

9.8. Mehrgitterverfahren

Bei der numerischen Approximation elliptischer Randwertprobleme durch Finite Elemente muss letztendlich ein lineares Gleichungssystem (LGS) der Form

$$(9.1) \quad A\hat{u} = b \quad A = a(\varphi_i, \varphi_j)_{i,j=1}^N \text{ Steifigkeitsmatrix} \\ b \in \mathbb{R}^N \text{ Lastvektor}$$

gelöst werden. Die Matrix $A \in \mathbb{R}^{N \times N}$ ist symmetrisch, positiv definit, groß (z.B. $N = 10^4 - 10^6$), aber dünn besetzt.

Im Prinzip kann (9.1) durch Cholesky-Zerlegung gelöst werden:

$$A = LL^T, \quad Lw = b, \quad L^T \hat{u} = w$$

Problem: $L \in \mathbb{R}^{N \times N}$ ist im Allgemeinen nicht dünn besetzt
 \Rightarrow hoher Aufwand, $O(N^3)$

Ziel: Lösung durch iterative Verfahren

Namens 2: CG-Verfahren, Krylov-Verfahren

Hier: Mehrgitterverfahren

Falls A durch Umnummerierung in die Form $A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$ mit Blockmatrizen A_{11}, A_{21}, A_{22} gebracht werden kann, zerfällt das LGS (9.1) in zwei kleinere LGS.

Wir nehmen daher o.B.d.A. an, dass eine solche Umnummerierung nicht existiert.

3.1 Klassische iterative Verfahren

(a) Grundidee: Zerlege $A = M - N$ so, dass M einfach zu invertieren ist.

$$A\hat{u} = b \iff M\hat{u} = N\hat{u} + b$$

Dies motiviert die Iteration

$$(3.2) \quad M\hat{u}^{k+1} = N\hat{u}^k + b \quad \text{bzw.} \quad \hat{u}^{k+1} = M^{-1}(N\hat{u}^k + b) \\ = \hat{u}^k + M^{-1}(b - A\hat{u}^k) \quad (*)$$

Klar: Falls $(\hat{u}^k)_k$ konvergiert, so gilt $\epsilon^k := \hat{u}^{k+1} - \hat{u}^k \rightarrow 0$
und $\epsilon^k = M^{-1}(b - A\hat{u}^k)$

~~$$A\hat{u} = b \iff M\hat{u} = N\hat{u} + b \iff \lim_{k \rightarrow \infty} (M\hat{u}^{k+1} - N\hat{u}^k) = \lim_{k \rightarrow \infty} (A\hat{u}^k - b)$$~~

$$\Rightarrow \hat{u} - \hat{u}^k = A^{-1}b - \hat{u}^k = A^{-1}(b - A\hat{u}^k) \stackrel{(*)}{=} A^{-1}M\epsilon^k \rightarrow 0$$

Frage: Wann konvergiert die Iteration? Konvergenzgeschwindigkeit?

(3.2) ist eine Fixpunktiteration mit Fixpunkt $\hat{u} = A^{-1}b$ und es gilt (128)

$$\hat{u}^{k+1} - \hat{u} \stackrel{(3.2)}{=} M^{-1}(N\hat{u}^k + b - M\hat{u}) \\ = M^{-1} \underbrace{(N\hat{u}^k + b - M\hat{u})}_{= -N\hat{u}} \\ = M^{-1}N(\hat{u}^k - \hat{u}) \\ = (M^{-1}N)^{k+1}(\hat{u}^0 - \hat{u})$$

Die Iteration konvergiert für beliebige Startwerte \hat{u}^0 also genau dann, wenn für alle Eigenwerte λ von $M^{-1}N$ $|\lambda| < 1$ gilt.

(c) SOR- und SSOR-Verfahren

Führe einen Relaxationsparameter ω in das Gauß-Seidel-Verfahren ein:

$$(9.3) \quad (D - \omega L) \hat{u}^{k+1} = \omega (R \hat{u}^k + b) - (\omega - 1) D \hat{u}^k$$

Für $\omega = 1 \Rightarrow$ Gauß-Seidel

$\omega > 1 \Rightarrow$ Overrelaxation

$\omega < 1 \Rightarrow$ Underrelaxation

Das SOR-Verfahren („successive overrelaxation“) verwendet (9.3) mit $\omega > 1$. Dadurch kann eine Verbesserung der Konvergenzgeschwindigkeit erreicht werden, doch die optimale Wahl von ω ist schwierig.

Bemerkung: Beim Gauß-Seidel- und SOR-Verfahren kann man die Rollen von L und R vertauschen.

Obwohl A symmetrisch ist, ist die Iterationsmatrix beim Jacobi-, Gauß-Seidel- oder SOR-Verfahren in der Regel nicht symmetrisch. Man kann das SOR-Verfahren jedoch symmetrisieren, indem jeder Iterationsschritt in zwei Halbschritte aufgeteilt wird.

Man erhält dann das symmetrische SOR-Verfahren (SSOR):

~~$$(D - \omega L) \hat{u}^{k+1/2} = \omega (R \hat{u}^k + b) - (\omega - 1) D \hat{u}^k$$~~

$$(D - \omega L) \hat{u}^{k+1/2} = \omega (R \hat{u}^k + b) - (\omega - 1) D \hat{u}^k$$

$$(D - \omega R) \hat{u}^{k+1} = \omega (L \hat{u}^{k+1/2} + b) - (\omega - 1) D \hat{u}^{k+1/2}$$

Man kann zeigen, dass die Iterationsmatrix dieses Verfahrens symmetrisch ist.

Satz 9.3 (Ostrowski-Reich-Theorem)

Sei $A \in \mathbb{R}^{N \times N}$ symmetrisch, $a_{ii} > 0$ und $\omega \in (0, 2)$. Das SOR- und das SSOR-Verfahren konvergieren genau dann, wenn A positiv definit ist.

(d) Ein Modellproblem: Poisson-Gleichung auf dem Einheitsquadrat (133)

$$\begin{aligned} -\Delta u &= f & \text{in } \Omega &= (0,1) \times (0,1) \\ u &= 0 & \text{auf } \Gamma \end{aligned}$$

Approximation durch finite Differenzen (vgl. 1.):

Wähle $N \in \mathbb{N}$, setze $h = \frac{1}{N}$, $x_n = nh$, $y_m = mh$. Löse

$$-(\Delta_h u_h)(x_n, y_m) = f(x_n, y_m) \quad \forall n, m = 1, \dots, N-1$$

wobei Δ_h die Approximation von Δ durch finite Differenzen ist (siehe 1.).

Die Eigenwerte λ_{kl} und Eigenvektoren v_{kl} von Δ_h sind

$$v_{kl}(x_n, y_m) = \sin\left(\frac{kn\pi}{N}\right) \sin\left(\frac{lm\pi}{N}\right)$$

$$\lambda_{kl} = \frac{2}{h^2} \left(\cos\left(\frac{k\pi}{N}\right) + \cos\left(\frac{l\pi}{N}\right) - 2 \right) \quad (\text{nachrechnen})$$

Δ_h ist isomorph zu einer Matrix $A_h \in \mathbb{R}^{(N-1)^2 \times (N-1)^2}$ mit den gleichen Eigenwerten

Die Iterationsmatrix $D^{-1}(L+R) \stackrel{= I - D^{-1}A}{=} \quad (134)$
des Jacobi-Verfahrens ist
isomorph zum Operator $I + \frac{h^2}{4} \Delta_h$. Er hat die Eigenwerte

$$1 + \frac{h^2}{4} \lambda_{kl} = \frac{1}{2} \left(\cos\left(\frac{k\pi}{N}\right) + \cos\left(\frac{l\pi}{N}\right) \right) \in (-1, 1)$$

und die Eigenvektoren v_{kl} ($k, l = 1, \dots, N-1$)

Für große N gilt

$$\max_{k, l = 1, \dots, N-1} \left| 1 + \frac{h^2}{4} \lambda_{kl} \right| = \frac{1}{2} \left(\cos\left(\frac{\pi}{N}\right) \cdot 2 \right) = 1 - \frac{\pi^2}{N^2} + O(N^{-4}) \approx 1$$

\Rightarrow sehr langsame Konvergenz

Aber: Für $1 \ll k, l \ll N-1$ ist $1 + \frac{h^2}{4} \lambda_{kl}$ deutlich kleiner als 1.
Das bedeutet, dass die mittelfrequenten Anteile des Fehlers schneller
verschwinden als die glatten Anteile.

Gedämpftes Jacobi-Verfahren:

$$D\hat{u}^{k+1} = \omega(L+R)\hat{u}^k + (1-\omega)D\hat{u}^k + \omega b, \quad \omega \in (0,1)$$

Die Eigenwerte der Iterationsmatrix $\omega D^{-1}(L+R) + (1-\omega)I = I - \omega D^{-1}A$ k sind

$$\lambda_{kl}^{\omega} = \frac{\omega}{2} \left(\cos\left(\frac{k\pi}{N}\right) + \cos\left(\frac{l\pi}{N}\right) \right) + (1-\omega)$$

Setze z.B. $\omega = 1/2$. Dann gilt $\lambda_{kl}^{\omega} \in (0,1)$ und

- falls $k \approx N, l \approx N$: $\lambda_{kl}^{\omega} \approx \frac{\omega}{2}(-1-1) + (1-\omega) = 1-2\omega = 0$
- falls $k \approx 1, l \approx 1$: $\lambda_{kl}^{\omega} \approx \frac{\omega}{2}(1+1) + (1-\omega) = 1$

Beim gedämpften Jacobi-Verfahren mit $\omega = \frac{1}{2}$ werden also alle hochfrequenten Anteile des Fehlers schnell reduziert.

Ein ähnliches Resultat gilt für Gauß-Seidel.

Fazit: Die hier vorgestellten Verfahren werden aufgrund der langsamen Konvergenz heute nicht mehr direkt als Löser für LGS eingesetzt.

Die Glättungseigenschaft kann jedoch zur Konstruktion von Mehrgitterverfahren ausgenutzt werden.

Grundidee: Führe einige Glättungsschritte mit Jacobi oder Gauß-Seidel durch. Danach ist der Fehler zwar noch groß, aber glatt. Also kann der Fehler auf einem größeren Gitter korrigiert werden (billiger!).