

Numerik partieller Differentialgleichungen II

Vorlesungsskript WS 2013/2014

Vorlesungsmitschrift von Jeronim Morina

Inhaltsverzeichnis

Vorwort

Dieses Dokument enthält die Mitschrift von Jeronim Morina zur Vorlesung “Numerik partieller Differentialgleichungen II” im Wintersemester 2014 bei Professor Klawonn. Wir können dem Leser weder Vollständigkeit noch Fehlerfreiheit (von dieser sind wir überzeugt, dass sie definitiv nicht gegeben ist) versprechen. Wir sind jedoch für Verbesserungsvorschläge dankbar, diese können an morina@jeronim.de geschickt werden.

Bonn, 26. Oktober 2013

Einleitung

Ziele:

- Schnellere Algorithmen für Parallelrechner entwickeln

Literaturangaben:

-

1 Iterationsverfahren für (große) lineare Gleichungssysteme

Gegeben: $Ax = b$, $A \in \mathbb{R}^{n \times n}$, A invertierbar, $x, b \in \mathbb{R}^n$ mit $x = A^{-1}b$ Ziel: Konstruktion eines iterativen Verfahrens zur Lösung von * der Form:

$$x^{(k+1)} := \varphi(x^{(k)}), \quad k \in \mathbb{N}, x^{(0)} \in \mathbb{R}^n \quad (1.0.1)$$

Iterationsfunktion $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ gegeben.

1.1 Das Richardson-Verfahren

Ansatz: Sei $M \in \mathbb{R}^{n \times n}$ invertierbar. Wir betrachten das folgende Splitting:

$$Mx + \underbrace{(A - M)x}_{=N} = b \quad (1.1.1)$$

und setzen

$$Mx^{(k+1)} + Nx^{(k)} = b, \quad k = 0, 1, 2, \dots \Leftrightarrow x^{(k+1)} = x^{(k)} + M^{-1}(b - Ax^{(k)}) = x^{(k)} + M^{-1}r^{(k)} \quad \text{mit } r^{(k)} := b - Ax^{(k)} \quad (1.1.2)$$

Dieses Iterationsverfahren lässt sich durch Einführen eines Relaxationsparameters $\alpha \in \mathbb{R}$ noch verallgemeinern.

$$x^{(0)} \in \mathbb{R}^n \text{ geg. } x^{(k+1)} := x^{(k)} + \alpha M^{-1}r^{(k)}, \quad k = 0, 1, 2, \dots \quad (1.1.3)$$

heißt (stationäres) Richardson-Verfahren. \Rightarrow Iterationsvorschrift: $\varphi(x) = Bx + c$ $B = I - \alpha M^{-1}A$, $C = \alpha M^{-1}b$

Für die Konvergenz kennen wir aus Numerik I folgenden Satz:

Satz 1.1. Das Iterationsverfahren (label) konvergiert genau dann, wenn

$$\rho(B) < 1 \quad (1.1.4)$$

wobei $\rho(B) := \max_{\lambda \in \sigma(B)} |\lambda|$ und $\sigma(B) = \{\lambda \in \mathbb{C} : \lambda \text{ EW von } B\}$

Hinreichende Konvergenzbedingung ist $\|B\| < 1$ wobei $\|\cdot\|$ eine beliebige Matrixnorm ist, die durch eine Vektornorm induziert wird.

$$\Rightarrow \|e^{(k)}\| \leq \eta^k \|e^{(0)}\|, \quad \eta := \|B\|, e^{(k)} = x - x^{(k)} \text{ Fehler im } k\text{-ten Schritt}$$

Satz 1.2. Das stationäre Richardson-Verfahren (label) mit $\alpha \neq 0$ konvergiert genau dann, wenn

$$2 * \frac{\operatorname{Re}(\lambda_i)}{\alpha |\lambda_i|^2} > 1, \quad \forall i = 1, \dots, n \quad (1.1.5)$$

wobei $\lambda_i \in \sigma(M^{-1}A)$, $i = 1, \dots, n$

Beweis. Wende (label itkonvergenz) auf $B = I - \alpha M^{-1}A$ an. Sei $\mu_i \in \sigma(B)$. Dann gilt

$$\Leftrightarrow \mu_i = 1 - \alpha \lambda_i, \quad \lambda_i \in \sigma(M^{-1}A) \rho(B) < 1 \Leftrightarrow |1 - \alpha \lambda_i| < 1, \quad \forall \lambda_i \in \sigma(M^{-1}A) \Leftrightarrow -2\alpha \operatorname{Re}(\lambda_i) + \alpha^2 |\lambda_i|^2 < 0$$

□

Satz 1.3. Voraussetzungen: $M^{-1}A$ habe nur positive reelle Eigenwerte $\lambda_i, i = 1, \dots, n$ (Invertierbarkeit nicht zwingend erforderlich), die wie folgt geordnet seien:

$$0 < \lambda_n \leq \lambda_{n-1} \leq \dots \leq \lambda_1$$

1 Dann ist das stationäre Richardson-Verfahren genau dann konvergent, wenn

$$0 < \alpha < \frac{2}{\lambda_1}$$

2 Der Spektralradius $\rho(B)$ ist minimal (und damit die Konvergenz am schnellsten), wenn

$$\alpha = \alpha_{opt} := \frac{2}{\lambda_1 + \lambda_n}$$

Dann gilt

$$\rho_{opt} = \min_{\alpha} (\rho(B_{\alpha})) = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} = \frac{\kappa - 1}{\kappa + 1} = 1 - \frac{2}{\kappa + 1}$$

mit $\kappa = \kappa(M^{-1}A) = \frac{\lambda_1}{\lambda_n}$ Konditionszahl

Beweis.

$$\mu_i \in \sigma(B) \Leftrightarrow \mu_i = 1 - \alpha \lambda_i, \quad \lambda_i \in \sigma(M^{-1}A)$$

Aus Satz 1.1.1: Richardson-Verfahren konvergiert $\Leftrightarrow |\mu_i| < 1, \quad \forall i = 1, \dots, n$

$$\Leftrightarrow |1 - \alpha \lambda_i| < 1 \Leftrightarrow -1 < 1 - \alpha \lambda_i < 1 \Leftrightarrow 0 < \alpha < \frac{2}{\lambda_i} \Leftrightarrow 0 < \alpha < \frac{2}{\lambda_1} = \min_i \frac{2}{\lambda_i}$$

mit $\rho(B) = \max_i |1 - \alpha \lambda_i| = \max \{|1 - \alpha \lambda_1|, |1 - \alpha \lambda_n|\}$ Der kleinste Spektralradius ρ_{opt} ergibt sich durch Schnittpunktbildung der beiden Geraden. Berechne den opt Schnittpunkt von $|1 - \alpha \lambda_1|$ und $|1 - \alpha \lambda_n|$ BILD

$$-(1 - \alpha \lambda_n) = 1 - \alpha \lambda_1 \Leftrightarrow \alpha_{opt} = \frac{2}{\lambda_1 + \lambda_n} \Rightarrow \rho_{opt} = 1 - \alpha_{opt} \lambda_n = 1 - \frac{2}{\lambda_1 + \lambda_n} \lambda_n = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$$

□

Die Konvergenz des stationären Richardson-Verfahrens ist also abhängig von der Konditionszahl von $M^{-1}A$ bzw. vom größten und kleinsten Eigenwert:

$$\lambda_1 := \max \lambda_i, \quad \lambda_i \in \sigma(M^{-1}A) \quad \lambda_n := \min \lambda_i$$

Bemerkung 1.4. Die Matrix M ist frei wählbar und man kann die Kondition des Systems damit a-priori beeinflussen. Daher nennt man sie auch Vorkonditionierer (engl. preconditioner) oder Vorkonditionierungsmatrix

- M sollte eine gute Approximation an A sein und $\kappa(M^{-1}A)$ sollte möglichst klein sein
- M^{-1} sollte sich einfach auf einem Vektor anwenden lassen, d.h. mit vertretbarem Rechenaufwand.

Bemerkung 1.5. Man spricht beim Verfahren $(^{**})$ auch vom vorkonditioniertem Richardson-Verfahren ($M \neq I$). Nachteil: Für optimale Konvergenz muss Information über λ_1 und λ_n vorliegen (die Ermittlung dieser, kann mitunter so teuer wie die Berechnung des ganzen Systems sein).

Algorithmus 1.6 (Vorkonditioniertes Richardson-Verfahren). Init: 1. Geg. $X^{(0)} \in \mathbb{R}^n$ (Startvektor), $\alpha \in \mathbb{R}$
 $\{0\}$ 2. Berechne das Startresiduum $r^{(0)} := b - Ax^{(0)}$ Iteration: while $\|r^{(k)}\| \leq \varepsilon \|r^{(0)}\|$ für gegebenes ε
 $y^{(k)} := M^{-1}r^{(k)}$

Konventionen: $y^{(k)} = M^{-1}r^{(k)}, x^{(k+1)} = Bx^{(k)} + C, \quad B = I - \alpha M^{-1}A, C = \alpha M^{-1}b$
 Voraussetzung: $M^{-1}A$ hat nur positive reelle Eigenwerte $0 < \lambda_n \leq \lambda_{n-1} \leq \dots \leq \lambda_1$
 Das vorkonditionierte Richardson-Verfahren konvergiert $\Leftrightarrow 0 < \alpha < \frac{2}{\lambda_1}$. Die optimale Konvergenzrate liegt dann bei $\alpha = \alpha_{opt} = \frac{2}{\lambda_1 + \lambda_n}$ M bzw. M^{-1} bezeichnet man als Vorkonditionierer. Unter der Konditionszahl $\kappa(M^{-1}A)$ verstehen wir folgenden Ausdruck:

$$\kappa\left(M^{-\frac{1}{2}}AM^{-\frac{1}{2}}\right) = \frac{\lambda_{\max}\left(M^{-\frac{1}{2}}AM^{-\frac{1}{2}}\right)}{\lambda_{\min}\left(M^{-\frac{1}{2}}AM^{-\frac{1}{2}}\right)} \underbrace{=}_{A, M \text{ s.p.d.}} \frac{\lambda_{\max}(M^{-1}A)}{\lambda_{\min}(M^{-1}A)}$$

1.2 Das Gradientenverfahren

Voraussetzung: $A \in \mathbb{R}^{n \times n}$, s.p.d. Dann löst $x \in \mathbb{R}^n \quad Ax = b \Leftrightarrow \phi(x) = \min_{y \in \mathbb{R}^n} \phi(y)$, mit $\phi(y) := \frac{1}{2}y^T Ay - y^T b$ Betrachte nun das Richardson-Verfahren, aber mit der Möglichkeit $\alpha = \alpha_k$ in jedem Schritt individuell zu wählen. Dann erhalten wir $\Rightarrow x^{(k+1)} = x^{(k)} + \alpha_k M^{-1}r^{(k)}, \quad r^{(k)} := b - Ax^{(k)} = Ax - Ax^{(k)} = Ae^{(k)}$ Wie zuvor ist: $e^{(k+1)} := e^{(k+1)}(\alpha_k) := \underbrace{(I - \alpha_k M^{-1}A)}_{=: B^{(k)}} e^{(k)}$ Sei

$$\|x\|_A := \sqrt{x^T Ax} = \sqrt{(x, x)_A}, \quad (x, y)_A := y^T Ax \quad (1.2.1)$$

Dann gilt

$$\|e^{(k+1)}\|_A^2 = (Ae^{(k+1)}, e^{(k+1)})_2 = (r^{(k+1)}, e^{(k+1)})_2 \quad (1.2.2)$$

1 Iterationsverfahren für (große) lineare Gleichungssysteme

$$\text{i } e^{(k+1)} = (I - \alpha_k M^{-1} A) e^{(k)} - \alpha_k M^{-1} A e^{(k)}$$

$$\text{ii } r^{(k+1)} = r^{(k)} - \alpha_k A M^{-1} A e^{(k)}$$

$$\Rightarrow \|e^{(k+1)}\|_A^2 \underbrace{=}_{y^{(k)}=M^{-1}r^{(k)}} (r^{(k)}, e^{(k)}) - \alpha_k ((Ay^{(k)}, e^{(k)}) + (r^{(k)}, M^{-1}Ae^{(k)})) + \alpha_k^2 (Ay^{(k)}, M^{-1}Ae^{(k)}) \underbrace{\Rightarrow}_{e^{(k)}A^{-1}r^{(k)}} \quad (1.2.3)$$

Das Residuum sollte klein werden. Deshalb minimieren wir den Fehler in der Energienorm.

Jetzt wählen wir α_k im $k+1$ Iterationsschritt so, dass der Fehler in der A -Norm minimiert wird, d.h. wir minimieren die Funktion

$$f(\alpha_k) := \|e^{(k+1)}(\alpha_k)\|_A^2$$

Betrachte

$$0 = \frac{\partial f}{\partial \alpha}(\alpha_k) = -2(y^{(k)}, r^{(k)}) + 2\alpha_k (Ay^{(k)}, y^{(k)}) \Leftrightarrow \alpha_k = \frac{y^{(k)}, r^{(k)}}{Ay^{(k)}, y^{(k)}} \quad (1.2.4)$$

$$\frac{\partial^2 f}{\partial \alpha_k^2}(\alpha_k) \underbrace{=}_{Aspd} 2(Ay^{(k)}, y^{(k)}) > 0 \quad (1.2.5)$$

α_k ist tatsächlich ein lokales Minimum

$$\|e^{(k+1)}\|_A^2 \underbrace{=}_{Def} (Ae^{(k+1)}, e^{(k+1)}) = \underbrace{=}_{Def} (A(x - x^{(k+1)}), (x - x^{(k+1)})) \underbrace{=}_{Ax=b} (b, x) - \underbrace{(b, x^{(k+1)}) - (x^{(k+1)}, b)}_{=-2(b, x^{k+1})} + (Ax^{(k+1)}, x^{(k+1)}) \quad (1.2.6)$$

Also ist die lokale Minimierung des Fehlers $e^{(k+1)}$ in der A -Norm äquivalent zur Minimierung des Funktionals $\phi(x^{(k+1)})(\alpha_k)$ unter allen Vektoren $X^{(k+1)}(\alpha_k)$ der Form $x^{(k)} + \alpha_k M^{-1} r^{(k)}$ durch Lösen 1-dim. Minimierungsprobleme.

Geometrische Veranschaulichung für $M = I$. Sei $\phi(x)$ die Höhenfunktion. \Rightarrow steilster Abstieg $-\nabla_x \phi(x) = -(Ax - b) = b - Ax \Rightarrow -\nabla_x \phi(x^{(k)}) = r^{(k)} \Rightarrow$ Gradientenwert: $X^{(k+1)} + \alpha_k r^{(k)} = x^{(k)} - \alpha_k \nabla_x \phi(x^{(k)})$. Das Gradientenverfahren minimiert lokal in Richtung des steilsten Abstiegs.

Satz 1.7 (Kantorowitsch Ungleichung). Sei $A \in \mathbb{R}^{n \times n}$ s.p.d. mit spektraler Konditionszahl $\kappa(A) := \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$. Dann gilt für jeden Vektore $0 \neq x \in \mathbb{R}^n$ die Ungleichung

$$\frac{(x^T A x)(x^T A^{-1} x)}{(x^T x)^2} \leq \frac{(\lambda_{\max} + \lambda_{\min})^2}{4\lambda_{\max}\lambda_{\min}} \quad (1.2.7)$$

Beweis. Die Eigenwerte von A seien geordnet $0 < \lambda_{\min} = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = \lambda_{\max}$. Diagonalisierung von A durch Orthonormales Q

$$Q^T A Q = D = \text{diag}_{i=1, \dots, n}(\lambda_i) \Rightarrow A^{-1} = (Q D Q^T)^{-1} = Q D^{-1} Q^T \Rightarrow \frac{(x^T A x)(x^T A^{-1} x)}{(x^T x)^2} = \frac{(x^T Q D Q^T x)(x^T Q D^{-1} Q^T x)}{(x^T x)^2} \quad (1.2.8)$$

Sei $g : \lambda \rightarrow \frac{1}{\lambda}$, dann liegen alle Punkte $(\lambda_i, \frac{1}{\lambda_i})$ auf dem Graphen von g . BILD $g''(\lambda) > 0 \quad \forall \lambda \in [\lambda_{\min}, \lambda_{\max}] \Rightarrow g$ ist konvexe Funktion. Daher liegen alle Punkte $(\lambda, \frac{1}{\lambda})$ mit $\lambda_{\min} < \lambda < \lambda_{\max}$ unterhalb der Geraden durch die Punkte $(\lambda_{\min}, \frac{1}{\lambda_{\min}})$ und $(\lambda_{\max}, \frac{1}{\lambda_{\max}})$. Der Punkt $(\sum_{i=1}^n z_i \lambda_i, \sum_{i=1}^n z_i \lambda_i^{-1})$ liegt in der konvexen Hülle der Punkte $(\lambda_i, \frac{1}{\lambda_i}), i = 1, \dots, n$. Daher liegen die konvexen Kombinationen $(\sum_{i=1}^n z_i \lambda_i, \sum_{i=1}^n z_i \lambda_i^{-1})$ in der schraffierten Fläche, insbesondere unterhalb der Geraden durch $(\lambda_{\min}, \frac{1}{\lambda_{\min}})$ und $(\lambda_{\max}, \frac{1}{\lambda_{\max}})$

$$\Rightarrow \lambda \rightarrow \frac{1}{\lambda_{\min}} + \frac{\frac{1}{\lambda_{\max}} - \frac{1}{\lambda_{\min}}}{\lambda_{\max} - \lambda_{\min}} (\lambda - \lambda_{\min}) = \frac{\lambda_{\min} + \lambda_{\max} - \lambda}{\lambda_{\max} - \lambda_{\min}} \Rightarrow \text{FEHLT}$$

Standardergebnis Analysis:

$$\max_{\lambda_{\min} \leq \lambda \leq \lambda_{\max}} \left(\lambda \frac{\lambda_{\max} + \lambda_{\min} - \lambda}{\lambda_{\max} \lambda_{\min}} \right) = \frac{(\lambda_{\min} + \lambda_{\max})^2}{4 \lambda_{\max} \lambda_{\min}} \quad (1.2.9)$$

Hieraus folgt die Behauptung mit $\lambda := \sum_{i=1}^n z_i \lambda_i \in (\lambda_{\min}, \lambda_{\max})$ □

Folgerung: Sei $M \in \mathbb{R}^{n \times n}$ spd, dann gilt

$$\frac{(M^{-1}Ax, x)_M}{(x, x)_M} \leq \frac{(\lambda_{\min} + \lambda_{\max})^2}{4 \lambda_{\min} \lambda_{\max}} \quad (1.2.10)$$

λ_{\min} = kleinster Eigenwert von $M^{-1}A$ und λ_{\max} = größter Eigenwert von $M^{-1}A$

Letzter Eintrag. Folgerung: $M \in \mathbb{R}^{n \times n}$, s.p.d., dann gilt:

$$\frac{(M^{-1}Ax, x)_M (A^{-1}Mx, x)_M}{(x, x)_M^2} \leq \frac{(\lambda_{\max} + \lambda_{\min})^2}{4 \lambda_{\max} \lambda_{\min}}$$

wobei $\lambda_{\max} = \lambda_{\max}(M^{-1/2}AM^{-1/2})$, $\lambda_{\min} = \lambda_{\min}(M^{-1/2}AM^{-1/2})$.

Es gilt:

$$A = Q^T D Q, \quad A^{1/2} := Q^T D^{1/2} Q, \quad D^{1/2} := \text{diag}_{i=1, \dots, n}(\sqrt{\lambda_i})$$

Beweis der Folgerung:

$$y := M^{1/2}x$$

$$\begin{aligned} & \frac{(M^{-1}Ax, x)_M (A^{-1}Mx, x)_M}{(x, x)_M^2} \\ & \stackrel{y=M^{1/2}x}{=} \frac{(AM^{-1/2}y, M^{-1/2}y)(A^{-1}M^{1/2}y, M^{1/2}y)}{(y, y)^2} \\ & = \frac{(M^{-1/2}AM^{-1/2}y, y)(M^{1/2}A^{-1}M^{1/2}y, y)}{(y, y)^2} \\ & \stackrel{\tilde{A}:=M^{-1/2}AM^{-1/2}}{=} \frac{(\tilde{A}y, y)(\tilde{A}^{-1}y, y)}{(y, y)^2} \\ & \stackrel{\text{Kant. Ungl.} + \tilde{A} \text{ s.p.d.}}{=} \frac{(\lambda_{\max}(\tilde{A}) + \lambda_{\min}(\tilde{A}))^2}{4 \lambda_{\max}(\tilde{A}) \lambda_{\min}(\tilde{A})} \end{aligned}$$

$$\lambda(\tilde{A}) = \lambda(M^{-1}A)$$

Es ist : $e^{(k)} := x - x^{(k)}$, $y^{(k)} := M^{-1}r^{(k)}$, $Ax = b$, $r^{(k)} = b - Ax^{(k)}$.

$$\frac{\|e^{(k)}\|_A^2 - \|e^{(k+1)}\|_A^2}{\|e^{(k)}\|_A^2} = \frac{(y^{(k)}, y^{(k)})_M}{(M^{-1}Ay^{(k)}, y^{(k)})_M} (A^{-1}My^{(k)}, y^{(k)})_M$$

$$\Leftrightarrow \|e^{(k+1)}\|_A^2 \leq \left(1 - \frac{4\lambda_{\max}\lambda_{\min}}{(\lambda_{\max} + \lambda_{\min})^2}\right) \|e^{(k)}\|_A^2 = \frac{(\lambda_{\max} - \lambda_{\min})^2}{(\lambda_{\max} + \lambda_{\min})^2} \|e^{(k)}\|_A^2$$

$$\Leftrightarrow \|e^{(k+1)}\|_A \leq \frac{(\lambda_{\max} - \lambda_{\min})}{\lambda_{\max} + \lambda_{\min}} \|e^{(k)}\|_A = \left(\frac{\kappa(M^{-1}A) - 1}{\kappa(M^{-1}A) + 1}\right) \|e^{(k)}\|_A \leq \left(\frac{\kappa(M^{-1}A) - 1}{\kappa(M^{-1}A) + 1}\right)^k \|e^{(0)}\|_A$$

Satz 1.8. Beim instationären (vorkonditionierten) Richardson-Verfahren/(vorkonditioniertes) Gradientenverfahren gilt für die Konvergenzrate folgende obere Schranke:

$$\kappa \frac{(M^{-1}A) - 1}{\kappa(M^{-1}A) + 1} \quad (1.2.11)$$

wobei $\kappa(M^{-1}A) := \kappa_2(M^{-1/2}AM^{-1/2}) = \frac{\lambda_{\max}(M^{-1/2}AM^{-1/2})}{\lambda_{\min}(M^{-1/2}AM^{-1/2})}$

Algorithmus 1.9 (vorkonditioniertes Gradientenverfahren). Initialisierung:

- 1 geg. Startvektor $x^{(0)} \in \mathbb{R}^n$
- 2 $r^{(0)} := b - Ax^{(0)}$ (Startresiduum)

Iteration: $k = 0, 1, 2, \dots$ solange Konvergenzkriterium erfüllt, z.B. $\|r^{(k)}\|_2 \leq \varepsilon \|r^{(0)}\|_2$

$$y^{(k)} := M^{-1}r^{(k)}$$

$$q^{(k)} := Ay^{(k)}$$

$$\alpha^{(k)} := \frac{(y^{(k)}, r^{(k)})}{(q^{(k)}, y^{(k)})}$$

$$x^{(k+1)} := x^{(k)} + \alpha^{(k)}y^{(k)}$$

$$r^{(k+1)} := r^{(k)} - \alpha^{(k)}q^{(k)}$$

Bemerkung 1.10. Man beachte, dass in dem vorliegenden Algorithmus in jeder Iteration jeweils mindestens eine Matrix-Vektor-Multiplikation mit A bzw. M^{-1} benötigt wird.

Das Verfahren der konjugierten Gradienten

Beim Gradientenverfahren haben wir zwei Phasen kennengelernt, um $x^{(k+1)}$ aus $x^{(k)}$ zu berechnen:

- Bestimmen der Suchrichtungen $y^{(k)}$
- Berechnen des lokalen Minimums von Φ in dieser Richtung

Bemerkung 1.11 (Beobachtung). 2) ist unabhängig von 1);

Ist eine beliebige Suchrichtung $p^{(k)}$ gegeben, so bestimme den Relaxationsparameter $\alpha^{(k)}$ derart, dass $\Phi(x^{(k)} + \alpha^{(k)}p^{(k)})$ minimal wird. Dies kann man wie beim Gradientenverfahren machen, d.h.

$$\alpha^{(k)} := \frac{(p^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})}$$

Frage: Kann man $p^{(k)}$ besser wählen als im Gradientenverfahren? (Wie will man Optimalität überhaupt definieren?) Dazu betrachten wir folgende Definition der Optimalität von Suchrichtungen (\rightarrow Die Suchrichtung werden orthogonal zu den Residuen gewählt)

Definition 1.12. Eine Richtung $x \in \mathbb{R}^n$ heißt optimal bzgl. einer Richtung $p \in \mathbb{R}^n \equiv \Phi(x) \leq \Phi(x + \lambda p) \quad \forall \lambda \in \mathbb{R} \quad \Phi(y) := \frac{1}{2}y^T Ay - b^T y$

Bem. 1.2.2.:

x optimal bzgl. $p \Leftrightarrow p \perp r := B - Ax$, wobei $p \perp r \Leftrightarrow (p, r) = 0$

Beweis: $\phi(x) = \frac{1}{2}x^T Ax - b^T x$

$$\lambda \in \mathbb{R} : \phi(x + \lambda p) = \phi(x) + \lambda \underbrace{(Ax - b, p)}_{=-r} + \frac{\lambda^2}{2} \underbrace{(Ap, p)}_{\geq 0}$$

Dann gilt:

$$\begin{aligned} \phi(x) &\leq \phi(x + \lambda p) \quad \forall \lambda \in \mathbb{R} \\ \Leftrightarrow 0 &\leq -\lambda(r, p) + \underbrace{\frac{\lambda^2}{2}(Ap, p)}_{\geq 0} \quad \forall \lambda \in \mathbb{R} \\ \Leftrightarrow (r, p) &= 0 \\ \Leftrightarrow r &\perp p \end{aligned}$$

Für das Gradientenverfahren mit $M = I$ gilt:

$x^{(k+1)}$ ist optimal bzgl. $r^{(k)}$ ($\alpha^{(k)}$ wird gerade so gewählt, dass $r^{(k+1)} \perp r^{(k)}$ ist).

Im Allgem. ist $x^{(k+2)}$ nicht mehr optimal bzgl. $r^{(k)}$.

Frage: Gibt es Abstiegsrichtungen die diese Optimalität erhalten?

Dazu sei $x^{(k+1)} = x^{(k)} + q$, wobei $x^{(k)}$ optimal sei bzgl. einer Richtung p , d.h. $p \perp r^{(k)}$. Wir verlangen nun, dass $x^{(k+1)}$ auch bzgl. p optimal sein soll:

$$r^{(k+1)} \perp p \Leftrightarrow 0 = (r^{(k+1)}, p) = (r^{(k)} - Aq, p) = -(Aq, p)$$

Also gilt:

$$r^{(k+1)} \perp p \Leftrightarrow (q, p)_A = (Aq, p) = 0 \Leftrightarrow q \text{ ist } A\text{-orth. zu } p$$

Bemerkung 1.13. Vorkonditionierte Verfahren sind für uns so wichtig, weil das Optimalitätskriterium der Orthogonalität in der Realität auf Grund von Rundungsfehlern (insbesondere bei schlecht konditionierten Problemen) diese verwischen würde. Ein Vorkonditionierer hilft uns dem entgegenzuwirken.

Bemerkung 1.14 (Folgerung). Um die Optimalität aufeinander folgender Iterationen zu gewährleisten, müssen die Abstiegsrichtungen gegenseitig A -orthogonal sein, d.h. $(p, q)_A = 0$. Dies nennt man A -konjugiert oder konjugiert zueinander. Ein Verfahren, welches konjugierte Abstiegsrichtungen verwendet, nennt man konjugiertes Verfahren.

Bemerkung 1.15 (vorgehensweise). Sei $p^{(0)} := r^{(0)}$, dann konstruiere Richtungen der Form

$$p^{(k+1)} := r^{(k+1)} - \beta^{(k)} p^{(k)}, \quad k = 0, 1, \dots \quad (1.2.12)$$

Wähle dabei $\beta^{(k)} \in \mathbb{R}$, so dass

$$(Ap^{(j)}, p^{(k+1)}) = 0 \quad \forall j = 0, 1, 2, \dots, k \quad (1.2.13)$$

d.h. $p^{(k+1)}$ soll konjugiert sein zu allen vorherigen Richtungen $p^{(j)}, j = 0, \dots, k$. Aus ref 1,2 und $j = k$ folgt $\beta^{(k)} = \frac{(p^{(k)}, r^{(k+1)})_A}{(p^{(k)}, p^{(k)})_A}$

1 Iterationsverfahren für (große) lineare Gleichungssysteme

Es bleibt zu Zeigen: ref 2 gilt dann auch für $j = 0, \dots, k$

Beweis. vollständige Induktion (Übung)

□

Algorithmus 1.16 (Konjugiertes Gradientenverfahren (cg-Verfahren)).

$$\alpha^{(k)} := \frac{(p^{(k)}, r^{(k)})}{(p^{(k)}, p^{(k)})}, \quad \beta^{(k)} := \frac{(p^{(k)}, r^{(k+1)})_A}{(p^{(k)}, p^{(k)})_A}$$

$$\text{mit } p^{(0)} := r^{(0)}, r^{(0)} = b - Ax^{(0)} \quad k = 0, 1, 2, \dots \quad p^{(0)} := r^{(0)} = b - Ax^{(0)}$$

berechne $\alpha^{(k)}$

$$x^{(k+1)} := x^{(k)} + \alpha^{(k)} p^{(k)} \quad (*) \quad r^{(k+1)} := r^{(k)} - \alpha^{(k)} A p^{(k)} \quad (**)$$

berechne $\beta^{(k)}$

$$p^{(k+1)} := r^{(k+1)} - \beta^{(k)} p^{(k)} \quad (***)$$

Es gilt auch (Übungen):

$$\alpha^{(k)} = \frac{(r^{(k)}, r^{(k)})}{(A p^{(k)}, p^{(k)})}, \quad \beta^{(k)} = \frac{(r^{(k+1)}, r^{(k+1)})}{(r^{(k)}, r^{(k)})}$$

Satz 1.17 (1.3.3). Sei A s.p.d. Dann konvergiert das cg-Verfahren und es gilt folgende Fehlerabschätzung:

$$\|e^{(k)}\|_A \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \|e^{(0)}\|, \quad k = 0, 1, \dots$$

mit $\kappa = \kappa(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$

Beweis. Folgt den Argumenten von Quaternioni/Valle Numerical Mathematics for PDE, Springer, S.48-50 und Quarteroni/Suer/Sater, Numerical Mathematics, S.154

$$1) \quad (*) \quad x^{(k+1)} = x^{(k)} + \alpha^{(k)} p^{(k)} \quad \underbrace{=}_{\text{Argument}} \quad x^{(0)} + \sum_{i=0}^k \alpha^{(i)} p^{(i)}$$

2) Sei $x = A^{-1}b$ die exakte Lösung, dann folgt aus 1)

$$(A p^{(j)}, x^{(k+1)} - x^{(0)}) = \sum_{i=0}^k \alpha^{(i)} \underbrace{(A p^{(j)}, p^{(i)})}_{=0, i \neq j} = \alpha^{(j)} (A p^{(j)}, p^{(j)}) \quad \underbrace{=}_{\text{Def. } \alpha^{(j)}} \quad (r^{(j)}, p^{(j)}) = (b - A x^{(j)}, p^{(j)}) = (A$$

3) Beh.:

$$\text{span}\{p^{(0)}, \dots, p^{(k)}\} = \text{span}\{r^{(0)}, A r^{(0)}, \dots, A^k r^{(0)}\} =: \underbrace{\mathcal{K}_{k+1}(r^{(0)})}_{\text{Krylovraum}}$$

Offensichtlich gilt: $\text{span}\{p^{(0)}, \dots, p^{(k)}\} = \text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(k)}\}$

Beweis. 3) mit vollst. Induktion;

IA: da $p^{(0)} = r^{(0)}$

' \subset ': IV: $\text{span}\{p^{(0)}, \dots, p^{(k)}\} = \mathcal{K}_{k+1}(r^{(0)})$

IS: IV $\Rightarrow Ap^{(k)} \in \mathcal{K}_{k+2}(r^{(0)})$

Aus (***) also $r^{(k+1)} = r^{(k)} - \alpha^{(k)} Ap^{(k)} \in \mathcal{K}_{k+2}(r^{(0)})$

$$(***) \Rightarrow p^{(k+1)} = r^{(k+1)} - \beta^{(k)} p^{(k)} \in \mathcal{K}_{k+2}(r^{(0)})$$

$$\Rightarrow \text{span}\{p^{(0)}, \dots, p^{(k+1)}\} \subset \mathcal{K}_{k+2}(r^{(0)})$$

' \supset '

IV $\text{span}\{p^{(0)}, \dots, p^{(k)}\} = \mathcal{K}_{k+1}(r^{(0)})$

IS: IV $\Rightarrow A^k r^{(0)} \in \text{span}\{p^{(0)}, \dots, p^{(k)}\}$

$$(**) \Rightarrow A^{k+1} r^{(0)} \in \text{span}\{r^{(0)}, \dots, r^{(k+1)}\} = \text{span}\{p^{(0)}, \dots, p^{(k+1)}\}$$

$$\Rightarrow ' \supset '$$

□

4) $2)+3) \Rightarrow x^{(k+1)} - x^{(k)}$ ist die Orthogonalprojektion des Anfangsfehlers $x - x^{(0)} =: e^{(0)}$ auf den Krylovraum $\mathcal{K}_{k+1}(r^{(0)})$.

5) es gilt mit 4) und $e^{(k+1)} = x - x^{(k+1)}$:

$$\begin{aligned} \|e^{(k+1)}\|_A &= \|(x - x^{(0)}) - (x^{(k+1)} - x^{(0)})\|_A \\ &= \min_{y \in \mathcal{K}_{k+1}(r^{(0)})} \|x - x^{(0)} - y\|_A = \min_{y \in \mathcal{K}_{k+1}(r^{(0)})} \|e^{(0)} - y\|_A \end{aligned}$$

Da $r^{(0)} = A(x - x^{(0)}) = Ae^{(0)}$ lässt sich ein bel. $y \in \mathcal{K}_{k+1}(r^{(0)})$ schreiben als:

$$y = \sum_{j=0}^k \tilde{\gamma}_j A^j r^{(0)} = \sum_{j=1}^{k+1} \gamma_j A^j e^{(0)} = \tilde{p}(A) e^{(0)}, \quad p \in \mathcal{P}_{k+1}$$

Also gilt $\|e^{(k+1)}\|_A = \min_{p \in \mathcal{P}_{k+1}^*} \|p(A) e^{(0)}\|_A$, wobei $\mathcal{P}_{k+1}^* = \{q \in \mathcal{P}_{k+1} | q(0) = 1\}$

6) A ist s.p.d. $\Rightarrow A = Q^T D Q$, Q orthogonale Matrix, $D = \text{diag}(\lambda_i)$, λ_i EV von A

$$\Rightarrow p(A) = Q^T p(D) Q$$

$$\Rightarrow \|e^{(k+1)}\|_A^2 = \min_{p \in \mathcal{P}_{k+1}^*} \|p(A) e^{(0)}\|_A^2 = \min_{p \in \mathcal{P}_{k+1}^*} \|Q^T p(D) Q e^{(0)}\|_A^2$$

$$\min_{p \in \mathcal{P}_{k+1}^*} \left(\underbrace{D Q Q^T}_{=I} p(D) Q e^{(0)} \underbrace{Q Q^T}_{=I} p(D) Q \underbrace{e^{(0)}}_{=z} \right)$$

$$\min_{p \in \mathcal{P}_{k+1}^*} (D p(D) z, p(D) z) = \sum_{i=1}^n \lambda_i z_i^2 (p(\lambda_i))^2$$

$$\leq \min_{p \in \mathcal{P}_{k+1}^*} \max_{\lambda_i \in \sigma(A)} (p(\lambda_i))^2 \underbrace{\sum_{i=1}^n \lambda_i z_i^2}_{=(P Q e^{(0)}, Q e^{(0)}) = \|e^{(0)}\|_A^2} = \left(\min_{p \in \mathcal{P}_{k+1}^*} \max_{\lambda_i \in \sigma(A)} (p(\lambda_i))^2 \right) \|e^{(0)}\|_A^2$$

$$\Leftrightarrow \frac{\|e^{(k+1)}\|_A}{\|e^{(0)}\|_A} \leq \min_{p \in \mathcal{P}_{k+1}^*} \max_{\lambda_i \in \sigma(A)} |p(\lambda)|$$

- 7) Als nächstes versuchen wir ein Polynom $q \in \mathcal{P}_{k+1}^*$ zu konstruieren, s.d. $|q(\lambda)|$ möglichst klein ist für $\lambda \in [\lambda_{min}, \lambda_{max}]$. Dazu führen wir die Tschebyscheff-Polynome ein (vgl. Martin Haute-Bourgois, Grundlagen d. Num. Math. und WissRech, Vieweg-Teubner, 3. Auflage S.284). Definition $T_k(x) := \cos(k \arccos(x))$, $k \geq 0, x \in [-1, 1]$ Tschebyscheff-Polynome k-ter Ordnung. $T_0(x) = 1, T_1(x) = x, t := \arccos(x)$ Behauptung: $T_{k-1}(x) + T_{k+1}(x) = \cos((k-1)t) + \cos((k+1)t)$ Somit gilt die rekursive Darstellung

$$T_0(x) = 1, T_1(x) = x, T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), k = 1, 2, \dots$$

$$\Rightarrow T_k \in \mathcal{P}_k$$

- 8) Es gilt auch die Darstellung

$$T_K(x) = \cosh(k \cdot \operatorname{arccosh}(x)) \text{ für } x \geq 1, T_0(x) = 1, T_1(x) = x$$

Für den Rest muss man die Rekursionsformel anwenden:

Setze für $x \geq 1$: $t := \operatorname{arccosh}(x) = \ln(x + \sqrt{x^2 - 1})$

Dann gilt

$$T_k(x) = \cosh(kt) = \frac{1}{2}(e^{kt} + e^{-kt}) \geq \frac{1}{2}(e^t)^k = \frac{1}{2}(x + \sqrt{x^2 - 1})^k$$

- 9) Wähle nun $q \in \mathcal{P}_{k+1}^*$ wie folgt

$$q(\lambda) := \frac{T_{k+1}\left(\frac{2\lambda - \lambda_{min} - \lambda_{max}}{\lambda_{min} - \lambda_{max}}\right)}{T_{k+1}\left(\frac{\lambda_{min} + \lambda_{max}}{\lambda_{max} - \lambda_{min}}\right)} \Rightarrow q(0) = 1, q \in \mathcal{P}_{k+1} \rightarrow q \in \mathcal{P}_{k+1}^*$$

Weiterhin gilt:

$$\max_{\lambda \in \sigma(A)} |q(\lambda)| \leq \frac{1}{T_{k+1}(\dots\dots\dots)}$$

□