

ITERATIONSVERFAHREN

Prof. Dr. Andreas Frommer



Sommersemester 2024

Bergische Universität Wuppertal
Fachbereich Mathematik

Inhaltsverzeichnis

1	Krylov-Unterräume	2
2	Allgemeine Betrachtungen...	9
3	Die Tschebyscheff-Polynome	14
4	Das Tschebyscheff - Verfahren	22
5	Analyse von Tschebyscheff- und...	28
6	Basen für Krylov-Unterräume	35

1 Krylov-Unterräume

1.1 Definition

Zu $r \in \mathbb{C}^n \setminus \{0\}$ und $A \in \mathbb{C}^{n \times n}$ definieren wir den *Krylov-Unterraum* der Stufe m als

$$\begin{aligned} K_m(A, r) &:= \text{span}\{r, Ar, A^2r, \dots, A^{m-1}r\} \\ &= \{y \in \mathbb{C}^n : y = \sum_{j=0}^{m-1} \alpha_j A^j r\} \\ &= \{y \in \mathbb{C}^n : y = p_{m-1}(A)r, \quad p_{m-1} \in \Pi_{m-1}\}, \end{aligned}$$

wobei Π_{m-1} der Vektorraum der Polynome vom Höchstgrad $m-1$ ist, d.h.

$$\Pi_{m-1} := \left\{ p : p = \sum_{j=0}^{m-1} a_j x^j \right\}.$$

Krylov-Unterräume sind geschachtelt. Da sie alle Teilräume von \mathbb{C}^n sind, werden sie irgendwann nicht mehr größer werden. Wir untersuchen dies genauer.

1.2 Lemma

Zu $r \in \mathbb{C}^n$ existiert ein minimales $m^* \in \mathbb{N}_0$, $m^* \leq n$ und ein Polynom $p_{m^*} \neq 0$ vom Grad m^* mit

1. $p_{m^*}(A)r = 0$ (dieses Polynom ist bis auf skalare Vielfache eindeutig; alle Polynome p mit $p(A)r = 0$ bilden ein Ideal).
2. Es gilt

$$\begin{aligned} \dim(K_m(A, r)) &\leq \dim(K_{m+1}(A, r)), \quad m = 1, \dots, m^*, \\ \dim(K_{\hat{m}}(A, r)) &= \hat{m}, \quad \hat{m} = 1, \dots, m^*, \\ K_{m^*}(A, r) &= K_{m^*+1}(A, r) = K_{m^*+2}(A, r) = \dots \end{aligned}$$

Beweis:

1. Das Minimalpolynom p_A von A erfüllt $p_A(A)r = 0$, also existiert auch ein Polynom minimalen Grades mit der geforderten Eigenschaft.
2. Es ist klar, dass

$$\dim(K_m(A, r)) \leq \dim(K_{m+1}(A, r)) \leq \dim(K_m(A, r)) + 1$$

gilt, da

$$K_{m+1}(A, r) = K_m(A, r) + \langle A^m r \rangle.$$

- Für

$$p_{m^*}(t) = \sum_{j=0}^{m^*} c_j t^j, \quad c_{m^*} \neq 0$$

folgt sofort aus $p_{m^*}(A)r = 0$

$$A^{m^*}r \in K_{m^*}(A, r) \Rightarrow K_{m^*}(A, r) = K_{m^*+1}(A, r).$$

Jedes Polynom p mit $\deg(p) = m > m^*$ besitzt eine Darstellung $p = q \cdot p_{m^*} + s$ mit $\deg(s) < m^*$. Also ist $p(A)r = s(A)r$ und damit $K_m(A, r) \subseteq K_{m^*}(A, r)$, also $K_m(A, r) = K_{m^*}(A, r)$.

- Es sei \tilde{m} der erste Index mit $K_{\tilde{m}}(A, r) = K_{\tilde{m}+1}(A, r)$, dann ist

$$\begin{aligned} A^{\tilde{m}}r \in K_{\tilde{m}} &\iff \exists \text{ Polynom } p \text{ vom Grad } \tilde{m} \text{ mit } \tilde{p}(A)r = 0 \\ &\implies \tilde{m} \geq m^* \Rightarrow \tilde{m} = m^*. \end{aligned}$$

□

1.3 Satz

Sei $A \in \mathbb{C}^{n \times n}$ regulär und $x^0 \in \mathbb{C}^n$, $r^0 = b - Ax^0$. Dann erfüllt die Lösung $x^* = A^{-1}b$ von

$$Ax = b$$

die Beziehung

$$\begin{aligned} x^* &\in x^0 + K_{m^*}(A, r^0), \\ x^* &\notin x^0 + K_m(A, r^0), \quad m < m^*. \end{aligned}$$

Beweis: Sei $p_{m^*}(t) = \sum_{j=0}^{m^*} c_j t^j$ wie in Lemma 1.2. Dann ist $c_0 \neq 0$, denn sonst gälte

$$\sum_{j=1}^{m^*} c_j A^j r = 0 \iff A \underbrace{\left(\sum_{j=1}^{m^*} c_j A^{j-1} r \right)}_{\neq 0} = 0,$$

im Widerspruch zur Regularität von A .

Es gilt also

$$\begin{aligned} p_{m^*}(A)r^0 = 0 &\iff A^{-1}p_{m^*}(A)r^0 = 0 \\ &\iff A^{-1} \sum_{j=0}^{m^*} c_j A^j r^0 = 0 \end{aligned}$$

$$\begin{aligned}
&\Longleftrightarrow A^{-1} \sum_{j=0}^{m^*} c_j A^j (b - Ax^0) = 0 \\
&\implies x^* = A^{-1}b = x^0 + \frac{1}{c_0} \sum_{j=1}^{m^*} c_j A^{j-1} r^0 \\
&\quad \in x^0 + K_{m^*}(A, r^0).
\end{aligned}$$

Die Annahme $x^* \in x^0 + K_{m^*-1}(A, r^0)$ führt auf ein Polynom $p \neq 0$ mit $\deg(p) < m^*$ und $p(A)r^0 = 0$, im Widerspruch zur Minimalität von m^* . \square

1.4 Definition

Ein *Krylov-Unterraum-Verfahren* (KUV) zur Lösung von $Ax = b$ ist ein Unterraumverfahren mit Startwert x^0 , Startresiduum $r^0 = b - Ax^0$ und

$$x^m \in x^0 + K_m(A, r^0).$$

Es ist also $x^m = x^0 + q_{m-1}(A)r^0$, $\deg(q_{m-1}) \leq m-1$ und

$$\begin{aligned}
r^m &= b - Ax^m = r^0 - Aq_{m-1}(A)r^0 \\
&= p_m(A)r^0, \quad p_m \in \bar{\Pi}_m = \{p \in \Pi_m : p(0) = 1\} \\
&\in K_{m+1}(A, r^0).
\end{aligned}$$

p_m und q_{m-1} (mit $p_m(t) = 1 - t \cdot q_{m-1}(t)$) heißen auch die zum KUV gehörigen *Verfahrenspolynome*.

Umgekehrt definiert jede Folge von Polynomen $q_{m-1} \in \Pi_{m-1}$ ein KUV mit $x^m = x^0 + q_{m-1}(A)r^0$ und ebenso jede Folge von Polynomen $p_m \in \bar{\Pi}_m$ (mit $r^m = p_m(A)r^0$). Grundsätzlich besteht immer der Zusammenhang

$$(1.1) \quad p_m(t) = 1 - tq_{m-1}(t).$$

Wir interpretieren nun einfache bekannte Verfahren als KUV.

Erinnerung: Es sei $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$, dann sind $D, L, U \in \mathbb{C}^{n \times n}$ gegeben durch

$$D = \begin{pmatrix} a_{1,1} & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & \ddots & \\ & 0 & & & \ddots \\ & & & & & a_{n,n} \end{pmatrix} \quad (\text{Diagonalteil})$$

$$L = \begin{pmatrix} 0 & & & & \\ -a_{2,1} & \ddots & & & \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ -a_{n,1} & \dots & \dots & -a_{n,n-1} & 0 \end{pmatrix} \quad \begin{array}{l} \text{(negativer)} \\ \text{linker unterer Dreiecksteil} \end{array}$$

$$U = \begin{pmatrix} 0 & -a_{1,2} & \dots & \dots & -a_{1,n} \\ & \ddots & & & \vdots \\ & & \ddots & & \vdots \\ & & & \ddots & -a_{n-1,n} \\ 0 & & & & 0 \end{pmatrix} \quad \begin{array}{l} \text{(negativer)} \\ \text{rechter oberer Dreiecksteil} \end{array}$$

Es ist also $A = D - L - U$.

Wir betrachten nun die drei Standard-Verfahren:

Jacobi: $x^{m+1} = D^{-1}((L + U)x^m + b)$

Gauß-Seidel: $x^{m+1} = (D - L)^{-1}(Ux^m + b)$

SOR: $x^{m+1} = (\frac{1}{\omega}D - L)^{-1}((\frac{1-\omega}{\omega}D + U)x^m + b), \quad \omega \in \mathbb{R} \setminus \{0\}$

und interpretieren sie folgendermaßen als KUV:

Jacobi: Setze $H := D^{-1}(L + U)$ und $A' = D^{-1}A$, $b' = D^{-1}b$. Mit $r^0 = b' - A'x^0$ erhalten wir

$$\begin{aligned} x^{m+1} &= Hx^m + b' \\ r^{m+1} &= b' - A'x^{m+1} = b' - A'(Hx^m + b') \\ &= b' - A'((I - A')x^m + b') \\ &= b' - A'(x^m + r^m) \\ &= r^m - A'r^m = (I - A')r^m \\ \Rightarrow r^m &= (I - A')^m r^0. \end{aligned}$$

Also ist das Jacobi-Verfahren ein KUV bezüglich $A'x = b'$ mit

$$\begin{aligned} p_m(t) &= (1 - t)^m, \\ q_{m-1}(t) &= \frac{1 - p_m(t)}{t} = \sum_{j=0}^{m-1} (-1)^j \binom{m}{j+1} t^j. \end{aligned}$$

Gauß-Seidel: ist KUV bezüglich $\underbrace{(D - L)^{-1} A}_{=A'} x = \underbrace{(D - L)^{-1} b}_{b'}$ mit

$$p_m(t) = (1 - t)^m,$$

$$q_{m-1}(t) = \frac{1 - p_m(t)}{t} = \sum_{j=0}^{m-1} (-1)^j \binom{m}{j+1} t^j.$$

SOR: ist KUV bezüglich $\underbrace{(\frac{1}{\omega} D - L)^{-1} A}_{=A'} x = \underbrace{(\frac{1}{\omega} D - L)^{-1} b}_{b'}$ mit

$$p_m(t) = (1 - t)^m,$$

$$q_{m-1}(t) = \frac{1 - p_m(t)}{t} = \sum_{j=0}^{m-1} (-1)^j \binom{m}{j+1} t^j.$$

Allgemein gilt folgender

1.5 Satz

Ein Iterationsverfahren

$$(1.2) \quad x^{m+1} = Hx^m + b'$$

zur Lösung von $Ax = b$ mit $A = M - N$, $H = M^{-1}N$, $b' = M^{-1}b$ ist ein KUV (bezüglich des Systems $A'x = b'$, $A' = M^{-1}A$) mit $p_m(t) = (1 - t)^m$.

Bemerkung: Die Verfahrenspolynome hängen bei diesen einfachen Iterationen nicht vom Startresiduum ab.

1.6 Definition

Ein KUV zur Lösung von $Ax = b$ mit regulärem A heißt *konvergent*, falls

$$\lim_{k \rightarrow \infty} r^k = 0 \iff \lim_{k \rightarrow \infty} x^k = x^* = A^{-1}b.$$

1.7 Satz

Ein Iterationsverfahren der Gestalt (1.2) ist konvergent, falls

$$\rho(I - M^{-1}A) < 1$$

gilt.

Beweis: Es gilt $r^m = (I - M^{-1}A)^m r^0$, wegen $\rho(I - M^{-1}A) < 1$ gilt $\lim_{k \rightarrow \infty} (I - M^{-1}A)^k = 0$, also $\lim_{k \rightarrow \infty} r^k = 0$. \square

1.8 Definition

Das *Richardson-Verfahren erster Ordnung* zur Lösung von $Ax = b$ ist die Iteration

$$x^{m+1} = (I - A)x^m + b = x^m + r^m.$$

Die Richardson-Iteration erster Ordnung ist im Fall $D = I$ mit der Jacobi-Iteration identisch. Für die Verfahrenspolynome gilt wieder

$$p_m(t) = (1 - t)^m.$$

1.9 Korollar

Das Richardson-Verfahren erster Ordnung zur Lösung von $Ax = b$ konvergiert, falls

$$\rho(I - A) < 1,$$

z.B. falls $\text{spek}(A) \subset (0, 2)$.

1.10 Definition

Das *relaxierte Richardson-Verfahren erster Ordnung* zur Lösung von $Ax = b$ ist die Iteration

$$x^{m+1} = x^m + \alpha r^m.$$

Hierbei ist $\alpha \in \mathbb{C}$ ein fester Parameter, der bei geeigneter Wahl die Konvergenz beschleunigt.

Die Verfahrenspolynome für das relaxierte Richardson-Verfahren sind

$$(1.3) \quad p_m(t) = (1 - \alpha t)^m.$$

1.11 Definition

Das *Richardson-Verfahren zweiter Ordnung* zur Lösung von $Ax = b$ ist bei gegebenem x^0 die Iteration

$$\begin{aligned} x^1 &= (I - A)x^0 + b, \\ x^{m+1} &= \alpha((I - A)x^m + b) + (1 - \alpha)x^{m-1}, \quad m \geq 1, \end{aligned}$$

mit $\alpha \in \mathbb{C}$ fest.

Interpretation als KUV:

$$\begin{aligned} r^1 &= (I - A)r^0 \quad \Rightarrow \quad p_1(t) = 1 - t \\ r^{m+1} &= b - Ax^{m+1} \\ &= b - \alpha A((I - A)x^m + b) - (1 - \alpha)Ax^{m-1} \\ &= \alpha r^m - \alpha Ar^m + (1 - \alpha)r^{m-1} \\ &= \alpha(I - A)r^m + (1 - \alpha)r^{m-1} \end{aligned}$$

also

$$p_{m+1}(t) = \alpha(1-t)p_m(t) + (1-\alpha)p_{m-1}(t), \quad m \geq 1.$$

Solche *3-Term-Rekursionen* sind typisch für „fortgeschrittene“ KUV. Für eine spätere Analyse halten wir fest, dass gilt

$$(1.4) \quad \begin{pmatrix} p_{m+1}(t) \\ p_m(t) \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha(1-t) & 1-\alpha \\ 1 & 0 \end{pmatrix}}_B \cdot \begin{pmatrix} p_m(t) \\ p_{m-1}(t) \end{pmatrix}.$$

2 Allgemeine Betrachtungen zur Konvergenz von Krylov-Unterraum-Verfahren

Für jedes Krylov-Unterraum-Verfahren zur Lösung von $Ax = b$, $A \in \mathbb{C}^{n \times n}$, $b \in \mathbb{C}^n$ gilt

$$r^m = b - Ax^m = p_m(A)r^0, \quad p_m \in \bar{\Pi}_m.$$

Ein KUV ist konvergent, wenn

$$\lim_{m \rightarrow \infty} p_m(A)r^0 = 0$$

gilt. Bei geeigneter Wahl von p_m wird $p_m(A)r^0 = 0$ bereits für ein endliches $m \leq n$; manche Verfahren erreichen dies auch tatsächlich. Andere Verfahren (wie z.B. das Richardson-Verfahren) benötigen eine unendliche Iterationszahl.

Man unterscheidet zwei verschiedene Ansätze zur Analyse von KUV:

1. parameterfreie KUV: Sie benötigen keine (detailliertere) Information über Eigenschaften von A (z.B. das Spektrum). In der Regel bilden sie $p_m(A)$ dann in expliziter Abhängigkeit von $r^0 \hookrightarrow$ später (vergleiche cg-Verfahren),
2. parameterabhängige KUV verwenden detailliertere Informationen über A , z.B. über die Lage des Spektrums. In der Regel bilden sie dann $p_m(A)$ unabhängig von r^0 . Vorteil: keine Innenproduktbildung, damit besser parallelisierbar; Nachteil: eventuell weniger effizient.

Frage: Wann gilt $\lim_{m \rightarrow \infty} p_m(A) = 0 \in \mathbb{C}^{n \times n}$?

Wäre A diagonalisierbar, d.h. es gibt eine invertierbare Matrix T und eine Diagonalmatrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, so dass

$$A = T\Lambda T^{-1},$$

dann gilt

$$p_m(A) = Tp_m(\Lambda)T^{-1}.$$

Damit ist dann

$$\lim_{m \rightarrow \infty} p_m(A) = 0 \iff \lim_{m \rightarrow \infty} p_m(\lambda) = 0 \quad \forall \lambda \in \text{spek}(A).$$

Ist A nicht diagonalisierbar, so werden weitere Bedingungen notwendig.

2.1 Satz

Es sei

$$A = T J T^{-1}$$

mit

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix} \quad J_\ell = \begin{pmatrix} \lambda_\ell & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_\ell \end{pmatrix} \in \mathbb{C}^{n_\ell \times n_\ell}$$

die Jordan-Normalform von A mit $\lambda_\ell \in \text{spek}(A) = \{\lambda_1, \dots, \lambda_k\}$ (aber Mehrfachnennungen möglich). Dann gilt

$$\lim_{m \rightarrow \infty} p_m(A) = 0 \iff \lim_{m \rightarrow \infty} p_m^{(\nu)}(\lambda_\ell) = 0 \quad \ell = 1, \dots, k, \quad \nu = 0, \dots, n_\ell - 1.$$

Beweis: Es gilt

$$p_m(A) = T p_m(J) T^{-1},$$

also

$$\lim_{m \rightarrow \infty} p_m(A) = 0 \iff \lim_{m \rightarrow \infty} p_m(J_\ell) = 0 \quad \ell = 1, \dots, k.$$

Wir betrachten also

$$J_\ell = \lambda_\ell \cdot I_{n_\ell} + S, \quad S = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{C}^{n_\ell \times n_\ell} \Rightarrow S^{n_\ell} = 0.$$

Damit ergibt sich (mit $\binom{k}{\nu} = 0$, falls $\nu > k$)

$$J_\ell^k = \sum_{\nu=0}^{n_\ell-1} \binom{k}{\nu} S^\nu \lambda_\ell^{k-\nu}.$$

Es sei nun $p_m(t) = \sum_{k=0}^m c_k t^k$, dann ist

$$\begin{aligned} p_m(J_\ell) &= \sum_{k=0}^m c_k \sum_{\nu=0}^{n_\ell-1} \binom{k}{\nu} S^\nu \lambda_\ell^{k-\nu} \\ &= \sum_{\nu=0}^{n_\ell-1} \frac{1}{\nu!} S^\nu \underbrace{\sum_{k=0}^m c_k \frac{k!}{(k-\nu)!} \lambda_\ell^{k-\nu}}_{p_m^{(\nu)}(\lambda_\ell)}. \end{aligned}$$

Hierin bezeichnet $p_m^{(\nu)}$ die ν -te Ableitung des Polynoms p_m .
In der Matrix S^ν ist genau die ν -te obere Nebendiagonale (mit Einsen) besetzt. Es gilt also

$$\lim_{m \rightarrow \infty} p_m(J_\ell) = 0 \iff \lim_{m \rightarrow \infty} p_m^{(\nu)}(\lambda_\ell) = 0 \quad \nu = 0, \dots, n_\ell - 1.$$

□

2.2 Korollar

Sei $A \in \mathbb{C}^{n \times n}$ diagonalisierbar (z.B. hermitesch oder normal¹). Dann gilt

$$\lim_{m \rightarrow \infty} p_m(A) = 0 \iff \lim_{m \rightarrow \infty} p_m(\lambda) = 0 \quad \forall \lambda \in \text{spek}(A).$$

Für diagonalisierbares A ist also die Größe $\max_{\lambda \in \text{spek}(A)} |p_m(\lambda)|$ ein Maß für die Größe von $p_m(A)$, sie ist sogar eine Norm von $p_m(A)$:

$$A = SDS^{-1}, \quad D = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Nehme

$$\|x\|_S = \|S^{-1}x\|_2,$$

dann ist

$$\begin{aligned} \|A\|_S &= \max_{\|x\|_S \neq 0} \frac{\|Ax\|_S}{\|x\|_S} = \max_{\|x\|_S \neq 0} \frac{\|DS^{-1}x\|_2}{\|S^{-1}x\|_2} \\ &= \max_{y \neq 0} \frac{\|Dy\|_2}{\|y\|_2} = \|D\|_2 \\ &= \max_{\lambda \in \text{spek}(A)} |\lambda| \end{aligned}$$

und analog

$$\|p_m(A)\|_S = \max_{\lambda \in \text{spek}(A)} |p_m(\lambda)|.$$

Minimierung von

$$\max_{\lambda \in \text{spek}(A)} |p_m(\lambda)|$$

minimiert also auch $\|p_m(A)\|_S$.

Bemerkung: $\|\cdot\|_S$ ist i.A. eine „schiefe“ Norm. Ist A jedoch normal, so kann man S unitär wählen, so dass $\|x\|_S = \|x\|_2$ und damit $\|A\|_S = \|A\|_2 = \max_{\lambda \in \text{spek}(A)} |p_m(\lambda)|$.

¹d.h. $AA^* = A^*A$, wobei A^* die Adjungierte von A bezeichnet

Im diagonalisierbaren Fall löst damit ein bestmögliches Iterationsverfahren für jedes m die MinMax-Aufgabe

$$(2.1) \quad \min_{p_m \in \overline{\Pi}_m} \max_{\lambda \in \text{spek}(A)} |p_m(\lambda)|.$$

Dies ist jedoch in den meisten Fällen praktisch nicht möglich, schon alleine weil man i.A. die Eigenwerte von A nicht alle kennt.

Parameterabhängige KUV verwenden deshalb nur partielle Informationen über $\text{spek}(A)$, um mittels dieser Informationen möglichst „gute“ Polynome $p_m(t)$ zu finden.

Es sei z.B. (mittels Satz von Gerschgorin o.Ä.) bekannt, dass

$$\text{spek}(A) \subset D(\zeta, \rho) = \{z \in \mathbb{C} : |z - \zeta| \leq \rho\}.$$

Wie lautet die Lösung von

$$\min_{p_m \in \overline{\Pi}_m} \max_{\lambda \in D(\zeta, \rho)} |p_m(\lambda)| ?$$

Beachte: $0 \notin D(\zeta, \rho)$ ist hier eine vernünftige zusätzliche Bedingung, da ansonsten die Lösung $p_m \equiv 1$ lautet für alle m , d.h. es resultiert ein nicht konvergentes KUV.

2.3 Satz

Sei $\rho > 0$, sei $\gamma \in \mathbb{C}$ mit $|\gamma| > \rho$. Dann löst das Polynom

$$p_m(z) = \frac{z^m}{\gamma^m}$$

die Aufgabe

$$\min_{p_m \in \Pi_m, p_m(\gamma)=1} \max_{\lambda \in D(0, \rho)} |p_m(\lambda)|,$$

der Wert ist $|\rho/\gamma|^m$.

Bemerkung: $|\rho/\gamma|$ ist klein, wenn γ weit vom Kreis $D(0, \rho)$ entfernt ist. Zum Beweis des Satzes benötigen wir den Satz von Rouché.

2.4 Lemma (Rouché)

Es seien f, g holomorph im Gebiet $\Omega \subset \mathbb{C}$. Es sei Γ ein einfach geschlossener, in Ω nullhomologer Weg (z.B. der Rand eines Kreises in \mathbb{C}), so dass gilt

$$|f(\zeta) - g(\zeta)| < |g(\zeta)| \quad \text{für alle } \zeta \in \Gamma.$$

Dann haben f und g gleich viele Nullstellen im Inneren von Γ .

Vergleiche REMMERT, Funktionentheorie I. Grundlehren, Springer, 4. Auflage; Seite 310f.

Beweis: Sei $p_m^*(z) = \frac{z^m}{\gamma^m}$. Sei $p_m \in \Pi_m$, $p_m(\gamma) = 1$ ein Polynom mit

$$\max_{\lambda \in D(0, \rho)} |p_m(\lambda)| < \max_{\lambda \in D(0, \rho)} |p_m^*(\lambda)|.$$

Dann gilt

$$|(p_m^*(z) - p_m(z)) - p_m^*(z)| = |p_m(z)| < \left| \frac{\rho}{\gamma} \right|^m = |p_m^*(z)| \text{ für } z \in \partial D(0, \rho).$$

Also besitzen $p_m^* - p_m$ und p_m^* gleich viele Nullstellen im Inneren von $D(0, \rho)$, nämlich m Stück. Außerdem gilt $(p_m^* - p_m)(\gamma) = 1 - 1 = 0$, also besitzt das Polynom $p := p_m^* - p_m \in \Pi_m$ mindestens $m + 1$ Nullstellen, ist also nach dem Fundamentalsatz der Algebra das Nullpolynom. \square

2.5 Aufgabe

Finde ausgehend von Satz 2.3 die Lösung von

$$\min_{p_m \in \Pi_m} \max_{\lambda \in D(1, \rho)} |p_m(\lambda)|.$$

Antwort: s. Übung.

Wir erhalten so die folgende Interpretation: Bezüglich der Information

$$\text{spek}(A) \subseteq D(1, \rho)$$

mit $\rho < 1$ ist das Richardson-Verfahren *optimal*.

3 Die Tschebyscheff-Polynome

3.1 Definition

Die *Tschebyscheff-Polynome* $T_m(z)$ sind gegeben durch die Rekursion

$$T_{m+1}(z) = 2zT_m(z) - T_{m-1}(z), \quad m \geq 1,$$

mit

$$T_0(z) = 1, \quad T_1(z) = z.$$

Es ist bekannt: Die $T_m(z)$ sind Orthogonalpolynome bezüglich des Innenproduktes

$$\langle p, q \rangle = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} p(x)q(x) dx, \quad \text{auf } \Pi = \text{Menge der Polynome.}$$

3.2 Satz

Die Lösung p_m der Minimierungsaufgabe

$$\min_{p_m \in \Pi_m, p_m(c)=1} \max_{\lambda \in [-1,1]} |p_m(\lambda)|,$$

für $c \notin [-1, 1]$ lautet

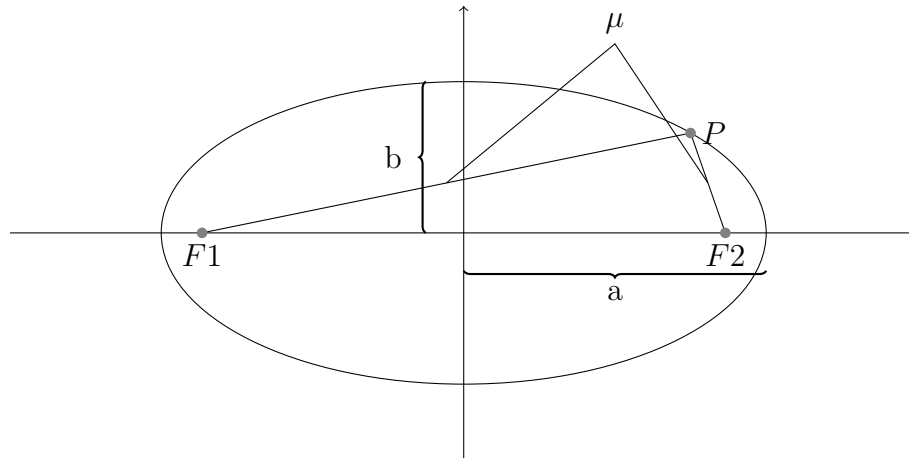
$$p_m(z) = \frac{1}{T_m(c)} T_m(z).$$

Beweis: Siehe Vorlesung: Parallele Algorithmen (Vorlesung im Wintersemester 2002/03, Abschnitt 4.4). \square

Wir interessieren uns nun für Optimalpolynome auf etwas allgemeineren Gebieten, nämlich Ellipsen.

3.3 Definition

Für $\rho \geq 1$ sei E_ρ die Ellipse mit Brennpunkten in -1 und $+1$ und mit großer Halbachse $a = (1/2) \cdot (\rho + 1/\rho)$ und kleiner Halbachse $b = (1/2) \cdot (\rho - 1/\rho)$. Die „Schnurlänge“ μ beträgt also $\rho + 1/\rho$.



Die hier gegebene Definition einer Ellipse über ihre Brennpunkte und den Parameter ρ ist nicht Standard. Man rechnet aber einfach nach, dass so tatsächlich eine Ellipse gegeben ist, indem man nachweist, dass die Summe der Abstände zu den beiden Brennpunkten für jeden Punkt auf E_ρ gerade die Schnurlänge μ ist.

3.4 Definition

Die *Joukowski-Abbildung* $J : \mathbb{C} \setminus \{0\} \longrightarrow \mathbb{C}$ ist definiert durch

$$J(w) = \frac{1}{2}(w + w^{-1}).$$

Mit Hilfe dieser Abbildung wollen wir Folgendes zeigen:

3.5 Lemma

Es ist $J(D(0, \rho)) = E_\rho$.

Beweis: Wir stellen als erstes die Gleichung für E_ρ auf ($z = x + iy$):

$$\frac{x^2}{\left(\frac{\rho+1/\rho}{2}\right)^2} + \frac{y^2}{\left(\frac{\rho-1/\rho}{2}\right)^2} = 1.$$

Eine Parametrisierung von E_ρ ist dann

$$x(t) = \frac{\rho + 1/\rho}{2} \cdot \cos t, \quad y(t) = \frac{\rho - 1/\rho}{2} \cdot \sin t$$

mit $t \in [0, 2\pi)$. Dann gilt

$$x^2(t) + y^2(t) = \left(\frac{\rho + 1/\rho}{2}\right)^2 \cos^2 t + \left(\frac{\rho - 1/\rho}{2}\right)^2 \sin^2 t$$

$$= \frac{1}{4} (\rho + (1/\rho))^2 - \sin^2 t.$$

Andererseits gilt mit $w = \rho e^{i\theta} \in D(0, \rho)$ für das Bild $z = J(w) = \frac{1}{2}(w + w^{-1})$

$$\begin{aligned} z\bar{z} &= \frac{1}{4}(w + w^{-1})(\bar{w} + \bar{w}^{-1}) \\ &= \frac{1}{4}(w\bar{w} + w^{-1}\bar{w} + w\bar{w}^{-1} + w^{-1}\bar{w}^{-1}) \\ &= \frac{1}{4}\left(\rho^2 + e^{-2i\theta} + e^{2i\theta} + \frac{1}{\rho^2}\right) \\ &= \frac{1}{4}\left(\rho^2 + 2\cos(2\theta) + \frac{1}{\rho^2}\right) \\ &= \frac{1}{4}\left(\rho^2 - 4\sin^2(\theta) + 2 + \frac{1}{\rho^2}\right) \\ &= \frac{1}{4}\left(\rho^2 + 2 + \frac{1}{\rho^2}\right) - \sin^2(\theta) \\ &= \frac{1}{4}\left(\rho + \frac{1}{\rho}\right)^2 - \sin^2(\theta). \end{aligned}$$

Vergleich mit der Parameterdarstellung ergibt einfach

$$\theta = t.$$

□

Die Gleichung $z = J(w) = \frac{1}{2}(w + w^{-1})$ hat zwei Lösungen w , die zueinander invers sind. Sei nun $w(z)$ eine dieser Lösungen z.B. die betragsmäßig größere.

3.6 Lemma

Es gilt

$$T_m(z) = \left[\frac{1}{2} (w^m(z) + w^{-m}(z)) \right].$$

Beweis: Wir beweisen dies natürlich über die Rekursionsformel. Für $m = 0, 1$ ist die Behauptung offensichtlich richtig und per Induktion gilt

$$\begin{aligned} T_{m+1}(z) &= z(w^m(z) + w^{-m}(z)) - \frac{1}{2}(w^{m-1}(z) + w^{-(m-1)}(z)) \\ &= \frac{1}{2}(w(z) + w^{-1}(z))(w^m(z) + w^{-m}(z)) \\ &\quad - \frac{1}{2}(w^{m-1}(z) + w^{-(m-1)}(z)) \\ &= \frac{1}{2}(w^{m+1}(z) + w^{-(m+1)}(z)). \end{aligned}$$

□

Betrachtet wird nun die MinMax-Aufgabe

$$(3.1) \quad \min_{p \in \Pi_m, p(\gamma)=1} \max_{\lambda \in E_\rho} |p(\lambda)| \quad \text{für } \gamma \notin E_\rho.$$

3.7 Satz

Sei c_m der Wert des Minimums in (3.1) und $\gamma = \frac{1}{2}(w_\gamma + w_\gamma^{-1})$, wobei w_γ die betragsmäßig größere der beiden möglichen Lösungen ist. Dann gilt

$$\frac{\rho^m}{|w_\gamma|^m} \leq c_m \leq \frac{1}{|w_\gamma^m + w_\gamma^{-m}|} \left(\rho^m + \left(\frac{1}{\rho} \right)^m \right),$$

wobei die obere Schranke für das Polynom

$$p_m^*(z) = \frac{1}{T_m(\gamma)} T_m(z)$$

erreicht wird.

Beweis: Wir stellen ein beliebiges Polynom $p_m \in \bar{\Pi}_m$ in der Basis der Tschebyscheff-Polynome dar, d.h.

$$p_m = \sum_{\nu=0}^m a_\nu T_\nu.$$

Wegen $p_m(\gamma) = 1$ gilt $1 = \sum_{\nu=0}^m a_\nu T_\nu(\gamma)$. Unter Verwendung von $w = w(z)$ wie im Lemma 3.6 ergibt sich

$$\begin{aligned} p_m(z) &= \sum_{\nu=0}^m a_\nu \left(\frac{1}{2} (w^\nu + w^{-\nu}) \right) \\ &= \sum_{\nu=0}^m a'_\nu \left(\frac{\frac{1}{2} (w^\nu + w^{-\nu})}{\frac{1}{2} (w_\gamma^\nu + w_\gamma^{-\nu})} \right). \end{aligned}$$

Hierin ist $\sum_{\nu=0}^m a'_\nu = 1$. Wählen wir speziell $a'_m = 1$, $a'_\nu = 0$, $\nu = 0, \dots, m-1$, also

$$p_m(z) = p_m^*(z) = \frac{1}{T_m(\gamma)} T_m(z) = \frac{w^m + w^{-m}}{w_\gamma^m + w_\gamma^{-m}},$$

so gilt wegen des Maximums-Prinzips

$$\max_{z \in E_\mu} |p_m^*(z)| = \max_{z \in \partial E_\mu} |p_m^*(z)|$$

$$\begin{aligned}
&= \max_{w \in \partial D(0, \rho)} \left| \frac{w^m + w^{-m}}{w_\gamma^m + w_\gamma^{-m}} \right| \\
&= \frac{1}{|w_\gamma^m + w_\gamma^{-m}|} \left[\rho^m + \left(\frac{1}{\rho} \right)^m \right].
\end{aligned}$$

Die letzte Gleichheit ist richtig, weil $|w^m + w^{-m}| \leq |w^m| + |w^{-m}| = |\rho^m| + |\rho^{-m}|$, und diese obere Schranke wird für $w = \rho$ auch angenommen.

Für die Bestimmung einer unteren Schranke sei das Optimalpolynom $\tilde{p}_m(z)$ dargestellt durch

$$\begin{aligned}
\tilde{p}_m(z) &= \sum_{\nu=0}^m b_\nu \left(\frac{w^\nu + w^{-\nu}}{w_\gamma^\nu + w_\gamma^{-\nu}} \right), \quad \sum_{\nu=0}^m b_\nu = 1 \\
&= \left(\frac{w^{-m}}{w_\gamma^{-m}} \right) \underbrace{\sum_{\nu=0}^m b_\nu \left(\frac{w^{\nu+m} + w^{-\nu+m}}{w_\gamma^{\nu+m} + w_\gamma^{-\nu+m}} \right)}_{\text{Polynom } q \text{ vom Grad } 2m \text{ in } w, q(w_\gamma) = 1}.
\end{aligned}$$

Nach Satz 2.3 gilt

$$\max_{w \in D(0, \rho)} \left| \sum_{\nu=0}^m b_\nu \left(\frac{w^{\nu+m} + w^{-\nu+m}}{w_\gamma^{\nu+m} + w_\gamma^{-\nu+m}} \right) \right| \geq \left(\frac{\rho}{|w_\gamma|} \right)^{2m},$$

falls w_γ die betragsmäßig größere Lösung ist (und damit außerhalb von $D(0, \rho)$ liegt). Also gilt

$$|\tilde{p}_m(z)| \geq \left| \frac{w^{-m}}{w_\gamma^{-m}} \right| \left(\frac{\rho}{|w_\gamma|} \right)^{2m} = \frac{|w^{-m}| \rho^{2m}}{|w_\gamma|^m} = \frac{\rho^m}{|w_\gamma|^m}.$$

□

3.8 Bemerkung

Für $\gamma \notin E_\rho$ gilt $|w_\gamma| < \rho^{-1}$ mit $\rho \geq 1$. Also ist sowohl $\lim_{m \rightarrow \infty} (1/\rho)^m = 0$ als auch $\lim_{m \rightarrow \infty} w_\gamma^{-m} = 0$. Satz 3.7 zeigt so, dass die Tschebyscheff-Polynome *asymptotisch optimal* sind, aber nicht notwendig optimal für endliches m .

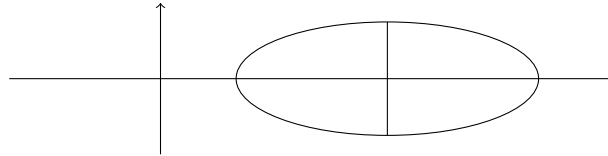
3.9 Aufgabe

Bestimme jeweils die (im Sinne von Satz 3.7) asymptotisch optimale Lösung der MinMax-Aufgabe

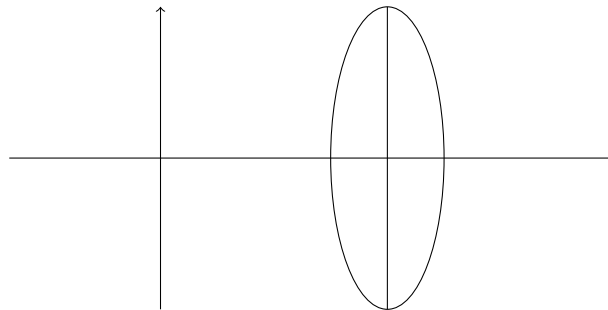
$$\min_{p_m \in \Pi_m} \max_{\lambda \in E} |p(\lambda)|$$

und den zugehörigen Wert des Minimums für die Fälle:

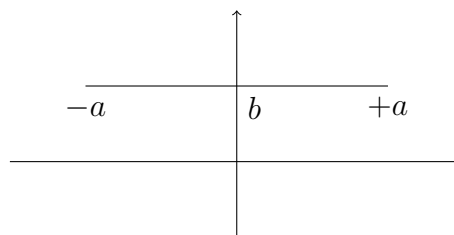
1. $E = a + bE_\rho$, wobei $a, b > 0$ und $a - b \cdot \frac{1}{2}(\rho + 1/\rho) > 0$



2. $E = a + be^{i\frac{\pi}{2}}E_\rho$, wobei $a, b > 0$ und $a - b \cdot \frac{1}{2}(\rho + \frac{1}{\rho}) > 0$



3. E_ρ ist ein Geradenstück parallel zur reellen Achse von $-a$ bis $+a$ mit Achsenabschnitt b auf der imaginären Achse.



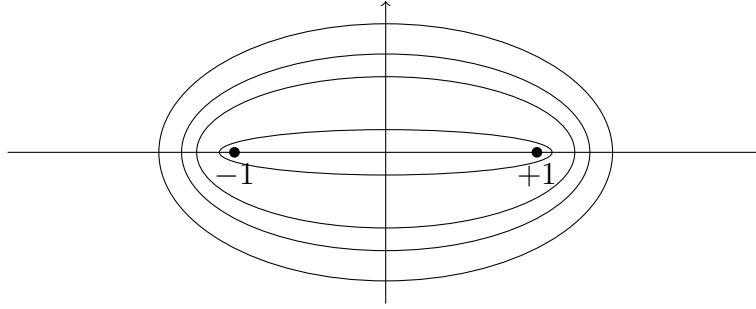
Sind in diesem Fall die Polynome auch optimal für endliches m ?

3.10 Bemerkung

Die Optimalpolynome im Sinne von Satz 3.7 sind unabhängig vom Parameter ρ der Ellipse E_ρ (solange γ immer noch außerhalb der Ellipse liegt).

Aus Satz 2.3 ist bekannt, dass für den Kreis $D(0, \rho)$ die MinMax-Aufgabe

$$\min_{p \in \Pi_m, p(\gamma)=1} \max_{\lambda \in D(0, \rho)} |p_m(\lambda)|$$

Abbildung 1: Ellipsen E_ρ für verschiedene Werte von ρ

durch das Polynom $p_m(z) = \frac{z^m}{\gamma^m}$ gelöst wird. Für dessen Ableitungen $p_m^{(k)}(z) = m(m-1) \cdot \dots \cdot (m-k+1) \frac{z^{m-k}}{\gamma^m}$ gilt

$$\begin{aligned} \max_{z \in D(0, \rho)} |p_m^{(k)}(z)| &= m \cdot \dots \cdot (m-k+1) \frac{\rho^{m-k}}{\gamma^m} \\ &= m \cdot \dots \cdot (m-k+1) \rho^{-k} \left(\frac{\rho}{\gamma} \right)^m \xrightarrow{m \rightarrow \infty} 0, \end{aligned}$$

wobei die Konvergenz sogar gleichmäßig ist auf $D(0, \rho)$. Dies ist wichtig mit Blick auf Satz 2.1

Für die Tschebyscheffpolynome und Ellipsen gilt ein entsprechendes Resultat.

3.11 Satz

Es sei $E_\rho = J(D(0, \rho))$ mit $\rho > 1$ eine Ellipse. Weiter sei p_m^* das m -te skalierte Tschebyscheffpolynom, $p_m^*(z) = \frac{w^m + w^{-m}}{w_\gamma^m + w_\gamma^{-m}}$, wobei $z = J(w) = \frac{1}{2}(w + w^{-1})$ und $J(w_\gamma) = \gamma$, mit γ außerhalb von E_ρ , d.h. $|w_\gamma| > \rho$. Dann gilt für alle $k \in \mathbb{N}$

$$\left| p_m^{*(k)}(z) \right| \xrightarrow{m \rightarrow \infty} 0 \quad \text{gleichmäßig auf } E_\rho.$$

Beweis: Beachte (im Vorgriff auf die Formulierung des Tschebyscheff-Verfahrens in Algorithmus 4.7): Im Zusammenhang mit Satz 2.1 besagt dieser Satz, dass Diagonalisierbarkeit für ein erfolgreiches Tschebyscheffverfahren nicht vorausgesetzt werden muss.

Mit

$$\begin{aligned} p_m^*(z) &= \frac{w^m + w^{-m}}{w_\gamma^m + w_\gamma^{-m}}, \\ \frac{dz}{dw} &= \frac{1}{2}(1 - w^{-2}) \quad \text{und} \quad \frac{dw}{dz} \Big|_{z=J(w)} = \frac{2}{1 - w^{-2}}, \end{aligned}$$

erhalten wir

$$\begin{aligned} \frac{dp_m^*(z)}{dw} \cdot \frac{dw}{dz} &= \frac{mw^{m-1} - mw^{-(m+1)}}{w_\gamma^m + w_\gamma^{-m}} \cdot \frac{2}{1 - w^{-2}} \\ &= 2 \frac{mw^{m+1} - mw^{-m+1}}{(w_\gamma^m + w_\gamma^{-m})(w^2 - 1)}. \end{aligned}$$

Der Term $s_m(w) = m \frac{w^{m+1}}{(w^2-1)(w_\gamma^m + w_\gamma^{-m})}$ wird für $|w| = \rho$ abgeschätzt durch

$$|s_m(w)| \leq m \frac{1}{\rho^2 - 1} \cdot \frac{\rho^{m+1}}{|w_\gamma^m + w_\gamma^{-m}|} \leq \frac{m}{\rho^2 - 1} \cdot \frac{\rho}{|w_\gamma/\rho|^m - |w_\gamma\rho|^{-m}}.$$

Hierin ist die rechte obere Schranke unabhängig von $w \in \partial D(0, \rho)$. Wegen des Maximumsprinzips gilt die Schranke sogar für alle $w \in D(0, \rho)$, also für alle $z \in E_\rho$. Aus

$$|w_\gamma/\rho| > 1, \quad |w_\gamma\rho| > 1,$$

erkennen wir mit der Regel von de l'Hôpital, dass die obere Schranke für $m \rightarrow \infty$ gegen 0 geht. Dies beweist den Satz für die erste Ableitung; für höhere Ableitungen geht der Beweis (induktiv) analog. \square

4 Das Tschebyscheff - Verfahren

4.1 Definition

Mit $E(f_1, f_2, \rho)$ bezeichnen wir die Ellipse mit den Brennpunkten f_1, f_2 und den Halbachsen $\frac{1}{4}|f_1 - f_2| \left(\rho + \frac{1}{\rho}\right)$ und $\frac{1}{4}|f_1 - f_2| \left(\rho - \frac{1}{\rho}\right)$, $\rho \geq 1$.

Wir betrachten nun folgende Situation:

Gegeben: $Ax = b$, A regulär.

Bekannt sei: $\text{spek}(A) \subseteq E(f_1, f_2, \rho)$ mit $0 \notin E(f_1, f_2, \hat{\rho})$.

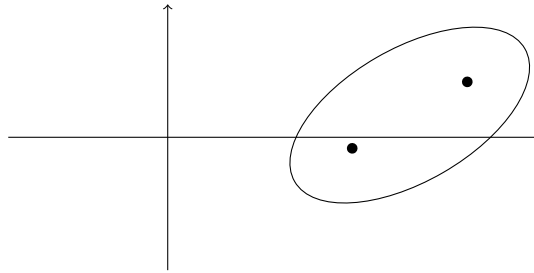


Abbildung 2: $\text{spek}(A)$ liegt in einer Ellipse, welche 0 nicht enthält

Mit Hilfe der Tschebyscheff-Polynome werden wir (asymptotisch optimale) Näherungslösungen für die MinMax-Aufgabe

$$(4.1) \quad \min_{p_m \in \Pi_m} \max_{\lambda \in E(f_1, f_2, \rho)} |p_m(\lambda)|$$

finden. Dafür benötigen wir zunächst die affine Transformation, welche unsere Standard-Ellipse $E_\rho = E(-1, 1, \rho)$ auf $E(f_1, f_2, \hat{\rho})$ mit $\hat{\rho} = \frac{|f_2 - f_1|}{2} \rho$ abbildet. Wir betrachten hierzu die affinen Grundabbildungen

$$D_\phi : z \mapsto e^{i\phi} z, \phi \in \mathbb{R} \quad (\text{Drehung})$$

$$T_a : z \mapsto z + a, a \in \mathbb{C} \quad (\text{Translation})$$

$$S_c : z \mapsto cz, c > 0 \quad (\text{Skalierung})$$

4.2 Lemma

Die Abbildung

$$\begin{aligned} \varphi &= T_{\frac{1}{2}(f_1+f_2)} \circ D_{\arg(f_2-f_1)} \circ S_{\frac{1}{2}|f_1-f_2|} \\ \varphi &: z \mapsto e^{i \arg(f_2-f_1)} \frac{1}{2} |f_2 - f_1| z + \underbrace{\frac{f_2 - f_1}{2}}_{\alpha} = \underbrace{\frac{f_1 + f_2}{2}}_{\beta} \end{aligned}$$

bildet die Ellipse $E(-1, 1, \rho)$ ab auf $E(f_1, f_2, \hat{\rho})$.

Damit können wir nun Tschebyscheff Polynome bezüglich allgemeiner Ellipsen $E(f_1, f_2, \hat{\rho})$ definieren.

4.3 Definition

Sei $E = E(f_1, f_2, \hat{\rho})$ eine Ellipse mit $0 \notin E$. Dann ist

$$T_m^E(z) = T_m\left(\frac{1}{\alpha}(z - \beta)\right)$$

mit $\alpha = \frac{1}{2}(f_2 - f_1)$ und $\beta = \frac{1}{2}(f_1 + f_2)$ das *Tschebyscheff-Polynom* vom Grad m bzgl. E .

Nach Satz 3.7 ist damit klar, dass

$$p_m(z) = \frac{T_m^E(z)}{T_m^E(0)}$$

eine approximative, asymptotisch optimale, Lösung für die MinMax-Aufgabe (4.1) ist.

4.4 Definition

Das zu p_m gehörige KUV heißt *Tschebyscheff-Verfahren*.

Wir werden nun dieses Verfahren konkret herleiten. Hierfür betrachten wir den Fall einer allgemeinen 3-Term-Rekursion

$$p_{m+1}(t) = \alpha_{m+1}(t + \beta_{m+1})p_m(t) + \gamma_{m+1}p_{m-1}(t), \quad m \geq 1.$$

Wegen $p_m(0) = 1$, $\forall m$, folgt $\gamma_m = 1 - \alpha_m\beta_m$ und insbesondere

$$p_0(t) = 1, \quad p_1(t) = \alpha_1\left(t + \frac{1}{\alpha_1}\right).$$

Es gilt also für die Residuen $r^m = b - Ax^m = p_m(A)r^0$

$$r^{m+1} = \alpha_{m+1}(A + \beta_{m+1}I)r^m + (1 - \alpha_{m+1}\beta_{m+1})r^{m-1}, \quad m \geq 1,$$

$$r^0 = r^0,$$

$$r^1 = \alpha_1\left(A + \frac{1}{\alpha_1}I\right)r^0 = \alpha_1Ar^0 + r^0.$$

Damit erhalten wir für die Iterierten unter Einsetzen von $r^m = b - Ax^m$ die Rekursion

$$x^0 = x^0,$$

$$x^1 = x^0 - \alpha_1r^0,$$

$$x^{m+1} = \alpha_{m+1}\beta_{m+1}x^m + (1 - \alpha_{m+1}\beta_{m+1})x^{m-1} - \alpha_{m+1}r^m.$$

Damit erhalten wir als allgemeine Struktur für ein KUV mit 3-Term-Rekursion für die Residuen den folgenden Algorithmus:

4.5 Algorithmus

wähle x^0 , setze $r^0 = b - Ax^0$
 $x^1 = x^0 - \alpha_1 r^0$
 $r^1 = \alpha_1 A r^0 + r^0$
for $m = 1, 2, \dots$ **do**
 $x^{m+1} = \alpha_{m+1} \beta_{m+1} x^m + (1 - \alpha_{m+1} \beta_{m+1}) x^{m-1} - \alpha_{m+1} r^m$
 $r^{m+1} = b - A x^{m+1}$
end for

(Alternativ könnte man r^{m+1} auch aus der 3-Term-Rekursion bestimmen)
 Jetzt geht es um die konkrete Bestimmung der Koeffizienten α_m, β_m beim Tschebyscheff-Verfahren.

Mit $E = E(f_1, f_2, \hat{\rho})$, $a = \frac{1}{2}(f_2 - f_1)$ und $b = \frac{1}{2}(f_2 + f_1)$ gilt

$$T_m^E(z) = T_m\left(\frac{1}{a}(z - b)\right).$$

Es gilt $T_0(z) = 1$, $T_1(z) = z$ und $T_{m+1}(z) = 2zT_m(z) - T_{m-1}(z)$. Also folgt

$$\begin{aligned}
 T_0^E(z) &= 1, \\
 T_1^E(z) &= \frac{1}{a}(z - b) \\
 \text{und} \quad T_{m+1}^E(z) &= T_{m+1}\left(\frac{1}{a}(z - b)\right) = \frac{2}{a}(z - b)T_m^E(z) - T_{m-1}^E(z).
 \end{aligned}$$

Die Residuenpolynome $p_m(z)$ für das Tschebyscheff-Verfahren sind

$$p_m(z) = \frac{T_m^E(z)}{T_m^E(0)}.$$

Wir benötigen also $c_m := T_m^E(0)$ und finden

$$(4.2) \quad \left\{ \begin{array}{l} c_0 = 1, \\ c_1 = -\frac{b}{a} \\ \text{und} \quad c_{m+1} = -\frac{2}{a}bc_m - c_{m-1}, \quad m \geq 1. \end{array} \right.$$

Damit erhalten wir

$$\begin{aligned}
 p_0(z) &= 1, \\
 p_1(z) &= -\frac{1}{b}(z - b) \\
 \text{und} \quad p_{m+1}(z) &= \frac{1}{c_{m+1}} \left(\frac{2}{a}(z - b)c_m p_m(z) - c_{m-1} p_{m-1}(z) \right), \quad m \geq 1.
 \end{aligned}$$

4.6 Bemerkung

Es gibt verschiedene andere Formulierungen, mit denen die Koeffizienten aufdatiert werden können. Die Werte c_m werden schnell sehr groß; deshalb ist es aus numerischen Gründen besser, die Berechnung von c_{m+1} zu vermeiden, und dafür die Quotienten $\delta_{m+1} = \frac{c_{m+1}}{c_m}$ aufzudatieren. Wir erhalten

$$\frac{c_{m+1}}{c_m} = -\frac{2}{a}b - \frac{c_{m-1}}{c_m},$$

und damit

$$\begin{aligned} \delta_1 &= c_1 \\ \text{und} \quad \delta_{m+1} &= -\frac{2}{a}b - \delta_m^{-1}, \quad m \geq 0. \end{aligned}$$

Wegen $\frac{c_{m-1}}{c_{m+1}} = 1 - \frac{2b}{a} \frac{c_m}{c_{m+1}}$ (dies gilt, da $p_m(0) = 1$ für alle m) ergibt sich

$$\begin{aligned} p_0(z) &= 1, \\ p_1(z) &= -\frac{1}{b}(z - b) \\ \text{und} \quad p_{m+1}(z) &= \frac{2}{a\delta_{m+1}}(z - b)p_m(z) - \left(1 - \frac{2b}{a\delta_{m+1}}\right)p_{m-1}(z). \end{aligned}$$

Die Koeffizienten der 3-Term-Rekursion sind also

$$\alpha_m = \frac{2}{a\delta_m}, \quad \beta_m = -b; \quad m \geq 2.$$

Damit erhalten wir den folgenden Algorithmus für das Tschebyscheff-Verfahren zur Lösung von $Ax = c$ (b ist schon vergeben ...):

4.7 Algorithmus

Require: $\text{spek}(A) \subseteq E(f_1, f_2, \rho) \not\equiv 0$
 setze $a = \frac{1}{2}(f_2 - f_1)$, $b := \frac{1}{2}(f_2 + f_1)$
 setze $\delta_1 = -b/a$
 wähle x^0
 setze $r^0 = c - Ax^0$
 $\alpha_1 = -2/b$
 $x^1 = x^0 - \alpha_1 r^0$
 $r^1 = \alpha_1 A r^0 + r^0$
for $m = 1, 2, \dots$ **do**
 $\delta_{m+1} = -\frac{2}{a}b - 1/\delta_m$
 $\alpha_{m+1} = 2/(a \cdot \delta_{m+1})$
 $x^{m+1} = (-\alpha_{m+1}b)x^m + (1 + \alpha_{m+1}b)x^{m-1} - \alpha_{m+1}r^m$
 $r^{m+1} = c - Ax^{m+1}$
end for

4.8 Bemerkung

Wegen $p_m(0) = 1$ muss auch gelten

$$1 = \frac{c_m}{c_{m+1}} \cdot \frac{2}{a}(-b) - \frac{c_{m-1}}{c_{m+1}},$$

was man mit (4.2) bestätigt.

4.9 Aufgabe

Implementiere das Tschebyscheff-Verfahren für das Modellproblem I für verschiedene Gittergrößen N .

Verwende dabei $\text{spek}(A) \subseteq [\xi, \Xi]$ mit

$$\xi = 4 - 4 \cos(\pi h), \quad h = \frac{1}{N+1},$$

und $\Xi = 4 + 4 \cos(\pi h).$

Die zugehörige Ellipse ist also

$$E(\xi, \Xi, 1)$$

- Plotte $\|r^m\|$ gegen m für verschiedene Werte von N .
- Teste auch 2Ξ statt Ξ (langsamere Konvergenz).
- Teste auch $\Xi - \xi$ statt Ξ (keine Konvergenz).
- Teste auch $\xi + 10 \cdot h^2$ statt ξ .
- Teste auch $\frac{\xi + \Xi}{2}$ statt ξ (langsamere Konvergenz).

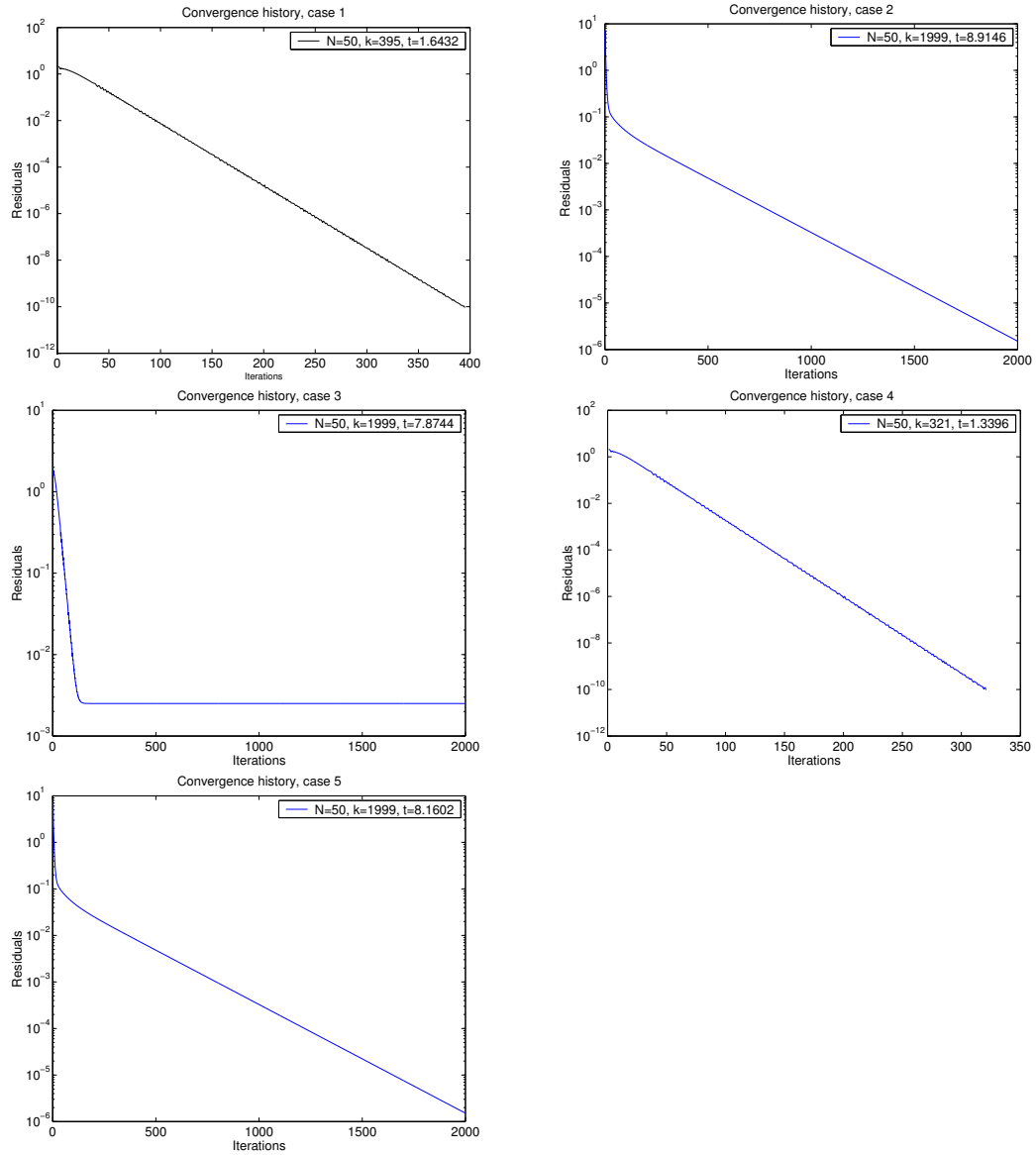


Abbildung 3: Numerische Resultate für das Modellproblem auf einem 50×50 -Gitter

5 Analyse von Tschebyscheff- und Richardson-Verfahren

5.1 Satz

Sei p_m das Polynom des Tschebyscheff-Verfahrens bzgl. $E(f_1, f_2, \rho) \not\equiv 0$ mit $\text{spek}(A) \subseteq E(f_1, f_2, \rho)$. Dann gilt

$$r^m = p_m(A)r^0$$

und damit für jede Norm $\|\cdot\|$

$$\|r^m\| \leq \|p_m(A)\| \cdot \|r^0\|.$$

1. Es sei A diagonalisierbar mit $A = SDS^{-1}$ (siehe 2.2) und $\|\cdot\|$ sei die Norm $\|\cdot\|_S$. Dann gilt

$$\|p_m(A)\| \leq \frac{\rho^m + \rho^{-m}}{|w_\gamma^m + w_\gamma^{-m}|},$$

wobei w_γ (betragsgrößte) Lösung von $\frac{1}{2}(w + w^{-1}) = -b/a$.

2. Ist in 1. sogar S orthogonal (z.B. weil A normal ist), so gilt sogar

$$\|p_m(A)\|_2 \leq \frac{\rho^m + \rho^{-m}}{|w_\gamma^m + w_\gamma^{-m}|}.$$

Beweis: Ist q_m die approximative Optimallösung für die Ellipse $E(-1, 1, \rho)$ mit $q_m(\gamma) = 1$, dann gilt nach Satz 3.7

$$\max_{t \in E(-1, 1, \bar{\rho})} |q_m(t)| = \frac{\bar{\rho}^m + \bar{\rho}^{-m}}{|w_\gamma^m + w_\gamma^{-m}|}.$$

Mit der affinen Transformation $z \mapsto az + b$ aus Lemma 4.2, welche $E(-1, 1, \rho)$ auf $E(f_1, f_2, \rho)$ abbildet, ergibt sich der Satz. \square

Besonders wichtig ist der Fall $\text{spek}(A) \subseteq [\xi, \Xi] \subseteq (0, +\infty)$ und A symmetrisch und positiv definit:

5.2 Korollar

Ist $\text{spek}(A) \subseteq [\xi, \Xi] \subseteq (0, +\infty)$ und A spd, dann gilt für die Residuen des Tschebyscheff-Verfahrens

$$\|r^m\|_2 \leq \|p_m(A)\|_2 \cdot \|r^0\|_2$$

mit

$$\|p_m(A)\|_2 \leq \frac{2c^m}{1 + c^{2m}} \quad \text{mit } c = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}, \quad \kappa = \frac{\Xi}{\xi}.$$

Beweis: Hier ist $f_1 = \xi$, $f_2 = \Xi$, $\rho = 1$, $a = \frac{1}{2}(\Xi - \xi)$, $b = \frac{1}{2}(\Xi + \xi)$ und $\gamma = -b/a$. Aus Satz 5.1 folgt so

$$\|p_m(A)\|_2 \leq \frac{2}{|w_\gamma^m + w_\gamma^{-m}|}$$

mit

$$\frac{1}{2}(w_\gamma + w_\gamma^{-1}) = -\frac{b}{a} = -\frac{\Xi + \xi}{\Xi - \xi} = -\frac{\kappa + 1}{\kappa - 1}.$$

Auflösen der quadratischen Gleichung ergibt

$$\begin{aligned} w_\gamma &= -\frac{\kappa + 1}{\kappa - 1} - \sqrt{\left(\frac{\kappa + 1}{\kappa - 1}\right)^2 - 1} \\ &= -\frac{\kappa + 1}{\kappa - 1} - \sqrt{\frac{4\kappa}{(\kappa - 1)^2}} \\ &= -\frac{\kappa + 1}{\kappa - 1} - \frac{2\sqrt{\kappa}}{\kappa - 1} \\ &= -\frac{(\sqrt{\kappa} + 1)^2}{(\sqrt{\kappa} - 1)(\sqrt{\kappa} + 1)} \\ &= -\frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1}. \end{aligned}$$

□

Zum Abschluss (und Vergleich) analysieren wir noch das Richardson-Verfahren 2. Ordnung. Dazu sei A spd mit $\text{spek}(A) \subseteq [\xi, \Xi]$.

5.3 Aufgabe

Für das relaxierte Richardson-Verfahren 1. Ordnung

$$x^{m+1} = x^m + \alpha r^m$$

wird für die zugehörigen Polynome $p_m(z) = (1 - \alpha z)^m$ die Größe

$$\max_{z \in [\xi, \Xi]} |p_m(z)|$$

minimal, wenn $\alpha = 2/(\xi + \Xi)$.

5.4 Bemerkung

Das relaxierte Richardson Verfahren 1. Ordnung $Ax = b$ ist äquivalent zum gewöhnlichen Richardson-Verfahren 1. Ordnung für

$$(\alpha A)x = \alpha b,$$

also

$$x^{m+1} = x^m + (\alpha b - \alpha A x^m).$$

Wir gehen ab jetzt davon aus, dass diese „Skalierung“ des LGS (mit $\alpha = \frac{2}{\xi + \Xi}$) bereits erfolgte, d.h. es gilt ab jetzt

$$\text{spek}(A) \subseteq [1 - \gamma, 1 + \gamma] \text{ mit } \gamma = \frac{\Xi - \xi}{\Xi + \xi} \in [0, 1).$$

Für das System

$$Ax = b$$

betrachten wir jetzt das Richardson-Verfahren 2. Ordnung

$$\begin{aligned} x^{m+1} &= \omega((I - A)x^m + b) + (1 - \omega)x^{m-1} \\ &= \omega(x^m + r^m) + (1 - \omega)x^{m-1} \end{aligned}$$

mit

$$\begin{aligned} r^{m+1} &= \omega r^m - \omega A r^m + (1 - \omega)r^{m-1}, \\ p_{m+1} &= \omega(1 - t)p_m(t) + (1 - \omega)p_{m-1}(t). \end{aligned}$$

Daraus folgt

$$\begin{pmatrix} p_{m+1}(t) \\ p_m(t) \end{pmatrix} = \underbrace{\begin{pmatrix} \omega(1-t) & 1-\omega \\ 1 & 0 \end{pmatrix}}_{B_\omega(t)} \begin{pmatrix} p_m(t) \\ p_{m-1}(t) \end{pmatrix}.$$

Wir wollen nun das „optimale“ ω ermitteln, d.h. wir suchen die Lösung ω von

$$(5.1) \quad \min_{\omega \in \mathbb{R}} \max_{t \in [1-\gamma, 1+\gamma]} \rho(B_\omega(t)).$$

Zur besseren Darstellung machen wir den Übergang $t \rightarrow 1 - t$, also

$$\begin{aligned} 1 - t &\longrightarrow t \in [-\gamma, \gamma], \\ B_\omega(t) &= \begin{pmatrix} \omega(1-t) & 1-\omega \\ 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \omega t & 1-\omega \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

Wir suchen damit die Lösung ω der Minimierungsaufgabe

$$\min_{\omega \in \mathbb{R}} \max_{t \in [-\gamma, \gamma]} \rho(B_\omega(t)).$$

Nun bestimmen wir die Eigenwerte λ von $B_\omega(t)$. Es gilt nämlich

$$\begin{aligned} (\lambda - \omega t)\lambda - (1 - \omega) &= 0 \\ \iff \lambda^2 - (\omega t)\lambda - (1 - \omega) &= 0, \end{aligned}$$

d.h. wir erhalten

$$\lambda_{1,2} = \frac{\omega t}{2} \pm \sqrt{\frac{\omega^2 t^2}{4} + (1 - \omega)}.$$

Wir machen zunächst eine Fallunterscheidung zur Berechnung von $\rho(B_\omega(t))$.

Fall I: $\frac{\omega^2 t^2}{4} + (1 - \omega) < 0$. Dann gilt

$$\begin{aligned} |\lambda_{1,2}|^2 &= \frac{(\omega t)^2}{4} + \frac{(\omega t)^2}{4} + (1 - \omega) \\ &= \frac{(\omega t)^2}{2} + (1 - \omega). \end{aligned}$$

Also gilt

$$(5.2) \quad \rho(B_\omega(t)) = \left(\frac{(\omega t)^2}{2} + (1 - \omega) \right)^{1/2}.$$

Fall II: $\frac{\omega^2 t^2}{4} + (1 - \omega) \geq 0$.

(a) $\omega t \geq 0$: In diesem Fall haben wir

$$\rho(B_\omega(t)) = \frac{\omega t}{2} + \sqrt{\frac{\omega^2 t^2}{4} + (1 - \omega)}.$$

(b) $\omega t \leq 0$: Dann gilt

$$\rho(B_\omega(t)) = \left| \frac{\omega t}{2} - \sqrt{\frac{\omega^2 t^2}{4} + (1 - \omega)} \right|.$$

In beiden Fällen (a) und (b) erhalten wir aber

$$(5.3) \quad \rho(B_\omega(t)) = \frac{|\omega t|}{2} + \sqrt{\frac{\omega^2 t^2}{4} + (1 - \omega)}.$$

Sei nun ω fest. Wir bestimmen $\max_{t \in [-\gamma, \gamma]} \rho(B_\omega(t)) =: \beta_\omega$.

1. *Fall*: $\omega \leq 1$. Es kommt nur Fall II vor. Aus (5.3) erhalten wir, da $\rho(B_\omega(t))$ eine gerade Funktion ist und monoton wachsend für $t \in [0, \gamma]$

$$\beta_\omega = \frac{|\omega|\gamma}{2} + \sqrt{\frac{\omega^2\gamma^2}{4} + (1 - \omega)}.$$

2. *Fall*: $\omega > 1$. Jetzt kommen Fall I und Fall II beide vor. Es gilt

$$\frac{\omega^2 t^2}{4} + (1 - \omega) \geq 0 \iff |t| \geq \frac{2\sqrt{\omega - 1}}{\omega}.$$

Dann ist

$$\rho(B_\omega(t)) = \frac{\omega|t|}{2} + \sqrt{\frac{\omega^2 t^2}{4} + (1 - \omega)}.$$

Im Fall $|t| \leq \frac{2\sqrt{\omega - 1}}{\omega}$ ist

$$\rho(B_\omega(t)) = \left(\frac{(\omega t)^2}{2} + (1 - \omega) \right)^{1/2},$$

also ist

$$\beta_\omega = \max \left\{ \max_{|t| \leq \frac{2\sqrt{\omega-1}}{\omega}} \left(\frac{(\omega t)^2}{2} + (1 - \omega) \right)^{1/2}, \max_{\frac{2\sqrt{\omega-1}}{\omega} \leq |t| \leq \gamma} \left(\frac{\omega|t|}{2} + \sqrt{\frac{(\omega t)^2}{4} + (1 - \omega)} \right) \right\}.$$

Hierin kommt der erste Term nur vor, wenn $\gamma \geq \frac{2\sqrt{\omega-1}}{\omega}$. Aus Monotoniegründen folgt

$$\begin{aligned} \max_{|t| \leq \frac{2\sqrt{\omega-1}}{\omega}} \left(\frac{(\omega t)^2}{2} + (1 - \omega) \right)^{1/2} &= \left(\frac{(\omega \frac{2\sqrt{\omega-1}}{\omega})^2}{2} + (1 - \omega) \right)^{1/2} \\ &= \sqrt{\omega - 1}, \\ \max_{\frac{2\sqrt{\omega-1}}{\omega} \leq |t| \leq \gamma} \left(\frac{\omega|t|}{2} + \sqrt{\frac{(\omega t)^2}{4} + (1 - \omega)} \right) &= \frac{\omega\gamma}{2} + \sqrt{\frac{(\omega\gamma)^2}{4} + (1 - \omega)}. \end{aligned}$$

Für $\gamma \geq \frac{2\sqrt{\omega-1}}{\omega}$ gilt

$$\frac{\omega\gamma}{2} + \sqrt{\frac{(\omega\gamma)^2}{4} + (1 - \omega)} \geq \frac{\omega \frac{2\sqrt{\omega-1}}{\omega}}{2} + \sqrt{\frac{(\omega \frac{2\sqrt{\omega-1}}{\omega})^2}{4} + (1 - \omega)}$$

$$= \sqrt{\omega - 1}.$$

Damit gilt

$$\beta_\omega = \begin{cases} \sqrt{\omega - 1} & \text{falls } \gamma < \frac{2\sqrt{\omega-1}}{\omega}, \\ \frac{\omega\gamma}{2} + \sqrt{\frac{(\omega\gamma)^2}{4} + (1 - \omega)} & \text{falls } \gamma \geq \frac{2\sqrt{\omega-1}}{\omega}. \end{cases}$$

Unter Beachtung von

$$\begin{aligned} & \gamma < \frac{2\sqrt{\omega-1}}{\omega} \\ \iff & \omega^2\gamma^2 < 4(\omega-1) \\ \iff & \frac{2-2\sqrt{1-\gamma^2}}{\gamma^2} < \omega < \frac{2+2\sqrt{1-\gamma^2}}{\gamma^2} \\ \iff & \frac{2}{1+\sqrt{1-\gamma^2}} < \omega < \frac{2}{1-\sqrt{1-\gamma^2}} \end{aligned}$$

erhalten wir durch Zusammenfassen von „1. Fall“ und „2. Fall“

$$\beta_\omega = \begin{cases} \sqrt{\omega - 1} & \text{falls } \frac{2}{1+\sqrt{1-\gamma^2}} < \omega < \frac{2}{1-\sqrt{1-\gamma^2}}, \\ \frac{\omega\gamma}{2} + \sqrt{\frac{(\omega\gamma)^2}{4} + (1 - \omega)} & \text{sonst.} \end{cases}$$

Für $\omega \geq 2/(1 - \sqrt{1 - \gamma^2})$ (≥ 2) ist $\frac{\omega\gamma}{2} \geq \sqrt{\omega - 1} \geq 1$, also $\beta_\omega \geq 1$. Für $\omega < 2/(1 + \sqrt{1 - \gamma^2})$ ist

$$\beta_\omega = \frac{\omega\gamma}{2} + \sqrt{\frac{(\omega\gamma)^2}{4} + (1 - \omega)}$$

monoton fallend. Die Größe β_ω wird also minimal für $\omega = \omega_0 = \frac{2}{1+\sqrt{1-\gamma^2}}$ und hat dann den Wert $\sqrt{\omega_0 - 1}$.

Zusammenfassend erhalten wir also:

5.5 Satz

Der optimale Parameter ω für das Richardson-Verfahren 2. Ordnung, welcher die Minimierungsaufgabe (5.1) löst, ist

$$\omega_0 = \frac{2}{1 + \sqrt{1 - \gamma^2}}.$$

Jetzt kann man für $w = w_0$ die Größe

$$\max_{t \in [1-\gamma, 1+\gamma]} |p_m(t)|$$

diskutieren unter Beachtung von $p_0 = 1$, $p_1(t) = 1 - t$.

5.6 Aufgabe

Bestimme für alle $t \in [1 - \gamma, 1 + \gamma]$ die Lösung der Differenzengleichung 2. Ordnung

$$p_{m+1}(t) = w_0 \cdot (1 - t)p_m(t) + (1 - w_0)p_{m-1}(t)$$

unter den Anfangsbedingungen $p_0 = 0$ und $p_1(t) = 1 - t$ und zeige damit

1. $\max_{t \in [1-\gamma, 1+\gamma]} |p_m(t)| = (\sqrt{w_0 - 1})^m (1 + m\sqrt{1 - \gamma^2}) =: \rho_m$, (Dies ist sehr aufwendig!²)
2. ρ_m fällt monoton in m ,
3. $\rho_m > c_m$, $c_m = \frac{2c^m}{1+c^{2m}}$ aus dem Tschebyscheff Verfahren mit $c = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$,
 $\kappa = \frac{1+\gamma}{1-\gamma}$,
4. es gilt $\lim_{n \rightarrow \infty} \frac{\rho_m}{c_m} = 1$.

5.7 Bemerkung

Aus $r^m = p_m(A)r^0$ folgt $e^m = p_m(A)e^0$ mit $e^m = A^{-1}b - x^m$ und damit

$$\|e^m\| \leq \|p_m(A)\| \|e^0\|.$$

Tschebyscheff- und Richardson-Verfahren machen eine Schranke für $\|p_m(A)\|$ klein. Dies wirkt sich demnach in den Abschätzungen nicht nur für das Residuum, sondern auch für den *Fehler* aus.

²Man liest also besser bei den Meistern nach: G. Golub, R. Varga: Chebyshev semi-iterative methods, successive overrelaxation iterative methods and second-order Richardson iterative methods, Part I and II, Numer. Math. **3**, 147-168 (1961).

6 Basen für Krylov-Unterräume

Gegeben: $A \in \mathbb{C}^{n \times n}$ nicht notwendigerweise hermitesch, A regulär,

$$K_m = \langle r, Ar, \dots, A^{m-1}r \rangle.$$

Gesucht: v^1, \dots, v^m mit „schönen Eigenschaften“ und

$$K_m = \langle v^1, \dots, v^m \rangle.$$

Erinnerung: Gram-Schmidt-Orthogonalisierung

Gegeben: $y_1, \dots, y_m \in \mathbb{C}^{n \times n}$.

6.1 Algorithmus (Gram-Schmidt-Orthogonalisierung, klassisch)

```

 $v^i = y^i, i = 1, \dots, m$ 
for  $i = 1, \dots, m$  do
  for  $j = 1, \dots, i - 1$  do
     $v^i = v^i - \langle y^i, v^j \rangle v^j$ 
  end for
   $v^i = \frac{1}{\|v^i\|} \cdot v^i$ 
end for

```

Es ist bekannt, dass diese Variante aber numerisch nicht stabil ist. Günstiger ist es, wenn man bei der Berechnung der Innenprodukte statt y_i den aktuellen Vektor v^i verwendet. Dies ist möglich, da zu jedem Zeitpunkt der j -Schleife der Vektor $y^i - v^i \in \langle v^1, \dots, v^{j-1} \rangle$ senkrecht auf v^j steht.

6.2 Algorithmus (modifizierte Gram-Schmidt-Orthogonalisierung)

```

 $v^i = y^i, i = 1, \dots, m$ 
for  $i = 1, \dots, m$  do
  for  $j = 1, \dots, i - 1$  do
     $v^i = v^i - \langle v^i, v^j \rangle v^j$ 
  end for
   $v^i = \frac{1}{\|v^i\|} \cdot v^i$ 
end for

```

Unser Ziel ist es, sukzessive eine ONB von $K_m(A, r)$ aufzubauen.

Idee:

$\langle v^1, \dots, v^m \rangle$ sei eine ONB in $K_m(A, r)$ mit $v^m \notin K_{m-1}(A, r)$. Dann gilt insbesondere $Av^m \in K_{m+1}(A, r) \setminus K_m(A, r)$. Orthogonalisiere also Av^m gegen v^1, \dots, v^m .

Durch Anwendung des modifizierten Gram-Schmidt-Verfahrens resultiert das *Arnoldi-Verfahren*.

6.3 Algorithmus (Arnoldi-Verfahren)

```

 $v^1 = r, h_{1,0} = \|r\|, v^1 = \frac{1}{h_{1,0}} r$ 
for  $i = 2, 3, \dots$  do
   $v^i = Av^{i-1}$ 
  for  $j = 1, \dots, i-1$  do
     $h_{j,i-1} = \langle v^i, v^j \rangle$ 
     $v^i = v^i - h_{j,i-1} v^j$ 
  end for
   $h_{i,i-1} = \|v^i\|$ 
   $v^i = \frac{1}{h_{i,i-1}} \cdot v^i$ 
end for

```

Wir führen folgende Kurznotation ein

$$(6.1) \quad AV_m = V_{m+1}H_{m+1,m}$$

mit

$$V_m = [v^1, \dots, v^m] \in \mathbb{C}^{n \times m}$$

$$H_{m+1,m} = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,m} \\ h_{2,1} & h_{2,2} & \ddots & h_{2,m} \\ & \ddots & \ddots & \vdots \\ & & \ddots & h_{m,m} \\ & & & h_{m+1,m} \end{pmatrix} \in \mathbb{C}^{(m+1) \times m}.$$

Im Falle $A = A^H$ vereinfacht sich der Arnoldi-Prozess, da die meisten Einträge in $H_{m+1,m}$ verschwinden. Es bezeichne dazu H_m die Matrix

$$H_m = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,m} \\ h_{2,1} & h_{2,2} & \ddots & h_{2,m} \\ & \ddots & \ddots & \vdots \\ & & h_{m-1,m} & h_{m,m} \end{pmatrix} \in \mathbb{C}^{m \times m}.$$

6.4 Lemma

Falls A hermitesch ist, ist H_m eine reelle, symmetrische Tridiagonalmatrix. Es gilt also $h_{i+1,i} = h_{i,i+1}$ für alle i und $h_{i,j} = 0$ für $i > j + 1$.

Beweis: Aus (6.1) folgt

$$\begin{aligned} V_m^H A V_m &= V_m^H V_{m+1} H_{m+1,m} \\ &= \begin{pmatrix} & 0 \\ I_m & \vdots \\ & 0 \end{pmatrix} H_{m+1,m} \\ &= H_m. \end{aligned}$$

Mit $V_m^H A V_m$ ist auch H_m hermitesch, besitzt also Tridiagonalgestalt. Außerdem sind dann die Diagonalelemente reell (und positiv). Die Elemente $h_{i,i-1}$ der ersten Nebendiagonalen ergeben sich nach Algorithmus 6.3 als $\|v^i\|$, sind also auch reell. \square

Damit erhalten wir aus dem Arnoldi-Verfahren das Lanczos-Verfahren zur Bestimmung einer ONB von $K_m(A, r)$ im Falle $A = A^H$.

6.5 Algorithmus (Lanczos-Verfahren)

```

 $v^1 = \frac{r}{\|r\|}, v^0 = 0, h_{1,0} = 0$ 
for  $i = 2, 3, \dots$  do
   $v^i = A v^{i-1}$ 
   $v^i = v^i - h_{i-1,i-2} v^{i-2}$ 
   $h_{i-1,i-1} = \langle v^i, v^{i-1} \rangle$ 
   $v^i = v^i - h_{i-1,i-1} v^{i-1}$ 
   $h_{i,i-1} = \|v^i\|$ 
   $v^i = \frac{1}{h_{i,i-1}} \cdot v^i$ 
end for

```

Eine alternative Bezeichnung für die Koeffizienten im Lanczos-Verfahren ist

$$h_{i-1,i-2} = \beta_{i-1}, \quad h_{i-1,i-1} = \alpha_{i-1}.$$

Damit lässt sich der Algorithmus wie folgt formulieren:

6.6 Algorithmus (Lanczos-Verfahren)

```

 $v^1 = \frac{r}{\|r\|}, v^0 = 0, \beta_1 = 0$ 
for  $i = 2, 3, \dots$  do
   $q^i = Av^{i-1} - \beta_{i-1}v^{i-2}$ 
   $\alpha_{i-1} = \langle q^i, v^{i-1} \rangle$ 
   $q^i = q^i - \alpha_{i-1}v^{i-1}$ 
   $\beta_i = \|q^i\|$ 
   $v^i = \frac{q^i}{\beta_i}$ 
end for

```

Hierin gilt:

$$(6.2) \quad \beta_i v^i = Av^{i-1} - \alpha_{i-1}v^{i-1} - \beta_{i-1}v^{i-2}.$$

6.7 Definition

Zwei Folgen v^1, \dots, v^m und $w^1, \dots, w^m \in \mathbb{C}^n$ heißen *bi-orthogonal*, falls gilt

$$\langle v^i, w^j \rangle = \delta_{i,j}.$$

6.8 Bemerkung

v^i, w^j sind nur bis auf reziproke Vielfache eindeutig bestimmt.

Im Falle $A \neq A^H$ konstruieren wir nun eine bi-orthogonale Basis v^1, \dots, v^m von $K_m(A, r)$ und w^1, \dots, w^m von $K_m(A^H, \tilde{r})$.

Man berechnet v^1, w^1 mit dem *unsymmetrischen* Lanczos-Verfahren.

Situation:

v^1, \dots, v^m sei eine Basis von $K_m(A, r)$
 w^1, \dots, w^m sei eine Basis von $K_m(A^H, \tilde{r})$
 die beiden Basen sind bi-orthogonal

und

$$v^m \notin K_{m-1}(A, r)$$

$$w^m \notin K_{m-1}(A^H, \tilde{r}).$$

Ansatz:

$$q^{m+1} = Av^m,$$

$$p^{m+1} = A^H w^m.$$

Wir setzen an:

$$(6.3) \quad \tilde{v}^{m+1} = q^{m+1} - \sum_{i=1}^m \alpha_i v^i,$$

$$(6.4) \quad \tilde{w}^{m+1} = p^{m+1} - \sum_{i=1}^m \beta_i w^i.$$

Bedingung für Bi-Orthogonalität:

$$\begin{aligned} \langle \tilde{v}^{m+1}, w^j \rangle &= \langle q^{m+1}, w^j \rangle - \alpha_j \stackrel{!}{=} 0, \\ \langle v^j, \tilde{w}^{m+1} \rangle &= \langle v^j, p^{m+1} \rangle - \bar{\beta}_j \stackrel{!}{=} 0. \end{aligned}$$

Damit erhalten wir

$$\begin{aligned} \alpha_j &= \langle q^{m+1}, w^j \rangle, \\ \beta_j &= \langle p^{m+1}, v^j \rangle. \end{aligned}$$

Für $j \leq m-2$ gilt allerdings weiter

$$\begin{aligned} \alpha_j &= \langle q^{m+1}, w^j \rangle = \langle Av^m, w^j \rangle \\ &= \langle v^m, A^H w^j \rangle \\ &= 0, \end{aligned}$$

da $A^H w^j \in K_{m-1}(A^H, \tilde{r})$. Analoges gilt für die β_j , d.h. es gilt

$$\alpha_j = \beta_j = 0 \text{ für } j = 1, \dots, m-2.$$

Zur Normierung sei

$$\begin{aligned} \delta_{m+1} &= \left(\langle \tilde{v}^{m+1}, \tilde{w}^{m+1} \rangle \right)^{\frac{1}{2}} \\ v^{m+1} &= \frac{1}{\delta_{m+1}} \tilde{v}^{m+1} \\ w^{m+1} &= \frac{1}{\bar{\delta}_{m+1}} \tilde{w}^{m+1} \end{aligned}$$

Schließlich weisen wir noch nach, dass die Koeffizienten α_{m-1} und β_{m-1} vom vorherigen Schritt schon bekannt sind (dafür verwenden wir kurzfristig einen zusätzlichen oberen Index):

$$\begin{aligned} \alpha_{m-1} &= \langle Av^m, w^{m-1} \rangle \\ &= \langle v^m, A^H w^{m-1} \rangle \end{aligned}$$

$$\begin{aligned}
&= \langle v^m, \tilde{w}^m + \beta_{m-1}^{(m-1)} w^{m-1} + \beta_{m-2}^{(m-1)} w^{m-2} \rangle \\
&= \langle v^m, \tilde{w}^m \rangle = \delta_m
\end{aligned}$$

$$\begin{aligned}
\beta_{m-1} &= \langle A^H w^m, v^{m-1} \rangle \\
&= \langle w^m, A v^{m-1} \rangle \\
&= \dots \\
&= \langle w^m, \tilde{v}^m \rangle = \bar{\delta}_m
\end{aligned}$$

Insgesamt erhalten wir damit folgendes Verfahren:

6.9 Algorithmus (unsymmetrisches Lanczos-Verfahren)

```

 $v^0 = w^0 = 0$ 
 $\delta_1 = (\langle r, \tilde{r} \rangle)^{\frac{1}{2}}$ 
 $v^1 = \frac{1}{\delta_1} \cdot r$ 
 $w^1 = \frac{1}{\bar{\delta}_1} \cdot \tilde{r}$ 
for  $m = 1, 2, \dots$  do
   $\alpha_m = \langle A v^m, w^m \rangle$ 
   $\tilde{v}^{m+1} = A v^m - \alpha_m v^m - \delta_m v^{m-1}$ 
   $\tilde{w}^{m+1} = A^H w^m - \bar{\alpha}_m w^m - \bar{\delta}_m w^{m-1}$ 
   $\delta_{m+1} = (\langle \tilde{v}^{m+1}, \tilde{w}^{m+1} \rangle)^{\frac{1}{2}}$ 
   $v^{m+1} = \frac{1}{\delta_{m+1}} \cdot \tilde{v}^{m+1}$ 
   $w^{m+1} = \frac{1}{\bar{\delta}_{m+1}} \cdot \tilde{w}^{m+1}$ 
end for

```

6.10 Bemerkung

Im unsymmetrischen Lanczos kann Folgendes passieren

$$\delta_m = 0,$$

obwohl $\tilde{v}^m \neq 0$ und $\tilde{w}^m \neq 0$.

Das nennt man *serious breakdown*, da das Verfahren abbricht, ohne dass eine Basis für den maximalen Krylov-Unterraum $K_{m^*}(A, r)$ oder $K_{m^{**}}(A^H, \tilde{r})$ berechnet wurde.

7 Krylov-Unterraum-Verfahren mit minimalem Residuum

Gegeben:

$A \in \mathbb{C}^{n \times n}$ mit $A \neq A^H$, A regulär.

Ziel:

Löse $Ax = b$ mit Startvektor x^0 und $r^0 = b - Ax^0$.

Idee:

Bestimme ONB von $K_m(A, r^0)$ mit Arnoldi-Prozess und errechne $x^m \in x^0 + K_m(A, r^0)$, so dass

$$\|b - Ax^m\|_2 = \min_{x \in x^0 + K_m(A, r^0)} \|b - Ax\|_2.$$

Ansatz:

$x = x^0 + V_m z^m$ mit $z^m = (z_1, \dots, z_m)^T \in \mathbb{C}^m$ sowie $V_m = [v^1, \dots, v^m]$.

Damit erhalten wir

$$b - Ax = r^0 - AV_m z^m = r^0 - V_{m+1} H_{m+1,m} z^m.$$

Mit $r^0 = h_{1,0} v^1$ und $h_{1,0} = \|r^0\|$ gilt nun

$$b - Ax = V_{m+1} \left[\begin{pmatrix} h_{1,0} \\ 0 \\ \vdots \\ 0 \end{pmatrix} - H_{m+1,m} z^m \right]$$

und somit

$$\|b - Ax\|_2 = \|h_{1,0} e^1 - H_{m+1,m} z^m\|_2.$$

Hierbei ist e^1 der erste Einheitsvektor in \mathbb{C}^{m+1} , $H_{m+1,m} \in \mathbb{C}^{(m+1) \times m}$ aber $z \in \mathbb{C}^m$.

Exkurs:

Zu lösen ist also ein überbestimmtes Least Squares Problem $Bx = c$ mit $B \in \mathbb{C}^{k \times \ell}$, $k > \ell$. Gesucht ist ein x , welches $\|Bx - c\|_2^2$ minimiert. Dies findet man unter Verwendung der QR-Faktorisierung.

Ansatz:

Faktorisiere $B = QR$ mittels QR-Zerlegung. Dann folgt

$$\begin{aligned}\|Bx - c\|_2^2 &= \|Q^H(Bx - c)\|_2^2 \\ &= \|Rx - Q^H c\|_2^2.\end{aligned}$$

Das Minimum wird erreicht für

$$x = R_1^{-1}((Q^H c)(1 : \ell)),$$

wobei

$$R = \begin{pmatrix} & R_1 & \\ 0 & \dots & \\ & & 0 \end{pmatrix}, \quad R_1 \in \mathbb{C}^{\ell \times \ell} \text{ obere Dreiecksmatrix.}$$

In unserem Falle ist $H_{m+1,m}$ eine obere Hessenberg Matrix. Ihre QR-Faktorisierung kann mittels Jacobi Rotationen bestimmt werden. Es kann sogar Q in der QR-Faktorisierung von $H_{m+1,m}$ auf die für $H_{m+2,m+1}$ aufdatiert werden.

Notation:

$$Q_{m+1} H_{m+1,m} = R_{m+1,m} \quad \text{QR-Faktorisierung}$$

mit

$$Q_{m+1} = J_m^{(m+1)} \dots J_1^{(m+1)},$$

wobei $J_j^{(m+1)}$ eine Jacobi Rotation in \mathbb{C}^{m+1} auf den Zeilen j und $j+1$ ist, welche das ‘untere Nebendiagonalelement’ $h_{j+1,j}$ eliminiert,

$$J_j^{(m+1)} = \begin{pmatrix} I & & \\ & \bar{c}_j & \bar{s}_j \\ & -s_j & c_j \\ & & & I \end{pmatrix}.$$

Um anzugeben, wie man $J_m^{(m+1)}$ bestimmt, beobachten wir zunächst

$$J_{m-1}^{(m+1)} \dots J_1^{(m+1)} H_{m+1,m} = J_{m-1}^{(m+1)} \dots J_1^{(m+1)} \left(\begin{array}{c|c} H_{m,m-1} & \begin{matrix} h_{1,m} \\ \vdots \\ h_{m,m} \end{matrix} \\ \hline 0 & h_{m+1,m} \end{array} \right)$$

$$= \begin{pmatrix} R_{m,m-1} & \eta^{(m)} \\ 0 & h_{m+1,m} \end{pmatrix}$$

mit

$$\eta^{(m)} = (r_{1,m}, \dots, r_{m-1,m}, \eta_{m,m})^T,$$

wobei

$$\begin{aligned} r_{1,m} &= \bar{c}_1 h_{1,m} + \bar{s}_1 h_{2,m}, \\ \eta_{1,m} &= -s_1 h_{1,m} + c_1 h_{2,m}, \\ r_{2,m} &= \bar{c}_2 \eta_{2,m} + \bar{s}_2 h_{3,m}, \\ \eta_{2,m} &= -s_2 \eta_{1,m} + c_2 h_{3,m}, \\ &\vdots \\ \eta_{m,m} &= -s_{m-1} \eta_{m-1,m} + c_{m-1} h_{m,m}. \end{aligned}$$

Damit erhalten wir für die Parameter c_m, s_m der Jacobi Rotation $J_m^{(m+1)}$, welche $(0, \dots, 0, \eta_{m,m}, h_{m+1,m})^T$ abbildet auf $(0, \dots, 0, r_{m,m}, 0)^T$ die Werte

$$(7.1) \quad c_m = \eta_{m,m} / \sqrt{|\eta_{m,m}|^2 + |h_{m+1,m}|^2}, \quad s_m = h_{m+1,m} / \sqrt{|\eta_{m,m}|^2 + |h_{m+1,m}|^2}$$

und

$$r_{m,m} = \bar{c}_m \eta_{m,m} + \bar{s}_m h_{m+1,m} = \sqrt{|\eta_{m,m}|^2 + |h_{m+1,m}|^2}.$$

Die Lösung des LS-Problems

$$\min \|h_{1,0}e^1 - H_{m+1,m}z^m\|_2$$

lautet damit

$$z^m = h_{1,0} R_m^{-1}((Q_{m+1}e^1)(1:m)), \quad \text{wobei } R_{m+1,m} = \begin{pmatrix} R_m \\ 0 \dots 0 \end{pmatrix},$$

und der Wert des Minimums ist

$$h_{1,0}(Q_{m+1}e^1)_{m+1}.$$

Dieser Wert kann ohne Kenntnis von z^m wie folgt bestimmt werden: Setze

$$Q_m e^1 = h_{1,0} \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{m-1} \\ \tilde{\epsilon}_m \end{pmatrix}.$$

Dann ist

$$Q_{m+1}e^1 = J_m^{(m+1)} \begin{pmatrix} Q_m & 0 \\ 0 & 1 \end{pmatrix} e^1 = J_m^{(m+1)} \cdot \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{m-1} \\ \tilde{\epsilon}_m \\ 0 \end{pmatrix} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \\ \tilde{\epsilon}_{m+1} \end{pmatrix},$$

mit

$$\begin{aligned} \epsilon_m &= \tilde{\epsilon}_m \bar{c}_m, \\ \tilde{\epsilon}_{m+1} &= \tilde{\epsilon}_m \bar{s}_m. \end{aligned}$$

Insbesondere gilt

$$\tilde{\epsilon}_{m+1} = \prod_{j=1}^m \bar{s}_j,$$

so dass wir für den Wert des Minimums den Ausdruck

$$h_{1,0}\tilde{\epsilon}_{m+1} = \|r^0\| \cdot \tilde{\epsilon}_{m+1}$$

erhalten.

Wir haben jetzt alle Zutaten für das Verfahren GMRES (Generalized Minimal Residual):

7.1 Algorithmus (GMRES (Saad, Schultz, 1986))

wähle x^0 , setze $r^0 = b - Ax^0$, $\rho_0 = h_{1,0} = \|r^0\|_2$
for $m = 1, 2, \dots$, bis ρ_m klein genug **do**
 bestimme nächsten Vektor v^m und Koeffizienten h_1^m, \dots, h_{m+1}^m aus Arnoldi-Prozess
 bestimme s_m, c_m nach (??)
 datiere ϵ_m und $\tilde{\epsilon}_{m+1}$ auf
 setze $\rho^m = \rho_{m-1} \cdot \bar{s}_m$ { $\rho_m = \|r^m\|$ }
 if ρ_m klein genug **then**
 löse $R_m z^m = (\epsilon_1, \dots, \epsilon_m)^T$
 setze $x^m = x^0 + V_m z^m$
 end if
end for

7.2 Bemerkung

Speicher- und Rechenaufwand (ohne MVMs) verhält sich wie $\mathcal{O}(mn + m^2)$
 \curvearrowright Verfahren wird schnell impraktikabel.

Es bezeichne x^m die GMRES-Iterierte nach jedem Schritt, also die Lösung von

$$\min_{x \in x^0 + K_m(A, r^0)} \|b - Ax\|_2.$$

Dann gilt natürlich

$$\|r^0\| \geq \|r^1\| \geq \dots$$

Das folgende Beispiel zeigt, dass hier lange Gleichheit herrschen kann.

7.3 Beispiel

Sei

$$A = \begin{pmatrix} 0 & & & 1 \\ 1 & \ddots & & 0 \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix}, \quad x^0 = 0, \quad b = e^1 = r^0.$$

Dann ist $K_m(A, r^0) = \text{span}\{e^1, \dots, e^m\}$, $e^i = i$ -ter Einheitsvektor. Lösung von

$$\min_{x \in K_m(A, r^0)} \|e^1 - Ax\|_2,$$

ist für $m = 1, 2, \dots, n-1$ wegen $Ax \in \text{span}\{e^2, \dots, e^{m+1}\}$ stets $x^m = 0$. Es gilt also für die GMRES-Iterierten

$$x^0 = x^1 = \dots = x^{n-1} = 0, \quad x^n = e^n = A^{-1}b$$

und

$$r^0 = r^1 = \dots = r^{n-1} = e^1,$$

also

$$\|r^0\| = \dots = \|r^{n-1}\| = 1.$$

Beachte: $\text{spek}(A) = \{e^{\frac{2\pi i k}{n}}, k = 0, \dots, n-1\}$. Ein so um die 0 verteiltes Spektrum ist für KUV prinzipiell schwierig, weil die Verfahrenspolynome ja $p(0) = 1$ erfüllen müssen.

Idee für „praktikable“ GMRES-Variante: „Restarted“GMRES=GMRES(k) mit restart Wert k .

7.4 Algorithmus (GMRES(k))

```

wähle  $x^0$ 
for  $m = 1, 2, \dots$  do
    führe  $k$  Schritte von GMRES aus mit Startvektor  $x^{m-1}$ 
    Ergebnis ist  $x^m$ 
end for

```

Es ist nun

$$x^m \in x^{m-1} + K_k(A, r^{m-1}), r^{m-1} = b - Ax^{m-1},$$

aber es gilt i.A. *nicht*

$$x^m \in x^0 + K_{km}(A, r^0).$$