

Kapitel 6

Iterationsverfahren für lineare Gleichungssysteme

Falls n sehr groß ist und falls die Matrix \mathbf{A} dünn besetzt ist (*sparse*), dann wählt man zur Lösung von $\mathbf{Ax} = \mathbf{b}$ im Allgemeinen iterative Verfahren.

Definition 6.1 Seien $\mathbf{A}, \mathbf{B}, \mathbf{C}$ $n \times n$ Matrizen. Das Paar (\mathbf{B}, \mathbf{C}) heißt *Zerlegung* (engl. *splitting*) von \mathbf{A} , falls \mathbf{B} invertierbar ist und $\mathbf{A} = \mathbf{B} - \mathbf{C}$ gilt.

Mittels einer Zerlegung definiert man ein Iterationsverfahren zur Lösung von $\mathbf{Ax} = \mathbf{b}$, nämlich mittels

$$\mathbf{B}\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{b}. \quad (6.1)$$

Falls $\mathbf{x}^{(k)} \rightarrow \mathbf{x}^*$ für $k \rightarrow \infty$, dann ist \mathbf{x}^* offenbar eine Lösung von $\mathbf{Ax} = \mathbf{b}$.

Man möchte \mathbf{B} so bestimmen, dass

- (i) die Iteration (6.1) möglichst schnell konvergiert;
- (ii) das Gleichungssystem (6.1) einfach lösbar ist.

Beispiel 6.2 Sei $\mathbf{A} = [a_{ij}]$ gegeben. Wir definieren

$$\mathbf{L} = \begin{bmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & 0 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ & 0 & a_{23} & \dots & a_{2n} \\ & & 0 & \ddots & \vdots \\ & & & \ddots & a_{n-1,n} \\ & & & & 0 \end{bmatrix}$$

und $\mathbf{D} = \text{diag}(a_{ii})$. Dann gilt offenbar $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$.

- (i) Verfahren von *Jacobi* (Gesamtschrittverfahren):
 $\mathbf{B} := \mathbf{D}$ und $\mathbf{C} := -\mathbf{L} - \mathbf{U}$ (notwendige Voraussetzung: $a_{ii} \neq 0$)
- (ii) Verfahren von *Richardson*:
 $\mathbf{B} := \mathbf{I}$ und $\mathbf{C} := \mathbf{I} - \mathbf{A}$ (immer definiert, billig, selten konvergent)
- (iii) Verfahren von *Gauß-Seidel* (Einzelschrittverfahren):
 $\mathbf{B} := \mathbf{D} + \mathbf{L}$ und $\mathbf{C} := -\mathbf{U}$ (notwendige Voraussetzung: $a_{ii} \neq 0$)

Bei allen drei Verfahren ist die Iteration (6.1) leicht lösbar.

Das Jacobi-Verfahren lautet: Berechne für $i = 1, \dots, n$

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

Das Gauß-Seidel-Verfahren lautet: Berechne für $i = 1, \dots, n$

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

Dabei stehen die *neuen* und *alten* Komponenten von \mathbf{x} stets im selben Vektor, z.B. im i -ten Schritt der $(k+1)$ -ten Iteration

$$\mathbf{x} = [x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}, \star, x_{i+1}^{(k)}, \dots, x_n^{(k)}]^\top.$$

Das Verfahren von Jacobi ist parallelisierbar, jenes von Gauß-Seidel nicht.

Wann konvergieren diese Verfahren? Wegen

$$\mathbf{B}\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{b}, \quad \mathbf{A} = \mathbf{B} - \mathbf{C}$$

und der geforderten Invertierbarkeit von \mathbf{B} gilt

$$\mathbf{x}^{(k+1)} = \mathbf{H}\mathbf{x}^{(k)} + \mathbf{d} \quad \text{mit} \quad \mathbf{H} = \mathbf{B}^{-1}\mathbf{C} \quad \text{und} \quad \mathbf{d} = \mathbf{B}^{-1}\mathbf{b}.$$

Das ist gerade eine Fixpunktaufgabe.

Satz 6.3 Die Iteration $\mathbf{x}^{(k+1)} = \mathbf{H}\mathbf{x}^{(k)} + \mathbf{d}$ konvergiert (d.h. $\mathbf{x}^{(k)} \rightarrow \mathbf{x}^*$), falls der Spektralradius von \mathbf{H} kleiner als 1 ist, d.h.

$$\rho(\mathbf{H}) := \max\{|\lambda| ; \lambda \text{ ist Eigenwert von } \mathbf{H}\} < 1. \quad (6.2)$$

Offensichtlich gilt dann $\mathbf{A}\mathbf{x}^* = \mathbf{b}$.

Beweis: Wegen

$$\mathbf{x}^{(k)} = \mathbf{H}^k \mathbf{x}^{(0)} + \left(\mathbf{H}^{k-1} + \dots + \mathbf{H} + \mathbf{I} \right) \mathbf{d} = \mathbf{H}^k \mathbf{x}^{(0)} + (\mathbf{I} - \mathbf{H})^{-1} (\mathbf{I} - \mathbf{H}^k) \mathbf{d}$$

kommt es auf die Potenzen von \mathbf{H} an. Falls (6.2) gilt, gehen die Potenzen gegen Null (Jordan'sche Normalform). \square

Auch für die Konvergenzgeschwindigkeit ist der Spektralradius von \mathbf{H} ausschlaggebend. Subtrahiert man die beiden Gleichungen

$$\mathbf{x}^{(k+1)} = \mathbf{H}\mathbf{x}^{(k)} + \mathbf{d} \quad \text{und} \quad \mathbf{x}^* = \mathbf{H}\mathbf{x}^* + \mathbf{d}$$

voneinander, so erhält man

$$\mathbf{x}^{(k+1)} - \mathbf{x}^* = \mathbf{H}(\mathbf{x}^{(k)} - \mathbf{x}^*) = \dots = \mathbf{H}^{k+1}(\mathbf{x}^{(0)} - \mathbf{x}^*).$$

Somit gilt

$$\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|}{\|\mathbf{x}^{(0)} - \mathbf{x}^*\|} \leq \|\mathbf{H}^k\|.$$

Bemerkung: Für beliebige (quadratische) Matrizen \mathbf{K} gilt

$$\varrho(\mathbf{K}) = \lim_{k \rightarrow \infty} \|\mathbf{K}^k\|^{1/k}.$$

Insbesondere ist $\varrho(\mathbf{K}) \leq \|\mathbf{K}\|$.

Satz 6.4 Sei \mathbf{A} eine $n \times n$ Matrix. Falls \mathbf{A} diagonaldominant ist, d.h.

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad i = 1, \dots, n,$$

dann konvergieren Jacobi- und Gauß-Seidel-Verfahren.

Beweis: Wie führen den Beweis für das Jacobi-Verfahren aus.

Wegen $\mathbf{B}^{-1}\mathbf{C} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})$ gilt

$$(\mathbf{B}^{-1}\mathbf{C})_{ij} = \begin{cases} -\frac{a_{ij}}{a_{ii}} & i \neq j, \\ 0 & i = j. \end{cases}$$

Somit ist

$$\|\mathbf{B}^{-1}\mathbf{C}\|_{\infty} = \max_i \sum_{j=1}^n |(\mathbf{B}^{-1}\mathbf{C})_{ij}| < 1,$$

und das Resultat folgt wegen $\varrho(\mathbf{B}^{-1}\mathbf{C}) \leq \|\mathbf{B}^{-1}\mathbf{C}\|$. □

Eine Matrix heißt *schwach* diagonaldominant, falls

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad i = 1, \dots, n,$$

und für *mindestens* ein i die strenge Ungleichung gilt.

Beispiel 6.5 Die Matrix

$$\mathbf{A} = \begin{bmatrix} 4 & -1 & 0 & \vdots & -1 & 0 & 0 & \vdots & 0 & 0 & 0 \\ -1 & 4 & -1 & \vdots & 0 & -1 & 0 & \vdots & 0 & 0 & 0 \\ 0 & -1 & 4 & \vdots & 0 & 0 & -1 & \vdots & 0 & 0 & 0 \\ \hdashline -1 & 0 & 0 & \vdots & 4 & -1 & 0 & \vdots & -1 & 0 & 0 \\ 0 & -1 & 0 & \vdots & -1 & 4 & -1 & \vdots & 0 & -1 & 0 \\ 0 & 0 & -1 & \vdots & 0 & -1 & 4 & \vdots & 0 & 0 & -1 \\ \hdashline 0 & 0 & 0 & \vdots & -1 & 0 & 0 & \vdots & 4 & -1 & 0 \\ 0 & 0 & 0 & \vdots & 0 & -1 & 0 & \vdots & -1 & 4 & -1 \\ 0 & 0 & 0 & \vdots & 0 & 0 & -1 & \vdots & 0 & -1 & 4 \end{bmatrix}$$

ist schwach diagonaldominant.

Falls die Matrix \mathbf{A} nur schwach diagonaldominant ist, braucht man zusätzliche *strukturelle* Eigenschaften von \mathbf{A} , um die Konvergenz des Einzel- und Gesamtschrittverfahrens zu garantieren.

Definition 6.6 Eine $n \times n$ Matrix \mathbf{A} heißt *zerfallend*, wenn es eine echte Teilmenge $J \subset \{1, 2, \dots, n\}$ gibt mit

$$a_{ij} = 0 \quad \text{für } i \in J, j \notin J.$$

Eine zerfallende Matrix hat nach Ummummerierung der Gleichungen und Variablen die Blockstruktur

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

mit \mathbf{A}_{11} und \mathbf{A}_{22} quadratisch. Damit kann die Lösung des Gleichungssystems auf die Lösung von zwei kleineren Systemen zurückgeführt werden. Für nicht-zerfallende Matrizen gilt folgender Satz (ohne Beweis):

Satz 6.7 Sei \mathbf{A} eine $n \times n$ Matrix. Falls \mathbf{A} nicht zerfallend und schwach diagonaldominant ist, dann konvergieren Jacobi- und Gauß-Seidel-Verfahren.

Gedämpfte Iteration. Es handelt sich dabei um eine Methode, den Spektralradius der Iterationsmatrix \mathbf{H} zu verkleinern. Wir definieren

$$\tilde{\mathbf{x}}^{(k+1)} = \omega(\mathbf{H}\tilde{\mathbf{x}}^{(k)} + \mathbf{d}) + (1 - \omega)\tilde{\mathbf{x}}^{(k)}, \quad \omega \in \mathbb{R}, \omega \neq 0$$

mit einem Dämpfungsparameter ω . Setzt man $\mathbf{H}_\omega := \omega\mathbf{H} + (1 - \omega)\mathbf{I}$, so erhält man

$$\tilde{\mathbf{x}}^{(k+1)} = \mathbf{H}_\omega\tilde{\mathbf{x}}^{(k)} + \omega\mathbf{d},$$

mit derselben Lösung wie das ursprüngliche Problem

$$\mathbf{x} = \mathbf{H}_\omega\mathbf{x} + \omega\mathbf{d} \quad \Leftrightarrow \quad \mathbf{x} = \omega\mathbf{H}\mathbf{x} + \mathbf{x} - \omega\mathbf{x} + \omega\mathbf{d} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{H}\mathbf{x} + \mathbf{d}.$$

Definition 6.8 Gegeben sei ein Iterationsverfahren $\mathbf{x}^{(k+1)} = \mathbf{H}\mathbf{x}^{(k)} + \mathbf{d}$. Dann heißt $\tilde{\mathbf{x}}^{(k+1)} = \mathbf{H}_\omega\tilde{\mathbf{x}}^{(k)} + \omega\mathbf{d}$ (zugehöriges) Relaxationsverfahren. Die Zahl ω heißt Relaxationsparameter. Man nennt den Fall $\omega > 1$ Überrelaxation und $\omega < 1$ Unterrelaxation.

Bei Einzelschrittverfahren (z.B. Gauß-Seidel) verwendet man die bereits berechneten Werte

$$x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)} \quad \text{anstelle von} \quad x_1^{(k)}, \dots, x_{i-1}^{(k)}$$

zur Relaxation. Man spricht dann von *sukzessiver* Relaxation.

Das SOR-Verfahren. Aus der Gauß-Seidel-Iteration erhält man

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

$$(\mathbf{D} + \mathbf{L})\tilde{\mathbf{x}}^{(k+1)} = \omega(-\mathbf{U}\tilde{\mathbf{x}}^{(k)} + \mathbf{b}) + (1 - \omega)\mathbf{D}\tilde{\mathbf{x}}^{(k)} + (1 - \omega)\mathbf{L}\tilde{\mathbf{x}}^{(k+1)}$$

$$(\mathbf{D} + \omega\mathbf{L})\tilde{\mathbf{x}}^{(k+1)} = \left((1 - \omega)\mathbf{D} - \omega\mathbf{U} \right) \tilde{\mathbf{x}}^{(k)} + \omega\mathbf{b}$$

Definition 6.9 Sei $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$ wie vorher. Das Iterationsverfahren

$$\mathbf{x}^{(k+1)} = \mathbf{H}_\omega\mathbf{x}^{(k)} + \mathbf{d}$$

mit $\mathbf{H}_\omega = (\mathbf{D} + \omega\mathbf{L})^{-1} \left((1 - \omega)\mathbf{D} - \omega\mathbf{U} \right)$ und $\mathbf{d} = \omega(\mathbf{D} + \omega\mathbf{L})^{-1}\mathbf{b}$ heißt SOR-Verfahren (sukzessive Überrelaxation).

Satz 6.10 Sei \mathbf{A} eine symmetrische, positiv definite Matrix und $\omega \in (0, 2)$. Dann konvergiert das SOR-Verfahren zur Lösung von $\mathbf{Ax} = \mathbf{b}$.

Beweis: Wir zeigen $\rho(\mathbf{H}_\omega) < 1$. Sei λ ein Eigenwert von \mathbf{H}_ω und $\mathbf{v} \neq \mathbf{0}$ ein zugehöriger Eigenvektor, $\mathbf{H}_\omega \mathbf{v} = \lambda \mathbf{v}$. Dann ist

$$\underbrace{((1 - \omega)\mathbf{D} - \omega\mathbf{U})}_{\mathbf{D} - \omega\mathbf{A} + \omega\mathbf{L}} \mathbf{v} = \lambda(\mathbf{D} + \omega\mathbf{L})\mathbf{v} \quad (6.3)$$

und weiters

$$-\omega\mathbf{A}\mathbf{v} = (\lambda - 1)(\mathbf{D} + \omega\mathbf{L})\mathbf{v}.$$

Wir setzen $\alpha = \mathbf{v}^*(\mathbf{D} + \omega\mathbf{L})\mathbf{v}$. Da \mathbf{A} positiv definit vorausgesetzt wurde, ist

$$(1 - \lambda)\alpha = \omega\mathbf{v}^*\mathbf{A}\mathbf{v} =: \beta > 0,$$

insbesondere $\lambda \neq 1$. Weiters ist (wegen der Symmetrie von \mathbf{A}) $\mathbf{L}^T = \mathbf{U}$ und somit

$$\bar{\alpha} = \overline{\mathbf{v}^*(\mathbf{D} + \omega\mathbf{L})\mathbf{v}} = \mathbf{v}^*(\mathbf{D} + \omega\mathbf{U})\mathbf{v}.$$

Somit ist wegen (6.3)

$$\begin{aligned} \lambda\alpha + \bar{\alpha} &= \mathbf{v}^*(\lambda(\mathbf{D} + \omega\mathbf{L}) + \mathbf{D} + \omega\mathbf{U})\mathbf{v} \\ &= \mathbf{v}^*(\mathbf{D} + \omega\mathbf{L} - \omega\mathbf{A} + \mathbf{D} + \omega\mathbf{U})\mathbf{v} = (2 - \omega)\mathbf{v}^*\mathbf{D}\mathbf{v} > 0. \end{aligned}$$

In der letzten Ungleichung wurde die Identität $d_{ii} = \mathbf{e}_i^T \mathbf{A} \mathbf{e}_i > 0$ verwendet.

Also ergibt sich

$$0 < \lambda\alpha + \bar{\alpha} = \beta \left(\frac{\lambda}{1 - \lambda} + \frac{1}{1 - \bar{\lambda}} \right) = \beta \frac{\lambda - \lambda\bar{\lambda} + 1 - \lambda}{|1 - \lambda|^2} = \beta \frac{1 - |\lambda|^2}{|1 - \lambda|^2}.$$

Somit ist $1 - |\lambda|^2 > 0$ und damit $|\lambda| < 1$. □

Wie lautet der optimale Relaxationsparameter, welcher einen minimalen Spektralradius von \mathbf{H}_ω ergibt? Diese Frage kann nur in Spezialfällen beantwortet werden (vgl. Literatur; Dissertation von Young).

SSOR-Verfahren. Falls die Matrix \mathbf{A} symmetrisch ist ($\mathbf{U} = \mathbf{L}^T$), dann ist die Iterationsmatrix

$$\mathbf{H} = -(\mathbf{D} + \mathbf{L})^{-1}\mathbf{L}^T$$

des Gauß-Seidel-Verfahrens nicht notwendig symmetrisch. Um Symmetrie zu erhalten, definiert man das Gauß-Seidel-Verfahren in Rückwärtsrichtung durch

$$(\mathbf{D} + \mathbf{U})\mathbf{x}^{(k+1)} = -\mathbf{L}\mathbf{x}^{(k)} + \mathbf{b}.$$

Man tauscht also \mathbf{U} mit \mathbf{L} aus und löst von unten nach oben auf. Das SSOR-Verfahren ist die Kombination aus einem SOR-Schritt in Vorwärtsrichtung mit einem in Rückwärtsrichtung. Es besitzt eine symmetrische Iterationsmatrix und wird vor allem als Vorkonditionierer verwendet.

Blockweise Iterationen. Eine Verallgemeinerung der bisher vorgestellten Iterationen sind Blockiterationen. Die Matrix \mathbf{A} sei in Blöcke zerlegt

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1n} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{A}_{n1} & \mathbf{A}_{n2} & \dots & \mathbf{A}_{nn} \end{bmatrix}.$$

Wir definieren wieder eine Zerlegung (in Blöcke)

$$\mathbf{L} = \begin{bmatrix} \mathbf{0} & & & & \\ \mathbf{A}_{21} & \mathbf{0} & & & \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{0} & & \\ \vdots & \vdots & \ddots & \ddots & \\ \mathbf{A}_{n1} & \mathbf{A}_{n2} & \dots & \mathbf{A}_{n,n-1} & \mathbf{0} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \mathbf{0} & \mathbf{A}_{12} & \mathbf{A}_{13} & \dots & \mathbf{A}_{1n} \\ & \mathbf{0} & \mathbf{A}_{23} & \dots & \mathbf{A}_{2n} \\ & & \mathbf{0} & \ddots & \vdots \\ & & & \ddots & \mathbf{A}_{n-1,n} \\ & & & & \mathbf{0} \end{bmatrix}.$$

und $\mathbf{D} = \text{blockdiag}[\mathbf{A}_{11}, \dots, \mathbf{A}_{nn}]$. Mit dieser Zerlegung lassen sich wie vorher Jacobi-, Gauß-Seidel und SOR-Verfahren definieren. Einziger Unterschied ist, dass nun alle Operationen blockweise definiert sind.

Beispiel 6.11 Das Block-Jacobi-Verfahren für das Gleichungssystem

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}$$

lautet

$$\begin{aligned} \mathbf{A}_{11}\mathbf{x}_1^{(k+1)} &= -\mathbf{A}_{12}\mathbf{x}_2^{(k)} + \mathbf{b}_1, \\ \mathbf{A}_{22}\mathbf{x}_2^{(k+1)} &= -\mathbf{A}_{21}\mathbf{x}_1^{(k)} + \mathbf{b}_2. \end{aligned}$$

6.1 Mehrgitterverfahren

Mehrgitterverfahren (*multigrid methods*) wurden in den späten 70er Jahren des 20. Jahrhunderts entdeckt. Es handelt sich um iterative Verfahren, die *strukturierte* Gleichungssysteme in N Unbekannten mit $\mathcal{O}(N)$ Operationen lösen. Die bis dahin bekannten iterativen Löser benötigten typischerweise $\mathcal{O}(N^{3/2})$ Operationen. Das Beschleunigungspotential der neuen Verfahren ist \sqrt{N} ; bei $N = 10^8$ bedeutet das z.B. 1 Minuten Rechenzeit anstatt 166 Stunden.

6.1.1 1-D Modellproblem

Wir betrachten das 1-D elliptische Randwertproblem

$$-u''(x) = f(x), \quad x \in (0, 1); \quad u(0) = u(1) = 0.$$

Zur numerischen Lösung zerlegen wir das Intervall $[0, 1]$ in $N + 1$ äquidistante Teile mit der Gitterweite $h = 1/(N + 1)$ und den Stützstellen $x_j = jh$ und approximieren

$$-u''(x_j) = \frac{-u(x_{j+1}) + 2u(x_j) - u(x_{j-1}))}{h^2} + \mathcal{O}(h^2).$$

Damit ergibt sich für die Unbekannten $w_j \approx u(x_j)$ das lineare Gleichungssystem $\mathbf{A}\mathbf{w} = \mathbf{b}$ mit $b_j = f(x_j)$. Die Systemmatrix \mathbf{A} hat Bandstruktur

$$\mathbf{A} = \frac{1}{h^2} \begin{bmatrix} 2 & -1 & 0 & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & -1 & 2 & -1 \\ 0 & \dots & \dots & 0 & -1 & 2 \end{bmatrix}.$$

(In Raumdimension 2 lautet diese Matrix $\mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}$. In Raumdimension 3 hat sie eine entsprechende Gestalt.)

6.1.2 Analyse der Jacobi-Iteration

Wir schreiben $\mathbf{A} = \mathbf{D} - \mathbf{C}$, wobei \mathbf{D} die Diagonale von \mathbf{A} bezeichnet. Die Jacobi-Iteration lautet

$$\mathbf{D}\mathbf{w}^{(k+1)} = \mathbf{C}\mathbf{w}^{(k)} + \mathbf{b},$$

beziehungsweise

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} - \underbrace{\mathbf{D}^{-1}(\mathbf{A}\mathbf{w}^{(k)} - \mathbf{b})}_{\text{Defekt von } \mathbf{w}^{(k)}}.$$

Wir betrachten im Folgenden eine gedämpfte Jacobi-Iteration

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} - \vartheta \mathbf{D}^{-1}(\mathbf{A}\mathbf{w}^{(k)} - \mathbf{b}), \quad 0 < \vartheta \leq 1. \tag{6.4}$$

Im Beispiel ist $\mathbf{D}^{-1} = \frac{h^2}{2} \mathbf{I}$. Mit $\omega = \vartheta/2$ lautet die Iteration (6.4) somit

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} - \omega h^2 (\mathbf{A}\mathbf{w}^{(k)} - \mathbf{b}), \quad 0 < \omega \leq 1/2.$$

Wir analysieren die Eigenschaften dieser Iteration mittels *diskrete* Fourieranalyse. Die Eigenwerte μ_j und die dazugehörigen Eigenvektoren \mathbf{v}_j von \mathbf{A} lauten

$$\mu_j = \frac{4}{h^2} \sin^2(jh\pi/2), \quad \mathbf{v}_j = \sqrt{2} \begin{bmatrix} \sin(j\pi h) \\ \sin(2j\pi h) \\ \vdots \\ \sin(Nj\pi h) \end{bmatrix},$$

wobei $j = 1, \dots, N$. Die Überprüfung erfolgt durch Nachrechnen mittels Sumsensätze der Winkelfunktionen

$$\begin{aligned} (\mathbf{A}\mathbf{v}_j)_n &= \frac{1}{h^2} [-1 \quad 2 \quad -1] \sqrt{2} \begin{bmatrix} \sin((n-1)j\pi h) \\ \sin(nj\pi h) \\ \sin((n+1)j\pi h) \end{bmatrix} \\ &= \frac{\sqrt{2}}{h^2} (-2 \sin(nj\pi h) (\cos(j\pi h) - 1)) \\ &= \frac{2\sqrt{2}}{h^2} \sin(nj\pi h) \cdot (2 - 2 \cos^2(j\pi h/2)) \\ &= \frac{4}{h^2} \sin^2(j\pi h/2) \sqrt{2} \sin(nj\pi h) = (\mu_j \mathbf{v}_j)_n. \end{aligned}$$

Hier wurde insbesondere die Relation

$$\cos 2\varphi = \cos^2 \varphi - \sin^2 \varphi = 2 \cos^2 \varphi - 1 = 1 - 2 \sin^2 \varphi$$

verwendet.

Die Iterationsmatrix \mathbf{H} der gedämpften Jacobi-Iteration lautet

$$\mathbf{H} = \mathbf{I} - \omega h^2 \mathbf{A}.$$

Diese Matrix hat dieselben Eigenvektoren wie \mathbf{A} , ihre Eigenwerte lauten jedoch

$$\ell_j = \ell_j(\omega) = 1 - \omega h^2 \mu_j = 1 - 4\omega \sin^2(j\pi h/2), \quad j = 1, \dots, N.$$

Abbildung 6.1 zeigt links die ersten vier (niederfrequenten) Eigenvektoren und rechts die letzten vier (hochfrequenten) Eigenvektoren von \mathbf{H} im Fall $N = 12$.

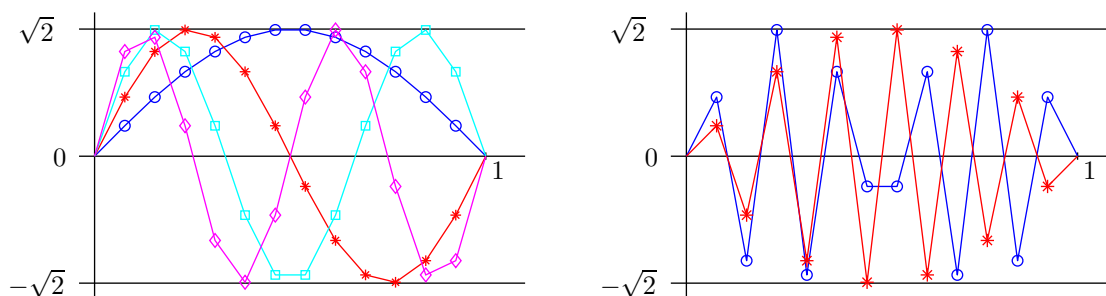


Abbildung 6.1: Eigenvektoren von \mathbf{H} für den Fall $N = 12$. Das linke Bild zeigt die niederfrequenten Vektoren \mathbf{v}_1 (blau; Kreise), \mathbf{v}_2 (rot; Sterne), \mathbf{v}_3 (cyan; Quadrate) und \mathbf{v}_4 (magenta; Rauten). Das rechte Bild zeigt die hochfrequenten Vektoren \mathbf{v}_{11} (blau; Kreise) und \mathbf{v}_{12} (rot; Sterne).

Wir betrachten nun die Fehler der Jacobi-Iteration, dargestellt in der Basis der Eigenvektoren. Sei

$$\mathbf{w}^{(0)} - \mathbf{w} = \sum_{j=1}^N \alpha_j \mathbf{v}_j.$$

Wegen $\mathbf{w}^{(k+1)} = \mathbf{H}\mathbf{w}^{(k)} + \mathbf{b}$ und $\mathbf{w} = \mathbf{H}\mathbf{w} + \mathbf{b}$ gilt

$$\mathbf{w}^{(1)} - \mathbf{w} = \mathbf{H}(\mathbf{w}^{(0)} - \mathbf{w}) = \mathbf{H} \sum_{j=1}^N \alpha_j \mathbf{v}_j = \sum_{j=1}^N \alpha_j \mathbf{H}\mathbf{v}_j = \sum_{j=1}^N \alpha_j \ell_j \mathbf{v}_j$$

und allgemein

$$\mathbf{w}^{(k)} - \mathbf{w} = \sum_{j=1}^N \alpha_j \ell_j^k \mathbf{v}_j = \sum_{j=1}^N \beta_j \mathbf{v}_j, \quad k \geq 0.$$

Diese Formel zusammen mit Abb. 6.2 zeigt für $\omega = 1/4$ folgendes: Für niedrige Frequenzen (d.h. kleines j) gilt $\beta_j \approx \alpha_j$; für hohe Frequenzen gilt hingegen: $|\beta_j| \ll |\alpha_j|$. Die hochfrequenten Fehler werden bei $\omega = 1/4$ somit ausgezeichnet gedämpft. Man sagt, das Verfahren besitzt eine sehr gute *Glättungseigenschaft* (*smoothing property*).

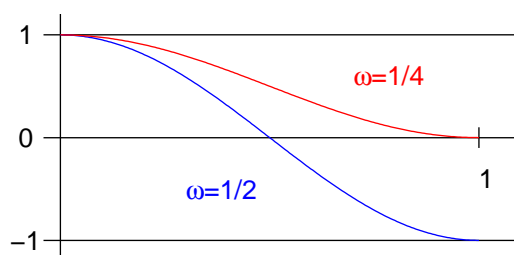


Abbildung 6.2: Eigenwerte $1 - 4\omega \sin^2(jh\pi/2)$ des ungedämpften Jacobi-Verfahrens ($\omega = 1/2$) und des mit $\omega = 1/4$ gedämpften Verfahrens als Funktion von $x = jh$, $0 < x < 1$.

Bemerkung 6.12 Die Analyse in 2-D und 3-D erfolgt ähnlich. In 2-D lauten die Eigenwerte und Eigenvektoren der Systemmatrix $\mu_j + \mu_k$ und $\mathbf{v}_j \otimes \mathbf{v}_k$ für $1 \leq j, k \leq N$ mit μ_j und \mathbf{v}_j aus dem 1-D Fall.

6.1.3 Zweigitterverfahren

Gesucht ist eine Lösung \mathbf{w} des linearen Gleichungssystems

$$\mathbf{A}\mathbf{w} = \mathbf{b} \quad (\text{zur Gitterweite } h).$$

Ausgehend von einem Startwert $\mathbf{w}^{(0)}$ führen wir einige gedämpfte Jacobi-Iterationen aus und erhalten \mathbf{w}^* , dessen Fehler $\mathbf{v} = \mathbf{w}^* - \mathbf{w}$ *glatt* ist (d.h. keine hochfrequenten Anteile besitzt). Offensichtlich ist \mathbf{v} gerade die gesuchte *Korrektur* zur exakten Lösung. Man beachte, dass \mathbf{v} i.A. sehr groß ist.

Für den Defekt \mathbf{d} der “Näherung” \mathbf{w}^* gilt

$$\mathbf{d} = \mathbf{A}\mathbf{w}^* - \mathbf{b} = \mathbf{A}\mathbf{w}^* - \mathbf{A}\mathbf{w} = \mathbf{A}\mathbf{v}.$$

Somit erfüllt die gesuchte Korrektur eine Gleichung vom selben Typ wie die ursprüngliche Gleichung. Die *entscheidende Idee* ist: eine glatte Funktion \mathbf{v} kann auch auf einem *größeren* Gitter gut approximiert werden.

Sei $\hat{h} = 2h$ und $\hat{\mathbf{A}}$ die zugehörige Systemmatrix für das größere Gitter. Das größere Gitter hat um den Faktor 2^d weniger Gitterpunkte, wobei $d = 1, 2, 3$ die Raumdimension bezeichnet. Man löst

$$\hat{\mathbf{A}}\hat{\mathbf{v}} = \hat{\mathbf{d}},$$

wobei $\hat{\mathbf{d}}$ die *Restriktion* (Einschränkung) von \mathbf{d} auf das größere Gitter ist

$$\hat{\mathbf{d}} = \mathcal{R}\mathbf{d}.$$

Beispiel 6.13 (i) Die *triviale* Restriktion lautet

$$\mathcal{R}\mathbf{d}(x) = \mathbf{d}(x) \quad \text{für } x = 2jh \text{ (auf dem groben Gitter)}.$$

Hier wurde die intuitive Kurzschreibweise $\mathbf{d}(kh) = d_k$ für Gitterfunktionen verwendet. Die obige Bedingung lautet daher $\hat{d}_j = d_{2j}$.

(ii) Eine bessere Restriktion erhält man mit

$$\mathcal{R}\mathbf{d}(x) = \frac{1}{4}(\mathbf{d}(x-h) + 2\mathbf{d}(x) + \mathbf{d}(x+h))$$

für $x = 2jh$ auf dem groben Gitter. \diamond

Denken wir uns für den Moment die Gleichung auf dem groben Gitter exakt gelöst, also

$$\widehat{\mathbf{v}} = \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{d}}.$$

Die gesuchte Korrektur ist offenbar nur auf dem groben Gitter definiert. Um eine Näherung an \mathbf{v} zu erhalten, muss man sie auf das feine Gitter fortsetzen. Diesen Prozess nennt man *Prolongation*

$$\mathbf{v}^* = \mathcal{P}\widehat{\mathbf{v}}.$$

Im Folgenden verwenden wir lineare Interpolation als Prolongation

$$\mathcal{P}\widehat{\mathbf{v}}(x) = \begin{cases} \widehat{\mathbf{v}}(x) & \text{für } x = j\widehat{h} = 2jh, \\ \frac{1}{2}(\widehat{\mathbf{v}}(x+h) + \widehat{\mathbf{v}}(x-h)) & \text{für } x = (2j+1)h, \end{cases}$$

vgl. auch Abb. 6.3.

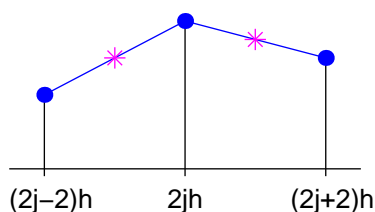


Abbildung 6.3: Prolongation durch lineare Interpolation.

Mit Hilfe der Prolongation der Grobgitterkorrektur ergibt sich die verbesserte Näherungslösung zu

$$\mathbf{w}_{\text{neu}} = \mathbf{w}^* - \mathbf{v}^*.$$

Algorithmus (Zweigitterverfahren). Gegeben sei $\mathbf{w}^{(k)}$.

(i) Glättungsschritt: Man berechnet \mathbf{w}^* mittels einiger gedämpfter Jacobi-Iterationen.

(ii) Grobgitterkorrektur:

Berechne Defekt am feinen Gitter $\mathbf{d} = \mathbf{A}\mathbf{w}^* - \mathbf{b}$

Restriktion des Defekts auf das grobe Gitter $\widehat{\mathbf{d}} = \mathcal{R}\mathbf{d}$

Löse Gleichung auf grobem Gitter $\widehat{\mathbf{v}} = \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{d}}$

Prolongation auf das feine Gitter $\mathbf{v}^* = \mathcal{P}\widehat{\mathbf{v}}$

Korrektur am feinen Gitter $\mathbf{w}^{(k+1)} = \mathbf{w}^* - \mathbf{v}^*.$

6.1.4 Konvergenz des Zweigitterverfahrens

Wir untersuchen obiges Zweigitterverfahren für $\omega = 1/4$. Die Zwei-Gitter-Iteration lautet

$$\mathbf{w}^{(k+1)} = \mathbf{M}\mathbf{w}^{(k)} + \mathbf{N}\mathbf{b}$$

mit der Iterationsmatrix

$$\mathbf{M} = (\mathbf{I} - \mathcal{P}\widehat{\mathbf{A}}^{-1}\mathcal{R}\mathbf{A})\mathbf{H}^\nu$$

bei ν gedämpften Jacobi-Iterationen. Man hat Konvergenz, falls $\rho(\mathbf{M}) < 1$.

Satz 6.14 (a) Der Spektralradius der Iterationsmatrix $\mathbf{M} = \mathbf{M}(\nu)$ erfüllt

$$\varrho(\mathbf{M}) \leq \max\{x(1-x)^\nu + x^\nu(1-x); 0 \leq x \leq 1/2\} =: \varrho_\nu < 1.$$

(b) Für die Spektralnorm $\|\mathbf{M}\| = \sqrt{\varrho(\mathbf{M}^T \mathbf{M})}$ gilt

$$\|\mathbf{M}\| \leq \max\left\{\sqrt{2x^2(1-x)^{2\nu} + 2x^{2\nu}(1-x)^2}; 0 \leq x \leq 1/2\right\} =: \zeta_\nu < 1.$$

Eine einfache Rechnung ergibt die nachstehende Tabelle:

ν	1	2	3	4
ϱ_ν	1/2	1/4	1/8	0.083
ζ_ν	1/2	1/4	0.15	0.116

Man zeigt weiters, dass für $\nu \rightarrow \infty$ gilt:

$$\varrho_\nu \approx \frac{1}{e\nu}, \quad \zeta_\nu \approx \frac{1}{2\nu}.$$

Beweis: Einfache (aber langwierige) Rechnung. Man verwendet dazu die angepasste Basis \mathbf{v}_j von Eigenvektoren. □

Wahl von ν . Der Aufwand für 2ν Glättungen ist in etwa doppelt so hoch wie für ν Glättungen. Somit sind *zwei* Iterationen mit ν Glättungsschritten besser als *eine* Iteration mit 2ν Glättungsschritten, falls

$$\zeta_\nu^2 \leq \zeta_{2\nu} \approx \frac{\zeta_\nu}{2}, \quad \text{d.h. } \zeta_\nu \lesssim \frac{1}{2}.$$

In der Praxis heißt das: Man verwendet 2 gedämpfte Jacobi-Iterationen.

6.1.5 Mehrgitterverfahren

Man wählt Gitterweiten

$$h_0 = H, \quad h_1 = \frac{H}{2}, \quad h_2 = \frac{H}{4}, \quad \dots, \quad h_\ell = \frac{H}{2^\ell}$$

und sucht für die feinste Gitterweite die Lösung des linearen Gleichungssystems

$$\mathbf{A}_\ell \mathbf{w}_\ell = \mathbf{b}_\ell.$$

(Der Index ℓ deutet hier an, dass dieses Gleichungssystem zur Gitterweite h_ℓ gehört.)

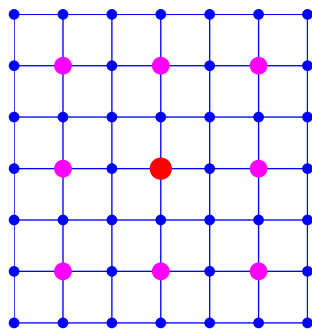


Abbildung 6.4: Drei ineinander geschachtelte Gitter mit 1, 9 und 49 Punkten.

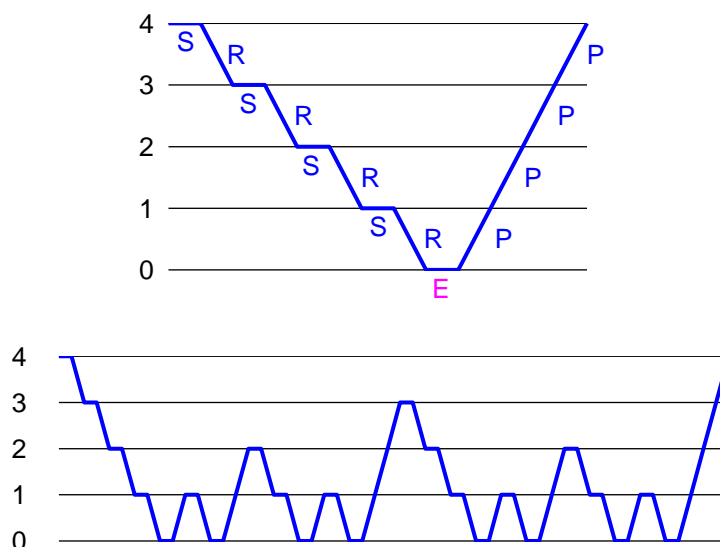


Abbildung 6.5: V-Zyklus (oben) und W-Zyklus (unten) einer Mehrgitteriteration. Hierbei bedeutet S Glättung, R Restriktion, E Auswertung und P Prolongation (nach Hackbusch, 1985).

Wendet man ein Zweigitterverfahren an, so benötigt man die Lösung von

$$\mathbf{A}_{\ell-1} \mathbf{w}_{\ell-1} = \mathbf{d}_{\ell-1}$$

(Defektkorrektur am Grobgitter). Diese Gleichung löst man wieder iterativ mit dem Zweigitterverfahren und erhält ein neues Gleichungssystem

$$\mathbf{A}_{\ell-2} \mathbf{w}_{\ell-2} = \mathbf{d}_{\ell-2}$$

auf einem noch größeren Gitter, usw. Auf dem größten Gitter löst man das System direkt (im Extremfall eine einzige Gleichung!). Je nachdem, ob man jeweils eine oder zwei Zweigitter-Iterationen pro Niveau ausführt, spricht man von einem V-Zyklus bzw. von einem W-Zyklus, vgl. Abb. 6.5.

Algorithmus. $\text{MGM}(\ell, \mathbf{w}, \mathbf{b})$

Falls $\ell = 0$, setze $\mathbf{w} = \mathbf{A}_0^{-1} \mathbf{b}$.

Sonst: \mathbf{w}^* = Ergebnis von ν Jacobi-Iterationen zum Startwert \mathbf{w}

$$\mathbf{d} = \mathcal{R}(\mathbf{A}_\ell \mathbf{w} - \mathbf{b})$$

$$\mathbf{v} = \mathbf{0}$$

für $j = 1, \dots, \gamma$: $\text{MGM}(\ell - 1, \mathbf{v}, \mathbf{d})$

$$\mathbf{w} = \mathbf{w}^* - \mathcal{P} \mathbf{v}$$

Beim V-Zyklus ist $\gamma = 1$, beim W-Zyklus ist $\gamma = 2$.

Aufwand des Verfahrens. Sei N die Anzahl der Gleichungen am feinsten Gitter. Glättung, Restriktion und Prolongation sind $\mathcal{O}(N)$ Operationen mit kleiner $\mathcal{O}(\dots)$ -Konstante, falls man mit *sparse matrix* Techniken arbeitet. Die Anzahl der Gleichungen nimmt pro Niveau etwa um den Faktor 2^{-d} ab, wobei d die Raumdimension bezeichnet. Weiters wird am Niveau ℓ das Niveau k gerade $\gamma^{\ell-k}$ ($1 \leq k \leq \ell$) aufgerufen. Somit ist der Gesamtaufwand der Iteration

$$C \left(N + \gamma \frac{N}{2^d} + \gamma^2 \frac{N}{2^{2d}} + \dots + \gamma^{\ell-1} \frac{N}{2^{(\ell-1)d}} \right) + C_0 = \mathcal{O}(N).$$

Die neue Konstante lautet abhängig von d und γ :

	$\gamma = 1$	$\gamma = 2$
$d = 1$	$2C$	ℓC
$d = 2$	$\frac{4}{3}C$	$2C$
$d = 3$	$\frac{8}{7}C$	$\frac{4}{3}C$

Das Multigridverfahren ist somit in beliebigen Raumdimensionen ein $\mathcal{O}(N)$ -Prozess.

Varianten verwenden andere Glättungsiterationen (meist Gauß-Seidel) oder Nachglätten (d.h. weitere Glättungsschritte nach der Grobgitterkorrektur) sowie andere Restriktionen und Prolongationen.

6.2 Das Verfahren der konjugierten Gradienten

Das Verfahren der konjugierten Gradienten geht auf Hestenes und Stiefel (1952) zurück. Es wird auch als cg-Verfahren bezeichnet (vom Englischen *conjugate gradients methods*).

Im ganzen Abschnitt sei \mathbf{A} eine reelle, symmetrische, positiv definite $n \times n$ Matrix (eine sg. spd-Matrix) und $\mathbf{b} \in \mathbb{R}^n$ ein Vektor. (Typischerweise ist hier n sehr groß und \mathbf{A} dünn besetzt.) Die Aufgabe

$$\text{löse } \mathbf{Ax} = \mathbf{b}$$

ist somit äquivalent zur Aufgabe

$$\text{minimiere } \frac{1}{2} \mathbf{x}^T \mathbf{Ax} - \mathbf{b}^T \mathbf{x}.$$

Wir bezeichnen im Folgenden mit $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt im \mathbb{R}^n . Die spd-Matrix \mathbf{A} definiert ein (dazu äquivalentes) Skalarprodukt $\langle \cdot, \cdot \rangle_{\mathbf{A}}$ mittels

$$\langle \mathbf{y}, \mathbf{z} \rangle_{\mathbf{A}} = \langle \mathbf{y}, \mathbf{Az} \rangle = \langle \mathbf{Ay}, \mathbf{z} \rangle$$

und eine Norm

$$\|\mathbf{y}\|_{\mathbf{A}} = \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle_{\mathbf{A}}},$$

die sg. Energienorm (bezüglich \mathbf{A}).

Definition 6.15 Sei \mathbf{A} eine reelle, symmetrische, positiv definite $n \times n$ Matrix und \mathbf{u}, \mathbf{v} zwei Vektoren im \mathbb{R}^n . Man nennt \mathbf{u} *\mathbf{A} -orthogonal*¹ zu \mathbf{v} (in Zeichen: $\mathbf{u} \perp_{\mathbf{A}} \mathbf{v}$) wenn gilt: $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{A}} = 0$.

¹In der älteren Literatur spricht man auch von *\mathbf{A} -konjugiert*.