

Numeric

Ruben Triwari

INHALTSVERZEICHNIS

CHAPTER

NORMEN UND KONDITIONSZAHL PAGE 2

Kapitel 1

Normen und Konditionszahl

1: Matrixnormen

Sei $A \in \mathbb{K}^{N \times N}$ beliebig.

Die Spaltensummennorm ist die maximale Summe der Spalten einer Matrix.

$$\|A\|_1 = \max_{n=1,2,\dots,N} \sum_{m=1}^N |a_{m,n}|$$

Die Zeilensummennorm ist die maximale Summe der Zeilen einer Matrix.

$$\|A\|_\infty = \max_{n=1,2,\dots,N} \sum_{m=1}^N |a_{n,m}|$$

Die Frobeniusnorm ist die Wurzel der Quadratsumme über alle Einträge einer Matrix.

$$\|A\|_F = \sqrt{\sum_{m=1}^N \sum_{n=1}^N |a_{n,m}|^2} = \sqrt{\sum_{n=1}^N \lambda_n(A^*A)}$$

Die zweite Gleichheit gilt, da

$$\text{spur}(A^*A) = \sum_{n=1}^N \sum_{m=1}^N \overline{a_{n,m}} a_{n,m} = \sum_{n=1}^N \sum_{m=1}^N |a_{n,m}|^2 = \|A\|_F^2$$

Spektralnrm:

$$\|A\|_2 = \max_{n=1,2,\dots,N} \sqrt{\lambda_n(A^*A)}$$

2: Induzierte Matrixnorm

Sei $\|\cdot\|$ eine Norm auf \mathbb{K}^N , dann ist die induzierte Matrixnorm:

$$\|A\| = \sup_{0 \neq x \in \mathbb{K}^N} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|.$$

Es gilt die Submultiplikativität:

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|.$$

Sowie:

$$\|\mathbb{1}\| = \sup_{\|x\|=1} \|\mathbb{1}x\| = \sup_{\|x\|=1} \|x\| = 1.$$

Wichtige induzierte Matrixnormen:

$$\|x\|_1 \leftrightarrow \|A\|_1$$

$$\|x\|_\infty \leftrightarrow \|A\|_\infty$$

$$\|x\|_2 \leftrightarrow \|A\|_2$$

3: Konditionszahl

Die Konditionszahl ist der Fehlerverstärker falls eine Störung in einer regulären Matrix $A \in \mathbb{K}^{N \times N}$ in einem linearen Gleichungssystem vorkommt. Sei $\|\cdot\|$ eine Matrixnorm, dann setze:

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|.$$

Es gilt $\forall \alpha \in \mathbb{K} \setminus \{0\}$:

$$\text{cond}(\alpha A) = \|\alpha A\| \cdot \|(\alpha A)^{-1}\| = |\alpha \cdot \alpha^{-1}| \cdot \|A\| \cdot \|A^{-1}\| = \text{cond}(A)$$

Falls $\|\cdot\|$ eine induzierte Matrixnorm ist, gilt wegen Submultiplikativität:

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\| \geq \|A \cdot A^{-1}\| = 1$$

Außerdem für eine induzierte Matrixnorm:

$$\text{cond}(A) = \frac{\sup_{\|x\|=1} \|Ax\|}{\inf_{\|x\|=1} \|Ax\|}$$

Daraus folgt für die Spektralnorm:

$$\text{cond}(A) = \sqrt{\frac{\max_n \lambda_n(A^*A)}{\min_n \lambda_n(A^*A)}}$$

4: Fehlerabschätzungen für lineare Gleichungssysteme

Sei $A, \Delta A \in \mathbb{K}^{N \times N}$ und $b, \Delta b, x, \Delta x \in \mathbb{K}^N$.

Mit $Ax = b$ und $A(x + \Delta x) = b + \Delta b$ gilt dann:

$$\|\Delta x\| \leq \|A^{-1}\| \cdot \|\Delta b\| \quad \text{und} \quad \frac{\|\Delta x\|}{\|x\|} \leq \text{cond}(A) \frac{\|\Delta b\|}{\|b\|}$$

Dies motiviert die Konditionszahl.

Falls nun $\|\Delta A\| < \frac{1}{\|A^{-1}\|}$ gilt dann $A + \Delta A$ regulär, mit

$$\|(A + \Delta A)^{-1}\| \leq \|A^{-1}\| \frac{1}{1 - \|\Delta A\| \cdot \|A^{-1}\|}$$

Daraus erhält man falls weiter $\|\Delta A\| < \frac{1}{\|A^{-1}\|}$ gilt und mit

$$(A + \Delta A)(x + \Delta x) = b + \Delta b.$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \text{cond}(A) \frac{1}{1 - \|\Delta A\| \cdot \|A^{-1}\|} \left(\frac{\Delta A}{A} + \frac{\Delta b}{b} \right)$$

5: Strikt diagonaldominante Matrizen

Sei $A \in \mathbb{K}^{N \times N}$ dann heißt die Matrix strikt diagonaldominant, wenn für alle $n = 1, 2, \dots, N$ gilt:

$$\sum_{m=1}^N |a_{n,m}| < |a_{n,n}|$$

Eigenschaften falls A strikt diagonaldominant:

- A regulär
- So ist bei Anwendung der L-R Zerlegung $a_{n,n}^{(n)} \neq 0 \quad \forall n$
- Die Pivoelementsuche liefert immer $a_{n,n}^{(n)}$

6: LR-Zerlegung

Sei $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \in \mathbb{K}^N$ ein Vektor, $n \in \{1, 2, \dots, N\}$ mit $y_n \neq 0$ und setze

$$l_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{y_{n+1}}{y_n} \\ \vdots \\ \frac{y_N}{y_n} \end{pmatrix}.$$

Wobei die ersten n Einträge null sind. Daraus erhält man die Matrix:

$$L_n := E_n - l_n \cdot e_n^* = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \frac{-y_{n+1}}{y_n} & 1 & \\ & & \vdots & & \ddots \\ & & \frac{-y_N}{y_n} & & & 1 \end{pmatrix}$$

Die y Einträge befinden sich in der n -ten Spalte und in den $(n+1) - N$ Zeilen. Diese Matrix hat folgende Eigenschaften:

$$L_n y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und} \quad L_n e_m = e_m \quad \text{für } m \neq n$$

Sei A eine beliebige Matrix mit $y = A_n$, dann setze:

$$R := L_{N-1} \cdots L_2 L_1 A \quad \text{und} \quad L := L_1^{-1} L_2^{-1} \cdots L_{N-1}^{-1}$$

Dann sind L und R Dreiecksmatrizen und für L gilt:

$$L = E_n + l_1 e_1^* + l_2 e_2^* + \cdots + l_{N-1} e_{N-1}^*$$

Jetzt kann das LGS einfach mit L und R gelöst werden:

$$Ax = L(Rx) = b \iff Lz = b \quad \text{und} \quad Rx = z$$

Dieser Algorithmus verwendet folgende Anzahl an Schritten:

$$\text{Rechenaufwand} = \frac{1}{3}N^3 - \frac{1}{3}N$$

7: Pivotsuche

Idee: Falls das Pivoelement bei der LR-Zerlegung $a_{n,n}^{(n)} = 0$ ist, tauschen wir die Zeilen oder Spalten um ein Pivotelement $a_{n,n}^{(n)} \neq 0$ erhalten.

Die Vertauschungen können mittels einer Permutationsmatrix durchgeführt werden:

$$P_n := \left(\begin{array}{ccc|ccc|ccc} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ \hline & & & 0 & 0 & \cdots & 0 & 1 & \\ & & & 0 & 1 & & & 0 & \\ & & & \vdots & & \ddots & & \vdots & \\ & & & 0 & & & 1 & 0 & \\ & & & 1 & 0 & \cdots & 0 & 0 & \\ \hline & & & & & & & & 1 & \ddots & \\ & & & & & & & & & \ddots & 1 \end{array} \right) \begin{array}{l} \leftarrow n \\ \leftarrow p \end{array}$$

$\uparrow \quad \uparrow$
 $n \quad p$

Die Matrix heißt elementare Permutationsmatrix und hat folgende Eigenschaften:

- $A \cdot P_n$ vertauscht n-te und p-te Zeile von A
- $P_n \cdot A$ vertauscht n-te und p-te Spalte von A
- $P_n^2 = E_n$

Der n-te Schritt mit Pivotsuche ist daher:

$$A^{(n+1)} = L_n P_n A^{(n)}$$

Dann gilt für R und L:

$$R := \tilde{L}_{N-1} \tilde{L}_{N-2} \cdots \tilde{L}_1 P A \quad \text{und} \quad L := \tilde{L}_1^{-1} \tilde{L}_2^{-1} \cdots \tilde{L}_{N-1}^{-1}$$

mit

$$\tilde{L}_n := P_{N-1} \cdots P_{n+1} L_n P_{n+1} \cdots P_{N-1}$$

Also gilt wieder:

$$PA = LR \quad \text{dann lässt man wieder} \quad Ly = Pb \quad \text{und} \quad Rx = y$$

Spaltenpivotsuche:

Das m-te Pivotelement mit dem höchsten Score wird gewählt mit der Spaltenmaximierungsstrategie:

$$\text{Pivoelement-Zeile} = \operatorname{argmax}_{n=m, \dots, N} = \frac{|a_{n,m}|}{\sum_{l=m}^N |a_{n,l}|}$$

8: LR-Zerlegung Algorithmus

Algorithmus:

1. Start: $A^{(1)} := A$, $L^{(1)} := 0$, $P^{(1)} := E_N$

2. Pivotsuche ergibt neues Element: Vertausche Zeilen von $A^{(n)}$, also berechne:
 $A'^{(n)} := P_{n,j} A^{(n)}$, $P^{(n)} := P_{n,j} P^{(n-1)}$

3. Pivotsuche gibt kein neues Element: $P^{(n)} := P^{(n-1)}$

4. Berechne: $l_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{a_{n+1,n}^{(n)}}{a_{n,n}^{(n)}} \\ \vdots \\ \frac{a_{N,n}^{(n)}}{a_{n,n}^{(n)}} \end{pmatrix}$, $L_n := \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & l_{n+1} & 0 & \\ & & \vdots & \ddots & \\ & & l_N & & 0 \end{pmatrix}$, $L_-^{(n)} := E_N - L^{(n)}$

5. Berechne: $A^{(n+1)} := L_-^{(n)} A^{(n)}$

6. Schluss: $R := A^{(N)}$, $L = E_N + L^{(1)} + \dots + L^{(N)}$, $P := P^{(N)}$

9: Cholesky Zerlegung

Sei $A \in \mathbb{K}^{N \times N}$ hermitesch und positiv definit, dann gibt es eine untere Dreiecksmatrix mit positiven Diagonalelementen und

$$A = LL^*.$$

Ansatz zur Berechnung:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix} = \begin{pmatrix} l_{1,1} & & & \\ l_{2,1} & l_{2,2} & & \\ \vdots & \vdots & \ddots & \\ l_{N,1} & l_{N,2} & \cdots & l_{N,N} \end{pmatrix} \cdot \begin{pmatrix} l_{1,1} & \overline{l_{1,2}} & \cdots & \overline{l_{1,N}} \\ & l_{2,2} & \cdots & \overline{l_{2,N}} \\ & & \ddots & \vdots \\ & & & l_{N,N} \end{pmatrix}$$

Mit Rechenaufwand:

$$\text{Rechenaufwand} = \frac{1}{6}N^3 + \mathcal{O}(N^2)$$

10: Householder-Transformation

Die Householder-Transformation bzgl. eines Vektors $v \in \mathbb{K}^N$ ist:

$$P = E_n - \frac{2}{v^*v} v v^* \in \mathbb{K}^{N \times N}$$

Die Householder-Transformation ist eine Spiegelung an der Ebene orthogonal zu v .
 Für P gilt folgende Eigenschaften:

- $Pv = -v$
- $Pw = w \quad \forall w \perp v$
- P hermitesch, also $P^* = P$
- P unitär, also $P^*P = E_N$

11: QR-Zerlegung

Sei $A \in \mathbb{K}^{M \times N}$ mit $M \geq N$ und $\text{rang}(A) = N$. Dann gibt es eine unitäre Matrix $Q \in \mathbb{K}^{M \times M}$ und eine obere Dreiecksmatrix $R \in \mathbb{K}^{M \times N}$ mit

$$A = QR$$

Wegen Q unitär kann das Gleichungssystem wie folgt gelöst werden:

$$Rx = Q^*b$$

Algorithmus:

Im n -ten Schritt haben wir $A^{(n)}$ bestimmt und sei $a_n = A_n^{(n)}$ die n -te Spalte von $A^{(n)}$.

1. Berechne $v_n := \frac{a_n}{\|a_n\|_2} + \delta_n e_1$ mit $\delta_n := \begin{cases} \frac{(a_n)_1}{|(a_n)_1|} & , (a_n)_1 \neq 0 \\ 1 & , \text{sonst} \end{cases}$
2. Berechne $\beta_n := \frac{2}{v_n^* v_n}$, $w_n := A^* v_n$
3. Berechne $P_n = E_N - \beta_n v_n v_n^* = \begin{pmatrix} E_n & 0 \\ 0 & P'_n \end{pmatrix} \implies P_n A^{(n)} = A^{(n)} - \beta_n v_n w_n^* = \begin{pmatrix} r_{n,n} & * \\ 0 & A^{(n+1)} \end{pmatrix}$
4. Schluss: $R := P_n \cdots P_1 A$ und $Q := P_1 \cdots P_n$

Bemerkungen:

- QR-Zerlegung nicht eindeutig für $M \neq N$.
- Die Konditionszahl von der QR Zerlegung ist gleich der von A , also $\text{cond}_2(A) = \text{cond}_2(RQ) = \text{cond}_2(R)$.
- Also besser als bei der LR Zerlegung mit $\text{cond}_2(A) \leq \text{cond}_2(L)\text{cond}_2(R)$, wobei die Konditionszahlen von L und R viel höher als A sein können.
- Auch besser als die Cholesky Zerlegung, da für die cholesky-Zerlegung gilt: $\text{cond}_2(A) \leq (\text{cond}_2(L'))^2$.
- QR-Zerlegung ist doppelt so aufwendig wie LR-Zerlegung und viermal so aufwendig wie cholesky-Zerlegung.
- Falls der Rang nicht voll ist kann wieder Vertauschungen von Spalten verwendet werden, um den Algorithmus trotzdem anzuwenden zu können.
- Der Algorithmus kann leicht abgewandelt werden dass die Pivoelemente immer ungleich Null sind.

$$\text{Rechenaufwand} = MN^2 - \frac{1}{3} + O(MN)$$

12: Lineare Ausgleichsrechnung

Sei $Ax = b$ mit $A \in \mathbb{K}^{M \times N}$, $x \in \mathbb{K}^N$, $b \in \mathbb{K}^M$ und $M > N$. Dieses LGS ist allgemein nicht lösbar, deswegen brauchen wir einen neuen Lösungsbegriff:

$$\inf_{x \in \mathbb{K}^N} \|Ax - b\|_2$$

Die Minimierer sind die Lösungen der Gaußschen Normalengleichung:

$$A^*Ax = A^*b$$

Also eine Alternative zu der QR-Zerlegung

13: Banachscher Fixpunktsatz

Sei M vollständig bzgl. einer Metrik d und sei $\phi : M \rightarrow M$ eine Abbildung, sodass es ein $q > 1$ gibt mit

$$\forall x, y \in M : d(\phi(x), \phi(y)) \leq q \cdot d(x, y)$$

Dann gibt es **genau ein** $\hat{x} \in M$ mit $\phi(\hat{x}) = \hat{x}$. Außerdem konvergiert für jedes $x^{(0)} \in M$ die Folge

$$x^{(n+1)} := \phi(x^{(n)}), \forall n \in \mathbb{N}_0$$

gegen \hat{x} und für $n \geq 1$ gilt:

- (a) Monotonie: $d(x^{(n)}, \hat{x}) \leq q \cdot d(x^{(n-1)}, \hat{x})$
- (b) A-priori-Schranke: $d(x^{(n)}, \hat{x}) \leq \frac{q^n}{1-q} \cdot d(x^{(0)}, x^{(1)})$
- (c) A-posteriori-Schranke: $d(x^{(n)}, \hat{x}) \leq \frac{q}{1-q} \cdot d(x^{(n)}, x^{(n-1)})$

14: Fixpunktsatz angewendet auf lineare Gleichungssysteme

Anwendung auf LGS $Ax = b$, zerlege in $A = M - N$. Daraus erhalten wir

$$Mx = Nx + b.$$

dass kann nun als ein Fixpunktproblem geschrieben werden:

$$x = Tx + c \quad \text{mit} \quad T = M^{-1}N \quad \text{und} \quad c = M^{-1}b$$

also $\phi(x) = Tx + c$.

Dieses Problem konvergiert falls $\|T\| < 1$, für eine induzierte Matrixnorm.

Gesamtschrittverfahren(Jacobi-Verfahren):

Hier ist $M = D$, $N = L + R$, wobei D Diagonalmatrix mit Einträgen ungleich Null, dann ist das wieder ein Fixpunktproblem mit:

$$\begin{aligned} x_{\text{neu}} &= D^{-1}(b + (L + R)x_{\text{alt}}) \\ \iff x_{\text{neu}} &= \underbrace{D^{-1}b}_{:=c} + \underbrace{D^{-1}(L + R)}_{:=T} x_{\text{alt}} \end{aligned}$$

oder konkret:

$$x_n^{(k+1)} = \frac{1}{a_{n,n}}(b_n - \sum_{\substack{m=1 \\ m \neq n}}^N a_{n,m} \cdot x_m^{(k)})$$

Einzelschrittverfahren(Gauß-Seidel-Verfahren):

Hier ist $M = D - L$ und $N = R$, daraus:

$$\begin{aligned} x &= (D - L)^{-1}(b - Rx) \\ \rightsquigarrow x^{(\text{neu})} &= (D - L)^{-1}(b - Rx^{(\text{alt})}) \\ \iff x^{(\text{neu})} &= D^{-1}(b + Lx^{(\text{neu})} - Rx^{(\text{alt})}) \end{aligned}$$

konkret:

$$x_n^{k+1} = \frac{1}{a_{n,n}}(b_n - \sum_{k=1}^{n-1} a_{n,m} x_m^{k+1} - \sum_{k=n+1}^N a_{n,m} x_m^k)$$

Falls A strikt diagonaldominant konvergieren beide Verfahren für jeden Startwert $x^{(0)} \in \mathbb{K}^N$

15: Konjugierte Gradienten

Sei $A \in \mathbb{K}^{N \times N}$ hermitesch und positiv definit. Das Verfahren berechnet die exakte Lösung wird aber oft vorher abgebrochen.

Algorithmus:

1. Start: Wähle $x^{(0)} \in \mathbb{K}^N$ und setze $d^{(0)} := r^{(0)} = b - Ax^{(0)}$. Ist $d^{(0)} = 0$ kann abgebrochen werden.

2. Berechne:

$$\alpha_{k-1} := \frac{d^{(k-1)*} r^{(k-1)}}{d^{(k-1)*} A d^{(k-1)}}, \quad \beta_{k-1} := \frac{d^{(k-1)*} A r^{(k)}}{d^{(k-1)*} A d^{(k-1)}} \quad \text{und} \quad r^{(k)} := b - Ax^{(k)}$$

3. Berechne:

$$x^{(k)} := x^{(k-1)} + \alpha_{k-1} d^{(k-1)} \quad \text{und} \quad d^{(k)} := r^{(k)} + \beta_{k-1} d^{(k-1)}$$

4. Schluss: $d^{(k)} = 0 \implies x^{(k)} = \hat{x}$ mit $A\hat{x} = b$

Das Verfahren konvergiert nach genau N Schritten.

Konvergenzrate:

$$\|x^{(k)} - \hat{x}\|_2 \leq 2 \operatorname{cond}(A) \left(\frac{\sqrt{\operatorname{cond}(A)} - 1}{\sqrt{\operatorname{cond}(A)} + 1} \right)^k \|x^0 - \hat{x}\|_2$$

16: Eigenwerteinschließung

Sei $A, \Delta A \in \mathbb{K}^{N \times N}$, $\|\cdot\|$ eine induzierte Matrixnorm und $\sigma(A) := \{\lambda \in \mathbb{C} : \lambda \text{ ist Eigenwert von } A\}$. Dann gilt (Bauer-Fike):

$$\min_{\lambda(A) \in \sigma(A)} |\lambda(A) - \mu(\Delta A)| \leq \operatorname{cond}(A) \cdot \|\Delta A\|.$$

Wenn A normal ($A^*A = AA^*$) ist, so gilt sogar

$$\min_{\lambda(A) \in \sigma(A)} |\lambda(A) - \mu(\Delta A)| \leq \|\Delta A\|_2.$$

Gerschgorin:

Sei

$$\mathcal{K}_n := \{\xi \in \mathbb{C} : |a_{n,n} - \xi| \leq \sum_{\substack{m=1 \\ m \neq n}}^N |a_{n,m}|\} \quad \text{und} \quad \mathcal{K}_n^* := \{\xi \in \mathbb{C} : |a_{n,n} - \xi| \leq \sum_{\substack{m=1 \\ m \neq n}}^N |a_{m,n}|\}.$$

Dann ist

$$\sigma(A) \subset \left(\bigcup_{n=1}^N \mathcal{K}_n \right) \cap \left(\bigcup_{n=1}^N \mathcal{K}_n^* \right).$$

Wenn $\mathcal{N}_1 \dot{\cup} \mathcal{N}_2 = \{1, 2, \dots, N\}$ mit

$$\left(\bigcup_{n \in \mathcal{N}_1} \mathcal{K}_n \right) \cap \left(\bigcup_{n \in \mathcal{N}_2} \mathcal{K}_n^* \right) = \emptyset.$$

Dann enthalten $\bigcup_{n \in \mathcal{N}_i} \mathcal{K}_n$, $i = 1, 2$ je genau $\#\mathcal{N}_i$ Eigenwerte.

Courant-Fischer:

Falls A hermitesch mit Eigenwerten $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Dann gilt für jeden Unterraum $\mathcal{L} \subset \mathbb{K}^N$ mit $\dim(\mathcal{L}) = n$, dass

$$\min_{0 \neq x \in \mathcal{L}} \frac{x^* A x}{x^* x} \leq \lambda_n \quad \text{und} \quad \max_{0 \neq x \in \mathcal{L}} \frac{x^* A x}{x^* x} \geq \lambda_{N-n+1}$$

mit Gleichheit $\mathcal{L} = \{x_1, \dots, x_n\}$ bzw. $\mathcal{L} = \{x_{N-n+1}, \dots, x_N\}$, wobei $(x_m)_{m=1}^N$ Eigenvektoren einer Orthonormalbasis sind.

17: Courant-Fischer Folgerung

Sei $A, \Delta A \in \mathbb{K}^{N \times N}$ hermitesch, dann

$$\lambda_n(A) + \lambda_N(\Delta A) \leq \lambda_n(A + \Delta A) \leq \lambda_n(A) + \lambda_1(\Delta A), \forall n = 1, \dots, N$$

und insbesondere

$$|\lambda_n(A + \Delta A) - \lambda_n(A)| \leq \|\Delta A\|_2 \forall n = 1, \dots, N$$

18: Potenzmethode

Motivation: Sei $A \in \mathbb{K}^{N \times N}$ diagonalisierbar, und $(v_n)_{n=1}^N$ eine Orthonormalbasis aus Eigenvektoren, so gilt für $x = \sum_{n=1}^N \xi_n v_n$, dass

$$A^k x = \sum_{n=1}^N \lambda_n^k \xi_n v_n.$$

Also dominiert der betragsgrößte Eigenwert für große k .

Sei $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_N|$ die Eigenwerte von A , $0 \neq y \in \mathbb{C}^N$, mit $A^* y = \overline{\lambda_1} y$ und sei $\tilde{z}^{(0)} \in \mathbb{C}^N$ mit $y^* \tilde{z}^{(0)} \neq 0$, dann:

$$z^{(k)} := \frac{A^k \tilde{z}^{(0)}}{\|A^k \tilde{z}^{(0)}\|}, \forall k \geq 0$$

so gilt

$$\limsup_{k \rightarrow \infty} \| \|Az^{(k)}\| - |\lambda_1| \| \leq \left(\frac{|\lambda_2|}{|\lambda_1|} \right)^k.$$

Außerdem gibt es genau ein Eigenvektor von λ_1 mit $\|v\| = 1$ und $\text{sgn}(y^* v) = \text{sgn}(y^* \tilde{z}^{(0)})$ und dann gilt:

$$\limsup_{k \rightarrow \infty} \| z^{(k)} - \text{sgn}(\lambda_1)^k v \| \leq \left(\frac{|\lambda_2|}{|\lambda_1|} \right)^k.$$

Algorithmus:

1. Start: Wähle $\tilde{z}^{(0)}$, sodass möglichst $y^* \tilde{z}^{(0)} \neq 0$
2. Berechne $z^k := \frac{\tilde{z}^{(k)}}{\|\tilde{z}^{(k)}\|}$
3. Berechne $\tilde{z}^{(k)} := Az^{(k-1)}$
4. Schluss Eigenwert: $|\lambda_1| \approx \|Az^{(k)}\|$
5. Schluss Eigenvektor: $\text{sgn}(\lambda_1)^k z^{(k)} \approx v$

Definition 1.0.1: Lokale Konvergenz

Ein Iterationsverfahren $x_{k+1} = \phi(x_k)$, mit $\phi(\hat{x}) = \hat{x}$ heißt *lokal Konvergent* gegen $\hat{x} \in \mathbb{K}$, falls es eine offene Menge $U \subset \mathcal{D}(\phi)$ mit $\hat{x} \in U$ gibt, sodass für jedes $x^{(0)} \in U$ das Iterationsverfahren definiert ist und gegen \hat{x} konvergiert.

Definition 1.0.2: Konvergenz Ordnung

Ein lokal gegen \hat{x} konvergentes Iterationsverfahren heißt konvergent *von der Ordnung* $p \geq 1$, falls es ein $\rho > 0$ und $C < \infty$ (mit $C < 1$ für $p = 1$) gibt mit:

$$\|x_{k+1} - \hat{x}\| \leq C \|x_k - \hat{x}\|^p, \quad \forall k \geq 0 \quad \text{und} \quad \|x_0 - \hat{x}\| < \rho$$

Example 1.0.1

Das Gesamtschrittverfahren(Banach), Einzelschrittverfahren(Banach) und CG-Verfahren konvergieren mit Ordnung 1.

19: Konvergenz von Iterationsverfahren

Falls $\phi : \mathcal{D}(\phi) \subset \mathbb{K}^N \rightarrow \mathbb{K}^N$ stetig differenzierbar mit \hat{x} im inneren von $\mathcal{D}(\phi)$ und es gibt eine Induzierte Matrixnorm mit $\|\phi'(x)\| < 1$ so ist die Fixpunktiteration lokal konvergent gegen \hat{x}

20: Ordnung von Iterationsverfahren

Die Funktion $\phi : \mathcal{D}(\phi) \subset \mathbb{K} \rightarrow \mathbb{K}$ sei p -mal stetig differenzierbar mit $p \geq 2$ und habe einen Fixpunkt \hat{x} im inneren von $\mathcal{D}(\phi)$. Ist

$$\phi'(x) = \dots = \phi^{(p-1)}(x) = 0$$

so ist die Fixpunktiteration lokal konvergent gegen \hat{x} mit Konvergenzordnung mindestens p . Ist zusätzlich $\phi^{(p)}(\hat{x}) \neq 0$, ist die Ordnung genau p .

21: Newton-Verfahren

Die Funktion $f : \mathcal{D}(f) \subset \mathbb{R} \rightarrow \mathbb{R}$ ist differenzierbar und sei $\hat{x} \in \mathbb{R}$ mit $f(\hat{x}) = 0$, dann heißt folgende Vorschrift *Newton-Verfahren*:

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}, \forall k \geq 0$$

Falls $f \in C^3[a, b]$ und $\hat{x} \in (a, b)$ mit $f(\hat{x}) = 0, f'(\hat{x}) \neq 0$. Dann konvergiert das Newton-Verfahren lokal quadratisch, also $p = 2$, gegen \hat{x} .

Sekantenverfahren:

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k)$$

Sei $-\infty < a < b \leq \infty$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig mit $f(a) = 0$ und in (a, b) stetig differenzierbar $f'(x) > 0$ für alle $x \in (a, b)$. Außerdem gelte

$$f(x) \geq f'(x)(x - a) \forall x \in (a, b).$$

Dann konvergiert das Newton-Verfahren für alle $x_0 \in (a, b)$ gegen a und die iterierte ist streng monoton fallend.

22: Polynom-Interpolation

Seien $x_0, \dots, x_N \in \mathbb{R}$ paarweise verschieden. Dann heißen $w(x) := \prod_{n=0}^N (x - x_n)$ das Knotenpolynom. Falls die Knoten angeordnet sind durch $x_0 < \dots < x_N$ heißt,

$$l_n(x_k) := \frac{w(x_k)}{(x_k - x_n)w'(x_k)} = \prod_{\substack{m=0 \\ m \neq n}}^N \frac{x_k - x_m}{x_n - x_m} = \delta_{n,k}$$

das Lagrange-Grundpolynom. Dann gibt es für alle $y_0, \dots, y_N \in \mathbb{K}$, genau ein Polynom p von $\text{Grad} \leq N$ mit

$$p(x_n) = y_n, \forall n = 0, \dots, N.$$

Dieses Polynom ist gegeben durch

$$p(x) = \sum_{n=0}^N y_n l_n(x)$$

Sei nun $f \in C^{N+1}[a, b]$ und p das Interpolationspolynom vom $\text{Grad} \geq N$. Dann gibt es zu jedem $x \in [a, b]$ ein $\xi \in \text{conv}(x, x_0, \dots, x_N)$ (Konvexhülle) mit

$$f(x) - p(x) = \frac{f^{(N+1)}(\xi)}{(N+1)!} w(x).$$

Insbesondere ist

$$|f(x) - p(x)| \leq \frac{\max_{\xi \in [a, b]} |f^{(N+1)}(\xi)|}{(N+1)!} |w(x)|.$$

Es gilt:

$$\min_{x_0, \dots, x_N \in [a, b]} \max_{x \in [a, b]} |(x - x_0) \dots (x - x_N)| = \frac{(b - a)^{N+1}}{2 \cdot 4^N}.$$

Die beste Wahl ist dann für die Knoten:

$$x_n = \frac{1}{2} \left((b - a) \cos \frac{(2n-1)\pi}{2(N+1)} + a + b \right), n = 0, \dots, N$$

Sei nun $K_0, \dots, K_M \in \mathbb{N}$. Dann gibt es für alle $(y_m^{(k)})_{\substack{m=0, \dots, M \\ k=0, \dots, K_m-1}}$ genau ein Polynom p von

$\text{Grad} \geq N := \left(\sum_{m=0}^M K_m \right) - 1$ mit

$$p^{(k)}(x_m) = y_m^{(k)}, \quad \text{für } m = 0, \dots, M, \quad k = 0, \dots, K_m - 1.$$

Falls $f \in C^{N+1}[a, b]$ und für jedes $x \in [a, b]$ gibt es ein $\xi \in \text{conv}(x, x_0, \dots, x_M)$ mit

$$f(x) - p(x) = \frac{f^{(N+1)}(\xi)}{(N+1)!} w(x)$$

wobei

$$w(x) := (x - x_0)^{K_0} \dots (x - x_M)^{K_M}.$$

Definition 1.0.3: Splines

Sei $[a, b]$ Intervall und Knoten $a = x_0 < x_1 < \dots < x_n = b$. Sei

$$h_n := x_n - x_{n-1}, n = 1, \dots, N$$

und

$$h := \max_n h_n$$

Ein Spline von Grad L ist eine Funktion $s \in C^{L-1}[a, b]$, die auf jeden Intervall $[x_{n-1}, x_n]$ mit einem Polynom von Grad $\geq L$ übereinstimmt. Die Menge aller Splines vom Grad L wird mit S_L oder $S_{L,\Delta}$ mit $\Delta = \{x_0, \dots, x_N\}$ bezeichnet.

S_L ist ein $(N + L)$ -dimensionaler Vektorraum.

Definition 1.0.4: L_2 Norm

$$\|f\|_{L_2(a,b)} := \left(\int_a^b |f(x)|^2 dx \right)^{\frac{1}{2}}$$

Definition 1.0.5: Sobolevraum

Der Sobolevraum $H^1(a, b)$ besteht aus allen Funktionen $f : [a, b] \rightarrow \mathbb{K}$ für die es ein $c \in \mathbb{K}$ und $g \in L^2(a, b)$ gibt mit

$$f(x) = c + \int_a^x g(t) dt \quad \forall x \in [a, b]$$

Periodischer Sobolevraum:

$$H_{\text{per}}^1 := \{f \in H^1(0, 2\pi) : f(0) = f(2\pi)\}$$

Für beliebige $s > 0$ kann dieser Charakterisiert werden durch:

$$H_{\text{per}}^s(0, 2\pi) = \left\{ f \in L^2(0, 2\pi) : \sum_{k \in \mathbb{Z}} (|k|^{2s} + 1) |\hat{f}_k|^2 < \infty \right\}$$

Mit Norm:

$$\|f\|_{H_{\text{per}}^s(0, 2\pi)} := \left(2\pi \sum_{k \in \mathbb{Z}} (|k|^{2s} + 1) |\hat{f}_k|^2 \right)^{\frac{1}{2}}$$

23: Lineare Splines

Hutfunktionen:

$$\Lambda_0(x) = \begin{cases} \frac{x-x_1}{x_0-x_1} & \text{falls } x_0 \leq x \leq x_1, \\ 0 & \text{falls } x_1 \leq x \leq x_N \end{cases}$$

$$\Lambda_n(x) = \begin{cases} 0 & \text{falls } x_0 \leq x \leq x_{n-1}, \\ \frac{x-x_{n-1}}{x_n-x_{n-1}} & \text{falls } x_{n-1} \leq x \leq x_n \\ \frac{x-x_{n+1}}{x_n-x_{n+1}} & \text{falls } x_n \leq x \leq x_{n+1}, \\ 0 & \text{falls } x_{n+1} \leq x \leq x_N \end{cases}$$

$$\Lambda_N(x) = \begin{cases} 0 & \text{falls } x_0 \leq x \leq x_{N-1}, \\ \frac{x-x_{N-1}}{x_N-x_{N-1}} & \text{falls } x_{N-1} \leq x \leq x_N \end{cases}$$

Für alle $y_0, \dots, y_N \in \mathbb{K}$ ist

$$s = \sum_{n=0}^N y_n \Lambda_n$$

der eindeutig bestimmte lineare Spline mit

$$s(x_n) = y_n, n = 0, \dots, N$$

Intuition: $\Lambda_n(x_m) = \delta_{n,m}, n, m = 0, \dots, N$

Sei nun $f \in C^2[a, b]$, dann gilt:

$$\max_{x \in [a, b]} |s(x) - f(x)| \leq \frac{h^2}{8} \max_{\xi \in [a, b]} |f''(\xi)|$$

Sei nun $f \in H^1(a, b)$, dann gilt

$$\|s - f\|_{L_2(a, b)} \leq \frac{h}{\pi} \|f'\|_{L_2(a, b)}$$

und für $f \in H^2(a, b)$

$$\|s - f\|_{L_2(a, b)} \leq \frac{h^2}{\pi^2} \|f''\|_{L_2(a, b)} \quad \text{und} \quad \|s' - f'\|_{L_2(a, b)} \leq \frac{h}{\pi} \|f''\|_{L_2(a, b)}$$

Hierbei wird die optimale konstante saturiert durch

$$f(x) = \sin \frac{\pi(x-a)}{b-a}.$$

24: Kubische Splines

Brachte Polynome Grad ≤ 3 zwischen den Knoten. Sei x_0, \dots, x_N und y_0, \dots, y_N dann

$$s(x_n) = y_n, \forall n = 1, \dots, N$$

Wir haben nicht genug Informationen, da S_3 Dimension $N + 3$ und wir haben nur $N + 1$ Werte. Setze:

$$s_n := s(x_n) \quad s'_n := s'(x_n) \quad \gamma_n := s''(x_n) \quad n = 0, \dots, N$$

Daraus ergibt sich folgendes Gleichungssystem:

In Matrixschreibweise:

$$\frac{1}{6} \begin{pmatrix} h_1 & 2(h_1 + h_2) & h_2 & & & \\ & h_2 & 2(h_2 + h_3) & h_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & h_{N-1} & 2(h_{N-1} + h_N) & h_N \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_N \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{N-1} \end{pmatrix},$$

wobei wir gesetzt haben

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{N-1} \end{pmatrix} := - \begin{pmatrix} -\frac{1}{h_1} & \frac{1}{h_1} + \frac{1}{h_2} & -\frac{1}{h_2} & & & \\ & -\frac{1}{h_2} & \frac{1}{h_2} + \frac{1}{h_3} & -\frac{1}{h_3} & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\frac{1}{h_{N-1}} & \frac{1}{h_{N-1}} + \frac{1}{h_N} & -\frac{1}{h_N} \end{pmatrix} \begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_N \end{pmatrix}$$

Die obigen Matrizen haben die Dimension $(n - 1) \times (n + 1)$, also nicht invertierbar. Aber mit den Randbedingungen:

- $s''(a) = s''(b) = 0$ natürliche Splines
- $s'(a) = y'_0$ und $s'(b) = y'_N$ vollständige kubische Splines

Mit der jeweiligen Bedingung sind obigen Matrizen Regulär, also können die Polynome des Splines bestimmt werden. Außerdem ist obiges LGS gut konditioniert:

$$\text{cond}_2(\dots) \leq 3 \frac{\max h_n}{\min h_n}$$

für äquidistante Knoten folgt

$$\text{cond}_2(\dots) \leq 3$$

Falls $f \in C^4[a, b]$ dann ist bzgl Maximumsnorm:

$$\max_{x \in [a, b]} |s(x) - f(x)| \leq \frac{5}{386} h^4 \max_{\xi \in [a, b]} |f^{(4)}(\xi)|$$

Falls $f \in H^4(a, b)$ dann ist bzgl L_2 -Norm:

$$\|s - f\|_{L_2(a, b)} \leq \frac{h^4}{\pi^4} \|f^{(4)}\|_{L_2(a, b)}, \quad \|s' - f'\|_{L_2(a, b)} \leq \frac{h^3}{\pi^3} \|f^{(4)}\|_{L_2(a, b)}, \quad \|s'' - f''\|_{L_2(a, b)} \leq \frac{h^2}{\pi^2} \|f^{(4)}\|_{L_2(a, b)}$$

25: Fouriertransformation

Fouriertransformation:

Für $f \in L^1(0, 2\pi)$ heißen die Zahlen

$$\hat{f}_k := \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} f(\theta) d\theta$$

die Fourierkoeffizienten von f und die Abbildung $f \mapsto (\hat{f})_{k \in \mathbb{Z}}$ heißt Fouriertransformation.

Sei nun $\mathcal{F}_N : \mathbb{C} \rightarrow \mathbb{C}$ die diskrete Fouriertransformation. Für einen Vektor $f = (f_0, \dots, f_{N-1})^T \in \mathbb{C}$ bezeichnet $\mathcal{F}f = ((\mathcal{F}_N f)_0, \dots, (\mathcal{F}_N f)_{N-1})^T \in \mathbb{C}$ Vektor mit Komponenten:

$$(\mathcal{F}_N f)_k := \frac{1}{N} \sum_{n=0}^{N-1} f_n e^{-2\pi i \frac{kn}{N}}, k = 0, \dots, N-1.$$

Schnelle Fouriertransformation:

Sei $N = 2M$ gerade, $\omega_N := e^{\frac{2\pi i}{N}}$, $\omega_M := w_N^2 = e^{\frac{2\pi i}{M}}$. Für $f = (f_0, \dots, f_{N-1})^T \in \mathbb{C}$ setze $f^{(+)}, f^{(-)} \in \mathbb{C}$:

$$f^{(+)}_m := f_m + f_{m+M}, \quad f^{(-)}_m := (f_m - f_{m+M})\omega_N^{-m}, \quad \forall m = 0, \dots, M-1.$$

Dann ist

$$(\mathcal{F}_N f)_{2m} = \frac{1}{2} (\mathcal{F} f^{(+)}_m), \quad (\mathcal{F}_N f)_{2m+1} = \frac{1}{2} (\mathcal{F} f^{(-)}_m) \quad \forall m = 0, \dots, M-1.$$

Sei $f \in L^2(0, 2\pi)$. Dann ist

$$\sum_{k \in \mathbb{Z}} |\hat{f}|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(\theta)|^2 d\theta < \infty$$

und $\sum_{k=-K}^K \hat{f}_k e^{ik\theta}$ konvergiert für $K \rightarrow \infty$ in L^2 gegen f .

Sei andersherum $(\alpha_k)_{k \in \mathbb{Z}}$ eine Folge mit $\sum_{k \in \mathbb{Z}} |\alpha_k|^2 < \infty$. Dann konvergiert $\sum_{k=-K}^K \alpha_k e^{ik\theta}$ für $K \rightarrow \infty$ in L^2 gegen eine Funktion f und es gilt $\hat{f}_k = \alpha_k, \forall k$

Sei $f \in H^s_{\text{per}(0, 2\pi)}$ mit $s < \frac{1}{2}$ dann konvergiert die Fourierreihe von f ($\sum_{k=-K}^K \hat{f}_k e^{ik\theta}$) gleichmäßig in $[0, 2\pi]$ gegen f und f ist stetig.

26: Trigonometrische Interpolation

Sei O.B.d.A: $L = 2\pi$, $0 \leq \theta_0 < \theta_1 \cdots < \theta_{N-1} \leq L$ und y_0, \dots, y_{N-1} . Zur Vereinfachung äquidistante Knoten:

$$\theta_n = \frac{2\pi n}{N}$$

Funktion aus:

$$\mathcal{T}_N := \left\{ \sum_{-\frac{N}{2}+1 \leq k \leq \frac{N}{2}} \alpha_k e^{ik\theta} : \alpha_k \in \mathbb{C}, \forall -\frac{N}{2}+1 \leq k \leq \frac{N}{2} \right\}$$

Bei N ungerade, gilt für die Indexmenge:

$$-\frac{N-1}{2} \leq k \leq \frac{N-1}{2}$$

$t \in \mathcal{T}_N$ das das Interpolationsproblem löst ist gegeben durch:

$$t(\theta) = \sum_{-\frac{N}{2}+1 \leq k \leq \frac{N}{2}} \alpha_k e^{ik\theta}$$

mit

$$\alpha_k = \frac{1}{N} \sum_{n=0}^{N-1} y_n e^{ik\theta_n}, \forall -\frac{N}{2}+1 \leq k \leq \frac{N}{2}$$

Fehlerabschätzung:

Sei $f \in H_{\text{per}}^s(0, 2\pi)$ mit $s > \frac{1}{2}$ und $t_N \in \mathcal{T}_N$ das interpolierende trigonometrisches Polynom. Dann gilt:

$$\|t_N - f\|_{L^2(0, 2\pi)} \leq 2^s \sqrt{1 + c_s} N^{-s} \|f\|_{H_{\text{per}}^s(0, 2\pi)},$$

mit $c_s := 2 \sum_{l=1}^{\infty} (2l-1)^{-2s}$