Kapitel 2: Finite Differenzenmethode für parabolische DGL

2.1 Ein Beispiel aus der Wirtschaftsmathematik

$V(t,s)$ = Wert einer europäischen Option zur Zeit t beim Aktienpreis s

erfüllt (BS) $\quad \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} + rs\frac{\partial V}{\partial s} - rV = 0 \qquad s \in (0,\infty) \quad t \in [0,T)$

Black-Scholes Gleichung

σ: Schwankungsmaß (Volatilität) für Aktienkurs

r: Rendite eines festverzinsten Papiers (Bond)

RB in t: „Endbedingung" $\quad V(s,T) = C(s) \qquad s \in (0,\infty)$

RB in s: $\quad V(0,t)$ = Optionspreis bei Pleite des Aktienunternehmens

$\lim_{s\to\infty} \frac{V(s,t)}{C(s)} = 1 \qquad$ Optionspreis bei ∞-hohem Aktienpreis

Für spezielle C (z.B. $C(S) = (S-K)^+$) ist Lösung analytisch bekannt

Für allgemeine C: (z.B. $C(S) = \min\{L, ((S-K)^+)^p\}$ $L>0$, $p>0$: bestimme Lösung numerisch

in beiden Fällen nutze Trafo $V(t,s) \Longleftrightarrow u(\tau, x)$

mit $\tau = \sigma^2 \frac{T-t}{2}$ Zeitumkehr & Skalierung $t \in (0, T_{max}]$

$x = \ln(S/K)$ exponentielle Verzerrung $x \in (-\infty, \infty)$

und $V(t,s) = K \exp\left(\frac{1-k}{2} x - \frac{(1+k)^2}{4} \tau\right) u(x,\tau)$ $k = \frac{2r}{\sigma^2}$

dann (Kettenregel): V löst BS genau dann, wenn u die <u>Diffusionsgleichung</u> löst

$$\partial_t u = \partial_x^2 u \qquad t \in (0, T_{max}], \quad x \in (-\infty, \infty)$$

mit den umgerechneten AW, RB

vor der Diskretisierung: $x \in (-\infty, \infty) \rightsquigarrow x \in [-A, A]$ endliches Gebiet

erfordert <u>künstliche</u> RB bei $x \in \{-A, A\}$

denn wahre RB bei $x \in \{-\infty, \infty\}$

Einfache (nicht sehr genaue) Approximation: ziehe RB von $\pm\infty$ nach $\pm A$

ergibt <u>Modellproblem</u>

$$\begin{cases} \partial_t u = \partial_x^2 u & (t,x) \in (0, T_{max}] \times (-A, A) \\ u_{|t=0} = \bar{u} & x \in [-A, A] \\ u = g & (t,x) \in (0, T_{max}] \times \{-A, A\} \end{cases}$$

2.2. Diskretisierung mit FDM und Linienmethode

Linienmethode: üblicher Zugang bei zeitabhängigen Problemen

Differentialoperator (DOP)

$$\partial_t u = L u$$

Schritt 1: diskretisiere alle Variablen außer t

Koordinaten der endlich-dimensionalen Approximation

$$\partial_t u_h = L_h u_h$$

diskrete Approximation des DOP

System gew. DGL

(z.B. Matrix wenn L linear)

Schritt 2: diskretisiere t - Variable (z.B. mit RK oder Mehrschritt Verfahren)

Durchführung im Modellproblem ...

Schritt 1: "räumliche" Diskretisierung

$$x_i = -A + i \cdot h \qquad i \in \mathbb{Z}, \quad h = \frac{2A}{M}, \quad M \in \mathbb{N}_{\geq 2}$$

$$\partial_x^2 u(t, x_i) = \frac{u(t, x_{i-1}) - 2u(t, x_i) + u(t, x_{i+1})}{h^2} + R_{i,h}$$

$|R_{i,h}| \leq C h^2$ wenn $u \in C^4$

mit $\tilde{u}(t) = \begin{bmatrix} \tilde{u}_1(t) \\ \vdots \\ \tilde{u}_{M-1}(t) \end{bmatrix}$ und $\tilde{u}_i \approx u(t, x_i)$

ergibt sich FD-Approximation: $\begin{cases} \frac{d}{dt}\tilde{u}(t) = A\, u(t) + b(t) \\ \tilde{u}(0) = \overline{u} \end{cases}$ mit $\overline{u}_i = \overline{u}(x_i)$

wobei $A = \frac{1}{h^2} \begin{bmatrix} -2 & +1 & & & & \\ +1 & -2 & +1 & & & \\ & +1 & -2 & +1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & +1 & -2 & +1 \\ & & & & +1 & -2 \end{bmatrix} \in \mathbb{R}^{(M-1)\times(M-1)}$ $\quad b(t) = \frac{1}{h^2} \begin{bmatrix} -g(t, -A) \\ 0 \\ \vdots \\ 0 \\ -g(t, +A) \end{bmatrix} \in \mathbb{R}^{M-1}$

Schritt 2: Diskretisierung der gew. DGL

beachte: Eigenwertstruktur von A beeinflusst Wahl der Diskretisierung

1) A ist symmetrisch $\Rightarrow$ $\sigma(A) \subset \mathbb{R}$

2) Lage der EW mit Gerschgorin ...

$K_i := \{ z \in \mathbb{C} \mid |z - A_{ii}| \leq \sum_{j \neq i} |A_{ij}| \}$ $\quad$ i-ter Gerschgorin Kreis von A

wir wissen: $\sigma(A) \subset \bigcup_{i=1}^{N} K_i =: G$

hier

$$\begin{bmatrix} -2/h^2 & 1/h^2 & & \\ 1/h^2 & -2/h^2 & 1/h^2 & \\ & \ddots & \ddots & \ddots \\ & & 1/h^2 & -2/h^2 \end{bmatrix} \begin{array}{l} \longrightarrow K_1 = \overline{B_{-2/h^2}(1/h^2)} \\ \Longrightarrow K_2, \dots, K_{N-1} = \overline{B_{-2/h^2}(2/h^2)} \\ \\ \longrightarrow K_N = K_1 \end{array}$$

$\leadsto$ $G = \overline{B_{-2/h^2}(2/h^2)}$ also $\sigma(A) \subset G \cap \mathbb{R} = [-4/h^2, 0]$

Wir wissen außerdem (Argumente wie in 1.5): $\operatorname{Kern} A = \{0\}$ also $0 \in \sigma(A)$

insgesamt $\sigma(A) \subset [-4/h^2, 0)$ d.h. <u>A negativ definit</u>

konkreter Eigenvektor: $v_i = \cos(\alpha x_i)$ $\alpha = \begin{cases} (M-1)\frac{\pi}{2A} & M \text{ gerade} \\ M \frac{\pi}{2A} & M \text{ ungerade} \end{cases}$

dann $v_0 := \cos(\alpha(-A)) = \cos(\alpha A) =: v_M = \cos\left(\frac{\text{ungerade}}{2}\pi\right) = 0$

also $-2v_1 + v_2 = v_0 - 2v_1 + v_2$

und $v_{M-2} - 2v_M = v_{M-2} - 2v_{M-1} + v_M$

allgemein gilt $v_{i-1} - 2v_i + v_{i+1} = \cos(\alpha x_i - \alpha h) - 2\cos(\alpha x_i) + \cos(\alpha x_i + \alpha h)$

Additionstheoreme $= \cos(\alpha x_i)\cos(\alpha h) + \sin(\alpha x_i)\sin(\alpha h) - 2\cos(\alpha x_i) + \cos(\alpha x_i)\cos(\alpha h) - \sin(\alpha x_i)\sin(\alpha h)$

$= 2(\cos(\alpha h) - 1)\, v_i$ also $(v_1, \ldots, v_{M-1})^T$ EV von A mit EW

$$\lambda = \frac{2}{h^2}(\cos(\alpha h) - 1) \underset{h = \frac{2A}{M}}{=} \frac{2}{h^2}\begin{cases} -2 & M \text{ ungerade} \\ -1 - \cos\left(\frac{\pi}{M}\right) & M \text{ gerade} \end{cases}$$

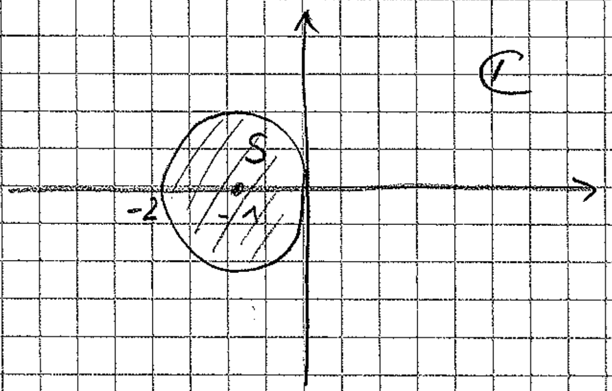
Für großes M : $\cos\left(\frac{\pi}{M}\right) \approx 1$

also $\lambda \approx -\frac{M^2}{A^2}$ sehr groß!

Stabilitätsbedingung zu Einschrittverfahren für $\frac{d}{dt}\tilde{u} = A\tilde{u}$:

$\Delta t \cdot \lambda \in S$ ← Stabilitätsgebiet des Verfahrens

explizites Euler-Verfahren



notw. Stabilitätsbedingung $\Delta t \frac{M^2}{A^2} \leq 2$

also $\underline{\Delta t \leq C h^2}$

solche Stabilitätsbedingungen gibt es für <u>jedes explizite</u> Verfahren (da S kompakt)

beachte Bedingung sehr ungünstig: Halbierung von h ⇝ <u>Vervierfachung</u> der benötigten Zeitschritte benutze deshalb <u>implizite</u>, A-stabile Verfahren ($\mathbb{C}^- \subset S$)

Beispiel: Θ-Verfahren für $\dot{u}(t) = F(t, u(t))$ $\quad u^n \approx u(t_n), \quad t_n = n\Delta t, \quad n \in \mathbb{N}_0$

$$\frac{u^{n+1} - u^n}{\Delta t} = \underbrace{\Theta F(t_{n+1}, u^{n+1}) + (1-\Theta) F(t_n, u^n)}$$

$\Theta \in [0,1]$: Konvexkombination aus explizitem und implizitem Euler-Verfahren

$\Theta = \frac{1}{2}$: Crank-Nicolson Verfahren

$\Theta > 0$: A-stabil, implizit

Spezialfall $\quad F(t, u(t)) = A u(t) + b(t)$

$$\underbrace{(I - \Theta \Delta t A)}_{B} u^{n+1} = \underbrace{(I + (1-\Theta)\Delta t A)}_{E} u^n + \underbrace{\Delta t \left(\Theta b(t_{n+1}) - (1-\Theta) b(t_n)\right)}_{d^n}$$

also $\begin{cases} B u^{n+1} = E u^n + d^n & n = 0, \dots, N-1, \quad N \cdot \Delta t = T_{max} \\ u^0 = \bar{u} \end{cases}$

2.3 Konvergenzuntersuchung

Lässt sich $\begin{cases} B u^{n+1} = \underline{E} u^n + d^n \\ u^0 = \bar{u} \end{cases}$ überhaupt durchführen?

D.h. ist B invertierbar?

$B = \underline{I} - \Theta \Delta t A$

$$\sigma(A) \subset \left[-\tfrac{4}{h^2}, 0\right)$$

$$\sigma(\Theta \Delta t A) \subset \left[-\frac{4\Theta\Delta t}{h^2}, 0\right] \swarrow \text{ wegen } \Theta = 0$$

$$\sigma(-\Theta \Delta t A) \subset \left[0, \frac{4\Theta\Delta t}{h^2}\right]$$

$$\sigma(B) \subset \left[1, 1 + \frac{4\Theta\Delta t}{h^2}\right]$$

also $0 \notin \sigma(B)$ also B invertierbar

Praxis: 1 LGS pro Zeitschritt

Theorie: $\begin{cases} u^{n+1} = B^{-1} \underline{E} u^n + B^{-1} d^n \\ u^0 = \bar{u} \end{cases}$

Sei v Lösung von

$$\mathrm{DP}\begin{cases} \partial_t v = \partial_x^2 v & \text{auf } (0, T_{max}] \times (-A, A) \\ v|_{t=0} = \bar{u} & \text{auf } [-A, A] \\ v = g & \text{auf } (0, T_{max}] \times \{-A, A\} \end{cases}$$

$v_i^n := v(t_n, x_i)$ exakte Lösung an den Gitterpunkten

Fehler: $e^n = v^n - u^n$

erfüllt $\mathcal{B} e^{n+1} = \mathcal{B} v^{n+1} - \mathcal{B} u^{n+1} = \mathcal{B} v^{n+1} - E u^n - d^n = E e^n + \tau^n$

$\tau^n = \mathcal{B} v^{n+1} - (E v^n + d^n)$ "Residuum" "lokaler Diskretisierungsfehler"

rekursiv:

$$\begin{aligned} e^n &= \mathcal{B}^{-1} E e^{n-1} + \mathcal{B}^{-1} \tau^{n-1} \\ &= \mathcal{B}^{-1} E \left(\mathcal{B}^{-1} E e^{n-2} + \mathcal{B}^{-1} \tau^{n-2} \right) + \mathcal{B}^{-1} \tau^{n-1} \\ &= (\mathcal{B}^{-1} E)^2 e^{n-2} + (\mathcal{B}^{-1} E)^1 \mathcal{B}^{-1} \tau^{n-2} + (\mathcal{B}^{-1} E)^0 \mathcal{B}^{-1} \tau^{n-1} \\ &= \dots = \underset{\substack{\uparrow \\ \text{Anfangsfehler} \\ (\text{hier } e^0 = 0)}}{(\mathcal{B}^{-1} E)^n e^0} + \underbrace{\sum_{k=0}^{n-1} (\mathcal{B}^{-1} E)^k \mathcal{B}^{-1} \tau^{n-1-k}}_{\text{akkumulierter lokaler Fehler}} \end{aligned}$$

Normabschätzung:

$$\|e^n\|_p \leq \underbrace{\|(B^{-1}E)^n\|}\,\|e^0\|_p + \sum_{k=0}^{n-1} \underbrace{\|(B^{-1}E)^k B^{-1}\|}\,\|r^{n-1-k}\|_p$$

benötige Matrixnormabschätzungen

$p = \infty$: nutze Maximumprinzip wie in Kapitel 1

$p = 2$: nutze Eigenwertinformation

Lemma 2.1: Sei $A \in \mathbb{R}^{n\times n}$. Dann gilt in der Matrixnorm zur euklidischen Norm auf $\mathbb{R}^n$ $\quad \|A\| = \sqrt{\rho(A^TA)}$ wobei ρ den Spektralradius bezeichnet.
Ist A symmetrisch, dann gilt sogar $\|A\| = \rho(A)$

Beweis: A^TA symmetrisch $\Rightarrow \exists\, U \in \mathbb{R}^{n\times n} \quad U^TU = I \;\wedge\; A^TAU = U\Lambda$ mit $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$

$$\|Ax\|_2^2 = \langle Ax, Ax\rangle = \langle A^TAx, x\rangle = \langle U\Lambda U^Tx, x\rangle = \langle \Lambda U^Tx, U^Tx\rangle$$

$$= \sum_{i=1}^n \lambda_i (U^Tx)_i^2 \leq \rho(A^TA)\, \|U^Tx\|_2^2 = \rho(A^TA)\, \|x\|_2^2 \quad \text{also} \quad \|A\| \leq \sqrt{\rho(A^TA)}$$

Ist x EV von A^TA zum EW $\lambda_{max} = \rho(A^TA)$, dann

$$\|Ax\|_2^2 = \langle A^TAx, x\rangle = \lambda_{max}\|x\|_2^2 \quad \text{also} \quad \|A\| \geq \sqrt{\rho(A^TA)}$$

wenn $A = A^T$ dann $\|A\| = \sqrt{\rho(A^TA)} = \sqrt{\rho(A^2)} = \sqrt{\rho(A)^2} = \rho(A)$ $\blacksquare$

Lemma 2.2 Sei A wie in Abschnitt 2.2 konstruiert und sei $\|\cdot\|$ die Matrixnorm zu $\|\cdot\|_2$.
Dann gilt für $B = I - \Delta t\, \Theta A$ dass $\|B^{-1}\| \leq 1$.
Ist entweder $\Theta \geq \frac{1}{2}$ oder $\Theta < \frac{1}{2}$ und $\Delta t \leq \frac{1}{2-4\Theta} h^2$ dann gilt $\|B^{-1}E\| \leq 1$
wobei $E = I + \Delta t (1-\Theta) A$.

Beweis: Wir wissen A ist symmetrisch und $\sigma(A) \subset [-\frac{4}{h^2}, 0)$
dann B symmetrisch und $\sigma(B) \subset [1, 1 + \alpha\Theta]$ $\qquad \alpha := \frac{4\Delta t}{h^2}$
also $\sigma(B^{-1}) \subset \left[\frac{1}{1+\alpha\Theta}, 1\right]$ also $\|B^{-1}\| \leq 1$
auch E symmetrisch und

$$\Theta E = \Theta I + (1-\Theta)\Delta t\, \Theta A = \underbrace{\Theta I + (1-\Theta) I}_{I} - (1-\Theta)\underbrace{(I - \Delta t \Theta A)}_{B}$$

also für $\Theta \neq 0$ $\quad E = \frac{1}{\Theta} I - \frac{1-\Theta}{\Theta} B$

d.h. $\quad B^{-1}E = \frac{1}{\Theta} B^{-1} - \frac{1-\Theta}{\Theta} I$

d.h. $\quad \sigma(B^{-1}E) \subset \left[\frac{1}{\Theta}\frac{1}{1+\alpha\Theta} - \frac{1-\Theta}{\Theta}, \frac{1}{\Theta} - \frac{1-\Theta}{\Theta}\right] = \left[1 - \frac{\alpha}{1+\alpha\Theta}, 1\right]$

Es ist $\quad 1 - \frac{\alpha}{1+\alpha\Theta} \geq -1 \iff \frac{\alpha}{1+\alpha\Theta} \leq 2 \iff \alpha \leq 2 + 2\alpha\Theta \iff \alpha\left(\frac{1}{2} - \Theta\right) \leq 1$

$\iff \Theta \geq \frac{1}{2}$ oder $\Theta < \frac{1}{2} \wedge \alpha \leq \frac{1}{\frac{1}{2} - \Theta}$

d.h. nach Voraussetzung $\sigma(B^{-1}E) \subset [-1,1]$ also $\|B^{-1}E\| \leq 1$ $\blacksquare$

Für die Raum-Zeit Norm $\|z\|_{\infty,2} = \max_{0 \leq n \leq N} \|z^n\|_2$ folgt dann

Lemma 2.3 Sei u_i^n die Lösung des diskreten Problems und v_i^n die exakte Lösung von DP an den Gitterpunkten mit Residuum $\tau^n = Bv^{n+1} - (Ev^n + d^n)$.

Dann gilt für den Fehler $e = v - u$

$$\|e\|_{\infty,2} \leq \|e^0\|_2 + N \|\tau\|'_{\infty,2} \quad \text{mit} \quad \|\tau\|'_{\infty,2} = \max_{0 \leq n \leq N-1} \|\tau^n\|_2$$

Beweis: Wir wissen $e^n = (B^{-1}E)^n e^0 + \sum_{k=0}^{n-1} (B^{-1}E)^k B^{-1} \tau^{n-1-k}$

also $\|e^n\| \leq \|(B^{-1}E)^n\| \, \|e^0\| + \sum_{k=0}^{n-1} \|(B^{-1}E)^k\| \, \|B^{-1}\| \, \|\tau\|'_{\infty,2}$

wegen $\|D^m\| \leq \|D\|^m$ und Lemma 2.2 folgt Beh. $\blacksquare$

Definition 2.4. Sei $\Omega = [0, T_{max}] \times [-A, A]$ und $v : \Omega \to \mathbb{R}$. Sind die partiellen Ableitungen $\partial_t^k \partial_x^l v$ stetig für $0 \le l \le L \in \mathbb{N}_0$ und $0 \le k \le K \in \mathbb{N}_0$, dann sagen wir $v \in C^{K,L}(\Omega)$.

Lemma 2.5. Sei $v \in C^{3,4}([0, T_{max}] \times [-A, A])$ die Lösung von DP und $\theta = \frac{1}{2}$. Dann gilt für das Residuum $\|r\|'_{\infty,2} \le \Delta t \, C \, (h^2 + \Delta t^2)$

Beweis:

$$r^n = B v^{n+1} - (E v^n - d^n)$$
$$= \Delta t \left(\frac{v^{n+1} - v^n}{\Delta t} - \left(\theta (A v^{n+1} + b^{n+1}) + (1-\theta)(A v^n + b^n) \right) \right)$$

Satz von Taylor

$$\left\| \frac{v^{n+1} - v^n}{\Delta t} - \left((\partial_t v)^n + \tfrac{1}{2} \Delta t (\partial_t^2 v)^n \right) \right\|_\infty \le C \Delta t^2$$

enthält Schranke von $\partial_t^3 v$

es ist

$$(A v^n + b^n)_i = (\Delta_h v)_i^n := \frac{v_{i-1}^n - 2 v_i^n + v_{i+1}^n}{h^2} \qquad i = 1, \dots, M-1$$

und

$$\| (\Delta_h v)^n - (\partial_x^2 v)_i^n \| \le C h^2$$

enthält Schranke von $\partial_x^4 v$

$$\| A v^{n+1} + b^{n+1} - \left((\Delta_h v)^n + \Delta t \, \Delta_h (\partial_t v) \right)^n \| \leq C \, \Delta t^2$$

C ← enthält Schranke von $\partial_t^2 v$

$$\| \tau^n \| \leq \Delta t \, \| (\partial_t v)^n + \tfrac{1}{2} \Delta t (\partial_t^2 v)^n - (\Delta_h v)^n - \Theta \Delta t \Delta_h (\partial_t v)^n \| + \Delta t \, C \, \Delta t^2$$

$$\leq \Delta t \, \| \underbrace{(\partial_t v)^n - (\partial_x^2 v)^n}_{=0} + \Delta t \underbrace{\left(\tfrac{1}{2} (\partial_t^2 v)^n - \Theta (\partial_x^2 \partial_t v)^n \right)}_{(*)\, =\, 0 \text{ wenn } \Theta = \frac{1}{2}} \| + \Delta t \, C (\Delta t^2 + h^2)$$

☒

Für $\Theta \neq \frac{1}{2}$ fällt $(*)$ nicht weg .. es genügt t-Entwicklung bis Ordnung 2

Satz 2.6 Sei $\Omega = [0, T_{max}] \times [-A, A]$ und $v \in C^{3,4}(\Omega)$ für $\Theta = \frac{1}{2}$ bzw $v \in C^{2,4}(\Omega)$ für $\Theta \neq \frac{1}{2}$ die exakte Lösung von DP. Ist u die diskrete Lösung, dann gibt es $C > 0$, so dass

$$\| v - u \|_{\infty, 2} \leq C (\Delta t^\gamma + h^2)$$

mit $\gamma = 1$ für $\Theta \neq \frac{1}{2}$ und $\gamma = 2$ für $\Theta = \frac{1}{2}$

Beweis: kombiniere Lemma 2.5 bzw Modifikation für $\Theta \neq \frac{1}{2}$ mit Lemma 2.3 ☒