Formelsammlung Numerische Mathematik für Maschinenbauer – SS 2018

- **Maschinengenauigkeit:** Es sei $\mathbb{M}(b, m, r, R)$ die Menge der Maschinenzahlen. Dann ist eps $:= \frac{b^{1-m}}{2}$ die relative Maschinengenauigkeit.

- **Fehlerverstärkung:** Es sei $f : \mathbb{R}^n \to \mathbb{R}$ hinreichend oft differenzierbar. Dann gilt:

$$\frac{f(\tilde{x}) - f(x)}{f(x)} \doteq \sum_{j=1}^{n} \phi_j(x) \cdot \frac{\tilde{x}_j - x_j}{x_j} \tag{1}$$

mit

$$\phi_j(x) = \frac{\partial f(x)}{\partial x_j} \cdot \frac{x_j}{f(x)}, \tag{2}$$

bzw.

$$\underbrace{\left|\frac{f(\tilde{x}) - f(x)}{f(x)}\right|}_{\text{rel. Fehler der Ausgabe}} \quad \dot{\leq} \quad \underbrace{\kappa_{\text{rel}}(x)}_{\text{Fehlerverstärkung}} \quad \underbrace{\sum_{j=1}^{n} \left|\frac{\tilde{x}_j - x_j}{x_j}\right|}_{\text{rel. Fehler der Eingabe}} \tag{3}$$

mit

$$\kappa_{\text{rel}}(x) = \kappa_{\text{rel}}^{\infty}(x) = \max_j |\phi_j(x)|. \tag{4}$$

- **Störungssätze:** Für $A \in \mathbb{R}^{n \times n}$ mit $\det(A) \neq 0$ sei $\kappa(A) := \kappa_{\|\cdot\|}(A) := \|A^{-1}\| \, \|A\|$.

  a) Nur die rechte Seite b ist gestört, d.h. $x + \Delta x$ ist die Lösung von $A(x + \Delta x) = b + \Delta b$. Dann gilt

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \, \frac{\|\Delta b\|}{\|b\|} \tag{5}$$

  Sei nun $\tilde{x}$ die Annäherung der Lösung x und $\tilde{r} = b - A\tilde{x}$ das Residuum. Dann gilt:

$$\kappa(A)^{-1} \frac{\|\tilde{r}\|}{\|b\|} \leq \frac{\|x - \tilde{x}\|}{\|x\|} \leq \kappa(A) \frac{\|\tilde{r}\|}{\|b\|} \tag{6}$$

  b) Es gelte $\kappa(A) \frac{\|\Delta A\|}{\|A\|} < 1$. Es sei $x + \Delta x$ die Lösung von $(A + \Delta A)(x + \Delta x) = b + \Delta b$. Dann gilt:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A) \frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right) \tag{7}$$

- **Cholesky-Verfahren:** $(A = L\,D\,L^T)$ Für aufeinanderfolgende Spalten, $k = 1, 2 \ldots, n$ erhält man für $d_{k,k}$ und $l_{i,k}$ $(i > k)$:

$$d_{k,k} = a_{k,k} - \sum_{j=1}^{k-1} l_{k,j}^2 d_{j,j}, \qquad l_{i,k} = \left(a_{i,k} - \sum_{j=1}^{k-1} l_{i,j} d_{j,j} l_{k,j} \right) / d_{k,k} \tag{8}$$

- **Givens-Rotation:** Grundaufgabe: Gegeben $(a, b)^T \in \mathbb{R}^2 \setminus \{0\}$. Finde $c, s \in \mathbb{R}$ mit

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}, \qquad \text{Lösung: } r := \pm\sqrt{a^2 + b^2},\ c := \frac{a}{r},\ s := \frac{b}{r} \tag{9}$$

- **Householder-Spiegelung:** Grundaufgabe: Zu $y \in \mathbb{R}^n$, $y \notin \text{span}(e^1)$, finde $v \in \mathbb{R}^n$ so, dass

$$Q_v y = \pm \|y\|_2 e^1 \text{ mit } Q_v = I - 2\frac{vv^T}{v^T v} \tag{10}$$

Lösung: mit $\alpha = \text{sign}(y_1)\|y\|_2$, $v = y + \alpha e^1$ gilt $Q_v y = -\alpha e^1$

- **Linearer Ausgleich:** ($A \in \mathbb{R}^{m\times n}$ mit Rang$(A) = n$: Finde x^* mit $\|A\,x^* - b\|_2 = \min_{x\in\mathbb{R}^n} \|A\,x - b\|_2$)

$$\kappa_2(A) := \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \Big/ \min_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \tag{11}$$

  Für die Kondition bezüglich Störungen in b gilt:

$$\frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \leq \frac{\kappa_2(A)}{\cos(\Theta)} \frac{\|\tilde{b} - b\|_2}{\|b\|_2}. \tag{12}$$

  Dabei ist Θ der Winkel zwischen Ax^* und b.

- **Banachscher Fixpunktsatz**
  Es sei X ein linear normierter Raum mit Norm $\|\cdot\|$. $E \subset X$ sei eine vollständige Teilmenge von X. Die Abbildung Φ sei eine Selbstabbildung auf E: $\Phi : E \to E$. Ferner sei Φ eine Kontraktion auf E: $\|\Phi(x) - \Phi(y)\| \leq L\|x - y\|$ für alle $x, y \in E$, mit $L < 1$. Dann gilt:
  1. Es existiert genau ein Fixpunkt x^* von Φ in E.
  2. Für beliebiges $x_0 \in E$ konvergiert $x_{k+1} = \Phi(x_k), k = 0, 1, 2, ...$ gegen den Fixpunkt x^*.
  3. A-priori-Fehlerabschätzung: $\|x_k - x^*\| \leq \frac{L^k}{1-L}\|x_1 - x_0\|$
  4. A-posteriori-Fehlerabschätzung: $\|x_k - x^*\| \leq \frac{L}{1-L}\|x_k - x_{k-1}\|$

  Hinreichendes Kriterium: Sei $X = \mathbb{R}^n$, E abgeschlossen und konvex sowie $\Phi \in C^1(E)$ mit $\Phi(E) \subset E$ und $\max_{x\in E} \|\Phi'(x)\| \leq L < 1$. Dann sind die Voraussetzungen des Banachschen Fixpunktsatzes erfüllt.

- **Newtonverfahren:** $x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k)$

  Durchführung des Newtonverfahrens für Systeme:
  Gegeben: Startwert x^0. Für $k = 0, 1, 2, \ldots$:
  - Löse $f'(x^k)s^k = -f(x^k)$
  - Setze $x^{k+1} = x^k + s^k$

- **Nichtlineare Ausgleichsrechnung:** gegeben $F : \mathbb{R}^n \to \mathbb{R}^m$ $(n \leq m)$ stetig differenzierbar, finde x^* mit $\|F(x^*)\|_2 = \min_{x\in\mathbb{R}^n} \|F(x)\|_2$.

  Gauß-Newton-Verfahren: Startwert x^0, für $k = 0, 1, 2, \ldots$:
  - Finde s^k mit minimaler 2-Norm so, dass $\|F'(x^k)\, s^k + F(x^k)\|_2 = \min_{s\in\mathbb{R}^n} \|F'(x^k)s + F(x^k)\|_2$.
  - Setze $x^{k+1} = x^k + s^k$.

  Falls $F'(x_k)$ vollen Rang hat, entspricht dies der Iterationsvorschrift

$$x_{k+1} = x_k - [F'(x_k)^T F'(x_k)]^{-1} F'(x_k)^T F(x_k)$$

  Levenberg-Marquardt-Verfahren: Startwert x^0, für $k = 0, 1, 2, \ldots$:
  - Finde s^k sodass $\|F'(x^k)\, s^k + F(x^k)\|_2^2 + \mu^2\|s^k\|_2^2 = \min_{s\in\mathbb{R}^n} \|F'(x^k)s + F(x^k)\|_2^2 + \mu^2\|s\|_2^2$ mit zu wählendem Parameter $\mu > 0$.
  - Setze $x^{k+1} = x^k + s^k$.

  Dies führt auf die Iterationsvorschrift $x_{k+1} = x_k - [F'(x_k)^T F'(x_k) + \mu^2 I]^{-1} F'(x_k)^T F(x_k)$

- **Interpolation**

Lagrange-Interpolationspolynom
Interpolationsproblem: zu gegebenen paarweise verschiedenen Stützstellen $x_0, \ldots, x_n$ und gegebenen Daten $f(x_0)$, …, $f(x_n)$, finde $P_n \in \Pi_n$ mit $P_n(x_j) = f(x_j)$, $j = 0, ..., n$.
Lagrange-Darstellung: $P_n(x) = \sum_{j=0}^{n} f(x_j) l_{j,n}(x)$, mit $l_{j,n}(x) = \prod_{k=0, k\neq j}^{n} \frac{x - x_k}{x_j - x_k}$

Neville-Aitken

$$P(f|x_0, ..., x_n)(x) = \frac{x - x_0}{x_n - x_0} P(f|x_1, ..., x_n)(x) + \frac{x_n - x}{x_n - x_0} P(f|x_0, ..., x_{n-1})(x) \qquad (13)$$

Rekursionsformel:

$$P_{i,0} := f(x_i) \qquad P_{i,k} := P(f|x_{i-k}, ..., x_i)(x), \quad 0 \leq k \leq i \leq n, \qquad (14)$$

$$P_{i,k} = P_{i,k-1} + \frac{P_{i,k-1} - P_{i-1,k-1}}{x_i - x_{i-k}}(x - x_i) \qquad \rightarrow \; P_{n,n} = P_n(x) \qquad (15)$$

Newtonsche Interpolationsformel

$$\begin{aligned} P(f|x_0, \ldots, x_n)(x) = [x_0]f + (x - x_0)[x_0, x_1]f + (x - x_0)(x - x_1)[x_0, x_1, x_2]f + \ldots \\ + (x - x_0) \cdots (x - x_{n-1})[x_0, \ldots, x_n]f \end{aligned} \qquad (16)$$

Dividierte Differenzen

$$p_{i,0} := [x_i]f = f(x_i) \qquad (17)$$

$$[x_0, ..., x_n]f = \frac{[x_1, ..., x_n]f - [x_0, ..., x_{n-1}]f}{x_n - x_0} \qquad (18)$$

$$p_{i,k} = \frac{p_{i,k-1} - p_{i-1,k-1}}{x_i - x_{i-k}} \qquad (19)$$

Eigenschaft $[x_i, ..., x_{k+i}]\, f = \frac{f^{(k)}(\xi)}{k!}$.

Horner-Schema
Gegeben: $p_{0,0}, ..., p_{n,n}$, x.

- $b_n := p_{n,n}$
- Für $k = n-1, n-2, ..., 0$:
  $b_k = p_{k,k} + (x - x_k)\, b_{k+1}$

Dann ist $P_n(x) = b_0$.

Interpolationsfehler
Seien $x_0, \ldots, x_n$ Stützstellen, $a := \min\{x_0, \ldots, x_n\}$, $b := \max\{x_0, \ldots, x_n\}$ und $x \in \mathbb{R}$.
Sei $I := [\min\{a, x\}, \max\{b, x\}]$. Für $f \in C^{n+1}(I)$ existiert $\xi \in I$ so, dass

$$f(x) - P(f|x_0, \ldots, x_n)(x) = (x - x_0) \cdots (x - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!}. \qquad (20)$$

Für $\hat{x} \in [a, b]$ gilt:

$$|f(\hat{x}) - P(f|x_0, ..., x_n)(\hat{x})| \leq \left| \prod_{j=0}^{n} (\hat{x} - x_j) \right| \max_{y \in [a,b]} \frac{\left| f^{(n+1)}(y) \right|}{(n+1)!} \qquad (21)$$

- **Numerische Integration**

Trapezregel:
Es sei $f \in C^2([t_{k-1}, t_k])$. Es gilt:
$\frac{h}{2}\left[f(t_{k-1}) + f(t_k)\right] = \int_{t_{k-1}}^{t_k} f(x)dx + \frac{f''(\xi_k)}{12}h^3$
für ein $\xi_k \in [t_{k-1}, t_k]$

summierte Trapezregel:
$T(h) = h\left[\frac{1}{2}f(a) + f(t_1) + \cdots + f(t_{n-1}) + \frac{1}{2}f(b)\right]$
mit $n\,h = b - a$, für $f \in C^2([a,b])$:
$\left|T(h) - \int_a^b f(x)dx\right| \leq \frac{b-a}{12}\,h^2 \max_{x\in[a,b]} |f''(x)|$

summierte Simpson-Regel:
$S(h) = \frac{h}{6}\left[f(t_0) + 4f(\frac{t_0+t_1}{2}) + 2f(t_1) + 4f(\frac{t_1+t_2}{2}) + 2f(t_2) + ... + 2f(t_{n-1}) + 4f(\frac{t_{n-1}+t_n}{2}) + f(t_n)\right]$

Newton-Cotes-Formeln
$h = d - c$,
für $m = 0$: $x_0 = c + \xi_0 h$ mit $\xi_0 := \frac{1}{2}$,
für $m > 0$: $x_j = c + \xi_j h$ mit $\xi_j := \frac{j}{m}$ für $j = 0, \ldots, m$

$$I_m(f) = \int_c^d P(f|x_0, ..., x_m)(x)dx = h\sum_{j=0}^{m} c_j f(x_j) = h\sum_{j=0}^{m} c_j f(c + \xi_j h)$$

Exaktheit vom Grade m: $\forall\, Q \in \Pi_m$ gilt $I_m(Q) = \int_c^d Q(x)dx$.

m	Name	ξ_j	c_j	$I_m(f) - \int_c^d f(x)dx$
0	Mittelpunktsregel	$\frac{1}{2}$	1	$-\frac{h^3}{24}\,f^{(2)}(\xi)$
1	Trapezregel	$0, 1$	$\frac{1}{2}, \frac{1}{2}$	$\frac{h^3}{12}\,f^{(2)}(\xi)$
2	Simpson-Regel	$0, \frac{1}{2}, 1$	$\frac{1}{6}, \frac{4}{6}, \frac{1}{6}$	$\frac{h^5}{2880}\,f^{(4)}(\xi)$

Gauß-Quadratur (auf $[c,d]$ mit $h = d - c$)

- $G_m(f) = h\sum_{i=0}^{m} \omega_i f(x_i)$, Exaktheitsgrad $2m + 1$
- positive Gewichte $\omega_i = \frac{1}{h}\int_c^d l_{im}(x)\,dx$
- Es ex. $\xi \in [c,d]$, so dass $\left|G_m(f) - \displaystyle\int_c^d f(x)dx\right| = \dfrac{((m+1)!)^4}{((2m+2)!)^3(2m+3)} h^{2m+3}|f^{(2m+2)}(\xi)|$