

1 Mehrgitterverfahren zur Lösung von Linearen Gleichungssystemen

1.1 Motivation

1.1.1 Das Modellproblem

Um die im folgenden vorgestellten Ideen besser veranschaulichen zu können betrachten wir als Modellproblem die Poisson-Differentialgleichung

$$-\Delta u = f$$

auf dem Gebiet $\Omega := (0, 1) \times (0, 1)$ mit $u(x) = g(x) \quad \forall x \in \partial\bar{\Omega}$. Multiplikation von unendlich oft differenzierbaren Funktionen $\gamma \in C_0^\infty(\Omega)$ und Integration ergeben:

$$\Leftrightarrow - \int_{\Omega} \Delta u \gamma \, d(x, y) = \int_{\Omega} f \gamma \, d(x, y) \quad \forall \gamma \in C_0^\infty(\Omega)$$

Partielle Integration (Satz von Green), sowie die Nullrandbedingungen für γ liefern dann:

$$\Leftrightarrow \int_{\Omega} \nabla u \nabla \gamma \, d(x, y) = \int_{\Omega} f \gamma \, d(x, y) \quad \forall \gamma \in C_0^\infty(\Omega)$$

Nun ist $a(u, \gamma) := \int_{\Omega} \nabla u \nabla \gamma \, d(x, y)$ eine elliptische und stetige Bilinearform auf $H(\Omega)_0^1$, sowie die rechte Seite $(f, \gamma)_{L^2}$ nach dem Satz von Riesz eine stetige Linearform auf $H(\Omega)_0^1$. Zerlegt man Ω nun in uniforme 3-Ecke der Größe h und definiert als S_h den endlich dimensionalen Vektorraum der Funktionen, die auf den Knotenpunkten $(x_j, y_j) \in \Omega_h$ leben, kann man als Basis von S_h die stückweise linearen Funktionen $(\Psi_i)_{i=1}^N$ wählen:

$$\Psi_i(x_j, y_j) := \delta_{ij}$$

Somit genügt es nun gegen jedes Element der Basis zu testen.

Die auf S_h eingeschränkte Lösung u besitzt bezüglich der endlichen Basis die Darstellung $u_h(x, y) := \sum_{k=1}^N u_{h,k} \Psi_k(x, y)$.

Es folgt also:

$$a(u, \Psi_i) = \sum_{j=1}^N u_{h,j} \cdot a(\Psi_j, \Psi_i) = (f, \Psi_i)_{L^2} \quad \forall 1 \leq i \leq N$$

1 Mehrgitterverfahren zur Lösung von Linearen Gleichungssystemen

Rechnet man die Integrale der Bilinearform aus ergibt sich für $z_i := (x_i, y_i) \in \Omega_h$:

$$\int_{\Omega} \nabla \Psi_i \nabla \Psi_j d(x, y) = \begin{cases} 4 & i = j \\ -1 & \|z_i - z_j\|_2 = h \\ 0 & \text{sonst} \end{cases}$$

Die rechte Seite definieren wir:

$$b_i := (f, \Psi_i)_{L^2}$$

Somit erhält man ein lineares Gleichungssystem $Au_h = b$ mit dem Stern:

$$\begin{bmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 1 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{bmatrix}$$

Dieser entspricht *zufällig* dem Stern, der sich auch durch finite Differenzen ergeben würde. Da es sich mit finiten Differenzen wesentlich leichter und übersichtlicher umgehen lässt, werden wir im folgenden so tun, als könnten wir Ω mit einem Gitter von $(n-1) \cdot (n-1) =: N$ inneren Punkten überziehen, welches man verfeinert, indem man zwischen je zwei Punkten einen weiteren einfügt und vergrößert indem man jeden zweiten Punkt weg lässt. Der Abstand zwischen zwei Punkten von Ω_h entspricht also immer der Maschenweite $h := \frac{1}{n-1}$

Desweiteren wollen wir unseren Vektor u_k , $1 \leq k \leq N$ zweifach indizieren, so dass u_{ij} , $1 \leq i, j \leq n-1$ den Funktionswerten $u_h(x_i, y_j)$, $\forall (x_i, y_j) \in \Omega_h$ entspricht.

1.1.2 Konvergenzrate und Fehlerreduktion

Fakt 1.1.1 *Ist M die zu einem Iterationsverfahren gehörige Iterationsmatrix, so entspricht die Konvergenzrate dem Spektralradius $\rho(M)$. D.h. in jedem Schritt wird der Iterations-Fehler $|x - x^{(k)}|$ und den Faktor $\rho(M)$ verkleinert.*

Für unser Modellproblem und das Richardsonverfahren gilt:

Fakt 1.1.2 *Die Eigenwerte der Matrix A sind*

$$\lambda_{i,j}(A) = \frac{4}{h^2} \left(\sin^2 \left(\frac{i\pi}{2n} \right) + \sin^2 \left(\frac{j\pi}{2n} \right) \right) \quad 1 \leq i, j \leq n-1$$

Wählen wir als Dämpfungsfaktor Θ für die Richardsoniteration:

$$\Theta = \frac{1}{\lambda_{max}} = \frac{h^2}{4 \cdot \max_{i,j} \left(\sin^2 \left(\frac{i\pi}{2n} \right) + \sin^2 \left(\frac{j\pi}{2n} \right) \right)} \approx \frac{h^2}{8}$$

so ergibt sich für $\rho(M_{Rich}) = \rho(E - \Theta A)$:

$$\rho(M_{rich}) = 1 - \Theta \lambda_{min} = 1 - \frac{\lambda_{min}}{\lambda_{max}} \approx 1 - \frac{1}{2} \cdot \min_{i,j} \left(\sin^2 \left(\frac{i\pi}{2n} \right) + \sin^2 \left(\frac{j\pi}{2n} \right) \right) \leq 1 - O(h^2)$$

Dieser Wert wird angenommen, wenn der Iterationsfehler $e^k := x - x^k$ ein Vielfaches des zu λ_{min} gehörigen Eigenvektors ist.

Fakt 1.1.3 Die zu den $\lambda_{i,j}$ gehörigen Eigenvektoren von A sind:

$$v_{i,j}(x, y) := \frac{1}{2} h \sin(ix) \cdot \sin(jy) \quad 1 \leq i, j \leq n-1, \quad (x, y) \in \Omega_h$$

Fakt 1.1.4 Die $v_{i,j}$ bilden eine Basis des \mathbb{R}^N und sind orthonormal.

Für die niedrigen Frequenzen des Fehlers (i, j nahe bei Null) konvergiert das Verfahren also am langsamsten.

Wir betrachten nun den Unterraum V_{osz} der „stark oszillierenden“ Komponenten:

$$V_{osz} := \text{span} \left\{ v_{i,j}, \max\{i, j\} \geq \frac{n}{2} \right\}$$

Mindestens einer der beiden Indizes i, j liegt also im „hochfrequenten“ Teil des Frequenzintervalls $[1; n-1]$.

Wir wollen nun zeigen, dass im hypothetischen Fall $e^{(k)} \in V_{osz}$, das Richardson-Verfahren schnellere Konvergenz ergäbe.

Lemma 1.1.1 $x^{(0)}$ sei eine Approximation von x , so dass gilt: $e^{(0)} := x^{(0)} - x \in V_{osz}$. Dann liegen alle weiteren Fehler $e^{(k)}$ ebenfalls in V_{osz} und es gilt:

$$\|e^{(k)}\|_2 \leq \frac{3}{4} \|e^{(k-1)}\|_2$$

Beweis Da die $v_{i,j}$ eine orthonormale Basis bilden, können wir $\|x\|_2$ für jedes $x \in \mathbb{R}^N$ darstellen als:

$$\|x\|_2 = \left\| \sum_{i,j} \gamma_{i,j} v_{i,j} \right\|_2 = \sqrt{\sum_{i,j} |\gamma_{i,j}|^2}$$

wegen $Mv_{ij} = (1 - \Theta\lambda_{ij})v_{ij}$ liefert also die Anwendung von M auf den Iterationsfehler $e^{(k)}$ mit den Koeffizienten ξ_{ij} :

$$\begin{aligned} \|e^{(k+1)}\|_2^2 &= \|Me^{(k)}\|_2^2 \\ &= \left\| \sum_{i,j} \xi_{ij} (1 - \Theta\lambda_{ij}) v_{ij} \right\|_2^2 \\ &\leq \sum_{i,j} |\xi_{ij}|^2 |1 - \Theta\lambda_{ij}|^2 \\ &\leq \max_{i,j} |1 - \Theta\lambda_{ij}| \sum_{i,j} |\xi_{ij}|^2 \\ &= \max_{i,j} |1 - \Theta\lambda_{ij}| \cdot \|e^{(k)}\|_2^2 \end{aligned}$$

Wenn man nun das für das Maximum für die in V_{osz} vorkommenden i, j bildet, wobei oBdA $i \geq \frac{n}{2}$ gelte, ergibt sich:

$$\max_{i,j} |1 - \Theta\lambda_{ij}| = \max_{i,j} \left| 1 - \frac{\lambda_{ij}}{\lambda_{max}} \right| = \max_{i,j} \left| 1 - \frac{\sin^2(\frac{i\pi}{2n}) + \sin^2(\frac{j\pi}{2n})}{2} \right| \leq \left| 1 - \frac{\sin(\frac{\pi}{4})}{2} \right| = \frac{3}{4}$$

Liegt der Fehler also in V_{osz} , so verringert die Richardson-Iteration *unabhängig* von h den Fehler um den Faktor $\frac{3}{4}$.

1.1.3 Die Glättungseigenschaft

Dieses Ergebnis hat jedoch leider keinerlei praktische Bedeutung, da sich die Voraussetzungen für Lemma 1.1.1 nur mit Aufwand herstellen lassen, der annähernd so groß ist, wie für die exakte Lösung von $Ax = b$.

Wir können jedoch festhalten:

Folgerung 1.1.1 *Der Startfehler $e^{(0)}$ sei aufgespalten in:*

$$e^{(0)} = e_{osz}^{(0)} + e_{glatt}^{(0)}, \quad e_{osz}^{(0)} \in V_{osz}, \quad e_{glatt}^{(0)} \in V_{glatt} := V_{osz}^\perp$$

Dann gilt nach k Iterationen für den Fehler $e^{(k)} = e_{osz}^{(k)} + e_{glatt}^{(k)}$:

$$e_{glatt}^{(k)} = M^k e_{glatt}^{(0)}, \quad e_{osz}^{(k)} = M^k e_{osz}^{(0)}$$

wobei der glatte Fehler nur sehr langsam verringert wird, für den oszillierenden Anteil jedoch nach dem Lemma gilt:

$$\|M^k e_{osz}^{(0)}\|_2 \leq \left(\frac{3}{4}\right)^k \|e_{osz}^{(0)}\|_2$$

Das Richardsonverfahren dämpft also den „stark oszillierenden“ Anteil des Fehlers wesentlich stärker als den „glatten“.

Gleiches gilt auch für die „Jacobi-Iteration“, das „Gauß-Seidel Verfahren“ oder „SOR“. Solche Verfahren bezeichnen wir im folgenden als „Glättungsiterationen“ und verwenden für sie in Anlehnung an das englische „Smoothing“ die Bezeichnung \mathcal{S} .

k Glättungsschritte angewandt auf einen Vektor $x^{(p)}$ werden also notiert als $x^{(p+k)} = \mathcal{S}^k x^{(p)}$

1.2 Das Mehrgitterverfahren

Da wir nun also ein Verfahren haben, das in V_{osz} den Fehler kräftig dämpft, wollen wir nun ein Verfahren Φ für den komplementären Raum V_{glatt} finden. Die Verknüpfung $\Phi \circ \mathcal{S}$ besäße dann also in beiden Unterräumen gute Konvergenz.

Dabei verwenden wir die Idee, dass sich glatte Funktionen schon auf einem groben Gitter recht gut annähern lassen.

1.2.1 Die Idee der Grobgitterkorrektur

Sei $\tilde{x}_l := \mathcal{S}^k x(0)_l$ die Näherungslösung nach einigen wenigen (k) Glättungsschritten auf der Gitterebene l . Dann ist

$\tilde{e}_l := x_l - \tilde{x}_l$ der exakte Fehler auf l und damit die Korrektur, mit der man aus \tilde{x} die exakte Lösung berechnen könnte:

$$A(\tilde{x} + \tilde{e}) = Ax = b$$

Für den Fehler gilt also:

$$A_l \tilde{e}_l = A_l(x_l - \tilde{x}_l) = A_l x_l - A_l \tilde{x}_l = b_l - A_l \tilde{x}_l =: d_l$$

Die rechte Seite d_l nennen wir „Defekt“ auf der Gitterebene l .

Nach den letzten Abschnitten ist \tilde{e}_l glatt, weshalb es möglich sein sollte \tilde{e}_l schon auf einem groben Gitter der Ebene $l-1$ anzunähern:

$$\tilde{e}_l \approx p \tilde{e}_{l-1}$$

wobei p die „Prolongation“ $V_{l-1} \rightarrow V_l$ bezeichnet.

Könnte man also das LGS

$$A_{l-1} e_{l-1} = b_{l-1}, \quad b_{l-1} := r d_l$$

wobei r die „Restriktion“ $V_l \rightarrow V_{l-1}$ bezeichnet, auf Ω_h exakt lösen, so wäre eine sehr glatte Approximation von x gegeben.

1.2.2 Restriktion und Prolongation

Um unser Problem auf einem größeren Gitter betrachten und die dort erhaltenen Ergebnisse später auch wieder auf dem feinen Gitter verwenden zu können, sind Operatoren „ r “ (Restriktion) und „ p “ (Prolongation) nötig.

Dabei seien die Vektoren x_l und b_l aus $V_l \cong \mathbb{R}^{N_l}$, mit $N_l := \#\{(x, y) : (x, y) \in \Omega_h\}$.

Restriktion

Die Restriktion r ist eine lineare und surjektive Abbildung

$$r : V_l \rightarrow V_{l-1}, \quad l \geq 0$$

welche die Feingitterfunktionen in Grobgitterfunktionen abbildet.

Die einfachste solche Abbildung ist im Fall $\Omega_{l-1} \subset \Omega_l$ die „triviale Restriktion“ r_{triv} , welche einfach die Punkte die nicht in Ω_{l-1} liegen wegfallen läßt. Im Fall des Modellproblems also:

$$r_{triv} u_l(x, y) = u_{l-1}(x, y) \quad \forall (x, y) \in \Omega_{l-1}$$

In der Praxis verwendet man jedoch meistens das gewichtete Mittel der Nachbarwerte oder Ähnliches.

Prolongation

Die Prolongation p ist eine lineare und injektive Abbildung

$$p : V_{l-1} \rightarrow V_l, \quad l \geq 0$$

1 Mehrgitterverfahren zur Lösung von Linearen Gleichungssystemen

Für die Wahl der Prolongation gibt es viele Möglichkeiten. Bei unserem Modellproblem bietet sich die stückweise lineare Interpolation zwischen den Gitterpunkten von Ω_{l-1} an: Dazu übertragen wir zunächst alle Punkte von Ω_{l-1} nach Ω_l mittels

$$p u_{l-1}(x, y) := u_{l-1}(x, y)$$

Dann interpolieren wir zeilenweise ($\forall(x, y) \in \Omega_{l-1} : \frac{x}{h_l} \text{ gerade} \wedge \frac{y}{h_l} \text{ ungerade}$):

$$p u_{l-1}(x, y) := \frac{1}{2}(u_{l-1}(x + h_l, y) + u_{l-1}(x - h_l, y))$$

Dann spaltenweise ($\forall(x, y) \in \Omega_{l-1} : \frac{x}{h_l} \text{ ungerade} \wedge \frac{y}{h_l} \text{ gerade}$):

$$p u_{l-1}(x, y) := \frac{1}{2}(u_{l-1}(x + h_l, y) + u_{l-1}(x - h_l, y))$$

Und zuletzt die verbleibenden Punkte ($\forall(x, y) \in \Omega_{l-1} : \frac{x}{h_l} \text{ ungerade} \wedge \frac{y}{h_l} \text{ ungerade}$):

$$p u_{l-1}(x, y) := \frac{1}{4}(u_{l-1}(x + h_l, y) + u_{l-1}(x - h_l, y) + u_{l-1}(x, y + h_l) + u_{l-1}(x, y - h_l))$$

Dieses Verfahren (9 Punkte Prolongation) läßt sich auch als Stern darstellen:

$$\begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

Die Matrizen für Restriktion und Prolongation sind immer schwach besetzt, d.h. man wird sie niemals wie eine vollbesetzte Matrix abspeichern, sondern sie in Form eines Operators (eine Funktion, die zu gegebenem Vektor x die Restriktierte rx bzw Prolongierte px zurück liefert) bereit halten.

1.2.3 Der Mehrgitter-Algorithmus

Im folgenden bezeichne $\mathcal{S} := \mathcal{S}_l$ den Glätter auf der Gitterebene l . p und r entsprechen jeweils den zur aktuellen Gitterebene gehörigen Prolongationen p_l , bzw Restriktionen r_l .

Wir betrachten eine gegebene Näherung $x^{l,k}$ der Lösung im k -ten Zyklus auf der Gitterebene $l \geq 1$:

Mehrgitter-Iteration MGI₁:

Schritt 1: *A priori Glättung*

Man führe ν_1 Glättungsschritte durch:

$$x^{l,k} = \mathcal{S}^{\nu_1} x^{l,k}$$

Schritt 2: *Grobitterkorrektur*

Man berechne den Defekt

$$d_l = b_l - A_l x^{l,k}$$

und anschließend b_{l-1} indem man den Defekt auf das nächst gröbere Gitter restringiert:

$$b_{l-1} = rd_l$$

Sei nun

$$A_{l-1}\tilde{y}^{l-1} = b_{l-1}$$

Für $l \geq 1$ berechne man eine Näherung y^{l-1} von \tilde{y}^{l-1} indem man μ Schritte von MGI_{l-1} mit dem Startwert $x^{l-1,0} = 0$ ausführt.

Ist $l = 0$ berechne man die Lösung y^{l-1} exakt und setze $\tilde{y}^{l-1} = y^{l-1}$.

Jetzt setzt man

$$x_{l,k} = x_{l,k} + py_{l-1}$$

Schritt 3: A posteriori Glättung

Man führe ν_2 Glättungsschritte durch:

$$x_{l,k} = S^{\nu_2} x^{l,k}$$

Die Wahl der p_l und r_l sowie der Parameter ν_1 , ν_2 und μ haben zum Teil erheblichen Einfluß auf die Konvergenzgeschwindigkeit. Hierauf soll auch nicht weiter eingegangen werden, es sei nur angemerkt, dass man eine Iteration mit $\mu = 1$ als „V-Cycle“ und mit $\mu = 2$ als „W-Cycle“ bezeichnet.

Desweiteren wählt man der Einfachheit halber oft $\nu_2 = 0$ - im Falle des „V-Cycles“ ist es jedoch besser $\nu_1 = \nu_2$ zu wählen.

1.2.4 Konvergenz