

Numerische Mathematik 2

Skript zur Vorlesung in JLU Gießen (2013)
 Mitgeschrieben und gel^AT_EXt von Julian Bergmann

Inhaltsverzeichnis

1	Gewöhnliche Differentialgleichungen	1
1.1	Einführung	1
1.1.1	Definition	1
1.1.2	DGL n-ter Ordnung	1
1.1.3	Satz von Peano	1
1.1.4	Definition	1
1.1.5	Satz von Picard Lindelöf	2
1.1.6	Satz	2
1.2	Einschrittverfahren	2
1.2.1	Eulerisches Polyzugverfahren	2
1.2.2	Definitionen	3
1.2.3	Taylor-Verfahren	5
1.2.4	Trapezregel	5
1.2.5	Die θ -Methode	6
1.2.6	Explizite Runge-Kutta-Verfahren	6
1.2.7	Das klassische Runge-Kutta-Verfahren	9
1.2.8	Spezielles Runge-Kutta-Verfahren	10
1.2.9	Implizite Runge-Kutta-Verfahren	11
1.2.10	Kollokationsverfahren	11
1.3	Mehrschrittverfahren	15
1.3.1	Beispiel: Adams-Verfahren	15
1.3.2	Allgemeine lineare Mehrschrittverfahren	19
1.3.3	Exkurs: Lineare Differenzgleichungen	21
1.3.4	Konvergenz	23
1.3.5	Schrittweitensteuerung	26
1.3.6	Prädiktor-Korrektor-Verfahren	27
2	Lösung großer, dünner LGS	28
2.1	motivierendes Beispiel	28
2.2	Gaußelimination/Cholesky-Zerlegung	29
2.2.1	Matrizen und Bäume	29
2.2.2	Satz von Parter	30
2.3	Iterative Verfahren für große LGS	31
2.3.1	Standard Matrixiterationen	31
2.3.2	Stein-Rosenberg	34
2.3.3	kurze Zusammenfassung	38
2.3.4	Nachiteration	39
2.3.5	Verfahren der konjugierten Gradienten	40
3	Graphik	43
3.1	Bernsteinpolynom	43
3.2	Bézier-Kurven	50
3.3	Auswertung eines polynoms in dwer Bézier-Darstellung	52
3.4	Bézier-Kurven Gradanhebung	54
3.5	Bézier-Splines	55
3.6	Spline-Funktionen	56

Beweise	60
Stichwortverzeichnis	78

1 Gewöhnliche Differentialgleichungen

1.1 Einführung

1.1.1 Definition

Ab nun gegeben sei ein Streifen $S = \{(x, y) : a \leq x \leq b, y \in \mathbb{R}^n\}$ mit $a, b \in \mathbb{R}$ und $f : S \rightarrow \mathbb{R}^n$ stetig.

Das *Anfangsproblem* AWP mit der DGL

$$y'(x) = f(x, y) \quad , \quad x \in [a, b] \quad (1.1.1)$$

ist die Aufgabe, eine differenzierbare Funktion $y : [a, b] \rightarrow \mathbb{R}^n$ zu finden, die in S die DGL 1.1.1 löst und für ein gegebenes $x_0 \in [a, b]$ und $y_0 \in \mathbb{R}^n$ einer Anfangsbedingung $y(x_0) = y_0$ genügt.

Beispiel 1

$$\begin{aligned} 1) \quad & y' = f(x) \quad , \quad y(x_0) = y_0 \\ & \Rightarrow y(x) = y_0 + \int_{x_0}^x f(t) dt \end{aligned}$$

$$\begin{aligned} 2) \quad & y' = \lambda y \quad , \quad y(x_0) = y_0 \quad \text{unabhängig von } x: \text{AWP ist autonom} \\ & \Rightarrow y(x) = y_0 e^{\lambda(x-x_0)} \end{aligned}$$

1.1.2 DGL n-ter Ordnung

$$\begin{aligned} y^{(n)} &= f(x, y, y', y'', \dots, y^{(n-1)}) \\ z_1 &= y \quad , \quad z_2 = y' \quad , \quad \dots \quad , \quad z_n = y^{(n-1)} \\ z &= (z_1, \dots, z_n)^T \quad , \quad z' = (z_2, \dots, z_n, f(x, z_1, \dots, z_n))^T \end{aligned}$$

1.1.3 Satz von Peano

Existenz und Eindeutigkeit der Lösung

Sei f stetig in $[a, b] \times U$, U Umgebung $\subset \mathbb{R}^n$
 Sei $c > 0$ und $S_c = \{(x, y), x \in [a, b], \|y - y_0\| \leq c\}$

Dann ex. eine Lösung des AWP für $x \in [x_0, \min\{b, \frac{b}{\mu}\}]$, $\mu = \max_{(x,y) \in S_c} \{\|f(x, y)\|\}$.

\Rightarrow Für eine stetige Funktion existiert die Lösung zumindest lokal.

1.1.4 Definition

f heißt *Lipschitz-stetig* (Lst), wenn

$$\forall x \in [a, b], y_1, y_2 \in \mathbb{R}^n \quad \exists L : \|f(x, y_1) - f(x, y_2)\| \leq L \|y_1 - y_2\|$$

1.1.5 Satz von Picard Lindelöf

Sei f Lst. auf S . Dann ex. zu jedem $x_0 \in [a, b]$ und jeden $y_0 \in \mathbb{R}^n$ genau eine Fu. mit

- 1) y st. und st. differenzierbar (dbar)
- 2) $y' = f(x, y)$, $x \in [a, b]$
- 3) $y(x_0) = y_0$

1.1.6 Satz

Sei f Lst. auf S , $x_0 \in [a, b]$ fest.

$y'(x, s)$ bezeichne Lösung des AWP's

$$y' = f(x, y) \quad , \quad y(x_0) = s.$$

Dann gilt

$$\|y(x, s_1) - y(x, s_2)\| \leq e^{L(x-x_0)} \|s_1 - s_2\|$$

für $x \in [x_0, b]$

Die Methoden können für den Fall $y \in \mathbb{R}^n$ ohne große Änderungen übertragen werden.

1.2 Einschrittverfahren

1.2.1 Eulerisches Polygonzugverfahren

$$y' = f(x, y), \quad y' = \lim_{h \rightarrow 0} \frac{y(x+h) - y(x)}{h}$$

$\Rightarrow y' \approx \frac{y(x+h) - y(x)}{h}$ mit kleinem h .

$$x_0, x_1 = x_0 + h, x_2 = x_1 + h, \dots, x_k = x_0 + kh, \quad k = 1, 2, 3, \dots$$

$$y(x_k) \approx y_k = y_{k-1} + hf(x_{k-1}, y_{k-1}) \Leftrightarrow \frac{y_k - y_{k-1}}{h} \approx f(x_{k-1}, y_{k-1})$$

Bemerkung:

Man kann auch mit variablen Schrittwerten arbeiten: $x_{k+1} = x_k + h_k$, $h_k > 0$

13-04-19

Beispiel 2

$$y' = y, \quad y(0) = 1, \quad y(x) = e^x, \quad h = 0.1$$

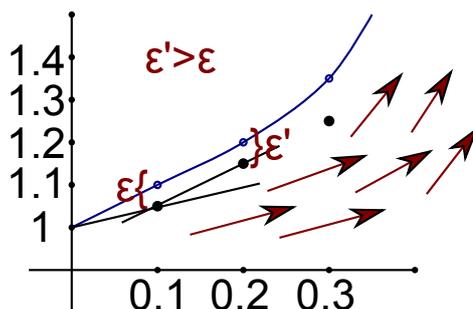
$$x_0 = 0, \quad x_1 = 0.1, \quad x_2 = 0.2, \quad x_3 = 0.3$$

$$y_0 = 1, \quad y_1 = y_0 + hf(x_0, y_0) = 1.1, \quad y_2 = 1.21, \quad y_3 = 1.331$$

$$y(0.1) \approx 1.1, \quad e^{0.1} = 1.1052$$

$$y(0.2) \approx 1.21, \quad e^{0.2} = 1.2214$$

$$y(0.3) \approx 1.331, \quad e^{0.3} = 1.3499$$



1.2.2 Definitionen

Definition:

Ein *Einschrittverfahren* zur Bestimmung der Näherungslösung $y_k \approx y(x_k)$ des AWP's $y' = f(x, y)$, $x \in [a, b]$, $y(x_0) = y_0$ hat die Form:

$$\begin{aligned}x_{k+1} &= x_k + h \\ y_{k+1} &= y_k + h\varphi(x_k, y_k, h), \quad k = 0, 1, \dots\end{aligned}$$

φ heißt *Zuwachsfunktion* oder *Verfahrensfunktion*.

Beispiel: Euler $\varphi(x, y, h) = f(x, y)$

Definition:

Der *lokale Diskretisierungsfehler* an der Stelle $(x, y(x))$ ist definiert durch

$$\tau(x, y(x), h) = \frac{y(x+h) - y(x)}{h} - \varphi(x, y(x), h)$$

Hier ist y die richtige Lösung.

$$y(x+h) = y(x) + h\varphi(x, y(x), h) - h\tau(x, y(x), h)$$

Bemerkung: Fehler, wenn man im Punkt $(x, y(x))$ die richtige Lösung hat und einen Schritt des Verfahrens macht.

Definition:

Ein *Einschrittverfahren* heißt *konsistent*, wenn

$$\lim_{h \rightarrow 0} \tau(x, y(x), h) = 0$$

für $x \in [a, b]$.

Das Verfahren besitzt die Konsistenzordnung $p > 0$, wenn $\tau(x, y(x), h) = O(p)$.

Beispiel: Das eulerische Verfahren ist konsistent und die Konsistenzordnung 1. Siehe Beweis 1 (Anhang)

Definition:

Die *Näherungslösung*: $y_{k,h}$

Die *exakte Lösung*: $y(x_k)$

Globaler Diskretisierungsfehler an der Stelle $x_k = e_{k,h} = y_{k,h} - y(x_k)$

Definition:

Ein *Einschrittverfahren* zur Bestimmung der Lösung des AWP's

$$y' = f(x, y), \quad x \in [a, b], \quad y(x_0) = y_0$$

heißt *konvergent*, wenn

$$\lim_{h \rightarrow 0} \max_{k: x_k \in [a, b]} |e_{k,h}| = 0$$

Das Verfahren hat die *Konvergenzordnung* $p > 0$, wenn

$$\max_{k: x_k \in [a, b]} |e_{k,h}| = O(h^p), \quad h \rightarrow 0$$

Beispiel 3: eulerische Polygonzugverfahren

Das eulerische Polygonzugverfahren ist konvergent mit Konvergenzordnung 1.

Beweis:

$$y_{k+1} = y_{k,h} + hf(x_k, y_{k,h})$$

$$y(x_{k+1}) = y(x_k) + hy'(x_k) + O(h^2)$$

$$e_{k+1,h} = e_{k,h} + h[f(x_k, y_{k,h}) - \underbrace{y'(x_k)}_{f(x_k, y(x_k))}] + O(h^2)$$

$$= e_{k,h} + h[f(x_k, y_{k,h}) - f(x_k, y(x_k))] + O(h^2)$$

$$\text{Da } f \text{ Lst., ist } |f(x_k, y_{k,h}) - f(x_k, y(x_k))| \leq L|y_{k,h} - y(x_k)| = L|e_{k,h}|$$

$$\Rightarrow |e_{k+1,h}| \leq |e_{k,h}| + hL|e_{k,h}| + O(h^2)$$

$$\Rightarrow |e_{k+1,h}| \leq (1 + hL)|e_{k,h}| + O(h^2)$$

Lemma:

Genügen die Zahlen ξ_k , $k = 0, 1, \dots, n$ einer Abschätzung der Form $|\xi_{k+1}| \leq (1 + \delta)|\xi_k| + \beta$, $\delta > 0$, $\beta \geq 0$, $k = 0, 1, \dots, n-1$, so gilt

$$|\xi_n| \leq e^{n\delta}|\xi_0| + \frac{e^{n\delta} - 1}{\delta}\beta$$

Siehe Beweis 2 (Anhang)

$$\Rightarrow |e_{k+1,h}| \leq (1 + hl)|e_{k,h}| + O(h^2)$$

$$\text{Nach dem Lemma folgt: } |e_{n,h}| \leq e^{nhL} \underbrace{|e_{0,h}|}_{=y_{0,h}-y(x_0)=0} + \frac{e^{((nh)\leq(b-a)L)-1}-1}{hL}O(h^2)$$

$$|e_{n,h}| \leq \frac{e^{L(b-a)}-1}{L}O(h) \text{ für alle } x_n \in [a, b]$$

$$\Rightarrow \max_{n: x_n \in [a, b]} |e_{n,h}| \leq \frac{e^{L(b-a)}-1}{L}O(h) \xrightarrow{h \rightarrow 0} 0$$

Die Konvergenzordnung ist somit 1.

Satz:

Ein *Einschrittverfahren* der *Konsistenzordnung* $p > 0$ dessen Verfahrensfunktion φ Lst. (bzgl. y) ist, ist konvergent mit der Konvergenzordnung p .

Siehe Beweis 3 (Anhang)

Bemerkung:

$$|e_{nh}| \leq \frac{e^{L(b-a)} - 1}{L} h$$

⇒ **Spezialfall Euler-Verfahren** (oft zu grob!)

Beispiel 4

$$y' = 100y, \quad y(0) = 1, \quad x \in [0, 1]$$

ex. Lsg.: $y(x) = e^{-100x}, \quad y(1) = e^{-100}$

$$\text{eul. Polygonzug: } y_{k+1} = y_k + hf(x_k, y_k) = y_k + h(-100y_k) = (1 - 100h)y_k$$

$$y(1) \approx y_n, \quad n = \frac{1}{h}, \quad h = 0.02, \quad n = 50$$

$$y_n = (1 - 2)^{50} = 1$$

$$|e_{n,h}| \leq \frac{e^{L(b-a)} - 1}{L} h$$

$$f(x, y) = -100y, \quad |f(x, y_1) - f(x, y_2)| < L|y_1 - y_2|, \quad L = \max_{x \in [a,b]} \left| \frac{\partial f}{\partial y}(x, y) \right| = 100$$

$$h = 0.02, \quad n = 50 \Rightarrow |e_{50,0.02}| \leq \dots \approx 10^{39}$$

1.2.3 Taylor-Verfahren

$$y' = f(x, y), \quad y(x_0) = y_0, \quad x \in [a, b]$$

$$f \in C^p(S) \Rightarrow y \in C^{p+1}([a, b]), \quad \text{Taylorlsg. existiert!}$$

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2} y''(x) + \dots + \frac{h^p}{p!} y^{(p)}(x) + R_h$$

$$R_h(x) = \frac{1}{p!} \int_0^h (h-\xi)^p y^{(p+1)}(\xi) d\xi$$

Das Verfahren:

$$x_{k+1} = x_k + h$$

$$y_{k+1} = y_k + \varphi_p(x_k, y_k, h) \text{ mit}$$

$$\varphi_p(x, y, h) = f(x, y) + \frac{h}{2!} Dd(x, y) + \dots + \frac{h^{p-1}}{p!} D^{p-1} f(x, y) \text{ mit } D: \text{ Differentiation nach } x.$$

$$Df(x, y) = f_x + f_y y' = f_x + f_y f = F$$

$$D^2 f(x, y) = D(f_x + f_y f)$$

$$= f_{xx} + f_{xy} y' + (f_{yx} + f_{yy} y') f + f_y (f_x + f_y y') = \underbrace{f_{xx} + 2f_{xy} f + f_{yy} f^2}_G + f_y F$$

$$\varphi_p(x, y, h) = f + \frac{h}{2} F + \frac{h^2}{6} (G + f_y F) \dots$$

Theoretisch: beliebige **Konsistenzordnung**, praktisch unbrauchbar

Satz:

Taylor-Verfahren hat **Konsistenzordnung** p

1.2.4 Trapezregel

Verfahren der 1. **Konsistenzordnung**.

$$y_{k+1} = x_k + h, \quad y_{k+1} = y_k + h \left(\frac{f(x_k, y_k) + f(\dots)}{2} \right), \quad k = 0, 1, 2, \dots$$

$$y_1 = y_0 + \frac{h}{2} (f(x_0, y_0) + f(x_1, y_1))$$

Implizite Methode ⇒ Jeder Schritt nur eine Gleichung (Vergleich: eul. Polzug: **explizit**).

Beispiel 5

$$y' = y, \quad y(0) = 1, \quad y(x) = e^x, \quad h = 0.1$$

$$\text{exakte Lösung: } y(0,1) = 1.10517, \quad y(0,2) = 1.22140, \quad y(0,3) = 1.34986$$

$$\text{mit Euler: } y_1 = 1.1, \quad y_2 = 1.21, \quad y_3 = 1$$

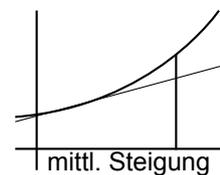
$$\text{Trapez: } y_0 = 1, \quad y_{k+1} = y_k + \frac{h}{2}(f(x_k, y_k) + f(x_{k+1}, y_{k+1})) = y_k + \frac{h}{2}y_k + \frac{h}{2}y_{k+1}$$

$$(1 - \frac{h}{2})y_{k+1} = (1 + \frac{h}{2})y_k$$

$$y_{k+1} = \frac{1 + \frac{h}{2}}{1 - \frac{h}{2}} y_k = \left(\frac{1 + \frac{h}{2}}{1 - \frac{h}{2}} \right)^{k+1}$$

$$h = 0.1 : \frac{1 + \frac{h}{2}}{1 - \frac{h}{2}} = \frac{21}{19}, \quad y_k = \left(\frac{21}{19} \right)^k$$

$$y_1 = 1.10526, \quad y_2 = 1.22161, \quad y_3 = 1.35020$$

**Satz:**

Trapezregel hat Konsistenzordnung 2

1.2.5 Die θ -Methode

$$0 < \theta < 1, \quad x_{k+1} = x_k + h$$

$$y_{k+1} = y_k + h(\theta f(x_k, y_k) + (1 - \theta)f(x_{k+1}, y_{k+1}))$$

$$\theta = 1 \Rightarrow \text{euler. Polzug}$$

$$\theta = \frac{1}{2} \Rightarrow \text{Trapez}$$

$$\theta = 0 \Rightarrow y_{k+1} = y_k + hf(x_{k+1}, y_{k+1})$$

 \Rightarrow impl. eul. Polzugimpl.: $f(x_{k+1}, y_{k+1})$ verschwindet nichtexpl.: $\theta = 1$

Konsistenzordnung:

$$\begin{aligned} \tau(x, y(x), h) &= \frac{y(x+h) - y(x)}{h} - \theta f(x, y) - (1 - \theta)f(x + h, y(x + h)) \\ &= \frac{1}{h} \left(hy'(x) + \frac{h^2}{2}y''(x) + \frac{h^3}{6}y'''(x) + O(h^4) \right) - \theta y' \\ &= (1 - \theta)(y' + hy'' + \frac{h^2}{2}y''' + O(h^4)) \\ &= h(\frac{1}{2} - (1 - \theta))y'' + h^2(\frac{1}{6} - \frac{1}{2}(1 - \theta))y''' + O(h^3) \\ &= h(\theta - \frac{1}{2})y'' + h^2(\frac{1}{2}\theta - \frac{1}{3})y''' + O(h^3) \end{aligned}$$

 \Rightarrow Ordnung 1 oder 2 (bei $\theta = \frac{1}{2}$)**1.2.6 Explizite Runge-Kutta-Verfahren**

13-04-26

Sei $s \in \mathbb{N}$. Ein explizites *s-stufiges Runge-Kutta-Verfahren* hat die Form

$$x_{k+1} = x_k + h, \quad y_{k+1} = y_k + h\varphi(x_k, y_k, h)_{k=0,1,\dots}$$

$$k_1(x, y) = f(x, y)$$

$$k_2(x, y) = f(x + c_2h, y_k + ha_{21}k_1(x, y))$$

$$k_3(x, y) = f(x + c_3h, y_k + h(a_{31}k_1(x, y) + a_{32}k_2(x, y)))$$

 \vdots

$$k_s(x, y) = f(x + c_sh, y_k + h \sum_{i=1}^{s-1} a_{si}k_i(x, y))$$

$$\begin{aligned}\varphi(x, y, h) &= \sum_{j=1}^s b_j k_j(x, y) \\ y_{k+1} &= y_k + h \sum_{j=1}^s b_j k_j\end{aligned}$$

Butcher-Tableau:

Die Koeffizienten

$$\begin{array}{c|ccc} c_1 = 0 & & & 0 \\ c_2 & a_{21} & & \\ c_3 & a_{31} & a_{32} & \\ \vdots & \vdots & & \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss-1} \\ \hline & b_1 & b_2 & \dots & b_{s-1} & b_s \end{array}$$

Butcher-Tableau: $\begin{array}{c|c} \vec{c} & A \\ \hline & \vec{b}^T \end{array}$

1) Wir wollen, dass die Methode mindestens Ordnung 1 hat:

$$\begin{aligned}\tau(x, y(x), h) &= \frac{y^{(k+h)} - y(x)}{h} - \varphi(x, y, h) \\ \frac{y^{(k+h)} - y(x)}{h} &= y'(x) + O(h) = f(x, y) + O(h) \\ \varphi(x, y, h) &= \sum_{j=1}^s b_j k_j(x, y) \\ &= \sum_{j=1}^s b_j f(x + c_j h, y + h \sum_{i=1}^{j-1} a_{ji} k_i(x, y)) \\ &= \sum_{j=1}^s b_j (f(x, y) + O(h)) = \sum_{j=1}^s b_j f(x, y) + O(h) \\ \tau(x, y(x), h) &= f(x, y) - \sum_{j=1}^s b_j f(x, y) + O(h) \\ &= \left(1 - \sum_{j=1}^s b_j\right) f(x, y) + O(h)\end{aligned}$$

$$\text{Ordnung 1} \Leftrightarrow \tau(x, y(x), h) = O(h)$$

$$\Leftrightarrow \sum_{j=1}^s b_j = 1$$

2) Wir wollen, dass $k_i(x, y(x))$ die Ableitungen $y'(x + c_i h)$ mindestens von Ordnung 2 approximieren, d.h.

$$k_i(x, y) = y'(x + c_i h) + O(h^2)$$

$$\begin{aligned}y'(x, c_i h) &= y'(x) + y''(x) c_i h + O(h^2) \\ y' &= f(x, y) \\ y'' &= f_x + f_y y' = f_x + f_y f \\ &= f(x, y) + h c_i (f_x(x, y) + f_y(x, y) f(x, y)) + O(h^2)\end{aligned}$$

$$\begin{aligned}k_i(x, y) &= f\left(x + c_i h, y + h \sum_{j=1}^{i-1} a_{ji} k_j\right) \\ &= f(x, y) + h \left(f_x c_i + f_y \sum_{j=1}^{i-1} a_{ji} k_j(x, y)\right)\end{aligned}$$

$$\begin{aligned}k_i &= y'(c_i h) + O(h^2) \Leftrightarrow \\ &= f(x, y) + h c_i (f_x(x, y) + f_y(x, y) f(x, y)) \\ &= f(x, y) + h \left(f_x(x, y) c_i + f_y(x, y) f(x, y) \sum_{j=1}^{i-1} a_{ij}\right)\end{aligned}$$

$$c_i f_x(x, y) + c_i f_y(x, y) f(x, y) = c_i f_x(x, y) + f_y(x, y) f(x, y) \sum_{j=1}^{i-1} a_{ij}$$

$$\Rightarrow c_i = \sum_{j=1}^{i-1} a_{ij}$$

Beispiel 6: für s=2

$$\begin{array}{c|cc} c_2 & a_{21} & \\ \hline & b_1 & b_2 \end{array}$$

$c_2 = a_{21} \quad b_1 + b_2 = 1$
 $\Rightarrow 2$ freie Parameter.

$$\begin{aligned} x_{k+1} &= x_k + h \\ k_1(x, y) &= f(x, y) \\ k_2(x, y) &= f(x + c_2 h, y + h a_{21} k_1(x, y)) \\ &= f(x + c_2 h, y + h a_{21} f(x, y)) \\ y_{k+1} &= y_k + h(b_1 f(x, y) + b_2 f(x_k + c_2 h, y_k + h a_{21} f(x_k, y_k))) \end{aligned}$$

Wir wollen ein Verfahren der Ordnung 2 konstruieren.

$$\begin{aligned} \frac{y(x+h)-y(x)}{h} &= y'(x) + \frac{h}{2} y''(x) + O(h^2) \\ &= f(x, y) + \frac{h}{2} (f_x(x, y) + f_y(x, y) f(x, y)) + O(h^2) \\ \varphi(x, y, h) &= b_1 f(x, y) + b_2 f(x + c_2 h, y + h a_{21} f(x, y)) \\ &= b_1 f(x, y) + b_2 (f(x, y) + f_x(x, y) c_2 h + f_y(x, y) h a_{21} f(x, y) + O(h^2)) \\ &= (b_1 + b_2) f(x, y) + b_2 c_2 f_x(x, y) h + b_2 a_{21} f_y(x, y) f(x, y) h + O(h^2) \\ &= f(x, y) + b_2 c_2 h f_x(x, y) + b_2 a_{21} h f_y(x, y) f(x, y) + O(h^2) \end{aligned}$$

Für Ordnung 2 muss roter Teil gleich sein!

Ordnung 2 $\Leftrightarrow \tau(x, y, h) = O(h^2)$

$$\Leftrightarrow \frac{y(x,y)-y(x)}{h} = \varphi(x, y, h) + O(h^2)$$

$$\begin{aligned} \Leftrightarrow f(x, y) + \frac{h}{2} f_x(x, y) + \frac{h}{2} f_y(x, y) f(x, y) \\ = f(x, y) + b_2 c_2 h f_x(x, y) + b_2 a_{21} h f_y(x, y) f(x, y) \end{aligned}$$

$$\Leftrightarrow b_2 c_2 = \frac{1}{2}$$

$$\frac{1}{2} = b_2 a_{21} \quad (\text{äquivalent zu oben, da } a_{21} = c_2)$$

$$b_1 + b_2 = 1, \quad c_2 = a_{21}, \quad b_2 c_2 = \frac{1}{2}$$

Zur Bestimmung der Parameter muss man ein nicht lineares Gliederungssystem lösen.

Eine einparametrische Lösungsmenge $\begin{array}{c|cc} 0 & & \\ \hline c_2 & c_2 & \\ \hline & 1 - \frac{1}{2c_2} & \frac{1}{2c_2} \end{array}$

Für jedes $c_2 \in (0, 1]$ ist das ein Runge-Kutta-Verfahren der Ordnung 2.

Beispiel 7: Runge-Kutta: Euler-Cauchy-Verfahren

$$\begin{array}{c|cc} c_2 = 1 & & \\ 0 & & \\ \hline 1 & 1 & \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \quad \Rightarrow \varphi(x, y) = \frac{f}{2} f(x, y) + \frac{1}{2} f(x + h, y + h f(x, y))$$

Beispiel 8: Runge-Kutta: verbessertes eul. Polygonzugverfahren

$$\begin{array}{c|cc}
 c_2 & & \frac{1}{2} \\
 0 & & \\
 \frac{1}{2} & \frac{1}{2} & \\
 \hline
 & 0 & 1
 \end{array}
 \Rightarrow \varphi(x + \frac{1}{2}h, y + \frac{1}{2}hf(x, y))$$

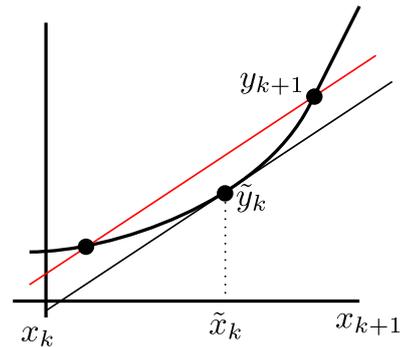
Andere Darstellung:

$$\left. \begin{array}{l}
 \tilde{x}_k = x_k + \frac{h}{2} \\
 \tilde{y}_k = y_k + \frac{h}{2}f(x_k, y_k) \\
 x_{k+1} = x_k + h \\
 y_{k+1} = y_k + hf(\tilde{x}_k, \tilde{y}_k)
 \end{array} \right\} \begin{array}{l}
 \text{Eulerische Polygonzugverfahren mit der} \\
 \text{Schrittweite } \frac{h}{2} \\
 \\
 \text{Wir rechnen die Ableitung in Punkte } (\tilde{x}_k, \tilde{y}_k) \\
 \text{als Approximation der Ableitung auf das} \\
 \text{Intervall}
 \end{array}$$

s=4 (vierstufige Verfahren)

$$\begin{array}{c|cccc}
 0 & & & & \\
 c_2 & a_{21} & & & \\
 c_3 & a_{31} & a_{32} & & \\
 c_4 & a_{41} & a_{42} & a_{43} & \\
 \hline
 & b_1 & b_2 & b_3 & b_4
 \end{array}$$

$$\begin{aligned}
 c_2 &= a_{21} \\
 c_3 &= a_{31} + a_{32} \\
 c_4 &= a_{41} + a_{42} + a_{43} \\
 b_1 + b_2 + b_3 + b_4 &= 1
 \end{aligned}$$



9 freie Parameter

Ein Verfahren der Ordnung 4 ist möglich!

7 Gleichungen durch Vergleich der Terme der Taylorentwicklung

⇒ Ein nicht-lineares Gleichungssystem mit einer zweiparametrischen Lösungsmenge.

1.2.7 Das klassische Runge-Kutta-Verfahren

$$\begin{aligned}
 x_{k+1} &= x_k + h \\
 k_1 &= f(x_k, y_k) \\
 k_2 &= f(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_1) \\
 k_3 &= f(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_2) \\
 k_4 &= f(x_k + h, y_k + hk_2) \\
 y_{k+1} &= y_k + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)
 \end{aligned}$$

$$\begin{array}{c|ccc}
 0 & & & \\
 \frac{1}{2} & \frac{1}{2} & & \\
 \frac{1}{2} & 0 & \frac{1}{2} & \\
 1 & 0 & 0 & 1 \\
 \hline
 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6}
 \end{array}$$

Beispiel 9

$$\begin{aligned}
y' &= y, \quad y(0) = 1, \quad h = 0.1 \\
x_0 &= 0, \quad y_0 = 1 \\
k_1 &= f(0, 1) = 1 \\
k_2 &= f(0.05, 1 + 0.05k_1) = 1 + 0.05 = 1.05 \\
k_3 &= f(0.05, 1 + 0.05k_2) = 1.0525 \\
k_4 &= f(0.1, 1 + 0.1k_3) = 1.10525 \\
y_1 &= y_0 + \frac{0.1}{6}(k_1 + k_2 + k_3 + k_4) = 1.105170833 \\
e^{0.1} &= 1.105170918
\end{aligned}$$

1.2.8 Spezielles Runge–Kutta–Verfahren

13-04-30

Vergleich mit Quadraturformel:

Sonderfall: $y' = f(x)$, $y(x_0) = y_0$

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(t) dt$$

$$\int_{x_k}^{x_{k+1}} f(t) dt \approx h \sum_{j=1}^s b_j f(x_k + c_j h, y(x_k + c_j h))$$

$y' = f(x, y)$, $y(x_0) = y_0$

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(t, y(t)) dt$$

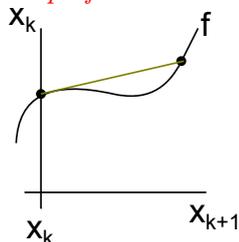
$$\int_{x_k}^{x_{k+1}} f(t, y(t)) dt \approx h \sum_{j=1}^s b_j f(x_k + c_j h, y(x_k + c_j h))$$

$$*: y(x_k) + \int_{x_k}^{x_k + c_j h} f(t, y(t)) dt \approx y(x_k) + h \sum_{i=1}^{j-1} a_{ji} \underbrace{f(x_k + c_i h, y(x_k + c_i h))}_{\approx k_i}$$

$$f(x_k + c_j h, y(x_k + c_j h)) \approx k_j = y(x_k) + h \sum_{i=1}^{j-1} a_{ji} k_i$$

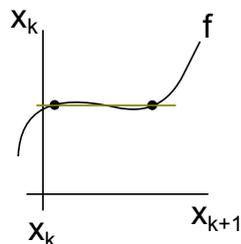
Sonderfall $y' = f(x)$, $y(x_0) = y_0$

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(t) dt$$

1) *Trapezformel:*

$$\int_{x_k}^{x_{k+1}} f(t) dt \approx \frac{h}{2} (f(x_k) + f(x_{k+1})) y_{k+1} = y_k + \frac{h}{2} (f(x_k) + f(x_{k+1}))$$

2) *Euler Cauchy Verfahren*



$$\int_{x_k}^{x_{k+1}} f(t) dt = hf\left(\frac{x_k+x_{k+1}}{2}\right) = hf\left(x_k + \frac{h}{2}\right)$$

$$y_{k+1} = y_k + hf\left(x_k + \frac{h}{2}\right)$$

Das *verbesserte Eulersche Polygon Verfahren*

3) *Simson-Regel:*

$$\int_{x_k}^{x_{k+1}} f(t) dt = \frac{h}{6} \left(f(x_k) + 4f\left(\frac{x_k+x_{k+1}}{2}\right) + f(x_{k+1}) \right)$$

$$y_{k+1} = y_k + \frac{h}{6} \left(f(x_k) + 4f\left(x_k + \frac{h}{2}\right) + f(x_{k+1}) \right)$$

Das *klassische Runge-Kutta-Verfahren*

$$k_1 = f(x_k, y_k)$$

$$k_2 = f\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_1\right)$$

$$k_3 = f\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_2\right)$$

$$k_4 = f\left(x_k + h, y_k + hk_3\right)$$

$$y_{k+1} = y_k + h \left(\frac{1}{6}k_1 + \frac{4}{6}k_2 + \frac{2}{6}k_3 + \frac{1}{6}k_4 \right)$$

1.2.9 Implizite Runge-Kutta-Verfahren

Butcher-Tableau:

c_1	a_{11}	a_{12}	\dots	$a_{1,s-1}$	a_{1s}
c_2	a_{21}	a_{22}	\dots	$a_{2,s-1}$	a_{2s}
\vdots					
c_s	a_{s1}	a_{s2}	\dots	$a_{s,s-1}$	a_{ss}
	b_1	b_2	\dots	b_{s-1}	b_s

$$x_{k+1} = x_k + h$$

$$k_j = f(x_k + c_j h, y_k + h \sum_{i=1}^s a_{ji} k_i)_{j=1, \dots, s} \quad y_{k+1} = y_k + h \sum_{j=1}^s b_j k_j$$

Ein System von s nicht-linearen Gleichungen mit s Unbekannten $k_j, j = 1, \dots, s$

Vorteil: ein Verfahren der Ordnung $2s$ ist möglich!

Genau wie bei den expliziten Verfahren haben wir $\sum_{j=1}^s b_j = 1$

und $\sum_{i=1}^s a_{ji} = c_j, \quad j = 1, \dots, s$

1.2.10 Kollokationsverfahren

Seien c_1, \dots, c_s gegeben und paarweise verschieden.

Finde ein Polynom $u \in \mathbb{P}_s$ mit $u(x_k) = y_k$ und

$$u'(x_k + c_j h) = f(x_k + c_j h, u(x_k + c_j h)), \quad j = 1, \dots, s$$

Setze $y_{k+1} = u(x_{k+1})$

Sei $\omega(x) = \prod_{j=1}^s (x - c_j)$ das Knotenpolynom

$$l_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^s \frac{x-c_j}{c_i-c_j}, \quad i = 1, \dots, s$$

die Lagrange-Grundpolynome.

$$l_i(c_j) = \delta_{ij}$$

Wir definieren:

$$l_j = \int_0^1 l_j(t) dt, \quad j = 1, \dots, s$$

$$a_{ji} = \int_0^{c_j} l_i(t) dt, \quad i, j = 1, \dots, s$$

Bemerkung:

Das Kollokationsverfahren ist identisch mit dem Runge-Kutta-Verfahren

mit dem Bucher-Tableau $\begin{array}{c|c} \vec{c} & \vec{A} \\ \hline & \vec{b}^T \end{array}$

Siehe Beweis 4 (Anhang)

Beispiel 10

$$s = 2, \quad c_1 = 0, \quad c_2 = \frac{2}{3}$$

$$\omega(x) = x(x - \frac{2}{3})$$

$$l_1(x) = \frac{x - \frac{2}{3}}{0 - \frac{2}{3}} = -\frac{3}{2}(x - \frac{2}{3}) = -\frac{3}{2}x + 1$$

$$l_2(x) = \frac{x}{\frac{2}{3}} = \frac{3}{2}x$$

$$b_1 = \int_0^1 l_1(t) dt = \int_0^1 (-\frac{3}{2}t + 1) dt = -\frac{3}{2} \cdot \frac{1}{2} + 1 = \frac{1}{4}$$

$$b_2 = \int_0^1 l_2(t) dt = \int_0^1 \frac{3}{2}t dt = \frac{3}{2} \cdot \frac{1}{2} = \frac{3}{4}$$

$$a_{11} = \int_0^{c_1} l_1(t) dt = \int_0^0 l_1(t) dt = 0$$

$$a_{12} = \int_0^{c_1} l_2(t) dt = 0$$

$$a_{21} = \int_0^{c_2} l_1(t) dt = \int_0^{\frac{2}{3}} (-\frac{3}{2}t + 1) dt = -\frac{3}{2} \cdot \frac{1}{2} \left(\frac{2}{3}\right)^2 + \frac{2}{3} = -\frac{1}{3} + \frac{2}{3} = \frac{1}{3}$$

$$a_{22} = \int_0^{c_2} l_2(t) dt = \int_0^{\frac{2}{3}} \frac{3}{2}t dt = \frac{3}{2} \cdot \frac{1}{2} \left(\frac{2}{3}\right)^2 = \frac{1}{3}$$

Aus dem Kollokationsverfahren wird $c_1 = 0$, $c_2 = \frac{2}{3}$ bekommt man das Runge-Kutta-Verfahren.

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \hline & \frac{1}{4} & \frac{3}{4} \end{array}$$

$$k_i = f(x_k, c_i h, y_k + h(a_{11}k_1 + a_{12}k_2)) = f(x_k, y_k)$$

$$k_2 = f(x_k + c_2 h, y_k + h(a_{21}k_1 + a_{22}k_2)) = f(x_k + \frac{2}{3}h, y_k + \frac{h}{3}(k_1 + k_2))$$

$$y_{k+1} = y_k + h(\frac{1}{4}k_1 + \frac{3}{4}k_2)$$

Bemerkung:

Es gibt implizite Runge–Kutta–Verfahren, die sich nicht durch ein Kollokationsverfahren ableiten lassen.

Beispiel 11

$$\begin{array}{c|cc} 0 & \frac{1}{4} & -\frac{1}{4} \\ \frac{2}{3} & \frac{1}{4} & \frac{5}{4} \\ \frac{3}{3} & \frac{1}{4} & \frac{12}{4} \\ \hline & \frac{1}{4} & \frac{3}{4} \end{array}$$

Die Koeffizienten sind eindeutig durch die c bestimmt. Hier ergeben die zusätzlichen 1. Zeilen–Werte zusätzliche mögliche Lösungen. Daher ist hier keine eindeutige Lösung mit dem Kollokationsverfahren ableitbar.

Satz:

Seien c_1, c_2, \dots, c_s gegeben und paarweise verschieden.

Ist das Knotenpolynom

$$\omega(x) = \prod_{j=1}^s (x - c_j)$$

orthogonal zu allen Polynomen vom Grad $\leq m-1$ (mit $m \leq s$), so hat das durch die Kollokation definierte Runge–Kutta–Verfahren die Ordnung $s+m$.

Orthogonalität: $\int_0^1 \omega(x)x^j dx = 0, \quad j = 0, 1, \dots, m-1$

Korollar:

Sind c_1, c_2, \dots, c_s die Nullstellen des Legendre–Polynoms vom Grad s , so hat das durch die Kollokation definierte Runge–Kutta–Verfahren die Ordnung $2s$.

Siehe Beweis 5 (Anhang), wobei die folgenden Sätze und Bemerkungen verwendet werden.

Beispiel 12: für $s=1$

Konstruiere ein Verfahren der Ordnung 2!

1) Legendre–Polynom vom Grad 1. (Orthogonal zu \mathbb{P}_0)

$$\int_0^1 (ax + b) dx = 0$$

$$\begin{aligned} \text{O.B.d.A. } a=1 &\Rightarrow \int_0^1 (x + b) dx = 0 = \frac{1}{2} + b \Rightarrow b = -\frac{1}{2} \\ &\Rightarrow p_1(x) = x - \frac{1}{2} \end{aligned}$$

2) Seine Nullstellen

$$c_1 = \frac{1}{2}$$

3) Die Kollokation:

$$l_1(x) = 1$$

$$b_1 = \int_0^1 l_1(x) dx = \int_0^1 1 dx = 1$$

$$a_{11} = \int_0^{c_1} l_1(x) dx = \int_0^{\frac{1}{2}} 1 dx = \frac{1}{2}$$

$$\frac{\frac{1}{2} \mid \frac{1}{2}}{1}$$

$$k_1 = f(x_k + \frac{1}{2}h, y_k + \frac{1}{2}hk_1)$$

$$y_{k+1} = y_k + hk_1$$

Diese Methode hat die Ordnung 2!

Alekseev–Gröbner–Lemma (adaptiert):

Sei u eine glatte Funktion mit

$$u(x_k) = y_k$$

Dann

$$u(x) - y(x) = \int_{x_k}^x \Phi(x, t, u(t))(u'(t) - f(t, u(t))) dt$$

wobei $\Phi(x, t, u(t))$ die Ableitung $\frac{dv}{dv_0}$ an der Stelle x der Lösung des AWP's

$$v' = f(x, v), \quad v(t) = v_0$$

mit $v_0 = u(t)$ bezeichnet.

Beispiel 13: zu Alekseev–Gröbner–Lemma

$$y' = y, \quad v' = v, \quad \text{mit } v(t) = v_0$$

Die Lösung $v(\tau) = v_0 e^{\tau-t}$

$$\frac{dv}{dv_0}(\tau) = e^{\tau-t}$$

$$\Phi(x, t, u(t)) = e^{\tau-t} \Big|_{\tau=x} = e^{x-t}$$

$$u(x) - g(x) = \int_{x_k}^x e^{x-t}(u'(t) - f(t, u(t))) dt = \int_{x_k}^x e^{x-t}(u'(t) - u(t)) dt$$

$$\text{z.B. } u(t) = t - \frac{y_k}{x_k} u(x) - g(x) = x \frac{y_k}{x_k} - y_k e^{x_k}$$

$$\begin{aligned} \int_{x_k}^x e^{x-t} \left(\frac{y_k}{x_k} - t \frac{y_k}{x_k} \right) dt &= \frac{y_k}{x_k} \int_{x_k}^x e^{x-t} (1-t) dt \\ &= -\frac{y_k}{x_k} \int_{x_k}^x (1-t) de^{x-t} = -\frac{y_k}{x_k} \left((1-t) e^{x-t} \Big|_{t=x_k}^{t=x} + \int_{x_k}^x e^{x-t} dt \right) \\ &= -\frac{y_k}{x_k} \left((1-x) - (1-x_k) e^{x-x_k} - 1 + e^{x-x_k} \right) \\ &= -\frac{y_k}{x_k} (-x + e^{x-x_k} x_k) = x \frac{y_k}{x_k} - y_k e^{x-x_k} \end{aligned}$$

Satz über Gauß–Quotienten:

Eine Quadraturformel

$$\int_a^b g(t) dt = \sum_{\nu=0}^h \lambda_\nu g(\tau_\nu) + R(g)$$

ist exakt für alle Polynome $p \in \mathbb{P}_{n+m}$ genau dann, wenn sie interpolierend ist und das Knotenpolynom

$$\omega(t) = \prod_{\nu=1}^n (t - \tau\nu)$$

orthogonal zu \mathbb{P}_{m-1} ist.

1.3 Mehrschrittverfahren

Einschrittverfahren:

zur Berechnung von y_{k+1} wird nur die Information im Punkt (x_k, y_k) benutzt.

Mehrschrittverfahren:

zur Berechnung von y_{k+s} wird die Information in den Punkten (x_k, y_k) , $(x_{k+1}, y_{k+1}), \dots, (x_{k+s-1}, y_{k+s-1})$ verwendet.

1.3.1 Beispiel: Adams-Verfahren

$$y' = f(x, y), \quad y(x_0) = y_0$$

$$\text{Ansatz: } y(x_{k+1}) = y(x_{k+q}) + \int_{x_{k+q}}^{x_{k+s}} f(x, y(x)) dx \quad \text{mit } 0 \leq q \leq s-1$$

Bezeichnung: $f_i := f(x_i, y_i)$

Idee: Wir ersetzen $f(x, y(x))$ durch ein Interpolationspolynom $p(x)$ und integrieren $p(x)$ statt $f(x, y(x))$.

Expliziter Fall: $p(x)$ interpoliert die Punkte $\{(x_{k+i}, f_{k+i})\}_{i=0}^{s-1}$

Impliziter Fall: $p(x)$ interpoliert die Punkte $\{(x_{k+i}, f_{k+i})\}_{i=0}^s$

$$\deg p = m = \begin{cases} s-1 & , \text{expl Fall} \\ s & , \text{impl Fall} \end{cases}$$

$$y_{k+s} = y_{k+q} + \int_{x_{k+q}}^{x_{k+s}} p(x) dx$$

Lagrange-Grundpolynome:

$$l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^m \frac{x - x_{k+j}}{x_{k+i} - x_{k+j}}, \quad i = 0, 1, \dots, m$$

$$p(x) = \sum_{i=0}^m f_{k+i} l_{k+i}(x)$$

$$y_{k+s} = y_{k+q} + \sum_{i=0}^m f_{k+i} \int_{x_{k+q}}^{x_{k+s}} l_{k+i}(x) dx$$

$$x_{k+s} = y_{k+q} + h \sum_{i=0}^m b_i f_{k+i}$$

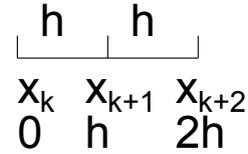
$$\text{mit } b_i = \frac{1}{h} \int_{x_{k+q}}^{x_{k+s}} l_{k+i}(x) dx$$

Beispiel 14: für s=2, q=1

$$y_{k+2} = y_{k+1} + \sum_{i=0}^1 f_{k+1} b_i$$

$$b_0 = \int_{\frac{h}{2}}^{2h} l_0(x) dx, \quad l_0(x) = \frac{x-h}{0-h}$$

$$b_1 = \int_h^{2h} l_1(x) dx, \quad l_1(x) = \frac{x-0}{h-0}$$



$$p(x) = f_{k+m} + [f_{k+m}, f_{k+m-1}](x - x_{k+m}) + \dots + [f_{k+m}, f_{k+m-1}, \dots, f_k](x - x_{k+m}) \dots (x - x_{k+1})$$

$$[f_{r+i}, f_{r+i-1}, \dots, f_r] = \frac{1}{h^i i!} \nabla^i f_{r+i}$$

mit ∇^i die rückwärts gewonnene Differenzen:

$$\nabla^0 f_r = f_r$$

$$\nabla^l f_r = \nabla^{l-1} f_r - \nabla^{l-1} f_{r-1} \quad l = 1, 2, \dots$$

Substitution: $t = \frac{1}{h}(x - k + m)$

$$ht = x - x_k + m, \quad x = x_{k+m} + ht$$

$$x = x_{k+s}$$

$$t = \frac{1}{h}(x_{k+s} - x_{k+m}) = \frac{1}{h}(x_0 + h(k+s) - (x_0 + h(k+m))) = s - m$$

$$p^*(t) = p(th + x_{k+m})$$

$$= f_{k+m} + th \frac{\nabla^1 f_{k+m}}{h} + t(t+1)h^2 \frac{\nabla^2 f_{k+m}}{2h^2} + \dots + t(t+1)\dots(t+m-1)h^m \frac{\nabla^m f_{k+m}}{m!h^m}$$

$$= \sum_{i=0}^m \binom{t+i-1}{i} \nabla^i f_{k+m}$$

$$y_{k+s} = y_{k+q} + \int_{x_{k+q}}^{x_{k+s}} p(x) dx$$

$$= y_{k+q} + h \int_{q-m}^{s-m} p^*(t) dt$$

$$= y_{k+q} + h \sum_{q-m}^{s-m} \sum_{i=0}^m \binom{t+i-1}{i} \nabla^i f_{k+m} dt$$

$$= y_{k+q} + h \sum_{i=0}^m b_i^* \nabla^i f_{k+m}$$

Mit $b_i^* = \int_{s-q}^{s-m} \binom{t+i-1}{i} dt$

Wenn man b_i^* kennt, kann man b_i folgendermaßen berechnen. $\nabla^i f_{k+m} = \sum_{r=0}^i (-1)^r \binom{i}{r} f_{k+m-r}$

$$\nabla f_{k+m} = f_{k+m} - f_{k+m-1} = \sum_{r=0}^1 (-1)^r \binom{1}{r} f_{k+m-r}$$

$$\nabla^2 f_{k+m} = \nabla^1 f_{k+m} - \nabla^1 f_{k+m-1}$$

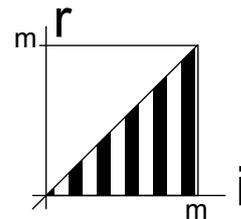
$$= (f_{k+m} - f_{k+m-1}) - (f_{k+m-1} - f_{k+m-2})$$

$$= f_{k+m} - 2f_{k+m-1} + f_{k+m-2}$$

$$= \sum_{r=0}^2 (-1)^r \binom{2}{r} f_{k+m-r}$$

Mit *: $y_{k+s} = y_{k+q} + h \sum_{i=0}^m b_i f_{k+i}$

$$= y_{k+q} + h \sum_{i=0}^m b_i^* \nabla^i f_{k+m}$$



$$\begin{aligned} \sum_{i=0}^m b_i^* \nabla^i f_{k+m} &= \sum_{i=0}^m b_i^* \sum_{r=0}^i (-1)^r \binom{i}{r} f_{k+m-r} \\ &= \sum_{r=0}^m f_{k+m-r} \underbrace{\sum_{i=r}^m (-1)^r \binom{i}{r} b_i^*}_{b_{m-r}} \\ b_{m-r} &= (-1)^r \sum_{i=r}^m \binom{i}{r} b_i^* \\ b_i^* &= \int_{s-q}^{s-m} \binom{t+i-1}{i} dt \\ \binom{t+i-1}{i} &= \frac{(t+i-1)(t+i-2)\dots(t+1)t}{i!} \\ &= (-1)^i \frac{(-t-i+1)(-t-i+2)\dots(-t-1)(-t)}{i!} \\ &= (-1)^i \binom{-t}{i} \\ (1+z)^{-t} &= \sum_{i=0}^{\infty} \binom{-t}{i} (-z)^i, \quad |z| \leq 1 \\ (1-z)^{-t} &= \sum_{i=0}^{\infty} \binom{-t}{i} (-z)^i \\ &= \sum_{i=0}^{\infty} (-1)^i \binom{-t}{i} z^i, \quad |z| < 1 \\ &= \sum_{i=0}^{\infty} \binom{t+i-1}{i} z^i, \quad |z| < 1 \end{aligned}$$

Lemma

Seien $a < b \leq 1$ und $0 < R < 1$

Dann lässt sich die Funktion

$$g(z, a, b) = \int_a^b (1-z)^{-t} dt = -\frac{(1-z)^{-t}}{\ln(1-z)} \Big|_{t=a}^{t=b}$$

in eine für $|z| \leq R$ gleichmäßig konvergierende Potenzreihe entwickeln:

$$g(z, a, b) = \sum_{i=0}^{\infty} b_i^*(a, b) z^i$$

$$\text{mit } b_i^*(a, b) = \int_a^b \binom{t+i-1}{i} dt$$

Siehe Beweis 6 (Anhang)

$$y_{k+s} = y_{k+q} + h \sum_{i=0}^m k_i^* \nabla^i f_{k+m}$$

$$\text{mit } k_i^* = \int_{q-m}^{s-m} \binom{t+i-1}{i} dt$$

k_i^* sind die Koeffizienten der Potenzreihenentwicklung von $g(z, q-m, s-m)$

Klassifikation der Methoden

$q = s-1 \rightarrow m = s-1$ (**explizit**): s-Schritt-Adams-Bashforth-Verfahren, $g(z, 0, 1)$

$q = s-1 \rightarrow m = s$ (**implizit**): s-Schritt-Adams-Moulton-Verfahren, $g(z, -1, 0)$

$q = s-2 \rightarrow m = s-1$ (**explizit**): Nyström-Verfahren, $g(z, -1, 1)$

$q = s - 2 \rightarrow m = s$ (**implizit**): Milne-Simpson-Verfahren, $g(z, -2, 0)$

Beispiel 15: Adams-Moulton-Verfahren

13-05-10

(implizit, $q = s - 1, m = s$)

$$\begin{aligned} g(z, q - m, s - m) &= g(z, -1, 0) \\ &= -\frac{(1-z)^{-t}}{\ln(1-z)} \Big|_{t=-1}^{t=0} \\ &= -\frac{1}{\ln(1-z)} + \frac{1-z}{\ln(1-z)} = -\frac{z}{\ln(1-z)} \\ -\frac{\ln(1-z)}{z} &= \frac{\sum_{h=1}^{\infty} \frac{z^h}{h}}{z} = \sum_{j=1}^{\infty} \frac{z^{j-1}}{j} \\ &= \sum_{j=0}^{\infty} \frac{z^j}{j+1} \end{aligned}$$

$$\begin{aligned} g(z, -1, 0) \cdot \left(-\frac{1}{2} \frac{\ln(1-z)}{z}\right) &= 1 \\ \left(\sum_{j=0}^{\infty} b_i^* z^i\right) \left(\sum_{j=0}^{\infty} \frac{z^j}{j+1}\right) &= 1 \end{aligned}$$

$$\begin{aligned} (b_0^* + b_1^* z + b_2^* z^2 + \dots)(1 + \frac{1}{2}z + \frac{1}{3}z^2 + \dots) \\ = b_0^* + z(\frac{1}{2}b_0^* + b_1^*) + z^2(\frac{1}{3}b_0^* + \frac{1}{2}b_1^* + b_2^*) + \dots \end{aligned}$$

$$b_0^* = 1, \quad b_1^* = -\frac{1}{2}b_0^*, \quad b_2^* = -\frac{1}{3}b_0^* - \frac{1}{2}b_1^*$$

⋮

$$b_i^* = -\sum_{j=0}^{i-1} \frac{b_j^*}{i+1-j}, \quad i = 1, 2, \dots$$

$$b_0^* = 1$$

$$b_1^* = -\frac{1}{2}b_0^* = -\frac{1}{2}$$

$$b_2^* = -\frac{1}{3} + \frac{1}{4} = -\frac{1}{12}$$

$$b_3^* = -\frac{1}{4} + \frac{1}{6} - \frac{1}{24} = -\frac{1}{24}$$

...

z.B. $s = 3$

$$y_{k+3} = y_{k+2} + h \sum_{i=0}^3 b_i^* \nabla^i f_{k+3}$$

$$b_{m-r} = (-1)^r \sum_{i=r}^m \binom{i}{r} b_i^*$$

$$\nabla^0 f_{k+3} = f_{k+3}$$

$$\nabla^1 f_{k+3} = f_{k+3} - f_{k+2}$$

$$\nabla^2 f_{k+3} = f_{k+3} - 2f_{k+2} + f_{k+1}$$

$$\nabla^3 f_{k+3} = f_{k+3} - 3f_{k+2} + 3f_{k+1} - f_k$$

$$y_{k+3} = y_{k+2} + h(1 \cdot \nabla^0 f_{k+3} - \frac{1}{2} \nabla^1 f_{k+3} - \frac{1}{12} \nabla^2 f_{k+3} - \frac{1}{24} \nabla^3 f_{k+3})$$

$$= y_{k+2} + h(f_{k+3} - \frac{1}{2}(f_{k+3} - f_{k+2}) - \frac{1}{12}(f_{k+3} - 2f_{k+2} + f_{k+1})$$

$$- \frac{1}{24}(f_{k+3} - 3f_{k+2} + f_{k+1} - f_k))$$

$$= y_{k+2} + h(\frac{1}{24}f_k - \frac{5}{24}f_{k+1} + \frac{19}{24}f_{k+2} + \frac{9}{24}f_{k+3})$$

$$\sum b_i = \frac{1}{24} - \frac{5}{24} + \frac{19}{24} + \frac{9}{24} = \frac{24}{24} 1$$

Die ersten fünf Koeffizienten b_i^*

	i=0	i=1	i=2	i=3	i=4
Adams–Bashforth	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$
Adams–Moulton	1	$-\frac{1}{2}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$
Nyström	2	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{29}{90}$
Milne–Simpson	2	-2	$\frac{1}{3}$	0	$-\frac{1}{90}$

1.3.2 Allgemeine lineare Mehrschrittverfahren

Definition

Ein *lineares Mehrschrittverfahren* hat die Form

$$\sum_{i=0}^s a_i y_{k+i} = h \sum_{i=0}^s b_i f_{k+i}$$

mit $a_i, b_i \in \mathbb{R}$ $a_s = 1$

Ein echtes s-Schrittverfahren liegt vor, wenn $|a_0| + |b_0| = 0$ ist.

$b_s \neq 0 \Rightarrow$ implizit

$b_s = 0 \Rightarrow$ explizit

Definition

Die Polynome

$$\varrho(\xi) = \sum_{i=0}^s a_i \xi^i \quad \text{und} \quad \sigma(\xi) = \sum_{i=0}^s b_i \xi^i$$

heißen *erstes* und *zweites charakteristisches Polynom*

Ein Verfahren ist durch das Paar (ϱ, σ) eindeutig festgelegt.

Konsistenz

Definition

Der *lokale Diskretisierungsfehler* eines Mehrschrittverfahrens ist

$$\tau(x, y(x), h) = \frac{1}{h} \sum_{i=0}^s a_i y(x + ih) - \sum_{i=0}^s b_i y'(x + ih)$$

wobei $y(x)$ die exakte Lösung des AWP's $y' = f(x, y)$, $y(x)$ gegeben.

Definition

Das Verfahren heißt *konsistent*, wenn

$$\tau(x, y(x), h) \rightarrow 0, \quad h \rightarrow 0$$

Das Verfahren besitzt die *Konsistenzordnung* $p > 0$, wenn

$$\tau(x, y(x), h) = O(h^p), \quad h \rightarrow 0$$

$$L(g, h) := \frac{1}{h} \sum_{i=0}^s (a_i g(ih) - h b_i g'(ih))$$

für $g \in \mathbb{C}^1([a, b])$

für $g(t) = y(x + t)$ ist $L(g, h)$ genau der *lokale Diskretisierungsfehler*.

Satz

Sei (ϱ, σ) ein s-Schrittverfahren.

Folgende Aussagen sind äquivalent

(i) Das Verfahren hat die Konsistenzordnung p

(ii) $L(t^m, 1) = 0$, $m = 0, 1, \dots, p$

$$\text{Hier ist } L(t^m, 1) = \frac{1}{h} \sum_{i=0}^s (a_i i^m - b_i m i^{m-1})$$

(iii) $w = 0$ ist $(p + 1)$ -fache Nullstelle von $\varrho(e^w) - w\sigma(e^w)$

(iv) $\xi = 1$ ist p -fache Nullstelle von $\frac{\varrho(\xi)}{\ln(\xi)} - \sigma(\xi)$

Siehe Beweis 7 (Anhang)

Beispiel 16: Anwendung von (ii)

Aufgabe: finde ein implizites 3-Schritt-Verfahren maximaler Ordnung mit

$$a_0 = a_1 = 0, \quad a_3 = 1$$

$$a_0 y_k + a_1 y_{k+1} + a_2 y_{k+2} + a_3 y_{k+3}$$

$$= h(b_0 f_k + b_1 f_{k+1} + b_2 f_{k+2} + b_3 f_{k+3})$$

$$a_2 y_{k+12} + y_{k+3} = h(b_0 f_k + b_1 f_{k+1} + b_2 f_{k+2} + b_3 f_{k+3})$$

5 freie Parameter \Rightarrow Ordnung 4 (hoffentlich) möglich.

Bedingungen: $L(t^m, 1) = 0, \quad m = 0, 1, 2, 3, 4, \dots$

$$L(t^m, 1) = \sum_{i=0}^3 (a_i t^{im} - b_i m i^{m-1})$$

$$= 3^m + a_2 \cdot 2^m - b_3 m \cdot 3^{m-1} - b_2 m \cdot 2^{m-1} - b_1 m - b_0 m 0^{m-1}$$

$$\Rightarrow L(t^0, 1) = 1 + a_2 = 0 \quad \Rightarrow \quad a_2 = -1$$

$$L(t^1, 1) = 3 + 2a_2 - b_3 - b_2 - b_1 - b_0 = 0 \text{ usw.}$$

\Rightarrow Lineares Gleichungssystem (5×5)

$$\text{Lösung: } b_0 = \frac{1}{24}, \quad b_1 = -\frac{5}{24}, \quad b_2 = \frac{19}{24}, \quad b_3 = \frac{3}{8}$$

Also: 3-schrittige Adams-Moulton-Verfahren

Korollar

13-05-14

Ein Mehrschrittverfahren ist genau dann für jedes AWP mit $f \in \mathbb{C}^1(S)$ konsistent, wenn die folgenden **Konsistenzbedingungen** erfüllt sind.

$$\varrho(1) = 0 \quad \text{und} \quad \varrho'(1) = \sigma(1)$$

Siehe Beweis 8 (Anhang)

Beispiel 17: 3-schrittiges Adams-Moulton-Verfahren

$$y_{k+3} = y_{k+2} + h\left(\frac{1}{24}f_k - \frac{5}{24}f_{k+1} + \frac{19}{24}f_{k+2} + \frac{9}{24}f_{k+3}\right)$$

$$\varrho(\xi) = \xi^3 - \xi^2$$

$$y_{k+3} - y_{k+2} \quad \text{mit} \quad y_l = \xi^l \quad \Rightarrow \quad \xi^{k+3} - \xi^{k+2}$$

$$\sigma(\xi) = \frac{9}{24}\xi^3 + \frac{19}{24}\xi^2 - \frac{5}{24}\xi + \frac{1}{24} \quad \text{Konsistenzbedingungen:}$$

$$\varrho(1) = 1^3 - 1^2 = 0$$

$$\varrho'(\xi) = 3\xi^2 - 2\xi$$

$$\varrho'(1) = 3 - 2 = 1$$

$$\sigma(1) = \frac{9}{24} + \frac{19}{24} - \frac{5}{24} + \frac{1}{24} = 1$$

$$\Rightarrow \sigma(1) = \varrho'(1)$$

\Rightarrow Das Verfahren ist konvergent

$$\begin{aligned}\Psi(w) &:= \rho(e^w) - w\sigma(e^w) \\ &= e^{3w} - e^{2w} - \frac{w}{24}(9e^{3w} + 19e^{2w} - 5e^w + 1) \\ \text{Wichtige Vielfachheit hat die Nullstelle } w &= 0? \\ \Psi(0) &= 1 - 1 = 0 \\ \Psi'(w) &= 3e^{3w} - 2e^{2w} - \frac{1}{24}(9e^{3w} + 19e^{2w} - 5e^w + 1) - \frac{w}{24}(27e^{3w} + 38e^{2w} - 5e^w) \\ \Psi'(0) &= 3 - 2 - \frac{1}{24}(9+19) - 5 + 1 = 0 \\ \Psi''(x) &= 9e^{3w} - 4e^{2w} - \frac{2}{24}(27e^{3w} + 38e^{2w} - 5e^w) - \frac{w}{24}(81e^{3w} + 74e^{2w} - 5e^w) \\ \Psi''(0) &= 9 - 4 - \frac{1}{12}(27 + 38 - 5) = 5 - \frac{1}{12}60 = 0 \\ \text{usw.}\end{aligned}$$

1.3.3 Exkurs: Lineare Differenzgleichungen

Definition

Eine Folge der Gleichungen

$$y_{k+s} + a_{s-1,k}y_{k+s-1} + \dots + a_{s-1,k}y_{k+1} + a_{0,k}y_k = f_k, \quad k = 0, 1, 2, \dots$$

mit $a_{j,k}, f_k \in \mathbb{C}$ heißt eine lineare *Differenzgleichung* der Ordnung s .

$f_k \equiv 0 \Rightarrow$ Die Gleichung heißt *homogen*, sonst *inhomogen*.

Hängen die Koeffizienten $a_{s-1,k}, \dots, a_{0,k}$ nicht von k ab, so heißt die Gleichung eine *lineare Differenzgleichung mit konstanten Koeffizienten*

Lineare Differenzgleichungen heißen auch *lineare Rekursionen*.

Awp y_0, y_1, \dots, y_{s-1} sind gegeben.

Dann kann man die Werte y_s, y_{s+1}, \dots rekursiv berechnen

Beispiel 18: Fibonacci-Zahlen

$$y_0 = 0, \quad y_1 = 1, \quad y_{k+2} = y_k + y_{k+1}, \quad k = 0, 1, \dots$$

$$y_2 = y_0 + y_1 = 1$$

$$y_3 = y_1 + y_2 = 2$$

$$y_4 = y_2 + y_3 = 3$$

$$y_5 = y_3 + y_4 = 5$$

Dies ist eine lineare Differenzgleichung 2. Ordnung mit konstanten Koeffizienten

Satz

Für eine homogene lineare Differenzgleichung der Ordnung s gilt:

- 1) Bei gegebenen Anfangswerten $y_0, y_1, \dots, y_{s-1} \in \mathbb{C}$ existiert genau eine Lösung $\{y_k\}_{k=0}^{\infty}$
- 2) Die Menge aller Lösungen bilden ein s -dimensionalen Teilraum im Raum aller komplexen Zahlen. Eine Basis dieses Teilraums bilden die s Lösungsfolgen $\{y_k^{(j)}\}_{k=0}^{\infty}$, $j = 0, 1, \dots, s-1$ mit den Anfangswerten $y_k^{(j)} = \delta_{kj}$, $k = 0, 1, \dots, s-1$, $j = 0, 1, \dots, s-1$
- 3) Ist $a_{0,k} \neq 0$ für alle k , so sind die Folgen $\{y_{k+i}^{(0)}\}_{k=0}^{\infty}$, $\{y_{k+i}^{(1)}\}_{k=0}^{\infty}$, \dots , $\{y_{k+i}^{(s-1)}\}_{k=0}^{\infty}$ für jedes i linear unabhängig.

Siehe Beweis 9 (Anhang)

Satz

Die Gesamtheit aller Lösungen einer homogenen ($f_k = 0$) linearen Differenzengleichung mit konstanten Koeffizienten ($a_{s-1,k}$ unabh. von k)

$$y_{k+s} + a_{s-1}y_{k+s-1} + \dots + a_1y_{k+1} + a_0y_k = 0$$

erhält man durch

$$y_k = \sum_{i=0}^m p_i(k) \cdot \lambda_i^k$$

wobei λ_i paarweise verschiedene Nullstellen des charakteristischen Polynoms

$$\varrho(t) = t^s + a_{s-1}t^{s-1} + \dots + a_1t + a_0$$

sind.

Außerdem ist $p_i(k)$ ein Polynom vom Grad $\leq \tau_i = 1$, wobei τ_i die Vielfachheit der Nullstelle λ_i ist.

Anmerkung:

In den Koeffizienten für die $p_i(k)$ stehen uns $\sum_{i=1}^m \tau_i = s$ Freiheitsgrade zur Verfügung.

Diese Koeffizienten sind z.B. durch die Anfangswerte y_0, y_1, \dots, y_{s-1} festgelegt.

Siehe Beweis 10 (Anhang)

Beispiel 19: Fibonacci-Zahlen

$$y_0 = 0, \quad y_1 = 1$$

$$y_{k+2} - y_{k+1} - y_k = 0$$

$$\varrho(\lambda) = \lambda^2 - \lambda - 1 = 0$$

$$\Rightarrow \lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$$

Allgemeine Lösung:

$$y_k = A \left(\frac{1+\sqrt{5}}{2} \right)^k + B \left(\frac{1-\sqrt{5}}{2} \right)^k$$

$$y_0 = A \left(\frac{1+\sqrt{5}}{2} \right)^0 + B \left(\frac{1-\sqrt{5}}{2} \right)^0 = A + B \stackrel{!}{=} 0$$

$$y_1 = A \left(\frac{1+\sqrt{5}}{2} \right)^1 + B \left(\frac{1-\sqrt{5}}{2} \right)^1 \stackrel{!}{=} 1$$

$$A = \frac{\det \begin{pmatrix} 0 & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix}}{\det \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \end{pmatrix}} = \frac{1}{\frac{1-\sqrt{5}}{2} - \frac{1+\sqrt{5}}{2}} = \frac{1}{\sqrt{5}}$$

$$\Rightarrow B = -\frac{1}{\sqrt{5}}$$

$$\Rightarrow y_k = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^k - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^k$$

Beispiel 20

$$y_{k+2} - 4y_{k+1} + 4y_k = 0$$

$$y_0 = 1, \quad y_1 = 4$$

$$\varrho(\lambda) = \lambda^2 - 4\lambda + 4 = 0 \Rightarrow \lambda_{1,2} = 2$$

Allgemeine Lösung $y_k = (Ak + B) \cdot 2^k$

$$y_0 = B \stackrel{!}{=} 1$$

$$y_1 = (A + B) \cdot 2 \stackrel{!}{=} 4 \Rightarrow A = 1$$

$$\Rightarrow y_k = (k + 1)2^k$$

Beispiel 21

$$y_{k+2} - 4y_{k+1} + 4y_k = 4^k$$

$$y_0 = 1, y_1 = 4$$

Lösung der homog.: $y_k^{(h)} = (Ak + B)2^k$

Ansatz für $y_k^{(p)}$: $y_k^{(p)} = C4^k$

$$\Rightarrow C4^{k+2} = 4C \cdot 4^{k+1} - 4C \cdot 4^k = 4^k$$

$$\Rightarrow 4C = 1 \Rightarrow c = \frac{1}{4}$$

$$y_k^{(p)} = \frac{1}{4}4^k = 4^{k-1}$$

$$\Rightarrow y_k = (Ak + B)2^k + 4^{k-1}$$

Beispiel 22: 2-Schritt-Verfahren

$$y_{k+2} - 4y_{k+1} + 3y_k = -2hf_{k+1}$$

Wir lösen das AWP mit $y' = 1$, $y(0) = 1$

Exakte Lösung: $y(x) = 1 + x$

$$y_0 = 1, y_1 = 1 + h$$

$$y_{k+2} - 4y_{k+1} + 3y_k = -2h$$

$$\varrho(\lambda) = \lambda^2 + 3 = 0 \Rightarrow \lambda_{1,2} = 2 \pm 1$$

$$y_k^{(h)} = A3^k + B1^k = A3^k + B$$

Ansatz: $y_k^{(p)} = Ck$

$$\Rightarrow C(k+2) - 4C(k+1) + 3Ck = -2h$$

$$Ck + 2C - 4Ck - 4C + 3Ck = -2h$$

$$-2C = -2h \Rightarrow C = h$$

$$y_k^{(p)} = hk$$

$$y_k = AB^k + B + hk$$

$$y_0 = A + B = 1$$

$$y_1 = 3A + B + h = 1 + h \Rightarrow A = 0, B = 1$$

$$\Rightarrow y_k = 1 + hk \text{ mit } hk = y_R$$

$$\Rightarrow y_R = 1 + x \text{ (exakte Lösung)}$$

1.3.4 Konvergenz

13-05-21

DefinitionDer *globale Diskretisierungsfehler*

$$\varepsilon(x_k, h) = y_k - y(x_k)$$

DefinitionEin Mehrschrittverfahren heißt *konvergent*, wenn

$$\max_{x_k \in [a, b]} |\varepsilon(x_k, h)| \rightarrow 0, h \rightarrow 0$$

Das Verfahren hat *Konvergenzordnung* $p > 0$, wenn

$$\max_{x_k \in [a,b]} |\varepsilon(x_k, h)| = O(h^p), h \rightarrow 0$$

Bemerkung

Konsistenz ist nicht hinreichend für die Konvergenz.

Beispiel 23: 2-Schrittverfahren

$$\begin{aligned} y_{k+2} - 4y_{k+1} + 3y_k &= -2hf_{k+1} \\ \varrho(\xi) &= \xi^2 - 4\xi + 3 & \sigma(\xi) &= -2\xi \\ \varrho(1) &= 1 - 4 + 3 = 0 \\ \varrho'(\xi) &= 2\xi - 4 \\ \varrho'(1) &= 2 - 4 = -2 = \sigma(1) \\ \Rightarrow \text{Das Verfahren ist konsistent.} \end{aligned}$$

Wir betrachten das AWP

$$y' = -y, y(0) = 1$$

Die exakte Lösung $y = e^{-x}$
 $x_k = kh, y_0 = 1, y_1 = e^{-h}$

$$\begin{aligned} f_{k+1} &= -y_{k+1} \\ y_{k+2} - 4y_{k+1} + 3y_k &= 2hy_{k+1} \\ y_{k+2} - (4 + 2h)y_{k+1} + 3y_k &= 0 \end{aligned}$$

$$\begin{aligned} \lambda^2 - (4 + 2h)\lambda + 3 &= 0 \\ \lambda_{1,2} &= 2 + h \pm \sqrt{(2 + h)^2 - 3} \\ &= 2 + h \pm \sqrt{4 + 4h + h^2 - 3} = 1 + 2h + O(h^2) \\ &= 2 + h \pm \sqrt{h^2 + 4h + 1} \end{aligned}$$

In der Tat. $1 + 4h + h^2 = 1 + 4h + 4h^2 + O(h^2)$

$$\begin{aligned} \lambda_{1,2} &= 2 + h \pm (1 + 2h) + O(h^2) \\ \lambda_1 &= 2 + h - 1 - 2h + O(h^2) \\ &= 1 - h + O(h^2) = e^{-h} + O(h^2) \end{aligned}$$

$$\lambda_2 = 2 + h + 1 + 2h + O(h^2) = 3 + 3h + O(h^2) = 3e^h + O(h^2)$$

$$y_k = A\lambda_1^k + B\lambda_2^k$$

$$y_0 = A + B = 1$$

$$y_1 = A\lambda_1 + B\lambda_2 = e^{-h}$$

$$\begin{aligned} A &= \frac{\det \begin{pmatrix} 1 & 1 \\ e^{-h} & \lambda_2 \end{pmatrix}}{\det \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}} = \frac{\lambda_2 - e^{-h}}{\lambda_2 - \lambda_1} \\ &= \frac{(3+3h)-(1-h)+O(h^2)}{(3+3h)-(1-h)+O(h^2)} = 1 + O(h^2) \end{aligned}$$

$$B = 1 - A = 1 - 1 + O(h^2) = O(h^2)$$

$$\begin{aligned} y_2 &= (1 + O(h^2))\lambda_1^2 + O(h^2)\lambda_2^2 \\ &= (1 + O(h^2))e^{-2h} + O(h^2)3^2 e^{2h} + O(h^2) \\ &= e^{-2h} + O(h^2)3^2 e^{2h} + O(h^2) \end{aligned}$$

$$x_k = kh \quad y_k = e^{-x_k} + \underbrace{O(h^2)3^{x_k/h} e^x}_{\rightarrow \infty, h \rightarrow 0} + O(h^2)$$

y_k ist unbeschränkt \Rightarrow konvergiert nicht

Definition

Wir sagen, dass das erste charakteristische Polynom $\varrho(\xi)$ die *Wurzelbedingung* erfüllt, wenn alle seine Nullstellen betragsmäßig ≤ 1 sind und die Nullstellen vom Betrag 1 einfach sind.

Solche Mehrschrittverfahren heißen auch *stabil*.

Konvergenzsatz von Dahlquist

Ein lineares Mehrschrittverfahren ist genau dann konvergent, wenn es konsistent ist und sein erstes charakteristisches Polynom die Wurzelbedingung erfüllt.

Beispiel 24: 2-Schrittverfahren

$$y_{k+2} - 4y_{k+1} + 3y_k = -2hf_{k+1}$$

$$\varrho(\xi) = \xi^2 - 4\xi + 3$$

$$\xi_{1,2} = 2 \pm \sqrt{4-3} = 2 \pm 1$$

$$\xi_1 = 1, \xi_2 = 3$$

\Rightarrow Die Wurzelbedingung ist nicht erfüllt \Rightarrow dieses Verfahren konvergiert nicht.

Satz

Ist ein lineares Mehrschrittverfahren konvergent, so erfüllt ϱ die Wurzelbedingung
Siehe Beweis 11 (Anhang)

Satz

Ist ein lineares Mehrschrittverfahren konvergent, so ist es auch konsistent.
Siehe Beweis 12 (Anhang)

Es bleibt zu zeigen, dass Konsistenz und Wurzelbedingung hinreichend für die Konvergenz sind.

Siehe Beweis 13 (Anhang) für explizite Verfahren

Satz

Wir betrachten das AWP

$$y' = f(x, y), \quad y(x_0) = y_0$$

mit $f \in C^p(\varrho)$, $p \geq 1$

Wenn ein lineares s -Schrittverfahren die Konsistenzordnung p besitzt, sein erstes charakteristisches Polynom die Wurzelbedingung erfüllt und für die Startwerte gilt

$$|y_j - y(x_j)| = O(h^p), \quad h \rightarrow 0, \quad j = 0, 1, \dots, s-1$$

dann hat das Verfahren die Konvergenzordnung p .

Beispiel 25: Einschnittverfahren

$$y_{k+1} = y_k + h\varphi(x_k, y_k, h)$$

Wurzelbedingung:

$$\varrho(\xi) = \xi - 1$$

$\xi = 1 \Rightarrow$ Die Wurzelbedingung ist automatisch erfüllt.

Beispiel 26: Adams–Bashforth, Adams–Moulton, Nyström, Milne–Simpson usw.

$$y_{k+s} = y_{k+q} + h(\dots)$$

$$\varrho(\xi) = \xi^s - \xi^q = \xi^q(\xi^{s-q} - 1)$$

$\xi = 0$ eine mehrfache Nullstelle, $s-q$ einfache Nullstellen $\sqrt[s-q]{1}$

\Rightarrow Die Wurzelbedingung ist automatisch erfüllt.

Die rechte Seite ist so konstruiert, dass man die Konsistenzordnung

$$m+1 = \begin{cases} s & , \text{explizit} \\ s+1 & , \text{implizit} \end{cases}$$

\Rightarrow Konvergenzordnung $m+1$



Erste Dahlquist Schranke

Die größte erreichbare Ordnung eines konvergenten s -Schrittverfahrens ist

$$\begin{cases} s & , \text{explizit} \\ 2\lfloor \frac{s}{2} + 1 \rfloor & , \text{implizit} \end{cases}$$

$$s = 2m \quad 2\lfloor m+1 \rfloor = 2m+2 = s+2$$

$$s = 2m+1 \quad 2\lfloor m+\frac{1}{2}+1 \rfloor = 2m+2 = s+1$$

1.3.5 Schrittweitensteuerung

Problematik

Große $h \rightarrow$ schnell, aber ungenau Kleine $h \rightarrow$ genau aber sehr viel Aufwand

Ansatz: variable Schrittweite h_k

$$x_{k+1} = x_k + h_k$$

Methode nach Milne

Wir suchen eine Lösung eines AWP mit einem Mehrschrittverfahren

$$\sum_{i=0}^s a_i y_{k+i} = h \sum_{i=0}^s b_i f_{k+i}, \quad a_s = 1$$

der Ordnung p .

Zur Überwachung benutzen wir ein zweites Verfahren der gleichen Ordnung p :

$$\sum_{i=q}^s \tilde{a}_i y_{k+i} = h \sum_{i=q}^s \tilde{b}_i f_{k+i}, \quad \tilde{a}_s = 1$$

mit $q \leq s-1$ (q kann auch negativ sein)

In jedem Schritt wird y_{k+s} aus y_{k+s-1}, \dots, y_k nach dem ersten Verfahren berechnet.

Gleichzeitig wird \tilde{y}_{k+s} aus $y_{k+s-1}, \dots, y_{k+q}$ nach dem zweiten Verfahren berechnet.

Konsistenzordnung $p \Rightarrow$

$$y_{k+s} - y(x_{k+s}) = Ch^{p+1}y^{(k+1)}(x_k) + O(h^{p+2})$$

$$\tilde{y}_{k+s} - y(x_{k+s}) = \tilde{C}h^{p+1}y^{(k+1)}(x_k) + O(h^{p+2})$$

Wir nehmen an, dass $C \neq \tilde{C}$

$$y_{k+s} - \tilde{y}_{k+s} \approx (C - \tilde{C})h^{p+1}y^{(k+1)}(x_k)$$

$$y^{(k+1)}(x_k) \approx \frac{1}{(C - \tilde{C})h^{p+1}}(y_{k+s} - \tilde{y}_{k+s})$$

$$y_{k+s} - y(x_{k+s}) \approx \frac{C}{C - \tilde{C}}(y_{k+s} - \tilde{y}_{k+s})$$

$$|y_{k+s} - y(x_{k+s})| \approx \left| \frac{C}{C - \tilde{C}} \right| |y_{k+s} - \tilde{y}_{k+s}|$$

Kleiner Fehler \Rightarrow vergrößere h

Großer Fehler \Rightarrow verkleinere h

Beispiel 27: Trapezregel Ordnung 2

$$y_{k+1} = y_k + \frac{h}{2}(f_k + f_{k+1})$$

Zur Überwachung: 2-Schritt-Adams-Bashforth: $y_{k+2} = y_{k+1} + \frac{h}{2}(3f_{k+1} - f_k)$

$$y_{x_{k+1}} = y(x_k) + hy'(x_k) + \frac{h^2}{2}y''(x_k) + \frac{h^3}{6}y'''(x_k) + O(h^4)$$

Trapezregel:

$$y_{k+1} = y_k + \frac{h}{2}(f(x_k, y_k) + f(x_{k+1}, y_{k+1})) = y_k + \frac{h}{2}(y'(x_k) + y'(x_{k+1}))$$

$$= y_k + \frac{h}{2}(y'(x_k) + y'(x_k) + hy''(x_k) + \frac{h^2}{2}y'''(x_k) + O(h^3))$$

$$= y_k + hy'(x_k) + \frac{h^2}{2}y''(x_k) + \frac{h^3}{4}y'''(x_k) + O(h^4)$$

$$y_{k+1} - y(x_{k+1}) = (\frac{1}{4} - \frac{1}{6})h^3y'''(x_k) + O(h^4) = \frac{1}{12}h^3y'''(x_k) + O(h^4)$$

$$\Rightarrow C = \frac{1}{12}$$

Adams-Bashforth-Verfahren:

$$\tilde{y}_{k+1} = y_k + \frac{h}{2}(3f_k - f_{k-1}) = y_k + \frac{h}{2}(3f(x_k, y_k) - f(x_{k-1}, y_{k-1}))$$

$$= y_k + \frac{h}{2}(3y'(x_k) - y'(x_{k-1}))$$

$$= y_k + \frac{h}{2}(3y'(x_k) - (y'(x_k) - hy''(x_k) + \frac{h^2}{2}y'''(x_k) + O(h^3)))$$

$$= y_k + hy'(x_k) + \frac{h^2}{2}y''(x_k) - \frac{h^3}{4}y'''(x_k) + O(h^4)$$

$$\tilde{y}_{k+s} - y(x_{k+s}) = (-\frac{1}{4} - \frac{1}{6})h^3y'''(x_k) + O(h^4) = -\frac{5}{12}h^3y'''(x_k) + O(h^4)$$

$$\Rightarrow \tilde{C} = -\frac{5}{12}$$

$$\left| \frac{C}{C - \tilde{C}} \right| = \frac{\frac{1}{12}}{\frac{1}{12} + \frac{5}{12}} = \frac{1}{6}$$

$$\Rightarrow |y_{k+1} - y(x_{k+1})| \approx \frac{1}{6} |y_{k+1} - \tilde{y}_{k+1}|$$

1.3.6 Prädiktor-Korrektor-Verfahren

Man kombiniert ein implizites und ein explizites Verfahren mit gleichem s .
(z.B. Adams-Moulton und Adams-Bashforth)

Explizites Verfahren \Rightarrow Prädiktor

Implizites Verfahren \Rightarrow Koorektor

Explizites Verfahren:

$$\sum_{i=0}^s a_i y_{k+i} = h \sum_{i=0}^{s-1} b_i f_{k+i}$$

Implizites Verfahren:

$$\sum_{i=0}^s \tilde{a}_i y_{k+i} = h \sum_{i=0}^s \tilde{b}_i f_{k+i} \quad \text{jeweils mit } a_s = \tilde{a}_s = 1$$

Algorithmus:

$$y_{k+s}^{(0)} = - \sum_{i=0}^{s-1} a_i y_{k+i} + h \sum_{i=0}^{s-1} b_i f_{k+i}$$

$$z_{k+s} = - \sum_{i=0}^{s-1} \tilde{a}_i y_{k+i} + h \sum_{i=0}^{s-1} \tilde{b}_i f_{k+i}$$

$$y_{k+s}^{(l)} = z_{k+s} + h \tilde{b}_s f(x_{k+s}, y_{k+s}^{(l+1)}), \quad l = 1, 2, \dots$$

2 Lösung großer, dünn besetzter, linearer Gleichungssysteme

2.1 motivierendes Beispiel

Laplace-Gleichung:

$$-\Delta u = 0, \quad x \in \Omega \subset \mathbb{R}^d$$

$$u|_{\partial\Omega} = g$$

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \dots + \frac{\partial^2}{\partial x_d^2}$$

Erstes Beispiel:

$$d = 1, \quad \Omega = [0, 1]$$

$$-y''(x) = 0, \quad y(0) = g_0, \quad y(1) = g_1$$

Methode der finiten Differenzen:

$$\text{Diskretisierung:} \quad 0 = x_0 < x_1 < \dots < x_n = 1, \quad x_i = ih, \quad h = \frac{1}{n}$$

$$y(x_i) \approx y_i$$

$$y'(x_i) \approx \frac{y_{i+1} - y_i}{x_{i+1} - x_i} = \frac{y_{i+1} - y_i}{h}$$

$$y''(x_i) \approx \frac{y'(x_i) - y'(x_{i-1}))}{x_i - x_{i-1}} \approx \frac{\frac{y_{i+1} - y_i}{h} - \frac{y_i - y_{i-1}}{h}}{h} = \frac{1}{h^2}(y_{i+1} - 2y_i + y_{i-1})$$

$$y_0 = g_0$$

$$-\frac{1}{h^2}(-y_0 + 2y_1 - y_2) = 0$$

$$-\frac{1}{h^2}(-y_1 + 2y_2 - y_3) = 0$$

⋮

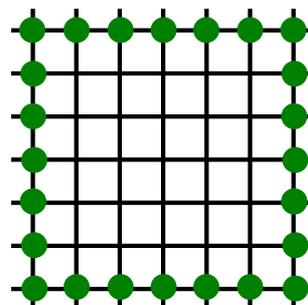
$$-\frac{1}{h^2}(-y_{n-2} + 2y_{n-1} - y_n) = 0$$

$$y_n = g_n$$

⇒ n Gleichungen mit $n - 1$ Unbekannten.

$$-\frac{1}{h^2} = \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & 1 & 0 & \dots & 0 \\ 0 & -1 & 2 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} f_1 \\ 0 \\ \vdots \\ 0 \\ f_{n-1} \end{pmatrix}$$

Bandmatrix (tridiagonale Matrix!)



2. Beispiel:

$$d = 2, \quad \Omega = [0, 1]^2$$

$$h = \frac{1}{n}, \quad x_{ij} = (ih, jh)$$

$x_{0,j}, x_{n,j}, x_{i,0}, x_{i,n}$ sind gegeben.

$$\Delta y = \frac{\partial^2 y}{\partial x_1^2} + \frac{\partial^2 y}{\partial x_2^2} = \frac{y_{i+1,j} - 2y_{i,j} + y_{i-1,j}}{h^2} + \frac{y_{i,j} - 2y_{i,j} + y_{i,j-1}}{h^2}$$

5-Punkte-Schema

$$-\Delta y = \frac{1}{h^2}(-y_{i+1,j} - y_{i-1,j} - y_{i,j+1} - y_{i,j-1} + 4y_{i,j})$$

$Y = (y_{ij})_{i,j=0}^n$ mit lexikographischer Ordnung

$AY = F$, A ist dünn besetzt.

$$A = \frac{1}{h^2} \begin{pmatrix} T & -I & & & & \\ -I & T & -I & & & \\ & \ddots & \ddots & \ddots & & \\ & & -I & T & -I & \\ & & & -I & T & \\ & & & & & \end{pmatrix}$$

mit $T = \begin{pmatrix} 4 & -1 & & & & \\ -1 & 4 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 4 & -1 & \\ & & & -1 & 4 & \end{pmatrix}$

2.2 Gaußelimination/Cholesky-Zerlegung

Zerlegung:

A sei symmetrisch und positiv definit.

$$A = LL^T$$

mit einer unteren Dreiecksmatrix L.

Algorithmus (*Cholesky-Zerlegung*):

$$\text{für } j = 1, \dots, n : l_{j,j} = \sqrt{a_{jj} - \sum_{i=1}^{j-1} l_{ji}^2}$$

$$\text{für } k = j + 1, \dots, n : l_{kj} = \frac{1}{l_{jj}} \left(a_{kj} - \sum_{i=1}^{j-1} l_{ki} l_{ji} \right)$$

2.2.1 Matrizen und Bäume

Ein *Graph* ist eine Menge von *Knoten* (Ecken) E und *Kanten* K: $g(E, K)$

Für eine dünn besetzte, symmetrische Matrix $A \in \mathbb{R}^n$ nehmen wir

$$E = \{1, 2, \dots, n\}$$

und definieren den Graph folgendermaßen: $(i, j) \in K \Leftrightarrow a_{ij} \neq 0$

Beispiel 1

1) Tridiagonalmatrix

$$\begin{array}{c} \begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \end{pmatrix} \end{matrix} \end{array}$$

$$a_{12} \neq 0 \Rightarrow (1, 2) \in K$$

$$(1, 3) \notin K \Rightarrow a_{1,3} = 0$$

monotoner Baum!



2) Pfeilspitzmatrizen

$$\begin{pmatrix} * & * & * & * & * & * \\ * & * & & & & \\ * & & * & & & \\ * & & & * & & \\ * & & & & * & \\ * & & & & & * \end{pmatrix}$$

nicht monotoner Baum!

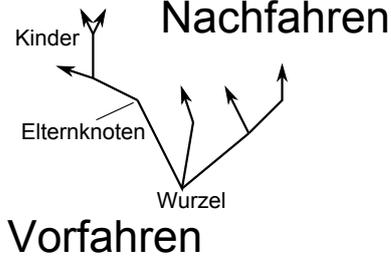
3) Zirkulante

$$\begin{pmatrix} * & * & & & * \\ * & * & * & & \\ & * & * & * & \\ & & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{pmatrix}$$

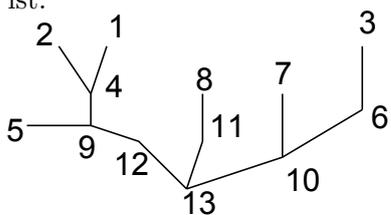
Kein Baum!

Ein *Baum* ist ein zusammenhängender Graph ohne Zyklen (d.h. jeweils zwei Knoten sind durch genau einen Pfad verbunden)

ein *gewurzelter Baum* ist ein gerichteter Baum mit einem ausgezeichneten Knoten (*Wurzel*), so dass jeder Knoten durch einen gerichteten Pfad von der Wurzel ausgehend erreichbar ist.



Ein gewurzelter Baum heißt *monoton*, falls jeder Knoten vor allen seinen Vorfahren nummeriert ist.



2.2.2 Satz von Parter

Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische, positiv definite Matrix und ihr Graph ein monotoner, gewurzelter Baum. Dann besitzt A eine Cholesky-Zerlegung $A = LL^T$ mit

$$l_{kj} = \frac{a_{kj}}{l_{jj}}, \quad 1 \leq j < k \leq n$$

(alle nicht diagonalen Elemente)

Insbesondere $a_{kj} = 0 \Rightarrow l_{kj} = 0$
 d.h. die Besetzungsstruktur bleibt erhalten.
 Siehe Beweis 14 (Anhang)

Beispiel 2

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

$$L = \begin{pmatrix} \sqrt{2} & 0 & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{\sqrt{3}}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & -\frac{\sqrt{2}}{\sqrt{3}} & \frac{2}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 0 & -\frac{2}{\sqrt{3}} & \frac{\sqrt{6}}{\sqrt{5}} \end{pmatrix}$$

$$j = 1 : l_{11} = \sqrt{a_{11}} = \sqrt{2}$$

$$l_{21} = \frac{a_{21}}{l_{11}} = -\frac{1}{\sqrt{2}}$$

$$l_{31} = \frac{a_{31}}{l_{11}} = 0$$

etc.

$$j = 2 : l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{2 - \frac{1}{2}} = \frac{\sqrt{3}}{\sqrt{2}}$$

$$j = 2 : l_{32} = \frac{1}{l_{22}}(a_{32} - l_{31}l_{21}) = -\frac{\sqrt{2}}{\sqrt{3}}$$

$$j = 2 : l_{42} = \frac{1}{l_{22}}(a_{42} - l_{41}l_{21}) = -\frac{\sqrt{2}}{\sqrt{3}}$$

$$l_{54} = \frac{1}{l_{44}}(a_{54} - l_{51}l_{41} - l_{52}l_{42} - l_{53}l_{43})$$

2.3 Iterative Verfahren für große LGS

13-05-31

2.3.1 Standard Matrixiterationen

Wir betrachten ein LGS $Ax = b$ mit $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ und $A = P - N$ wobei P eine „einfach“ (Tridiagonal) zu invertierende Matrix und N „irgendeine“ Matrix ist.

wir verwenden die Iteration $Px_{k+1} = Nx_k + b$, wobei P eine „einfache“ Faktorisierung hat bzw. $Px_{k+1} = cx_k$ lösbar ist.

Matrixiteration:

$$x_{k+1} = P^{-1}Nx_k + P^{-1}b = Hx_k + V$$

Satz

Das Iterationsverfahren $x_{k+1} = Hx_k + V$ konvergiert genau dann für jeden Startwert x_1 , wenn der Spektralradius von H kleiner als 1 ist, also $\rho(H) < 1$ mit $\rho(H) = \{|\lambda| \mid \lambda \text{ gr. EW v. } H\}$
Siehe Beweis 15 (Anhang)

Beispiel 3: unvollständige LR-Faktorisierung

Sei $A = P - N$ mit $P = LR$ (leicht zu berechnen, z.B. symmetr. + tridiagonal), N aber nicht leicht zu faktorisieren.

Wir iterieren:

$$\begin{aligned} Px_{k+1} &= Nx_k + b \\ LRx_{k+1} &= Nx_r + b \quad \text{für } k = 0, 1, \dots \end{aligned}$$

und lösen $L \underbrace{Rx_{k+1}}_{y_{k+1}} = c_k$, wobei man c_k bekommt mit $Ly_{k+1} = c_k$, $Rx_{k+1} = y_{k+1}$

Satz von Householder–John

Sei $A \in \mathbb{R}^{n \times n}$ symm., sowie A und $P^+P^T - A$ pos. definit.

Dann ist $\rho(H) < 1$

Siehe Beweis 16 (Anhang)

Beispiel 4

Tridiagonalmatrix:

$$\begin{pmatrix} \alpha_1 & \beta_1 & \cdots & 0 \\ \beta_1 & \alpha_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_{n-1} & \beta_{n-1} \\ 0 & \cdots & \beta_{n-1} & \alpha_n \end{pmatrix}$$

angenommen A pos def. und

a) $P = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$, dann gilt:

$$P + P^T - A = \begin{pmatrix} \alpha & -\beta_1 & 0 & & 0 \\ -\beta_1 & \alpha_2 & -\beta_2 & & \\ 0 & -\beta_2 & \alpha_3 & & \\ & & & \ddots & \\ 0 & & & & \alpha_n \end{pmatrix}$$

also $y^T(P + P^T - A)y = x^T Ax > 0$, mit $x_j = y_j(-1)^j$,

denn die Einträge der quadr. Form.

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

die mit $-\beta_k$ auftreten, sind genau die, bei denen $j = \pm i$ ist.

b) $P = \begin{pmatrix} \alpha_1 & 0 & & & 0 \\ \beta_1 & \alpha_2 & & & \\ 0 & \beta_2 & \alpha_3 & & \\ & & & \ddots & \\ 0 & & & & \alpha_n \end{pmatrix}$ dann gilt:

$$P + P^T - A = \begin{pmatrix} \alpha_1 & & & & \\ & \ddots & & & \\ & & & & \\ & & & & \alpha_n \end{pmatrix} \text{ pos. def., da } \alpha_i > 0, \text{ für alle } i.$$

Beispiel 5

$$\text{Sei } A = \underbrace{D}_{\text{diagonal}} - \underbrace{L_0}_{\text{links unten}} - \underbrace{R_0}_{\text{rechts oben}} = \begin{pmatrix} * & & -R_0 \\ & D & \\ -L_0 & & * \end{pmatrix}$$

Somit können wir für die Matrix H versch. Verfahren anwenden:

$$\begin{aligned} D &:= H = D^{-1}(L + R) && \text{Jacobi} \\ \tilde{L} &:= H = (D - L)^{-1}R && \text{Gauss - Seidel} \\ \tilde{L}\omega &:= H(D - \omega L)^{-1}((1 - \omega)D + \omega R) && \text{S.O.R.} \\ &\text{mit Parameter } \omega \in (0, 2) \end{aligned}$$

Mit D,L,R wie oben folgt dann im Einzelnen:

$$\begin{aligned} P = D & \quad N = L + R && \text{Jacobi} \\ P = D - L & \quad N = R && \text{Gauss - Seiden} \\ P = \frac{1}{\omega}D - L & \quad N = (\frac{1}{\omega} - 1)D + R && \text{SOR} \end{aligned}$$

(SOR: successive over relaxation)

Bemerkung

Ist A strikt diagonaldominant, d.h. $\sum_{\substack{k=i \\ k \neq j}} |a_{jk}| < |a_{ji}|$ für $j = 1, \dots, n$

So sind $\rho(B)$, $\rho(\tilde{L})$, jeweils kl. 1, d.h. die ersten beiden Verfahren konvergieren.

Jacobi-Verfahren:

$$\begin{aligned} Dx_{k+1} &= (L_0 + R_0)x_k + b \\ x_{k+1} &= \underbrace{D^{-1}(L_0 + R_0)x_k}_H + \underbrace{D^{-1}b}_c \\ x_{k+1} &= Hx_k + c \quad H = D^{-1}(L_0 + R_0) \\ L &:= D^{-1}L_0, \quad R = D^{-1}R_0, \quad H = L + R = B \end{aligned}$$

Gauß-Seidel:

$$(D - l_0)^{-1}R_0 = H = \tilde{L}$$

SOR:

$$(D - \omega l_0)^{-1}((1 - \omega)D + \omega R_0) = \tilde{L}_\omega$$

Konvergenz für die Verfahren

Idee:

Für jedes $H \in \{B, \tilde{L}, \tilde{L}_\omega\}$, Konvergenzbeweise über Fixpunktsätze gehen immer: über Spektralradius.

Vorteil: universell

Nachteil: numerisch aufwendig.

Einfache Fälle: A strikt diagonaldominant.

$$\text{d.h. } |a_{jj}| > \sum_{\substack{k=1 \\ k \neq j}}^n |a_{jk}| \forall 1 \leq j \leq n \quad \Rightarrow \rho(B), \quad \rho(\tilde{L}) < 1$$

2.3.2 Stein–Rosenberg

Sei D invertierbar und $B \geq 0$ komponentenweise. Dann gelten entweder
 $\varrho(B) = \varrho(\tilde{L}) = 0 \vee \varrho(B) = \varrho(\tilde{L}) = 1 \vee \varrho(\tilde{L}) < \varrho(B) < 1 \vee \varrho(\tilde{L}) > \varrho(B) > 1$

Satz

Sei A tridiagonal. Dann gilt

$$\lambda \in \sigma(B) \Rightarrow \lambda^2 \in \sigma(\tilde{L})$$

$$\mu \in \sigma(\tilde{L}) \setminus \{0\} \Rightarrow \pm\sqrt{\mu} \in \sigma(B)$$

Bemerkung

Vgl Stein–Rosenberg $\varrho(B)$ mit einem größten EW $\lambda \Rightarrow \lambda^2 \in \sigma(\tilde{L})$

Beispiel 6

$$A = \begin{pmatrix} 3 & 261 & \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

$$D = \begin{pmatrix} 3 & & \\ & 3 & \\ & & 3 \end{pmatrix} \quad L_0 = \begin{pmatrix} 0 & & \\ -2 & 0 & \\ -1 & -2 & 0 \end{pmatrix}$$

$$R_0 = \begin{pmatrix} 0 & -2 & -1 \\ & 0 & -2 \\ & & 0 \end{pmatrix}$$

$$B = D^{-1}(L_0 + R_0)$$

$$B = \begin{pmatrix} 0 & -\frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & 0 & -\frac{2}{3} \\ -\frac{1}{3} & -\frac{2}{3} & 0 \end{pmatrix} \quad \left| \begin{array}{ccc} -\lambda & -\frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & -\lambda & -\frac{2}{3} \\ -\frac{1}{3} & -\frac{2}{3} & -\lambda \end{array} \right|$$

$$\chi(\lambda) = -\lambda^3 - \frac{4}{27} + \lambda\left(\frac{1}{9} + \frac{4}{9} + \frac{4}{9}\right) = -\lambda^3 + \lambda - \frac{4}{27}$$

$$\varrho(B) = 1.12409... > 1 \quad \text{Jacobi divergiert}$$

Gauß–Seidel:

$$\tilde{L} = (D - l_0)^{-1}R_0 = \begin{pmatrix} 3 & 0 & 0 \\ 2 & 3 & 0 \\ 1 & 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -2 & -1 \\ & 0 & -2 \\ & & 0 \end{pmatrix}$$

$$\varrho(\tilde{L}) \approx 0.45049 < 1 \quad \text{Gauß–Seidel konvergiert}$$

Beispiel 7

$$A = \begin{pmatrix} 3 & 2 & 1 \\ -2 & 3 & 2 \\ -1 & -2 & 3 \end{pmatrix} \quad D = \begin{pmatrix} 3 & & \\ & 3 & \\ & & 3 \end{pmatrix}$$

$$R = \begin{pmatrix} 0 & -2 & -1 \\ 0 & -2 & \\ 0 & & \end{pmatrix} \quad L_0 = \begin{pmatrix} 0 & & \\ 2 & 0 & \\ 1 & 2 & 0 \end{pmatrix}$$

$$B = \left(\begin{array}{ccc|ccc} 0 & -\frac{2}{3} & -\frac{1}{3} & -\lambda & -\frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & 0 & -\frac{2}{3} & \frac{2}{3} & -\lambda & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & 0 & \frac{1}{3} & \frac{2}{3} & -\lambda \end{array} \right)$$

$$\chi = -\lambda^3 + \frac{4}{9} - \frac{4}{9} - \frac{1}{9}\lambda = -\lambda^3 - \lambda = -\lambda(\lambda^2 + 1)$$

$$\sigma(B) = \{0, i, -i\}, \quad \rho(B) = 1$$

$$\tilde{L} = \begin{pmatrix} 3 & 0 & 0 \\ -2 & 3 & 0 \\ -1 & -2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -2 & -1 \\ & 0 & -2 \\ & & 0 \end{pmatrix}$$

$$\sigma(\tilde{L}) = \{0, \frac{1}{54}(23 \pm \sqrt{97})\}, \quad \sigma 8\tilde{L} < 1$$

Beispiel 8

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

Ausrechnen d. Eigenwerte: Übung.

$$\rho(\tilde{L}) = \frac{1}{54}(32 \pm \sqrt{1393}), \quad \rho(B) > 1$$

Beide Verfahren divergent.

SOR: Verfahren: Konvergenz und Wahl des Parameters ω (damit Spektralradius minimiert wird)

Besetzungsparameter der Matrix $A \in \mathbb{R}^{n \times n}$. Betrachte $\mathcal{K} \subset \mathbb{N} \times \mathbb{N}$ mit

$$(k, l) \in \mathcal{K} \Rightarrow a_{kl} \neq 0, \quad k \leq l$$

$$A = (a_{kl})_{k,l}^n$$

Definition: Ordnungsvektor

$j \in \mathbb{N}^n$ heißt Ordnungsvektor zu der Matrix $A \in \mathbb{R}^{n \times n}$, falls

$$|j_k - j_l| = 1 \quad \forall (k, l) \in \mathcal{K}$$

j heißt (sogar) **kompatibler Ordnungsvektor**, falls j Ordnungsvektor ist und

$$j_k - j_l = +1, \text{ falls } k > l \text{ und } \quad j_k - j_l = -1, \text{ falls } k < l \quad \forall (k, l) \in \mathcal{K}$$

Lemma

Falls $A \in \mathbb{R}^{n \times n}$ einen Ordnungsvektor besitzt, gibt es eine Permutationsmatrix $P \in \mathbb{Z}^{n \times n}$ s.d. PAP^{-1} einen kompatiblen Ordnungsvektor besitzt.

Siehe Beweis 17 (Anhang)

Ziel: Spektren (Spektralradien) für $B, \tilde{L}, \tilde{L}_\omega$ in Beziehung bringen für Matrizen die Ordnungsvektoren haben

Beispiel 9

$$\begin{pmatrix} * & 0 & * & 0 & * \\ 0 & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & 0 & 0 & * & * \\ * & 0 & 0 & * & * \end{pmatrix}$$

besitzt den Ordnungsvektor $j = (2, 2, 1, 2, 3)^T$, der aber nicht kompatibel ist (Beispiel $a_{31} \neq 0$)

Betrachte Permutation:

$$\pi^{-1}(1) = 3, \quad \pi^{-1}(2) = 1, \quad \pi^{-1}(3) = 2, \quad \pi^{-1}(4) = 4, \quad \pi^{-1}(5) = 5$$

Der neue Ordnungsvektor ist $(1, 2, 2, 2, 3)^T = i$

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

$$PAP^{-1} = \begin{pmatrix} * & * & * & 0 & 0 \\ * & * & 0 & 0 & * \\ * & 0 & * & 0 & 0 \\ 0 & 0 & 0 & * & * \\ * & * & 0 & * & * \end{pmatrix}$$

Lemma

Habe A einen kompatiblen Ordnungsvektor.

Sei $A = D - L_0 - R_0$. Dann ist

$$g(s, t) = \det(tL_0 + t^{-1}R_0 - sD)$$

mit $s \in \mathbb{R}$, $t \in \mathbb{R}$, $t \neq 0$ unabhängig von t .

Siehe Beweis 18 (Anhang)

Satz:

Habe $A \in \mathbb{R}^{n \times n}$ einen kompatiblen Ordnungsvektor. B sei die Iterationsmatrix des Jacobi-Verfahrens.

Dann gilt:

1. $\mu \in \sigma(B) \Rightarrow -\mu \in \sigma(B)$
2. $\mu \in \sigma(B) \wedge 0 < \omega < 2 \wedge \lambda$ erfüllt $\lambda + \omega - 1 = \omega\mu\sqrt{\lambda} \Rightarrow \lambda \in \sigma(\tilde{L}\omega)$
3. falls $0 < \omega < 2$, $\lambda \in \sigma(\tilde{L}\omega)$, so gibt es ein $\mu \in \sigma(B)$ mit $\lambda + \omega - 1 = \omega\mu\sqrt{\lambda}$

Siehe Beweis 19 (Anhang)

Korollar

Habe $A \in \mathbb{R}^{n \times n}$ einen kompatiblen Ordnungsvektor. Dann gilt:

$$\varrho(\tilde{L}) = \varrho(B)^2$$

Siehe Beweis 20 (Anhang)

Satz

Mit $\varrho(\tilde{L}_\omega) < 1 \Rightarrow 0 < \omega < 2$

Daher muss wie vorher A einen kompatiblen Ordnungsvektor haben.

Siehe Beweis 21 (Anhang)

Sei $\lambda \in \sigma(\tilde{L}_\omega)$, es gibt ein $\mu \in \sigma(B)$ s.d. $p(\sqrt{\lambda}) = 0$ mit $p(z) = z^2 - \omega\mu\lambda + \omega - 1$

Cohn-Schur-Kriterium über die Lage der Nullstellen eines quadratischen Polynoms:

Beide Nullstellen von $\alpha z^2 + \beta z + \gamma$ ($\alpha, \beta, \gamma \in \mathbb{C}$) liegen in $\overline{B(0)}$ genau dann, wenn $|\alpha|^2 \geq |\gamma|^2$, $(|\alpha|^2 - |\gamma|^2)^2 \geq (\alpha\bar{\beta} - \beta\bar{\gamma})^2$

Zusatz: Für Nullstellen in $\overset{\circ}{B}_1(0)$ nimmt man das gleiche Kriterium, aber $|\alpha| > |\gamma|^2$

Hier: beide Nullstellen in $\overset{\circ}{B}_1(0)$ wenn

$$(\omega - 1)^2 < 1 \text{ und } \mu^2 < 1 \text{ d.h. } \varrho(B) < 1$$

$$\varrho(B) < 1 \text{ und } (\omega - 1)^2 < 1 \text{ (bzw. } |\omega - 1| < 1 \text{ bzw. } 0 < \omega < 2) \Leftrightarrow \varrho(\tilde{L}_\omega) < 1$$

Satz:

Habe $A \in \mathbb{R}^{n \times n}$ einen kompatiblen Ordnungsvektor, $\sigma(B) \subset \mathbb{R}$, $\bar{\mu} = \varrho(B) < 1$

Dann gilt

$$\varrho(\tilde{L}_\omega) > \varrho(\tilde{L}_{\omega_{opt}}) \quad \forall 0 < \omega < 2, \quad \omega \neq \omega_{opt}$$

$$\text{wobei } \omega_{opt} = \frac{2(1-\sqrt{1-\bar{\mu}^2})}{\bar{\mu}^2} = 1 + \left(\frac{\bar{\mu}}{1+\sqrt{1-\bar{\mu}^2}} \right)^2 \in (1, 2)$$

$$\text{Dabei } \varrho(\tilde{L}_{\omega_{opt}}) = \omega_{opt} - 1 = \left(\frac{\bar{\mu}}{1+\sqrt{1-\bar{\mu}^2}} \right)^2 \in (0, 1)$$

Siehe Beweis 22 (Anhang)

Beispiel 10

$$\begin{array}{ll} \varrho(B) = \frac{3}{5} = 0.6 & \text{Jacobi} \\ \varrho(\tilde{L}_1) = 0.36 & \text{Gauß-Seidel} \\ \varrho(\tilde{L}_{\omega_{opt}}) = \frac{1}{9} & \text{SOR mit opt. Parameter} \end{array}$$

Bemerkung:

Die Matrix A besitzt einen kompatiblen Ordnungsvektor \Leftrightarrow

$$\{1, 2, \dots, n\} = S_1 \cup S_2 \quad \text{und} \quad S_1 \cap S_2 = \emptyset$$

$$(k, l) \in K \Leftrightarrow (k \in S_1 \wedge l \in S_2) \vee (l \in S_2 \wedge k \in S_1)$$

2.3.3 kurze Zusammenfassung

Fixpunktiteration:

$$\begin{aligned}\vec{x} &= T\vec{x} + \vec{u}, \quad \vec{x}, \text{vecu} \in \mathbb{R}^n(\mathbb{C}^n) \\ T &\in \mathbb{R}^{n \times n} \quad (\mathbb{C}^{n \times n}) \\ I - T &\text{ regulär.}\end{aligned}$$

Iterationsverfahren:

$$\begin{aligned}\vec{x}^{(0)} &\in \mathbb{R}^n \text{ beliebig (der Startvektor)} \\ \vec{x}^{(k+1)} &= T\vec{x}^{(k)} + \vec{u}, \quad k = 0, 1, 2, \dots\end{aligned}$$

Satz

Das Iterationsverfahren konvergiert für jeden Startvektor $\vec{x}_0 \in \mathbb{R}^n$ gegen die Lösung genau dann, wenn $\rho(T) < 1$

$$\rho(T) = \max\{|\lambda| : \lambda \text{ ist EW von } T\} = \text{der } \textit{Spektralradius}$$

Lemma

Ist $\|\cdot\|$ eine natürliche Matrixnorm so gilt $\rho(T) < \|T\|$.

Wenn $\|\vec{x}\|$ gegeben in \mathbb{R}^n , dann ist $\|T\| = \max_{\|\vec{x}\|=1} \|T\vec{x}\|$ natürliche Matrixnorm.

Siehe Beweis 23 (Anhang)

Sei nun $A\vec{x} = \vec{b}$, A regulär,

$$A = D - L - R = \begin{pmatrix} D & -R & -R \\ -L & D & -R \\ -L & -L & D \end{pmatrix}$$

Wir gehen dabei davon aus, dass D regulär ist (d.h. alle Diagonalelemente von A sind von 0 verschieden).

Lemma

Sei $A \in \mathbb{R}^{n \times n}$ regulär, so gibt es eine Permutationsmatrix derart, dass alle Diagonalelemente von PA von 0 verschieden sind.

Siehe Beweis 24 (Anhang)

O.B.d.A. sind alle Diagonalelemente von A von Null verschieden:

$$\begin{aligned}A\vec{x} &= \vec{b} \\ (D - L - R)\vec{x} &= \vec{b} \\ D\vec{x} &= (L + R)\vec{x} + \vec{b} \\ \vec{x} &= \underbrace{D^{-1}(L + R)}_T \vec{x} + \underbrace{D^{-1}\vec{b}}_{\vec{u}}\end{aligned}$$

$$\vec{x} = T\vec{x} + \vec{u}$$

$$\text{mit } T = B = D^{-1}(L + R) \text{ und } \vec{u} = D^{-1}\vec{b}$$

„ Gesamtschrittverfahren“, „ Jacobi-Verfahren“

$\vec{x}^{(0)} \in \mathbb{R}^n$ beliebig.

$$\vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{u}, \quad k = 0, 1, 2, \dots$$

Um zum Gauß-Seidel-Verfahren zu kommen:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(- \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i \right), \quad i = 1, \dots, n$$

Aber für $j = 1, \dots, i-1$ haben wir schon $x_j^{(k+1)}$ berechnet

$$\Rightarrow \text{wir verwendensie an der Stelle von } x_i^{(k)}, \quad j = 1, \dots, i-1$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(- \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_j \right)$$

„ Einzelschrittverfahren“, „ Gauß-Seidel-Verfahren“.

$$a_{ij} x_i^{(k+1)} + \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} = - \sum_{j=i+1}^{n+1} a_{ij} x_j^{(k)} + b_i$$

$$(D - L)\vec{x}^{(k+1)} = R\vec{x}^{(k)} + \vec{b}$$

$$\vec{x}^{(k+1)} = (D - L)^{-1} R\vec{x}^{(k)} + (D - L)^{-1} \vec{b}$$

$$R = \tilde{L} = (D - L)^{-1} R$$

Satz

Sei A strikt diagonaldominant (d.h. $|a_{ii}| > \sum_{k \neq i} |a_{ik}|$, $i = 1, \dots, n$) dann konvergiert das Jacobi-Verfahren.

Satz

Sei A hermetisch und positiv definit, so ist das Gauß-Seidel-Verfahren konvergent.

Relaxation

$$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}} \left(- \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(j)} + b_j \right)$$

$$x_i^{(k+1)} = x_i^{(k)} + \omega(\tilde{x}_i^{(k+1)} - x_i^{(k)})$$

$\omega = 1 \Rightarrow$ Gauß-Seidel

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \omega\tilde{x}_i^{(k+1)}$$

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \omega \frac{1}{a_{ii}} \left(- \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i \right)$$

$$a_{ii} x_i^{(k+1)} + \omega \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} = (1 - \omega)a_{ii} x_i^{(k)} - \omega \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i$$

$$(D - \omega L)\vec{x}^{(k+1)} = ((1 - \omega)D - \omega R)\vec{x}^{(k)} + \vec{b}$$

$$\vec{x}^{(k+1)} = (D - \omega L)^{-1} ((1 - \omega)D + \omega R)\vec{x}^{(k)} + (D - \omega L)^{-1} \vec{b}$$

$$T = \tilde{L}_\omega = (D - \omega L)^{-1} ((1 - \omega)D + \omega R)$$

Satz von Kahan

$$\varrho(\tilde{L}_\omega) \geq |\omega - 1|$$

und $\varrho(\tilde{L}_\omega) = |\omega - 1|$ genau dann, wenn alle Eigenwerte von \tilde{L}_ω den Betrag $|\omega - 1|$ haben.

Satz

Sei A hermetisch und positiv definit. Dann $\varrho(\tilde{L}_\omega) < 1 \Leftrightarrow \omega \in (0, 2)$

2.3.4 Nachiteration

$A\vec{x} = \vec{b}$ wurde durch eine direkte Methode gelöst (z.B. LR-Zerlegung).

Ziel: Einflüsse von Rundungsfehlern durch Nachiteration zu dämpfen.

$$A \approx \tilde{L}\tilde{R}$$

$$A = \tilde{L}\tilde{R} + F \text{ mit } F \text{ Fehlermatrix.}$$

$\|F\|_\infty$ ist sehr klein.

$$A\vec{x} = \vec{b}$$

$$\Rightarrow (\tilde{L}\tilde{R} + F)\vec{x} = \vec{b}$$

$$\tilde{L}\tilde{L}\vec{x} = -F\vec{x} + \vec{b} = -(A - \tilde{L}\tilde{R})\vec{x} + \vec{b}$$

$$\tilde{L}\tilde{R}\vec{x} = \tilde{L}\tilde{R}\vec{x} - \underbrace{(A\vec{x} - \vec{b})}_{\text{Residuum}}$$

Da F sehr klein ist, gehen wir davon aus, dass \tilde{L} und \tilde{R} regulär sind.
 $\Rightarrow \vec{x} = \vec{x} - \tilde{R}^{-1}\tilde{L}^{-1}(A\vec{x} - \vec{b})$

$$\tilde{x} = T\vec{x} + \vec{u}$$

$$\text{mit } T = I - \tilde{R}^{-1}\tilde{L}^{-1}A, \quad \vec{u} = \tilde{R}^{-1}\tilde{L}^{-1}\vec{b}$$

Iterationsverfahren:

$\vec{x}^{(0)}$ = die Näherungslösung, die man mit dem direkten Verfahren bekommen hat.

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - \tilde{R}^{-1}\tilde{L}^{-1}(A\vec{x}^{(k)} - \vec{b}), \quad k = 0, 1, \dots$$

$\vec{x}^{(0)}$ ist eine sehr gute Näherung.

$$T = I - \underbrace{\tilde{R}^{-1}\tilde{L}^{-1}A}_{\approx A^{-1}}$$

$\Rightarrow \rho(T)$ ist sehr klein

\Rightarrow Das Verfahren konvergiert sehr schnell.

In der Regel sind 1 oder 2 Iterationen ausreichend.

Algorithmus:

$\vec{x}^{(0)}$ = mit der direkten Methode berechnete Näherungslösung.

Für $k=0,1,2,\dots$:

1. Berechne $\vec{r}^{(k)} = A\vec{x}^{(k)} - \vec{b}$ mit doppelter Stellenzahl.

2. Löse $\tilde{L}\vec{v}^{(k)} = -\vec{r}^{(k)}$ (einfache Genauigkeit)

$$\tilde{R}\vec{z}^{(k)} = \vec{v}^{(k)}$$

3. $\vec{x}^{(k+1)} = \vec{x}^{(k)} + \vec{z}^{(k)}$

2.3.5 Verfahren der konjugierten Gradienten

(auch *CG-Verfahren*)

$$A\vec{x} = \vec{b}, \quad A \in \mathbb{R}^{n \times n}, \quad \vec{b} \in \mathbb{R}^n$$

A ist symmetrisch und positiv definit.

Betrachte $f: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(\vec{x}) = \frac{1}{2}\vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$$

Bedingungen für ein Minimum:

$f(\vec{x}^*) \leq f(\vec{x})$ für alle $\vec{x} \in \mathbb{R}^n$ genau dann wenn

$$1. \nabla f(\vec{x}^*) = \left(\frac{\partial f(\vec{x})}{\partial x_i} \right)_{i=1}^n \Big|_{\vec{x}=\vec{x}^*} = \vec{0}$$

$$2. \nabla^2 f(\vec{x}^*) = \left(\frac{\partial^2 f(\vec{x})}{\partial x_i \partial x_j} \right)_{i,j=1}^n \Big|_{\vec{x}=\vec{x}^*} \text{ ist positiv definit}$$

$$f(\vec{x}) = \frac{1}{2} \sum_{i,j=1}^n a_{ij} x_i x_j - \sum_{i=1}^n b_i x_i$$

$$\frac{\partial f(\vec{x})}{\partial x_i} = \sum_{j=1}^n a_{ij} x_j - b_i$$

$$\nabla f(\vec{x}) = A\vec{x} - \vec{b}$$

$$\frac{\partial^2 f(\vec{x})}{\partial x_i \partial x_j} = a_{ij}$$

$$\nabla^2 f(\vec{x}) = A \text{ (positiv definit)}$$

$\vec{x}^* = \arg \min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}) \Leftrightarrow A\vec{x}^* = \vec{b}$ Folgende Aussagen sind äquivalent.

1. $A\vec{x}^* = \vec{b}$

2. \vec{x}^* ist das Minimum der Funktion $f(\vec{x}) = \frac{1}{2}\vec{x}^T A\vec{x} - \vec{b}^T \vec{x}$

$\vec{x}^{(0)} \in \mathbb{R}^n$ Startwert.

$\vec{x}^{(k+1)} = \vec{x}^{(k)} + \lambda_k \vec{d}^{(k)}, \quad k = 0, 1, \dots$

$\vec{d}^{(k)}$ ist die **Suchrichtung**.

13-06-18

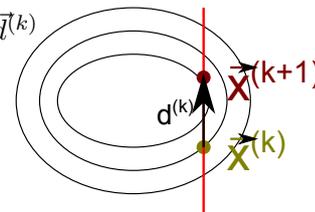
Sei $\vec{d}^{(k)}$ gegeben, dann bestimmen wir λ_k folgendermaßen:

$g_k(\lambda) = f(\vec{x}^{(k)} + \lambda \vec{d}^{(k)})$

$g_k(\lambda) = \frac{1}{2}(\vec{x}^{(k)})^T A\vec{x}^{(k)} + \lambda(\vec{x}^{(k)})^T A\vec{d}^{(k)} + \frac{1}{2}\lambda^2(\vec{d}^{(k)})^T A\vec{d}^{(k)} - \vec{b}^T \vec{x}^{(k)} - \lambda\vec{b}^T \vec{d}^{(k)}$

$g'_k(\lambda) = (\vec{x}^{(k)})^T A\vec{d}^{(k)} + \lambda(\vec{d}^{(k)})^T A\vec{d}^{(k)} - \vec{b}^T \vec{d}^{(k)} = 0$

$\lambda = -\frac{(\vec{x}^{(k)})^T A\vec{d}^{(k)} - \vec{b}^T \vec{d}^{(k)}}{(\vec{d}^{(k)})^T A\vec{d}^{(k)}}$



Residuen: $\vec{r}^{(k)} = A\vec{x}^{(k)} - \vec{b}$

$\lambda_k = -\frac{(A\vec{x}^{(k)} - \vec{b})^T \vec{d}^{(k)}}{(\vec{d}^{(k)})^T A\vec{d}^{(k)}} = -\frac{(\vec{r}^{(k)})^T \vec{d}^{(k)}}{(\vec{d}^{(k)})^T A\vec{d}^{(k)}}$

$$\lambda_k = -\frac{(\vec{r}^{(k)})^T \vec{d}^{(k)}}{(\vec{d}^{(k)})^T A\vec{d}^{(k)}}$$

Wahl der Suchrichtung: Die **Suchrichtungen** sollen linear unabhängig sein, z.B. orthogonal. wir verwenden die A-Orthogonalität.

Ein Skalarprodukt auf \mathbb{R}^n :

$\langle \vec{z}_1, \vec{z}_2 \rangle_A = (\vec{z}_1)^T A\vec{z}_2, \quad \vec{z}_1, \vec{z}_2 \in \mathbb{R}^n$

Das ist ein Skalarprodukt, da A positiv definit ist.

$\|z\|_A = \sqrt{\langle z, z \rangle_A}$

$\vec{x}^{[0]} \in \mathbb{R}^n$ Startwert.

$\vec{x}^{[k+1]} = \vec{x}^{[k]} + \lambda_k \vec{d}^{[k]}, \quad k = 0, 1, \dots$

$\vec{d}^{[k]}$ ist die Suchrichtung. $(\langle \vec{z}, \vec{z} \rangle_A)^{1/2} = (\vec{z}^T A\vec{z})^{1/2}$

$\vec{z}_1 \perp_A \vec{z}_2 \Leftrightarrow \langle \vec{z}_1, \vec{z}_2 \rangle_A = (\vec{z}_1)^T A\vec{z}_2 = 0$

Wir wählen die Suchrichtungen $\vec{d}^{(0)}, \vec{d}^{(1)}, \vec{d}^{(2)}, \dots, \vec{d}^{(n-1)}$, so dass $\vec{d}^{(j)} \perp_A \vec{d}^{(i)} \forall i \neq j$

Gram-Schmitt-Orthogonalisierung:

Seien $\{\vec{u}_0, \vec{u}_1, \dots, \vec{u}_{n-1}\}$ linear unabhängige Vektoren. Dann

$\vec{d}^{(0)} = \vec{u}_0$

⋮

$\vec{d}^{(i)} = \vec{u}_i + \sum_{k=0}^{i-1} \beta_{ik} \vec{d}^{(k)}, \quad i = 1, \dots, n-1$

$\langle \vec{d}^{(k)}, \vec{d}^{(l)} \rangle_A = 0, \quad k = 0, 1, \dots, i-1$

$\langle \vec{d}^{(k)}, \vec{d}^{(i)} \rangle_A = \langle \vec{d}^{(k)}, \vec{u}_i \rangle_A + \sum_{j=0}^{i-1} \beta_{ij} \langle \vec{d}^{(k)}, \vec{d}^{(j)} \rangle_A$

$= \langle \vec{d}^{(k)}, \vec{u}_i \rangle_A + \beta_{ik} \langle \vec{d}^{(k)}, \vec{d}^{(k)} \rangle_A$

$\beta_{ik} = -\frac{\langle \vec{d}^{(k)}, \vec{u}_i \rangle_A}{\langle \vec{d}^{(k)}, \vec{d}^{(k)} \rangle_A}$

$$\beta_{ik} = -\frac{(\vec{u}_i)^T A\vec{d}^{(k)}}{(\vec{d}^{(k)})^T A\vec{d}^{(k)}}$$

Fehler: $\vec{e}^{(i)} = \vec{x}^{(i)} - \vec{x}^*$ (\vec{x}^* ist die Lösung)

$A\vec{e}^{(i)} = A\vec{x}^{(i)} - A\vec{x}^* = A\vec{x}^{(i)} - \vec{b}$

Satz:

$$\vec{e}^{(n)} = \vec{0}$$

d.h. dass das Verfahren in höchstens n Schritten die Lösung liefert.
Siehe Beweis 25 (Anhang)

$$D_i = \text{span}\{\vec{d}^{(0)}, \dots, \vec{d}^{(i-1)}\}$$

$$\vec{e}^{(i)} = \vec{e}^{(0)} + D_i$$

Satz

$$\|\vec{e}^{(i)}\|_A \leq \|\vec{z}\|_A$$

für alle $\vec{z} \in \vec{e}^{(0)} + D_i$
Siehe Beweis 26 (Anhang)

Lemma

$$(\vec{r}^{(j)})^T \vec{d}^{(i)} = 0, \quad i = 0, \dots, j-1$$

Siehe Beweis 27 (Anhang)

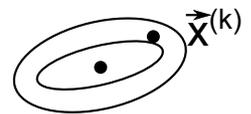
Wahl der Suchrichtung im CG-Verfahren

$$f(\vec{x}) = \frac{1}{2} \vec{x}^T A \vec{x} - \vec{b}^T \vec{x}$$

$$\nabla f(\vec{x}) = A \vec{x} - \vec{b}$$

$$\nabla f(\vec{x}^{(k)}) = \vec{r}^{(k)}$$

\Rightarrow Wir wählen $-\vec{r}^{(k)}$ als ursprüngliche Suchrichtung.



Wir wählen $\{\vec{u}_0, \vec{u}_1, \dots, \vec{u}_{n-1}\} = \{-\vec{r}^{(0)}, -\vec{r}^{(1)}, \dots, \vec{r}^{(n-1)}\}$ mit
 $\vec{d}^{(0)} = -\vec{r}^{(0)}$

\vdots

$$\vec{d}^{(i)} = -\vec{r}^{(i)} + \sum_{k=0}^{i-1} \beta_{ik} \vec{d}^{(k)}$$

$$\beta_{ik} = -\frac{(\vec{r}^{(i)})^T A \vec{d}^{(k)}}{(\vec{d}^{(k)})^T A \vec{d}^{(k)}}$$

Lemma

$$\beta_{ik} = 0, \quad k = 0, 1, \dots, i-2$$

Siehe Beweis 28 (Anhang)

Algorithmus (CG-Verfahren)

- 1) Wähle $\vec{x}^{(0)}$ (den Startvektor)

$$\vec{r}^{(0)} = A \vec{x}^{(0)} - \vec{b}$$

$$\vec{d}^{(0)} = -\vec{r}^{(0)}$$
- 2) Für $k = 0, 1, \dots, n-1$

$$\lambda_k = \frac{(\vec{r}^{(k)})^T \vec{d}^{(k)}}{(\vec{d}^{(k)})^T A \vec{d}^{(k)}}$$

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} + \lambda_k \vec{d}^{(k)}$$

$$\vec{r}^{(k+1)} = \vec{r}^{(k)} - \lambda_k A \vec{d}^{(k)}$$

$$\vec{d}^{(k+1)} = -\vec{r}^{(k+1)} - \frac{(\vec{r}^{(k+1)})^T A \vec{d}^{(k)}}{(\vec{d}^{(k)})^T A \vec{d}^{(k)}} \vec{d}^{(k)}$$

In Computerrechnungen:

$$\text{Stopp wenn } \|\vec{e}^{(k)}\|_A < \varepsilon$$

wobei ε eine vorgegebene Toleranz ist.

Fehler: $\vec{e}^{(i)} = \vec{x}^{(i)} - \vec{x}^*$ (\vec{x}^* ist genaue Lösung)

A symmetrisch und positiv definit

$\Rightarrow A$ hat reelle positive Eigenwerte.

$\lambda_{min}, \lambda_{max}$ bezeichnen den kleinsten bzw. größten EW.

Die Konditionszahl $cond_2(A) = \frac{\lambda_{max}}{\lambda_{min}}$

Satz:

$$\|\vec{e}^{(k)}\|_A \leq 2 \left(\frac{\sqrt{cond_2(A)-1}}{\sqrt{cond_2(A)+1}} \right)^k \|\vec{e}^{(0)}\|_A$$

bzw.

$$\|\vec{e}^{(k)}\|_A \leq 2 \left(\frac{\sqrt{\lambda_{max}} - \sqrt{\lambda_{min}}}{\sqrt{\lambda_{max}} + \sqrt{\lambda_{min}}} \right)^k \|\vec{e}^{(0)}\|_A$$

Siehe Beweis 29 (Anhang)

$cond_2(A)$ groß verursacht große Probleme. Deswegen...

Vorkonditionierung

Idee: statt $A\vec{x} = \vec{b}$

lösen wir $M^{-1}A\vec{x} = M^{-1}\vec{b}$

mit einer symmetrischen positiv definiten M , $M^{-1}A$ besser konditioniert.

$M^{-1}A$ ist i.A. nicht symmetrisch, deswegen:

$M = HH^T$, wobei H regulär ist., $M^{-1} = H^{-T}H^{-1}$ (z.B. Cholesky)

$M^{-1}A$ und $H^{-1}AH^{-T}$ haben die gleichen EW.

Siehe Beweis 30 (Anhang)

Beispiel 11: Jacobi-Vorkonditionierung

$$M = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$$

$$M^{-1}A = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

Beispiel 12: Cholesky-Vorkonditionierung

Unvollständige Cholesky-Faktorisierung:

Normal: $A = LL^T$

Dastatt nehmen wir $A \approx HH^T$

H ist wie L, wobei $l_{ij} = 0$, überall, wo $a_{ij} = 0$

3 Graphik

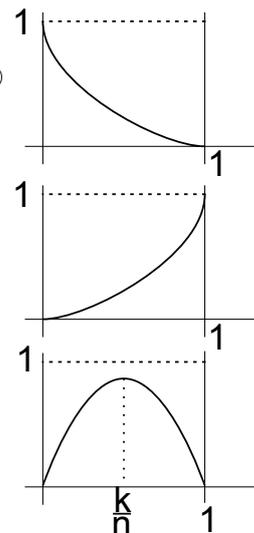
3.1 Bernsteinpolynom

Definition

$n \in \mathbb{N}$. Die Polynome

$$p_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}, \quad k = 0, 1, \dots, n$$

heißen *Bernstein-Polynome* bzw. *Bernstein-Grundpolynome* vom Grad n .



Verabredung

Wir verstehen $p_{n,k} \equiv 0$, wenn $k < 0$ oder $k > n$ ist!

Satz

Für $n \in \mathbb{N}, k = 0, 1, \dots, n$ gilt:

- 1) $p_{n,k}$ hat eine k -fache Nullstelle im Punkt $x = 0$.
- 2) $p_{n,k}$ hat eine $(n-k)$ -fache Nullstelle im Punkt $x = 1$.
- 3) $p_{n,k}$ ist positiv für $x \in (0, 1)$,
 streng monoton wachsend für $x \in [0, \frac{k}{n}]$,
 streng monoton fallend für $x \in [\frac{k}{n}, 1]$
 und hat somit das Maximum auf dem Intervall $[0, 1]$ im Punkt $x = \frac{k}{n}$.
- 4) $\{p_{n,k}\}_{k=0}^n$ sind linear unabhängig und bilden somit eine Basis des Raums \mathbb{P}_n der Polynome vom Grad $\leq n$.

Siehe Beweis 31 (Anhang)

Lemma

- 1) $\sum_{k=0}^n p_{n,k}(x) = 1$
- 2) $\sum_{k=0}^n \frac{k}{n} p_{n,k}(x) = x$
- 3) $\sum_{k=0}^n \left(\frac{k}{n}\right)^2 p_{n,k}(x) = \frac{n-1}{n} x^2 + \frac{1}{n} x$

Siehe Beweis 32 (Anhang)

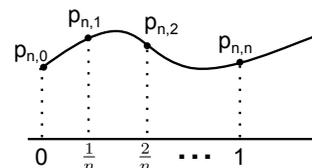
Definition

Der *Bernstein-Operator* der Ordnung n ist der Operator

$$B_n : C([0, 1]) \rightarrow \mathbb{P}$$

$$(B_n f)(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) p_{n,k}(x)$$

Ziel: einen konstruktiven Beweis des Approximationssatzes von Weierstraß zu geben.



Approximationssatz von Weierstraß

Sei $f \in C([0, 1])$. Dann gibt es zu jedem $\varepsilon > 0$ ein Polynom p derart, dass

$$\max_{x \in [0,1]} |f(x) - p(x)| < \varepsilon$$

$$\|g\| = \|g\|_{C([0,1])} = \max_{x \in [0,1]} |g(x)| \text{ die Norm auf } C([0, 1]).$$

Die Menge aller Polynome ist dicht im Raum $C([0, 1])$

Definition

Ein linearer Operator $L : C([0, 1]) \rightarrow C([0, 1])$ heißt *positiv*, falls

$f(x) \geq 0, x \in [0, 1] \Rightarrow (Lf)(x) \geq 0, x \in [0, 1]$ B_n ist ein linearer positiver Operator.

Satz von

Sei $\{L_n\}_{n=1}^{\infty}$ eine Folge der positiven linearen Operatoren,

$$L_n : C([0, 1]) \rightarrow C([0, 1]).$$

Seien $e_k(x) = x^k, k = 0, 1, 2, \dots$ die *Testfunktionen*

L_n positiv $\Rightarrow g < h \Rightarrow L_n g \leq L_n h$ Beweis: $h - g \geq 0, L_n(h - g) = L_n h - L_n g \geq 0$

Folgerung:

$$-|f| \leq f \leq |f|$$

$$L_n(-|f|) \leq L_n f \leq L_n(|f|)$$

$$\Rightarrow |L_n f| \leq L_n(|f|)$$

Gilt

$$\lim_{n \rightarrow \infty} \|(L_n e_k) - e_k\| = 0, k = 0, 1, 2, \dots$$

so gilt

$$\lim_{n \rightarrow \infty} \|(L_n f) - f\| = 0 \quad \forall f \in C([0, 1])$$

Siehe Beweis 33 (Anhang)

$$|(L_n f)(x) - f(x)| \leq \underbrace{L_n(|f(t) - f(x)|)}_{< 2\epsilon/3}(x) + \underbrace{M \|L_n e_0 - e_0\|}_{\leq \epsilon/3} \quad \underline{28.06.13}$$

$$\text{Setze } C = \max\left\{3M, \frac{24M}{\delta^2}\right\}$$

$$\exists N = N(\epsilon) \quad \forall n > N \quad \|L_n e_k - e_k\| < \frac{\epsilon}{C}, \quad k = 0, 1, 2, \dots$$

insbesondere

$$e_k - \frac{\epsilon}{C} < L_n e_k < e_k + \frac{\epsilon}{C}, \quad k = 0, 1, 2$$

Wir betrachten $|f(t) - f(x)|$, $t \in [0, 1]$ die Variable (x -fest)

$$\text{1. Fall: } |t - x| < \delta$$

$$\Rightarrow |f(t) - f(x)| < \frac{\epsilon}{6}$$

$$\text{2. Fall: } |t - x| \geq \delta$$

$$\frac{|t - x|}{\delta} \geq 1 \quad \frac{(t - x)^2}{\delta^2} \geq 1$$

$$|f(t) - f(x)| \leq 2M \leq 2M \frac{(t - x)^2}{\delta^2}$$

$$\text{Also, } |f(t) - f(x)| < \frac{\epsilon}{6} + \frac{2M}{\delta^2} (t - x)^2, \quad t \in [0, 1]$$

$$|f(t) - f(x)| < \frac{\epsilon}{6} e_0(t) + \frac{2M}{\delta^2} (t^2 - 2xt + x^2)$$

$$= \frac{\epsilon}{6} e_0(t) + \frac{2M}{\delta^2} (e_2(t) - 2xe_1(t) + x^2 e_0(t))$$

Aus der Positivität von L_n folgt, dass

$$L_n(|f(t) - f(x)|)(x) < L_n\left[\frac{\epsilon}{6} e_0(t) + \frac{2M}{\delta^2} (e_2(t) - 2xe_1(t) + x^2 e_0(t))\right](x)$$

Für den ersten Term:

$$1 - \frac{\epsilon}{C} < (L_n e_0)(x) < 1 + \frac{\epsilon}{C} < 2 \quad (\text{da } \epsilon \text{ nicht zu groß ist})$$

$$L_n(|f(t) - f(x)|)(x) < \frac{\epsilon}{3} + \frac{2M}{\delta^2} \left((L_n e_2)(x) - 2x(L_n e_1)(x) + x^2(L_n e_0)(x) \right) \quad -x$$

$$\left((L_n e_0)(x) < 1 + \frac{\epsilon}{C}; (L_n e_1)(x) < x - \frac{\epsilon}{C}; (L_n e_2)(x) < x^2 + \frac{\epsilon}{C} \right)$$

$$< \frac{\epsilon}{3} + \frac{2M}{\delta^2} \left(x^2 + \frac{\epsilon}{C} - 2x\left(x - \frac{\epsilon}{C}\right) + x^2\left(1 + \frac{\epsilon}{C}\right) \right)$$

$$= \frac{\epsilon}{3} + \frac{2M}{\delta^2} \left(x^2 + \frac{\epsilon}{C} - 2x^2 + 2x \frac{\epsilon}{C} + x^3 + x^2 \frac{\epsilon}{C} \right) = \frac{\epsilon}{3} + \frac{2M}{\delta^2} \frac{\epsilon}{C} \underbrace{(1 + 2x + x^2)}_{=(x+1)^2 \leq 24}$$

$$> \frac{\epsilon}{3} + \epsilon \cdot \frac{8M}{C\delta^2} \leq$$

für $x \in [0, 1]$

$$\text{parabolisch} \left(c \geq \frac{24M}{\delta^2} \Rightarrow \frac{1}{c} \leq \frac{\delta^2}{24M} \right)$$

$$\leq \frac{\epsilon}{3} + \epsilon \frac{8M}{\delta^2} + \frac{\delta^2}{24M} = \frac{\epsilon}{3} + \frac{\epsilon}{3} = \frac{2\epsilon}{3}$$

$$n > N \Rightarrow \|L_n e_0 - e_0\| \leq \frac{\epsilon}{3}$$

$$c \geq 3M \Rightarrow \frac{1}{c} \leq \frac{1}{3M}$$

Also haben wir für den zweiten Term die Abschätzung

$$M \|L_n e_0 - e_0\| \leq M \cdot \frac{\epsilon}{3M} = \frac{\epsilon}{3}$$

Ende des Beweises:

$$|(L_n f)(x) - f(x)| < \frac{2\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

Diese Abschätzung ist gleichmäßig in x

$$\Rightarrow \|L_n f - f\| < \epsilon$$

Wir haben bewiesen, dass es für jedes (kleines) $\epsilon > 0$ ein $N > 0$ gibt derart, dass

$$\forall n > N \quad \|L_n f - f\| < \epsilon$$

$$\text{Folglich,} \quad \lim_{n \rightarrow \infty} \|L_n f - f\| = 0$$

□

Satz: $\lim_{n \rightarrow \infty} \|B_n f - f\| = 0$ für alle $f \in C([0, 1])$

Zur Erinnerung: $B_n f(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) P_{n,k}(x)$ und

$$P_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

Beweis: Nach dem Satz von Korovkin genügt es, die Konvergenz für die drei Testfunktionen $e_0(x) = 1$, $e_1(x) = x$, $e_2(x) = x^2$

zu zeigen.

$$(B_n e_0)(x) = \sum_{k=0}^n P_{n,k}(x) = e_0(x)$$

(Reproduktion)

$$(B_n e_1)(x) = \sum_{k=0}^n \frac{k}{n} P_{n,k}(x) = x = e_1(x)$$

(Reproduktion)

Dann

Bem: B_n reproduziert lineare Funktionen.

$$(B_n e_2)(x) = \sum_{k=0}^n \left(\frac{k}{n}\right)^2 P_{nk}(x) = \frac{n-1}{n} x^2 + \frac{1}{n} x \rightarrow x^2 = e_2(x), n \rightarrow \infty$$

Nach dem Satz von Korovkin,

$$B_n f \rightarrow f, n \rightarrow \infty$$

für alle $f \in C([0,1])$. □

Somit haben wir auch den Approximationssatz von Weierstraß bewiesen.

$P_n = B_n f$ ist ein Polynom vom Grad n und $P_n \rightarrow f, n \rightarrow \infty$, gleichmäßig.

id

Bem Geometrische Eigenschaften von f bleiben bei der Approximation mittels Bernstein-Polynome erhalten.

Satz

1) Ist f monoton, so ist auch $B_n f$ monoton

2) Ist f konvex (konkav), so ist auch $B_n f$ konvex (konkav).

Lemma: $P'_{nk}(x) = n(P_{n-1,k-1}(x) - P_{n-1,k}(x))$

Bew:

$$P_{nk}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

$$P'_{nk}(x) = \frac{n!}{k!(n-k)!} k x^{k-1} (1-x)^{n-k} - \frac{n!}{k!(n-k)!} x^k (n-k) (1-x)^{n-k-1}$$

$$= \frac{n!}{(k-1)!(n-k)!} x^{k-1} (1-x)^{(n-1)-(k-1)} - \frac{n!}{k!(n-1-k)!} x^k (1-x)^{(n-1)-k}$$

$$= n \cdot \frac{(n-1)!}{(k-1)!(n-1-k)!} x^{k-1} (1-x)^{(n-1)-(k-1)} - n \cdot \frac{(n-1)!}{k!(n-1-k)!} x^k (1-x)^{(n-1)-k}$$

$$= n(P_{n-1,k-1}(x) - P_{n-1,k}(x))$$

□

Lemma: Sei $p(x) = \sum_{k=0}^n c_k P_{nk}(x)$.

Dann

$$p^{(\nu)}(x) = \frac{n!}{(n-\nu)!} \sum_{k=0}^{n-\nu} \Delta^\nu c_k P_{n-\nu,k}(x), \quad \nu = 0, 1, \dots, n$$

$\Delta^\nu c_k$ sind die vorwärts genommene Differenzen:

$$\Delta^0 c_k = c_k$$

$$\Delta^1 c_k = c_{k+1} - c_k$$

$$\Delta^\nu c_k = \Delta^{\nu-1} c_{k+1} - \Delta^{\nu-1} c_k, \quad \nu = 2$$

Bemerkung:

$$\text{Bem.: } p \in \mathbb{P}_n \Rightarrow p^{(\nu)}(x) \equiv 0, \quad \nu \geq n+1$$

Beweis:

$$\nu = 0 \quad p(x) = \sum_{k=0}^n c_k P_{n,k}(x) \quad \checkmark$$

$$\nu = 1 \quad p'(x) = \sum_{k=0}^n c_k P'_{n,k}(x)$$

$$= n \sum_{k=0}^n c_k (P_{n-1,k-1}(x) - P_{n-1,k}(x))$$

$$= n \left[\sum_{k=0}^n c_k P_{n-1,k-1}(x) - \sum_{k=0}^n c_k P_{n-1,k}(x) \right]$$

$$= n \left[\sum_{k=1}^n c_k P_{n-1,k-1}(x) - \sum_{k=0}^{n-1} c_k P_{n-1,k}(x) \right]$$

$P_{n-1,-1} \equiv 0$ $P_{n-1,n} \equiv 0$

$k_{\text{neu}} = k-1$

$$= n \left[\sum_{k=0}^{n-1} c_{k+1} P_{n-1,k}(x) - \sum_{k=0}^{n-1} c_k P_{n-1,k}(x) \right]$$

$$= n \sum_{k=0}^{n-1} (c_{k+1} - c_k) P_{n-1,k}(x)$$

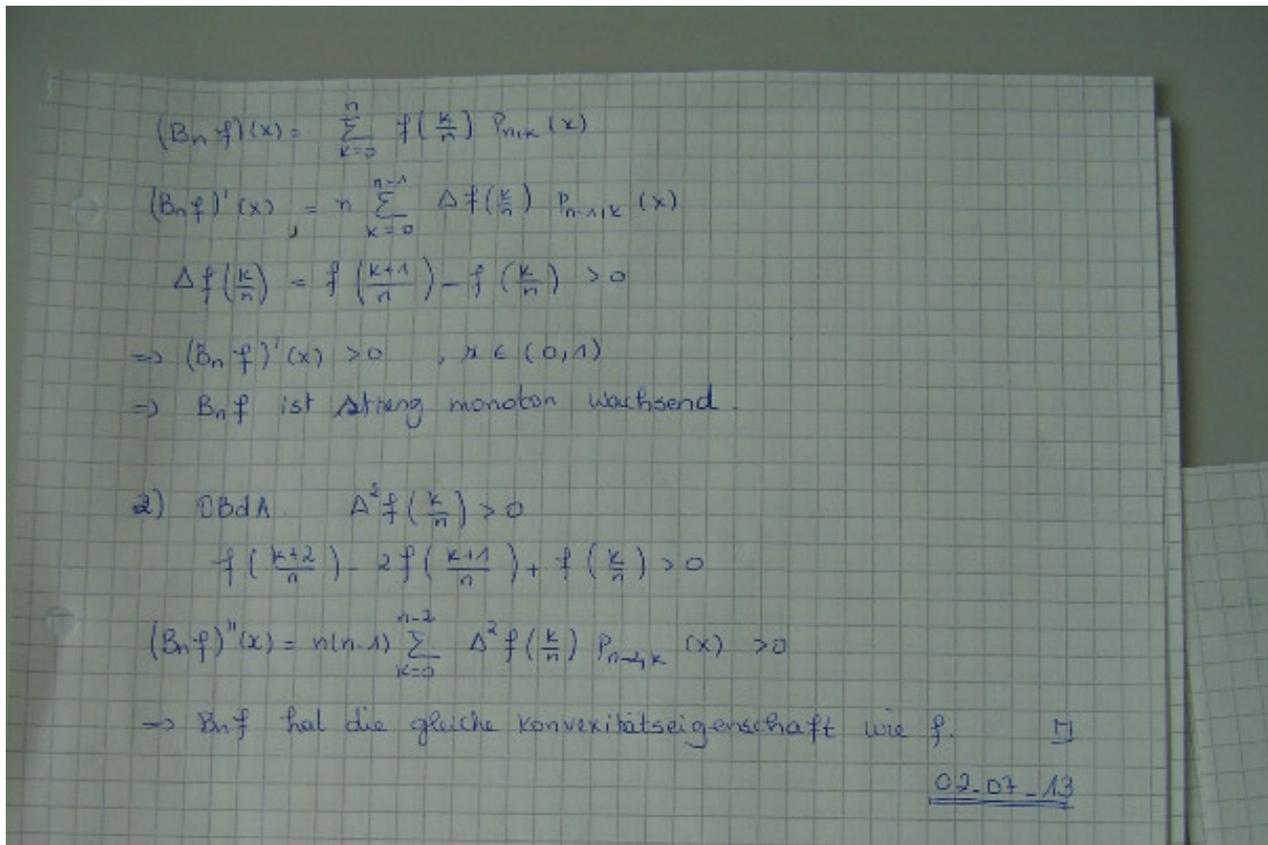
$$= n \sum_{k=0}^{n-1} \Delta c_k \cdot P_{n-1,k}(x) \quad \checkmark$$

$\nu \geq 2 \rightarrow$ Beweis per Induktion

Beweis des Satzes über die Erhaltung geometrischer Eigenschaften

1) OBD A ist f string monoton wachsend

Zu zeigen: Bnf ist string monoton wachsend



13-07-02

3.2 Bézier-Kurven

Definition

Gegeben seien $\beta_0, \beta_1, \dots, \beta_n \in \mathbb{R}^d$.

Die parametrische Kurve

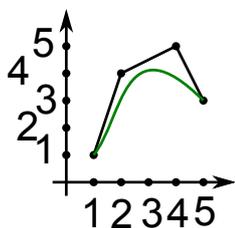
$$p(t) = \sum_{k=0}^n \beta_k p_{n,k}(t), \quad t \in [0, 1]$$

heißt eine **Kurve in Bézier-Darstellung** oder **Bézier-Kurve**.

Die Punkte $\beta_0, \beta_1, \dots, \beta_n$ heißen **Bézier-Punkte** oder **Kontrollpunkte**.

Das Polygonzug durch die Punkte $\beta_0, \beta_1, \dots, \beta_n$ heißt **Bézie-Polygon** oder **Kontrollpolygon**.

Beispiel 1



Die Kontrollpunkte $\binom{1}{1}$, $\binom{2}{4}$, $\binom{4}{5}$, $\binom{5}{3}$

$$p(t) = \binom{1}{1} p_{3,0}(t) + \binom{2}{4} p_{3,1}(t) + \binom{4}{5} p_{3,2}(t) + \binom{5}{3} p_{3,3}(t)$$

$$= \binom{1}{1} (1-t)^3 + \binom{2}{4} 3t(1-t)^2 + \binom{4}{5} 3t^2(1-t) + \binom{5}{3} t^3, \quad t \in [0, 1]$$

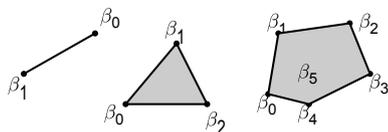
$$p(0) = \binom{1}{1}, \quad p(1) = \binom{5}{3}$$

Definition

Konvexe Hülle der Punkte $\beta_0, \beta_1, \dots, \beta_n \in \mathbb{R}^d$
 $\text{conv}(\beta_0, \beta_1, \dots, \beta_n) = \{x \in \mathbb{R}^d : x = \sum_{j=0}^n \lambda_j \beta_j, \text{ mit } \lambda_j \geq 0, \sum_{j=0}^n \lambda_j = 1\}$

Siehe Beweis 34 (Anhang)

Beispiel 2



$$\begin{aligned} x &\in \text{conv}\{\beta_0, \beta_1\} \\ &\Rightarrow x \in \lambda_0 \beta_0 + \lambda_1 \beta_1 \\ &\text{mit } \lambda_0, \lambda_1 \geq 0 \text{ m } \lambda_0 + \lambda_1 = 1 \\ x &= \lambda_0 \beta_0 + (1 - \lambda_0) \beta_1, \quad \lambda_0 \in [0, 1] \end{aligned}$$

Eigenschaft

- 1) $p(0) = \beta_0, p(1) = \beta_n$.
(Die Bézier-Kurve interpoliert die Randpunkte)
Siehe Beweis 35 (Anhang)
- 2) $p'(0) = n(\beta_1 - \beta_0), p'(1) = n(\beta_n - \beta_{n-1})$
Siehe Beweis 36 (Anhang)

Bemerkung

$p(t)$ folgt der Geometrie des Kontrollpunktes. (Auch die Monotonie, die Konvexität usw. bleiben erhalten).

$p \in \mathbb{P}_n$

Monom-Darstellung: $p(t) = \sum_{\nu=0}^n \frac{p^{(\nu)}(0)}{\nu!} t^\nu$

Bézier-Darstellung: $p(t) = \sum_{k=0}^n \beta_k p_{n,k}(t)$

Aus der Bézier-Darstellung zur Monom-Darstellung.

$$\frac{p^{(\nu)}(0)}{\nu!} = \frac{n!}{(n-\nu)! \nu!} \sum_{k=0}^{n-\nu} \Delta^\nu \beta_k p_{n-\nu,k}(0) = \binom{n}{\nu} \Delta^\nu \beta_0$$

\Rightarrow

$$p(t) = \sum_{\nu=0}^n \binom{n}{\nu} \Delta^\nu \beta_0 t^\nu$$

Beispiel 3

$$\begin{aligned} p(0) &= \Delta^0 \beta_0 = \beta_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ p'(0) &= 3 \Delta^1 \beta_0 = 3(\beta_1 - \beta_0) = 3 \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \end{pmatrix} \\ \frac{p''(0)}{2!} &= \binom{3}{2} \Delta^2 \beta_0 = 3(\beta_2 - 2\beta_1 + \beta_0) = 3 \left(\begin{pmatrix} 4 \\ 5 \end{pmatrix} - 2 \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 3 \\ -6 \end{pmatrix} \\ \frac{p^{(3)}(0)}{3!} &= \Delta^3 \beta_0 = \beta_3 - 3\beta_2 + 3\beta_1 - \beta_0 = \begin{pmatrix} 5 \\ 3 \end{pmatrix} - 3 \begin{pmatrix} 4 \\ 5 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix} \\ p(t) &= \begin{pmatrix} -2 \\ -1 \end{pmatrix} t^3 + \begin{pmatrix} 3 \\ -6 \end{pmatrix} t^2 + \begin{pmatrix} 3 \\ 9 \end{pmatrix} t + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

Aus der Monom-Darstellung zur Bézier-Darstellung:

$$\binom{n}{\nu} \Delta^\nu \beta_0 = \frac{p^{(\nu)}(0)}{\nu!}$$

$$\beta_0 = p(0)$$

$$n(\beta_1 - \beta_0) = p'(0) \Rightarrow \beta_1 = \beta_0 + \frac{p'(0)}{n} \text{ usw.}$$

Beispiel 4

$$\beta_0 = p(0) = \binom{1}{1}$$

$$3(\beta_1 - \beta_0) = p'(0) = \binom{3}{9}$$

$$\beta_1 = \beta_0 + \frac{1}{3} \binom{3}{9} = \binom{1}{1} + \binom{1}{3} = \binom{2}{4}$$

$$3(\beta_2 - 2\beta_1 + \beta_0) = \binom{3}{-6}$$

$$\beta_2 = 2\beta_1 - \beta_0 + \frac{1}{3} \binom{3}{-6} = 2 \binom{2}{4} - \binom{1}{1} + \binom{1}{-2} = \binom{4}{5}$$

$$\beta_3 - 3\beta_2 + 3\beta_1 - \beta_0 = \binom{-2}{-1}$$

$$\beta_3 = 3\beta_2 - 3\beta_1 + \beta_0 + \binom{-2}{-1} = 3 \binom{4}{5} - 3 \binom{2}{4} + \binom{1}{1} + \binom{-2}{-1} = \binom{5}{3}$$

3.3 Auswertung eines polynoms in dwer Bézier-Darstellung

Lemma

$$p_{n,0}(t) = (1-t)p_{n-1,0}(t)$$

$$p_{n,k}(t) = (1-t)p_{n-1,k}(t) + tp_{n-1,k-1}(t), \quad k = 1, \dots, n-1$$

$$p_{n,n}(t) = tp_{n-1,n-1}(t)$$

Auswertungsschema:

$t \in [0, 1]$ fest.

$$\begin{aligned} p(t) &= \sum_{k=0}^n \beta_k p_{n,k}(t) \\ &= \sum_{k=0}^n \beta_k [(1-t)p_{n-1,k}(t) + tp_{n-1,k-1}(t)] \\ &= \sum_{k=0}^{n-1} \beta_k (1-t)p_{n-1,k}(t) + \sum_{k=1}^n \beta_k t p_{n-1,k-1}(t) \\ &= \sum_{k=0}^{n-1} \beta_k (1-t)p_{n-1,k}(t) + \sum_{k=0}^{n-1} \beta_{k+1} t p_{n-1,k}(t) \\ &= \sum_{k=0}^{n-1} [\beta_k (1-t) + \beta_{k+1} t] p_{n-1,k}(t) \\ &= \sum_{k=0}^{n-1} \beta_k^{(1)} p_{n-1,k}(t), \quad \text{mit } \beta_k^{(1)} = \beta_k (1-t) + \beta_{k+1} t \end{aligned}$$

Rekursiv:

$$p(t) = \sum_{k=0}^{n-2} \beta_k^{(2)} p_{n-2,k}(t)$$

$$\text{mit } \beta_k^{(2)} = \beta_k^{(1)} (1-t) + \beta_{k+1}^{(1)} t$$

usw.

$$p(t) = \beta_0^{(n)}$$

Satz: *Algorithmus von de Casteljau*

Ein Polynom p in der Bézier-Darstellung $p(t) = \sum_{k=0}^n \beta_k p_{n,k}(t)$

lässt sich an einer Stelle t folgendermaßen auswerten:

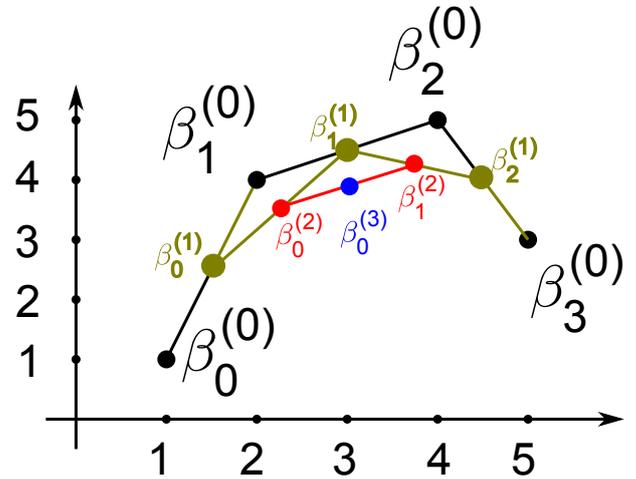
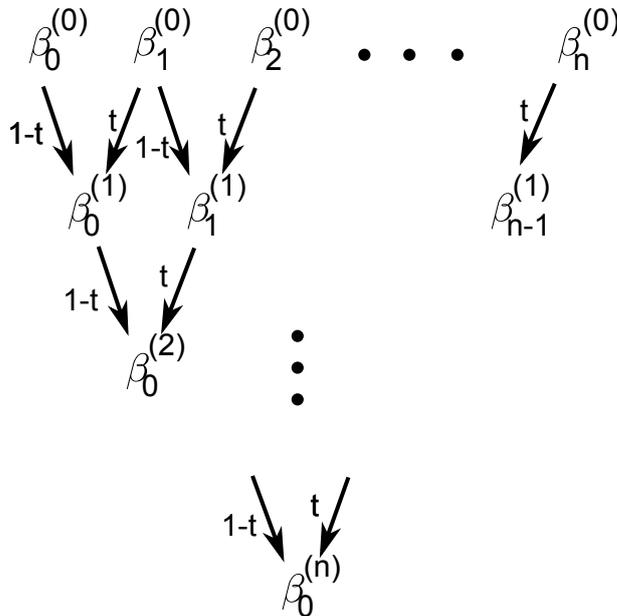
$$\beta_k^{(0)} = \beta_k, \quad k = 0, 1, \dots, n$$

Für $m = 1, 2, \dots, n$:

Für $k = 0, 1, \dots, n - m$

$$\beta_k^{(m)} = (1 - t)\beta_k^{(m-1)} + t\beta_{k+1}^{(m-1)}$$

Dann gilt: $p(t) = \beta_0^{(n)}$



Wendet man das Schema an, erhält man im Beispiel $\binom{3}{3.875}$

Auswertung der Polynome in der Monom-Darstellung

$$p(t) = a_0t^n + a_1t^{n-1} + \dots + a_n$$

Die Anzahl der Operationen:

„x“: $n + (n - 1) + \dots + 0 = \frac{n(n+1)}{2}$

„+“: n

$$\begin{aligned} p(t) &= a_0t^n + a_1t^{n-1} + a_2t^{n-2} + \dots + a_n \\ &= (a_0t + a_1)t^{n-1} + a_2t^{n-2} + \dots + a_n \\ &= ((a_0t + a_1)t + a_2)t^{n-2} + \dots + a_n \\ &= (\dots(((a_0t + a_1)t + a_2)t + a_3)\dots) + a_n \end{aligned}$$

Die Anzahl der Operationen:

„x“: n

„+“: n

Das Horner-Schema

$$b_0 = a_0$$

$$b_1 = a_0t + a_1$$

$$b_2 = b_1t + a_2$$

⋮

$$b_n = b_{n-1}t + a_n$$

Dann gilt $b_n = p(t)$

a_0	a_1	a_2	\dots	a_n
t_0	$-$	b_0t_0	b_1t_0	\dots
	b_0	b_1	$b_2 \dots$	b_n

$$p(t) = (t - t_0)(b_0t^{n-1} + b_1t^{n-2} + \dots + b_{n-1}) + b_n = \underbrace{b_0}_{a_1} t^n + \underbrace{(b_1 - b_0t_0)}_{a_1} t^{n-1} + \dots + \underbrace{(b_n - b_{n-1}t_0)}_{a_n}$$

$$p(t) = (t - t_0)q(t) + b_n \quad \text{mit } b_n = p(t_0)$$

Wiederholte Anwendung des Horner-Schemas: $q(t) = (t - b_0)r(t) + c_{n-1}$, $c_{n-1} = q(t_0)$

$$p(t) = (t - t_0)q(t) + b_n$$

$$p'(t) = q(t) + (t - t_0)q'(t)$$

$$p'(t_0) = q(t_0)$$

Anwendung: *Newton-Horner-Verfahren* zur Bestimmung von Nullstellen von Polynomen.

$$X_n = X_{n-1} - \frac{p(X_{n-1})}{p'(X_{n-1})}$$

$$p'(t) = q(t) + (t - t_0)q'(t)$$

$$p''(t) = q'(t) + q'(t) + (t - t_0)q''(t)$$

$$p''(t_0) = 2q'(t_0)$$

$$\text{Aber : } q(t) = (t - t_0)r(t) + q(t_0)$$

$$q'(t_0) = 2r(t_0)$$

$$\Rightarrow r(t_0) = \frac{1}{2}p''(t_0)$$

Im k -ten Schritt des *vollständigen Horner-Schemas* wertet man $\frac{p^{(k)}(t_0)}{k!}$ aus!

$$\Rightarrow p(t) = \sum_{k=0}^n \frac{p^{(k)}(t_0)}{k!} (t - t_0)^k$$

Beispiel 5

13-07-05

$$p(t) = 2t^6 - t^5 + t^3 + 2t^2 - 2t + 4$$

$$p(0.5) = ?$$

	2	-1	0	1	2	-2	4
$t_0 = 0.5$	-	1	0	0	0.5	1.25	-0.375
	2	0	0	1	2.5	-0.75	3.625

$t_0 = 0.5$	-	1	0.5	0.25	0.625	1.5625
	2	1	0.5	1.25	3.125	0.8125

$t_0 = 0.5$	-	1	1	0.75	1
	2	2	1.5	2	4.125

$t_0 = 0.5$	-	1	1.5	1.5
	2	3	3	3.5

$t_0 = 0.5$	-	1	2
	2	4	5

$t_0 = 0.5$	-	1
	2	5

$t_0 = 0.5$	-
	2

$$\Rightarrow \phi(0.5) = 3.625$$

$$\Rightarrow p'(0.5) = 0.8125$$

$$p(t) = (t - 0.5)(2t^5 + t^2 + 2.5t - 0.75) + 3.625$$

$$p(t) = 2(t - 0.5)^6 + 5(t - 0.5)^5 + 5(t - 0.5)^4 + 3.5(t - 0.5)^3 + 4.125(t - 0.5)^2 + 0.8125(t - 0.5) + 3.625$$

3.4 Bézier-Kurven Gradanhebung

$\{p_{n,k}\}_{k=0}^n$ ist eine Basis der Polynome vom Grad $\leq n$

Fragestellung: Wie kann man Polynome vom Grad $< n$ in der Basis des Raums \mathbb{P}_n darstellen?

In der Monom-Basis:

$$p(t) \in \mathbb{P}_{n-1}$$

$$p(t) = a_{n-1}t^{n-1} + a_{n-2}t^{n-2} + \dots + a_0$$

$$\Rightarrow p(t) = 0t^n + a_{n-1}t^{n-1} + \dots + a_0$$

In der Basis der Bernstein-Grundpolynome: man muss die Kontrollpunkte umrechnen.

Lemma

$$p_{n,k}(t) = \frac{n+1-k}{n+1}p_{n+1,k}(t) + \frac{k+1}{n+1}p_{n+1,k+1}(t)$$

Siehe Beweis 37 (Anhang)

Satz

Gegeben seien $n+1$ Kontrollpunkte $\beta_0, \beta_1, \dots, \beta_n$. Definiert man

$$\hat{\beta}_0 = \beta_0$$

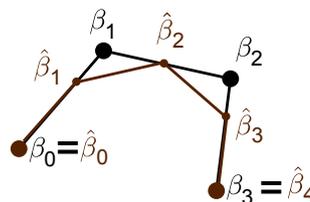
$$\hat{\beta}_{n+1} = \beta_n$$

$$\text{hat } \beta_k = \frac{k}{n+1}\beta_{k-1} + \frac{k+1}{n+1}\beta_k, \quad k = 1, \dots, n$$

So gilt:

$$\sum_{k=0}^n \beta_k p_{n,k}(t) = \sum_{k=0}^{n+1} \hat{\beta}_k p_{n+1,k}(t)$$

Siehe Beweis 38 (Anhang)



Lemma

$$p_{n,k}(t) = \frac{n+1-k}{n+1}p_{n+1,k}(t) + \frac{k+1}{n+1}p_{n+1,k+1}(t)$$

3.5 Bézier-Splines

Satz

Eine zusammengesetzte Bézier-Kurve

$$s(t) = \begin{cases} \sum_{k=0}^n \beta_{k,0} p_{n,k}(t) & , t \in [0, 1) \\ \sum_{k=0}^n \beta_{k,1} p_{n,k}(t-1) & , t \in [1, 2) \end{cases}$$

ist genau dann r -mal stetig differenzierbar, wenn

$$\Delta^m \beta_{n-m,0} = \Delta^m \beta_{0,1}, \quad m = 0, 1, \dots, r$$

Siehe Beweis 39 (Anhang)

Beispiel 6

$$s(t) = \begin{cases} p_{3,3}(t) & , t \in [0, 1) \\ p_{3,0}(t-1) + 2p_{3,1}(t-1) + 4p_{3,2}(t-1) + 4p_{3,3}(t-1) & , t \in [1, 2) \\ 4p_{3,0}(t-2) + 4p_{3,1}(t-2) + 2p_{3,2}(t-2) + p_{3,3}(t-2) & , t \in [2, 3) \\ 0 & , \text{sonst} \end{cases}$$

$$s(0+) = p_{3,3}(t)|_{t=0} = t^3|_{t=0}$$

$$s(0+) = s'(0+) = s''(0+) = 0$$

aber $s^{(3)}(0+) = 6 \neq 0 \Rightarrow s$ ist höchstens 2-mal stetig differenzierbar.

$$\begin{array}{l|l} \beta_0 & \delta\beta_0 = \beta_1 - \beta_0 \\ \beta_1 & \delta\beta_1 = \beta_2 - \beta_1 \quad \Delta^2\beta_0 \\ \beta_2 & \delta\beta_2 = \beta_3 - \beta_2 \quad \Delta^2\beta_1 \\ \beta_3 & \end{array} \quad \Delta^3\beta_0$$

$t = 0$	0		
	0		
	0	0	$\Rightarrow 2$ mal stetig differenzierbar.
	0	0	
	0	1	
	1	1	
$t = 1$	1		
	1		
	2	1	$\Rightarrow 2$ mal stetig differenzierbar.
	2		
	4	-2	
	0		
$t = 2$	4		
	4		
	0		
	4	-2	
	-2		
	2	-1	
	-1		
	1		

3.6 Spline-Funktionen

13-07-09

Spezielle Spline Funktionen sind uns bereits begegnet B_{m-1} mit

$$B_{m-1}(t) = \sum_{\nu=0}^m \omega_{\nu}(x_{\nu} - t)_{+}^{m-1}, \quad \omega_{\nu} := \prod_{\substack{k=0 \\ k \neq \nu}}^m \frac{1}{x_{\nu} - x_k}$$

Sie bestehen stückweise aus Polynomen vom Grad $m-1$ und sind insgesamt $(m-2)$ -mal stetig differenzierbar.

Wir beschränken uns im Folgenden auf äquidistante Knoten.

Definition

eine Funktion $S : \mathbb{R} \rightarrow \mathbb{R}$ erfülle die folgenden Bedingungen:

1. S stimmt auf jedem Intervall $[\nu, \nu + 1)$, $\nu \in \mathbb{Z}$ mit einem Polynom $p_{\nu} \in \Pi_m$, $m \in \mathbb{N}$ überein.
2. Es gilt: $S \in C^{m-1}(\mathbb{R})$

Dann heißt S *polynomialer Spline* vom Grad m mit dem Knoten $\nu \in \mathbb{Z}$.

Definition

Ist $\tilde{S} : [k, l) \rightarrow \mathbb{R}$ die Restriktion eines polynomiellen Spline S vom Grad m auf das Intervall $[k, l)$ mit $k, l \in \mathbb{Z}$, $k < l$, so heißt \tilde{S} *polynomieller Spline* vom Grad m auf dem Intervall $[k, l)$ mit Knoten ν , $\nu = k, \dots, l$

Bemerkung

Wichtige Fälle in der Anwendung:

$m=1$ Polynomzug

$m=2$ quadratisch

$m=3$ kubisch

Bemerkung

1. Für einen polynomiellen Spline S vom Grad m mit Knoten $\nu \in \mathbb{Z}$ können wir die Darstellung

$$S(t) = \sum_{\mu=0}^m \beta_{\nu,\mu} p_{m,\mu}(t - \nu) \quad , \quad t \in [\nu, \nu + 1) \quad , \quad \nu \in \mathbb{Z}$$

2. Für einen polynom. Spline S auf $[k, l)$ benutzen wir die entsprechende Darstellung

$$S(t) = \sum_{\mu=0}^m \beta_{\nu,\mu} p_{m,\mu}(t - \nu) \quad , \quad t \in [\nu, \nu + 1) \quad , \quad \nu \in \{k, k + 1, \dots, l - 1\}$$

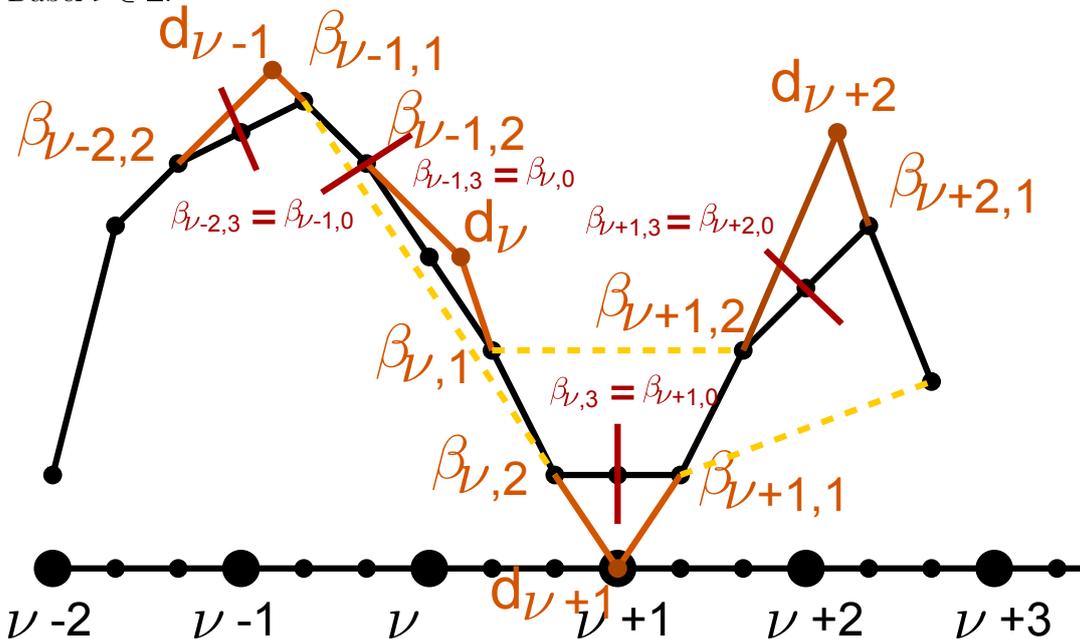
Für $m \geq 1$ ist wegen der Definition sogar $t \in [\nu, \nu + 1]$ und obige Darstellung gilt sogar in ganz $[k, l]$

Untersuche nun den besonderen wichtigen Fall kubischer Splines:

Für einen kubischen Spline genügt in jedem Knoten den Folgenden drei Bedingungen

- $\beta_{\nu-1,3} = \beta_{\nu,0}$
- $\beta_{\nu-1,3} - \beta_{\nu-1,2} = \beta_{\nu,1} - \beta_{\nu,0}$
- $2\beta_{\nu-1,2} - \beta_{\nu-1,1} = 2\beta_{\nu,1} - \beta_{\nu,2}$

Dabei $\nu \in \mathbb{Z}$.



Bemerkung

Zweimalige stetige Diffbarkeit in der Anschlussstelle bedeutet (außer dass Wert und Richtung an dieser Stelle übereinstimmen), dass die Hilfspunkte, die durch Extrapolation an den mittleren Bézier-Punkten entstehen, übereinstimmen.

definieren eine Hilfsgröße $d_\nu := 2\beta_{\nu,1} - \beta_{\nu,2}$ $\nu \in \mathbb{Z}$

Dann ist Bed (3) gleichwertig mit $d_\nu = 2\beta_{\nu-1,2} - \beta_{\nu-1,1}$, $\nu \in \mathbb{Z}$

Betrachte wegen $S(\nu) = \beta_{\nu,0}$, $S''(\nu) = m(m-1)\Delta^2\beta_{\nu,0}$ für kritische Splines auch

$$d_\nu = S(\nu) - \frac{S''(\nu)}{6}, \quad \nu \in \mathbb{Z}$$

gilt.

Bemerkung

Geg. ein kubischer Spline. Man kann leicht zeigen, dass

1. Die Geraden durch die Punkte $\begin{pmatrix} \nu-\frac{1}{3} \\ \beta_{\nu-1,2} \end{pmatrix}$ und $\begin{pmatrix} \nu-\frac{1}{3} \\ \beta_{\nu,1} \end{pmatrix}$ und durch die Punkte $\begin{pmatrix} \nu-\frac{2}{3} \\ \beta_{\nu-1,1} \end{pmatrix}$ und $\begin{pmatrix} \nu+\frac{2}{3} \\ \beta_{\nu,2} \end{pmatrix}$ sind parallel.

2. Die Punkte $\begin{pmatrix} \nu+\frac{1}{3} \\ \beta_{\nu,1} \end{pmatrix}$ und $\begin{pmatrix} \nu+\frac{2}{3} \\ \beta_{\nu,2} \end{pmatrix}$ liegen auf der Verbindungsgeraden von $\begin{pmatrix} \nu \\ d_\nu \end{pmatrix}$ und $\begin{pmatrix} \nu-1 \\ d_{\nu-1} \end{pmatrix}$

Betrachte den kubischen Spline zu den Knoten $\nu = 0, \dots, k$.

Definiere d_ν für $\nu = 0, \dots, k-1$ wie oben und ersetze

$$d_k := 2\beta_{k-1,2} - \beta_{k-1,1}$$

$$\beta_{k,0} := \beta_{k-1,3}$$

Damit gilt wegen $S''(k) = 6\Delta^2\beta_{k-1,1}$ auch hier

$$d_\nu = \beta_{\nu,0} - \frac{S''(\nu)}{6}, \quad \nu = 0, \dots, k$$

Dem Bézier-Polynom ist ein Polygonzug überlagert, dessen Ebenen durch d_ν bestimmt sind. Dabei stimmen das Bézier-Polynom und der d_ν -Polynomzug auf den inneren Intervallen $[\nu + \frac{1}{3}, \nu + \frac{2}{3}]$, $\nu = 0, \dots, k-1$ überein.

Bemerkung

Die Bézier-Koeffizienten $\beta_{\nu,\mu}$, $\mu = 0, \dots, 3$, $\nu = 0, \dots, k-1$ eines kubischen Splines

$$S(t) = \sum_{\mu=0}^3 \beta_{\nu,\mu} p_{3,\mu}(t-\nu), \quad t \in [\nu, \nu+1], \quad \nu = 0, \dots, k-1$$

sind durch die Anschlussbedingungen

$$\begin{array}{cccc} & \beta_{\nu-1,3} & & -\beta_{\nu,0} = 0 \\ -\beta_{\nu-1,2} & +\beta_{\nu-1,3} & -\beta_{\nu,1} & +\beta_{\nu,0} = 0 \\ -\beta_{\nu-1,1} & +2\beta_{\nu-1,2} & -2\beta_{\nu,1} & +\beta_{\nu,2} = 0 \end{array}, \quad \nu = 1, \dots, k-1 \text{ miteinander verknüpft.}$$

Dies sind die $3(k+1)$ homogenen linearen Gleichungen für die $4k$ Koeffizienten $\beta_{\nu,\mu}$ für $\mu = 0, 1, 2, 3$ und $\nu = 0, \dots, k-1$. Dieses homog. lin. Glnsystem hat einen Lösungsraum V mit

$$\dim V \geq 4k - 3(k-1) = k+3$$

Da $\beta_{0,0}$ und $\beta_{k-1,3} = \beta_{k,0}$ in den Anschlussbedingungen nicht vorkommen, können wir die Elemente von V mit Hilfe von $\beta_{0,0}$ und $\beta_{k,0}$ parametrisieren.

Wir zeigen nun, dass die Bézier-Koeffizienten $\beta_{0,1} > \beta_{\nu,0}$ für $\nu = 0, \dots, k$ und $\beta_{k-1,2}$ eine Parametrisierung der Elemente von V ergeben.

Satz

Zu jedem Wert

$$\beta := (\beta_{0,0}, \beta_{0,1}, \beta_{0,2}, \beta_{1,0}, \dots, \beta_{k-1,2}, \beta_{k,0})^T \in \mathbb{R}^{n+3}$$

gibt es genau einen kubischen Spline S auf $[0, b)$ mit

$$S(t) = \sum_{\mu=0}^3 \beta_{\mu,\nu} p_{3,\mu}(t-\nu), \quad t \in [\nu, \nu+1], \quad \nu = 0, \dots, k-1$$

Von den $4k$ Polynomen $\beta_{\nu,\mu}$ sind also durch die Anschlussbedingungen $4k - (k + 3) = 3k - 3$ Parameter festgelegt.

Siehe Beweis 40 (Anhang)

Lemma

$$1. \quad \sum_{\nu=-1}^{k+1} \beta_{\nu,3}(t) = 1$$

$$2. \quad \sum_{\nu=-1}^{k+1} \nu \beta_{\nu,3}(t) = t$$

$$3. \quad \sum_{\nu=-1}^{k+1} (\nu^2 - \frac{1}{3}) \beta_{\nu,3}(t) = t^2$$

Siehe Beweis 41 (Anhang)

$$\text{Also: } S(t) = \sum_{\nu=-1}^{k+1} p_{\nu} \beta_{\nu,3}(t), \quad t \in [0, k]$$

$p_{-1}, p_0, p_1, \dots, p_{k+1} \in \mathbb{R}^2$ für eine parametr. Kurve die Kontrollpunkte (*de Boor-Punkte*)
 \Rightarrow ganze Theorie für Bézier-Kurve

Beweise

1 Beweis (Abschnitt 1.2.2):

$$\begin{aligned}\tau(x, y(x), h) &= \frac{y(x, y), y(x)}{h} - \underbrace{f(x, y(x))}_{y'(x)} \\ y(x+h) &= y(x) + hy'(x) + \frac{h^2}{2}y''(x) + O(h^3) \\ \Rightarrow \tau(x, y(x), h) &= \frac{1}{h}(hy'(x) + \frac{h^2}{2}y''(x) + O(h^3)) - y'(x) \\ &= \frac{h}{2}y''(x) + O(h^2) = O(h)\end{aligned}$$

2 Beweis (Abschnitt 1.2.2):

$$\begin{aligned}|\xi_n| &\leq (1+\delta)|\xi_{n-1}| + \beta \leq (1+\delta)^2|\xi_{n-2}| + (1+(1+\delta))\beta \\ &\leq (1+\delta)^3|\xi_{n-3}| + (1+(1+\delta) + (1+\delta)^2)\beta \\ &\leq (1+\delta)^n|x_{i_0}| + (1+(1+\delta) + (1+\delta)^2 + \dots + (1+\delta)^{n-1})\beta \\ &= (1+\delta)^n|\xi_0| + \frac{(1+\delta)^n - 1}{\delta}\beta \text{ (Summenformel)} \\ &\leq e^{n\delta}|\xi_0| + \frac{e^{n\delta} - 1}{\delta}\beta\end{aligned}$$

Da $0 < 1 + \delta < e^\delta$, $\delta > 0$

3 Beweis (Abschnitt 1.2.2):

$$\begin{aligned}y_{k+1, h} &= y_{k, h} + h\varphi(x_k, y_k, h) \\ \tau(x_k, y(x_k), h) &= \frac{y(x_{k+1}) - y(x_k)}{h} - \varphi(x_k, y(x_k), h) \\ y(x_{k+1}) &= y(x_k) + h\varphi(x_k, y(x_k), h) + h\tau(x_k, y(x_k), h) \\ e_{k+1, h} &= e_{k, h} + h(\varphi(x_k, y_k, h) - \varphi(x_k, y(x_k), h)) - h\tau(x_k, y(x_k), h) \\ \text{Lstigkeit von } \varphi &: |\varphi(x_k, y_k, h) - \varphi(x_k, y(x_k), h)| \leq M|y_k - y(x_k)| = M|e_{k, h}| \\ \text{Konsistenz} &\Rightarrow |\tau(x_k, y(x_k), h)| = O(h^p) \Rightarrow |\tau(x_k, y(x_k), h)| \leq Nh^p, \quad 0 < h \leq h_0 \\ |e_{k+1, h}| &\leq |e_{k, h}| + hM|e_{k, h}| + Nh^{p+1} = (1+hM)|e_{k, h}| + Nh^{p+1}\end{aligned}$$

Nach dem Lemma:

$$\begin{aligned}|e_{n, h}| &\leq e^{nhM} \underbrace{|e_{n, h}|}_{=0} + \frac{e^{nhM} - 1}{hM}Nh^{p+1} \\ &\leq \frac{e^{M(b-a)} - 1}{M}Nh^p \\ \max_{n: x_n \in [a, b]} |e_{n, h}| &\leq \frac{e^{M(b-a)} - 1}{M}Nh^p\end{aligned}$$

\Rightarrow Konvergenz ist von der Ordnung p

4 Beweis (Abschnitt 1.2.10):

Sei $p \in \mathbb{P}_{s-1}$ das Interpolationspolynom zu den Werten $u'(x_k + c_j h)$, $j = 1, \dots, s$
d.h. $p(x_k + c_j h) = u'(x_k + c_j h)$, $j = 1, \dots, s$

$$p(x) = \sum_{i=1}^s u'(x_k + c_i h) l_i\left(\frac{x-x_k}{h}\right)$$

In der Tat:

$$\begin{aligned}p(x_k + c_j h) &= \sum_{i=1}^s u'(x_k + c_i h) l_i\left(\frac{x_k + c_j h - x_k}{h}\right) \\ &= \sum_{i=1}^s u'(x_k + c_i h) l_i(c_j) \\ &= u'(x_k + c_j h)\end{aligned}$$

$p, u' \in \mathbb{P}_{s-1}$ und $p(x_k + c_j h) = u'(x_k + c_j h)$ für $j = 1, \dots, s$

$\Rightarrow d = p - u' \in \mathbb{P}_{s-1}$, $d(x_k + c_j h) = 0$, $j = 1, \dots, s$

$\Rightarrow d \equiv 0 \Rightarrow p \equiv u'$

$$\text{Also } u'(x) = \sum_{i=1}^s \underbrace{u'(x_k + c_i h)}_{f(x_k + c_i h, u(x_k + c_i h))} l_i\left(\frac{x-x_k}{h}\right)$$

Def. Kollokationsverfahren

$$u'(x) = \sum_{i=1}^s f(x_k + c_i h, u(x_k + c_i h)) l_i\left(\frac{x-x_k}{h}\right)$$

$$\begin{aligned} u(x) &= u(x_k) + \int_{x_k}^x u'(t) dt \\ &= y_k + \int_{x_k}^x \sum_{i=1}^s f(x_k + c_i h, u(x_k + c_i h)) l_i\left(\frac{t-x_k}{h}\right) dt \\ &= y_k + \sum_{i=1}^s f(x_k + c_i h, u(x_k + c_i h)) \int_{x_k}^x l_i\left(\frac{t-x_k}{h}\right) dt \end{aligned}$$

jetzt: $t_{neu} = \frac{t-x_k}{h}$

$\Rightarrow x_k \rightarrow 0$, $x \rightarrow \frac{x-x_k}{h}$, $dt = h dt_{neu}$

$$u(x) = y_k + h \sum_{i=1}^s f(x_k + c_i h, u(x_k, c_i h)) \int_0^{\frac{x-x_k}{h}} l_i(t) dt$$

Insbesondere

$$\begin{aligned} u(x_k + c_j h) &= y_k + h \sum_{i=1}^s f(x_k + c_i h, u(x_k + c_i h)) \int_0^{c_j} l_i(t) dt \\ &= y_k + h \sum_{i=1}^s a_{ji} f(x_k, c_i h, u(x_k, c_i h)) \end{aligned}$$

Wir setzen $k_i := f(x_k, c_i h, u(x_k, c_i h))$

Dann gilt:

$$\begin{aligned} k_j &= f(x_k + c_j h, u(x_k, c_j h)) \\ &= f(x_k + c_j h, y_k + h \sum_{i=1}^s a_{ji} k_i) \end{aligned}$$

Wie im Runge-Kutta-Verfahren: $y_{k+1} = u(x_{k+1})$.

$$\begin{aligned} y_{k+1} &= x_k + h \sum_{i=1}^s \underbrace{f(x_k + c_i h, u(x_k + c_i h))}_{k_i} \int_0^h l_i(t) dt \\ &= y_k + h \sum_{i=1}^s k_i \int_0^1 l_i(t) dt \\ &= y_k + h \sum_{i=1}^s b_i k_i \end{aligned}$$

wie im Runge-Kutta-Verfahren 1

5 Beweis (Abschnitt 1.2.10):

Wir betrachten einen Schritt und wir gehen von der „richtigen“ Lösung des AWP's $y' = f(x, y)$, $y(x_k) = y_k$ aus. $u(x)$ ist das Polynom aus dem Kollokationsverfahren.

$$\begin{aligned} y_{k+1} - y(x_{k+1}) &= u(x_{k+1}) - y(x_{k+1}) \\ &= \int_{x_k}^{x_{k+1}} \Phi(x_{k+1}, t, u(t)) (u'(t) - f(t, u(t))) dt \\ &= \int_{x_k}^{x_{k+1}} \Phi(x_{k+1}, t, u(t)) d(t, u(t)) dt \end{aligned}$$

mit $d(t, u(t)) = u'(t) - f(t, u(t))$

Für $\int_{x_k}^{x_{k+1}} \Phi(x_{k+1}, t, u(t)) d(t, u(t)) dt$ verwenden wir die Quadraturformel mit Knoten $x_k + c_1 h, x_k + c_2 h, \dots, x_k + c_s h$

Orthogonalität \Rightarrow diese Quadraturformel ist exakt für \mathbb{P}_{s-1+m} .

$$y_{k+1} - y(x_{k+1}) = \sum_{j=1}^s \lambda_j \Phi(x_{k+1}, x_k + c_j h, u(x_k + c_j h)) \cdot d(x_k + c_j h, u(x_k + c_j h)) + R(\Phi d)$$

*:= 0, da $d(x_k + c_j h, u(x_k + c_j h)) = u'(x_k + c_j h) - f(x_k + c_j h, u(x_k + c_j h)) = 0$ per Konstruktion!

$$\Rightarrow y_{k+1} - y(x_k) = R(\Phi d)$$

$$\Phi d = p_{s-1+m}(t) + O(t^{s+m})$$

$$R(\Phi d) = R(p_{s-1+m}) + R(O(t^{s+m}))$$

*:= 0, da quadraturformel exakt für \mathbb{P}_{s-1+m} ist!

$$y_{k+1} - y(x_{k+1}) = R(O(t^{s+m}))$$

$$\begin{aligned} R(O(t^{s+m})) &= \int_{x_k}^{x_{k+1}} O(t^{s+m}) dt \\ &= O((x_{k+1} - x_k)^{s+m-1}) = O(h^{s+m-1}) \end{aligned}$$

Folglich $y_{k+1} - y(x_{k+1}) = O(h^{s+m-1})$, $h \rightarrow 0$

Ende des Beweises:

$$\tau(x_k, y(x_k), h) = \frac{y(x_{k+1}) - y(x_k)}{h} - \varphi(x_k, y(x_k), h)$$

$$y_{k+1} = y_k + h\varphi(x_k, y(x_k), h)$$

$$\varphi(x_k, y(x_k), h) = \frac{y_{k+1} - y_k}{h}$$

Hier jeweils $y(x_k) = y_k$!

$$\begin{aligned} \text{Damit } \tau(x_k, y(x_k), h) &= \frac{y(x_{k+1}) - y_k}{h} - \frac{y_{k+1} - y_k}{h} \\ &= \frac{y(x_{k+1}) - y_{k+1}}{h} = \frac{O(h^{s+m+1})}{h} = O(h^{s+m}), \quad h \rightarrow 0 \end{aligned}$$

6 Beweis (Abschnitt 1.3.1):

$$\begin{aligned} 1) \int_a^b (1-z)^{-t} dt &= \int_a^b e^{-t \ln(1-z)} dt \\ &= \frac{1}{-\ln(1-z)} e^{-t \ln(1-z)} \Big|_{t=a}^{t=b} \\ &= -\frac{(1-z)^{-t}}{\ln(1-z)} \Big|_{t=a}^{t=b} \end{aligned}$$

2) Wir zeigen, dass die Reihe

$$(1-z)^{-t} = \sum_{i=0}^{\infty} \binom{t+i-1}{i} z^i$$

gliedweise integriert werden kann $|z| \leq R$

wir wollen zeigen, dass die Reihe gleichmäßig im Kreis $\{z : |z| \leq R\}$ konvergiert.

i) $t \in [0, 1]$

$$\begin{aligned} \left| \frac{t+j-1}{j} \right| &= \left| 1 + \frac{t-1}{j} \right| \leq 1 \\ \binom{t+i-1}{i} &= \frac{(t+i-1)(t+i-2)\dots t}{i(i-1)\dots 1} = \frac{t+i-1}{i} \frac{t+i-2}{i-1} \dots \frac{t}{1} \end{aligned}$$

$$= \prod_{j=1}^i \frac{t+j-1}{j} \Rightarrow \left| \binom{t+i-1}{i} \right| \leq 1$$

$$\sum_{i=0}^{\infty} \left| \binom{t+i-1}{i} z^i \right| \leq \sum_{i=0}^{\infty} R^i$$

\Rightarrow Die Reihe konvergiert gleichmäßig

ii) $a \leq t < 0$

$$j \geq a+1 \quad \left| \frac{t+j-1}{j} \right| = \left| 1 + \frac{t-1}{j} \right| \leq 1$$

$$0 \geq t \geq a : \quad |t-1| \leq a+1 : \quad \frac{|t-1|}{j} \leq 1, \quad -1 \leq \frac{t-1}{j} \leq 0$$

$$\text{Für } 1 \leq j < a : \quad \left| \frac{t+j-1}{j} \right| = \left| 1 + \frac{t-1}{j} \right|$$

$$1 \geq 1 + \frac{t-1}{j} = 1 - \frac{(1-t)}{j} \geq 1 - (1-t) = t \geq a$$

$$\begin{aligned} \left| \frac{t+j-1}{j} \right| &\leq \begin{cases} 1 & , j \geq a \\ a & , 1 \leq j < a \end{cases} \\ \left| \binom{t+i-1}{i} \right| &= \prod_{j=1}^i \left| \frac{t+j-1}{j} \right| \\ &= \prod_{j=1}^{[a]+1} \underbrace{\left| \frac{t+i-1}{j} \right|}_{\leq a} \prod_{j=[a]+2}^i \underbrace{\left| \frac{t+i-1}{j} \right|}_{\leq 1} \leq a^{[a]+i} \\ \sum_{i=0}^{\infty} \left| \binom{t+i-1}{i} z^i \right| &\leq ia^{[a]+1} \sum_{i=0}^{\infty} R^i \\ \text{Für alle } a \leq t \leq b &\text{ konvergiert die Reihe gleichmäßig in } t, \text{ da} \\ \sum_{i=0}^{\infty} \left| \binom{t+i-1}{i} z^i \right| &\leq \max\{1, a^{[a]+1}\} \sum_{i=0}^{\infty} R^i \end{aligned}$$

3) Wir integrieren gliedweise:

$$\begin{aligned} g(z, a, b) &= \int_a^b \sum_{i=0}^{\infty} \binom{t+i-1}{i} z^i dt \\ &= \sum_{i=0}^{\infty} \int_a^b \binom{t+i-1}{i} dt z^i = \sum_{i=0}^{\infty} b_i^* z^i \end{aligned}$$

7 Beweis (Abschnitt 1.3.2):

1) (i) \Leftrightarrow (ii)

$$\begin{aligned} y(x+ih) &= \sum_{m=0}^p \frac{y^{(m)}(x)}{m!} i^m h^m + O(h^{p+1}) \\ y'(x+ih) &= \sum_{m=0}^{p-1} \frac{y^{(m+1)}(x)}{m!} i^m h^m + O(h^p) \\ &= \sum_{m=1}^p \frac{y^{(m)}(x)}{(m-1)!} i^{m-1} h^{m-1} + O(h^p) \\ \tau(x, y(x), h) &= \frac{1}{h} \sum_{i=0}^s a_i y(x+ih) - \sum_{i=0}^s b_i y'(x+ih) \\ &= \frac{1}{h} \sum_{i=0}^s \left(a_i \sum_{m=0}^p \frac{y^{(m)}(x)}{m!} i^m h^m - b_i \sum_{m=1}^p \frac{y^{(m)}(x)}{m!} m i^{m-1} h^m \right) + O(h^p) \\ &= \frac{1}{h} \sum_{i=0}^s a_i y(x) + \frac{1}{h} \sum_{m=1}^p \frac{y^{(m)}(x)}{m!} h^m \cdot \sum_{i=0}^s (a_i i^m - b_i m i^{m-1}) + O(h^p) \end{aligned}$$

$$* = \varrho(1)$$

$$* = L(t^m, 1)$$

$$g = t^m$$

$$L(t^m, h) = \frac{1}{h} \sum_{i=0}^s (a_i (ih)^m - h b_i m (ih)^{m-1})$$

$$L(t^m, 1) = \frac{1}{h} \sum_{i=0}^s (a_i i^m - h b_i m i^{m-1})$$

$$\tau(x, y(x), h) = \frac{1}{h} \varrho(1) y(x) + \frac{1}{h} \sum_{m=1}^p \frac{y^{(m)}(x)}{m!} h^m L(t^m, 1) + O(h^p)$$

$$L(t^0, 1) = \sum_{i=0}^s a_i \cdot 1^0 = \sum_{i=0}^s a_i = \varrho(1), \quad g = 1, \quad g' = 0$$

$$\tau(x, y(x), h) = \frac{1}{h} \sum_{m=0}^p \frac{y^{(m)}(x)}{m!} h^m L(t^m, 1) + O(h^p)$$

Folgt: Ordnung $p \Leftrightarrow L(t^m, 1) = 0$ für $m = 0, 1, \dots, p$: (i) \Leftrightarrow (ii)

1) (ii) \Leftrightarrow (iii)

$$\tau(x, y(x), h) = L(y(\underbrace{x}_{\text{fest}} + t)) = \frac{1}{h} \sum_{m=0}^p h^m \frac{y^{(m)}(x)}{m!} L(t^m, 1) + O(h^p)$$

Speziell für $y = e^x$ an der Stelle $x = 0$:

$$L(e^t, h) = \frac{1}{h} \sum_{m=0}^p \frac{h^m}{m!} L(t^m, 1) + O(h^p)$$

Andererseits:

$$\begin{aligned} L(e^t, h) &= \frac{1}{h} \sum_{i=0}^s (a_i e^{ih} - h b_i e^{ih}) \\ &= \frac{1}{h} \sum_{i=0}^s a_i (e^h)^i - \sum_{i=0}^s b_i (e^h)^i \\ &= \frac{1}{h} \varrho(e^h) - \sigma(e^h) \end{aligned}$$

$$\varrho(e^h) - \sigma(e^h) = \sum_{m=0}^p \frac{h^m}{m!} L(t^m, 1) + O(h^{p+1})$$

$$(ii) \Leftrightarrow \varrho(e^h) - h\sigma(e^h) = O(h^{p+1})$$

$$\Leftrightarrow h = 0 \text{ ist } (p+1)\text{-fache Nullstelle}$$

2) (iii) \Leftrightarrow (iv)

Transformation: $\xi = e^w$

$$w = \ln(\xi), \quad w = 0 \Rightarrow \xi = 1$$

$$\varrho(e^w) - w\sigma(e^w) = \varrho(\xi) - \ln(\xi)\sigma(\xi)$$

$$\varrho(1) = L(t^0, 1) = 0$$

$$\Rightarrow \frac{\varrho(\xi)}{\ln(\xi)} \text{ ist analytisch} \Rightarrow \frac{\varrho(\xi)}{\ln(\xi)} - \sigma(\xi)$$

$$\xi = 1 \text{ ist } p\text{-fache Nullstelle.}$$

8 Beweis (Abschnitt 1.3.2):

Konsistenz $\Leftrightarrow w = 0$ ist mindestens 2-fache Nullstelle von $\varrho(e^w) - w\sigma(e^w)$

$$\varrho(e^w) - w\sigma(e^w)|_{w=0} = \varrho(1) \overset{0}{0}$$

$$(\varrho(e^w) - w\sigma(e^w))'|_{w=0} = (\varrho'(e^w)e^w - \sigma(e^w) - w\sigma'(e^w)e^w)|_{w=0} = \varrho(1) - \sigma(1) \overset{0}{=} 0$$

9 Beweis (Abschnitt 1.3.3):

1) klar.

2) $\{y_k\}, \{\tilde{y}_k\}$ Lösungen $\Rightarrow \{y_k + \tilde{y}_k\}$ ist auch Lösung
 \Rightarrow Die Menge aller Lösungen bilden einen linearen Teilraum

Die Lösungen $\{y_k^{(j)}\}_{k=0}^\infty$:

$$\{y_k^{(0)}\} = \{1, \underbrace{0, \dots, 0}_s, y_s^{(0)}, y_{s+1}^{(0)}, \dots\}$$

$$\{y_k^{(1)}\} = \{0, 1, \dots, 0, y_s^{(1)}, y_{s+1}^{(1)}, \dots\}$$

...

$$\{y_k^{(s-1)}\} = \{0, \dots, 0, 1, y_s^{(s-1)}, y_{s+1}^{(s-1)}, \dots\}$$

sind linear unabhängig \Rightarrow die Dimension des Raums der Lösungen ist $\geq s$

Ist $\{x_k\}_{k=0}^\infty$ eine Lösung.

$$\text{Betrachte } z_k = \sum_{j=0}^{s-1} x_j y_k^{(j)}$$

$\{z_k\}$ ist auch eine Lösung (als Linearkombination von Lösungen)

Es gilt: $x_k = z_k$ für $k = 0, 1, \dots, s-1$

Da aber die Lösung durch die ersten s Werte bestimmt ist, gilt $x_k = z_k$ für alle k

$\Rightarrow \{y_k^{(j)}\}, j = 0, 1, \dots, s-1$ bilden eine Basis des Raumes aller Lösungen.

Insbesondere ist die Dimension dieses Raumes gleich s .

3) $\{y_k\}$ eine Lösung.

$$\vec{y}_k = \begin{pmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+s-1} \end{pmatrix} \in \mathbb{C}^s$$

$$\vec{y}_{k+1} = A_k \vec{y}_k \text{ mit}$$

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{0,k} & -a_{1,k} & -a_{2,k} & \cdots & -a_{s-1,k} \end{pmatrix}$$

$$y_{k+s} = -a_{0,k}y_k - a_{1,k}y_{k+1} - \dots - a_{s-1,k}y_{k+s-1}$$

$$\det(A_k) = (-1)^{s+1}(-a_{0,k}) \det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (-1)^s a_{0,k}$$

Wegen $a_{0,k} \neq 0$ sind alle Matrizen A_k regulär

$$\vec{y}_i = \begin{pmatrix} y_i \\ y_{i+1} \\ \vdots \\ y_{i+s-1} \end{pmatrix} = A_{i-1}A_{i-2}\dots A_1A_0\vec{y}_0$$

10 Beweis (Abschnitt 1.3.3):

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{s-1} \end{pmatrix} = A$$

Dann gilt: $\vec{y}_k = A^k \vec{y}_0$

Wähle T so, dass $T^{-1}AT = J$ die Jordanische Normalform

Unterscheide drei Fälle:

1) Alle λ_i einfach

$$\Rightarrow J = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_s \end{pmatrix} = \text{diag}(\lambda_i)_{i=1}^s$$

$$\text{Dann gilt: } \vec{y}_k = A^k \vec{y}_0 = T J^k T^{-1} \vec{y}_0 = T \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_s^k \end{pmatrix} T^{-1} \vec{y}_0$$

$$\Rightarrow y_k = \sum_{i=1}^s \beta_i \lambda_i^k$$

2. Allgemeiner Fall:

$$J = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_m \end{pmatrix} \text{ mit } J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix}$$

$$\Rightarrow J_i^k = \begin{pmatrix} \lambda_i^k & \binom{k}{1} \lambda_i^{k-1} & \binom{k}{2} \lambda_i^{k-2} & \cdots \\ 0 & \lambda_i^k & \binom{k}{1} \lambda_i^{k-1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ mit } \binom{k}{j} = 0 \text{ für } j > k$$

$$\text{d.h. } J_i^1 = \begin{pmatrix} \lambda_i & 1 & \cdots \\ \vdots & \lambda_i & 1 \\ \vdots & \vdots & \ddots \end{pmatrix}, J_i^2 = \begin{pmatrix} \lambda_i^2 & 2\lambda_i & 1 & \cdots \\ \vdots & \vdots & \ddots & \end{pmatrix}$$

$$\Rightarrow j^k = \begin{pmatrix} J_1^k & & & \\ & J_2^k & & \\ & & \ddots & \\ & & & J_m^k \end{pmatrix}$$

Alle Koeffizienten von J^k haben die Form
 „ $\lambda_i^k \dots$ Polynom in k vom Grad $\leq \tau_i - 1$ “
 $\Rightarrow \vec{y}_k = T J^k T^{-1} \vec{y}_0$
 $\Rightarrow y_k$ haben die gewünschte Form.

11 Beweis (Abschnitt 1.3.4):

Betrachte das AWP $y_0 = 0 \quad y(0) = 0$
 Die exakte Lösung $y \equiv 0$
 $y(x, h)$ bezeichnet die Näherungslösung.
 Konvergenz $\Rightarrow y(x, h) \rightarrow 0 \quad h \rightarrow 0$ für jedes feste x .

x fest, $x \neq x_0$
 $h_k = \frac{x}{k}$
 $y_k = y(x, h_k)$

Für dieses k gilt: $\sum_{i=0}^s a_i y_{k+i} = 0$

y_k ist die Lösung dieser Differenzgleichung.

$\varrho(\lambda) = \sum_{i=0}^s a_i \lambda^i$ = das erste charakteristische Polynom des Verfahrens.

$\lambda_1, \lambda_2, \dots, \lambda_s$ sind die Nullstellen.

1) Sei λ eine Nullstelle, dann ist $y_k = h \lambda^k$ eine Lösung.

$$y_k = y(x, h_k) = h \lambda^{\frac{x}{h}} \rightarrow 0, h \rightarrow 0 \\ \Leftrightarrow |\lambda| \leq 1$$

2) Sei λ eine mehrfache Nullstelle.

Dann ist $y_k = \sqrt{h} k \lambda^k$ eine Lösung.

$$y_k = y(x, h_k) = \sqrt{h} \frac{x}{h} \lambda^{\frac{x}{h}} = x \frac{1}{\sqrt{h}} \lambda^{\frac{x}{h}} \rightarrow 0, h \rightarrow 0 \\ \Leftrightarrow |\lambda| < 1$$

12 Beweis (Abschnitt 1.3.4):

1) Betrachte $y' = 0$, $y(0) = 1$

Die exakte Lösung $y \equiv 1$

$$\sum_{i=0}^s a_i y_{k+i} = 0 \text{ für die Lösung.}$$

Konvergenz $\Rightarrow \lim_{h \rightarrow 0} y(x, h) = 1$ für jedes x .

x fest, $h_k = \frac{x}{k}$, $k = \frac{x}{h_k}$

$$\lim_{h \rightarrow 0} y_{\frac{x}{h} + i} = 1$$

$$\lim_{h \rightarrow 0} \sum_{i=0}^s a_i y_{\frac{x}{h} + i} = 0$$

$$\Rightarrow \sum_{i=0}^s a_i = 0 = \varrho(1)$$

2) Betrachte $y' = 1$, $y(0) = 0$

Die exakte Lösung: $y = x$

Konvergenz \Rightarrow die Wurzelbedingung

$\Rightarrow \lambda = 1$ ist eine einfache Nullstelle

$\varrho'(1) \neq 0$

Setze $M = \frac{\sigma(1)}{\varrho'(1)}$.

Ziel: zu zeigen, dass $M = 1$ ist.

Zu den Startwerten $y_k = Mkh$, $k = 0, \dots, s-1$ liefert das Verfahren eine Folge y_k

mit $\sum_{i=0}^s a_i y_{k+1} = h \sum_{i=0}^s b_i f_{k+i} = h \sum_{i=0}^s b_i = h\sigma(1)$

Wir betrachten die Differenzgleichung

$$\sum_{i=0}^s a_i y_{k+i} = h\sigma(1)$$

mit den Startwerten $y_k = Mkh$, $k = 0, \dots, s-1$

Die Lösung ist $y_k = Mkh$, $k \in \mathbb{N}_0$

In der Tat:

$$\begin{aligned} \sum_{i=0}^s a_i y_{k+i} &= \sum_{i=0}^s a_i M(k+i)h \\ &= Mkh \underbrace{\sum_{i=0}^s a_i}_{=\varrho(1)=0} + Mh \underbrace{\sum_{i=0}^s a_i i}_{\sum_{i=0}^{s-1} y_i} \\ &= h\sigma(1) \end{aligned}$$

$\Rightarrow y_k = Mkh$ ist eine Lösung.

x fest, $h_k = \frac{x}{k}$

$y(x, h) = Mx \rightarrow x$, $h \rightarrow 0$

$\Leftrightarrow M = 1$

$\Rightarrow \varrho'(1) = \sigma(1)$

\Rightarrow Konsistenz

13 Beweis (Abschnitt 1.3.4):

Wir betrachten das Verfahren

$$\sum_{i=0}^s a_i y_{k+1} = h \sum_{i=0}^{s-1} b_i f_{k+i}, \quad a_s = 1$$

Wir nehmen an, dass die Startwerte y_0, \dots, y_{s-1} gegen die exakte Lösung konvergieren:

$y_j - y(x_j) = o(1)$, $h \rightarrow 0$, $j = 0, \dots, s-1$

Wir definieren eine Hilfsfolge c_l

$$c_0 = 1$$

$$c_1 + a_{s-1}c_0 = 0$$

$$c_2 + a_{s-1}c_1 + a_{s-2}c_0 = 0$$

\vdots

$$c_s + a_{s-1}c_1 + \dots + a_0c_0 = 0$$

und für $l > s$:

$$c_l + a_{s-1}c_{l-1} + a_{s-2}c_{l-2} + \dots + a_0c_{l-s} = 0$$

$\{c_l\}$ ist die Lösung einer Differenzgleichung mit dem charakteristischen Polynom ϱ

$\Rightarrow c_l$ ist beschränkt (Wurzelbedingung)

$$\sum_{k=0}^l c_{l-k} \sum_{i=0}^s a_i y_{k+i} = y_{l+s} + \sum_{j=0}^{s-1} \beta_j y_j \quad \beta_j \text{ beschränkt.}$$

$$\begin{aligned}
\text{In der Tat: } & \sum_{k=0}^l c_{l-k} \sum_{i=0}^s a_i y_{k+i} = \\
& c_l(a_0 y_0 + a_1 y_1 + \dots + a_s y_s) \\
& + c_{l-1}(a_0 y_1 + a_1 y_2 + \dots + a_s y_{s+1}) \\
& \vdots \\
& + c_0(a_0 y_l + a_1 y_{l+1} + \dots + a_s y_{s+l}) \\
& = c_l a_0 y_0 + (c_l a_1 + c_{l-1} a_0) y_1 \\
& + (c_l a_{s-1} + \dots + c_1 a_0) y_{s-1} \\
& + \underbrace{(c_l a_s + c_{l-1} a_{s-1} + \dots + c_{l-s} a_0)}_{=0} y_s \\
& + \underbrace{(c_{l+1} a_s + c_l a_{s-1} + \dots + c_{l-s+1} a_0)}_{=0} y_{s+1} \\
& \vdots \\
& + \underbrace{(c_1 a_s + c_0 a_{s-1})}_{=0} y_{l+s-1} \\
& + \underbrace{a_s}_{=1} \underbrace{c_0}_{=1} y_l + s \\
\Rightarrow & \sum_{k=0}^l c_{l-k} \sum_{i=0}^s a_i y_{k+l} = y_{l+s} + \sum_{j=0}^{s-1} \beta_j y_j \\
& \beta_j \text{ sind beschränkt}
\end{aligned}$$

$$\begin{aligned}
\sum_{k=0}^l c_{l-k} \sum_{i=0}^s a_i y_{k+i} &= h \sum_{k=0}^l c_{l-k} \sum_{i=0}^{s-1} b_i f_{k+i} \\
y_{s+l} &= h \sum_{k=0}^l c_{l-k} \sum_{i=0}^{s-1} b_i f_{k+i} - \sum_{j=0}^{s-1} \beta_j y_j
\end{aligned}$$

Für die exakte Lösung gilt:

$$\sum_{i=0}^s a_i y(x_{k+i}) = h \sum_{i=0}^{s-1} b_i f(x_{k+i}, y(x_{k+i})) + O(1)h \quad \text{wegen der Konsistenz}$$

$$\sum_{k=0}^l c_{l-k} \sum_{i=0}^s a_i y(x_{k+i}) = h \sum_{k=0}^l c_{l-k} \sum_{i=0}^{s-1} b_i f(x_{k+i}, y(x_{k+i})) + O(1)h$$

$$y(x_{l+s}) = h \sum_{k=0}^l c_{l-k} \sum_{i=0}^{s-1} b_i f(x_{k+i}, y(x_{k+i})) - \sum_{j=0}^{s-1} \beta_j y(x_j) + O(1)h$$

Für globalen Fehler:

$$\begin{aligned}
e_{l+s,h} &= y_{l+s} - y(x_{l+s}) \\
&= h \sum_{k=0}^l c_{l-k} \sum_{i=0}^{s-1} b_i (f(x_{k+i}, y_{k+i}) - f(x_{k+i}, y(x_{k+i}))) - \sum_{j=0}^{s-1} \beta_j \underbrace{(y_j - y(x_j))}_{\leq L|y_{k+i} - y(x_{k+i})} + O(1)h
\end{aligned}$$

$$|\cdot| \leq L|y_{k+i} - y(x_{k+i})| = L e_{k+i,h} \quad \text{da } f \text{ lipschitz-stetig ist.}$$

$$\Rightarrow |e_{l+s,h}| \leq hM \sum_{i=0}^{l+s-1} |e_{i,h}| + O(1)$$

Mit einer Konstante M.

Nach Lemma (Anfang d. Vorlesung):

$$|e_{n,h}| \leq \frac{1-(hM)^n}{1-hM} O(1) \rightarrow 0, \quad h \rightarrow 0$$

14 Beweis (Abschnitt 2.2.2):

Induktion in j:

Induktionsanfang: $j = 1$

$$\begin{aligned}
l_{11} &= \sqrt{a_{11}} \\
l_{21} &= \frac{a_{21}}{l_{11}} \\
&\vdots \\
l_{n1} &= \frac{a_{n1}}{l_{11}} \text{ (immer so)}
\end{aligned}$$

Induktionsschritt: $j = q - 1 \rightarrow j = q$

Nach Induktionsannahme ist die Besetzungsstruktur für die ersten $q-1$ Spalten erhalten.

$$\begin{aligned}
a_{kj} = 0 &\Rightarrow l_{kj} = 0 \text{ für } 1 \leq j \leq q - 1 \\
l_{qq} &= \sqrt{a_{qq} - \sum_{i=1}^{q-1} l_{qi}^2} \\
l_{kq} &= \frac{1}{l_{qq}} \left(a_{kq} - \sum_{i=1}^{q-1} l_{ki} l_{qi} \right), \quad q + 1 \leq k \leq n
\end{aligned}$$

Zu zeigen: $\sum_{i=1}^{q-1} l_{ki} l_{qi} = 0$

Der Baum ist monoton \Rightarrow Für jedes i , $1 \leq i \leq q - 1$ gibt es genau ein k_i mit $i + 1 \leq k_i \leq n$, sodass $(i, k_i) \in K$

Sei $q + 1 \leq k \leq n$

1) $l_{ki} \neq 0$ für ein i , $1 \leq i \leq q - 1$
 $\Rightarrow a_{ki} \neq 0 \Rightarrow (i, k) \in K$ (k ist ein Elternknoten) $\Rightarrow (i, q) \notin K$
 $\Rightarrow a_{qi} = 0 \Rightarrow l_{qi} = 0$

2) Analog: $l_{qi} \neq 0 \Rightarrow l_{ki} = 0$

Es folgt, dass $\sum_{i=1}^{q-1} l_{ki} l_{qi} = 0$

15 Beweis (Abschnitt 2.3.1):

Notwendigkeit, siehe Numerik 1!

Sei also $\rho(H) < 1$, also $\lim_{k \rightarrow \infty} H^k = 0$

Damit gilt dann:

$$x_k = H^k x_0 + \underbrace{(H^{k-1} + H^{k-2} + \dots + H^1 + H^0)}_{=(I-H)^{-1}(I-(H^k=0))} V$$

Also gilt dann:

$$\lim_{k \rightarrow \infty} x_k = (I - H)^{-1} I V = (I - H)^{-1} V = (I - P^{-1} N)^{-1} P^{-1} b = (P - N)^{-1} b = A^{-1} b$$

16 Beweis (Abschnitt 2.3.1):

Sei $\lambda \in \sigma(H)$ EW von H , Z.z.: $|\lambda| < 1$, denn dann ist $\rho(H) < 1$.

Sei v EV von H zu λ , dann ist

$$Hv = (P^{-1} N)v = (P^{-1}(P - A))v = (I - P^{-1} A)v = \lambda v$$

Multipliziere dies von links mit P , dann folgt:

$$(P - A)v = \lambda P v$$

$$\Leftrightarrow Pv - \lambda P v = Av$$

$$\Leftrightarrow (1 - \lambda)P v = Av$$

Da A pos. def., insbes regulär ist, folgt $\lambda \neq 1$.

Weiter:

$$\begin{aligned} R \ni v^T A v &= (1 - \lambda) v^T P v \\ &= (1 - \bar{\lambda}) v^T P v \end{aligned}$$

Es folgt:

$$\begin{aligned} \left(\frac{1}{1-\lambda} + \frac{1}{1-\bar{\lambda}}\right) v^T A v &= v^T (P + P^T - A) v \\ &= \frac{1-|\lambda|^2}{|1-\lambda|^2} \quad \text{mit } \lambda = \bar{\lambda}, \text{ da alle Werte reell.} \\ &= \frac{1-|\lambda|^2}{|1-\lambda|^2} \underbrace{v^T A v}_{>0} = \underbrace{v^T (P + P^T - A) v}_{>0} \end{aligned}$$

17 Beweis (Abschnitt 2.3.2):

Sei j Ordnungsvektor und $i_k = j_{\pi(k)}$, π Permutation. Wähle π derart, dass $i_1 \leq i_2 \leq \dots \leq i_n$.
Dann ist i kompatibler Ordnungsvektor für PAP^{-1} .

18 Beweis (Abschnitt 2.3.2):

$H = tL_0 + t^{-1}R_0 - sD$ hat bzgl. Ordnungsspektren die gleiche Besetzungsstruktur wie A
 $\Rightarrow H$ hat auch kompatiblen OV.

$$H = (h_{ij})_{ij=1}^n$$

$$g(s, t) = \det(H) = \sum_{\pi \in \prod_n} (-1)^{|\pi|} \prod_{i=1}^n h_{i, \pi(i)}$$

$$h_{ij} = -t^{\sigma_{i-j}} s^{1-|\sigma_{i-j}|} a_{ij}$$

$$\sigma_m = \begin{cases} 0 & , m = 0 \\ +1 & , m > 0 \\ -1 & , m < 0 \end{cases}$$

$$g(s, t) = (-1)^n \sum_{\pi \in \prod_n} (-1)^{|\pi|} t^{d_L(\pi) - d_R(\pi)} s^{n - d_L(\pi) - d_R(\pi)} \prod_{i=1}^n a_{i, \pi(i)}$$

$$d_L(\pi) = |\{l \mid l > \pi(l)\}|$$

$$d_R(\pi) = |\{l \mid l < \pi(l)\}|$$

Zu zeigen: $d_L(\pi) = d_R(\pi)$

$$d_L(\pi) = \sum_{\substack{l > \pi(l) \\ l=1}}^n j_l - j_{\pi(l)}$$

$$d_R(\pi) = \sum_{\substack{l < \pi(l) \\ l=1}}^n j_{\pi(l)} - j_l$$

$$d_L(\pi) - d_R(\pi) = \sum_{l=1}^n j_l - \sum_{l=1}^n j_{\pi(l)} = 0$$

19 Beweis (Abschnitt 2.3.2):

$$\begin{aligned} \det(L_0 + R_0 - \mu D) &= g(\mu, 1) \stackrel{\text{Lemma}}{=} g(\mu, -1) \\ &= \det(-L_0 - R_0 - \mu D) = (-1)^n \det(L_0 + R_0 + \mu D) \end{aligned} \quad \text{Sei } \mu \in \sigma(B), \text{ d.h.}$$

$$\begin{aligned} \det(D^{-1}(L_0 + R_0) - \mu I) &= 0 \\ &= \det(D^{-1}) \det(L_0 + R_0 - \mu D) = 0 \\ &= (-1)^n \det(D^{-1}) \det(L_0 + R_0 + \mu D) = 0 \\ &= (-1)^n \det(\underbrace{D^{-1}(L_0 + R_0) + \mu I}_{=B}) = 0 \end{aligned}$$

$$\Rightarrow -\mu \in \sigma(B)$$

Für 2. bzw. für die SOR-Matrix $\det(I - \omega L_0) = 1$

$$\begin{aligned}\det(\tilde{L}_\omega - \lambda I) &= \det((I - \omega L_0)^{-1}(\omega R_0 + (1 - \omega)I) - \lambda I) \\ &= \det(\omega R_0 + (1 - \omega)I - \lambda(I - \omega L_0)) \\ &= \det(\omega R_0 + \lambda\omega L_0 - (\lambda + \omega - 1)I)\end{aligned}$$

$\lambda = 0$ (Fall 1):

Für $\lambda = 0$ muss $\det(\omega R_0 + (1 - \omega)I) = (1 - \omega)^n = 0$.

$\omega = 1$: Gleichung $\lambda + \omega - 1 = \omega\mu\sqrt{\lambda}$ gilt für $\omega - 1 = \lambda = 0$

$\lambda \neq 0$ (Fall 2):

$$\det(\tilde{L}_\omega - \lambda I) = \det(\omega R_0 + \lambda\omega L_0 - (\lambda + \omega - 1)I)$$

Mit $t = \sqrt{\lambda}$ und unserem Lemma gilt:

$$\det(\tilde{L}_\omega - \lambda) = \det(\omega\sqrt{\lambda}\left(\frac{1}{\sqrt{\lambda}}R_0 + \sqrt{\lambda}L_0 - \frac{\lambda + \omega - 1}{\omega\sqrt{\lambda}}I\right))$$

Nach der Voraussetzung von 2.):

$$\dots = \det(\omega\sqrt{\lambda}\left(\frac{1}{\sqrt{\lambda}}R_0 + \sqrt{\lambda}L_0 - \mu I\right))$$

$$= \det(\omega\sqrt{\lambda}(R_0 + L_0 - \mu I)) = 0 \Rightarrow \mu \in \sigma(B)$$

20 Beweis (Abschnitt 2.3.2):

Mit $\omega = 1$ wird SOR Verfahren zum Gauß-Seidel-Verfahren:

$$\lambda = \mu\sqrt{\lambda} \quad (\lambda = 0)$$

$$\mu = \sqrt{\lambda} \quad \lambda = \mu^2, \quad \mu \in \sigma(B)$$

21 Beweis (Abschnitt 2.3.2):

$$\varrho(\tilde{L}_\omega) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$$

$$\prod_{k=1}^n |\lambda_k| = |\det(\tilde{L}_\omega)| = |1 - \omega|^n$$

$$\varrho(\tilde{L}_\omega) = \max |\lambda_k| \geq \left(\prod_{k=1}^n |\lambda_k|\right)^{1/2} = |1 - \omega|$$

$$\text{Ist } \varrho(\tilde{L}_\omega) < 1 \Rightarrow |1 - \omega| < 1 \Rightarrow 0 < \omega < 2$$

22 Beweis (Abschnitt 2.3.2):

$$\begin{aligned}\mu &\in \sigma(B), \quad \mu \leq \bar{\mu} \\ \lambda + \omega - 1 &= \omega\mu\sqrt{\lambda}\end{aligned}$$

$$\text{Lösung: } \lambda_{1,2} = \frac{1}{4}(\omega\mu \pm \sqrt{(\omega\mu)^2 - 4(\omega - 1)})^2$$

$$\text{Betrachte } \lambda_2 = \frac{1}{4}F(\omega, |\mu|)^2$$

$$F(\omega, t) = \omega t + \sqrt{(\omega t)^2 - 4(\omega - 1)}$$

$F(\omega, t)$ wächst streng monoton in t :

$$\varrho(\tilde{L}_\omega) = \begin{cases} \frac{1}{4}(\omega\bar{\mu} + \sqrt{\omega^2\bar{\mu}^2 - 4(\omega - 1)}) & , 0 < \omega < \omega_{opt} \\ \omega - 1 & \omega_{opt} \leq \omega < 2 \end{cases}$$

Für $\omega_{opt} \leq \omega < 2$ ist die Diskriminante $\omega^2\bar{\mu}^2 - 4(\omega - 1)$ negativ, dann $|\lambda_{1,2}| = \left|\frac{1}{4}(\omega\mu \pm i\sqrt{4(\omega - 1) - (\omega\mu)^2})\right| = |\omega - 1|$

$$\text{Für } \omega = \omega_{opt} = 1 + \left(\frac{\bar{\mu}}{1 + \sqrt{1 - \bar{\mu}^2}}\right)^2$$

sind $\omega - 1$ und

$$\frac{1}{4}(\omega\bar{\mu} + \sqrt{\omega^2\bar{\mu}^2 - 4(\omega - 1)})$$

$$\text{gleich. und zwar } \varrho(\tilde{L}_{\omega_{opt}}) = \omega_{opt} - 1 = \left(\frac{\bar{\mu}}{1 + \sqrt{1 - \bar{\mu}^2}}\right)^2$$

23 Beweis (Abschnitt 2.3.3):

Sei λ ein Eigenwert und \vec{x} der dazugehörige Eigenvektor mit $\|\vec{x}\| = 1$.
 $|\lambda| = |\lambda|\|\vec{x}\| = \|A\vec{x}\| \leq \|A\|\|\vec{x}\| = \|A\|$

Für jeden Eigenwert gilt: $|\lambda| \leq \|A\|$
 $\Rightarrow \varrho(T) \leq \|A\|$

24 Beweis (Abschnitt 2.3.3):

Induktion in n :

$n=1$: offensichtlich

$n \geq 1$: Wir entwickeln $\det(A)$ nach der ersten Spalte:

$$\det(A) = \sum_{i=1}^n (-1)^i a_{i1} \det(A_{i1}) \neq 0$$

$\Rightarrow \exists k : a_{k1} \neq 0$ und $\det A_{k1} \neq 0$

Vertausche die erste und k -te Zeile!

$$PA = \begin{pmatrix} * & & & \\ & * & * & * \\ & * & PA & * \\ & * & \text{reg.} & * \end{pmatrix}$$

Die Aussage gilt für $PA_{ki} \in \mathbb{R}^{(n-1) \times (n-1)}$ per Induktionsannahme.

25 Beweis (Abschnitt 2.3.5):

$\{\vec{d}^{(0)}, \vec{d}^{(1)}, \dots, \vec{d}^{(n-1)}\}$ bilden eine Basis des Raums \mathbb{R}^n .

$\vec{e}^{(0)} = \sum_{j=0}^{n-1} \delta_j \vec{d}^{(j)}$ mit δ_j wie folgt:

$$\langle \vec{d}^{(k)}, \vec{e}^{(0)} \rangle_A = \sum_{j=0}^{n-1} \delta_j \langle \vec{d}^{(k)}, \vec{d}^{(j)} \rangle_A = \delta_k \langle \vec{d}^{(k)}, \vec{d}^{(k)} \rangle_A$$

$$\delta_k = \frac{\langle \vec{d}^{(k)}, \vec{e}^{(0)} \rangle_A}{\langle \vec{d}^{(k)}, \vec{d}^{(k)} \rangle_A} = \frac{\langle \vec{d}^{(k)}, \vec{e}^{(0)} + \sum_{i=0}^{k-1} \lambda_i \vec{d}^{(i)} \rangle_A}{\langle \vec{d}^{(k)}, \vec{d}^{(k)} \rangle_A} = \frac{\langle \vec{d}^{(k)}, \vec{e}^{(k)} \rangle_A}{\langle \vec{d}^{(k)}, \vec{d}^{(k)} \rangle_A}$$

$$\vec{e}^{(0)} = \vec{x}^{(0)} - \vec{x}^*$$

$$\vec{e}^{(k)} = \vec{x}^{(k)} - \vec{x}^*$$

$$\vec{x}^{(k)} = \vec{x}^{(0)} + \lambda_0 \vec{d}^{(0)} + \dots + \lambda_{k+1} \vec{d}^{(k+1)}$$

$$\delta_k = \frac{\langle \vec{d}^{(k)}, \vec{e}^{(k)} \rangle_A}{\langle \vec{d}^{(k)}, \vec{d}^{(k)} \rangle_A} = \frac{(\vec{d}^{(k)})^T \overbrace{A \vec{e}^{(k)}}^{\vec{r}^{(k)}}}{(\vec{d}^{(k)})^T A \vec{d}^{(k)}} = -\lambda_k$$

$$\vec{e}^{(0)} = - \sum_{k=0}^{n-1} \lambda_k \vec{d}^{(k)}$$

$$\vec{e}^{(k)} = \vec{e}^{(0)} + \sum_{i=0}^{k-1} \lambda_i \vec{d}^{(i)} = - \sum_{i=0}^{n-1} \lambda_i \vec{d}^{(i)} + \sum_{i=0}^{k-1} \lambda_i \vec{d}^{(i)}$$

$$\vec{e}^{(k)} = - \sum_{i=k}^{n-1} \lambda_i \vec{d}^{(i)}$$

Insbesondere gilt $\vec{e}^{(n)} = \vec{0}$

26 Beweis (Abschnitt 2.3.5):

$$\begin{aligned}\vec{z} &= \vec{e}^{(0)} + \sum_{j=0}^{i-1} \delta_j \vec{d}^{(j)} \\ &= - \sum_{j=0}^{n-1} \lambda_j \vec{d}^{(j)} + \sum_{j=0}^{i-1} \delta_j \vec{d}^{(j)} \\ &= - \sum_{j=i}^{n-1} \lambda_j \vec{d}^{(j)} - \sum_{j=0}^{i-1} (\lambda_j - \delta_j) \vec{d}^{(j)} \\ \|\vec{z}\|_A^2 &= \sum_{j=i}^{n-1} \lambda_j^2 + \sum_{j=0}^{i-1} (\lambda_j - \delta_j)^2 \geq \sum_{j=i}^{n-1} \lambda_j^2 = \|\vec{e}^{(i)}\|_A^2\end{aligned}$$

27 Beweis (Abschnitt 2.3.5):

$$\begin{aligned}\vec{r}^{(j)} &= A\vec{e}^{(j)} = A \left(- \sum_{k=j}^{n-1} \lambda_k \vec{d}^{(k)} \right) = - \sum_{k=j}^{n-1} \lambda_k A \vec{d}^{(k)} \\ (\vec{r}^{(j)})^T \vec{d}^{(i)} &= - \sum_{k=j}^{n-1} \lambda_k (A \vec{d}^{(k)})^T \vec{d}^{(i)} = - \lambda_k \sum_{k=j}^{n-1} (\vec{d}^{(k)})^T A \vec{d}^{(i)} = 0\end{aligned}$$

28 Beweis (Abschnitt 2.3.5):

$$\begin{aligned}\vec{r}^{(k-1)} &= \vec{r}^{(k)} + \lambda_k A \vec{d}^{(k)} \\ (\vec{r}^{(i)})^T \vec{r}^{(k+1)} &= (\vec{r}^{(i)})^T \vec{r}^{(k)} + \lambda_k (\vec{r}^{(i)})^T A \vec{d}^{(k)} \\ (\vec{r}^{(i)})^T \vec{r}^{(j)} &= 0 \text{ für } j = 0, 1, \dots, i-1\end{aligned}$$

$$\text{weil } \vec{r}^{(i)} = - \sum_{k=i}^{n-1} \lambda_k A \vec{d}^{(k)}$$

$$\text{und } (\vec{r}^{(i)})^T \vec{r}^{(j)} = - \sum_{k=i}^{n-1} \lambda_k (\vec{d}^{(k)})^T A \vec{r}^{(j)}$$

$$\Rightarrow (\vec{r}^{(i)})^T A \vec{d}^{(k)} = 0 \Rightarrow \beta_{ij} = 0$$

$$\vec{d}^{(i)} = -\vec{r}^{(i)} - \frac{(\vec{r}^{(i)})^T A \vec{d}^{(i-1)}}{(\vec{d}^{(i-1)})^T A \vec{d}^{(i-1)}} \vec{d}^{(i-1)}$$

Dabei gilt $(\vec{r}^{(j)})^T \vec{r}^{(j)} = 0$, $i = 0, 1, \dots, j-1$

$$\vec{d}^{(i)} = -\vec{r}^{(i)} + \sum_{k=0}^{i-1} \beta_{ik} \vec{d}^{(k)}$$

$$\underbrace{(\vec{r}^{(j)})^T \vec{d}^{(i)}}_{=0} = -(\vec{r}^{(j)})^T \vec{r}^{(i)} + \sum_{k=0}^{i-1} \beta_{ik} \underbrace{(\vec{r}^{(j)})^T \vec{d}^{(k)}}_{=0}$$

$$\Rightarrow (\vec{r}^{(j)})^T \vec{r}^{(i)} = 0$$

$$\vec{d}^{(0)} = -\vec{r}^{(0)}$$

$$\vec{d}^{(i)} = -\vec{r}^{(i)} - \frac{(\vec{r}^{(i)})^T A \vec{d}^{(i-1)}}{(\vec{d}^{(i-1)})^T A \vec{d}^{(i-1)}} \vec{d}^{(i-1)}, \quad i = 1, 2, \dots, n-1$$

29 Beweis (Abschnitt 2.3.5):

$$\vec{e}^{(i)} \in \vec{e}^{(0)} + \text{span}\{\vec{r}^{(0)}, A\vec{r}^{(0)}, \dots, A^{i-1}\vec{r}^{(0)}\}$$

$$= \vec{e}^{(0)} + \text{span}\{A\vec{e}^{(0)}, A^2\vec{e}^{(0)}, \dots, A^i\vec{e}^{(0)}\}$$

$$\Rightarrow \vec{e}^{(i)} = \left(I + \sum_{j=1}^i \psi_j A^j \right) \vec{e}^{(0)} \text{ mit irgendwelchen Koeffizienten } \psi_j$$

$$\vec{e}^{(i)} = P_i(A) \vec{e}^{(0)} \text{ mit einem Polynom } P_i \text{ vom Grad } i \text{ mit } P_i(0) = 1$$

Seien $\lambda_1, \dots, \lambda_n > 0$ die Eigenwerte und $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ Orthonormale Eigenvektoren.

$$\vec{e}^{(0)} = \sum_{j=1}^n \xi_j \vec{v}_j \text{ (einer Linearkombination von } \vec{v}_1, \dots, \vec{v}_n)$$

$$\begin{aligned}\vec{e}^{(i)} &= P_i(A)\vec{e}^{(0)} = \sum_{j=1}^n \xi_j P_i(A)\vec{v}_j \\ &= \sum_{j=1}^n \xi_j P_i(\lambda_j)\vec{v}_j\end{aligned}$$

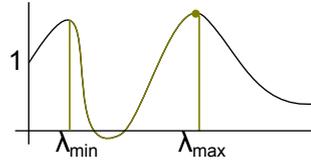
$$\|\vec{e}^{(0)}\|_A^2 = \sum \xi_j^2$$

$$\|\vec{e}^{(i)}\|_A^2 = \sum \xi_j^2 |P_i(\lambda_j)|^2 \leq \max_{\lambda_j} |P_i(\lambda_j)|^2 \cdot \|\vec{e}^{(0)}\|_A^2$$

$$\|\vec{e}^{(i)}\|_A \leq \max_{\lambda_j} |P_i(\lambda_j)| \cdot \|\vec{e}^{(0)}\|_A$$

$$\|\vec{e}^{(i)}\|_A \leq \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |P_i(\lambda)| \cdot \|\vec{e}^{(0)}\|_A$$

$$\|\vec{e}^{(i)}\|_A \leq \min_{\substack{P_i \text{ v. grad } i \\ \text{mit } P_i(0)=1}} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |P_i(\lambda)| \|\vec{e}^{(0)}\|_A \quad \text{Chebyshev-Polynome 1. Art.}$$



$$T_i(t) = \cos(i \arccos(t))$$

$$= \frac{1}{2}((t + \sqrt{t^2 - 1})^i + (t - \sqrt{t^2 - 1})^i), \quad t \in [-1, 2]$$

i Nullstellen im Intervall $(-1, 1)$

$i + 1$ Extrema im Intervall $[-1, 1]$

$$T_i(\cos(\frac{k\pi}{i})) = (-1)^k, \quad k = 0, 1, \dots, i$$

Problem bei $|P_i(\lambda)|$: Min wird erreicht für

$$P_i^*(\lambda) = \frac{T_i(\frac{\lambda_{\max} + \lambda_{\min} - 2\lambda}{\lambda_{\max} - \lambda_{\min}})}{T_i(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}})}$$

$$P_i^*(0) = 1$$

$$P_i^*(\lambda_{\min}) = \frac{T_i(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}})}{T_i(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}})} = \frac{T_i(1)}{T_i(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}})}$$

$$P_i^*(\lambda_{\max}) = \frac{T_i(\frac{\lambda_{\min} - \lambda_{\max}}{\lambda_{\max} - \lambda_{\min}})}{T_i(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}})} = \frac{T_i(-1)}{T_i(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}})}$$

P_i^* löst die Minimierungsaufgabe.

Beweis: Sei Q_i ein Polynom Grad i mit $Q_i(0) = 1$

$$\text{und } \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |q_i(\lambda)| < \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |P_i(\lambda)|$$

$$R_i(\lambda) = P_i(\lambda) - Q_i(\lambda)$$

$$R_i(0) = 1 - 1 = 0$$

und $R_i(\lambda)$ hat i Nullstellen im Intervall $[\lambda_{\min}, \lambda_{\max}]$

R_i hat $i + 1$ verschiedene Nullstellen \Rightarrow Widerspruch.

$$\min_{\substack{P_i \text{ v. grad } i \\ P_i(0)=1}} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |P_i(\lambda)| = A$$

$$= \frac{1}{T_i(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}})} = (T_i(\frac{\text{cond}_2(A)+1}{\text{cond}_2(A)-1}))^{-1}$$

$$= 2 \left((t + \sqrt{t^2 - 1})^i + (t - \sqrt{t^2 - 1})^i \right)^{-1}$$

$$\text{mit } t = \frac{\text{cond}_2(A)+1}{\text{cond}_2(A)-1}$$

$$\begin{aligned}t + \sqrt{t^2 - 1} &= \frac{\text{cond}_2(A)+1}{\text{cond}_2(A)-1} + \frac{\sqrt{(\text{cond}_2(A)+1)^2 - (\text{cond}_2(A)-1)^2}}{\text{cond}_2(A)-1} \\ &= \frac{\text{cond}_2(A)+1 + 2\sqrt{\text{cond}_2(A)}}{\text{cond}_2(A)-1} = \frac{(\text{cond}_2(A)+1)^2}{(\sqrt{\text{cond}_2(A)+1})(\sqrt{\text{cond}_2(A)-1})} = \frac{\sqrt{\text{cond}_2(A)+1}}{\sqrt{\text{cond}_2(A)-1}}\end{aligned}$$

$$t - \sqrt{t^2 - 1} = \frac{\sqrt{\text{cond}_2(A)+1}}{\sqrt{\text{cond}_2(A)-1}}$$

$$A = 2 \left(\left(\frac{\sqrt{\text{cond}_2(A)-1}}{\sqrt{\text{cond}_2(A)+1}} \right)^i + \left(\frac{\sqrt{\text{cond}_2(A)+1}}{\sqrt{\text{cond}_2(A)-1}} \right)^i \right)^{-1} \leq 2 \left(\left(\frac{\sqrt{\text{cond}_2(A)+1}}{\sqrt{\text{cond}_2(A)-1}} \right)^i \right)^{-1} = 2 \left(\frac{\sqrt{\text{cond}_2(A)-1}}{\sqrt{\text{cond}_2(A)+1}} \right)^i$$

$$\Rightarrow \|\vec{e}^{(i)}\|_A \leq 2 \left(\frac{\sqrt{\text{cond}_2(A)-1}}{\sqrt{\text{cond}_2(A)+1}} \right)^i \|\vec{e}^{(0)}\|$$

Sei λ ein Eigenwert von $M^{-1}A$ mit dem Eigenvektor \vec{v} .

Dann ist λ auch ein Eigenwert von $H^{-1}AH^{-T}$ mit dem Eigenvektor $H^T\vec{v}$.
 $(H^{-1}AH^{-T})(H^T\vec{v}) = H^{-1}A\vec{v} = H^T \underbrace{H^{-T}H^{-1}}_{M^{-1}} A\vec{v} = H^T \underbrace{M^{-1}A\vec{v}}_{\lambda\vec{v}} = \lambda(H^T\vec{v})$

Statt $A\vec{x} = \vec{b}$ lösen wir $H^{-1}AH^{-T}\vec{y} = H^{-1}\vec{b}$ und dann $H^T\vec{x} = \vec{y}$.

31 Beweis (Abschnitt 3.1):

1),2) $p_{n,k}(x) > 0$, $x \in (0,1) \Rightarrow$ offensichtlich.

$$\begin{aligned} 3) \quad p_{n,k}(x) &= \binom{n}{k} x^k (1-x)^{n-k} \\ p'_{n,k}(x) &= \binom{n}{k} k x^{k-1} (1-x)^{n-k} - \binom{n}{k} x^k (n-k) (1-x)^{n-k-1} \\ &= \binom{n}{k} x^{k-1} (1-x)^{n-k-1} (k(1-x) - (n-k)x) \\ &= \binom{n}{k} x^{k-1} (1-x)^{n-k-1} (k-nx) \begin{cases} = 0 & , x = \frac{k}{n} \\ > 0 & , x \in (0, \frac{k}{n}) \\ < 0 & , x \in (\frac{k}{n}, 1) \end{cases} \end{aligned}$$

4) Lineare Unabhängigkeit

$$\sum_{k=0}^n \alpha_k p_{n,k}(x) \equiv 0$$

Zu zeigen: $\alpha_k = 0$, $k = 0, 1, \dots, n$

$$\begin{aligned} \sum_{k=0}^n \alpha_k \binom{n}{k} x^k (1-x)^{n-k} &= 0, \quad x \in (0,1) \quad | : (1-x)^n \\ \sum_{k=0}^n \alpha_k \binom{n}{k} \left(\frac{x}{1-x}\right)^k &= 0, \quad x \in (0,1) \end{aligned}$$

Variablentransformation: $s = \frac{x}{x-1} \Rightarrow s \in (0, \infty)$

$$\sum_{k=0}^n \alpha_k \binom{n}{k} s^k = 0, \quad s \in (0, \infty)$$

$\{s^k\}_{k=0}^n$ sind linear unabhängig.

$\Rightarrow \alpha_k = 0$, $k = 0, 1, \dots, n$

$\{p_{n,k}\}_{k=0}^n$ sind $n+1$ linear unabhängige Polynome vom Grad n , $\dim \mathbb{P} = n+1$

$\Rightarrow \{p_{n,k}\}_{k=0}^n$ bilden eine Basis.

32 Beweis (Abschnitt 3.1):

$$1) \quad \sum_{k=0}^n p_{n,k}(x) = \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = (x + (1-x))^n = 1$$

$$\begin{aligned} 2) \quad \sum_{k=0}^n \frac{k}{n} p_{n,k}(x) &= \sum_{k=0}^n \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} \\ &= \sum_{k=0}^n \frac{k}{n} \cdot \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k} \\ &= \sum_{k=1}^n \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} x^{k-1} (1-x)^{(n-1)-(k-1)} \\ &= x \sum_{k=1}^n \binom{n-1}{k-1} x^{k-1} (1-x)^{(n-1)-(k-1)} \\ &= x \sum_{k=0}^{n-1} \binom{n-1}{k} x^k (1-x)^{(n-1)-k} = x \sum_{k=0}^{n-1} p_{n-1,k}(x) = x \end{aligned}$$

$$\begin{aligned} 3) \quad \sum_{k=0}^n \frac{k(k-1)}{n(n-1)} p_{n,k}(x) &= x^2 \\ &= \sum_{k=0}^n \frac{k(k-1)}{n(n-1)} \cdot \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k} \\ &= \sum_{k=2}^n \frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} x^{k-2} (1-x)^{(n-2)-(k-2)} \\ &= x^2 \sum_{k=0}^{n-2} \binom{n-2}{k} x^k (1-x)^{(n-2)-k} = x^2 \sum_{k=0}^{n-2} p_{n-2,k}(x) = x^2 \end{aligned}$$

$$\begin{aligned} \left(\frac{k}{n}\right)^2 &= \frac{1}{n^2} (k(k-1) + k) = \frac{n-1}{n} \cdot \frac{k(k-1)}{n(n-1)} + \frac{1}{n} \cdot \frac{k}{n} \\ \sum_{k=0}^n \left(\frac{k}{n}\right)^2 p_{n,m}(x) \dots &= \frac{n-1}{n} \sum_{k=0}^n \frac{k(k-1)}{n(n-1)} p_{n,k}(x) + \frac{1}{n} \sum_{k=0}^n \frac{k}{n} p_{n,k}(x) = \frac{n-1}{n} x^2 + \frac{1}{n} x \end{aligned}$$

33 Beweis (Abschnitt 3.1):

Bezeichnung: $M = \|f\|$.

$\varepsilon > 0$ fest, $x \in [0, 1]$ fest.

f ist stetig auf $[0, 1]$.

$\Rightarrow f$ ist gleichmäßig stetig.

$\Rightarrow \exists \delta > 0 : |f(t_1) - f(t_2)| < \frac{\varepsilon}{6}$

Für alle $t_1, t_2, \in [0, 1]$ mit $|t_1 - t_2| < \delta$

$$\begin{aligned} |(L_n f)(x) - f(x)| &= |(L_n f)(x) - f(x)(L_n e_0)(x) + f(x)(L_n e_0)(x) - f(x)e_0(x)| \\ &\leq |(L_n f)(x) - f(x)(L_n e_0)(x)| + |f(x)| \cdot |(L_n e_0)(x) - e_0(x)| \\ &= |L_n[f(\cdot) - f(x)e_0(\cdot)](x)| + M|(L_n e_0)(x) - e_0(x)| \\ &\leq L_n(|f(\cdot) - f(x)|)(x) + M\|L_n e_0 - e_0\| \end{aligned}$$

34 Beweis (Abschnitt 3.2):

$$p(t) = \sum_{k=0}^n \beta_k p_{n,k}(t)$$

Für ein festes t nehmen wir $\lambda_k = p_{n,k}(t)$

Für $t \in [0, 1]$ gilt: $p_{n,k}(t) \geq 0$ und $\sum_{k=0}^n p_{n,k} = 1$

35 Beweis (Abschnitt 3.2):

$$p_{n,k}(t) = \binom{n}{k} t^k (1-t)^{n-k}$$

$$p_{n,k}(0) = \begin{cases} 1 & , k = 0 \\ 0 & , k = 1, \dots, n \end{cases}$$

$$p(0) = \sum_{k=0}^n \beta_k p_{n,k}(0) = \beta_0, \quad t = 1 \text{ entsprechend.}$$

36 Beweis (Abschnitt 3.2):

$$p^{(\nu)}(t) = \frac{n!}{(n-\nu)!} \sum_{k=0}^{n-\nu} \Delta^\nu \beta_k p_{n-\nu,k}(t)$$

$$p'(0) = n \sum_{k=0}^{n-1} \Delta \beta_k p_{n-1,k}(0) = n \Delta \beta_0 = n(\beta_1 - \beta_0)$$

$t=1$ entsprechend.

37 Beweis (Abschnitt 3.4):

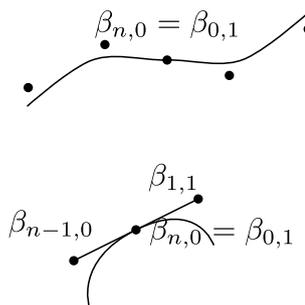
$$\begin{aligned} &\frac{n+1-k}{n+1} \binom{n+1}{k} t^k (1-t)^{n+1-k} + \frac{k+1}{n+1} \binom{n+1}{k+1} t^{k+1} (1-t)^{(n+1)-(k+1)} \\ &= \frac{(n+1-k)(n+1)!}{(n+1)k!(n+1-k)!} t^k (1-t)^{n-k} (1-t) + \frac{(k+1)(n+1)!}{(n+1)(k+1)!(n-k)!} t^{k+1} (1-t)^{n-k} t \\ &= \frac{n!}{k!(n-k)!} t^k (n-t)^{n-k} (1-t) + \frac{n!}{k!(n-k)!} t^{k+1} (1-t)^{n-k} t \\ &= p_{n,k}(t)(1-t+t) = p_{n,k}(t) \end{aligned}$$

38 Beweis (Abschnitt 3.4):

$$\begin{aligned}
\sum_{k=0}^n \beta_k p_{n,k} &= \sum_{k=0}^n \beta_k \left(\left(1 - \frac{k}{n+1}\right) p_{n+1,k} + \frac{k+1}{n+1} p_{n+1,k+1} \right) \\
&= \sum_{k=0}^n \beta_k \left(1 - \frac{k}{n+1}\right) p_{n+1,k}(t) + \sum_{k=0}^n \beta_k \frac{k+1}{n+1} p_{n+1,k+1}(t) \\
&= \sum_{k=0}^n \beta_k \left(t - \frac{k}{n+1}\right) p_{n+1,k} + \sum_{k=1}^{n+1} \beta_{k-1} \frac{k}{n+1} p_{n+1,k} \\
&= \beta_0 p_{n+1,0} + \sum_{k=1}^n \left(\beta_k \left(1 - \frac{k}{n+1}\right) + \beta_{k-1} \frac{k}{n+1} \right) p_{n+1,k} + \beta_n p_{n+1,n+1} \\
&= \sum_{k=0}^{n+1} \hat{\beta}_k p_{n+1,k}
\end{aligned}$$

39 Beweis (Abschnitt 3.5):

$$\begin{aligned}
s \in C &: \beta_{n,0} = \beta_{0,1} \\
s \in C^1 &: s(1-) = s(1+) \Leftrightarrow \beta_{n,0} = \beta_{0,1} \\
s'(1-) &= s'(1-) \\
\text{Für } t \in [0, 1): & \\
s'(t) &= n \sum_{k=0}^{n-1} \Delta \beta_{k,0} p_{n-1,k}(t) \\
s'(t-) &= n \Delta \beta_{n-1,0} \\
\text{Für } t \in [1, 2): & \\
s'(t) &= n \sum_{k=0}^{n-1} \Delta \beta_{k,1} p_{n-1,k}(t+1) \\
s'(1+) &= n \Delta \beta_{0,1} \\
s \in C^1 &\Leftrightarrow \beta_{n,0} = \beta_{0,1} \\
&\Delta \beta_{n-1,0} = \Delta \beta_{0,1}
\end{aligned}$$

**40 Beweis** (Abschnitt 3.6):

$$\begin{aligned}
d_\nu &= 2\beta_{\nu,1} - \beta_{\nu,2} \\
d_k &= 2\beta_{k-1,2} - \beta_{k-1} , \quad \nu = 0, \dots, k-1 \\
\text{Sind die Anschlussbedingungen, die die Stetigkeit der 2. Ableitungen in } (0, k) &\text{ sichern, gleichwertig zu}
\end{aligned}$$

$$d_\nu = 2\beta_{\nu-1,2} - \beta_{\nu-1,1} , \quad \nu = 1, \dots, k-1.$$

Einerseits gilt dann $d_{\nu-1} + 2d_\nu = 3\beta_{\nu-1,2}$, $\nu = 1, \dots, k$
 und andererseits $2d_\nu + d_{\nu+1} = 3\beta_{\nu,1}$, $\nu = 0, \dots, k-1$
 Durch Addition folgt:

$$d_{\nu-1} + 4d_\nu + d_{\nu+1} = 3(\beta_{\nu-1,2} + \beta_{\nu,1}) , \quad \nu = 1, \dots, k-1$$

stetigkeit der 1. Ableitung ergibt:

$$\beta_{\nu-1,2} + \beta_{\nu,1} = \beta_{\nu-1,3} + \beta_{\nu,0}$$

also:

$$d_{\nu-1} + 4d_\nu + d_{\nu+1} = 3(\beta_{\nu,3} + \beta_{\nu,0}) , \quad \nu = 1, \dots, k-1$$

Mi den Stetigkeitsbedingungen $\beta_{\nu-1,3} = \beta_{\nu,0}$, $\nu = 1, \dots, k-1$ folgt schließlich:

$$d_{\nu-1} + 4d_\nu + d_{\nu+1} = 6\beta_{\nu,0} , \quad \nu = 1, \dots, k-1$$

Die Anschlussbed. führen also auf das Glng. System:

$$\begin{pmatrix} 2 & 1 & & & 0 \\ 1 & 4 & 1 & & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & 4 & 1 \\ 0 & & & & 1 & 2 \end{pmatrix}$$

Man sieht leicht, dass das Gleichungssystem diagonaldominant und folglich eindeutig lösbar ist.

Zu gegebenen $\beta_{0,1}, \beta_{1,0}, \dots, \beta_{k-1,0}, \beta_{k-1,2}$ erhält man Hilfsgrößen d_0, \dots, d_k . Daraus ergibt sich dann die übrige Bézier-Koeffizienten aus

$$\beta_{\nu-1,2} = \frac{1}{3}(d_{\nu-1} + 2d_{\nu}), \quad \nu = 1, \dots, k$$

$$\beta_{\nu,1} = \frac{1}{3}(2d_{\nu} + d_{\nu+1}), \quad \nu = 0, \dots, k-1$$

Außerdem kann man $\beta_{0,0}$ und $\beta_{k-1,3}$ frei wählen, da sie von d_0, \dots, d_k unabhängig sind.

41 Beweis (Abschnitt 3.6):

1. Angewendet auf $S(t) = 1$, $t \in [0, k]$ ergibt sich mit:

$$S(t) = \sum_{\nu=1}^{k+1} \alpha_{\nu} \beta_{\nu,3}(t) \quad \text{kubische B-Splines, ganzzahlige Knoten}$$

Dann gilt:

$$\alpha_{-1} = 2S(0) + \frac{2}{3}S''(0) - S(1) + \frac{1}{6}S''(1)$$

$$\alpha_{\nu} = S(\nu) - \frac{1}{6}S''(\nu), \quad \nu = 0, \dots, k$$

$$\alpha_{k+1} = 2S(k) + \frac{2}{3}S''(k) - S(k-1) + \frac{1}{6}S''(k-1)$$

$$\Rightarrow 1 = \sum_{\nu=-1}^{k+1} \beta_{\nu,3}(t)$$

2. $S(t) = t$, $t \in [0, k]$, $S''(t) = 0$

$$\alpha_{\nu} = S(\nu) - \frac{1}{6}S''(\nu) = \nu, \quad \nu = 0, \dots, k$$

$$\alpha_{-1} = 2S(0) - S(1) = -1$$

$$\alpha_{k+1} = 2S(k) - S(k-1) = 2k - (k-1) = k+1$$

$$\Rightarrow t = \sum_{\nu=-1}^{k+1} \nu \beta_{\nu,3}(t)$$

3. $S(t) = t^2$, $t \in [0, k]$, $S''(t) = 0$

$$\alpha_{\nu} = S(\nu) - \frac{1}{6}S''(\nu) = t^2 - \frac{1}{3}, \quad \nu = 0, \dots, k$$

$$\alpha_{-1} = 2S(0) + \frac{2}{3}S''(0) - S(1) - \frac{1}{6}S''(1)$$

$$= (-1)^2 + \frac{1}{3}$$

$$\alpha_{k+1} = 2S(k) + \frac{2}{3}S''(k) - S(k-1) + \frac{2}{3}S''(k-1)$$

$$= (k+1)^2 - \frac{1}{3}$$

Stichwortverzeichnis

- θ -Methode, 6
- ., 21
- 3-schrittiges Adams–Moulton–Verfahren, 20
- 5-Punkte–Schema, 28

- Adams–Bashforth, 18
- Adams–Moulton, 18
- Adams–Verfahren, 15
- Alekseev–Gröbner–Lemma, 14
- Algorithmus von de Casteljau, 52
- Anfangsproblem, 1

- Bézie–Polygon, 50
- Bézier–Kurve, 50
- Bézier–Punkte, 50
- Bézier–Splines, 55
- Baum, 30
- Bernstein–Grundpolynome, 44
- Bernstein–Operator, 44
- Bernstein–Polynome, 44
- Bernsteinpolynom, 43
- Butcher–Tableau, 7

- CG–Verfahren, 40
- Cholesky–Vorkonditionierung, 43
- Cholesky–Zerlegung, 29
- Cohn–Schur–Kriterium, 37

- de Boor–Punkte, 59
- DGL n-ter Ordnung, 1
- Differenzgleichung, 21

- Einschrittverfahren, 2–4
- Einzelschrittverfahren, 39
- Erste Dahlquist Schranke, 26
- erstes charakteristisches Polynom, 19
- Euler Cauchy Verfahren, 10
- Eulerisches Polyzugverfahren, 2
- exakte Lösung, 3
- Existenz und Eindeutigkeit der Lösung, 1
- explizite Methode, 5
- Explizite Runge–Kutta–Verfahren, 6
- Expliziter Fall, 15

- Fixpunktiteration, 38

- Gauß–Quotienten, 14
- Gauß–Seidel, 33
- Gauß–Seidel–Verfahren, 39
- Gaußelimination/Cholesky–Zerlegung, 29
- Gesamtschrittverfahren, 38
- gewurzelter Baum, 30
- globale Diskretisierungsfehler, 23

- Globaler Diskretisierungsfehler, 3
- Gradanhebung, 54
- Gram–Schmitt–Orthogonalisierung:, 41
- Graph, 29

- homogen, 21
- Horner–Schema, 53
- Householder–John, 32

- Implizite Methode, 5
- implizite Runge–Kutta–Verfahren, 11
- Impliziter Fall, 15
- inhomogen, 21
- Iterative Verfahren für große LGS, 31

- Jacobi–Verfahren, 38
- Jacobi–Vorkonditionierung, 43
- Jacobi–Verfahren, 33

- Kanten, 29
- klassische Runge–Kutta–Verfahren, 11
- Knoten, 29
- kompatibler Ordnungsvektor, 35
- konsistent, 3, 19
- Konsistenzbedingungen, 20
- Konsistenzordnung, 4, 5
- Konsistenzsordnung, 19
- Kontrollpolygon, 50
- Kontrollpunkte, 50
- konvergent, 3, 23
- Konvergenzordnung, 3, 24
- Konvergenzsatz von Dahlquist, 25
- Konvexe Hülle, 51
- Korovkin, 45
- Kurve in Bézier–Darstellung, 50

- Lösung großer, dünn besetzter, linearer
Gleichungssysteme, 28
- Lineare Differenzgleichungen, 21
- lineare Rekursionen, 21
- lineares Mehrschrittverfahren, 19
- Lipschitz–stetig, 1
- lokale Diskretisierungsfehler, 19
- lokale Diskretisierungsfehler, 3, 19

- Methode der finiten Differenzen, 28
- Milne, 26
- Milne–Simpson, 18

- Näherungslösung, 3
- Newton–Horner–Verfahren, 54
- Nyström, 18

Ordnungsvektor, 35
Orthogonalität, 13

Peano, 1
Picard Lindelöf, 2
polynomialer Spline, 56
polynomieller Spline, 56
positiver linearer Operator, 45
Prädiktor–Korrektor–Verfahren, 27

Relaxation, 39
Residuum, 40

s–stufiges Runge–Kutte–Verfahren, 6
Satz von Kahan, 39
Schrittweitensteuerung, 26
Simson–Regel, 11
SOR, 33
Spektralradius, 38
Spezialfall Euler–Verfahren, 5
Spline–Funktionen, 56
stabil, 25
Standard Matrixiterationen, 31
Stein–Rosenberg, 34
Suchrichtung, 41
Suchrichtungen, 41

Taylor–Verfahren, 5
Testfunktionen, 45
Trapezformel, 10
Trapezregel, 5

Unvollständige Cholesky–Faktorisierung, 43
unvollständige LR–Faktorisierung, 31

verbesserte Eulersche Polygon Verfahren, 11
Verfahren der konjugierten Gradienten, 40
Verfahrensfunktion, 3
Vergleich Quadraturformel, 10
vollständigen Horner–Schemas, 54
Vorkonditionierung, 43

Wahl der Suchrichtung im CG–Verfahren, 42
Wurzel, 30
Wurzelbedingung, 25

Zuwachsfunktion, 3
zweites charakteristisches Polynom, 19