18./19. VL Einführung in die Numerische Mathematik, 12.6.09, 17.6.09

Krylov-Raum-Verfahren zur Lösung linearer Gleichungssysteme

Ziel ist weiterhin die iterative Lösung des linearen Gleichungssystems

$$A\vec{x} = \vec{b} \quad , \ A \ (n\times n)\text{-Matrix, regulär}, \ \vec{b}\in\mathbb{R}^n,$$

mit der eindeutigen Lösung $\vec{x}_* = A^{-1}\vec{b}$.

Hierzu betrachten wir mit

$$\{0\} \subset D_1 \subset D_2 \subset \ldots \subset \mathbb{R}^n \qquad (5.21)$$

zunächst eine Folge von linearen Unterräumen, die noch präzisiert wird. Im Folgenden werden Ansätze zur Bestimmung von Vektorfolgen $\vec{x}_k \in D_k$, $k=1,2,\ldots$ betrachtet (mit dem letztendlichen Ziel mit dieser Folge die exakte Lösung $\vec{x}_*$ zu erreichen).

Definition 5.15

a) Für gegebene Ansatzräume (5.21) hat der Ansatz des orthogonalen Residuums zur Bestimmung von Vektoren $\vec{x}_1, \vec{x}_2, \ldots \in \mathbb{R}^n$ die Form

$$\left.\begin{aligned} &\vec{x}_k \in D_k \\ &A\vec{x}_k - \vec{b} \in D_k^\perp \end{aligned}\right\} k = 1,2,\ldots \qquad (5.22)$$

b) Der Ansatz des minimalen Residuums zur Bestimmung von Vektoren $\vec{x}_1, \vec{x}_2, \ldots \in \mathbb{R}^n$ hat für gegebene Ansatzräume die Form

$$\left.\begin{aligned} &\vec{x}_k \in D_k \\ &\|A\vec{x}_k - \vec{b}\|_2 = \min_{\vec{x}\in D_k} \|A\vec{x} - \vec{b}\|_2 \end{aligned}\right\} k = 1, 2, \dots \qquad (5.23)$$

(hier bezeichnet $M^\perp$ wie üblich das orthogonale Komplement von M, also

$$M^\perp = \{\vec{y} \in \mathbb{R}^n : \vec{y}^T \vec{x} = 0 \text{ f.a. } \vec{x} \in M\}).$$

Bei der Wahl spezieller Ansatzräume (5.21) werden die sogenannten Krylovräume von Bedeutung sein.

Definition 5.16
Zu gegebener Matrix $A \in \mathbb{R}^{n\times n}$ und einem Vektor $\vec{b} \in \mathbb{R}^n$ ist die Folge der Krylovräume durch

$$K_k(A, \vec{b}) = \mathrm{span}\{\vec{b}, A\vec{b}, \dots, A^{k-1}\vec{b}\} \subset \mathbb{R}^n, \quad k = 0, 1, \dots$$

erklärt.

Bemerkung
Im Folgenden werden die in Definition 5.15 angegebenen Ansätze mit den speziellen Räumen $D_k = K_k(A, \vec{b})$ betrachtet, wobei wir den Schwerpunkt auf den Ansatz (5.22) legen.

Der Ansatz des orthogonalen Residuums (5.22) für symmetrische, pos. definite Matrizen $A \in \mathbb{R}^{n\times n}$

Für pos. definite, symmetrische Matrizen soll nun Existenz und Eindeutigkeit von Vektoren $\vec{x}_k$ für (5.22) diskutiert werden. Dazu werden die Skalarprodukte und Normen

$$\langle \vec{x}, \vec{y} \rangle_2 = \vec{x}^T \vec{y} \quad , \vec{x}, \vec{y} \in \mathbb{R}^n$$

$$\langle \vec{x}, \vec{y} \rangle_A := \vec{x}^T A \vec{y} \quad , \vec{x}, \vec{y} \in \mathbb{R}^n, \ \|\vec{x}\|_A = \langle \vec{x}, \vec{x} \rangle_A^{1/2}$$

betrachtet (Nachweis, dass $\langle \cdot, \cdot \rangle_A$, $\|\cdot\|_A$ Skalarprodukt und Norm im Falle einer pos. def., symm. Matrix A sind, ist als Übung zu führen).

Satz 5.17

Zu gegebener symm., pos. definiter Matrix $A \in \mathbb{R}^{n \times n}$ sind für $k = 1, 2, \ldots$ die Vektoren $\vec{x}_k$ aus dem Ansatz des orthogonalen Residuums (5.22) – mit allg. Ansatzräumen D_k gemäß (5.21) – eindeutig bestimmt, und es gilt

$$\|\vec{x}_k - \vec{x}_*\|_A = \min_{\vec{x} \in D_k} \|\vec{x} - \vec{x}_*\| \quad , k = 1, 2, \ldots \qquad (5.24)$$

Beweis

Eindeutigkeit:

Sei k fest gewählt. Für $\vec{x}_k, \hat{\vec{x}}_k$ mit der Eigenschaft (5.22) gilt

$$\langle \underbrace{A(\vec{x}_k - \hat{\vec{x}}_k)}_{\in D_k^\perp}, \underbrace{\vec{x}_k - \hat{\vec{x}}_k}_{\in D_k} \rangle_2 = 0 \Rightarrow \vec{x}_k = \hat{\vec{x}}_k .$$

Existenz:

Mit einer beliebigen Basis $\vec{d}_0, \vec{d}_1, \ldots, \vec{d}_{m-1}$ von D_k setzt man

$$\vec{x}_k = \sum_{j=0}^{m-1} \alpha_j \vec{d}_j \qquad (5.25)$$

an und erhält

$$\vec{x}_k \text{ genügt (5.22)} \iff A\vec{x}_k - \vec{b} \in D_k^{\perp}$$

$$\iff \langle A\vec{x}_k - \vec{b}, \vec{d}_k \rangle_2 = 0, \; k = 0, \dots, m-1, \qquad (5.26)$$

$$\iff \sum_{j=0}^{m-1} \langle A\vec{d}_j, \vec{d}_k \rangle_2 \alpha_j = \langle \vec{b}, \vec{d}_k \rangle_2 \qquad (5.27)$$

$$k = 0, \dots, m-1$$

(5.27) ist ein lin. Gleichungssystem von m Gleichungen für die Koeff. $\alpha_0, \dots, \alpha_{m-1}$. Da $\vec{x}_k$ mit (5.22) eindeutig bestimmt ist (wurde schon gezeigt), ist das Gleichungssystem (5.27) eindeutig lösbar, woraus die Existenz von $\vec{x}_k$ folgt.

Minimaleigenschaft (5.24):

Für $\vec{x} \in D_k$ findet man

$$\|\vec{x} - \vec{x}_*\|_A^2 = \|\vec{x}_k - \vec{x}_* + \vec{x} - \vec{x}_k\|_A^2$$

$$= \|\vec{x}_k - \vec{x}_*\|_A^2 + 2\underbrace{\langle \underbrace{A(\vec{x}_k - \vec{x}_*)}_{\in D_k^{\perp}}, \underbrace{\vec{x} - \vec{x}_k}_{\in D_k} \rangle_2}_{= 0} + \|\vec{x} - \vec{x}_k\|_A^2$$

$$\geq \|\vec{x}_k - \vec{x}_*\|_A^2 .$$

$\square$

Der Ansatz des orthog. Residuums (5.22) für

<u>gegebene A-konjugierte Basen</u>

Mit dem Beweis von Satz 5.17 ist bereits eine Möglichkeit zur Bestimmung von $\vec{x}_k$ für (5.22) ausgehend von einer Basis $\vec{d}_0, \dots, \vec{d}_{m-1}$ für D_k mit dem Gleichungssystem (5.27) aufgezeigt worden. Im Folgenden wird ein Spezialfall behandelt, bei dem (5.27) Diagonalgestalt hat.

Definition 5.18

Es sei $A \in \mathbb{R}^{n\times n}$ symm. und pos. definit. Gegebene Vektoren $\vec{d}_0, \vec{d}_1, \dots, \vec{d}_{n-1} \in \mathbb{R}^n \setminus \{0\}$ heißen A-konjugiert, falls

$$\langle A\vec{d}_i, \vec{d}_j \rangle_2 = \langle \vec{d}_i, \vec{d}_j \rangle_A = 0 \quad \text{für } i \neq j \text{ gilt.}$$

Bemerkung

Falls eine A-konjugierte Basis von D_k gegeben ist, hat (5.27) Diagonalgestalt und damit ist $\vec{x}_k$ gemäß Ansatz (5.25) sehr einfach berechenbar.

Satz 5.19

Für eine gegebene symmetr., pos. definite Matrix $A \in \mathbb{R}^{n\times n}$ und A-konjugierte Vektoren $\vec{d}_0, \vec{d}_1, \dots$ gelte $D_k = \mathrm{span}\{\vec{d}_0, \dots, \vec{d}_{k-1}\}$, $k = 1, 2, \dots$

Dann erhält man für den Ansatz des orthog. Residuums (5.22) die folgenden Darstellungen für $k = 1, 2, \dots$

$$\vec{x}_k = \sum_{j=0}^{k-1} \alpha_j \vec{d}_j \quad \text{mit } \alpha_j = -\frac{\langle \vec{r}_j, \vec{d}_j \rangle_2}{\langle A\vec{d}_j, \vec{d}_j \rangle_2} \qquad (5.28)$$

$$\vec{r}_j := A\vec{x}_j - \vec{b}, \; j \geq 1, \; \vec{r}_0 = -\vec{b}. \qquad (5.29)$$

Beweis

Folgt unmittelbar für $k = n$ aus (5.25) – (5.27). $\square$

Bemerkung

a) Aus (5.28) folgt die Unabhängigkeit der α_j von k und damit gilt

$$\vec{x}_{k+1} = \vec{x}_k + \alpha_k \vec{d}_k \,,\; \vec{r}_{k+1} = \vec{r}_k + \alpha_k A \vec{d}_k \,, \qquad (5.30)$$
$$k = 0, 1, \dots ; \; \vec{x}_0 = 0$$

b) Aufgrund der ersten Identität von (5.30) bezeichnet man $\vec{d}_k$ als Suchrichtung und α_k als Schrittweite.

c) Außerdem wird mit (5.30) klar, dass eine simultane Berechnung der Suchrichtungen und Lösungsapproximationen $\vec{x}_k$ in der Reihenfolge
$$\vec{d}_0, \vec{x}_1, \vec{d}_1, \vec{x}_2, \dots$$
möglich ist.
In der Praxis wird im Fall $D_k = K_k(A, \vec{b})$ auch so vorgegangen, was im Folgenden behandelt werden soll.

19. VL Einführung in die Num. Mathematik, 17.6.09

Das CG-Verfahren für pos. definite, symm. Matrizen

Definition 5.20:
Zu gegebener symm., pos. definiter Matrix $A \in \mathbb{R}^{N \times N}$ ist das <u>Verfahren der konjugierten Gradienten</u> gegeben durch den Ansatz (5.22) mit der speziellen Wahl
$$D_k = K_k(A, \vec{b}) \,, \quad k = 0, 1, \dots \qquad (5.31)$$

Dieses Verfahren bezeichnet man auch kurz als CG-Verfahren.

Bemerkung

Zur konkreten Bestimmung der Krylovapproximationen fehlen uns nur noch geeignete Suchrichtungen, am besten A-konjugierte Suchrichtungen $\vec{d}_0, \vec{d}_1, \ldots$ Das soll nun geschehen.

Der folgende Hilfssatz behandelt die Berechnung A-konjugierter Suchrichtungen in $K_k(A, \vec{b})$ für $k = 0, 1, \ldots$

Ausgehend von den Notationen des Satzes 5.19 wird für den fixierten Index k dabei so vorgegangen, dass – ausgehend von einer bereits konstruierten A-konjugierten Basis $\vec{d}_0, \ldots, \vec{d}_{k-1}$ für $K_k(A, \vec{b})$ – eine A-konjugierte Basis für $K_{k+1}(A, \vec{b})$ gewonnen wird durch eine Gram-Schmidt-Orthogonalisierung der Vektoren $\vec{d}_0, \ldots, \vec{d}_{k-1}, -\vec{r}_k \in \mathbb{R}^n$ bzgl. des Skalarproduktes $\langle \cdot, \cdot \rangle_A$.

Wie sich im Beweis von Hilfssatz 5.21 herausstellt, genügt hier eine Gram-Schmidt-Orthogonalisierung der beiden Vektoren $\vec{d}_{k-1}, -\vec{r}_k \in \mathbb{R}^n$.

Hilfssatz 5.21

Zu gegebener symm., pos. definiter Matrix $A \in \mathbb{R}^{n \times n}$ und mit den Notationen des Satzes 5.19 seien die Suchrichtungen speziell wie folgt gewählt:

$\vec{d}_0 = \vec{b}$, sowie

$$\vec{d}_k = -\vec{r}_k + \beta_{k-1}\vec{d}_{k-1}, \quad \beta_{k-1} = \frac{\langle A\vec{r}_k, \vec{d}_{k-1}\rangle_2}{\langle A\vec{d}_{k-1}, \vec{d}_{k-1}\rangle_2}, \quad k = 1, \dots, k_*-1 \tag{5.32}$$

wobei k_* den ersten Index mit $\vec{r}_{k_*} = 0$ bezeichnet. Mit dieser Wahl sind die Vektoren $\vec{d}_0, \vec{d}_1, \dots, \vec{d}_{k_*-1} \in \mathbb{R}^n$ A-konjugiert und es gilt

$$\operatorname{span}\{\vec{d}_0, \dots, \vec{d}_{k-1}\} = \operatorname{span}\{\vec{b}, \vec{r}_1, \dots, \vec{r}_{k-1}\} = K_k(A, \vec{b}), \quad k = 1, \dots, k_* \tag{5.33}$$

Beweis

Vollst. Induktion über $k = 1, 2, \dots, k_*$ zum Nachweis der A-Konjugiertheit der Vektoren $\vec{d}_0, \vec{d}_1, \dots, \vec{d}_{k-1} \in \mathbb{R}^n$ und der Formeln (5.32)

wegen
$$\operatorname{span}\{\vec{d}_0\} = \operatorname{span}\{\vec{b}\} = K_1(A, \vec{b})$$
ist der Induktionsanfang gemacht.

Im Folgenden sei angenommen, dass (5.32) ein System von A-konjugierten Vektoren mit der Eigenschaft (5.33) liefert mit einem fixierten Index $1 \le k \le k_*-1$.

Gemäß dem Ansatz des orthogonalen Residuums (S.22) gilt $\vec{r}_k \in K_k(A,\vec{b})^\perp$ und im Fall $\vec{r}_k \neq 0$ sind damit die Vektoren $\vec{d}_0, \dots, \vec{d}_{k-1}, -\vec{r}_k$ lin. unabh. Eine Gram-Schmidt-Orthogonalisierung dieser Vektoren bzgl. des Skalarproduktes $\langle \cdot,\cdot \rangle_A$ liefert den Vektor

$$\vec{d}_k = -\vec{r}_k + \sum_{j=0}^{k-1} \frac{\langle A\vec{r}_k, \vec{d}_j \rangle_2}{\langle A\vec{d}_j, \vec{d}_j \rangle_2} \vec{d}_j$$

$$\overset{(*)}{=} -\vec{r}_k + \beta_{k-1} \vec{d}_{k-1} \qquad (S.34)$$

wobei (*) aus den Eigenschaften

$$K_{k-1}(A,\vec{b}) \subset K_k(A,\vec{b}) \quad \text{sowie} \quad \vec{r}_k \in K_k(A,\vec{b})^\perp$$

folgt, also

$$\langle A\vec{r}_k, \vec{d}_j \rangle_2 = \langle \vec{r}_k, A\vec{d}_j \rangle_2 = 0 \quad , j = 0,1,\dots,k-2 .$$

Nach Konstruktion sind die Vektoren $\vec{d}_0, \dots, \vec{d}_{k-1}, \vec{d}_k$ A-konjugiert und es gilt

$$\operatorname{span}\{\vec{d}_0, \dots, \vec{d}_{k-1}, \vec{d}_k\} = \operatorname{span}\{\vec{b}, \vec{r}_1, \vec{r}_2, \dots, \vec{r}_k\}.$$

Aufgrund der 2. Formel in (S.30) gilt noch

$$\operatorname{span}\{\vec{b}, \vec{r}_1, \vec{r}_2, \dots, \vec{r}_k\} \subseteq K_{k+1}(A,\vec{b})$$

so dass aus Dimensionsgründen auch hier notwendigerweise Gleichheit vorliegt. $\square$

Bemerkung

Mit dem durch Hilfssatz S.21 beschriebenen Abbruch wird gleichzeitig die Lösung von $A\vec{x} = \vec{b}$ geliefert, es gilt also $\vec{x}_{k_*} = \vec{x}_*$. Dabei gilt notwendigerweise

$$k_* \leq n,$$

denn aufgrund der lin. Unabhängigkeit der beiden Vektorsysteme in (5.33) erhält man

$$\dim \mathcal{K}_k = k$$

für $k = 0, 1, \ldots, k_*$.

Im folgenden Hilfssatz werden Darstellungen für die Schrittweiten gezeigt, wie sie auch in numerischen Implementierungen verwendet werden.

Hilfssatz 5.22

In der Situation des Hilfssatzes 5.21 gelten die Darstellungen

$$\alpha_k = \frac{\|\vec{r}_k\|_2^2}{\langle A\vec{d}_k, \vec{d}_k\rangle_2}, \quad k = 0, 1, \ldots, k_* - 1, \qquad (5.35)$$

$$\beta_{k-1} = \frac{\|\vec{r}_k\|_2^2}{\|\vec{r}_{k-1}\|_2^2}, \quad k = 1, \ldots, k_* - 1 \quad (\vec{r}_0 := -\vec{b}) \qquad (5.36)$$

Beweis

Mit $\vec{r}_k \in \mathcal{K}_k(A, \vec{b})^\perp$ sowie der Beziehung (5.34) für die Suchrichtung $\vec{d}_k$ erhält man

$-\langle \vec{r}_k, \vec{d}_k\rangle_2 = \|\vec{r}_k\|_2^2$ und zusammen mit (5.28) ergibt dies (5.35). Diese Darstellung (5.35) für α_k zusammen mit der Identität $\vec{r}_k = \vec{r}_{k-1} + \alpha_{k-1} A\vec{d}_{k-1}$ aus (5.30) liefert

$$\|\vec{r}_k\|_2^2 = \underbrace{\langle \vec{r}_k, \vec{r}_{k-1}\rangle_2}_{=0} + \alpha_{k-1}\langle \vec{r}_k, A\vec{d}_{k-1}\rangle_2 = \beta_{k-1}\|\vec{r}_{k-1}\|_2^2$$ und

damit gilt für β_{k-1} die Beziehung (5.36). $\square$

Konvergenzgeschwindigkeit des CG-Verfahrens

Wir haben bisher festgestellt, dass das CG-Verfahren mit $\vec{x}_{k_*} = \vec{x}_*$ nach k_* Schritten die Lösung ergibt. k_* kann aber sehr groß sein und deshalb interessiert auch der Fehler im k-ten Schritt ($k=1,2,\dots$). Hilfreich ist

Hilfssatz 5.23
Zu gegebener symm. pos. definiter Matrix $A \in \mathbb{R}^{n\times n}$ sei $(\lambda_j, \vec{v}_j)_{j=1,\dots,n}$ ein vollst. System von Eigenwerten $\lambda_j > 0$ und zugehörigen Eigenvektoren $\vec{v}_j \in \mathbb{R}^n$, also gilt $A\vec{v}_j = \lambda_j \vec{v}_j$, $\quad \vec{v}_k^T \vec{v}_j = \delta_{kj}$, $k,j=1,\dots,n$.
Mit der Entwicklung

$$\vec{x} = \sum_{j=1}^{n} c_j \vec{v}_j \in \mathbb{R}^n$$

gelten für jedes Polynom p die folgenden Darstellungen

$$p(A)\vec{x} = \sum_{j=1}^{n} c_j p(\lambda_j) \vec{v}_j \tag{5.37}$$

$$\|p(A)\vec{x}\|_2 = \left(\sum_{j=1}^{n} c_j^2 p(\lambda_j)^2\right)^{1/2}, \quad \|p(A)\vec{x}\|_A = \left(\sum_{j=1}^{n} c_j^2 \lambda_j p(\lambda_j)^2\right)^{1/2}$$

Speziell gilt also

$$m^{1/2}\|\vec{x}\|_2 \le \|\vec{x}\|_A \le M^{1/2}\|\vec{x}\|_2, \quad \vec{x} \in \mathbb{R}^n, \tag{5.39}$$

$$\left(m := \min_{1\le j\le n} \lambda_j, \quad M := \max_{1\le j\le n} \lambda_j\right).$$

Beweis

Mit der angegebenen Entwicklung für $\vec{x} \in \mathbb{R}^n$ gilt

$$A^\nu \vec{x} = \sum_{j=1}^n c_j \lambda_j^\nu \vec{v}_j, \quad \nu = 0, 1, \ldots$$

und daraus folgt (5.37). Weiter berechnet man

$$\|p(A)\vec{x}\|_2 = \left\langle \sum_{k=1}^n c_k p(\lambda_k)\vec{v}_k, \sum_{j=1}^n c_j p(\lambda_j)\vec{v}_j \right\rangle_2^{1/2} =$$

$$= \left(\sum_{k,j=1}^n c_k c_j p(\lambda_k) p(\lambda_j) \underbrace{\langle \vec{v}_k, \vec{v}_j \rangle_2}_{=\delta_{kj}} \right)^{1/2}$$

$$= \left(\sum_{j=1}^n c_j^2 p(\lambda_j)^2 \right)^{1/2},$$

und analog erhält man

$$\|p(A)\vec{x}\|_A = \left\langle A\left(\sum_{k=1}^n c_k p(\lambda_k)\vec{v}_k\right), \sum_{j=1}^n c_j p(\lambda_j)\vec{v}_j \right\rangle_2^{1/2} =$$

$$= \left\langle \sum_{k=1}^n c_k \lambda_k p(\lambda_k)\vec{v}, \sum_{j=1}^n c_j p(\lambda_j)\vec{v}_j \right\rangle_2^{1/2} =$$

$$= \left(\sum_{k,j=1}^n c_k c_j \lambda_k p(\lambda_k) p(\lambda_j) \underbrace{\langle \vec{v}_k, \vec{v}_j \rangle_2}_{=\delta_{kj}} \right)^{1/2} =$$

$$= \left(\sum_{j=1}^n c_j^2 \lambda_j p(\lambda_j)^2 \right)^{1/2}.$$

$\square$

Es gilt nun noch den Fehler $\|\vec{x}_k - \vec{x}_*\|_A$, den man im k-ten Schritt des CG-Verfahrens macht, abzuschätzen. Einmal gilt der

Satz 5.24

Zu einer gegebenen symm., pos. definiten Matrix $A \in \mathbb{R}^{n\times n}$ gelten für das CG-Verfahren die folgenden Fehlerabschätzungen:

$$\|\vec{x}_k - \vec{x}_*\|_A \le \left(\inf_{p\in\Pi_k,\, p(0)=1} \sup_{\lambda\in\sigma(A)} |p(\lambda)| \right) \|\vec{x}_*\|_A \tag{5.40}$$

$k = 0, 1, \ldots, k_*$.

Beweis

Für jedes Polynom $p \in \Pi_k$ mit $p(0)=1$ ist $q(t) := (1-p(t))/t$ ein Polynom vom Grad höchstens $k-1$, und damit gilt mit $\vec{x} := q(A)\vec{b}$ Folgendes:

$$\vec{x} \in K_k(A,\vec{b}), \quad \vec{x} - \vec{x}_* = -p(A)\vec{x}_*.$$

Mit Hilfssatz 5.23 und der Entwicklung $\vec{x}_* = \sum_{j=1}^n c_j \vec{v}_j \in \mathbb{R}^n$ erhält man

$$\|\vec{x}_k - \vec{x}_*\|_A \underset{(5.24)}{\le} \underbrace{\|\vec{x} - \vec{x}_*\|_A}_{=\|p(A)\vec{x}_*\|_A} = \left(\sum_{j=1}^n c_j^2 \lambda_j p(\lambda_j)^2\right)^{1/2} \le$$

$$\le \sup_{\lambda\in\sigma(A)} |p(\lambda)| \left(\sum_{j=1}^n c_j^2 \lambda_j\right)^{1/2} = \sup_{\lambda\in\sigma(A)} |p(\lambda)|\, \|\vec{x}_*\|_A$$

$\square$

Zur quantitativen Präzisierung der Abschätzung (5.40) des Satzes 5.24 benutzen wir die hier nicht bewiesene Eigenschaft der Tschebyscheff-Polynome erster Art $T_0, T_1, \dots$

$$T_k\left(\frac{\varkappa+1}{\varkappa-1}\right) \geq \frac{1}{2}\left(\frac{\sqrt{\varkappa}+1}{\sqrt{\varkappa}-1}\right)^k \quad \text{für } \varkappa \in \mathbb{R},\ \varkappa > 1. \tag{5.41}$$

Es gilt der

Satz 5.25

Zu einer gegebenen symm., pos. definiten Matrix $A \in \mathbb{R}^{n\times n}$ gelten für das CG-Verfahren die Fehlerabschätzungen

$$\|\vec{x}_k - \vec{x}_*\|_A \leq 2\gamma^k \|\vec{x}_*\|_A, \quad k = 0, 1, \dots \tag{5.42}$$

$$\|\vec{x}_k - \vec{x}_*\|_2 \leq 2\sqrt{\varkappa_A}\,\gamma^k \|\vec{x}_*\|_2, \quad k = 0, 1, \dots$$

mit $\varkappa_A = \operatorname{cond}_2(A)$, $\gamma = \dfrac{\sqrt{\varkappa_A}-1}{\sqrt{\varkappa_A}+1}$.

Beweis.

Satz 5.24 wird im Fall $\varkappa_A > 1$, d.h. $M > m$, angewandt mit dem Polynom

$$p(\lambda) = \frac{T_k[(M+m-2\lambda)/(M-m)]}{T_k[(M+m)/(M-m)]}, \quad \lambda \in \mathbb{R},$$

wobei m und M den kleinsten und größten Eigenwert von A bezeichnen. Offensichtlich ist $p \in \Pi_k$ und $p(0) = 1$, wegen $\sigma(A) \subset [m, M]$ und

$$\max_{m \leq \lambda \leq M} |p(\lambda)| = \left|T_k\left(\frac{M+m}{M-m}\right)\right|^{-1} = \left|T_k\left(\frac{\varkappa_A+1}{\varkappa_A-1}\right)\right|^{-1} \underset{(5.41)}{\leq} 2\gamma^k$$

folgt aus (5.40) die erste Abschätzung, also (5.42).

Die zweite Abschätzung des Satzes ist eine unmittelbare Konsequenz aus der ersten Abschätzung unter der Nutzung der Normäquivalenz (5.39).

$\square$

GMRES-Verfahren

Lässt man die Voraussetzung der Symmetrie und pos. Definitheit der Matrix A fallen und fordert nur die Regularität, dann ist ein CG-Verfahren zur Lösung von $A\vec{x}=\vec{b}$ nicht möglich. Eine Alternative ist das GMRES-Verfahren:

Definition 5.26

Das GMRES-Verfahren ist definiert durch den Ansatz des minimalen Residuums (5.23) mit der speziellen Wahl $D_k = K_k(A,\vec{b})$, es gilt also

$$\vec{x}_k \in K_k(A,\vec{b}), \quad \|A\vec{x}_k-\vec{b}\|_2 = \min_{\vec{x}\in K_k(A,\vec{b})} \|A\vec{x}-\vec{b}\|_2, \quad k=0,\dots,k_*.$$

Bemerkung

Die Abkürzung "GMRES" hat ihren Ursprung in der Bezeichnung "generalized minimal residual method".

Detaillierte Konstruktionsmethoden für die Approximationen $\vec{x}_k$ beim GMRES-Verfahren werden in Plato beschrieben.