

Proseminar Numerische Verfahren

DIE TSCHEBYSCHJEFF-BESCHLEUNIGUNG

Nikolaus Schäfer

Dozent: Prof. Dr. Rannacher

18. März 2018

Inhaltsverzeichnis

1. Einleitung	3
2. Die Tschebyscheff-Beschleunigung	4
2.1. Anwendung auf die Tschebyscheff-Beschleunigung	9
2.2. Fazit	11

1. Einleitung

In der Mathematik und speziell in dem Teilgebiet der numerischen Mathematik müssen sehr häufig Probleme der Art

$$Ax = b \tag{1}$$

gelöst werden, mit einer reellen quadratischen Matrix $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ und einem Vektor $b = (b_j)_{j=1}^n \in \mathbb{R}^n$.

Wenn wir Probleme mit $n \gg 1000$ betrachten, dann wird neben der Komplexität des Algorithmus auch der Speicherplatz zunehmend wichtig. Deswegen werden lineare Gleichungssysteme mit solch hoher Dimension iterativ gelöst. Die direkten Verfahren zur Lösung linearer Gleichungssysteme (Gauß-Elimination, Cholesky-Zerlegung etc.) gehen von beliebigen vollbesetzten Matrizen aus. Diese Eigenschaft steht jedoch in zahlreichen, wichtigen Fällen Problemen $Ax = b$ gegenüber bei denen, neben der Höhe der Dimension n , die Matrix A eine spezielle Struktur besitzt (z.B. Bandstruktur, Symmetrie). Diese Struktur sollte von den Lösungsalgorithmen ausgenutzt werden. Die einfachste Lösung für das Problem (1) erhalten wir indem wir die Gleichung in eine Fixpunktgleichung umformen:

$$Ax = b \Leftrightarrow Cx = Cx - Ax + b \Leftrightarrow x = (I - C^{-1}A)x + C^{-1}b, \tag{2}$$

wobei $C \in GL(n)$ eine geeignete reguläre Matrix ist. Dann, startend mit einem Anfangswert x_0 , benutzt man eine simple Fixpunktiteration:

$$x^t = \underbrace{(I - C^{-1}A)}_{=:B} x^{t-1} + \underbrace{C^{-1}b}_{=:c}, \quad t = 1, 2, \dots \tag{3}$$

Zur Lösung dieser Gleichung haben wir schon zahlreiche Verfahren kennengelernt (Jacobi-Verfahren, Gauß-Seidel-Verfahren, SOR-Verfahren). Diese Verfahren interessieren sich jedoch nur für das gerade aktuelle Element der Folgen der Iterierten. Im t -ten Schritt eines Iterationsverfahrens also für x^{t-1} . Die restlichen Iterierten x^0, \dots, x^{t-2} werden nicht berücksichtigt. Es gibt aber auch Verfahren, bei denen alle Iterierten x^0, \dots, x^{t-1} in die Berechnung einer verbesserten Approximation y^t der Lösung einfließen können. Eines dieser Verfahren ist die **Tschebyscheff-Beschleunigung**. Es wird sich herausstellen, dass dieses semiiterative Verfahren, mit gewissen Annahmen über die Matrizen A und B , ähnlich effizient wie die bisher untersuchten iterativen Verfahren implementiert werden kann, aber zum Teil deutlich bessere Konvergenzeigenschaften besitzt.

2. Die Tschebyscheff-Beschleunigung

Für diese Art der Konvergenzbeschleunigung der Fixpunktiteration

$$x^t = Bx^{t-1} + c, \quad t = 1, 2, \dots$$

zum Problem (1) betrachten wir eine reelle, symmetrische Koeffizientenmatrix A und eine reelle diagonalisierbare Iterationsmatrix B ¹. Angenommen die Fixpunktiteration konvergiert gegen die Lösung $x \in \mathbb{R}^n$ des Gleichungssystems

$$Ax = b \Leftrightarrow x = Bx + c, \quad (4)$$

also $\text{spr}(B) < 1$ ², dann bleibt die Frage, wie man aus der Folge der Vektoren x^t eine zweite (sekundäre) Folge y^t konstruieren kann, die möglichst schneller gegen die Lösung x konvergiert, als x^t selbst. Die Idee der Tschebyscheff-Beschleunigung ist es eine Linearkombination

$$y^t = \sum_{s=0}^t \gamma_s^t x^s \quad t \geq 1 \quad (5)$$

aus den Werten x^0, \dots, x^t zu konstruieren. Bei geeigneter Wahl der Koeffizienten γ_s^t soll die Ersatzfolge $\{y^0, y^1, \dots\}$ schneller konvergieren als die Ausgangsfolge $\{x^0, x^1, \dots\}$.

Falls $x^0 = \dots = x^t = x$ bereits die Lösung ist, so fordern wir, dass auch $y^t = x$, woraus sofort folgt, dass

$$\sum_{s=0}^t \gamma_s^t = 1, \quad (6)$$

Für den Fehler $y^t - x$ gilt insbesondere, dass

$$\begin{aligned} y^t - x &= \sum_{s=0}^t \gamma_s^t (x^s - x) = \sum_{s=0}^t \gamma_s^t (Bx^{s-1} + c - (Bx + c)) \\ &= \sum_{s=0}^t \gamma_s^t B(x^{s-1} - x) = \dots = \sum_{s=0}^t \gamma_s^t B^s (x^0 - x) = p_t(B)(x^0 - x) \end{aligned} \quad (7)$$

mit einem Polynom $p_t \in P_t$ von Grad t gegeben durch

$$p_t(z) = \sum_{s=0}^t \gamma_s^t z^s, \quad p_t(1) = 1 \quad (8)$$

Da diese Iteration von einer Folge von Iterationsmatrizen $p_t(B)$, $t \geq 1$, bestimmt wird, müssen

¹Die Anforderungen werden später noch größer. Zunächst muss aber B diagonalisierbar sein, damit der spektrale Abbildungssatz (Begründung für Gleichung (9)) für reelle Koeffizientenmatrizen gilt

²Konvergenzkriterien für das Jacobi-Verfahren und das Gauß-Seidel-Verfahren, die $\text{spr}(B) < 1$ implizieren, werden in Bemerkung 2.1 nochmals aufgegriffen

wir, um das Konvergenzverhalten zu untersuchen, den Spektralradius von $p_t(B)$ anschauen.

Es gilt:

$$\lambda(p_t(B)) = p_t(\lambda(B)) \quad (9)$$

Nach dem Fixpunktsatz (siehe Rannacher [1]³ 3.1) konvergiert das Verfahren umso schneller, je kleiner der Spektralradius von $p_t(B)$ ist. Wir betrachten also das folgende Optimierungsproblem:

$$\min_{p_t \in P_t, p_t(1)=1} \max_{\lambda \in \sigma(B)} |p_t(\lambda)| \quad (10)$$

Oft ist es aber so, dass man die Eigenwerte von B nicht kennt. Meistens kann man höchstens eine Einschränkung, $\text{spr}(B) \leq 1 - \delta$ mit einem kleinen $\delta > 0$, vornehmen. Deswegen müssen wir das Optimierungsproblem abschwächen auf:

$$\text{spr}(p_t(B)) \leq \min_{p_t \in P_t, p_t(1)=1} \max_{|x| \leq 1-\delta} |p_t(x)| \quad (11)$$

Dieses ist eindeutig lösbar, wenn $\sigma(B) \subseteq \mathbb{R}$.

Um dies zu gewährleisten nehmen wir im folgenden an, neben der Symmetrie der reellen Koeffizientenmatrix $A = L + D + L^T$, dass die **Iterationmatrix B ähnlich zu einer symmetrischen Matrix** ist, da dann gilt: B ist diagonalisierbar und $\sigma(B) \subseteq \mathbb{R}$

Bemerkung 2.1.

Im Allgemeinen ist die Iterationsmatrix B selbstverständlich keine symmetrische Matrix. Sie muss noch nicht mal ähnlich zu einer symmetrischen Matrix sein (z.B. bei der Gauß-Seidel-Methode mit $H_1 = -(D + L)^{-1}L^T$).

Die Iterationsmatrix der Jacobi-Methode mit $J = -D^{-1}(L + L^T)$ ist zwar symmetrisch, aber damit das Jacobi-Verfahren auch konvergiert muss die Systemmatrix A strikt diagonaldominant sein.

Eine weitere interessante Anmerkung ist, dass in vielen Problemen der Praxis symmetrisch positiv definite Systemmatrizen betrachtet werden. Dies reicht aber für die Konvergenz des Jacobi-Verfahrens nicht aus. Als Abhilfe kann man nach Möglichkeit⁴ dann das symmetrische Gauß-Seidel-Verfahren benutzen, da das Gauß-Seidel Verfahren für eine symmetrisch positiv definite Systemmatrix A immer konvergiert.

Die Lösung des Optimierungsproblems (11) basiert auf den bereits bekannten Tschebyscheff-Polynomen.

³[1] R. Rannacher: *Numerical Linear Algebra*,

Lecture Notes, Heidelberg University, <http://numerik.uni-hd.de/~lehre/notes/>

⁴sofern die Iterationsmatrix ähnlich zu einer symmetrischen Matrix ist

Die Tschebyscheff-Polynome:

Die (reellen) Tschebyscheff-Polynome lassen sich wie folgt darstellen (siehe z.B. Stoer & Burlisch [2] ⁵ oder Rannacher [3] ⁶):

$$T_t(x) = \begin{cases} (-1)^t \cosh(t \operatorname{arccosh}(-x)), & x \leq -1 \\ \cos(t \arccos(x)), & -1 \leq x \leq 1 \\ \cosh(t \operatorname{arccosh}(x)), & x \geq 1 \end{cases} \quad (12)$$

Das t -te Tschebyscheff-Polynom ist durch die folgende Rekursion definiert:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{t+1}(x) = 2xT_t(x) - T_{t-1}(x), \quad t \geq 1 \quad (13)$$

In einigen Fällen, wie im Beweis des folgenden Theorems, kann die globale Repräsentation der Tschebyscheff-Polynome nützlich sein:

$$T_t(x) = \frac{1}{2}[(x + \sqrt{x^2 - 1})^t + (x - \sqrt{x^2 - 1})^t], \quad x \in \mathbb{R} \quad (14)$$

Mit Hilfe dieser Polynome lässt sich nun unser Minimierungsproblem lösen:

Theorem 2.2. (Minmax-Problem)

Sei $[a, b] \subseteq \mathbb{R}$ ein nicht-leeres Intervall and sei $c \in \mathbb{R} \setminus [a, b]$.

Dann ist das Minimum

$$\min_{p \in P_t, p(c)=1} \max_{a \leq x \leq b} |p(x)| \quad (15)$$

eindeutig durch die folgenden Polynome gegeben:

$$p(x) := C_t(x) = \frac{T_t(1 + 2\frac{x-b}{b-a})}{T_t(1 + 2\frac{c-b}{b-a})}, \quad x \in [a, b] \quad (16)$$

Für $a < b < c$ gilt dann:

$$\min_{p \in P_t, p(c)=1} \max_{a \leq x \leq b} |p(x)| = \frac{1}{T_t(1 + 2\frac{c-b}{b-a})} = \frac{2\gamma^t}{1 + \gamma^{2t}} \leq 2\gamma^t, \quad (17)$$

$$\text{wobei, } \gamma = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}, \quad \kappa := \frac{c-a}{c-b}$$

⁵[2] J. Stoer and R. Burlisch: *Numerische Mathematik 1/2*, Springer, 2007 (10. Edition)

⁶[3] R. Rannacher: *Numerische Mathematik 0 (Einführung in die Numerische Mathematik)*, Lecture Notes, Heidelberg, University, <http://numerik.uni-hd.de/~lehre/notes/>

Beweis.

1. Da sich durch eine affine Transformation die Maximumsnorm nicht ändert, versuchen wir zunächst eine affine Transformation zu finden, damit wir uns auf das Intervall $[-1, 1]$ beschränken können.

Betrachte die affine Abbildung

$$\begin{aligned}\varphi : [a, b] &\xrightarrow{\cong} [-1, 1] \\ x &\mapsto y = \varphi(x) = 2\frac{x-a}{b-a} - 1 = \frac{2x-a-b}{b-a}\end{aligned}$$

mit der Umkehrabbildung

$$\begin{aligned}\alpha : [-1, 1] &\xrightarrow{\cong} [a, b], \\ y &\mapsto f(y) = \frac{1-y}{2}a + \frac{1+y}{2}b.\end{aligned}$$

Ist nun $p_t \in P_t$ mit $\deg(p_t) = t$ und führendem Koeffizienten 1 die Lösung des Minimax-problems

$$\min_{p \in P_t, p(c)=1} \max_{-1 \leq x \leq 1} |p(x)|$$

so ist $\hat{p}_t(x) := p_t(\alpha(y))$ die Lösung des ursprünglichen Problems **(15)** mit führendem Koeffizienten $\frac{(b-a)^t}{2^t}$.

Deshalb können wir uns also auf den Fall $[a, b] = [-1, 1]$, $c \in \mathbb{R} \setminus [-1, 1]$ einschränken.

2. Nun gilt:

$$C_t(x) = \frac{T_t(1 + 2\frac{x-1}{1-(-1)})}{T_t(1 + 2\frac{c-1}{1-(-1)})} = \frac{T_t(x)}{T_t(c)} = \tilde{C}T_t(x), \text{ wobei } \tilde{C} = T_t(c)^{-1} \quad (18)$$

Das Tschebyscheff-Polynom $T_t(x) = \cos(t \arccos(x))$ nimmt die Werte ± 1 für ein $k \in \{0, \dots, t\}$ genau dann an, wenn

$$t \cdot \arccos(x) = k \cdot \pi \Leftrightarrow \arccos(x) = \frac{k\pi}{t} \Leftrightarrow x_k = \cos(k\pi/t)$$

Darüber hinaus alterniert $T_t(x)$ zwischen 1 und -1 , was bedeutet, dass $T_t(x_k)$ und $T_t(x_{k+1})$ unterschiedliche Vorzeichen haben. Insbesondere folgt daraus, dass auch $C_t(x_k)$ und $C_t(x_{k+1})$ unterschiedliche Vorzeichen haben.

Da $C_t(x)$ also $t+1$ mal die Werte $\pm \tilde{C}$ annimmt und stetig ist gilt insbesondere, dass $C_t(x)$ t Nullstellen hat.

Weiterhin folgt aus $\max_{x \in [-1,1]} |T_t(x)| = 1$, dass $\max_{x \in [-1,1]} |C_t(x)| = |\tilde{C}|$.

3. Wir wollen nun durch einen Widerspruch zeigen, dass $|\tilde{C}|$ tatsächlich das eindeutige Minimum für ein Polynom t -ten Grades ist.

Wir nehmen also an es gäbe ein Polynom $q \in P_t$, sodass

$$\max_{x \in [-1,1]} |q(x)| < \max_{x \in [-1,1]} |C_t(x)| = |\tilde{C}| \text{ und } q(c) = 1.$$

Definiere $r := C_t - q$.

Da $\text{sgn}(r(x_k)) = \text{sgn}(C_t(x_k))$ für $k \in \{0, \dots, t\}$, wechselt r im Intervall $[-1, 1]$ das Vorzeichen t -mal. Somit hat r mindestens t Nullstellen in $[-1, 1]$.

Darüber hinaus ist $r(c) = C_t(c) - q(c) = 0$. Also muss $r \in P_t$ sogar mindestens $t + 1$ Nullstellen haben, was bedeutet, dass r das Nullpolynom sein muss, da $\deg(r) \leq t$. Dies stellt aber einen Widerspruch dar, da dann $q = C_t$ gilt und somit auch $\max_{x \in [-1,1]} |q(x)| = |\tilde{C}|$.

Nach Definition von $T_t(x)$ gilt $|T_t(x)| \leq 1 \forall x \in [-1, 1]$. Somit gilt:

$$\max_{x \in [a,b]} |C_t(x)| = \frac{1}{T_t(1 + 2\frac{c-b}{b-a})} \quad (19)$$

Nun bleibt zu zeigen, dass $\frac{1}{T_t(1 + 2\frac{c-b}{b-a})} = \frac{2\gamma^t}{1 + \gamma^{2t}} \leq 2\gamma^t$

Wir verwenden die Darstellung

$$T_t(x) = \frac{1}{2}[(x + \sqrt{x^2 - 1})^t + (x - \sqrt{x^2 - 1})^t], \quad x \in \mathbb{R}$$

Da

$$\begin{aligned} 1 + 2\frac{c-b}{b-a} &= \frac{b-a}{b-a} + 2\frac{c-b}{b-a} = \frac{2c-b-a}{b-a} = \frac{c-b+c-a}{b-a} = \frac{c-b}{b-a} \frac{c-a}{b-a} \\ &= \left(\frac{c-a}{c-b} + 1\right) \cdot \frac{c-b}{b-a} = \frac{\frac{c-a}{c-b} + 1}{\frac{b-a}{c-b}} = \frac{\frac{c-a}{c-b} + 1}{\frac{-(-b)-a+c-c}{c-b}} = \frac{\frac{c-a}{c-b} + 1}{\frac{c-a}{c-b} - \frac{c-b}{c-b}} = \frac{\frac{c-a}{c-b} + 1}{\frac{c-a}{c-b} - 1} = \frac{\kappa + 1}{\kappa - 1} \end{aligned}$$

und

$$\frac{\kappa + 1}{\kappa - 1} + \sqrt{\left(\frac{\kappa + 1}{\kappa - 1}\right)^2 - 1} = \frac{\kappa + 1}{\kappa - 1} + \frac{2\sqrt{\kappa}}{\kappa - 1} = \frac{(\sqrt{\kappa} + 1)^2}{\kappa - 1} = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$$

gilt:

$$\begin{aligned}
\frac{1}{T_t(1+2\frac{c-b}{b-a})} &= \frac{1}{T_t(\frac{\kappa+1}{\kappa-1})} = \frac{1}{\frac{1}{2}[(\frac{\sqrt{\kappa+1}}{\sqrt{\kappa-1}})^t + (\frac{\sqrt{\kappa-1}}{\sqrt{\kappa+1}})^t]} \\
&= \frac{2(\frac{\sqrt{\kappa-1}}{\sqrt{\kappa+1}})^t}{[(\frac{\sqrt{\kappa+1}}{\sqrt{\kappa-1}})^t + (\frac{\sqrt{\kappa-1}}{\sqrt{\kappa+1}})^t] \cdot (\frac{\sqrt{\kappa-1}}{\sqrt{\kappa+1}})^t} = \frac{2\gamma^t}{1 + \underbrace{\gamma^{2t}}_{>0}} \leq 2\gamma^t, \text{ da } 1 + \gamma^{2t} > 1
\end{aligned} \tag{20}$$

□

2.1. Anwendung auf die Tschebyscheff-Beschleunigung

Weiterhin nehmen wir an, dass B ähnlich zu einer symmetrischen Matrix ist, und dass $\text{spr}(B) < 1$. Zusätzlich nehmen wir nun an, dass wir den Parameter $\text{spr}(B) := \rho \in (-1, 1)$ kennen, sodass $\sigma(B) \subset [-\rho, \rho]$ ⁷. Man beachte nun, dass $T_t(1) = \cos(t \cdot \arccos(1)) = \cos(0) = 1$. Mit $a = -\rho$, $b = \rho$ und $c = 1$ können wir also die Polynome $p_t = C_t$ aus (16) benutzen um die y^t aus (5) zu definieren.

Allerdings müsste man durch diese Art der Auswertung der y^t alle Iterierten x^t speichern, was für große Probleme oft nicht möglich ist. Als Abhilfe kann man sich nun die Rekursionsformel (14) der Tschebyscheff-Polynome zur Hand nehmen, die die jeweiligen Iterierten y^t überträgt.

Die Polynome $p = C_t$ aus (16) genügen der Rekursionsformel durch:

$$p_0(x) = 1, \quad p_1(x) = \frac{T_1(1 + 2\frac{x-\rho}{\rho-(-\rho)})}{T_1(1 + 2\frac{1-\rho}{\rho-(-\rho)})} = \frac{T_1(1 + \frac{x}{\rho} - \frac{\rho}{\rho})}{T_1(1 + \frac{1}{\rho} - \frac{\rho}{\rho})} = \frac{T_1(x/\rho)}{T_1(1/\rho)} = \frac{x/\rho}{1/\rho} = x \tag{21}$$

und

$$\begin{aligned}
p_{t+1} &= \frac{T_{t+1}(x/\rho)}{T_{t+1}(1/\rho)} = \frac{\frac{2x}{\rho}T_t(x/\rho) - T_{t-1}(x/\rho)}{\mu_{t+1}} = \frac{\frac{2x}{\rho}\frac{\mu_t T_t(x/\rho)}{\mu_t} - \frac{\mu_{t-1} T_{t-1}(x/\rho)}{\mu_{t-1}}}{\mu_{t+1}} \\
&= \frac{\frac{2x}{\rho}\mu_t p_t(x) - \mu_{t-1} p_{t-1}(x)}{\mu_{t+1}}, \quad t \geq 1, \quad \mu_t := T_t(1/\rho)
\end{aligned} \tag{22}$$

Außerdem gelten die folgenden wichtigen Relationen:

$$\mu_0 = 1, \quad \mu_1 = 1/\rho, \quad \mu_{t+1} = \frac{2}{\rho}\mu_t - \mu_{t-1} \tag{23}$$

Diese folgen direkt aus der Definition der Rekursionsformel von $T_t(x)$.

Mit diesen Vorbereitungen können wir nun die Tschebyscheff-Beschleunigung implementieren.

⁷ ρ lässt sich z.B. durch die Potenzmethode recht schnell und effektiv schätzen

Mit $x := \lim_{t \rightarrow \infty} x^t$, $e^0 := x^0 - x$ erhalten wir für den Fehler $y^t - x = \tilde{e}^t = p_t(B)e^0$:

$$\begin{aligned} y^{t+1} &= x + \tilde{e}^{t+1} = x + p_{t+1}(B)e^0 = x + 2\frac{\mu_t}{\rho\mu_{t+1}}Bp_t(B)e^0 - \frac{\mu_{t-1}}{\mu_{t+1}}p_{t-1}(B)e^0 \\ &= x + 2\frac{\mu_t}{\rho\mu_{t+1}}B\tilde{e}^t - \frac{\mu_{t-1}}{\mu_{t+1}}\tilde{e}^{t-1} = x + 2\frac{\mu_t}{\rho\mu_{t+1}}B(y^t - x) - \frac{\mu_{t-1}}{\mu_{t+1}}(y^{t-1} - x) \\ &= 2\frac{\mu_t}{\rho\mu_{t+1}}By^t - \frac{\mu_{t-1}}{\mu_{t+1}}y^{t-1} + \frac{1}{\mu_{t+1}}(\mu_{t+1} - \frac{2}{\rho}\mu_t B + \mu_{t-1})x \end{aligned}$$

Durch die Verwendung von (23) gilt:

$$\begin{aligned} \frac{1}{\mu_{t+1}}(\mu_{t+1} - \frac{2}{\rho}\mu_t B + \mu_{t-1})x &= \frac{1}{\mu_{t+1}}(\frac{2}{\rho}\mu_t - \mu_{t-1} - \frac{2}{\rho}\mu_t B + \mu_{t-1})x \\ &= \frac{2\mu_t}{\rho\mu_{t+1}}(1 - B)x = \frac{2\mu_t}{\rho\mu_{t+1}}(x - Bx) \end{aligned}$$

Aus der Fixpunktrelation $x = Bx + c$ folgt außerdem $c = x - Bx$ und somit können wir x aus der Rekursion entfernen und erhalten:

$$y^{t+1} = 2\frac{\mu_t}{\rho\mu_{t+1}}By^t - \frac{\mu_{t-1}}{\mu_{t+1}}y^{t-1} + 2\frac{\mu_t}{\rho\mu_{t+1}}c, \quad y^0 = x^0, \quad y^1 = x^1 = Bx^0 + c \quad (24)$$

Nun sehen wir, dass die Tschebyscheff-Beschleunigung für $x^t = Bx^{t-1} + c$ daraus besteht die Rekursionsformeln (23) und (24) auszuwerten. Die Kosten dieser Auswertung ist ähnlich zu der von $x^t = Bx^{t-1} + c$ bei der die kostenaufwändigste Rechnung die Matrixmultiplikation By^t ist.

Um den Beschleunigungseffekt zu quantifizieren, schreiben wir die sekundäre Iteration in der Form:

$$y^t - x = \sum_{s=0}^t \gamma_s^t (x^s - x) = p_t(B)(x^0 - x)$$

wobei γ_s^t die Koeffizienten des Polynoms p_t sind. Es gilt:

$$p_t(x) = C_t(x) = \frac{T_t(x/\rho)}{T_t(1/\rho)}$$

Durch (17) aus Theorem 2.2 folgt nun:

$$\text{spr}(p_t(B)) = \max_{\lambda \in \sigma(B)} |p_t(\lambda)| = \frac{2\gamma^t}{1 + \gamma^{2t}} \leq 2\gamma^t, \gamma := \underbrace{\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}}_{< 1, \text{ da } \sqrt{\kappa} > 1}, \kappa := \frac{1 + \rho}{1 - \rho}$$

Für das asymptotische Konvergenzverhalten der ersten Iteration gilt:

$$\limsup_{t \rightarrow \infty} \left(\frac{\|x^t - x\|}{\|x^0 - x\|} \right)^{\frac{1}{t}} = \text{spr}(B) \leq \rho = 1 - \delta \quad (25)$$

und für die sekundäre Iteration mit den optimierten y^t gilt vergleichsweise

$$\begin{aligned} \limsup_{t \rightarrow \infty} \left(\frac{\|y^t - x\|}{\|x^0 - x\|} \right)^{\frac{1}{t}} &\leq \gamma = \frac{\sqrt{\frac{1+\rho}{1-\rho}} - 1}{\sqrt{\frac{1+\rho}{1-\rho}} + 1} = \frac{\sqrt{\frac{1+(1-\delta)}{1-(1-\delta)}} - 1}{\sqrt{\frac{1+(1-\delta)}{1-(1-\delta)}} + 1} = \frac{\frac{\sqrt{2-\delta}}{\sqrt{\delta}} - \frac{\sqrt{\delta}}{\sqrt{\delta}}}{\frac{\sqrt{2-\delta}}{\sqrt{\delta}} + \frac{\sqrt{\delta}}{\sqrt{\delta}}} \\ &= \frac{\sqrt{2-\delta} - \sqrt{\delta}}{\sqrt{2-\delta} + \sqrt{\delta}} = \underbrace{\frac{\sqrt{2-\delta}}{\sqrt{2-\delta} + \sqrt{\delta}}}_{\leq 1} - \underbrace{\frac{1}{\sqrt{2-\delta} + \sqrt{\delta}}}_{=: c' < 1} \sqrt{\delta} \leq 1 - c' \sqrt{\delta} \end{aligned} \quad (26)$$

was uns für $0 < \delta \ll 1$ eine erhebliche Verbesserung bei der Konvergenzgeschwindigkeit gibt.

2.2. Fazit

Liegt also ein konvergentes Iterationsverfahren vor mit einer symmetrischen Systemmatrix A und einer Iterationsmatrix B , die ähnlich zu einer symmetrischen Matrix ist, so kann die Tschebyscheff-Beschleunigung verwendet werden um eine erhebliche Verbesserung der Konvergenzgeschwindigkeit zu erzielen. Der Nachteil dieses Verfahrens liegt darin, dass wir den Parameter $\rho = \text{spr}(B)$ wissen bzw. schätzen müssen. Obwohl das SOR-Verfahren⁸, die Tschebyscheff-Beschleunigung und das CG-Verfahren nahezu gleich aufwändig sind, ist das CG-Verfahren viel praktischer, da bei diesem Verfahren kein Parameter geschätzt werden muss.

⁸Das SOR-Verfahren beruht auf einem Parameter ω der basierend auf $\text{spr}(B_J)$ (= Spektralradius der Iterationsmatrix der Jacobi-Methode) geeignet gewählt werden kann, um die Konvergenz zu beschleunigen