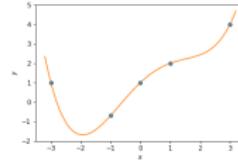
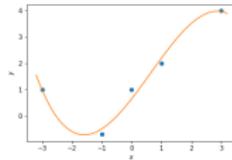
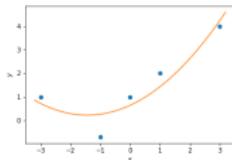
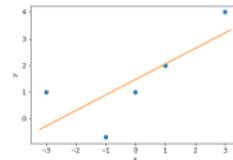
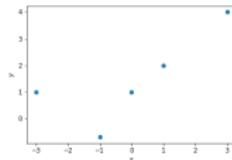
20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

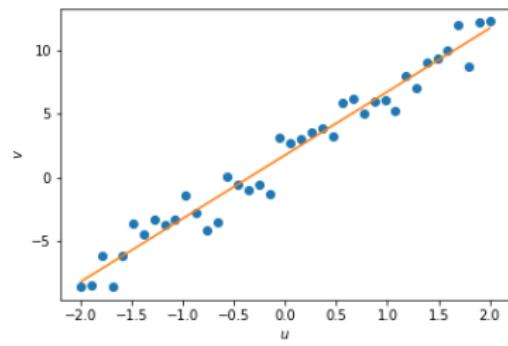
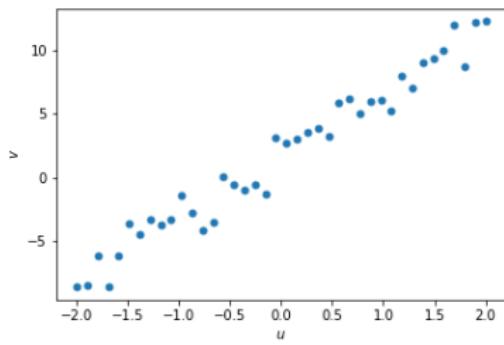
- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Wahl des Modells



Die Daten im Bild oben links werden von links oben nach rechts unten mit Polynomen von wachsendem Grad approximiert. Ab Polynomgrad 4 erreichen wir Interpolation.

Lineares Ausgleichsproblem



Daten ohne und mit Ausgleichsgerade

Lineares Ausgleichsproblem

Lineares Ausgleichsproblem

Gegeben: $(u_i, v_i) \in \mathbb{R}^2, \quad i = 1, \dots, n$, z.B. gemessene Punkte.

Gesucht: Gerade, die die Punkte (u_i, v_i) „am besten“ approximiert.

Idee: Wähle Gerade $g(u) := x_0 + x_1 u$ die den quadratischen Approximationsfehler minimiert, d.h.

$$\sum_{i=1}^n |v_i - g(u_i)|^2 = \sum_{i=1}^n (v_i - x_0 - x_1 u_i)^2 \longrightarrow \min \quad \star$$

oder $\underbrace{\|b - Ax\|_2}_{\text{lineares Ausgleichsproblem}} \longrightarrow \min$ mit $A = \begin{pmatrix} 1 & \cdots & 1 \\ u_1 & \cdots & u_n \end{pmatrix}^T$,

$b = (v_1, \dots, v_n)^T$, $x = (x_0, x_1)^T$. So können z.B. Messfehler in b ausgeglichen werden (siehe obige Abbildung).

Lösungen des Problems heißen **Kleinste-Quadrat-Lösungen** (auf Englisch **least squares solutions**).

Lineares Ausgleichsproblem

Anmerkung: Die Äquivalenz der beiden Formulierungen in Gleichung (★) lässt sich wie folgt nachvollziehen:

$$\sum_{i=1}^n (v_i - (x_0 + x_1 u_i))^2 = \left\| \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} - \begin{pmatrix} 1 & u_1 \\ \vdots & \vdots \\ 1 & u_n \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \right\|_2^2 = \|b - Ax\|_2^2$$

Da weiterhin $\|\cdot\|^2 \rightarrow \|\cdot\|$ eine monotone Umformung ist, gilt, dass die quadratische Minimierung $\|b - Ax\|_2^2 \rightarrow \min$ exakt dieselben Lösungen wie $\|b - Ax\|_2 \rightarrow \min$ besitzt.

Lineares Ausgleichsproblem

Definition 2.1 (Lineares Ausgleichsproblem)

Seien $A \in \mathbb{C}^{m \times n}$, $m \geq n$, $b \in \mathbb{C}^m$. Dann heißt das Minimierungsproblem:
Gesucht sei $x \in \mathbb{C}^n$ mit

$$x = \arg \min_{y \in \mathbb{R}^n} \|b - Ay\|_2 \quad (2.1)$$

lineares Ausgleichsproblem (LAGP).

Das Ausgleichproblem heißt **linear**, da die Parameter x nur linear in den Defekt $Ax - b$ eingehen.

Ausgleichspolynome

Wird statt der Ausgleichsgerade ein AusgleichsPolynom vom Grad k

$$g(u) = \sum_{i=0}^k a_i u^i$$

zu den Daten $(u_i, v_i) \in \mathbb{R}^2, i = 1, \dots, n$ gewünscht, so ergibt sich folgendes lineares Ausgleichsproblem:

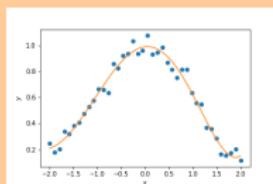
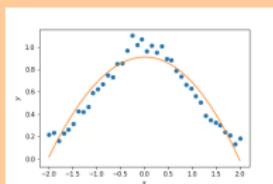
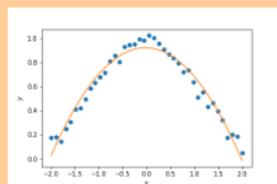
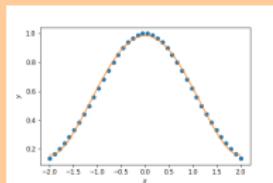
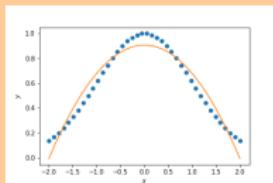
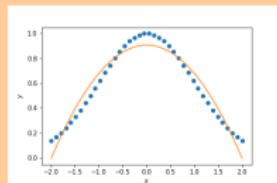
$$x = \arg \min_{y \in \mathbb{R}^n} \|b - Ay\|_2$$

mit

$$A = \begin{pmatrix} 1 & u_1 & u_1^2 & \cdots & u_1^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_n & u_n^2 & \cdots & u_n^k \end{pmatrix}, x = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix}, b = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Beispiel Approximation der Gaußschen Glockenkurve

Beispiel (Gaußsche Glockenkurve)



Approximation der Gaußschen Glockenkurve durch Polynome unterschiedlichen Grades: Links Grad 2, Mitte Grad 3 und Rechts Grad 4.
Die Daten in der zweiten Zeile wurde zufällig verauscht.

Lineares Ausgleichsproblem

Satz 2.1 (Lösung des linearen Ausgleichsproblems)

Sei $A \in \mathbb{C}^{m \times n}$, $b \in \mathbb{C}^m$. Jede Lösung x des linearen Ausgleichsproblems

$$x = \arg \min_{y \in \mathbb{R}^n} \|b - Ay\|_2 \quad (2.2)$$

ist eine Lösung der „Gaußschen Normalengleichung“

$$A^*Ax = A^*b. \quad (2.3)$$

Umgekehrt ist jede Lösung x der „Gaußschen Normalengleichung“ auch eine Lösung des linearen Ausgleichsproblems.

Die Matrix A^*A

Lemma 2.2 (Trivialer Nullraum)

Die Matrix A^*A ist hermitesch und positiv semidefinit, d.h. $x^*A^*Ax \geq 0$.
 A^*A ist genau dann positiv definit, wenn der Nullraum $\text{Kern}(A)$ trivial ist,
d.h. $\text{Kern}(A) = 0$.

Ferner ist $\text{Kern}(A^*A) = \text{Kern}(A)$ und $\text{Bild}(A^*A) = \text{Bild}(A^*) = \text{Kern}(A)^\perp$.

Beweis des Lemmas

Beweis: (Lemma 2.2)

- $(A^*A)^* = A^*(A^*)^* = A^*A$
- $x^*A^*Ax = \|Ax\|_2^2 \geq 0 \quad \forall x \in \mathbb{C}^n$, d.h. A^*A ist positiv semidefinit.
- Sei nun $x \in \text{Kern}(A^*A)$.

$$A^*Ax = 0 \Rightarrow x^*A^*Ax = 0 \Rightarrow \|Ax\|_2^2 = 0 \Rightarrow Ax = 0.$$

D.h. $\text{Kern}(A^*A) \subseteq \text{Kern}(A)$ und wegen $Ax = 0 \Rightarrow A^*Ax = 0$
 $\text{Kern}(A^*A) = \text{Kern}(A)$.

- Wenn $\text{Kern}(A) = \{0\}$ ist $\text{Kern}(A^*A) = 0$ und somit für alle $x \in \mathbb{C}^n \setminus \{0\}$ $x^*A^*Ax > 0$, d.h. A^*A ist positiv definit.
- Weiter gilt $\dim \text{Bild}(A^*A) = n - \dim \text{Kern}(A^*A) = n - \dim \text{Kern}(A) = \text{Rang}(A) = \dim \text{Bild}(A^*)$. Damit ist $\text{Bild}(A^*A) = \text{Bild}(A^*)$.
- Zu $\text{Bild}(A^*) = \text{Kern}(A)^\perp$: Sei $z \in \text{Bild}(A^*)$, $x \in \text{Kern}(A)$ beliebig.
 Dann existiert ein $y \in \mathbb{C}^m$ mit $z = A^*y$ und es gilt

$$x^*z = x^*A^*y = (Ax)^*y = 0 \cdot y = 0,$$

d.h. $\text{Kern}(A)$ und $\text{Bild}(A^*)$ sind orthogonal und wegen
 $\dim \text{Kern}(A) + \dim \text{Bild}(A^*) = n$ folgt $\text{Bild}(A^*) = \text{Kern}(A)^\perp$.



Beweis des Satzes

Beweis: (Satz 2.1) Setze für $x \in \mathbb{C}^n$

$$\phi(x) = \frac{1}{2} \|b - Ax\|_2^2 = \frac{1}{2} (b - Ax)^*(b - Ax).$$

Dann gilt für jedes \hat{x} mit $A^*A\hat{x} = A^*b$

$$\phi(x) - \phi(\hat{x}) = \frac{1}{2} (x - \hat{x})^* A^* A (x - \hat{x}) \quad (\text{quadratische Ergänzung}).$$

Da A^*A positiv semi-definit (2.2), ist folgt

$$\phi(x) - \phi(\hat{x}) \geq 0 \iff \phi(x) \geq \phi(\hat{x}) \quad \forall \hat{x} \text{ mit } A^*A\hat{x} = A^*b,$$

d.h. das Minimum von ϕ wird an den \hat{x} mit $A^*A\hat{x} = A^*b$ angenommen.
 Solche \hat{x} existieren, da A^*b zu $Bild(A^*) = Bild(A^*A)$ gehört.
 Es bleibt zu zeigen, dass Gleichheit genau dann gilt, wenn eine Lösung x des Ausgleichsproblems auch die Normalengleichung erfüllt.

Beweis des Satzes 2.1 (Fortsetz.)

(Fortsetzung des Beweises von Satz 2.1)

Um zu zeigen, dass Gleichheit genau dann gilt, wenn eine Lösung x des Ausgleichsproblems auch die Normalengleichung erfüllt, betrachten wir ein x , welches $\phi(x)$ minimiert. Mit einem wie oben gewählten \hat{x} gilt:

$$\begin{aligned} \phi(x) - \phi(\hat{x}) &= 0 \\ \Leftrightarrow \frac{1}{2}(x - \hat{x})^* A^* A (x - \hat{x}) &= 0 \\ \Leftrightarrow \frac{1}{2} \|A(x - \hat{x})\|_2^2 &= 0 \\ \Leftrightarrow (x - \hat{x}) &\in \text{Kern}(A). \end{aligned}$$

Unter Ausnutzung von $(x - \hat{x}) \in \text{Kern}(A)$ setzen wir in die Normalengleichung \hat{x} ein:

$$\begin{aligned} A^* b &= A^* A \hat{x} \\ &= A^* (A \hat{x} + \underbrace{A(x - \hat{x})}_0) \\ &= A^* A (\hat{x} + (x - \hat{x})) \\ &= A^* A x \end{aligned}$$

Residuum

Bemerkung 2.1 ()

Das *Residuum* (Rest) $r = b - A\hat{x}$ gehört zu $\text{Kern}(A^*)$:

$$A^*r = A^*b - A^*A\hat{x} = 0.$$

Da nach dem Lemma 2.2 $\text{Bild}(A^*) = \text{Kern}(A)^\perp$ gilt, ist auch $\text{Bild}(A) = \text{Kern}(A^*)^\perp$ beziehungsweise $\text{Kern}(A^*) = \text{Bild}(A)^\perp$, d.h. r steht orthogonal auf $\text{Bild}(A)$.

Alternativer Beweis für den reellen Fall

Bemerkung 2.2 (Minimierung mit Methoden der Differentialrechnung)

Im reellen Fall kann das Minimum in Gleichung 2.1 alternativ mit Methoden der Differentialrechnung bestimmt werden.

Sei dazu $d \in \mathbb{R}^n$, so berechnen wir die Richtungsableitung nach d an der Stelle x gemäß

$$\frac{\partial}{\partial d} \Phi(x) = \lim_{t \rightarrow 0} \frac{\Phi(x + td) - \Phi(x)}{t}.$$

Wegen

$$\begin{aligned} \Phi(x + td) - \Phi(x) &= \frac{1}{2}(b - Ax - tAd)^*(b - Ax - tAd) - \frac{1}{2}(b - Ax)^*(b - Ax) \\ &= t(Ad)^*(Ax - b) + \frac{1}{2}t^2(Ad)^*Ad. \end{aligned}$$

folgt

$$\frac{\partial}{\partial d} \Phi(x) = (Ad)^*(Ax - b)$$

Nullsetzen liefert eine notwendige Bedingung für ein Minimum.

Beispiel: Normalengleichung

Beispiel (Normalengleichung)

Wir versuchen die zugehörige Normalengleichung zu folgendem Problem zu finden:

$$A^* A x = A^* b$$

mit $A = \begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 4 \\ 1 & 7 \end{pmatrix}$ und $b = \begin{pmatrix} 1 \\ 2 \\ 6 \\ 4 \end{pmatrix}$. Dazu berechnen wir:

$$A^* A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 3 & 4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 4 \\ 1 & 7 \end{pmatrix} = \begin{pmatrix} 4 & 14 \\ 14 & 74 \end{pmatrix}$$

$$A^* b = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 3 & 4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 13 \\ 58 \end{pmatrix}$$

$$(A^* A)^{-1} = \frac{1}{4 \cdot 74 - 14 \cdot 14} \cdot \begin{pmatrix} 74 & -14 \\ -14 & 4 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix}.$$

Anwendung: Normalengleichung

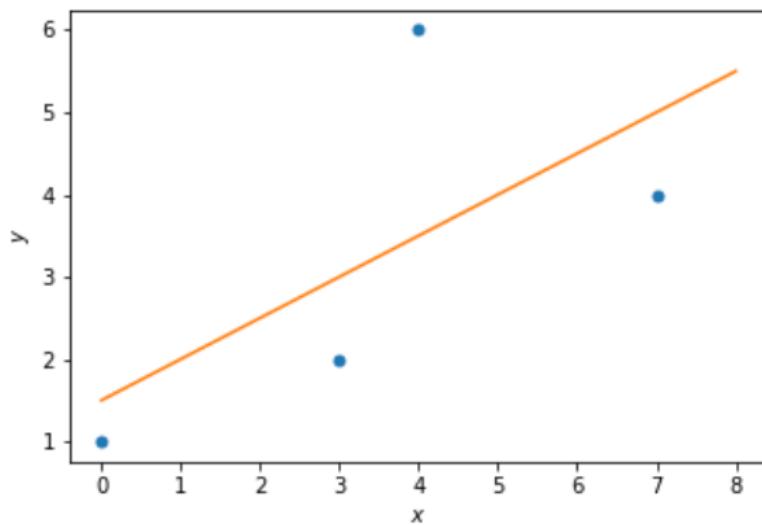


Abbildung 2.4: Daten als Punkte und Ausgleichsgerade

Tikhonov Regularisierung (Phillips '62, Tikhonov '63)

Bemerkung 2.3 (Tikhonov Regularisierung)

Phillips und Tikhonov minimieren nicht nur die kleinsten Fehlerquadrate, sondern schlagen vor, gleichzeitig die Norm des Lösungsvektors klein zu machen und betrachten für $\lambda > 0$, $A \in \mathbb{C}^{m \times n}$, $\mathbb{I} = \text{Id}_n \in \mathbb{R}^{n \times n}$.

$$\hat{x} = \arg \min_{x \in \mathbb{R}^n} (\|Ax - b\|_2^2 + \lambda \|x\|_2^2) \quad (2.4)$$

Diese Problem kann auch als gewöhnliches Kleinste Quadrate Problem geschrieben werden:

$$\hat{x} = \arg \min_{x \in \mathbb{R}^n} \left(\left\| \begin{pmatrix} A \\ \sqrt{\lambda} \mathbb{I} \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} \right\|_2^2 \right)$$

Tikhonov Regularisierung (Phillips '62, Tikhonov '63)

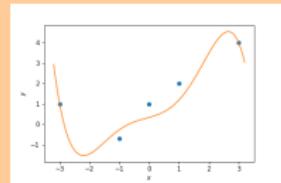
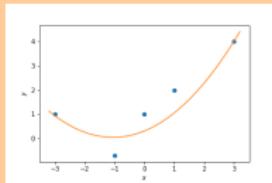
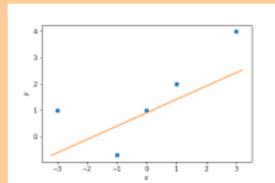
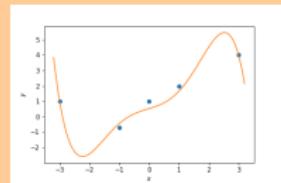
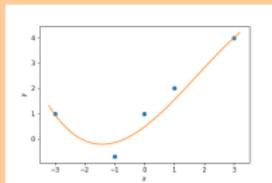
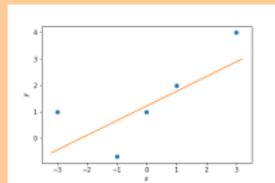
Bemerkung 2.4 (Tikhonov Regularisierung)

Für die zugehörige Normalengleichung ergibt sich

$$\begin{aligned}(A^* - \sqrt{\lambda}Id) \begin{pmatrix} A \\ \sqrt{\lambda}Id \end{pmatrix} x &= (A^* - \sqrt{\lambda}Id) \begin{pmatrix} b \\ 0 \end{pmatrix} \\ \Leftrightarrow (A^*A + \lambda Id)x &= A^*b\end{aligned}$$

Lineare Ausgleichsrechnung

Beispiel (Thikhonov Regularisierung)



In den Bildern der Spalten Approximation der Daten jeweils mit Polynomgrad 1,3 und 5. In der oberen Reihe mit Regularisierung $\lambda = 1$. Untere Zeile mit $\lambda = 5$.

Kapitel 3:

Singulärwertzerlegung (SVD)

(Singular Value Decomposition)

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Wiederholung: Spektralwertzerlegung

Definition 3.1 (**Selbstadjungierter Endomorphismus**)

Sie $V, \langle \cdot, \cdot \rangle$ ein euklidischer Vektorraum. Ein Endomorphismus $f : V \rightarrow V$ heißt selbstadjungiert (bzgl. des gegebenen Skalarprodukt $\langle \cdot, \cdot \rangle$), falls für beliebige $v, w \in V$ gilt:

$$\langle f(v), w \rangle = \langle v, f(w) \rangle$$

Bemerkung 3.1 (**Matrixdarstellung selbstadjungierter Endomorphismen**)

Sei $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V und $A = DM_B(f)$, so gilt

$$f \text{ selbstadjungiert} \Leftrightarrow A^t = A$$

Wiederholung: Spektralsatz für euklidische Vektorräume

Satz 3.1 (Spektralsatz)

Sei $f : V \rightarrow V$ ein selbstadjungierter Endomorphismus eines n -dimensionalen euklidischen Vektorraums V . Dann besitzt V eine Orthonormalbasis, die aus Eigenvektoren von f besteht.

Korollar 3.2 (Diagonalisierung hermitischer Matrizen)

Sei $A \in \mathbb{C}^{n \times n}$ hermitisch, d.h. $A^* = A$. Dann gibt es eine unitäre Matrix T , d.h. $T^* \cdot T = Id_n$, so dass

$$T^* A T = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Unitäres Skalarprodukt und unitärer Vektorraum

Definition 3.2 (Unitäres Skalarprodukt und unitärer Vektorraum)

Ist V ein komplexer Vektorraum, so heißt eine Abb. von $V \times V \rightarrow \mathbb{C}$ mit

$$(v, w) \mapsto v \cdot w = \langle v, w \rangle$$

ein unitäres Skalarprodukt, wenn für alle $v, v' \in V$ und $\lambda \in \mathbb{C}$ die folgenden Eigenschaften erfüllt sind:

- 1 $(u + v') \cdot w = v \cdot w + v' \cdot w$ und $(\lambda v) \cdot w = \lambda(v \cdot w)$ (Linearität im ersten Argument),
- 2 $v \cdot w = \overline{v \cdot w}$ (hermitesch)
- 3 $v \cdot v \geq 0$ und $v \cdot v = 0 \Leftrightarrow v = 0$ (positive Definitheit).

Ist $\langle \cdot, \cdot \rangle$ ein unitäres Skalarprodukt in V , so nennt man V einen unitären Vektorraum.

Die Singulärwertzerlegung

Sei p der Rang von A ($=$ Rang von A^*A) und seien $\lambda_1, \dots, \lambda_n$ die absteigend sortierten Eigenwerte von A^*A , d.h.

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > \lambda_{p+1} = \dots = \lambda_n = 0,$$

sowie $v_1, \dots, v_n \in \mathbb{C}^n$ eine Basis orthonormaler Eigenvektoren ($A^*Av_i = \lambda_i v_i$ $\forall i = 1, \dots, n$, $v_i^*v_j = \delta_{ij}$, $i, j = 1, \dots, n$). Eine solche *Spektralzerlegung* existiert, da A^*A hermitesch und positiv semidefinit ist.

Nun definieren wir Vektoren $u_i \in \mathbb{C}^m$ durch

$$u_i = \frac{1}{\sqrt{\lambda_i}} Av_i, \quad i = 1, \dots, p.$$

Für $1 \leq i, j \leq p$ gilt dann

$$u_i^* u_j = \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}} (Av_i)^* (Av_j) = \frac{1}{\sqrt{\lambda_i \lambda_j}} v_i^* (A^* A v_j) = \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} v_i^* v_j = \delta_{ij}.$$

Die Singulärwertzerlegung

Die Vektoren u_i , $i = 1, \dots, p$, bilden also eine Orthonormalbasis des gesamten Bildraumes $\text{Bild}(A)$, denn

$$\dim \text{Bild}(A) = n - \dim \text{Kern}(A) = n - \dim \text{Kern}(A^* A) = n - (n - p) = p.$$

Diese Basis kann durch weitere $m - p$ Vektoren u_{p+1}, \dots, u_m zu einer orthonormalen Basis des \mathbb{C}^m erweitert werden (vgl. Basisergänzungssatz). Des Weiteren gilt

$$A^* u_i = \begin{cases} \frac{1}{\lambda_i} A^* A v_i = \frac{\lambda_i}{\sqrt{\lambda_i}} v_i = \sqrt{\lambda_i} v_i & \forall i = 1, \dots, p, \\ 0 & \forall i \geq p + 1, \dots, m. \end{cases}$$

Graphische Darstellung

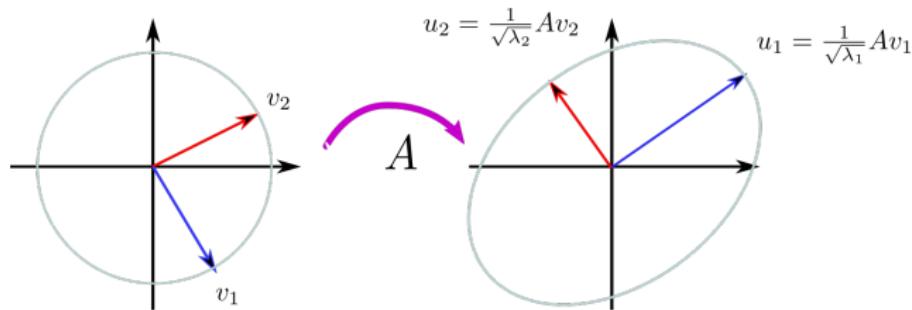


Abbildung 3.5: Grapische Darstellung der SVD

Singulärwertzerlegung

Definition 3.3 (Singulärwertzerlegung)

Sei $A \in \mathbb{C}^{m \times n}$ eine Matrix mit Rang p . Ein System

$$\{\sigma_i, u_j, v_k : i = 1, \dots, p, j = 1, \dots, m, k = 1, \dots, n\} \quad (3.1)$$

mit $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p > 0$ und Orthonormalbasen $\{u_j\}_{j=1}^m, \{v_k\}_{k=1}^n$ des \mathbb{C}^m bzw. \mathbb{C}^n , wobei

$$\begin{aligned} Av_i &= \sigma_i u_i, & A^* u_i &= \sigma_i v_i, & \forall i = 1, \dots, p, \\ Av_k &= 0, & A^* u_j &= 0, & j, k > p, \end{aligned}$$

heißt *Singulärwertzerlegung* von A . Die σ_i heißen Singulärwerte von A . Ihre Quadrate σ_i^2 sind die von 0 verschiedenen Eigenwerte von A^*A .

Singulärwertzerlegung

Satz 3.3 (Existenz der Singulärwertzerlegung)

Jede Matrix $A \in \mathbb{C}^{m \times n}$ mit Rang p besitzt eine Singulärwertzerlegung.

Zur Erinnerung:

Singulärwertzerlegung von $A \in \mathbb{C}^{m \times n}$: Ein System

$$\{\sigma_i, u_j, v_k : i = 1, \dots, p, j = 1, \dots, m, k = 1, \dots, n\}$$

mit $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p > 0$ und Orthonormalbasen $\{u_j\}_{j=1}^m, \{v_k\}_{k=1}^n$ des \mathbb{C}^m bzw. \mathbb{C}^n , wobei

$$\begin{aligned} Av_i &= \sigma_i u_i, & A^* u_i &= \sigma_i v_i, & i &= 1, \dots, p, \\ Av_k &= 0, & A^* u_i &= \sigma_i v_i, & \forall i &= 1, \dots, p, \\ Av_k &= 0, & A^* u_i &= 0, & j, k &> p. \end{aligned}$$

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

SVD in Matrixnotation

Seien $U \in \mathbb{C}^{m \times m}$ und $V \in \mathbb{C}^{n \times n}$ mit

$$U = \begin{pmatrix} & & & \\ u_1 & | & u_2 & | & \dots & | & u_m & | \\ & & & & & & & \end{pmatrix}, \quad V = \begin{pmatrix} & & & \\ v_1 & | & v_2 & | & \dots & | & v_n & | \\ & & & & & & & \end{pmatrix}$$

und $\Sigma = \begin{pmatrix} \sigma_1 & & 0 & 0 \\ & \ddots & & 0 \\ 0 & & \sigma_p & 0 \\ 0 & 0 & & 0 \end{pmatrix}$.

Dann gilt:

$$AV = U\Sigma \Leftrightarrow A = U\Sigma V^* \Leftrightarrow A^* = V\Sigma^* U^* \Leftrightarrow A = \sum_{i=1}^p \sigma_i u_i v_i^*. \quad (3.2)$$

Schematische Darstellung

$$\Sigma \begin{pmatrix} n \\ A \end{pmatrix} = \Sigma \begin{pmatrix} m \\ U \end{pmatrix} \Sigma \begin{pmatrix} n \\ \Sigma \end{pmatrix} \Sigma \begin{pmatrix} n \\ V^T \end{pmatrix}$$

Abbildung 3.6: Schematische Darstellung der SVD

 $m > n :$

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

 $m = n :$

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \end{pmatrix}$$

 $m < n :$

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & \sigma_n & 0 \end{pmatrix}$$

Geometrische Interpretation

Zur geometrischen Interpretation beschränken wir uns auf reelle Matrizen.

Sei S die Einheitssphäre im \mathbb{R}^n , $A \in \mathbb{R}^{m \times n}$, $m \geq n$. A habe vollen Rang. Seien dann $\sigma_1, \dots, \sigma_n$ die Längen der Halbachsen von AS (siehe Abbildung) und $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$ die Singulärwerte von A .

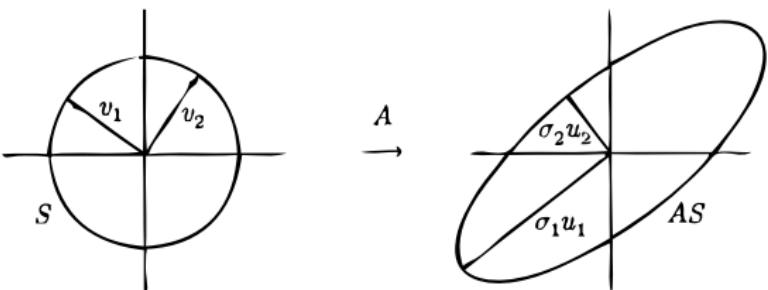


Abbildung 3.7: Ellipse mit gestreckten bzw.
gestauchten Hauptachsen

Geometrische Interpretation

Die normierten Einheitsvektoren entlang der Hauptachsen $\{u_1, \dots, u_n\}$ sind die *Linkseigenvektoren*. Die *Rechtseigenvektoren* von A sind die Einheitsvektoren $\{v_1, \dots, v_n\} \in S$, die durch $Av_j = \sigma_j u_j$ ($\forall j \in \{1, \dots, n\}$) auf die Hauptachse abgebildet werden.

In Matrixschreibweise:

$$\left(\begin{array}{c|ccccc} & A & & & & \\ \hline & & & & & \end{array} \right) \cdot \left(\begin{array}{c|c|c|c} v_1 & v_2 & \dots & v_n \\ \hline & & & \end{array} \right) = \left(\begin{array}{c|c|c|c} u_1 & u_2 & \dots & u_n \\ \hline & & & \end{array} \right) \cdot \left(\begin{array}{cccc} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{array} \right)$$

Geometrische Interpretation

Korollar 3.4 (Bild der Einheitkugel)

Das Bild der Einheitskugel bezüglich der 2-Norm $\|\cdot\|_2$ unter einer beliebigen $(m \times n)$ -Matrix ist ein verallgemeinertes Ellipsoid.

Wenn die Vektoren $v = \sum_{i=1}^n a_i v_i$ die Einheitssphäre des \mathbb{R}^n durchlaufen, d.h. $\sum_{i=1}^n a_i^2 = 1$, dann durchlaufen ihre Bilder $Av = \sum_{i=1}^n \sigma_i a_i u_i$ ein verallgemeinertes Ellipsoid, denn es gilt

$$\sum_{i=1}^n \frac{1}{\sigma_i^2} \sigma_i^2 a_i^2 = \sum_{i=1}^n a_i^2 = 1.$$

Dies ist die Gleichung eines verallgemeinerten Ellipsoids mit Scheitelpunkten $\pm \sigma_i e_i$, wobei e_i der i -te Basisvektor ist.

Beispiel Singulärwertzerlegung

Beispiel (Rezept zur Berechnung einer SVD)

Die Singulärwertzerlegung $A = U\Sigma V^T$ einer Matrix $A \in \mathbb{R}^{m \times n}$ kann man folgendermaßen berechnen:

- 1 Berechne $B = A^T A$.
- 2 Berechne die Eigenwerte von B :
 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > \lambda_{k+1} = \dots = \lambda_n = 0$.
- 3 Die normierten Eigenvektoren $\{v_1, \dots, v_n\}$ sind eine Orthonormal-Basis des \mathbb{R}^n , wobei v_i ein EV zum EW λ_i ist. Dann ist $V := [v_1, \dots, v_n]$. Es gilt: $V^T = V^{-1}$.
- 4 Bilde die Diagonalmatrix Σ aus den Singulärwerten von A : $\sigma_i = \sqrt{\lambda_i}$.
- 5 Finde die Matrix U durch Berechnen der Vektoren $u_i = \frac{1}{\sqrt{\lambda_i}} A v_i$ für $i \leq k$. Ergänze diese Vektoren zu einer ON-Basis des \mathbb{R}^m (vgl. Basisergänzungssatz in Kapitel 1).

Beispiel Bildkompression

Beispiel (Beispiel Bildkompression)

Gegeben sei ein Grauwertbild, dessen Speicherbedarf reduziert werden soll.

Lösung:

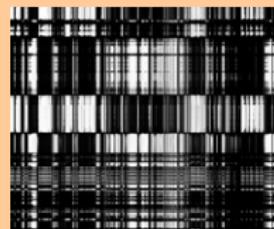
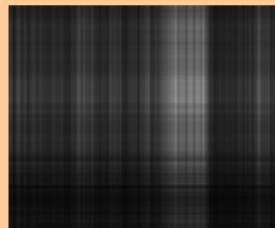
- 1 Fasse das Grauwertbild mit $n \times m$ Pixeln als Matrix $A \in \mathbb{R}^{n \times m}$ auf.
- 2 Berechne die SVD von $A = U\Sigma V^*$.
- 3 Berechne eine **Rank-k Approximation** von A :

$$\hat{A}(k) := \sum_{i=1}^k \sigma_i u_i v_i^t$$

- 4 Speichere Matrix als Bild.

Beispiel Bildkompression

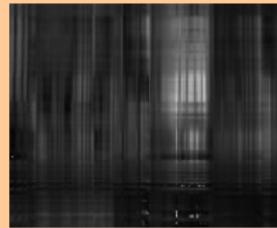
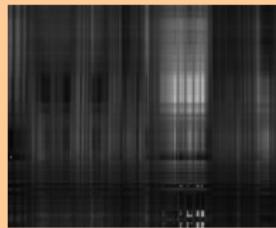
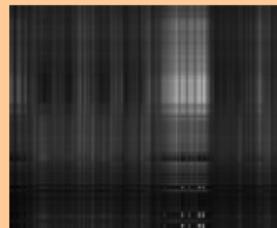
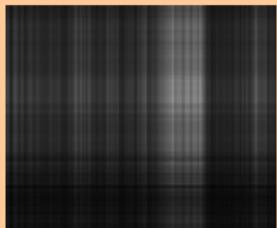
Beispiel (Beispiel Bildkompression)



Die ersten fünf Hauptkomponenten des links oben gezeigten Bilds aus dem Kölner Dom.

Beispiel Bildkompression

Beispiel (Beispiel Bildkompression)



Oben links eine Rang 50 Approximation des Bilds aus dem Kölner Dom.
Folgend Rang-1 bis Rang-5 Approximationen.

Matrix Approximation

Theorem (Spektralnorm) Sei $A \in \mathbb{C}^{m \times n}$ und σ_1 der größte Singulärwert von A . Dann gilt

$$\|A\|_2 = \sup \frac{\|Ax\|_2}{\|x\|_{(2)}} = \sigma_1.$$

Theorem (Eckart-Young Theorem) Sei $A \in \mathbb{C}^{m \times n}$ ein Matrix vom Rang r und sei $B \in \mathbb{C}^{m \times n}$ eine Matrix vom Rang k . Für jedes $k \leq r$ und jede Rang- k Approximation $\hat{A}(k) := \sum_{i=1}^k \sigma_i u_i v_i^t$ von A gilt

$$\begin{aligned}\hat{A}(k) &= \arg \min_{\{B \in \mathbb{C}^{m \times n} | \text{rang}(B)=k\}} \|A - B\|_2, \\ \|A - \hat{A}(k)\|_2 &= \sigma_{k+1}.\end{aligned}$$

Für die Beweise verweisen wir auf die Übungen bzw. die Literatur.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Zur Erinnerung: Eigenwertzerlegung beliebiger Matrizen

Definition 3.4 (Eigenwertzerlegung)

Eine Zerlegung der Form

$$A = V \Lambda V^{-1} \quad (3.3)$$

heißt *Eigenwertzerlegung* (oder *Spektralzerlegung*) von A .

Satz 3.5 (Diagonalisierbarkeit)

Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann diagonalisierbar, wenn sie n linear unabhängige Eigenvektoren besitzt.

Zur Erinnerung: Eigenwertzerlegung beliebiger Matrizen

Bemerkung 3.2 (Eigenwertzerlegung in Matrixschreibweise)

Besitzt $A \in \mathbb{C}^{n \times n}$ n linear unabhängige Eigenvektoren v_i , so gilt für alle $i = 1, \dots, n$

$$Av_i = \lambda_i v_i$$

In Gleichung (3.3) erhalten wir damit

$$A = \left(\begin{array}{c|c|c|c} v_1 & v_2 & \dots & v_n \end{array} \right) \underbrace{\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}}_{\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)} \left(\begin{array}{c|c|c|c} v_1 & v_2 & \dots & v_n \end{array} \right)^{-1}$$

Unterschiede zwischen SVD und Eigenwertzerlegung

SVD

$$A = U\Sigma V^*$$

vs.

EVD

$$A = V\Lambda V^{-1}$$

- SVD verwendet zwei Basen (U, V)
EVD verwendet eine Basis (V)
- SVD verwendet orthonormale Basen
Basen der EVD sind nicht zwingend orthonormal
- SVD existiert für jede Matrix A
EVD existiert nicht für jede Matrix A
- SVD ist interessant zum Studium der Matrix A selbst
EVD ist interessant zum Studium von Matrixpotenzen A^k, e^{tA} :
 $Av_i = \lambda_i v_i \Leftrightarrow A^k v_i = \lambda_i^k v_i$

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Pseudoinverse

Definition 3.5 (Pseudoinverse)

Sei $A = U\Sigma V^*$ die Singulärwertzerlegung von $A \in \mathbb{C}^{m \times n}$. Dann heißt

$$A^+ = V\Sigma^+U^* \in \mathbb{C}^{n \times m}$$

mit

$$\Sigma^+ = \begin{pmatrix} \sigma_1^{-1} & & 0 \\ & \ddots & 0 \\ 0 & \dots & \sigma_p^{-1} & 0 \\ & & 0 & 0 \end{pmatrix} \in \mathbb{C}^{n \times m}$$

Pseudoinverse oder *Moore-Penrose-Inverse* von A . Man kann nachrechnen:

$$A^+ = \sum_{i=1}^p \sigma_i^{-1} v_i u_i^*,$$

$$\text{Kern}(A^+) = \text{Kern}(A^*) = \text{Bild}(A)^\perp, \quad \text{Bild}(A^+) = \text{Bild}(A^*) = \text{Kern}(A)^\perp.$$

Pseudoinverse und Lineare Ausgleichsrechnung

Satz 3.6 (Lösung des Ausgleichsproblems)

Der Vektor A^+b ist die eindeutig bestimmte Lösung des linearen Ausgleichsproblems

$$\|b - Ax\|_2 \longrightarrow \min$$

(vgl. Gleichung (\star) am Anfang von Kapitel 4) mit minimaler euklidischer Norm.

Wir beweisen die Aussage in mehreren Schritten:

Beweis:

- 1 $\forall b \in \mathbb{C}^m$ gilt $AA^+b - b \in \text{Kern}(A^+)$:

$$A^+AA^+ = V\Sigma^+U^*U\Sigma V^*\Sigma^+U^* = V\Sigma^+U^* = A^+$$

Daher gilt: $A^+(AA^+b - b) = A^+AA^+b - A^+b = A^+b - A^+b = 0$.

- 2 A^+b erfüllt die Normalengleichung $A^*A(A^+b) = A^*b$:

$$A^*A(A^+b) - A^*b = A^* \underbrace{(AA^+b - b)}_{\in \text{Kern}(A^+) = \text{Kern}(A^*)} = 0$$

Pseudoinverse und Lineare Ausgleichsrechnung

1 ...

2 ...

3 Ist z eine weitere Lösung der Normalengleichung, so gilt:

$$A^* A z = A^* A (A^+ b) = A^* b \text{ und somit}$$

$$w = A^+ b - z \in \text{Kern}(A^* A) = \text{Kern}(A)$$

Andererseits ist $A^+ b \in \text{Bild}(A^+) = \text{Kern}(A)^\perp$. Daher ist

$$z = \underbrace{A^+ b}_{\text{Kern}(A)^\perp} - \underbrace{w}_{\text{Kern}(A)} \text{ eine orthogonale Zerlegung und nach Pythagoras}$$

gilt

$$\|z\|_2^2 = \|A^+ b\|_2^2 + \|w\|_2^2 \geq \|A^+ b\|_2^2$$

mit Gleichheit genau dann, wenn $\|w\|_2^2 = 0$, d.h. für $z = A^+ b$.



Pseudoinverse und Lineare Ausgleichsrechnung

Korollar 3.7 (Trivialer Nullraum)

Ist $\text{Kern}(A) = \{0\}$, so gilt $A^+ = (A^*A)^{-1}A^*$.

Beweis: Die Lösung des linearen Ausgleichsproblems $\|b - Ax\| \rightarrow \min$ ist laut Normalengleichung gegeben durch $(A^*A)^{-1}A^*b$, aber auch durch A^+b . Da dies für beliebiges b gilt folgt die Behauptung. □

Anschaulich ist die Pseudoinverse A^+ also eine Matrix, die die Lösungen des linearen Ausgleichsproblems zur Matrix A für beliebiges b beschreibt. Entsprechend stellt AA^+ die orthogonale Projektion auf das Bild von A dar.

Kapitel 4: QR-Zerlegung

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Projektionsmatrix

Definition 4.1 (Projektionsmatrix)

Eine quadratische Matrix $P \in \mathbb{C}^{n \times n}$ heißt *Projektionsmatrix*, falls

$$P^2 = P.$$

Bemerkung 4.1 (Bild und Kern)

Sei $v \in \mathbb{C}^n$, $P \in \mathbb{C}^{n \times n}$.

1 Falls $v \in \text{Bild}(P)$, so gilt $Pv = v$.

2 Sei $v \in \mathbb{C}^n$. Dann liegt

$$Pv - v \in \text{Kern}(P).$$

Denn:

$$P(Pv - v) = P^2v - Pv = Pv - Pv = 0.$$

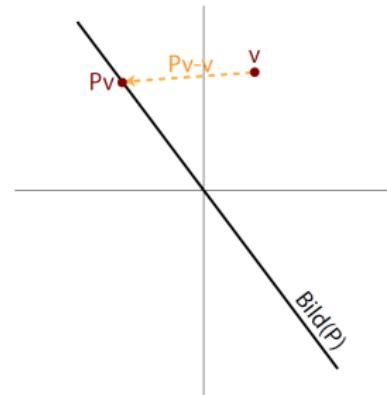


Abbildung 4.8: Beispiel einer Projektion.

Komplementäre Projektionsmatrix

Satz 4.1 (Komplementäre Projektionsmatrix)

Sei $P \in \mathbb{C}^{n \times n}$ eine Projektionsmatrix. Dann ist auch $I - P$ eine Projektionsmatrix. $I - P$ heißt *komplementäre Projektionsmatrix*.

Beweis: Da P eine Projektionsmatrix ist, gilt zunächst $P^2 = P$. Daher gilt für $I - P$:

$$(I - P)^2 = I - 2P + \underbrace{P^2}_{=P} = I - P.$$



Wir können unter anderem folgende Beobachtungen machen:

- 1 $I - P$ projiziert auf $\text{Kern}(P)$, d.h. $\text{Bild}(I - P) = \text{Kern}(P)$.
- 2 Es gilt: $\text{Kern}(I - P) = \text{Bild}(P)$.
- 3 Es ist $\text{Kern}(I - P) \cap \text{Kern}(P) = \{0\}$ (und damit auch $\text{Bild}(I - P) \cap \text{Bild}(P) = \{0\}$).

(Beweis der Beobachtungen zu komplementären Matrizen:)

Beweis:

1 Z.z.: $\text{Bild}(I - P) = \text{Kern}(P)$:

(i) Es gilt zum einen die Mengeninklusion $\text{Bild}(I - P) \supseteq \text{Kern}(P)$, denn:

$$\forall v \in \text{Kern}(P) : v = Iv - \underbrace{Pv}_{=0} = (I - P)v \in \text{Bild}(I - P).$$

(ii) Zum anderen gilt aber auch die Inklusion $\text{Bild}(I - P) \subseteq \text{Kern}(P)$, denn:

$$\forall v \in \text{Bild}(I - P) : (I - P)v = v - Pv \in \text{Kern}(P).$$

2 Z.z.: $\text{Kern}(I - P) = \text{Bild}(P)$:

Da $(I - P)$ eine Projektionsmatrix ist (Satz 4.1), folgt:

$$\text{Kern}(I - P) = \text{Bild}(I - (I - P)) = \text{Bild}(P).$$

3 Z.z. $\text{Kern}(I - P) \cap \text{Kern}(P) = \{0\}$:

$$\text{Es gilt } \forall v \in \text{Kern}(I - P) \cap \text{Kern}(P): 0 = (I - P)v = v - \underbrace{Pv}_{=0} = v.$$



Wir haben also gezeigt, dass eine Projektion den \mathbb{C}^n in zwei Teilräume teilt. Sind umgekehrt S_1, S_2 zwei Teilräume des \mathbb{C}^n mit $S_1 \cap S_2 = \{0\}$ und $S_1 \oplus S_2 = \mathbb{C}^n$. Dann gibt es eine Projektion P , sodass

$$\text{Bild}(P) = S_1 \quad \text{und} \quad \text{Kern}(P) = S_2.$$

P heißt dann Projektion auf S_1 entlang S_2 .

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Orthogonale Projektion

Definition 4.2 (Orthogonale Projektion)

Eine *orthogonale Projektion* ist eine Projektion, die entlang eines Unterraums S_2 auf S_1 projiziert, wobei S_1 und S_2 orthogonal sind.

Theorem 4.2 (Orthogonale Projektion)

Eine Projektion ist genau dann orthogonal, wenn gilt: .

$$P = P^*.$$

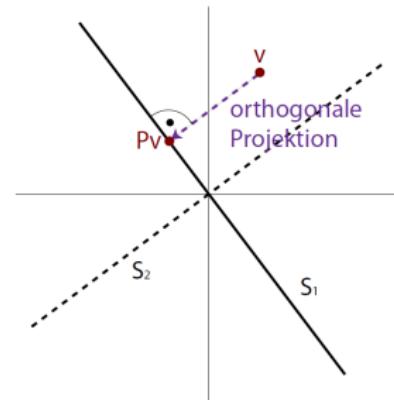


Abbildung 4.9: Beispiel einer orthogonalen Projektion.

Beweis von Theorem 4.2

Beweis: Sei $P = P^*$, dann gilt für beliebige $x, y \in \mathbb{C}^n$:

$$\underbrace{(Px)^*}_{\in \text{Bild}(P)=S_1} \quad \overbrace{(I - P)y}^{\in \text{Bild}(I-P)=S_2} = x^* P^*(I - P^2)y = 0.$$

D.h. Px und $(I - P)y$ sind orthogonal zueinander.

Sei umgekehrt P eine orthogonale Projektion auf S_1 entlang S_2 , wobei S_1 und S_2 orthogonal sind und $\dim S_1 = k$.

Dann wählen wir eine Orthonormalbasis $\{q_1, \dots, q_n\}$ des \mathbb{C}^n , sodass $\{q_1, \dots, q_k\}$ eine Basis von S_1 und $\{q_{k+1}, \dots, q_n\}$ eine Basis von S_2 ist .

Dann gilt

$$\begin{aligned}\forall j \leq k : \quad Pq_j &= q_j \quad \text{und} \\ \forall j > k : \quad Pq_j &= 0.\end{aligned}$$

(Fortsetzung des Beweises von Theorem 4.2)

Sei nun Q eine unitäre Matrix mit den Spalten q_j . Dann gilt

$$Q^*PQ = \begin{pmatrix} 1 & 0 & \dots & & \dots & 0 \\ 0 & 1 & & & & \vdots \\ \vdots & & \ddots & & & \\ & & & 1 & 0 & \\ \vdots & & & & \ddots & \vdots \\ 0 & \dots & & & \dots & 0 \end{pmatrix} =: \Sigma.$$

Die Matrix Q^*PQ ist diagonal, wobei die ersten k Einträge 1 sind.

Damit haben wir (wegen $Q^*Q = I$) mit $P = Q\Sigma Q^*$ eine SVD (Kap. 5) von P , für die gilt:

$$P^* = (Q\Sigma Q^*)^* = Q\Sigma Q^* = P.$$



- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Projektion mit orthonormaler Basis

Sei $\{q_1, \dots, q_n\}$ eine Menge orthonormaler Vektoren des \mathbb{C}^m und \hat{Q} die zugehörige $(m \times n)$ -Matrix. Dann gilt für alle $v \in \mathbb{C}^m$:

$$v = r + \sum_{i=1}^n (q_i q_i^*) v,$$

mit $r \in \langle q_1, \dots, q_n \rangle^\perp$
 und $\|q_i\| = 1 \quad \forall i = 1, \dots, n.$

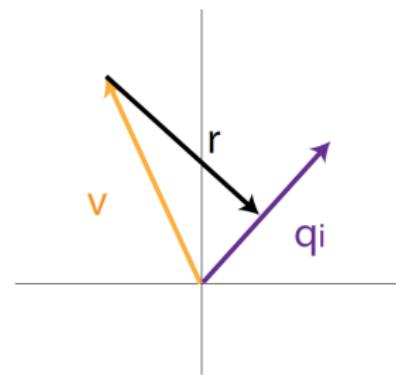


Abbildung 4.10: Projektion mit orthonormaler Basis.

Projektion mit orthonormaler Basis

Daher ist die Abbildung $v \mapsto \sum_{i=1}^n (q_i q_i^*) v$ eine orthogonale Projektion auf $\text{Bild}(\hat{Q})$:

$$\begin{pmatrix} w \\ \vdots \\ w \end{pmatrix} = \begin{pmatrix} & Q \\ Q & \end{pmatrix} \cdot \begin{pmatrix} & Q^* \\ Q^* & \end{pmatrix} \begin{pmatrix} v \\ \vdots \\ v \end{pmatrix}.$$

Somit ist jedes Produkt $\hat{Q} \hat{Q}^*$ eine Projektion auf den Spaltenraum von \hat{Q} .

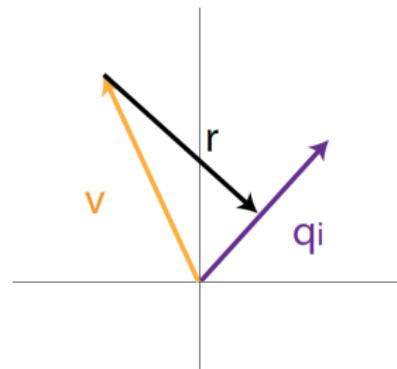


Abbildung 4.11:
Schematische Darstellung
der vollständigen
QR-Faktorisierung.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Idee der QR-Zerlegung

Seien zunächst $A \in \mathbb{C}^{m \times n}$ eine Matrix mit vollem Rang (1.3) und a_j , $j = 1, \dots, n$, deren Spaltenvektoren. Wir betrachten die geschachtelten Unterräume von \mathbb{C}^m , die von den Spalten a_j aufgespannt werden:

$$\underbrace{\langle a_1 \rangle}_{\dim \leq 1} \subseteq \underbrace{\langle a_1, a_2 \rangle}_{\dim \leq 2} \subseteq \underbrace{\langle a_1, a_2, a_3 \rangle}_{\dim \leq 3} \subseteq \dots \underbrace{\langle a_1, a_2, a_3, \dots, a_n \rangle}_{\dim \leq n} \quad (4.2)$$

Dabei ist $\langle a_1, \dots, a_j \rangle$ der von den Vektoren a_1, \dots, a_j aufgespannte Unterraum.

Konstruiere nun eine Folge *orthonormaler Vektoren* q_1, q_2, \dots etc., die diese verschachtelten Unterräume aufspannen, d.h.:

$$\langle q_1, q_2, \dots, q_j \rangle = \langle a_1, a_2, \dots, a_j \rangle \quad \text{für } j = 1, \dots, n. \quad (4.3)$$

Hinführung zur reduzierten QR-Zerlegung

Somit haben wir mit (4.2) und (4.3)

$$\left(\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_n \end{array} \right) = \underbrace{\left(\begin{array}{c|c|c|c} q_1 & q_2 & \dots & q_n \end{array} \right)}_{\hat{Q} \in \mathbb{C}^{m \times n}} \underbrace{\left(\begin{array}{ccccc} r_{11} & r_{12} & \dots & \dots & r_{1n} \\ 0 & r_{22} & & & \vdots \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & r_{nn} \end{array} \right)}_{\hat{R} \in \mathbb{C}^{n \times n}}, \quad (4.4)$$

wobei die Einträge auf der Diagonalen ungleich null sind, also $r_{kk} \neq 0 \forall k = 1, \dots, n$.

Reduzierte QR-Faktorisierung

Ausgeschrieben erhält man demnach für die einzelnen Spalten

$$\begin{aligned} a_1 &= r_{11}q_1 \\ a_2 &= r_{12}q_1 + r_{22}q_2 \\ a_3 &= r_{13}q_1 + r_{23}q_2 + r_{33}q_3 \\ &\vdots \\ a_n &= r_{1n}q_1 + r_{2n}q_2 + \dots + r_{nn}q_n \end{aligned} \tag{4.5}$$

und als Matrix

$$A = \hat{Q}\hat{R},$$

wobei \hat{Q} eine $(m \times n)$ -Matrix mit orthonormalen Spalten und \hat{R} eine rechte obere $(n \times n)$ -Dreiecksmatrix ist.

Definition 4.3 (Reduzierte QR-Faktorisierung)

Die in Gleichung (4.4) dargestellte Repräsentation heißt *reduzierte QR-Faktorisierung*.

Vollständige QR-Zerlegung

Bei der *vollständigen QR-Faktorisierung* von $A \in \mathbb{C}^{m \times n}$ werden $(m - n)$ orthonormale Spalten zu \hat{Q} hinzugefügt, sodass $\hat{Q} \in \mathbb{C}^{m \times m}$ unitär wird. Gleichzeitig werden $(m - n)$ Zeilen mit Nullen an die rechte obere Dreiecksmatrix angehängt.

Schematisch sieht dies wie folgt aus:

$$\begin{matrix} & n \\ m & \end{matrix} \boxed{A} = \boxed{\begin{matrix} & n \\ & \hat{Q} \\ m & \end{matrix}} \quad \boxed{\begin{matrix} q_{n+1} & \dots & q_m \\ m & \end{matrix}} \quad \boxed{\begin{matrix} R \\ \vdots \\ 0 \\ \dots \\ m \end{matrix}}$$

Q

Abbildung 4.12: Projektion mit orthonormaler Basis

QR-Faktorisierung

Theorem 4.3 (QR-Faktorisierung)

Jede Matrix $A \in \mathbb{C}^{m \times n}$, $m \geq n$, besitzt eine QR-Faktorisierung.

Eine mögliche Anwendung der QR-Zerlegung ist das Lösen von Gleichungssystemen:

Bemerkung 4.3 ()

Ein Gleichungssystem kann wie folgt gelöst werden

Gegeben: $A \in \mathbb{C}^{m \times m}$, A nicht singulär, und $b \in \mathbb{C}^m$

Gesucht: x mit $Ax = b$

Lösung: QR-Faktorisierung

(i) Berechne die QR-Faktorisierung $A = QR$:

$$Ax = QRx = b \Leftrightarrow Rx = Q^*b$$

(ii) $Rx = Q^*b$ wird einfach durch Rücksubstitution gelöst.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Vorwärts- und Rückwärtssubstitution

Gegeben seien die Matrizen $L \in \mathbb{C}^{n \times n}$ und $U \in \mathbb{C}^{n \times n}$ sowie der Vektor $b \in \mathbb{C}^n$. Dabei sei L eine untere Dreiecksmatrix und U eine obere Dreiecksmatrix, wobei die Diagonaleinträge $L_{i,i}$ und $U_{i,i}$, $i \in \{1, \dots, n\}$ ungleich 0 sind.

Gesucht sind Methoden zum Lösen der Gleichungssysteme $Lx = b$ und $Ux = b$.

Man unterscheidet dabei zwischen *Vorwärts-* und *Rückwärtssubstitution*, jenachdem ob die Matrix *untere* oder *obere* Dreiecksgestalt hat.

Vorwärts- und Rückwärtssubstitution

Vorwärtssubstitution:

$$Lx = b$$

L ist eine *untere* Dreiecksmatrix

Algorithmus 4.1 (Vorwärtssubstitution)

for $i = 1 \rightarrow n$ **do**

$$x(i) = b(i)/L_{i,i}$$

$$b(i+1 : n) = b(i+1 : n) - L_{(i+1):n,i} \cdot x(i)$$

end for

Fall $n = 2$:

$$\begin{pmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Dann berechnet man:

$$x_1 = b_1/L_{11}$$

$$x_2 = (b_2 - L_{21}x_1)/L_{22}$$

Allgemein:

$$x_i = \left(b_i - \sum_{j=1}^{i-1} L_{ij}x_j \right) / L_{ii}$$

Vorwärts- und Rückwärtssubstitution

Fall $n = 2$:

$$\begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Rückwärtssubstitution:

$$Ux = b$$

U ist eine *obere* Dreiecksmatrix

Dann berechnet man:

$$x_2 = b_2/U_{22}$$

$$x_1 = (b_1 - U_{12}x_2)/U_{11}$$

Allgemein:

$$x_i = \left(b_i - \sum_{j=i+1}^n U_{ij}x_j \right) / U_{ii}$$

Algorithmus 4.2 (Rückwärtssubstitution)

```
for i = n → 1 do
    x(i) = b(i)/Ui,i
```

```
b(1 : i - 1) = b(1 : i - 1) - U1:(i-1),i · x(i)
end for
```

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

1 Vektoren und Matrizen

2 Vektorräume

3 Normen

4 Lineare Ausgleichsrechnung

5 Einführung in die SVD

6 Die Singulärwertzerlegung in Matrixnotation

7 Unterschiede zwischen SVD und Eigenwertzerlegung

8 Singulärwertzerlegung und Lineare Ausgleichsrechnung

9 Projektoren und Projektionsmatrizen

- Orthogonale Projektionen

- Projektion mit orthonormaler Basis

10 Die QR-Zerlegung

- Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren

- Das modifizierte Gram-Schmidt-Verfahren

- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation

- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix

- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

■ Beispiele für Stabilitäten

- Rückwärtsstabilität der Subtraktion
- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Gram-Schmidt-Orthonormalisierung

Wie konstruiert man eine orthogonale Matrix Q mit den Spalten $\{q_1, \dots, q_n\}$?

Idee: *Gram-Schmidt-Orthogonalisierung*.

Seien a_1, a_2, \dots, a_n wie in Gleichung (4.2) als Spalten der Matrix $A \in \mathbb{C}^{m \times n}$ gegeben. Im j -ten Schritt soll $q_j \in \langle a_1, \dots, a_j \rangle$ berechnet werden, sodass

$$q_j \perp \langle q_1, \dots, q_{j-1} \rangle \quad \text{und} \quad \|q_j\| = 1.$$

Lösung: Setze $q_j := \frac{v_j}{\|v_j\|}$, wobei

$$v_j = a_j - (q_1^* a_j) q_1 - (q_2^* a_j) q_2 - \dots - (q_{j-1}^* a_j) q_{j-1}. \quad (4.6)$$

Stellen wir nun jeweils die i -te die Gleichung in (4.5) nach q_i um, so ergibt sich:

$$\begin{aligned} q_1 &= \frac{a_1}{r_{11}} \\ q_2 &= \frac{a_2 - r_{12}q_1}{r_{22}} \\ q_3 &= \frac{a_3 - r_{13}q_1 - r_{23}q_2}{r_{33}} \\ &\vdots \\ q_n &= \frac{a_n - \sum_{i=1}^{n-1} r_{in}q_i}{r_{nn}}. \end{aligned} \tag{4.7}$$

Die Koeffizienten in (4.6) lauten daher wie folgt:

$$r_{ij} := \begin{cases} q_i^* a_j, & (i \neq j) \\ \left\| a_j - \sum_{l=1}^{j-1} r_{lj} q_l \right\|_2, & (i = j) \end{cases} \tag{4.8}$$

Algorithmus 4.3 (Klassisches Gram-Schmidt-Verfahren (CGS))

```

for  $j = 1 \rightarrow n$  do
     $v_j = a_j$ 
    for  $i = 1 \rightarrow j - 1$  do
         $r_{ij} = q_i^* a_j$ 
         $v_j = v_j - r_{ij} q_i$ 
    end for
     $r_{jj} = \|v_j\|_2$ 
     $q_j = v_j / r_{jj}$ 
end for

```

Beispiel (Der Fall $n = 2$)

Die obige Abbildung veranschaulicht das Verfahren für zwei Vektoren:

a_2 ließe sich wie folgt errechnen:

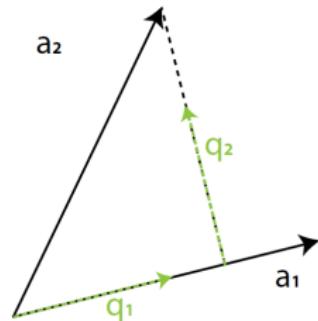
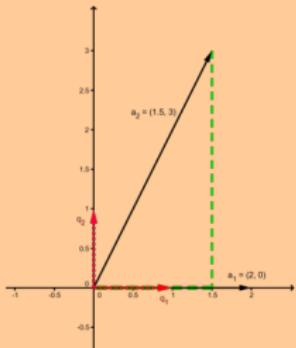


Abbildung 4.13:
Beispiel für das
klassische Gram-
Schmidt-Verfahren

$$\begin{aligned} q_1 &= \frac{a_1}{\|a_1\|} \\ q_2 &= a_2 - q_1 q_1^* a_2 \\ a_2 &= q_1 q_1^* a_2 + q_2 q_2^* a_2 \end{aligned}$$

Beispiel (CGS für zwei Vektoren im \mathbb{R}^2)

Betrachte die beiden Vektoren $a_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ und $a_2 = \begin{pmatrix} 1.5 \\ 3 \end{pmatrix}$:



$$\begin{aligned} q_1 &= \frac{a_1}{\|a_1\|} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ q'_2 &= a_2 - \underbrace{\langle a_2, q_1 \rangle}_{=1.5 \cdot 1 + 3 \cdot 0 = 1.5} q_1 \\ &= \begin{pmatrix} 1.5 \\ 3 \end{pmatrix} - 1.5 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} \end{aligned}$$

Abbildung 4.14:

GS im \mathbb{R}^2

$$q_2 = \frac{q'_2}{\|q'_2\|} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Der Vektor a_2 lässt sich als Linearkombination von q_1 und q_2 darstellen (gestrichelte grüne Linie):

$$a_2 = q_1 \langle q_1, a_2 \rangle + q_2 \langle q_2, a_2 \rangle = 1.5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Gram-Schmidt-Verfahren in Matrixschreibweise

Ist der j -te Spaltenvektor $a_j \notin \langle q_1, \dots, q_{j-1} \rangle$, so berechnet sich q_j nach Gleichung 4.6 gemäß

$$q_j = \frac{a_j - (q_1^* a_j) q_1 - (q_2^* a_j) q_2 - \dots - (q_{j-1}^* a_j) q_{j-1}}{\|a_j - (q_1^* a_j) q_1 - (q_2^* a_j) q_2 - \dots - (q_{j-1}^* a_j) q_{j-1}\|_2}$$

Schreiben wir die Skalarprodukte $q_i^* a_j$ rechts der Vektoren q_i und Klammern wir a_j nach rechts aus so erhalten wir

$$q_j = \frac{\left(I - q_1 q_1^* - q_2 q_2^* - \dots - q_{j-1} q_{j-1}^* \right) a_j}{\left\| \left(I - q_1 q_1^* - q_2 q_2^* - \dots - q_{j-1} q_{j-1}^* \right) a_j \right\|_2}$$

Bezeichnen wir mit \hat{Q}_{j-1} die $(m \times (j-1))$ -Matrix ist, welche die $(j-1)$ ersten Spalten von \hat{Q} enthält, d.h.

$$\hat{Q}_{j-1} = \left(\begin{array}{c|c|c|c|c} q_1 & q_2 & \cdot & \cdot & q_{j-1} \\ \hline \end{array} \right),$$

so erhalten wir

$$\begin{aligned} q_j &= \frac{\left(I - q_1 q_1^* - q_2 q_2^* - \dots - q_{j-1} q_{j-1}^* \right) a_j}{\left\| \left(I - q_1 q_1^* - q_2 q_2^* - \dots - q_{j-1} q_{j-1}^* \right) a_j \right\|_2} \\ &= \frac{\left(I - \hat{Q}_{j-1} \hat{Q}_{j-1}^* \right) a_j}{\left\| \left(I - \hat{Q}_{j-1} \hat{Q}_{j-1}^* \right) a_j \right\|_2}. \end{aligned}$$

Die Matrix $\hat{Q}_{j-1}\hat{Q}_{j-1}^*$ projiziert orthogonal auf $\langle q_1, \dots, q_{j-1} \rangle$ und somit ist

$$P_j = I - \hat{Q}_{j-1}\hat{Q}_{j-1}^* \quad (4.9)$$

die $(m \times m)$ -Matrix vom Rang $m - (j - 1)$, die \mathbb{C}^m orthogonal auf den zu $\langle q_1, \dots, q_{j-1} \rangle$ orthogonalen Unterraum projiziert. Für $j = 1$ ist $P_1 = I$. Mit diesen Projektionsmatrizen lässt sich das klassische Gram-Schmidtverfahren als Folge von Zuweisungen schreiben:

$$q_1 = \frac{P_1 a_1}{\|P_1 a_1\|_2}, \quad q_2 = \frac{P_2 a_2}{\|P_2 a_2\|_2}, \dots, \quad q_n = \frac{P_n a_n}{\|P_n a_n\|_2}. \quad (4.10)$$

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- **Das modifizierte Gram-Schmidt-Verfahren**
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Vom klassischen zum modifizierten GS-Verfahren

Das klassische Gram-Schmidt-Verfahren ist numerisch instabil, d.h. maschinell errechnete Ergebnisse können von den tatsächlichen (exakten) Lösungen teilweise deutlich abweichen. Zur Stabilisierung wurde das Verfahren daher wie folgt modifiziert: anstatt wie der klassische Algorithmus 4.3 eine einzige orthogonale Projektion vom Rang $m - (j - 1)$ zu berechnen

$$q_j = \frac{P_j a_j}{\|P_j a_j\|_2}, \quad (4.11)$$

berechnet das modifizierte Verfahren dieselbe Projektion von P_j als Folge von $j - 1$ Projektionen

$$P_{\perp q_i} = I - q_i q_i^*, i = 1, \dots, j - 1 \quad (4.12)$$

vom Rang $m - 1$ auf die jeweils zu q_1, \dots, q_{j-1} orthogonalen Unterräume.

Ausmultiplizieren und Ausnutzen der Orthogonalität von q_i und $q_j, j \neq i$ liefert

$$\begin{aligned} P_{\perp q_{j-1}} \dots P_{\perp q_2} P_{\perp q_1} &= (I - q_{j-1} q_{j-1}^*) \cdots (I - q_2 q_2^*)(I - q_1 q_1^*) \\ &= I - q_{j-1} q_{j-1}^* - \cdots - q_2 q_2^* - q_1 q_1^* \\ &= P_j. \end{aligned}$$

Das heißt eine zu Gleichung (4.11) äquivalente Aussage ist

$$q_j = \frac{P_{\perp q_{j-1}} \dots P_{\perp q_2} P_{\perp q_1} a_j}{\|P_{\perp q_{j-1}} \dots P_{\perp q_2} P_{\perp q_1} a_j\|}. \quad (4.13)$$

Der modifizierte Algorithmus berechnet also v_j auf die folgende Weise:

$$\begin{aligned}
 v_j^{(1)} &= a_j \\
 v_j^{(2)} &= P_{\perp q_1} v_j^{(1)} = (I - q_1 q_1^*) v_j^{(1)} = v_j^{(1)} - q_1 q_1^* v_j^{(1)} \\
 v_j^{(3)} &= P_{\perp q_2} v_j^{(2)} = v_j^{(2)} - q_2 q_2^* v_j^{(2)} \\
 &\vdots \\
 v_j = v_j^{(j)} &= P_{\perp q_{j-1}} v_{j-1}^{(j-1)} = v_{j-1}^{(j-1)} - q_{j-1} q_{j-1}^* v_{j-1}^{(j-1)}
 \end{aligned} \tag{4.14}$$

Dabei ist $v_j^{(i)}$ der Vektor v_j in der i -ten Iteration des Algorithmus, d.h. nach der $(i-1)$ -ten Projektion zur Berechnung von q_j .

Modifiziertes Gram-Schmidt-Verfahren (MGS)

Algorithmus 4.4 (Modifiziertes Gram-Schmidt-Verfahren (MGS))

```
for i = 1 → n do
     $v_i = a_i$ 
end for
for i = 1 → n do
     $r_{ii} = \|v_i\|_2$ 
     $q_i = v_i / r_{ii}$ 
    for j = i + 1 → n do
         $r_{ij} = q_i^* v_j$ 
         $v_j = v_j - r_{ij} q_i$ 
    end for
end for
```

Vergleich: CGS vs. MGS

Tabelle: Vergleich von klassischem und modifiziertem GS-Verfahren

Klassisches GS-Verfahren	Modifiziertes GS-Verfahren
<ul style="list-style-type: none">■ numerisch instabil■ 1 Projektion vom Rang $m - (j - 1)$■ $v_j = P_j a_j$	<ul style="list-style-type: none">■ numerisch stabiler■ $j - 1$ Projektionen vom Rang $m - 1$■ $v_j = P_{\perp q_{j-1}} \dots P_{\perp q_2} P_{\perp q_1} a_j$

Komplexität der Gram-Schmidt-Algorithmen

Satz 4.4 (Komplexität der Gram-Schmidt-Algorithmen)

Die Algorithmen (4.3) (klassisches Gram-Schmidt-) und (4.4) (modifiziertes Gram-Schmidt-Verfahren) benötigen $O(mn^2)$ Operationen, wobei eine Addition, eine Multiplikation, eine Wurzeloperation, eine Division und eine Subtraktion jeweils eine Operation darstellen.

Beweis: Wir betrachten hier den Fall des modifizierten Gram-Schmidt-Verfahrens (für CGS analog):

```

for  $i = 1 \rightarrow n$  do
    for  $j = i+1 \rightarrow n$  do       $m$  Multiplikationen +  $(m-1)$  Additionen
         $r_{ij} = q_i^* v_j$            $m$  Multiplikationen +  $m$  Subtraktionen
         $v_j = v_j - r_{ij} q_i$ 
    end for                   $\Rightarrow 4m - 1$  Operationen
end for

```

Der Gesamtaufwand beträgt somit also:

$$\begin{aligned}
 \sum_{i=1}^n \sum_{j=i+1}^n (4m-1) &= (4m-1) \sum_{i=1}^n \sum_{j=i+1}^n 1 \\
 &= (4m-1) \sum_{i=1}^n (n-i) = (4m-1) \sum_{i=0}^{n-1} i \\
 &= (4m-1) \frac{n(n-1)}{2} \approx 2mn^2
 \end{aligned}$$

Geometrische Anschauung der Komplexität

In der äußeren Schleife von Algorithmus 4.3 wird die gesamte Matrix bearbeitet, indem ein Vielfaches der ersten Spalte von der anderen Spalte abgezogen wird. Im zweiten Schritt, muss nur noch eine Untermatrix bearbeitet werden, da ein Vielfaches der zweiten Spalte von der Spalte 3, ..., n abgezogen wird usw.

Dies bedeutet, dass für die Gram-Schmidt-Orthogonalisierung $\approx n^2m$ Operationen benötigt werden:
 $\approx n^2m$ Operationen $\hat{=}$ Volumen der Figur.

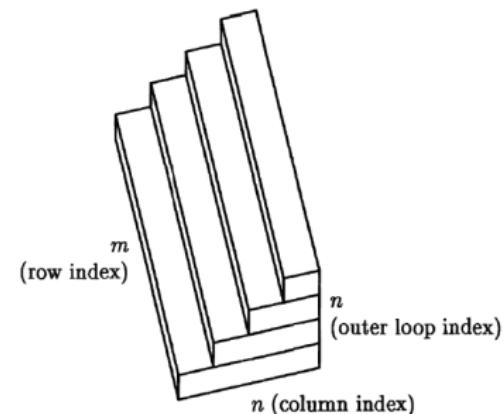


Abbildung 4.15: Geometrische Anschauung

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- **Gram-Schmidt als Dreiecksorthonormalisierung**

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Gram-Schmidt als Dreiecksorthonormalisierung

Jeder äußere Schritt des modifizierten Gram-Schmidt-Algorithmus kann als Multiplikation mit einer quadratischen oberen Dreiecksmatrix von rechts interpretiert werden.

Schritt 1:

$$\left(\begin{array}{c|c|c|c} v_1^{(1)} & v_2^{(1)} & \dots & v_n^{(1)} \end{array} \right) \underbrace{\left(\begin{array}{cccc} \frac{1}{r_{11}} & -\frac{r_{12}}{r_{11}} & -\frac{r_{13}}{r_{11}} & \dots \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{array} \right)}_{R_1} = \left(\begin{array}{c|c|c|c} q_1 & v_2^{(2)} & v_3^{(2)} & \dots \\ & v_2^{(2)} & v_3^{(2)} & \dots \\ & & v_3^{(2)} & \dots \\ & & & v_n^{(2)} \end{array} \right)$$

Schritt 2:

$$R_2 = \begin{pmatrix} 1 & & & \\ \frac{1}{r_{22}} & 1 & & \\ & & 1 & \\ & & & \ddots \end{pmatrix}$$

Schritt 3:

$$R_3 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \frac{1}{r_{33}} & \\ & & & 1 \\ & & & & \ddots \end{pmatrix}$$

Am Ende der Iteration wurden folgende Multiplikationen ausgeführt:

$$A \underbrace{R_1 R_2 \cdots R_n}_{\hat{R}^{-1}} = \hat{Q}$$

Dabei ist \hat{R}^{-1} eine reelle rechte obere Dreiecksmatrix.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Householder-Triangularisierung

Wir betrachten nun eine Alternative, die QR-Zerlegung einer Matrix $A \in \mathbb{C}^{m \times n}$ zu berechnen:

Anstatt A von rechts mit einer oberen Dreiecksmatrix zu multiplizieren, wird A von links mit unitären Matrizen Q_k multipliziert, sodass

$$\underbrace{Q_n \cdots Q_2 Q_1}_{Q^*} A = R \quad (4.15)$$

eine obere Dreiecksmatrix wird. Da $Q = Q_1^* Q_2^* \cdots Q_n^*$ auch unitär ist, erhält man so ebenfalls eine vollständige QR-Zerlegung.

Dieses Verfahren wird *Householder-Triangularisierung* genannt.

Insgesamt haben wir damit zur Berechnung einer QR-Faktorisierung nun zwei Methoden kennen gelernt:

- (i) **Gram-Schmidt:** Dreiecksorthonormalisierung
- (ii) **Householder:** orthogonale Triangularisierung

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Idee der Householder-Transformation

(Alston Scott Householder, 1958)

Schematisch lässt sich die Householder-Transformation wie folgt darstellen:

$$\underbrace{\begin{pmatrix} x & x & x \\ x & x & x \end{pmatrix}}_A \xrightarrow{Q_1} \underbrace{\begin{pmatrix} x & x & x \\ 0 & x & x \end{pmatrix}}_{Q_1 A} \xrightarrow{Q_2} \underbrace{\begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & x \\ 0 & 0 & x \end{pmatrix}}_{Q_2 Q_1 A} \xrightarrow{Q_3} \underbrace{\begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{Q_3 Q_2 Q_1 A}$$

Konkret bedeutet dies: Q_k verändert die Zeilen k, \dots, m . Dabei bildet Q_k eine Linearkombination dieser Zeilen, wobei die Nullen in den Spalten $1, \dots, k - 1$ erhalten bleiben. Nach n Schritten ist die Matrix $Q_n \dots Q_1 A$ eine obere Dreiecksmatrix.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
 - Orthogonale Projektionen
 - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
 - Einschub: Vorwärts- und Rückwärtssubstitution

11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

14 Floating Point Arithmetik

15 Stabilität

- Beispiele für Stabilitäten
 - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts $Q^* b$ (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen $Ax = b$ mittels QR-Faktorisierung

16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

17 Cholesky-Faktorisierung

18 Eigenwerte und Eigenvektoren

19 Schur-Faktorisierung

20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Householder-Spiegelungen

Definition 4.4 (Householder-Transformation)

Die Matrix

$$H := I - \frac{2}{v^* v} v v^* \quad \in \mathbb{C}^{m \times m} \quad \text{mit } v \in \mathbb{C}^m \setminus \{0\}$$

heißt *Householder-Transformation*.

Lemma 4.5 (Householder-Transformation)

Die Householder-Transformation ist eine hermitesch unitäre Matrix mit

$$Hv = -v, \quad Hw = w \quad \forall w \in \{v\}^\perp.$$

Beweis:

- H ist nach Definition hermitesch. ✓
- H ist unitär:

$$\begin{aligned}
 H^*H &= H^2 = I - \frac{4}{v^*v}vv^* + \frac{4}{(v^*v)^2}v(v^*v)v^* \\
 &= I - \frac{4}{v^*v}vv^* + \frac{4}{v^*v}vv^* \\
 &= I.
 \end{aligned}$$

- Sei $w \perp v$. Dann gilt:

$$\begin{aligned}
 Hw &= Iw - \frac{2}{v^*v} \underbrace{vv^*w}_{=0} = w, \\
 Hv &= Iv - \frac{2}{v^*v}vv^*v = v - 2v = -v.
 \end{aligned}$$



Householder-Transformation: Geometrische Anschauung

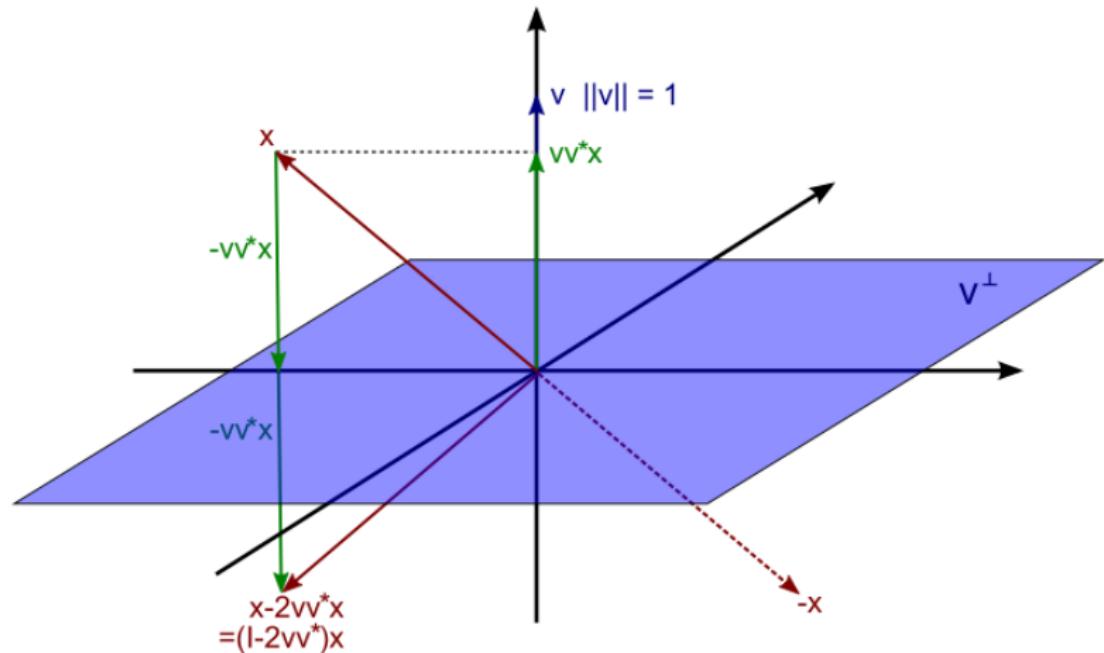


Abbildung 4.16: Geometrische Anschauung der Householder-Transformation

Bemerkung 4.4 ()

Die Householder-Transformation ist unter der $\|\cdot\|_2$ invariant:

$$\|Hx\|_2^2 = (Hx)^* Hx = x^* H^* Hx = x^* x = \|x\|_2^2$$

Vertiefung des Householder-Verfahrens

Mit Hilfe der Householder-Transformation konstruieren wir die unitären Matrizen Q_k wie folgt:

$$Q_k = \begin{pmatrix} I & 0 \\ 0 & H \end{pmatrix},$$

wobei I eine $(k - 1) \times (k - 1)$ Einheitsmatrix und H eine $(m - k + 1) \times (m - k + 1)$ unitäre Matrix ist, welche die Nullen in der k -ten Spalte einführt. H wird als Householder-Matrix (Definition 4.4) gewählt. Sei dazu $x \in \mathbb{C}^{m-k+1}$ der Vektor mit den Einträgen k, \dots, m aus der k -ten Spalte der Matrix $Q_{k-1} Q_{k-2} \cdots Q_1 A$. Dann soll H zu der folgenden Abbildung führen:

$$x = \begin{pmatrix} x_k^{(k)} \\ x_{k+1}^{(k)} \\ \vdots \\ x_m^{(k)} \end{pmatrix} \xrightarrow{H} Hx = \begin{pmatrix} \pm \|x\|_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \pm \|x\|_2 e_1.$$

Daher ist:

$$Hx = x - \frac{2}{v^* v} v(v^* x) = \xi e_1,$$

$$\text{mit } |\xi| = \|x\|_2,$$

d.h. also, dass v ein Vielfaches von $x - \xi e_1$ sein muss.

x wird also durch diese Transformation wie gewünscht gespiegelt (siehe Abbildung rechts).

Damit bei dieser Subtraktion keine Auslöschung auftritt, wählen wir

$$\xi = \begin{cases} -\frac{x_1}{|x_1|} \|x\|_2 & \text{für } x_1 \neq 0 \\ -\|x\|_2 & \text{für } x_1 = 0 \end{cases}$$

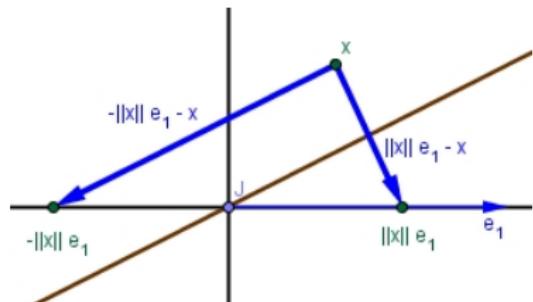


Abbildung 4.17: Geometrische
Anschaugung

Nun wählen wir v so, dass x auf ξe_1 abgebildet wird, d.h.

$$v = \lambda(x - \xi e_1).$$

Setzen wir nun Werte für ξ ein und wählen $\lambda = \frac{1}{\|x\|_2}$, $\|x\|_2 \neq 0$, so ergibt sich:

$$\begin{aligned} v &= \lambda(x - \xi e_1) \\ &= \frac{1}{\|x\|_2} \left(x + \frac{x_1 \|x\|_2}{|x_1|} e_1 \right) \\ &= \frac{1}{|x_1| \|x\|_2} (|x_1| x + x_1 \|x\|_2 e_1) \quad \text{für } x_1 \neq 0 \end{aligned}$$

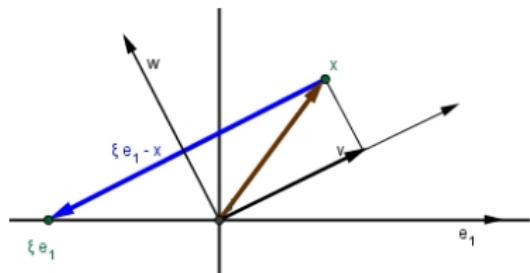


Abbildung 4.18: Geometrische Anschauung

In beiden Fällen gilt:

$$\begin{aligned} v^*x &= \|x\|_2 + |x_1| \\ v^*v &= 2 + 2 \frac{|x_1|}{\|x\|_2}. \end{aligned}$$

Analog ergibt sich

$$\begin{aligned} v &= \lambda(x - \xi e_1) \\ &= \frac{1}{\|x\|_2} (x + \|x\|_2 e_1) \\ &= \frac{x}{\|x\|_2} + e_1 \quad \text{für } x_1 = 0 \end{aligned}$$

Mit diesem v ergibt sich in der Tat für $x_1 \neq 0$

$$\begin{aligned}
 Hx = x - \frac{2}{v^* v} v(v^* x) &= x - 2 \frac{\|x\|_2 + |x_1|}{2 + 2 \frac{|x_1|}{\|x\|_2}} \frac{|x_1| x + x_1 \|x\|_2 e_1}{|x_1| \|x\|_2} \\
 &= x - \|x\|_2 \frac{\|x\|_2 + |x_1|}{\|x\|_2 + |x_1|} \frac{|x_1| x + x_1 \|x\|_2 e_1}{|x_1| \|x\|_2} \\
 &= x - \frac{|x_1| x + x_1 \|x\|_2 e_1}{|x_1| \|x\|_2} \\
 &= x - x - \frac{x_1}{|x_1|} \|x\|_2 e_1 \\
 &= -\frac{x_1}{|x_1|} \|x\|_2 e_1
 \end{aligned}$$

und analog für $x_1 = 0$

$$Hx = x - \|x\|_2 v = -\|x\|_2 e_1.$$

Householder QR-Zerlegung

~~~ Diese Überlegungen führen zu folgendem Algorithmus:

Wir benutzen eine Matlab-ähnliche Notation zur Kennzeichnung von Teilen einer Matrix:  $A_{[i:j,k:l]}$  ist die  $(j - i + 1) \times (l - k + 1)$  Untermatrix mit linker oberer Ecke  $a_{ij}$  und unterer rechter Ecke  $a_{kl}$ . Falls die Untermatrix ein Vektor ist, so schreiben wir  $A_{[i,k:l]}$  oder  $A_{[i:j,k]}$  und wir definieren  $\text{sign}(x_1) = +1$ , falls  $x_1 \geq 0$  und  $\text{sign}(x_1) = -1$  falls  $x_1 < 0$ .

## Algorithmus 4.5 (Householder QR-Zerlegung)

**for**  $i = 1 \rightarrow n$  **do**

$x = A_{[i:m,i]}$  ( $i$ -te Spalte der  $(m - i + 1) \times (n - i + 1)$  Teilmatrix)

$\hat{v}_i = \text{sign}(x_1)\|x\|_2 e + x$  (mit  $e = (1, 0, \dots, 0)^T \in \mathbb{R}^{m-i+1}$ )

$$v_i = \frac{\hat{v}_i}{\|\hat{v}_i\|_2}$$

$$A_{[i:m,i:n]} = A_{[i:m,i:n]} - 2v_i v_i^* A_{[i:m,i:n]}$$

**end for**

# Householder QR-Zerlegung

## Bemerkung 4.5 (Berechnung von Q ist optional)

Der Algorithmus 4.5 berechnet die obere Dreiecksmatrix  $R$  der Faktorisierung  $A = QR$ .  $Q$  selbst wird nicht explizit berechnet. Der Grund dafür ist, dass die Berechnung von  $Q$  zusätzlichen Aufwand bedeuten würde, der meist nicht unbedingt notwendig ist.

## Beispiel (Anwendung der QR-Zerlegung)

Lösung von  $Ax = b$  durch QR-Zerlegung. Dabei sei  $A$  quadratisch und habe vollen Rang.

$$Ax = QRx = b$$

Strategie: Berechne  $y = Q^*b$  und löse dann  $Rx = y$ . Wende dann sukzessive die Operation wie bei der Berechnung von  $R$  an.

Im vorangehenden Beispiel benötigt man die explizite Berechnung von  $Q$ , die im Algorithmus 4.5 jedoch nicht gegeben ist.

$Q$  kann mittels folgender Algorithmen berechnet werden:

### Algorithmus 4.6 (Berechnung von $Q^* b$ )

```

for  $k = 1 \rightarrow n$  do
     $b_{k:m} = b_{k:m} - 2v_k v_k^* b_{k:m}$ 
end for

```

### Algorithmus 4.7 (Berechnung eines Produktes $Qy$ )

```

for  $k = n \rightarrow 1$  do
     $y_{k:m} = y_{k:m} - 2v_k v_k^* y_{k:m}$ 
end for

```

Somit kann man  $Q$  auf zwei Arten berechnen: Berechnung von  $QI$

- 1 mittels  $Q^* e_1, Q^* e_2, \dots, Q^* e_m$  und anschließendem Konjugieren oder
- 2 mittels  $Qe_1, Qe_2, \dots, Qe_m$ .

# Komplexität der Householder-Faktorisierung

## Satz 4.6 (Komplexität der Householder-Faktorisierung)

Die Householder-Faktorisierung benötigt  $mn^2 - \frac{1}{3}n^3 + O(mn)$  Operationen.

Die Komplexität liegt also wie beim Gram-Schmidt-Verfahren in  $O(mn^2)$  und kann daher anschaulich mit dem Volumen einer Pyramide verglichen werden (vgl. oben).

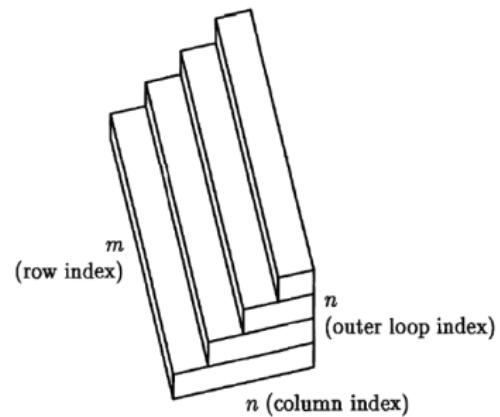


Abbildung 4.19: Geometrische  
Anschauung

## Bemerkung 4.6 (Anwenden der QR-Zerlegung)

- 1 Die QR-Zerlegung ist anwendbar auf das Ausgleichsproblem

$$\|b - Ax\|_2 = \|Q^*(b - Ax)\|_2 = \|Q^*b - Q^*QRx\|_2 = \|c - Rx\|_2$$

mit  $c = Q^*b$ . Zerlegen wir  $R$  und  $c$  in

$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, R_1 \in \mathbb{C}^{n \times n}, c_1 \in \mathbb{C}^n,$$

dann ist nach Pythagoras

$$\|b - Ax\|_2^2 = \|c - Rx\|_2^2 = \|c_1 - R_1x\|_2^2 + \|c_2\|_2^2 \geq \|c_2\|_2^2,$$

mit Gleichheit genau dann, wenn  $c_1 = R_1x \Leftrightarrow x = R_1^{-1}c$ .

- 2 Die QR-Zerlegung (mit der Householder-Transformation) gehört zu den numerisch stabilsten Algorithmen der Numerischen Linearen Algebra. Der Grund liegt darin, dass unitäre Transformationen keinerlei Fehlerverstärkung hervorrufen.

# Kapitel 5:

## Kondition und Stabilität

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Absoluter und relativer Fehler

Betrachte zunächst die Auswertung einer reellwertigen Funktion:

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto f(x); \\ \tilde{f} \text{ Algorithmus zur Berechnung von } f.$$

Dann heißt

$$\|\tilde{f}(x) - f(x)\| \quad \text{absoluter Fehler} \quad \text{und} \quad (5.1)$$

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} \quad \text{relativer Fehler.} \quad (5.2)$$

Aufgrund von *Daten- oder Rundungsfehlern* wird die Funktion  $f$  i.A. nicht an der Stelle  $x$ , sondern an der Stelle  $\tilde{x} = x + \Delta x$  ausgewertet.  
 Wie wirkt sich der Fehler aber auf das Ergebnis aus?

# Auswirkung des Fehlers

Bezeichnen wir mit

$$\Delta y = f(x + \Delta x) - f(x) \quad (5.3)$$

den *fortgepflanzten absoluten Fehler*, so gilt nach dem Mittelwertsatz für  $f \in \mathcal{C}^1$

$$\Delta y = f(x + \Delta x) - f(x) = f'(\xi)\Delta x, \quad \text{wobei } \xi \in [x, x + \Delta x].$$

Ist die Ableitung *Lipschitz-stetig*, dann gilt sogar

$$\Delta y = f'(x)\Delta x + O(|\Delta x|^2).$$

Dabei heißt  $a_\varepsilon \in O(b_\varepsilon)$ , falls  $\exists C > 0$  mit  $|a_\varepsilon| \leq C b_\varepsilon \quad \forall \varepsilon$  aus einer vereinbarten Grundmenge  $E \subset \mathbb{R}^+$ . Das heißt aber  $\frac{a_\varepsilon}{b_\varepsilon} \rightarrow \varepsilon_0$ . Falls  $a_\varepsilon = O(b_\varepsilon)$  und  $b_\varepsilon = O(a_\varepsilon)$ , so schreiben wir auch  $a_\varepsilon \sim b_\varepsilon$ .

# Maß für die Fehlerverstärkung

Vernachlässigen wir den quadratischen Term, so ist

$$K_{abs} = |f'(x)| \quad (5.4)$$

ein Maß für die *Fehlerverstärkung* des *absoluten Eingabefehlers*.

Üblicherweise ist der *relative Fehler* von größerer Bedeutung:

$$\frac{\Delta y}{y} \approx \frac{f'(x)}{f(x)} \Delta x = f'(x) \frac{x}{f(x)} \frac{\Delta x}{x}.$$

# Absolute und relative Konditionszahl

## Definition 5.1 (**Absolute und relative Konditionszahl**)

Die Zahl

$$K_{abs} = |f'(x)| \quad (5.5)$$

heißt *absolute Konditionszahl* des Problems  $x \mapsto f(x)$ . Für  $x \cdot f(x) \neq 0$  ist

$$K_{rel} = \left| \frac{f'(x) \cdot x}{f(x)} \right| \quad (5.6)$$

die *relative Konditionszahl* des Problems.

## Bemerkung 5.1 (**Gut und schlecht konditioniert**)

Die Konditionszahlen beschreiben die Verstärkung des absoluten bzw. relativen Fehlers. Ein Problem heißt *schlecht konditioniert*, falls eine der beiden Konditionszahlen deutlich größer als 1 ist. Ansonsten heißt es *gut konditioniert*.

Beispiel ( $f(x) = x + a$ )

$$f(x) = x + a, \quad f'(x) = 1 \quad \Rightarrow K_{abs} = |f'(x)| = 1$$

Für die relative Konditionszahl ergibt sich

$$K_{rel} = \left| \frac{f'(x) \cdot x}{f(x)} \right| = \left| \frac{x}{x + a} \right|.$$

$K_{rel}$  wird groß, falls  $|x + a| \ll |x|$ , also wenn  $x \approx -a$ . Diesen schlecht konditionierten Fall bezeichnet man als Auslöschung: Für  $a = -1$ ,  $x = 1,000001$  und  $\Delta x = 0,001$  ist

$$f(x) = x + a = 0,000001 \quad \text{und} \quad f(x + \Delta x) + a = 0,001001.$$

Der absolute Fehler ist also 0,001, d.h. gleich dem Eingangsfehler. Der relative Fehler ist

$$\frac{\Delta y}{y} = \frac{0,001}{0,000001} \approx 10^6 \frac{0,001}{1,000001},$$

Beispiel ( $f(x) = ax$ )

Betrachte die Funktion  $f(x) = ax$ . Die absolute und relative Konditionszahl lautet

$$\begin{aligned} K_{abs} &= |f'(x)| = |a|, \\ K_{rel} &= \left| \frac{f'(x) \cdot x}{ax} \right| = 1. \end{aligned}$$

In diesem Fall ist die absolute Konditionszahl schlecht, falls  $|a| \gg 1$ . Der relative Fehler bleibt fest.

Beispiel ( $f(x) = x^2 - 2x + 1 = a_2x^2 + a_1x + a_0$ )

Nullstellenbestimmung:  $f(x) = x^2 - 2x + 1 = a_2x^2 + a_1x + a_0$ . Die Nullstellen sind abhängig von den Koeffizienten  $a_i$ .

Betrachte die Abhängigkeit von  $a_0$ :

$f(x) = x^2 - 2x + 0,9999 = (x - 0,99)(x - 1,01)$ . D.h. die Wurzeln des Polynoms ändern sich in der Größenordnung der Wurzel der Koeffizienten.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- **Kondition einer Matrix**
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Kondition einer Matrix

Betrachte das Problem  $x \mapsto F(x)$ , wenn  $x$  und  $F(x) \in \mathbb{C}^n$ . Wir beschränken uns auf den Spezialfall, in dem ein LGS der Form  $Az = b$  zu lösen ist, wobei  $A \in \mathbb{C}^{n \times n}$  invertierbar ist. In diesem Fall ist  $F(b) = A^{-1}b$ . Bei Eingangsfehler  $\Delta b$  ergibt sich

$$z = A^{-1}b, \quad z + \Delta z = A^{-1}(b + \Delta b) = A^{-1}b + A^{-1}\Delta b,$$

d.h. die berechnete Lösung  $z + \Delta z$  enthält den fortgepflanzten Fehler  $\Delta z = A^{-1}\Delta b$ .

# Kondition einer Matrix

Sind  $\|\cdot\|_M$  und  $\|\cdot\|$  ein verträgliches Matrix-/Vektornormpaar, d.h.  
 $\|Ax\| \leq \|x\| \|A\|_M \quad \forall A \in \mathbb{C}^{m \times n}$  and  $\forall x \in \mathbb{C}^n$ , dann folgt:

$$\frac{\|\Delta z\|}{\|z\|} = \frac{\|A^{-1}\Delta b\|}{\|z\|} \leq \|A^{-1}\|_M \frac{\|\Delta b\|}{\|b\|} \frac{\|Az\|}{\|z\|} \leq \|A^{-1}\|_M \|A\|_M \frac{\|\Delta b\|}{\|b\|} \quad (5.7)$$

Das heißt, ein relativer Eingangsfehler in der Größenordnung  $\frac{\|\Delta b\|}{\|b\|}$  führt zu  
 einem relativen Fehler der Größenordnung  $\frac{\|\Delta z\|}{\|z\|} \leq \|A^{-1}\|_M \|A\|_M \frac{\|\Delta b\|}{\|b\|}$  in der  
 Lösung.

# Kondition einer Matrix

## Definition 5.2 (Kondition einer Matrix)

Der Faktor

$$K_M(A) := \|A^{-1}\|_M \cdot \|A\|_M \quad (5.8)$$

wird als *Kondition der Matrix A* bezüglich der Norm  $\|\cdot\|_M$  bezeichnet.

## Bemerkung 5.2 ()

Falls  $\|\cdot\|_M$  durch eine Vektornorm induziert wird, kann man Beispiele für  $b$  und  $Ab$  konstruieren, für die in (5.7) Gleichheit herrscht.

# Kondition der Berechnung $b = Ax$

## Satz 5.1 (Kondition der Berechnung $b = Ax$ )

Das Problem der Berechnung  $b = Ax$ , gegeben  $x$  und eine nicht-singuläre Matrix  $A \in \mathbb{C}^{m \times m}$ , hat ebenfalls die Kondition

$$K_M(A) := \|A\|_M \|A^{-1}\|_M \quad (5.9)$$

bezüglich Störungen von  $x$ .

*Beweis:*

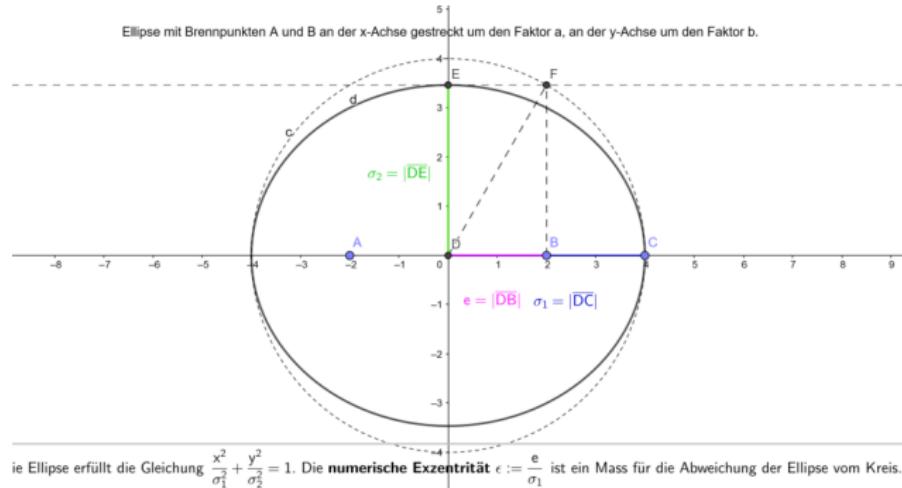
$$\begin{aligned} \frac{\|\Delta b\|}{\|b\|} &= \frac{\|A\Delta x\|}{\|b\|} \\ &\leq \frac{\|A\|_M \|\Delta x\|}{\|b\|} = \|A\|_M \frac{\|\Delta x\|}{\|x\|} \frac{\|A^{-1}b\|}{\|b\|} \\ &\leq \|A\|_M \|A^{-1}\|_M \frac{\|\Delta x\|}{\|x\|} \end{aligned}$$

# Kondition der Berechnung $b = Ax$

## Bemerkung 5.3 (Kondition und Singulärwerte in nicht singulären Fall)

Falls  $\|\cdot\|_M = \|\cdot\|_2$  und  $A$  nicht singulär ist, gilt wegen

$\|A\|_2 = \|U\Sigma V^*\|_2 = \|\Sigma\|_2$  insbesondere  $\|A\|_2 = \sigma_1$  und  $\|A^{-1}\|_2 = \frac{1}{\sigma_n}$  und daher  $K(A) = \frac{\sigma_1}{\sigma_n}$ .  $\epsilon = \frac{1}{K(A)} \sqrt{K(A)^2 - 1}$  kann als Exzentrizität der Ellipse, die das Bild der Einheitskugel unter  $A$  ist, interpretiert werden.



# Kondition der Berechnung $b = Ax$

## Bemerkung 5.4 (Kondition und Singulärwerte im singulären Fall)

Für eine Matrix  $A \in \mathbb{C}^{m \times n}$ ,  $m \geq n$ , mit vollem Rang wird die Kondition mit Hilfe der *Pseudoinversen* definiert:

$$K(A) = \|A\| \|A^+\|.$$

Für die  $\|\cdot\|_2$  gilt in diesem Fall

$$K(A) = \frac{\sigma_1}{\sigma_n}.$$

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- **Kondition eines Gleichungssystems**

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Kondition eines Gleichungssystems

Wie ändert sich  $b = A^{-1}x$ , wenn  $A$  um  $\Delta A$  variiert?

Sei dazu  $A$  nicht singulär. Betrachte dazu die Gleichung

$$\begin{aligned}
 & (A + \Delta A)(x + \Delta x) = b \\
 \Leftrightarrow & \underbrace{Ax}_b + \Delta Ax + A\Delta x + \underbrace{\Delta A\Delta x}_{\approx 0} = b \\
 & \text{d.h. } \Delta Ax \approx -A\Delta x \\
 \Leftrightarrow & \|\Delta x\| \leq \|A^{-1}\| \|\Delta A\| \|x\| \\
 \Leftrightarrow & \frac{\|\Delta x\|}{\|x\|} \leq \|A^{-1}\| \|A\| \frac{\|\Delta A\|}{\|A\|}.
 \end{aligned}$$

( $\Delta A$  benötigt ein anderes  $\Delta x$ )

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Motivation

Reelle Zahlen können bekanntlicher Weise mit einem Computer nicht exakt dargestellt werden. Dies hat zur Folge, dass sich Rechenfehler wegen der Ungenauigkeit der Darstellung anhäufen können und zu extremen Abweichungen vom tatsächlichen Ergebnis führen.

## Definition 5.3 (**Floating Point Zahlen**)

Die diskrete endliche Menge aller in einem Rechner darstellbaren Zahlen bezeichnen wir als *Floating Point Zahlen*:

$$F = \left\{ x \in \mathbb{R} \mid \begin{array}{l} \exists \beta \geq 2 \text{ und } e \geq 1 \text{ mit } m \in \mathbb{N}, 1 \leq m \leq \beta^t : \\ x = 0 \text{ oder } x = \pm \left( \frac{m}{\beta^t} \right) \beta^e \end{array} \right\},$$

wobei  $t$  die *Präzision* ist und in IEEE die Werte 24 bzw. 53 annimmt.  $\beta$  heißt *Basis* und nimmt normalerweise den Wert 2 an.

# Floating Point Zahlen

## Bemerkung 5.5 ()

- 1 Wir können  $m$  auf den Bereich  $\beta^{t-1} \leq m \leq \beta^t - 1$  beschränken und damit die Wahl von  $m$  eindeutig machen.
- 2 Der Term  $\left(\pm \frac{m}{\beta^t}\right)$  heißt *Mantisse*,  $e$  heißt *Exponent*.
- 3 Relative Genauigkeit bei *single precision*:  $2^{-24} \approx 5,96 \times 10^{-8}$   
Relative Genauigkeit bei *double precision*:  $2^{-53} \approx 1,11 \times 10^{-16}$

# Modellierung der Rechenarithmetik

**Modellierung der Rechenarithmetik** Zur Modellierung der Rechenarithmetik machen wir folgende beiden Modellannahmen:

- 1  $x \bullet y := \square(x \circ y)$ , wobei  $x, y \in F$ ,  $\circ$  die mathematische Grundoperation,  $\square x$  die Rundung von  $x \in \mathbb{R}$  zur nächstgelegenen Maschinenzahl und  $\bullet$  die Realisierung dieser Grundoperation auf dem Rechner ist.
- 2 Die *Rechneroperation* soll den tatsächlichen Wert innerhalb einer *maximalen relativen Genauigkeit* bestimmen, d.h.  
 $\forall x \in \mathbb{R}, \exists \varepsilon : |\varepsilon| \leq \varepsilon_M$  mit

$$\square x = x(1 + \varepsilon). \quad (5.10)$$

Dabei ist  $\varepsilon_M = \inf\{x > 0 : 1 \oplus x \neq 1\}$ .

# Underflow und Overflow

## Bemerkung 5.6 (Machine Epsilon)

- 1 Der genaue Wert von  $\varepsilon_M$  ist rechnerabhängig. In der Regel ist  $\varepsilon_M = 2^{-d}$  für ein  $d > 0$ .
- 2 Die Gleichung (5.10) wird falsch, wenn eine von Null verschiedene Zahl  $x$  auf Null gerundet wird. Diese Situation nennt man *Underflow*. Ähnlich spricht man von *Overflow*, wenn das Rechenergebnis größer als die darstellbaren Zahlen wird (also über den gegebenen Zahlenbereich hinausreicht).  
Sofern weder Overflow noch Underflow auftreten, ist der relative Rundungsfehler nach diesem Modell beschränkt durch

$$\frac{|\square x - x|}{|x|} \leq \varepsilon_M. \quad (5.11)$$

# Realisierung von Elementaroperationen

Unter beiden Modellannahmen 1. und 2. sind alle Elementaroperationen auf dem Rechner in der folgenden Weise realisiert:

$$\forall x, y \in \mathbb{R}, \exists \varepsilon : |\varepsilon| \leq \varepsilon_M, \text{ s.d.} \\ x \bullet y = (x \circ y)(1 + \varepsilon) \quad (5.12)$$

d.h. jede Operation der Floating Point Arithmetik ist *genau* bis auf einen *relativen Fehler* in der Größenordnung der Maschinengenauigkeit.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

### ■ Beispiele für Stabilitäten

- Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Gute Realisierungsalgorithmen

## Notation:

Sei  $f : \mathbb{R} \rightarrow F$  eine Funktion, die  $f(x)$  unter Verwendung von Fließkommazahlen berechnet, d.h.  $f$  beschreibt einen Realisierungsalgorithmus.

## Genauigkeit:

Betrachten wir die Implementierung von  $f$  zur Lösung eines Problems  $x \mapsto f(x) = y$ ,  $x \in D(f) \subset \mathbb{R}$ . Seien dazu  $x$  und  $y$  von Null verschieden.  
 $f$  ist ein *guter* Realisierungsalgorithmus, wenn

$$\left| \frac{f(x) - f(x)}{f(x)} \right| \leq c_V K_{rel} \varepsilon_M \quad (5.13)$$

mit einem mäßig großen  $c_V > 0$ , das von  $x$  unabhängig ist.

# Vorwärtsstabilität

## Definition 5.4 (Vorwärtsstabilität)

Diese Form der Stabilitätsanalyse heißt *Vorwärtsanalyse* und  $f$  heißt *vorwärts stabil*, wenn Gleichung (5.13) erfüllt ist, d.h.

$$\left| \frac{f(x) - f(x)}{f(x)} \right| \leq c_V K_{rel} \varepsilon_M$$

## Bemerkung 5.7 (Problem der Vorwärtsanalyse)

Die Vorwärtsanalyse ist häufig sehr schwierig, da die Abhängigkeit von der Konditionszahl subtil ist. Die Lösung des Problems ist die Rückwärtsanalyse:

$$\frac{\Delta y}{y} \approx \frac{f'(x)}{f(x)} \Delta x = \underbrace{f'(x) \frac{x}{f(x)}}_{K_{rel}} \frac{\Delta x}{x} \quad (5.14)$$

# Rückwärtsstabilität

## Definition 5.5 (Rückwärtsanalyse)

Bei der *Rückwärtsanalyse* interpretiert man die berechnete Näherung als exakte Lösung mit gestörten Eingangsdaten, d.h.  $f(x) = f(x + \Delta x)$  und untersucht  $|\Delta x|$ .

Gibt es mehrere Urbilder  $x + \Delta x$ , so wählt man das mit kleinster Störung  $\Delta x$ . Gilt dann

$$\left| \frac{\Delta x}{x} \right| \leq c_R \varepsilon_M, \quad (5.15)$$

mit einem mäßig großen  $c_R > 0$ , das von  $x$  unabhängig ist, so heißt  $f$  *rückwärts stabil*.

# Rückwärtsstabilität

## Bemerkung 5.8 (Rückwärtsstabilität)

Für einen rückwärts stabilen Algorithmus gilt nach Gleichung (5.14) mit  $\tilde{x} = x + \Delta x$ :

$$\left| \frac{f(x) - f(\tilde{x})}{f(x)} \right| = \left| \frac{f(\tilde{x}) - f(x)}{f(x)} \right| \leq K_{rel} \left| \frac{\tilde{x} - x}{x} \right| \leq c_R K_{rel} \varepsilon_M.$$

Bis auf den Einfluss des Approximationsfehlers in (5.14) ist damit jeder rückwärts stabile Algorithmus auch vorwärts stabil. Die Umkehrung gilt i.A. nicht.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten

- Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Rückwärtsstabilität der Subtraktion

## Theorem 5.2 (Rückwärtsstabilität der Subtraktion)

Die Subtraktion ist rückwärts stabil.

**Beweis:** Zur Erinnerung:

$$\forall x \in \mathbb{R}, \exists \varepsilon, |\varepsilon| \leq \varepsilon_M, \text{ s.d. } \square(x) = x(1 + \varepsilon) \quad (\text{siehe (5.10)})$$

$$\forall x, y \in \mathbb{R}, \exists \varepsilon, |\varepsilon| \leq \varepsilon_M, \text{ s.d. } x \bullet y = (x \circ y)(1 + \varepsilon) \quad (\text{siehe (5.12)})$$

Beweis der Rückwärtsstabilität:

$$\begin{aligned} (\square x \square \square y) &= (x(1 + \varepsilon_1) - y(1 + \varepsilon_2))(1 + \varepsilon_3) \\ &= x(1 + \varepsilon_1 + \varepsilon_3 + \varepsilon_1 \varepsilon_3) - y(1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_2 \varepsilon_3) \\ &= x(1 + \varepsilon_4) - y(1 + \varepsilon_5) \end{aligned}$$

wobei  $|\varepsilon_4|, |\varepsilon_5| \leq 2\varepsilon_M + O(\varepsilon_M^2)$ , d.h.  $(\square x \square \square y) = \tilde{x} - \tilde{y}$  mit  
 $\frac{|\tilde{x} - x|}{|x|} = O(\varepsilon_M)$ ,  $\frac{|\tilde{y} - y|}{|y|} = O(\varepsilon_M)$ . Das heißt:  $(\square x \square \square y)$  ist rückwärts stabil.

$$\begin{pmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{22} & \dots & r_{2m} \\ \ddots & \vdots & & r_{mm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

### Theorem 5.3 (Rückwärtsstabilität der Rückwärtssubstitution)

Die Rückwärtssubstitution ist rückwärts stabil, d.h. für die berechnete Lösung  $\tilde{x}$  gilt:

$$(R + \delta R)\tilde{x} = b \quad (5.16)$$

für eine obere Dreiecksmatrix  $\delta R \in \mathbb{C}^{m \times m}$  mit  $\frac{\|\delta R\|}{\|R\|} = O(\varepsilon_M)$ .

Insbesondere gilt  $\forall i, j$ :

$$\frac{|\delta r_{ij}|}{|r_{ij}|} \leq m\varepsilon_M + O(\varepsilon_M^2). \quad (5.17)$$

Aus der komponentenweise Rückwärtsstabilität (5.17) folgt die normweise Stabilität (5.16).

# Beweis: Rückwärtsstabilität der Rückwärtssubstitution

**Beweis:** Wir beweisen die Aussage von Theorem 5.3 für  $m = 1, 2, 3$ .  
 $m = 1$ :

$$\tilde{x}_1 = b_1 \boxdot r_{11} = \frac{b_1}{r_{11}}(1 + \varepsilon_1) \quad \text{für } |\varepsilon_1| \leq \varepsilon_M.$$

Setze  $\varepsilon'_1 := \frac{-\varepsilon_1}{1+\varepsilon_1}$ , dann folgt

$$\tilde{x}_1 = \frac{b_1}{r_{11}(1 + \varepsilon'_1)} \quad \text{für } |\varepsilon'_1| \leq \varepsilon_M + O(\varepsilon_M^2),$$

wobei die Abschätzung für  $|\varepsilon'_1|$  aus der Taylor-Entwicklung folgt.  
 Damit gilt für  $(R + \delta R)\tilde{x} = b$ :

$$(r_{11} + \underbrace{r_{11}\varepsilon'_1}_{\delta r_{11}})\tilde{x}_1 = b_1 \quad \text{für } \frac{|\delta r_{11}|}{|r_{11}|} \leq \varepsilon_M + O(\varepsilon_M^2).$$

$m = 2$ :

$$\tilde{x}_2 = b_2 \boxminus r_{22} = \frac{b_2}{r_{22}(1 + \varepsilon_1)} \quad \text{für } |\varepsilon_1| \leq \varepsilon_M + O(\varepsilon_M^2).$$

Im zweiten Schritt haben wir mehrere Floating Point Operationen zu berücksichtigen:

$$\begin{aligned}\tilde{x}_1 &= (b_1 \boxminus (\tilde{x}_2 \boxtimes r_{12})) \boxplus r_{11} \\ &= (b_1 \boxminus (b_1 - \tilde{x}_2 \cdot r_{12}(1 + \varepsilon_2))) \boxplus r_{11}, \quad |\varepsilon_2| \leq \varepsilon_M \\ &= \frac{(b_1 - \tilde{x}_2 \cdot r_{12}(1 + \varepsilon_2))(1 + \varepsilon_3)}{r_{11}} \cdot (1 + \varepsilon_4), \quad |\varepsilon_3|, |\varepsilon_4| \leq \varepsilon_M\end{aligned}$$

$$\begin{aligned}\varepsilon'_{3,4} &= \frac{-\varepsilon_{3,4}}{1 + \varepsilon_{3,4}} \\ &= \frac{b_1 - \tilde{x}_2 r_{12}(1 + \varepsilon_2)}{r_{11}(1 + \varepsilon'_3)(1 + \varepsilon'_4)} \quad |\varepsilon'_3|, |\varepsilon'_4| \leq \varepsilon_M + O(\varepsilon_M^2) \\ &= \frac{b_1 - \tilde{x}_2 r_{12}(1 + \varepsilon_2)}{r_{11}(1 + 2\varepsilon_5)} \quad |\varepsilon_5| \leq \varepsilon_M + O(\varepsilon_M^2),\end{aligned}$$

Damit können wir die Rückwärtsanalyse nun als  $(R + \delta R)\tilde{x} = b$  schreiben mit

$$\frac{|\delta R|}{|R|} = \begin{pmatrix} \frac{|\delta r_{11}|}{|r_{11}|} & \frac{|\delta r_{12}|}{|r_{12}|} \\ 0 & \frac{|\delta r_{22}|}{|r_{22}|} \end{pmatrix} = \begin{pmatrix} 2|\varepsilon_5| & |\varepsilon_2| \\ 0 & |\varepsilon_1| \end{pmatrix} \leq \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \varepsilon_M + O(\varepsilon_M^2),$$

wobei Betrag und Division in  $|\delta R|/|R|$  komponentenweise zu verstehen sind.

$m = 3$ :

Die ersten zwei Schritte erfolgen wie oben. Dabei ist  $\varepsilon_5 = \varepsilon_3$ .

$$\begin{aligned}\tilde{x}_1 &= [(b_1 \boxminus (\tilde{x}_2 \boxtimes r_{12})) \boxminus (\tilde{x}_3 \boxtimes r_{13})] \boxdot r_{11} \\ &= \frac{[(b_1 - \tilde{x}_2 r_{12}(1 + \varepsilon_4))(1 + \varepsilon_6) - \tilde{x}_3 r_{13}(1 + \varepsilon_5)](1 + \varepsilon_7)}{r_{11}(1 + \varepsilon'_8)} \\ &= \frac{[(b_1 - \tilde{x}_2 r_{12}(1 + \varepsilon_4))(1 + \varepsilon_6) - \tilde{x}_3 r_{13}(1 + \varepsilon_5)]}{r_{11}(1 + \varepsilon'_8)(1 + \varepsilon'_7)} \\ &= \frac{(b_1 - \tilde{x}_2 r_{12}(1 + \varepsilon_4)) - \tilde{x}_3 r_{13}(1 + \varepsilon_5)(1 + \varepsilon'_6)}{r_{11}(1 + \varepsilon'_6)(1 + \varepsilon'_8)(1 + \varepsilon'_7)}.\end{aligned}$$

Somit gilt für  $(R + \delta R)\tilde{x} = b$ :

$$\frac{|\delta R|}{|R|} = \begin{pmatrix} \frac{|\delta r_{11}|}{|r_{11}|} & \frac{|\delta r_{12}|}{|r_{12}|} & \frac{|\delta r_{13}|}{|r_{13}|} \\ 0 & \frac{|\delta r_{22}|}{|r_{22}|} & \frac{|\delta r_{23}|}{|r_{23}|} \\ 0 & 0 & \frac{|\delta r_{33}|}{|r_{33}|} \end{pmatrix} \leq \begin{pmatrix} 3 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \varepsilon_M + O(\varepsilon_M^2)$$

$m \geq 3$ :

Die Analyse für größere  $m$  funktioniert analog.



# Rückwärtsstabilität des Produkts $Q^*b$ (ohne Beweis)

Bei der Berechnung von  $Q^*b$  treten Fehler auf. Anstatt  $Q^*b$  genau zu berechnen, wird ein  $\tilde{y}$  berechnet.

## Theorem 5.4 (Rückwärtsstabilität des Produkts $Q^*b$ )

Die Operation  $Q^*b$  ist rückwärts stabil, d.h.  $\exists \delta Q, \|\delta Q\| = O(\varepsilon_M)$  mit

$$(Q + \delta Q)\tilde{y} = b. \quad (5.18)$$

Das heißt, das Resultat der Berechnung von Householder-Reflektoren in Floating Point Arithmetik entspricht einer Multiplikation mit einer Matrix

$$(Q + \delta Q)^{-1}.$$

# Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)

## Theorem 5.5 (Rückwärtsstabilität der Householder-Triangularisierung)

Sei  $A = QR$ ,  $A \in \mathbb{C}^{m \times n}$ , eine mittels Householder-Transformation berechnete Faktorisierung von  $A$ . Sei weiterhin  $\tilde{R}$  die mit Floating-Point-Genauigkeit berechnete obere Dreiecksmatrix und  $\tilde{Q} = \tilde{Q}_1 \tilde{Q}_2 \dots \tilde{Q}_n$ , wobei  $\tilde{Q}_k$  der exakt unitäre Reflektor des entsprechenden Vektors  $\tilde{v}_k$  in Floating-Point-Arithmetik ist.

Dann gilt

$$\tilde{Q}\tilde{R} = A + \delta A, \quad \frac{\|\delta A\|}{\|A\|} = O(\varepsilon_M) \quad (5.19)$$

für  $\delta A \in \mathbb{C}^{m \times n}$ .

# Algorithmus zur Lösung von $Ax = b$

## Algorithmus 5.1 (Lösung von $Ax = b$ mittels QR-Faktorisierung)

$$\begin{aligned} QR &= A \\ y &= Q^* b \\ x &= R^{-1} y \end{aligned}$$

## Theorem 5.6 (Rückwärtsstabilität der Lösung von $Ax = b$ mittels QR-Faktorisierung)

Der Algorithmus 5.1 ist rückwärts stabil, d.h.

$$(A + \delta A)\tilde{x} = b, \quad \frac{\|\delta A\|}{\|A\|} = O(\varepsilon_M) \quad (5.20)$$

für ein  $\delta A \in \mathbb{C}^{n \times n}$

**Beweis:** Die Gleichungen (5.16) und (5.18) implizieren:

$$\begin{aligned}
 b = (\tilde{Q} + \delta Q)(\tilde{R} + \delta R)\tilde{x} &= [\underbrace{\tilde{Q}\tilde{R}}_{=A+\delta A} + (\delta Q)\tilde{R} + \tilde{Q}\delta R + (\delta Q)(\delta R)]\tilde{x} \\
 &= [A + \underbrace{\delta A + \delta Q\tilde{R} + \tilde{Q}\delta R + \delta Q\delta R}_{\Delta A}]\tilde{x} \\
 &= [A + \Delta A]\tilde{x}.
 \end{aligned}$$

Z.z.: Jeder Term von  $\Delta A$  ist relativ zu  $A$  klein.

Nach Theorem 5.5 gilt

$$\tilde{Q}\tilde{R} = A + \delta A, \quad \tilde{Q} \text{ unitär}, \quad \frac{\|\delta A\|}{\|A\|} = O(\varepsilon_M)$$

$$\frac{\|\tilde{R}\|}{\|A\|} \leq \|\tilde{Q}^*\| \frac{\|A + \delta A\|}{\|A\|} = O(1), \quad \text{falls } \varepsilon_M \rightarrow 0.$$

Daraus erhalten wir

$$\frac{\|\delta Q\tilde{R}\|}{\|A\|} \leq \|\delta Q\| \frac{\|\tilde{R}\|}{\|A\|} = O(\varepsilon_M)$$

und

$$\frac{\|\tilde{Q}\delta R\|}{\|A\|} \leq \|\tilde{Q}\| \frac{\|\delta R\|}{\|\tilde{R}\|} \frac{\|\tilde{R}\|}{\|A\|} = O(\varepsilon_M).$$

Schließlich gilt noch

$$\frac{\|(\delta Q)(\delta R)\|}{\|A\|} \leq \|\delta Q\| \frac{\|\delta R\|}{\|A\|} = O(\varepsilon_M^2).$$

## Theorem 5.7 (Genauigkeit der Lösung)

Die mit Algorithmus 5.1 berechnete Lösung erfüllt

$$\frac{\|\tilde{x} - x\|}{\|x\|} = O(K(A) \cdot \varepsilon_M).$$

# Kapitel 6:

## LU-Zerlegung

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# LU-Faktorisierung: Idee

Sei  $A \in \mathbb{C}^{m \times m}$ . Im Folgenden wollen wir die Matrix  $A$  in eine linke untere und eine rechte obere Dreiecksmatrix zerlegen.

Idee: Bringe  $A$  auf Dreiecksgestalt, indem, ähnlich wie bei der QR-Zerlegung, Nullen unterhalb der Diagonalen in den Spalten erzeugt werden. Dies lässt sich mit einer Folge von *unteren Dreiecksmatrizen*  $L_k$  durchführen:

$$\underbrace{L_{m-1} \dots L_2 L_1}_{L^{-1}} A = U,$$

$$A = L_1^{-1} L_2^{-1} \cdots L_{m-1}^{-1} U = LU$$

$U$  obere Dreiecksmatrix (**upper**)

$L$  untere Dreiecksmatrix (**lower**)

Im Deutschen wird anstatt der hier gewählten englischen Bezeichnung **LU** häufig die Bezeichnung **LR** gewählt, wobei **L** für die untere **linke** Dreiecksmatrix und **R** für die obere **rechte** Dreiecksmatrix steht.

# LU-Faktorisierung: Schematische Dartstellung

Schematische Darstellung:

$$\begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \xrightarrow{L_1} \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{pmatrix} \xrightarrow{L_2} \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{pmatrix} \xrightarrow{L_3} \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{pmatrix}$$

Methoden:

Wir kennen die folgenden Methoden, eine Matrix  $A$  in eine Dreiecksmatrix umzuwandeln:

**Gram-Schmidt:**  $A = QR$  Orthogonalisierung

**Householder:**  $A = QR$  orthogonale Triangulierung

**Gauß-Elimination:**  $A = LU$  Triangularisierung (Gauß)

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

### ■ LU-Faktorisierung (ohne Pivotisierung)

- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

Im  $k$ -ten Schritt des Eliminationsprozesses muss  $L_k$  so gewählt werden, dass alle Zeilen von  $A$  mit Index  $1, \dots, k$  erhalten bleiben und der  $k$ -te Spaltenvektor der Matrix  $A$  wie folgt verändert wird:

$$L_k a_k = L_k \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{k,k} \\ a_{k+1,k} \\ \vdots \\ a_{m,k} \end{pmatrix} = \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{k,k} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Dazu erhalten wir die ersten  $m$  Zeilen und ziehen für  $k < j \leq m$  von der  $j$ -ten Zeile jeweils das  $l_{j,k}$ -fache der  $k$ -ten Zeile ab, wobei

$$l_{j,k} = \frac{a_{j,k}}{a_{k,k}}, \quad k < j \leq m$$

Die Matrix  $L_k$  hat damit folgende Form:

$$L_k = \begin{pmatrix} 1 & & & & \\ \vdots & \ddots & & & \\ & & 1 & & \\ & -l_{k+1,k} & & 1 & \\ & \vdots & & & \ddots \\ & -l_{m,k} & & & & 1 \end{pmatrix}$$

Durch diese Konstruktion erhalten wir in der  $k$ -ten Spalte Nullen unterhalb der Diagonalen:

Beispiel:

$$L_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -l_{3,2} & 1 & 0 & 0 \\ 0 & -l_{4,2} & 0 & 1 & 0 \\ 0 & -l_{5,2} & 0 & 0 & 1 \end{pmatrix}$$

Definieren wir den Vektor  $l_k$  durch

$$l_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ l_{k+1,k} \\ \vdots \\ l_{m,k} \end{pmatrix}, \text{ wobei } l_{j,k} = \frac{a_{j,k}}{a_{k,k}}, k < j \leq m, l_{j,k} = 0, j \leq k,$$

so können wir die Matrix  $L_k$  schreiben als

$$L_k = I - l_k e_k^*. \quad (6.1)$$

Bei der Berechnung der Matrix  $L$  gilt unter anderem Folgendes:

- Da  $e_k^* I_k = 0$ , gilt:

$$(I - I_k e_k^*)(I + I_k e_k^*) = I - I_k e_k^* I_k e_k^* = I,$$

d.h.:  $(I - I_k e_k^*)^{-1} = (I + I_k e_k^*).$

- 

$$\begin{aligned} L_k^{-1} L_{k+1}^{-1} &= (I + I_k e_k^*)(I + I_{k+1} e_{k+1}^*) \\ &= I + I_k e_k^* + I_{k+1} e_{k+1}^* + I_k \underbrace{e_k^* I_{k+1}}_{=0} e_{k+1}^* \\ &= I + I_k e_k^* + I_{k+1} e_{k+1}^* \end{aligned}$$

Also berechnet sich  $L$  wie folgt:

$$L = L_1^{-1} L_2^{-1} \cdots L_{m-1}^{-1} = \begin{pmatrix} 1 & 0 & 0 & \cdot & 0 \\ l_{2,1} & 1 & 0 & \cdot & 0 \\ l_{3,1} & l_{3,2} & \ddots & 0 & 0 \\ & & \ddots & 1 & 0 \\ l_{m,1} & l_{m,2} & & l_{m,m-1} & 1 \end{pmatrix}$$

# Algorithmus der LU-Zerlegung

## Algorithmus 6.1 (LU-Zerlegung)

```
U = A  
L = I  
for k = 1 → m - 1 do  
    for j = k + 1 → m do  
         $l_{j,k} = u_{j,k} / u_{k,k}$   
         $u_{j,k:m} = u_{j,k:m} - l_{j,k} u_{k,k:m}$   
    end for  
end for
```

## Theorem 6.1 (Komplexität der LU-Zerlegung)

Der obige Algorithmus 6.1 benötigt etwa  $\frac{2}{3}m^3$  Operationen.

# Algorithmus der LU-Zerlegung

## Bemerkung 6.1 ()

- 1 Der Aufwand zur Berechnung der LU-Zerlegung mit  $\frac{2}{3}m^3$  Operationen ist um den Faktor 2 kleiner als der Aufwand der QR-Zerlegung, die  $\frac{4}{3}m^3$  Operationen benötigt. Aber die QR Zerlegung ist stabiler und daher bekannter.
- 2  $Ax = b$  kann durch eine LU-Zerlegung gelöst werden.

$$LUx = b$$

$Ly = b$       Forward Substitution ( $\approx m^2$  Operationen)

$Ux = y$       Back Substitution ( $\approx m^2$  Operationen)

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Motivation der LU-Faktorisierung mit Pivotisierung

## Problem:

Der Algorithmus 6.1 (LU-Zerlegung ohne Pivotisierung) funktioniert nicht, falls  $x_{k,k} = 0$ : Betrachte dazu die Matrix  $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ .

Außerdem wird der Algorithmus instabil, falls  $x_{k,k} \ll x_{j,k}$  für ein  $j$  mit  $k < j \leq m$ .

Da  $x_{kk}$  bei der Berechnung von  $l_k$  eine wichtige Rolle spielt, heißt  $x_{k,k}$  *Pivotelement*.

# Zulassen von Permutationen

Im  $k$ -ten Schritt werden Vielfache der  $k$ -ten Zeile von den Zeilen  $k + 1, \dots, m$  der aktuellen Matrix abgezogen und so Nullen in der  $k$ -ten Spalte dieser Zeilen erzeugt:

$$\begin{pmatrix} x & x & x & x & x \\ x_{k,k} & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \longrightarrow \begin{pmatrix} x & x & x & x & x \\ x_{k,k} & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{pmatrix}$$

Lässt man Permutationen von Zeilen und Spalten zu, so gibt es keinen Grund, die  $k$ -te Zeile und  $k$ -te Spalte im  $k$ -ten Schritt zu bearbeiten.  
(siehe Beispiel nächste Slide)

# Zulassen von Permutationen

Beispiel:  $k = 2, i = 4$ :

$$\begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ x_{i,k} & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \longrightarrow \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ x_{i,k} & \times & \times & \times \\ 0 & \times & \times & \times \end{pmatrix}$$

... oder  $k = 2, i = 4, j = 3$ :

$$\begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & x_{i,k} & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \longrightarrow \begin{pmatrix} \times & \times & \times & \times & \times \\ \times & 0 & \times & \times \\ \times & 0 & \times & \times \\ \times & x_{i,k} & \times & \times \\ \times & 0 & \times & \times \end{pmatrix}$$

# Komplettes vs. partielleles Pivotieren

Wird vor diesen Schritten das Element  $x_{ij}$  an die Position  $x_{kk}$  durch Zeilen- und Spaltenpermutation verschoben, so verläuft die Elimination wie im Standardfall. Leider ist der Aufwand für das Suchen eines guten Pivotelements im  $k$ -ten Schritt  $O((m - k)^2)$ , sodass insgesamt  $O(m^3)$  Suchoperationen durchgeführt werden müssen (**komplettes Pivotieren**). In der Praxis wird deshalb **partielles Pivotieren** eingesetzt. Dabei werden nur Zeilen vertauscht. Man sucht dazu das betragsgrößte Element in den  $(m - k + 1)$  Subdiagonaleinträgen der  $k$ -ten Spalte ( $O(m - k)$  Operationen, ergibt  $O(m^2)$  für die gesamte LU-Zerlegung). Sind bei der Spaltenpivotisierung alle Elemente der Spalte 0, so ist die Matrix singulär.

# Partielles Pivotieren

$$\begin{array}{c}
 \left( \begin{array}{ccccc} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x_{i,k} & x & x & x & x \\ x & x & x & x & x \end{array} \right) \xrightarrow{P_1} \left( \begin{array}{ccccc} x & x & x & x & x \\ x_{i,k} & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{array} \right) \xrightarrow{L_1} \left( \begin{array}{ccccc} x & x & x & x & x \\ x_{i,k} & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{array} \right) \\
 \text{Pivot auswählen} \qquad \qquad \qquad \text{Zeilen vertauschen} \qquad \qquad \qquad \text{Elimination}
 \end{array}$$

Mit diesem Verfahren erhält man eine *obere Dreiecksmatrix*  $U$  nach  $(m - 1)$  Schritten:

$$L_{m-1}P_{m-1} \cdots L_2P_2L_1P_1A = U,$$

wobei die Matrizen  $P$  Permutationsmatrizen sind.

# Exkurs: Permutationsmatrizen

Eine Permutationsmatrix ist eine binäre Matrix, die in jeder Zeile und in jeder Spalte genau einen 1-Eintrag hat.

## Definition 6.1 (Permutationsmatrix)

Sei  $\pi$  eine Permutation der Länge  $n$ . Eine Permutationsmatrix  $P$  zur Permutation  $\pi$  ist dann gegeben durch

$$P_\pi := (e_{\pi(1)} | \dots | e_{\pi(n)}).$$

Ein Beispiel für eine Permutationsmatrix: Vertauschen der 2-ten und 4-ten Zeile:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}}_P \begin{pmatrix} x & x & x & x \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \\ x & x & x & x \\ x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} \end{pmatrix} = \begin{pmatrix} x & x & x & x \\ x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} \\ x & x & x & x \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \end{pmatrix}$$

# Exkurs: Permutationsmatrizen

Allgemeines Schema für das Vertauschen der Zeilen  $i$  und  $j$ :

$$P_{i,j} = \begin{pmatrix} & & 1 & & & \\ & & \ddots & & & \\ & & & 1 & & \\ \hline i \rightarrow & & 0 & & 1 & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \\ \hline j \rightarrow & & 1 & & 0 & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

# Exkurs: Permutationsmatrizen

## Bemerkung 6.2 (Permutationsmatrizen)

- 1 Es gilt:  $P_{i,j}^2 = I$ .  
(Dies entspricht einem zweimaligen Vertauschen der gleichen Zeilen)
- 2 Multiplikation einer Matrix  $A$  mit einer Permutationsmatrix von links vertauscht Zeilen.
- 3 Multiplikation einer Matrix  $A$  mit einer Permutationsmatrix von rechts vertauscht Spalten.

# LU-Zerlegung mit Permutationsmatrizen

## Lemma 6.2 ()

Seien  $k < i$ ,  $P_i$  und  $L_k$  wie oben definiert. Dann ist  $P_i L_k = L'_k P_i$ , wobei  $L'_k$  bis auf eine Vertauschung von  $I_{i,k}$  und  $I_{j_i,k}$  wieder die Form von Gleichung (6.1) hat, d.h.

$$L'_k = I - I'_k e_k^*.$$

**Beweis:**

$$P_i L_k = P_i L_k P_i^2 = (P_i L_k P_i) P_i$$

$$(P_i L_k P_i) = \begin{pmatrix} \ddots & & & & \\ & 1 & & & \\ & -l_{k+1,k} & 1 & & \\ & \vdots & & \ddots & \\ & -l_{j_i,k} & & 0 & 1 \\ & \vdots & & & \\ & -l_{i_i,k} & & 1 & 0 \\ & \vdots & & & \ddots \end{pmatrix} \cdot P_i$$

$$= \begin{pmatrix} \ddots & & & & \\ & 1 & & & \\ & -l_{k+1,k} & 1 & & \\ & \vdots & & \ddots & \\ & -l_{j_i,k} & & & 1 \\ & \vdots & & & \\ & -l_{i_i,k} & & & 1 \\ & \vdots & & & \ddots \end{pmatrix} = L'_k$$

Setzt man nun

$$L'_k = P_{m-1} \cdots P_{k+1} L_k P_{k+1} \cdots P_{m-1},$$

so ist  $L'_k$  eine linke, untere Dreiecksmatrix und es gilt

$$\begin{aligned} L_{m-1} P_{m-1} \cdots L_2 P_2 L_1 P_1 A &= U \\ \Leftrightarrow \underbrace{(L'_{m-1} \cdots L'_2 L'_1)}_{=:L^{-1}} \underbrace{(P_{m-1} \cdots P_2 P_1)}_{=:P} A &= U. \end{aligned}$$

### Beispiel ()

$$\begin{aligned} L'_3 &= L_3, & L'_2 &= P_3 L_2 P_3, & L'_1 &= P_3 P_2 L_1 P_2 P_3 \\ L'_3 L'_2 L'_1 P_3 P_2 P_1 &= L_3 \underbrace{(P_3 L_2 P_3)}_{=:I} \underbrace{(P_3 P_2 L_1 P_2 P_3)}_{=:I} P_1 &= L_3 P_3 L_2 P_2 L_1 P_1 \end{aligned}$$

# LU-Faktorisierung (mit Pivotisierung)

Damit erhalten wir die allgemeine *LU*-Faktorisierung

$$PA = LU.$$

## Algorithmus 6.2 (LU-Faktorisierung mit Pivotisierung)

$$U = A, L = I, P = I$$

**for**  $k = 1 \rightarrow m - 1$  **do**

    Wähle  $i \geq k$  mit maximalem  $|u_{ik}|$ .

    Vertausche die Zeilen  $u_{k,k:m}$  und  $u_{i,k:m}$ .

    Vertausche die Zeilen  $l_{k,1:k-1}$  und  $l_{i,1:k-1}$ .

    Vertausche die Zeilen  $p_{k,:}$  und  $p_{i,:}$ .

**for**  $j = k + 1 \rightarrow m$  **do**

$$l_{j,k} = u_{j,k}/u_{k,k}$$

$$u_{j,k:m} = u_{j,k:m} - l_{j,k} u_{k,k:m}$$

**end for**

**end for**

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Cholesky-Faktorisierung: Motivation

Die LU-Zerlegung zerlegt eine Matrix  $A \in \mathbb{C}^{m \times m}$  in zwei verschiedene Matrizen  $L$  und  $U$ . Es liegt nahe zu fragen, ob zwischen diesen beiden Matrizen ein Zusammenhang besteht.

Die Cholesky-Zerlegung zerlegt eine Matrix  $A \in \mathbb{C}^{m \times m}$  in zwei Matrizen  $L$  und  $U$ , wobei gilt, dass  $L^* = U$ .

## Definition 6.2 (Cholesky-Zerlegung)

Eine Zerlegung einer hermitischen positiv definiten Matrix  $A \in \mathbb{C}^{m \times m}$  der Form

$$A = R^*R,$$

wobei  $R$  eine rechte obere Dreiecksmatrix ist, nennt man *Cholesky-Zerlegung*.

Im Folgenden wollen wir eine solche Zerlegung herleiten. Sei dazu  $A \in \mathbb{C}^{m \times m}$  hermitesch (im rellen Fall symmetrisch) und positiv definit, d.h.  $\forall x \neq 0 : x^*Ax > 0$ .

# Cholesky-Faktorisierung: Herleitung

Im Folgenden wollen wir eine solche Zerlegung herleiten. Sei dazu  $A \in \mathbb{C}^{m \times m}$  hermitesch (im rellen Fall symmetrisch) und positiv definit, d.h.  $\forall x \neq 0 : x^* A x > 0$ .

Idee: symmetrische Gauß-Elimination

Sei  $A$  von der Form

$$A = \begin{pmatrix} 1 & \omega^* \\ \omega & K \end{pmatrix},$$

wobei  $\omega$  ein  $(m - 1)$ -dimensionaler  $\mathbb{C}$ -Vektor und  $K$  eine hermitesche  $(m - 1) \times (m - 1)$ -Matrix ist.

Nun zerlegen wir  $A$  wie bei der LU-Zerlegung, d.h. wir erhalten

$$\begin{pmatrix} 1 & \omega^* \\ \omega & K \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \omega & I \end{pmatrix} \begin{pmatrix} 1 & \omega^* \\ 0 & K - \omega\omega^* \end{pmatrix}$$

(Vielfache der ersten Zeile werden von den anderen Zeilen subtrahiert).

# Cholesky-Faktorisierung: Herleitung

Anstatt mit der normalen Gauß-Elimination weiter zu machen, wird bei der Cholesky-Faktorisierung die Matrix symmetrisch gehalten, indem Nullen in der ersten Zeile eingeführt werden:

$$\begin{pmatrix} 1 & \omega^* \\ 0 & K - \omega\omega^* \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & K - \omega\omega^* \end{pmatrix}}_{\text{Adjungierte von } \begin{pmatrix} 1 & 0 \\ \omega & I \end{pmatrix}},$$

d.h.

$$\begin{pmatrix} 1 & \omega^* \\ \omega & K \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \omega & I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & K - \omega\omega^* \end{pmatrix} \begin{pmatrix} 1 & \omega^* \\ 0 & I \end{pmatrix}.$$

# Cholesky-Faktorisierung: Herleitung

Nun verallgemeinern wir: Sei  $A$  von der Form (ähnlich wie oben)

$$A = \begin{pmatrix} a_{11} & \omega^* \\ \omega & K \end{pmatrix}.$$

Da  $A$  positiv definit ist, muss  $a_{11}$  größer 0 sein.

Wir setzen nun  $\alpha = \sqrt{a_{11}}$  und erhalten

$$A = \begin{pmatrix} \alpha & 0 \\ \frac{\omega}{\alpha} & I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & K - \frac{\omega\omega^*}{a_{11}} \end{pmatrix} \begin{pmatrix} \alpha & \frac{\omega^*}{\alpha} \\ 0 & I \end{pmatrix} =: R_1^* A_1 R_1.$$

Rekursive Anwendung auf die (wieder hermitesche und positiv definite) Untermatrix  $K - \frac{\omega\omega^*}{a_{11}}$  liefert schließlich

$$A = \underbrace{R_1^* R_2^* \dots R_m^*}_{R^*} \underbrace{R_m \dots R_2 R_1}_R = R^* R, \quad r_{jj} > 0,$$

wobei  $R$  eine obere Dreiecksmatrix ist.

# Algorithmus und Komplexität der Cholesky-Zerlegung

## Algorithmus 6.3 (Cholesky-Zerlegung)

$$R = A$$

Setze Einträge von  $R$  unter der Diagonalen auf 0.

**for**  $k = 1 \rightarrow m$  **do**

**for**  $j = k + 1 \rightarrow m$  **do**

$$R_{j,j:m} = R_{j,j:m} - R_{k,j:m} \bar{R}_{k,j}/R_{k,k}$$

**end for**

$$R_{k,k:m} = R_{k,k:m} / \sqrt{R_{k,k}}$$

**end for**

## Satz 6.3 (Komplexität der Cholesky-Zerlegung)

Die Cholesky-Zerlegung benötigt  $\approx \frac{1}{3}m^3$  Operationen und hat somit eine Komplexität von  $O(m^3)$ .

# Existenz der Cholesky-Zerlegung

## Satz 6.4 (Cholesky-Zerlegung)

Jede hermitesche positiv definite Matrix  $A \in \mathbb{C}^{m \times m}$  hat eine Cholesky-Zerlegung.

## Bemerkung 6.3 (Laufzeitvergleich von LU- und Cholesky-Zerlegung)

Der Aufwand zur Berechnung einer Cholesky-Zerlegung ist um den Faktor 2 kleiner als der Aufwand einer LU-Zerlegung. Die Cholesky-Zerlegung existiert jedoch nur für hermitesche positiv-definite Matrizen, die LU-Zerlegung (mit Pivotisierung) existiert für jede quadratische Matrix.

# Kapitel 7:

## Eigenwertprobleme

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Eigenwerte und Eigenvektoren

## Definition 7.1 (Eigenwert und Eigenvektor)

Seien  $A \in \mathbb{C}^{m \times m}$  und  $x \in \mathbb{C}^m \setminus \{0\}$ . Dann heißt  $x$  *Eigenvektor* (EV) von  $A$  und  $\lambda \in \mathbb{C}$  der dazugehörige *Eigenwert* (EW), falls

$$Ax = \lambda x.$$

Der dazugehörige *Eigenraum* ist gegeben durch  $E_\lambda := \{x \in \mathbb{C}^m | Ax = \lambda x\}$ .

Die Eigenwerte einer Matrix (oder einer linearen Abbildung) kann man mittels des *charakteristischen Polynoms* bestimmen:

$$p_A(z) = \det(A - zI).$$

Zur Erinnerung:  $g_A(\lambda) = \dim E_\lambda$  ist die *geometrische Multiplizität* (oder geometrische Vielfachheit) von  $\lambda$ . Die *algebraische Multiplizität* (oder algebraische Vielfachheit)  $a_A$  von  $\lambda$  ist die Potenz des Faktors  $(z - \lambda)$  im charakteristischen Polynom. d.h.  $a_A(\lambda_i) = l_i$  für  $p_A = (z - \lambda_i)^{l_i}$  mit  $\lambda_i$  EW von  $A$ .

# Eigenwerte und Eigenvektoren

## Beispiel (Charakteristisches Polynom)

Betrachte die Matrix

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

Das charakteristische Polynom von  $A$  ist:

$$p_A(z) = \det(A - zI) = (1 - z)^2.$$

Die Matrix hat folgende Vielfachheiten:

$$\begin{aligned} g_A(\lambda) &= 1, \\ a_A(\lambda) &= 2. \end{aligned}$$

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Schur-Faktorisierung

## Definition 7.2 (**Schur-Faktorisierung**)

Eine *Schur-Faktorisierung* einer Matrix  $A \in \mathbb{C}^{m \times m}$  ist eine Faktorisierung der Form

$$A = QTQ^*,$$

wobei  $Q$  unitär und  $T$  eine obere Dreiecksmatrix ist.

## Bemerkung 7.1 ()

$A$  und  $T$  besitzen dasselbe charakteristische Polynom, dieselben Eigenwerte  $\lambda$  und die gleichen algebraischen und geometrischen Multiplizitäten:

$$\begin{aligned} p_A(z) &= \det(QTQ^* - zI) = \det(Q(T - zI)Q^*) \\ &= \det(Q)\det(T - zI)\det(Q^*) = \det(T - zI) = p_T(z) \end{aligned}$$

Damit müssen die Eigenwerte von  $A$  notwendigerweise auf der Diagonalen von  $T$  erscheinen ( $\det(Q)\det(Q^*) = |\det(Q)|^2 = 1$ ).

# Schur-Faktorisierung

## Satz 7.1 (Schur-Faktorisierung)

Jede quadratische Matrix  $A \in \mathbb{C}^{m \times m}$  besitzt eine Schur-Faktorisierung.

**Beweis:** Induktion über die Dimension  $m$  von  $A$ .

$m = 1$ : trivial.

Sei daher  $m \geq 2$ :

Sei  $x$  ein Eigenvektor von  $A$  mit EW  $\lambda$ . Es gilt:

$$Ax = \lambda x \Leftrightarrow A \frac{x}{\|x\|} = \frac{\lambda x}{\|x\|}.$$

$x$  sei also normalisiert und bilde die erste Spalte einer unitären Matrix  $U$ :

$$U = [x, y].$$

# Beweis des Satzes

Dann sind  $x$  und  $y$  orthogonal und es ist  $x * \lambda x = \lambda \|x\|_2^2 = \lambda$ . Es folgt:

$$\begin{aligned} U^* A U &= (x, y)^* A (x, y) \\ &= (x, y)^* (\lambda x, Ay) \\ &= \begin{pmatrix} \lambda & x^* Ay \\ 0 & y^* Ay \end{pmatrix} = \begin{pmatrix} \lambda & B \\ 0 & C \end{pmatrix} \end{aligned}$$

Nach Induktionsvoraussetzung existiert eine Schur-Faktorisierung  $VTV^*$  von  $C$ .

Nun setzen wir  $Q = U \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}$ . Dann gilt (Schur-Faktorisierung):

# Beweis des Satzes

$$\begin{aligned}
 \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}^* U^* A U \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}^* \underbrace{\begin{pmatrix} \lambda & B \\ 0 & C \end{pmatrix}}_{\begin{pmatrix} \lambda & BV \\ 0 & CV \end{pmatrix}} \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 \\ 0 & V^* \end{pmatrix} \begin{pmatrix} \lambda & BV \\ 0 & CV \end{pmatrix} \\
 &= \begin{pmatrix} \lambda & BV \\ 0 & V^* CV \end{pmatrix}
 \end{aligned}$$

□

# Frobenius Begleitmatrix

## Bemerkung 7.2 (Frobenius Begleitmatrix eines Polynoms)

Das charakteristische Polynom der  $m \times m$  Matrix

$$M = \begin{pmatrix} 0 & \dots & 0 & -a_0 \\ 1 & \ddots & \vdots & \vdots \\ \ddots & 0 & -a_{m-2} & \\ 0 & 1 & -a_{m-1} & \end{pmatrix}$$

ist

$$p(z) = a_0 + a_1 z + \dots + a_{m-1} z^{m-1} + z^m$$

Ist umgekehrt  $p$  ein beliebiges monisches Polynom vom Grad  $m$  mit den Koeffizienten  $a_0, \dots, a_{m-1}$ , dann nennt man die Matrix  $M$  die **Frobenius Begleitmatrix** von  $p$ .

# Frobenius Begleitmatrix

Zum Beweis betrachtet man  $M' = M - zI$  und entwickeln  $\det(M')$  nach der letzten Spalte. Für  $m = 4$  ergibt sich beispielsweise:

$$\begin{aligned}
 \det(M') &= \det \begin{pmatrix} -z & 0 & 0 & -a_0 \\ 1 & -z & 0 & -a_1 \\ 0 & 1 & -z & -a_2 \\ 0 & 0 & 1 & -z-a_3 \end{pmatrix} \\
 &= (-1)^m \left[ (-1)(-a_0) \det \begin{pmatrix} 1 & -z & 0 \\ 0 & 1 & -z \\ 0 & 0 & 1 \end{pmatrix} \right. \\
 &\quad + (-1)^2(-a_1) \det \begin{pmatrix} -z & 0 & 0 \\ 0 & 1 & -z \\ 0 & 0 & 1 \end{pmatrix} + (-1)^3(-a_2) \det \begin{pmatrix} -z & 0 & 0 \\ 1 & -z & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &\quad \left. + (-1)^4(-z - a_3) \det \begin{pmatrix} -z & 0 & 0 \\ 1 & -z & 0 \\ 0 & 1 & -z \end{pmatrix} \right] \\
 &= (-1)^4(a_0 + a_1z + a_2z^2 + a_3z^3 + z^4)
 \end{aligned}$$

# Problem der Eigenwertberechnung

Da für jedes  $m \geq 5$  ein Polynom  $p(z)$  mit Grad  $m$  existiert, das rationale Koeffizienten besitzt und eine rationale Wurzel  $v$  (d.h.  $(p(v) = 0)$ ) mit der Eigenschaft, dass  $v$  nicht als Term mit Radikalen, d.h. mit Additionen, Subtraktionen, Multiplikationen, Divisionen und  $k$ -ten Wurzeln aufgelöst werden kann (Theorem von Abel-Ruffini, Galoistheorie), können wir die exakten Wurzeln eines Polynoms nicht mit einer endlichen Zahl von Rechenschritten berechnen, d.h. wir sind auch nicht in der Lage, die entsprechenden Eigenwerte der Matrix exakt zu berechnen. Das bedeutet:

Jeder Eigenwertlöser muss iterativ sein!

Das beste, das wir erzielen können, ist eine hohe *Konvergenzrate*.

# Schur-Faktorisierung

## Bemerkung 7.3 (Schur-Faktorisierung)

Die meisten Algorithmen zur Eigenwertberechnung versuchen eine Schur-Faktorisierung zu bestimmen, indem eine Folge elementarer unitärer Ähnlichkeitstransformationen

$$X \mapsto Q_j^* X Q_j$$

durchgeführt wird, s.d.

$$Q_j^* \dots Q_2^* Q_1^* A Q_1 Q_2 \dots Q_j$$

gegen eine obere Dreiecksmatrix konvergiert.

Falls  $A$  hermitesch ist, dann ist  $\dots Q_j^* \dots Q_2^* Q_1^* A Q_1 Q_2 \dots Q_j \dots$  ebenfalls hermitesch und daher diagonal. Man kann also dieselben Algorithmen anwenden.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Einleitung

Obwohl sich viele der folgenden Konzepte auf allgemeine Matrizen verallgemeinern lassen, beschränken wir uns im folgenden der Einfachheit halber auf reelle symmetrische Matrizen, d.h.

$$A = A^T \in \mathbb{R}^{m \times m}, x \in \mathbb{R}^m, x^* = x^T \text{ und } \|x\| = \sqrt{x^T x}$$

Insbesondere besitzt  $A$  reelle Eigenwerte und eine vollständige Menge orthogonaler Eigenwerte.

Für den allgemeineren Fall weisen wir auf [?], Kap. 25-27.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

### ■ Rayleigh-Quotient

- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Motivation

**Gesucht:** Verfahren zur Schätzung von Eigenwerten

**Ansatz:** Zu gegebenem  $x \in \mathbb{R}^m \setminus \{0\}$  wird  $\alpha \in \mathbb{R}$  gesucht, das einem EW am ähnlichsten ist, d.h. das den folgenden Ausdruck minimiert:

$$\|Ax - \alpha x\|_2 \xrightarrow{\alpha} \min$$

Dies ist ein  $m \times 1$ -Least-Squares-Problem der Form

$$x\alpha = Ax,$$

wobei  $x$  die Matrix,  $\alpha$  der Vektor und  $Ax$  die rechte Seite ist. Es gilt:

$$\begin{aligned} \alpha \text{ Lösung des Problems} &\Leftrightarrow x^* x\alpha = x^* Ax \\ &\Leftrightarrow \alpha = \frac{x^* Ax}{x^* x} \end{aligned}$$

Damit ist  $r(x) := \alpha$  eine natürliche Schätzung für einen EW, falls  $x$  kein exakter EV ist.

# Rayleigh-Quotient

## Definition 7.3 (Rayleigh-Quotient)

Für einen Vektor  $x \in \mathbb{R}^m$  ist der Rayleigh-Quotient

$$r(x) := \frac{x^T A x}{x^T x}. \quad (7.1)$$

## Bemerkung 7.4 ()

Falls  $x$  ein EV ist, dann gilt

$$r(x) = \frac{x^T A x}{x^T x} = \frac{\lambda x^T x}{x^T x} = \lambda$$

# Gradient des Rayleigh-Quotienten

Wir betrachten nun den Gradienten von  $r(x)$  (siehe [?] S. 54) da wir so eine Charakterisierung von EV erhalten können:

$$\begin{aligned}
 \frac{\partial r(x)}{\partial x_j} &= \frac{f'g}{g^2} - \frac{g'f}{g^2} && \text{Quotientenregel} \\
 &= \frac{\frac{\partial}{\partial x_j}(x^T A x)}{x^T x} - \frac{(x^T A x) \frac{\partial}{\partial x_j}(x^T x)}{(x^T x)^2} \\
 &= \frac{2(Ax)_j}{x^T x} - \frac{(x^T A x)(2x_j)}{(x^T x)^2} \\
 &= \frac{2}{x^T x} (Ax - r(x)x)_j \\
 \Rightarrow \nabla r(x) &= \frac{2}{x^T x} (Ax - r(x)x)
 \end{aligned}$$

# Gradient des Rayleigh-Quotienten

## Bemerkung 7.5 (Eigenwertapproximation durch Rayleigh-Quotient)

- 1 Ist  $\nabla r(x) = 0$  für  $x \neq 0$ , dann ist  $x$  ein EV und  $r(x)$  der entsprechende EW, d.h. die EV von  $A$  sind *stationäre Punkte* von  $r(x)$  und die EW von  $A$  die Werte von  $r(x)$  an diesen Stellen.
- 2 Da  $\nabla r(q_j) = 0$  für alle EV  $q_j$ ,  $j = 1, \dots, m$ , von  $A$  gilt, folgt aus der Taylorapproximation

$$r(x) = r(q_j) + \nabla r(q_j)(q_j - x) + O(\|q_j - x\|^2),$$

dass

$$r(x) - r(q_j) = O(\|q_j - x\|^2),$$

d.h. der Rayleigh-Koeffizient ist eine quadratisch genaue Approximation eines Eigenwertes.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- **Potenziteration (Power Iteration)**
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Potenziteration

**Grundidee:** Sei  $A$  eine reelle symmetrische  $n \times n$  Matrix mit  $n$  betragsmäßig verschiedenen (daher reellen) Eigenwerten  $\lambda_i$ , die wir folgt indiziert sind:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| \geq 0.$$

Wir betrachte die Folge

$$\frac{x}{\|x\|}, \frac{Ax}{\|Ax\|}, \frac{A^2x}{\|A^2x\|}, \frac{A^3x}{\|A^3x\|}, \dots$$

Dann ist der Grenzwert dieser Folge der Eigenvektor zum betragsmäßig größte Eigenwert.

# Potenziteration

Um das zu einzusehen sei  $q_i$ ,  $\|q_i\| = 1$ ,  $i = 1, \dots, n$  jeweils der zu  $\lambda_i$  gehörende Eigenvektor von  $A$ .

Dann kann jeder Vektor  $v \in \mathbb{R}^n$  in diese Eigenbasis entwickelt werden:

$$v = \sum_{i=1}^n a_i q_i \text{ und es gilt } A^k v = \sum_{i=1}^n \lambda_i^k a_i q_i \quad \forall k \geq 1.$$

Konkret ergibt sich

$$\begin{aligned} v^{(0)} &= a_1 q_1 + a_2 q_2 + \dots + a_n q_n \quad q_i \text{ EV} \\ v^{(k)} &= c_k A^k v^{(0)} \quad (\text{wegen der Normierung}) \\ &= c_k (a_1 \lambda_1^k q_1 + a_2 \lambda_2^k q_2 + \dots + a_m \lambda_m^k q_m) \\ &= c_k \lambda_1^k \left( a_1 q_1 + a_2 \left( \frac{\lambda_2}{\lambda_1} \right)^k q_2 + \dots + a_m \left( \frac{\lambda_m}{\lambda_1} \right)^k q_m \right), \end{aligned}$$

d.h. für große  $k$  gilt

$$A^k v \approx \lambda_1^k a_1 q_1.$$

# Von Mises Potenzmethode

Verwenden wir den Rayleigh-Quotienten zur Schätzung des Eigenwerts von  $v^{(k)}$  so erhalten wir folgenden Algorithmus zur Eigenvektorschätzung zum betragsmäßig größten Eigenwert:

## Algorithmus 7.1 (Potenziteration)

Sei  $v^{(0)}$  ein vektor mit  $\|v^{(0)}\|_2 = 1$ .

**for**  $k = 1, 2, \dots$  **do**

$$w = Av^{(k-1)}$$

$$v^{(k)} = w / \|w\| \quad (\text{Normalisierung})$$

$$\lambda^{(k)} = (v^{(k)})^* Av^{(k)} \quad (\text{Rayleigh - quotient})$$

**end for**

## Theorem 7.2 (Konvergenz der Potenziteration)

Angenommen  $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_m| \geq 0$  sind sortierte EW,  $q_1 \in E_{\lambda_1}$  mit  $\|q_1\| = 1$  und  $q_1^T v^{(0)} \neq 0$ . Dann gilt

$$\|v^{(k)} - \pm q_1\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right), \quad |\lambda^{(k)} - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right) \quad \text{für } k \rightarrow \infty$$

Das Vorzeichen bei  $\pm$  wird hier in jedem Schritt so gewählt, dass das Ergebnis kleiner ist.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- **Inverse Iteration (Iteration der Potenzmethode)**
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Inverse Iteration

## Von der Potenziteration zur Potenzmethode:

Mit der Potenziteration kann der betragsmäßig größte Eigenwert berechnet werden. Wie berechnet man jedoch andere Eigenwerte?

Sei  $\mu \in \mathbb{R}$ , wobei  $\mu$  kein EW von  $A$  ist. Dann besitzt  $(A - \mu I)^{-1}$  dieselben Eigenvektoren wie  $A$ , denn:

$$Av = \lambda v \Leftrightarrow (A - \mu I)v = (\lambda - \mu)v \Leftrightarrow (\lambda - \mu)^{-1}v = (A - \mu I)^{-1}v$$

Die Eigenwerte von  $(A - \mu I)^{-1}$  sind also  $(\lambda_j - \mu)^{-1}$ , wobei die  $\{\lambda_j\}_j$  Eigenwerte von  $A$  sind. Ist nun  $\mu$  bereits nahe eines EW  $\lambda_J$  von  $A$ , dann ist  $(\lambda_J - \mu)^{-1}$  größer als  $(\lambda_j - \mu)^{-1} \forall j \neq J$ . Daher führt die Anwendung der Potenzmethode auf  $(A - \mu I)^{-1}$  zu einer Konvergenz gegen  $q_J$ . Diese Idee heißt *inverse Iteration*.

# Inverse Iteration

## Theorem 7.3 (Potenzmethode (Inverse Iteration))

Sei  $\lambda_J$  der EW, der am nächsten zu  $\mu$  liegt,  $\lambda_K$  der zweitnächste, d.h.

$$|\mu - \lambda_J| < |\mu - \lambda_K| \leq |\mu - \lambda_j| \quad \forall j \neq J.$$

Des Weiteren sei  $q_J^T v^{(0)} \neq 0$ . Dann gilt für die Iteration der Potenzmethode

$$\|v^{(k)} - (\pm q_J)\| = O\left(\left|\frac{\mu - \lambda_J}{\mu - \lambda_K}\right|^k\right), \quad |\lambda^{(k)} - \lambda_J| = O\left(\left|\frac{\mu - \lambda_J}{\mu - \lambda_K}\right|^{2k}\right)$$

(für  $k \rightarrow \infty$  und gleicher Bedeutung des  $\pm$ -Zeichens wie oben.)

**Beweis:** Folgt aus der Konvergenz der Potenziteration indem man  $A$  durch  $(A - \mu I)^{-1}$  ersetzt.

□

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- **Rayleigh-Quotient-Iteration**

## 21 QR-Verfahren

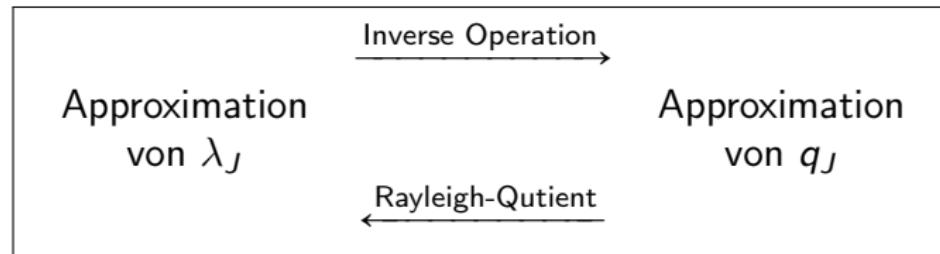
- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Rayleigh-Quotient-Iteration

Bisher haben wir folgende Methoden betrachtet:

- 1 EW-Schätzung aus einer EV-Schätzung (Rayleigh-Quotient)
- 2 EV-Schätzung aus EW-Schätzung (Inverse Iteration)

Jetzt kombinieren wir die beiden Ideen zur *Rayleigh-Quotient-Iteration*:



# Algorithmus: Rayleigh-Quotient-Iteration

## Algorithmus 7.2 (Rayleigh-Quotient-Iteration)

Sei  $v^{(0)}$  ein Vektor mit  $\|v^{(0)}\|_2 = 1$ .

$$\lambda^{(0)} = (v^{(0)})^* A v^{(0)} \quad (\text{Rayleigh-Quotient})$$

**for**  $k = 1, 2, \dots$  **do**

$$\text{solve } (A - \lambda^{(k-1)} I) w = v^{(k-1)}$$

$$v^{(k)} = w / \|w\| \quad (\text{Normalisierung})$$

$$\lambda^{(k)} = (v^{(k)})^* A v^{(k)} \quad (\text{Rayleigh-Quotient})$$

**end for**

# Konvergenz der Rayleigh-Quotient-Iteration

## Theorem 7.4 (Konvergenz der Rayleigh-Quotient-Iteration)

Die Rayleigh-Quotient-Iteration konvergiert gegen ein Eigenwert/Eigenvektor-Paar für alle Startvektoren  $v^{(0)} \neq 0$ . Falls sie konvergiert, ist die Konvergenz kubisch, d.h. falls  $\lambda_J$  ein Eigenwert von  $A$  und  $v^{(0)}$  hinreichend dicht zum Eigenvektor  $q_J$  liegt, dann gilt

$$\|v^{(k+1)} - (\pm q_J)\| = O(\|v^{(k)} - (\pm q_J)\|^3)$$

und

$$|\lambda^{(k+1)} - \lambda_J| = O(|\lambda^{(k)} - \lambda_J|^3)$$

(Hier ohne Beweis. Siehe [?], Kapitel 27.3. )

# Aufwand der Verfahren im Vergleich

Tabelle: Aufwand der Verfahren im Vergleich

## Power + Inverse Iteration

- Konvergenz: Quadratisch
- Jeder Schritt:  $O(m^2)$  Operationen
- $\Rightarrow O(m^2)$  Aufwand

## Rayleigh-Quotient-Iteration

- Konvergenz: Kubisch
- Jeder Schritt: Lösen eines Gleichungssystems:  $O(m^3)$  Operationen
- $\Rightarrow O(m^3)$  Aufwand
- Lösen des Gleichungssystems kann auf  $O(m^2)$  reduziert werden (QR- oder LU-Zerlegung)

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

## $k$ -ter Iterationsschritt

Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix und  $(\mu_k)_{k \geq 0}$  eine Folge reeller Zahlen, sogenannte '*Shifts*'. Dann berechnet man im  $k$ -ten Iterationsschritt eine QR-Zerlegung (vgl. Kapitel 9)

$$A_k - \mu_k I = Q_k R_k, \quad (7.2)$$

wobei  $Q_k$  unitär und  $R_k$  eine reelle obere Dreiecksmatrix ist und setzt

$$A_{k+1} := R_k Q_k + \mu_k I.$$

# QR-Verfahren

## Lemma 7.5 (QR-Verfahren)

Seien  $\mu_k$ ,  $Q_k$ ,  $R_k$  wie oben definiert und sei  $A_0 := A$ . Dann gelten die folgenden Identitäten:

- (a)  $A_{k+1} = Q_k^* A_k Q_k$  (d.h. insbesondere, dass die Matrizen  $A_k$  dieselben EW wie  $A$  besitzen)
- (b)  $A_{k+1} = (Q_0 Q_1 Q_2 \dots Q_k)^* A (Q_0 Q_1 Q_2 \dots Q_k)$
- (c)  $\prod_{j=0}^k (A - \mu_j I) = \underbrace{(Q_0 Q_1 Q_2 \dots Q_k)}_Q \underbrace{(R_k R_{k-1} \dots R_0)}_R$

**Beweis:**(a) Für  $A_{k+1}$  gilt:

$$\begin{aligned}
 A_{k+1} &= R_k Q_k + \mu_k I \\
 &= \underbrace{Q_k^* Q_k}_{id} R_k Q_k + \mu_k \underbrace{Q_k^* Q_k}_{id} \\
 &= Q_k^* \underbrace{(Q_k R_k + \mu_k I)}_{A_k} Q_k \\
 &= Q_k^* A_k Q_k
 \end{aligned}$$

(b) Folgt direkt aus (a) mittels vollständiger Induktion.

(c) Induktion über  $k$ :

$$\underline{k = 0:} \quad \underbrace{A}_{A_0} - \mu I = Q_0 R_0$$

$$\underline{k \rightarrow k + 1:}$$

(c) (Fortsetzung des Beweises)  
 $\underline{k \rightarrow k+1}$ :

$$\begin{aligned}
 Q_{k+1} R_{k+1} &= A_{k+1} - \mu_{k+1} I \\
 \stackrel{(b)}{=} &(Q_0 \dots Q_k)^* A (Q_0 \dots Q_k) - \mu_{k+1} (Q_0 \dots Q_k)^* (Q_0 \dots Q_k) \\
 &= (Q_0 \dots Q_k)^* (A - \mu_{k+1} I) (Q_0 \dots Q_k)
 \end{aligned}$$

$$\begin{aligned}
 \Rightarrow Q_0 \dots Q_k Q_{k+1} R_{k+1} &= (A - \mu_{k+1} I) (Q_0 \dots Q_k) \\
 \Rightarrow Q_0 \dots Q_k Q_{k+1} R_{k+1} R_k \dots R_0 &= (A - \mu_{k+1} I) \underbrace{(Q_0 \dots Q_k)(R_k \dots R_0)}_{\text{Ind.-Ann. } \prod_{j=0}^k (A - \mu_j I)} \\
 &= \prod_{j=0}^{k+1} (A - \mu_j I)
 \end{aligned}$$

□

# Konvergenz des QR-Verfahrens

## Satz 7.6 (Konvergenz des QR-Verfahrens)

Die Matrizen  $A_k$  werden im Verlauf der Iteration zu oberen Dreiecksmatrizen.

1. Wir betrachten zunächst den Fall ohne *Shifts*:  $\mu_k = 0 \ \forall k \in \mathbb{N}_0$ . Es ist

$$A_{k+1} = \underbrace{(Q_0 \dots Q_k)}_{Q_k} \underbrace{(R_k \dots R_0)}_{R_k} = A^{k+1} \quad (7.3)$$

eine QR-Zerlegung von  $A_{k+1}$ . Vergleicht man die jeweils erste Spalte dieser Matrizen gleichzeitig, so sieht man

$$\underbrace{A_{k+1} e_1}_{\text{1. Spalte}} = Q_k r_{11}^{(k)} e_1 = r_{11}^{(k)} q_1^{(k)}.$$

Nach der Potenzmethode (Algorithmus ??) konvergiert  $A^{k+1} e_1$  gegen den EV zum dominanten EW. Nach (b) ist

$$\begin{aligned} A_{k+1} &= Q_k^* A Q_k \\ A_{k+1} e_1 &= Q_k^* \underbrace{A q_1^{(k)}}_{\lambda_1 q_1^{(k)}} \approx \lambda_1 Q_k^* q_1^{(k)} \xrightarrow{\text{ausmultiplizieren}} \lambda_1 e_1 \end{aligned}$$

Damit hat  $A_{k+1}$  in etwa folgende Gestalt:

$$A_{k+1} \approx \left( \begin{array}{c|ccc} \lambda_1 & & \cdots & \\ 0 & & & \\ \vdots & & & \\ 0 & & \cdots & \end{array} \right).$$

2. Aus Gleichung (7.3) folgt für eine invertierbare Matrix  $A$  wegen der Orthogonalität von  $Q_k$

$$\begin{aligned} Q_k^* &= R_k(A^{k+1})^{-1} \\ q_n^{(k)} = e_n^* Q_k^* &= e_n^* R_k(A^{k+1})^{-1} = r_{nn}^{(k)} e_n^*(A^{k+1})^{-1}. \end{aligned}$$

Der Vektor  $q_n^{(k)}$  (letzte Spalte von  $Q^k$ ) ist also das mit  $r_{nn}^{(k)}$  multiplizierte Ergebnis der inversen Iteration, also eine Näherung des linken EV zum betragskleinsten EW  $\lambda_n$  von  $A$ . Daher folgt aus (b)

$$e_n^* A_{k+1} = e_n^* Q_k^* A Q_k = q_n^{(k)} A Q_k \approx \lambda_n q_n^{(k)*} Q_k = \lambda_n e_n$$

wobei  $A_{k+1} \approx \begin{pmatrix} \lambda_1 & \dots & \dots & \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$

Wir vermuten daher, dass  $A_k$  für  $k \rightarrow \infty$  gegen eine obere Dreiecksmatrix konvergiert.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Rayleigh-Quotienten-Iteration

3. Nach der vorangehenden Argumentation ist  $e_n$  ein Näherungsvektor der letzten Spalte von  $A_k$ .

Bilden wir also den Rayleigh-Quotienten  $\mu_k = e_n^* A_k e_n$  (rechtes unteres Element von  $A_k$ ) und führen einen Schritt der inversen Iteration bzgl. des linken EV aus, so ergibt dies

$$\underbrace{e_n^* (A_k - \mu_k I)^{-1}}_{\substack{Q_k R_k \\ R_k^{-1} \underbrace{Q_k^{-1}}_{Q_k^*}}} = e_n^* R_k^{-1} Q_k^* = \frac{1}{r_{nn}^{(k)}} e_n^* Q_k^* = \frac{1}{r_{nn}^{(k)}} q_n^{(k)*}, \quad (7.4)$$

wobei  $r_{nn}^{(k)}$  das rechte untere Element von  $R_k$  und  $q_n^{(k)}$  die hinterste Spalte von  $Q_k$  bezeichnet.

(Beachte:  $R_k^{-1}$  ist wieder eine Dreiecksmatrix, deren Diagonaleinträge die Kehrwerte der entsprechenden Diagonaleinträge von  $R_k$  sind.)

Mit anderen Worten: Ein Schritt der Rayleigh-Quotienten-Iteration ergibt gerade die hinterste Spalte von  $Q_k$  als neue Näherung an den linken EV von  $A_k$  zu  $\lambda_n$ .

Darüber hinaus erkennt man aus Lemma 7.5 Teil (a) sofort, dass das rechte untere Element von  $A_{k+1}$  der zugehörige Rayleigh-Quotient, also der nächste *Shift*  $\mu_{k+1}$  aus der Rayleigh-Quotienten-Iteration ist:

$$\mu_{k+1} = q_n^{(k)*} A_k q_n^{(k)} = e_n^* \underbrace{Q_k^* A_k Q_k}_{A_{k+1}} e_n = e_n^* A_{k+1} e_n \quad (7.5)$$

Wählt man also als *Shift*  $\mu_k$  in Gleichung (7.2) das  $(n, n)$ -te Element von  $A_k$ , dann darf man wie bei der Rayleigh-Quotienten-Iteration quadratische oder gar kubische Konvergenz dieser Elemente gegen den kleinsten EW  $\lambda_n$  von  $A$  erwarten.

*Shifts* dienen also der Konvergenz-Beschleunigung des QR-Verfahrens.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Zusammenfassung: QR-Verfahren mit und ohne Shifts

## Algorithmus 7.3 (QR-Verfahren ohne Shifts)

```

 $A^{(0)} = A$ 
for  $k = 1, 2, \dots$  do
     $Q^{(k)} R^{(k)} = A^{(k-1)}$ 
     $A^{(k)} = R^{(k)} Q^{(k)}$ 
end for

```

Die Matrix  $A \in \mathbb{C}^{m \times m}$  wird zunächst mittels QR-Faktorisierung in  $Q$  und  $R$  zerlegt. Dann werden  $R$  und  $Q$  in umgekehrter Reihenfolge multipliziert.

## Algorithmus 7.4 (QR-Verfahren mit einfachen Shifts)

```

 $A^{(0)} = A$ 
for  $k = 1, 2, \dots$  do
    Wähle einen Shift  $\mu^{(k)}$ 
     $Q^{(k)} R^{(k)} = A^{(k-1)} - \mu^{(k)} I$ 
     $A^{(k)} = R^{(k)} Q^{(k)} + \mu^{(k)} I$ 
end for

```

Statt die QR-Zerlegung von  $A$  zu berechnen, wird nun eine QR-Zerlegung von  $A - \mu^{(k)} I$  berechnet.

# Beziehungen zur Potenziteration: Simultane Iteration

## Algorithmus 7.5 (Simultane Iteration)

```
Wähle  $\bar{Q}^{(0)} = I$ 
for  $k = 1, 2, \dots$  do
     $Z = A\bar{Q}^{(k-1)}$ 
     $Z = \bar{Q}^{(k)}R^{(k)}$ 
     $A^{(k)} = (\bar{Q}^{(k)})^T A \bar{Q}^{(k)}$ 
end for
```

## Algorithmus 7.6 (QR-Verfahren ohne Shifts)

```
 $A^{(0)} = A$ 
for  $k = 1, 2, \dots$  do
     $A^{(k-1)} = Q^{(k)}R^{(k)}$ 
     $A^{(k)} = R^{(k)}Q^{(k)}$ 
     $\bar{Q}^{(k)} = Q^{(1)}Q^{(2)} \dots Q^{(k)}$ 
end for
```

# Beziehungen zur Potenziteration: Simultane Iteration

Definiert man sich für beide Algorithmen eine weitere Matrix  
 $\bar{R}^{(k)} := R^{(k)} R^{(k-1)} \dots R^{(1)}$ , so gilt folgender Satz:

## Satz 7.7 (Simultane Iteration und QR-Verfahren ohne Shifts)

Die beiden obigen Algorithmen erzeugen eine identische Folge von Matrizen  
 $\bar{R}^{(k)}, \bar{Q}^{(k)}, A^{(k)}$  mit

$$A^k = \bar{Q}^{(k)} \bar{R}^{(k)} \text{ mit der Projektion } A^{(k)} = (\bar{Q}^{(k)})^T A \bar{Q}^{(k)}.$$

(Hier ohne Beweis. Siehe [?], Kapitel 28.3. )

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Implementierung des QR-Verfahrens: das Hessenberg-Verfahren

Für beliebige Matrizen  $A \in \mathbb{K}^{n \times n}$  ist das QR-Verfahren sehr aufwendig, denn jede Iteration benötigt  $O(n^3)$  Operationen. Daher wird die Matrix  $A$  zunächst auf eine Obere Hessenberg-Form transformiert.

## Definition 7.4 ((Obere) Hessenbergmatrix)

Eine (obere) Hessenbergmatrix ist eine quadratische Matrix  $H \in \mathbb{C}^{n \times n}$ , deren Einträge unterhalb der ersten Nebendiagonalen gleich Null sind:

$$\begin{pmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \dots & h_{1,n} \\ h_{2,1} & h_{2,2} & h_{2,3} & \dots & h_{2,n} \\ 0 & h_{3,2} & h_{3,3} & \dots & h_{3,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{n,n-1} & h_{n,n} \end{pmatrix}$$

# Das Hessenberg-Verfahren

## Das zweistufige Hessenberg-Verfahren:

- 1 Berechnung einer oberen Hessenbergmatrix ( $O(m^3)$  Operationen)
- 2 Iteration zur Berechnung einer oberen Dreiecksmatrix. ( $O(m)$  Iterationen, um Maschinengenauigkeit zu erreichen, jede Iteration mit  $O(m^2)$  Aufwand.)

# Hessenbergverfahren: Schematische Darstellung

Das Verfahren lässt sich schematisch folgendermaßen darstellen:

$$\underbrace{\begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{pmatrix}}_{A \neq A^*} \xrightarrow{\text{Phase 1}} \underbrace{\begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{pmatrix}}_H \xrightarrow{\text{Phase 2}} \underbrace{\begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{pmatrix}}_T$$

$O(m^3) \quad O(m)O(m^2)$

Falls  $A$  hermitesch ist, so ist die Hessenbergmatrix *tridiagonal*:

$$\underbrace{\begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{pmatrix}}_{A = A^*} \xrightarrow{\text{Phase 1}} \underbrace{\begin{pmatrix} x & x & & & \\ x & x & x & & \\ & x & x & x & \\ & & x & x & x \\ & & & x & x \end{pmatrix}}_T \xrightarrow{\text{Phase 2}} \underbrace{\begin{pmatrix} x & & & & \\ & x & & & \\ & & x & & \\ & & & x & \\ & & & & x \end{pmatrix}}_D$$

$O(m^3) \quad O(m)O(m^2)$

# Algorithmus des Hessenbergverfahrens

Die Hessenberg-Form lässt sich analog zur Dreiecksform in der QR-Zerlegung durch Householder-Spiegelungen herstellen (vgl. Algorithmus 4.5). Die  $k$ -te Spiegelung wird dabei so modifiziert, dass sie das  $k$ -te Diagonalelement unberührt lässt. Um eine Ähnlichkeitstransformation zu erhalten, wird die unitäre Spiegelungsmatrix anders als bei der QR-Zerlegung sowohl von links als auch von rechts angewandt. Daraus resultiert folgender Algorithmus:

## Algorithmus 7.7 (Householder-Reduzierung auf Hessenberg-Form)

```

for  $k = 1 \rightarrow m - 2$  do
     $x = A_{k+1:m,k}$ 
     $v_k = \text{sign}(x_1) \|x\|_2 e_1 + x$ 
     $v_k = \frac{v_k}{\|v_k\|_2}$ 
     $A_{k+1:m,k:m} = A_{k+1:m,k:m} - 2v_k v_k^* A_{k+1:m,k:m}$ 
     $A_{1:m,k+1:m} = A_{1:m,k+1:m} - 2A_{1:m,k+1:m} v_k v_k^*$ 
end for

```

# Komplexität des Hessenbergverfahrens

## Satz 7.8 (Komplexität des Hessenbergverfahren)

Das Hessenbergverfahren benötigt  $\frac{10}{3}m^3$  Operationen.

Die Komplexitätsklasse des Verfahrens ist also  $O(m^3)$

## Satz 7.9 (Stabilität des Hessenbergverfahren)

Die Berechnung der Hessenbergform einer Matrix ist wie die Berechnung einer QR-Faktorisierung rückwärtsstabil.

- 1 Vektoren und Matrizen
- 2 Vektorräume
- 3 Normen
- 4 Lineare Ausgleichsrechnung
- 5 Einführung in die SVD
- 6 Die Singulärwertzerlegung in Matrixnotation
- 7 Unterschiede zwischen SVD und Eigenwertzerlegung
- 8 Singulärwertzerlegung und Lineare Ausgleichsrechnung
- 9 Projektoren und Projektionsmatrizen
  - Orthogonale Projektionen
  - Projektion mit orthonormaler Basis
- 10 Die QR-Zerlegung
  - Einschub: Vorwärts- und Rückwärtssubstitution

## 11 Das Gram-Schmidt-Verfahren

- Das klassische Gram-Schmidt-Verfahren
- Das modifizierte Gram-Schmidt-Verfahren
- Gram-Schmidt als Dreiecksorthonormalisierung

## 12 Householder-Triangularisierung

- Idee der Householder-Transformation
- Householder-Spiegelungen

## 13 Kondition

- Kondition einer Matrix
- Kondition eines Gleichungssystems

## 14 Floating Point Arithmetik

## 15 Stabilität

- Beispiele für Stabilitäten
  - Rückwärtsstabilität der Subtraktion

- Rückwärtsstabilität der Rückwärtssubstitution
- Rückwärtsstabilität des Produkts  $Q^* b$  (ohne Beweis)
- Rückwärtsstabilität der Householder-Triangularisierung (ohne Beweis)
- Lösen von Gleichungssystemen  $Ax = b$  mittels QR-Faktorisierung

## 16 LU-Faktorisierung (Gauß-Elimination)

- LU-Faktorisierung (ohne Pivotisierung)
- LU-Faktorisierung (mit Pivotisierung)

## 17 Cholesky-Faktorisierung

## 18 Eigenwerte und Eigenvektoren

## 19 Schur-Faktorisierung

## 20 Iterationsverfahren

- Rayleigh-Quotient
- Potenziteration (Power Iteration)
- Inverse Iteration (Iteration der Potenzmethode)
- Rayleigh-Quotient-Iteration

## 21 QR-Verfahren

- Beschleunigung des Verfahrens: Rayleigh-Quotienten-Iteration
- Zusammenfassung: QR-Verfahren mit und ohne Shifts
- Hessenberg: Zweistufiges Verfahren
- Berechnung einer Singulärwertzerlegung (SVD)

# Berechnung einer SVD mittels Eigenwertzerlegung

Die SVD  $A = U\Sigma V^*$  einer  $(m \times n)$ -Matrix  $A$ ,  $m \geq n$ , lässt sich aus einer EW-Zerlegung von  $A^*A$  berechnen:

$$A^*A = V\Sigma^*\Sigma V^* \quad (7.6)$$

Man kann also die SVD einer Matrix wie folgt bestimmen:

- 1 Berechne  $A^*A$ .
- 2 Berechne die EW-Zerlegung von  $A^*A = V\Lambda V^*$ .
- 3 Sei  $\Sigma$  die  $(m \times n)$  nicht-diagonale Wurzel von  $\Lambda$ .
- 4 Löse das LGS  $U\Sigma = AV$  für eine unitäre Matrix  $U$  (z.B. durch eine QR-Faktorisierung).

# Problem der SVD-Berechnung

Problem: Instabilität des Algorithmus:

- 1 Falls  $A^*A$  und  $\delta B$  gestört werden, so sind Änderungen der EW durch die 2-Norm von  $\delta B$  beschränkt:

$$|\lambda_k(A^*A + \delta B) - \lambda_k(A^*A)| \leq \|\delta B\|_2.$$

- 2 Ähnliches gilt für die Singulärwerte:

$$|\sigma_k(A + \delta A) - \sigma_k A| \leq \|\delta A\|_2.$$

Ein rückwärts-stabiler Algorithmus zur Berechnung der Singulärwerte  $\tilde{\sigma}_k$  erfüllt :

$$\begin{aligned}\tilde{\sigma}_k &= \sigma_k(A + \delta A), \quad \frac{\|\delta A\|}{\|A\|} = O(\varepsilon_M) \\ \Rightarrow |\tilde{\sigma}_k - \sigma_k| &= O(\varepsilon_M \cdot \|A\|)\end{aligned}$$

Es gilt für die Berechnung von  $\lambda_k(A^*A)$ :

$$\begin{aligned}|\tilde{\lambda}_k - \lambda_k| &= O(\varepsilon_M \cdot \|A^*A\|) = O(\varepsilon_M \cdot \|A\|^2) \\ \Rightarrow |\tilde{\sigma}_k - \sigma_k| &= O\left(\frac{|\tilde{\lambda}_k - \lambda_k|}{\sqrt{\lambda_k}}\right) = O\left(\varepsilon_M \cdot \frac{\|A\|^2}{\sigma_k}\right)\end{aligned}$$

Dies stellt kein Problem dar für dominante Singulärwerte mit  $\sigma_1 \approx \|A\|$ , ist jedoch problematisch für  $\sigma_k \ll \|A\|$ .

Angenommen,  $A$  ist quadratisch ( $m = n$ ). Betrachte die hermitesche Matrix

$$H = \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix}.$$

Wegen  $A = U\Sigma V^* \Leftrightarrow AV = U\Sigma \Leftrightarrow A^*U = V\Sigma^* = V\Sigma$  erhält man folgende EW-Zerlegung von  $H$ :

$$\begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix} \begin{pmatrix} V & V \\ U & -U \end{pmatrix} = \begin{pmatrix} V & V \\ U & -U \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix}.$$

D.h. die Singulärwerte von  $A$  sind die Absolutwerte der EW von  $H$  und die Singulärvektoren von  $A$  können aus den Eigenvektoren von  $H$  entnommen werden.

Im Gegensatz zur Berechnung der EW-Zerlegung von  $A^*A$  oder  $AA^*$  ist dieser Ansatz rückwärts-stabil. Die Standard-Algorithmen zur Berechnung der SVD basieren daher auf dieser Idee.

(Für Beispiele siehe [?], Kapitel 31.)

# Kapitel 8: Differenzierbare Funktionen

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Differenzierbare: Vorbemerkung

**Vorbemerkung:**  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt *differenzierbar* an einer Stelle  $a$ , falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h}$$

existiert. Gleichwertig damit ist die Existenz einer (von  $a$  abhängigen) linearen Abbildung  $L : \mathbb{R} \rightarrow \mathbb{C}$ , für die gilt

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a) - Lh}{\|h\|} = 0.$$

Dabei ist dann  $Lh = f'(a)h$ . Das heißt aber, dass der Zuwachs  $f(a + h) - f(a)$  durch  $Lh$  so gut approximiert wird, dass der Fehler schneller als  $\|h\|$  gegen 0 geht.

# Differenzierbareit: Definition

## Definition 8.1 (differenzierbar)

Eine Funktion  $f : U \rightarrow \mathbb{C}$  auf einer offenen Menge  $U \subset \mathbb{R}^n$  heißt *differenzierbar im Punkt  $a \in U$* , wenn es eine lineare Abbildung  $L : \mathbb{R}^n \rightarrow \mathbb{C}$  gibt derart, dass

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a) - Lh}{\|h\|} = 0. \quad (8.1)$$

Die Funktion  $f$  heißt *differenzierbar auf  $U$* , wenn sie in jedem Punkt  $x \in U$  differenzierbar ist.

## Bemerkung 8.1 ()

In (8.1) ist es gleichgültig, welche Norm verwendet wird, da alle Normen auf  $\mathbb{R}^n$  zueinander äquivalent sind.

# Umformulierung mittels Rest

Oft formuliert man (8.1) anhand des durch

$$f(a + h) - f(a) = Lh + R(h) \quad (8.2)$$

definierten Restes  $R(h)$ ; sie lautet dann

$$\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0. \quad (8.3)$$

# Eindeutigkeit des Differentials

## Lemma 8.1 (Eindeutigkeit des Differentials)

Die Gleichung (8.1) wird von höchstens einer linearen Abbildung erfüllt.

**Beweis:** Ist  $L^*$  eine weitere lineare Abbildung, so gilt für jeden Vektor  $v$  mit  $\|v\| = 1$

$$\begin{aligned} & \lim_{t \searrow 0} \frac{(f(a + tv) - f(a) - Ltv) - (f(a + tv) - f(a) - L^*tv)}{\|tv\|} \\ &= \lim_{t \searrow 0} \frac{(L^* - L)(tv)}{|t|} = (L^* - L)(v) = 0 \end{aligned}$$

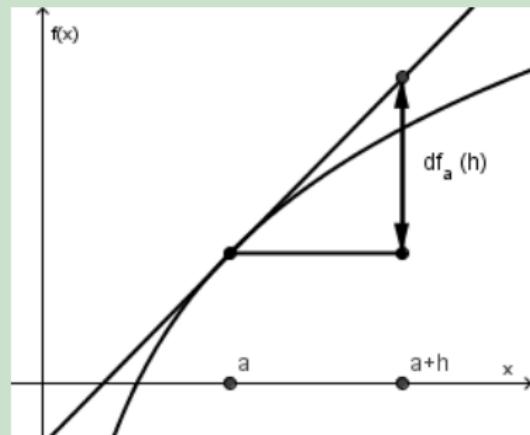
Da die Menge der Einheitsvektoren den  $\mathbb{R}^n$  aufspannt, folgt  $L = L^*$ .



# Differential

## Bemerkung 8.2 (Differential)

Die eindeutig bestimmte lineare Abbildung  $L$  heißt *Differential* oder *Linearisierung* von  $f$  im Punkt  $a$  und wird mit  $df(a)$  oder  $d_f(a)$  bezeichnet (siehe Abbildung). In alten Büchern wird das Differential auch als *totales Differential* bezeichnet.



# Ableitung

Sei  $\{e_1, \dots, e_n\}$  die Standardbasis des  $\mathbb{R}^n$ . Wegen der Linearität von  $df(a)$  gilt dann für jeden Vektor  $h = (h_1, \dots, h_n)^t \in \mathbb{R}^n$

$$df(a)h = \sum_{\nu=1}^n (df(a)e_\nu)h_\nu. \quad (8.4)$$

Dann ergibt sich mit nachfolgender Definition

$$df(a)h = f'(a)h. \quad (8.5)$$

## Definition 8.2 (Ableitung)

Den Zeilenvektor

$$f'(a) := (df(a)e_1, \dots, df(a)e_n) \quad (8.6)$$

nennen wir *Ableitung von  $f$  in  $a$* .

# Tangentialhyperebene

Die affin-lineare Funktion

$$Tf(x; a) := f(a) + f'(a)(x - a) \quad (8.7)$$

heißt *lineare Approximation* von  $f$  in  $a$  und bei reellem  $f$  heißt deren Graph

$$\{(x, x_{n+1}) \in \mathbb{R}^{n+1} | x_{n+1} = Tf(x; a)\} \quad (8.8)$$

die *Tangentialhyperebene* an den Graphen von  $f$  in  $(a, f(a))$ .

# Stetigkeit differenzierbarer Funktionen

## Satz 8.2 (Stetigkeit differenzierbarer Funktionen)

Eine in  $a$  differenzierbare Funktion ist auch in  $a$  stetig.

(Zum Beweis: In (8.2) gilt  $Lh \rightarrow 0$  und  $R(h) \rightarrow 0$  für  $h \rightarrow 0$ .)

Zur Erinnerung:

Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt *stetig* im Punkt  $a \in D$ , falls

$$\lim_{x \rightarrow a} f(x) = f(a).$$

# Beispiele differenzierbarer Funktionen

## Beispiel (1-zeilige Matrix)

Sei  $f(x) := Ax + b$ ,  $A$  eine 1–zeilige Matrix und  $b \in \mathbb{C}$ . Die durch  $Lh := Ah$  definierte lineare Abbildung erfüllt (8.1):

$$\lim_{h \rightarrow 0} \frac{A(x + h) + b - (Ax + b) - Ah}{\|h\|} = 0 \quad \forall x \in \mathbb{R}^n.$$

$f$  ist also in jedem Punkt  $a$  differenzierbar und es gilt

$$\begin{aligned} df_a h &= Ah \\ f'(a) &= A. \end{aligned}$$

# Beispiele differenzierbarer Funktionen

## Beispiel (symmetrische $(n \times n)$ -Matrix)

Sei  $f(x) := x^t A x$ ,  $A = (a_{ik})$  eine symmetrische  $(n \times n)$ -Matrix. Dann gilt:

$$f(a + h) - f(a) = 2a^t Ah + h^t Ah.$$

$Lh := 2a^t Ah$  definiert eine lineare Abbildung mit  $R(h) = h^t Ah$ .

Sei  $\sigma := \sum_{i,k=1}^n |a_{ik}|$ . Dann gilt  $|R(h)| = |h^t Ah| \leq \sigma \|h\|_\infty^2$  und damit

$$\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|_\infty} = 0.$$

$f$  ist also in jedem Punkt  $a$  differenzierbar und es gilt

$$df(a)h = 2a^t Ah, \quad f'(a) = 2a^t A.$$

## 22 Differenzierbare Funktionen

### ■ Darstellung des Differentials durch Richtungsableitungen

- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Richtungsableitungen: Vorbemerkung

Sei  $f$  eine differenzierbare Funktion. Die Werte  $df(a)h$ ,  $h \in \mathbb{R}^n$  sollen mit Hilfe von *Richtungsableitungen* ermittelt werden.

Für alle  $t \in \mathbb{R}$  mit hinreichend kleinem Betrag gilt zunächst

$$f(a + th) = f(a) + df(a)th + R(th).$$

Da  $R$  die Bedingung (8.3) erfüllt, gilt:

$$df(a)h = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t}. \quad (8.9)$$

# Partielle Ableitungen

## Definition 8.3 (Partielle Ableitungen)

Sei  $f : U \rightarrow \mathbb{C}$  eine (nicht notwendig differenzierbare) Funktion in einer Umgebung  $U$  von  $a$ . Dann versteht man unter der Ableitung von  $f$  im Punkt  $a$  in Richtung des Vektors  $h \in \mathbb{R}^n$  im Existenzfall den Grenzwert

$$D_h f(a) = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t}.$$

Die Ableitungen in den Richtungen  $e_1, \dots, e_n$  der Standardbasis heißen *partielle Ableitungen* von  $f$  und  $f$  heißt *partiell differenzierbar* in  $a$ , wenn alle partiellen Ableitungen  $D_{e_1} f(a), \dots, D_{e_n} f(a)$  existieren.  
Weitere Bezeichnungen für die partiellen Ableitungen sind:

$$D_{e_1} f(a) = \partial_1 f(a) = \frac{\partial f}{\partial x_1}(a) = f_{x_1}(a)$$

# Richtungsableitungen

## Satz 8.3 (Richtungsableitungen)

Eine in  $a$  differenzierbare Funktion  $f$  hat dort Richtungsableitungen in jeder Richtung. Sie ist dort insbesondere partiell differenzierbar. Ihr Differential in  $a$  hat für jeden Vektor  $h = (h_1, \dots, h_n)^t \in \mathbb{R}^n$  den Wert

$$df(a)h = f'(a)h = \partial_h f(a) = \sum_{\nu=1}^n \partial_\nu f(a)h_\nu \quad (8.10)$$

und ihre Ableitung  $f'(a)$  ist die 1–zeilige Matrix

$$f'(a) = (\partial_1 f(a), \dots, \partial_n f(a))$$

# Richtungsableitungen

**Beweis:** Die Existenz ist mit der Herleitung von (8.9) gezeigt. Die Formeln sind wegen  $df(a)e_\nu = \partial_\nu f(a)$  identisch mit (8.4), (8.6) und (8.5). Berechnung der partiellen Ableitungen:

Die Definition  $\partial_\nu f(a) = \lim_{t \rightarrow 0} \frac{f(a+te_\nu) - f(a)}{t}$  mit  $a = (a_1, \dots, a_n)^t$  läuft darauf hinaus, in  $f(x_1, \dots, x_n)$  alle Variablen  $x_k$  bis auf die  $\nu$ -te konstant  $= a_k$  zu setzen und die dann nur von  $x_\nu$  abhängige Funktion als Funktion einer Variablen zu differenzieren. □

Beispiel ( $f(x, y) = x^2 + y^2$ )

Die Richtungsableitungen der Funktion  $f(x, y) = x^2 + y^2$  sind  $\partial f_x(a, b) = 2a$  und  $\partial f_y(a, b) = 2b$ .

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- **Hauptkriterium für Differenzierbarkeit**
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Motivation

Um eine Funktion  $f$  auf Differenzierbarkeit in  $a$  zu untersuchen, klärt man zunächst, ob sie partiell differenzierbar ist. Im positiven Fall prüft man weiter, ob die einzige als Differential in Frage kommende lineare Abbildung

$$L : \mathbb{R}^n \longrightarrow \mathbb{C},$$

$$Lh = \sum_{\nu=1}^n \partial_\nu f(a)h$$

die Bedingung (8.1) erfüllt.

# Hinreichende Bedingung

## Bemerkung 8.3 (Hinreichende Bedingung)

Die bloße Existenz der partiellen Ableitungen impliziert i.a. nicht die Differenzierbarkeit.

Betrachte die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , definiert durch

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

$f$  ist im Nullpunkt nicht stetig, also erst recht nicht differenzierbar. Speziell in  $(0, 0)$  hat  $f$  wegen  $f(x, 0) = 0$  und  $f(0, y) = 0$  die partiellen Ableitungen  $\partial_x f(0, 0) = 0$  und  $\partial_y f(0, 0) = 0$ . In den Punkten  $(0, y)$ ,  $y \neq 0$  gilt  $\partial_x f(0, y) = 1/y$ .  $\partial_x f$  ist im Nullpunkt also unstetig,  $\partial_y f$  ebenso.

## Bemerkung 8.4 ()

Auch die Existenz aller Richtungsableitungen hat nicht die Differenzierbarkeit zur Folge.

Betrachte dazu die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , definiert durch

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

Es ist

$$f(tx, ty) = \frac{t^2 x^2 t y}{t^2 x^2 + t^2 y^2} = t f(x, y).$$

Diese Geraden haben im Nullpunkt Ableitungen in jede Richtung  $h = (h_1, h_2)$ :

$$\partial_h f(0, 0) = \lim_{t \rightarrow 0} \frac{f(th_1, th_2) - f(0, 0)}{t} = f(h_1, h_2).$$

# Hinreichende Bedingung

(Fortsetzung der Bemerkung)

Insbesondere sind die partiellen Ableitungen  $\partial_x f(0,0) = 0$  und  $\partial_y f(0,0) = 0$ . Als Differential kommt also höchstens  $L = (0,0)$  in Frage. Damit ist (8.1) aber nicht erfüllt, da  $\forall (h_1, h_1)$ :

$$\frac{f(h_1, h_1) - f(0,0) - L(h_1, h_1)}{\|(h_1, h_1)\|_\infty} = \frac{h_1^3}{2h_1^2|h_1|} = \pm \frac{1}{2} \neq 0$$

gilt. Folglich ist  $f$  im Nullpunkt nicht differenzierbar. Man prüft leicht nach, dass die partielle Ableitung im Nullpunkt unstetig ist. (In den Punkten  $(0, y)$ ,  $y \neq 0$  gilt  $\frac{\partial f}{\partial x}(0, y) = \frac{1}{y}$ .  $\frac{\partial f}{\partial x}$  ist im Nullpunkt also unstetig.)

# Differenzierbarkeitskriterium

## Theorem 8.4 (Differenzierbarkeitskriterium)

Existieren in einer Umgebung  $U$  von  $a \in \mathbb{R}^n$  alle partiellen Ableitungen  $\partial_1 f, \dots, \partial_n f$  und sind diese im Punkt  $a$  stetig, so ist  $f$  in  $a$  differenzierbar.

**Beweis:** Wir dürfen  $f$  als reell voraussetzen, da ein komplexes  $f$  genau dann differenzierbar ist, wenn  $\operatorname{Re} f$  und  $\operatorname{Im} f$  differenzierbar sind.

Wir zeigen dann, dass die Linearform  $L : \mathbb{R}^n \longrightarrow \mathbb{R}$  mit  $Lh = \sum_{\nu=1}^n \partial_\nu f(a) h_\nu$  die Bedingung (8.1) erfüllt.

Sei  $Q$  ein offener achsenparalleler Quader in  $U$  mit  $a \in Q$ . Jeder Punkt  $a + h \in Q$  kann mit  $a$  durch stückweise achsenparallelen Streckung in  $Q$  verbunden werden. Man setze dazu  $a_0 := a$ ,  $a_\nu = a_{\nu-1} + h_\nu e_\nu$ ,  $\nu = 1, \dots, n$ . Insbesondere ist dann  $a_n = a + h$  und

$$f(a + h) - f(a) = \sum_{\nu=1}^n (f(a_\nu) - f(a_{\nu-1})).$$

Die Differenzen der Summe werden gemäß dem Mittelwertsatz der Differentialrechnung umgeformt. Betrachte dazu die Funktionen  $\varphi_\nu : [0, h_\nu] \longrightarrow \mathbb{R}$ ,  $\varphi_\nu(t) := f(a_{\nu-1} + te_\nu)$ . Mit diesen Funktionen gilt

$$f(a_\nu) - f(a_{\nu-1}) = \varphi_\nu(h_\nu) - \varphi_\nu(0).$$

Da  $f$  partiell differenzierbar ist, sind die Funktionen  $\varphi_\nu$  differenzierbar und es gilt

$$\varphi'_\nu(t) = \partial_\nu f(a_{\nu-1} + te_\nu).$$

Nach dem Mittelwertsatz gibt es ein  $\tau_\nu \in [0, h_\nu]$ , sodass  $\varphi_\nu(h_\nu) - \varphi_\nu(0) = h_\nu \varphi'(\tau_\nu)$ . Mit  $\xi_\nu := a_{\nu-1} + \tau_\nu e_\nu$  folgt nun  $f(a_\nu) - f(a_{\nu-1}) = h_\nu \partial_\nu f(\xi_\nu)$ . Damit ergibt sich

$$f(a+h) - f(a) - Lh = \sum_{\nu=1}^n (\partial_\nu f(\xi_\nu) - \partial_\nu f(a)) h_\nu$$

und weiter

$$|f(a+h) - f(a) - Lh| \leq \|h\|_\infty \sum_{\nu=1}^n |\partial_\nu f(\xi_\nu) - \partial_\nu f(a)|.$$

Für  $h \rightarrow 0$  gilt  $\xi_\nu \rightarrow a$ ,  $\nu = 1, \dots, n$  und wegen der Stetigkeit der partiellen Ableitungen in  $a$  erhält man also

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Lh}{\|h\|_\infty} = \lim_{h \rightarrow 0} \sum_{\nu=1}^n |\partial_\nu f(\xi_\nu) - \partial_\nu f(a)| = 0.$$

## Beispiel (Differentiation rotationssymmetrischer Funktionen)

Es sei  $F : I \longrightarrow \mathbb{C}$  eine Funktion auf einem Intervall  $I \subset [0, \infty)$ .

Mit  $F$  erhält man auf der Kugelschale

$K(I) := \left\{ x \in \mathbb{R}^n \mid \|x\|_2 = \sqrt{\sum_{\nu=1}^n x_\nu^2} \in I \right\}$  eine Funktion  $f$  durch  
 $f(x) := F(\|x\|_2)$ .

Es sei nun  $I$  offen und  $F$  stetig differenzierbar. Damit ist auch  $K(I)$  offen und  $f$  hat an jeder Stelle  $x \in K(I)$ ,  $x \neq 0$ , die partiellen Ableitungen

$$\partial_\nu f(x) = F'(\|x\|_2) \cdot \frac{x_\nu}{\|x\|_2}, \quad \nu = 1, \dots, n.$$

Diese sind offensichtlich stetig. Somit ist  $f$  an jeder von 0 verschiedenen Stelle  $x \in K(I)$  differenzierbar und hat dort die Ableitung

$$f'(x) = \frac{F'(\|x\|_2)}{\|x\|_2} \cdot x^t \tag{8.11}$$

# Stetig differenzierbar

## Definition 8.4 (stetig differenzierbar)

Eine differenzierbare Funktion  $f : U \rightarrow \mathbb{C}$  auf einer offenen Menge  $U \subset \mathbb{R}^n$  heißt *stetig differenzierbar auf  $U$* , wenn  $df : U \rightarrow L(\mathbb{R}^n, \mathbb{C})$  stetig ist. Dies ist gleichwertig zur Stetigkeit der Ableitung

$$f' : U \rightarrow \mathbb{C}^n, \quad x \mapsto (\partial_1 f(x), \dots, \partial_n f(x)).$$

Mit dem Differenzierbarkeitskriterium folgt, dass eine Funktion  $f : U \rightarrow \mathbb{C}$  genau dann stetig differenzierbar ist, wenn alle  $n$  partiellen Ableitungen  $\partial_1 f, \dots, \partial_n f$  auf  $U$  existieren und stetig sind.

Den Vektorraum der stetig differenzierbaren Funktionen auf  $U$  bezeichnet man mit  $\mathcal{C}^1(U)$ .

# Gradient

## Definition 8.5 (Gradient)

Auf dem  $\mathbb{R}^n$  sei ein Skalarprodukt gegeben. Dann kann man jede Linearform (lineare Abbildung)  $L : \mathbb{R}^n \rightarrow \mathbb{R}$  mit Hilfe eines eindeutig bestimmten Vektors  $g \in \mathbb{R}^n$  darstellen:

$$Lh = \langle g, h \rangle \quad \forall h \in \mathbb{R}^n.$$

Ist  $L$  das Differential einer in  $a$  differenzierbaren reellwertigen Funktion  $f$ , so heißt  $g$  *Gradient von  $f$  in  $a$  bzgl.  $\langle \cdot, \cdot \rangle$* . Er wird mit  $\text{grad}f(a)$  bezeichnet. Bzgl.  $\langle \cdot, \cdot \rangle$  ist er der durch

$$df(a)h = \partial_h f(a) = \langle \text{grad}f(a), h \rangle$$

eindeutig bestimmte Vektor im  $\mathbb{R}^n$ .

# Gradient (Standardskalarprodukt)

## Bemerkung 8.5 ()

Im Fall des Standardskalarprodukts ist  $\text{grad}f(a)$  nach (8.10) der Spaltenvektor

$$\text{grad}f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} =: \nabla f(a) \in \mathbb{R}^n.$$

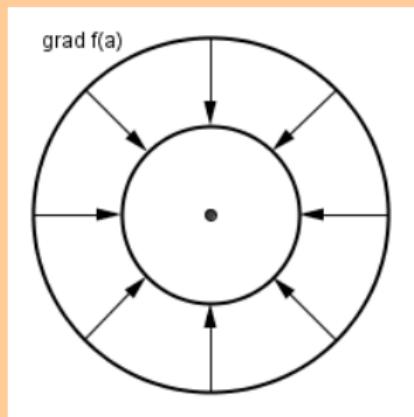
( $\nabla f(a)$  wird gesprochen “nabla  $f$  von  $a$ “.)

## Beispiel (Gradient einer rotationssymmetrischen Funktionen)

Die in (8.11) definierte rotationssymmetrische Funktion  $f$  hat im Fall einer reellen  $\mathcal{C}^1$ -Funktion  $F$  im Punkt  $a \neq 0$  den Gradienten

$$\text{grad } f(a) = \frac{F'(\|a\|_2)}{\|a\|_2} \cdot a.$$

Dieser ist im Fall  $F'(\|a\|_2) > 0$  zum Ortsvektor  $\overrightarrow{0a}$  parallel und im Fall  $F'(\|a\|_2) < 0$  antiparallel.



# Maximalitätseigenschaft

$\|\cdot\|$  bezeichne die zum Skalarprodukt gehörige Norm. Aufgrund der Cauchy-Schwarz-Ungleichung gibt es einen Winkel  $\varphi$  zwischen den Vektoren  $\text{grad}f(a)$  und  $h$  derart, dass

$$df(a)h = \partial_h f(a) = \langle \text{grad}f(a), h \rangle = \|\text{grad}f(a)\| \cdot \|h\| \cdot \cos \varphi. \quad (8.12)$$

Nach dieser Darstellung zeichnet sich der Gradient durch folgende Maximalitätseigenschaft aus:

- 1 Seine Länge  $\|\text{grad}f(a)\|$  ist das Maximum aller Richtungsableitungen  $\partial_h f(a)$  nach den Einheitsvektoren

$$\|\text{grad}f(a)\| = \max \{ \partial_h f(a) \mid \|h\| = 1 \} =: M$$

- 2 Im Fall  $M \neq 0$  gibt es genau einen Einheitsvektor  $\hat{h}$  mit  $\partial_{\hat{h}} f(a) = M$  und mit diesem ist  $\text{grad}f(a) = M\hat{h}$ . Der Gradient zeigt also in die Richtung des stärksten Anstiegs der Funktion im Punkt  $a$ .

# Rechenregeln für differenzierbare Funktionen

## Satz 8.5 (Rechenregeln)

Es gelten folgende algebraische Regeln:

Seien  $f, g : U \rightarrow \mathbb{C}$  differenzierbar in  $a \in U$ . Dann sind auch  $f + g$  und  $f \cdot g$  in  $a$  differenzierbar und es gilt

$$d(f + g)(a) = df(a) + dg(a)$$

$$d(f \cdot g)(a) = df(a) \cdot g(a) + f(a) \cdot dg(a)$$

Ist zusätzlich  $f(a) \neq 0$ , so ist auch  $\frac{1}{f}$  in  $a$  differenzierbar mit

$$d\left(\frac{1}{f}\right)(a) = -\frac{df(a)}{f^2(a)} \quad (\text{Quotientenregel}).$$

# Rechenregeln für differenzierbare Funktionen

## Bemerkung 8.6 ()

Für die Ableitungen in Satz (8.5) gelten also dieselben Regeln wie im Fall  $n = 1$ :

$$(f + g)'(a) = f'(a) + g'(a)$$

$$(f \cdot g)'(a) = f'(a) \cdot g(a) + g'(a) \cdot f(a)$$

$$\left(\frac{1}{f}\right)'(a) = -\frac{f'(a)}{f^2(a)}$$

## Lemma 8.6 (Rechenregeln für Differenzierbarkeit auf $U$ )

Sind  $f$  und  $g$  in  $U$  stetig differenzierbar, dann sind auch  $f + g$ ,  $f \cdot g$  und  $f/g$  in  $\{x \in U \mid g(x) \neq 0\}$  differenzierbar.

**Beweis:** (Quotientenregel) Zeige, dass die Linearform  $-df(a)/f^2(a)$  die Bedingung (8.1) erfüllt, d.h.

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Lh}{\|h\|} = 0.$$

Für hinreichend kurze Vektoren  $h \in \mathbb{R}^n$  ist auch  $f(a+h) \neq 0$  und es gilt

$$\begin{aligned} & \frac{1}{\|h\|} \left( \frac{1}{f(a+h)} - \frac{1}{f(a)} + \frac{df(a)h}{f^2(a)} \right) \\ &= \frac{-1}{f(a)f(a+h)} \left( \underbrace{\frac{f(a+h) - f(a) - df(a)h}{\|h\|}}_{\substack{\xrightarrow{h \rightarrow 0} 0 \\ \text{Diffbarkeit von } f \text{ in } a}} + \underbrace{\frac{f(a) - f(a+h)}{f(a)} \cdot \frac{df(a)h}{\|h\|}}_{\frac{df(a)h}{\|h\|} \text{ bleibt beschränkt für } \|h\| \rightarrow 0} \right) \end{aligned}$$



### Folgerung 8.7 (Rationale Funktionen)

Jede rationale Funktion ist in ihrem Definitionsbereich stetig differenzierbar.

# Kettenregel (1. Version)

## Satz 8.8 (Kettenregel (1. Version))

Es sei  $\gamma = \gamma_1, \dots, \gamma_n : I \longrightarrow U$  differenzierbar in  $t_0$  und  $f : U \longrightarrow \mathbb{C}$  differenzierbar in  $a = \gamma(t_0)$ . Dann ist  $f \circ \gamma$  differenzierbar in  $t_0$  und hat dort die Ableitung

$$\frac{d(f \circ \gamma)}{dt}(t_0) = df(a)\dot{\gamma}(t_0) = f'(a)\dot{\gamma}(t_0) = \sum_{i=1}^n \partial_i f(a) \cdot \dot{\gamma}_i(t_0).$$

Mit Hilfe des Gradienten lautet die Formel nach Definition 8.5

$$\frac{d(f \circ \gamma)}{dt}(t_0) = \langle \text{grad}f(a), \dot{\gamma}(t_0) \rangle$$

**Beweis:** Für  $k \in \mathbb{R}$ ,  $h \in \mathbb{R}^n$  mit hinreichend kleinen Beträgen gilt nach Voraussetzung

$$\begin{aligned}\gamma(t_0 + k) &= \gamma(t_0) + \dot{\gamma}(t_0)k + r_1(k)|k|, & \lim_{k \rightarrow 0} r_1(k) &= 0 \\ f(a + h) &= f(a) + df(a)h + r_2(h)\|h\|, & \lim_{h \rightarrow 0} r_2(h) &= 0\end{aligned}$$

Setzt man  $h := \gamma(t_0 + k) - \gamma(t_0)$ , so folgt

$$f(\gamma(t_0 + k)) = f(\gamma(t_0)) + df\gamma(t_0)\dot{\gamma}(t_0)k + R(k), \quad (8.13)$$

wobei

$$R(k) := df(a)r_1(k)|k| + r_2(\gamma(t_0 + k) - \gamma(t_0))\|\dot{\gamma}(t_0)k + r_1(k)|k|\|.$$

Offensichtlich gilt  $\lim_{k \rightarrow 0} \frac{R(k)}{k} = 0$ . Damit folgt die Behauptung aus (8.13).



# Kettenregel: Beispiel

## Beispiel (Kettenregel)

Sei  $f$  eine differenzierbare Funktion auf  $\mathbb{R}^2$ . Wir betrachten ihre Komposition  $F := f \circ \rho_2$  mit der Polarkoordinatenabbildung

$$F(r, \varphi) = f(r \cos \varphi, r \sin \varphi).$$

Differenziert man  $F$  bei festgehaltenem  $\varphi$  nach  $r$ , erhält man die partielle Ableitung nach  $r$  (nach  $\varphi$  entsprechend). Es ergibt sich

$$F_r(r, \varphi) = f_x(r \cos \varphi, r \sin \varphi) \cdot \cos \varphi + f_y(r \cos \varphi, r \sin \varphi) \cdot \sin \varphi$$

$$F_\varphi(r, \varphi) = f_x(r \cos \varphi, r \sin \varphi) \cdot (-r \sin \varphi) + f_y(r \cos \varphi, r \sin \varphi) \cdot r \cos \varphi$$

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- **Orthogonalität von Gradient und Nullmenge**
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Orthogonalität von Gradient und Nullmenge

Sei  $f : U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  eine differenzierbare Funktion und  $\gamma : I \rightarrow U$  eine differenzierbare Kurve, die in einer Niveaumenge von  $f$  verläuft, d.h. es ist  $f(\gamma(t)) = c$  für eine Konstante  $c$  und  $\forall t \in I$ .

Dann steht der Gradient von  $f$  im Punkt  $\gamma(t)$  senkrecht auf dem Tangentialvektor  $\dot{\gamma}(t)$ :

$$\text{grad}f(\gamma(t)) \perp \dot{\gamma}(t), \quad t \in I.$$

**Beweis:** Wegen  $c = f \circ \gamma(t)$  gilt

$$0 = \frac{d}{dt}(f \circ \gamma)(t) = \langle \text{grad}f(\gamma(t)), \dot{\gamma}(t) \rangle = 0$$



# Mittelwertsatz

## Theorem 8.9 (Mittelwertsatz)

Sei  $f$  eine reelle differenzierbare Funktion in einer offenen Menge  $U \subset \mathbb{R}^n$ . Ferner seien  $a, b \in U$  Punkte, deren Verbindungsstrecke in  $U$  liegt. Dann gibt es einen Punkt  $\xi \in [a, b]$  mit

$$f(a) - f(b) = df(\xi)(b - a) = f'(\xi)(b - a)$$

**Beweis:** Setze  $\gamma(t) := a + t(b - a)$ ,  $t \in [0, 1]$  und betrachte  $F := F \circ \gamma : [0, 1] \longrightarrow \mathbb{R}$ . Dann gilt  $F(1) - F(0) = f(b) - f(a)$ . Nach der Kettenregel ist  $F$  differenzierbar und nach dem Mittelwertsatz der Differentialrechnung in einer Variablen  $\exists \tau \in [0, 1]$ , sodass

$$F(1) - F(0) = F(\tau) = df(\gamma(\tau))(b - a).$$

Somit leistet der Punkt  $\xi := \gamma(\tau)$  das Gewünschte.



# Folgerung aus dem Mittelwertsatz

## Korollar 8.10 (Folgerung aus dem Mittelwertsatz)

Sei  $U \subset \mathbb{R}^n$  eine zusammenhängende offene Menge. Hat eine Funktion  $f : U \rightarrow \mathbb{C}$  überall die Ableitung 0, so ist sie konstant.

**Beweis:** Es genügt, die Behauptung für ein reelles  $f$  zu zeigen. Seien  $a, b$  beliebige Punkte in  $U$ . Dazu wähle nun Punkte  $a_0 := a, a_1, \dots, a_k = b$  derart, dass die Strecken  $[a_{i-1}, a_i]$  in  $U$  liegen. Anwendung des MWS bei jeder Strecke ergibt wegen  $f' = 0$ :  $f(a) = f(a_1), \dots, f(a_{k-1}) = f(b)$ . □

# Schranksatz

## Theorem 8.11 (Schranksatz)

Eine  $\mathcal{C}^1$ -Funktion  $f : U \rightarrow \mathbb{C}$  auf einer offenen Menge  $U$  ist auf jeder kompakten konvexen Teilmenge  $K \subset U$  Lipschitz-stetig. D.h. mit

$$\|f'\|_K := \max_{\xi \in K} \|f'(\xi)\|_{1,K} = \max_{\xi \in K} (|\partial_1 f(\xi)| + \dots + |\partial_n f(\xi)|)$$

gilt für beliebige  $x, y \in K$

$$|f(x) - f(y)| \leq \|f'\|_K \cdot \|y - x\|_\infty.$$

# Beweis des Schrankensatzes

**Beweis:** Mit  $x, y \in K$  liegt auch die Strecke  $[x, y]$  in  $K$ . Folglich ist die Funktion  $F := f \circ \gamma : [0, 1] \longrightarrow \mathbb{C}$  mit  $\gamma(t) = x + t(y - x)$  definiert. Nach dem Schrankensatz für Funktionen *einer* Veränderlichen gilt daher

$$|f(x) - f(y)| = |F(1) - F(0)| \leq \|\dot{F}\|_{[0,1]}.$$

Die Kettenregel ergibt die Ableitung

$$|\dot{F}(t)| \leq \sum_{i=1}^n |\partial_i f(\gamma(t))| \cdot |y_i - x_i|.$$

Damit folgt die Behauptung. □

# Höhere Ableitungen

## Definition 8.6 (Höhere Ableitungen)

Die partiellen Ableitungen  $\partial_1 f, \dots, \partial_n f$  einer Funktion können ihrerseits partiell differenzierbar sein. Darum heißen die Funktionen

$$\partial_{ij} f := \partial_i(\partial_j f)$$

*partielle Ableitungen 2. Ordnung* von  $f$ .

Weitere Bezeichnungen sind  $f_{x_j x_i}$  und  $\frac{\partial^2 f}{\partial x_j \partial x_i}$ .

# Höhere Ableitungen

## Beispiel ( $f(x, y) = x^y$ )

Betrachte die Funktion  $f(x, y) := x^y$  auf  $\mathbb{R}_+ \times \mathbb{R}$ . Die partiellen Ableitungen 1. Ordnung sind

$$f_x(x, y) = yx^{y-1}, \quad f_y(x, y) = x^y \ln x$$

und die partiellen Ableitungen 2. Ordnung

$$\begin{aligned} f_{xx}(x, y) &= y(y-1)x^{y-2} & f_{xy}(x, y) &= x^{y-1}(1 + y \ln x) \\ f_{yx}(x, y) &= x^{y-1}(1 + y \ln x) & f_{yy}(x, y) &= x^y (\ln x)^2. \end{aligned}$$

## Bemerkung 8.7 ()

In diesem Beispiel ist  $f_{xy} = f_{yx}$ . Im Allgemeinen ist jedoch  $\partial_{ij}f \neq \partial_{ji}f$ . Es kann sogar vorkommen, dass nur eine der partiellen Ableitungen  $\partial_{ij}f$  oder  $\partial_{ji}f$  existiert. Dies ist jedoch nicht der Fall, wenn eine der partiellen Ableitungen  $\partial_{ij}f$  oder  $\partial_{ji}f$  stetig ist.

# Satz von Schwarz

## Satz 8.12 (Schwarz)

Die Funktion  $f$  besitze in einer Umgebung von  $a \in \mathbb{R}^n$  die partiellen Ableitungen  $\partial_i f$ ,  $\partial_j f$  und  $\partial_{ji} f$ . Ferner sei  $\partial_{ji} f$  in  $a$  stetig. Dann existiert auch  $\partial_{ij} f(a)$  und es gilt

$$\partial_{ij} f(a) = \partial_{ji} f(a)$$

Für den Beweis des Satzes benötigen wir das folgende Lemma. (Wir verwenden im Beweis ein Analogon des MWS.)

Sei  $Q \subset \mathbb{R}^2$  ein Rechteck mit den Ecken  $(a, b), (a + h, b + k)$ ,  $h, k \neq 0$ ,  $\varphi : Q \rightarrow \mathbb{R}$ . Für  $\varphi$  sei

$$D_{Q\varphi} := \varphi(a + h, b + k) - \varphi(a + h, b) - \varphi(a, b + k) + \varphi(a, b).$$

# Hilfslemma für den Beweis des Satzes von Schwarz

## Lemma 8.13 ()

Sei  $\varphi$  reell und besitze auf  $Q$  die partiellen Ableitungen  $Q_1\varphi$  und  $Q_2\varphi$ . Dann existiert ein Tupel  $(\xi, \eta) \in Q$  mit

$$D_{Q\varphi} = hk \cdot \partial_{21}\varphi(\xi, \eta).$$

### Beweis: (des Lemmas)

$$u(x) := \varphi(x, b + k) - \varphi(x, b)$$

$$D_{Q\varphi} = u(a + h) - u(a) = hu'(\xi)$$

$$= h(\partial_1\varphi(\xi, b + k) - \partial_1\varphi(\xi, b)) = hk\partial_{21}\varphi(\xi, \eta)$$



Nun können wir den Satz von Schwarz beweisen:

# Beweis des Satzes von Schwarz

**Beweis:** Es genügt, ein reelles  $f$  zu betrachten. Man setze für  $(x, y)$  aus einer Umgebung  $V \subset \mathbb{R}^2$  von  $(0, 0)$

$$\varphi(x, y) := f(a + x\mathbf{e}_i + y\mathbf{e}_j).$$

Bei geeigneter Wahl von  $V$  existieren die partiellen Ableitungen  $\partial_1 \varphi, \partial_2 \varphi$  und  $\partial_{21} \varphi$ . Ferner ist  $\partial_{21} \varphi$  im Punkt  $(0, 0)$  stetig.

Es ist zu zeigen:  $\partial_{12} \varphi$  existiert in  $(0, 0)$  und  $\partial_{12} \varphi(0, 0) = \partial_{21} \varphi(0, 0)$  (\*).

Sei dazu  $\varepsilon > 0$ . Man wähle eine Umgebung  $V' \subset V$  von  $(0, 0)$ , sodass für  $(x, y) \in V'$  die Abschätzung  $|\partial_{21} \varphi(x, y) - \partial_{21} \varphi(0, 0)| < \varepsilon$  gilt und weiter wähle man ein achsenparalleles Rechteck  $Q \subset V'$  mit den gegenüberliegenden Ecken  $(0, 0)$  und  $(h, k)$ ,  $(h, k) \neq 0$ .

# Beweis des Satzes von Schwarz

Nach dem obigen Lemma ist dann

$$\left| \frac{D_Q \varphi}{h \cdot k} - \partial_{21} \varphi(0, 0) \right| < \varepsilon.$$

Wegen

$$\frac{D_Q \varphi}{h \cdot k} = \frac{1}{h} \left( \frac{\varphi(h, k) - \varphi(h, 0)}{k} - \frac{\varphi(0, k) - \varphi(0, 0)}{k} \right)$$

folgt mit  $k \rightarrow 0$

$$\left| \frac{\partial_2 \varphi(h, 0) - \partial_2 \varphi(0, 0)}{h} - \partial_{21} \varphi(0, 0) \right| \leq \varepsilon$$

für alle hinreichend kleinen  $|h| \neq 0$ . Damit ist (\*) bewiesen. □

# Stetige Differenzierbarkeit für höhere Ableitungen

## Definition 8.7 (*k*-mal stetig differenzierbar)

Sei  $U \subset \mathbb{R}^n$  offen. Eine Funktion  $f : U \longrightarrow \mathbb{C}$  heißt *k-mal stetig differenzierbar* oder auch  $\mathcal{C}^k$ -Funktion,  $k \geq 1$ , wenn alle partiellen Ableitungen  $\partial_{i_1} f, \dots, \partial_{i_k} f$  *k*-ter Ordnung auf  $U$  existieren und stetig sind. Den Vektorraum der  $\mathcal{C}^k$ -Funktionen auf  $U$  bezeichnet man mit

$$\mathcal{C}^k(U).$$

Auf Grund des Satzes von Schwarz (Satz 8.12) spielt bei einer  $\mathcal{C}^k$ -Funktion die Reihenfolge der partiellen Ableitungen  $\partial_{i_1} f, \dots, \partial_{i_k} f$  keine Rolle. Schließlich definiert man

$$\mathcal{C}^\infty(U) := \bigcap_k \mathcal{C}^k(U).$$

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- **Differentiale höherer Ordnung**
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Herleitung höherer Differentiale

Auf Grund des Satzes von Schwarz (Satz 8.12) kann man einer in einer Umgebung eines Punktes  $a \in \mathbb{R}^n$   $p$ -mal stetig differenzierbaren Funktion  $f$  in Verallgemeinerung des Differentials eine symmetrische,  $p$ -fach lineare Abbildung

$$d^p f(a) : \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_p \longrightarrow \mathbb{C}$$

zuordnen.

Wir betrachten zunächst die Fälle  $p = 1$  und  $p = 2$ :

# Differentiale 1. und 2. Ordnung

$$p = 1 : \quad a \mapsto df(a) : \mathbb{R}^n \longrightarrow \mathbb{C}$$

$$df(a)u = (\partial_1 f(a), \dots, \partial_n f(a)) \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

$$p = 2 : \quad (u, v) \in \mathbb{R}^n \times \mathbb{R}^n$$

$$d^{(2)}f(a)(u, v) := \partial_u(\partial_v f)(a) \quad (8.14)$$

# Differentiale 1. und 2. Ordnung

Diese Definition ist aus folgenden Gründen sinnvoll: Es gilt

$v = (v_1, \dots, v_n) \Rightarrow \partial_v f(x) = \sum_{i=1}^n \partial_i f(x) v_i$ . Die Funktion  $\partial_v f$  ist in einer

Umgebung von  $a$  stetig differenzierbar, da die Summanden  $\partial_1 f, \dots, \partial_n f$  diese Eigenschaft haben.

$\partial_v f$  besitzt also Richtungsableitungen und es gilt

$$\partial_u(\partial_v f(a)) = \sum_{i,j=1}^n \partial_{ij} f(a) u_i v_j. \quad (8.15)$$

$(u, v) \mapsto \partial_u \partial_v f(a)$  ist linear in jeder Variablen  $u, v$  und nach dem Satz von Schwarz (Satz 8.12) symmetrisch.

# Definition: Differential 2. Ordnung

## Definition 8.8 (Differential zweiter Ordnung)

Der Ausdruck in (8.14) bzw. in seiner Variante (8.15) heißt *Differential zweiter Ordnung von  $f$  in  $a$* .

Bezüglich der Standardbasis des  $\mathbb{R}^n$  besitzt das Differential zweiter Ordnung folgende Matrixdarstellung:

$$f''(a) = H_f(a) = \begin{pmatrix} \partial_{11}f(a) & \dots & \partial_{1n}f(a) \\ \vdots & & \vdots \\ \partial_{n1}f(a) & \dots & \partial_{nn}f(a) \end{pmatrix} \quad (8.16)$$

Für diese Matrix gilt  $d^2f(a)(u, v) = u^t f''(a)v$ .

## Definition 8.9 (Hesse-Matrix)

Die Matrix in (8.16) heißt *Hesse-Matrix* oder zweite Ableitung von  $f$  in  $a$ .

# Definition: Differentiale höherer Ordnung

Analog zur obigen Darstellung definiert man Differentiale höherer Ordnung:

## Definition 8.10 (Differentiale höherer Ordnung)

Für beliebige  $p \geq 1$  definiert man  $d^p f(a)$  wie folgt:

$$\partial^p f(a)(v^1, \dots, v^p) := \partial_{v^1}, \dots, \partial_{v^p} f(a). \quad (8.17)$$

Die dadurch erklärte Abbildung  $d^p f(a)$  ist invariant gegen Vertauschung der Variablen  $v^1, \dots, v^p$  und linear in jeder einzelnen Variablen. Sie hat die Darstellung

$$d^p f(a)(v^1, \dots, v^p) = \sum_{i_1=1}^n \dots \sum_{i_p=1}^n \partial_{i_1} \dots \partial_{i_p} f(a) v_{i_1}^1 \dots v_{i_p}^p. \quad (8.18)$$

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- **Die Taylor-Approximation**
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Die Taylor-Approximation

Sei  $f : U \rightarrow \mathbb{R}$  eine  $\mathcal{C}^{p+1}$ -Funktion auf einer offenen Menge  $U \subset \mathbb{R}^n$ . Weiter seien  $a, x \in U$  Punkte, deren Verbindungsstrecke in  $U$  liegt. (Wir führen dies auf den 1-dimensionalen Fall zurück.) Wir betrachten die Funktion  $F : [0, 1] \rightarrow \mathbb{R}$ ,

$$F(t) := f(a + th), \quad h := x - a.$$

Es gilt  $f(a) = F(0)$ ,  $f(x) = F(1)$ .  $F$  ist eine  $\mathcal{C}^{p+1}$ -Funktion auf  $[0, 1]$ . Nach der Taylorformel für Funktionen einer Veränderlichen gilt somit

$$F(1) = F(0) + F'(0) + \frac{1}{2!}F''(0) + \dots + \frac{1}{p!}F^{(p)}(0) + R_{p+1},$$

wobei das Restglied nach Lagrange mit einem  $\tau \in [0, 1]$  in der Form

$$R_{p+1} = \frac{1}{(p+1)!}F^{(p+1)}(\tau)$$

dargestellt werden kann.

# Die Taylor-Approximation

Die Ableitungen  $F^{(k)}$  berechnen wir durch wiederholte Anwendung der Kettenregel:

$$F'(t) = \sum_{i=1}^n \partial_i f(a + th) \cdot h_i$$

$$F''(t) = \sum_{i=1}^n \sum_{j=1}^n \partial_j \partial_i f(a + th) \cdot h_i h_j$$

 $\vdots$ 

$$F^{(p)}(t) = \sum_{i_1=1}^n \dots \sum_{i_p=1}^n \partial_{i_1} \dots \partial_{i_p} f(a + th) \cdot h_{i_1} \dots h_{i_p}$$

# Die Taylor-Approximation

Wir stellen  $F^{(k)}(t)$  mit Hilfe des Differentials  $d^{(k)}f(a)$  dar. Dazu führen wir allgemein für einen Vektor  $x \in \mathbb{R}^n$  folgende Bezeichnung ein:

$$d^{(k)}f(a)x^k := d^{(k)}f(a)\underbrace{(x, \dots, x)}_{k-\text{mal}},$$

Dies wird nach (8.18) komponentenweise zu

$$d^{(k)}f(a)x^k = \sum_{i_1=1}^n \dots \sum_{i_k=1}^n \partial_{i_1} \dots \partial_{i_k} f(a) x_{i_1} \dots x_{i_k}.$$

$d^{(k)}f(a)x^k$  ist ein homogenes Polynom vom Grad  $k$ . Damit gilt

$$F^{(k)}(t) = d^{(k)}f(a + th)h^k.$$

Schließlich setzen wir noch  $d^{(0)}f(a)x^0 := f(a)$  und wir können die Taylorapproximation wie folgt definieren.

# Die Taylor-Approximation

## Definition 8.11 (Taylorapproximation)

Die *Taylorapproximation der Ordnung p von f in a* ist

$$T_p f(x; a) := \sum_{k=0}^p \frac{1}{k!} d^{(k)} f(a) (x - a)^k.$$

## Bemerkung 8.8 ()

$T_1 f(x; a)$  ist die bereits in (8.7) eingeführte lineare Approximation.

# Taylorformel (mit Rest)

## Satz 8.14 (Taylorformel mit Rest)

Es sei  $f : U \rightarrow \mathbb{R}$  eine  $\mathcal{C}^{p+1}$ -Funktion. Sind  $a, x \in U$  Punkte, deren Verbindungsstrecke in  $U$  liegt, so gilt

$$f(x) = T_p f(x; a) + R_{p+1}(x; a),$$

wobei das Restglied mit einem geeigneten Punkt  $\xi \in [a, x]$  in der Form

$$R_{p+1}(x; a) = \frac{1}{(p+1)!} d^{p+1} f(\xi)(x-a)^{p+1}$$

dargestellt werden kann.

# Qualitative Taylorformel

## Korollar 8.15 (Qualitative Taylorformel)

Ist  $f : U \rightarrow \mathbb{R}$  eine  $\mathcal{C}^p$ -Funktion, so gilt an jeder Stelle  $a \in U$  für  $x \rightarrow a$

$$f(x) = T_p f(x; a) + o(\|x - a\|^p),$$

d.h. es gilt

$$\lim_{x \rightarrow a} \frac{f(x) - T_p f(x; a)}{\|x - a\|^p} = 0.$$

## Bemerkung 8.9 ()

$T_p f(x; a)$  stellt also ein Polynom eines Grades  $\leq p$  dar, welches  $f$  in der Nähe von  $a$  derart gut approximiert, dass der Fehler  $f(x) - T_p f(x; a)$  für  $x \rightarrow a$  schneller gegen Null geht als  $\|x - a\|^p$ .

**Beweis:** Zu  $\varepsilon > 0$  wähle man eine Kugel  $K_r(a) \subset U$  so, dass für  $y \in K_r(a)$

$$\frac{1}{p!} \sum_{i_1=1}^n \dots \sum_{i_p=1}^n |\partial_{i_1} \dots \partial_{i_p} f(y) - \partial_{i_1} \dots \partial_{i_p} f(a)| < \varepsilon$$

gilt (Stetigkeit von  $d^p f$ ). Für jeden Vektor  $h \in \mathbb{R}^n$  erhält man dann wegen  
 $|h_{i_1} \cdot \dots \cdot h_{i_p}| \leq \|h\|_\infty^p$

$$\left| \frac{1}{p!} \left( d^{(p)} f(y) - d^{(p)} f(a) \right) h^p \right| \leq \varepsilon \|h\|_\infty^p.$$

Zu jedem  $x \in K_r(a)$  wähle man nun weiter einen Punkt  $\xi \in [a, x[$ , mit dem die Taylorformel mit Rest gilt:

$$\begin{aligned} f(x) &= T_{p-1} f(x; a) + \frac{1}{p!} d^p f(\xi) (x - a)^p \\ &= T_p f(x; a) + \frac{1}{p!} (d^p f(\xi) - d^p f(a)) (x - a)^p, \end{aligned}$$

d.h.  $|f(x) - T_p f(x; a)| = \left| \frac{1}{p!} (d^p f(\xi) - d^p f(a)) (x - a) \right| \leq \varepsilon \|x - a\|_\infty^p.$

# Beispiele für Taylorpolynome

## Beispiel (Taylorpolynom 2. Ordnung)

Mit  $df(a)h = f'(a)h$  und  $d^2f(a)h = h^t f''(a)h$  ergibt sich:

$$\begin{aligned}T_2 f(x; a) &= f(a) + f'(a)(x - a) + \frac{1}{2}(x - a)^t f''(a)(x - a) \\&= f(a) + \sum_{i=1}^n \partial_i f(a)(x_i - a_i) + \frac{1}{2} \sum_{i,j=1}^n \partial_{ij} f(a)(x_i - a_i)(x_j - a_j)\end{aligned}$$

Beispiel ( $f(x, y) = x^y$ )

Sei  $f(x, y) = x^y$ . Gesucht:  $T_p f((x, y); (1, 1))$ .

Dazu bestimmen wir zunächst die partiellen Ableitungen erster und zweiter Ordnung:

$$f_x(x, y) = yx^{y-1}$$

$$f_y(x, y) = x^y \ln x$$

$$f_{xx}(x, y) = y(y-1)x^{y-2}$$

$$f_{yy}(x, y) = x^y (\ln x)^2$$

$$f_{xy}(x, y) = f_{yx}(x, y) = x^{y-1}(1 + \ln x)$$

Auswertungen der Ableitungen im Punkt  $(1, 1)$  ergeben:

$$f(1, 1) = 1, \quad f'(1, 1) (f_x(1, 1), f_y(1, 1)) = (1, 0), \quad f''(1, 1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Damit ergibt sich:

$$\begin{aligned} T_p f((x, y); (1, 1)) &= 1 + (1, 0) \begin{pmatrix} x-1 \\ y-1 \end{pmatrix} + \underbrace{(x-1, y-1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x-1 \\ y-1 \end{pmatrix}}_{(y-1, x-1)} \\ &= 1 + (x-1) + (y-1)(x-1) \end{aligned}$$

# Taylorreihen

## Definition 8.12 (Taylorreihen)

Es sei  $f \in \mathcal{C}^\infty(U)$ . Dann heißt die Reihe

$$\sum_{k=0}^n \frac{1}{k!} d^{(k)} f(a) (x-a)^k$$

*Taylorreihe von  $f$  im Punkt  $a \in U$ .*

Die Reihe konvergiert genau dann gegen  $f(x)$ , wenn  $\lim_{k \rightarrow \infty} R_k(x; a) = 0$ .

$f$  heißt *reell-analytisch* in  $U$ , wenn jeder Punkt  $a \in U$  eine Umgebung hat, in der  $f$  durch die Taylorreihe in  $a$  dargestellt wird.

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Schmiegequadrik

## Definition 8.13 (Schmiegequadrik)

Sei  $f$  eine reelle  $\mathcal{C}^2$ -Funktion in einer Umgebung von  $a \in \mathbb{R}^n$ . Ist  $f''(a)$  nicht die Nullmatrix, so beschreibt die quadratische Gleichung

$$x_{n+1} = T_2 f(x; a) = f(a) + f'(a)(x - a) + \frac{1}{2}(x - a)^t f''(a)(x - a)$$

eine Quadrik im  $\mathbb{R}^{n+1}$ . Diese heißt wegen  $f(x) - T_2 f(x; a) = o(\|x - a\|^2)$  *Schmiegequadrik* an den Graphen von  $f$  in  $(a, f(a))$ .

# Schmiegequadrik

Die Schmiegequadrik hat im Punkt  $(a, f(a))$  dieselbe Tangentialhyperebene wie der Graph und auch dieselbe Krümmung, wie in der Differentialgeometrie gezeigt wird.

Im Fall  $n = 2$  kann man durch Koordinatentransformation jede Schmiegequadrik in eine der folgenden Normalformen bringen:

(E)  $z = \pm(x^2 + y^2)$  elliptisches Paraboloid

(H)  $z = x^2 - y^2$  hyperbolisches Paraboloid

(P)  $z = \pm x^2$  parabolischer Zylinder

Eine Transformation in eine der Formen (E), (H), (P) ist genau dann möglich, wenn die Hesse-Matrix  $f''(a)$  definit, indefinit bzw. singulär, aber  $\neq 0$  ist.

# Definitheit

## Bemerkung 8.10 (Definitheitskriterium)

Eine quadratische Form  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $Q(x) = x^t A x$ , und ihre darstellende Matrix  $A$  heißen

|                           |                                                      |            |
|---------------------------|------------------------------------------------------|------------|
| positiv definit, wenn     | $Q(x) > 0 \quad \forall x \neq 0$                    | $Q > 0$    |
| negativ definit, wenn     | $Q(x) < 0 \quad \forall x \neq 0$                    | $Q < 0$    |
| positiv semidefinit, wenn | $Q(x) \geq 0$                                        | $Q \geq 0$ |
| negativ semidefinit, wenn | $Q(x) \leq 0$                                        | $Q \leq 0$ |
| indefinit, wenn           | $Q$ sowohl positive als auch negative Werte annimmt. | $Q \geq 0$ |

Diese Fälle lassen sich wie folgt charakterisieren:

- $Q > 0 \iff$  alle EW sind  $> 0$
- $Q < 0 \iff$  alle EW sind  $< 0$
- $Q \geq 0 \iff$  alle EW sind  $\geq 0$
- $Q \leq 0 \iff$  alle EW sind  $\leq 0$
- $Q \geq 0 \iff Q$  hat positive und negative EW

**Bemerkung 8.11 ()**

Im Fall  $n = 2$  hat man folgendes Kriterium:

|                                                        |                 |                                          |
|--------------------------------------------------------|-----------------|------------------------------------------|
| $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist | positiv definit | $\Leftrightarrow \det A > 0$ und $a > 0$ |
|                                                        | negativ definit | $\Leftrightarrow \det A > 0$ und $a < 0$ |
|                                                        | semidefinit     | $\Leftrightarrow \det A \geqslant 0$     |
|                                                        | indefinit       | $\Leftrightarrow \det A < 0$             |

**Bemerkung 8.12 ()**

Sei  $f$  wieder eine reelle  $C^2$ -Funktion in einer Umgebung von  $a \in \mathbb{R}^n$ . Ihr Graph heißt im Punkt  $(a, f(a))$

|              |                                                             |
|--------------|-------------------------------------------------------------|
| elliptisch   | $\Leftrightarrow f''(a)$ ist (positiv oder negativ) definit |
| hyperbolisch | $\Leftrightarrow f''(a)$ ist nicht singulär und indefinit   |
| parabolisch  | $\Leftrightarrow f''(a)$ ist singulär und $\neq 0$          |

Ein hyperbolischer Punkt heißt auch *Sattelpunkt*.

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- **Lokale Minima und Maxima**

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Lokale Extrema

## Definition 8.14 (Lokale Extrema)

Sei  $f$  eine reelle Funktion auf  $X \subset \mathbb{R}^n$ . Man sagt  $f$  habe in  $a \in X$  ein *lokales Maximum* bzw. *Minimum*, wenn es in  $X$  eine Umgebung  $V$  von  $a$  gibt, sodass

$$f(x) \leq f(a) \quad \text{bzw.} \quad f(x) \geq f(a) \quad \forall x \in U$$

gilt. Kann  $V$  so gewählt werden, dass sogar

$$f(x) < f(a) \quad \text{bzw.} \quad f(x) > f(a) \quad \forall x \in U \setminus \{a\}$$

gilt, so heißt  $a$  *isoliertes Maximum* bzw. *Minimum*.

# Notwendiges Kriterium für Extrema

## Satz 8.16 (Notwendiges Kriterium)

Sei  $U \subset \mathbb{R}^n$  offen. Hat  $f : U \rightarrow \mathbb{R}$  in  $a$  ein lokales Extremum und ist  $f$  in  $a$  partiell differenzierbar, so gilt

$$\partial_1 f(a) = \dots = \partial_n f(a) = 0. \quad (8.19)$$

Ist  $f$  in  $a$  differenzierbar, so besagt (8.19)  $df(a) = 0$ .

**Beweis:** Die durch  $F(t) = f(a + t e_k)$  in einem hinreichend kleinem Intervall um  $0 \in \mathbb{R}$  erklärte Funktion hat in  $t = 0$  ein lokales Extremum. Also ist  $F'(0) = 0$  und damit folgt  $\partial_k f(a) = F'(0) = 0$ . □

# Stationarität

## Definition 8.15 (stationär)

Eine in  $a$  differenzierbare Funktion heißt *stationär in  $a$* , wenn  $df(a) = 0$ . Nach dem soeben bewiesenen Satz hat eine differenzierbare Funktion auf einer offenen Menge höchstens an stationären Stellen lokale Extrema.

# Hinreichendes Kriterium

## Satz 8.17 (Hinreichendes Kriterium)

Es sei  $U \subset \mathbb{R}^n$  eine offene Menge und  $f : U \rightarrow \mathbb{R}$  eine  $\mathcal{C}^2$ -Funktion mit  $f'(a) = 0$ . Dann gilt:

- $f''(a) > 0 \Leftrightarrow f$  hat in  $a$  ein isoliertes Minimum
- $f''(a) < 0 \Leftrightarrow f$  hat in  $a$  ein isoliertes Maximum
- $f''(a) \geq 0 \Leftrightarrow f$  hat in  $a$  kein lokales Extremum.

Im indefiniten Fall gibt es Geraden  $g_1, g_2$  durch den Punkt  $a$  derart, dass  $f|_{U \cap g_1}$  in  $a$  ein isoliertes Maximum und  $f|_{U \cap g_2}$  in  $a$  ein isoliertes Minimum besitzt.

**Beweis:** Sei zunächst  $f''(a) > 0$ . Wegen  $f'(a) = 0$  gilt nach der qualitativen Taylorformel (Korollar 8.15) für hinreichend kleine Vektoren  $h$

$$f(a + h) = f(a) + \frac{1}{2}h^t f''(a)h + R(h),$$

wobei  $R(h)/\|h\|^2 \rightarrow 0$  für  $h \rightarrow 0$ . Die Funktion  $h \mapsto h^t f''(a)h$  hat auf der Einheitssphäre  $\{x \mid \|x\| = 1\}$  wegen  $f''(a) > 0$  ein positives Minimum  $m$ . Da jeder Vektor  $h$  das  $\|h\|$ -fache eines Einheitsvektors ist, folgt für alle  $h$

$$h^t f''(a)h \geq m\|h\|^2.$$

Wir wählen nun eine Kugel  $K_\varepsilon(a) \subset U$  so klein, dass  $|R(h)| \leq 1/4m\|h\|^2$  für  $\|h\| < \varepsilon$  gilt. Für  $a + h \in K_\varepsilon(a)$  erhalten wir dann

$$f(a + h) \geq f(a) + \frac{m}{4}\|h\|^2.$$

Danach nimmt  $f$  innerhalb  $K_\varepsilon(a)$  genau im Punkt  $a$  ein Minimum an. Im Fall  $f''(a) > 0$  ist die Behauptung damit bewiesen.  $f''(a) < 0$  wird durch den Übergang zu  $-f$  analog behandelt.

Sei schließlich  $f''(a)$  indefinit. Wir wählen Vektoren  $v$  und  $w$  mit  $v^t f''(a)v > 0$  bzw.  $w^t f''(a)w < 0$  und betrachten die Funktionen

$$\begin{aligned} F_v(t) &= f(a + tv), \\ F_w(t) &= f(a + tw), \end{aligned}$$

die in geeigneten Intervallen um  $0 \in \mathbb{R}$  definiert sind. Ihre ersten und zweiten Ableitungen in 0 sind

$$\begin{array}{ll} F'_v(0) = f'(a)v = 0 & F''_v(0) = v^t f''(a)v > 0, \\ F'_w(0) = f'(a)w = 0 & F''_w(0) = w^t f''(a)w < 0. \end{array}$$

Somit hat  $F_v$  in 0 ein isoliertes lokales Minimum und  $F_w$  ein isoliertes lokales Maximum.  $f$  hat daher in  $a$  kein lokales Extremum. □

# Kapitel 9:

# Differenzierbare Abbildungen

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Differenzierbare Abbildungen: Vorbemerkung

Es seien  $X, Y$  endlich-dimensionale normierte Vektorräume über einem Körper  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  und  $f : U \rightarrow Y$  eine Abbildung auf einer offenen Menge  $U \subset X$ . Besonders wichtig ist dabei der Fall, in dem  $X = \mathbb{K}^n$  und  $Y = \mathbb{K}^m$ :

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : \mathbb{K}^n \supset U \rightarrow \mathbb{K}^m, \quad (\text{Standardfall}) \quad (9.1)$$

Weiter verwenden wir auf dem Vektorraum  $L(X, Y)$  der  $\mathbb{K}$ -linearen Abbildungen von  $X$  und  $Y$  die induzierte Operatornorm.

Die Endlichkeit der Dimensionen der Vektorräume impliziert, dass jede lineare Abbildung von  $X$  nach  $Y$  stetig ist und dass  $X, Y$  und  $L(X, Y)$  vollständig normierte Räume sind.

# Differenzierbarkeit von Abbildungen

## Definition 9.1 (Differenzierbarkeit)

$f : U \rightarrow Y$  heißt *differenzierbar im Punkt  $a \in U$* , genauer  $\mathbb{K}$ -differenzierbar, wenn es eine  $\mathbb{K}$ -lineare Abbildung  $L : X \rightarrow Y$  gibt derart, dass der durch

$$f(a + h) = f(a) + L(h) + R(h)$$

erklärte Rest  $R$  die Bedingung

$$\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0 \tag{9.2}$$

erfüllt.

## Bemerkung 9.1 (Analogon zu differenzierbaren Funktionen)

Wie für Funktionen zeigt man, dass es höchstens eine solche Abbildung  $L$  gibt. Diese heißt *Differential* oder *Linearisierung von  $f$  in  $a$*  und wird mit  $df(a)$  bezeichnet.

Bzgl. Basen in  $X$  und  $Y$  kann  $df(a)$  als Matrix dargestellt werden. Diese heißt *Funktionalmatrix* oder auch *Ableitung von  $f$  in  $a$*  (bzgl. der Basen) und wird mit  $f'(a)$  bezeichnet. Im Fall  $\dim X = \dim Y$  heißt die Determinante von  $f'(a)$  *Funktionaldeterminante von  $f$  in  $a$* .

### Beispiel ( $f(x) := Ax + b$ )

Eine affine Abbildung  $f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$

$$f(x) := Ax + b, \quad A \in \mathbb{K}^{m \times n}, b \in \mathbb{K}^m$$

ist in jedem Punkt differenzierbar. Ihre Ableitung ist die Matrix  $A$  und das Differential die durch  $h \mapsto Ah$  gegebene lineare Abbildung  $\mathbb{K}^n \longrightarrow \mathbb{K}^m$ :

$$f'(a) = A, \quad df(a)h = Ah \quad \forall h \in \mathbb{K}^n.$$

# Reduktionslemma

## Lemma 9.1 (Reduktionslemma)

Eine Abbildung  $f = (f_1, f_2) : U \longrightarrow Y_1 \times Y_2$  in eine direkte Summe ist genau dann differenzierbar im Punkt  $a \in U$ , wenn dort  $f_1 : U \longrightarrow Y_1$  und  $f_2 : U \longrightarrow Y_2$  differenzierbar sind. Gegebenenfalls ist

$$df(a) = (d_1 f(a), d_2 f(a)). \quad (9.3)$$

**Beweis:**  $f_1, f_2$  seien in  $a$  differenzierbar. Dann gilt für  $i = 1, 2$ :

$$f_i(a + h) = f_i(a) + df_i(a)h + R_i(h),$$

wobei  $R_i$  die Bedingung (9.2) erfüllt. Wir setzen  $Lh := (df_1(a)h, df_2(a)h)$ .  $L$  ist eine lineare Abbildung  $X \longrightarrow Y_1 \times Y_2$  und mit ihr gilt

$$f(a + h) = f(a) + L(h) + (R_1(h), R_2(h)).$$

$R(h) := (R_1(h), R_2(h))$  erfüllt die Bedingung (9.2). Also ist  $f$  differenzierbar und hat das Differential  $L$ . Analog zeigt man die Umkehrung.

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- **Funktionalmatrix**
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Funktionalmatrix

## Korollar 9.2 (Funktionalmatrix)

Die Abbildung (9.1) ist genau dann in  $a \in U$  differenzierbar, wenn dort jede der Komponentenfunktionen  $f_1, \dots, f_m$  differenzierbar ist. Gegebenenfalls gilt für  $h \in \mathbb{K}^n$

$$df(a)h = f'(a)h,$$

wobei die Funktionalmatrix  $f'(a)$  folgende Gestalt hat:

$$f'(a) = \begin{pmatrix} f'_1(a) \\ \vdots \\ f'_m(a) \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & & \vdots \\ \partial_1 f_m(a) & \dots & \partial_n f_m(a) \end{pmatrix} \quad (9.4)$$

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- **Differenzierbarkeitskriterium für Abbildungen**
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Definitionen

Im Folgenden werden wir für allgemeine Abbildungen eine Analogon zu differenzierbaren Funktionen schaffen.

Dazu werden die meisten Aussagen in ähnlicher Weise hergeleitet, wie dies für differenzierbare Funktionen gemacht wird.

## Theorem 9.3 (Differenzierbarkeitskriterium)

Eine Abbildung  $f = (f_1, \dots, f_m) : U \longrightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$ , ist in  $a \in U$   $\mathbb{R}$ -differenzierbar, wenn alle partiellen Ableitungen  $\partial_\nu f_\mu$ ,  $\nu = 1, \dots, n$ ,  $\mu = 1, \dots, m$  in einer Umgebung von  $a$  existieren und im Punkt  $a$  stetig sind.

Man definiert weiterhin die Begriffe *Differential*, *Richtungsableitungen* u.ä. in Anlehnung an die Definitionen entsprechenden Definitionen für differenzierbare Funktionen aus Kapitel 23.

# Differential, Richtungsableitungen

## Definition 9.2 (Richtungsableitungen)

Das Differential einer in  $a$  differenzierbaren Abbildung kann wie für Funktionen mit Hilfe von *Richtungsableitungen* berechnet werden. In Verallgemeinerung von (8.9) gilt

$$df(a)h = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t} =: \partial_h f(a).$$

$\partial_h f(a)$  heißt *Ableitung von  $f$  in Richtung  $h$  im Punkt  $a$* . Die Ableitungen in Richtung einer fest gewählten Basis  $e_1, \dots, e_n$  für  $X$  heißen die partiellen Ableitungen bzgl. der Basis und werden wieder mit  $\partial_1 f(a), \dots, \partial_n f(a)$  bezeichnet.

# Funktionalmatrix und Richtungsableitungen

## Definition 9.3 (Funktionalmatrix und Richtungsableitungen)

Mit der Funktionalmatrix  $f'(a)$  bzgl. der Basen in  $X$  und  $Y$  hat man für die Richtungsableitungen die Darstellung

$$\partial_h f(a) = f'(a)h.$$

Insbesondere ist im Standardfall  $\partial_\nu f(a) = f'(a)e_\nu$  gleich der  $\nu$ -ten Spalte in der Funktionalmatrix:

$$f'(a) = (\partial_1 f(a), \dots, \partial_n f(a)) \quad \text{mit} \quad \partial_\nu f(a) = \begin{pmatrix} \partial_\nu f_1(a) \\ \vdots \\ \partial_\nu f_m(a) \end{pmatrix}.$$

# Stetig differenzierbare Abbildungen

## Definition 9.4 (stetig differenzierbar)

Eine differenzierbare Abbildung  $f : U \rightarrow Y$  auf einer offenen Menge  $U \subset X$  heißt *stetig differenzierbar in  $U$* , wenn ihr Differential  $df : U \rightarrow L(X, Y)$ ,  $x \mapsto df(x)$ , stetig ist.

## Bemerkung 9.2 (Stetigkeitstest)

$df : U \rightarrow L(X, Y)$  ist genau dann stetig, wenn für jeden Vektor  $h \in X$  die Abbildung  $U \rightarrow Y$ ,  $x \mapsto df(x)h$ , stetig ist.

### Bemerkung 9.3 (Ergänzung zum Reduktionslemma)

Das Reduktionslemma kann wie folgt ergänzt werden: Eine Abbildung  $f : U \longrightarrow Y_1 \times Y_2$  in eine direkte Summe ist genau dann stetig differenzierbar, wenn ihre beiden Komponenten  $f_i : U \longrightarrow Y_i$ ,  $i = 1, 2$ , stetig differenzierbar sind.

Für den Standardfall impliziert diese Ergänzung: Die Abbildung (9.1) ist genau dann stetig differenzierbar, wenn alle Komponentenfunktionen  $f_1, \dots, f_m$  stetig differenzierbar sind.

### Beispiel (Polarkoordinatenabbildung)

Wir betrachten die Polarkoordinatenabbildung  $P_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ ,

$$(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}:$$

$$P'_2(r, \varphi) = \begin{pmatrix} \partial_1 f_1 & \partial_2 f_1 \\ \partial_1 f_2 & \partial_2 f_2 \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

# $k$ -mal stetig differenzierbare Abbildungen

## Definition 9.5 ( $k$ -mal stetig differenzierbar)

Eine Abbildung  $f = f_1, \dots, f_m : U \longrightarrow \mathbb{K}^n$ ,  $U \subset \mathbb{K}^m$ , heißt  $k$ -mal stetig differenzierbar in  $U$ , wenn alle Komponentenfunktionen  $f_1, \dots, f_m$   $k$ -mal stetig differenzierbar sind.

Den Raum der  $k$ -mal stetig differenzierbaren Abbildungen  $U \longrightarrow \mathbb{K}^m$  bezeichnet man mit

$$\mathcal{C}^k(U, \mathbb{K}^m).$$

Man setzt außerdem

$$\mathcal{C}^\infty(U, \mathbb{K}^m) := \bigcap_{k=1}^{\infty} \mathcal{C}^k(U, \mathbb{K}^m).$$

# Kettenregel für differenzierbare Abbildungen

## Satz 9.4 (Kettenregel)

$X, Y, Z$  seien normierte Vektorräume und  $V$  offen in  $X$ ,  $U$  offen in  $Y$ . In  $V \xrightarrow{g} U \xrightarrow{f} Z$  sei  $g$  differenzierbar in  $a$  und  $f$  differenzierbar in  $b := g(a)$ . Dann ist  $f \circ g$  differenzierbar in  $a$  und es gilt

$$d(f \circ g)(a) = df(b) \circ dg(a). \quad (9.5)$$

Für Ableitungen bedeutet das

$$(f \circ g)'(a) = f'(b) \cdot g'(a). \quad (9.6)$$

Sind  $f, g$  stetig differenzierbar, dann auch  $f \circ g$ .

**Beweis:** (Siehe Kapitel 3 in [?])



# Ein Beispiel: Tangentialvektoren

Eine differenzierbare Abbildung  $f : U \rightarrow \mathbb{K}^m$ ,  $U \subset \mathbb{K}^n$  offen, ordnet einer differenzierbaren Kurve  $\gamma : I \rightarrow U$  die Bildkurve

$$f \circ \gamma : I \rightarrow \mathbb{K}^m$$

zu. Diese ist nach der Kettenregel ebenfalls differenzierbar und hat für  $t_0 \in I$  den Tangentialvektor

$$\frac{d}{dt}(f \circ \gamma)(t) = df(\gamma(t_0))\dot{\gamma}(t_0) = f'(\gamma(t_0))\dot{\gamma}(t_0). \quad (9.7)$$

Tangentialvektoren werden also durch das Differential bzw. mittels Funktionalmatrix abgebildet.

Anwendung auf zu den Basisvektoren parallele Koordinaten:

Für  $a \in U$  sei  $\varepsilon_i(t) = a + te_i$ ,  $t$  aus einem Intervall um 0, sodass  $\varepsilon_i(t) \in U$ .

Die Bildkurve  $f \circ \varepsilon_i$  hat für  $t = 0$  im Punkt  $f(a)$  den Tangentialvektor  $f'(a) \cdot e_i$ . Dies ist gerade der  $i$ -te Spaltenvektor von  $f'(a)$ . Die Kurven  $f \circ \varepsilon_1, \dots, f \circ \varepsilon_n$  heißen die von  $f$  erzeugten Koordinatenlinien durch  $f(a)$ .

Betrachte die Polarkoordinatenabbildung  $P_2 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$(r, \varphi) \mapsto \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$ . Diese bildet die Gerade  $g_{\varphi_0} : r \mapsto (r, \varphi_0)$  auf die

Gerade durch den Nullpunkt ab und die Gerade  $\tilde{g}_{r_0} : \varphi \mapsto (r_0, \varphi)$  auf Kreise durch den Nullpunkt. Die Spalten der Funktionalmatrix

$$P'_2(r_0, \varphi_0) = \begin{pmatrix} \cos \varphi_0 & -r_0 \sin \varphi_0 \\ \sin \varphi_0 & r_0 \cos \varphi_0 \end{pmatrix}$$

sind im Bildpunkt  $P_2(r_0, \varphi_0)$  Tangentialvektoren an  $P_2 \circ g_{\varphi_0}$  bzw.  $P_2 \circ \tilde{g}_{r_0}$ .

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Motivation: Extrema unter Nebenbedingungen

Gegeben sei eine Funktion  $f : U \rightarrow \mathbb{R}$  und weitere Funktionen  $\varphi_1, \dots, \varphi_k : U \rightarrow \mathbb{R}$  auf einer Menge  $U \subset \mathbb{R}^n$ . Sei  $M$  die Nullstellenmenge von  $\varphi = (\varphi_1, \dots, \varphi_k) : U \rightarrow \mathbb{R}^k$ :

$$M = \{x \in U | \varphi(x) = 0\}. \quad (9.8)$$

Gesucht werden Punkte  $x_0 \in M$  mit  $f(x) \leq f(x_0) \forall x \in M$  oder  $f(x) \geq f(x_0) \forall x \in M$ . Solche Punkte heißen *Maximal-* bzw. *Minimalpunkte von  $f$  auf  $M$*  oder auch *unter der Nebenbedingung  $\varphi = 0$* .

Das führt uns zu folgender notwendigen Bedingung für Maxima und Minima, falls  $M$  eine Mannigfaltigkeit ist:

# Lagrange-Multiplikatoren

## Satz 9.5 (Multiplikatorenregel von Lagrange)

$f$  und  $\varphi = (\varphi_1, \dots, \varphi_k)$  seien stetig differenzierbar auf einer offenen Menge  $U \subset \mathbb{R}^n$ .  $\varphi'(x)$  habe in jedem Punkt  $x \in M$  den Rang  $k$ . Dann gilt:

Ist  $x_0 \in M$  ein Extrempunkt von  $f$  auf  $M$ , so ist  $f'(x_0)$  eine Linearkombination von  $\varphi'_1(x_0), \dots, \varphi'_k(x_0)$ : Es gibt Zahlen  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ , sogenannte *Lagrange-Multiplikatoren*, mit

$$f'(x_0) = \sum_{i=1}^k \lambda_i \varphi'_i(x_0). \quad (9.9)$$

Im euklidischen  $\mathbb{R}^n$  bedeutet (9.9)

$$\text{grad} f(x_0) = \sum_{i=1}^k \lambda_i \text{grad} \varphi_i(x_0). \quad (9.10)$$

# Multiplikatorenregel von Lagrange

**Beweis:** (Hier skizzieren wir nur die Idee für den Beweis. Einen ausführlichen Beweis findet man im Kapitel 3.6 in [?])

Wir zeigen (9.10) und dass jeder Tangentialvektor  $v \in T_{x_0} M_0$  auf  $\text{grad}f(x)$  senkrecht steht. Zu  $v \in T_{x_0} M$   $\exists$  eine stetig differenzierbare Kurve  $\alpha : (-\varepsilon, \varepsilon) \rightarrow M$  mit  $\alpha(0) = x_0$  und  $\dot{\alpha}(0) = v$ . Die durch  $F(t) = f(\alpha(t))$  definierte Funktion  $F : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$  hat in  $t = 0$  ein lokales Extremum, d.h.

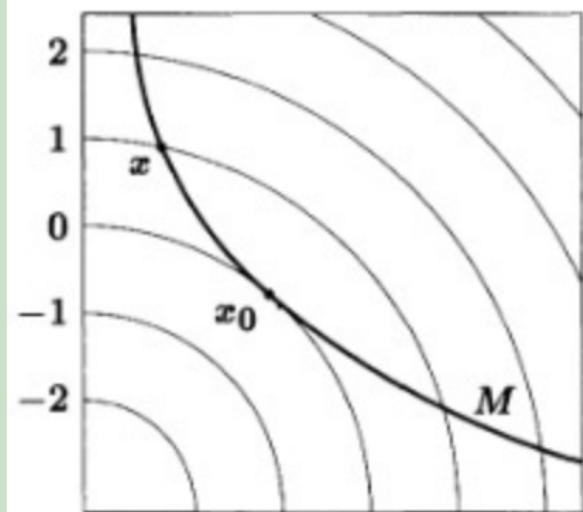
$$0 = \dot{F}(0) = f'(\alpha(0)) \cdot \dot{\alpha}(0) = \langle \text{grad}f(x_0), v \rangle = 0.$$



# Multiplikatorenregel und Niveaulinien

## Bemerkung 9.4 (Niveaulinien)

Betrachte den Fall  $f'(x_0) \neq 0$  und  $k = 1$ . In diesem Fall besagt die Multiplikatorenregel, dass sich die linearen Approximationen von  $f$  und  $\varphi$  im Punkt  $x_0$  berühren. Die Notwendigkeit ist leicht einzusehen (siehe Abbildung rechts):



## Beispiel (Punkt mit minimalem Abstand zu $(1, 0, 0)$ )

Gesucht ist der Punkt der Ebene  $z = x + y$ , der vom Punkt  $(1, 0, 0)$  den kleinsten euklidischen Abstand hat.

Zielfunktion:  $f(x, y, z) := (x - 1)^2 + y^2 + z^2$

Nebenbedingung:  $\varphi(x, y, z) = x + y - z = 0$

$$f'(x, y, z) = \lambda \varphi'(x, y, z)$$

$$(\partial_1 f, \partial_2 f, \partial_3 f)(x, y, z) = (\partial_1 \varphi, \partial_2 \varphi, \partial_3 \varphi)(x, y, z)$$

$\varphi$  hat vollen Rang  $\forall x$ .

Wir lösen nun das Gleichungssystem (siehe unten) und erhalten

$$(x, y, z) = \left( \frac{2}{3}, -\frac{1}{3}, \frac{1}{3} \right).$$

$\left( \frac{2}{3}, -\frac{1}{3}, \frac{1}{3} \right)$  ist der einzige Punkt, der für den minimalen Abstand in Frage kommt. Somit ist  $\left( \frac{2}{3}, -\frac{1}{3}, \frac{1}{3} \right)$  die Lösung dieser Aufgabe.

# Vervollständigung des Beispiels

Lösen des Gleichungssystems:

$$\textcircled{1} \quad 2(x - 1) = \lambda$$

$$\textcircled{2} \quad 2y = \lambda$$

$$\textcircled{3} \quad 2z = -\lambda$$

$$\textcircled{4} \quad x + y - z = 0$$


---

$$\textcircled{2} + \textcircled{3} = \textcircled{5}: \quad z = -y$$

$$\textcircled{5} + \textcircled{4} = \textcircled{6}: \quad y = -\frac{x}{2}$$

$$\textcircled{1} + \textcircled{2} = \textcircled{7}: \quad -x = 2x - 2$$

$$\frac{2}{3} = x \Rightarrow y = -\frac{1}{3} \Rightarrow z = \frac{1}{3}$$

**Bemerkung:**

Der Vektor  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 2/3 \\ -1/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ -1/3 \end{pmatrix}$  steht senkrecht auf der Ebene  $x + y - z = 0$ .

# Kapitel 10:

## Nichtlineare Gleichungen

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

### ■ Konvergenzbegriffe

- Iterative Verfahren zur Lösung linearer Gleichungssysteme

### ■ Konvergenzgeschwindigkeit einer Folge

### ■ Nullstellenbestimmung reeller Funktionen

- Das eindimensionale Newtonverfahren
- Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Vorbemerkungen

$$F, \sigma : D(F) \subset \mathbb{C}^n \longrightarrow \mathbb{C}^n$$

- Durch die Transformation  $F(x) := \sigma(x) - y$  kann jede nichtlineare Gleichung  $\sigma(x) = y$  in eine Nullstellenaufgabe  $F(x) = 0$  überführt werden.
- Da nichtlineare Gleichungen in der Regel nicht geschlossen gelöst werden können, ihre Lösungen also nicht in endlich vielen Schritten berechenbar sind, kommen fast ausschließlich *Iterationsverfahren* zur Approximation der Lösung zur Anwendung.
- Für Iterationsverfahren im  $\mathbb{C}^n$  verwenden wir wieder die Notation  $x^{(k)}$  mit hochgestelltem, geklammerten Iterationsindex. Im Eindimensionalen verwenden wir einen tiefgestellten Iterationsindex.

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

### ■ Konvergenzbegriffe

- Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Berechnung der $\nu$ -ten Wurzel einer Zahl $a \in \mathbb{C} \setminus \{0\}$

Iterationsverfahren zur Berechnung der  $\nu$ -ten Wurzel einer Zahl  $a \in \mathbb{C} \setminus \{0\}$ :

$$x_{k+1} = \frac{1}{\nu} \left( (\nu - 1)x_k + \frac{a}{x_k^{\nu-1}} \right), \quad k = 0, 1, \dots \quad (10.1)$$

Wenn die Folge konvergiert, erfüllt der Grenzwert die Fixpunktgleichung

$$\nu x = (\nu - 1)x + \frac{a}{x^{\nu-1}} \quad \Leftrightarrow \quad x^\nu = a.$$

Unklar ist, ob diese Folge überhaupt konvergiert und wenn ja, gegen welche Wurzel sie konvergiert.

Für  $\nu = 2$  (*klassisches Heronverfahren*) sieht dies folgendermaßen aus:

$$x_{k+1} = \frac{1}{2} \left( x_k + \frac{a}{x_k} \right)$$

Nun führt die Transformation

$$z_k = \frac{x_k - \sqrt{a}}{x_k + \sqrt{a}} \quad (10.2)$$

auf die Rekursion

$$z_{k+1} = \frac{\frac{1}{2}x_k + \frac{a}{2x_k} - \sqrt{a}}{\frac{1}{2}x_k + \frac{a}{2x_k} + \sqrt{a}} = \frac{x_k^2 + a - 2\sqrt{a}x_k}{x_k^2 + a + 2\sqrt{a}x_k} = \frac{(x_k - \sqrt{a})^2}{(x_k + \sqrt{a})^2} = z_k^2. \quad (10.3)$$

Man bestimmt den folgenden Grenzwert:

$$\lim_{k \rightarrow \infty} z_k = \lim_{k \rightarrow \infty} z_0^{(2^k)} = \begin{cases} 0 & |z_0| < 1 \\ 1 & |z_0| = 1 \\ \infty & |z_0| > 1 \end{cases}$$

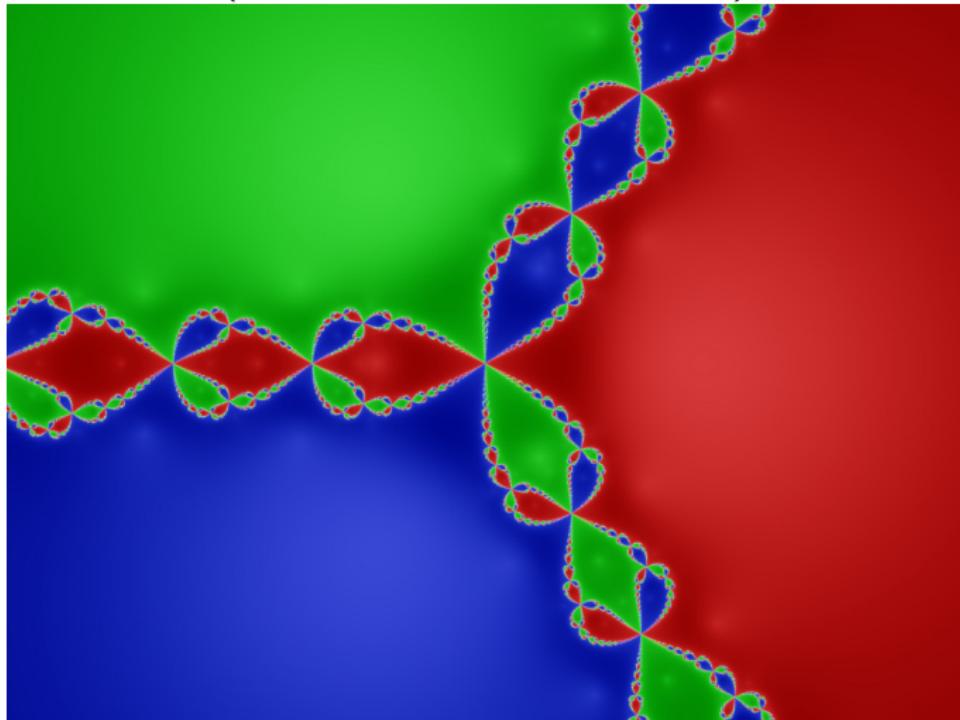
D.h.  $x_k$  konvergiert gegen  $\sqrt{a}$ , falls  $|z_0| < 1$  und gegen  $-\sqrt{a}$ , falls  $|z_0| > 1$ .  
 Für  $|z_0| = 1$  liegt keine Konvergenz vor:

$$\begin{aligned} |z| = \left| \frac{x_0 - \sqrt{a}}{x_0 + \sqrt{a}} \right| \geq 1 &\Leftrightarrow |x_0|^2 - 2\operatorname{Re}(\bar{x}_0)\sqrt{a} + |a| = |x_0 - \sqrt{a}|^2 \\ &\geq |x_0 + \sqrt{a}|^2 = |x_0|^2 + 2\operatorname{Re}(\bar{x}_0)\sqrt{a} + |a| \\ &\Leftrightarrow 0 \geq \operatorname{Re}(\bar{x}_0)\sqrt{a} \end{aligned}$$

Für positives  $a$  ergibt sich somit Konvergenz gegen diejenige Wurzel von  $a$ , die dasselbe Vorzeichen hat wie der Realteil von  $x_0$ .

Bereits für den Fall  $\nu = 3$  ergibt sich eine scheinbar chaotische Abhängigkeit vom Startwert  $x_0$ .

## (Bsp. in der komplexen Ebene)



$$f(z) = z^3 - 1$$

# Lokale Konvergenz

## Definition 10.1 (lokal konvergent)

Ein Iterationsverfahren  $x^{(k+1)} = \Phi(x^{(k)})$  mit  $\Phi : D(\Phi) \subset \mathbb{C}^n \longrightarrow \mathbb{C}^n$  heißt *lokal konvergent gegen  $\hat{x} \in \mathbb{C}^n$* , falls eine Umgebung  $U \subset D(\Phi)$  um  $\hat{x} \in U$  existiert, sodass  $\forall x^{(0)} \in U$  die resultierende Folge  $\{x^{(k)}\}$  gegen  $\hat{x}$  konvergiert. In diesem Fall spricht man von einem *anziehenden Fixpunkt  $\hat{x}$  von  $\Phi$* . Das Iterationsverfahren heißt *global konvergent*, wenn  $U$  der gesamte Raum  $\mathbb{C}^n$  ist.

## Bemerkung 10.1 (Konvergenz des Heronverfahrens)

Das Heronverfahren ( $\nu = 2, \nu = 3$ ) ist für jede  $\nu$ -te Wurzel von  $a \in \mathbb{C}$  lokal konvergent.

# Kontraktionen

Kontraktionen sind spezielle Abbildungen, die interessant sind, weil sie genau einen Fixpunkt haben, der sich iterativ approximieren lässt.

## Definition 10.2 (Kontraktion)

$\Phi : D \subset \mathbb{C}^n \rightarrow \mathbb{C}^n$  heißt *Kontraktion* wenn für alle  $x, y \in D$  gilt

$$\|\Phi(x) - \Phi(y)\| \leq q \cdot \|x - y\|$$

für ein festes  $q \in \mathbb{R}$  mit  $0 < q < 1$ . Der Faktor  $q$  heißt *Lipschitz-Konstante*.

## Beispiel (Affin lineare Kontraktion)

Die Abbildung  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  mit  $\Phi(x) := A \cdot x + b$  ist genau dann eine Kontraktion wenn für die induzierte Matrixnorm gilt  $\|A\| < 1$ , denn

$$\|\Phi(x) - \Phi(y)\| = \|A \cdot (x - y)\| \leq \|A\| \cdot \|x - y\|.$$

# Banachscher Fixpunktsatz

## Satz 10.1 (Banachscher Fixpunktsatz)

Sei  $D \subset \mathbb{C}^n$  und  $\Phi : D \rightarrow D$  eine Kontraktion mit Lipschitz-Konstante  $q$ .

Dann hat  $\Phi$  genau einen Fixpunkt  $\hat{x} \in D$ .

Für  $x^{(0)} \in D$  konvergiert die *Fixpunktiteration*  $x^{(k+1)} := \Phi(x^{(k)})$  gegen diesen Fixpunkt, d.h.

$$\hat{x} = \lim_{k \rightarrow \infty} x^{(k)}.$$

Für alle  $k \in \mathbb{N}$  gelten die folgenden Fehlerabschätzungen:

- 1** Monotonie:  $\|x^{(k)} - \hat{x}\| \leq q \cdot \|x^{(k-1)} - \hat{x}\|$
- 2** A-priori Schranke:  $\|x^{(k)} - \hat{x}\| \leq \frac{q^k}{1-q} \cdot \|x^{(1)} - x^{(0)}\|$
- 3** A-posteriori Schranke:  $\|x^{(k)} - \hat{x}\| \leq \frac{q}{1-q} \cdot \|x^{(k)} - x^{(k-1)}\|$

# Beweis des Fixpunktsatzes

**Beweis:** Per Induktion sieht man:

$$\|x^{(n+1)} - x^{(n)}\| \leq q^n \cdot \|x^{(1)} - x^{(0)}\|$$

Es folgt für  $k, p \in \mathbb{N}$ :

$$\begin{aligned} \|x^{(k+p)} - x^{(k)}\| &\leq \sum_{i=1}^p \|x^{(k+i)} - x^{(k+i-1)}\| \leq q^k \cdot \sum_{i=1}^p q^{i-1} \cdot \|x^{(1)} - x^{(0)}\| \\ &\leq \frac{q^k}{1-q} \cdot \|x^{(1)} - x^{(0)}\| \end{aligned} \tag{10.4}$$

Dies ist eine Nullfolge, d.h.  $x^{(k)}$  ist eine Cauchy-Folge. Somit existiert ein Grenzwert  $\hat{x} = \lim_{k \rightarrow \infty} x^{(k)}$ . Dieser ist ein Fixpunkt, denn

$$\Phi(\hat{x}) = \Phi\left(\lim_{k \rightarrow \infty} x^{(k)}\right) = \lim_{k \rightarrow \infty} \Phi(x^{(k)}) = \lim_{k \rightarrow \infty} x^{(k+1)} = \hat{x}.$$

# Beweis des Fixpunktsatzes

Für einen weiteren Fixpunkt  $\tilde{x} \in D$  gilt

$$\|\tilde{x} - \hat{x}\| = \|\Phi(\tilde{x}) - \Phi(\hat{x})\| \leq q \cdot \|\tilde{x} - \hat{x}\|,$$

d.h. es ist  $\|\tilde{x} - \hat{x}\| = 0$  und somit ist  $\hat{x} = \tilde{x}$  der einzige Fixpunkt von  $\Phi$ . Es bleibt die Abschätzungen zu zeigen:

$$\|x^{(k)} - \hat{x}\| = \|\Phi(x^{(k-1)}) - \Phi(\hat{x})\| \leq q \cdot \|x^{(k-1)} - \hat{x}\|$$

$$\|x^{(k)} - \hat{x}\| = \lim_{p \rightarrow \infty} \|x^{(k)} - x^{(k+p)}\| \stackrel{(10.4)}{\leq} \frac{q^k}{1-q} \cdot \|x^{(1)} - x^{(0)}\|$$

$$(1-q) \cdot \|x^{(k)} - \hat{x}\| \leq \|x^{(k)} - \hat{x}\| - \|x^{(k+1)} - \hat{x}\|$$

$$\leq \|x^{(k+1)} - x^{(k)}\| \leq q \cdot \|x^{(k)} - x^{(k-1)}\|$$



# Gleichungssysteme iterativ lösen

Wir zerlegen  $A$  in drei Matrizen  $L, U, D \in \mathbb{C}^{n \times n}$ :

$$A = \underbrace{L}_{\text{Einträge unter,}} + \underbrace{U}_{\text{über}} + \underbrace{D}_{\text{und auf der Diagonalen.}}$$

## Beispiel (Gesamtschrittverfahren (alias Jacobi-Verfahren))

Ein Fixpunkt von  $\Phi(x) := D^{-1} \cdot (b - (L + U) \cdot x)$  löst  $A \cdot x = b$ :

$$x = \Phi(x) \Leftrightarrow D \cdot x = b - (L + U) \cdot x \Leftrightarrow A \cdot x = b$$

Für  $\|D^{-1} \cdot (L + U)\| < 1$  ist  $\Phi$  eine Kontraktion. Als Algorithmus:

**for**  $k = 0, 1, \dots$  ( $k$  Iterationsindex)

**for**  $i = 1$  to  $n$

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

# Gleichungssysteme iterativ lösen

## Beispiel (Einzelschrittverfahren (alias Gauß-Seidel-Verfahren))

Ein Fixpunkt von  $\Phi(x) := (D + L)^{-1} \cdot (b - U \cdot x)$  löst  $A \cdot x = b$ :

$$x = \Phi(x) \Leftrightarrow (D + L) \cdot x = b - U \cdot x \Leftrightarrow A \cdot x = b$$

Für  $\|(D + L)^{-1} \cdot U\| < 1$  ist  $\Phi$  eine Kontraktion. Als Algorithmus:

**for**  $k = 0, 1, \dots$  ( $k$  Iterationsindex)

**for**  $i = 1$  to  $n$

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

# Gleichungssysteme iterativ lösen

Iterative Verfahren zum Lösen von Gleichungssystemen sind nicht immer anwendbar, aber sie sind:

- kaum anfällig für Rundungsfehler,
- effektiv wenn nur ein Vektor  $b \in \mathbb{C}^n$  genutzt wird,
- sehr schnell, wenn  $A$  eine dünn besetzte Matrix ist.

# Nichtlineare Kontraktionen

## Satz 10.2 (Kriterium für Kontraktionen)

Sei  $\Phi : D \rightarrow \mathbb{C}^n$  stetig differenzierbar und sei  $\hat{x} \in D$  ein Fixpunkt von  $\Phi$  mit  $\|\Phi'(\hat{x})\| < 1$ . Dann existiert ein  $r > 0$  so dass  $\Phi$  auf dem Ball  $B_r(\hat{x})$  vom Radius  $r$  um  $\hat{x}$  eine Kontraktion ist.

**Beweis:** Da  $\|\Phi'(x)\|$  stetig ist, existiert ein  $r > 0$  und ein  $q < 1$  so dass  $\|\Phi'(x)\| \leq q$  für alle  $x \in B_r(\hat{x})$ . Nach dem Mittelwertsatz der Differentialrechnung im  $\mathbb{C}^n$  wissen wir für alle  $x, y \in B_r(\hat{x})$ :

$$\Phi(y) - \Phi(x) = \int_0^1 \Phi'(x + t \cdot (y - x)) \cdot (y - x) dt$$

$$\Rightarrow \|\Phi(x) - \Phi(y)\| \leq \int_0^1 \|\Phi'(x + t \cdot (y - x))\| \cdot \|y - x\| dt \leq q \cdot \|x - y\|$$

Insbesondere ist  $\|\Phi(x) - \hat{x}\| \leq q \cdot \|x - \hat{x}\| \leq r$ , d.h.  $\Phi(x) \in B_r(\hat{x})$ .

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Definition: Konvergenzordnung

## Definition 10.3 (Konvergenzordnung)

Für eine reelle nichtnegative Nullfolge  $\{\varepsilon_k\}_{k \in \mathbb{N}}$  wird

$$K = \limsup_{k \rightarrow \infty} \varepsilon_k^{1/k} \quad (10.5)$$

als *asymtotischer Konvergenzfaktor* definiert. Die Folge  $\{\varepsilon_k\}$  heißt *sublinear*, *linear*, bzw. *superlinear*, je nachdem ob  $K = 1$ ,  $0 < K < 1$  oder  $K = 0$  ist. Gilt im superlinear konvergenten Fall zudem

$$\varepsilon_{k+1} \leq c \cdot \varepsilon_k^p \quad \text{für ein } p > 1, c > 0 \text{ und fast alle } k \in \mathbb{N},$$

dann hat die Folge die *Konvergenzordnung*  $p$ .

Entsprechend wird die Terminologie für konvergente Folgen  $\{x^{(k)}\} \subset \mathbb{K}^n$  mit Grenzwert  $\hat{x}$  über  $\varepsilon_k := \|x^{(k)} - \hat{x}\|$  eingeführt.

# Beispiel für quadratische Konvergenz

## Bemerkung 10.2 (Konvergenzgeschwindigkeit)

Als Faustregel erwartet man, dass sich mit Konvergenzordnung  $p$  die Anzahl der korrekten Dezimalstellen bei jeder Iteration ver- $p$ -facht.

## Beispiel (Konvergenzordnung $p = 2$ (quadratische Konvergenz))

Heronverfahren mit  $\nu = 2, a > 0$ :

$$z_k = \frac{x_k - \sqrt{a}}{x_k + \sqrt{a}}, \quad z_{k+1} = \left( \frac{x_k - \sqrt{a}}{x_k + \sqrt{a}} \right)^2 = z_k^2 \quad (10.6)$$

Für  $\{x_k\}$  des Heronverfahrens ergibt sich unter der Voraussetzung  $\sqrt{a}/2 \leq x_k \leq 2\sqrt{a}$ , dass auch  $x_{k+1} \in \left[ \frac{\sqrt{a}}{2}, 2\sqrt{a} \right]$  und es folgt aus (10.6):

$$|x_{k+1} - \sqrt{a}| = |z_{k+1}(x_{k+1} + \sqrt{a})| = \frac{|x_{k+1} + \sqrt{a}|}{|x_k + \sqrt{a}|^2} |x_k - \sqrt{a}|^2 \leq \frac{\frac{3}{2}\sqrt{a}}{\frac{9}{4}a} |x_k - \sqrt{a}|^2$$

# Lokale Konvergenz der Fixpunktiteration

## Satz 10.3 (Lokale Konvergenz der Fixpunktiteration)

Die Funktion  $\Phi : D(\Phi) \subset \mathbb{C}^n \longrightarrow \mathbb{C}^n$  sei stetig differenzierbar und habe einen Fixpunkt  $\hat{x}$  in  $D(\Phi)$ . Ferner sei  $\|\cdot\|$  eine Norm in  $\mathbb{C}^n$  und  $\|\cdot\|_M$  eine verträgliche Matrixnorm in  $\mathbb{C}^{n \times n}$  mit  $\|\Phi'(\hat{x})\|_M < 1$ .

Dann ist  $\Phi$  in einer Umgebung  $U$  von  $\hat{x}$  eine Kontraktion und die Fixpunktiteration

$$x^{(k+1)} = \Phi(x^{(k)}), \quad k = 0, 1, 2, \dots$$

ist lokal konvergent gegen  $\hat{x}$ .

**Beweis:** Dass  $\Phi$  lokal eine Kontraktion ist haben wir bereits bewiesen. Mit dem Banachschen Fixpunktsatz folgt die Konvergenz. □

# Lokal superlineare Konvergenz

## Satz 10.4 (Lokal superlineare Konvergenz)

Die Funktion  $\Phi : D(\Phi) \subset \mathbb{C}^n \longrightarrow \mathbb{C}^n$  sei  $p$ -mal stetig differenzierbar und habe einen Fixpunkt  $\hat{x} \in D(\Phi)$ . Ferner sei  $p \geq 2$ ,

$$\begin{aligned} d^{(1)}\Phi(\hat{x}) &= 0, \dots, d^{(p-1)}\Phi(\hat{x}) = 0 \\ \text{und} \quad d^{(p)}\Phi(\hat{x}) &\neq 0. \end{aligned}$$

Dann ist die Fixpunktiteration  $x_{k+1} = \Phi(x_k)$  lokal superlinear konvergent gegen  $\hat{x}$  und die Konvergenzordnung ist genau  $p$ .

**Beweis:** Es gilt  $\|\Phi'(\hat{x})\| = 0 < 1$  und somit folgt die lokale Konvergenz aus dem vorigen Satz. Nun entwickeln wir  $\Phi$  um  $\hat{x}$  in ein Taylorpolynom:

$$\begin{aligned}\Phi(x) &= \Phi(\hat{x}) + \sum_{i=1}^p \frac{1}{i!} \cdot d^{(i)}\Phi(\hat{x})(x - \hat{x})^i + o(\|x - \hat{x}\|^p) \\ &= \Phi(\hat{x}) + \frac{1}{p!} \cdot d^{(p)}\Phi(\hat{x})(x - \hat{x})^p + o(\|x - \hat{x}\|^p)\end{aligned}$$

Dann existiert eine von  $d^{(p)}\Phi(\hat{x})$  abhängige Konstante  $c > 0$  so dass:

$$\begin{aligned}\|x^{(k+1)} - \hat{x}\| &= \|\Phi(x^{(k)}) - \Phi(\hat{x})\| \\ &= \frac{1}{p!} \cdot \|d^p\Phi(\hat{x})(x^{(k)} - \hat{x})^p\| + o(\|x^{(k)} - \hat{x}\|^p) \\ &\leq \frac{1}{p!} \cdot c \cdot \|x^{(k)} - \hat{x}\|^p + o(\|x^{(k)} - \hat{x}\|^p)\end{aligned}$$



## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Lösen nichtlinearer Gleichungen

- *Erinnerung:* Jede Gleichung kann als Nullstellenproblem aufgefasst werden:

$$f_1(x) = f_2(x) \Leftrightarrow f_1(x) - f_2(x) = 0$$

# Lösen nichtlinearer Gleichungen

- *Erinnerung:* Jede Gleichung kann als Nullstellenproblem aufgefasst werden:

$$f_1(x) = f_2(x) \Leftrightarrow f_1(x) - f_2(x) = 0$$

- Entsprechende Lösungsalgorithmen sind:
  - auf sehr viele Probleme anwendbar.
  - bei hinreichender Differenzierbarkeit und niedriger Dimension sehr effizient.
  - weniger effektiv als spezialisierte Algorithmen.

# Lösen nichtlinearer Gleichungen

- *Erinnerung:* Jede Gleichung kann als Nullstellenproblem aufgefasst werden:

$$f_1(x) = f_2(x) \Leftrightarrow f_1(x) - f_2(x) = 0$$

- Entsprechende Lösungsalgorithmen sind:
  - auf sehr viele Probleme anwendbar.
  - bei hinreichender Differenzierbarkeit und niedriger Dimension sehr effizient.
  - weniger effektiv als spezialisierte Algorithmen.
- Eigenwertberechnung ist z.B. ein Nullstellenproblem:

$$A \cdot v = \lambda \cdot v \Leftrightarrow (A - \lambda \cdot I) \cdot v = 0$$

- Allerdings sind hier spezialisierte Algorithmen effektiver.

# Herleitung des Newtonverfahrens

- Sei  $D(f) \subset \mathbb{C}$  offen und  $f : D(f) \rightarrow \mathbb{C}$  stetig differenzierbar.
- Wir betrachten die Nullstellenaufgabe  $f(x) = 0$ .

# Herleitung des Newtonverfahrens

- Sei  $D(f) \subset \mathbb{C}$  offen und  $f : D(f) \rightarrow \mathbb{C}$  stetig differenzierbar.
- Wir betrachten die Nullstellenaufgabe  $f(x) = 0$ .
- *Idee:* Formuliere ein Fixpunktproblem:

$$x = x + g(x) \cdot f(x) =: \Phi(x)$$

- $g : D(f) \rightarrow \mathbb{C} \setminus \{0\}$  wird unten gewählt.

$$x = \Phi(x) \Leftrightarrow f(x) = 0$$

# Herleitung des Newtonverfahrens

- Sei  $D(f) \subset \mathbb{C}$  offen und  $f : D(f) \rightarrow \mathbb{C}$  stetig differenzierbar.
- Wir betrachten die Nullstellenaufgabe  $f(x) = 0$ .
- Idee: Formuliere ein Fixpunktproblem:

$$x = x + g(x) \cdot f(x) =: \Phi(x)$$

- $g : D(f) \rightarrow \mathbb{C} \setminus \{0\}$  wird unten gewählt.

$$x = \Phi(x) \Leftrightarrow f(x) = 0$$

- Sei  $\hat{x} \in D(f)$  eine Nullstelle.
- Wir wollen lokal quadratische Konvergenz, also  $\Phi'(\hat{x}) = 0$ :

$$\Phi'(\hat{x}) = 1 + g'(\hat{x}) \cdot \underbrace{f(\hat{x})}_{=0} + g(\hat{x}) \cdot f'(\hat{x}) = 0$$

$$\Rightarrow g(\hat{x}) = \frac{-1}{f'(\hat{x})}$$

# Das Newtonverfahren

Wählen wir speziell  $g := \frac{-1}{f'}$ , so erhalten wir das *Newtonverfahren*:

$$x_{k+1} = \Phi(x_k) = x_k - \frac{1}{f'(x_k)} \cdot f(x_k)$$

# Das Newtonverfahren

Wählen wir speziell  $g := \frac{-1}{f'}$ , so erhalten wir das *Newtonverfahren*:

$$x_{k+1} = \Phi(x_k) = x_k - \frac{1}{f'(x_k)} \cdot f(x_k)$$

## Bemerkung 10.3 (Geometrische Anschauung)

$x_{k+1}$  ist der Schnitt der Tangente an den Graph von  $f$  im Punkt  $(x_k, f(x_k))$  mit der  $x$ -Achse.

$$y(x) = f(x_k) + f'(x_k) \cdot (x - x_k)$$

# Das Newtonverfahren

Wählen wir speziell  $g := \frac{-1}{f'}$ , so erhalten wir das *Newtonverfahren*:

$$x_{k+1} = \Phi(x_k) = x_k - \frac{1}{f'(x_k)} \cdot f(x_k)$$

## Bemerkung 10.3 (Geometrische Anschauung)

$x_{k+1}$  ist der Schnitt der Tangente an den Graph von  $f$  im Punkt  $(x_k, f(x_k))$

$$y(x) = f(x_k) + f'(x_k) \cdot (x - x_k)$$

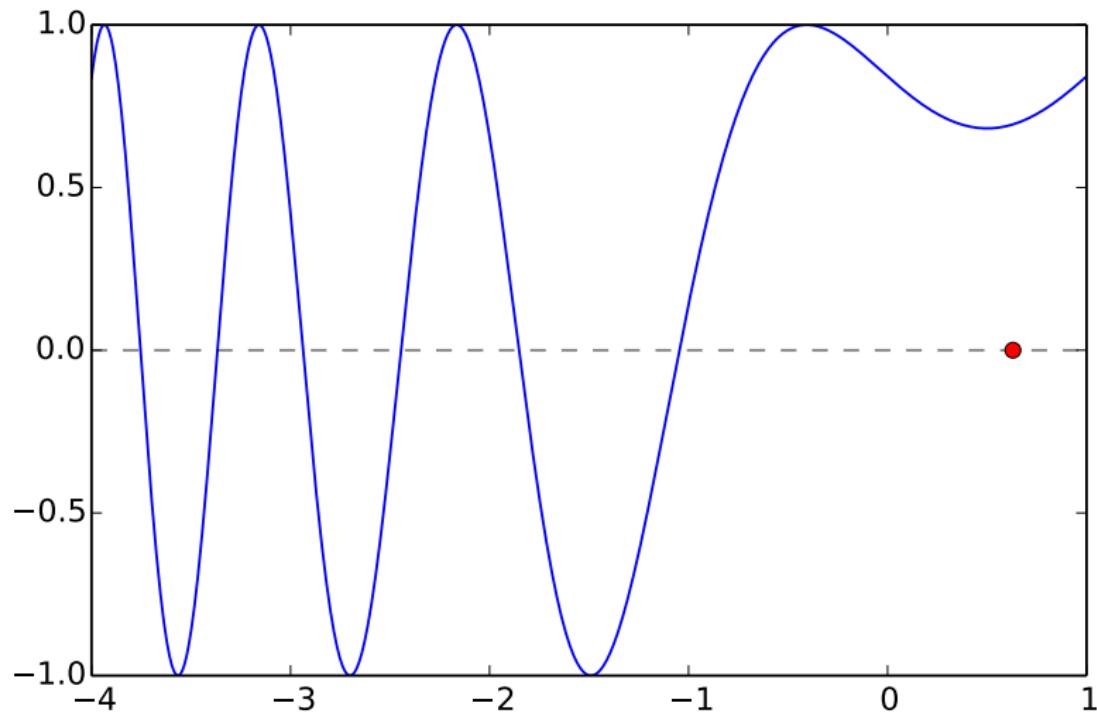
mit der  $x$ -Achse.

Das Newton-Verfahren alterniert also zwischen zwei Schritten:

- 1 Approxiere  $f$  durch seine Taylorapproximation erster Ordnung.
- 2 Berechne die Nullstelle der Taylorapproximation.

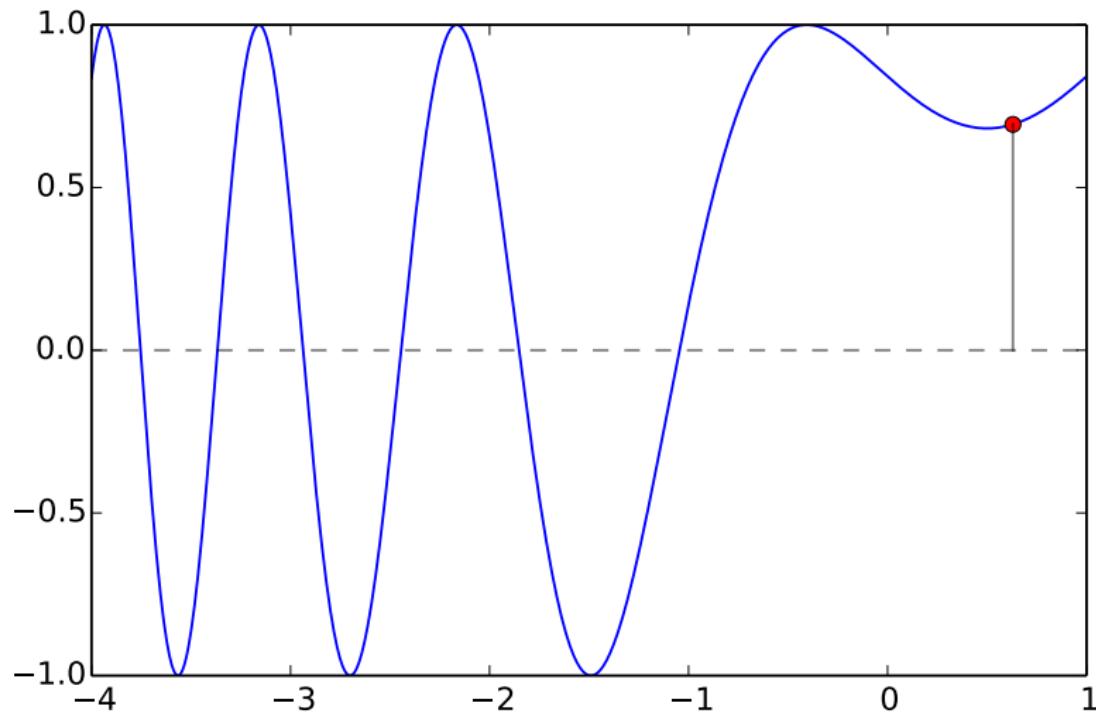
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



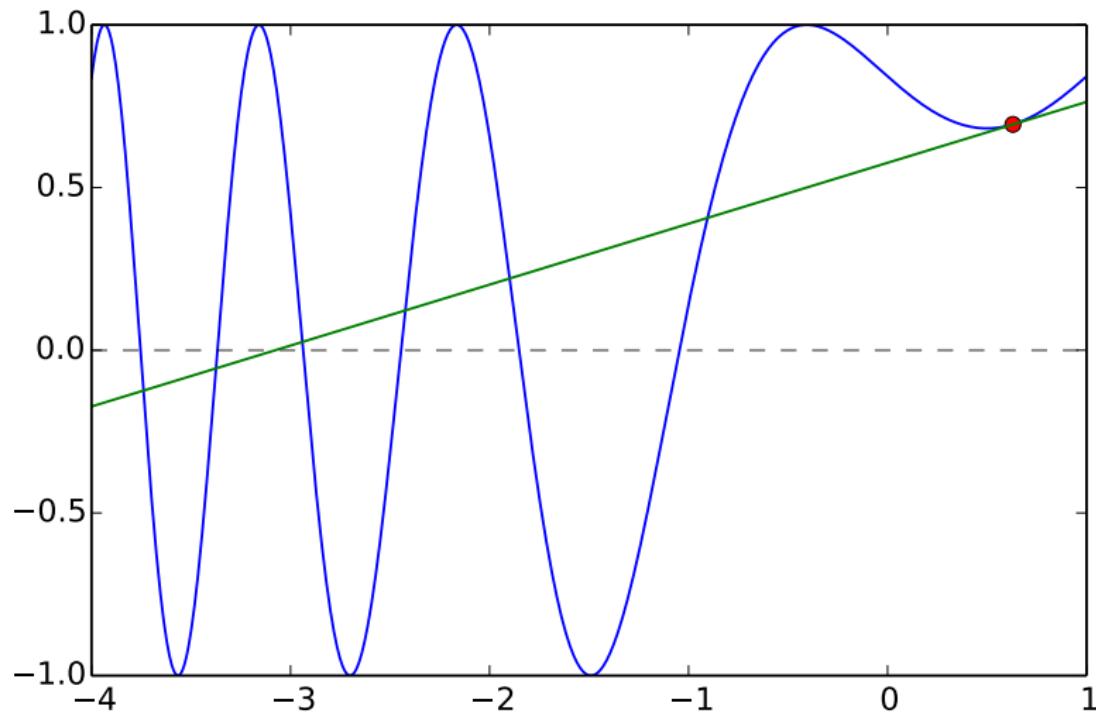
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



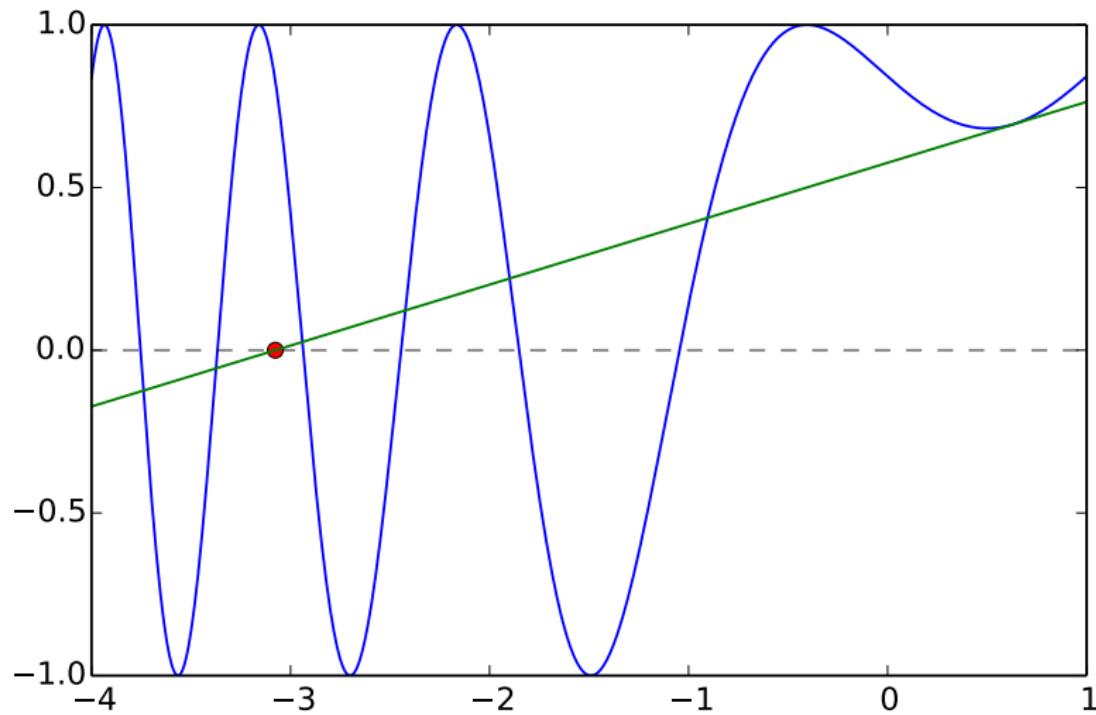
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



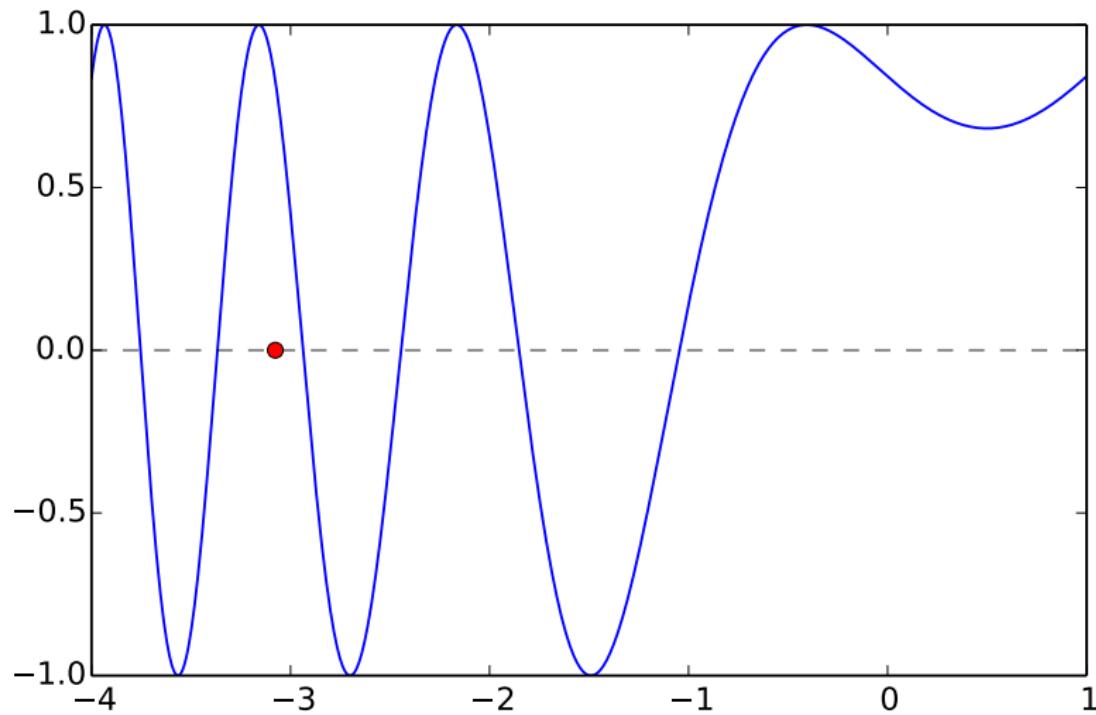
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



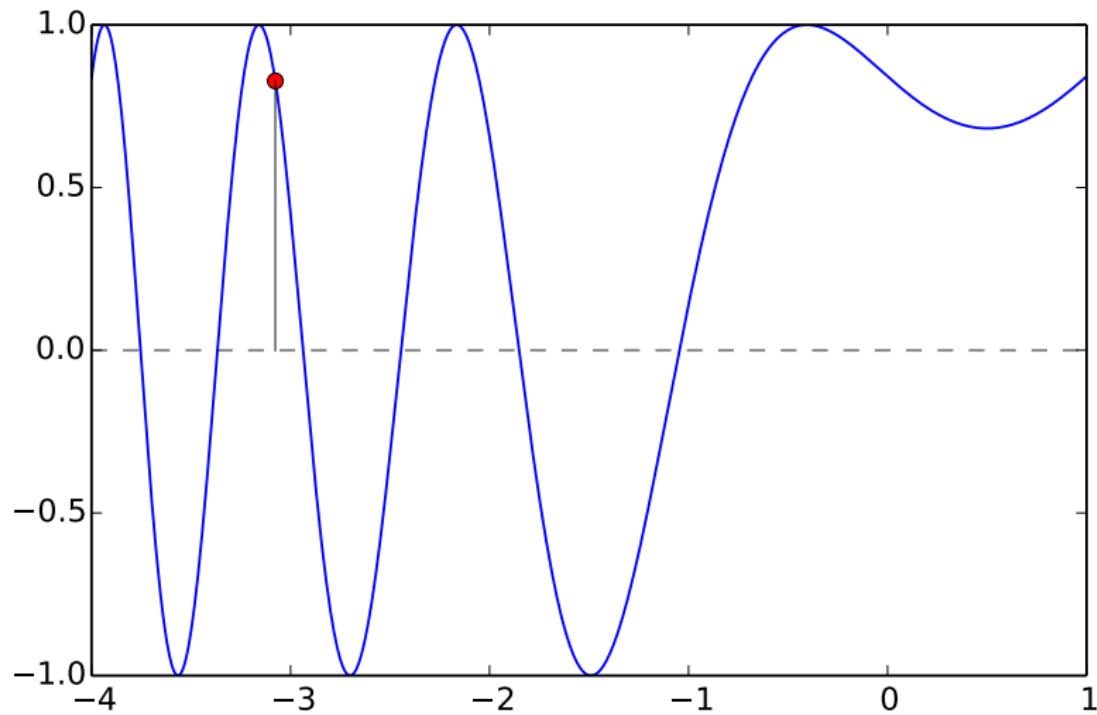
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



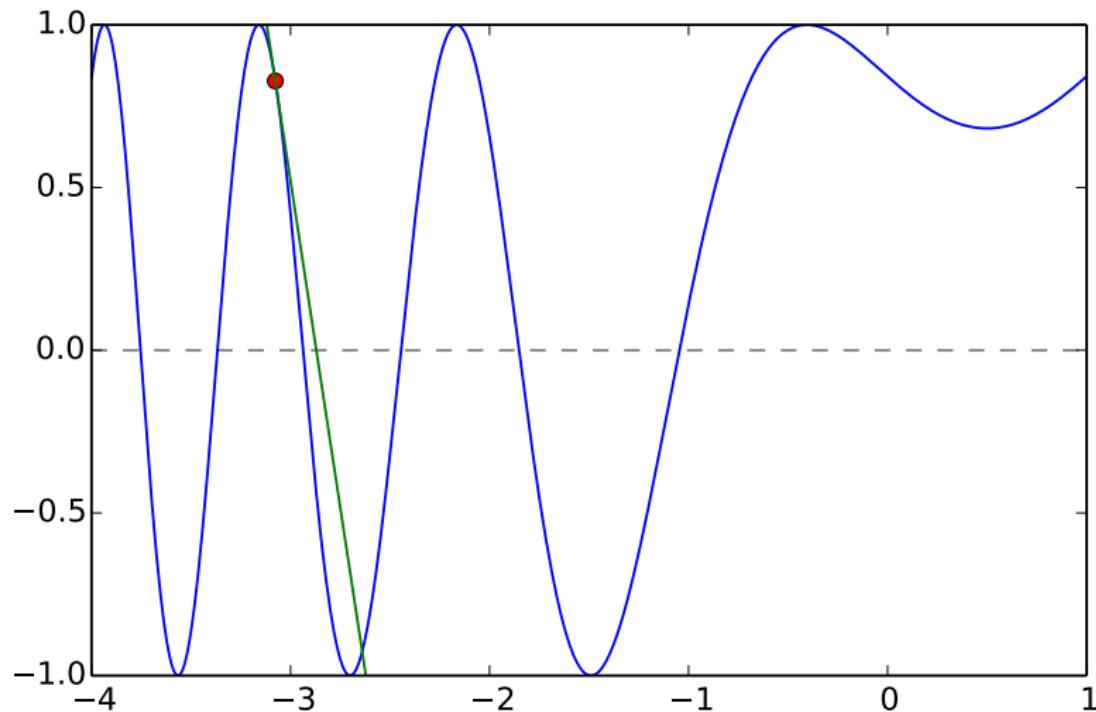
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



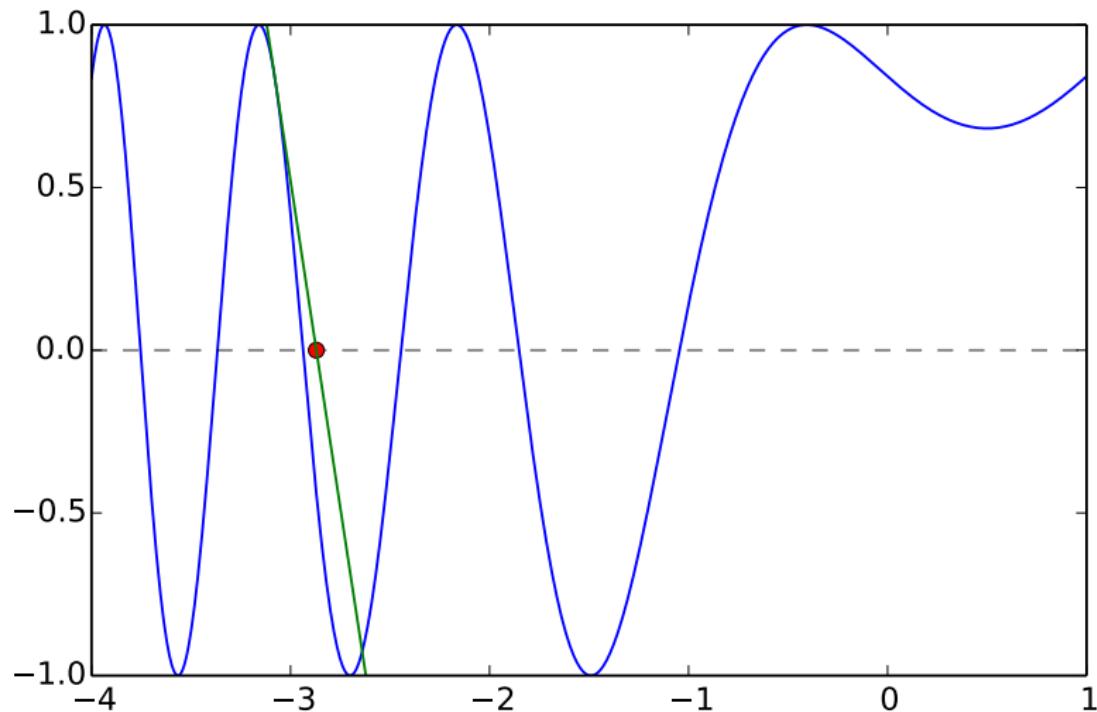
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



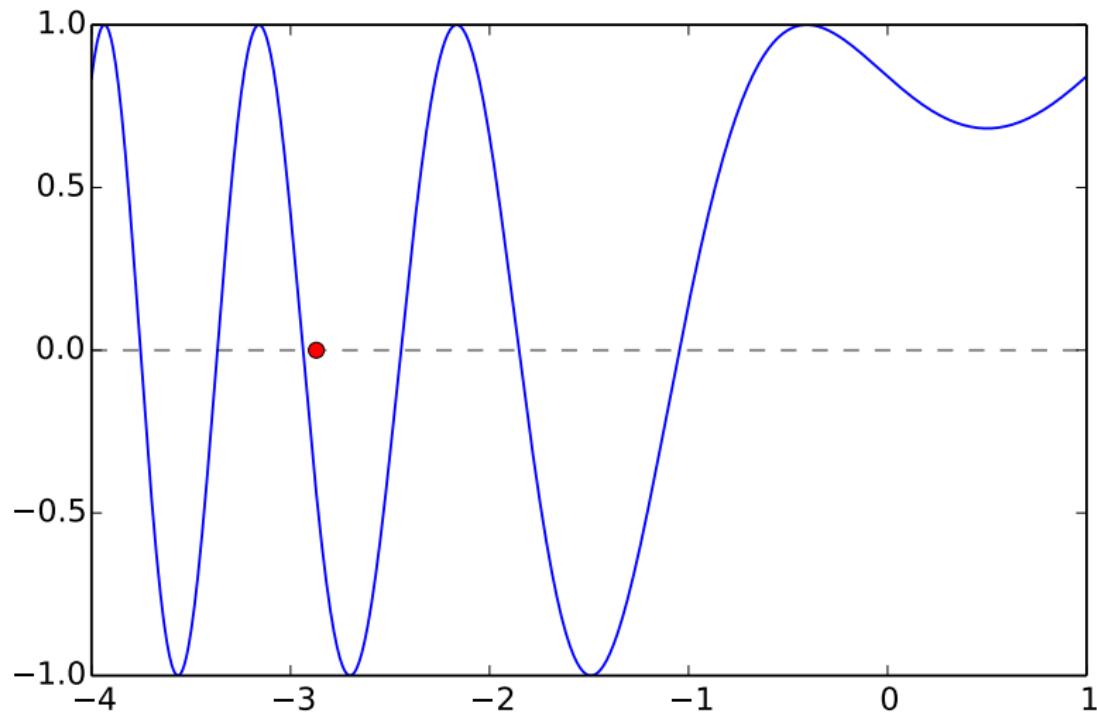
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



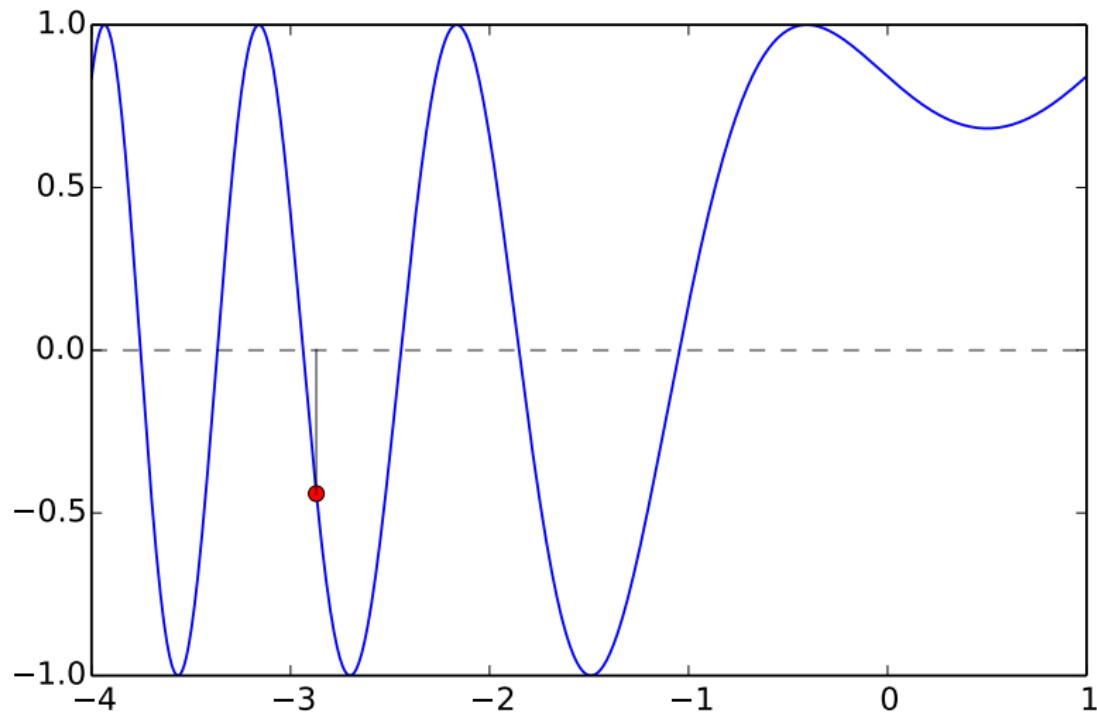
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



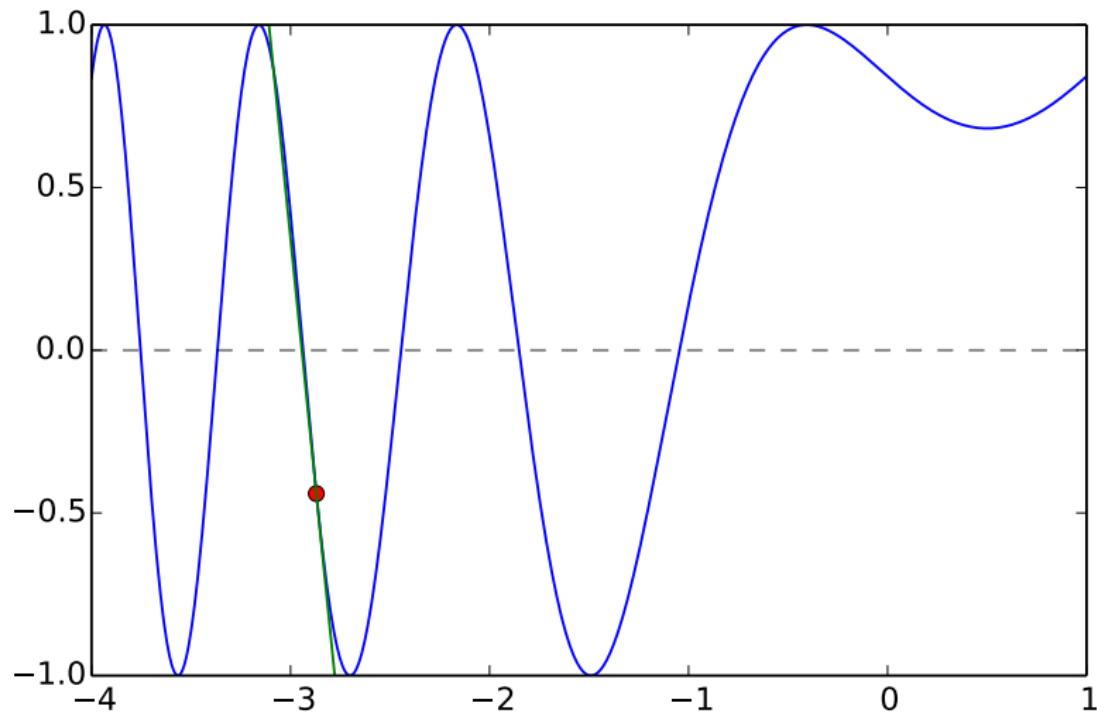
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$

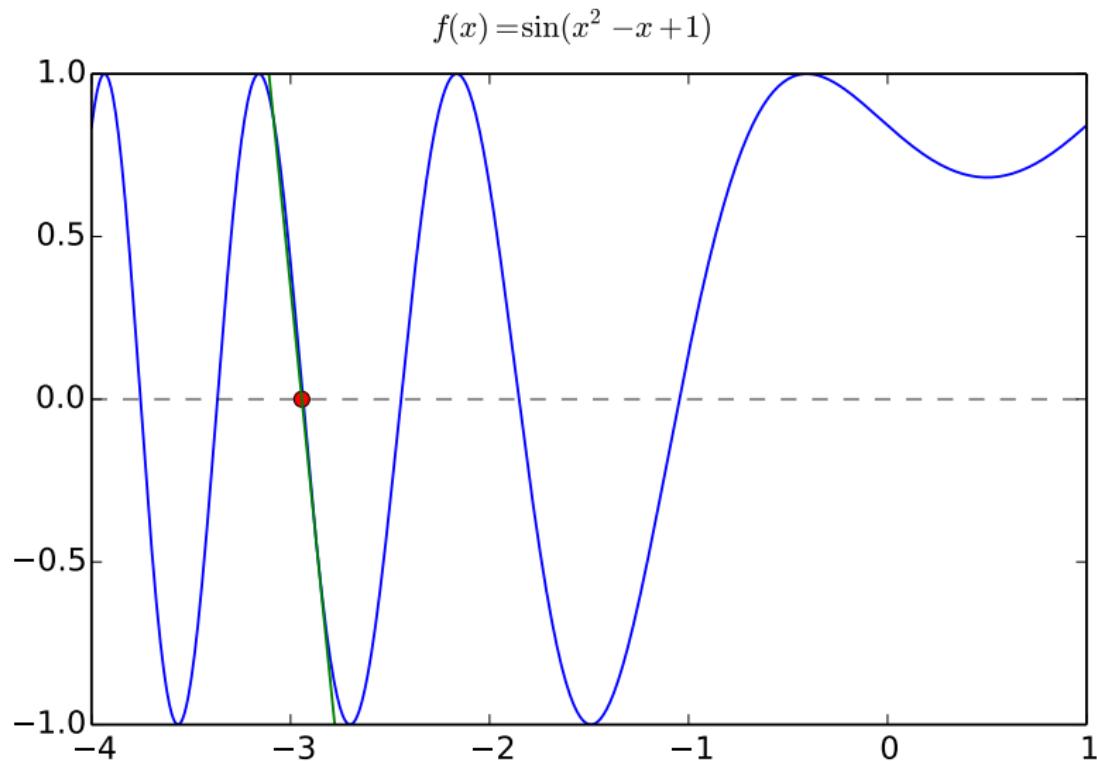


# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$

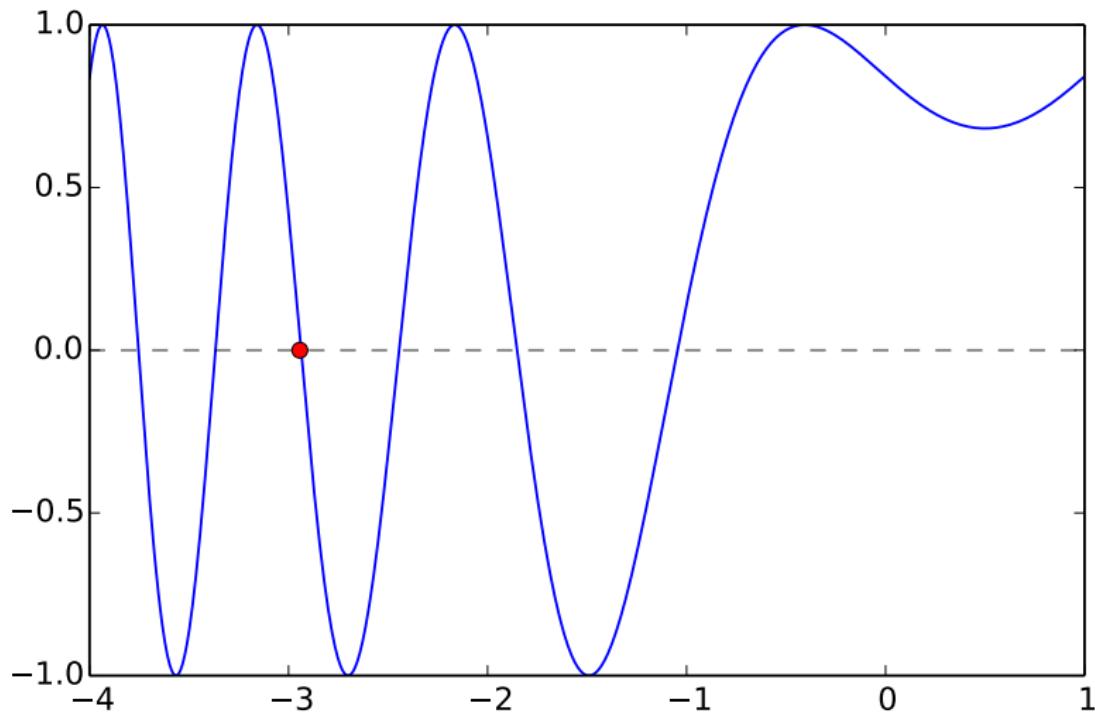


# Beispiel für das Newtonverfahren



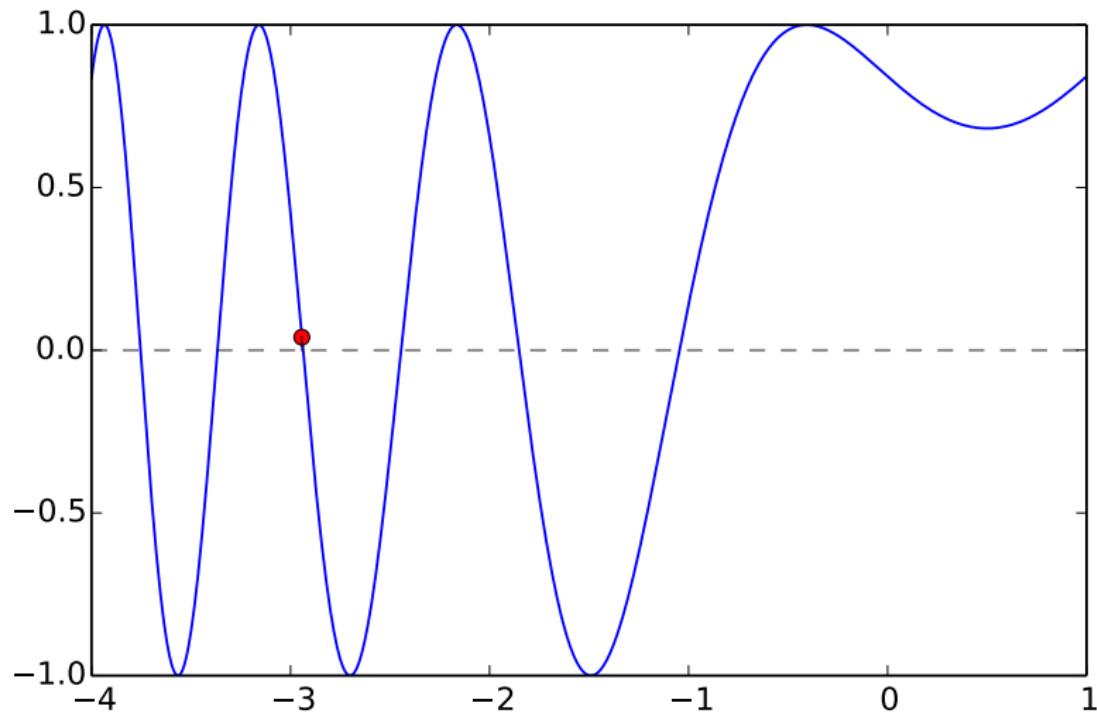
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



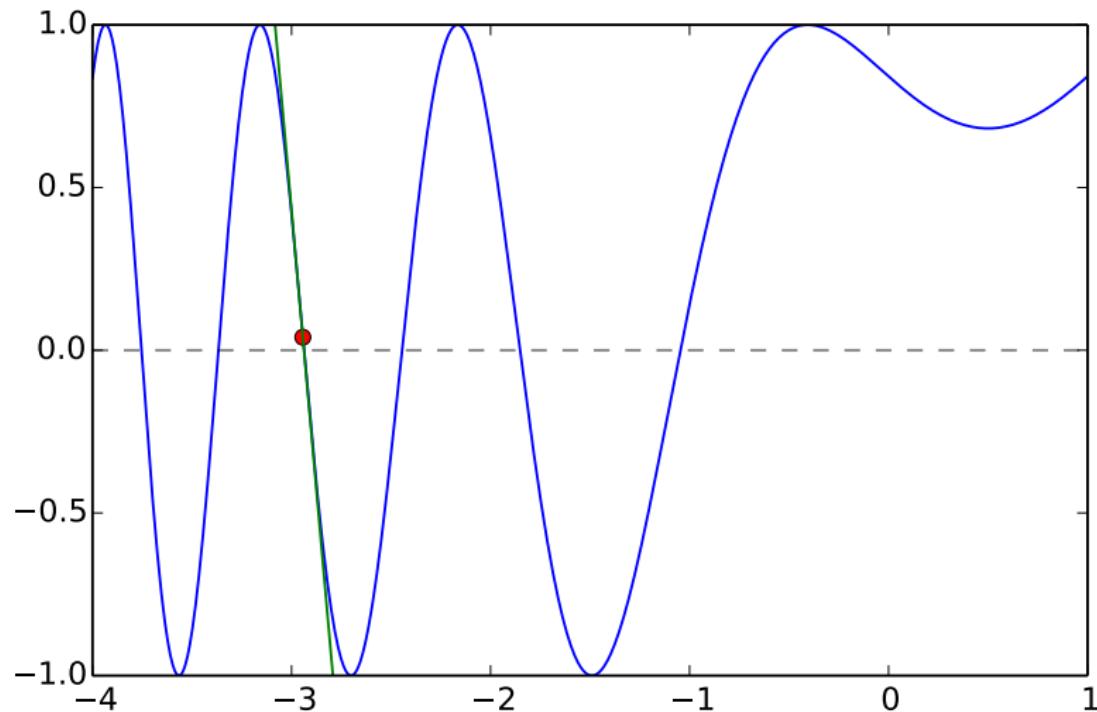
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



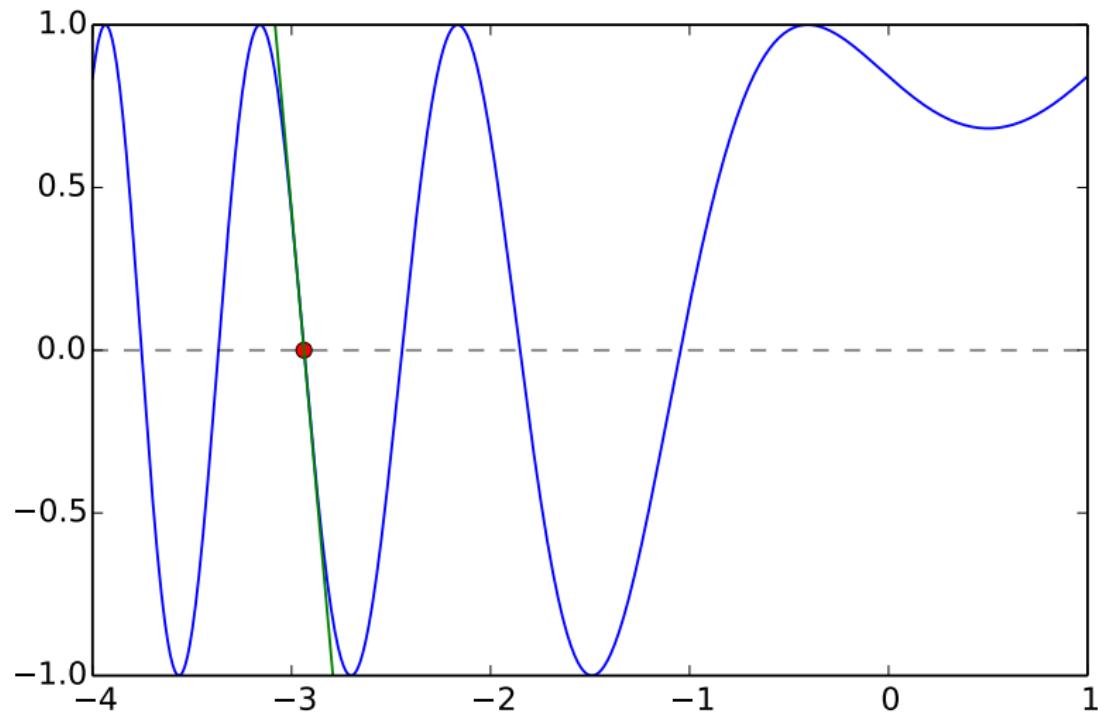
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



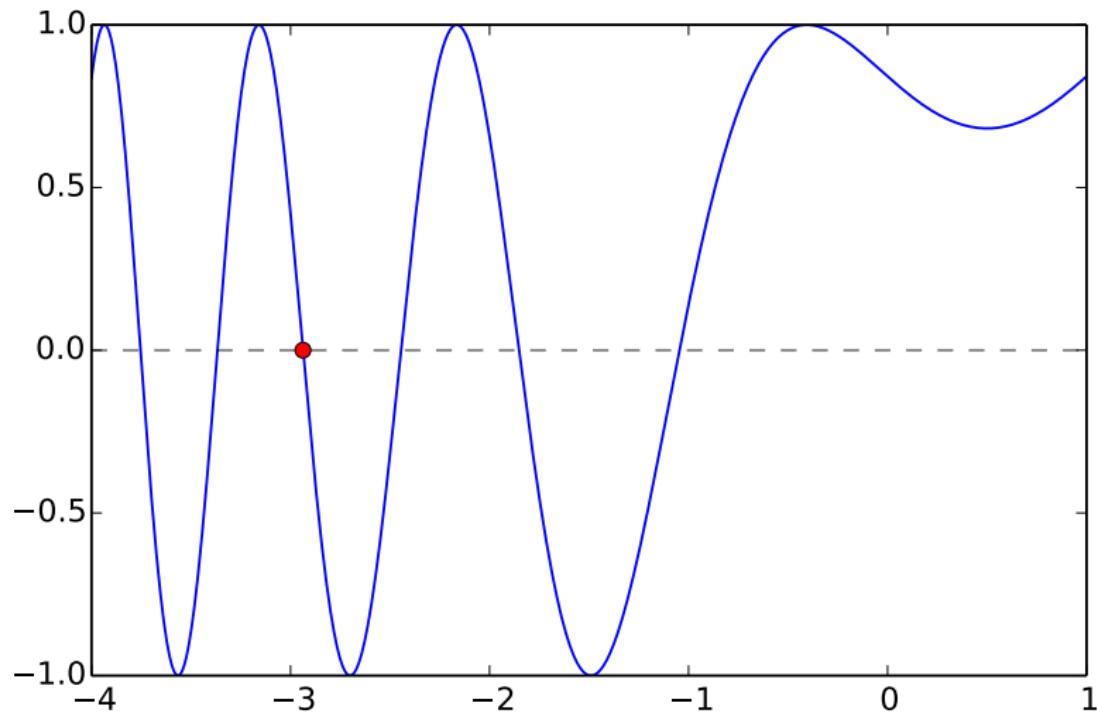
# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



# Beispiel für das Newtonverfahren

$$f(x) = \sin(x^2 - x + 1)$$



# Beispiel für das Newtonverfahren

Beispiel ( $f(x) = x^\nu - a$ )

Die Nullstellen der Funktion

$$f(x) = x^\nu - a, \quad \nu \in \mathbb{N} \setminus \{1\}, a \in \mathbb{R}^+$$

sind die  $\nu$ -ten Wurzeln der Zahl  $a$ . Das Newtonverfahren ergibt

$$x_{k+1} = x_k - \frac{x_k^\nu - a}{\nu x_k^{\nu-1}} = \frac{\nu-1}{\nu} x_k + \frac{a}{\nu} x_k^{1-\nu}, \quad k = 0, 1, \dots$$

Das entspricht dem Heronverfahren.

# Mindestens lokale Konvergenz des Newtonverfahrens

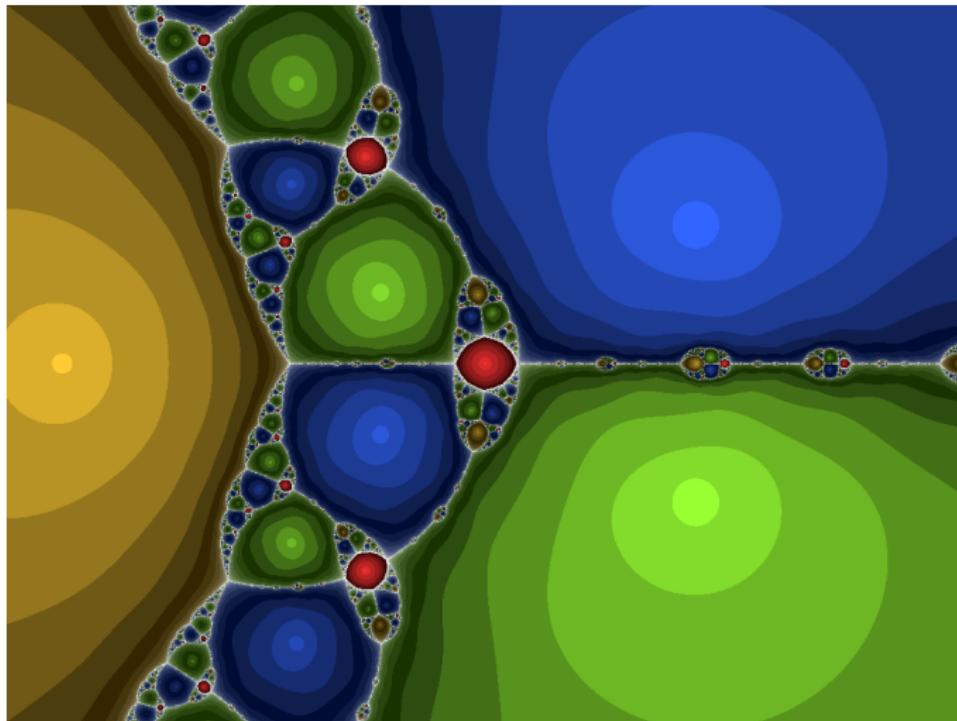
## Satz 10.5 (Lokal quadratische Konvergenz des Newtonverfahrens)

Sei  $f : D(f) \rightarrow \mathbb{C}$  zwei mal stetig differenzierbar und  $\hat{x} \in D(f)$  mit  $f(\hat{x}) = 0$  und  $f'(\hat{x}) \neq 0$ .

Dann konvergiert das Newtonverfahren (mindestens) lokal quadratisch gegen  $\hat{x}$ .

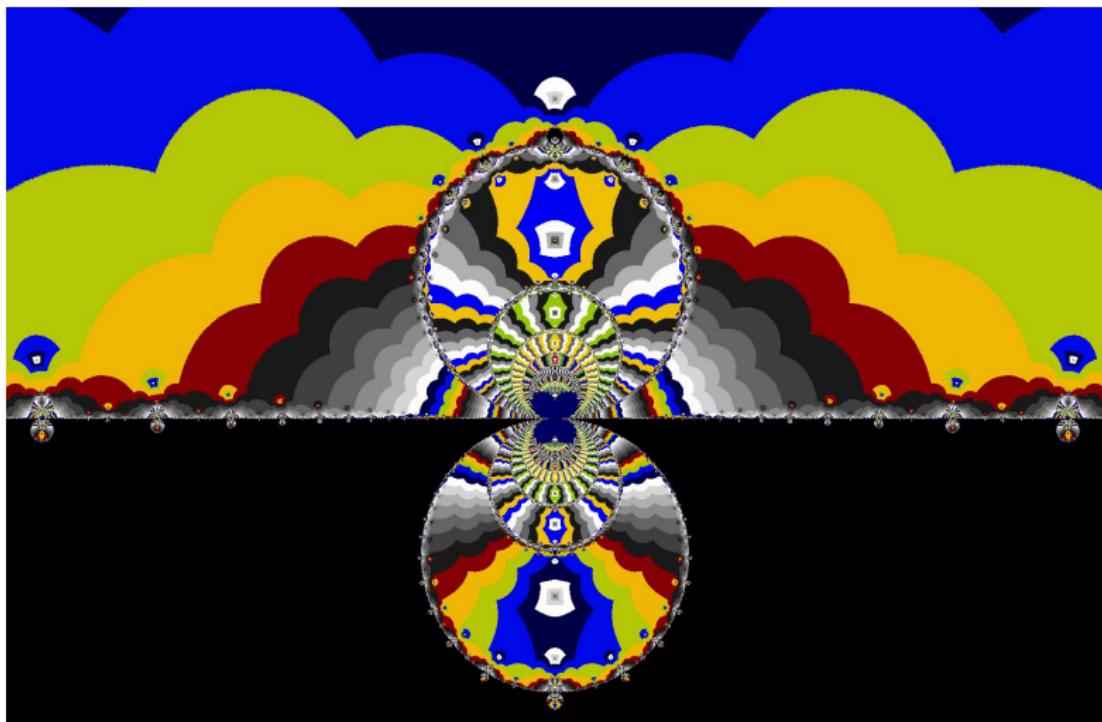
**Beweis:** Nach Konstruktion gilt  $\Phi(\hat{x}) = \hat{x}$  und  $\Phi'(\hat{x}) = 0$ . Die Behauptung folgt damit aus dem Satz über lokal superlineare Konvergenz. □

## Newton-Fraktale (Quelle Wikimedia)



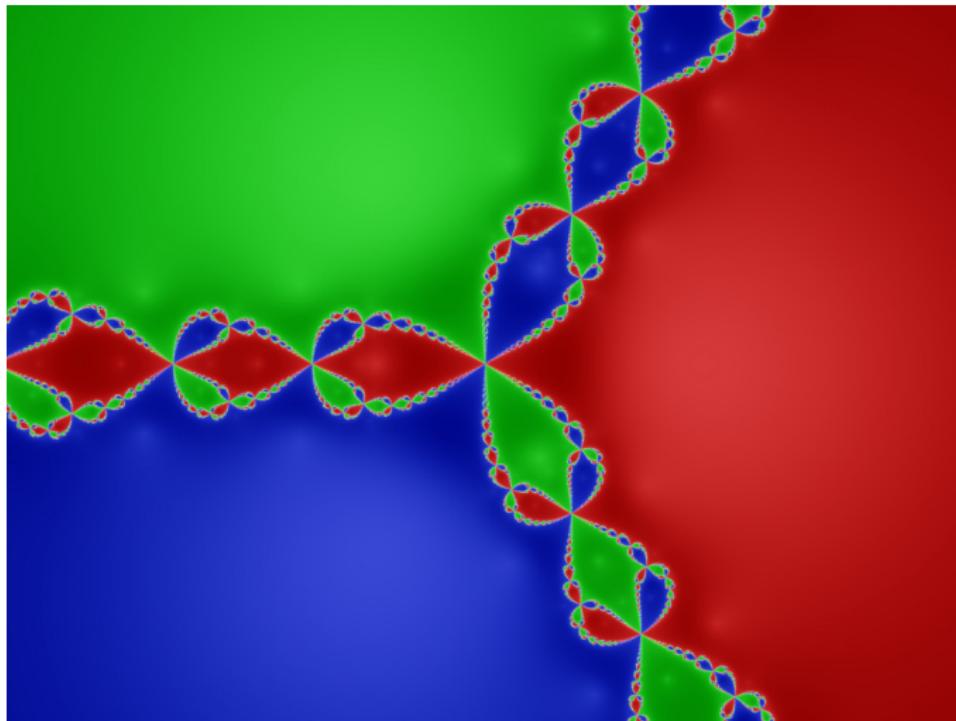
$$f(z) = z^3 - 2 \cdot z + 2$$

## Newton-Fraktale (Quelle Wikimedia)



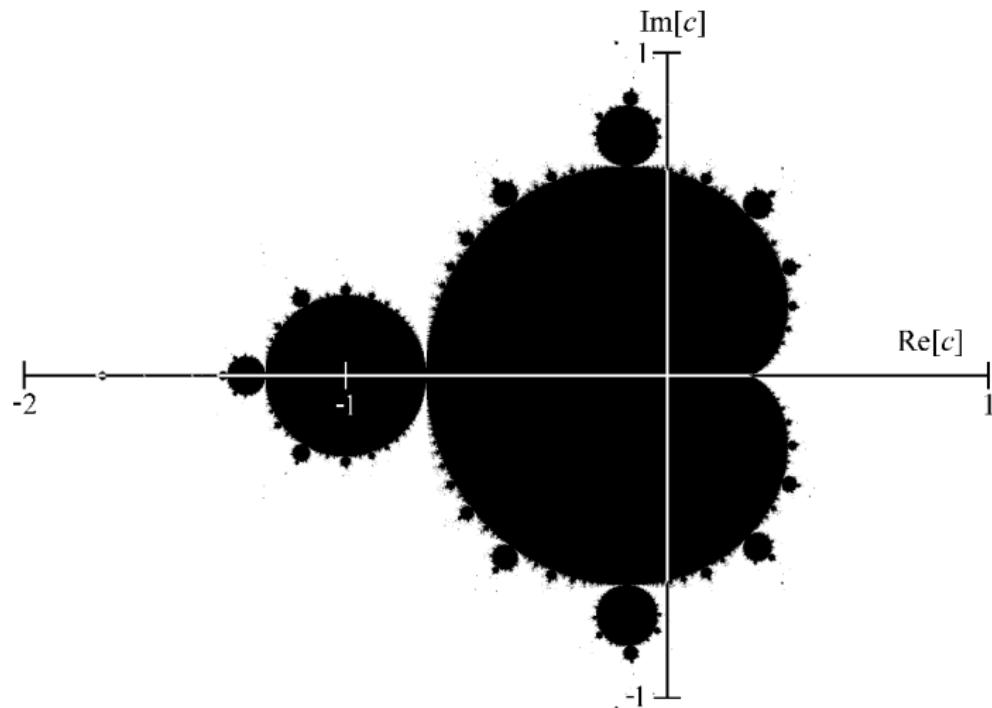
$$f(z) = \cosh(z) - 1$$

## Newton-Fraktale (Quelle Wikimedia)



$$f(z) = z^3 - 1$$

## Newton-Fraktale (Quelle Wikimedia)



$$z_{k+1} := z_k^2 + c \text{ (Mandelbrot)}$$

# Initialisierung des Newtonverfahrens

- Initialisierung beeinflusst Ergebnis auf chaotische Weise.
- Gute Initialisierung führt zu sehr schneller (quadratischer) Konvergenz gegen naheliegende Nullstelle.
- Schlechte Initialisierung führt zu langsamer oder keiner Konvergenz und evtl. zu unerwünschten Lösungen.

# Initialisierung des Newtonverfahrens

- Initialisierung beeinflusst Ergebnis auf chaotische Weise.
- Gute Initialisierung führt zu sehr schneller (quadratischer) Konvergenz gegen naheliegende Nullstelle.
- Schlechte Initialisierung führt zu langsamer oder keiner Konvergenz und evtl. zu unerwünschten Lösungen.
- Möglichst viel Wissen über das Problem für Initialisierung verwenden.
- Heuristisch eine Näherungslösung als Initialisierung berechnen.
- Systematisch mehrere Initialisierungen ausprobieren.

# Initialisierung des Newtonverfahrens

- Initialisierung beeinflusst Ergebnis auf chaotische Weise.
- Gute Initialisierung führt zu sehr schneller (quadratischer) Konvergenz gegen naheliegende Nullstelle.
- Schlechte Initialisierung führt zu langsamer oder keiner Konvergenz und evtl. zu unerwünschten Lösungen.
- Möglichst viel Wissen über das Problem für Initialisierung verwenden.
- Heuristisch eine Näherungslösung als Initialisierung berechnen.
- Systematisch mehrere Initialisierungen ausprobieren.
- Nicht vergessen:  $f$  muss (zwei mal) stetig differenzierbar sein.
  - ... und  $f'$  korrekt berechnet.

# Das Newtonverfahren im $\mathbb{C}^n$

- Sei  $D(f) \subset \mathbb{C}^n$  offen und  $f : D(f) \rightarrow \mathbb{C}^n$  stetig differenzierbar.
- Wir betrachten die Nullstellenaufgabe  $f(x) = 0 \in \mathbb{C}^n$ .

# Das Newtonverfahren im $\mathbb{C}^n$

- Sei  $D(f) \subset \mathbb{C}^n$  offen und  $f : D(f) \rightarrow \mathbb{C}^n$  stetig differenzierbar.
- Wir betrachten die Nullstellenaufgabe  $f(x) = 0 \in \mathbb{C}^n$ .
- Approximiere  $f$  durch die Taylorapproximation erster Ordnung.
- Löse Nullstellenaufgabe für die Taylorapproximation (lineares Gleichungssystem):

$$f(x^{(k)}) + f'(x^{(k)}) \cdot (x^{(k+1)} - x^{(k)}) = 0$$

- Dabei muss  $f'(x^{(k)})$  regulär sein.

# Das Newtonverfahren im $\mathbb{C}^n$

- Sei  $D(f) \subset \mathbb{C}^n$  offen und  $f : D(f) \rightarrow \mathbb{C}^n$  stetig differenzierbar.
- Wir betrachten die Nullstellenaufgabe  $f(x) = 0 \in \mathbb{C}^n$ .
- Approximiere  $f$  durch die Taylorapproximation erster Ordnung.
- Löse Nullstellenaufgabe für die Taylorapproximation (lineares Gleichungssystem):

$$f(x^{(k)}) + f'(x^{(k)}) \cdot (x^{(k+1)} - x^{(k)}) = 0$$

- Dabei muss  $f'(x^{(k)})$  regulär sein.
- Alle Berechnungen übertragen sich vom eindimensionalen Fall:

$$x^{(k+1)} = \Phi(x^{(k)}) = x^{(k)} - (f'(x^{(k)}))^{-1} \cdot f(x^{(k)})$$

$$f(\hat{x}) = 0$$

$$\Rightarrow \Phi(\hat{x}) = \hat{x} - 0 = \hat{x}$$

$$\wedge \Phi'(\hat{x}) = I - 0 - (f'(\hat{x}))^{-1} \cdot f'(\hat{x}) = 0$$

# Lokal quadratische Konvergenz des Newtonverfahrens

## Satz 10.6 (Lokal quadratische Konvergenz des Newtonverfahrens)

Sei  $f : D(f) \rightarrow \mathbb{C}^n$  zwei mal stetig differenzierbar und  $\hat{x} \in D(f)$  mit  $f(\hat{x}) = 0$  und  $f'(\hat{x})$  regulär.

Dann konvergiert das Newtonverfahren (mindestens) lokal quadratisch gegen  $\hat{x}$ .

**Beweis:** Nach Konstruktion gilt  $\Phi(\hat{x}) = \hat{x}$  und  $\Phi'(\hat{x}) = 0$ . Die Behauptung folgt damit aus dem Satz über lokal superlineare Konvergenz. □

# Kapitel 11:

## Nichtlineare Ausgleichsprobleme

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Nichtlineare Ausgleichsprobleme

- Sei nun  $m > n$ ,  $D(f) \subset \mathbb{R}^n$  offen und  $f : D(f) \rightarrow \mathbb{R}^m$  stetig differenzierbar.
- Das Nullstellenproblem  $f(x) = 0$  ist überbestimmt.
  - $m$  Gleichungen aber nur  $n$  Unbekannte.
  - Es ist davon auszugehen, dass keine Nullstelle existiert.

# Nichtlineare Ausgleichsprobleme

- Sei nun  $m > n$ ,  $D(f) \subset \mathbb{R}^n$  offen und  $f : D(f) \rightarrow \mathbb{R}^m$  stetig differenzierbar.
- Das Nullstellenproblem  $f(x) = 0$  ist überbestimmt.
  - $m$  Gleichungen aber nur  $n$  Unbekannte.
  - Es ist davon auszugehen, dass keine Nullstelle existiert.
- Betrachte nichtlineares Ausgleichsproblem:

$$\Phi(x) = \frac{1}{2} \|f(x)\|_2^2 = \frac{1}{2} f(x) \cdot f(x) \rightarrow \min$$

- Finde  $x \in D(f)$ , das  $\Phi(x) = \frac{1}{2} \|f(x)\|_2^2 = \frac{1}{2} f(x) \cdot f(x)$  minimiert.

# Nichtlineare Ausgleichsprobleme

## Erinnerung Die Bedingungen

$$\Phi'(\hat{x}) = 0 \text{ und} \quad (11.1)$$

$$\Phi''(\hat{x}) \text{ positiv definit} \quad (11.2)$$

sind gemäß (8.19) und (8.17) notwendig und hinreichend dafür, daßin  $\hat{x}$  in lokales Minimum von  $\Phi$  vorliegt. Dabei ist  $\Phi'(x) = \nabla \Phi \in \mathbb{R}^n$  der Gradient von  $\Phi$  und

$$\Phi''(x) = \left[ \frac{\partial^2}{\partial x_i \partial x_j} \Phi(x) \right]_{ij} \in \mathbb{R}^{n \times n}$$

die (symmetrische) Hessematrix. Ist zumindest die erste Gleichung erfüllt, so heißt  $\hat{x}$  stationärer Punkt.

# Nichtlineare Ausgleichsprobleme

Zur Lösung des nichtlinearen Ausgleichsproblems gibt es grob zwei Verfahrensklassen

- **Gradientenverfahren (Abstiegsverfahren)**, die in jedem Iterationsschritt das Funktional  $\Phi(x)$  in einem eindimensionalen affinen Raum minimieren, und
- **Newton-artige Verfahren**, bei denen  $\Phi$  durch eine lokale Linearisierung ersetzt wird.

# Gauß-Newton Verfahren

- Approximiere  $f$  durch die Taylorapproximation erster Ordnung.
- Löse Minimierungsaufgabe für die Taylorapproximation (lineares Ausgleichsproblem):

$$\|f(x^{(k)}) + f'(x^{(k)}) \cdot (x^{(k+1)} - x^{(k)})\|_2^2 \rightarrow \min$$

- Dabei muss  $f'(x^{(k)})$  vollen Rang haben.

# Gauß-Newton Verfahren

- Approximiere  $f$  durch die Taylorapproximation erster Ordnung.
- Löse Minimierungsaufgabe für die Taylorapproximation (lineares Ausgleichsproblem):

$$\|f(x^{(k)}) + f'(x^{(k)}) \cdot (x^{(k+1)} - x^{(k)})\|_2^2 \rightarrow \min$$

- Dabei muss  $f'(x^{(k)})$  vollen Rang haben.
- Problem: Das Gauß-Newton Verfahren ist i.A. nicht lokal konvergent.
- Wenn  $f'(x^{(k)})$  schlecht konditioniert ist kann  $x^{(k+1)}$  weit von  $x^{(k)}$  entfernt sein.
- Dann ist die Taylorapproximation nicht aussagekräftig.

# Gauß-Newton Verfahren

- Approximiere  $f$  durch die Taylorapproximation erster Ordnung.
- Löse Minimierungsaufgabe für die Taylorapproximation (lineares Ausgleichsproblem):

$$\|f(x^{(k)}) + f'(x^{(k)}) \cdot (x^{(k+1)} - x^{(k)})\|_2^2 \rightarrow \min$$

- Dabei muss  $f'(x^{(k)})$  vollen Rang haben.
- Problem: Das Gauß-Newton Verfahren ist i.A. nicht lokal konvergent.
- Wenn  $f'(x^{(k)})$  schlecht konditioniert ist kann  $x^{(k+1)}$  weit von  $x^{(k)}$  entfernt sein.
- Dann ist die Taylorapproximation nicht aussagekräftig.
- Gauß-Newton sollte **nicht** genutzt werden.

# Levenberg-Marquardt Verfahren

- Idee: Beschränke die Länge von  $h^{(k)} := x^{(k+1)} - x^{(k)}$ .
- Somit darf  $x^{(k+1)}$  nur in einer kleinen Kugel um  $x^{(k)}$  liegen.
  - Dort sollte die Taylorapproximation sinnvoll sein.
  - Die Kugel wird *Trust-Region* genannt.

# Levenberg-Marquardt Verfahren

- Idee: Beschränke die Länge von  $h^{(k)} := x^{(k+1)} - x^{(k)}$ .
- Somit darf  $x^{(k+1)}$  nur in einer kleinen Kugel um  $x^{(k)}$  liegen.
  - Dort sollte die Taylorapproximation sinnvoll sein.
  - Die Kugel wird *Trust-Region* genannt.
- Um  $x^{(k+1)} = x^{(k)} + h^{(k)}$  zu berechnen minimiere

$$\|f(x^{(k)}) + f'(x^{(k)}) \cdot h^{(k)}\|_2^2$$

unter der Nebenbedingung  $\|h^{(k)}\|_2 \leq \rho_k \in \mathbb{R}_+$ .

# Optimierung in Trust-Region

## Satz 11.1 (Optimierung in Trust-Region)

Wenn der Vektor  $h^{(k)}$

$$\|f(x^{(k)}) + f'(x^{(k)}) \cdot h^{(k)}\|_2^2 \quad (11.3)$$

unter der Nebenbedingung  $\|h^{(k)}\|_2 \leq \rho_k$  minimiert, existiert ein  $\lambda \geq 0$  mit

$$(f'(x^{(k)})^T \cdot f'(x^{(k)}) + \lambda \cdot I) \cdot h^{(k)} = -(f'(x^{(k)}))^T \cdot f(x^{(k)}). \quad (11.4)$$

Falls  $\|h^{(k)}\|_2 < \rho_k$  gilt  $\lambda = 0$ .

# Optimierung in Trust-Region

**Beweis:** Fall 1, für die optimale Lösung gilt  $\|h^{(k)}\|_2 < \rho_k$ :

Dann ist  $h^{(k)}$  lokales Minimum von (11.3) und somit Lösung des linearen Ausgleichsproblems  $\|f(x^{(k)}) + f'(x^{(k)}) \cdot h^{(k)}\|_2^2 \rightarrow \min$ . Setzen wir  $\lambda = 0$ , so wird (11.4) zur Normalengleichung für (11.3) und ist somit erfüllt:

$$(f'(x^{(k)}))^T \cdot f'(x^{(k)}) \cdot h^{(k)} = -(f'(x^{(k)}))^T \cdot f(x^{(k)})$$

Fall 2, für die optimale Lösung gilt  $\|h^{(k)}\|_2 = \rho_k$ :

Wir verwenden die Multiplikatorenregel von Lagrange.

$$g : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$h \mapsto \|f(x^{(k)}) + f'(x^{(k)}) \cdot h\|_2^2$$

$$\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$h \mapsto \|h\|_2^2$$

$h^{(k)}$  minimiert  $g(h^{(k)})$  unter der Nebenbedingung  $\varphi(h^{(k)}) = \rho_k^2$ .

# Optimierung in Trust-Region

$$\partial_i \varphi(h) = \partial_i \sum_{j=1}^n h_j^2 = 2 \cdot h_i$$

$$\begin{aligned}\partial_i g(h) &= \sum_{j=1}^n \partial_i (f_j(x^{(k)}) + f'_j(x^{(k)}) \cdot h)^2 \\ &= \sum_{j=1}^n \partial_i f_j(x^{(k)}) \cdot 2 \cdot (f_j(x^{(k)}) + f'_j(x^{(k)}) \cdot h) \\ &= 2 \cdot (f'(x^{(k)})^T \cdot (f(x^{(k)}) + f'(x^{(k)}) \cdot h))_i\end{aligned}$$

Nach der Multiplikatorenregel von Lagrange existiert dann ein  $\lambda \in \mathbb{R}$  mit:

$$\begin{aligned}\text{grad } g(h^{(k)}) &= -\lambda \cdot \text{grad } \varphi(h^{(k)}) \\ \Leftrightarrow f'(x^{(k)})^T \cdot (f(x^{(k)}) + f'(x^{(k)}) \cdot h^{(k)}) &= -\lambda \cdot h^{(k)} \\ \Leftrightarrow (f'(x^{(k)})^T \cdot f'(x^{(k)}) + \lambda \cdot I) \cdot h^{(k)} &= -f'(x^{(k)})^T \cdot f(x^{(k)})\end{aligned}$$

# Optimierung in Trust-Region

Um  $\lambda \geq 0$  zu zeigen betrachten wir die Eigenwertzerlegung  $V \cdot \Lambda \cdot V^T$  der symmetrischen, positiv semi-definiten Matrix  $f'(x^{(k)})^T \cdot f'(x^{(k)})$ . Dabei ist  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  und  $V$  orthogonal. Es gilt:

$$\begin{aligned}\|h^{(k)}\|_2^2 &= \|-(f'(x^{(k)})^T \cdot f'(x^{(k)}) + \lambda \cdot I)^{-1} \cdot f'(x^{(k)})^T \cdot f(x^{(k)})\|_2^2 \\ &= \|V \cdot (\Lambda + \lambda \cdot I)^{-1} \cdot V^T \cdot f'(x^{(k)})^T \cdot f(x^{(k)})\|_2^2 \\ &= \sum_{i=1}^n \frac{1}{(\lambda_i + \lambda)^2} \cdot (V^T \cdot f'(x^{(k)})^T \cdot f(x^{(k)}))_i^2\end{aligned}$$

Da  $\lambda_i \geq 0$  ist dieser Ausdruck streng monoton fallend in  $\lambda$ . Wenn die Gleichung also für  $\lambda < 0$  erfüllt wäre, so würde eine Lösung mit  $\|h^{(k)}\| < \rho_k$  existieren und Fall 1 greift.



# Berechnen von $h^{(k)}$

Wir können  $h^{(k)}$  also durch die folgenden Schritten berechnen:

- Löse  $\|f(x^{(k)}) + f'(x^{(k)}) \cdot h^{(k)}\|_2^2 \rightarrow \min$  ohne Nebenbedingung.

# Berechnen von $h^{(k)}$

Wir können  $h^{(k)}$  also durch die folgenden Schritten berechnen:

- Löse  $\|f(x^{(k)}) + f'(x^{(k)}) \cdot h^{(k)}\|_2^2 \rightarrow \min$  ohne Nebenbedingung.
- Falls  $\|h^{(k)}\|_2 > \rho_k$ :
  - Berechne  $\lambda$  durch Lösen der nichtlinearen Gleichung  $\|h^{(k)}\|_2 = \rho_k$  mit  $h^{(k)}$  wie in (11.4).
    - z.B. mit dem Newton-Verfahren.
  - Ermittle neues  $h^{(k)}$  mit  $\|h^{(k)}\|_2 = \rho_k$  durch Lösen des linearen Gleichungssystems (11.4).

# Armijo-Goldstein-Kriterium

- Es bleibt die Größe der Trust-Region  $\rho_k \in \mathbb{R}_+$  zu wählen.
  - Zu kleines  $\rho_k$  führt zu kleinen Schritten und langsamer Konvergenz.
  - Zu großes  $\rho_k$  kann zu divergentem Verhalten führen.

# Armijo-Goldstein-Kriterium

- Es bleibt die Größe der Trust-Region  $\rho_k \in \mathbb{R}_+$  zu wählen.
  - Zu kleines  $\rho_k$  führt zu kleinen Schritten und langsamer Konvergenz.
  - Zu großes  $\rho_k$  kann zu divergentem Verhalten führen.
- Sei  $\mathcal{E}(x) := \|f(x)\|_2^2$  das Fehlerfunktional.
- Gemäß Taylorapproximation gilt:

$$\begin{aligned}\mathcal{E}(x^{(k)} + h) - \mathcal{E}(x^{(k)}) &= \mathcal{E}'(x^{(k)}) \cdot h + o(\|h\|_2) \\ &= 2 \cdot f(x^{(k)})^T \cdot f'(x^{(k)}) \cdot h + o(\|h\|_2)\end{aligned}$$

# Armijo-Goldstein-Kriterium

- Es bleibt die Größe der Trust-Region  $\rho_k \in \mathbb{R}_+$  zu wählen.
  - Zu kleines  $\rho_k$  führt zu kleinen Schritten und langsamer Konvergenz.
  - Zu großes  $\rho_k$  kann zu divergentem Verhalten führen.
- Sei  $\mathcal{E}(x) := \|f(x)\|_2^2$  das Fehlerfunktional.
- Gemäß Taylorapproximation gilt:

$$\begin{aligned}\mathcal{E}(x^{(k)} + h) - \mathcal{E}(x^{(k)}) &= \mathcal{E}'(x^{(k)}) \cdot h + o(\|h\|_2) \\ &= 2 \cdot f(x^{(k)})^T \cdot f'(x^{(k)}) \cdot h + o(\|h\|_2)\end{aligned}$$

- Betrachte Quotient aus tatsächlicher und geschätzter Verbesserung:

$$\mu_k := \frac{\|f(x^{(k)} + h^{(k)})\|_2^2 - \|f(x^{(k)})\|_2^2}{2 \cdot f(x^{(k)})^T \cdot f'(x^{(k)}) \cdot h^{(k)}}$$

- Wähle Konstanten  $0 < \mu_- < \mu_+ < 1$ .
- $\mu_k < \mu_-$ :  $\rho_k$  muss verkleinert werden.
- $\mu_+ < \mu_k$ :  $\rho_k$  kann vergrößert werden.

# Levenberg-Marquardt-Verfahren

## Algorithmus 11.1 (Levenberg-Marquardt-Verfahren)

Gegeben:  $0 < \mu_- < \mu_+ < 1$ ,  $\rho_0 \in \mathbb{R}_+$ ,  $x^{(0)} \in D(f)$

**for**  $k = 0, 1, \dots$  **do**

Minimiere  $\|f(x^{(k)}) + f'(x^{(k)}) \cdot h^{(k)}\|_2^2$  mit  $\|h^{(k)}\| \leq \rho_k$

$$\mu_k := \frac{\|f(x^{(k)}) + h^{(k)}\|_2^2 - \|f(x^{(k)})\|_2^2}{2 \cdot f(x^{(k)})^T \cdot f'(x^{(k)}) \cdot h^{(k)}}$$

**if**  $\mu_k < \mu_-$  **then**

$$x^{(k+1)} := x^{(k)}$$

$$\rho_{k+1} := \frac{\rho_k}{2}$$

**else**

$$x^{(k+1)} := x^{(k)} + h^{(k)}$$

**if**  $\mu_k > \mu_+$  **then**

$$\rho_{k+1} := 2 \cdot \rho_k$$

**end if**

**end if**

**end for**

# Konvergenz von Levenberg-Marquardt

## Satz 11.2 (Konvergenz von Levenberg-Marquardt)

Sei  $f : D(f) \rightarrow \mathbb{R}^m$  stetig differenzierbar. Sei  $x^{(0)} \in D(f)$  und sei  $U \subset D(f)$  kompakt mit

$$\{x \in D(f) \mid \|f(x)\|_2 \leq \|f(x^{(0)})\|_2\} \subset U.$$

Dann gilt für die Iterierten des Levenberg-Marquardt-Verfahrens

$$\lim_{k \rightarrow \infty} \mathcal{E}'(x^{(k)}) = 0$$

wobei wieder  $\mathcal{E}(x) := \|f(x)\|_2^2$  das Fehlerfunktional bezeichnet.

# Konvergenz von Levenberg-Marquardt

## Satz 11.2 (Konvergenz von Levenberg-Marquardt)

Sei  $f : D(f) \rightarrow \mathbb{R}^m$  stetig differenzierbar. Sei  $x^{(0)} \in D(f)$  und sei  $U \subset D(f)$  kompakt mit

$$\{x \in D(f) \mid \|f(x)\|_2 \leq \|f(x^{(0)})\|_2\} \subset U.$$

Dann gilt für die Iterierten des Levenberg-Marquardt-Verfahrens

$$\lim_{k \rightarrow \infty} \mathcal{E}'(x^{(k)}) = 0$$

wobei wieder  $\mathcal{E}(x) := \|f(x)\|_2^2$  das Fehlerfunktional bezeichnet.

**Beweis:** Siehe "Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens", 3. Auflage von Martin Hanke-Bourgeois, S. 191 ff. □

# Stolpersteine bei iterativer Optimierung

- Der Satz sagt *nicht*, dass wir ein globales Minimum finden.
  - Die Folge  $x^{(k)}$  ist nicht zwangsläufig konvergent, nur  $\mathcal{E}'(x^{(k)})$  ist es.
  - Wenn, dann konvergiert sie zunächst nur gegen einen kritischen Punkt.
  - Selbst wenn es ein Minimum ist, muss es nicht das globale Minimum sein.

# Stolpersteine bei iterativer Optimierung

- Der Satz sagt *nicht*, dass wir ein globales Minimum finden.
  - Die Folge  $x^{(k)}$  ist nicht zwangsläufig konvergent, nur  $\mathcal{E}'(x^{(k)})$  ist es.
  - Wenn, dann konvergiert sie zunächst nur gegen einen kritischen Punkt.
  - Selbst wenn es ein Minimum ist, muss es nicht das globale Minimum sein.
- Auch hier sollte man  $x^{(0)}$  möglichst gut wählen.

# Stolpersteine bei iterativer Optimierung

- Der Satz sagt *nicht*, dass wir ein globales Minimum finden.
  - Die Folge  $x^{(k)}$  ist nicht zwangsläufig konvergent, nur  $\mathcal{E}'(x^{(k)})$  ist es.
  - Wenn, dann konvergiert sie zunächst nur gegen einen kritischen Punkt.
  - Selbst wenn es ein Minimum ist, muss es nicht das globale Minimum sein.
- Auch hier sollte man  $x^{(0)}$  möglichst gut wählen.
- Der Satz sagt nichts über die Konvergenzordnung.
- In der Praxis hat das Verfahren aber gute Konvergenzeigenschaften.

# Gradientenverfahren

Gradientenverfahren approximieren das Minimum des nichtlinearen Ausgleichsproblems  $\hat{x}$  ausgehend von einem Startwert  $x^{(0)}$  durch eine Iterationsfolge  $x^{(k)}$ , bei der sich  $x^{(k+1)}$  aus  $x^{(k)}$  durch die Wahl einer **Suchrichtung**  $d^{(k)}$  und einer **Schrittweite**  $\alpha_k > 0$  ergibt:

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}.$$

Suchrichtung und Schrittweite werden dabei so bestimmt, dass eine Abstiegsbedingung  $\Phi(x^{(k+1)}) < \Phi(x^{(k)})$  erfüllt ist.

# Methode des steilsten Abstiegs

Als Suchrichtung an der Stelle  $x^{(k)}$  bietet sich  $d^{(k)} = -\nabla\Phi(x^{(k)})$  an.  
Wegen der Cauchy-Schwarz-Ungleichung gilt für eine Suchrichtung  $d \in \mathbb{R}^n$  mit  $\|d\| = 1$

$$\frac{\partial \Phi}{\partial d}(x) = \nabla\Phi(x) \cdot d \leq \|\nabla\Phi(x)\| \|d\| = \|\nabla\Phi(x)\|$$

Setzt man  $d = \nabla\Phi(x)/\|\nabla\Phi(x)\|$ , so ergibt sich Gleichheit, d.h. die Wahl  $d^{(k)} = -\nabla\Phi(x^{(k)})$  ist optimal. Daher nennt man das resultierende Verfahren auch die **Methode des steilsten Abstiegs**.

# Methode des steilsten Abstiegs

Mit  $\Phi(x) = \frac{1}{2}f(x) \cdot f(x)$  ergibt sich für das nichtlineare Ausgleichsproblem

$$\nabla \Phi(x) = \sum_{j=1}^m f_j(x) \nabla f_j(x) = f'(x)^t f(x)$$

und wir erhalten die Iterationsfolge

$$x^{(k+1)} = x^{(k)} - \alpha_k f'(x^{(k)})^t f(x^{(k)})$$

# Rosenbrockfunktion

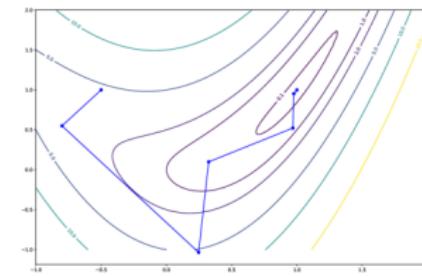
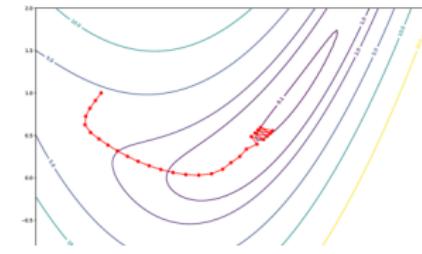
## Beispiel (Rosenbrockfunktion)

Ein klassisches Beispiel zum Testen von Optimierungsverfahren ist die *Rosenbrockfunktion*

$$\Phi(x) = \begin{pmatrix} 4(x_2 - x_1^2) \\ x_2 - 1 \end{pmatrix}$$

Wie an den Niveaulinien zu sehen ist, handelt es sich um ein langes bananenförmiges Tal mit einem Minimum bei  $(1, 1)^t$ .

Oben: Steilster Abstieg mit konstanter Schrittweite und . Die Gradienten von  $\Phi$  stehen senkrecht auf den elliptisch geformten Niveaulinien dieser Funktion und weisen nur in Ausnahmefällen in eine effiziente Abstiegsrichtung. Unten: Iterationen des Newton-Verfahrens. Jeweils mit Startwert  $x^{(0)} = (-0.5, 1)^t$ .



# Wahl der Suchrichtung

Die pro Schritt optimale Abstiegsrichtung braucht auf lange Sicht nicht optimal sein. In der Praxis werden daher auch alternative Suchrichtungen um den negativen Gradienten verwendet:

$$c \|\nabla \Phi(x^{(k)})\|_2 \leq \|d^{(k)}\|_2 \leq C \|\nabla \Phi(x^{(k)})\| \quad (11.5)$$

mit festen Konstanten  $c, C > 0$  und

$$\cos \angle(\nabla \Phi(x^{(k)}), d^{(k)}) = \frac{\nabla \Phi(x^{(k)})^t d^{(k)}}{\|\nabla \Phi(x^{(k)})\|_2 \|d^{(k)}\|_2} \leq -\delta \quad (11.6)$$

mit festem  $\delta \in (0, 1]$ .

# Wahl der Schrittweite

Zur Steuerung der Schrittweite ist es naheliegend, in jedem Schritt das eindimensionale Optimierungsproblem

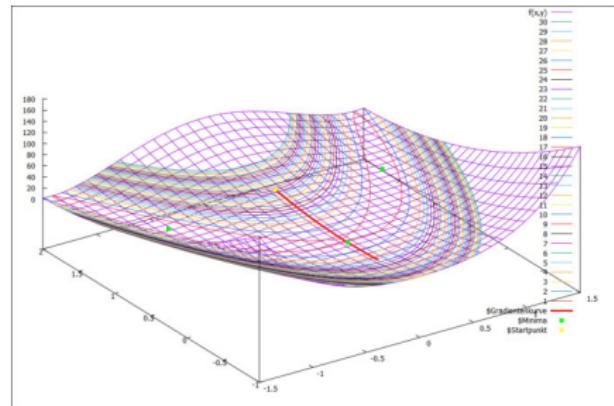
$$\min_{\alpha \in [0, \infty)} \Phi(x^{(k)} + \alpha d^{(k)}) \quad (11.7)$$

exakt zu lösen und dazu eine Nullstelle  $\hat{\alpha} \in [0, \infty)$  von

$$\nabla \Phi(x^{(k)} + \alpha d^{(k)})^t \cdot d^{(k)} \quad (11.8)$$

zu suchen.

Ein solches  $\hat{\alpha}$  existiert, wenn wir voraussetzen, dass die Menge  $M_0 = \{x \in \mathbb{R}^n | \Phi(x) \leq \Phi(x_0)\}$  kompakt ist. Falls  $\Phi$  zweimal stetig differenzierbar ist, kann man dazu z.B. das Newton-Verfahren einsetzen. Allerdings ist dabei nicht garantiert, dass wir die nächste Nullstelle finden.



# Wahl der Schrittweite

## Armijo-Schrittweiten

Ein in der Praxis angewandtes Verfahren verwendet *Armijo-Schrittweiten*. Dabei wird ausgehend von einer maximalen Schrittweite  $\alpha = 1$  die Schrittweite sukzessive um einen Faktor  $\beta$ , z.B.  $\beta = 1/2$ , verkleinert, bis die *Armijo-Goldstein-* Abstiegsbedingung erfüllt ist:

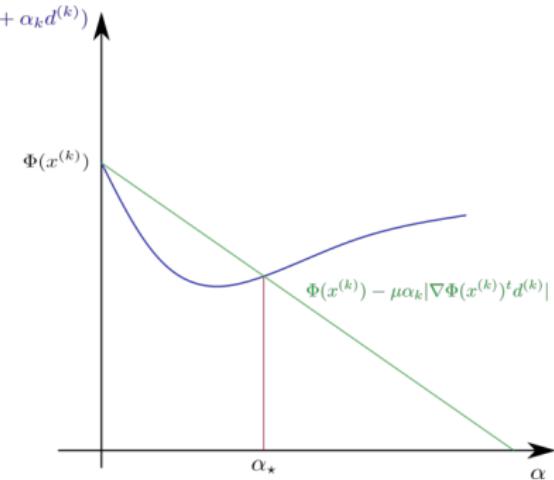


Abbildung 11.21: Armijo-Schrittweiten

$$\Phi(x^{(k)} + \alpha d^{(k)}) \leq \Phi(x^{(k)}) - \mu \alpha_k |\nabla \Phi(x^{(k)})^t d^{(k)}| \quad (11.9)$$

# Algorithmus steilster Abstieg

## Algorithmus 11.2 (Allgemeines Abstiegsverfahren)

**Input:** Eine Funktion  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$  und ein  $\mu \in (0, 1)$ ,  $\mu$  etwa 0.5  
wähle  $x^{(0)}$  und Suchrichtung  $d^{(0)}$

**for**  $k = 0, 1, 2, \dots$  **do**

$\alpha = 1$  // Schrittweite initialisieren

**while**  $\Phi(x^{(k)} + \alpha d^{(k)}) > \Phi(x^{(k)}) - \mu\alpha \|\text{grad}\Phi(x^{(k)}) * d^{(k)}\|$  **do**

$\alpha = \alpha/2$

**end while**

$x^{(k+1)} = x^{(k)} + \alpha d^{(k)}$

    wähle neue Suchrichtung  $d^{(k+1)}$  unter Beachtung der  
    Einschränkungen 11.5, 11.6.

**end for**

# Konvergenz von Abstiegsverfahren

## Satz 11.3 (Konvergenz von Abstiegsverfahren)

Die Funktion  $f$  sei in einer offenen Menge  $\mathcal{U} \subset \mathcal{D}(f)$  stetig differenzierbar mit Lipschitz-stetiger Ableitung  $f'$ . Ferner enthalte  $\mathcal{U}$  den Startvektor  $x^{(0)}$  sowie die gesamte Menge

$$\mathcal{M}(x_0) = \{x \in \mathcal{D}(f) : \Phi(x) \leq \Phi(x^{(0)})\}.$$

Falls die Suchrichtung  $d^{(k)}$  in jedem Iterationsschritt die Bedingung 11.5, 11.6 erfüllt, dann gilt für die Iterierten von Algorithmus 10, daß

$$\nabla \Phi(x^{(k)}) \rightarrow 0, \quad k \rightarrow \infty.$$

# Konvergenz von Abstiegsverfahren

Der Satz besagt nicht, dass die Folge  $x^{(k)}$  konvergent ist. Im allgemeinen konvergieren die Iterierten lediglich gegen eine stationären Punkt von  $\Phi$ . Hat  $\Phi$  nur einen stationären Punkt  $\hat{x}$  und ist die Menge  $\mathcal{M}(x_0)$  zudem beschränkt und  $f$  hinreichend glatt, dann konvergiert die Folge  $\{x^{(k)}\}$  gegen das Minimum  $\hat{x}$ .

# Konvergenz von Abstiegsverfahren

Der Satz besagt nicht, dass die Folge  $x^{(k)}$  konvergent ist. Im allgemeinen konvergieren die Iterierten lediglich gegen eine stationären Punkt von  $\Phi$ . Hat  $\Phi$  nur einen stationären Punkt  $\hat{x}$  und ist die Menge  $\mathcal{M}(x_0)$  zudem beschränkt und  $f$  hinreichend glatt, dann konvergiert die Folge  $\{x^{(k)}\}$  gegen das Minimum  $\hat{x}$ .

**Beweis:** Siehe "Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens", 3. Auflage von Martin Hanke-Bourgeois, S. 181 f.



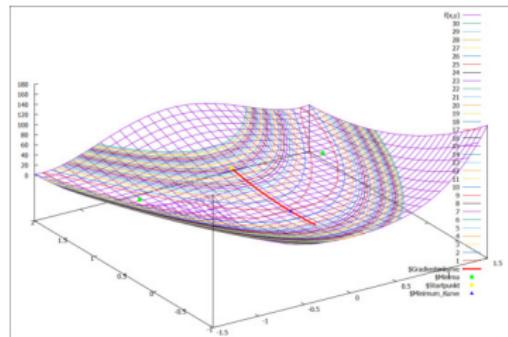
# Abstiegsverfahren

Gegeben sei die Funktion  $f(x_1, x_2) = (x_1, x_2^2 - 1, x_1(x_2^2 - 1))^t$  für das nichtlineare Ausgleichsproblem, s.d.

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

Das Minimum von  $\Phi$  ist Null und wird für  $x_1 = 0$  und  $x_2 = \pm 1$  angenommen. Hat eine Iterierte  $x^{(k)}$  des Algorithmus 10 die Koordinaten  $(x_1, 0)^t$ , so ergibt sich

$$d^{(k)} = -(2x_1, 0)^t = -2x^{(k)}.$$



Nach Satz 11.3 konvergiert daher  $x^{(k)}$  gegen Null für  $k \rightarrow \infty$ . Der Nullpunkt ist aber nur Sattelpunkt von  $\Phi$ !

# Kapitel 12:

## Abschlussbemerkungen

## 22 Differenzierbare Funktionen

- Darstellung des Differentials durch Richtungsableitungen
- Hauptkriterium für Differenzierbarkeit
- Orthogonalität von Gradient und Nullmenge
- Differentiale höherer Ordnung
- Die Taylor-Approximation
- (Geometrische) Bedeutung der zweiten Ableitung
- Lokale Minima und Maxima

## 23 Differenzierbare Abbildungen

- Funktionalmatrix
- Differenzierbarkeitskriterium für Abbildungen
  - Ein Beispiel: Tangentialvektoren
- Extrema unter Nebenbedingungen

## 24 Nichtlineare Gleichungen

- Konvergenzbegriffe
  - Iterative Verfahren zur Lösung linearer Gleichungssysteme
- Konvergenzgeschwindigkeit einer Folge
- Nullstellenbestimmung reeller Funktionen
  - Das eindimensionale Newtonverfahren
  - Das Newtonverfahren im  $\mathbb{C}^n$

## 25 Nichtlineare Ausgleichsprobleme

## 26 Abschlussbemerkungen

# Was wir *nicht* gelernt haben

- Numerische Integration
- Numerische Lösung gewöhnlicher und partieller Differentialgleichungen

# Was wir *nicht* gelernt haben

- Numerische Integration
- Numerische Lösung gewöhnlicher und partieller Differentialgleichungen
- Interpolation durch stückweise definierte Funktionen
- Effiziente Algorithmen für dünn besetzte Matrizen
- Weitere Optimierungsalgorithmen

# Was wir *nicht* gelernt haben

- Numerische Integration
- Numerische Lösung gewöhnlicher und partieller Differentialgleichungen
- Interpolation durch stückweise definierte Funktionen
- Effiziente Algorithmen für dünn besetzte Matrizen
- Weitere Optimierungsalgorithmen
- Implementierungen für riesige Datenmengen:
  - Out-of-Core und auf der GPU
- ...

# Was wir *nicht* gelernt haben

- Numerische Integration
- Numerische Lösung gewöhnlicher und partieller Differentialgleichungen
- Interpolation durch stückweise definierte Funktionen
- Effiziente Algorithmen für dünn besetzte Matrizen
- Weitere Optimierungsalgorithmen
- Implementierungen für riesige Datenmengen:
  - Out-of-Core und auf der GPU
  - ...
- Wo gibt es denn sowas?
  - Foundations of Graphics
  - Spezialvorlesungen der Mathematiker
  - Nach Bedarf in anderen Veranstaltungen
  - In der Literatur



Martin Hanke-Bourgeois.

*Grundlagen der numerischen Mathematik und des wissenschaftlichen Rechnens.*

Springer, 2009.



Konrad Königsberger.

*Analysis 2.*

Springer Verlag, 1997.



Lloyd N. Trefethen and David Bau III.

*Numerical Linear Algebra.*

Society for Industrial and Applied Mathematics, Philadelphia, 1997.