

Numerische Lineare Algebra

Oliver Ernst

Professur Numerische Mathematik

Wintersemester 2017/18



Mathematik!
TU Chemnitz

① Einleitung

- 1.1 Lineare Gleichungssysteme
- 1.2 Matrixfunktionen
- 1.3 Modellreduktion
- 1.4 Eigenwertaufgaben

② Krylov-Unterraumverfahren

- 2.1 Projektionen
- 2.2 Orthogonale Projektionsverfahren
- 2.3 Krylov-Unterräume
- 2.4 Basen von Krylov-Räumen

③ Lineare Gleichungssysteme

- 3.1 Lösungsstrategien
- 3.2 Selbstadjungierte Probleme
- 3.3 Das Verfahren der konjugierten Gradienten
- 3.4 Das LSQR-Verfahren für lineare Ausgleichsprobleme
- 3.5 Vorkonditionierung

④ Matrixfunktionen

- 4.1 Erste Definition mithilfe der Jordanschen Normalform

- 4.2 Hermitesche Polynominterpolation
- 4.3 Alternative Darstellungen von Matrixfunktionen
- 4.4 Resolventenintegrale
- 4.5 Ein Beispiel

5 Krylov-Verfahren für Matrixfunktionen

- 5.1 Schranken für $\|f(\mathbf{A})\|$
- 5.2 Fehlerschranken für Krylov-Verfahren
- 5.3 Die Konvergenz des CG-Verfahrens

6 Das QR-Verfahren für Eigenwertaufgaben

- 6.1 Reduktion auf Hessenberg-Gestalt
- 6.2 Vektoriteration
- 6.3 QR-Iteration

- ① Einleitung
- ② Krylov-Unterraumverfahren
- ③ Lineare Gleichungssysteme
- ④ Matrixfunktionen
- ⑤ Krylov-Verfahren für Matrixfunktionen
- ⑥ Das QR-Verfahren für Eigenwertaufgaben

- 1 Einleitung
- 2 Krylov-Unterraumverfahren
 - 2.1 Projektionen
 - 2.2 Orthogonale Projektionsverfahren
 - 2.3 Krylov-Unterräume
 - 2.4 Basen von Krylov-Räumen
- 3 Lineare Gleichungssysteme
- 4 Matrixfunktionen
- 5 Krylov-Verfahren für Matrixfunktionen
- 6 Das QR-Verfahren für Eigenwertaufgaben

Definition 2.1

Zwei Teilräume \mathcal{N} und \mathcal{R} des \mathbb{C}^n heißen **komplementär**, wenn

- $\mathbb{C}^n = \mathcal{N} + \mathcal{R}$, d.h. zu jedem $\mathbf{v} \in \mathbb{C}^n$ existieren Vektoren $\mathbf{n} \in \mathcal{N}$ sowie $\mathbf{r} \in \mathcal{R}$ mit $\mathbf{v} = \mathbf{n} + \mathbf{r}$, und
- $\mathcal{N} \cap \mathcal{R} = \{\mathbf{0}\}$.

In diesem Fall nennt man \mathbb{C}^n die **direkte Summe** von \mathcal{N} und \mathcal{R} , geschrieben $\mathbb{C}^n = \mathcal{N} \oplus \mathcal{R}$.

Beachte: bei einer direkten Summe ist obige Darstellung $\mathbf{v} = \mathbf{n} + \mathbf{r}$ eindeutig bestimmt.

Beispiel 2.2

Ein beliebiger Unterraum $\mathcal{U} \subset \mathbb{C}^n$ bildet zusammen mit seinem **orthogonalen Komplement**

$$\mathcal{U}^\perp := \{\mathbf{x} \in \mathbb{C}^n : \mathbf{x} \perp \mathbf{u} \text{ für alle } \mathbf{u} \in \mathcal{U}\}$$

eine direkte Summe von \mathbb{C}^n .

Definition 2.3

Eine lineare Abbildung (Matrix) $\mathbf{P} \in \mathbb{C}^{n \times n}$ heißt **idempotent**, falls $\mathbf{P}^2 = \mathbf{P}$.

Lemma 2.4

Die Matrix $\mathbf{P} \in \mathbb{C}^{n \times n}$ sei idempotent. Dann bilden ihr Nullraum $\mathcal{N}(\mathbf{P})$ und ihr Bildraum $\mathcal{R}(\mathbf{P})$ komplementäre Unterräume.

Korollar 2.5

- (a) Eine idempotente Matrix ist durch Vorgabe ihres Kerns und Bildraums eindeutig festgelegt.
- (b) Umgekehrt gibt es zu jedem Paar \mathcal{N}, \mathcal{R} komplementärer Unterräume des \mathbb{C}^n genau eine idempotente Matrix \mathbf{P} mit $\mathcal{N}(\mathbf{P}) = \mathcal{N}$ und $\mathcal{R}(\mathbf{P}) = \mathcal{R}$.

Definition 2.6

Wir nennen die Abbildung aus Korollar 2.5 die **(schiefe) Projektion auf \mathcal{R} längs \mathcal{N}** oder auch die **(schiefe) Projektion auf \mathcal{R} orthogonal zu $\mathcal{S} := \mathcal{N}^\perp$** und schreiben $P = P_{\mathcal{R}, \mathcal{S}}$.

Man beachte, dass die Unterräume \mathcal{R} und $\mathcal{S} = \mathcal{N}^\perp$ wegen $\mathbb{C}^n = \mathcal{R} \oplus \mathcal{N}$ dieselbe Dimension besitzen. Die lineare Abbildung bzw. die Matrix $P_{\mathcal{R}, \mathcal{S}}$ ist durch $P_{\mathcal{R}, \mathcal{S}}^2 = P_{\mathcal{R}, \mathcal{S}}$ sowie

$$P_{\mathcal{R}, \mathcal{S}} \mathbf{x} \in \mathcal{R} \text{ und } \mathbf{x} - P_{\mathcal{R}, \mathcal{S}} \mathbf{x} \perp \mathcal{S} \quad \text{für alle } \mathbf{x} \in \mathbb{C}^n$$

eindeutig festgelegt.

Lemma 2.7

Seien \mathcal{R} und \mathcal{S} Unterräume des \mathbb{C}^n derselben Dimension m . Dann sind die folgenden Aussagen zueinander äquivalent:

- (a) Die Projektion $P_{\mathcal{R}, \mathcal{S}}$ auf \mathcal{R} orthogonal zu \mathcal{S} existiert.
- (b) $\mathbb{C}^n = \mathcal{R} \oplus \mathcal{S}^\perp$.
- (c) $\mathcal{R} \cap \mathcal{S}^\perp = \{\mathbf{0}\}$.
- (d) Für beliebige Basen $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ von \mathcal{R} und $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\}$ von \mathcal{S} ist die Matrix

$$\mathbf{S}^H \mathbf{R} := [\mathbf{s}_1 \ \mathbf{s}_2 \ \dots \ \mathbf{s}_m]^H [\mathbf{r}_1 \ \mathbf{r}_2 \ \dots \ \mathbf{r}_m] \in \mathbb{C}^{m \times m}$$

invertierbar.^a

- (e) Es gibt Basen $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ von \mathcal{R} und $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\}$ von \mathcal{S} mit

$$(\mathbf{r}_i, \mathbf{s}_j) = \delta_{i,j} = \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j \end{cases} \quad (i, j = 1, 2, \dots, m).$$

Man spricht von **biorthogonalen Basen**.

^aOrthogonalität wird hier über das Euklidische Innenprodukt erklärt. Legt man ein anderes Innenprodukt $\langle \cdot, \cdot \rangle$ zugrunde, so muss die Matrix $\mathbf{S}^H \mathbf{R}$ durch die Matrix $[(\mathbf{r}_i, \mathbf{s}_j)]_{1 \leq i, j \leq m}$ ersetzt werden.

Ein wichtiger Spezialfall ergibt sich für identische Räume

$$\mathcal{R} = \mathcal{S} = \mathcal{U}.$$

(Beachte, dass $\mathbb{C}^n = \mathcal{U} \oplus \mathcal{U}^\perp$ für jeden Unterraum \mathcal{U} des \mathbb{C}^n gilt).

Wir schreiben dann $P_{\mathcal{U}}$ anstelle von $P_{\mathcal{U}, \mathcal{U}}$ und nennen $P_{\mathcal{U}}$ die **Orthogonalprojektion auf \mathcal{U}** . Sie ist durch $P_{\mathcal{U}}^2 = P_{\mathcal{U}}$ sowie

$$P_{\mathcal{U}} \mathbf{x} \in \mathcal{U} \text{ und } \mathbf{x} - P_{\mathcal{U}} \mathbf{x} \perp \mathcal{U} \quad \text{für alle } \mathbf{x} \in \mathbb{C}^n$$

eindeutig bestimmt.

Korollar 2.8

Seien $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ und $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\}$ Basen der m -dimensionalen Unterräume \mathcal{R} bzw. \mathcal{S} . Mit

$$\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \cdots \quad \mathbf{r}_m] \quad \text{und} \quad \mathbf{S} = [\mathbf{s}_1 \quad \mathbf{s}_2 \quad \cdots \quad \mathbf{s}_m]$$

seien die zugehörigen $(n \times m)$ -Matrizen bezeichnet. Existiert die Projektion $P_{\mathcal{R}, \mathcal{S}}$, so ist sie gegeben durch

$$\mathbf{x} \mapsto \mathbf{R}(\mathbf{S}^H \mathbf{R})^{-1} \mathbf{S}^H \mathbf{x}, \quad \mathbf{x} \in \mathbb{C}^n.$$

Ist $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ eine Basis des Unterraums \mathcal{U} und bezeichnet

$$\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m]$$

die zugehörige $(n \times m)$ -Matrix, so ist die Orthogonalprojektion $P_{\mathcal{U}}$ gegeben durch

$$\mathbf{x} \mapsto \mathbf{U}(\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \mathbf{x}, \quad \mathbf{x} \in \mathbb{C}^n.$$

Lemma 2.9

Sei $\mathbf{x} \in \mathbb{C}^n$. Dann ist

$$\|\mathbf{x} - P_{\mathcal{U}} \mathbf{x}\| < \|\mathbf{x} - \mathbf{u}\| \quad \text{für alle } \mathbf{u} \in \mathcal{U}, \mathbf{u} \neq P_{\mathcal{U}} \mathbf{x},$$

d.h. $P_{\mathcal{U}} \mathbf{x}$ ist die eindeutig bestimmte Bestapproximation an \mathbf{x} aus \mathcal{U} bezüglich der Norm $\|\cdot\|$.

- 1 Einleitung
- 2 Krylov-Unterraumverfahren
 - 2.1 Projektionen
 - 2.2 Orthogonale Projektionsverfahren
 - 2.3 Krylov-Unterräume
 - 2.4 Basen von Krylov-Räumen
- 3 Lineare Gleichungssysteme
- 4 Matrixfunktionen
- 5 Krylov-Verfahren für Matrixfunktionen
- 6 Das QR-Verfahren für Eigenwertaufgaben

- Sei in diesem Abschnitt \mathcal{U} stets ein Unterraum des \mathbb{C}^n , sei $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ eine Basis von \mathcal{U} und sei schließlich $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_m] \in \mathbb{C}^{n \times m}$.
- Weil \mathbf{U} vollen Spaltenrang besitzt, ist $\mathbf{U}^H \mathbf{U}$ invertierbar und

$$\mathbf{U}^\dagger := (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \in \mathbb{C}^{m \times n}$$

ist wohldefiniert.

- Die Matrix \mathbf{U}^\dagger ist die **Moore-Penrose-Inverse** von \mathbf{U} , also die eindeutig bestimmte Matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$, welche folgende vier Bedingungen erfüllt:

$$\mathbf{U} \mathbf{X} \mathbf{U} = \mathbf{U}, \quad \mathbf{X} \mathbf{U} \mathbf{X} = \mathbf{X}, \quad (\mathbf{X} \mathbf{U})^H = \mathbf{X} \mathbf{U}, \quad (\mathbf{U} \mathbf{X})^H = \mathbf{U} \mathbf{X}.$$

- In unserem Fall ist $\mathbf{U}^\dagger \mathbf{U} = \mathbf{I}_m$ (offensichtlich) und

$$\mathbf{U} \mathbf{U}^\dagger = \mathbf{U} (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H$$

die Orthogonalprojektion auf \mathcal{U} (vgl. Korollar 2.8).

- Für spätere Zwecke halten wir fest, dass für jede invertierbare Matrix $\mathbf{T} \in \mathbb{C}^{m \times m}$ gilt

$$(\mathbf{UT})^\dagger = \mathbf{T}^{-1} \mathbf{U}^\dagger \quad (2.1)$$

denn

$$\begin{aligned} (\mathbf{UT})^\dagger &= ((\mathbf{UT})^H (\mathbf{UT}))^{-1} (\mathbf{UT})^H = (\mathbf{T}^H \mathbf{U}^H \mathbf{UT})^{-1} \mathbf{T}^H \mathbf{U}^H \\ &= \mathbf{T}^{-1} (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{T}^{-H} \mathbf{T}^H \mathbf{U}^H = \mathbf{T}^{-1} (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H = \mathbf{T}^{-1} \mathbf{U}^\dagger. \end{aligned}$$

- Sei nun $\mathbf{A} \in \mathbb{C}^{n \times n}$. Wir nennen

$$\mathbf{A}_U := \mathbf{U}^\dagger \mathbf{A} \mathbf{U} \in \mathbb{C}^{m \times m} \quad (2.2)$$

den **Rayleigh-Quotienten** von \mathbf{A} bez. \mathbf{U} .

Bemerkung 2.10

Die Bezeichnung „Rayleigh-Quotient“ für A_U bedarf einer Erklärung: Ist \mathcal{U} eindimensional, also $U = u$ mit $u \neq 0$, so ist

$$A_U = \frac{u^H A u}{u^H u} = (u/\|u\|)^H A (u/\|u\|)$$

eine skalare Größe, d.h. ein Rayleigh-Quotient von A im klassischen Sinne. Die Menge dieser Rayleigh-Quotienten, also

$$W(A) := \left\{ \frac{u^H A u}{u^H u} : u \neq 0 \right\} = \{u^H A u : \|u\| = 1\}, \quad (2.3)$$

wird der **Wertebereich** von A genannt. Die Menge $W(A) \subset \mathbb{C}$ ist kompakt und konvex. Sie enthält die Eigenwerte von A .

Lemma 2.11 (Eigenschaften des Rayleigh-Quotienten)

Der Rayleigh-Quotient \mathbf{A}_U (2.2) von \mathbf{A} bezüglich U besitzt folgende Eigenschaften:

- (a) Die Matrix \mathbf{A}_U stellt die lineare Abbildung

$$P_{\mathcal{U}} \mathbf{A} |_{\mathcal{U}} : \mathcal{U} \rightarrow \mathcal{U},$$

also die Einschränkung von $P_{\mathcal{U}} \mathbf{A}$ auf \mathcal{U} bezüglich der Basis U dar.

- (b) Ist $\{\mathbf{u}_{m+1}, \mathbf{u}_{m+2}, \dots, \mathbf{u}_n\}$ eine Basis von \mathcal{U}^\perp und bezeichnet \hat{U} die Matrix $\begin{bmatrix} U & \mathbf{u}_{m+1} & \cdots & \mathbf{u}_n \end{bmatrix} \in \mathbb{C}^{n \times n}$, deren Spalten eine Basis des \mathbb{C}^n bilden, so repräsentiert

$$\begin{bmatrix} \mathbf{A}_U & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \in \mathbb{C}^{n \times n}$$

die lineare Abbildung $P_{\mathcal{U}} \mathbf{A} P_{\mathcal{U}} : \mathbb{C}^n \rightarrow \mathbb{C}^n$ bez. der Basis \hat{U} .

- (c) Sind die Spalten von U und V Basen des Unterraums \mathcal{U} , d.h. gibt es eine invertierbare Matrix $T \in \mathbb{C}^{m \times m}$ mit $V = UT$, so sind die Rayleigh-Quotienten \mathbf{A}_U und \mathbf{A}_V ähnliche Matrizen, genauer gilt $T \mathbf{A}_V = \mathbf{A}_U T$.

Definition 2.12

Es seien $\mathbf{A} \in \mathbb{C}^{n \times n}$, \mathcal{U} ein Unterraum des \mathbb{C}^n mit den Spalten von $\mathbf{U} \in \mathbb{C}^{n \times m}$ als Basis und \mathbf{A}_U der Rayleigh-Quotient von \mathbf{A} bez. \mathcal{U} .

- (a) Seien außerdem $f : \mathbb{C} \supseteq D \rightarrow \mathbb{C}$ eine Funktion, sodass $f(\mathbf{A})$ definiert ist, und $\mathbf{b} \in \mathbb{C}^n$. Existiert $f(\mathbf{A}_U)$, so nennen wir

$$\mathbf{U} f(\mathbf{A}_U) \mathbf{U}^\dagger \mathbf{b} = \mathbf{U} f(\mathbf{U}^\dagger \mathbf{A} \mathbf{U}) \mathbf{U}^\dagger \mathbf{b}$$

die **Rayleigh-Ritz-Approximation** an $f(\mathbf{A})\mathbf{b}$ aus \mathcal{U} .

- (b) Ist (ϑ, \mathbf{x}) ein Eigenpaar des Rayleigh-Quotienten \mathbf{A}_U , d.h.

$$\mathbf{x} \in \mathbb{C}^m \quad \text{mit} \quad \mathbf{A}_U \mathbf{x} = \vartheta \mathbf{x},$$

so nennen wir $(\vartheta, \mathbf{U}\mathbf{x})$ ein **Ritz-Paar** von \mathbf{A} bezüglich \mathcal{U} . Die Zahl ϑ heißt **Ritz-Wert** von \mathbf{A} bezüglich \mathcal{U} , der Vektor $\mathbf{U}\mathbf{x} \in \mathbb{C}^n$ heißt **Ritz-Vektor** von \mathbf{A} bez. \mathcal{U} .

Bemerkung 2.13

Sowohl Rayleigh-Ritz-Approximationen wie auch Ritz-Paare hängen vom Unterraum \mathcal{U} , aber nicht von der gewählten Basis ab:

Die Spalten von U und die Spalten von V seien jeweils Basen von \mathcal{U} . Es gibt daher eine invertierbare Matrix $T \in \mathbb{C}^{m \times m}$ mit $V = UT$. Wir wissen bereits, dass dies

$$V^\dagger = T^{-1}U^\dagger \quad \text{und} \quad A_V = T^{-1}A_U T$$

(siehe Lemma 2.11) impliziert. Damit ergibt sich

$$\begin{aligned} Vf(A_V)V^\dagger b &= UTf(T^{-1}A_U T)T^{-1}U^\dagger b = UTT^{-1}f(A_U)TT^{-1}U^\dagger b \\ &= Uf(A_U)U^\dagger. \end{aligned}$$

Da aus $A_V y = \vartheta y$ sofort $A_U(Ty) = \vartheta(Ty)$ folgt und $UTy = Vy$ gilt, sind auch Ritz-Paare basisunabhängig.

Lemma 2.14

Seien $\mathbf{A} \in \mathbb{C}^{n \times n}$, \mathcal{U} ein Unterraum des \mathbb{C}^n und $(\vartheta, \mathbf{y} = \mathbf{U}\mathbf{x})$ ein Ritz-Paar von \mathbf{A} bezüglich \mathcal{U} . Dann gilt

$$(\mathbf{A}\mathbf{y} - \vartheta\mathbf{y}, \mathbf{u}) = 0 \quad \text{oder} \quad (\mathbf{A}\mathbf{y}, \mathbf{u}) = (\vartheta\mathbf{y}, \mathbf{u}) \quad \text{für alle } \mathbf{u} \in \mathcal{U}.$$

- Lemma 2.14 besagt, dass das **Residuum**, welches mit dem Ritz-Paar (ϑ, \mathbf{y}) assoziiert ist, also $\mathbf{A}\mathbf{y} - \vartheta\mathbf{y}$, orthogonal zum Approximationsraum \mathcal{U} ist.
- Man kann nicht erwarten, dass analoge **Galerkin-Bedingungen** für beliebige Rayleigh-Ritz-Approximationen an $f(\mathbf{A})\mathbf{b}$ gelten, allein schon deshalb, weil es i. Allg. unmöglich ist, Näherungen an $f(\mathbf{A})\mathbf{b}$ ein Residuum zuzuordnen.

Krylov-Unterraumverfahren

Orthogonale Projektionsverfahren

- Die Funktion $f(\zeta) = 1/\zeta$ stellt eine Ausnahme dar: $f(\mathbf{A})\mathbf{b}$ ist die Lösung des linearen Gleichungssystems $\mathbf{A}\mathbf{x} = \mathbf{b}$ und für jede Näherung $\tilde{\mathbf{x}}$ an $\mathbf{A}^{-1}\mathbf{b}$ ist das Residuum $\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}$ definiert.

Lemma 2.15

Es seien $\mathbf{A} \in \mathbb{C}^{n \times n}$ und $\mathbf{A}_U \in \mathbb{C}^{m \times m}$ invertierbar. Dann gilt für die Rayleigh-Ritz-Approximation $\mathbf{x}_m = \mathbf{U}\mathbf{A}_U^{-1}\mathbf{U}^\dagger\mathbf{b}$ an $\mathbf{A}^{-1}\mathbf{b}$

$$(\mathbf{b} - \mathbf{A}\mathbf{x}_m, \mathbf{u}) = 0 \quad \text{oder} \quad (\mathbf{A}\mathbf{x}_m, \mathbf{u}) = (\mathbf{b}, \mathbf{u}) \quad \text{für alle } \mathbf{u} \in \mathcal{U}.$$

Krylov-Unterraumverfahren

Orthogonale Projektionsverfahren

- Einen weiteren wichtigen Spezialfall, bei dem ein Residuum eingeführt werden kann, stellt die Matrixexponentialfunktion dar, denn $\mathbf{y}(t) = \exp(t\mathbf{A})\mathbf{b}$ löst das Anfangswertproblem $\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t)$, $\mathbf{y}(0) = \mathbf{b}$.
- Zur Beurteilung der Qualität einer Näherung $\mathbf{x}_m(t) \approx \mathbf{y}(t)$ bietet sich die Untersuchung des Residuums $\mathbf{r}(t) = \mathbf{x}'_m(t) - \mathbf{A}\mathbf{x}_m(t)$ an.

Lemma 2.16

Für die Rayleigh-Ritz-Approximation

$$\mathbf{x}_m(t) = \mathbf{U} \exp(t\mathbf{A}_U) \mathbf{U}^\dagger \mathbf{b} \approx \exp(t\mathbf{A}) \mathbf{b}$$

gilt

$$(\mathbf{x}'_m(t) - \mathbf{A}\mathbf{x}_m(t), \mathbf{u}) = 0 \text{ oder } (\mathbf{x}'_m(t), \mathbf{u}) = (\mathbf{A}\mathbf{x}_m(t), \mathbf{u}) \quad \text{für alle } \mathbf{u} \in \mathcal{U}.$$

- 1 Einleitung
- 2 Krylov-Unterraumverfahren
 - 2.1 Projektionen
 - 2.2 Orthogonale Projektionsverfahren
 - 2.3 Krylov-Unterräume
 - 2.4 Basen von Krylov-Räumen
- 3 Lineare Gleichungssysteme
- 4 Matrixfunktionen
- 5 Krylov-Verfahren für Matrixfunktionen
- 6 Das QR-Verfahren für Eigenwertaufgaben

Krylov-Unterraumverfahren

Krylov-Räume

Zu gegebener quadratischer Matrix $A \in \mathbb{C}^{n \times n}$ und Vektor $b \in \mathbb{C}^n$ wird die lineare Hülle der Vektoren

$$b, Ab, A^2b, \dots, A^{m-1}b, \quad m \in \mathbb{N},$$

als **Krylov-Unterraum** der Ordnung m bezeichnet, kurz $\mathcal{K}_m(A, b)$.

- Die Bezeichnung geht zurück auf den russischen Schiffbauingenieur **Alexei Nikolaevich Krylov*** (1863–1945).
- Krylov-Unterraumverfahren wurden unter die **zehn wichtigsten numerischen Algorithmen** des 20. Jahrhunderts eingestuft.



A.N. Krylov

*Nicht zu verwechseln mit den russischen Mathematikern Nikolai Mitrofanovich Krylov (1879–1955) bzw. Vladimir Ivanovich Krylov (1902–1994).

Entscheidende Beobachtung: für jeden Vektor $\mathbf{b} \in \mathbb{C}^n$ existiert $d \in \{0, 1, \dots, n\}$ sodass $\mathcal{K}_d(\mathbf{A}, \mathbf{b})$ invariant ist unter \mathbf{A} , d.h. $\mathbf{A}\mathcal{K}_d(\mathbf{A}, \mathbf{b}) \subset \mathcal{K}_d(\mathbf{A}, \mathbf{b})$.

Definition 2.17

Zu $\mathbf{A} \in \mathbb{C}^{n \times n}$ und $\mathbf{b} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ heißt die eindeutig bestimmte Zahl $d = d(\mathbf{A}, \mathbf{b})$ definiert durch

$$d := \min_{1 \leq m \leq n} \{ \mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^m \mathbf{b} \text{ linear abhängig} \} \quad (2.4)$$

der **Grad von \mathbf{A} bezüglich \mathbf{b}** .

Lemma 2.18

Ist $\mathbf{b} \in \mathbb{C}^n$ ein Vektor vom Grad $d \geq 1$ bezüglich $\mathbf{A} \in \mathbb{C}^{n \times n}$, so gelten

- (a) $\dim \mathcal{K}_m(\mathbf{A}, \mathbf{b}) = m$ für $m = 1, 2, \dots, d$.
- (b) $d - 1 \leq \dim \mathbf{A}\mathcal{K}_d(\mathbf{A}, \mathbf{b}) \leq d$ und, falls \mathbf{A} regulär, $\mathbf{A}\mathcal{K}_d(\mathbf{A}, \mathbf{b}) = \mathcal{K}_d(\mathbf{A}, \mathbf{b})$.
- (c) Ist $\mathbf{A}\mathcal{K}_d(\mathbf{A}, \mathbf{b}) = \mathcal{K}_d(\mathbf{A}, \mathbf{b})$, so liegt \mathbf{b} im Bildraum von \mathbf{A} .

- Eine äquivalente Charakterisierung des m -ten Krylov-Raums von \mathbf{A} mit Startvektor \mathbf{b} ist

$$\mathcal{K}_m(\mathbf{A}, \mathbf{b}) = \{p(\mathbf{A})\mathbf{b} : p \in \mathcal{P}_{m-1}\}.$$

- Nach dem oben Gesagten bilden die Krylov-Räume eine aufsteigende Kette

$$\{\mathbf{0}\} = \mathcal{K}_0(\mathbf{A}, \mathbf{b}) \subset \mathcal{K}_1(\mathbf{A}, \mathbf{b}) \subset \cdots \subset \mathcal{K}_{d-1}(\mathbf{A}, \mathbf{b}) \subset \mathcal{K}_d(\mathbf{A}, \mathbf{b}),$$

die beim invarianten Raum $\mathcal{K}_d(\mathbf{A}, \mathbf{b})$ abbricht. Man bezeichnet den Grad von \mathbf{A} bezüglich \mathbf{b} daher auch als **Abbruchindex** der Krylov-Folge..

Lemma 2.19 (Eigenschaften des Abbruchindex $d(\mathbf{A}, \mathbf{b})$)

Für den Abbruchindex $d(\mathbf{A}, \mathbf{b})$ gelten:

- (a) $\mathcal{K}_{m-1}(\mathbf{A}, \mathbf{b}) \subsetneq \mathcal{K}_m(\mathbf{A}, \mathbf{b})$ genau dann, wenn $m \leq d$.
- (b) $\mathcal{K}_m(\mathbf{A}, \mathbf{b}) = \mathcal{K}_{m+1}(\mathbf{A}, \mathbf{b})$ genau dann, wenn $m \geq d$.
- (c) $\mathcal{K}_m(\mathbf{A}, \mathbf{b})$ ist genau dann \mathbf{A} -invariant, d.h. $\mathbf{A}\mathcal{K}_m(\mathbf{A}, \mathbf{b}) \subseteq \mathcal{K}_m(\mathbf{A}, \mathbf{b})$, wenn $m \geq d$.
- (d) d ist der Grad des Minimalpolynoms $m_{\mathbf{A}, \mathbf{b}}$ von \mathbf{b} bez. \mathbf{A} ($m_{\mathbf{A}, \mathbf{b}}$ ist das monische Polynom kleinsten Grades, welches $m_{\mathbf{A}, \mathbf{b}}(\mathbf{A})\mathbf{b} = \mathbf{0}$ erfüllt.)

Äquivalente Charakterisierung des Grads von \mathbf{A} bezüglich \mathbf{b} :

$$d(\mathbf{A}, \mathbf{b}) = \min\{m \in \mathbb{N} : \mathcal{K}_m(\mathbf{A}, \mathbf{b}) = \mathcal{K}_{m+1}(\mathbf{A}, \mathbf{b})\}$$

Lemma 2.20 (Invarianzeigenschaften bei Krylov-Räumen)

Die Krylov-Räume $\mathcal{K}_m(A, \mathbf{b})$ besitzen folgende Invarianzeigenschaften:

- (a) $\mathcal{K}_m(\alpha \mathbf{A}, \beta \mathbf{b}) = \mathcal{K}_m(\mathbf{A}, \mathbf{b})$ für alle $\alpha, \beta \in \mathbb{C} \setminus \{0\}$.
- (b) $\mathcal{K}_m(\mathbf{A} - \sigma \mathbf{I}, \mathbf{b}) = \mathcal{K}_m(\mathbf{A}, \mathbf{b})$ für alle $\sigma \in \mathbb{C}$.
- (c) $\mathcal{K}_m(\mathbf{TAT}^{-1}, \mathbf{Tb}) = \mathbf{T} \mathcal{K}_m(\mathbf{A}, \mathbf{b})$ für alle invertierbaren $\mathbf{T} \in \mathbb{C}^{n \times n}$.

- 1 Einleitung
- 2 Krylov-Unterraumverfahren
 - 2.1 Projektionen
 - 2.2 Orthogonale Projektionsverfahren
 - 2.3 Krylov-Unterräume
 - 2.4 Basen von Krylov-Räumen
- 3 Lineare Gleichungssysteme
- 4 Matrixfunktionen
- 5 Krylov-Verfahren für Matrixfunktionen
- 6 Das QR-Verfahren für Eigenwertaufgaben

Basen von Krylov-Räumen

Die Monombasis

- Die Menge $\{\mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^{m-1}\mathbf{b}\}$ ist (falls $m \leq d$) die kanonische Basis des Krylov-Raums $\mathcal{K}_m(\mathbf{A}, \mathbf{b})$.
- Die Vektoren $\mathbf{A}^k\mathbf{b}$ streben aber für $k \rightarrow \infty$ (fast immer) gegen einen Eigenvektor von \mathbf{A} , der zum betragsgrößten Eigenwert von \mathbf{A} gehört.
- Die Vektoren $\mathbf{A}^k\mathbf{b}$ werden also zunehmend linear abhängig und die Basis $\{\mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^{m-1}\mathbf{b}\}$ ist für numerische Berechnungen daher ungeeignet.

Basen von Krylov-Räumen

Erzeugung einer Orthonormalbasis

- Wir werden mit Hilfe des Gram-Schmidt-Verfahrens eine **geschachtelte Orthonormalbasis** $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d\}$ von $\mathcal{K}_d(\mathbf{A}, \mathbf{b})$ konstruieren.
- „Geschachtelt“ bedeutet, dass $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ für jedes $m = 1, 2, \dots, d$ eine Basis von $\mathcal{K}_m(\mathbf{A}, \mathbf{b})$ bilden oder äquivalent, dass

$$\mathbf{v}_m \in \mathcal{K}_m(\mathbf{A}, \mathbf{b}) \setminus \mathcal{K}_{m-1}(\mathbf{A}, \mathbf{b}).$$

- Auch hier ist es aus numerischer Sicht nicht ratsam, im m -ten Schritt den Vektor $\mathbf{A}^{m-1}\mathbf{b}$ gegen die bereits berechneten Basiselemente $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{m-1}$ zu orthogonalisieren.
- Wegen $\mathcal{K}_m(\mathbf{A}, \mathbf{b}) = \mathcal{K}_{m-1}(\mathbf{A}, \mathbf{b}) + \text{span}\{\mathbf{A}\mathbf{v}_{m-1}\}$ können wir stattdessen auch $\mathbf{A}\mathbf{v}_{m-1}$ gegen $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{m-1}$ orthogonalisieren.
- Analog zum Gram-Schmidt-Verfahren: berechne Orthogonalprojektion \mathbf{p} von $\mathbf{A}\mathbf{v}_{m-1}$ auf $\mathcal{K}_{m-1}(\mathbf{A}, \mathbf{b}) = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{m-1}\}$, bilde Differenz $\mathbf{A}\mathbf{v}_{m-1} - \mathbf{p} \perp \mathcal{K}_{m-1}(\mathbf{A}, \mathbf{b})$, und normiere diesen Vektor abschließend auf Länge 1.

Kompakte Darstellung:

$$\mathbf{v}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|}, \quad \mathbf{v}_m = \frac{(I - P_{\mathcal{K}_{m-1}(\mathbf{A}, \mathbf{b})}) \mathbf{A} \mathbf{v}_{m-1}}{\|(I - P_{\mathcal{K}_{m-1}(\mathbf{A}, \mathbf{b})}) \mathbf{A} \mathbf{v}_{m-1}\|} \quad m = 2, 3, \dots, d.$$

Dieser Algorithmus wird **Arnoldi-Prozess** [Arnoldi, 1951] genannt (vgl. Algorithmus 1).

Beachte:

- $h_{k+1,k} > 0$ für $k = 1, 2, \dots, d-1$ und
- $h_{d+1,d} = 0$, denn $h_{k+1,k} = 0$ gilt genau dann, wenn

$$\mathbf{A} \mathbf{v}_k \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \mathcal{K}_k(\mathbf{A}, \mathbf{b}),$$

also genau dann, wenn $k = d$ erfüllt ist.

Algorithmus 1: Klassisches Arnoldi-Verfahren.

Gegeben: $A \in \mathbb{C}^{n \times n}$, $b \in \mathbb{C}^n$, $b \neq 0$

```
1  $\beta \leftarrow \|b\|$ ,  $v_1 \leftarrow b/\beta$ .
2 for  $k = 1, 2, \dots$  do
3    $w \leftarrow Av_k$ 
4   for  $j = 1, 2, \dots, k$  do
5      $h_{j,k} \leftarrow (w, v_j)$ 
6    $w \leftarrow w - \sum_{j=1}^k h_{j,k} v_j$ 
7    $h_{k+1,k} \leftarrow \|w\|$ 
8   if  $h_{k+1,k} = 0$  then
9      $d \leftarrow k$ 
10    break
11  else
12     $v_{k+1} \leftarrow w/h_{k+1,k}$ 
```

- In der Praxis wird aus Stabilitätsgründen oft das **modifizierte Arnoldi-Verfahren** verwandt (vgl. Algorithmus 2).
- Klassisches und modifiziertes Arnoldi-Verfahren liefern in exakter Arithmetik dieselben Resultate.
- Beim klassischen Arnoldi-Verfahren wird \mathbf{v}_{m+1} im Wesentlichen (d.h. bis auf Normierung) durch eine Projektion $P_{\mathcal{W}_m} \mathbf{A} \mathbf{v}_m$ bestimmt. Dabei ist $P_{\mathcal{W}_m} = I - P_{\mathcal{K}_m(\mathbf{A}, \mathbf{b})}$, was $\mathcal{W}_m = \mathcal{K}_m(\mathbf{A}, \mathbf{b})^\perp$ bedeutet.
- Nun ist mit $\mathcal{V}_j = \{\mathbf{v}_j\}^\perp$ aber

$$\mathcal{W}_m = \mathcal{K}_m(\mathbf{A}, \mathbf{b})^\perp = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}^\perp = \mathcal{V}_1 \cap \mathcal{V}_2 \cap \dots \cap \mathcal{V}_m$$

und es ist leicht zu sehen (vgl. Übungsaufgabe), dass

$$P_{\mathcal{W}_m} = P_{\mathcal{V}_m} \cdots P_{\mathcal{V}_2} P_{\mathcal{V}_1}$$

gilt.

- Beim modifizierten Arnoldi-Verfahren wird (wieder bis auf Normierung) \mathbf{v}_{m+1} durch $P_{\mathcal{V}_m} \cdots P_{\mathcal{V}_2} P_{\mathcal{V}_1} \mathbf{A} \mathbf{v}_m$ berechnet.

Algorithmus 2: Modifiziertes Arnoldi-Verfahren.

Gegeben: $A \in \mathbb{C}^{n \times n}$, $b \in \mathbb{C}^n$, $b \neq 0$

```
1  $\beta \leftarrow \|b\|$ ,  $v_1 \leftarrow b/\beta$ .
2 for  $k = 1, 2, \dots$  do
3    $w \leftarrow Av_k$ 
4   for  $j = 1, 2, \dots, k$  do
5      $h_{j,k} \leftarrow (w, v_j)$ 
6      $w \leftarrow w - h_{j,k}v_j$ 
7    $h_{k+1,k} \leftarrow \|w\|$ 
8   if  $h_{k+1,k} = 0$  then
9      $d \leftarrow k$ 
10    break
11  else
12     $v_{k+1} \leftarrow w/h_{k+1,k}$ 
```

Krylov-Unterraumverfahren

Basen von Krylov-Räumen

- Rechnet man in Gleitpunktarithmetik, so wird auch beim modifizierten Arnoldi-Verfahren die Orthogonalität der Vektoren v_m mit zunehmendem m allmählich verloren gehen.
- Die Genauigkeit lässt sich durch **Reorthogonalisierung** verbessern (vgl. Algorithmus 3), was allerdings mit erheblichem Mehraufwand verbunden ist.
- Man kann die Reorthogonalisierung mehrmals durchführen, sollte sich dabei aber an die Kahan'sche Devise „twice is enough“ halten [B. N. Parlett, 1997, S. 115)]. Typischerweise ist ein Reorthogonalisierungsschritt ausreichend.

Krylov-Unterraumverfahren

Basen von Krylov-Räumen

Algorithmus 3: Modifiziertes Arnoldi-Verfahren mit Reorthogonalisierung.

Gegeben: $A \in \mathbb{C}^{n \times n}$, $b \in \mathbb{C}^n$, $b \neq 0$

```
1  $\beta \leftarrow \|b\|$ ,  $v_1 \leftarrow b/\beta$ .
2 for  $k = 1, 2, \dots$  do
3    $w \leftarrow Av_k$ 
4   for  $j = 1, 2, \dots, k$  do
5      $h_{j,k} \leftarrow (w, v_j)$ 
6      $w \leftarrow w - h_{j,k}v_j$ 
7   for  $i = 1, 2, \dots, \ell$  do
8     for  $j = 1, 2, \dots, k$  do
9        $\tilde{h}_{j,k} \leftarrow (w, v_j)$ 
10       $h_{j,k} \leftarrow h_{j,k} + \tilde{h}_{j,k}$ 
11       $w \leftarrow w - \tilde{h}_{j,k}v_j$ 
12    $h_{k+1,k} \leftarrow \|w\|$ 
13   if  $h_{k+1,k} = 0$  then
14      $d \leftarrow k$ , break
15   else
16      $v_{k+1} \leftarrow w/h_{k+1,k}$ 
```

Beispiel 2.21

Sei $\mathbf{A} \in \mathbb{R}^{183 \times 183}$ die Matrix FS_183_6 aus der Harwell-Boeing-Sammlung, eine nichtsymmetrische Matrix mit

$$\|\mathbf{A}\|_2 = 1.2 \cdot 10^9 \quad \text{und} \quad \text{cond}_2(\mathbf{A}) = 1.7 \cdot 10^{11}.$$

Wir wählen $\mathbf{b} = [1, 1, \dots, 1]^T$ und berechnen mit verschiedenen Varianten des Arnoldi-Verfahrens Orthonormalbasen

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} \quad \text{von} \quad \mathcal{K}_k(\mathbf{A}, \mathbf{b}), \quad k = 1, 2, \dots, 183.$$

Als Maß für die Orthonormalität dieser Basen bestimmen wir

$$\|\mathbf{I}_k - \mathbf{V}_k^T \mathbf{V}_k\|_2, \quad \text{wobei} \quad \mathbf{V}_k = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_k] \in \mathbb{R}^{183 \times k}.$$

Die Ergebnisse sind in Abbildung 3 dargestellt. □

Krylov-Unterraumverfahren

Basen von Krylov-Räumen

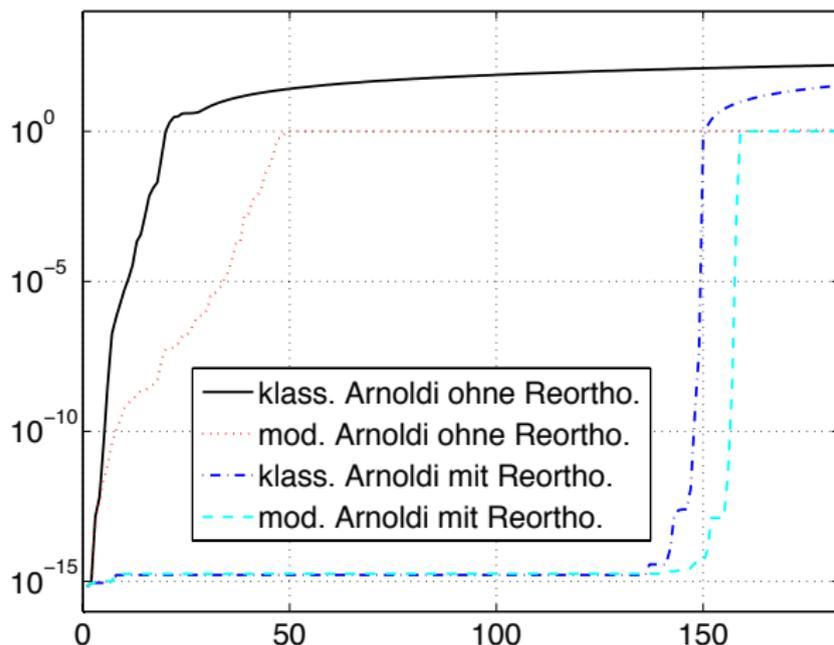


Abbildung 3: Verlust an Orthogonalität bei verschiedenen Implementierungen des Arnoldi-Verfahrens im Beispiel 2.21: Die dargestellten Kurven zeigen $\|I_k - V_k^T V_k\|_2$ gegen k für $k = 1, 2, \dots, 183 = \dim(\mathbf{A})$.

- Entscheidene Beziehung im Arnoldi-Verfahren:

$$\mathbf{A}\mathbf{v}_k = h_{1,k}\mathbf{v}_1 + h_{2,k}\mathbf{v}_2 + \cdots + h_{k,k}\mathbf{v}_k + h_{k+1,k}\mathbf{v}_{k+1}, \quad k = 1, \dots, d. \quad (2.5)$$

Für $k = d$ ist $h_{d+1,d} = 0$, also $\mathbf{A}\mathbf{v}_d = h_{1,d}\mathbf{v}_1 + h_{2,d}\mathbf{v}_2 + \cdots + h_{d,d}\mathbf{v}_d$.

- Schreibe (2.5) für $k = 1, 2, \dots, m$ in die Spalten einer $n \times m$ -Matrix:

$$\mathbf{A}\mathbf{V}_m = \mathbf{V}_{m+1}\mathbf{H}_{m+} = \mathbf{V}_m\mathbf{H}_m + h_{m+1,m}\mathbf{v}_{m+1}\mathbf{e}_m^\top \quad (2.6)$$

(für $m = d$ gilt $\mathbf{A}\mathbf{V}_d = \mathbf{V}_d\mathbf{H}_d$) mit $\mathbf{V}_m = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_m] \in \mathbb{C}^{n \times m}$ ($\mathbf{V}_m^\mathbf{H}\mathbf{V}_m = \mathbf{I}_m$) und einer oberen **Hessenberg-Matrix**

$$\mathbf{H}_{m+} = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,m} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,m} \\ & h_{3,2} & \cdots & h_{3,m} \\ & & \ddots & \vdots \\ & & & h_{m+1,m} \end{bmatrix} \in \mathbb{C}^{(m+1) \times m}.$$

$\mathbf{H}_m = [\mathbf{I}_m \quad \mathbf{0}] \mathbf{H}_{m+} \in \mathbb{C}^{m \times m}$ ist die quadratische obere Hessenberg-Matrix, die entsteht, wenn man die letzte, also die $(m+1)$ -te Zeile von \mathbf{H}_{m+} streicht, $\mathbf{e}_m \in \mathbb{R}^m$ ist der m -te Einheitsvektor im \mathbb{R}^m .

Lemma 2.22 (Eigenschaften der Hessenberg-Matrix \mathbf{H}_m)

Für die Einträge $h_{j,k}$ der Hessenberg-Matrix \mathbf{H}_m aus (2.6) gilt

$$h_{j,k} = (\mathbf{A}\mathbf{v}_k, \mathbf{v}_j), \quad 1 \leq j, k \leq m.$$

Die Nebendiagonalelemente $h_{k+1,k}$ sind positiv für $k < d$. Der Rang von \mathbf{H}_m beträgt mindestens $m - 1$ und es gilt $\mathcal{R}(\mathbf{H}_m) + \text{span}\{\mathbf{e}_1\} = \mathbb{C}^m$.

- Man nennt (2.6) eine **Arnoldi-Zerlegung** von \mathbf{A} . Sie wird im Folgenden eine wichtige Rolle spielen. Insbesondere sind die Krylov-Räume

$$\mathcal{K}_m(\mathbf{A}, \mathbf{v}) \subseteq \mathbb{C}^n \quad \text{und} \quad \mathcal{K}_m(\mathbf{H}_m, \mathbf{e}_1) \subseteq \mathbb{C}^m$$

isomorph ($\mathbf{e}_1 \in \mathbb{R}^m$ bezeichnet den ersten Einheitsvektor).

- Ist $m \ll n$, so erfordern Operationen im zweiten Raum natürlich deutlich weniger Aufwand.

Lemma 2.23

Für jedes Polynom $p(\lambda) = \alpha_0 + \alpha_1\lambda + \dots + \alpha_m\lambda^m \in \mathcal{P}_m$ gilt

$$p(\mathbf{A})\mathbf{b} = \beta \mathbf{V}_m p(\mathbf{H}_m) \mathbf{e}_1 + \alpha_m \beta \prod_{j=1}^m h_{j+1,j} \mathbf{v}_{m+1}$$

mit $\beta = \|\mathbf{b}\|$ und den Größen \mathbf{V}_m , \mathbf{v}_{m+1} , \mathbf{H}_m und $h_{j+1,j}$ aus der Arnoldi-Zerlegung (2.6). Insbesondere ist

$$p(\mathbf{A})\mathbf{b} = \beta \mathbf{V}_m p(\mathbf{H}_m) \mathbf{e}_1 \quad \text{für alle } p \in \mathcal{P}_{m-1}.$$

Lemma 2.23

Für jedes Polynom $p(\lambda) = \alpha_0 + \alpha_1\lambda + \dots + \alpha_m\lambda^m \in \mathcal{P}_m$ gilt

$$p(\mathbf{A})\mathbf{b} = \beta \mathbf{V}_m p(\mathbf{H}_m) \mathbf{e}_1 + \alpha_m \beta \prod_{j=1}^m h_{j+1,j} \mathbf{v}_{m+1}$$

mit $\beta = \|\mathbf{b}\|$ und den Größen \mathbf{V}_m , \mathbf{v}_{m+1} , \mathbf{H}_m und $h_{j+1,j}$ aus der Arnoldi-Zerlegung (2.6). Insbesondere ist

$$p(\mathbf{A})\mathbf{b} = \beta \mathbf{V}_m p(\mathbf{H}_m) \mathbf{e}_1 \quad \text{für alle } p \in \mathcal{P}_{m-1}.$$

Zur Berechnung der Arnoldi-Zerlegung (2.6) sind erforderlich

- m Multiplikationen mit \mathbf{A} ,
- $\frac{1}{2}m(m+1) = O(m^2)$ Innenprodukte
- und ebensoviele Vektoraufdatierungen der Form $\mathbf{x} \leftarrow \mathbf{x} + \alpha\mathbf{y}$.

Außerdem muss man die $m+1$ Basisvektoren $\{\mathbf{v}_m\}_{j=1}^{m+1}$ speichern, ein nicht unerheblicher Aufwand.

Krylov-Unterraumverfahren

Arnoldi-Zerlegung, selbstadjungierter Fall

Für spezielle Matrizen A reduzieren sich Arbeits- und Speicheraufwand beträchtlich:

- Lemma 2.22: $h_{j,m} = (Av_m, v_j)$ für alle j und m .
- Ist A selbstadjungiert, d.h. $(Av, w) = (v, Aw)$ für alle v und w ⁴, so folgt

$$h_{j,m} = (Av_m, v_j) = (v_m, Av_j) = \overline{(Av_j, v_m)} = \bar{h}_{m,j}, \quad (2.7)$$

also ist H_m Hermitesch, und somit tridiagonal.

- Da $h_{m+1,m} > 0$ (vgl. Lemma 2.22) besitzt H_m nur positive Nebendiagonalelemente.
- Die Hauptdiagonalelemente $h_{m,m} = (Av_m, v_m)$ liegen im Wertebereich der selbstadjungierten Matrix A , sind also auch reell (dass sie positiv sind, kann nur garantiert werden, wenn A zusätzlich positiv definit ist).

⁴Ist (\cdot, \cdot) das Euklidische Innenprodukt, so ist A genau dann Hermitesch.

Krylov-Unterraumverfahren

Arnoldi-Zerlegung, selbstadjungierter Fall

Fassen wir zusammen:

Ist A selbstadjungiert, so ist die Hessenberg-Matrix H_m aus (2.6) eine reelle symmetrische Tridiagonalmatrix. Wir werden – dem allgemeinen Usus folgend – die Matrizen H_{m+} und H_m dann mit T_{m+} bzw. T_m bezeichnen:

$$T_{m+} = \begin{bmatrix} \gamma_1 & \delta_2 & 0 & \dots & 0 \\ \delta_2 & \gamma_2 & \delta_3 & \ddots & \vdots \\ 0 & \delta_3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \delta_m \\ 0 & \dots & 0 & \delta_m & \gamma_m \\ \hline 0 & \dots & & 0 & \delta_{m+1} \end{bmatrix} \in \mathbb{R}^{(m+1) \times m} \quad (2.8)$$

($m = 1, 2, \dots, d$, $\delta_{d+1} = 0$).

Größen, die ohnehin verschwinden, müssen nicht berechnet werden. Daher vereinfacht sich der Arnoldi-Prozess zum **Hermiteischen** oder **symmetrischen Lanczos-Prozess**:

Algorithmus 4: Hermiteischer Lanczos-Prozess

Gegeben: Innenprodukt (\cdot, \cdot) auf \mathbb{C}^n mit induzierter Norm $\|\cdot\|$, $\mathbf{A} \in \mathbb{C}^{n \times n}$ selbstadjungiert bez. (\cdot, \cdot) , $\mathbf{b} \in \mathbb{C}^n$, $\mathbf{b} \neq \mathbf{0}$.

- 1 $\delta_1 \leftarrow \|\mathbf{b}\|$, $\mathbf{v}_1 \leftarrow \mathbf{b}/\delta_1$, $\mathbf{v}_0 \leftarrow \mathbf{0}$
 - 2 **for** $k = 1, 2, \dots, d - 1$ **do**
 - 3 $\mathbf{w} \leftarrow \mathbf{A}\mathbf{v}_k$
 - 4 $\gamma_k \leftarrow (\mathbf{w}, \mathbf{v}_k)$
 - 5 $\mathbf{w} \leftarrow \mathbf{w} - \gamma_k \mathbf{v}_k - \delta_k \mathbf{v}_{k-1}$
 - 6 $\delta_{k+1} \leftarrow \|\mathbf{w}\|$
 - 7 $\mathbf{v}_{k+1} \leftarrow \mathbf{w}/\delta_{k+1}$
-

Praktische Konsequenz der Tridiagonalform von \mathbf{T}_m :

- Orthogonalisierung von $\mathbf{A}\mathbf{v}_k$ gegen $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ erfordert nur Operationen mit \mathbf{v}_k und \mathbf{v}_{k-1} .
- Die übrigen Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-2}$ sind nicht beteiligt und müssen daher auch nicht gespeichert werden.
- Der Hermiteische Lanczos-Prozess erzeugt also eine geschachtelte Orthonormalbasis von $\mathcal{K}_d(\mathbf{A}, \mathbf{b})$ mit konstantem Speicher- und Arbeitsaufwand pro Schritt.

Allerdings: Beim Lanczos-Verfahren geht die Orthogonalität noch schneller verloren als beim Arnoldi-Verfahren, was nicht überrascht, weil hier $\mathbf{A}\mathbf{v}_k$ explizit nur gegen \mathbf{v}_k und \mathbf{v}_{k-1} orthogonalisiert wird.

Beispiel 2.24

Betrachte $\mathbf{A} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$ mit

$$\lambda_j = \lambda_1 + \frac{j-1}{n-1}(\lambda_n - \lambda_1)\rho^{n-j}, \quad j = 1, 2, \dots, n.$$

$\rho \in (0, 1]$ bestimmt die Verteilung der Eigenwerte von \mathbf{A} im Spektralintervall $[\lambda_1, \lambda_n]$ (für $\rho = 1$ sind die Eigenwerte äquidistant; je kleiner ρ ist, umso mehr konzentrieren sie sich am linken Rand λ_1). Diese Matrizenfamilie dient häufig zur Illustration von Effekten rundungsbehafteter Rechnung auf das Lanczos-Verfahren (z.B. in [Strakoš, 1991] oder in [Meurant und Strakoš, 2006]).

In unserer Rechnung: $n = 30$, $\rho = 0.8$, $\lambda_1 = 0.1$, $\lambda_n = 100$ sowie $\mathbf{b} = [1, 1, \dots, 1]^T$.

Wir sehen in Abbildung 4 die Norm $\|\mathbf{I}_j - \mathbf{V}_j^T \mathbf{V}_j\|_2$ ($= 0$ in exakter Arithmetik) als Funktion von j für verschiedene Orthogonalisierungsverfahren. Insbesondere der Lanczos-Prozess erzeugt Vektoren, die alles andere als orthogonal sind. Das impliziert aber auch, dass $\mathbf{V}_j^T \mathbf{A} \mathbf{V}_j$ keine Tridiagonalmatrix ist und schon gar nicht mit \mathbf{T}_j übereinstimmt: Auf der rechten Seite ist $\|\mathbf{T}_j - \mathbf{V}_j^T \mathbf{A} \mathbf{V}_j\|_2$ (theoretisch ebenfalls $= 0$) geplottet. Abbildung 5 zeigt, dass die berechnete Matrix $\mathbf{V}_{30}^T \mathbf{A} \mathbf{V}_{30}$ keineswegs tridiagonal ist.

Krylov-Unterraumverfahren

Hermiteischer Lanczos-Prozess

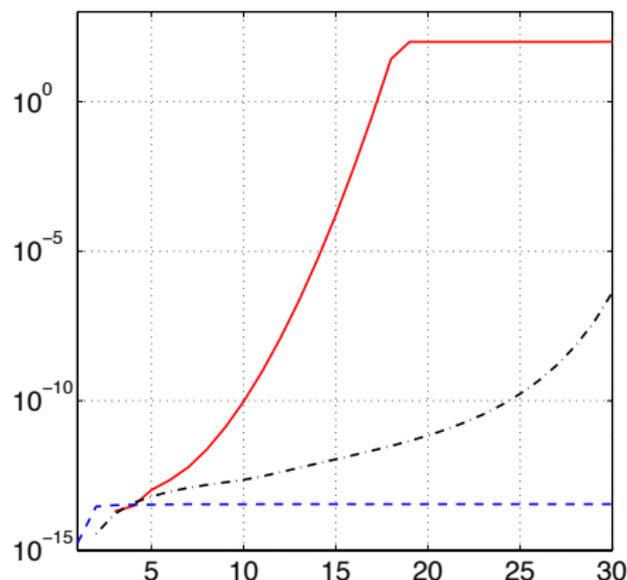
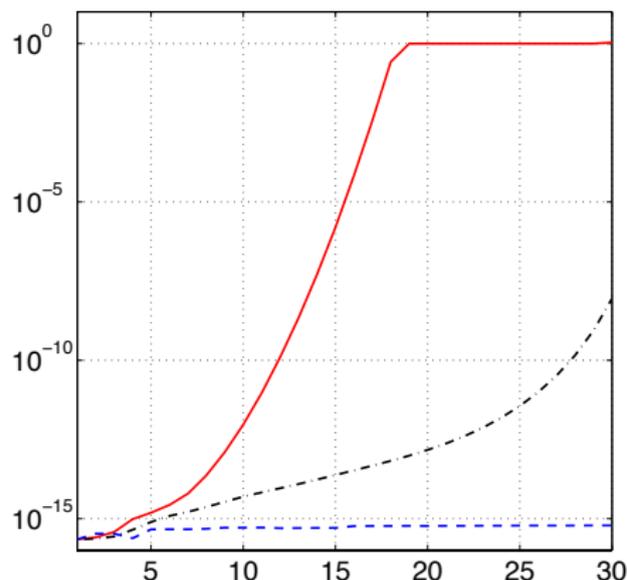


Abbildung 4: Orthogonalitätsverlust $\|I_j - V_j^T V_j\|_2$ (links) und Verlust der Projektionseigenschaft $\|T_j - V_j^T A V_j\|_2$ (rechts) beim Lanczos- (—) und modifizierten Arnoldi-Verfahren (— ·) sowie beim modifizierten Arnoldi-Verfahren mit einem Reorthogonalisierungsschritt (— —), vgl. Beispiel 2.24.

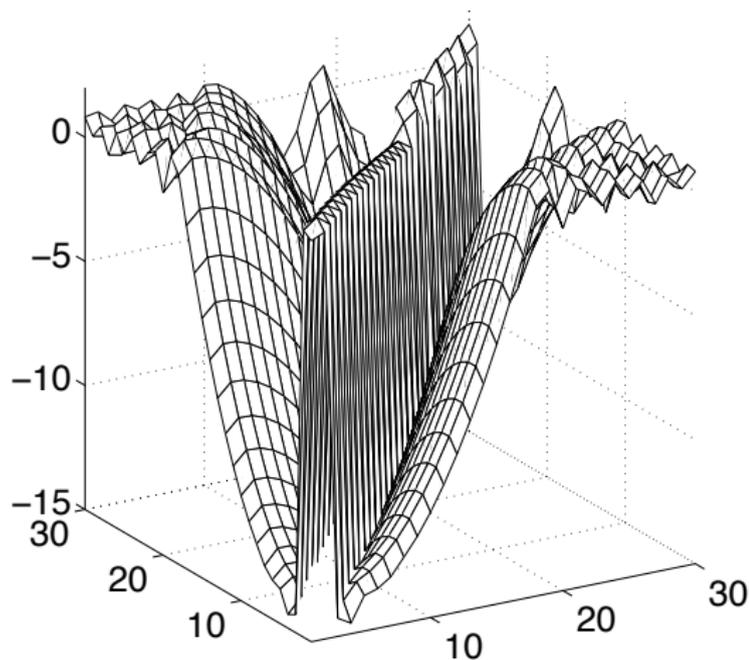


Abbildung 5: $\log_{10} |\mathbf{V}_{30}^T \mathbf{A} \mathbf{V}_{30}|$, vgl. Beispiel 2.24.

- Der Lanczos–Prozess verliert durch Rundungsfehlereinfluss also alle seine theoretischen Eigenschaften.
- Zur Berechnung orthonormaler Basen von Krylov–Räumen ist er sicher ungeeignet. Man hüte sich aber vor dem Kurzschluss, dass er für numerische Zwecke unbrauchbar ist.
- Überraschenderweise ist er ein zuverlässiger und effizienter Algorithmus zur Lösung linearer Gleichungssysteme, zur Berechnung von Matrixfunktionen und zur Berechnung gewisser Eigenwerte bei großen schwach besetzten Matrizen.

- Entscheidende Gleichung beim Hermiteschen Lanczos-Prozesses: dreistufige Rekursion⁵ (vgl. dazu (2.5))

$$\delta_{m+1} \mathbf{v}_{m+1} = (\mathbf{A} - \gamma_m \mathbf{I}_n) \mathbf{v}_m - \delta_m \mathbf{v}_{m-1}. \quad (2.9)$$

- Naheliegende Frage: gibt es neben selbstadjungierten noch andere Matrizen, die zu kurzen Rekursionen für die orthonormalen Basisvektoren der Krylov-Räume führen?
- Entscheidender Punkt bei selbstadjungierten Matrizen \mathbf{A} : für $j < m - 1$ gilt

$$h_{j,m} = (\mathbf{A} \mathbf{v}_m, \mathbf{v}_j) = (\mathbf{v}_m, \mathbf{A} \mathbf{v}_j) = 0$$

weil $\mathbf{v}_m \perp \mathcal{K}_{j+1}(\mathbf{A}, \mathbf{b}) \ni \mathbf{A} \mathbf{v}_j$. Dies lässt sich leicht verallgemeinern:

⁵Diese erinnert – wie wir noch sehen werden – nicht ohne Grund an die dreistufige Rekursionsformel, der reelle Orthogonalpolynome genügen.

- Gilt mit einem Polynom p vom Grad s

$$(\mathbf{A}\mathbf{v}, \mathbf{w}) = (\mathbf{v}, p(\mathbf{A})\mathbf{w}) \quad \text{für alle } \mathbf{v}, \mathbf{w} \in \mathbb{C}^n \quad (2.10)$$

d.h. ist \mathbf{A}^* , die Adjungierte von \mathbf{A} bez. (\cdot, \cdot) , durch $p(\mathbf{A})$ gegeben, dann folgt für $j < m - s$

$$h_{j,m} = (\mathbf{A}\mathbf{v}_m, \mathbf{v}_j) = (\mathbf{v}_m, \mathbf{A}^*\mathbf{v}_j) = (\mathbf{v}_m, p(\mathbf{A})\mathbf{v}_j) = 0,$$

weil $\mathbf{v}_m \perp \mathcal{K}_{j+s}(\mathbf{A}, \mathbf{b}) \ni p(\mathbf{A})\mathbf{v}_j$.

- Das bedeutet, dass \mathbf{H}_m **Bandstruktur** besitzt (nur die s oberen Nebendiagonalen enthalten von Null verschiedene Elemente) und führt sofort zu einer $(s + 2)$ -stufigen Rekursion für die Elemente der Krylov-Orthonormalbasis

$$h_{m+1,m}\mathbf{v}_{m+1} = \mathbf{A}\mathbf{v}_m - h_{m,m}\mathbf{v}_m - h_{m-1,m}\mathbf{v}_{m-1} - \dots - h_{m-s,m}\mathbf{v}_{m-s}. \quad (2.11)$$

- Wann die Beziehung (2.10) erfüllt ist, wann also

$$\mathbf{A}^* = p(\mathbf{A}) \quad \text{für ein } p \in \mathcal{P}_s$$

gilt, soll jetzt untersucht werden.

- Zunächst ist klar, dass \mathbf{A} dann notwendig normal sein muss, denn \mathbf{A} und $p(\mathbf{A})$ kommutieren.
- Umgekehrt kann man zeigen, dass es zu jeder normalen Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ stets ein Polynom $p \in \mathcal{P}_{n-1}$ gibt mit $\mathbf{A}^* = p(\mathbf{A})$ (Übungsaufgabe).
- Eine normale Matrix \mathbf{A} heißt **s-normal**, wenn s der minimale Grad eines Polynoms $p \neq 0$ ist, für das $\mathbf{A}^* = p(\mathbf{A})$ gilt.
- Ist \mathbf{A} s -normal und gilt $\mathbf{A}^* = p(\mathbf{A})$ für $p \in \mathcal{P}_s$, so folgt mit der (immer geltenden) Beziehung $\bar{p}(\mathbf{A}^*) = \mathbf{A}$ sofort $\bar{p}(p(\mathbf{A})) = \mathbf{A}$ und somit

$$q(\mathbf{A}) = \mathbf{O} \quad \text{mit dem Polynom } q(\lambda) = \bar{p}(p(\lambda)) - \lambda \in \mathcal{P}_{s^2}, \quad p \neq 0.$$

- Ist $s \geq 2$, so besitzt q den genauen Grad s^2 , der Grad des Minimalpolynoms von \mathbf{A} ist also höchstens s^2 . Insbesondere hat \mathbf{A} höchstens s^2 verschiedene Eigenwerte.
- [Khavinson und Świątek, 2003] konnten diese Schranke erheblich verschärfen: Ist p ein Polynom vom Grad $s \geq 2$, so hat die Gleichung $\bar{\lambda} = p(\lambda)$ höchstens $3s - 2$ Lösungen.
- Für uns heißt das, dass eine s -normale Matrix ($s \geq 2$) höchstens $3s - 2$ Eigenwerte besitzt. Damit ist der Fall $s \geq 2$ (aber d „klein“) für praktische Anwendungen uninteressant. (Nur wenige Matrizen erfüllen dies.)
- Es bleibt der Fall $s = 1$. Neben selbstadjungierten ($p(\lambda) = \lambda$) gehören auch **anti-selbstadjungierte Matrizen** ($(\mathbf{A}\mathbf{v}, \mathbf{w}) = -(\mathbf{v}, \mathbf{A}\mathbf{w})$ für alle \mathbf{v} und \mathbf{w} , also $p(\lambda) = -\lambda$) zu dieser Klasse der 1-normalen Matrizen. Außerdem ist eine Matrix 1-normal, wenn sie die Form $\mathbf{A} = \exp(i\theta)(\alpha\mathbf{I}_n + \mathbf{B})$ besitzt mit $\theta, \alpha \in \mathbb{R}$ und $\mathbf{B} \in \mathbb{R}^{n \times n}$ anti-selbstadjungiert.
- Man kann außerdem zeigen, dass damit alle Matrizen \mathbf{A} erfasst sind, die eine Gleichung $\mathbf{A}^* = p(\mathbf{A})$ mit $p \in \mathcal{P}_1$ erfüllen (vgl. Übungsaufgabe).

- [Faber und Manteuffel, 1984] bewiesen, dass die Basisvektoren \mathbf{v}_m genau dann für alle \mathbf{b} einer $(s + 2)$ -stufige Rekursion der Form (2.11) genügen, wenn \mathbf{A} s -normal ist (wir haben hier nur eine Richtung gezeigt).
- Da der Fall $s \geq 2$ keine praktische Bedeutung besitzt, halten wir fest, dass kurze Rekursionen i.W. nur für selbstadjungierte und anti-selbstadjungierte Matrizen existieren.
- Der Vollständigkeit halber sei erwähnt, dass es auch kurze Rekursionen anderer Struktur gibt:

$$h_{m+1,m} \mathbf{v}_{m+1} = \sum_{j=0}^{s_1} g_{m-j,m} \mathbf{A} \mathbf{v}_{m-j} - \sum_{j=0}^{s_2} h_{m-j,m} \mathbf{v}_{m-j}.$$

Prototyp ist der isometrische Arnoldi-Prozess (siehe [Gragg, 1993]) für isometrische Matrizen \mathbf{A} , d.h. $(\mathbf{A}\mathbf{v}, \mathbf{A}\mathbf{w}) = (\mathbf{v}, \mathbf{w})$ für alle \mathbf{v} und \mathbf{w} , die – wenn (\cdot, \cdot) das Euklidische Innenprodukt ist – genau die unitären sind – auch wenn er aus Stabilitätsgründen anders implementiert werden muss. Er spielt hier keine große Rolle, weil es unsinnig ist, Gleichungssysteme mit isometrischer Koeffizientenmatrix iterativ zu lösen.

- Von [Barth und Manteuffel, 2000] stammt das Resultat, dass eine Rekursion dieser Bauart für die Basisvektoren \mathbf{v}_m gilt, wenn es Polynome $q_k \in \mathcal{P}_k$ und $p_\ell \in \mathcal{P}_\ell$ gibt mit

$$\mathbf{A}^* q_k(\mathbf{A}) - p_\ell(\mathbf{A}) = \mathbf{B},$$

wobei \mathbf{B} eine Matrix niedrigen Rangs, sagen wir r , ist ($s_1 = k + r$, $s_2 = \ell + r$).

- Beispiele für solche Matrizen sind
 - isometrische Matrizen:

$$q_k(\lambda) = \lambda, \quad p_\ell(\lambda) = 1,$$

- die Klasse

$$\rho \mathbf{U} + \alpha \mathbf{I}, \quad \mathbf{U} \text{ isometrisch}, \quad \alpha, \rho \in \mathbb{C},$$

siehe [Jagels und Reichel, 1994]) sowie

- Störungen niedrigen Rangs selbstadjungierter (oder anti-selbstadjungierter) Matrizen (siehe [Beckermann und Reichel, 2008]).