5 Iterationsverfahren für lineare Gleichungssysteme

Klassische Iterationsverfahren

Sei $A \in \mathbb{R}^{N\times N}$ und $b \in \mathbb{R}^N$. Wir wollen nun das LGS $Ax = b$ iterativ lösen. Dazu betrachten wir die Komponenten $m = 1,\dots,N$:

$$\sum_{n=1}^{N} A[m,n]x[n] = b[m].$$

Wir formen um und erhalten für $x[m]$:

$$A[m,m]x[m] = b[m] - \sum_{n\neq m} A[m,n]x[n].$$

Falls $A[m,m] \neq 0$ erhalten wir also

$$x[m] = \left(b[m] - \sum_{n\neq m} A[m,n]x[n] \right) / A[m,m].$$

Daraus können wir das *Gesamtschritt-Verfahren (Jacobi-Verfahren)* ableiten:

S0) Wähle $x^0 \in \mathbb{R}^N,\ \varepsilon > 0$, setze $k = 0$.

S1) Falls $|Ax^k - b|_2 < \varepsilon$ STOP.

S2) Berechne $x^{k+1}[m] = (b[m] - \sum\limits_{n\neq m} A[m,n]x^k[n])/A[m,m]$ für $m = 1,...,N$.

S3) Setze $k := k+1$, gehe zu S1).

Dieses Verfahren kann man mit der Idee abändern, dass bei der Berechnung von x_m die x_n für $n < m$ ja schon bekannt sind. Dementsprechend kann man auch gleich mit diesen Werten weiterrechnen.
Dies führt zum *Einzelschritt-Verfahren (Gauß-Seidel-Verfahren)*:

S0) Wähle $x^0 \in \mathbb{R}^N,\ \varepsilon > 0$, setze $k = 0$.

S1) Falls $|Ax^k - b|_2 < \varepsilon$ STOP.

S2) Berechne $x^{k+1}[m] = \left(b[m] - \sum\limits_{n=1}^{m-1} A[m,n]x^{k+1}[n] - \sum\limits_{n=m+1}^{N} A[m,n]x^k[n] \right) / A[m,m]$ für $m = 1,...,N$.

S3) Setze $k := k+1$, gehe zu S1).

Um die Konvergenz von Iterationsverfahren zu untersuchen formulieren wir sie in Matrixform. Beachte, dass für invertierbare Matrizen $B \in \mathbb{R}^{N,N}$ die Aufgabe $Ax = b$ äquivalent ist zu $BAx = Bb$. Ein solches B heißt *Vorkonditionierer* von A, da B möglichst so gewählt wird, dass BA eine kleinere Kondition als A hat. Wir formulieren dies um in eine Fixpunktaufgabe, also $x = Bb - BAx + x$ und erhalten so

$$x = (I - BA)x + Bb = x + B(b - Ax)\,.$$

Wir bestimmen nun B für das Jacobi- und das Gauß-Seidel-Verfahren. Dazu zerlegen wir

$$A = L + D + R$$

mit der strikt unteren Dreiecksmatrix $L = \mathrm{lower}(A)$, der Diagonalmatrix $D = \mathrm{diag}(A)$ und der strikt oberen Dreiecksmatrix $R = \mathrm{upper}(A)$. Damit folgt für diese beiden Verfahren:

Gesamtschrittverfahren
Es gilt:

$$Dx^{k+1} = b - (L+R)x^k = b - (L+D+R)x^k + Dx^k.$$

Somit ist $D(x^{k+1} - x^k) = b - Ax^k$ und schließlich

$$x^{k+1} = x^k + D^{-1}(b - Ax^k),$$

also $B_J = D^{-1}$.

Einzelschrittverfahren
Es gilt

$$Dx^{k+1} = b - Lx^{k+1} - Rx^k,$$

woraus $(D+L)x^{k+1} = b - Rx^k = b - (L+D+R)x^k + (L+D)x^k = b - Ax^k + (D+L)x^k$ folgt und somit

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

also $B_{GS} = (D+L)^{-1}$.

(5.1) Satz
Seien $A, B \in \mathbb{R}^{N\times N}$ mit $\rho(I - BA) < 1$.
Dann ist A invertierbar und für alle $b \in \mathbb{R}^N$, $x^0 \in \mathbb{R}^N$ konvergiert die Iteration

$$x^{k+1} = x^k + B(b - Ax^k) \qquad k = 0, 1, 2, \dots\ ,$$

so dass gilt

$$\lim_{k\to\infty} x^k = A^{-1}b. \quad \text{und} \quad A^{-1} = B\sum_{k=0}^{\infty}\left(I_N - AB\right)^k.$$

(5.2) Satz
Sei $K \in \mathbb{R}^{N\times N}$ und $\varepsilon > 0$. Dann existiert eine Norm $\|\cdot\|$ mit $\|K\| \le \rho(K) + \varepsilon$.

Beweis. Aus der linearen Algebra wissen wir, dass eine reguläre Matrix $U \in \mathbb{C}^{N\times N}$ existiert, sodass

$$U^{-1}KU = \mathrm{diag}(\lambda_1, \dots, \lambda_N) + R$$

mit einer strikt oberen Dreiecksmatrix R, wobei λ_n $(n = 1, \dots, N)$ die Eigenwerte von K sind. Zu $\delta = \frac{\varepsilon}{\varepsilon + \|R\|_\infty} \in (0,1]$ definiere $D = \mathrm{diag}(1, \delta^1, \delta^2, \dots)$. Damit gilt
$D^{-1} = \mathrm{diag}(1, \delta^{-1}, \delta^{-2}, \dots)$ und

$$\begin{aligned} D^{-1}RD &= (\delta^{n-m}R[m,n])_{m,n=1\dots,N} \\ &= \begin{pmatrix} 0 & \delta R[1,2] & \delta^2 R[1,3] & \dots \\ & 0 & \delta R[2,3] & \dots \\ 0 & & \ddots & \end{pmatrix} \end{aligned}$$

Somit folgt

$$\|D^{-1}RD\|_\infty \le \delta\|R\|_\infty = \varepsilon - \varepsilon\delta = \varepsilon\underbrace{(1-\delta)}_{\in[0,1)} < \varepsilon\,. \tag{1}$$

Nun definiere die Vektornorm $|x| = |D^{-1}U^{-1}x|_\infty$ und erhalte

$$\begin{aligned}
\|K\| &= \sup_{x\neq 0}\frac{|Kx|}{|x|} = \sup_{x\neq 0}\frac{|D^{-1}U^{-1}Kx|_\infty}{|D^{-1}U^{-1}x|_\infty} \\
&= \sup_{y\neq 0}\frac{|D^{-1}U^{-1}KUDy|_\infty}{|y|_\infty} = \|D^{-1}(\mathrm{diag}(\lambda_n)+R)D\|_\infty \\
&\le \underbrace{\|\mathrm{diag}(\lambda_n)\|_\infty}_{=\rho(K)} + \|D^{-1}RD\|_\infty \overset{(1)}{\le} \rho(K)+\varepsilon.
\end{aligned}$$

□

Beweis. (Beweis von 5.1) Zu einer Matrix $K = I_N - BA$ wähle $|\cdot|$, $\|\cdot\|$ wie in 5.2 zu $\varepsilon = \frac{1-\rho(K)}{2} > 0$. Es gilt $\|K\| \le \rho(K)+\varepsilon = 1-\frac{\varepsilon}{2} < 1$. Damit ist A invertierbar; die Inverse ist mit der Neumannschen Reihe $A^{-1} = \sum\limits_{k=0}^{\infty}(I_N-BA)^kB$ darstellbar.

Für $x = A^{-1}b$ gilt dann

$$\begin{aligned}
x^{k+1}-x &= x^k + B(b-Ax^k) - x = x^k - x + B(Ax - Ax^k) \\
&= x^k - x + BA(x-x^k) = \underbrace{(I_N-BA)}_{=K}(x^k-x).
\end{aligned}$$

Und somit

$$|x^{k+1}-x| \le \|I_N-BA\|\,|x^k-x| \le \|I_N-BA\|^{k+1}|x^0-x| \longrightarrow 0 \quad (k\to\infty).$$

□

Bemerkung
Falls $\rho(I_N-BA)\ge 1$, dann existiert $q\in\mathbb{R}^N$ mit

$$|(I_N-BA)q|_2 = \rho(I_N-BA)|q|_2 \ge |q|_2 \neq 0$$

Wähle $b=0$, $x^0=q$, dann konvergiert (x^k) nicht gegen 0_N.

(5.3) Satz
Sei $A = L+D+L^T \in \mathbb{R}^{N\times N}$ symmetrisch und positiv definit. Dann konvergiert das Gauß-Seidel Verfahren mit $B=(L+D)^{-1}$ in der Energienorm $|x|_A := \sqrt{x^TAx}$.

Beweis. Sei dazu $A = L+D+L^T$ mit $L = \mathrm{lower}(A)$, $D = \mathrm{diag}(A)$ und $K = I_N - BA$, wobei $B = (D+L)^{-1}$. Zeige nun $\|K\|_A = \sup\limits_{x\neq 0}\frac{|Kx|_A}{|x|_A} < 1$, dann gilt $\rho(K)\le\|K\|_A<1$.

Es gilt

$$\begin{aligned}
|Kx|_A^2 &= (x-BAx)^TA(x-BAx) \\
&= x^T(A - ABA - AB^TA + AB^TABA)x \\
&= x^TAx - x^TA(B^T+B-B^TAB)Ax.
\end{aligned}$$

Weiter gilt
$$A + D = L + D + D + L^T = B^{-1} + B^{-T},$$
sodass
$$\begin{aligned} B^T + B - B^T AB &= B^T + B - B^T (A + D - D)B \\ &= B^T + B + B^T DB - B^T (B^{-1} + B^{-T})B \\ &= B^T DB\,. \end{aligned}$$

Da $A^T B^T DBA$ positiv definit ist, existiert ein $\alpha > 0$ mit $x^T(A^T B^T DBA)x \geq \alpha x^T Ax$. Es gilt also
$$|Kx|_A^2 = |x|_A^2 - x^T(AB^T DBA)x \leq |x|_A^2 - \alpha|x|_A^2 \leq (1-\alpha)|x|_A^2$$
und somit
$$\|K\|_A = \sup_{x \neq 0_N} \frac{|Kx|_A}{|x|_A} = \sup_{|x|_A=1} |Kx|_A \leq 1 - \alpha < 1\,.$$
□

Bemerkung

1. *Häufig wird die symmetrische Variante verwendet.*
2. *Beim SOR*[1]*-Verfahren ist* $B = \omega(D - \omega L)^{-1}$ *mit* $\omega \in (1,2)$.

(5.4) Definition
Sei $A \in \mathbb{R}^{N\times N}$. *Die Matrix A heißt* stark diagonaldominant, *wenn sie diagonaldominant ist* ($|A[n,n]| \geq \sum_{k\neq n} |A[n,k]|, n = 1,\dots,N$) *und wenn für ein* $j \in \{1,\dots,N\}$, $|A[j,j]| > \sum_{k\neq j} |A[j,k]|$ *gilt.*

(5.5) Satz
Sei $A \in \mathbb{R}^{N\times N}$ *stark diagonaldominant und irreduzibel. Dann gilt:*

a) *A ist regulär und das Jacobi-Verfahren konvergiert.*

b) *Falls* $A[n,n] > 0$, *dann ist A positiv definit.*

c) *Falls* $A[n,n] > 0, A[n,k] \leq 0\,(n \neq k)$ *ist, dann gilt* $A^{-1}[n,k] \geq 0$.

(5.6) Lemma
Wenn A irreduzibel ist, dann ist der Matrixgraph $\Gamma = \{(k,n) : A[k,n] \neq 0\}$ *zusammenhängend: zu jedem Paar* $n \neq j$ *existiert eine Folge* $j = j_0, j_1, \dots, j_R = n$ *mit* $A[j_1, j_0] \neq 0, A[j_2, j_1] \neq 0, \dots, A[j_R, j_{R-1}] \neq 0$.

Beweis. (nur skizziert) Ohne Einschränkung sei $j < n$.
Definiere $I = \{k : \text{es ex. ein Weg in } \Gamma \text{ von } j \text{ zu } k\}$.
Annahme: $n \notin I$; ohne Einschränkung sei $I = \{1,\dots,K\}$.
Dann gilt $A[m,k] = 0$ für $k \in I$, $k \leq K$, $m > K$.
Es folgt: A ist irreduzibel mit $A[K+1:N, 1:K] = 0$. Das ist ein Widerspruch.
□

[1] Successive Over-Relaxation, dt. Überrelaxation

Beweis. (Beweis von 5.5) Sei A irreduzibel. Dann gilt $\sum_{k=1}^N |A[n,k]| > 0$, mit 5.4 folgt $|A[n,n]| > 0$ für $n = 1, \ldots, N$. Somit sind $B = D^{-1}$, $K = I_N - D^{-1}A$ wohldefiniert.
Zu zeigen ist nun $\rho(K) < 1$. Sei dazu $\mu \in \mathbb{C}$ Eigenwert mit Eigenvektor $w \in \mathbb{C}^N \setminus \{0_N\}$ und $Kw = \mu w$; ohne Einschränkung sei $|w|_\infty = 1 = |w[n]|$ für ein $n \in \{1, \ldots, N\}$. Aus $\mu w = Kw$ folgt

$$\begin{aligned}
|\mu||w[n]| &= \left| \sum_{k=1}^N K[n,k]w[k] \right| \\
&\leq |A[n,n]|^{-1} \sum_{k \neq n} |A[n,k]| \\
&\leq 1
\end{aligned}$$

Annahme: $|\mu| = 1$. Zu j wähle $j = j_0, j_1, \ldots, j_k = n$ mit $A[j_1, j] \neq 0$. Es folgen

$$|w[j]| = |\mu w[j]| = |(Kw)[j]| \leq |A[j,j]|^{-1} \sum_{k \neq j} |A[j,k]| \leq 1$$

und

$$\begin{aligned}
|w[j_1]| = |\mu w[j_1]| &\leq |A[j_1, j_1]|^{-1} \left| \sum_{k \neq j_1} |A[j_1,k]w[k]| \right| \\
&\leq |A[j_1, j_1]|^{-1} \left| \sum_{k \neq j_1, j} |A[j_1,k]| + |A[j_1, j_1]w[j]| \right| \\
&< |A[j_1, j_1]|^{-1} \sum_{k \neq j_1} |A[j_1,k]| \\
&\leq 1
\end{aligned}$$

Induktiv folgt $|w[j_r]| < 1$, also $|w[n]| < 1$; dies ist ein Widerspruch zur Annahme.

zu a) Aus $\rho(K) = \rho(I_N - D^{-1}A) < 1$ folgt $(I_N - K)^{-1} = \sum_{k \geq 0} K^k$, also $A^{-1} = K^k D^{-1}$.

zu b) Mit dem Satz von Gerschgorin und der Diagonaldominanz von A folgt $\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda - A[n,n]| \leq |A[n,n]|, n = 1, \ldots, N\}$. Da A invertierbar ist, folgt $0 \notin \sigma(A)$; es ist also $\operatorname{Re} \sigma(A) > 0$, somit ist A positiv definit.

zu c) Aus $K = I_N - D^{-1}A = D^{-1}(D - A)$ folgt $K[n,k] \geq 0$.
$A^{-1} = \sum K^k D^{-1}$ zeigt $A^{-1}[n,k] \geq 0$.

$\square$

Krylovraum–Verfahren

Sei $B \in \mathbb{R}^{N\times N}$ ein Vorkonditionierer zu $A \in \mathbb{R}^{N,N}$. Wir betrachten wieder die lineare Iteration

$$x^{k+1} = x^k + B(b - Ax^k) = x^k + Br^k,$$

wobei r^k das *Residuum* $r^k = b - Ax^k$ sei. Für das Residuum folgt die Gleichung

$$r^{k+1} = b - Ax^{k+1} = b - A(x^k + Br^k) = (I_N - AB)r^k,$$

und somit

$$r^k = (I_N - AB)^k r^0.$$

(5.7) Definition
Zu $C \in \mathbb{R}^{N\times N}$ *und* $d \in \mathbb{R}^N$ *heißt* $\mathscr{K}_k(C,d) := \mathrm{span}\{d, Cd, \dots, C^{k-1}d\} \subset \mathbb{R}^N$ *der k-te* Krylovraum. *Es gilt:*

a) $\dim \mathscr{K}_k(C,d) \le k$

b) $\mathscr{K}_k(C,d) = \{P(C)d : P \in \mathbb{P}_k\}$

c) $x^k - x \in \mathscr{K}_k(BA, x^0 - x)$

d) $x^k \in \mathscr{K}_k(AB, r^0)$

e) $x^{k+1} \in x^0 + \mathscr{K}_k(BA, Br^0)$

(5.8) Lemma
Sei $Ax = b$ *und* $x^0 \in \mathbb{R}^N$ *Startwert,* $r^0 = b - Ax^0$*, A, B regulär. Falls* $\dim \mathscr{K}_k(AB, r^0) < k)$ *ist, dann gilt*

$$x \in x^0 + \mathscr{K}_{k-1}(BA, Br^0)$$

Beweis. Seien $r^0, ABr^0, \dots, (AB)^{k-1}r^0$ linear abhängig. Dann existieren $\alpha_0, \dots, \alpha_{k-1} \in \mathbb{R}$ mit $\alpha_0 r^0 + \alpha_1 ABr^0 + \ldots + \alpha_{k-1}(AB)^{k-1}r^0 = 0$. Ohne Einschränkung sei $\alpha_0 = 1$. Dann gilt

$$\begin{aligned} & b - Ax^0 + \alpha_1 ABr^0 + \ldots + \alpha_{k-1}(AB)^{k-1}r^0 = 0 \\ \Longrightarrow\ & b = A\underbrace{(x^0 - \alpha_1 Br^0 - \ldots - \alpha_{k-1}(BA)^{k-2}Br^0)}_{=x} \end{aligned}$$

□

Auf folgender Idee basieren die Krylovraum-Verfahren:

(1) Wähle ein geeignetes Skalarprodukt in $\mathbb{R}^N$ mit zugehöriger Norm.

(2) Konstruiere eine Orthonormalbasis $v^1, \dots, v^k$ von $\mathscr{K}_k$ bezüglich des gewählten Skalarproduktes.

(3) Approximiere $x = A^{-1}b$ mit minimalem Fehler (bzw. minimalem Residuum) in $x^0 + \mathscr{K}_k$.

Gram-Schmidt-Verfahren

Wir suchen eine Orthonormalbasis bzgl. eines Skalarprodukts $\langle\cdot,\cdot\rangle$ bzw. Norm $|x| = \sqrt{\langle\cdot,\cdot\rangle}$ des oben definierten Krylovraumes. Dazu bietet sich zunächst das Gram-Schmidt-Verfahren an. Die ersten zwei Schritte des Krylovraumverfahrens lauten also

S0) Wähle $x^0 \in \mathbb{R}^N$, $\varepsilon > 0$. Berechne $r^0 = b - Ax^0$, $z^1 = Br^0$, $h_{10} = |z^1|$ und $v^1 = \frac{1}{h_{10}} z^1$ (falls $|h_{10}| > \varepsilon$).

Setze $k = 1$.

S1) Berechne

$$\begin{aligned} w^k &= BAv^k \\ z^{k+1} &= w^k - \sum_{j=1}^{k} h_{jk} v^j \text{ mit } h_{jk} = \langle v^j, w^k\rangle \\ v^{k+1} &= \frac{1}{h_{k+1,k}} z^{k+1} \text{ mit } h_{k+1,k} = |z^{k+1}|. \end{aligned}$$

Dann ist $v^1,...,v^k$ eine Orthonormalbasis des Krylovraums

$$\mathcal{K}_k(BA, Br^0) = \operatorname{span}\{v^1,...,v^k\} = \operatorname{span}\{w^1,...,w^k\} = \operatorname{span}\{z^1,...,z^k\}$$

Es ist $|v^k| = 1$ und $\langle v^2, v^1\rangle = h_{21}\langle z^2, v^1\rangle = h_{21}(\langle w^1, v^1\rangle - h_{11}) = 0$.
Induktiv folgt $\langle v^k, v^j\rangle = 0$ für $j < k$.
Wir schreiben $BAv^k = w^k = z^{k+1} + \sum_{j=1}^{k} h_{jk} v^j = \sum_{j=1}^{k+1} h_{jk} v^j$, also

$$BAV_k = V_{k+1} H_k$$

mit $V_k = (v^1|\dots|v^k) \in \mathbb{R}^{N\times k}$ und $H_k = (h_{jm}) \in \mathbb{R}^{k+1,k}$.
Beachte, dass H_k obere Hessenbergform hat.

Bemerkung
Falls $h_{k+1,k} = 0$ ist, dann ist $w^k = BAv^k \in \operatorname{span}\{v^1,\dots,v^k\}$.
Es gilt dann $\mathcal{K}_{k+1}(BA, Br^0) = \mathcal{K}_k(BA, Br^0)$.

GMRES-Verfahren

Für das GMRES[2]-Verfahren wähle $\langle v,w\rangle_V = v^T w$. Somit gilt $V_k^T V_k = I_k$. Dieses Verfahren baut auf folgender Beobachtung auf:

(5.9) Satz
Es gilt

$$\min_{z\in x^0+\mathscr{K}_k(BA,Br^0)} |B(b-Az)|_2 = \min_{y\in\mathbb{R}^k} |h_{10}e^1 - H_k y|_2.$$

mit $e^1 = (1,0,\dots,0)^T \in \mathbb{R}^{k+1}$, $h_{10} = |Br^0|_2$, $H_k \in \mathbb{R}^{k+1,k}$.

Beweis. Es gilt

$$\begin{aligned}\min_{z\in x^0+\mathscr{K}_k(BA,Br^0)} |B(b-Az)|_2 &= \min_{y\in\mathbb{R}^k} |B(b-A(x^0+V_k y))|_2 \\ &= \min_{y\in\mathbb{R}^k} |Br^0 - BAV_k y|_2 \\ &= \min_{y\in\mathbb{R}^k} |h_{10}v^1 - V_{k+1}H_k y|_2 = \min_{y\in\mathbb{R}^k} |V_{k+1}(h_{10}e^1 - H_k y)|_2 \\ &= \min_{y\in\mathbb{R}^k} |h_{10}e^1 - H_k y|_2\end{aligned}$$

□

Algorithmus für das GMRES-Verfahren:

S0) Wähle $x^0 \in \mathbb{R}^N$, $\varepsilon > 0$. Berechne $r^0 = b - Ax^0$, $z^1 = Br^0$, $h_{10} = |z^1|_2$ und $v^1 = \frac{1}{h_{10}} z^1$. Setze $k = 1$.

S1) Berechne $w^k = BAv^k$, $z^{k+1} = w^k - \sum_{j=1}^k h_{jk} v^j$, $h_{jk} = (w^k)^T v^j$, $v^{k+1} = \frac{1}{h_{k+1,k}} z^{k+1}$, $h_{k+1,k} = |z^{k+1}|$. (D.h. $BAV_k = V_{k+1}H_k$.)

S2) Berechne $y^k \in \mathbb{R}^k$ mit
$$\rho_k = |h_{10}e^1 - H_k y^k|_2 = \min!$$
Dabei ist $H_k = (h_{jm})_{j=1,\dots,k+1,\ m=1,\dots,k} \in \mathbb{R}^{k+1,k}$.

S3) Wenn $\rho_k < \varepsilon$, setze
$$x^k = x^0 + \sum_{j=1}^k y_j^k v^j$$
und STOP.

S4) Setze $k := k+1$ und gehe zu S1).

Bemerkung
(1) Das Ausgleichsproblem in S2) ist effizient lösbar: berechne eine QR-Zerlegung von $H_k \in \mathbb{R}^{(k+1)\times k}$ *mit k Givensrotationen.*

(2) Ist $|B(b-Ax^k)|_2 = |Br^k|_2 = \rho_k$, *dann gilt* $r^k|_2 \le \|B^{-1}\|_2\rho_k$ *und es folgt* $|x - x^k| \le \|(BA)^{-1}\|_2\rho_k$

[2] **g**eneralized **m**inimal **res**idual

(3) Aus $\rho_k > 0$ folgt $r^k \neq 0$, also $x^k \neq x$. Somit ist $\dim \mathscr{K}_k(BA, Br^0) = k$, es gilt also $h_{k+1,k} \neq 0$.
Das GMRES-Verfahren ist wohldefiniert.

(4) Das Gram-Schmidt-Verfahren ist numerisch nicht stabil (beginne nach k_{max} wieder bei S0)).

(5.10) Satz
Sei BA positiv definit mit $z^T(BA)z \geq \alpha z^T z$, $\alpha > 0$, und sei $\|BA\|_2 \leq C$. Dann konvergiert das GMRES-Verfahren, und es gilt

$$|x^k - x|_2 \leq \kappa_2(BA)\sqrt{1-\frac{\alpha^2}{C^2}}^{\,k}\,|x^0 - x|$$

Beweis. Es ist

$$\begin{aligned}
Br^k &= B(b - Ax^k) = B(b - A(x^0 + V_k y^k)) \\
&= B(r^0 - AQ_k(BA)Br^0) \qquad (Q_k \in \mathbb{P}_{k-1}) \\
&= (I_N - BAQ_k(BA))Br^0 \\
&= P_k(BA)Br^0
\end{aligned}$$

mit $P_k(t) = 1 - tQ_k(t)$, d.h. $P_k \in \mathbb{P}_k$ mit $P(0) = 1$. Es gilt

$$\begin{aligned}
|Br^k|_2 &= \min_{P \in \mathbb{P}_k, P(0)=1} |P(BA)Br^0|_2 \\
&\leq \left| \left(I_N - \frac{\alpha}{C^2}BA\right)^k Br^0 \right|_2 \\
&\leq \left(\sqrt{1-\frac{\alpha^2}{C^2}}\right)^k |Br^0|_2,
\end{aligned}$$

denn

$$\begin{aligned}
\left| \left(I_N - \frac{\alpha}{C^2}BA\right)^k z \right|_2^2 &= z^T z - \frac{\alpha}{C^2} z^T BAz - \frac{\alpha}{C^2}(BAz)^T z + \frac{\alpha^2}{C^2}|BAz|_2^2 \\
&\leq z^T z - 2\frac{\alpha^2}{C^2} z^T z + \frac{\alpha^2}{C^2}|z|_2^2 \\
&= \left(1 - \frac{\alpha^2}{C^2}\right)|z|_2^2
\end{aligned}$$

Damit folgt

$$\begin{aligned}
|x^k - x|_2 &= |A^{-1}r^k|_2 = |A^{-1}B^{-1}Br^k|_2 \\
&\leq \|A^{-1}B^{-1}\|_2 |Br^k|_2 \\
&\leq \|(BA)^{-1}\|_2 \left(\sqrt{1-\frac{\alpha^2}{C^2}}\right)^k |Br^0|_2 \\
&\leq \|AB\|_2 |x - x^0|_2
\end{aligned}$$

□

Bemerkung

1. *α ist der kleinste Eigenwert von $\frac{1}{2}(BA + A^T B^T)$.*
2. *Falls A, B symmetrisch sind, ist $\kappa_2(BA) = \frac{C}{\alpha}$.*
3. *Falls A, B symmetrisch und positiv definit sind existiert eine Abschätzung mit $\sqrt{\kappa_2(BA)}$.*

Das cg-Verfahren

Seien $A, B \in \mathbb{R}^{N\times N}$ symmetrisch und positiv definit und wähle als Skalarprodukt das Energieskalarprodukt $\langle v, w\rangle_A = v^T A w$ mit zugeordneter Energienorm $|x|_A = \sqrt{\langle x, x\rangle_A}$. Bestimme x^k mit dem kleinsten Fehler in der Energienorm; genauer: zum Startwert x^0 berechne $r^0 = b - Ax^0$, $\mathcal{K}_k(BA, Br^0)$ und eine Orthonormalbasis $V_k = (v^1|\dots|v^k)$ mit $V_k^T A V_k = I_k$. Berechne $y^k \in \mathbb{R}^k$ und $x^k = x^0 + V_k y^k$ mit

$$|x^k - x|_A \le |x^0 + V_k y - x|_A \quad \forall y \in \mathbb{R}^k$$

(5.11) Satz
Für $y^k = V_k^T r^0$ und $x^k = x^0 + V_k y^k$ gilt

$$|x^k - x|_A \le |x^0 + V_k y - x|_A \quad \forall y \in \mathbb{R}^k$$

Beweis. Es gilt

$$\begin{aligned}
F(y) &= \frac{1}{2}|x^0 - V_k y - x|_A^2 \\
&= \frac{1}{2}(x^0 - V_k y - x)^T A (x^0 - V_k y - x) \\
&= \frac{1}{2}(x^0 - V_k y - x)^T (AV_k y - r^0) \\
&= \frac{1}{2} y^T V_k^T A V_k y - y^T V_k^T r^0 - \frac{1}{2}(x^0 - x)^T r^0
\end{aligned}$$

also impliziert $F(y^k) = \min!$ dass $\nabla F(y^k) = 0$ ist und somit $y^k - V_k^T r^0 = 0$.

□

Das Ziel ist es, einen Algorithmus zu entwerfen, in welchem nicht alle k Vektoren abgespeichert werden.
Eine erste Version ist gegeben durch:

S0) Wähle x^0, $\varepsilon > 0$, definiere $r^0 = b - Ax^0$, $w^0 = z^0 = Br^0$, $h_{10} = |z^0|_A$, $v^1 = \frac{1}{h_{10}} z^0$. Setze $k = 1$.

S1) Falls $|r^{k-1}| < \varepsilon$: Stop.

S2) Berechne

$$\begin{aligned}
w^{k+1} &= BAv^k \\
z^{k+1} &= w^{k+1} - \sum_{j=1}^{k} h_{jk} v^j, \quad h_{jk} = (v^j)^T A w^{k+1} \\
v^{k+1} &= \frac{1}{h_{k+1,k}} z^{k+1}, \quad h_{k+1,k} = |z^{k+1}|_A
\end{aligned}$$

S3) Berechne

$$x^{k+1} = x^k + \left((r^0)^T v^{k+1}\right) v^{k+1}$$
$$r^{k+1} = r^k - (r^0)^T v^{k+1} A v^{k+1}$$

S4) Setze $k := k+1$ und gehe zu S1)

Beobachtung Aus $h_{k+1,k}v^{k+1} = z^{k+1} = BAv^k - \sum_{j=1}^k h_{jk}v^j$ folgt $V_{k+1}H_k = BAV_k$ mit $V_k = (v^1|\dots|v^k)$ und $H_k[j,n] = h_{jn}$ symmetrischer Hessenbergmatrix, denn es gilt

$$H_k[j,k] = h_{jk} = \langle BAv^k, v^j\rangle_A = \left(BAv^k\right)^T Av^j = (v^k)^T(ABA)v^j = h_{kj}$$

Somit ist $h_{jk} = 0$ für $j = 0,\dots,k-2$ und es gilt

$$z^{k+1} = w^{k+1} - h_{kk}v^k - h_{k-1,k}v^{k-1}$$

Im k-ten Schritt werden $v^1,\dots,v^{k-2}$ nicht mehr benötigt.

(5.12) Lemma
Es gilt $\langle x^k - x, v^j\rangle_A = 0 = (r^k)^T v^j$ *für* $j = 1,\dots,k$.

Beweis. Es gilt

$$\begin{aligned}\langle x^k - x, v^j\rangle &= \langle x^0 + V_k y^k - x, v^j\rangle_A \\ &= \langle V_k y^k, v^j\rangle_A - \langle x - x^0, v^j\rangle_A \\ &= \sum_{n=1}^k y^k[n]\langle v^n, v^j\rangle_A - (r^0)^T v^j \\ &= y^k[j] - (r^0)^T v^j \\ &= 0\end{aligned}$$

□

(5.13) Folgerung
Für $j < k$ *gilt*

$$\begin{aligned}\langle Br^0, v^j\rangle_A &= \langle BA(x - x^k), v^j\rangle_A \\ &= (x - x^k)^T ABAv^j \\ &= \langle x - x^k, BAv^j\rangle_A \\ &= 0\end{aligned}$$

Also gilt

$$Br^k = \langle Br^k, v^{k+1}\rangle_A v^{k+1} + \langle Br^k, v^k\rangle_A v^k \tag{2}$$

(5.14) Lemma
Sei $d^k = \langle Br^{k-1}, v^k\rangle_A v^k$, *dann gilt*

$$x^k = x^{k-1} + \alpha_k d^k$$
$$d^{k+1} = Br^0 + \frac{\rho_k}{\rho_{k-1}} d^k$$

mit $\rho_k = (Br^k)^T r^k$ *und* $\alpha_k = \frac{\rho_{k-1}}{|d^k|_A^2}$.

Beweis. Für $k = 1$ gilt $d^1 = Br^0$. Weiter

$$\begin{aligned}(Br^k)^T r^k &= (d^{k+1})^T r^k \\ &= \langle d^{k+1}, x - x^k \rangle_A \\ &= \langle d^{k+1}, x - x^0 \rangle_A \\ &= (d^{k+1})^T r^0\end{aligned}$$

Also gilt

$$x^k = x^{k-1} + \left((r^0)^T v^k\right) v^k = x^{k-1} + \frac{(r^0)^T d^k}{|d^k|_A^2} d^k = x^{k-1} + \alpha_k d^k$$

Es gilt

$$Br^k = d^{k+1} + \frac{1}{|d^k|_A^2} \langle Br^k, d^k \rangle_A d^k$$

und schliesslich

$$\alpha_k \langle Br^k, d^k \rangle_A = (Br^k)^T (\alpha_k A d^k) = (Br^k)(r^{k-1} - r^k) = -(B^k)^T r^k = -\rho_k$$

$\square$

Der Algorithmus für das cg-Verfahren lautet somit

S0) Wähle x^0, $r^0 = b - Ax^0$, $d^1 = y^0 = Br^0$, $\rho_0 = (y^0)^T r^0$, $\varepsilon > 0$. Setze $k := 0$.

S1) Falls $\sqrt{\rho_k} < \varepsilon$: Stop

S2) Setze $k := k + 1$ und berechne

$$\begin{aligned}u^k &= Ad^k \\ \alpha_k &= \frac{\rho_{k+1}}{(u^k)^T d^k} \\ x^k &= x^{k-1} + \alpha_k d^k \\ r^k &= r^{k-1} - \alpha_k u^k \\ y^k &= Br^k \\ \rho_k &= (y^k)^T r^k \\ d^{k+1} &= y^k + \frac{\rho_k}{\rho_{k-1}} d^k\end{aligned}$$

Gehe zu S1)

(5.15) Satz
Für das cg-Verfahren gilt die Fehlerabschätzung

$$|x^k - x|_A \leq 2 \left(\frac{\sqrt{\kappa(BA)} - 1}{\sqrt{\kappa(BA)} + 1} \right)^k |x^0 - x|_A .$$

(5.16) Satz (ohne Beweis)
Es gilt

$$\min_{P\in\mathbb{P}_k, P(0)=1} \max_{t\in[a,b]} |P(t)| \leq \left(\frac{\sqrt{\frac{b}{a}}-1}{\sqrt{\frac{b}{a}}+1}\right)^k$$

Beweis. (Beweis von 5.15) Es gilt

$$\begin{aligned}
|x-x^k|_A &= \min_{y\in\mathbb{R}^k} |x^0+V_k y - x|_A \\
&= \min_{Q\in\mathbb{P}_{k-1}} |x^0 - Q(BA)Br^0 - x|_A \quad r^0 = A(x-x^0) \\
&= \min_{Q\in\mathbb{P}_{k-1}} |(I_N - Q(BA)BA)(x^0-x)|_A \quad P(t) = 1-tQ(t) \\
&= \min_{P\in\mathbb{P}_k, P(0)=1} \underbrace{|P(BA)(x^0-x)|_A}_{\leq \|P(BA)\|_A |x^0-x|_A}
\end{aligned}$$

Behauptung: $\|P(BA)\|_A \leq \max_{t\in\sigma(BA)} |P(t)|$
Aus $A = LL^T$ folgt $\sigma(BA) = \sigma(BLL^T) = \sigma(L^T BL)$, also $\|P(BA)\|_A = \|P(L^T BL)\|_2$. Somit gilt

$$\|P(BA)\|_A = \|P(L^T BL)\|_2 = \max_{\mu\in\sigma(P(L^T BL))} |\mu| = \max_{\lambda\in\sigma(L^T BL)} |P(\lambda)|$$

□