

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

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

$$\mathbf{Ax} = \mathbf{b}$$

lineare Gleichungssysteme

$$\|\mathbf{b} - \mathbf{Ax}\| \rightarrow \min$$

Ausgleichsrechnung

$$\mathbf{Ax} = \lambda \mathbf{x}$$

Eigenwertprobleme

$$Ax = b$$

lineare Gleichungssysteme

$$\|b - Ax\| \rightarrow \min$$

Ausgleichsrechnung

$$Ax = \lambda x$$

Eigenwertprobleme

Sind diese Probleme nicht längst gelöst?

$$Ax = b$$

lineare Gleichungssysteme

$$\|b - Ax\| \rightarrow \min$$

Ausgleichsrechnung

$$Ax = \lambda x$$

Eigenwertprobleme

Sind diese Probleme nicht längst gelöst?

- Grundbausteine des wissenschaftlichen Rechnens

$$Ax = b$$

lineare Gleichungssysteme

$$\|b - Ax\| \rightarrow \min$$

Ausgleichsrechnung

$$Ax = \lambda x$$

Eigenwertprobleme

Sind diese Probleme nicht längst gelöst?

- Grundbausteine des wissenschaftlichen Rechnens
- schnellere Algorithmen, neue Rechnerarchitekturen

$$Ax = b$$

lineare Gleichungssysteme

$$\|b - Ax\| \rightarrow \min$$

Ausgleichsrechnung

$$Ax = \lambda x$$

Eigenwertprobleme

Sind diese Probleme nicht längst gelöst?

- Grundbausteine des wissenschaftlichen Rechnens
- schnellere Algorithmen, neue Rechnerarchitekturen
- Spezialfälle: Strukturen ausnutzen
(dünn besetzt, symmetrisch, positiv-definit, symplektisch, Toeplitz, etc.)

$$Ax = b$$

lineare Gleichungssysteme

$$\|b - Ax\| \rightarrow \min$$

Ausgleichsrechnung

$$Ax = \lambda x$$

Eigenwertprobleme

Sind diese Probleme nicht längst gelöst?

- Grundbausteine des wissenschaftlichen Rechnens
- schnellere Algorithmen, neue Rechnerarchitekturen
- Spezialfälle: Strukturen ausnutzen
(dünn besetzt, symmetrisch, positiv-definit, symplektisch, Toeplitz, etc.)
- nichtlineare Verallgemeinerungen, e.g.

$$\|b - f(x)\| \rightarrow \min,$$

$$f : \mathbb{C}^N \rightarrow \mathbb{C}^N,$$

$$F(\lambda, \alpha)x = 0,$$

$$F : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}^{N \times N}.$$

$$Ax = b$$

lineare Gleichungssysteme

$$\|b - Ax\| \rightarrow \min$$

Ausgleichsrechnung

$$Ax = \lambda x$$

Eigenwertprobleme

Sind diese Probleme nicht längst gelöst?

- Grundbausteine des wissenschaftlichen Rechnens
- schnellere Algorithmen, neue Rechnerarchitekturen
- Spezialfälle: Strukturen ausnutzen
(dünn besetzt, symmetrisch, positiv-definit, symplektisch, Toeplitz, etc.)
- nichtlineare Verallgemeinerungen, e.g.

$$\|b - f(x)\| \rightarrow \min,$$

$$f : \mathbb{C}^N \rightarrow \mathbb{C}^N,$$

$$F(\lambda, \alpha)x = 0,$$

$$F : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}^{N \times N}.$$

- Matrixfunktionen, Modellreduktion.

① Einleitung

1.1 Lineare Gleichungssysteme

1.2 Matrixfunktionen

1.3 Modellreduktion

1.4 Eigenwertaufgaben

② Krylov-Unterraumverfahren

③ Lineare Gleichungssysteme

④ Matrixfunktionen

⑤ Krylov-Verfahren für Matrixfunktionen

⑥ Das QR-Verfahren für Eigenwertaufgaben

$$\mathbf{Ax} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{C}^{N \times N}, \quad \mathbf{b} \in \mathbb{C}^N.$$

- Bereits aus Grundvorlesung Numerik bekannt: **direkte** Verfahren wie Gauß-Elimination, Cholesky-Zerlegung, spezielle Varianten für dünn besetzte Matrizen, Toeplitz-, Vandermonde,
- Diskretisierung partieller Differentialgleichungen (FD, FE, FV, Spektral-Verfahren) oder Integralgleichungen (evtl. nach vorausgehender **Linearisierung**: direkte Verfahren zu aufwendig hinsichtlich **Speicherbedarf** oder **Rechenzeit**).
- Alternative: **Iterationsverfahren**. Erzeuge Folge von Näherungen $\{\mathbf{x}_m\}_{m \in \mathbb{N}_0}$ mit $\mathbf{x}_m \rightarrow \mathbf{A}^{-1}\mathbf{b}$, ($m \rightarrow \infty$).
- Einzig erforderliche Operation: Matrix-Vektor-Produkt

$$\mathbf{x} \rightarrow \mathbf{y} = \mathbf{Ax},$$

oft ohne explizite Darstellung von \mathbf{A} .

Die Matrix

$$A = \begin{bmatrix} 11. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 22. & 0. & 24. & 25. & 0. & 0. & 0. \\ 31. & 32. & 33. & 0. & 0. & 0. & 37. & 0. \\ 41. & 0. & 0. & 44. & 0. & 0. & 0. & 0. \\ 0. & 0. & 53. & 0. & 55. & 56. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 66. & 0. & 68. \\ 71. & 0. & 0. & 0. & 75. & 0. & 77. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0. & 87. & 88. \end{bmatrix} \in \mathbb{R}^{8 \times 8} \quad (1.1)$$

lässt sich vollständig aus den drei folgenden Vektoren rekonstruieren:

$$\begin{aligned} AA &= [11., 22., 24., 25., 31., 32., 33., 37., 41., 44., \\ &\quad 53., 55., 56., 66., 68., 71., 75., 77., 87., 88.], \\ JA &= [1, 2, 4, 5, 1, 2, 3, 7, 1, 4, 3, 5, 6, 6, 8, 1, 5, 7, 7, 8], \\ IA &= [1, 2, 5, 9, 11, 14, 16, 19, 21]. \end{aligned}$$

Einleitung

Beispiel: CSR-Format für dünn besetzte Matrizen

- AA : von Null verschiedene Einträge von A ; Länge: $\text{nnz}(A)$.
- $JA \in \mathbb{N}^{\text{nnz}(A)}$: Spaltenindices dieser Einträge.
- $IA \in \mathbb{N}^{N+1}$: Anfang der k -ten Zeile, $k = 1, \dots, N$.
- Beispiel: $IA(5) = 11$ bedeutet erster Nichtnulleintrag der 5. Zeile gegeben durch $AA(11) = 53$ mit Spaltenindex $JA(11) = 3$.
- Letzter Eintrag von IA : $IA(N + 1) = \text{nnz}(A) + 1$.
- Speicherbedarf: $\text{nnz}(A)$ reelle, $N + 1 + \text{nnz}(A)$ ganze Zahlen.
Vergleich: N^2 reelle Zahlen.
- Fortran77-Implementierung von $y \leftarrow Ax$ (vgl. etwa [Saad, 1996, Kapitel 3]):

```
C MATRIX-VECTOR PRODUCT Y=A*X, WHERE A IS STORED IN CSR-FORMAT
  DO I = 1,N
    K1 = IA(I)
    K2 = IA(I+1)-1
    Y(I) = DOTPRODUCT(AA(K1:K2),X(JA(K1:K2)))
  END DO
```

Einleitung

Beispiel: CSR-Format für dünn besetzte Matrizen

Für $A \in \mathbb{R}^{8 \times 8}$ aus (1.1) ergibt sich so $y \leftarrow Ax$ komponentenweise aus

I	K1	K2	Y(I)
1	1	1	11. * X(1)
2	2	4	22. * X(2) + 24. * X(4) + 25. * X(5)
3	5	8	31. * X(1) + 32. * X(2) + 33. * X(3) + 37. * X(7)
4	9	10	41. * X(1) + 44. * X(4)
5	11	13	53. * X(3) + 55. * X(5) + 56. * X(6)
6	14	15	66. * X(6) + 68. * X(8)
7	16	18	71. * X(1) + 75. * X(5) + 77. * X(7)
8	19	20	87. * X(7) + 88. * X(8)

- Rechenaufwand: $\text{nnz}(A)$ Multiplikationen, $\text{nnz}(A) - N$ Additionen (plus Indexberechnungen sowie Speicherzugriff). Vergleich: $O(N^2)$ Operationen.
- Matrix-Vektorprodukt in der Regel aufwendigste Komponente bei iterativen Lösungsverfahren.

Wir betrachten die Randwertaufgabe

$$-(\sigma(x)u'(x))' = f(x) \quad (0 < x < 1), \quad u(0) = u(1) = 0. \quad (1.2)$$

- Modell für viele physikalische Phänomene: Potentialströmung, elektrostatisches Potential, Wärmeleitung etc.
- u entspricht dem Potential oder der Temperatur.
- f : Quelle (stückweise stetig).
- $\sigma > 0$: (ortsabhängige) Leitfähigkeit, Koeffizient in Differentialoperator. (Stückweise stetig differenzierbar).
- Sachgemäß gestelltes Problem.

Diskretisierung mittels finiter Differenzen: Gitterpunkte¹

$$0 = x_0 < x_1 < \dots < x_n < x_{n+1} = 1.$$

Gitterweiten $h_j := x_{j+1} - x_j$ ($j = 0, 1, \dots, n$) nicht notwendig konstant.
Differenzengleichungen

$$-\sigma_{j+\frac{1}{2}} \frac{u(x_{j+1}) - u(x_j)}{h_j} + \sigma_{j-\frac{1}{2}} \frac{u(x_j) - u(x_{j-1}))}{h_{j-1}} = \frac{f_j^- h_{j-1} + f_j^+ h_j}{2} + \tau_j,$$

mit Fehlertermen $\{\tau_j\}_{j=1}^n$ sowie den Abkürzungen

$$\begin{aligned} \sigma_{j+\frac{1}{2}} &:= \sigma \left(x_j + \frac{1}{2} h_j \right), & \sigma_{j-\frac{1}{2}} &:= \sigma \left(x_j - \frac{1}{2} h_{j-1} \right), \\ f_j^+ &:= \lim_{x \rightarrow x_j^+} f(x), & f_j^- &:= \lim_{x \rightarrow x_j^-} f(x). \end{aligned}$$

¹Wir setzen voraus, dass sich die Unstetigkeitsstellen von f und σ unter diesen Punkten befinden.

Einleitung

Beispiel: 1D Randwertaufgabe

Die Koeffizientenfunktion σ geht in $\mathbf{A}(\sigma)$ ein durch den Vektor von Funktionswerten

$$\boldsymbol{\sigma} = [\sigma_{1/2}, \sigma_{1+1/2}, \dots, \sigma_{n+1/2}]^T \in \mathbb{R}^{n+1}.$$

Im Spezialfall äquidistanter Gitterpunkte ist

$$\mathbf{A} = \frac{1}{h} \text{tridiag}(-\sigma_{j-1/2}, \sigma_{j-1/2} + \sigma_{j+1/2}, -\sigma_{j+1/2}).$$

Ist außerdem f stetig so vereinfacht sich die rechte Seite zu

$$\mathbf{f} = h[f_j]_{j=1}^n, \quad f_j = f(x_j).$$

Abb. 1 zeigt, die Koeffizientenfunktion

$$\begin{aligned} \sigma(x) = & 1 + 3 \exp(-100(x - .18)^2) + 1.5 \exp(-50(x - .6)^2) \\ & + 2 \exp(-100(x - .8)^2) \end{aligned} \quad (1.4)$$

(oben) sowie die resultierende Lösung u der Randwertaufgabe (1.2) (unten). Dabei wurde $f(x) \equiv 1$ gewählt.

Einleitung

Beispiel: 1D Randwertaufgabe

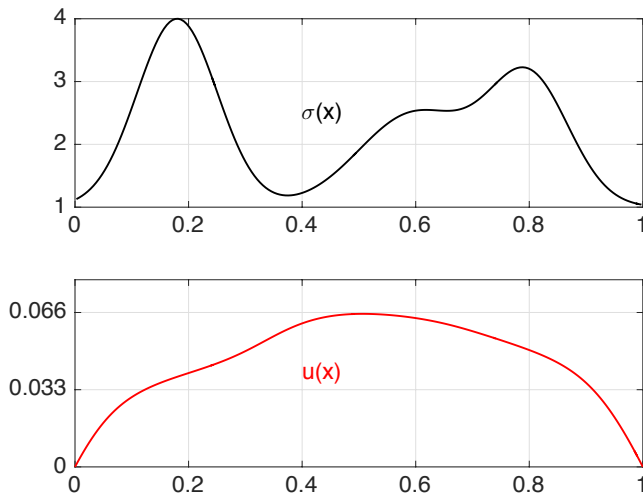


Abbildung 1: Koeffizientenfunktion σ aus (1.4) und Lösung u von (1.2).

Bemerkung 1.1

Wenn die Lösung u von (1.2) genügend glatt ist, kann man zeigen, dass

$$\max_{1 \leq j \leq n} |u(x_j) - u_j| \leq C \left[\max_{0 \leq j \leq n} h_j \right]^2$$

(die Konstante C hängt nur von u ab). Insbesondere streben die Lösungen der linearen Gleichungssysteme (1.3) gegen die Lösung der Randwertaufgabe (1.2), wenn die Schrittweite $h_{\max} = \max_{0 \leq j \leq n} h_j$ gegen 0 geht.

Einleitung

Beispiel: 1D Randwertaufgabe

Satz 1.2 (Eigenschaften der Koeffizientenmatrix $\mathbf{A}(\boldsymbol{\sigma})$)

Die Matrix $\mathbf{A} = \mathbf{A}(\boldsymbol{\sigma}) \in \mathbb{R}^{n \times n}$ ist für jeden komponentenweise positiven Vektor $\boldsymbol{\sigma}$ (wir schreiben $\boldsymbol{\sigma} > \mathbf{0}$) symmetrisch positiv definit und daher invertierbar. Insbesondere ist das lineare Gleichungssystem (1.3) eindeutig lösbar.

Der Beweis von Satz 1.2 erfordert einige Vorbereitungen:

Satz 1.3 (Gerschgorin (1931))

Für $\mathbf{A} = [a_{i,j}] \in \mathbb{C}^{n \times n}$ werden die **Gerschgorin-Kreisscheiben** wie folgt definiert

$$\mathbb{D}_i := \left\{ \zeta \in \mathbb{C} : |\zeta - a_{i,i}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{i,j}| \right\}, \quad i = 1, 2, \dots, n.$$

Dann gilt $\Lambda(\mathbf{A}) \subseteq \bigcup_{i=1}^n \mathbb{D}_i$ für das Spektrum $\Lambda(\mathbf{A})$ von \mathbf{A} .

Einleitung

Beispiel: 1D Randwertaufgabe

Wenden wir den Satz von Gerschgorin auf die Matrix $\mathbf{A} = \mathbf{A}(\sigma)$ aus (1.3) an ($\sigma > 0$), so sehen wir, dass alle Eigenwerte nichtnegativ sind. (Natürlich besitzt die symmetrische Matrix \mathbf{A} nur reelle Eigenwerte.) Um auszuschliessen, dass $\lambda = 0$ ein Eigenwert ist, benötigen wir eine Verschärfung von Satz 1.3.

Definition 1.4 (Reduzible und irreduzible Matrizen)

Eine Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ heißt **reduzibel**, wenn es eine Permutationsmatrix \mathbf{P} gibt (vgl. Aufgabe 2, 1. Übungsblatt), sodass \mathbf{PAP}^T die Blockform

$$\mathbf{PAP}^T = \left[\begin{array}{c|c} \mathbf{A}_{1,1} & \mathbf{O} \\ \hline \mathbf{A}_{2,1} & \mathbf{A}_{2,2} \end{array} \right],$$

mit quadratischen Diagonalklöcken besitzt. Gibt es keine solche Permutationsmatrix, so heißt \mathbf{A} **irreduzibel**.

Nur mithilfe dieser Definition ist es schwer zu entscheiden, ob eine gegebene Matrix reduzibel oder irreduzibel ist. Daher führen wir weitere Hilfsmittel ein.

Definition 1.5 (Graph einer Matrix)

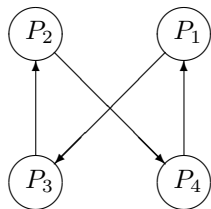
Wir ordnen der Matrix $\mathbf{A} = [a_{i,j}] \in \mathbb{C}^{n \times n}$ die n Knoten P_1, P_2, \dots, P_n zu. Für jedes Element $a_{i,j} \neq 0$ wird P_i mit P_j durch eine *gerichtete Kante* verbunden. Die Gesamtheit der Knoten und Kanten bildet den **gerichteten Graphen** $G(\mathbf{A})$ von \mathbf{A} .

Der gerichtete Graph $G(\mathbf{A})$ heißt **vollständig zusammenhängend**, wenn es für jedes geordnete Paar (P_i, P_j) von Knoten einen gerichteten Weg (d.h. eine Folge von gerichteten Kanten) gibt, der von P_i nach P_j führt.

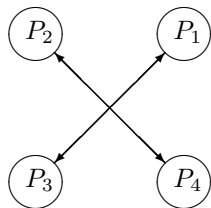
Einleitung

Beispiel: 1D Randwertaufgabe

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \text{ besitzt den Graphen } G(\mathbf{A}_1) =$$



$$\mathbf{A}_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \text{ besitzt den Graphen } G(\mathbf{A}_2) =$$



$G(\mathbf{A}_1)$ ist vollständig zusammenhängend ist, während $G(\mathbf{A}_2)$ nicht vollständig zusammenhängend ist.

Einleitung

Beispiel: 1D Randwertaufgabe

Lemma 1.6 (Charakterisierung irreduzibler Matrizen)

Die Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist genau dann irreduzibel, wenn ihr gerichteter Graph $G(\mathbf{A})$ vollständig zusammenhängend ist.

Satz 1.7 (Taussky (1948))

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ irreduzibel mit den Gerschgorin-Kreisscheiben $\mathbb{D}_1, \mathbb{D}_2, \dots, \mathbb{D}_n$. Liegt ein Eigenwert λ von \mathbf{A} auf dem Rand der Vereinigung aller Gerschgorin-Kreisscheiben, dann liegt er auf dem Rand jeder einzelnen Gerschgorin-Kreisscheibe:

$$\lambda \in \Lambda(\mathbf{A}), \lambda \in \partial \left(\bigcup_{i=1}^n \mathbb{D}_i \right) \Rightarrow \lambda \in \partial \mathbb{D}_i \quad \forall i = 1, 2, \dots, n.$$

Einleitung

Beispiel: 1D Randwertaufgabe

- Die Matrix $\mathbf{A}(\boldsymbol{\sigma})$ aus (1.3) ist für $\boldsymbol{\sigma} > \mathbf{0}$ irreduzibel, denn ihr Graph ist offensichtlich vollständig zusammenhängend.
- Wäre $\lambda = 0$ ein Eigenwert von \mathbf{A} , so müsste er auf dem Rand jeder Gerschgorin Kreisscheibe liegen, was aber nicht der Fall ist ($0 \notin \mathbb{D}_1$ und $0 \notin \mathbb{D}_n$).
- Also besitzt die Matrix \mathbf{A} nur positive Eigenwerte und ist damit – wie in Lemma 1.2 behauptet – positiv definit.

1 Einleitung

1.1 Lineare Gleichungssysteme

1.2 Matrixfunktionen

1.3 Modellreduktion

1.4 Eigenwertaufgaben

2 Krylov-Unterraumverfahren

3 Lineare Gleichungssysteme

4 Matrixfunktionen

5 Krylov-Verfahren für Matrixfunktionen

6 Das QR-Verfahren für Eigenwertaufgaben

Beispiel: Lösung des LGS $\mathbf{Ax} = \mathbf{b}$ mit $\mathbf{A} \in \mathbb{C}^{n \times n}$ invertierbar gegeben durch

$$\mathbf{x} = f(\mathbf{A})\mathbf{b}, \quad f: \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}, \quad \zeta \mapsto f(\zeta) = \frac{1}{\zeta}.$$

Beispiel: Exponentialfunktion einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ (beliebig) definiert durch

$$\exp(\mathbf{A}) = e^{\mathbf{A}} := \sum_{j=0}^{\infty} \frac{1}{j!} \mathbf{A}^j.$$

Satz 1.8

Die eindeutig bestimmte Lösung der Anfangswertaufgabe

$$\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t), \quad t > 0, \quad \mathbf{y}(0) = \mathbf{b}, \quad (1.5)$$

mit gegebenen $\mathbf{A} \in \mathbb{C}^{n \times n}$ und $\mathbf{b} \in \mathbb{C}^n$ ist die Funktion

$$\mathbf{y}(t) = \exp(t\mathbf{A})\mathbf{b}, \quad t \geq 0.$$

Die Lösung von (1.5) zur Zeit $t \geq 0$ lässt sich also ebenfalls ausdrücken durch die Anwendung einer Matrixfunktion auf einen Vektor:

$$\mathbf{y}(t) = f(\mathbf{A})\mathbf{b} \quad \text{mit} \quad f(\zeta) = f(\zeta; t) = \exp(t\zeta).$$

Die Zeit t spielt hier die Rolle eines Parameters.

Korollar 1.9

Die inhomogene Form der AWA (1.5), nämlich

$$\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t) + \mathbf{c}, \quad \mathbf{y}(0) = \mathbf{b}, \quad \mathbf{c} \in \mathbb{C}^n,$$

besitzt die eindeutig bestimmte Lösung $\mathbf{y}(t) = \exp(t\mathbf{A})\mathbf{b} + t\varphi_1(t\mathbf{A})\mathbf{c}$ mit

$$\varphi_1(\zeta) := \sum_{j=1}^{\infty} \frac{\zeta^{j-1}}{j!} = \frac{\exp(\zeta) - 1}{\zeta}.$$

Durch Hinzunahme der „Grundoperation“ $\mathbf{b} \mapsto f(\mathbf{A})\mathbf{b}$ ($f = \exp, \varphi_1$) lassen sich also einige Anfangswertaufgaben exakt lösen, vgl. [Hochbruck & Ostermann, 2010].

Beispiel: Instationäre Wärmeleitungsgleichung für $u = u(x, t)$:

$$\begin{aligned}u_t - (\sigma u_x)_x &= 0, & (x, t) &\in (0, 1) \times (0, t_{\text{end}}], \\u(x, 0) &= u_0(x), & x &\in (0, 1), \\u(0, t) &= u(1, t) = 0, & t &\in (0, t_{\text{end}}].\end{aligned}$$

Diskretisierung im Ort mit Differenzenschema wie im stationären 1D Beispiel ergibt System gewöhnlicher Differentialgleichungen

$$\mathbf{D} \begin{bmatrix} u'_1(t) \\ u'_2(t) \\ \vdots \\ u'_n(t) \end{bmatrix} + \mathbf{A}(\boldsymbol{\sigma}) \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_n(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \\ \vdots \\ u_n(0) \end{bmatrix} = \begin{bmatrix} u_0(x_1) \\ u_0(x_2) \\ \vdots \\ u_0(x_n) \end{bmatrix},$$

mit

$$\mathbf{D} := \text{diag} \left(\frac{h_0 + h_1}{2}, \frac{h_1 + h_2}{2}, \dots, \frac{h_{n-1} + h_n}{2} \right) \in \mathbb{R}^{n \times n}$$

für den zeitlichen Verlauf $u(x_j, t)$, $t \in (0, t_{\text{end}})$ der Näherungslösung an den Gitterpunkten $\{x_j\}_{i=1}^n$. ($\mathbf{A}(\boldsymbol{\sigma})$ wie in (1.2).)

Kompakter:

$$\mathbf{D} \mathbf{u}'(t) + \mathbf{A}(\boldsymbol{\sigma}) \mathbf{u}(t) = \mathbf{0}, \quad \mathbf{u}(0) = \mathbf{u}_0. \quad (1.6)$$

Die Lösung von (1.6) ist nach Satz 1.8

$$\mathbf{u}(t) = \exp(-t\mathbf{D}^{-1}\mathbf{A}(\boldsymbol{\sigma})) \mathbf{u}_0.$$

Abb. 2 zeigt sie für die Koeffizientenfunktion σ aus (1.4), die Anfangsbedingung

$$u_0(x) = x(1-x) \frac{2\delta(x-.5)^2}{\delta^2 + (x-.5)^4} \quad \text{mit } \delta = 0.01$$

und $t_{\max} = 1/2$.

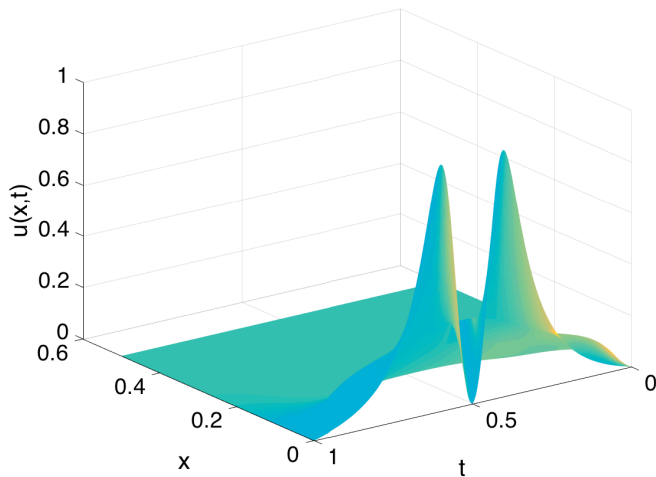


Abbildung 2: Eine Lösung der instationären Wärmeleitungsgleichung.

1 Einleitung

1.1 Lineare Gleichungssysteme

1.2 Matrixfunktionen

1.3 Modellreduktion

1.4 Eigenwertaufgaben

2 Krylov-Unterraumverfahren

3 Lineare Gleichungssysteme

4 Matrixfunktionen

5 Krylov-Verfahren für Matrixfunktionen

6 Das QR-Verfahren für Eigenwertaufgaben

- Ziel bei **Modellreduktion** [engl. *model order reduction* oder *reduced-order modeling*] ist es, die Komplexität eines mathematischen Modells zu reduzieren, ohne sein Verhalten wesentlich zu verändern.
- Beschreibt das Modell ein System oder einen Prozess, so soll das reduzierte Modell gewisse Aspekte dieses Systems oder Prozesses zumindest näherungsweise beschreiben.
- Transformation eines hochdimensionalen Problems auf ein niedrigdimensionales, wobei das reduzierte Problem gewisse Eigenschaften mit dem Originalsystem teilen soll.

Z.B. beschreibt

$$\begin{aligned} \mathbf{u}'(t) &= \mathbf{A}\mathbf{u}(t) + \mathbf{B}\mathbf{z}(t), & \mathbf{u}(t) &\in \mathbb{R}^n, \mathbf{z}(t) \in \mathbb{R}^m, \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{u}(t), & \mathbf{y}(t) &\in \mathbb{R}^\ell, \end{aligned} \tag{1.7}$$

ein **lineares zeit-invariantes** System (engl. *linear descriptor system*).

- $\mathbf{u}(t)$ nennt man den **Zustand**, seine Dimension n ist die **Ordnung** des Systems,
- $\mathbf{z}(t)$ den **Eingang** (engl. *input*) und
- $\mathbf{y}(t)$ den **Ausgang** (engl. *output*) des Systems.
- Einfachster Fall: $m = \ell = 1$, *single-input, single-output (SISO) system*.
Gegensatz: *multiple-input, multiple-output (MIMO) system*

- Simulation des Systems: berechne Übergang $z(t) \mapsto y(t)$.
Kosten proportional zu n^3 (es sei denn A ist schwach besetzt oder es stehen schnelle Löser für das Differentialgleichungssystem zur Verfügung)².
- Wenn die Ordnung n sehr groß ist und die Dimensionen der Eingangs- und Ausgangssignale klein sind (d.h. $\ell \ll n$, $m \ll n$), ist die Wahrscheinlichkeit hoch, dass das System hochgradig **redundant** ist.
- Ziel: System niedriger Ordnung

$$\begin{aligned} \mathbf{u}'_r(t) &= \mathbf{A}_r \mathbf{u}_r(t) + \mathbf{B}_r z(t), & \mathbf{x}_r(t) &\in \mathbb{R}^r, \\ \mathbf{y}_r(t) &= \mathbf{C}_r \mathbf{u}_r(t), & \mathbf{y}_r(t) &\in \mathbb{R}^\ell, \end{aligned} \tag{1.8}$$

also $r \ll n$, konstruieren, das das Input-Output-Verhalten des Originalsystems möglichst genau wiedergibt.

²Genauer: Anzahl Zeitschritte $\times n^2 [n^3]$ bei expliziten [impliziten] Zeitschrittverfahren.

- **Laplace-Transformation** angewandt auf lineares System (1.7) ($\mathbf{u}(0) = \mathbf{0}$):

$$\begin{aligned}s \hat{\mathbf{u}}(s) &= \mathbf{A} \hat{\mathbf{u}}(s) + \mathbf{B} \hat{\mathbf{z}}(s), \\ \hat{\mathbf{y}}(s) &= \mathbf{C} \hat{\mathbf{u}}(s)\end{aligned}$$

bzw.

$$\hat{\mathbf{y}}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} \hat{\mathbf{z}}(s) =: \mathbf{H}(s) \hat{\mathbf{z}}.$$

- Die Matrix $\mathbf{H}(s) \in \mathbb{R}^{\ell \times m}$ heißt **Übertragungsfunktion** des Systems.
- Analog besitzt das reduzierte System eine Übertragungsfunktion $\mathbf{H}_r \in \mathbb{R}^{\ell \times m}$.
- Wenn (1.8) eine gute Approximation an (1.7) ist, wird der Abstand

$$\|\mathbf{H}_r(s) - \mathbf{H}(s)\| \quad (\text{in einer geeigneten Norm})$$

klein sein.

Weiteres Beispiel: erneut instationäre Wärmeleitungsgleichung

$$\begin{aligned}u_t - (\sigma u_x)_x &= 0, & (x, t) &\in (0, 1) \times (0, t_{\text{end}}], \\u(x, 0) &= I\delta(x), & x &\in (0, 1), \quad I > 0 \\u_x(0, t) &= u(1, t) = 0, & t &\in (0, t_{\text{end}}].\end{aligned}$$

AB (Dirac-Impuls) entspricht Punktquelle an der Stelle $x = 0$.
Diskretisierung wie in vorangegangenen Beispielen.

Wegen der Neumann-RB jetzt auch $u(0, t)$ unbekannt; zusätzliche Gleichung

$$\frac{\sigma_{1/2}}{h_0}u_0 - \frac{\sigma_{1/2}}{h_0}u_1 = 0,$$

(was natürlich $u_0 = u_1$ bedeutet).

Neue Matrix:

$$B(\boldsymbol{\sigma}) = \left[\begin{array}{c|cccc} \frac{c_{1/2}}{h_0} & -\frac{c_{1/2}}{h_0} & 0 & \dots & 0 \\ \hline -\frac{c_{1/2}}{h_0} & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right] \in \mathbb{R}^{(n+1) \times (n+1)}$$

$A(\boldsymbol{\sigma})$

mit $A(\boldsymbol{\sigma})$ wie oben. Semidiskrete Version der **Anfangsrandwertaufgabe**:

$$D\mathbf{u}'(t) - B(\boldsymbol{\sigma})\mathbf{u} = \mathbf{0}, \quad \mathbf{u}(0) = \frac{I}{h_0} \mathbf{e}_1$$

mit

$$D = \frac{1}{2} \text{diag}(h_0, h_0 + h_1, h_1 + h_2, \dots, h_{n-1} + h_n) \in \mathbb{R}^{(n+1) \times (n+1)}$$

und dem ersten Einheitsvektor $\mathbf{e}_1 \in \mathbb{R}^{n+1}$.

Betrachte Koeffizientenvektor $\boldsymbol{\sigma} > \mathbf{0}$ als Input, den Wert von

$$u(0, t) = u_0(t) = \mathbf{e}_1^\top \mathbf{u}(t)$$

als den Output unseres Systems.

Input-Output-Abbildung:

$$\boldsymbol{\sigma} \mapsto u_0(t) = \frac{I}{h_0} \mathbf{e}_1^\top \exp(-t\mathbf{D}^{-1}\mathbf{B}(\boldsymbol{\sigma})) \mathbf{e}_1.$$

① Einleitung

1.1 Lineare Gleichungssysteme

1.2 Matrixfunktionen

1.3 Modellreduktion

1.4 Eigenwertaufgaben

② Krylov-Unterraumverfahren

③ Lineare Gleichungssysteme

④ Matrixfunktionen

⑤ Krylov-Verfahren für Matrixfunktionen

⑥ Das QR-Verfahren für Eigenwertaufgaben

Eigenwerte treten klassischerweise (aber nicht ausschließlich) in Anwendungen auf, bei denen **Schwingungsphänomene** von Interesse sind.

- Strukturdynamik
- Magnetohydrodynamik
- Elektrische Schaltkreise
- Quantenchemie
- Chemische Reaktionskinetik, insbesondere Verbrennung
- Steuerungs- und Regelungstechnik, dynamische Systeme
- Volkswirtschaftslehre
- Modellreduktion, Data Mining, UQ

Bei der Analyse des Schwingungsverhaltens von Konstruktionen (z.B. Flugzeuge, Gebäude) ergibt sich nach Diskretisierung im Ort (z.B. mittels finiter Elemente (FE)), im einfachsten Fall eine gedämpfte Schwingungsgleichung

$$M\ddot{u} + C\dot{u} + Ku = F$$

für die zeitliche Evolution der FE-Freiheitsgrade u . Hierbei ist M die Massenmatrix, K die Steifigkeitsmatrix, C modelliert Dämpfung oder gyroskopische Effekte, und F bezeichnet eine äußere Kraft.

Die **Eigenschwingungen** der Struktur ergeben sich aus Lösungen $u(t) = ve^{i\omega t}$ für $C = O$, $F = 0$, was auf

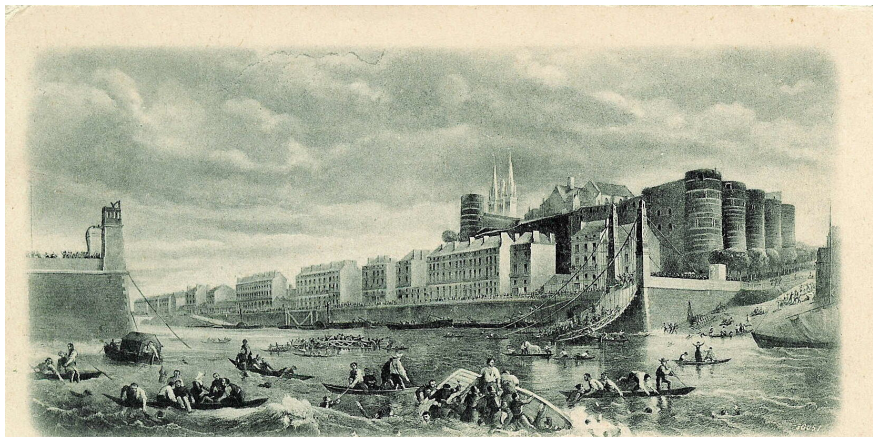
$$Kv = \omega^2 Mv,$$

d.h. ein verallgemeinertes Eigenwertproblem führt.

In technischen Systemen versucht man, Anregungen der Eigenschwingungen der Struktur zu unterbinden. Berühmte Beispiele: Brücken, z.B. Broughton Suspension Bridge, Manchester (1831), Angers (1850), Tacoma Narrows Bridge (1940), Millenium Bridge, London (2000).

Einleitung

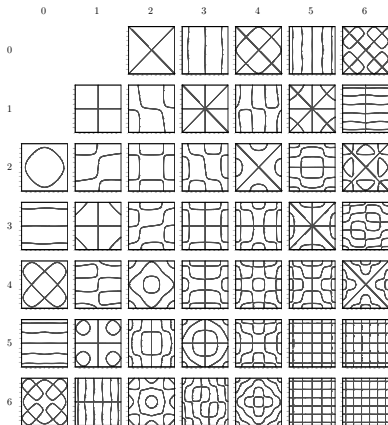
Eigenwertaufgaben in der Strukturdynamik



Catastrophe du Pont suspendu, actuellement Pont de la Basse chaine en l'année 1850 (le 11^{ème} Leger)
Angers



Die Klangfiguren von Chladni



Die Lösung der Anfangswertaufgabe

$$\dot{\mathbf{u}} = \mathbf{A}\mathbf{u}, \quad \mathbf{u}(0) = \mathbf{u}_0, \quad \mathbf{u} = \mathbf{u}(t) \in \mathbb{C}^n, \quad \mathbf{A} \in \mathbb{C}^{n \times n}, \quad (1.9)$$

(homogenes System linearer gewöhnlicher Dgln. mit konstanten Koeffizienten)
lautet bekanntlich

$$\mathbf{u}(t) = \exp(t\mathbf{A})\mathbf{u}_0, \quad t \geq 0.$$

Das Verhalten von \mathbf{u} für $t \rightarrow \infty$ kann man mit Hilfe der **Jordanschen Normalform**

$$\mathbf{J} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \text{diag}(\mathbf{J}_1, \dots, \mathbf{J}_k), \quad \mathbf{J}_\ell = \lambda_q \mathbf{I} + \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 \end{bmatrix}$$

ermitteln.

Satz 1.10

Für die Lösung \mathbf{u} von (1.9) gilt

(a) Genau dann gilt $\mathbf{u}(t) \rightarrow \mathbf{0}$ für $t \rightarrow \infty$ (**stabiler Fall**), wenn

$$\operatorname{Re} \lambda < 0 \quad \forall \lambda \in \Lambda(\mathbf{A}).$$

(b) Die Lösung ist unbeschränkt für $t \rightarrow \infty$ (**instabiler Fall**), wenn

$$\operatorname{Re} \lambda > 0 \quad \text{für mindestens ein } \lambda \in \Lambda(\mathbf{A}).$$

(c) Gilt

$$\max_{\lambda \in \Lambda(\mathbf{A})} \operatorname{Re} \lambda = 0,$$

so bleibt die Lösung genau dann beschränkt, wenn alle Eigenwerte $\lambda \in \Lambda(\mathbf{A})$ mit $\operatorname{Re} \lambda = 0$ halbeinfach sind.

Erinnerung: Man nennt einen Eigenwert **halbeinfach** (engl. *semisimple*), wenn seine algebraische Vielfachheit mit seiner geometrischen übereinstimmt.

Die Stabilitätsanalyse nichtlinearer dynamischer Systeme

$$\dot{\mathbf{u}} = \mathbf{f}(\mathbf{u}), \quad \mathbf{u}(0) = \mathbf{u}_0, \quad \mathbf{f} : \mathbb{C}^n \rightarrow \mathbb{C}^n, \quad (1.10)$$

geschieht meist durch Linearisierung um den (einen) stationären Zustand $\bar{\mathbf{u}} := \lim_{t \rightarrow \infty} \mathbf{u}(t)$. Sofern ein solcher existiert ist dieser Lösung der Gleichung $\mathbf{f}(\mathbf{u}) = \mathbf{0}$.

Das dynamische System (1.10) heißt **lokal stabil** in $\bar{\mathbf{u}}$, falls $\epsilon > 0$ existiert mit

$$\lim_{t \rightarrow \infty} \mathbf{u}(t) = \bar{\mathbf{u}}, \quad \text{sofern } \|\mathbf{u}(0) - \bar{\mathbf{u}}\| < \epsilon.$$

In vielen Fällen läßt sich die Frage nach der lokalen Stabilität von (1.10) durch Analyse der Linearisierung

$$\mathbf{f}(\mathbf{u}) \approx \mathbf{f}(\bar{\mathbf{u}}) + \mathbf{A}(\mathbf{u} - \bar{\mathbf{u}}), \quad \mathbf{A} := \mathbf{f}'(\bar{\mathbf{u}})$$

klären, d.h. durch die Realteile der Eigenwerte von \mathbf{A} .

Einleitung

Beispiel: Chemische Reaktionskinetik

Die Belousov-Zhabotinsky-Reaktion ist ein Beispiel für einen sog. **chemischen Oszillator**, bei dem sich zeitliche Oszillationen in einem chemischen Reaktionsmechanismus zeigen.



BZ-Reaktion in einer Petri-Schale, Wellenfront in gelb markiert.

Ein mathematisches Modell der BZ-Reaktion ist der sog. **Brusselator**³, einer Evolutionsgleichung der örtlichen Variation (in einer Raumkoordinate $r \in (0, 1)$) der Konzentration zweier miteinander reagierender Spezies x und y :

$$\partial_t x = D_1 \partial_{rr} x + A - (B + 1)x + x^2 y, \quad x(0, t) = x(1, t) = A.$$

$$\partial_t y = D_2 \partial_{rr} y + Bx - x^2 y, \quad y(0, t) = y(1, t) = \frac{B}{A},$$

$$x(r, 0) = x_0(r), \quad y(r, 0) = y_0(r).$$

³Brussels + Oszillator, Ilya Prigogine FU Brüssel

Einleitung

Beispiel: Chemische Reaktionskinetik

Ein stationärer Zustand ist gegeben durch

$$\bar{x} = A, \quad \bar{y} = \frac{B}{A}.$$

Die Jacobi-Matrix an dieser Stelle ist gegeben durch

$$\mathbf{J} = \begin{bmatrix} D_1 \partial_{rr} + B - 1 & A^2 \\ -B & D_2 \partial_{rr} - A^2 \end{bmatrix},$$

was nach Ortsdiskretisierung eine große, dünn besetzte Matrix ergibt.

Bifurkationsproblem: Ab welchem Wert von B setzt periodisches Verhalten ein? Hierbei überqueren zwei konjugiert-komplexe Eigenwerte von \mathbf{J} die imaginäre Achse (**Hopf-Bifurkation**).

Gegeben seien k Messungen (Realisierungen) $\{\mathbf{x}_j\}_{j=1}^k$ eine vektorwertigen Zufallsvariablen (Zufallsvektor) $\mathbf{x} \in \mathbb{R}^n$. Die mittelwertbereinigte Datenmatrix \mathbf{X} sei definiert durch

$$\mathbf{X} := [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k] - \bar{\mathbf{x}} \mathbf{e}^T, \quad \bar{\mathbf{x}} := \frac{1}{k} \sum_{j=1}^k \mathbf{x}_j, \quad \mathbf{e} := [1, 1, \dots, 1]^T.$$

- Die (Orthonormalbasis aus) Eigenvektoren der symmetrisch positiv semidefiniten Matrix $\mathbf{X}\mathbf{X}^T \in \mathbb{R}^{n \times n}$ stellt n paarweise unkorrelierte Größen dar, deren Varianz durch den zugehörigen Eigenwert gegeben ist.
- Die Gesamtvarianz ist gegeben durch die Summe der Eigenwerte.
- Dieses statistische Verfahren wird als **Hauptkomponentenanalyse** (principal components analysis, PCA) bezeichnet.
- Verallgemeinerungen hiervon sind als **Karhunen-Loève Entwicklung**, **Proper Orthogonal Decomposition (POD)** bekannt.