

Lanczos-Verfahren bei exponentiell
abfallenden Eigenwerten

Diplomarbeit

Anne Kielmann

Betreuer: Prof. Dr. Dr. Wolfgang Hackbusch

Christian-Albrechts-Universität zu Kiel

Institut für Informatik

Dezember 2008

Inhaltsverzeichnis

1	Einleitung	2
1.1	Motivation	2
1.2	Inhalt und Aufbau	4
2	Grundlagen	6
2.1	Eigenwertprobleme	6
2.2	Invariante Unterräume	10
3	Lanczos-Verfahren	13
3.1	Rayleigh-Ritz-Verfahren	13
3.2	Krylovräume	18
3.3	Symmetrisches Lanczos-Verfahren	24
4	Konvergenzbetrachtungen	28
4.1	Allgemeine Fehleranalyse für das Lanczosverfahren	28
4.1.1	Fehlerabschätzung der Eigenwerte mit Hilfe der Ritz-Werte	28
4.1.2	Fehlerabschätzung der Eigenwerte und Eigenvektoren mit Hilfe der Ritz-Vektoren	37
4.2	Analyse für exponentiell abfallende Eigenwerte	46
4.2.1	Konvergenzbetrachtung mit Hilfe der Ritz-Werte	46
4.2.2	Konvergenzbetrachtung mit Hilfe der Ritz-Vektoren	62
4.2.3	Konvergenzbetrachtung für abweichende Eigenwerte	64
5	Anwendung des Lanczos-Verfahrens	67
5.1	Testmatrizen und Implementierung	67
5.2	Praktische Probleme des Verfahrens	71
5.2.1	Orthogonalitätsverlust und doppelte Eigenwerte	71
5.2.2	Problemanalyse	74
5.2.3	Modifiziertes Lanczos-Verfahren	78
5.3	Numerische Ergebnisse	80
A	Grafiken Kapitel 4 und 5	91
A.1	Grafiken zur Theorie	91
A.2	Grafiken zum modifizierten Lanczos-Verfahren	95
A.3	Vergleich zwischen Theorie und Praxis	98

Inhaltsverzeichnis

Abbildungsverzeichnis	104
Literaturverzeichnis	106

Notation

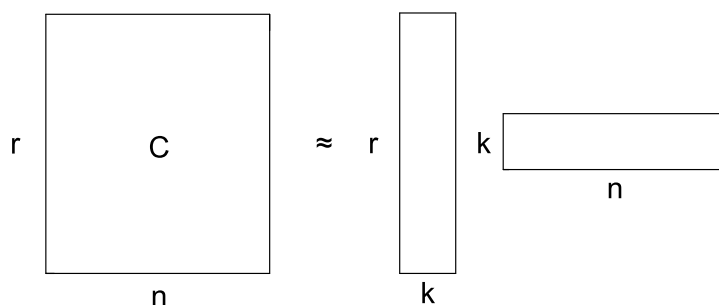
A, B, C	Matrizen
$I_{m \times m}$	$m \times m$ Einheitsmatrix
D, Σ	Diagonalmatrizen
U, V	orthogonale Matrizen
Q	Matrix mit orthonormalen Spalten
A^{-1}	Inverse der Matrix A
A^T	Transponierte der Matrix A
a, \dots, z	Vektoren
α, β, γ	Skalare
$\langle \cdot, \cdot \rangle$	euklidisches Skalarprodukt
$\ \cdot \ $	euklidische Norm
\mathcal{P}_n	Raum der Polynome mit Grad $\leq n$
S^m	m -dimensionaler Unterraum des \mathbb{R}^n
V	euklidischer Vektorraum
V^\perp	Senkrechttraum zu V

1 Einleitung

1.1 Motivation

Durch Diskretisierung von Integralgleichungen entstehen häufig hochdimensionale und voll besetzte Matrizen. Da dies einen hohen Speicherplatzbedarf und größeren Rechenaufwand bedeutet, ist man an einer Approximation durch Matrizen mit kleinerem Speicherverbrauch interessiert. Gelingt es zum Beispiel, eine gegebene Matrix $C \in \mathbb{R}^{r \times n}$ als Produkt zweier Matrizen aus $\mathbb{R}^{r \times k}$ bzw. $\mathbb{R}^{k \times n}$ sinnvoll darzustellen, so hat man für $k < r, n$ anstatt einem Speicherplatzbedarf von rn nur noch $k(r+n)$. Gilt also $k \ll r, n$, so wird sehr viel weniger Speicher benötigt.

Abbildung 1.1: Speicherplatzbedarf



Die Frage ist, wie man eine solche approximative Darstellung der Matrix C erhält, so dass der Informationsverlust gering bleibt und sich der Approximationsfehler steuern lässt.

In Hinblick auf die in dieser Arbeit untersuchten Matrizen mit exponentiell abfallenden Eigenwerten wird zur Beantwortung dieser Frage die Singulärwertzerlegung von C betrachtet. Gelte also

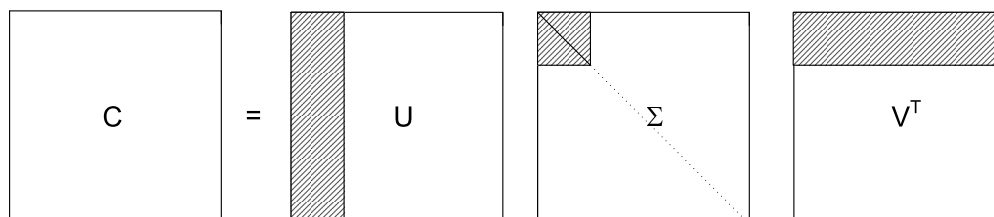
$$C = U\Sigma V^T \in \mathbb{R}^{r \times n}$$

wobei U und V orthonormale Matrizen sind und Σ eine Diagonalmatrix, welche die Eigen- bzw. Singulärwerte $\sigma_1 \geq \dots \geq \sigma_{\min\{r,n\}} \geq 0$ der Matrix C enthält.

Unter der Voraussetzung schnell abfallender, zumindest aber nach oben durch eine abfallende Folge beschränkter Singulärwerte liegt nun Idee nahe, die Singulärwerte ab einer bestimmten (genügend kleinen) Größe zu vernachlässigen. Ist beispielsweise zu $k < r, n$ der $k+1$ -te Wert σ_{k+1} kleiner als die gewünschte Genauigkeit, so werden nur noch $\sigma_1, \dots, \sigma_k$ betrachtet, und der Rest durch Nullen ersetzt. Da sich das geänderte Σ_k nun

also nur noch bis auf die ersten k Diagonalelemente von der Nullmatrix unterscheidet, sind auch bei den beiden Matrizen U und V^T nur noch die ersten k Spalten bzw. die ersten k Zeilen von Bedeutung.

Abbildung 1.2: Singulärwertzerlegung



Die so im Rang reduzierten Matrizen beschreiben eine kompakte Darstellung der Matrix C , das heißt, es müssen von der Diagonalmatrix Σ nur noch die ersten k Einträge gespeichert werden, und von den Matrizen U und V^T nur die ersten k Spalten beziehungsweise Zeilen. Diese Matrizen mit kleinerer Dimension seien mit C_k , Σ_k , U_k und V_k bezeichnet. Fasst man nun Σ_k und U_k als Produkt in einer Matrix zusammen, so erhält man die gewünschte Form der Darstellung aus Abbildung 1.1 mit einem Speicherplatzbedarf von $k(r+n)$.

Es bleibt die Frage nach dem dabei entstandenen Genauigkeitsverlust der Ausgangsmatrix C . Um den Fehler einzuschätzen, wird die Norm der Differenzmatrix $C - C_k$ untersucht. Nach Theorem 2.5.3 in [2] gilt für die Matrix C_k der folgende Satz:

Satz 1.1. Sei $C \in \mathbb{R}^{r \times n}$ gegeben mit der obigen Singulärwertzerlegung $C = U \Sigma V^T$, $\Sigma = \text{diag}\{\sigma_1, \dots, \sigma_{\min\{r,n\}}\}$. Sei $k < \text{rang}(C)$ und seien u_i und v_i die i -ten Spalten von U bzw. V . Die Matrix C_k hat dann die folgende Darstellung

$$C_k = \sum_{i=1}^k \sigma_i u_i v_i^T$$

Dann gilt für die Norm von $C - C_k$:

$$\min_{\text{rang}(B)=k} \|C - B\|_2 = \|C - C_k\|_2 = \sigma_{k+1}$$

Beweis . Siehe [2], Theorem 2.5.3. □

Besitzt C schnell abfallende Eigenwerte, so ist schon für kleine k , also $k \ll r, n$, eine gute Approximation der Form C_k zu erwarten, denn der Approximationsfehler σ_{k+1} ist entsprechend klein. Die Ausgangsmatrix auf diese Art zu komprimieren ist somit besonders vorteilhaft für Matrizen mit exponentiell abfallenden Eigen- bzw. Singulärwerten.

Um solch eine Speicher und Aufwand sparende Matrixapproximation von C_k überhaupt vornehmen zu können, muss jedoch die Singulärwertzerlegung bis zur Stufe k bekannt sein. Das heisst, es müssen Singulärwerte und linke und rechte Singulärvektoren berechnet werden, jedoch nicht das gesamte Spektrum. Um die Singulärwertzerlegung zu erhalten, kann auch die Eigenwertzerlegung von $A := C^T C = V \Sigma^2 V^T \in \mathbb{R}^{n \times n}$ betrachtet werden.

Die Berechnung der Matrix U bzw. U_k ist für die Niedrig-Rang-Darstellung C_k von C aufgrund der folgenden Nebenrechnung nicht notwendig, wenn V_k bekannt ist:

Setze $X := C V_k$ und $Y := V_k$, dann folgt für $X Y^T$:

$$\begin{aligned} X Y^T &= C U_k V_k^T = U \Sigma V^T V_k V_k^T \\ &= U \Sigma \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \hline & & & 0 \end{pmatrix} V_k^T = U \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_k & \\ \hline & & & 0 \end{pmatrix} V_k^T \\ &= U_k \Sigma_k \underbrace{V_k^T}_{=Y^T} = C_k \end{aligned}$$

Das heißt, es gilt $X = U_k \Sigma_k$. Damit erhält man allein mit den Matrizen $X = C V_k$ und $Y = V_k$ die gewünschte Speicherplatz sparende Darstellung der Ausgangsmatrix C aus Abbildung 1.1.

Über das *symmetrische Lanczos-Verfahren* ist bekannt, dass es sich im Allgemeinen nicht zur Berechnung eines vollständigen Eigenwertproblems eignet, aber gute Konvergenzeigenschaften für einige Eigenpaare zu den extremen Eigenwerten aufweist. Da zum Beispiel für das beschriebene Problem der Datenkomprimierung mit Hilfe der Singulärwertzerlegung auch nur einige, nämlich die k größten, Eigenwerte überhaupt von Interesse sind, ist die Berechnung der Eigenpaare mit dem Lanczos-Algorithmus ein vielversprechender Lösungsansatz. Das Lanczos-Verfahren ist ein Iterationsverfahren, welches auf der Basis von Krylovräumen Eigenraum-Approximationen der Ausgangsmatrix entwickelt.

1.2 Inhalt und Aufbau

Diese Arbeit beschäftigt sich mit der Anwendung des Lanczos-Verfahrens auf Matrizen mit exponentiell abfallenden Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n$. Da dieses Verfahren nur für symmetrische Matrizen definiert ist, wird das ursprüngliche Singulärwertproblem $C = U \Sigma V^T \in \mathbb{R}^{r \times n}$ in das Eigenwertproblem

$$A := C^T C = V D V^T \in \mathbb{R}^{n \times n}$$

umformuliert. Dabei ist $D = \Sigma^T \Sigma$ eine Diagonalmatrix, die die Singulärwerte zum Quadrat auf der Diagonalen enthält ($\sigma_i^2 = \lambda_i$). $V \in \mathbb{R}^{n \times n}$ ist eine orthonormale Matrix, deren Spalten die Eigenvektoren von A bilden.

Thema der Diplomarbeit sind Matrizen A mit Eigenwerten λ_i , für die die folgende Gleichung gilt:

$$\lambda_i = c_1 e^{-c_2(i-1)^\alpha} \quad \text{mit } c_1, c_2 \geq 0 \text{ und } \alpha \in (0, 1]$$

Der Parameter α entspricht häufig dem Kehrwert der Raumdimension des ursprünglichen Problems, zum Beispiel der zu diskretisierenden Integralgleichung.

Der Kern dieser Arbeit beschäftigt sich mit dem Verhalten des Lanczos-Verfahrens angewandt auf die speziellen Matrizen mit obigen Eigenwerten, und seinen Konvergenzeigenschaften.

Dazu werden in Kapitel 2 und 3 zunächst Grundlagen und Herleitung des Lanczos-Verfahrens vorgestellt. Im vierten Kapitel werden die allgemeinen Konvergenzaussagen des Verfahrens entwickelt und die entstandenen Fehlerabschätzungen für die vorgegebenen Eigenwerte untersucht. Es wird sich herausstellen, dass die ursprünglichen Konvergenzbetrachtungen unter der Voraussetzung exponentiell abfallender Eigenwerte vereinfacht werden können und gute Eigenpaar- Approximationen mit dem Lanczos-Algorithmus zu erwarten sind. Im letzten Kapitel werden dann die Ergebnisse des programmierten Verfahrens präsentiert. Diese bestätigen die Erwartungen der Theorie und liefern für bestimmte Parameter α, c_1, c_2 schon nach wenigen Iterationen sogar sehr gute Eigenpaar-Approximationen.

Hat man eine Matrix gegeben, die beispielsweise durch die Diskretisierung einer Integralgleichung entstanden ist, und deren Singulärwerte ab einer gewissen Stufe durch eine exponentiell abfallende Folge beschränkt sind, so ist die Berechnung einer Approximation ihrer Singulärwertzerlegung mit Hilfe des Lanczos-Verfahrens erfolgversprechend.

Danksagung:

Für die Vergabe des Themas und Betreuung danke ich Herrn Prof. Dr. Dr. Wolfgang Hackbusch.

Herrn Prof. Dr. Steffen Börm möchte ich herzlich für fachliche und insbesondere auch seelische Unterstützung danken, für die er immer Zeit gefunden hat.

Für die fachliche (und nicht fachliche) Diskussion und für sein Engagement danke ich Herrn Jens Burmeister.

Außerdem bedanke ich mich herzlich bei meinem Vater, Herrn Dr. Jürgen Kielmann, für das gründliche Korrekturlesen dieser Arbeit.

2 Grundlagen

In diesem Kapitel werden einige grundlegende Definitionen und Lemmata vorgestellt, welche einen Einstieg zur Lösung des späteren Eigenwertproblems geben sollen und in den nachfolgenden Kapiteln vorausgesetzt werden.

2.1 Eigenwertprobleme

Die Problemstellung dieser Arbeit ist die Singulärwertzerlegung einer gegebenen Matrix $C = U\Sigma V^T \in \mathbb{R}^{r \times n}$ mit $r, n \in \mathbb{N}$. O.B.d.A. sei $n < r$ und $\text{Rang } C = n$, U, V orthonormal und Σ eine Diagonalmatrix in $\mathbb{R}^{r \times n}$, welche die Singulärwerte $\sigma_1 \geq \dots \geq \sigma_n$ auf der Diagonalen enthält. Um diese Zerlegung mit Hilfe des später vorgestellten symmetrischen Lanczos-Verfahren zu berechnen, wird jedoch im Folgenden die symmetrische Matrix $A = C^T C$ betrachtet. Die Aufgabe ist nun also das numerische Lösen des Eigenproblems $A = V D V^T \in \mathbb{R}^{n \times n}$, wobei V die Matrix der Eigenvektoren von A ist und $D = \Sigma^T \Sigma$ Diagonalmatrix mit den Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n$, so dass $\lambda_i = \sigma_i^2$.

Da die Dimension der betrachteten Matrizen sehr hoch ist und analytische Ansätze nicht mehr verfolgt werden können, wird versucht, die Eigenwerte und Eigenvektoren mit numerischen Methoden zu approximieren. Ein erster und wichtiger Begriff in der Frage nach Eigenwerten mit zugehörigen Eigenvektoren, auf den in den späteren Kapiteln häufig zurückgegriffen werden wird, ist der **Rayleigh-Quotient**:

Definition 2.1. *Der Rayleigh-Quotient zu einer Matrix $A \in \mathbb{R}^{n \times n}$ und einem Vektor $x \neq 0, x \in \mathbb{R}^n$ ist definiert durch*

$$R(x, A) := \frac{\langle Ax, x \rangle}{\langle x, x \rangle}$$

Der Rayleigh-Quotient ist aufgrund der folgenden Bemerkung ein wichtiges Instrument zur Berechnung von Eigenwerten und Eigenvektoren von Matrizen.

Bemerkung 2.1. *Für jeden Eigenvektor x von A mit Eigenwert λ ist*

$$R(x, A) = \lambda$$

Um festzustellen, wie gut eine Eigenvektorapproximation ist, ist es wichtig, den Winkel zwischen zwei Vektoren und zwischen Vektoren und Vektorräumen zu beschreiben.

Definition 2.2. Der Winkel ϕ zwischen zwei Vektoren $x, y \in \mathbb{R}^n$ mit $x, y \neq 0$ wird beschrieben durch

$$\cos \angle(x, y) = \cos \phi := \frac{|\langle x, y \rangle|}{\|x\| \|y\|}$$

und

$$\sin \angle(x, y) = \sin \phi := \sqrt{1 - \cos^2 \phi}$$

Definition 2.3. Der Winkel ψ zwischen einem Vektor $x \in \mathbb{R}^n \setminus \{0\}$ und einem Unterraum X von \mathbb{R}^n ist definiert durch:

$$\cos \angle(x, X) = \cos \psi := \max_{y \in X} \cos \angle(x, y)$$

und

$$\sin \angle(x, X) = \sin \psi := \min_{y \in X} \sin \angle(x, y)$$

Da für die spätere Konvergenzanalyse auch der Tangens eines Winkels ψ benötigt wird, folgt noch die Definition von $\tan \psi$ mit Hilfe von \sin und \cos .

Definition 2.4. Für zwei Vektoren $x, y \in \mathbb{R}^n$ und $X \subseteq \mathbb{R}^n$ sei $\tan \angle(x, y)$ beziehungsweise $\tan \angle(x, X)$ definiert durch:

$$\tan \angle(x, y) := \frac{\sin \angle(x, y)}{\cos \angle(x, y)}$$

und

$$\tan \angle(x, X) := \min_{y \in X} \frac{\sin \angle(x, y)}{\cos \angle(x, y)}$$

Auch zwischen der Norm einer Matrix A und ihren Eigenwerten besteht ein Zusammenhang. Hierzu werden zwei weitere Begriffe eingeführt:

Definition 2.5. Sei $A \in \mathbb{R}^{n \times n}$. Das Spektrum von A wird bezeichnet mit $\sigma(A) := \{\lambda \mid \lambda \text{ Eigenwert von } A\}$. Sei außerdem der Spektralradius von A definiert durch $\rho(A) := \max \{|\lambda| \mid \lambda \in \sigma(A)\}$.

Lemma 2.1. Sei $A \in \mathbb{R}^{n \times n}$. Dann gilt:

$$\|A\| = \sqrt{\rho(A^T A)}$$

Ist A eine normale Matrix, also $A^T A = A A^T$, dann gilt für die Spektralnorm von A :

$$\|A\| = \rho(A)$$

Insbesondere gilt dies also für symmetrische Matrizen.

Beweis . Siehe [3], Satz 2.9.5 □

Eine weitere Beziehung zwischen den Eigenwerten einer Matrix und dem Rayleigh-Quotienten ist in dem folgenden Lemma festgehalten:

Lemma 2.2. Sei $A \in \mathbb{R}^{n \times n}$, $A = A^T$ mit Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n$ und zugehörigen orthonormalen Eigenvektoren x^1, \dots, x^n und sei $X^k := \text{span}\{x^1, \dots, x^k\}$. Dann gilt:

$$\min_{x \in X^k} R(x, A) = \lambda_k = \max_{y \in X^{k-1}^\perp} R(y, A)$$

Bemerkung 2.2. Eigenvektoren symmetrischer Matrizen können immer orthonormal gewählt werden. Für $A \in \mathbb{R}^{n \times n}$ bilden sie eine Orthonormalbasis (kurz ONB) des \mathbb{R}^n .

Beweis von Lemma 2.2. Sei

$$x = \sum_{i=1}^k \alpha_i x^i \in X^k, \quad x \neq 0 \quad \text{mit} \quad \alpha_i \in \mathbb{R}$$

Da die Eigenvektoren x^1, \dots, x^n orthonormal sind, gilt für die Norm von x

$$\|x\|^2 = \left\langle \sum_{i=1}^k \alpha_i x^i, \sum_{l=1}^k \alpha_l x^l \right\rangle = \sum_{i=1}^k \langle \alpha_i x^i, \alpha_i x^i \rangle = \sum_{i=1}^k \alpha_i^2$$

Weiter ist wegen der Voraussetzung $\lambda_k \leq \lambda_1, \dots, \lambda_{k-1}$

$$R(x, A) = \frac{\langle Ax, x \rangle}{\langle x, x \rangle} = \frac{\langle \sum_{i=1}^k \lambda_i \alpha_i x^i, \sum_{l=1}^k \alpha_l x^l \rangle}{\|x\|^2} = \frac{\sum_{i=1}^k \lambda_i \alpha_i^2}{\|x\|^2} \geq \frac{\lambda_k \sum_{i=1}^k \alpha_i^2}{\sum_{i=1}^k \alpha_i^2} = \lambda_k$$

Da x beliebig gewählt wurde, folgt die linke Gleichung des Lemmas.

Sei nun $0 \neq y \in X^{k-1}^\perp = \text{span}\{x^k, \dots, x^n\}$ also

$$y = \sum_{i=k}^n \beta_i x^i \quad \text{mit} \quad \beta_i \in \mathbb{R}$$

Analog zum ersten Teil gilt wegen $\lambda_k \geq \lambda_{k+1}, \dots, \lambda_n$

$$\|y\|^2 = \sum_{i=k}^n \beta_i^2 \quad \text{und} \quad R(y, A) = \frac{\langle Ay, y \rangle}{\langle y, y \rangle} = \frac{\sum_{i=k}^n \lambda_i \beta_i^2}{\sum_{i=k}^n \beta_i^2} \leq \lambda_k$$

Damit ist auch der rechte Teil der Gleichung gezeigt. □

Mit Hilfe von Lemma 2.2 wird nun das **Min-Max-Theorem** bewiesen, welches insbesondere für Kapitel 4, in dem die Konvergenz des Lanczos-Verfahrens betrachtet wird, von Bedeutung ist.

Satz 2.1 (Min-Max-Theorem). *Sei A wie in Lemma 2.2 mit Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n$. Dann gilt:*

$$\max_{\substack{W \subseteq \mathbb{R}^n \\ \dim W = i}} \min_{\substack{y \in W \\ y \neq 0}} R(y, A) = \lambda_i = \min_{\substack{V \subseteq \mathbb{R}^n \\ \dim V = i-1}} \max_{\substack{x \in V^\perp \\ x \neq 0}} R(x, A)$$

Beweis . Seien V und W beliebige Unterräume des \mathbb{R}^n mit $\dim V = i-1$ und $\dim W = i$. Nach dem Dimensionssatz für Vektorräume gilt

$$(2.1) \quad \dim V + \dim V^\perp = n \quad \text{also} \quad \dim V^\perp = n - (i-1) \quad \text{und}$$

$$(2.2) \quad \dim(V^\perp + W) = \dim V^\perp + \dim W - \dim(V^\perp \cap W)$$

Da V^\perp und W in \mathbb{R}^n enthalten sind, kann $\dim(V^\perp + W)$ höchstens n sein. Also folgt aus (2.1) und (2.2): $n - (i-1) + i - \dim(V^\perp \cap W) \leq n$. Daraus folgt

$$(2.3) \quad \dim(V^\perp \cap W) \geq 1$$

Wegen (2.3) existiert mindestens ein $z \in (V^\perp \cap W)$ mit $z \neq 0$. Also gilt für dieses z :

$$\max_{\substack{x \in V^\perp \\ x \neq 0}} R(x, A) \geq R(z, A) \geq \min_{\substack{y \in W \\ y \neq 0}} R(y, A)$$

Insbesondere gilt also

$$(2.4) \quad \max_{\substack{x \in V^\perp \\ x \neq 0}} R(x, A) \geq \min_{\substack{y \in W \\ y \neq 0}} R(y, A)$$

Da die Vektorräume beliebig gewählt wurden, gilt (2.4) auch für:

$$(2.5) \quad \min_{\substack{V \subseteq \mathbb{R}^n \\ \dim V = i-1}} \max_{\substack{x \in V^\perp \\ x \neq 0}} R(x, A) \geq \max_{\substack{W \subseteq \mathbb{R}^n \\ \dim W = i}} \min_{\substack{y \in W \\ y \neq 0}} R(y, A)$$

Setze nun $V^* := X^{i-1}$ also $V^{*\perp} = \text{span}\{x^i, \dots, x^n\}$ und $W^* := X^i = \text{span}\{x^1, \dots, x^i\}$. Dann gilt mit (2.5) und Lemma 2.2:

$$\lambda_i = \max_{x \in V^{*\perp}} R(x, A) \geq \min_{\substack{V \subseteq \mathbb{R}^n \\ \dim V = i-1}} \max_{\substack{x \in V^\perp \\ x \neq 0}} R(x, A) \geq \max_{\substack{W \subseteq \mathbb{R}^n \\ \dim W = i}} \min_{\substack{y \in W \\ y \neq 0}} R(y, A) \geq \min_{y \in W^*} R(y, A) = \lambda_i$$

Damit ist die Gleichheit gezeigt und somit die Behauptung. □

Mit Hilfe von Lemma 2.1 wird noch das folgende Korollar bewiesen, welches für den Beweis der Optimalität des im nächsten Abschnitt vorgestellten Rayleigh-Ritz-Verfahrens gebraucht wird.

Korollar 2.1. Seien $A, B, C \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit. Gilt $A = B + C$ so folgt $\|A\| \geq \|B\|$ und $\|A\| \geq \|C\|$

Beweis . Aus $A = B + C$ folgt $Ax = Bx + Cx$ für alle $x \in \mathbb{R}^n$, und somit gilt auch:

$$\langle Ax, x \rangle = \langle Bx, x \rangle + \langle Cx, x \rangle$$

Da A, B und C positiv definit sind, sind alle obigen Skalarprodukte größer als Null für ein $x \neq 0$. Also ist

$$(2.6) \quad \langle Ax, x \rangle \geq \langle Bx, x \rangle \quad \text{für alle } x \in \mathbb{R}^n$$

Insbesondere folgt also auch:

$$(2.7) \quad R(x, A) \geq R(x, B) \quad \text{für alle } x \in \mathbb{R}^n$$

Weiter gilt mit der Eigenschaft der positiven Definitheit von A und B , dass auch alle Eigenwerte von A und B größer als Null sind. Somit gilt insbesondere $0 < \lambda_{A_{max}} := \rho(A)$ und $0 < \lambda_{B_{max}} := \rho(B)$. Außerdem gilt wegen der Symmetrie von A und B nach Lemma 2.1: $\|A\| = \rho(A)$ beziehungsweise $\|B\| = \rho(B)$.

Seien x_A und x_B die Eigenvektoren zu $\lambda_{A_{max}}$ beziehungsweise zu $\lambda_{B_{max}}$, so dass gilt:

$$(2.8) \quad \rho(A) = R(x_A, A) \quad \text{und} \quad \rho(B) = R(x_B, B)$$

Es folgt mit (2.7) und (2.8):

$$\|A\| \stackrel{\text{Lem.2.1}}{=} \rho(A) \stackrel{(2.8)}{=} R(x_A, A) \stackrel{(2.7)}{\geq} R(x_B, B) \stackrel{(2.8)}{=} \rho(B) \stackrel{\text{Lem.2.1}}{=} \|B\|$$

Analog folgt die Behauptung für C . □

2.2 Invariante Unterräume

Da das im nächsten Kapitel vorgestellte Rayleigh-Ritz-Verfahren (und insbesondere das Lanczos-Verfahren) zu einer gegebenen Matrix $A \in \mathbb{R}^{n \times n}$ Unterräume des \mathbb{R}^n kleinerer Dimension benutzt, um das Eigenwertproblem von A zu approximieren, wird hier eine bezüglich Eigenwertproblemen wichtige Klasse von Unterräumen vorgestellt.

Definition 2.6 (Invarianter Unterraum). Seien $m, n \in \mathbb{N}, m < n$ und S^m ein m -dimensionaler Unterraum des \mathbb{R}^n . Gilt

$$AS^m = \{ As \mid s \in S^m \} \subset S^m$$

so ist S^m ein unter A **invarianter Unterraum**.

Bemerkung 2.3. Insbesondere ist jeder Eigenraum einer Matrix A ein invarianter Unterraum.

Für das nächste wichtige Lemma, welches den Bezug der invarianten Unterräume zu Eigenwertproblemen herstellt, wird zuerst der folgende Korollar bewiesen:

Korollar 2.2. Sei A symmetrisch, $A \in \mathbb{R}^{n \times n}$ und S^m invarianter Unterraum. Sei $v \in S$ ein Eigenvektor und $V = \text{span}\{v\}$. Dann gilt für $V^\perp := \{w \in S^m \mid w \perp V\}$:

$$AV^\perp \subseteq V^\perp$$

Beweis. Sei $w \in V^\perp$. Zu zeigen ist $Aw \in V^\perp$. Da A symmetrisch ist, können $n - 1$ zu v orthogonale, normierte Eigenvektoren q_2, \dots, q_n von A gefunden werden, so dass A zerlegt werden kann in $A = QDQ^T$. Dabei sei Q die Matrix der Eigenvektoren und D Diagonalmatrix mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ von A ist. Wegen der Orthogonalität von v und q_2, \dots, q_n gilt $V^\perp = \text{span}\{q_2, \dots, q_n\}$. Sei o.B.d.A. $Q = [v, q_2, \dots, q_n]$. Da $w \perp V$ ist, folgt:

$$\begin{aligned} Aw &= (QDQ^T)w = (QD)[v, q_2, \dots, q_n]^T w = (QD) \begin{pmatrix} 0 \\ q_2^T w \\ \vdots \\ q_n^T w \end{pmatrix} \\ &= Q \begin{pmatrix} 0 \\ \lambda_2 q_2^T w \\ \vdots \\ \lambda_n q_n^T w \end{pmatrix} = \begin{pmatrix} \sum_{i=2} q_{i1} \lambda_i q_i^T w \\ \vdots \\ \sum_{i=2} q_{in} \lambda_i q_i^T w \end{pmatrix} \in \text{span}\{q_2, \dots, q_n\} = V^\perp \end{aligned}$$

□

Mit Hilfe von Korollar 2.2 kann nun das folgende Lemma bewiesen werden:

Lemma 2.3. Ein unter $A = A^T \in \mathbb{R}^{n \times n}$ invarianter Unterraum $S^m \subset \mathbb{R}^n$ hat eine Basis aus Eigenvektoren von A .

Beweis. Induktion über die Dimension m von S^m :

Induktionsvoraussetzung:

Sei $m = 1$. Dann ist $S^m = \text{span}\{s\}$ mit $s \in \mathbb{R}^n$. Da S^m invariant ist, gilt $As \in \text{span}\{s\} = \{\lambda s \mid \lambda \in \mathbb{R}\}$. Also ist s Eigenvektor von A .

Induktionsannahme:

Gelte die Behauptung für $S^m \subset \mathbb{R}^n$ mit $\dim S^m \leq m$, $m < n$.

Induktionschluß:

Sei $\dim S = m + 1$. Da $AS^m \subseteq S^m$ nach Voraussetzung gilt, gilt auch:

$$(2.9) \quad p(A)S^m \subseteq S^m \quad \text{für alle } p \in \mathcal{P}_n$$

Setze

$$(2.10) \quad p_j(x) := \prod_{\substack{i=1 \\ i \neq j}}^n \frac{x - \lambda_i}{\lambda_j - \lambda_i}, \quad \text{also } p_j \in \mathcal{P}_n \text{ und } p_j(\lambda_i) = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j \end{cases}$$

Da $\dim S^m > 0$, existiert ein $s \in S^m \setminus \{0\}$. Weil die Eigenvektoren x^1, \dots, x^n von A eine ONB von \mathbb{R}^n bilden (siehe Bem. 2.2), kann s dargestellt werden durch

$s = \sum_{i=1}^n \alpha_i x^i$ für $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Es gilt für ein beliebiges Polynom $p \in \mathcal{P}_n$:

$$p(A)s = p(A) \sum_{i=1}^n \alpha_i x^i = \sum_{i=1}^n \alpha_i p(A)x^i = \sum_{i=1}^n \alpha_i p(\lambda_i)x^i$$

Wegen $s \neq 0$ existiert mindestens ein $j \in \{1, \dots, n\}$ mit $\alpha_j \neq 0$. Dann gilt für dieses j nach (2.10) und (2.9):

$$p_j(A)s = \sum_{i=1}^n \alpha_i p_j(\lambda_i)x^i = \alpha_j \lambda_j x^j \in S^m$$

Also enthält S^m mindestens den Eigenvektor x^j von A , also auch $V := \text{span} \{x^j\}$. Das heißt, für V^\perp kann Korollar 2.2 angewendet werden, und es folgt:

$$(2.11) \quad AV^\perp \subseteq V^\perp$$

Aufgrund der Dimensionsformel gilt $\dim V^\perp = \dim S^m - \dim V = m - 1$. Es folgt aus (2.11) und der Induktionsannahme: V^\perp wird aufgespannt von Eigenvektoren von A . Aus

$$S^m = V \oplus V^\perp$$

folgt die Behauptung. □

3 Lanczos-Verfahren

In diesem Kapitel wird die Herleitung des Lanczos-Verfahrens beschrieben. Dazu wird zuerst das Rayleigh-Ritz-Verfahren vorgestellt, welches in Zusammenhang mit den darauffolgend erklärten Krylovräumen das Lanczos-Verfahren bildet. Das Kapitel endet mit dem allgemein bekannten Lanczos-Algorithmus. Später wird der Algorithmus für das spezielle Eigenwertproblem dieser Arbeit noch leicht abgewandelt und in Kapitel 5 dargestellt.

3.1 Rayleigh-Ritz-Verfahren

Wie die bereits in der Einleitung beschrieben, entstehen in der Praxis durch Diskretisierung von Integralgleichungen häufig hochdimensionale voll besetzte Matrizen. Man ist aber meist nicht an der gesamten Matrix interessiert, sondern nur an bestimmten Singulärwerten und -Vektoren, um die Singulärwertzerlegung der Ausgangsmatrix zu approximieren. Sei im folgenden eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ für ein sehr großes n gegeben. A besitze die Eigenwerte $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ und die zugehörigen Eigenvektoren x^1, \dots, x^n . In dieser Arbeit werden Matrizen mit exponentiell abfallenden Eigenwerten betrachtet, so dass für die Approximation der Eigenwertzerlegung aufgrund der Größe von A nicht alle Eigenpaare erwünscht sind, sondern nur einige der größten Eigenwerte mit zugehörigen Eigenvektoren. Eine Matrixzerlegung, z.B. mit dem QR-Verfahren (Siehe z.B. [7], Kapitel 9.4), ist hier also nicht unbedingt von Vorteil, da nicht alle Eigenpaare benötigt werden und die dabei verwendeten Matrixberechnungen aufgrund der Größe von A nicht empfehlenswert sind.

Allgemein existieren in der Praxis häufig Probleme, zu denen die Matrix A nicht vollständig bekannt ist, sondern nur das Matrix-Vektorprodukt Ax , $x \in \mathbb{R}^n$. In diesem Fall ist eine QR-Zerlegung nicht möglich, während das Lanczos-Verfahren nur die Operation Ax benötigt, und somit trotzdem angewendet werden kann.

Die Überlegungen des vorangegangenen Absatzes führen zu der Idee, die in Kapitel 2.2 vorgestellten **invarianten Unterräume** zu benutzen:

Hat man einen unter A invarianten Unterraum $S^m \subset \mathbb{R}^n$ gegeben, so enthält dieser nach Lemma 2.3 eine Basis aus Eigenvektoren von A .

Hat dieser Unterraum S^m die Dimension $m \ll n$, so können aus ihm m Eigenpaare von A berechnet werden. Da m üblicherweise sehr viel kleiner ist als n , sind diese Berechnungen mit relativ wenig Aufwand verbunden.

Aus dieser Idee entstehen zwei Fragen:

1. Wie erhält man überhaupt einen passenden Unterraum?
2. Hat man einen beliebigen Unterraum gegeben, wie kann man feststellen, ob dieser invariant unter A ist?

Die erste Frage wird im nächsten Abschnitt dieses Kapitels durch die sogenannten **Krylovräume** beantwortet.

Im Folgenden wird zunächst die zweite Frage behandelt. Hierzu wird versucht, die Frage nach der Invarianz von einem gegebenen m -dimensionalen Unterraum S^m zu präzisieren.

Sei $A \in \mathbb{R}^{n \times n}$ und sei zu $S^m \subset \mathbb{R}^n$ eine Matrix $F = [f_1, \dots, f_m]$, $F \in \mathbb{R}^{n \times m}$ gegeben, so dass

$$\text{Bild } F = \{ Fx \mid x \in \mathbb{R}^m \} = \text{span} \{ f_1, \dots, f_m \} = S^m$$

ist. Nach der Definition 2.6 ist S^m genau dann invariant unter A , wenn $A S^m \subseteq S^m$ gilt. Da die Spalten von F eine Basis von S^m bilden, ist diese Definition gleichbedeutend mit

$$(3.1) \quad Af_j = \sum_{i=1}^m f_i b_{ij} \quad \text{für alle } i, j \in \{1, \dots, m\}, \quad \text{wobei } b_{ij} \in \mathbb{R}$$

Um diese Gleichungen in einer einfacheren Weise zu beschreiben, wird (3.1) in eine Matrixgleichung umformuliert. (3.1) ist äquivalent zu:

$$(3.2) \quad \boxed{A} \boxed{F} = \boxed{F} \boxed{B}$$

Korollar 3.1. *Wenn die Matrix $F \in \mathbb{R}^{n \times m}$ vollen Rang hat, so existiert genau eine Lösung für Gleichung (3.2), nämlich $B = (F^T F)^{-1} F^T A F$. Ist $\text{Rang}(F) < m$, so existieren viele Lösungen von (3.2).*

Bemerkung 3.1. *Für eine rechteckige Matrix $Q \in \mathbb{R}^{n \times m}$ mit $m \leq n$, deren Spaltenvektoren orthonormal sind, gilt $Q^T Q = I_m$. Eine Matrix für die dies gilt, sei im Folgenden als orthonormal bezeichnet, auch wenn für sie **nicht** $Q^T = Q^{-1}$ gilt.*

Hat man eine orthonormale Matrix $Q \in \mathbb{R}^{n \times m}$ gefunden, welche (3.2) mit $F = Q$ erfüllt, so ergeben sich einige Vereinfachungen des Invarianztests in (3.2) gegenüber Korollar 3.1.

Korollar 3.2. *Sei $Q \in \mathbb{R}^{n \times m}$ eine orthonormale Matrix und es gelte $AQ - QB = 0$ für $B \in \mathbb{R}^{m \times m}$. Dann gilt $B = Q^T A Q$.*

Korollar 3.3. *Unter den Voraussetzungen von Korollar 3.2 gilt: Ist A symmetrisch, so ist auch $B = Q^T A Q$ symmetrisch.*

Beweis . Es gilt:

$$B^T = (Q^T A Q)^T = Q^T A^T Q = Q^T A Q = B$$

□

Warum ist nun die Darstellung (3.2) eine Vereinfachung in Bezug auf das Eigenwertproblem einer Matrix A , welches mit Hilfe eines invarianten Unterraums gelöst werden soll? Die Antwort gibt das folgende Lemma:

Lemma 3.1. *Gelte für gegebenes $A \in \mathbb{R}^{n \times n}$, $F \in \mathbb{R}^{n \times m}$ mit $m \leq n$, $\text{Rang}(F) = m$ und $B \in \mathbb{R}^{m \times m}$ die Gleichung (3.2), also $A F - F B = 0$. Dann folgt:
Ist $y \in \mathbb{R}^m$ ein Eigenvektor der Matrix B zum Eigenwert λ , so ist $F y$ ein Eigenvektor von A zum selben Eigenwert λ .*

Beweis . Sei $y \in \mathbb{R}^m$ und $\lambda \in \mathbb{R}$ mit $B y = \lambda y$. Nach Voraussetzung gilt $A F - F B = 0$, also ist auch $A F y - F B y = 0$. Daraus folgt

$$A F y - F \lambda y = 0 \quad \text{also} \quad A (F y) = \lambda (F y)$$

□

Bemerkung 3.2 (Approximative invariante Unterräume). *Nach Korollar 3.2 und 3.3 ist es sinnvoll, einen Unterraum zu bestimmen, der von einer Orthonormalbasis aufgespannt wird. Normalerweise hat man jedoch keinen invarianten Unterraum zur Verfügung. Also wird versucht, einen geeigneten Unterraum zu finden, der wenigstens eine Näherung an einen invarianten Unterraum darstellt. Dieser beinhaltet dann relativ gute Eigenvektorapproximationen.*

Aufgrund von Bemerkung 3.2 beziehen sich die folgenden Überlegungen und Algorithmen dieses Kapitels auf **approximative** invariante Unterräume.

In diesem Zusammenhang wird auch der Begriff der Residuumsmatrix eingeführt:

Definition 3.1. *Zu gegebenen Matrizen $A \in \mathbb{R}^{n \times n}$ und $F \in \mathbb{R}^{n \times m}$ bezeichne*

$$R(M) := A F - F M$$

*die **Residuumsmatrix** zu $M \in \mathbb{R}^{m \times m}$*

Hat man einen geeigneten Unterraum S^m bestimmt, so zeigt der folgende Algorithmus, wie man daraus die gewünschten Eigenpaar-Approximationen von A erhält.

Der Rayleigh-Ritz-Algorithmus :

Sei ein Unterraum der Dimension m , $S^m \subset \mathbb{R}^n$ gegeben mit $S^m = \text{span}\{s_1, \dots, s_m\}$ und $m \ll n$. Der Rayleigh-Ritz-Algorithmus verfährt wie folgt:

1. Orthonormalisiere s_1, \dots, s_m zu q_1, \dots, q_m und bilde $Q = [q_1, \dots, q_m] \in \mathbb{R}^{n \times m}$
2. Berechne $B := Q^T A Q \in \mathbb{R}^{m \times m}$
3. Löse das Eigenwertproblem für die kleinere Matrix B .
Dies kann zum Beispiel mit Hilfe des QR-Verfahrens geschehen, vergleiche [7]. Seien μ_i die Eigenwerte und $y^i \in \mathbb{R}^m$ die zugehörigen (orthonormalen) Eigenvektoren für $i = 1, \dots, m$.
4. Berechne $z^i := Q y^i$ für $i \in \{1, \dots, m\}$
5. Gib (μ_i, z^i) , $i \in \{1, \dots, m\}$ zurück

Definition 3.2. Die Matrix $B = Q^T A Q$ heißt **Ritz-Matrix**, die Eigenwerte μ_i von B sind die **Ritz-Werte** und die Vektoren $z^i = Q y^i$ werden als **Ritz-Vektoren** bezeichnet.

Warum benutzt man nun gerade dieses Verfahren um aus einem Unterraum Eigenpaare zu approximieren?

Das Rayleigh-Ritz-Verfahren (kurz : RR-Verfahren) liefert mit dem Ritz-Paar (μ_i, z^i) die besten Näherungen von Eigenpaaren von A , welche aus dem gegebenen Unterraum S^m erreicht werden können (vgl. [8], Seite 235 und [6], Seite 133).

Nach dem Min-Max-Theorem (Satz 2.1) gilt für die Eigenwerte von A :

$$\lambda_i = \max_{\substack{W \subset \mathbb{R}^n \\ \dim W = i}} \min_{\substack{x \in W \\ x \neq 0}} R(x, A)$$

Die beste Approximation der Eigenwerte von A aus dem kleiner dimensionalen Unterraum $S^m \subset \mathbb{R}^n$ sind also

$$\theta_i := \max_{\substack{W \subset S^m \\ \dim W = i}} \min_{\substack{s \in W \\ s \neq 0}} R(s, A)$$

Die Optimalität der Ritz-Werte gibt somit das folgende Lemma wieder:

Lemma 3.2. Seien μ_1, \dots, μ_m die mit RR-Verfahren aus $S^m \subset \mathbb{R}^n$ berechneten Ritz-Werte. Dann gilt:

$$\mu_i = \theta_i$$

Beweis . Sei Q die Matrix, deren Spalten eine ONB von S^m bilden, also $S^m = \{Qa \mid a \in \mathbb{R}^m\}$. Das heißt, für jedes $x \in S^m$ gilt $x = Qa$ für ein $a \in \mathbb{R}^m$. Damit folgt:

$$\begin{aligned} \theta_i &= \max_{\substack{W \subset S^m \\ \dim W = i}} \min_{\substack{x \in W \\ x \neq 0}} R(x, A) \\ &= \max_{\substack{W \subset S^m \\ \dim W = i}} \min_{\substack{x = Qa \in W \\ x \neq 0}} R(Qa, A) \end{aligned}$$

$$\begin{aligned}
 &= \max_{\substack{S \subseteq \mathbb{R}^m \\ \dim S = i}} \min_{\substack{a \in S \\ a \neq 0}} R(a, Q^T A Q) \\
 &= \mu_i
 \end{aligned}$$

nach dem Min-Max-Theorem 2.1

□

Auch in anderer Weise ist das Rayleigh-Ritz-Verfahren optimal, nämlich in Bezug auf die Matrix Q , deren Spaltenvektoren eine Basis von S^m bilden. Den Beweis dafür liefert das folgende Lemma.

Lemma 3.3. Sei $Q \in \mathbb{R}^{n \times m}$ orthonormal und $B = Q^T A Q$. Dann gilt

$$\|R(B)\| \leq \|R(M)\| \quad \text{für alle } M \in \mathbb{R}^{m \times m}$$

Beweis . Sei $M \in \mathbb{R}^{m \times m}$. Betrachte zuerst:

$$\begin{aligned}
 (3.3) \quad R(M)^T R(M) &= (AQ - QM)^T (AQ - QM) \\
 &= (Q^T A - M^T Q^T)(AQ - QM) && \text{da } A \text{ symmetrisch} \\
 &= Q^T A^2 Q - Q^T A Q M - M^T Q^T A Q + M^T \underbrace{Q^T Q}_{=I} M \\
 &= Q^T A^2 Q - BM - M^T B + M^T M && \text{da } B = Q^T A Q
 \end{aligned}$$

Da B nach Korollar 3.3 symmetrisch ist, gilt außerdem:

$$(3.4) \quad (B - M)^T (B - M) = B^2 - BM - M^T B + M^T M$$

Für B ergibt sich aus (3.3), indem man dort M durch B ersetzt:

$$(3.5) \quad R(B)^T R(B) = Q^T A^2 Q - B^2$$

Aus (3.3) folgt nun durch einsetzen von (3.4) und (3.5):

$$(3.6) \quad R(M)^T R(M) = (B - M)^T (B - M) + R(B)^T R(B)$$

Da alle Matrizenprodukte in (3.6) symmetrisch positiv definit sind, folgt nach Korollar 2.1 eine Abschätzung für die Norm:

$$(3.7) \quad \|R(B)^T R(B)\| \leq \|R(M)^T R(M)\|$$

Nach Lemma 2.1 ergibt sich somit

$$\begin{aligned}
 \|R(B)^T R(B)\| &= \rho(R(B)^T R(B)) = \|R(B)\|^2 \quad \text{und} \\
 \|R(M)^T R(M)\| &= \rho(R(M)^T R(M)) = \|R(M)\|^2
 \end{aligned}$$

Mit (3.7) folgt somit die Behauptung.

□

Aus Lemma 3.2 geht hervor, dass die Ritz-Werte μ_i optimal gewählt wurden. Leider erfüllen die Ritz-Vektoren z^i nicht diese Eigenschaft der Bestapproximation (vgl. [8], Seite 238). Es gilt für sie aber zumindest das folgende Lemma.

Lemma 3.4. *Seien (μ_i, z^i) Ritz-Paare des Unterraums S^m mit $z^i = Qy^i$, y^i Eigenvektoren zu den Eigenwerten μ_i der Ritz-Matrix. Definiere den **Residuumvektor** r zu einem Ritz-Vektor z^i durch $r(z^i) := Az^i - \mu_i z^i$. Es gilt*

$$r(z^i) \perp S^m$$

Beweis . Gelte $S^m = \text{span}\{q_1, \dots, q_m\}$, q_i orthonormal. Dann ist zu zeigen: $\langle r(z^i), q_i \rangle = 0$. Es gilt:

$$\begin{aligned} \langle r(z^i), q_i \rangle = 0 &\Leftrightarrow Q^T r(z^i) = 0 \\ &\Leftrightarrow Q^T (Az^i - \mu_i z^i) = 0 \\ &\Leftrightarrow Q^T Az^i = \mu_i Q^T z^i \\ &\Leftrightarrow Q^T A(Qy^i) = \mu_i Q^T (Qy^i) \\ &\Leftrightarrow (Q^T A Q)y^i = \mu_i y^i \end{aligned}$$

Die letzte Gleichung gilt nach Punkt 3 des Rayleigh-Ritz-Algorithmus, denn die μ_i sind gerade die Eigenwerte zu den Eigenvektoren y^i der Ritz-Matrix $B = Q^T A Q$. \square

Das Rayleigh-Ritz-Verfahren beruht auf einem vorhandenen m -dimensionalen Unterraum S^m . Es bleibt immer noch die Frage, wie man einen solchen Unterraum erhält, welcher die gewünschte Dimension m hat, mit wenig Aufwand zu bestimmen ist und welcher vielleicht schon einige Eigenvektorapproximationen enthält. Diese Frage behandelt das nächste Unterkapitel.

3.2 Krylovräume

Sei auch hier im Folgenden $A = A^T \in \mathbb{R}^{n \times n}$ mit Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ und zugehörigen Eigenvektoren x^1, \dots, x^n .

Bei der Vektoriteration (vgl. [4], [8]) werden die Vektoren $x, Ax, \dots, A^m x$ für einen Startvektor $x \in \mathbb{R}^n$ nacheinander berechnet. Wenn x nicht orthogonal zu x^1 ist, konvergiert diese Folge gegen x^1 , also gegen den Eigenvektor mit dem betragsmäßig größten Eigenwert. Die Verwendung von Krylovräumen basiert auf der Idee, die einzelnen Folgenglieder $A^k x$ nicht wieder zu überschreiben, sondern den von ihnen aufgespannten Unterraum zu berechnen.

Definition 3.3 (Krylovraum). *Zu $m \in \mathbb{N}$ und $x \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n \times n}$ ist der **Krylovraum** definiert durch $\mathcal{K}_m(x, A) := \text{span}\{x, Ax, \dots, A^{m-1}x\}$*

Warum ist nun gerade der Krylovraum vorteilhaft zur Verwendung als Unterraum-Approximation des RR-Verfahrens? Die nachfolgenden Überlegungen und Eigenschaften erklären die Wahl der Krylovräume.

Für ein großes n und $m \in \mathbb{N}$ mit $m \ll n$ ist zu vermuten, dass die Vektoren $A^k x$ linear unabhängig sind. Deshalb ist $\dim \mathcal{K}_m(x, A) = m$ zu erwarten.

Weil die Eigenvektoren x^1, \dots, x^n von A eine (orthonormale) Basis des \mathbb{R}^n bilden, kann ein beliebiges $x \in \mathbb{R}^n$ dargestellt werden als

$$x = \sum_{i=1}^n \alpha_i x^i \quad \text{mit } \alpha_i \in \mathbb{R}$$

Da der Krylovraum $\mathcal{K}_m(x, A)$ von den Vektoren $A^k x$, $k \in \{0, 1, \dots, m-1\}$ aufgespannt wird, betrachte:

$$A^k x = A^k \sum_{i=1}^n \alpha_i x^i = \sum_{i=1}^n \alpha_i \lambda_i^k x^i$$

Das heißt, für große Eigenwerte λ_i wird für wachsendes k der Anteil in Richtung der Eigenvektoren x^i verstärkt, während der Anteil von Eigenvektoren zu kleinen Eigenwerten abnimmt. Aufgrund der in diesem Kapitel vorausgesetzten abfallenden Eigenwerte $\lambda_1 \geq \dots \geq \lambda_n$ von A beinhaltet $\mathcal{K}_m(x, A)$ Näherungen an die Eigenvektoren zu größeren Eigenwerten.

Eine andere Sichtweise, die Verwendung von Krylovräumen zu erklären, ist, den Gradienten des Rayleigh-Quotienten $R(x, A)$ zu betrachten.

Nach der Definition des Rayleigh-Quotienten gilt für einen Eigenvektor x_λ und Eigenwert λ :

$$R(x_\lambda, A) = \lambda$$

Aufgrund von Lemma 2.2 sind die Eigenvektoren von A gerade die Extremstellen von $R(x, A)$. Das heißt, die Ableitung von $R(x, A)$ ist Null für x_λ .

Setze $f(x) := \langle Ax, x \rangle$ und $g(x) := \langle x, x \rangle$. Für den Gradienten des Rayleigh-Quotienten gilt:

$$(3.8) \quad \text{grad } R(x, A) = \text{grad } \frac{f(x)}{g(x)} = \frac{g(x)(\text{grad } f(x)) - (\text{grad } g(x))f(x)}{g^2(x)}$$

Es gilt da A symmetrisch ist:

$$\text{grad } f(x) = \text{grad } \langle Ax, x \rangle = Ax + A^T x = 2Ax \quad \text{und}$$

$$\text{grad } g(x) = \text{grad } \langle x, x \rangle = 2x$$

Eingesetzt in (3.8) ergibt sich:

$$\text{grad } R(x, A) = \frac{2}{\langle x, x \rangle} (Ax - R(x, A)x)$$

Diese letzte Gleichung zeigt, dass es eine gute Idee ist, den nächsten Basisvektor des Unterraums in Richtung Ax zu suchen.

Also erfüllen die Krylovräume schon zwei der am Ende des letzten Abschnitts geforderten Eigenschaften. Ein weiterer Vorteil ist, dass zur Berechnung eines Krylovraums für beliebiges m nur der Vektor x und die Matrix A benötigt werden. Hat man $\mathcal{K}_k(x, A) = \text{span} \{x, Ax, \dots, A^{k-1}x\}$ bestimmt, erhält man $\mathcal{K}_{k+1}(x, A)$ durch Berechnen von $A(A^{k-1}x)$. Es gilt also

Bemerkung 3.3. $A\mathcal{K}_m(x, A) \subset \mathcal{K}_{m+1}(x, A)$ für alle $m \in \mathbb{N}$.

Außerdem erfüllen Krylovräume noch die folgenden Eigenschaften.

Korollar 3.4 (Skalierungsinvarianz). *Es gilt $\mathcal{K}_m(\alpha x, \beta A) = \mathcal{K}_m(x, A)$ für alle $\alpha, \beta \in \mathbb{R}^n \setminus \{0\}$.*

Beweis . Es ist

$$\begin{aligned} \mathcal{K}_m(\alpha x, \beta A) &= \text{span} \{ \alpha x, \alpha(\beta A)x, \dots, \alpha(\beta A)^{m-1}x \} \\ &= \text{span} \{ \alpha x, \alpha\beta Ax, \dots, \alpha\beta^{m-1}A^{m-1}x \} \\ &= \mathcal{K}_m(x, A) \end{aligned}$$

□

Korollar 3.5 (Translationsinvarianz). *Es gilt $\mathcal{K}_m(x, A - \gamma I) = \mathcal{K}_m(x, A)$ für alle $\gamma \in \mathbb{R}^n$*

Beweis . Induktion über die Dimension des Krylovraumes $\mathcal{K}_m(x, A)$:

Induktionsvoraussetzung:

Für $\dim \mathcal{K}_m(x, A) = 1$ ist $\mathcal{K}_1(x, A - \gamma I) = \text{span} \{x\} = \mathcal{K}_1(x, A)$

Induktionsannahme:

Gelte $\mathcal{K}_m(x, A - \gamma I) = \mathcal{K}_m(x, A)$ für ein $m > 1$.

Induktionsschluß:

Betrachte

$$\begin{aligned} \mathcal{K}_{m+1}(x, A - \gamma I) &= \text{span} \{x, \dots, (A - \gamma I)^m x\} \\ &= \text{span} \{ \mathcal{K}_m(x, A - \gamma I), (A - \gamma I)^m x \} \\ &= \text{span} \{ \mathcal{K}_m(x, A), (A - \gamma I)^m x \} \quad \text{nach I. Annahme} \end{aligned}$$

Nach der binomischen Formel gilt

$$(3.9) \quad (A - \gamma I)^m x = \sum_{i=0}^m (-\gamma)^{(m-i)} \binom{m}{i} A^i x$$

Da $\binom{m}{m} = 1 \neq 0$ ist, gilt $(A - \gamma I)^m \in \mathcal{K}_{m+1}(x, A)$, und es folgt:

$$(3.10) \quad \mathcal{K}_{m+1}(x, A - \gamma I) = \text{span} \{ \mathcal{K}_m(x, A), (A - \gamma I)^m x \} \subseteq \mathcal{K}_{m+1}(x, A)$$

Aus (3.9) folgt außerdem auch

$$A^m x = (A - \gamma I)^m x - \underbrace{\sum_{i=0}^{m-1} (-\gamma)^{(m-i)} \binom{m}{i} A^i x}_{\in \mathcal{K}_m(x, A)}$$

Damit gilt nach Induktionsannahme:

$$A^m x \in \text{span} \{ (A - \gamma I)^m x, \mathcal{K}_m(x, A) \} = \mathcal{K}_{m+1}(x, A - \gamma I)$$

Es folgt

$$(3.11) \quad \mathcal{K}_{m+1}(x, A) = \text{span} \{ \mathcal{K}_m(x, A), A^m x \} \subseteq \mathcal{K}_{m+1}(x, A - \gamma I)$$

Aus (3.10) und (3.11) folgt die Behauptung. □

Korollar 3.6 (Basistransformation). *Für jede reguläre Matrix P ist $\mathcal{K}_m(Px, PAP^{-1}) = P\mathcal{K}_m(x, A)$.*

Beweis . Es gilt

$$\begin{aligned} y \in \mathcal{K}_m(Px, PAP^{-1}) &\Leftrightarrow y = \sum_{i=0}^{m-1} \alpha_i (PAP^{-1})^i Px \quad \text{für } \alpha_i \in \mathbb{R} \\ &= \sum_{i=0}^{m-1} \alpha_i PA^i x = P \sum_{i=0}^{m-1} \alpha_i A^i x \\ &\Leftrightarrow y \in P\mathcal{K}_m(x, A) \end{aligned}$$

□

Aufgrund der Tatsache, dass jedes $y \in \mathcal{K}_m(x, A)$ beschrieben werden kann als

$$y = \sum_{i=0}^{m-1} \alpha_i A^i x = p(A)x \quad \text{für } p \in \mathcal{P}_{m-1} \quad \text{und } \alpha_i \in \mathbb{R},$$

existiert offensichtlich ein Zusammenhang zwischen dem Raum $\mathcal{K}_m(x, A)$ und Polynomen $p \in \mathcal{P}_{m-1}$. Es gilt sogar folgende Aussage:

Lemma 3.5. *Sei zu $m \in \mathbb{N}$ und $x \in \mathbb{R}^n$ $\dim \mathcal{K}_m(x, A) = m$. Dann ist die Abbildung $\phi : \mathcal{P}_{m-1} \rightarrow \mathcal{K}_m(x, A)$ mit $\phi(p) \mapsto p(A)x$ ein Isomorphismus.*

Beweis . Aufgrund der Definition von ϕ ist ϕ offensichtlich ein surjektiver Homomorphismus. Da außerdem $\dim \mathcal{K}_m(x, A) = m = \dim \mathcal{P}_{m-1}$ gilt, folgt auch die Injektivität von ϕ . \square

Den wichtigen Zusammenhang zwischen Polynomen und den aus einem Krylovraum berechneten Ritz-Paaren (μ_i, z^i) zeigt das nächste Lemma.

Lemma 3.6. *Sei zu $A = A^T$ und $x \in \mathbb{R}^n$ der Krylovraum $\mathcal{K}_m(x, A)$ mit $\dim \mathcal{K}_m(x, A) = m$ gegeben. Dann gilt für die Rayleigh-Ritz-Paare (μ_k, z^k) und für alle Polynome $p \in \mathcal{P}_{m-1}$:*

$$p(A)x \perp z^k \Leftrightarrow p(\mu_k) = 0 \quad \text{für alle } k \in \{1, \dots, m\}$$

Beweis . Definiere zunächst die beiden Mengen

$$S_k = \{y \in \mathcal{K}_m(x, A) \mid z^k \perp y\} \text{ und } N_k = \{p(A)x \mid p \in \mathcal{P}_{m-1}, p(\mu_k) = 0\}.$$

Nach Lemma 3.5 existiert für alle $y \in \mathcal{K}_m(x, A)$ ein $p \in \mathcal{P}_{m-1}$, so dass $y = p(A)x$. Also ist zu beweisen: $S_k = N_k$.

Zeige zuerst $N_k \subseteq S_k$. Sei dazu y aus N_k , das heißt $y = p(A)x$ mit $p \in \mathcal{P}_{m-1}$ und $p(\mu_k) = 0$. p ist also von der Form $p(\xi) = (\xi - \mu_k)q(\xi)$ wobei $q \in \mathcal{P}_{m-2}$ ist. Also gilt

$$(3.12) \quad q(A)x \in \mathcal{K}_{m-1}(x, A) \subset \mathcal{K}_m(x, A)$$

Es ist $z^k \perp y$ zu zeigen. Es gilt:

$$(3.13) \quad \begin{aligned} \langle z^k, y \rangle &= \langle z^k, p(A)x \rangle = \langle z^k, (A - \mu_k I)q(A)x \rangle = \langle (A - \mu_k I)z^k, q(A)x \rangle && \text{da } A = A^T \\ &= \langle r(z^k), q(A)x \rangle = 0 && \text{nach (3.12) und Lemma 3.4} \end{aligned}$$

Somit folgt:

$$(3.14) \quad N_k \subseteq S_k$$

Aufgrund der Definition von N_k ist

$$N_k = \{(A - \mu_k I)q(A)x \mid q \in \mathcal{P}_{m-2}\} = (A - \mu_k I)\mathcal{K}_{m-1}(x, A).$$

Also ist N_k ein $m - 1$ dimensionaler Unterraum von $\mathcal{K}_m(x, A)$, da $\dim \mathcal{K}_m(x, A) = m$ nach Voraussetzung. Sei $Z_k := \text{span}\{z^k\}$, dann ist $S_k = \mathcal{K}_m(x, A) \cap Z_k^\perp$ ein Unterraum von $\mathcal{K}_m(x, A)$. Weil $\dim \mathcal{K}_m(x, A) = m$ gilt, folgt mit der Dimensionsformel

$$\dim Z_k + \dim S_k = 1 + \dim S_k = m$$

Also ist S_k ebenfalls $m - 1$ dimensionaler Unterraum des $\mathcal{K}_m(x, A)$. Mit (3.14) folgt $S_k = N_k$ und somit die Behauptung. \square

Für Schritt 1 des Rayleigh-Ritz-Verfahrens, welcher (wie im nächsten Abschnitt erklärt) auf der Basis von Krylovräumen stattfinden wird, benötigt man eine ONB des Unterraums S^m . Das heißt, es wird nun ein Algorithmus gesucht, der zu $k \in \mathbb{N}$ für $\mathcal{K}_k(x, A) = \text{span}\{x, Ax, \dots, A^{k-1}x\}$ eine ONB $\{q_1, \dots, q_k\}$ mit $\mathcal{K}_k(x, A) = \text{span}\{q_1, \dots, q_k\}$ produziert. Der erste Algorithmus beruht auf der Definition der Krylovräume und dem Gram-Schmidt-Orthogonalisierungsverfahren (vergleiche [7]). Er wird im Folgenden beschrieben, wobei $\dim \mathcal{K}_m(x, A) = m$ vorausgesetzt sei.

Algorithmus 3.1.

```

für k = 1 bis m
    berechne  $q := A^{k-1}x$ 
     $q = q - \sum_{j=1}^{k-1} \langle q, q_j \rangle q_j$    Orthogonalisierung gegen  $q_1, \dots, q_{k-1}$ 
    falls  $q = 0$    Abbruch    $q_k$  wären dann linear abhängig
     $q_k := q / \|q\|$    Normalisierung
stop
    
```

Bemerkung 3.4. *Unter der Annahme, dass der Startvektor x nicht bereits in einem invarianten Unterraum liegt, ist der Test, ob $q = 0$ gilt, in exakter Arithmetik nicht nötig, da der Krylovraum nach Voraussetzung Dimension m hat. Somit müssten die Vektoren q_1, \dots, q_m linear unabhängig sein. Da aber aufgrund von Rundungsfehlern keine exakte Arithmetik zu erwarten ist, muß diese Abfrage in $\|q\| < \epsilon$ mit $\epsilon \in \mathbb{R}_{>0}$ geändert werden ($q = 0 \Leftrightarrow \|q\| = 0$).*

Der Grund dafür, dass eine ONB berechnet wird, ist unter anderem die Aussage von Korollar 3.3. Diese besagt, dass die Matrix $Q^T A Q$ zu einem symmetrischen A selbst symmetrisch ist. Im nachfolgenden Lemma wird gezeigt, dass für eine symmetrische Matrix A in Verbindung mit Krylovräumen diese Matrix sogar tridiagonal ist.

Lemma 3.7. *Sei $\dim \mathcal{K}_m(x, A) = m$, A symmetrisch und sei $Q_m = [q_1, \dots, q_m]$ die Matrix, die als Spaltenvektoren die obige ONB von $\mathcal{K}_m(x, A)$ enthält. Dann hat die Matrix $Q_m^T A Q_m$ strikte Tridiagonalgestalt.*

Beweis . Es gilt

$$(3.15) \quad (Q_m^T A Q_m)_{ij} = \langle A q_j, q_i \rangle$$

Für $q_j \in \mathcal{K}_j(x, A) = \text{span}\{q_1, \dots, q_j\}$ gilt nach Bemerkung (3.3): $A q_j \in \mathcal{K}_{j+1}(x, A) = \text{span}\{q_1, \dots, q_{j+1}\}$. Da die q_1, \dots, q_m eine ONB bilden, gilt $q_i \perp \mathcal{K}_{j+1}(x, A)$ für alle $i \in \{j+2, \dots, m\}$.

Also gilt auch $q_i \perp Aq_j$. Damit ist

$$\langle Aq_j, q_i \rangle = 0 \quad \text{für alle } i \in \{j + 2, \dots, m\}.$$

Da nach Korollar 3.3 die Matrix $Q_m^T A Q_m$ symmetrisch ist, folgt auch

$$\langle Aq_j, q_i \rangle = 0 \quad \text{für alle } i \in \{1, \dots, j - 2\}.$$

Damit ist die Tridiagonalgestalt gezeigt. Zu beweisen ist noch

$$\langle Aq_j, q_i \rangle \neq 0 \quad \text{für } j = i - 1$$

Annahme: Es gelte $\langle Aq_j, q_i \rangle = 0$ für alle $j \leq i - 1$.

Es ist

$$A^{i-1}x \in AK_{i-1}(x, A) = \text{span} \{Aq_1, \dots, Aq_{i-1}\}$$

Wegen der Annahme $\langle Aq_j, q_i \rangle = 0$ für alle $j \leq i - 1$ folgt $q_i \perp A^{i-1}x$. Außerdem gilt $q_i \perp \mathcal{K}_{i-1}(x, A)$, so dass insgesamt folgt:

$$(3.16) \quad q_i \perp \text{span} \{\mathcal{K}_{i-1}(x, A), A^{i-1}x\} = \mathcal{K}_i(x, A)$$

Da $q_i \in \mathcal{K}_i(x, A)$ nach Definition gilt, ist (3.16) ein Widerspruch zur Annahme. Damit ist auch die strikte Tridiagonalgestalt gezeigt. \square

3.3 Symmetrisches Lanczos-Verfahren

Die vorangegangenen Überlegungen beschäftigten sich mit dem Rayleigh-Ritz-Verfahren allgemein und den Krylovräumen, die als Unterräume zur Approximation mit dem RR-Verfahren verwendet werden können. Wie erhält man aus diesen Informationen einen Algorithmus, der RR-Verfahren und Krylovräume vereint und möglichst wenig Aufwand benötigt? Diese Frage führt nun zu dem (später für das zu lösende Singulärwertproblem noch leicht abgewandelten) *symmetrischen Lanczos-Algorithmus*.

Das Ziel dieses Algorithmus ist es, die Tridiagonalgestalt der Matrix $Q^T A Q$ auszunutzen und diese gleichzeitig zu berechnen. Das Besondere wird sein, dass dafür nicht nur ein Krylovraum benutzt wird, sondern fortlaufend zur gewünschten Größe der Ritz-Matrix eine Folge von Krylovräumen und somit eine Folge von tridiagonalen Matrizen produziert wird. Der Trick hierbei ist, dass nicht in jedem Schritt neue Matrizen berechnet werden, sondern die Struktur der Krylovräume so ausgenutzt wird, dass immer nur eine neue Zeile und Spalte an die vorige Matrix angehängt werden muß.

Sei dazu angenommen die Matrix $T_m := Q_m^T A Q_m$ aus Lemma 3.7 sei bereits berechnet. Dann ist T_m von der Form:

Zu Beginn dieses Kapitels wurde bemerkt, dass der Lanczos-Algorithmus Krylovräume benutzt, um außer der ONB auch gleichzeitig die Einträge der tridiagonalen Matrizen T_k , welche $k \times k$ Hauptuntermatrizen von T_m sind, zu bestimmen. Das heißt, es wird das RR-Verfahren mit Krylovräumen durchgeführt. Also wird nach einem Algorithmus gesucht, der parallel zur Berechnung der ONB eines gegebenen Krylovraums auch die Werte α_i, β_j mit $i \in \{1, \dots, k\}, j \in \{1, \dots, k-1\}$ von T_k berechnet. Dazu dient folgendes Korollar.

Korollar 3.7. *Es gilt für $i < k$*

$$(3.19) \quad Aq_i - q_{i-1}\beta_{i-1} - q_i\alpha_i = q_{i+1}\beta_i$$

Definition 3.4. *Setze $r_i := Aq_i - q_{i-1}\beta_{i-1} - q_i\alpha_i = \beta_i q_{i+1}$*

Beweis (von Korollar 3.7). Wie zu Beginn dieses Abschnitts beschrieben gilt $T_k = Q_k^T A Q_k$, und somit $\alpha_i = \langle Aq_i, q_i \rangle$ und $\beta_i = \langle Aq_i, q_{i+1} \rangle$ bzw. $\beta_{i-1} = \langle Aq_{i-1}, q_i \rangle$. Betrachtet man die zweite und dritte Zeile von Algorithmus 3.2, so gilt:

$$\begin{aligned} q &:= Aq_i - \langle Aq_i, q_{i-1} \rangle q_{i-1} - \langle Aq_i, q_i \rangle q_i \\ &= Aq_i - \beta_{i-1} q_{i-1} - \alpha_i q_i = r_i \end{aligned} \quad \text{Definition 3.4}$$

$$(3.20) \quad q_{i+1} = r_i / \|r_i\|$$

Nach der Definition 3.4 von r_i und (3.20) bleibt zu zeigen: $\beta_i = \|r_i\|$. Da in Algorithmus 3.2 für $r_i = 0$ abgebrochen wird, kann $\|r_i\| \neq 0$ angenommen werden. Betrachte zunächst

$$\begin{aligned} (3.21) \quad \langle Aq_i - r_i, r_i \rangle &= \langle \beta_{i-1} q_{i-1} + \alpha_i q_i, Aq_i - q_{i-1}\beta_{i-1} - q_i\alpha_i \rangle \\ &= \beta_{i-1} \langle q_{i-1}, Aq_i \rangle - \beta_{i-1}^2 - 0 + \alpha_i \langle q_i, Aq_i \rangle - 0 - \alpha_i^2 \quad \text{da } q_i \text{ ONB} \\ &= \beta_{i-1}^2 - \beta_{i-1}^2 + \alpha_i^2 - \alpha_i^2 = 0 \quad \text{Definition von } \alpha_i, \beta_i \end{aligned}$$

Aus (3.21) folgt:

$$(3.22) \quad \|r_i\|^2 = \langle r_i, r_i \rangle = \langle Aq_i - r_i, r_i \rangle + \langle r_i, r_i \rangle = \langle Aq_i, r_i \rangle$$

Mit (3.22) und (3.20) folgt für β_i :

$$(3.23) \quad \beta_i = \langle Aq_i, q_{i+1} \rangle = \langle Aq_i, \frac{r_i}{\|r_i\|} \rangle = \frac{1}{\|r_i\|} \langle Aq_i, r_i \rangle = \|r_i\|$$

□

Der Vektor r_i hat eine besondere Bedeutung, denn es gilt aufgrund der Gleichung aus Korollar 3.7 die folgende Aussage:

Korollar 3.8.

$$AQ_k - Q_k T_k = [0, \dots, 0, r_k]$$

Aus Korollar 3.8 geht insbesondere hervor:

$$(3.24) \quad R(T_k) = r_k e_k^T r_k \quad \text{mit} \quad r_k = \|q_k\| q_{k+1}$$

wobei e_k der k -te Einheitsvektor sei. Das heißt, es ist genau dann ein invarianter Unterraum erreicht, wenn r_k der Nullvektor ist, also insbesondere wenn $\|r_k\| = 0$ gilt. Mit Hilfe von (3.23), $\alpha_i = \langle Aq_i, q_i \rangle$ und der Gleichung aus Korollar (3.7) kann somit ein Algorithmus aufgestellt werden, der rekursiv eine ONB von $\mathcal{K}_m(x, A)$ und die Tridiagonaleinträge α_i, β_i berechnet.

Algorithmus 3.3. Symmetrischer Lanczos-Algorithmus

Setze $q_0 = 0, q_1 = x/\|x\|, \beta_0 = 0, \alpha_1 = \langle Aq_1, q_1 \rangle$
für $k = 1$ **bis** $m - 1$
berechne $r_k := Aq_k - \beta_{k-1}q_{k-1} - \alpha_k q_k$ Orthogonalisierung gegen q_k, q_{k-1}
 falls $\|r_k\| < \epsilon \|Aq_k\|$ **Abbruch** q_k wären dann linear abhängig
 $\beta_k := \|r_k\|$
 $q_{k+1} := r_k/\beta_k$ Normalisierung
 $\alpha_{k+1} := \langle Aq_{k+1}, q_{k+1} \rangle$
stop

Bemerkung 3.5. Sei A symmetrisch und sei $\dim \mathcal{K}_m(x, A) = m$. Der oben erklärte Lanczos-Algorithmus benötigt zur Durchführung m Matrix-Vektor-Multiplikationen und $2m$ Skalarprodukte. Als Startvektor kann ein Zufallsvektor gewählt werden, oder z.B $x = (1, \dots, 1) \in \mathbb{R}^n$.

Die Abfrage $\|r_k\| < \epsilon$ wurde zu $\|r_k\| < \epsilon \|Aq_k\|$ verstärkt, um das ϵ unabhängig von der Matrix A wählen zu können. Durch den Vergleich mit $\|Aq_k\|$ bleibt das Abbruchkriterium skalierungsinvariant.

4 Konvergenzbetrachtungen

In diesem Kapitel werden zuerst allgemeine Fehlerabschätzungen der mit dem Lanczos-Verfahren berechneten Ritz-Paare entwickelt, und diese dann danach für die speziellen Matrizen mit exponentiell abfallenden Eigenwerten untersucht.

4.1 Allgemeine Fehleranalyse für das Lanczosverfahren

Sei im Folgenden:

$A \in \mathbb{R}^{n \times n}$, $A = A^T$ mit den Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ und den zugehörigen Eigenvektoren x^1, \dots, x^n .

Definiere $X^j := \text{span} \{ x^i \mid 1 \leq i \leq j \}$ und $X^{j\perp} := \text{span} \{ x^i \mid j < i \leq n \}$.

Sei x Startvektor des Lanczos-Verfahrens mit $\|x\| = 1$ und $\dim \mathcal{K}_m(x, A) = m$, $m \leq n$ für ein vorgegebenes, festes m . Sei $Q = [q_1, \dots, q_m] \in \mathbb{R}^{n \times m}$ die im Lanczos-Verfahren berechnete orthonormale Matrix deren Spalten die Lanczos-Vektoren bilden.

Seien μ_1, \dots, μ_m die Eigenwerte der Ritz-Matrix $B = Q^T A Q \in \mathbb{R}^{m \times m}$ (also μ_1, \dots, μ_m Ritz-Werte) mit zugehörigen Eigenvektoren y^1, \dots, y^m .

Sei ψ_j der Winkel zwischen x und X^j . Sei $x^{(j)} \in X^j$ so, dass $x^{(j)} \cos \psi_j$ die orthogonale Projektion von x auf X^j und $x^{\perp(j)} \sin \psi_j$ mit $x^{\perp(j)} \in X^{j\perp}$ die orthogonale Projektion von x auf $X^{j\perp}$ ist. Sei außerdem $\|x^{(j)}\|_2 = \|x^{\perp(j)}\| = 1$, so dass gilt:

$$x = x^{(j)} \cos \psi_j + x^{\perp(j)} \sin \psi_j$$

wobei $\psi_j := \angle(x, X^j)$

Es wird zunächst eine Fehlerabschätzung der Eigenwerte entwickelt, in die die berechneten Ritz-Werte einfließen.

4.1.1 Fehlerabschätzung der Eigenwerte mit Hilfe der Ritz-Werte

Wie in den vorangegangenen Kapiteln beschrieben, ist der Rayleigh-Quotient ein wichtiges Hilfsmittel um Eigenwerte zu berechnen und zu approximieren. Außerdem wurde der Zusammenhang zwischen Krylovraum $\mathcal{K}_m(x, A)$, aus dem mit Hilfe des Lanczos-Verfahrens die Eigenpaare approximiert werden, und Polynomraum \mathcal{P}_{m-1} herausgearbeitet. Der erste Schritt in Richtung einer Abschätzung des Fehlers zwischen Eigenwerten und Ritz-Werten beschreibt den Abstand des j -ten Eigenwerts zu dem Rayleigh-Quotienten $R(p(A)x, A)$, für ein $p \in \mathcal{P}_{m-1}$. Weil $p(A)x$ ein Element des Krylovraums

$\mathcal{K}_m(x, A)$ ist, stellt das folgende Lemma eine erste obere Schranke des Approximationsfehlers der Eigenwerte dar.

Lemma 4.1. *Sei $p \in \mathcal{P}_{m-1}$ mit $p(\lambda_j) \neq 0$ und $\cos \angle(x, X^j) \neq 0$. Dann gilt:*

$$R(p(A)x, \lambda_j I - A) = \frac{\langle (\lambda_j I - A)p(A)x, p(A)x \rangle}{\|p(A)x\|} \leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j) \|p(A)x^{\perp(j)}\|}{\cos \angle(x, X^j) p(\lambda_j)} \right]^2$$

Bemerkung 4.1. *Lemma 4.1 wird als obere Schranke für den Fehler zwischen den ersten j Ritz-Werten und Eigenwerten dienen. Indem der Rayleigh-Quotient $R(p(A)x, \lambda_j I - A)$ nach unten gegen den Abstand der Ritz-Werte μ_j zu den Eigenwerten λ_j abgeschätzt wird, kann mit Hilfe von Lemma 4.1 der Fehler eingegrenzt werden.*

Außerdem ist die Bedingung $\cos \angle(x, X^j) \neq 0$ hervorzuheben. Würde $\cos \angle(x, X^j) = 0$ gelten, so entspräche dies $x \perp X^j$. Das heißt, da keine Anteile von X^j in x enthalten wären, ist ein solcher Startvektor ungeeignet zur Berechnung von X^j .

Beweis von Lemma 4.1. Aus $x = X^j \cos \psi_j + X^{j\perp} \sin \psi_j$ folgt $p(A)x = p(A)X^j \cos \psi_j + p(A)X^{j\perp} \sin \psi_j$. Da X^j und $X^{j\perp}$ von Eigenvektoren aufgespannt werden und somit invariant unter A sind, gilt:

$$(4.1) \quad p(A)X^j \in X^j \quad \text{und} \quad p(A)X^{j\perp} \in X^{j\perp}$$

Der Beweis des Lemmas erfolgt in mehreren Schritten. Zunächst wird die linke Seite, also $R(p(A)x, (\lambda_j I - A))$, ausgerechnet. Für den Rayleigh-Quotienten gilt:

$$\begin{aligned} R(p(A)x, (\lambda_j I - A)) &= \frac{\langle (\lambda_j I - A)p(A)x, p(A)x \rangle}{\|p(A)x\|^2} \\ &= \frac{\overbrace{\langle (\lambda_j I - A)(p(A)X^j \cos \psi_j + p(A)X^{j\perp} \sin \psi_j), p(A)X^j \cos \psi_j + p(A)X^{j\perp} \sin \psi_j \rangle}^{=: Z}}{\|p(A)x\|^2} \end{aligned}$$

Wegen $X^j \perp X^{j\perp}$ folgt:

$$(4.2) \quad \begin{aligned} Z &= \langle (\lambda_j I - A)p(A)X^j \cos \psi_j, p(A)X^j \cos \psi_j \rangle \\ &\quad + \langle (\lambda_j I - A)p(A)X^{j\perp} \sin \psi_j, p(A)X^{j\perp} \sin \psi_j \rangle \\ &= \langle (\lambda_j I - A)p(A)X^j, p(A)X^j \rangle \cos^2 \psi_j + \langle (\lambda_j I - A)p(A)X^{j\perp}, p(A)X^{j\perp} \rangle \sin^2 \psi_j \end{aligned}$$

Das heißt, es gilt

$$(4.3) \quad R(p(A)x, (\lambda_j I - A)) =$$

$$\frac{\langle (\lambda_j I - A)p(A)x^{(j)}, p(A)x^{(j)} \rangle \cos^2 \psi_j + \langle (\lambda_j I - A)p(A)x^{\perp(j)}, p(A)x^{\perp(j)} \rangle \sin^2 \psi_j}{\|p(A)x\|^2}$$

Um den Zähler in (4.3) weiter auszurechnen und abzuschätzen wird nun wegen (4.1) das Skalarprodukt $\langle z, (\lambda_j I - A)z \rangle$ für Vektoren z aus X^j und $X^{j\perp}$ betrachtet.

Für alle $z = \sum_{i=1}^j a_i x^i \in X^j$ gilt:

$$\begin{aligned} \langle z, (\lambda_j I - A)z \rangle &= \langle z, \lambda_j z - A \left(\sum_{i=1}^j a_i x^i \right) \rangle = \langle z, \lambda_j \sum_{i=1}^j a_i x^i - \sum_{i=1}^j a_i \lambda_i x^i \rangle \\ &= \langle z, \sum_{i=1}^j (\lambda_j - \lambda_i) a_i x^i \rangle = \sum_{i=1}^j (\lambda_j - \lambda_i) |a_i|^2 \|x^i\|^2 \\ &\leq 0 \qquad \qquad \qquad \text{da } \lambda_i \geq \lambda_j \text{ für } i \leq j \end{aligned}$$

Damit folgt für den ersten Teil des Zählers

$$(4.4) \quad \langle p(A)x^{(j)}, (\lambda_j I - A)p(A)x^{(j)} \rangle \cos^2 \psi_j \leq 0, \quad \text{da } p(A)x^{(j)} \in X^j.$$

Für alle $z = \sum_{i=j+1}^n a_i x^i \in X^{j\perp}$ gilt:

$$\begin{aligned} \langle z, (\lambda_j I - A)z \rangle &= \sum_{i=j+1}^n (\lambda_j - \lambda_i) |a_i|^2 \|x^i\|^2 \\ &\leq \sum_{i=j+1}^n (\lambda_j - \lambda_n) |a_i|^2, \qquad \text{da } \lambda_n \leq \lambda_i \text{ für alle } i, \text{ also} \\ &= (\lambda_j - \lambda_n) \left(\sum_{i=j+1}^n |a_i|^2 \right) \\ &= (\lambda_j - \lambda_n) \|z\|^2 \end{aligned}$$

Wegen $p(A)x^{\perp(j)} \in X^{j\perp}$ folgt also für den zweiten Teil des Zählers von (4.3):

$$(4.5) \quad \langle p(A)x^{\perp(j)}, (\lambda_j I - A)p(A)x^{\perp(j)} \rangle \sin^2 \psi_j \leq (\lambda_j - \lambda_n) \|p(A)x^{\perp(j)}\|^2 \sin^2 \psi_j$$

Insgesamt folgt aus (4.3) bis (4.5) die Ungleichung:

$$(4.6) \quad R(p(A)x, \lambda_j I - A) \leq (\lambda_j - \lambda_n) \left[\frac{\|p(A)x^{\perp(j)}\| \sin \psi_j}{\|p(A)x\|} \right]^2$$

Für den Startvektor $x = \sum_{i=1}^n a_i x^i \in X^n := \text{span}\{x^1, \dots, x^n\}$ gilt, da die Eigenvektoren eine ONB bilden:

$$x = \sum_{i=1}^n \langle x, x^i \rangle x^i, \text{ also } p(A)x = \sum_{i=1}^n p(\lambda_i) \langle x, x^i \rangle x^i$$

Da vorausgesetzt wurde, dass der Startvektor x normiert ist, folgt mit

$$\cos^2 \angle(x, x^i) = |\langle x, x^i \rangle|^2 / (\|x\| \|x^i\|)^2 = \langle x, x^i \rangle^2$$

die Abschätzung

$$(4.7) \quad \|p(A)x\|^2 = \sum_{i=1}^n p(\lambda_i)^2 \cos^2 \angle(x, x^i) \geq p^2(\lambda_j) \cos^2 \angle(x, x^j)$$

Somit folgt aus (4.6) und (4.7) die Behauptung:

$$R(p(A)x, \lambda_j I - A) \leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j) \|p(A)x^{\perp(j)}\|}{\cos \angle(x, x^j) p(\lambda_j)} \right]^2 \text{ für } j \in \{1, \dots, m\}$$

□

Damit wurde eine obere Schranke für den Abstand des Rayleigh-Quotienten eines Vektors aus dem betrachteten Krylovraum und λ_j entwickelt. Das nächste Lemma liefert nun eine untere Schranke des Abstandes von λ_j zu dem betrachteten Rayleigh-Quotienten. Zusammen mit Lemma 4.1 ergibt sich dann ein erster Überblick über den absoluten Fehler des Ritz-Werts gegenüber echtem Eigenwert.

Lemma 4.2. *Sei $p \in \mathcal{P}_{m-1}$ mit $p(\lambda_j) \neq 0$, und sei außerdem $p(\mu_i) = 0$ für $i \in \{1, \dots, j-1\}$ und $j \leq m$. Dann gilt:*

$$0 \leq \lambda_j - \mu_j \leq R(p(A)x, \lambda_j I - A) \quad \text{für } j \in \{1, \dots, m\}$$

Beweis. Da die orthonormale Matrix Q aus den Spaltenvektoren q_1, \dots, q_m besteht und der Krylovraum $\mathcal{K}_m(x, A)$ von q_1, \dots, q_m aufgespannt wird, ist jeder Vektor $y \in \mathcal{K}_m(x, A)$ von der Form $Qv \in \mathbb{R}^n$, für ein $v \in \mathbb{R}^m$. Definiere $W^k := \text{span}\{y^1, \dots, y^k\}$, $1 \leq k \leq m$, $W^0 := \{0\}$, wobei y^1, \dots, y^k Eigenvektoren zu μ_1, \dots, μ_k sind, und $V^k := \text{span}\{z^1, \dots, z^k\}$, $1 \leq k \leq m$, $V^0 := \{0\}$, wobei $z^i = Qy^i$ die Ritz-Vektoren sind. Es gilt $V^k = QW^k$ und für alle $y = Qv \in \mathcal{K}_m(x, A)$ folgt, da Q orthonormal ist:

$$(4.8) \quad R(y, A) = \frac{\langle Ay, y \rangle}{\|y\|^2} = \frac{\langle A(Qv), (Qv) \rangle}{\|Qv\|^2} = \frac{\langle \overbrace{(Q^T A Q)}^{=B} v, v \rangle}{\|v\|^2} = R(v, B)$$

Der linke Teil des Lemmas folgt für $j \in \{1, \dots, m\}$ mit:

$$\begin{aligned}
 \lambda_j &= \max_{\substack{V \subseteq \mathbb{R}^n \\ \dim V = j}} \min_{x \neq 0 \in V} R(x, A) && \text{nach Lemma 2.1 (Min-Max-Theorem)} \\
 &\geq \max_{\substack{V \subseteq \mathcal{K}_m(x, A) \\ \dim V = j}} \min_{x \neq 0 \in V} R(x, A) && \text{da } \mathcal{K}_m(x, A) \subseteq \mathbb{R}^n \\
 &= \max_{\substack{W \subseteq \mathbb{R}^m \\ \dim W = j}} \min_{y \neq 0 \in W} R(y, B) && \text{nach (4.8)} \\
 &= \mu_j && \text{nach Lemma 2.1 (Min-Max-Theorem)}
 \end{aligned}$$

Nach Lemma 3.6 gilt für die Ritzpaare $(\mu_1, z^1), \dots, (\mu_m, z^m)$ aus $\mathcal{K}_m(x, A)$ und $q \in \mathcal{P}_{m-1}$:

$$q(A)x \perp z^i \Leftrightarrow q(\mu_i) = 0$$

Da nach Voraussetzung $p(\mu_i) = 0$ ist für alle $i \in \{1, \dots, j-1\}$, folgt $p(A)x \perp z^i$, also $p(A)x \in QW^{(j-1)\perp}$

Weiter ergibt sich mit der Definition des Rayleigh-Quotienten:

$$\begin{aligned}
 \mu_j &= \max_{y \neq 0 \in W^{(j-1)\perp}} R(y, B) && \text{nach Lemma 2.2} \\
 &= \max_{y \neq 0 \in W^{(j-1)\perp}} \frac{\langle A(Qy), Qy \rangle}{\langle Qy, Qy \rangle} && \text{da } Q \text{ orthonormal ist, gilt } \langle Qy, Qy \rangle = \langle y, y \rangle \\
 &= \max_{x \neq 0 \in QW^{(j-1)\perp}} R(x, A) && \text{nach (4.8)} \\
 &\geq R(p(A)x, A) && \text{da } p(A)x \in QW^{j-1\perp}
 \end{aligned}$$

Nun folgt aus

$$R(p(A)x, \lambda_j I - A) = \lambda_j - R(p(A)x, A) \geq \lambda_j - \mu_j \geq 0$$

die Behauptung. □

Aus Lemma 4.2 und Lemma 4.1 folgt insgesamt:

$$(4.9) \quad 0 \leq \lambda_j - \mu_j \leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j) \|p(A)x^{\perp(j)}\|}{\cos \angle(x, x^{(j)}) p(\lambda_j)} \right]^2 \quad \text{für } j \in \{1, \dots, m\}$$

Damit ist eine erste Fehlerabschätzung der mit dem Lanczos-Verfahren approximierten Eigenwerte entstanden. Um diese noch zu verbessern, wird die rechte Seite der Ungleichung (4.9) genauer betrachtet. Der Winkel zwischen (zufällig gewähltem) Startvektor x

und Eigenraum X^j ist nicht beeinflussbar, aber die Ungleichung gilt für beliebige Polynome $p \in \mathcal{P}_{m-1}$, welche die Voraussetzungen $p(\mu_i) = 0$ für $i \in \{1, \dots, j\}$ und $p(\lambda_j) \neq 0$ erfüllen. Um eine verbesserte Fehlerabschätzung für die Ritz-Werte zu erhalten, kann $p \in \mathcal{P}_{m-1}$ so gewählt werden, dass

$$(4.10) \quad \frac{\|p(A)x^{\perp(j)}\|}{p(\lambda_j)}$$

möglichst klein ist. Da die Ritz-Werte μ_1, \dots, μ_j Nullstellen von p sind, ist p von der Form:

$$(4.11) \quad p(t) = q(t) \prod_{i=1}^{j-1} (\mu_i - t) \quad \text{mit } q \in P_{m-j}, \quad \text{also}$$

$$p(\lambda_j) = q(\lambda_j) \prod_{i=1}^{j-1} (\mu_i - \lambda_j)$$

Um den obigen Bruch (4.10) zu minimieren, betrachte zuerst den Zähler. Es gilt:

$$(4.12) \quad \begin{aligned} \|p(A)x^{\perp(j)}\| &= \|q(A) \prod_{i=1}^{j-1} (\mu_i I - A)x^{\perp(j)}\| \\ &\leq \|q(A)x^{\perp(j)}\| \left\| \prod_{i=1}^{j-1} (\mu_i I - A) \right\| \\ &\leq \|q(A)x^{\perp(j)}\| \prod_{i=1}^{j-1} \|\mu_i I - A\| \end{aligned}$$

Es bleibt noch $\|\mu_i I - A\|$ abzuschätzen.

Dazu betrachte zunächst $\|(\mu_i I - A)v\|$ für einen Vektor v .

Für alle $v = \sum_{k=1}^n c_k x^k \in \mathbb{R}^n$ und alle μ_i mit $1 \leq i \leq j-1$ ist

$$\begin{aligned} \|(\mu_i I - A)v\| &= \left\| \mu_i \sum_{k=1}^n c_k x^k - \sum_{k=1}^n c_k \lambda_k x^k \right\| \\ &= \|(\mu_i - \lambda_1)c_1 x^1 + \dots + (\mu_i - \lambda_n)c_n x^n\| \\ &= \sqrt{\sum_{k=1}^n (\mu_i - \lambda_k)^2 \langle c_k x^k, c_k x^k \rangle} \\ &= \sqrt{\sum_{k=1}^n (\mu_i - \lambda_k)^2 c_k^2} \end{aligned} \quad \text{da } \langle x^k, x^k \rangle = 1$$

$$\begin{aligned}
 &\leq \sqrt{\sum_{k=1}^n (\mu_i - \lambda_n)^2 c_k^2} && \mu_1, \dots, \mu_m \in [\lambda_n, \lambda_1] \text{ nach Satz 2.1} \\
 &= \sqrt{(\mu_i - \lambda_n)^2 \sum_{k=1}^n c_k^2} \\
 &= (\mu_i - \lambda_n) \|v\|
 \end{aligned}$$

Damit folgt für $\|\mu_i I - A\|$:

$$\|\mu_i I - A\| = \max_{\|v\|=1} \|\mu_i v - Av\| \leq \max_{\|v\|=1} (\mu_i - \lambda_n) \|v\| = \mu_i - \lambda_n$$

Für die abzuschätzende Norm $\|p(A)x^{\perp(j)}\|$ aus (4.10) folgt mit (4.12) und obiger Ungleichung:

$$(4.13) \quad \|p(A)x^{\perp(j)}\| \leq \|q(A)x^{\perp(j)}\| \prod_{i=1}^{j-1} (\mu_i - \lambda_n)$$

Mit Gleichung (4.9) und der Definition von p in (4.11) gilt somit:

$$(4.14) \quad 0 \leq \lambda_j - \mu_j \leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j) \prod_{i=1}^{j-1} (\mu_i - \lambda_n) \|q(A)x^{\perp(j)}\|}{\cos \angle(x, x^{(j)}) \prod_{i=1}^{j-1} (\mu_i - \lambda_j) q(\lambda_j)} \right]^2$$

Das neue Ziel zur Verbesserung der Fehlerschranke muß nun also lauten:

$$(4.15) \quad \text{Minimiere } \frac{\|q(A)x^{\perp(j)}\|}{|q(\lambda_j)|}$$

Da $x^{\perp(j)} \in X^{j\perp}$ ist, existieren a_{j+1}, \dots, a_n , mit $x^{\perp(j)} = \sum_{i=j+1}^n a_i x^i$, so dass $\|x^{\perp(j)}\| = 1$ gilt.

$$\begin{aligned}
 \|q(A)x^{\perp(j)}\| &= \left\langle \sum_{i=j+1}^n q(\lambda_j) a_i x^i, \sum_{i=j+1}^n q(\lambda_j) a_i x^i \right\rangle \\
 &= \sum_{i=j+1}^n q(\lambda_j)^2 |a_i|^2 \underbrace{\|x^i\|^2}_{=1} \\
 &\leq \max_{l \in \{j+1, \dots, n\}} q(\lambda_l)^2 \sum_{i=j+1}^n |a_i|^2
 \end{aligned}$$

$$\begin{aligned} &\leq \max_{t \in [\lambda_n, \lambda_{j+1}]} q(t)^2 \underbrace{\|x^{\perp(j)}\|^2}_{=1} && \text{es gilt } \sum_{i=j+1}^n |a_i|^2 = \|x^{\perp(j)}\|^2, \text{ da } x^i \text{ ONB} \\ &= \max_{t \in [\lambda_n, \lambda_{j+1}]} q(t)^2 \end{aligned}$$

Die Minimierungsaufgabe aus (4.15) ergibt sich nun also zu:

$$(4.16) \quad \text{Minimiere } \max_{t \in [\lambda_n, \lambda_{j+1}]} \frac{|q(t)|}{|q(\lambda_j)|} \text{ über alle } q \in P_{m-j}$$

Da zur Lösung dieser Aufgabe die später eingeführten *Tschebyscheff-Polynome* verwendet werden, welche besondere Eigenschaften auf dem Intervall $[-1, 1]$ besitzen, benötigt man zur Vereinfachung von (4.16) die folgenden Definitionen:

Definition 4.1. Sei $\phi : \mathbb{R} \rightarrow [-1, 1]$ gegeben durch:

$$\phi(t) = 1 - 2 \left(\frac{\lambda_{j+1} - t}{\lambda_{j+1} - \lambda_n} \right)$$

Bemerkung 4.2. Die Abbildung ϕ ist offensichtlich linear, und da λ_n auf -1 und λ_{j+1} auf 1 abgebildet wird, ist ϕ eine bijektive Abbildung von $[\lambda_n, \lambda_{j+1}]$ auf $[-1, 1]$.

Definition 4.2. Setze

$$\gamma_j := \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n}$$

Bemerkung 4.3. γ_j beschreibt den Abstand von λ_j zum Rest des Spektrums. Da $\gamma_j > 0$ nach Definition ist, gilt:

$$\phi(\lambda_j) = 1 - 2 \frac{\lambda_{j+1} - \lambda_j}{\lambda_{j+1} - \lambda_n} = 1 + 2\gamma_j > 1$$

Die Minimierungsaufgabe (4.16) ist mit $q = \pi \circ \phi$ nun äquivalent zu:

$$(4.17) \quad \begin{aligned} &\text{Minimiere } \max_{\xi \in [-1, 1]} \frac{|\pi(\xi)|}{|\pi(\xi_0)|} \text{ über alle } \pi \in P_{m-j} \\ &\text{mit } \xi_0 = \phi(\lambda_j) > 1 \end{aligned}$$

Um (4.17) zu lösen, werden folgende Polynome eingeführt:

Definition 4.3. Die *Tschebyscheff-Polynome* $T_m \in \mathcal{P}_{m-1}$ sind für alle $m \in \mathbb{N}_0$ definiert durch:

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_m(x) &= 2x T_{m-1}(x) - T_{m-2}(x) \end{aligned} \quad \text{für alle } x \in \mathbb{R}$$

Satz 4.1. Für $\xi_0 \in \mathbb{R}_{>1}$ gilt:

$$\min_{\pi \in \mathcal{P}_{m-j}} \max_{\xi \in [-1,1]} \frac{|\pi(\xi)|}{|\pi(\xi_0)|} \quad \text{ist gegeben durch } p(\xi) := \frac{T_k(\xi)}{T_k(\xi_0)}$$

wobei T_k das k -te Tschebyscheff-Polynom ist.

Beweis. Siehe [4], Satz 6.6.1. □

Bemerkung 4.4. Die Tschebyscheffpolynome sind klein innerhalb des Intervalls $[-1, 1]$ und groß außerhalb. Sie erfüllen deshalb die gesuchten Eigenschaften der Minimierungsaufgabe (4.17), da (wie in Bemerkung 4.3 und 4.2 gezeigt) $\phi(\lambda_j) > 1$ und $|\phi(t)| \leq 1$ für alle $t \in [\lambda_n, \lambda_{j+1}]$ gilt, so dass der Nenner in (4.17) wächst und der Zähler klein bleibt.

Lemma 4.3. Für die Tschebyscheff-Polynome aus Satz 4.1 gilt:

$$(4.18) \quad T_k(x) = \cos(k \arccos(x)) \quad \text{für } |x| < 1$$

Beweis. Siehe [1], Lemma 3.7. □

Sei nun p gewählt wie in Satz 4.1. Aus $\phi(t) \in [-1, 1]$ für alle $t \in [\lambda_n, \lambda_{j+1}]$ und Lemma 4.3 folgt insbesondere:

$$\max_{t \in [\lambda_n, \lambda_{j+1}]} |p(\phi(t))| \leq 1$$

Da außerdem nach Bemerkung 4.3 $\phi(\lambda_j) = 1 + 2\gamma_j$ gilt, folgt mit $q(t) := p(\phi(t))$ eingesetzt in Gleichung (4.14) nun der abschließende Satz.

Satz 4.2. Für die Ritz-Werte μ_1, \dots, μ_m des Lanczos-Verfahrens und Startvektor x gilt:

$$0 \leq \lambda_j - \mu_j \leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j)}{\cos \angle(x, x^j)} \frac{1}{T_{m-j}(1 + 2\gamma_j)} \prod_{i=1}^{j-1} \frac{\mu_i - \lambda_n}{\mu_i - \lambda_j} \right]^2$$

mit $\gamma_j = \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n}$

Bemerkung 4.5. Satz 4.2 liefert eine Abschätzung des absoluten Fehlers zwischen j -tem Ritz-Wert und Eigenwert. Für $j = 1$ ist der Grad des Tschebyscheff-Polynoms am größten, außerdem entfällt das Produkt über die Ritz-Werte. Wie in Abschnitt 4.2 gezeigt werden wird, ist für die in dieser Arbeit betrachteten exponentiell abfallenden Eigenwerte auch γ_j für $j = 1$ am größten. Allgemein ist für kleine j der Nenner der Klammer in 4.2 groß und das Produkt relativ klein. Für wachsendes j wächst auch die obere Schranke in Satz 4.2, deshalb ist das Lanczos-Verfahren besonders für die Suche nach den extremen (hier: größten) Eigenwerten (samt Eigenvektoren) geeignet.

Da Satz 4.2 von der Konvergenz der Ritz-Werte μ_1, \dots, μ_{j-1} abhängt, wird im folgenden Abschnitt neben einer Winkelabschätzung der Ritz-Vektoren noch ein weiterer Satz zur Konvergenz der Ritz-Werte entwickelt, welcher die Eigenwerte anstelle der Ritz-Werte in dem Produkt der obigen Abschätzung verwendet. Dafür wird aber der Fehlerwinkel der Ritz-Vektoren benutzt, um rekursiv eine Fehlerabschätzung der Ritz-Werte zu erhalten.

4.1.2 Fehlerabschätzung der Eigenwerte und Eigenvektoren mit Hilfe der Ritz-Vektoren

Im Folgenden wird nun der Fehlerwinkel zwischen den Ritz-Vektoren und den Eigenvektoren analysiert.

Da der Krylovraum $\mathcal{K}_m(x, A)$ Eigenvektornäherungen $\tilde{x}_1, \dots, \tilde{x}_m$ beinhaltet, ist es sinnvoll den Winkel zwischen einem Eigenvektor x^j und $\mathcal{K}_m(x, A)$ zu betrachten.

Definiere dazu analog zu (4.11) ein Polynom:

$$\tilde{p}(t) = \tilde{q}(t) \prod_{i=1}^{j-1} (\lambda_i - t) \text{ mit } \tilde{q} \in P_{m-j} \text{ also}$$

$$\tilde{p}(\lambda_j) = \tilde{q}(\lambda_j) \prod_{i=1}^{j-1} (\lambda_i - \lambda_j)$$

Der Unterschied zwischen p aus (4.11) und \tilde{p} besteht also nur darin, dass die echten Eigenwerte anstatt der Ritz-Werte die Nullstellen bilden. Der Startvektor x sei nun orthogonal zerlegt in $u \in X^{j-1}$, $v \in \text{span}\{x^j\}$ und $w \in X^{j\perp}$ durch:

$$x = u \cos \angle(x, X^{j-1}) + x^j \cos \angle(x, x^j) + w \sin \angle(x, X^j)$$

mit $\|u\| = \|w\| = 1$. Da das Polynom \tilde{p} so gewählt ist, dass $\lambda_1, \dots, \lambda_{j-1}$ Nullstellen sind und $u \in X^{j-1} = \text{span}\{x^1, \dots, x^{j-1}\}$ gilt, ist $\tilde{p}(A)u = 0$. Damit folgt:

$$(4.19) \quad \begin{aligned} \tilde{p}(A)x &= \tilde{p}(A)u \cos \angle(x, X^{j-1}) + \tilde{p}(A)x^j \cos \angle(x, x^j) + \tilde{p}(A)w \sin \angle(x, X^j) \\ &= \tilde{p}(A)x^j \cos \angle(x, x^j) + \tilde{p}(A)w \sin \angle(x, X^j) \end{aligned}$$

Für kleine Winkel stimmen ψ , $\tan \psi$ und $\sin \psi$ wegen der Reihenentwicklung des Tangens bis auf $\mathcal{O}(\psi^3)$ -Terme überein. Da $\tilde{p}(A)x \in \mathcal{K}_m(x, A)$ ist, wird im Folgenden $\tan \angle(\tilde{p}(A)x, x^j)$ betrachtet, um den Winkel zwischen $\mathcal{K}_m(x, A)$ und x^j einzuschätzen.

Korollar 4.1. *Es gilt*

$$|\tan \angle(\tilde{p}(A)x, x^j)| = \left| \frac{\sin \angle(x, X^j)}{\cos \angle(x, x^j)} \frac{\|\tilde{p}(A)w\|}{\tilde{p}(\lambda_j)} \right|$$

Beweis. Betrachte zunächst den Tangens des Winkels zum Quadrat:

$$\begin{aligned}
 (4.20) \quad \tan^2 \angle(\tilde{p}(A)x, x^j) &= \frac{\sin^2 \angle(\tilde{p}(A)x, x^j)}{\cos^2 \angle(\tilde{p}(A)x, x^j)} \\
 &= \frac{1 - \cos^2 \angle(\tilde{p}(A)x, x^j)}{\cos^2 \angle(\tilde{p}(A)x, x^j)} \\
 &= \frac{\|\tilde{p}(A)x\|^2 - \langle \tilde{p}(A)x, x^j \rangle^2}{\langle \tilde{p}(A)x, x^j \rangle^2}
 \end{aligned}$$

Die letzte Gleichung in (4.20) gilt aufgrund der Definition des Kosinus:

$$\cos \angle(\tilde{p}(A)x, x^j) = \frac{|\langle \tilde{p}(A)x, x^j \rangle|}{\|\tilde{p}(A)x\| \underbrace{\|x^j\|}_{=1}}$$

Weil $w \in X^{j^\perp}$ ist, und somit senkrecht auf u und x^j steht, gilt mit (4.19) weiter:

$$(4.21) \quad \langle \tilde{p}(A)x, \tilde{p}(A)x \rangle = p(\lambda_j)^2 \cos^2 \angle(x, x^j) + \|\tilde{p}(A)w\|^2 \sin^2 \angle(x, X^j) \quad \text{und}$$

$$(4.22) \quad \langle \tilde{p}(A)x, x^j \rangle^2 = p(\lambda_j)^2 \cos^2 \angle(x, x^j)$$

Mit (4.21) und (4.22) eingesetzt in (4.20) ergibt sich die Behauptung. □

Aus Korollar 4.1 resultiert wieder eine Minimierungsaufgabe für das Polynom

$$\tilde{p}(t) = \tilde{q}(t) \prod_{i=1}^{j-1} (\lambda_i - t), \text{ um den Fehler besser abzuschätzen.}$$

Wähle nun analog zu Satz 4.2 und Definition 4.1 : $\tilde{q}(t) = T_{m-j}(\phi(t))$.

Wegen $w \in X^{j^\perp}$ existieren $w_{j+1}, \dots, w_n \in \mathbb{R}$ so, dass $w = \sum_{i=j+1}^n w_i x^i$ mit $\|w\| = 1$ gilt.

Für die Norm $\|\tilde{p}(A)w\|$, welche nach oben abgeschätzt werden soll, folgt:

$$\begin{aligned}
 \|\tilde{p}(A)w\| &= \left\| \sum_{i=j+1}^n \tilde{p}(\lambda_i) w_i x^i \right\| \\
 &\leq \max_{j+1 \leq k \leq n} |\tilde{p}(\lambda_k)| \underbrace{\left\| \sum_{i=j+1}^n w_i x^i \right\|}_{=\|w\|} \\
 &= \max_{j+1 \leq k \leq n} |\tilde{p}(\lambda_i)| \quad \text{da } \|w\| = 1 \\
 &= \max_{j+1 \leq k \leq n} \left| \prod_{i=1}^{j-1} (\lambda_k - \lambda_i) T_{m-j}(\phi(\lambda_k)) \right| \\
 &\leq \max_{j+1 \leq k \leq n} \left| \prod_{i=1}^{j-1} (\lambda_k - \lambda_i) \max_{j+1 \leq k \leq n} |T_{m-j}(\phi(\lambda_k))| \right|
 \end{aligned}$$

$$\leq \max_{j+1 \leq k \leq n} \left| \prod_{i=1}^{j-1} (\lambda_k - \lambda_i) \right| \leq \left| \prod_{i=1}^{j-1} (\lambda_n - \lambda_i) \right|$$

Die letzte Ungleichung folgt nach Bemerkung 4.4, welche besagt, dass $T_{m-j}(\phi(t)) \leq 1$ für $\lambda_n \leq t \leq \lambda_{j+1}$ gilt.

Mit $\tilde{p}(\lambda_j) = \tilde{q}(\lambda_j) \prod_{i=1}^{j-1} (\lambda_i - \lambda_j)$ folgt nun:

$$(4.23) \quad \left| \tan \angle(\tilde{p}(A)x, x^j) \right| \leq \left| \frac{\sin \angle(x, X^j)}{\cos \angle(x, X^j)} \frac{1}{T_{m-j}(\phi(\lambda_j))} \prod_{i=1}^{j-1} \frac{\lambda_j - \lambda_n}{\lambda_i - \lambda_j} \right|$$

Für den Winkel zwischen x^j und $\mathcal{K}_m(x, A)$ gilt nach Definition 2.4:

$$\begin{aligned} \left| \tan \angle(x^j, \mathcal{K}_m(x, A)) \right| &= \left| \frac{\sin \angle(x^j, \mathcal{K}_m(x, A))}{\cos \angle(x^j, \mathcal{K}_m(x, A))} \right| \\ &= \min_{y \in \mathcal{K}_m(x, A)} \frac{|\sin \angle(x^j, y)|}{|\cos \angle(x^j, y)|} \\ &= \min_{y \in \mathcal{K}_m(x, A)} |\tan \angle(x^j, y)| \\ &\leq |\tan \angle(x^j, p(A)x)| \end{aligned}$$

Mit (4.23) folgt somit:

$$(4.24) \quad \left| \tan \angle(\mathcal{K}_m(x, A), x^j) \right| \leq \left| \frac{\sin \angle(x, X^j)}{\cos \angle(x, X^j)} \frac{1}{T_{m-j}(\phi(\lambda_j))} \prod_{i=1}^{j-1} \frac{\lambda_j - \lambda_n}{\lambda_i - \lambda_j} \right|$$

Bemerkung 4.6. *Unter der Annahme, dass die Eigenwerte schon relativ gut konvergiert sind (also $\mu_i \approx \lambda_i$ für $i \in \{1, \dots, j-1\}$), ist anhand obiger Gleichung in Vergleich mit Satz 4.2 zu erkennen, dass sich der Fehler der Eigenwerte ungefähr quadratisch zum Fehler der Eigenvektoren verhält. Verdeutlicht wird dies auch in dem unten folgenden Satz 4.3.*

Da der Fehler zwischen λ_j und μ_j in Satz 4.2 von den Ritz-Werten μ_1, \dots, μ_{j-1} abhängt, wird nun eine weitere Abschätzung entwickelt, mit dem Ziel eine Aussage über den Fehler zwischen Eigenwert und Ritz-Wert zu erhalten, die unabhängig von den vorher berechneten Ritz-Werten ist. Außerdem fehlt noch ein konkreter Zusammenhang zwischen Ritz-Vektor und Eigenvektor, da bisher nur der Winkel zwischen Krylovraum und Eigenvektor in (4.24) betrachtet wurde. Das Ergebnis ergibt zusammengefasst den nächsten Satz, der im Folgenden in einzelnen Schritten bewiesen wird. Allerdings stellt dieser eine nicht ganz befriedigende Lösung des Problems dar, weil nun der Fehlerwinkel der Ritz-Vektoren, welcher ebenfalls mit der Konvergenz der μ_i zusammenhängt, mit in die Abschätzung der Ritz-Werte einfließt.

Satz 4.3. Für die Ritz-Werte μ_1, \dots, μ_j und Ritz-Vektoren z^1, \dots, z^{j-1} mit $j \in \{1, \dots, m\}$ des Lanczos-Verfahrens und Startvektor x gilt:

$$0 \leq \lambda_j - \mu_j \leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j)}{\cos \angle(x, x^j)} \frac{1}{T_{m-j}(1 + 2\gamma_j)} \prod_{i=1}^{j-1} \frac{\lambda_i - \lambda_n}{\lambda_i - \lambda_j} \right]^2 + \sum_{i=1}^{j-1} (\lambda_i - \lambda_n) \sin^2 \angle(z^i, x^i)$$

mit $\gamma_j = \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n}$ und

$$\sin^2 \angle(z^i, x^i) \leq \frac{1}{\lambda_i - \lambda_{i+1}} \left[\lambda_i - \mu_i + \sum_{k=1}^{i-1} (\lambda_k - \lambda_n) \sin^2 \angle(z^k, x^k) \right]$$

Bemerkung 4.7. Vergleicht man die obere Gleichung des Satzes mit Gleichung (2.4), so ist zu erkennen, dass in die Abschätzung des Approximationsfehlers der Eigenwerte die obere Schranke des Fehlerwinkels zwischen $\mathcal{K}_m(x, A)$ und Eigenvektor einfließt. Bemerkung 4.6 wird also bestätigt. Außerdem ist zu bemerken, dass in dem Produkt aus Satz 4.2 die Ritz-Werte μ_i durch die Eigenwerte λ_i ersetzt sind und somit nicht mehr in der Eigenwertabschätzung auftauchen. Dafür kommt die Summe über den Abstand der Eigenwerte multipliziert mit dem Winkel der vorherigen Ritz-Vektoren und Eigenvektoren hinzu.

Der erste Teil von Satz 4.3 folgt im Wesentlichen aus dem nächsten Lemma:

Lemma 4.4. Sei S^m m -dimensionaler Unterraum des \mathbb{R}^n . Sei $s \in S^m$ gegeben mit $\langle s, x^i \rangle = 0$ und $\|s\| = 1$. Sei $\phi_i := \angle(x^i, z^i)$ der Winkel zwischen i -tem Eigenvektor x^i von A und aus S^m berechnetem Ritz-Vektor z^i , $i \in 1, \dots, j-1$. Dann gilt:

$$\mu_j \geq R(s, A) - \sum_{i=1}^{j-1} (\mu_j - \lambda_n) \sin^2 \phi_i$$

Beweis. Die Ritz-Vektoren z^1, \dots, z^m bilden eine ONB von $S_m = \text{span}\{z^1, \dots, z^m\}$. Also läßt sich s darstellen durch

$$s = \sum_{i=1}^{j-1} s_i z^i + \underbrace{\sum_{i=j}^m s_i z^i}_{=:t} \quad s_i \in \mathbb{R} \text{ für } i \in \{1, \dots, m\}$$

Da die Ritz-Vektoren orthogonal sind, gilt

$$(4.25) \quad \langle t, z^i \rangle = 0 \quad \text{für } i = 1, \dots, j-1.$$

Deshalb folgt für s_i und t :

$$(4.26) \quad |\langle s, z^i \rangle| = \left| \left\langle t + \sum_{k=1}^{j-1} s_k z^k, z^i \right\rangle \right| = \left| \langle t, z^i \rangle + \left\langle \sum_{k=1}^{j-1} s_k z^k, z^i \right\rangle \right| = |\langle s_i z^i, z^i \rangle| = |s_i|$$

Mit Hilfe der folgenden Gleichung kann die Norm von t abgeschätzt werden. Es ist

$$1 = \|s\|^2 = \left\langle t + \sum_{i=1}^{j-1} s_i z^i, t + \sum_{i=1}^{j-1} s_i z^i \right\rangle = \langle t, t \rangle + \left\langle \sum_{i=1}^{j-1} s_i z^i, \sum_{i=1}^{j-1} s_i z^i \right\rangle$$

und deshalb folgt für $\|t\|^2$:

$$(4.27) \quad \|t\|^2 = 1 - \left\| \sum_{i=1}^{j-1} s_i z^i \right\|^2 \leq 1$$

Wie in der nächsten Rechnung zu sehen sein wird, besteht ein Zusammenhang zwischen $|s_i|$ und dem Winkel ϕ_i , also wird $|s_i|$ genauer betrachtet.

Nach Voraussetzung ist $\langle s, x^i \rangle = 0$, also auch $\langle s, x^i \cos \phi_i \rangle = 0$, deshalb folgt mit Hilfe der Cauchy-Schwarzschen Ungleichung und (4.26):

$$(4.28) \quad |s_i| = |\langle s, z^i \rangle| = |\langle s, z^i \rangle - \langle s, x^i \cos \phi_i \rangle| = |\langle s, z^i - x^i \cos \phi_i \rangle| \leq \|s\| \|z^i - x^i \cos \phi_i\|$$

Für die Norm des Differenzvektors von z^i und $x^i \cos \phi_i$ gilt weiter:

$$\begin{aligned} \|z^i - x^i \cos \phi_i\| &= \|z^i - x^i \langle x^i, z^i \rangle\| \\ &= \sqrt{\langle z^i, z^i \rangle - 2 \langle z^i, x^i \langle x^i, z^i \rangle \rangle + \langle x^i \langle x^i, z^i \rangle, \langle x^i \langle x^i, z^i \rangle \rangle} \\ &= \sqrt{1 - 2 \langle x^i, z^i \rangle \langle x^i, z^i \rangle + \langle x^i, z^i \rangle^2} \\ &= \sqrt{1 - \langle x^i, z^i \rangle^2} = |\sin \phi_i| \quad \text{da } \sin \phi_i = \sqrt{1 - \cos^2 \phi_i} \end{aligned}$$

Nach Voraussetzung ist $\|s\| = 1$, und es gilt mit (4.28) und obiger Rechnung:

$$(4.29) \quad |s_i| \leq \|s\| \|z^i - x^i \cos \phi_i\| = |\sin \phi_i|$$

Da die Ungleichung

$$R(s, A) \leq \mu_j + \sum_{i=1}^{j-1} (\mu_j - \lambda_n) \sin^2 \phi_i$$

gezeigt werden soll, wird nun der Rayleigh-Quotient $R(s, A)$ berechnet. Seien dazu y^1, \dots, y^m die Eigenvektoren der Ritz-Matrix $B = Q^T A Q$, also $z^i = Q y^i$. Es gilt:

$$(4.30) \quad R(s, A) = \frac{\langle A s, s \rangle}{\|s\|^2} = \underbrace{\langle A t, t \rangle}_{=1} + \underbrace{\left\langle A \sum_{i=1}^{j-1} s_i z^i, t \right\rangle}_{=0} + \underbrace{\left\langle A t, \sum_{i=1}^{j-1} s_i z^i \right\rangle}_{=0} + \left\langle A \sum_{i=1}^{j-1} s_i z^i, \sum_{i=1}^{j-1} s_i z^i \right\rangle$$

$$\begin{aligned}
 &= \langle At, t \rangle + \sum_{i=1}^{j-1} s_i^2 \langle Az^i, z^i \rangle = \langle At, t \rangle + \sum_{i=1}^{j-1} s_i^2 \langle AQy^i, Qy^i \rangle \\
 &= \langle At, t \rangle + \sum_{i=1}^{j-1} s_i^2 \langle By^i, y^i \rangle = \langle At, t \rangle + \sum_{i=1}^{j-1} s_i^2 \mu_i
 \end{aligned}$$

Weil der n -te Eigenwert λ_n in der Summe über $\sin \phi_i$ in der zu beweisenden Ungleichung auftaucht, wird Gleichung (4.30) auf die Matrix $A - \lambda_n I$ angewendet:

$$\begin{aligned}
 R(s, A - \lambda_n I) &= \langle (A - \lambda_n I)t, t \rangle + \sum_{i=1}^{j-1} s_i^2 \langle (A - \lambda_n I)z^i, z^i \rangle \\
 &= \langle (A - \lambda_n I)t, t \rangle + \sum_{i=1}^{j-1} (s_i^2 \langle Az^i, z^i \rangle - \lambda_n \langle z^i, z^i \rangle) \\
 &= \langle (A - \lambda_n I)t, t \rangle + \sum_{i=1}^{j-1} (\mu_i - \lambda_n) s_i^2 && \text{da } \|z^i\| = 1 \\
 &\leq \frac{\langle (A - \lambda_n I)t, t \rangle}{\|t\|^2} + \sum_{i=1}^{j-1} (\mu_i - \lambda_n) s_i^2 && \text{nach (4.27)} \\
 &\leq R(t, A - \lambda_n I) + \sum_{i=1}^{j-1} (\mu_i - \lambda_n) \sin^2 \phi_i && \text{nach (4.29)}
 \end{aligned}$$

Damit folgt:

$$\begin{aligned}
 R(s, A) - \lambda_n &= R(s, A - \lambda_n I) \leq R(t, (A - \lambda_n I)) + \sum_{i=1}^{j-1} (\mu_i - \lambda_n) \sin^2 \phi_i \\
 &= R(t, A) - \lambda_n + \sum_{i=1}^{j-1} (\mu_i - \lambda_n) \sin^2 \phi_i
 \end{aligned}$$

Also gilt:

$$(4.31) \quad R(t, A) \geq R(s, A) - \sum_{i=1}^{j-1} (\mu_i - \lambda_n) \sin^2 \phi_i$$

Um die Behauptung des Lemmas zu erhalten, muß nun nur noch gezeigt werden, dass $\mu_j \geq R(t, A)$ ist. Dies folgt analog zu dem Beweis von Lemma 4.2 mit $z^i = Qy^i$. Sei hierzu der Raum W^{j-1} wieder definiert durch $W^{j-1} = \text{span}\{y^1, \dots, y^{j-1}\}$, $W^0 = \{0\}$.

Nach (4.25) ist $t \in QW^{j-1\perp}$, also gilt:

$$\mu_j = \max_{\substack{y \neq 0 \\ y \in W^{j-1\perp}}} R(y, B) = \max_{\substack{x \neq 0 \\ x \in QW^{j-1\perp}}} R(x, A) \geq R(t, A)$$

□

Um die Abschätzung der Eigenwerte aus Satz 4.3 zu erhalten, müssen nur noch die schon gezeigte Abschätzung aus Satz 4.2 und Lemma 4.4 in Zusammenhang gebracht werden. Mit Hilfe von Lemma 4.4 und Lemma 4.1 folgt für ein Polynom $p \in \mathcal{P}_{m-1}$:

$$\begin{aligned} 0 \leq \lambda_j - \mu_j &\leq \lambda_j - [R(s, A) - \sum_{i=1}^{j-1} (\mu_i - \lambda_n) \sin^2 \phi_i] && \text{Lemma 4.4} \\ &= R(s, \lambda_j I - A) + \sum_{i=1}^{j-1} (\mu_i - \lambda_n) \sin^2 \phi_i \\ &\leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j) \|p(A)x^{\perp(j)}\|}{\cos \angle(x, x^j) p(\lambda_j)} \right]^2 + \sum_{i=1}^{j-1} (\mu_i - \lambda_n) \sin^2 \phi_i && \text{Lemma 4.1} \end{aligned}$$

Analog zum Beweis von Satz 4.2 ergibt sich der erste Teil von Satz 4.3 mit:

$$p := \tilde{q}(t) \tilde{p}(t) = \tilde{q}(t) \prod_{i=1}^{j-1} (\lambda_i - t)$$

Es bleibt noch der zweite Teil von Satz 4.3 zu zeigen. Wieder erfolgt der erste Schritt dazu mit Hilfe des Rayleigh-Quotienten. Diesmal wird der Abstand zwischen Eigenwert und Rayleigh-Quotient des Ritz-Vektors betrachtet, um die Konvergenz der Ritz-Vektoren zu untersuchen.

Lemma 4.5. *Für den Rayleigh-Quotienten $R(z^j, \lambda_j I - A)$ gilt:*

$$\lambda_j - \mu_j = R(z^j, (\lambda_j I - A)) = \sum_{i=1}^n (\lambda_j - \lambda_i) \cos^2 \angle(x^i, z^j)$$

Beweis. Die linke Seite der Gleichung folgt mit $z^j = Qy^j$ sofort aus der nächsten Rechnung. Hierbei wird ausgenutzt, dass die Matrix Q orthonormal ist, und y^j die normierten Eigenvektoren zu den Eigenwerten μ_j von $Q^T A Q$ sind.

$$\begin{aligned} (4.32) \quad R(z^j, (\lambda_j I - A)) &= \frac{\langle (\lambda_j I - A)z^j, z^j \rangle}{\underbrace{\langle z^j, z^j \rangle}_{=1}} = \langle \lambda_j z^j, z^j \rangle - \langle A z^j, z^j \rangle \\ &= \lambda_j \langle Qy^j, Qy^j \rangle - \langle (AQ)y^j, Qy^j \rangle \end{aligned}$$

$$\begin{aligned} &= \lambda_j \langle y^j, y^j \rangle - \langle (Q^T A Q) y^j, y^j \rangle \\ &= \lambda_j - \mu_j \end{aligned}$$

Betrachte nun die rechte Seite der Gleichung. Sei hierzu $\phi_{ij} := \angle(x^i, z^j)$ und $\phi_j := \angle(x^j, z^j)$. Da die Eigenvektoren x^1, \dots, x^n von A eine Orthonormalbasis des \mathbb{R}^n bilden, und auch z^j normiert ist, gilt:

$$(4.33) \quad z^j = \sum_{i=1}^n x^i \langle x^i, z^j \rangle \quad \text{und} \quad \langle x^i, z^j \rangle^2 = \cos^2 \phi_{ij}$$

Daraus folgt für den Rayleigh-Quotienten:

$$\begin{aligned} (4.34) \quad R(z^j, \lambda_j I - A) &= \langle \lambda_j z^j, z^j \rangle - \langle A z^j, z^j \rangle \\ &= \langle \lambda_j \sum_{i=1}^n x^i \langle x^i, z^j \rangle, z^j \rangle - \langle A \sum_{i=1}^n x^i \langle x^i, z^j \rangle, z^j \rangle \\ &= \langle \lambda_j \sum_{i=1}^n x^i \langle x^i, z^j \rangle, z^j \rangle - \langle \sum_{i=1}^n \lambda_i x^i \langle x^i, z^j \rangle, z^j \rangle \\ &= \langle \sum_{i=1}^n (\lambda_j - \lambda_i) x^i \langle x^i, z^j \rangle, z^j \rangle \\ &= \sum_{i=1}^n (\lambda_j - \lambda_i) \langle x^i \langle x^i, z^j \rangle, \sum_{k=1}^n x^k \langle x^k, z^j \rangle \rangle \\ &= \sum_{i=1}^n (\lambda_j - \lambda_i) \langle x^i, z^j \rangle^2 \\ &= \sum_{i=1}^n (\lambda_j - \lambda_i) \cos^2 \phi_{ij} \end{aligned}$$

Somit sind beide Gleichungen der Behauptung gezeigt. □

Nach Lemma 4.5 gilt also insbesondere:

$$(4.35) \quad \lambda_j - \mu_j = \sum_{i=1}^n (\lambda_j - \lambda_i) \cos^2 \phi_{ij}$$

Aus dieser Gleichung folgt der nächste Schritt in Richtung der Winkelabschätzung aus Satz 4.3. Es gilt nach (4.35):

$$\lambda_j - \mu_j - \sum_{i=1}^{j-1} (\lambda_j - \lambda_i) \cos^2 \phi_{ij} = \sum_{i=j+1}^n (\lambda_j - \lambda_i) \cos^2 \phi_{ij}$$

Also ist aufgrund der Voraussetzung $\lambda_1 \geq \dots \geq \lambda_n$ auch:

$$\lambda_j - \mu_j + \sum_{i=1}^{j-1} \underbrace{(\lambda_i - \lambda_j)}_{>0} \cos^2 \phi_{ij} = \sum_{i=j+1}^n \underbrace{(\lambda_j - \lambda_i)}_{>0} \cos^2 \phi_{ij} \geq (\lambda_j - \lambda_{j+1}) \sum_{i=j+1}^n \cos^2 \phi_{ij}$$

Um die Fehlerabschätzung endgültig zu zeigen, muss anstelle des Kosinus des Winkels der Sinus in die Gleichung (4.36) eingebracht werden. Dazu wird ausgenutzt, dass die Ritz-Vektoren orthonormal sind. Es gilt:

$$1 = \|z^j\|^2 = \left\langle \sum_{i=1}^n x^i \langle x^i, z^j \rangle, \sum_{i=1}^n x^i \langle x^i, z^j \rangle \right\rangle = \sum_{i=1}^n \cos^2 \phi_{ij}$$

Zusammen mit $\sin^2 \phi_j = 1 - \cos^2 \phi_j$ folgt

$$(4.36) \quad \sum_{i=j+1}^n \cos^2 \phi_{ij} = \underbrace{\sum_{i=1}^n \cos^2 \phi_{ij}}_{=1} - \cos^2 \phi_j - \sum_{i=1}^{j-1} \cos^2 \phi_{ij} = \sin^2 \phi_j - \sum_{i=1}^{j-1} \cos^2 \phi_{ij}$$

Mit (4.36) und (4.36) ergibt sich insgesamt:

$$(4.37) \quad \lambda_j - \mu_j + \sum_{i=1}^{j-1} (\lambda_i - \lambda_j) \cos^2 \phi_{ij} \geq (\lambda_j - \lambda_{j+1}) \left(\sin^2 \phi_j - \sum_{i=1}^{j-1} \cos^2 \phi_{ij} \right)$$

Da der Sinus des Winkels zwischen j -tem Ritz-Vektor z^j und Eigenvektor x^j in Satz 4.3 abgeschätzt werden soll, wird die Gleichung nach $\sin^2 \phi_j$ aufgelöst.

$$(4.38) \quad \begin{aligned} \sin^2 \phi_j &\leq \frac{1}{(\lambda_j - \lambda_{j+1})} \left[(\lambda_j - \mu_j) + \sum_{i=1}^{j-1} (\lambda_i - \lambda_j) \cos^2 \phi_{ij} + \sum_{i=1}^{j-1} (\lambda_j - \lambda_{j+1}) \cos^2 \phi_{ij} \right] \\ &= \frac{1}{(\lambda_j - \lambda_{j+1})} \left[(\lambda_j - \mu_j) + \sum_{i=1}^{j-1} (\lambda_i - \lambda_{j+1}) \cos^2 \phi_{ij} \right] \end{aligned}$$

Um rekursiv den Fehlerwinkel bestimmen zu können, wird $\cos \phi_{ij}$ auch auf der rechten Seite durch $\sin \phi_i$ abgeschätzt. Es gilt:

$$\begin{aligned} \cos \phi_{ij} &= |\langle x^i, z^j \rangle| \\ &= |\langle x^i - \cos \phi_i z^i, z^j \rangle| && \text{da } \langle z^i, z^j \rangle = 0 \\ &\leq \|x^i - \cos \phi_i z^i\| \|z^j\| && \text{nach Cauchy-Schwarz-Ungleichung} \\ &= \sqrt{\langle x^i, x^i \rangle - 2 \cos \phi_i \langle x^i, z^i \rangle + \cos^2 \phi_i \langle z^i, z^i \rangle} \end{aligned}$$

$$\begin{aligned} &= \sqrt{1 - 2 \cos^2 \phi_i + \cos^2 \phi_i} = \sqrt{1 - \cos^2 \phi_i} \\ &= \sin \phi_i \end{aligned}$$

Ersetze nun also in (4.38) $\cos \phi_{ij}$ durch $\sin \phi_i$, dann folgt

$$\sin^2 \angle(z^i, x^i) \leq \frac{1}{\lambda_i - \lambda_{i+1}} \left[\lambda_i - \mu_i + \sum_{k=1}^{i-1} (\lambda_k - \lambda_n) \sin^2(z^k, x^k) \right].$$

Somit ist auch der zweite Teil des Satzes 4.3 bewiesen.

Satz 4.3 ist jedoch im Vergleich mit Satz 4.2 relativ kompliziert, und deshalb schwerer zu analysieren. Weil die obige Winkelsumme zusätzlich in die Konvergenzaussage eingeht, und damit auch der erhoffte Vorteil durch die Eliminierung der μ_i aus der Eigenwertabschätzung nicht ganz so groß ist, wird im nächsten Kapitel hauptsächlich Satz 4.2 für Konvergenzuntersuchungen der speziellen Matrizen mit exponentiell abfallenden Eigenwerten verwendet.

4.2 Analyse für exponentiell abfallende Eigenwerte

In diesem Kapitel wurden bisher nur allgemeine Konvergenzaussagen des symmetrischen Lanczos-Verfahrens gezeigt. Thema dieser Arbeit sind Matrizen mit exponentiell abfallenden Eigenwerten. In den folgenden Überlegungen geht es um die Frage, wie sich die Konvergenzaussagen des vorigen Abschnitts für die speziellen Eigenwerte verhalten. Es werden also die vorangegangenen Sätze genauer betrachtet mit dem Ziel, diese für die vorgegebenen Eigenwerte auszuwerten und eventuell zu vereinfachen.

Im folgenden werden die Eigenwerte

$$(4.39) \quad \lambda_k := c_1 e^{-c_2 (k-1)^\alpha} \quad \text{mit } \alpha \in (0, 1], c_1, c_2 > 0$$

betrachtet.

4.2.1 Konvergenzbetrachtung mit Hilfe der Ritz-Werte

Im Folgenden wird Satz 4.2 für die Eigenwerte aus (4.39) untersucht.

Wesentlich für die Konvergenzaussage in Satz 4.2 ist das Tschebyscheffpolynom $T_{m-j}(1 + 2\gamma_j)$ im Nenner.

Um die Größenordnung des Fehlers einschätzen zu können, wird im Folgenden

$$\frac{1}{T_{m-j}(1 + 2\gamma_j)}$$

untersucht. Als Motivation für die nächsten Schritte, welche sich mit der Größe γ_j beschäftigen, wird die folgende Formel aus [8], Seite 378 (vergleiche auch [4], Seite 46), wiedergegeben.

$$(4.40) \quad T_{m-j}(1 + 2\gamma_j) \approx \begin{cases} \frac{1}{2}(1 + 2\sqrt{\gamma_j} + 2\gamma_j)^{m-j} & \text{für } 0 < \gamma_j < 0.1 \\ \frac{1}{2}e^{2k\sqrt{\gamma_j}} & \text{für } (m-j)\sqrt{\gamma_j} > 1 \end{cases}$$

Aus der Formel (4.40) ist zu erkennen, dass $T_{m-j}(1 + 2\gamma_j)$ für wachsendes γ_j ansteigt, und somit die Fehlerschranke kleiner wird.

Betrachte also zuerst γ_j für die Eigenwerte aus (4.39)

$$\lambda_k = c_1 e^{-c_2 (k-1)^\alpha}$$

Es gilt:

$$(4.41) \quad \begin{aligned} \gamma_j &:= \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n} \\ &= \frac{c_1 e^{-c_2 (j-1)^\alpha} - c_1 e^{-c_2 j^\alpha}}{c_1 e^{-c_2 j^\alpha} - c_1 e^{-c_2 (n-1)^\alpha}} \\ &= \frac{e^{c_2 (j^\alpha - (j-1)^\alpha)} - 1}{1 - \frac{1}{e^{c_2 ((n-1)^\alpha - j^\alpha)}}} \end{aligned}$$

Bemerkung 4.8. *Offensichtlich ist γ_j unabhängig von c_1 , also auch die Größe des Tschebyscheff-Polynoms an der Stelle $1 + 2\gamma_j$. Aufgrund seiner Definition ist γ_j und somit auch der Wert des Tschebyscheff-Polynoms abhängig von den Abständen der Eigenwerte zueinander. Liegen diese eng beieinander, so ist γ_j kleiner als für weit auseinander liegende Eigenwerte.*

Da in der Praxis sehr große Matrizen auftreten, darf im Folgenden n als hinreichend groß angenommen werden. Außerdem zeigt sich, dass häufig Matrizen behandelt werden, für die gilt, dass der Parameter α gleich dem Kehrwert der Raumdimension ist. Das heißt es gilt zum Beispiel $\alpha = 1$, $\frac{1}{2}$ oder $\frac{1}{3}$. Es gelten für γ_j die folgenden Ungleichungen:

$$(4.42) \quad \gamma_1 = \frac{e^{c_2} - 1}{1 - \frac{e^{c_2}}{e^{c_2 (n-1)^\alpha}}} > e^{c_2} - 1 \quad \forall \alpha \in (0, 1]$$

$$(4.43) \quad \gamma_j = \frac{e^{c_2 (j^\alpha - (j-1)^\alpha)} - 1}{1 - \frac{e^{c_2 j^\alpha}}{e^{c_2 (n-1)^\alpha}}} > e^{c_2 (j^\alpha - (j-1)^\alpha)} - 1 =: \hat{\gamma}_j \quad \text{für } j > 1$$

Im Fall $\alpha = 1$ gilt damit wie für γ_1 :

$$(4.44) \quad \gamma_j > e^{c_2} - 1$$

Allgemein gilt:

Korollar 4.2. Sei $\hat{\gamma}_j$ wie in 4.43 definiert.

Aus $\alpha \rightarrow 0$ folgt $\hat{\gamma}_j \rightarrow 0$

Beweis. Es gilt für $j \geq 2$:

$$\int_{j-1}^j \alpha x^{\alpha-1} dx = [x^\alpha]_{j-1}^j = j^\alpha - (j-1)^\alpha$$

Für $x \in [j-1, j]$, es gilt also insbesondere $x \geq 1$, ist

$$\lim_{\alpha \rightarrow 0} \int_{j-1}^j \alpha x^{\alpha-1} dx = \int_{j-1}^j \lim_{\alpha \rightarrow 0} \alpha x^{\alpha-1} dx = \int_{j-1}^j \lim_{\alpha \rightarrow 0} \frac{\alpha x^\alpha}{x} dx = \int_{j-1}^j 0 dx = 0$$

Es folgt:

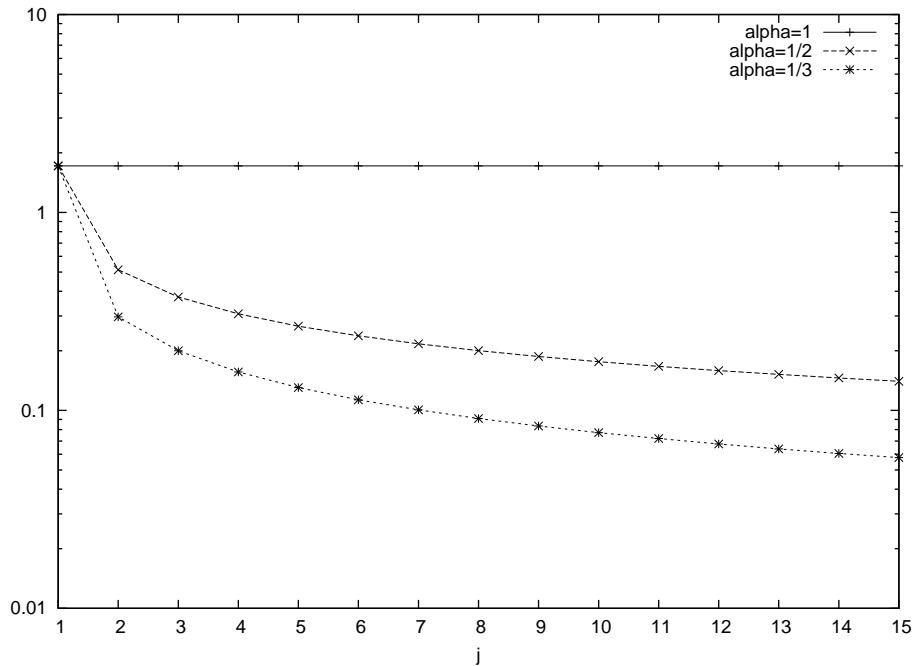
$$\lim_{\alpha \rightarrow 0} e^{c_2(j^\alpha - (j-1)^\alpha)} - 1 = 0$$

□

Bemerkung 4.9. Für $j > 1$ gilt: Je kleiner α ist, desto kleiner ist auch γ_j , und desto dichter liegen die Eigenwerte zusammen. Wie in (4.43) zu erkennen, ist γ_1 für großes n (zum Beispiel wie in den hier untersuchten Fällen für $n = 1000$), unabhängig von α .

Die anschließende Grafik zeigt γ_j in Abhängigkeit von α für $\alpha = 1, 1/2$ und $1/3$. Außerdem ist $n = 1000$, und $c_2 = 1$ gesetzt. γ_1 , beziehungsweise γ_j für $\alpha = 1$ liegt hier bei ungefähr $e - 1 = 1,718$.

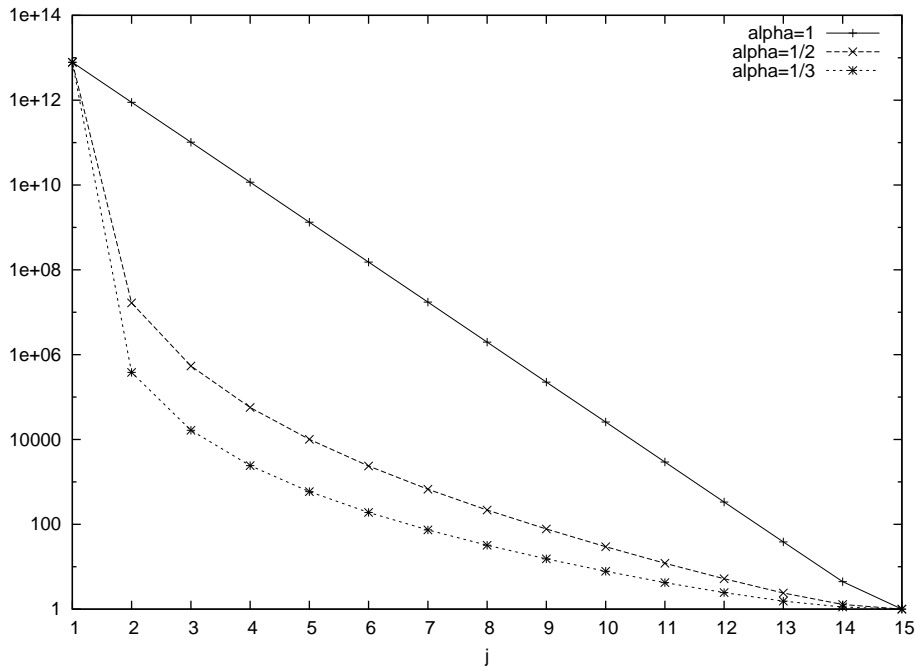
Abbildung 4.1: γ_j für $j = 1$ bis $m = 15$ und $\alpha = 1, 1/2, 1/3$



Bemerkung 4.10. Da also das Tschebyscheff-Polynom $T_{m-j}(1 + 2\gamma_j)$ groß ist für ein großes γ_j , gilt wegen Bemerkung 4.9 somit auch: Je kleiner α ist, umso kleiner ist auch $T_{m-j}(1 + 2\gamma_j)$.

Die folgende Grafik zeigt $T_{m-j}(1 + 2\gamma_j)$ für $n = 1000$ und $c_2 = 1$.

Abbildung 4.2: $T_{m-j}(1 + 2\gamma_j)$ für $j = 1$ bis 15 und $\alpha = 1, 1/2, 1/3$



Was kann man nun konkret über das Tschebyscheff-Polynom T_{m-j} an der Stelle $1 + 2\gamma_j$ aussagen? Eine Hilfestellung zur Berechnung gibt folgendes Lemma:

Lemma 4.6. Es gilt für $|\xi| > 1$ und $k \in \mathbb{N}$:

$$T_k(\xi) = \frac{1}{2} \left[\left(\xi + \sqrt{\xi^2 - 1} \right)^k + \left(\xi + \sqrt{\xi^2 - 1} \right)^{-k} \right]$$

Beweis . Siehe [1], Lemma 3.8. □

Um nun $T_{m-j}(1 + 2\gamma_j)$ mit Hilfe von Lemma 4.6 auszuwerten, berechne zuerst $\xi^2 - 1$ für $\xi := 1 + 2\gamma_j$:

$$(4.45) \quad (1 + 2\gamma_j)^2 - 1 = \left(1 + 2 \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n} \right)^2 - 1$$

$$\begin{aligned}
 &= 1 + 4 \left(\frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n} \right) + 4 \left(\frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n} \right)^2 - 1 \\
 &= 4 \left(\frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n} + \frac{(\lambda_j - \lambda_{j+1})^2}{(\lambda_{j+1} - \lambda_n)^2} \right) \\
 &= 4 \left(\frac{(\lambda_j - \lambda_{j+1})(\lambda_{j+1} - \lambda_n) + (\lambda_j - \lambda_{j+1})^2}{(\lambda_{j+1} - \lambda_n)^2} \right) \\
 &= 4 \frac{(\lambda_j - \lambda_{j+1})(\lambda_j - \lambda_n)}{(\lambda_{j+1} - \lambda_n)^2} > 0 \quad \text{da } \lambda_1 \geq \dots \geq \lambda_n
 \end{aligned}$$

Nun kann die Klammer $\xi + \sqrt{\xi^2 - 1}$ aus Lemma 4.6 mit $\xi = 1 + 2\gamma_j$ ausgerechnet werden. Mit (4.45) gilt:

$$\begin{aligned}
 (4.46) \quad &1 + 2\gamma_j + \sqrt{(1 + 2\gamma_j)^2 - 1} \\
 &= 1 + 2\gamma_j + \sqrt{4 \frac{(\lambda_j - \lambda_{j+1})(\lambda_j - \lambda_n)}{(\lambda_{j+1} - \lambda_n)^2}} \\
 &= 1 + 2 \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n} + \frac{2}{\lambda_{j+1} - \lambda_n} \sqrt{(\lambda_j - \lambda_{j+1})(\lambda_j - \lambda_n)} \\
 &= \frac{(\lambda_{j+1} - \lambda_n) + 2(\lambda_j - \lambda_{j+1}) + 2\sqrt{(\lambda_j - \lambda_{j+1})(\lambda_j - \lambda_n)}}{\lambda_{j+1} - \lambda_n} \\
 &= \frac{(\lambda_j - \lambda_{j+1}) + (\lambda_j - \lambda_n) + 2\sqrt{(\lambda_j - \lambda_{j+1})}\sqrt{(\lambda_j - \lambda_n)}}{\lambda_{j+1} - \lambda_n} \\
 &= \frac{(\sqrt{\lambda_j - \lambda_{j+1}} + \sqrt{\lambda_j - \lambda_n})^2}{\lambda_{j+1} - \lambda_n} \\
 &= \left(\frac{\lambda_j - \lambda_n}{\lambda_{j+1} - \lambda_n} \right) \left(1 + \frac{\sqrt{\lambda_j - \lambda_{j+1}}}{\sqrt{\lambda_j - \lambda_n}} \right)^2 \\
 &= \left(\frac{\lambda_j - \lambda_n}{(\lambda_j - \lambda_n) + (\lambda_{j+1} - \lambda_j)} \right) \left(1 + \frac{\sqrt{\lambda_j - \lambda_{j+1}}}{\sqrt{\lambda_j - \lambda_n}} \right)^2 \\
 &= \left(\frac{\lambda_{j+1} - \lambda_j}{\lambda_j - \lambda_n} + 1 \right)^{-1} \left(1 + \frac{\sqrt{\lambda_j - \lambda_{j+1}}}{\sqrt{\lambda_j - \lambda_n}} \right)^2
 \end{aligned}$$

Zur Vereinfachung der Formel definiere:

$$\kappa_j := \frac{\lambda_j - \lambda_n}{\lambda_j - \lambda_{j+1}}.$$

Dann folgt mit κ_j eingesetzt in (4.46):

$$\begin{aligned}
 (4.47) \quad 1 + 2\gamma_j + \sqrt{(1 + 2\gamma_j)^2 - 1} &= \left(1 - \frac{1}{\kappa_j}\right)^{-1} \left(1 + \frac{1}{\sqrt{\kappa_j}}\right)^2 \\
 &= \left(\left(1 - \frac{1}{\sqrt{\kappa_j}}\right)\left(1 + \frac{1}{\sqrt{\kappa_j}}\right)\right)^{-1} \left(1 + \frac{1}{\sqrt{\kappa_j}}\right)^2 \\
 &= \frac{\sqrt{\kappa_j} + 1}{\sqrt{\kappa_j} - 1}
 \end{aligned}$$

Außerdem setze:

$$h_j := \frac{\sqrt{\kappa_j} - 1}{\sqrt{\kappa_j} + 1}$$

Insgesamt folgt für die Berechnung des Tschebyscheff-Polynoms aus Satz 4.2 nach (4.46) und Lemma 4.6:

$$(4.48) \quad T_{m-j}(1 + 2\gamma_j) = \frac{1}{2} \left[h_j^{m-j} + \frac{1}{h_j^{m-j}} \right]$$

Da (4.48) quadriert und im Nenner der Abschätzung auftaucht, wird der Kehrwert von (4.48) betrachtet. Es gilt:

$$(4.49) \quad \frac{1}{T_{m-j}(1 + 2\gamma_j)} = \frac{2}{[h_j^{m-j} + h_j^{-(m-j)}]} = \frac{2h_j^{m-j}}{1 + h_j^{2(m-j)}} \leq 2h_j^{m-j}$$

beziehungsweise

$$(4.50) \quad \frac{1}{T_{m-j}(1 + 2\gamma_j)^2} \leq 4h_j^{2(m-j)}$$

Das Tschebyscheff-Polynom $T_{m-j}(1 + 2\gamma_j)$ ist groß für ein großes γ_j . Aus der Definition von γ_j und κ_j folgt die Beziehung:

$$(4.51) \quad \gamma_j = \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n} = \frac{\lambda_j - \lambda_{j+1}}{\lambda_j - \lambda_n + \lambda_{j+1} - \lambda_j} = \frac{1}{\kappa_j - 1}$$

Aus diesem Zusammenhang der beiden Größen folgt: Für $\kappa_j \approx 1$ ist $\gamma_j \approx \infty$.

Nach (4.49) ist zu erkennen, dass für kleines h_j auch $1/T_{m-j}(1 + 2\gamma_j)$ klein ist. Außerdem gilt für $\kappa_j \approx 1$:

$$h_j = \frac{\sqrt{\kappa_j} - 1}{\sqrt{\kappa_j} + 1} \approx 0$$

Betrachte nun $\kappa_j = \frac{\lambda_j - \lambda_n}{\lambda_{j+1} - \lambda_{j+1}}$ für die speziellen Eigenwerte

$$\lambda_k = c_1 e^{-c_2(k-1)^\alpha} \text{ für } \alpha \in (0, 1] \text{ und } c_1, c_2 > 0.$$

Bemerkung 4.11. *Offensichtlich ist auch κ_j unabhängig von c_1 , und somit auch h_j .*

Aus der Definition von κ_j ergeben sich die folgenden Ungleichungen.

$$(4.52) \quad 1 < \kappa_1 = \frac{e^{c_2} - \frac{1}{e^{c_2(n-1)^\alpha}}}{e^{c_2} - 1} < \frac{e^{c_2}}{e^{c_2} - 1}$$

$$(4.53) \quad 1 < \kappa_j = \frac{e^{c_2(j^\alpha - (j-1)^\alpha)} - e^{c_2(j^\alpha - (n-1)^\alpha)}}{e^{c_2(j^\alpha - (j-1)^\alpha)} - 1} < \frac{e^{c_2(j^\alpha - (j-1)^\alpha)}}{e^{c_2(j^\alpha - (j-1)^\alpha)} - 1} \quad \text{für } j > 1$$

Im Fall $\alpha = 1$ gilt damit wie für κ_1 :

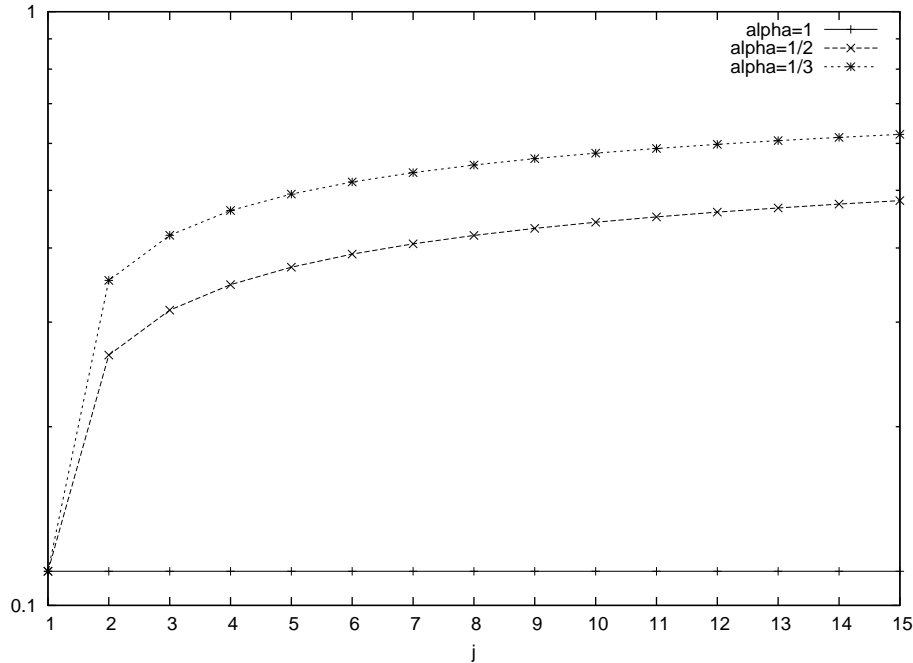
$$(4.54) \quad \kappa_j < \frac{e^{c_2}}{e^{c_2} - 1}$$

Aus dem Beweis von Korollar 4.2 folgt für $j > 1$:

Bemerkung 4.12. *Für kleine α wächst κ_j und somit auch h_j . Da aufgrund der vorherigen Überlegungen die Konvergenz für ein $\kappa_j \approx 1$ besonders gut ist, verschlechtert sich die Konvergenz für kleine α . Wie in den nächsten Abbildungen zu sehen ist, sind aber auch für kleinere α zumindest für die ersten Eigenwerte immer noch relativ gute Approximationen zu erwarten. Auch für κ_1 gilt: κ_1 ist für $n \ll m$ praktisch unabhängig von α .*

Die folgende Grafik zeigt h_j für $\alpha = 1, 1/2$ und $1/3$, wobei die Matrixdimension $n = 1000$ ist.

Abbildung 4.3: h_j für $j = 1$ bis 15 und $\alpha = 1, 1/2, 1/3$



Das bisher in Satz 4.2 betrachtete Tschebyscheff-Polynom im Nenner sorgt für die Konvergenz der Eigenwerte. Für μ_j mit $j > 1$ kommt jedoch noch das Produkt

$$(4.55) \quad \prod_{i=1}^{j-1} \frac{\mu_i - \lambda_n}{\mu_i - \lambda_j}$$

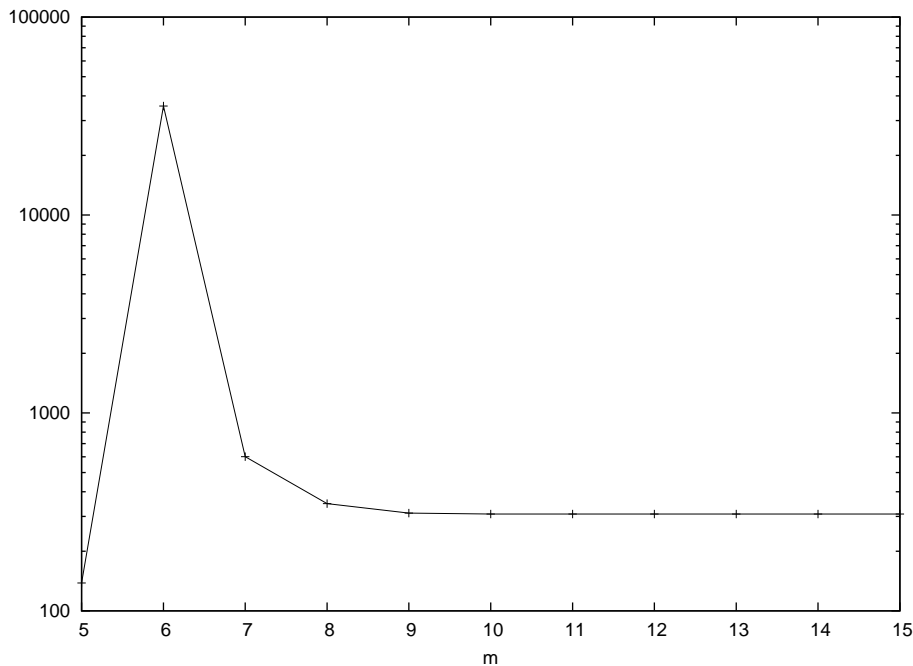
hinzu. Dieses Produkt ist dafür verantwortlich, dass die (später noch graphisch dargestellte) Fehlerschranke aus Satz 4.2 für kleine α und spätere Eigenwerte zu Beginn sehr große Werte annimmt. Dies erklärt sich durch den Nenner, der aus dem Produkt über $\mu_i - \lambda_j$ besteht. Für fortlaufendes j wächst auch das Produkt. Je kleiner α in der Formel (4.39) gewählt wird, desto dichter liegen die Eigenwerte und infolgedessen auch die Ritz-Werte und die Eigenwerte beieinander. Das bedeutet, der Nenner $\mu_i - \lambda_j$ ist zum Beispiel für $i = j - 1, j - 2$ sehr viel kleiner als der Zähler $\mu_i - \lambda_n$, da $\lambda_n \approx 0$ für große n . Desweiteren ist bei den (in Kapitel 5 beschriebenen) Matrizen, auf die das Lanczos-Verfahren angewandt wurde, zu beobachten, dass der zuletzt neu hinzugekommene Ritz-Wert μ_{j-1} zu Beginn seiner Iteration nahe bei Null liegt. Deshalb ist es auch möglich, dass μ_{j-1} kleiner als λ_j ist. Dies ist ebenfalls auf die für kleine α sehr geringen Abstände der Eigenwerte, und infolgedessen langsame Konvergenz der Ritz-Werte zurückzuführen. Durch dieses Schwanken von μ_{j-1} um λ_{j-1} kann es passieren, dass beim ersten Auftauchen von μ_{j-1} der Abstand zu λ_j größer ist, als im nächsten Iterationsschritt.

Das bedeutet, dass der Nenner von (4.55) größer, und somit das Produkt insgesamt kleiner ist als im nächsten Schritt. μ_{j-1} konvergiert erst langsam gegen λ_{j-1} , und infolgedessen ist der Abstand zwischen μ_{j-1} und λ_j in den nächsten Iterationen immer noch sehr klein, und das Produkt sehr groß. Diese Überlegungen erklären den (in der nächsten Grafik für das Beispiel λ_5 , $\alpha = 1/2$, $n = 1000$) dargestellten anfänglichen Anstieg und die Größe des Produktes zu μ_j in seinen ersten Iterationsschritten, welches dann in den darauffolgenden Schritten langsam kleiner wird und sich schließlich konstant verhält. Da in die Fehlerabschätzung das quadrierte Produkt einfließt, wird auch in der folgenden Grafik das Produkt zum Quadrat gezeichnet. In dem Beispiel ist zu sehen, dass schon bei der Berechnung des Fehlers von μ_5 (für Eigenwerte nach Formel (4.39) mit $\alpha = 1/2$) das Produkt

$$(4.56) \quad \left(\prod_{i=1}^4 \frac{\mu_i - \lambda_{1000}}{\mu_i - \lambda_5} \right)^2$$

einen Wert von über 30000 annimmt. Auch das zuvor beschriebene Phänomen, dass der im ersten Schritt schlecht approximierte Ritz-Wert für einen kleineren Startwert des Produkts sorgt, wird deutlich.

Abbildung 4.4: Produktquadrat (4.56) aus Satz 4.2 zur Fehlerabschätzung von λ_5



Wie in der oberen Grafik zu sehen ist, pendelt sich das betrachtete Produkt bei wachsender Iterationszahl bei einer konstanten Größe ein. Um unter der Voraussetzung exponentiell abfallender Eigenwerte eine Fehlerabschätzung zu erhalten, die eine Abschätzung

für die Konvergenzrate unabhängig von den vorangegangenen Ritz-Werten liefert, wird Satz 4.2 noch einmal untersucht. Das dort definierte γ_j sowie die Winkel zwischen x und x_j beziehungsweise X^j sind offensichtlich unabhängig von den Ritz-Werten. Das Ziel ist also, das Produkt (4.55) durch eine Konstante auszudrücken, in der μ_1, \dots, μ_{j-1} nicht vorkommen. Wenn dies gezeigt werden kann, so muß nur die Anzahl der Iterationen groß genug gewählt werden. Durch das Tschebyscheff-Polynom im Nenner der Konvergenzaussage wird dann der Approximationsfehler mit wachsender Iterationszahl verringert.

Wie sich schon bei der genaueren Betrachtung des Produkts (4.56) herausgestellt hat, existieren (nach wenigen Iterationsschritten) Ritz-Werte μ_i , für die **nicht** gilt: $\lambda_{i+1} \leq \mu_i$. Deshalb ist es nicht möglich, eine obere Schranke für den Nenner

$$\prod_{i=1}^{j-1} (\mu_i - \lambda_j)$$

von (4.55) anzugeben. Der Kurvenverlauf in Abbildung 4.4, nach dem anzunehmen ist, dass das Produkt (4.55) (in Folge der Konvergenz der Ritz-Werte) relativ schnell gegen das Produkt über die echten Eigenwerte konvergiert, gibt den Anstoß zu folgendem Lemma.

Lemma 4.7. *Seien μ_1, \dots, μ_j die mit dem Lanczos-Verfahren berechneten Ritz-Werte der Matrix $A \in \mathbb{R}^{n \times n}$, mit $1 \leq j \leq k < n$. Definiere*

$$\tilde{K}_j^{(m)} := \prod_{i=1}^{j-1} \frac{\mu_i - \lambda_n}{\mu_i - \lambda_j}$$

wobei μ_1, \dots, μ_{j-1} die ersten $j - 1$ Ritz-Werte des m -ten Iterationsschrittes sind. Außerdem setze

$$K_j := \prod_{i=1}^{j-1} \frac{\lambda_i - \lambda_n}{\lambda_i - \lambda_j}$$

Dann gilt für alle $j \in \{1, \dots, k\}$:

$$\lim_{m \rightarrow n} \tilde{K}_j^{(m)} = K_j$$

Beweis. Durch Induktion über k :

Induktionsvoraussetzung: Sei $k = 1$.

Nach Satz 4.2 gilt

$$0 \leq \lambda_1 - \mu_1 \leq (\lambda_1 - \lambda_n) \left[\frac{\sin \angle(x, X^1)}{\cos \angle(x, x^1)} \frac{1}{T_{m-1}(1 + 2\gamma_1)} \right]^2$$

Also ist $\tilde{K}_j^{(m)} = 1 = K_j$ für alle $j \leq k$ und alle $m \in \mathbb{N}$

Induktionsannahme: Gelte die Behauptung für ein $j \in \{1, \dots, k\}$.

Induktionsschluß: Betrachte $\tilde{K}_j^{(m)}$ für $1 \leq j \leq k+1$. Nach Definition von $\tilde{K}_j^{(m)}$ und K_j gilt:

$$\tilde{K}_{k+1}^{(m)} = \tilde{K}_k^{(m)} \cdot \frac{\mu_k - \lambda_n}{\mu_k - \lambda_{k+1}} \quad \text{und} \quad K_{k+1} = K_k \cdot \frac{\lambda_k - \lambda_n}{\lambda_k - \lambda_{k+1}}$$

Wird die Induktionsannahme auf $\tilde{K}_k^{(m)}$ angewendet, so folgt:

$$\begin{aligned} \lim_{m \rightarrow n} \tilde{K}_{k+1}^{(m)} &= \lim_{m \rightarrow n} \left(\tilde{K}_k^{(m)} \cdot \frac{\mu_k - \lambda_n}{\mu_k - \lambda_{k+1}} \right) \\ &= \lim_{m \rightarrow n} \tilde{K}_k^{(m)} \cdot \lim_{m \rightarrow n} \frac{\mu_k - \lambda_n}{\mu_k - \lambda_{k+1}} \\ &\stackrel{I.A.}{=} K_k \cdot \lim_{m \rightarrow n} \frac{\mu_k - \lambda_n}{\mu_k - \lambda_{k+1}} \end{aligned}$$

Für die Ritz-Werte μ_1, \dots, μ_k des m -ten Iterationsschrittes gilt nach Satz 4.2

$$0 \leq \lambda_j - \mu_j \leq (\lambda_j - \lambda_n) \left[\frac{\sin \angle(x, X^j)}{\cos \angle(x, x^j)} \frac{1}{T_{m-j}(1 + 2\gamma_j)} \cdot \tilde{K}_j^{(m)} \right]^2$$

Nach Induktionsannahme ist $\tilde{K}_j^{(m)} = K_j$ für alle $j < k$, und somit folgt die Konvergenz der Ritz-Werte μ_1, \dots, μ_k aus dem obigen Satz, da für wachsendes m das Tschebyscheff-Polynom im Nenner immer größer wird. Das heißt, es gilt:

$$\lim_{m \rightarrow n} \frac{\mu_k - \lambda_n}{\mu_k - \lambda_{k+1}} = \frac{\lambda_k - \lambda_n}{\lambda_k - \lambda_{k+1}}$$

Damit folgt für $\tilde{K}_{k+1}^{(m)}$:

$$\lim_{m \rightarrow n} \tilde{K}_{k+1}^{(m)} = K_k \cdot \lim_{m \rightarrow n} \frac{\mu_k - \lambda_n}{\mu_k - \lambda_{k+1}} = K_k \cdot \frac{\lambda_k - \lambda_n}{\lambda_k - \lambda_{k+1}} = K_{k+1}$$

Damit ist die Konvergenz von $\tilde{K}_k^{(m)}$ gegen K_k gezeigt. □

Mit Hilfe des obigen Lemmas und den vorangegangenen Überlegungen folgt nun eine vereinfachte Konvergenzaussage für die nach (4.39) gebildeten Eigenwerte.

Insbesondere kann der letzte Eigenwert λ_n in der Abschätzung vernachlässigt werden, da dieser für ein großes n dicht bei Null liegt. So erhält man aus Satz 4.2 eine speziell für exponentiell abfallende Eigenwerte gültige Aussage über den relativen Fehler.

Satz 4.4. Seien $\lambda_1 \geq \dots \geq \lambda_n$ definiert wie in (4.39), und seien $\mu_1, \dots, \mu_j, 1 \leq j \leq k$ die Ritz-Werte des m -ten Schritt des Lanczos-Verfahrens. Sei $m_0 < m$ der Iterationsschritt, für den gilt: $\tilde{K}_j^{(m_0)} = K_j$ mit $\tilde{K}_j^{(m_0)}$ und K_j aus Lemma 4.7.

$$0 \leq \lambda_j - \mu_j \leq \lambda_j \cdot K_j^2 w_j^2 \cdot 4 h_j^{2(m-j)}$$

Mit den Konstanten

$$h_j = \frac{\sqrt{\kappa_j} - 1}{\sqrt{\kappa_j + 1}}, \quad \kappa_j = \frac{e^{c_2(j^\alpha - (j-1)^\alpha)} - \frac{1}{e^{c_2((n-1)^\alpha - j^\alpha)}}}{e^{c_2(j^\alpha - (j-1)^\alpha)} - 1}, \quad w_j = \frac{\sin \angle(x, X^j)}{\cos \angle(x, x^j)}$$

wobei $1 < \kappa_j < \frac{e^{c_2(j^\alpha - (j-1)^\alpha)}}{e^{c_2(j^\alpha - (j-1)^\alpha)} - 1}$

Beweis . Man schätze $1/T_j(1 + 2\gamma_j)$ aus Satz 4.2 mit Hilfe von (4.50) ab und verwende (4.53) und Lemma 4.7. □

Bemerkung 4.13. Die in Satz 4.4 gegebene obere Schranke ist für ein hinreichend großes m unabhängig von der Matrixdimension n und den Ritz-Werten μ_1, \dots, μ_{j-1} . Sie beschreibt den relativen Fehler des Eigenwerts λ_j .

Für den Spezialfall $\alpha = 1$ ergibt sich die neue Fehlerabschätzung zu:

Korollar 4.3. Seien die Voraussetzungen wie in Satz 4.4. Für die Ritz-Werte μ_1, \dots, μ_j zu den Eigenwerten nach (4.39) mit $\alpha = 1$ gilt:

$$0 \leq \lambda_j - \mu_j \leq \frac{1}{e^{j-1}} \cdot K_j^2 w_j^2 * 4 h_j^{2(m-j)}$$

Mit den Konstanten

$$h_j = \frac{\sqrt{\kappa_j} - 1}{\sqrt{\kappa_j + 1}}, \quad \kappa_j = \frac{e^{c_2} - \frac{1}{e^{c_2(n-1-j)}}}{e^{c_2} - 1}, \quad w_j = \frac{\sin \angle(x, X^j)}{\cos \angle(x, x^j)}$$

wobei $1 < \kappa_j < \frac{e^{c_2}}{e^{c_2} - 1}$

Insbesondere gilt für λ_1 :

$$0 \leq \lambda_1 - \mu_1 \leq 4 h_1^{2(m-1)} w_1^2$$

In dem vorangegangenen Abschnitt wurde die Konvergenztheorie für die vorgegebenen Eigenwerte untersucht und einige Vereinfachungen und Abschätzungen vorgestellt. Aber wie sieht die Konvergenzaussage des Lanczos-Verfahrens für die Eigenwerte aus (4.39) tatsächlich in Zahlen aus? Um zum Beispiel die Abschätzung aus Satz 4.2 konkret zu berechnen, muss auch w_j bestimmt werden. Das heißt, der Kosinus des Winkels zwischen dem Startvektor x und Eigenvektor x^j beziehungsweise der Sinus des Winkels zwischen x und Eigenraum X^j muss ausgerechnet werden. Dazu dient das folgende Lemma.

Lemma 4.8. Zu beliebigem $x \in \mathbb{R}^n$, $x \neq 0$, sei z_x mit Hilfe der Eigenvektoren x^i von A definiert durch:

$$z_x := \sum_{i=1}^j \langle x, x^i \rangle x^i.$$

Dann gilt:

$$\cos \angle(x, X^j) = \frac{\|z_x\|}{\|x\|}$$

Beweis . Nach Definition 2.3 beziehungsweise Definition 2.2 ist

$$(4.57) \quad \cos \angle(x, X^j) = \max_{z \in X^j} \cos \angle(x, z) = \max_{z \in X^j} \frac{|\langle x, z \rangle|}{\|x\| \|z\|}$$

Weil x^1, \dots, x^n eine ONB des \mathbb{R}^n bilden, gilt für $x \in \mathbb{R}^n$ und $z \in X^j = \text{span} \{x^1, \dots, x^j\}$:

$$(4.58) \quad x = \sum_{i=1}^n \langle x, x^i \rangle x^i \quad \text{und} \quad z = \sum_{k=1}^j \langle z, x^k \rangle x^k$$

Mit den beiden Gleichungen aus (4.58) folgt, da die Eigenvektoren x^i orthonormal sind:

$$(4.59) \quad \begin{aligned} \langle x, z \rangle &= \left\langle \sum_{i=1}^n \langle x, x^i \rangle x^i, \sum_{k=1}^j \langle z, x^k \rangle x^k \right\rangle = \sum_{k=1}^j \langle \langle x, x^k \rangle x^k, \langle z, x^k \rangle x^k \rangle \\ &= \sum_{k=1}^j \langle x, x^k \rangle \langle z, x^k \rangle = \sum_{k=1}^j \langle z, \langle x, x^k \rangle x^k \rangle \\ &= \left\langle z, \sum_{k=1}^j \langle x, x^k \rangle x^k \right\rangle = \langle z, z_x \rangle \end{aligned} \quad \text{nach Definition von } z_x$$

Mit $\cos \angle(z_x, z) = \frac{\langle z_x, z \rangle}{\|z_x\| \|z\|}$ und (4.59) kann (4.57) weiter ausgerechnet werden:

$$(4.60) \quad \cos \angle(x, X^j) = \max_{z \in X^j} \frac{|\langle x, z \rangle|}{\|x\| \|z\|} = \frac{1}{\|x\|} \max_{z \in X^j} \frac{|\langle z_x, z \rangle|}{\|z\|} = \frac{\|z_x\|}{\|x\|} \max_{z \in X^j} \cos \angle(z_x, z)$$

Da $\cos \angle(z_x, z) \leq 1$ für alle $z \in X^j$ ist, und $z_x \in X^j$ mit $\cos \angle(z_x, z) = 1$ gilt, folgt:

$$\max_{z \in X^j} \cos \angle(z_x, z) = \cos \angle(z_x, z_x) = 1$$

Mit (4.59) gilt somit die Behauptung. □

Um die Fehlerabschätzung für ein $j \in \{1, \dots, m\}$ berechnen zu können, muss

$$\frac{\sin^2 \angle(x, X^j)}{\cos^2 \angle(x, x^j)}$$

ausgerechnet werden. Für den Nenner ist nach Definition das Skalarprodukt $\langle x, x^j \rangle$ und die Norm von x zu berechnen ($\|x_j\| = 1$). Für den Zähler wird obiges Lemma verwendet. Es gilt

$$\sin^2 \angle(x, X^j) := 1 - \cos^2 \angle(x, X^j) \stackrel{\text{Lemma 4.8}}{=} 1 - \frac{\langle z_x, z_x \rangle}{\langle x, x \rangle}$$

wobei z_x wie in Lemma 4.8 definiert sei, und somit leicht mit Hilfe der Parsevalschen Gleichung zu bestimmen ist.

Die folgenden Abbildungen zeigen nun den nach Satz 4.2 entstandenen relativen Fehler der Eigenwerte λ_1 , λ_3 und λ_7 , und zwar für die auch bisher betrachteten verschiedenen Werte von $\alpha = 1, 1/2, 1/3$ und $c_1 = c_2 = 1$. Da wie schon gezeigt der Parameter c_1 keinen Einfluss auf die Konvergenz hat, werden nur Beispiele zu $c_1 = 1$ betrachtet. Weitere Bilder zu $c_2 \neq 1$ befinden sich im Anhang. Die dafür benötigten Ritz-Werte stammen aus dem in Kapitel 5 beschriebenen Algorithmus 5.3.

Da $\lambda_1 = 1$ und $\lambda_2 = \frac{1}{e^{c_2}}$ für alle α gilt, ist die Konvergenz für λ_1 in allen drei Fällen gleich gut, da insbesondere der Abstand zwischen λ_1 und λ_2 gleichbleibend ist. Deshalb kann die zu λ_1 gehörende Fehlerkurve als Vergleich zu den später konvergierenden Eigenwerten λ_3 und λ_7 dienen.

Es wurde für den Fall $\alpha = 1$ das Verhalten der 3 Eigenwerte für die Iterationsschritte $m = 1$ bis $m = 23$ betrachtet, da das verwendete Verfahren aufgrund von numerischer Instabilität an diesem Punkt abbricht (Erläuterung siehe Kapitel 5). In den anderen Fällen wurde die Konvergenz für $m = 45$ Iterationsschritte betrachtet, welche auf der x -Achse abgebildet sind.

In allen drei Grafiken ist exponentielle Konvergenz zu erkennen (die Achse der Fehlerschranken ist logarithmisch dargestellt). Bei $\alpha = 1$ konvergieren auch die kleineren Eigenwerte λ_3 und λ_7 genauso schnell wie λ_1 .

Eine deutliche Verschlechterung ergibt sich hingegen bei der Betrachtung der Eigenwerte λ_3 und λ_7 in den Fällen $\alpha = 1/2$ und $\alpha = 1/3$ im Gegensatz zu λ_1 . Die vorangegangene Theorie, die langsamere Konvergenz für kleineres α vermuten lässt, wird hier bestätigt.

Abbildung 4.5: Fehlerschranke λ_1 , λ_3 und λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1$

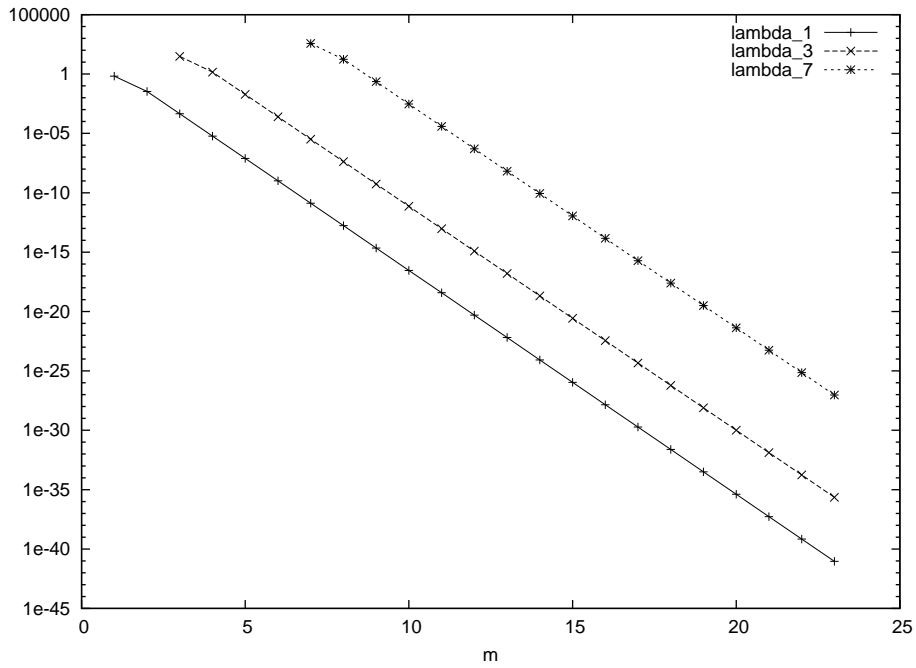


Abbildung 4.6: Fehlerschranke λ_1 , λ_3 und λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/2$

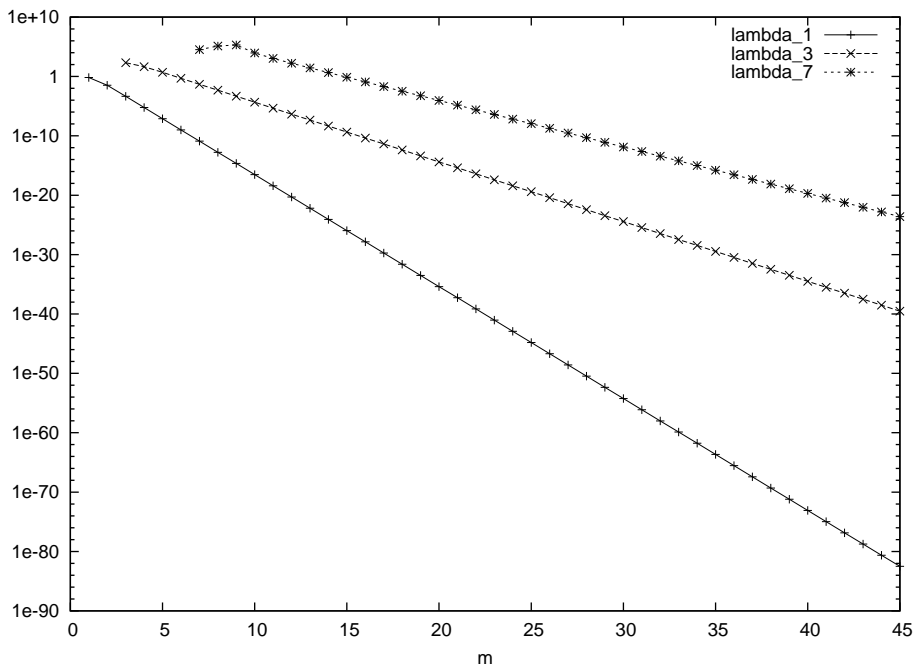
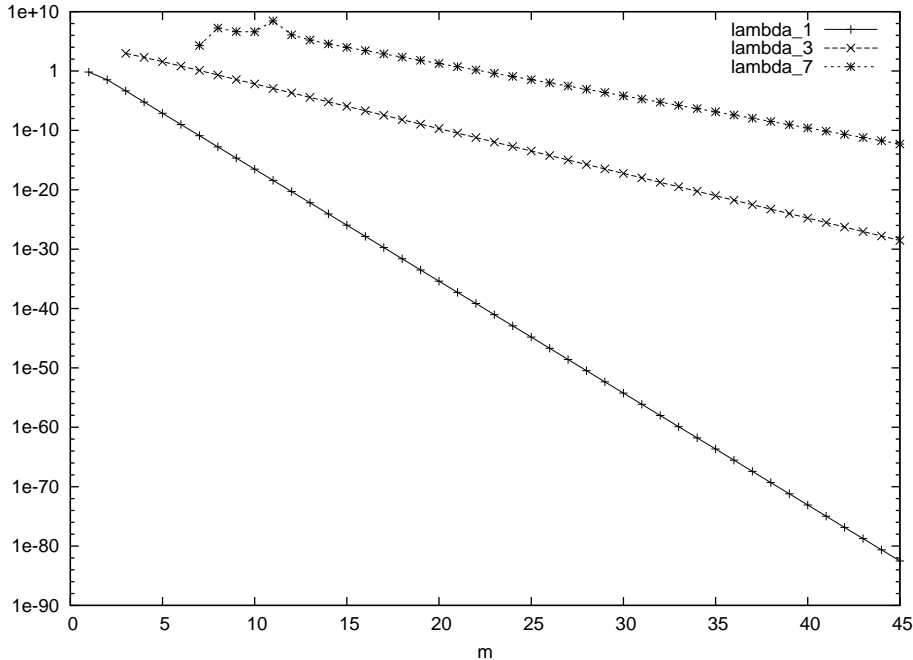


Abbildung 4.7: Fehlerschranke λ_1 , λ_3 und λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/3$



An den beiden letzten Grafiken ist zu bemerken, dass zu Beginn der Berechnung insbesondere für λ_7 und $\alpha = 1/3$ sehr unterschiedliche große Werte auftauchen. Diese sind auf das schon beschriebene Produkt

$$\prod_{i=1}^{j-1} \frac{\mu_i - \lambda_n}{\mu_i - \lambda_j}$$

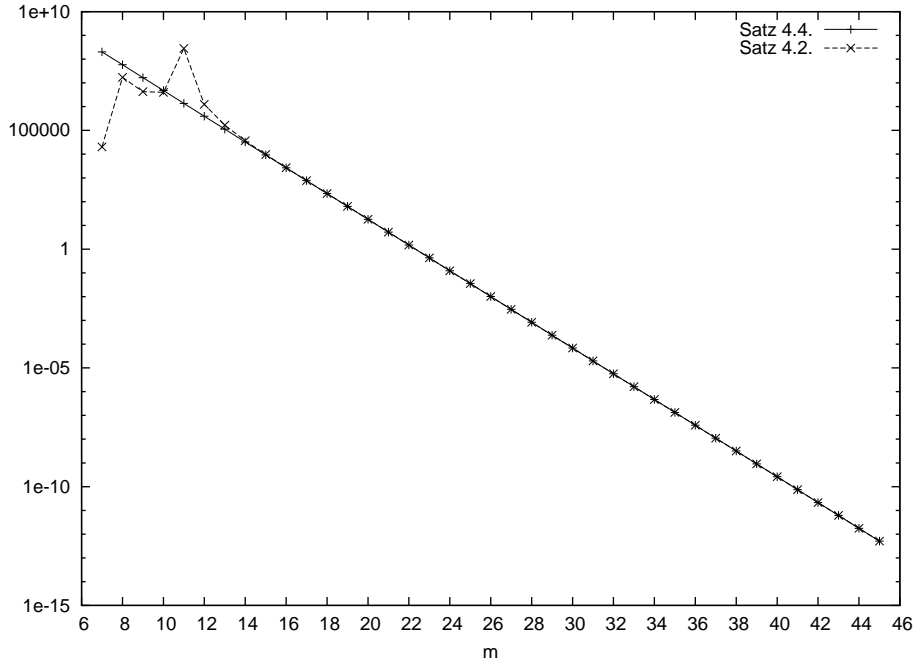
welches für $j > 1$ in Satz 4.2 entsteht, zurückzuführen. Wie festgestellt wurde, konvergiert jedoch das Produkt bei fortlaufender Iteration zu der Konstanten

$$(4.61) \quad K_j := \prod_{i=1}^{j-1} \frac{\lambda_i - \lambda_n}{\lambda_i - \lambda_j}$$

Insgesamt ist also gute Konvergenz zu erwarten, besonders für $\alpha = 1$.

Dass die in Satz 4.4 vereinfachte Konvergenzaussage keine Verschlechterung der Abschätzung aus Satz 4.2 darstellt, zeigt der Vergleich der Kurven der zugehörigen oberen Schranken in der nächsten Grafik. Diese ist für das Beispiel λ_7 mit $\alpha = 1/3$ gezeichnet. Dadurch, dass das in Lemma 4.7 definierte \tilde{K}_j aus Satz 4.2 in Satz 4.4 durch das (ebenfalls in Lemma 4.7 definierte) K_j ersetzt wurde, ist die Kurve zur Schranke aus Satz 4.4 monoton fallend, im Gegensatz zum stark schwankenden Graph der ursprünglichen Konvergenzaussage.

Abbildung 4.8: Fehlerschrankenvergleich von λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/3$



4.2.2 Konvergenzbetrachtung mit Hilfe der Ritz-Vektoren

Da in dem Abschnitt zur allgemeinen Konvergenz auch Satz 4.3 bewiesen wurde, wird nun zumindest für den Spezialfall $\alpha = 1$ auch diese Abschätzung noch einmal für die exponentiell abfallenden Eigenwerte untersucht. Es werden dieselben Größen h_j und w_j wie in Korollar 4.3 verwendet. Das Produkt besteht hier jedoch auch für den Iterationsbeginn aus den echten Eigenwerten $\lambda_1, \dots, \lambda_j$ und ist nicht abhängig von den Ritz-Werten μ_1, \dots, μ_{j-1} . Da aber dafür noch die Summe über $(\lambda_i - \lambda_n) \sin^2 \angle(x^i, z^i)$ hinzukommt, liefert Satz 4.3 eine recht komplizierte Fehlerabschätzung der Eigenwerte. Unter der Voraussetzung $\alpha = 1$ lassen sich einige Teile dieses Satzes jedoch vereinfachen.

Betrachte zunächst das Produkt, welches in 4.3 quadratisch vorkommt:

$$(4.62) \quad \prod_{k=1}^{j-1} \frac{\lambda_k - \lambda_n}{\lambda_k - \lambda_j} < \prod_{k=1}^{j-1} \frac{\lambda_k}{\lambda_k - \lambda_j} = \prod_{k=1}^{j-1} \frac{c_1 e^{-c_2 (k-1)^\alpha}}{c_1 (e^{-c_2 (k-1)^\alpha} - e^{-c_2 (j-1)^\alpha})} = \prod_{k=1}^{j-1} \frac{1}{1 - \frac{e^{c_2 (k-1)^\alpha}}{e^{c_2 (j-1)^\alpha}}}$$

Für $\alpha = 1$ gilt also:

$$(4.63) \quad \prod_{k=1}^{j-1} \frac{\lambda_k - \lambda_n}{\lambda_k - \lambda_j} < \prod_{k=1}^{j-1} \frac{1}{1 - e^{c_2 (k-j)}}$$

In Satz 4.3 kommt die Summe über $(\lambda_i - \lambda_n) \sin^2 \angle(x^i, z^i)$ sowohl in der Abschätzung der Eigenwerte, als auch in der des Fehlerwinkels vor. Auch hier ist die Voraussetzung $\alpha = 1$ eine Vereinfachung. Es gilt:

$$(4.64) \quad \sum_{i=1}^{j-1} (\lambda_i - \lambda_n) \sin^2 \angle(x^i, z^i) \leq \sum_{i=1}^{j-1} (\lambda_i - \lambda_n) < \sum_{i=1}^{j-1} \lambda_i$$

Für die Summe über die Eigenwerte gilt:

$$(4.65) \quad \begin{aligned} \sum_{i=1}^{j-1} \lambda_i &= \sum_{i=1}^{j-1} \frac{c_1}{e^{c_2(i-1)}} && \text{nach (4.39)} \\ &= c_1 \sum_{i=0}^{j-2} \frac{1}{e^{c_2 i}} && \text{Geometrische Reihe} \\ &= c_1 \frac{1 - \frac{1}{e^{c_2(j-1)}}}{1 - \frac{1}{e^{c_2}}} && \text{Wert der geometrischen Reihe} \\ &= c_1 \frac{e^{c_2} - e^{-c_2 j}}{e^{c_2} - 1} \end{aligned}$$

Betrachte nun insgesamt Satz 4.3 für $\alpha = 1$.

Satz 4.5. Sei $\alpha = 1$ und w_j, h_j wie in Satz 4.4. Dann gilt für die Eigenwerte λ_j aus (4.39):

$$\begin{aligned} 0 \leq \lambda_j - \mu_j &\leq \lambda_j \cdot L_j^2 w_j^2 \cdot 4 h_j^{2(m-j)} + c_1 \left(\frac{e^{c_2} - e^{-c_2 j}}{e^{c_2} - 1} \right) \\ \text{und } \sin^2 \angle(z^i, x^i) &\leq \frac{e^{c_2 i}}{e^{c_2} - 1} \left[e^{-c_2(i-1)} - \frac{\mu_i}{c_1} + \left(\frac{e^{c_2} - e^{-c_2(i-1)}}{e^{c_2} - 1} \right) \right] \\ \text{mit } L_j &= \prod_{k=1}^{j-1} \frac{1}{1 - e^{c_2(k-j)}} \end{aligned}$$

Beweis . Wie in Satz 4.4 kann $1/T_j(1 + 2\gamma_j)$ durch (4.50) abgeschätzt werden. Das Produkt der Eigenwerte in Satz 4.3 wird durch die rechte Seite der Ungleichung (4.62) ersetzt. Die Summe $\sum_{i=1}^{j-1} (\lambda_i - \lambda_n) \sin^2 \angle(x^i, z^i)$ aus Satz 4.3 wird unter Verwendung von (4.64) und (4.65) abgeschätzt. Da $\lambda_n \approx 0$ für große n , kann auch in diesem Satz λ_n vernachlässigt werden. Außerdem wurden in der Abschätzung des Winkels die Eigenwerte nach ihrer Definition (4.39) mit $\alpha = 1$ eingesetzt. □

Das Ziel der Entwicklung von Satz 4.3 war es, die Ritz-Werte aus der Konvergenzaussage der Eigenwerte zu eliminieren. Dies ist zwar gelungen, aber nun hat man den Nachteil, dass der Fehler des Winkels zwischen den Ritz-Vektoren und Eigenvektoren mit eingeht. Da außerdem gezeigt wurde, dass das Produkt (4.55) aus Satz 4.2 gegen das Produkt mit den echten Eigenwerten aus Satz 4.3 konvergiert, ist Satz 4.2 für die Konvergenzbetrachtungen vorzuziehen.

4.2.3 Konvergenzbetrachtung für abweichende Eigenwerte

Bisher sind nur Eigenwerte betrachtet worden, die genau auf den von den Parametern c_1, c_2, α vorgegebenen Grenzen liegen:

$$\lambda_j = \frac{c_1}{e^{c_2(j-1)^\alpha}}$$

Was für eine Konvergenz kann für Eigenwerte, die in einer Umgebung um die vorgegebenen Grenzen liegen, erreicht werden? Im Folgenden wird ein konkretes Beispiel entwickelt, mit Eigenwerten, die nicht mehr ganz exponentiell abfallend sind. Betrachte also Eigenwerte $\tilde{\lambda}_j$, die in einer ϵ -Umgebung der ursprünglichen Eigenwerte liegen. Da diese in Abhängigkeit von c_1, c_2, α und wachsendem j schnell abfallen, ist die Umgebung relativ zu λ_j zu wählen, das heißt, nun werden Eigenwerte betrachtet, für die gilt:

$$\lambda_j(1 - \epsilon) \leq \tilde{\lambda}_j \leq \lambda_j(1 + \epsilon) \quad \text{für ein } \epsilon > 0$$

Die obige Definition ist gleichbedeutend mit: $\tilde{\lambda}_j \in [\lambda_j(1 - \epsilon), \lambda_j(1 + \epsilon)]$. Damit die Eigenwerte $\tilde{\lambda}_j$ sinnvoll in Bezug auf die Problemstellung exponentiell abfallender Eigenwerte gewählt sind, muß gewährleistet sein, dass nicht $\tilde{\lambda}_{j+1} > \tilde{\lambda}_j$ gilt. Das heißt, ϵ muß die folgende Ungleichung erfüllen:

$$\lambda_j(1 - \epsilon) - \lambda_{j+1}(1 + \epsilon) > 0$$

Durch Einsetzen der exponentiell abfallenden Eigenwerte λ_j aus (4.39) folgt aus obiger Ungleichung eine obere Schranke für ϵ :

$$(4.66) \quad \epsilon < \frac{e^{c_2 j^\alpha} - e^{c_2(j-1)^\alpha}}{e^{c_2 j^\alpha} + e^{c_2(j-1)^\alpha}} =: S_j \quad \text{für alle } j \in \{1, \dots, m\}$$

Bemerkung 4.14. Aus obiger Definition ist zu erkennen, dass die Schranke S_j kleiner wird für wachsendes j .

Für den Spezialfall $\alpha = 1$ folgt aus (4.66):

$$(4.67) \quad \epsilon < \frac{e^{c_2} - 1}{e^{c_2} + 1} =: S$$

Es zeigt sich, dass ϵ für $\alpha = 1$ unabhängig von j ist.

Was bedeuten die abgeänderten Eigenwerte $\tilde{\lambda}_j$ für die Konvergenz des Lanczos-Verfahrens? Da diese, wie in Satz 4.2 zu erkennen ist, im Wesentlichen von dem Wert κ_j abhängt, wird im Folgenden $\tilde{\kappa}_j$ für die nun nicht mehr ganz exponentiell abfallenden Eigenwerte betrachtet. Für die Konvergenzrate $\tilde{\kappa}_j$ ergibt sich:

$$(4.68) \quad \tilde{\kappa}_j = \frac{\tilde{\lambda}_j - \tilde{\lambda}_n}{\tilde{\lambda}_j - \tilde{\lambda}_{j+1}} \leq \frac{\lambda_j(1 + \epsilon) - \lambda_n(1 - \epsilon)}{\lambda_j(1 - \epsilon) - \lambda_{j+1}(1 + \epsilon)}$$

Bemerkung 4.15. Nach (4.68) gilt offensichtlich $\tilde{\kappa}_j \rightarrow \kappa_j$ für $\epsilon \rightarrow 0$.

Um ein konkretes Beispiel mit den neu definierten Eigenwerten zu konstruieren, an dem die Konvergenz des Lanczos-Verfahrens untersucht werden kann, wird der Einfachheit halber wieder der Spezialfall $\alpha = 1$ betrachtet. Dazu muss ϵ nach (4.67) gewählt werden. Setze:

$$(4.69) \quad \epsilon^* := \frac{1}{2}S = \frac{1}{2} \frac{e^{c_2} - 1}{e^{c_2} + 1}$$

Für die Werte $\lambda_j = c_1 e^{-c_2(j-1)}$ und nach Ungleichung (4.68) ergibt sich für die Kondition $\tilde{\kappa}_j$ der Eigenwerte $\tilde{\lambda}_j$:

$$\begin{aligned} \tilde{\kappa}_j &< \frac{c_1 e^{-c_2(j-1)}(1 + \epsilon^*) - c_1 e^{-c_2(n-1)}(1 - \epsilon^*)}{c_1 e^{-c_2(j-1)}(1 - \epsilon^*) - c_1 e^{-c_2 j}(1 + \epsilon^*)} \\ &< \frac{c_1 e^{-c_2(j-1)}(1 + \epsilon^*)}{c_1 e^{-c_2(j-1)}(1 - \epsilon^*) - c_1 e^{-c_2 j}(1 + \epsilon^*)} \\ &= \frac{e^{c_2}(1 + \epsilon^*)}{e^{c_2}(1 - \epsilon^*) - (1 + \epsilon^*)} = \frac{e^{c_2}(1 + \frac{1}{2} \frac{e^{c_2} - 1}{e^{c_2} + 1})}{e^{c_2}(1 - \frac{1}{2} \frac{e^{c_2} - 1}{e^{c_2} + 1}) - (1 + \frac{1}{2} \frac{e^{c_2} - 1}{e^{c_2} + 1})} \\ &= \frac{e^{c_2}(2(e^{c_2} + 1)) + (e^{c_2} - 1)}{e^{c_2}(2(e^{c_2} + 1) - (e^{c_2} - 1)) - (2(e^{c_2} + 1) + e^{c_2} - 1)} \\ &= \frac{3e^{2c_2} + e^{c_2}}{e^{2c_2} - 1} \\ &= \frac{3e^{c_2} + 1}{e^{c_2} - e^{-c_2}} \end{aligned}$$

Vergleicht man diese obere Schranke mit $\frac{e^{c_2}}{e^{c_2} - 1}$, der Schranke des ursprünglichen κ_j für $\alpha = 1$ (siehe Ungleichung (4.52)), so ist wegen

$$\frac{e^{c_2}}{e^{c_2} - 1} < \frac{3e^{2c_2} + 1}{e^{c_2} - e^{-c_2}} \quad \text{für } c_2 > 0$$

eine Verschlechterung der Konvergenz zu erwarten.

Mit der Betrachtung der Eigenwerte $\tilde{\lambda}_j \in [\lambda_j (1 - \epsilon), \lambda_j (1 + \epsilon)]$ ist der theoretische Teil zu der Frage nach der Konvergenz von exponentiell abfallenden Eigenwerten, welche mit dem Lanczos-Verfahren berechnet wurden, abgeschlossen. Da wichtige Bestandteile dieser Konvergenztheorie nicht ohne die Verwendung von bereits praktisch mit dem Lanczos-Algorithmus berechneten Werten auskommen, wird in dem folgenden Kapitel die praktische Anwendung des Lanczos-Algorithmus untersucht. Wie dort festgestellt werden wird, sind die tatsächlich erreichten Ergebnisse zu den speziellen Eigenwerten sogar noch sehr viel besser, als es die vorangegangenen Grafiken vermuten lassen.

5 Anwendung des Lanczos-Verfahrens

In diesem abschließenden Kapitel werden Testfälle für das spezielle Singulärwertproblem dieser Arbeit diskutiert, und der verwendete Lanczos-Algorithmus beschrieben. Des Weiteren werden praktische Probleme, die bei der Anwendung des Lanczos-Verfahrens auftreten können, analysiert und infolgedessen eine verbesserte Variante des Algorithmus entwickelt. Anschließend werden die Ergebnisse des verbesserten Verfahrens grafisch dargestellt und mit der theoretischen Fehleranalyse aus Kapitel 4 verglichen.

5.1 Testmatrizen und Implementierung

Das ursprüngliche Problem dieser Arbeit sind Matrizen mit Singulärwertzerlegung

$$C = U \Sigma V^T \in \mathbb{R}^{r \times n}$$

wobei $U \in \mathbb{R}^{r \times r}$ und $V \in \mathbb{R}^{n \times n}$ orthonormale Matrizen sind und $\Sigma \in \mathbb{R}^{r \times n}$ eine Diagonalmatrix ist. Weiter soll für die Singulärwerte $\sigma_k = \Sigma_{kk}$ mit $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_{\min\{r,n\}} \geq 0$ gelten:

$$(5.1) \quad \sigma_k^2 = c_1 e^{-c_2 (k-1)^\alpha} =: \lambda_k \quad \text{mit } \alpha \in (0, 1], c_1, c_2 > 0$$

Hierbei seien λ_k die Eigenwerte der Matrix

$$A = C^T C = V D V^T \in \mathbb{R}^{n \times n} \quad \text{mit } D = \Sigma^T \Sigma = \text{diag}\{\sigma_1^2, \dots, \sigma_n^2\} = \text{diag}\{\lambda_1, \dots, \lambda_n\}.$$

Die orthonormalen Spalten der Matrix V sind die Eigenvektoren von A . Das Ziel ist nun, die Eigenpaare der Matrix A mit dem Lanczos-Verfahren zu berechnen, um zu testen, wie gut das Verfahren für die speziellen Eigenwerte funktioniert.

Wie erhält man nun geeignete Matrizen C , die die obigen Bedingungen erfüllen, um das Lanczos-Verfahren anzuwenden?

Gesucht sind Matrizen beliebiger Dimension, zu denen die speziellen Singulärwerte aus Σ vorgegeben werden können und außerdem die orthonormale Matrix V mit den Eigenvektoren von A bekannt ist.

Die hier verwendete Idee ist, eine Diagonalmatrix $\Sigma \in \mathbb{R}^{r \times n}$ mit beliebigem r und n vorzugeben, auf deren Diagonalen bereits die Singulärwerte $\sigma_k = \sqrt{\lambda_k}$ nach ihrer Definition (5.1) stehen. Dann kann von links mit einer orthonormalen Matrix $U \in \mathbb{R}^{r \times r}$ und von rechts mit einer orthonormalen Matrix $V \in \mathbb{R}^{n \times n}$ multipliziert werden.

Auf die so entstandene Matrix C , beziehungsweise $A = C^T C$, kann dann das im Folgenden vorgestellte Lanczos-Verfahren angewendet werden. Mit Hilfe der nun bekannten Eigenwerte und Eigenvektoren der Matrix kann auch der Approximationsfehler, den die praktische Anwendung des Lanczos-Verfahrens mit sich bringt, untersucht werden.

Da eine Diagonalmatrix leicht aufzustellen ist, sind nun orthonormale Matrizen gesucht, die möglichst nur durch die Vorgabe ihrer Dimension r beziehungsweise $n \in \mathbb{N}$ bestimmt sind. Die Suche führt zu den Tschebyscheff-Knoten und Tschebyscheff-Polynomen, die ein System von orthogonalen Funktionen bilden. Sei dazu zu $n \in \mathbb{N}$:

$$T_i(x) = \cos(i \arccos(x)) \quad \text{für } |x| < 1, i \in \{0, \dots, n-1\}$$

das i -te Tschebyscheff-Polynom und seien

$$x_j = \cos\left(\frac{j+0.5}{n}\pi\right), \quad j \in \{0, \dots, n-1\}$$

die Nullstellen des n -ten Tschebyscheff-Polynoms. Betrachte nun die Matrix:

$$T^n \in \mathbb{R}^{n \times n} \quad \text{mit } T_{ij}^n = T_i(x_j) \quad \text{für } i, j \in \{0, \dots, n-1\}$$

Für die Matrix $D_T^n := T_{ij}^{nT} T_{ij}^n \in \mathbb{R}^{n \times n}$ gilt das folgende Korollar:

Korollar 5.1.

$$D_{T_{ij}^n}^n = \begin{cases} 0 & \text{für } i \neq j \\ \frac{n}{2} & \text{für } i = j > 0 \\ n & \text{für } i = j = 0 \end{cases}$$

Beweis . Siehe [5], Satz 3.37. □

Also entpuppt sich T^n zwar noch nicht als eine orthonormale Matrix, aber immerhin ergibt das Produkt mit der Transponierten T^n eine reguläre Diagonalmatrix:

$$(5.2) \quad T_{ij}^{nT} T_{ij}^n = D_T^n = \begin{pmatrix} n & & & \\ & \frac{n}{2} & & \\ & & \ddots & \\ & & & \frac{n}{2} \end{pmatrix}$$

Leider gilt Korollar 5.1 nicht für die Umkehrung $T^n T^{nT}$, und da nach Matrizen gesucht wird, die $T^T = T^{-1}$ erfüllen, muß T^n noch weiter bearbeitet werden. Betrachte dazu die Diagonalmatrix

$$\Delta_n := (D_T^n)^{-1} = \begin{pmatrix} \frac{1}{n} & & & \\ & \frac{2}{n} & & \\ & & \ddots & \\ & & & \frac{2}{n} \end{pmatrix}$$

und

$$\Delta_n^{1/2} := \begin{pmatrix} \sqrt{\frac{1}{n}} & & & \\ & \sqrt{\frac{2}{n}} & & \\ & & \ddots & \\ & & & \sqrt{\frac{2}{n}} \end{pmatrix}$$

Aus der Gleichung (5.2) und der Definition von Δ_n beziehungsweise $\Delta_n^{1/2}$ folgt nach Multiplikation von $\Delta_n^{1/2}$ von beiden Seiten:

$$I_{n \times n} = D_T^n \Delta_n = \Delta_n^{1/2} D_T^n \Delta_n^{1/2}$$

Da $\Delta_n^{1/2}$ symmetrisch ist, folgt:

$$I_{n \times n} = (T^n \Delta_n^{1/2})^T (T^n \Delta_n^{1/2})$$

Weil $\Delta_n^{1/2}$ und T^n (nach Gleichung (5.2)) regulär sind, ist auch $T^n \Delta_n^{1/2}$ regulär, und somit gilt auch:

$$I_{n \times n} = (T^n \Delta_n^{1/2}) (T^n \Delta_n^{1/2})^T$$

Das heißt, mit $T^n \Delta_n^{1/2}$ hat man eine orthonormale Matrix gefunden, welche allein durch die Dimension n bestimmt ist. Nun lassen sich mit Hilfe der Diagonalmatrix $\Sigma \in \mathbb{R}^{r \times n}$ und der Konstruktion:

$$C = (T^r \Delta_r^{1/2}) \Sigma (T^n \Delta_n^{1/2})^T \in \mathbb{R}^{r \times n}$$

zu beliebigem r und $n \in \mathbb{N}$ Singulärwertzerlegungen erzeugen, so dass die symmetrische Matrix A

$$A = C^T C = (T^n \Delta_n^{1/2}) D (T^n \Delta_n^{1/2})^T \in \mathbb{R}^{n \times n}$$

die Eigenwerte $\lambda_1, \dots, \lambda_n$ nach Bedingung (5.1) auf der Diagonalmatrix $D = \Sigma^T \Sigma \in \mathbb{R}^{n \times n}$ enthält, und die Eigenvektoren die Spalten von $T^n \Delta_n^{1/2}$ bilden. Die Eigenpaare werden dann mit Hilfe des nachfolgend vorgestellten Algorithmus 5.3 berechnet.

Die verwendete Programmiersprache ist C. Für Operationen wie zum Beispiel Matrix-Vektor-Multiplikationen, Skalarprodukte und auch für das Lösen des Eigenproblems der Ritz-Matrix wurde die *Sun Performance Library* benutzt. Diese stellt hierzu einen speziellen, auf dem QR-Verfahren basierenden Algorithmus für symmetrische tridiagonale Matrizen zur Verfügung.

Das bereitgestellte QR-Verfahren gibt die Eigenwerte (und orthonormale Eigenvektoren) von $B \in \mathbb{R}^{m \times m}$ zwar der Größe ihrer Beträge nach sortiert aus, aber die kleinen Eigenwerte zuerst, also $\lambda_m, \dots, \lambda_1$. Da dies nicht der ursprünglichen Reihenfolge $\lambda_1 \geq \dots \geq \lambda_m$

entspricht, wird der Algorithmus nicht auf A , sondern auf $-A$ angewendet und auf Grund dessen werden die Absolutbeträge ausgegeben.

In Anlehnung an den Lanczos-Algorithmus 3.3 aus Kapitel 3 wird zunächst noch einmal der Algorithmus für $-A = -C^T C$ in Pseudo-Code verfasst. Hierzu sei noch bemerkt, dass für jeden Vektor $q \in \mathbb{R}^l$ folgendes gilt:

$$\langle Aq, q \rangle = \langle C^T Cq, q \rangle = \langle Cq, Cq \rangle$$

Algorithmus 5.1.

Setze $q_0 = 0$, $q_1 = x/\|x\|$, $\beta_0 = 0$ und $\alpha_1 = -\langle Cq_1, Cq_1 \rangle$
für $k = 1$ **bis** $m - 1$
 berechne $r_k := -C^T Cq_k - \beta_{k-1}q_{k-1} - \alpha_k q_k$
 falls $\|r_k\| < \epsilon \|Aq_k\|$ **Abbruch**
 $\beta_k := \|r_k\|$
 $q_{k+1} := r_k/\beta_k$
 $\alpha_{k+1} := -\langle Cq_{k+1}, Cq_{k+1} \rangle$
stop

In Anbetracht hochdimensionaler Matrizen ist möglichst zu vermeiden, das Produkt $C^T C \in \mathbb{R}$ zu berechnen. Deshalb wird der Algorithmus 5.1 noch einmal umformuliert zu:

Algorithmus 5.2.

Setze $q_0 = 0$, $q_1 = x/\|x\|$, $\beta_0 = 0$, $c_1 = Cq_1$ und $\alpha_1 = -\langle c_1, c_1 \rangle$
für $k = 1$ **bis** $m - 1$
 berechne $r_k := -C^T c_k - \beta_{k-1}q_{k-1} - \alpha_k q_k$ mit $C^T c_k = Aq_k$
 falls $\|r_k\| < \epsilon \|Aq_k\|$ **Abbruch**
 $\beta_k := \|r_k\|$
 $q_{k+1} := r_k/\beta_k$
 $c_{k+1} := Cq_{k+1}$
 $\alpha_{k+1} := -\langle c_{k+1}, c_{k+1} \rangle$ mit $\alpha = \langle Aq_{k+1}, q_{k+1} \rangle$
stop

5.2 Praktische Probleme des Verfahrens

Im Folgenden werden Probleme, die bei der praktischen Anwendung des Lanczos-Verfahrens auftreten, dargestellt. Sie entstehen durch endliche Computer-Arithmetik, welche in der Theorie nicht berücksichtigt wird. Am Ende der später folgenden Analyse dieser Probleme wird ein abgewandelter Lanczos-Algorithmus präsentiert, welcher zumindest in Hinsicht auf das Modellproblem der exponentiell abfallenden Eigenwerte eine Lösung bietet.

5.2.1 Orthogonalitätsverlust und doppelte Eigenwerte

Die Konvergenzanalyse aus Kapitel 4 läßt auf sehr gute Konvergenz des in Kapitel 3 beschriebenen Lanczos-Verfahrens hoffen. Die nachfolgenden praktischen Ergebnisse bestätigen dies für die extremen (hier: größten) Eigenwerte, aber aufgrund begrenzter Rechengenauigkeit tauchen auch Probleme auf. Es ist zu beobachten, dass nach einer gewissen Anzahl von Lanczos-Iterationsschritten die in der Theorie vorausgesetzte Orthogonalität der berechneten Lanczos-Vektoren verloren geht. Da dieses und das nachfolgend beschriebene Phänomen der mehrfach auftauchenden Ritz-Werte insbesondere für den Testfall $\alpha = 1$ zu beobachten ist, beziehen sich die Beispiele dieses Abschnittes der Einfachheit halber auf Testmatrizen $A = C^T C \in \mathbb{R}^{n \times n}$ mit vorgegebenen Eigenwerten

$$\lambda_j = e^{-(j-1)}$$

und Dimension $n = 1000$. In der folgenden Tabelle sind die Skalarprodukte zwischen den ersten zwei Lanczos-Vektoren (zweite Spalte) und dem zuletzt aus dem m -ten Iterationsschritt berechneten Lanczos-Vektor q_m aufgelistet. Es ist deutlich zu erkennen, dass schon bei $m = 10$ die Vektoren q_1 und q_m praktisch nicht mehr orthogonal sind. Die Orthogonalität zwischen q_{m-1} und q_m ist jedoch nicht gestört, wie in der letzten Spalte der Tabelle dokumentiert ist, während sich der Orthogonalitätsverlust zwischen erstem und letztem Lanczos-Vektor in jedem Schritt verstärkt. Diese Vektoren sind somit schon fast linear abhängig.

Aus Platzgründen wird in dieser und in den folgenden Tabellen die Notation verwendet: $1.000_{-j} \hat{=} 1.000 \cdot 10^{-j}$

Skalarprodukte der Lanczos-Vektoren:

m	$\langle q_1, q_m \rangle$	$\langle q_m, q_{m+1} \rangle$
1	1.000	1.063 ₋₁₅
2	1.063 ₋₁₅	-7.572 ₋₁₅
3	-5.328 ₋₁₅	1.484 ₋₁₄
4	-2.531 ₋₁₄	-4.139 ₋₁₄
5	1.073 ₋₁₃	6.641 ₋₁₄
6	-7.247 ₋₁₂	-2.001 ₋₁₃
7	1.055 ₋₉	6.741 ₋₁₃
8	-5.803 ₋₇	-1.648 ₋₁₂
9	8.009 ₋₄	1.095 ₋₁₂
10	-7.421 ₋₁	-2.763 ₋₁₅
11	2.228 ₋₁	1.433 ₋₁₃
12	2.865 ₋₁	-1.246 ₋₁₃
13	-2.391 ₋₃	3.486 ₋₁₃
14	-1.836 ₋₁	-4.130 ₋₁₃
15	2.399 ₋₃	2.488 ₋₁₂
16	6.492 ₋₂	-3.660 ₋₁₄
17	7.748 ₋₁	6.758 ₋₁₅

Da nur aufgrund der Orthogonalität der Lanczos-Vektoren die Tridiagonalgestalt der Ritz-Matrix

$$B = Q^T A Q \quad \text{mit} \quad Q = [q_1, \dots, q_m]$$

gegeben ist, und aufgrund obiger Tabelle nicht einmal der volle Rang der Matrix Q gesichert ist, muß wegen des Orthogonalitätsverlusts auch der Zusammenhang zwischen Ritz-Werten und echten Eigenwerten in Frage gestellt werden.

Es zeigt sich unter den selben Voraussetzungen wie für den Orthogonalitätstest das folgende Phänomen: Betrachtet man für jeden Iterationsschritt m die $j < m$ größten Ritz-Werte μ_1, \dots, μ_j , so ist zu beobachten, dass diese zu Beginn des Verfahrens relativ schnell gegen die echten Eigenwerte $\lambda_1, \dots, \lambda_j$ konvergieren. Wird aber der Iterationsschritt erreicht, bei dem der Orthogonalitätsverlust einsetzt, so taucht ein weiterer Ritz-Wert auf, welcher keinen neuen Eigenwert λ_m approximiert, sondern den größten, schon konvergierten Eigenwert λ_1 . Nach einigen weiteren Schritten passiert dies erneut, und zusätzlich wird auch der nächst größere Eigenwert λ_2 nochmals durch einen weiteren Ritz-Wert approximiert. Für fortlaufende Iteration erhält man so eine Reihe mehrfach auftretender Ritz-Werte, welche zu den (einfachen) extremen Eigenwerten von A gehören.

Dieses Verhalten ist in der folgenden Tabelle dokumentiert. Sie zeigt, dass direkt nach dem Eintreten des Orthogonalitätsverlustes, also in diesem Fall bei $m = 10$, der erste

Eigenwert λ_1 (in diesem Beispiel $\lambda_1 = 1$) doppelt auftaucht. Außerdem werden in den darauffolgenden Schritten auch die nächst größeren Eigenwerte doppelt approximiert, bis bei $m = 17$ λ_1 sogar dreifach erscheint.

Da die Eigenwerte $\lambda_{11}, \dots, \lambda_{17}$ und die zugehörigen Ritz-Werte schon sehr klein sind, sind in der Tabelle ab $m = 11$ aus Platzgründen keine neuen Ritz-Werte (μ_{11} bis μ_{17}) eingetragen.

Eigenwertberechnung mit dem Lanczos-Verfahren

m	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7	μ_8	μ_9	μ_{10}
1	6.35 ₋₁									
2	9.89 ₋₁	4.74 ₋₂								
3	1.00	3.46 ₋₁	6.14 ₋₃							
4	1.00	3.68 ₋₁	1.28 ₋₁	9.78 ₋₄						
5	1.00	3.68 ₋₁	1.35 ₋₁	3.90 ₋₂	2.95 ₋₄					
6	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.59 ₋₂	6.91 ₋₅				
7	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.12 ₋₃	1.41 ₋₅			
8	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.73 ₋₃	2.16 ₋₃	4.11 ₋₆		
9	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.47 ₋₃	7.54 ₋₄	1.63 ₋₆	
10	1.00	9.14 ₋₁	3.68 ₋₁	1.35 ₋₁	4.98 ₋₂	1.83 ₋₂	6.74 ₋₃	2.47 ₋₃	7.53 ₋₄	1.63 ₋₆
11	1.00	1.00	3.68 ₋₁	1.35 ₋₁	4.98 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.13 ₋₄	3.02 ₋₄
12	1.00	1.00	3.68 ₋₁	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.06 ₋₄
13	1.00	1.00	3.68 ₋₁	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄
14	1.00	1.00	3.68 ₋₁	3.68 ₋₁	1.35 ₋₁	1.35 ₋₁	4.98 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃
15	1.00	1.00	3.68 ₋₁	3.68 ₋₁	1.35 ₋₁	1.35 ₋₁	4.98 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃
16	1.00	1.00	3.68 ₋₁	3.68 ₋₁	1.35 ₋₁	1.35 ₋₁	5.05 ₋₂	4.98 ₋₂	1.83 ₋₂	6.74 ₋₃
17	1.00	1.00	9.97 ₋₁	3.68 ₋₁	3.68 ₋₁	1.35 ₋₁	1.35 ₋₁	4.98 ₋₂	4.97 ₋₂	1.83 ₋₂

Insbesondere gilt dasselbe Verhalten auch für die zugehörigen Ritz-Vektoren. Eigentlich müßte das Verfahren in obigem Fall abbrechen, da das Abbruchkriterium des Algorithmus (3.3) die Norm des neuen Lanczos-Vektors überprüft. Ist diese zu klein, so bedeutet dies lineare Abhängigkeit der Lanczos-Vektoren, und somit wäre ein invarianter Unterraum erreicht. Um Rundungsfehler abzufangen und Skalierungsinvarianz der Matrix zu sichern, wurde die Abfrage nach der Norm bereits geändert. Es wird die Differenz zur Norm des vorangegangenen Lanczos-Vektors geprüft. Wegen der endlichen Arithmetik ist jedoch auch diese Abfrage nicht ausreichend, so dass immer mehr Iterationsschritte ausgeführt werden, wodurch die Berechnung des Krylovraums wieder von neuem startet und somit weitere Kopien der extremen Eigenwerte und zugehörigen Eigenvektoren berechnet werden. Der genauere Zusammenhang zwischen Orthogonalitätsverlust der Lanczos-Vektoren, Konvergenz eines Ritz-Paares und dem Auftreten doppelter Ritz-Paare wird in dem nächsten Abschnitt untersucht.

5.2.2 Problemanalyse

Die zuvor beschriebenen Beispiele des Orthogonalitätsverlusts und den mehrfach approximierten Eigenwerten zeigen ein allgemeines Problem des Lanczos-Verfahrens, welches auch in [8], [5] und [2] geschildert wird. In [5], Seite 142 ist dies auch graphisch dargestellt.

Im Folgenden werden diese Probleme genauer betrachtet. Es wird sich herausstellen, dass ein konkreter Zusammenhang zwischen der Konvergenz eines Ritz-Paares und dem Orthogonalitätsverlust zwischen einem der ersten und dem neu berechneten Lanczos-Vektors besteht. Den wesentlichen Hintergrund liefert das (auch in [8] beschriebene) *Theorem von Paige*. Um die Aussagen dieses Theorems verstehen zu können, müssen vorher noch einige Begriffe und Notationen erklärt werden.

Die Konvergenz der Ritz-Vektoren spielt eine wichtige Rolle in den folgenden Überlegungen bezüglich des Verhaltens des Lanczos-Verfahrens. Deshalb liegt die Frage nahe, wie die Konvergenz der Vektoren überhaupt festzustellen ist. Wie in Algorithmus 5.2 beschrieben, können in jedem Schritt des Verfahrens Eigenwerte der Ritz-Matrix berechnet werden. Besteht Bedarf, so könnten auch alle Ritz-Vektoren in jedem Schritt bestimmt werden. Um j Ritz-Vektoren zu berechnen, müssen aber pro Schritt j Matrix-Vektor-Multiplikationen durchgeführt werden. Das folgende Korollar zeigt, dass es auch ohne explizites Ausrechnen möglich ist, eine Aussage über die Konvergenz der Ritz-Vektoren zu erhalten:

Korollar 5.2. Sei $T_j = Q_j^T A Q_j \in \mathbb{R}^{j \times j}$ die tridiagonale Matrix des j -ten Lanczos-Schrittes, aus der die Eigenpaare (μ_i, y^i) , $i \in \{1, \dots, j\}$ berechnet wurden. Dann gilt für einen Ritz-Vektor $z^i = Q_j y^i$ und β_j aus dem Algorithmus 3.3:

$$\|Az^i - z^i \mu_i\| = \beta_j |y_j^i|$$

Beweis . Sei $e_j \in \mathbb{R}^j$ der j -te Einheitsvektor. Es gilt:

$$\begin{aligned} \|Az^i - z^i \mu_i\| &= \|AQ_j y^i - Q_j y^i \mu_i\| \\ &= \|(AQ_j - Q_j T_j) y^i\| && \text{da } y^i \mu_i = T_j y^i \\ &= \|(\beta_j q_{j+1} e_j^T) y^i\| && \text{nach Korollar 3.8} \\ &= \beta_j |e_j^T y^i| = \beta_j |y_j^i| && \text{da } \|q_{j+1}\| = 1 \end{aligned}$$

□

Bemerkung 5.1. Aus Korollar 5.2 folgt, dass der Ritz-Vektor z^i im j -ten Schritt schon gut konvergiert ist, wenn $\beta_j |y_j^i|$ klein ist. Insbesondere folgt, dass ein Ritz-Vektor unabhängig von β_j auch dann schon gut konvergiert sein kann, wenn $|y_j^i|$ hinreichend klein ist.

Da im Folgenden das Verhalten des Lanczos-Algorithmus in der Praxis betrachtet werden soll, muß berücksichtigt werden, dass sich durch die endliche Rechner-Arithmetik

Ungenauigkeiten in der Berechnung der Matrizen und Vektoren des Lanczos-Verfahrens ergeben.

Wie schon in der Tabelle der Skalarprodukte zu beobachten ist, sind zum Beispiel die Lanczos-Vektoren q_1, \dots, q_j im j -ten Schritt nicht mehr zwangsläufig orthogonal, insbesondere gilt also nicht mehr $Q_j^T Q_j = I_{j \times j}$. Um die Anwesenheit von Rundungsfehlern in den nachfolgenden Überlegungen zu verdeutlichen, werden die betroffenen Größen, unter anderem Q_j, T_j und auch die Ritz-Vektoren z^i durch \tilde{Q}_j, \tilde{T}_j und \tilde{z}^i gekennzeichnet. Da die Matrix \tilde{T}_j , aus der die Eigenpaare $(\tilde{\mu}_i, \tilde{y}^i)$ berechnet werden, sehr viel kleinere Dimension als n hat, sei jedoch zur Vereinfachung angenommen, dass zur Lösung dieses Eigenproblems nicht noch weitere Rundungsfehler hinzukommen. Das heißt es gelte:

$$(5.3) \quad \tilde{T}_j = Y_j \Lambda Y_j^T$$

wobei Y_j die Matrix der Eigenvektoren $\tilde{y}^1, \dots, \tilde{y}^j$ von \tilde{T}_j sei und Λ Diagonalmatrix mit den Eigenwerten $\tilde{\mu}_1, \dots, \tilde{\mu}_j$.

Weiter sei aufgrund der Orthogonalisierung des neuen Lanczos-Vektors gegen die beiden Vorgänger in Algorithmus 3.3 angenommen, dass gilt:

$$(5.4) \quad \tilde{q}_{i+1}^T \tilde{q}_i = 0 \quad \text{für alle } i \in \{1, \dots, j-1\} \quad \text{und} \quad \tilde{r}_j^T \tilde{q}_j = 0.$$

Um das schon erwähnte Ziel, das Theorem von Paige, einführen zu können, müssen die beiden wichtigen Gleichungen,

$$(5.5) \quad A Q_j - Q_j T_j = r_j e_j^T$$

und

$$(5.6) \quad Q_j^T Q_j = I_{j \times j}$$

die die Theorie des Lanczos-Verfahrens ergibt, an die nun veränderte Situation mit Rundungsfehlern angepasst werden. Sie werden nun beschrieben durch

$$(5.7) \quad A \tilde{Q}_j - \tilde{Q}_j \tilde{T}_j = r_j e_j^T + F_j$$

und

$$(5.8) \quad I_{j \times j} - \tilde{Q}_j^T \tilde{Q}_j = C_j^T + \Delta_j + C_j$$

F_j, C_j und Δ_j stellen also die Störung der oberen Gleichungen dar, wobei C_j strikte obere Dreiecksmatrix ist und Δ_j Diagonalmatrix.

Mit den nun eingeführten Matrizen kann der nächste Satz vorgestellt werden, dessen Aussagen nachfolgend noch vereinfacht diskutiert werden.

Satz 5.1 (Paige's Theorem). *Gelten die Gleichungen (5.3) - (5.8). Sei K_j der strikte obere Dreiecksanteil von $F_j^T \tilde{Q}_j - \tilde{Q}_j^T F_j$ und N_j der von $\Delta_j \tilde{T}_j - \tilde{T}_j \Delta_j$.*

Sei $G_j = Y_j^T (K_j + N_j) Y_j$ mit $G_{j,i,k} = (\gamma_{ik})$. Dann gilt für die Ritz-Vektoren $\tilde{z}^1, \dots, \tilde{z}^j$ und den Lanczos-Vektor \tilde{q}_{j+1} :

$$\tilde{z}^{iT} \tilde{q}_{j+1} = \frac{\gamma_{ii}}{\beta_j |\tilde{y}_j^i|}$$

und für $i \neq k$ gilt

$$(\tilde{\mu}_i - \tilde{\mu}_k) \tilde{z}^{iT} \tilde{z}^k = \gamma_{ii} \left(\frac{\tilde{y}_j^k}{\tilde{y}_j^i} \right) - \gamma_{kk} \left(\frac{\tilde{y}_j^i}{\tilde{y}_j^k} \right) - (\gamma_{ik} - \gamma_{ki})$$

Beweis . Siehe [8], Theorem 13.4.1. □

Um das Verständnis des Satzes zu erleichtern, wird das nächste Korollar vorgestellt, welches ermöglicht, die Details des Satzes handlicher darzustellen.

Nach [8] gilt

Korollar 5.3. *Für die Elemente γ_{ik} der Matrix G_j aus Satz 5.1 gilt:*

$$|\gamma_{ik}| \leq \varepsilon \|A\|$$

wobei ε die Maschinengenauigkeit ist.

Mit Hilfe des Korollars lassen sich beide Teile von Satz 5.1 stark vereinfachen. Für die Ritz-Vektoren gilt nun:

$$(5.9) \quad \tilde{z}^{iT} \tilde{q}_{j+1} \leq \frac{\varepsilon \|A\|}{\beta_j |\tilde{y}_j^i|}$$

Die erste Schlußfolgerung, die aus (5.9) gezogen werden kann, folgt nun mit Korollar 5.2. Demnach gilt $\|Az^i - z^i \mu_i\| = \beta_j |y_j^i|$ in exakter Arithmetik. Auch in dieser Formel müssten also Rundungsfehler berücksichtigt werden, nach [8], Seite 293 sind diese aber harmlos. Das bedeutet, für einen Ritz-Vektor \tilde{z}^i , der schon relativ gut konvergiert ist, ist $\beta_j |y_j^i|$ sehr klein. Aus (5.9) ergibt sich damit folgende wichtige Bemerkung:

Bemerkung 5.2. *Ist ein Ritz-Vektor \tilde{z}^i für ein $i \in \{1, \dots, j\}$ bereits konvergiert, so müssen \tilde{z}^i und \tilde{q}_{j+1} nicht orthogonal sein, und \tilde{q}_{j+1} hat somit eine nicht-triviale Komponente in Richtung eines Eigenvektors.*

Aufgrund der obigen Bemerkung lässt sich nun auch leicht der angekündigte Zusammenhang zwischen der Konvergenz der Ritz-Paare und Orthogonalitätsverlust der Lanczos-Vektoren erklären: Die Ritz-Vektoren $\tilde{z}^i = \tilde{Q}_j \tilde{y}^i$, $i \in \{1, \dots, j\}$, des j -ten Iterationsschrittes sind in dem Krylovraum $\tilde{\mathcal{K}}_j(x, A) = \text{span} \{\tilde{q}_1, \dots, \tilde{q}_j\}$ enthalten. Nach Konstruktion der \tilde{q}_i müsste eigentlich $q_{j+1} \perp \mathcal{K}_j(x, A)$ gelten. Nach Bemerkung 5.2 kann es passieren, dass schon konvergierte Ritz-Vektoren \tilde{z}^i nicht orthogonal auf \tilde{q}_{j+1} stehen. In diesem Fall können auch nicht alle vorherigen Lanczos-Vektoren $\tilde{q}_1, \dots, \tilde{q}_j$ orthogonal auf q_{j+1} stehen. Diese Erkenntnis führt zur nächsten Bemerkung:

Bemerkung 5.3. *Der Orthogonalitätsverlust der Lanczos-Vektoren \tilde{q}_i tritt ein, wenn einer der Ritz-Vektoren \tilde{z}^i konvergiert.*

Doch wie erklärt sich nun der beobachtete Zusammenhang zwischen dem Orthogonalitätsverlust und dem Auftreten doppelter Eigenwertapproximationen?

Durch Konvergenz der Ritz-Vektoren kann es passieren, dass die Lanczos-Vektoren nicht orthogonal aufeinander stehen, und wie in der Tabelle zu den Skalarprodukten zu sehen war, verschlimmert sich dieser Effekt bei fortlaufender Iteration zunehmend, so dass einige Lanczos-Vektoren schon fast linear abhängig sind. Dadurch startet sich die Berechnung des selben Krylovraums von neuem, es werden keine neuen Basisvektoren des nächst größeren Raums bestimmt, sondern die schon vorhandenen Vektoren werden erneut berechnet. Eine Konsequenz ist das Erscheinen mehrfacher Ritz-Paare, welches durch genaueres Betrachten des Satzes 5.1 verdeutlicht wird.

Wie schon mit Hilfe von Korollar 5.2 festgestellt, gilt für einen beinahe konvergierten Ritz-Vektor \tilde{z}^i , dass $\beta_j |\tilde{y}_j^i|$ sehr klein ist. Da auch festgestellt wurde, dass β_j aufgrund von Rundungsfehlern nicht unbedingt klein sein muß, ist also in diesem Fall das j -te (also das letzte) Element des Eigenvektors \tilde{y}^i sehr klein.

Für den Ritz-Vektor \tilde{z}^i bedeutet das:

$$\tilde{z}^i = \tilde{Q}_j \tilde{y}^i = \begin{pmatrix} \sum_{k=1}^j \tilde{q}_1^k \tilde{y}_k^i \\ \vdots \\ \sum_{k=1}^j \tilde{q}_j^k \tilde{y}_k^i \end{pmatrix} \tilde{y}_j^i \text{ klein} \approx \begin{pmatrix} \sum_{k=1}^{j-1} \tilde{q}_1^k \tilde{y}_k^i \\ \vdots \\ \sum_{k=1}^{j-1} \tilde{q}_j^k \tilde{y}_k^i \end{pmatrix}$$

Das bedeutet, die j Ritz-Vektoren $\tilde{z}^1, \dots, \tilde{z}^j$ befinden sich in dem nun (fast) nur noch $(j-1)$ -dimensionalen Krylovraum $\tilde{\mathcal{K}}_{j-1}(x, A)$, und sind (fast) linear abhängig. Da die Ritz-Vektoren Eigenvektoren approximieren, entstehen mehrere *Ritz-Vektoren* zu einem einfachen Eigenwert von A . Insbesondere gilt:

Bemerkung 5.4. *Auch die Ritz-Vektoren verlieren Orthogonalität, wenn sie beginnen zu konvergieren.*

Der zweite Teil aus Satz 5.1 liefert zusammen mit Korollar 5.3 auch eine Erklärung der mehrfachen Ritz-Werte zum selben Eigenwert:

Da $|\gamma_{i,k}| \leq \varepsilon \|A\|$ für alle $i, k \in \{1, \dots, j\}$ gilt, kann die rechte Seite der zweiten Gleichung des Satzes sehr klein werden. Auf der linken Seite steht die Differenz zweier Ritz-Werte $\tilde{\mu}_i$ und $\tilde{\mu}_k$ (für $i \neq k$) multipliziert mit dem Skalarprodukt der zugehörigen Ritz-Vektoren \tilde{z}^i und \tilde{z}^k . Aufgrund von Bemerkung 5.4 ist dieses Skalarprodukt aber nicht zwingend nahe bei Null, also müssen $\tilde{\mu}_i$ und $\tilde{\mu}_k$ dicht beieinander liegen.

Da die Ritz-Werte und -Vektoren zu den größten Eigenwerten am schnellsten konvergieren, erklärt sich somit auch die Häufigkeit der Ritz-Werte, die λ_1 approximieren.

Zusammengefasst ergibt sich aus den vorangegangenen Überlegungen die abschließende Bemerkung:

Bemerkung 5.5. *Sobald die Ritz-Paare anfangen zu konvergieren, kann auch die Orthogonalität der Lanczos-Vektoren und Ritz-Vektoren verloren gehen, und infolgedessen erscheinen mehrere Kopien der (einfachen) extremen Eigenwerte der Matrix A .*

Das relativ schnelle Auftauchen dieses Phänomens (für $\alpha = 1$ bereits nach $m = 10$ Schritten, für $\alpha = 1/2, 1/3$ erst etwas später) zeigt also unter anderem, wie gut die vorgegebenen exponentiell abfallenden Eigenwerte mit dem Lanczos-Verfahren berechnet werden.

Eine Folge der mehrfachen Ritz-Paare ist die schwierige Zuordnung der Ritz-Werte zu den echten Eigenwerten. Wie bekommt man also das Problem in den Griff? Eine Lösung bietet das im nächsten Abschnitt beschriebene veränderte Lanczos-Verfahren.

5.2.3 Modifiziertes Lanczos-Verfahren

Das Problem der mehrfach approximierten Eigenwerte wird durch den Orthogonalitätsverlust hervorgerufen. Das eigentliche Problem des Algorithmus ist jedoch, dass dieser in dem Fall der (fast) linearen Abhängigkeit der Lanczos-Vektoren nicht abbricht (β_j ist nicht klein genug).

Da die mangelnde Orthogonalität zuerst zwischen den ersten Lanczos-Vektoren und den neu berechneten auftritt, liegt die Idee nahe, die Begründung dafür in dem Weglassen der Orthogonalisierung gegen die weiter als 2 Schritte zurückliegenden Lanczos-Vektoren zu suchen. In exakter Arithmetik war dies gerade als Vorteil hervorgehoben worden, da Rechenaufwand gespart wird. Dies ist aber auch die Begründung dafür, dass nicht abgebrochen wird, weil die fehlende Orthogonalität (welche durch ein zu kleines β_j festgestellt werden würde) gar nicht bemerkt wird, da sie aufgrund der Theorie vorausgesetzt wurde. Gerade in dem Beispiel exponentiell abfallender Eigenwerte werden, wie schon in den ersten Kapiteln begründet, meist nur einige wenige der größten Eigenwerte überhaupt gesucht. Da die Konvergenztheorie aus Kapitel 4 diesen Eigenwerten gute Konvergenz verspricht, und dies auch in den bisherigen numerischen Ergebnissen bestätigt scheint, wird nun der folgende Lösungsweg zur Änderung des Lanczos-Algorithmus vorgeschlagen:

Orthogonalisierung von q_{j+1} gegen alle vorangegangenen Lanczos-Vektoren q_1, \dots, q_j . Zur Bestimmung der Ritz-Vektoren wird die Matrix Q_m mit den Spaltenvektoren q_1, \dots, q_m sowieso gespeichert, und m wird in der Regel aus den oben genannten Gründen nicht sehr groß sein. Also ist der zusätzliche Aufwand für die komplette Orthogonalisierung der Lanczos-Vektoren, zum Beispiel mit dem Verfahren von Gram-Schmidt, gering. Allerdings gilt nun nicht mehr $\beta_k = \|q_{k+1}\|$, so dass β_k mit Hilfe der Formel:

$$\beta_k = \langle Aq_{k+1}, q_k \rangle = \langle Cq_{k+1}, Cq_k \rangle$$

berechnet werden muß.

Hier also der abgewandelte Lanczos-Algorithmus mit Gram-Schmidtschen-Orthogonalisierungs-Verfahren.

Algorithmus 5.3 (Modifiziertes Lanczos-Verfahren).

Setze $q_0 = 0$, $q_1 = x/\|x\|$, $\beta_0 = 0$, $c = Cq_1$ und $\alpha_1 = -\langle c, c \rangle$
für $k = 1$ **bis** $m - 1$
berechne $r_k := -C^T c$ $r_k = Aq_k$
 $r_k = r_k - \sum_{i=1}^k \langle r_k, q_i \rangle q_i$ Orthogonalisierung gegen q_1, \dots, q_k
falls $\|r_k\| < \epsilon \|Aq_k\|$ **Abbruch**
 $q_{k+1} := r_k / \beta_k$
 $c := Cq_{k+1}$
 $\alpha_{k+1} := -\langle c, c \rangle$
 $\beta_k = \langle c, Cq_k \rangle$
stop

Wie in den nächsten zwei Tabellen dokumentiert ist, werden mit dem modifizierten Lanczos-Algorithmus keine Eigenwerte doppelt oder mehrfach berechnet. Bis ungefähr zum 20. Iterationsschritt (für $\alpha = 1$, ansonsten später), und insbesondere bis $m = 10$, wie in der Tabelle zu sehen, ist auch kein wesentlicher Orthogonalitätsverlust zu erkennen. Dann verlieren die Vektoren auch bei diesem Algorithmus merklich an Orthogonalität. Ist der 23. Schritt erreicht, so bricht der neue Algorithmus ab, da diesmal aufgrund der genaueren Berechnung die Norm des neuen Lanczos-Vektors trotz Rundungsfehlern sehr klein ist. Das heißt, der modifizierten Lanczos-Algorithmus bricht ab, wenn die Lanczos-Vektoren nahezu linear abhängig sind, und begeht somit nicht denselben Fehler wie der vorherige Algorithmus.

Skalarprodukte der modifizierten Lanczos-Vektoren:

m	$\langle q_1, q_m \rangle$
1	1.000
2	1.943 ₋₁₅
3	3.124 ₋₁₅
4	1.008 ₋₁₄
5	4.528 ₋₁₄
6	1.096 ₋₁₃
7	3.678 ₋₁₃
8	1.467 ₋₁₂

m	$\langle q_1, q_m \rangle$
9	4.846 ₋₁₂
10	1.320 ₋₁₁
11	2.788 ₋₁₁
12	8.656 ₋₁₁
13	4.090 ₋₁₀
14	7.560 ₋₁₀
15	2.803 ₋₀₉
16	1.280 ₋₀₈

Eigenwerte des modifizierten Lanczos-Verfahrens

m	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7	μ_8	μ_9	μ_{10}
1	6.35 ₋₁									
2	9.89 ₋₁	4.74 ₋₂								
3	1.00	3.46 ₋₁	6.15 ₋₃							
4	1.00	3.68 ₋₁	1.28 ₋₁	9.78 ₋₄						
5	1.00	3.68 ₋₁	1.35 ₋₁	3.90 ₋₂	2.95 ₋₄					
6	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.59 ₋₂	6.91 ₋₅				
7	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.12 ₋₃	1.41 ₋₅			
8	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.73 ₋₃	2.16 ₋₃	4.11 ₋₆		
9	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.47 ₋₃	7.53 ₋₄	1.63 ₋₆	
10	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.06 ₋₄	2.30 ₋₄	5.96 ₋₇
11	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄	3.32 ₋₄	1.03 ₋₄
12	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄	3.32 ₋₄	1.23 ₋₄
13	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄	3.32 ₋₄	1.23 ₋₄
14	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄	3.32 ₋₄	1.23 ₋₄
15	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄	3.32 ₋₄	1.23 ₋₄
16	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄	3.32 ₋₄	1.23 ₋₄
17	1.00	3.68 ₋₁	1.35 ₋₁	4.95 ₋₂	1.83 ₋₂	6.74 ₋₃	2.48 ₋₃	9.12 ₋₄	3.32 ₋₄	1.23 ₋₄

5.3 Numerische Ergebnisse

Alle folgenden Abbildungen beziehen sich auf die unter Abschnitt 5.1 definierten Testmatrizen, deren Eigenwerte nach der Formel (5.1) gebildet wurden. Die Eigenwerte und zugehörigen Eigenvektoren wurden mit dem modifizierten Lanczos-Algorithmus berechnet. Die Eigenwerte wurden dabei wieder über den Parameter α , der im Wesentlichen über den exponentiellen Abfall der Eigenwerte entscheidet, variiert. In allen Grafiken ist der relative Fehler dargestellt. Die x -Achse zeigt die Anzahl der Iterationsschritte an. Bis auf die Bilder, die den Winkel zwischen Eigenvektoren und Ritz-Vektoren darstellen, ist in allen Zeichnungen die y -Achse logarithmiert dargestellt.

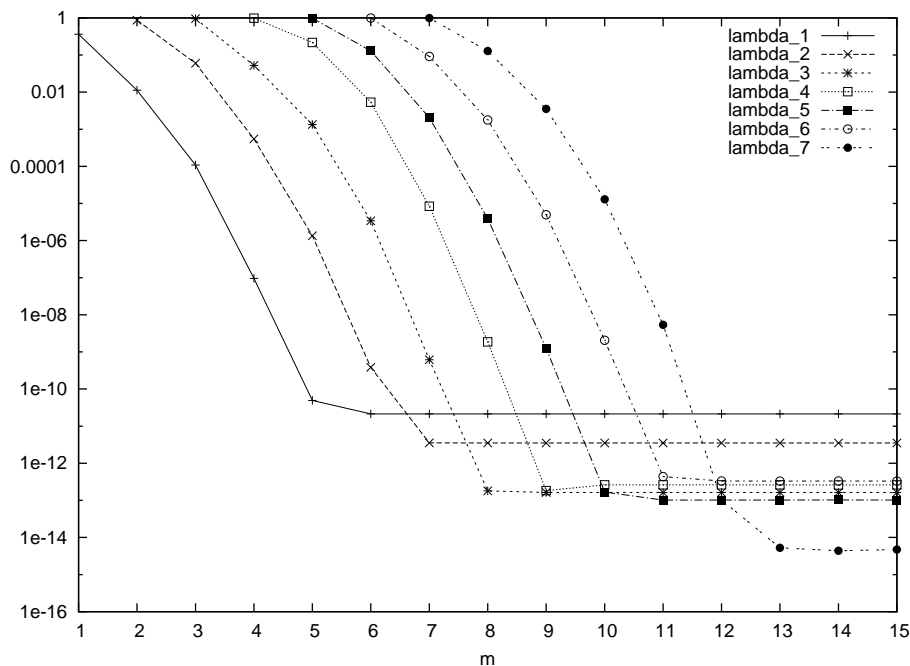
Weitere Kurven, in denen außer α auch der Parameter c_2 verändert (dass heißt $c_2 \neq 1$) wurde, befinden sich im Anhang. Da c_1 , wie in Kapitel 4 festgestellt, keinen Einfluss auf die Konvergenz der Eigenwerte hat, wurde c_1 in allen Rechnungen gleich 1 gewählt. Auch die Daten (zum praktischen Teil) der Grafiken im Anhang stammen aus dem modifizierten Lanczos-Verfahren.

Auf den nächsten drei Abbildungen ist der relative Fehler der ersten sieben mit dem Lanczos-Verfahren berechneten Eigenwerte zu sehen, jeweils für $\alpha = 1$, $\alpha = 1/2$ und $\alpha = 1/3$ und $c_1 = c_2 = 1$. Wie nach den Überlegungen aus Kapitel 4 zu erwarten war, konvergieren die Eigenwerte λ_i , $i = 2, \dots, 7$, langsamer für $\alpha \neq 1$.

Die erste Grafik (zu $\alpha = 1$) zeigt den Verlauf von λ_1 bis λ_7 bei $m = 15$ Iterationsschritten. Es ist zu erkennen, dass λ_1 bereits nach 6 Schritten bis auf eine Genauigkeit von 10^{-11} konvergiert ist.

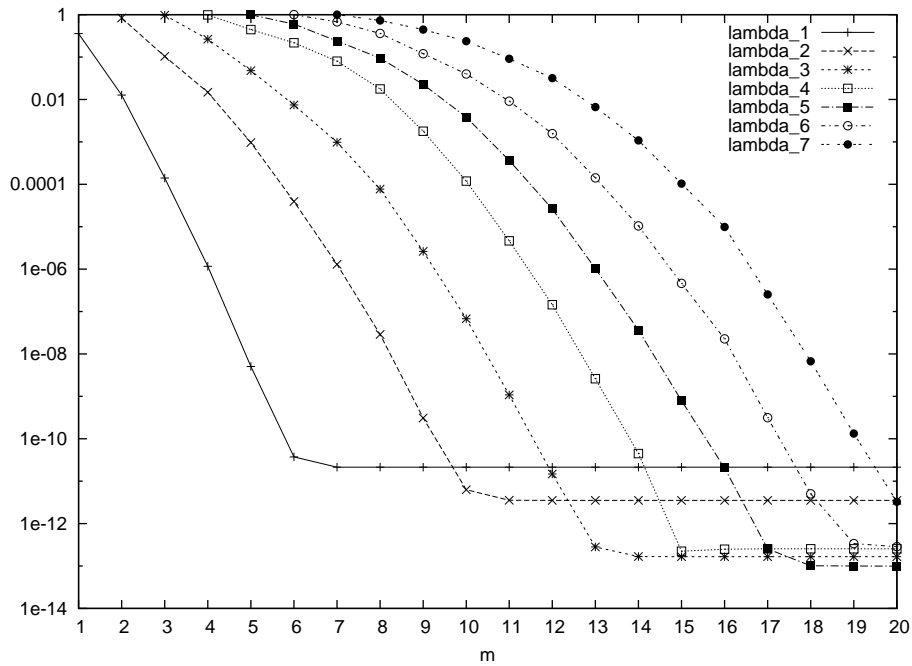
Da die Maschinengenauigkeit bei 10^{-16} liegt, ist dies ein sehr gutes Ergebnis, auch wenn die späteren Eigenwerte eine Genauigkeit von ungefähr 10^{-13} erreichen. Auch die Eigenwerte λ_2 bis λ_7 konvergieren bereits nach 6 weiteren Iterationsschritten (nach ihrer ersten Berechnung). Also bestätigt auch die Praxis sehr gute (exponentielle) und insbesondere gleich gute Konvergenz aller Eigenwerte für $\alpha = 1$.

Abbildung 5.1: Relativer Fehler $\lambda_1, \dots, \lambda_7$ für $c_1 = c_2 = 1$ und $\alpha = 1$



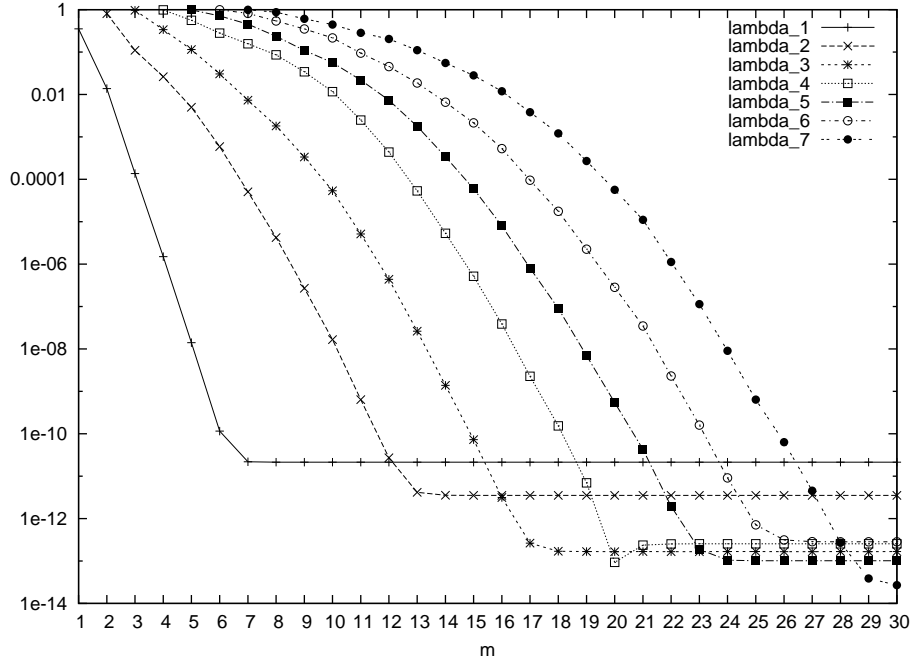
Die folgende Grafik zeigt den Fehlerverlauf der ersten sieben Eigenwerte, welche mit $\alpha = 1/2$ berechnet wurden. Hier sind $m = 20$ Iterationsschritte durchgeführt worden, da die Eigenwerte bis auf λ_1 langsamer konvergieren. Zum Beispiel ist zu beobachten, dass λ_3 erst nach 13 Schritten eine Genauigkeit von 10^{-13} erreicht, also ungefähr doppelt so viele Iterationen benötigt wie für den Fall $\alpha = 1$. Trotzdem gilt auch für $\alpha = 1/2$, dass der relative Fehler der Eigenwerte sehr schnell abfällt.

Abbildung 5.2: Relativer Fehler $\lambda_1, \dots, \lambda_7$ für $c_1 = c_2 = 1$ und $\alpha = 1/2$



Betrachtet man nun die Fehlerkurve der Eigenwerte $\lambda_2, \dots, \lambda_7$ zu $\alpha = 1/3$, so ist insbesondere im Vergleich zur Kurve von λ_1 eine Verschlechterung zu erkennen. λ_3 konvergiert in diesem Fall erst nach 17-Schritten.

Abbildung 5.3: Relativer Fehler $\lambda_1, \dots, \lambda_7$ für $c_1 = c_2 = 1$ und $\alpha = 1/3$



Da mit dem modifizierten Lanczos-Verfahren nicht nur die Eigenwerte, sondern auch die Eigenvektoren der Matrix $C^T C$ aus 5.1 berechnet wurden, wird nun der Winkel zwischen echtem Eigenvektor und berechnetem Ritz-Vektor untersucht. Die nächsten drei Grafiken zeigen den Kosinus des Winkels zwischen den ersten 7 Ritz-Vektoren z^i und zugehörigen Eigenvektoren x^i , analog zu der Darstellung der Eigenwerte jeweils für den Parameter $\alpha = 1$, $\alpha = 1/2$ und $\alpha = 1/3$, sowie $c_1 = c_2 = 1$.

Es ist genau wie in den Bildern zur Konvergenz der Eigenwerte zu erkennen, dass sich die Konvergenz für kleineres α verlangsamt. Hierbei ist zu bemerken, dass für kleine Fehlerwinkel ϕ der Kosinus $\cos \phi$ nahe bei 1 liegen muß, da $\cos 0 = 1$ gilt.

Abbildung 5.4: $\cos \angle(x^1, z^1), \dots, \cos \angle(x^7, z^7)$, für $c_1 = c_2 = 1$ und $\alpha = 1$

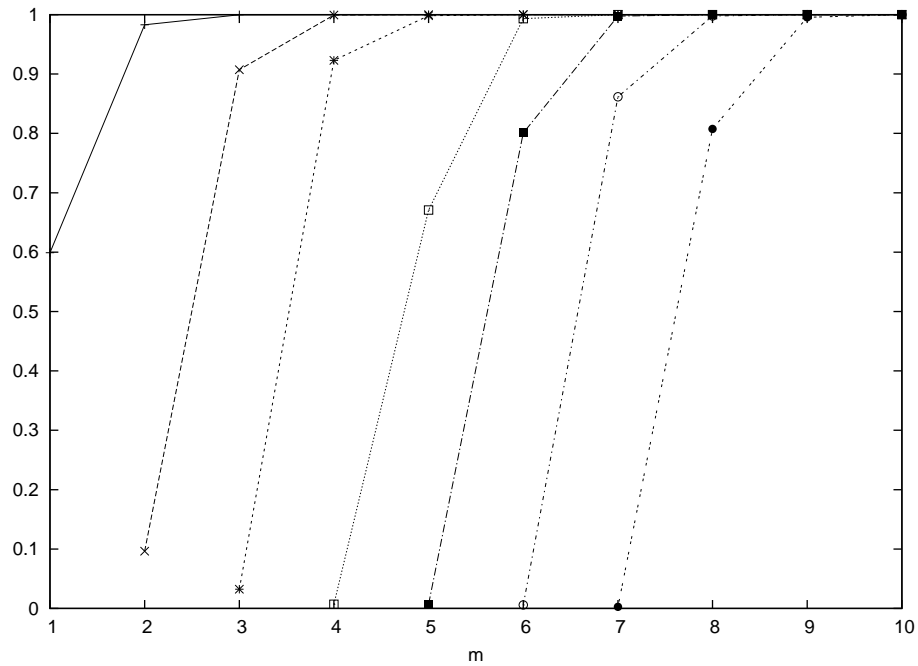


Abbildung 5.5: $\cos \angle(x^1, z^1), \dots, \cos \angle(x^7, z^7)$, für $c_1 = c_2 = 1$ und $\alpha = 1/2$

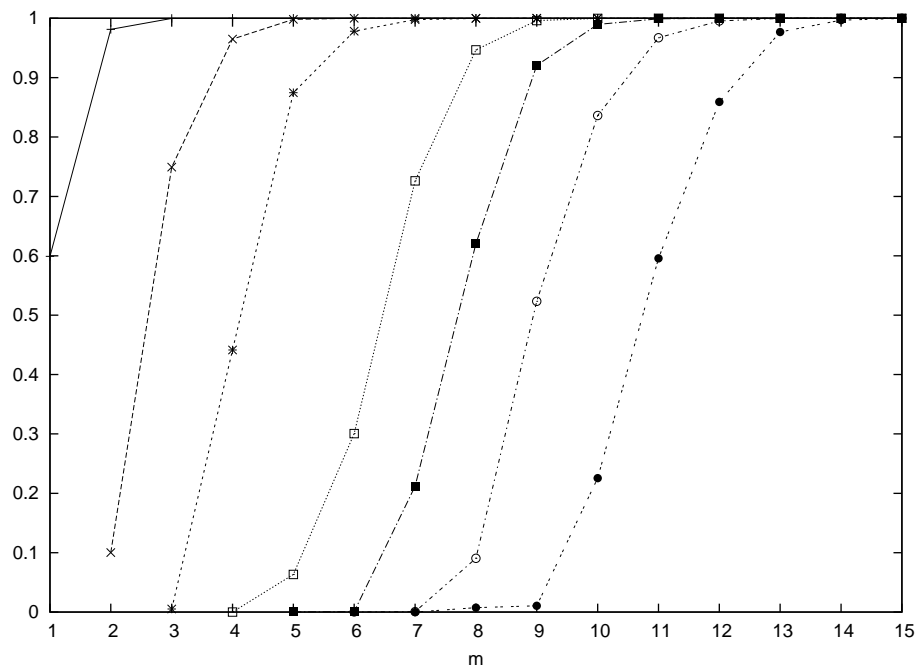
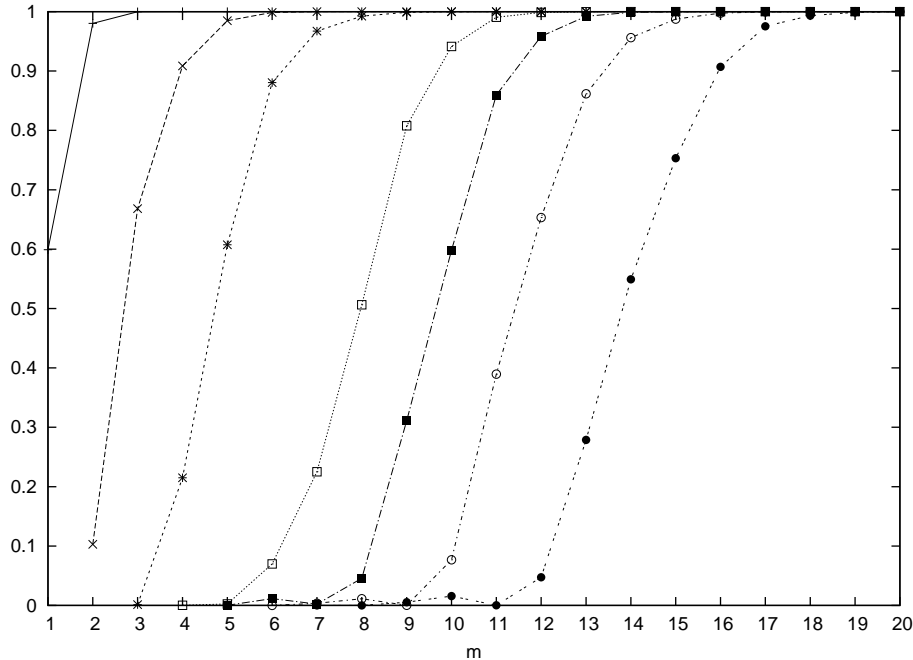


Abbildung 5.6: $\cos \angle(x^1, z^1), \dots, \cos \angle(x^7, z^7)$, für $c_1 = c_2 = 1$ und $\alpha = 1/3$



Nun werden die Ergebnisse der Eigenwertapproximation des Lanczos-Verfahrens mit den theoretischen Konvergenzaussagen aus Kapitel 4 verglichen. Hier wird wie auch schon in den theoretischen Fehleranalysen der Schwerpunkt auf die Eigenwerte λ_1 , λ_3 und λ_7 gelegt, und zwar für die auch bisher betrachteten verschiedenen Werte für α .

Da an den obigen Abbildungen zu sehen ist, dass die Ritz-Werte in der Praxis sehr schnell konvergieren, wird im Folgenden mit der theoretischen Fehlerschranke aus Satz 4.2 verglichen, die anstelle des Produkts (4.55) über die Ritz-Werte das konstante Produkt

$$\prod_{i=1}^{j-1} \frac{\lambda_i - \lambda_n}{\lambda_i - \lambda_j}$$

über die Eigenwerte aus Satz 4.4 verwendet.

Die nächste Grafik zeigt den Verlauf des Fehlers bei der Berechnung des ersten Eigenwerts λ_1 im Vergleich zur theoretischen Fehlerkurve der Abbildung 4.2.1 für $\alpha = 1$. Es ist zu sehen, dass bis zum 7. Iterationsschritt der bei der praktischen Anwendung entstandene Approximationsfehler sogar unter der (theoretisch) berechneten Fehlerschranke liegt. Ein ähnliches Ergebnis zeigt sich für die folgenden Eigenwerte λ_3 und λ_7 , die mit den Parametern $\alpha = 1$ und $c_1, c_2 = 1$ berechnet wurden. Je größer das betrachtete j ist, um so deutlicher ist der Unterschied der theoretischen oberen Schranke und dem tatsächlichen, besseren Ergebnis in der Praxis zu erkennen.

Abbildung 5.7: Relativer Fehler λ_1 für $c_1 = c_2 = 1$ und $\alpha = 1$

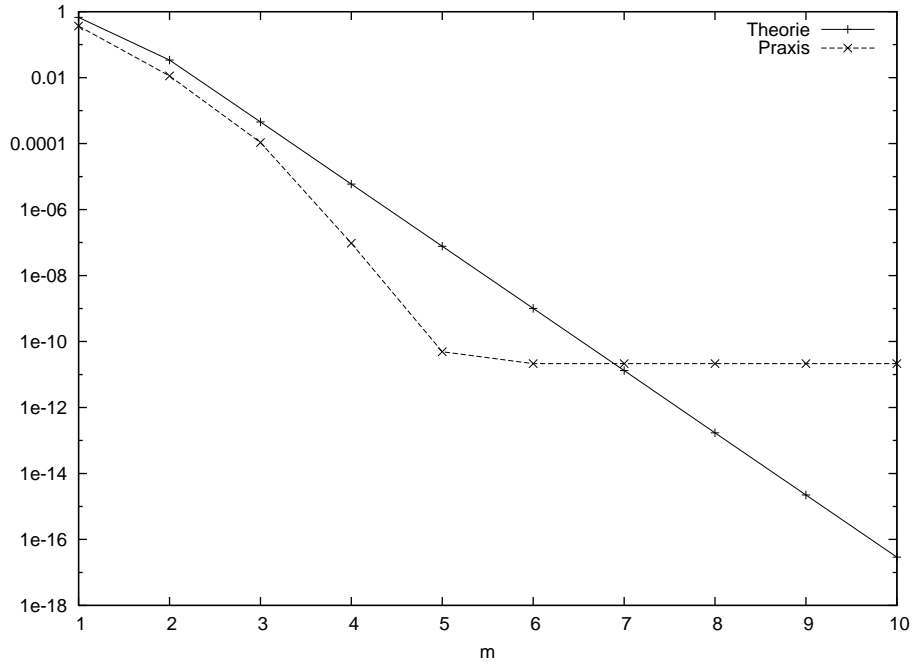


Abbildung 5.8: Relativer Fehler λ_3 für $c_1 = c_2 = 1$ und $\alpha = 1$

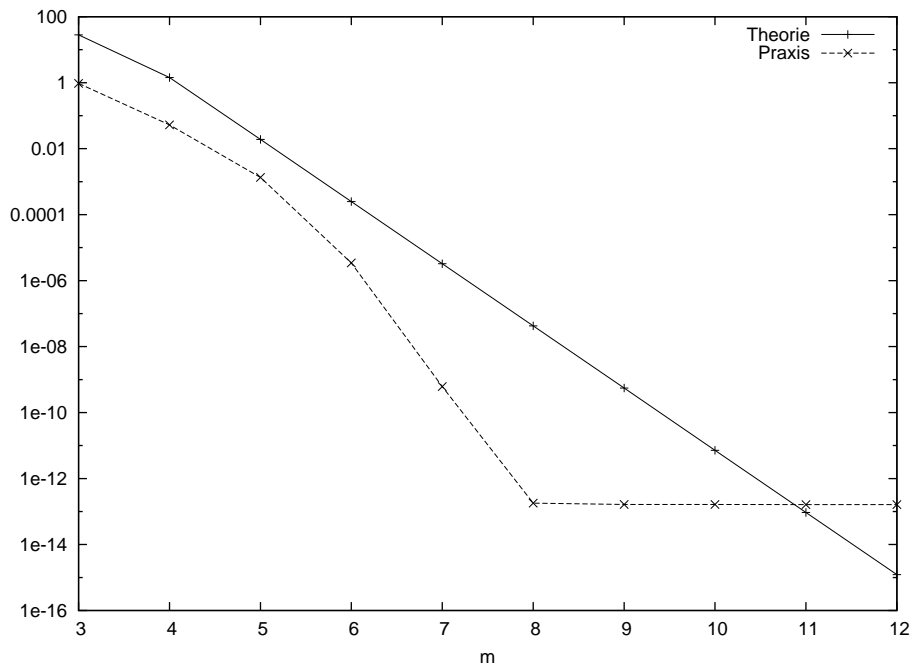
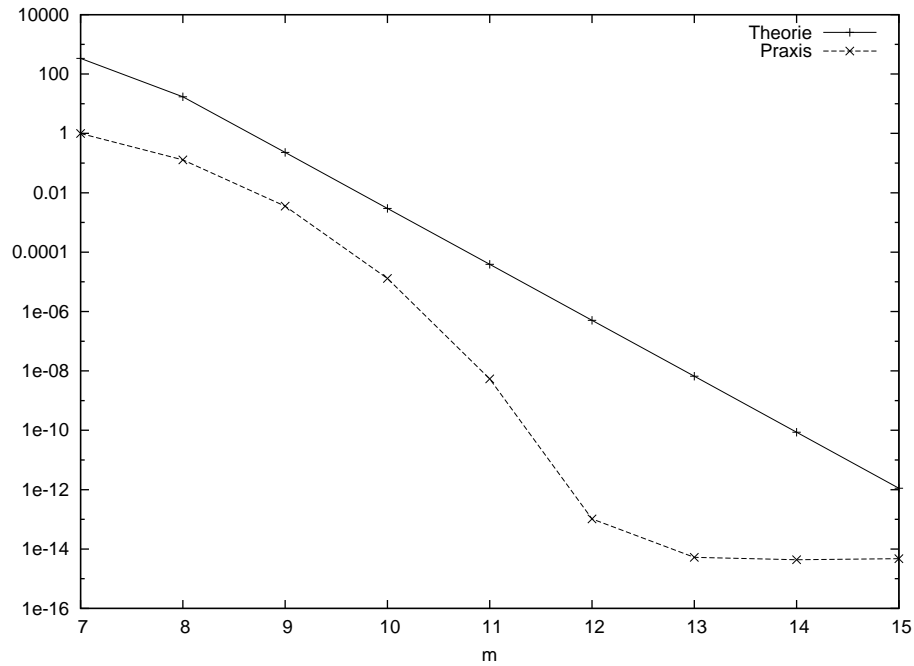


Abbildung 5.9: Relativer Fehler λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1$



Da λ_1 unabhängig von α konvergiert, werden in den folgenden Bildern zu $\alpha = 1/2$ und $\alpha = 1/3$ nur die Eigenwerte λ_3 und λ_7 betrachtet. Hier wird der Unterschied zu dem tatsächlichen praktischen Fehler besonders deutlich: Bishin zu 20 beziehungsweise 30 Iterationsschritten liegt der Fehler der mit dem modifizierten Lanczos-Verfahren berechneten Werte weit unter der in Satz 4.2 angegebenen oberen Schranke.

Abbildung 5.10: Vergleich zwischen Theorie und Praxis: λ_3 für $c_1 = c_2 = 1$ und $\alpha = 1/2$

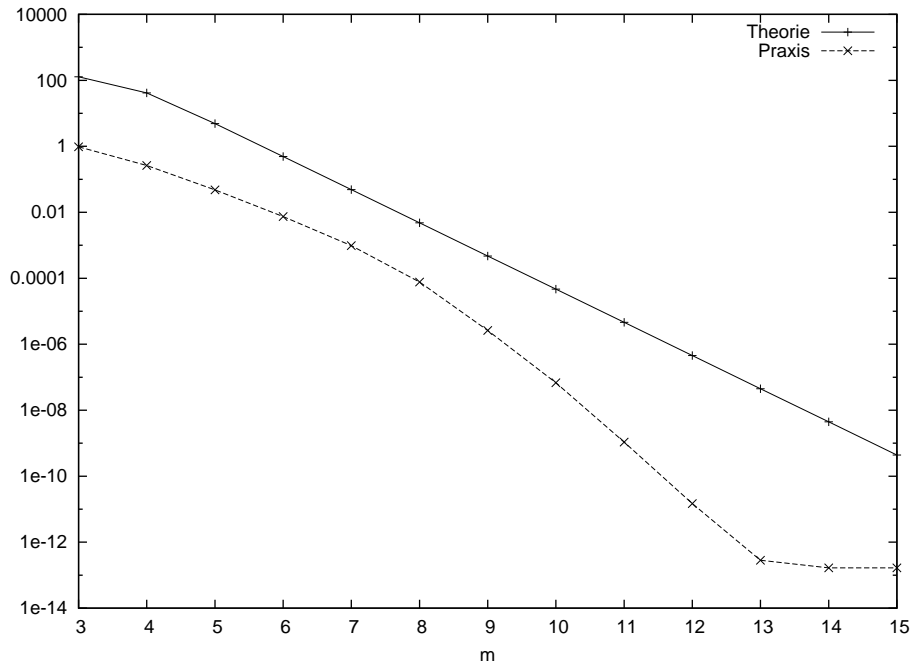


Abbildung 5.11: Vergleich zwischen Theorie und Praxis: λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/2$

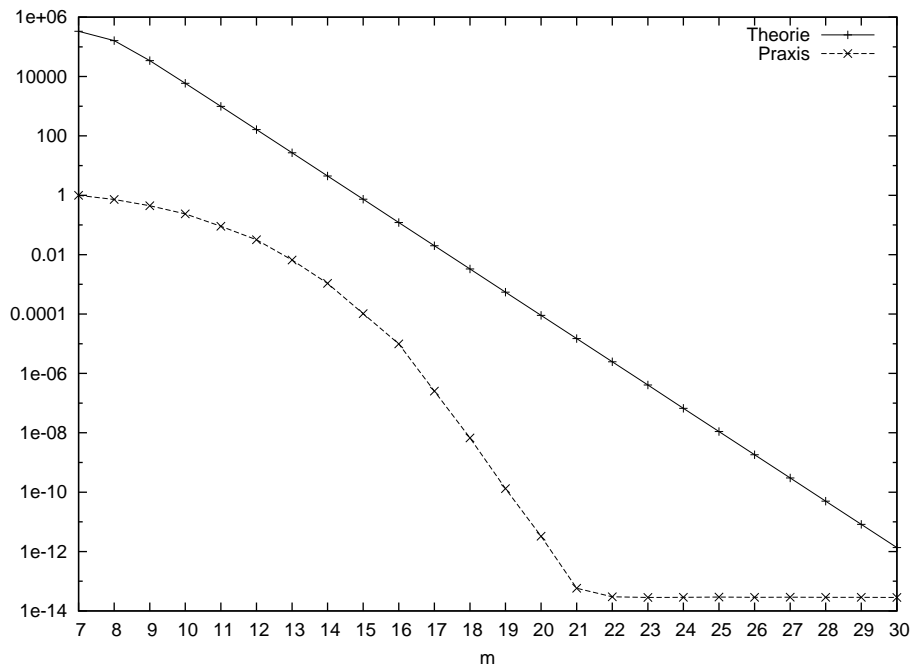


Abbildung 5.12: Vergleich zwischen Theorie und Praxis: λ_3 für $c_1 = c_2 = 1$ und $\alpha = 1/3$

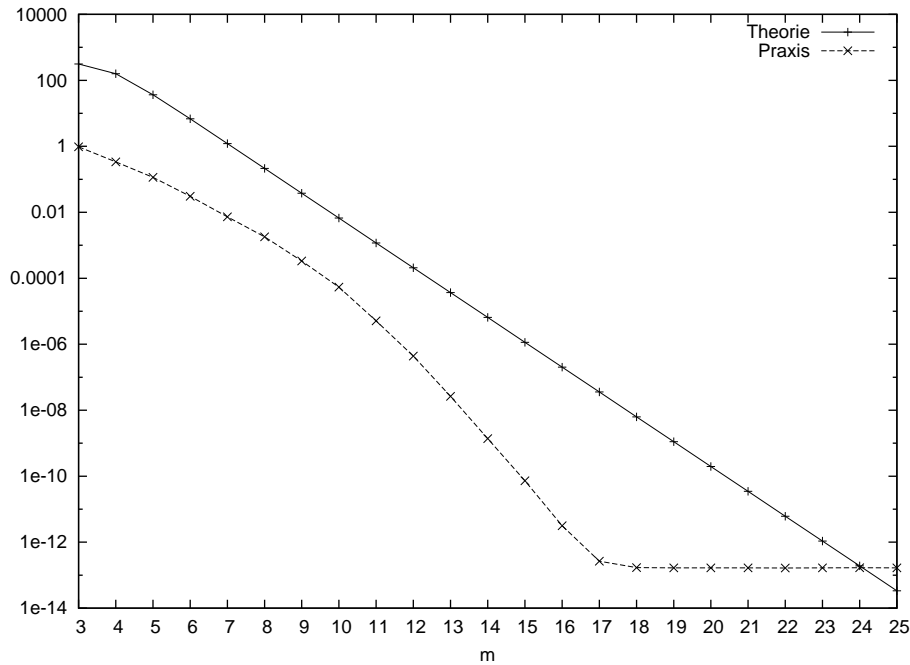
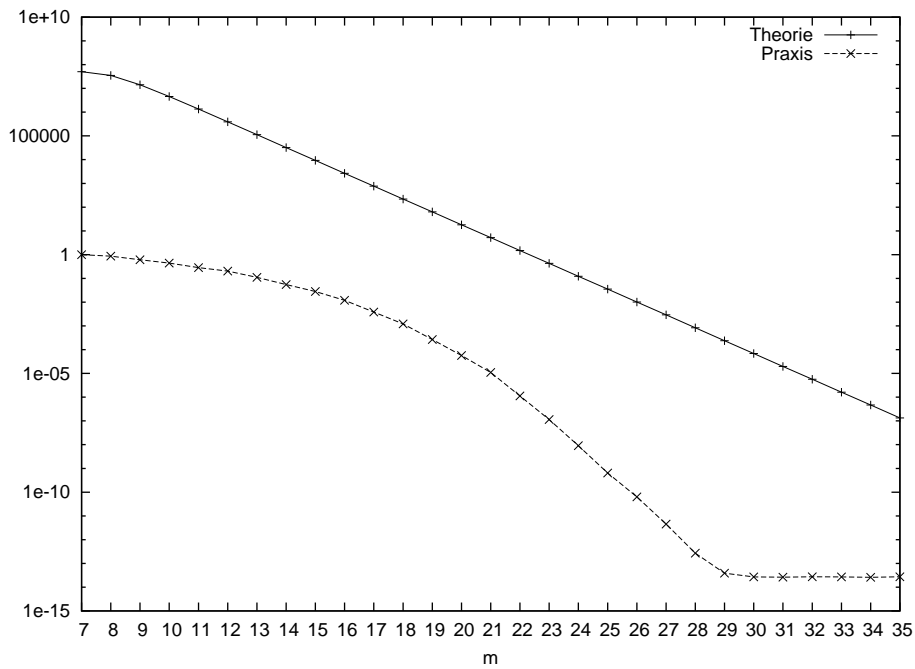


Abbildung 5.13: Vergleich zwischen Theorie und Praxis: λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/3$



Abschließend ist also zu bemerken, dass sich das (modifizierte) Lanczos-Verfahren sehr gut zur Berechnung von exponentiell abfallenden Eigenwerten eignet. Sowohl die theoretischen Ergebnisse aus Kapitel 4, als auch insbesondere die hier (und im Anhang) gezeigten praktischen Ergebnisse bestätigen dies.

Im nun folgenden Anhang wurden alle Eigenwerte nach der Formel (5.1) berechnet, wobei der Parameter c_1 immer gleich 1 gesetzt wurde. Der Parameter α wurde, wie in den vorangehenden Kapiteln, über 1, $1/2$ und $1/3$ variiert.

Die Grafiken des Anhangs präsentieren die Ergebnisse für $c_2 = 1/2$ und $c_2 = 3$. Für alle praktischen Ergebnisse wurde das modifizierte Lanczos-Verfahren verwendet. Da in dem Fall $\alpha = 1$ und $c_2 = 3$ das modifizierte Lanczos-Verfahren nach 8 Schritten abbricht, ist in den Grafiken zur Theorie und zum Fehlervergleich nur der Approximationsfehler zu den Eigenwerten λ_1 und λ_3 gezeichnet.

In den Grafiken wurde zum Vergleich der theoretischen oberen Schranke und den praktischen Ergebnissen wegen der schon gezeigten guten Konvergenz der Eigenwerte wieder Satz 4.2 mit dem Produkt über die echten Eigenwerten verwendet.

A Grafiken Kapitel 4 und 5

A.1 Grafiken zur Theorie

Abbildung A.1: Fehlerschranke aus Satz 4.2 für λ_1, λ_3 mit $\alpha = 1, c_2 = 3$

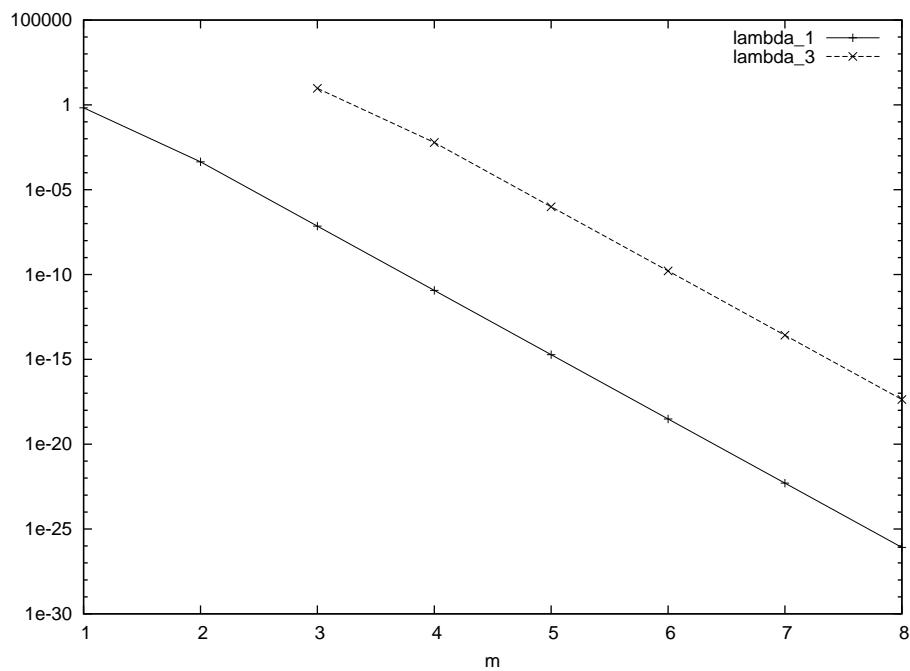


Abbildung A.2: Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1, c_2 = 1/2$

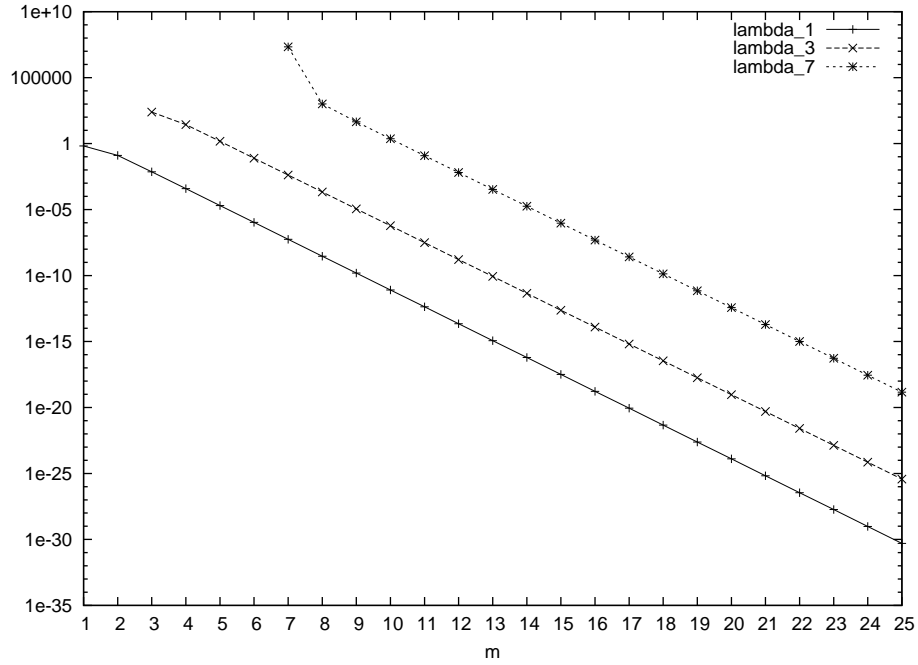


Abbildung A.3: Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/2, c_2 = 3$

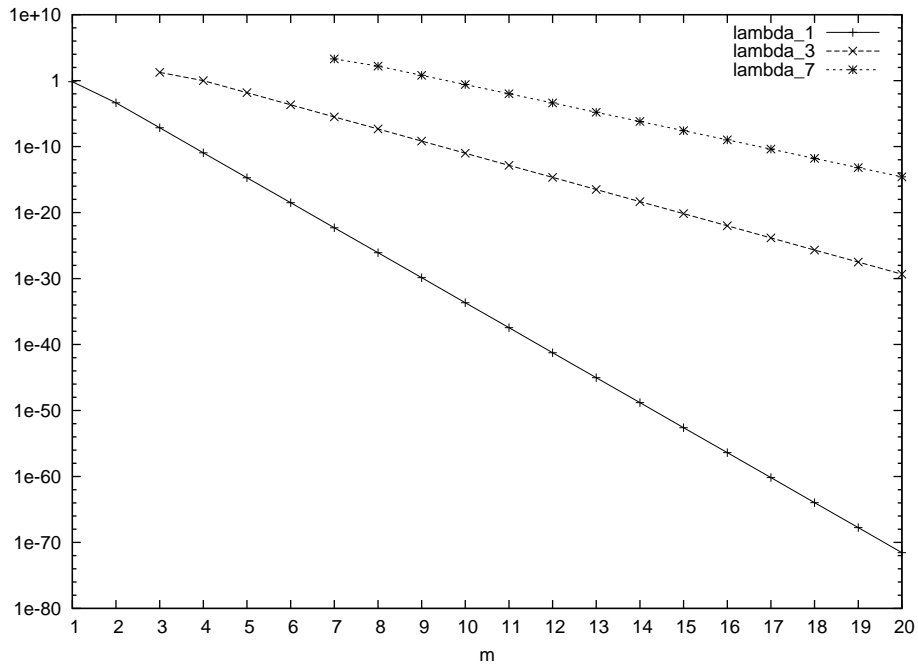


Abbildung A.4: Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/2, c_2 = 0,5$

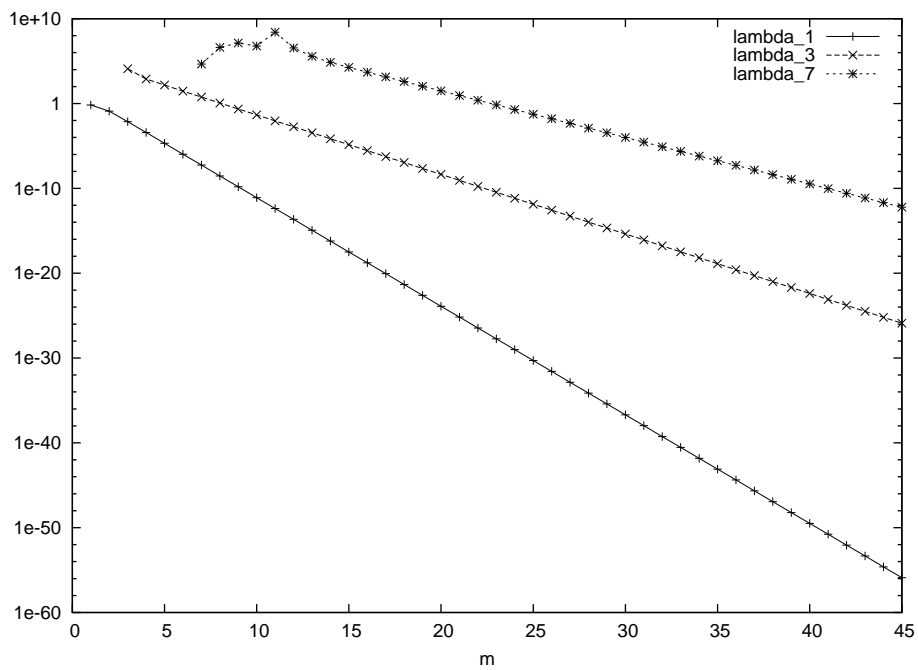


Abbildung A.5: Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/3, c_2 = 3$

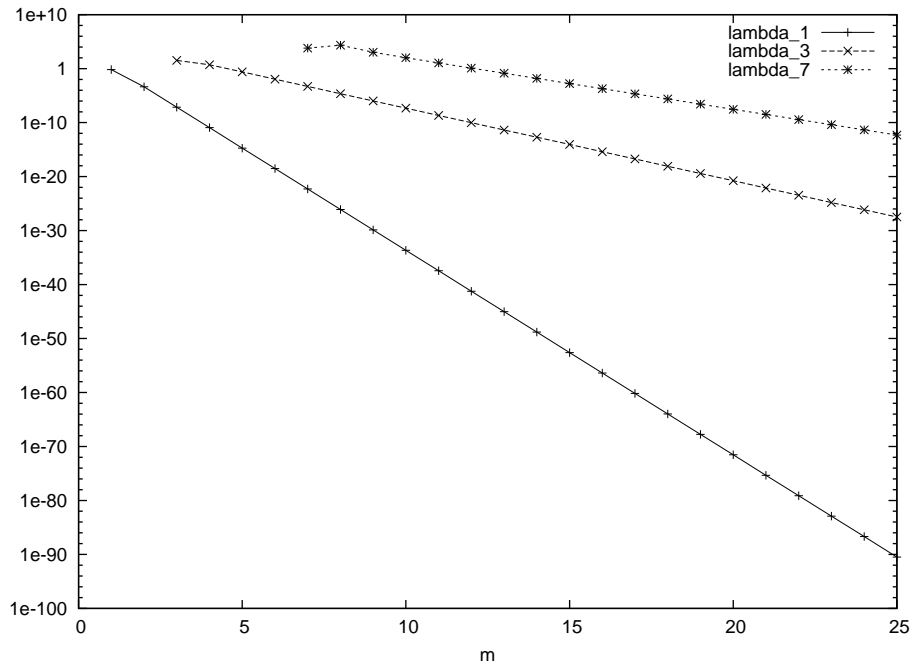
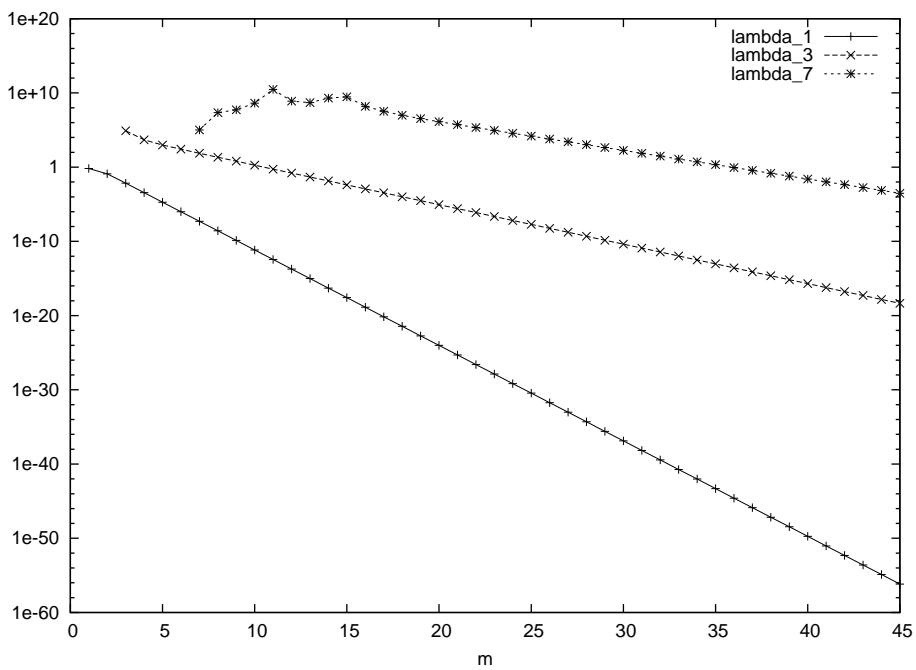


Abbildung A.6: Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/3, c_2 = 0,5$



A.2 Grafiken zum modifizierten Lanczos-Verfahren

Abbildung A.7: Relativer Fehler λ_1 bis λ_6 für $\alpha = 1$ und $c_2 = 3$

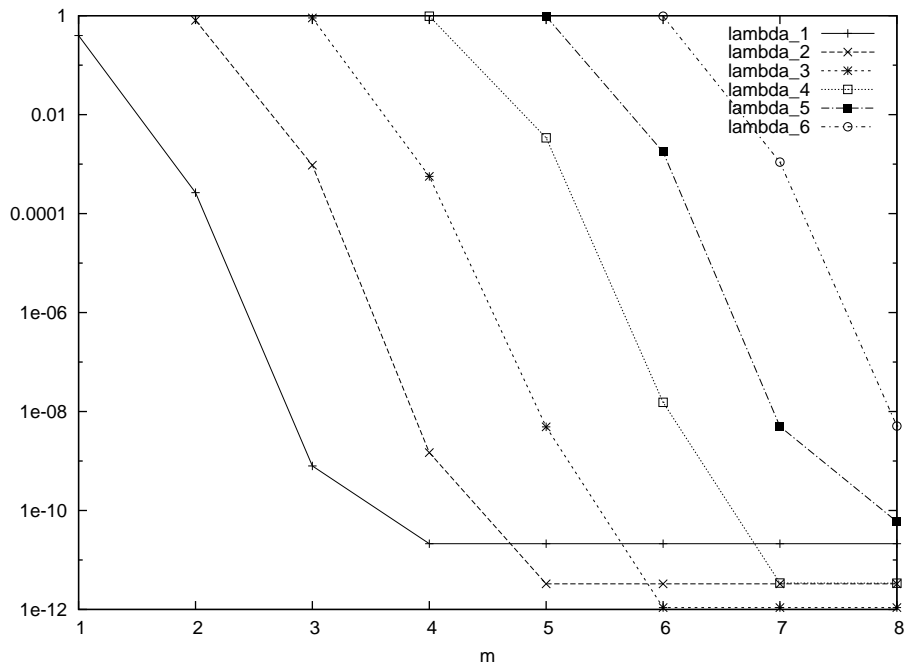


Abbildung A.8: Relativer Fehler λ_1 bis λ_7 für $\alpha = 1$ und $c_2 = 1/2$

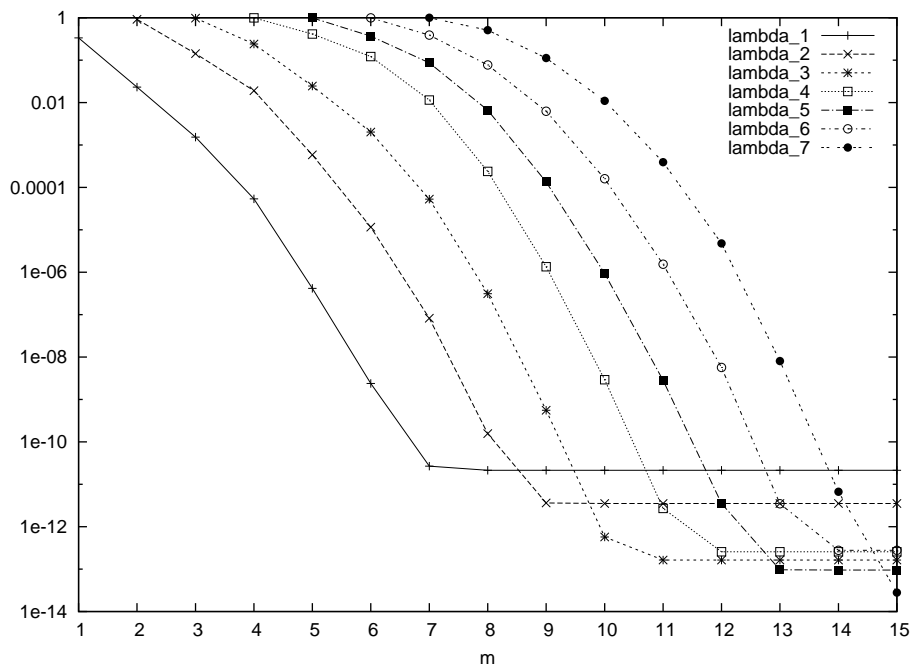


Abbildung A.9: Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/2$ und $c_2 = 3$

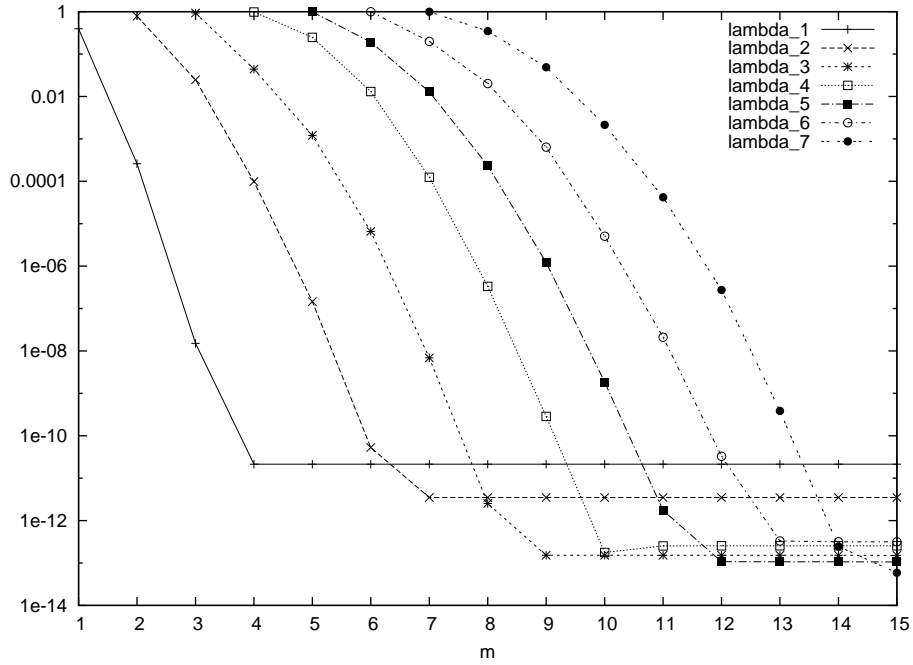


Abbildung A.10: Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/2$ und $c_2 = 1/2$

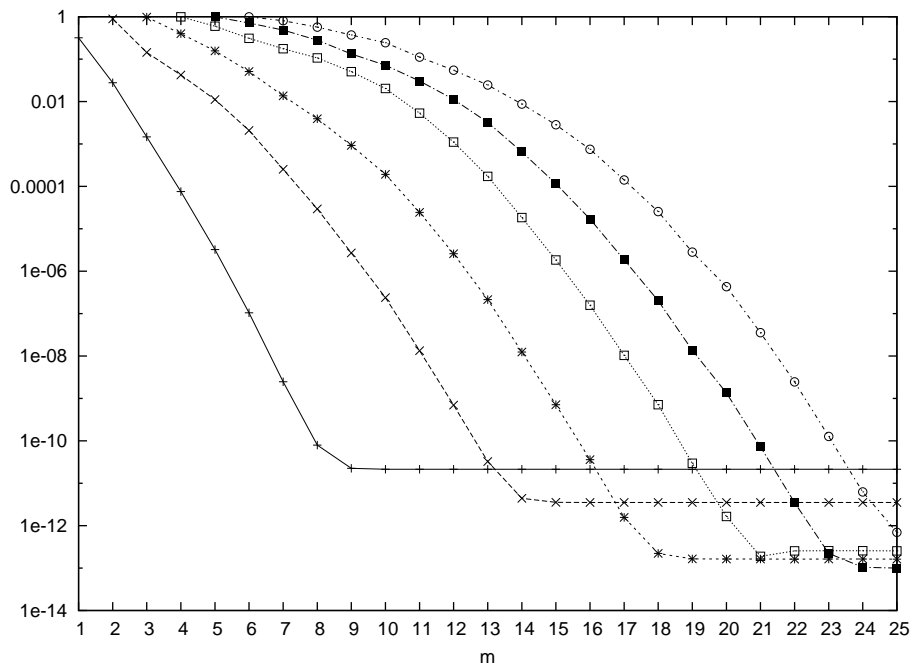


Abbildung A.11: Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/3$ und $c_2 = 3$

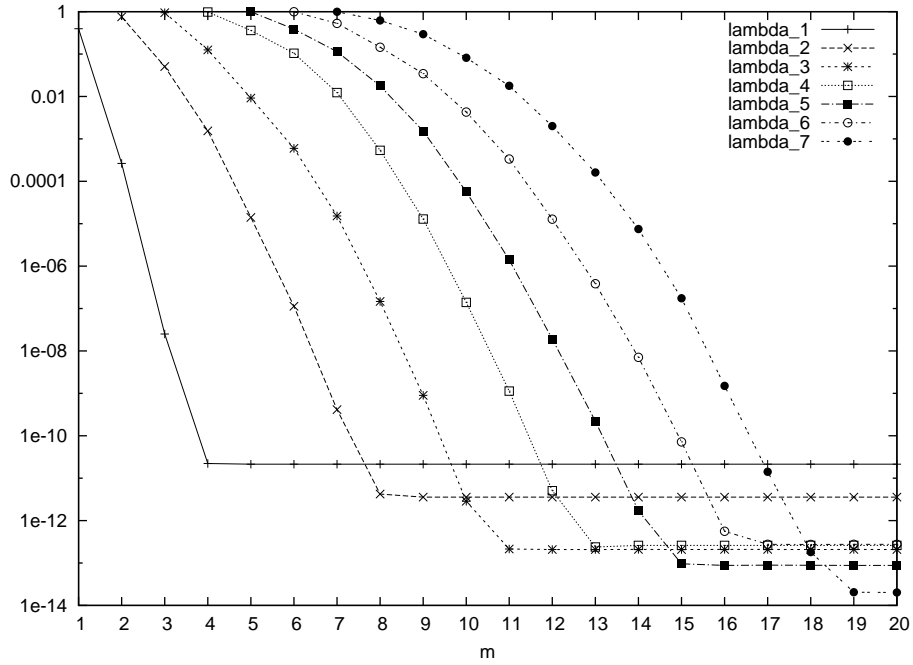
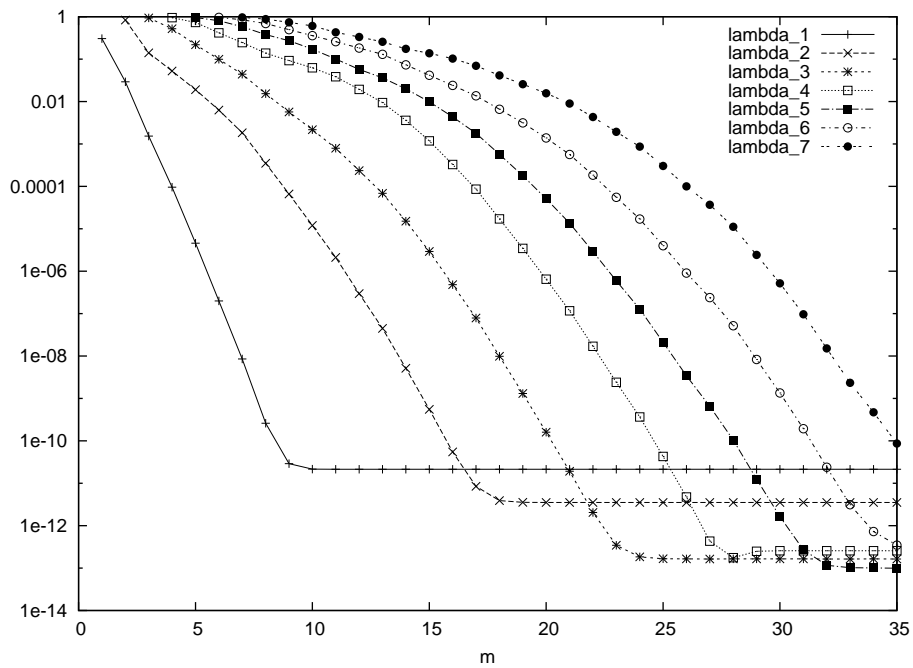


Abbildung A.12: Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/3$ und $c_2 = 1/2$



A.3 Vergleich zwischen Theorie und Praxis

Abbildung A.13: Fehlervergleich λ_3 für $\alpha = 1$ und $c_2 = 3$

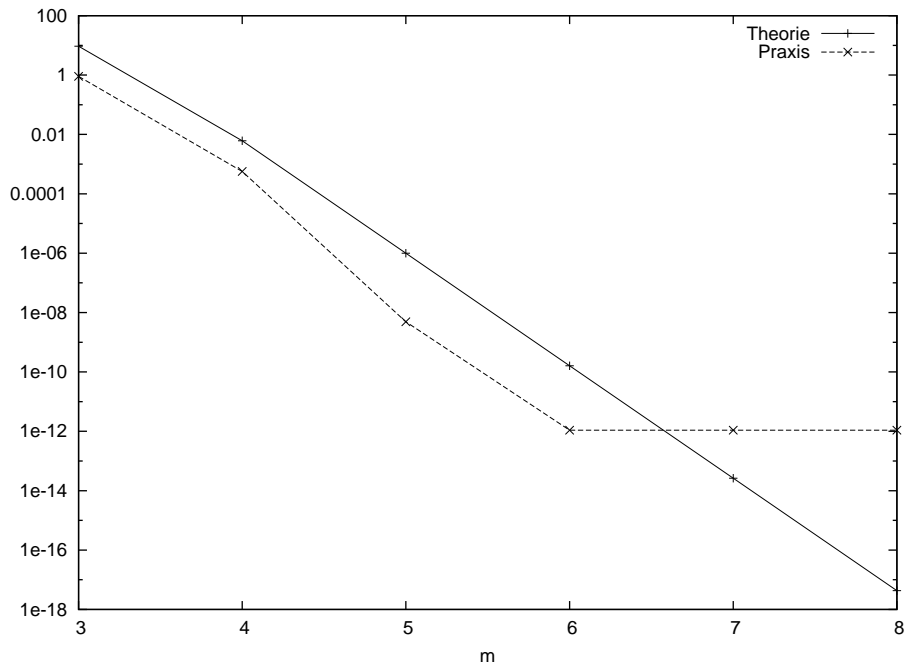


Abbildung A.14: Fehlervergleich λ_3 für $\alpha = 1$ und $c_2 = 1/2$

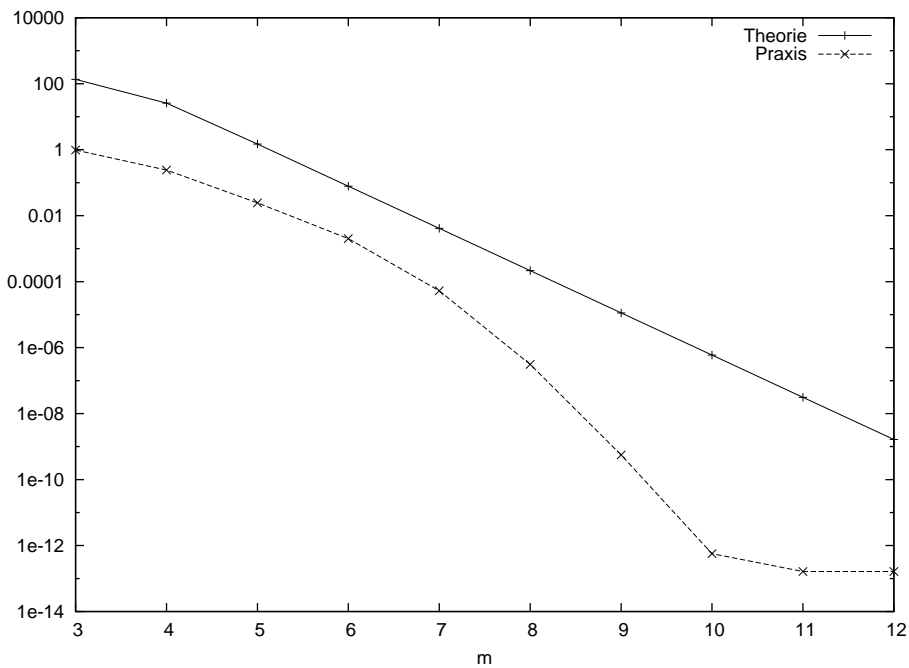


Abbildung A.15: Fehlervergleich λ_7 für $\alpha = 1$ und $c_2 = 1/2$

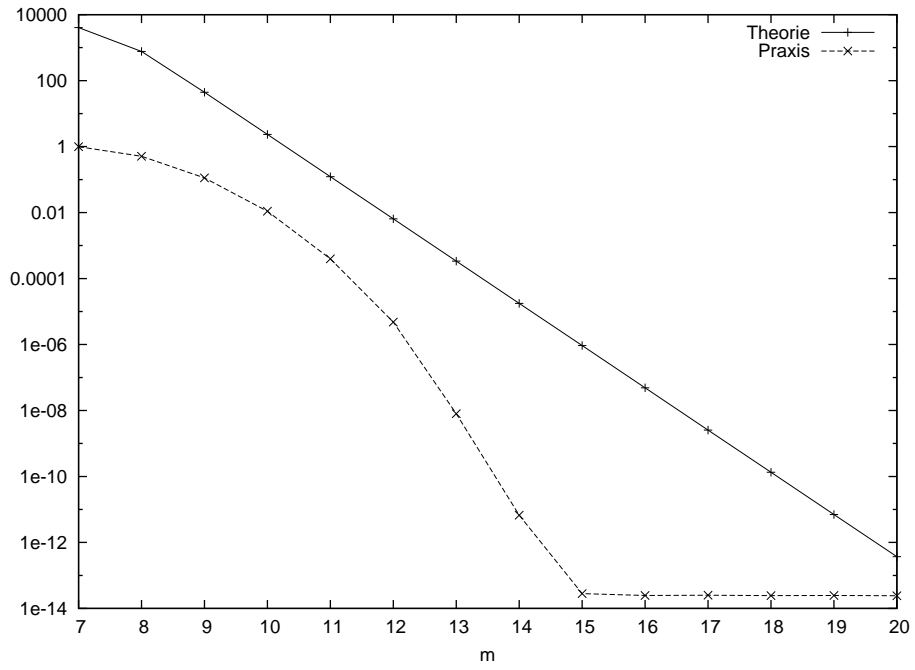


Abbildung A.16: Fehlervergleich λ_3 für $\alpha = 1/2$ und $c_2 = 3$

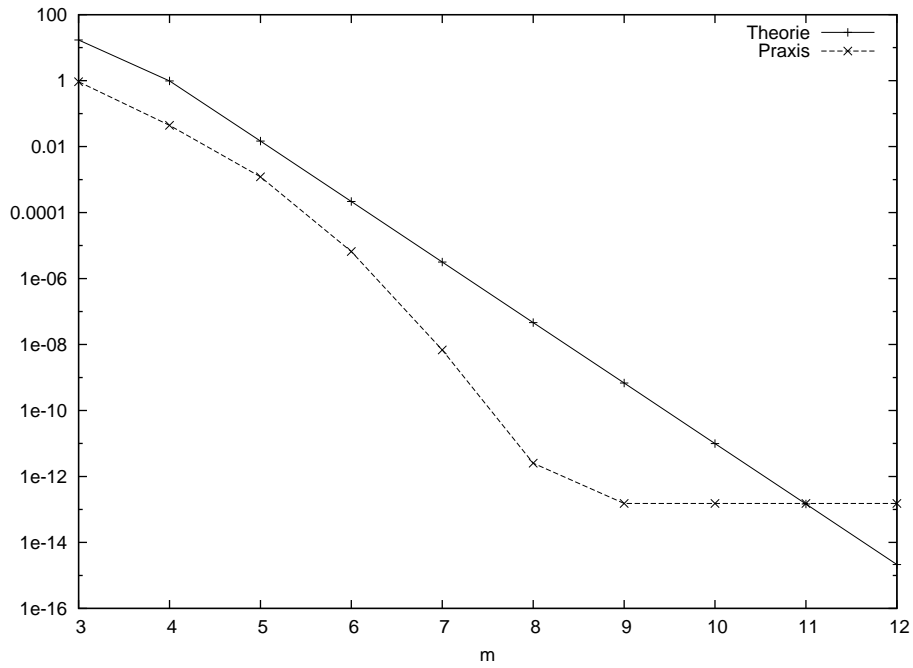


Abbildung A.17: Fehlervergleich λ_3 für $\alpha = 1/2$ und $c_2 = 1/2$

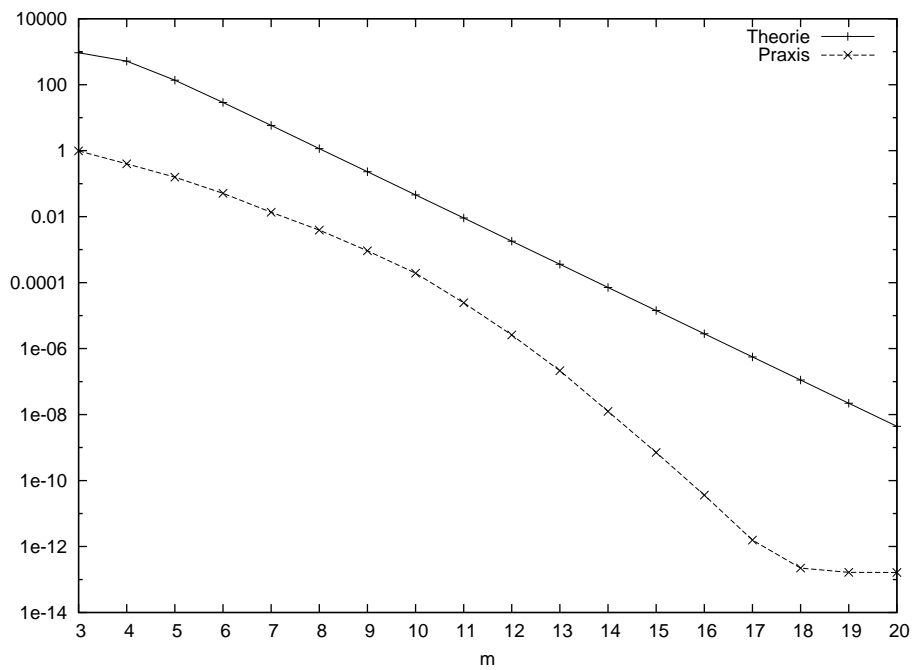


Abbildung A.18: Fehlervergleich λ_7 für $\alpha = 1/2$ und $c_2 = 3$

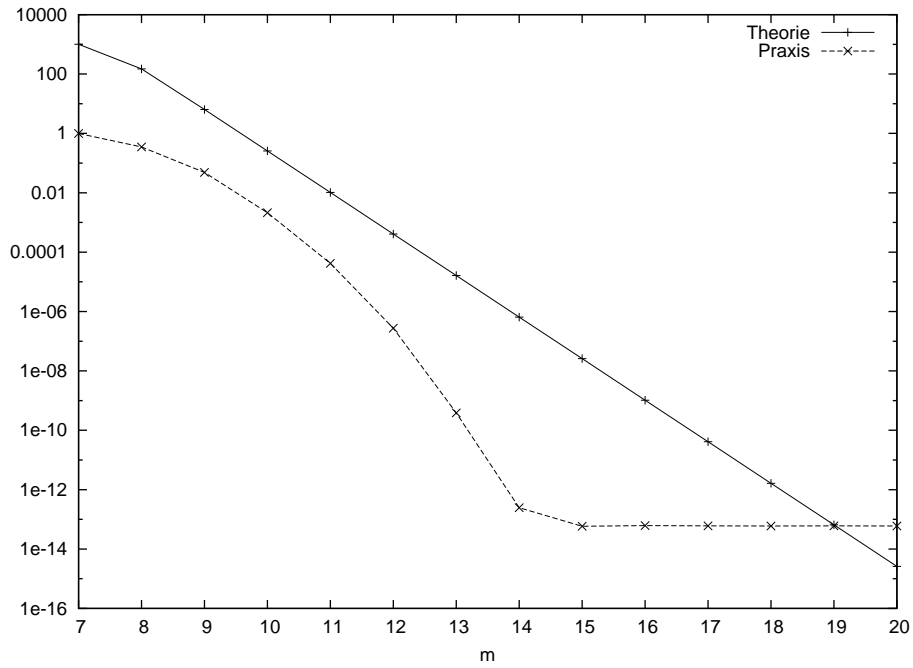


Abbildung A.19: Fehlervergleich λ_7 für $\alpha = 1/2$ und $c_2 = 1/2$

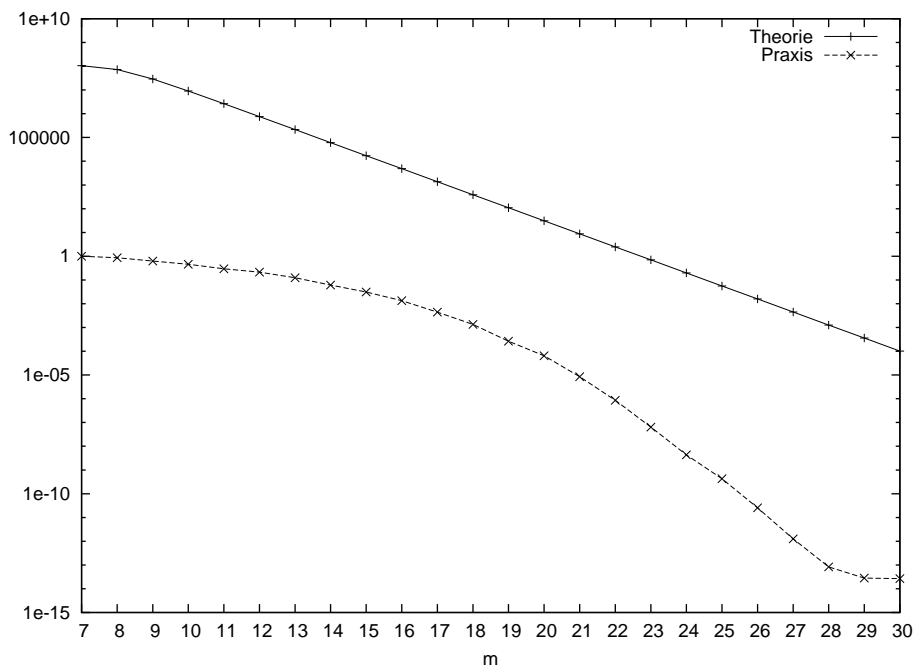


Abbildung A.20: Fehlervergleich λ_3 für $\alpha = 1/3$ und $c_2 = 3$

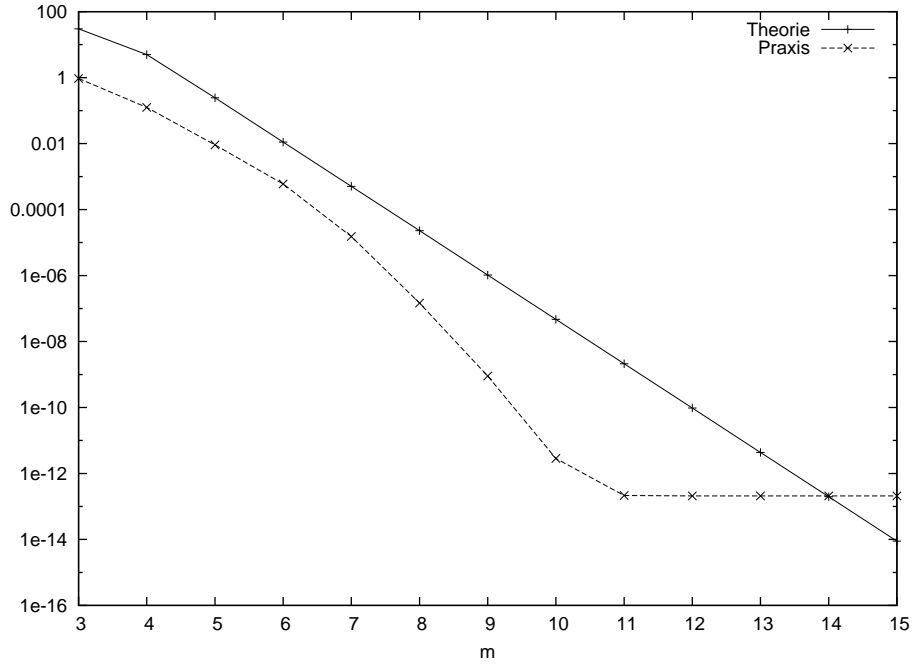


Abbildung A.21: Fehlervergleich λ_3 für $\alpha = 1/3$ und $c_2 = 1/2$

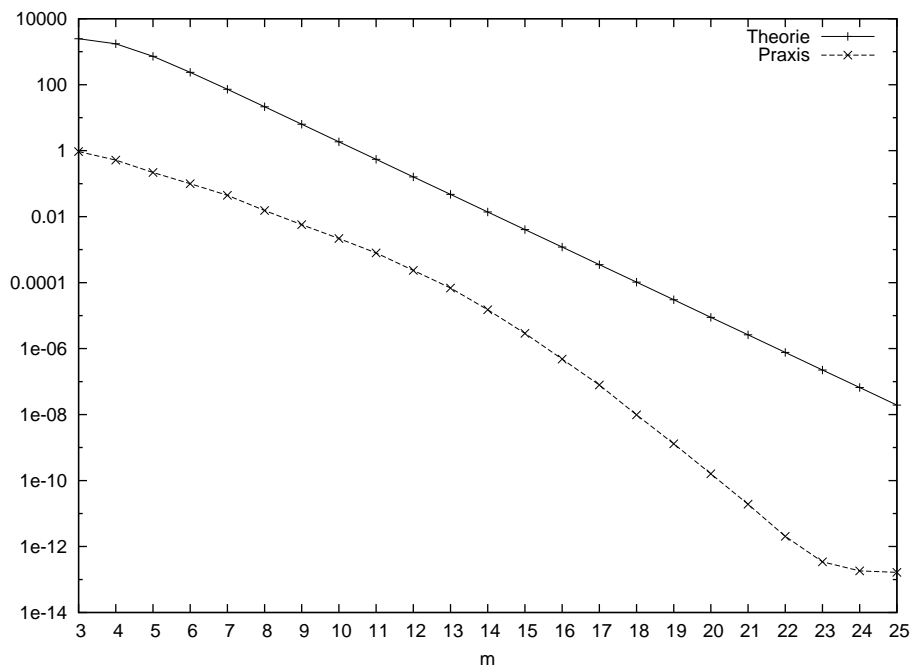


Abbildung A.22: Fehlervergleich λ_7 für $\alpha = 1/3$ und $c_2 = 3$

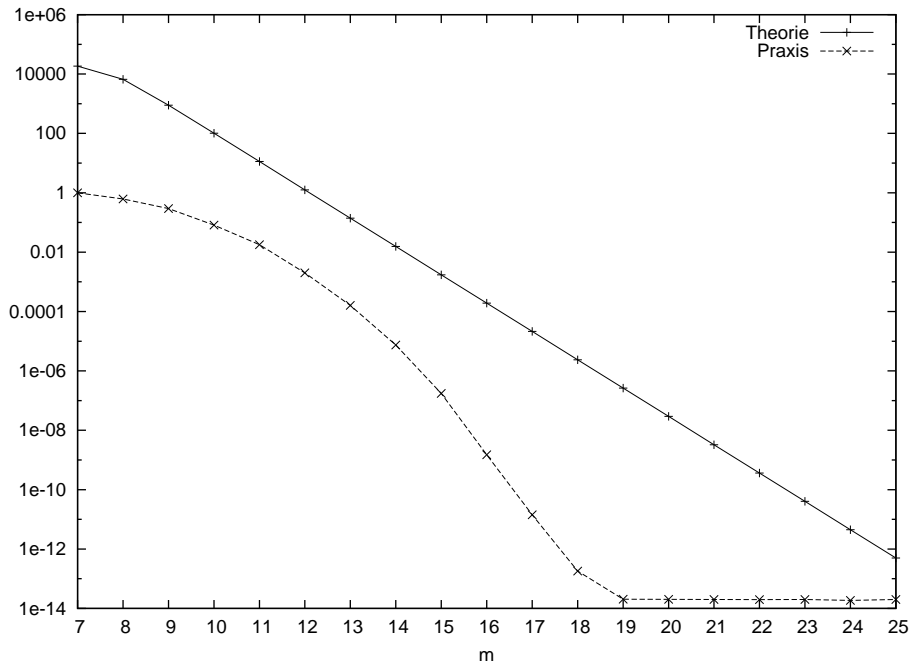
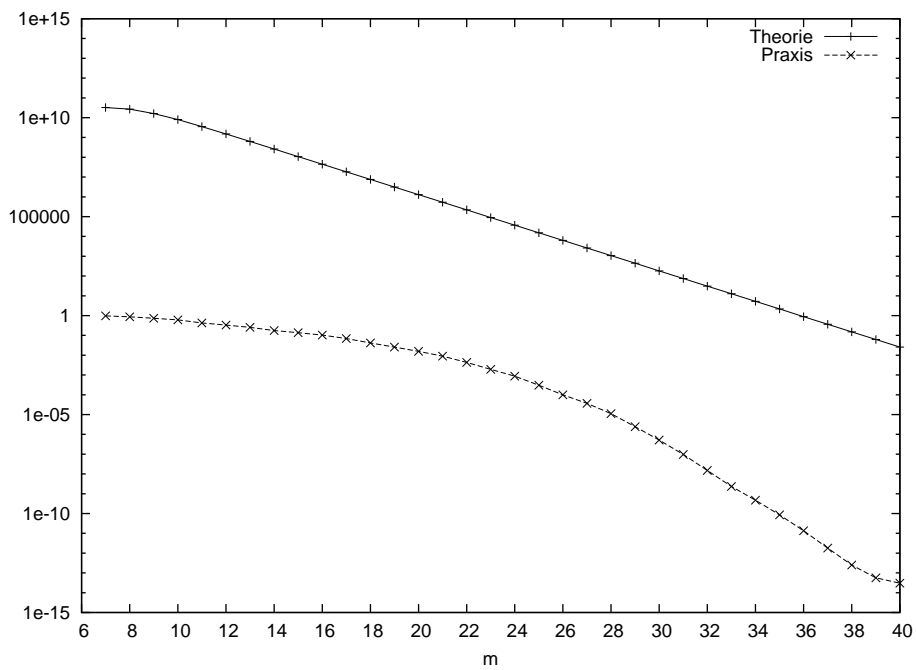


Abbildung A.23: Fehlervergleich λ_7 für $\alpha = 1/3$ und $c_2 = 1/2$



Abbildungsverzeichnis

1.1	Speicherplatzbedarf	2
1.2	Singulärwertzerlegung	3
4.1	γ_j für $j = 1$ bis $m = 15$ und $\alpha = 1, 1/2, 1/3$	48
4.2	$T_{m-j}(1 + 2\gamma_j)$ für $j = 1$ bis 15 und $\alpha = 1, 1/2, 1/3$	49
4.3	h_j für $j = 1$ bis 15 und $\alpha = 1, 1/2, 1/3$	53
4.4	Produktquadrat (4.56) aus Satz 4.2 zur Fehlerabschätzung von λ_5	54
4.5	Fehlerschranke λ_1, λ_3 und λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1$	60
4.6	Fehlerschranke λ_1, λ_3 und λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/2$	60
4.7	Fehlerschranke λ_1, λ_3 und λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/3$	61
4.8	Fehlerschrankenvergleich von λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/3$	62
5.1	Relativer Fehler $\lambda_1, \dots, \lambda_7$ für $c_1 = c_2 = 1$ und $\alpha = 1$	81
5.2	Relativer Fehler $\lambda_1, \dots, \lambda_7$ für $c_1 = c_2 = 1$ und $\alpha = 1/2$	82
5.3	Relativer Fehler $\lambda_1, \dots, \lambda_7$ für $c_1 = c_2 = 1$ und $\alpha = 1/3$	83
5.4	$\cos \angle(x^1, z^1), \dots, \cos \angle(x^7, z^7)$, für $c_1 = c_2 = 1$ und $\alpha = 1$	84
5.5	$\cos \angle(x^1, z^1), \dots, \cos \angle(x^7, z^7)$, für $c_1 = c_2 = 1$ und $\alpha = 1/2$	84
5.6	$\cos \angle(x^1, z^1), \dots, \cos \angle(x^7, z^7)$, für $c_1 = c_2 = 1$ und $\alpha = 1/3$	85
5.7	Relativer Fehler λ_1 für $c_1 = c_2 = 1$ und $\alpha = 1$	86
5.8	Relativer Fehler λ_3 für $c_1 = c_2 = 1$ und $\alpha = 1$	86
5.9	Relativer Fehler λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1$	87
5.10	Vergleich zwischen Theorie und Praxis: λ_3 für $c_1 = c_2 = 1$ und $\alpha = 1/2$	88
5.11	Vergleich zwischen Theorie und Praxis: λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/2$	88
5.12	Vergleich zwischen Theorie und Praxis: λ_3 für $c_1 = c_2 = 1$ und $\alpha = 1/3$	89
5.13	Vergleich zwischen Theorie und Praxis: λ_7 für $c_1 = c_2 = 1$ und $\alpha = 1/3$	89
A.1	Fehlerschranke aus Satz 4.2 für λ_1, λ_3 mit $\alpha = 1, c_2 = 3$	91
A.2	Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1, c_2 = 1/2$	92
A.3	Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/2, c_2 = 3$	93
A.4	Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/2, c_2 = 0,5$	93
A.5	Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/3, c_2 = 3$	94
A.6	Fehlerschranke aus Satz 4.2 für $\lambda_1, \lambda_3, \lambda_7$ mit $\alpha = 1/3, c_2 = 0,5$	94
A.7	Relativer Fehler λ_1 bis λ_6 für $\alpha = 1$ und $c_2 = 3$	95
A.8	Relativer Fehler λ_1 bis λ_7 für $\alpha = 1$ und $c_2 = 1/2$	95
A.9	Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/2$ und $c_2 = 3$	96

A.10 Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/2$ und $c_2 = 1/2$	96
A.11 Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/3$ und $c_2 = 3$	97
A.12 Relativer Fehler λ_1 bis λ_7 für $\alpha = 1/3$ und $c_2 = 1/2$	97
A.13 Fehlervergleich λ_3 für $\alpha = 1$ und $c_2 = 3$	98
A.14 Fehlervergleich λ_3 für $\alpha = 1$ und $c_2 = 1/2$	98
A.15 Fehlervergleich λ_7 für $\alpha = 1$ und $c_2 = 1/2$	99
A.16 Fehlervergleich λ_3 für $\alpha = 1/2$ und $c_2 = 3$	100
A.17 Fehlervergleich λ_3 für $\alpha = 1/2$ und $c_2 = 1/2$	100
A.18 Fehlervergleich λ_7 für $\alpha = 1/2$ und $c_2 = 3$	101
A.19 Fehlervergleich λ_7 für $\alpha = 1/2$ und $c_2 = 1/2$	101
A.20 Fehlervergleich λ_3 für $\alpha = 1/3$ und $c_2 = 3$	102
A.21 Fehlervergleich λ_3 für $\alpha = 1/3$ und $c_2 = 1/2$	102
A.22 Fehlervergleich λ_7 für $\alpha = 1/3$ und $c_2 = 3$	103
A.23 Fehlervergleich λ_7 für $\alpha = 1/3$ und $c_2 = 1/2$	103

Literaturverzeichnis

- [1] Steffen Börm. *Iterative Lösungsverfahren für große lineare Gleichungssysteme*. Skript, 2007.
- [2] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. Johns Hopkins University Press, 1996.
- [3] Wolfgang Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. Teubner Studienbücher, 1991.
- [4] Wolfgang Hackbusch. *Numerische Lösung von Eigenwertaufgaben*. Skript, 2005.
- [5] Norbert Köckler Hans-Rudolf Schwarz. *Numerische Mathematik*. Teubner, 2004.
- [6] W. Kahan and B.N. Parlett. How far should you go with the lanczos process? In *Sparse Matrix Computations*. Academic Press Inc., 1975.
- [7] Gerhard Opfer. *Numerische Mathematik für Anfänger*. vieweg, 2002.
- [8] Beresford N. Parlett. *The Symmetric Eigenvalueproblem*. siam: Classics In Applied Mathematics, 2002.

Erklärung

Hiermit versichere ich, die vorliegende Diplomarbeit
„Lanczos-Verfahren bei exponentiell abfallenden Eigenwerten“ selbstständig und nur
unter Verwendung der angegebenen Literatur angefertigt zu haben.

15.12.2008