

# Höhere Mathematik 1

## Rechnerarithmetik

### Gleitkommazahlen / Maschinenzahlen

- Darstellung:  
 $M = \{x \in \mathbb{R} | x = \pm 0.m_1m_2m_3 \dots m_n \cdot B^{e_{l-1} \dots e_1 e_0 - bias}\} \cup \{0\}$
- Basis  $B$ , Mantissenlänge  $n$ , Exponentenbereich  $[e_{\min}, e_{\max}]$
- Normalisierte Zahl:  $m_1 \neq 0$
- Bias:  $bias = B^{l-1} - 1$  für  $l$ -stellige Exponentendarstellung

## Fehleranalyse

- Gerundete Zahl:  $\tilde{x} = rd(x)$
- Absoluter Rundungsfehler:  $|\tilde{x} - x| \leq \frac{1}{2}B^{e-n+1}$
- Relative Rundungsfehler:  $\left| \frac{\tilde{x}-x}{x} \right| \leq \text{eps}$
- Maschinengenauigkeit:  $\text{eps} = \max_{x \in M} \left| \frac{rd(x)-x}{x} \right| = \frac{1}{2}B^{1-n}$

## Fehlerfortpflanzung bei Funktionen

- Absoluter Fehler:  $|f(\tilde{x}) - f(x)| \approx |f'(x)| \cdot |\tilde{x} - x|$
- Relativer Fehler:  $\left| \frac{f(\tilde{x})-f(x)}{f(x)} \right| \approx \left| \frac{x \cdot f'(x)}{f(x)} \right| \cdot \left| \frac{\tilde{x}-x}{x} \right|$
- Kondition (Je kleiner desto besser):  $K = \left| \frac{x \cdot f'(x)}{f(x)} \right|$

## Nullstellenprobleme

### Fixpunktiteration

- Fixpunktgleichung:  $F(x) = x$
- Fixpunkte ( $\bar{x}$ ) sind Nullstellen von  $f(x) = F(x) - x$
- Iterationsvorschrift:  $x_{n+1} = F(x_n)$
- Konvergenz:  $|F'(\bar{x})| < 1 \rightarrow$  anziehender Fixpunkt
- Divergenz:  $|F'(\bar{x})| > 1 \rightarrow$  abstoßender Fixpunkt

### Banach'scher Fixpunktsatz

- Sei  $F : [a, b] \rightarrow [a, b]$  und es existiert eine Konstante  $0 < \alpha < 1$  mit  
$$|F(x) - F(y)| \leq \alpha |x - y| \quad \forall x, y \in [a, b]$$
- Dann besitzt  $F$  genau einen Fixpunkt  $\bar{x} \in [a, b]$  und die Iteration  $x_{n+1} = F(x_n)$  konvergiert für jeden Startwert  $x_0 \in [a, b]$  gegen  $\bar{x}$ . (Lipschitz-Stetigkeit mit Lipschitz-Konstante  $\alpha$ )
- a-priori-Fehlerabschätzung:

$$|x_n - \bar{x}| \leq \frac{\alpha^n}{1 - \alpha} |x_1 - x_0|$$

- a-posteriori-Fehlerabschätzung:

$$|x_n - \bar{x}| \leq \frac{\alpha}{1 - \alpha} |x_n - x_{n-1}|$$

- Umformung:  $\left| \frac{F(x)-F(y)}{x-y} x - y \right| \leq \alpha$

## Newton-Verfahren

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Für den Startwert  $x_0$  muss gelten, dass  $f(x_0) \neq 0$  und  $f'(x_0) \neq 0$ . Ein geeigneter Startwert kann durch die Nullstellenbestimmung der Ableitungsfunktion gefunden werden.

### Vereinfachte Newton-Verfahren

- Fixierter Ableitungswert:  $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}$
- Sekantenverfahren:  $x_{n+1} = x_n - f(x_n) \cdot \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$

## Fehlerabschätzung

### Konvergenzordnung

- Sei  $(x_n)$  eine konvergierende Folge gegen  $\bar{x}$
- Die Folge hat die Konvergenzordnung  $q$ , wenn eine Konstante  $c > 0$  existiert, so dass

$$|x_{n+1} - \bar{x}| \leq c \cdot |x_n - \bar{x}|^q$$

- $q = 1$  lineare Konvergenz,  $q = 2$  quadratische Konvergenz

### Nullstellensatz von Bolzano

- Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig mit  $f(a) \cdot f(b) < 0$  oder  $f(a) \geq 0 \geq f(b)$
- Dann existiert mindestens ein  $\bar{x} \in (a, b)$  mit  $f(\bar{x}) = 0$

## Lineare Gleichungssysteme

### LR-Zerlegung

- Zerlegung einer Matrix  $A$  in das Produkt  $A = L \cdot R$
- $L$ : untere Dreiecksmatrix mit Einsen auf der Diagonale
- $R$ : obere Dreiecksmatrix
- Lösen von  $A \cdot x = b$  durch Lösen von  $L \cdot y = b$  und  $R \cdot x = y$
- Voraussetzung: Keine Null-Pivotelemente während der Zerlegung / kein Zeilentausch notwendig (sonst  $P \cdot A = L \cdot R$ )

### Beispiel

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -3 & -2 \\ 5 & 1 & 4 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -5 & -3 & 1 \end{pmatrix}}_L \cdot \underbrace{\begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & 6 \end{pmatrix}}_R$$

### Permutationsmatrix

- Vertauscht Zeilen oder Spalten einer Matrix
- $P \cdot A$ : Vertauschen der Zeilen von  $A$
- $A \cdot P$ : Vertauschen der Spalten von  $A$
- $P_1, P_2$  sind Permutationsmatrizen  $\Rightarrow P = P_1 \cdot P_2$  ist auch eine Permutationsmatrix
- $P^{-1} = P^T$

## QR-Zerlegung

- Zerlegung einer Matrix  $A$  in das Produkt  $A = Q \cdot R$
- $Q$ : orthogonale Matrix ( $Q^T \cdot Q = I$ )
- $R$ : obere Dreiecksmatrix

### Householder-Matrizen

- $H$  ist orthogonal und symmetrisch ( $H^T = H$ )
- $H$  spiegelt an der Hyperebene orthogonal zu  $u$
- $H = H^T = H^{-1} \Rightarrow H \cdot H = I_n$
- Wenn  $u$  ein normierter Vektor ist, dann gilt:

$$H = I_n - 2uu^T$$

### QR-Zerlegung mit Householder-Matrizen

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 4 & -2 & 6 \\ 3 & 1 & 0 \end{pmatrix}, b = \begin{pmatrix} 9 \\ -4 \\ 9 \end{pmatrix}$$

1.  $v_1 = a_1 + \text{sign}(a_{11}) \cdot |a_1| \cdot e_1$
2.  $u_1 = \frac{v_1}{\|v_1\|_2}$
3.  $H_1 = I - 2u_1u_1^T$
4.  $A^{(1)} = H_1 \cdot A, b^{(1)} = H_1 \cdot b$

$$A^{(1)} = \begin{pmatrix} -5.0990 & 0.5883 & -4.5107 \\ 0 & -2.9258 & 3.6976 \\ 0 & 0.3056 & -1.7268 \end{pmatrix}$$

Rekursiv mit  $A^{(1)}$  ( $2 \times 2$ -Matrix) durchführen bis  $A^{(n-1)}$  obere Dreiecksmatrix ist.

### Vektor- und Matrixnormen

- Vektornormen:

$$\text{1-Norm, Summennorm: } \|x\|_1 = \sum_{i=1}^n |x_i|$$

$$\text{2-Norm, euklidische Norm: } \|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$$

$$\text{Unendlich-Norm, Maximumnorm: } \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

- Matrixnormen ( $n \times m$ -Matrizen):

$$\text{1-Norm, Spaltensummennorm: } \|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|$$

$$\text{2-Norm, Spektralnrm: } \|A\|_2 = \sqrt{\rho(A^T A)}$$

$$\text{Unendlich-Norm, Zeilensummennorm: } \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|$$

### Spektralradius

- $\rho(A) = \max\{|\lambda| | \lambda \text{ ist Eigenwert von } A\}$
- Für jede Matrixnorm gilt:  $\rho(A) \leq \|A\|$

Fehlerrechnung Vektoren

- $\|\cdot\|$ : gewählte Matrixnorm
- Konditionszahl:  $cond(A) = \|A\| \cdot \|A^{-1}\|$
- Abschätzung des relativen Lösungsfehlers:

$$\frac{\|\tilde{x} - x\|}{\|x\|} \leq cond(A) \cdot \frac{\|\tilde{b} - b\|}{\|b\|}, \text{ falls } \|b\| \neq 0$$

- Je größer  $cond(A)$ , desto schlechter ist das LGS konditioniert

Fehlerrechnung Matrizen

- Gegeben:  $\tilde{A} = A + E$  mit  $E$  als Störmatrix ( $E = A - \tilde{A}$ )
- Abschätzung des relativen Lösungsfehlers:

$$\frac{\|\tilde{x} - x\|}{\|x\|} \leq \frac{cond(A)}{1 - cond(A) \cdot \frac{\|E\|}{\|A\|}} \cdot \left( \frac{\|E\|}{\|A\|} + \frac{\|b - \tilde{b}\|}{\|b\|} \right)$$

- Gilt nur, wenn  $cond(A) \cdot \frac{\|E\|}{\|A\|} < 1$

Aufwandsabschätzung

- LR-Zerlegung:  $\frac{2}{3}n^3$  Operationen
- Vorwärts- und Rückwärtseinsetzen:  $2n^2$  Operationen
- QR-Zerlegung mit Householder:  $2n^3$  Operationen
- Vorwärts- und Rückwärtseinsetzen:  $2n^2$  Operationen
- Allgemein:

$$\sum_{i=1}^n i = \frac{n(n+1)}{2} \approx \frac{n^2}{2}$$

und

$$\sum_{i=1}^n i^2 = \frac{1}{6}n + \frac{1}{2}n^2 + \frac{1}{3}n^3 \approx \frac{1}{3}n^3$$

Iterative Verfahren für LGS

Jacobi-Verfahren

- Zerlegung: D ist die Diagonalmatrix von A, L die untere Dreiecksmatrix und R die obere Dreiecksmatrix

$$\begin{aligned} Ax &= b \\ (L + D + R)x &= b \\ Dx &= -(L + R)x + b \\ x &= D^{-1}(L + R)x + D^{-1}b \end{aligned}$$

- Iterationsvorschrift:  $x^{(k+1)} = D^{-1}(L + R)x^{(k)} + D^{-1}b$
- Iterationsmatrix:  $B_J = D^{-1}(L + R)$

Gauß-Seidel-Verfahren

- Iterationsvorschrift:  $x^{(k+1)} = (D - L)^{-1}Rx^{(k)} + (D - L)^{-1}b$
- Iterationsmatrix:  $B_{GS} = (D - L)^{-1}R$

Konvergenz und Abschätzung

- Beide Verfahren konvergieren, wenn  $\|B\| < 1$
- Symmetrisch positiv definite Matrizen (nur Gauß-Seidel)
- Siehe Fixpunktiteration ( $\alpha = \|B\|$ )

Diagonaldominanz

- Eine Matrix ist diagonaldominant wenn eines der folgenden Kriterien erfüllt ist:
  - $\forall i = 1, \dots, n : |a_{ii}| > \sum_{j=1, j \neq i} |a_{i,j}|$   
(Zeilensummenkriterium)
  - $\forall j = 1, \dots, n : |a_{jj}| > \sum_{i=1, i \neq j} |a_{i,j}|$   
(Spaltensummenkriterium)
- Beide Verfahren konvergieren für diagonaldominante Matrizen

Komplexe Zahlen

Darstellung

- Normalform / kartesische Form:  $z = x + iy, x, y \in \mathbb{R}$
- Polardarstellung / Trigonometrische Form:  $z = r(\cos \varphi + i \sin \varphi)$
- Eulersche Darstellung:  $z = re^{i\varphi}$
- Betrag:  $|z| = r = \sqrt{x^2 + y^2}$
- Argument:  $\varphi = \arg(z) = \tan^{-1}(\frac{y}{x}) = \cos^{-1}(\frac{x}{r}) = \sin^{-1}(\frac{y}{r})$

Rechenregeln

- Addition:  $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$
- Multiplikation:  $z_1 \cdot z_2 = (x_1 + iy_1)(x_2 + iy_2)$
- Multiplikation in Polardarstellung:

$$z_1 \cdot z_2 = r_1 r_2 [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)]$$

- Division:

$$\frac{z_1}{z_2} = \frac{z_1 \cdot z_2^*}{z_2 \cdot z_2^*}, \quad z_2 \neq 0$$

- Division in Polardarstellung:

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} [\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)], \quad z_2 \neq 0$$

- Potenzieren in der Polardarstellung:

$$z^n = r^n [\cos(n\varphi) + i \sin(n\varphi)] = r^n e^{in\varphi}$$

Wurzeln

- $z = r(\cos \varphi + i \sin \varphi) = re^{i\varphi}$
- n-te Wurzel:

$$\begin{aligned} z_k &= r^{1/n} \left[ \cos \left( \frac{\varphi + 2k\pi}{n} \right) + i \sin \left( \frac{\varphi + 2k\pi}{n} \right) \right] = re^{i \frac{\varphi + 2k\pi}{n}} \\ k &= 0, 1, \dots, n - 1 \end{aligned}$$