Mitschrieb

Numerik

Emma Bach

Basierend auf:

Vorlesungen Numerik I + II von
Prof. Dr. Patrick Dondl

2025-11-26

Inhalt

Chapter 1

Aufgabenstellung

In der Numerik beschäftigt man sich mit der praktischen Berechnung von Lösungen mathematischer Probleme.

Beispiel 1.0.1. Berechne $\int_0^1 e^{-x^2} dx$!

Beispiel 1.0.2. Berechne $\sin(20)$!

Beispiel 1.0.3. Berechne $\sqrt{753}$!

Beispiel 1.0.4. Berechne $\min_{x \in [0,1]} F(x)$, für eine geeignete Funktion F!

Beispiel 1.0.5. Berechne x, sodass $f(x) = 0$!

Beispiel 1.0.6. Berechne $x \in \mathbb{R}^n$, sodass $Ax = b$!

Definition 1.0.7. Eine Mathematische Aufgabe in der Numerik besteht im Finden einer Lösung von

$$F(x, d) = 0$$

für gegebenes Datum d und gegebene Funktion F.

Typischerweise können in akzeptabler Zeit keine exakten Lösungen gefunden werden, sondern nur Approximationen. Insbesondere stehen statt den vollen Mengen $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ etc. auch nur endlich viele **Maschinenzahlen** zur Verfügung - arbiträre reelle Zahlen benötigen unendlich viel Speicher! Rechenoperationen sind dementsprechend Fehlerbehaftet, es gibt Rundungsfehler. Außerdem gibt es in reellen Anwendungen oft **Modellfehler** und **Datenfehler**.

Eine Grundlegende Idee in der Numerik ist es deshalb, eine gute Balance zwischen Exaktheit und Aufwand der Berechnung zu finden.

Beispiel 1.0.8. Die Berechnung der Determinante einer Matrix mittels Laplaceschem Entwicklungssatz benötigt $O(n!)$ Rechenoperationen. Die Determinante mit diesem Verfahren zu berechnen, dauert sehr viel länger, als das Universum alt ist.

Besser: Matrix (approximativ) auf Dreiecksgestalt bringen und die Diagonalelemente multiplizieren.

Definition 1.0.9. Eine Mathematische Aufgabe heißt **wohlgestellt**, wenn zu geeigneten Daten d eindeutige Lösungen x existieren, und diese stetig von d abhängt. Andernfalls ergibt die Suche nach einer numerischen Lösung wenig Sinn. Für wohlgestellte Probleme existiert eine Lösungsfunktion φ, sodass $x = \varphi(d)$ das Problem löst, d.h. $f(\varphi(d), d) = 0$.

Definition 1.0.10. Ein numerischer Algorithmus zur näherungsweisen Lösung einer wohlgestellten Aufgabe φ ist eine Abbildung $\tilde{\varphi}$, die durch Hintereinanderausführung möglicherweise fehlerbehafteter elementarer Rechenoperationen definiert ist, also

$$\tilde{\varphi} = f_j \circ f_{j-1} \circ \ldots \circ f_1$$

Definition 1.0.11. Der **Aufwand** eines Verfahrens $\tilde{\varphi}$ ist die Anzahl der benötigten elementaren Rechenschritte. Typischerweise interessiert uns nicht die exakte Anzahl an Schritten, sondern nur die Größenordnung.

Proposition 1.0.12. *Das Gaußverfahren hat Aufwand* $\mathcal{O}(n^3)$.

Chapter 2

Numerische Lineare Algebra

2.1 Matrixfaktorisierung

2.1.1 Dreiecksmatrizen

Definition 2.1.1. Eine Matrix $L \in \mathbb{R}^{n \times n}$ heißt **untere Dreiecksmatrix**, falls $\forall i < j : l_{ij} = 0$.

Definition 2.1.2. Eine Matrix $U \in \mathbb{R}^{n \times n}$ heißt **obere Dreiecksmatrix**, falls $U^\top$ eine untere Dreiecksmatrix ist.

Definition 2.1.3. Eine Dreiecksmatrix heißt **normalisiert**, falls alle ihre Diagonaleinträge 1 sind.

Definition 2.1.4. Eine Matrix heißt **regulär**, wenn sie invertierbar ist.

Lemma 2.1.5. *Die quadratischen oberen (bzw. unteren) Dreiecksmatrizen bilden unter Matrixmultiplikation eine Gruppe.*

Lineare Gleichungssysteme mit regulärer Dreiecksmatrix lassen sich leicht lösen. Sei $U \in \mathbb{R}^{n \times n}$ eine reguläre obere Dreiecksmatrix und $b \in \mathbb{R}^n$. Wir berechnen $x \in \mathbb{R}^n$ folgendermaßen:

1. for i = n : -1 : 1:
   (a) $x_i = \left(b_i - \sum_{j=i+1}^{n} u_{ij} x_j \right) \cdot \frac{1}{u_{ii}}$
2. end.

Der Aufwand dieses Verfahrens ist $\mathcal{O}(n^2)$. Ein analoger Algorithmus existiert für untere Dreiecksmatrizen.

2.1.2 LU-Zerlegung

Falls für eine reguläre Matrix $A \in \mathbb{R}^{n \times n}$ eine Zerlegung $A = LU$ in eine untere Dreiecksmatrix U und eine obere Dreiecksmatrix L gegeben ist, so lässt sich das lineare Gleichungssystem $Ax = b$ in zwei Schritten lösen:

1. Löse $Ly = b$.
2. Löse $Ux = y$.

Definition 2.1.6. Eine Faktorisierung $A = LU$ mit unterer Dreiecksmatrix $L \in \mathbb{R}^{n\times n}$ und oberer Dreiecksmatrix $U \in \mathbb{R}^{n\times n}$ heißt ***LU*-Zerlegung** von A. Die Zerlegung heißt **normalisiert**, falls L normalisiert ist.

Satz 2.1.7. *Für jede reguläre Matrix $A \in \mathbb{R}^{n\times n}$ sind äquivalent:*

1. *Es existiert eine eindeutige normalisierte LU-Zerlegung.*
2. *Alle Untermatrizen $A_k = (a_{ij})_{(i,j)\in(1,\dots,k)^2}$ sind regulär.*

Beweis.

→ Ist A regulär, so sind auch L und U regulär. Damit sind von L und U alle Diagonaleinträge nicht null. Somit sind auch die Untermatrizen L_k und U_k regulär, somit auch die Untermatrizen $A_k = L_k U_k$.

← Für $n = 1$ ist die Aussage klar. Sei nun angenommen, die Aussage gelte für Matrizen der Größe $(n-1)\times(n-1)$. Damit existieren Matrizen L_{n-1}, U_{n-1}, sodass $A_{n-1} = L_{n-1}U_{n-1}$ eine normalisierte LU-Zerlegung ist. Seien nun $\binom{b}{a_{nn}}$ die letzte Spalte und $(c^\top, a_{nn})$ die letzte Zeile von A. Die Aussage ist bewiesen, wenn geeignete $l, u \in \mathbb{R}^{n-1}$ und $r \in \mathbb{R}$ existieren, sodass

$$\begin{aligned}
\begin{pmatrix} A_{n-1} & b \\ c^\top & a_{nn} \end{pmatrix} &= \begin{pmatrix} L_{n-1} & 0 \\ l^\top & 1 \end{pmatrix} \begin{pmatrix} U_{n-1} & u \\ 0 & r \end{pmatrix} \\
&= \begin{pmatrix} L_{n-1}U_{n-1} & L_{n-1}u \\ (U_{n-1}^\top l)^\top & l^\top u + r \end{pmatrix} \\
&= \begin{pmatrix} A_{n-1} & L_{n-1}u \\ (U_{n-1}^\top l)^\top & l^\top u + r \end{pmatrix}
\end{aligned}$$

Wir brauchen also $L_{n-1}u = b$, $U_{n-1}^\top l = c$, und $a_{nn} = l^\top u + r$. Durch Regularität von L_{n-1} und U_{n-1} existieren eindeutige Lösungen u, l, der ersten beiden Gleichungen, somit ist auch r festgelegt.

□

Korollar 2.1.8.

- *Jede positiv definite Matrix besitzt eine eindeutige LU-Zerlegung.*
- *Jede strikt diagonaldominante Matrix, also jede Matrix A mit $\sum_{j\in 1,\dots,n, i\neq j} |a_i j| < |a_{ii}|$ besitzt eine eindeutige LU-Zerlegung.*
- *Die Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ besitzt keine LU-Zerlegung.*

- *Die Nullmatrix besitzt zwar LU-Zerlegungen, diese sind aber nicht eindeutig.*

Lemma 2.1.9. *Falls $A = LU$ eine normalisierte LU-Zerlegung von A ist, so gilt*

$$a_{ik} = u_{ik} + \sum_{j=1}^{i-1} l_{ij} u_{jk}$$

und

$$a_{ki} = l_{ki} u_{ii} + \sum_{j=1}^{i-1} l_{kj} u_{ji}$$

> *Beweis.* Es gilt $l_{ij} = 0$ für $j > i$ und $l_{ii} = 1$. Es gilt außerdem $u_{ij} = 0$ für $j < i$. Die Formeln folgen direkt aus der Definition des Matrixprodukts. $\square$

Diese Formeln lassen sich für $i \leq k$ nach u_{ik} auflössen und für $k > i$ nach l_{ki} auflösen. Wir erhalten folgenden Algorithmus:

---

for $i = 1, i \leq n, i++$ **do**
  for $k = i, k \leq n, k++$ **do**
    $u_{ik} \leftarrow a_{ik} - \sum_{j=1}^{i-1} l_{ij} u_{jk}$
  end for
  for $k = i+1, k \leq n, k++$ **do**
    $l_{ki} \leftarrow \frac{1}{u_{ii}} \cdot \left(a_{ki} - \sum_{j=1}^{i-1} l_{kj} u_{ji} \right)$
  end for
end for

Proposition 2.1.10. *Der Rechenauftrag dieses Algorithmus beträgt $O(n^3)$.*

Proposition 2.1.11. *Es ist nicht mehr Speicher nötig, als sowieso für A benötigt wird. Die Einträge von A können im Speicher einfach sukzessiv durch die jeweiligen Einträge von L bzw. U ersetzt werden.*

Definition 2.1.12. Ein numerisches Problem φ heißt **schlecht Konditioniert**, wenn kleine Unterschiede in der Eingabe zu großen Unterschieden in der korrekten Lösung führen, also wenn

$$\frac{|\varphi(\tilde{x}) - \varphi(x)|}{|\varphi(x)|} >> \frac{|\tilde{x} - x|}{|x|}$$

Ansonsten heißt die Aufgabe **gut konditioniert**.

Definition 2.1.13. Ein Verfahren $\tilde{\varphi}$ heißt **numerisch instabil**, wenn eine Störung $\tilde{x}$ existiert, sodass der durch Rundungsfehler verursachte relative Fehler erheblich größer ist als der rein durch die Störung verursachte Fehler.

Beispiel 2.1.14. Sei

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{pmatrix}$$

Für $\varepsilon \in \mathbb{R}^+$. Es gilt

$$A^{-1} = \begin{pmatrix} 1+\frac{1}{\varepsilon} & -\frac{1}{\varepsilon} \\ -\frac{1}{\varepsilon} & \frac{1}{\varepsilon} \end{pmatrix}$$

$$A^{-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \qquad A^{-1}\begin{pmatrix} 1 \\ 1+\varepsilon \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Eine kleine Störung in den Daten b des linearen Gleichungssystems $Ax = b$ führt zu Problemen der Größenordnung 1!!! So können wir keine Numerik machen!!! Dieses Problem ist **schlecht konditioniert**.

Beispiel 2.1.15. Sei

$$A = \begin{pmatrix} \varepsilon & 1 \\ 1 & 0 \end{pmatrix},$$

also

$$A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -\varepsilon \end{pmatrix}.$$

So haben wir kein Problem bei der Berechnung von $A^{-1}b$ - die Aufgabe ist gut konditioniert. Sagen wir nun, wir versuchen, das Gleichungssystem effizient durch LU-Zerlegung zu lösen. Wir sehen, LU-Zerlegung von A ist jedoch gegeben durch

$$A = \begin{pmatrix} 1 & 0 \\ \frac{1}{\varepsilon} & 1 \end{pmatrix} \begin{pmatrix} \varepsilon & 1 \\ 0 & \frac{1}{\varepsilon} \end{pmatrix},$$

und die Berechnung von $L^{-1}b$ und $U^{-1}b$ führt nun wieder zu großen Rundungsfehlern. Aus unserer Idee entsteht also ein **instabiler Algorithmus.**

Satz 2.1.16. Cholesky-Zerlegung: *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit. So existiert eine eindeutige untere Dreiecksmatrix L, sodass*

$$A = LL^\top.$$

und $l_{ii} > 0$

Beweis. Für $n = 1$ ist die Suche durch $l_{11} = \sqrt{a_{11}}$ erledigt.

Die Untermatrix A_{n-1} ist immer ebenfalls positiv definit und symmetrisch. Sei also $A_{n-1} = L_{n-1}L_{n-1}^\top$. Wir setzen $\binom{b}{a_{nn}}^\top$ als die letzte Zeile von A. Dann müssen wir zum Beweis des Satzes einen Vektor $c \in \mathbb{R}^{n-1}$ und ein $\alpha \geq 0$ finden, sodass

$$\begin{pmatrix} A_{n-1} & b \\ b^\top & a_{nn} \end{pmatrix} = \begin{pmatrix} L_{n-1} & 0 \\ c^\top & \alpha \end{pmatrix} \begin{pmatrix} L_{n-1}^\top & c \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} A_{n-1} & L_{n-1}c \\ (L_{n-1} - c)^\top & \alpha^2 + c^\top c \end{pmatrix}$$

Dies ist nach Annahme äquivalent zu $L_{n-1}c = b$ und $c^\top c + \alpha^2 = a_{nn}$

Da L regulär ist existiert ein eindeutiges c, welches die erste Gleichung erfüllt. Es gilt:

$$\det A = \det \begin{pmatrix} L_{n-1} & 0 \\ c^\top & \alpha \end{pmatrix} \cdot \det \begin{pmatrix} L_{n-1}^\top & c \\ 0 & \alpha \end{pmatrix} = \alpha^2 (\det L_{n-1})^2$$

Da $\det A > 0$ und $\det L_{n-1} \geq 0$ bekommen wir $\alpha > 0$, sodass $c^\top c + \alpha^2 = a_{nn}$ ebenfalls eine eindeutige positive Lösung hat. □

Lemma 2.1.17. *Für $A = LL^\top$ gilt:*

$$a_{ik} = \begin{cases} l_{ik}l_{kk} + \sum_{j=i}^{k-1} l_{ij}l_{ki} & i > k \\ l_{kk}^2 + \sum_{j=1}^{k-1} l_{kj}^2 & i = k \end{cases}$$

Beweis. Matrixmultiplikation ohne triviale Summanden. □

Cholesky-Zerlegung:

for $k = 1, i \leq n, i++$ **do**
    $l_{kk} = \sqrt{a_{kk} - \sum_{j=1}^{k-1} l_{kj}^2}$
    for $i = k+1, i \leq n, i++$ **do**
        $l_{ik} = (a_{kk} - \sum_{j=1}^{k-1} l_{ij}l_{kj})\frac{1}{l_{kk}}$
    end for
end for

Proposition 2.1.18. *Der Aufwand ist wieder $\mathcal{O}(n^3)$, allerdings mit kleineren Konstanten.*

Proposition 2.1.19. *Lösung von $Ax = b$ für $A = LL^\top$ wie gehabt durch $Ly = b$ und $L^\top x = y$.*

2.1.3 Matrixnormen

Bekannt sind die üblichen Vektornormen auf $\mathbb{R}^n$, insbesondere

$$\|\vec{v}\|_p = \left(\sum_{j=1}^{n} |v_j|^p\right)^{\frac{1}{p}}$$

und

$$\|\vec{v}\|_\infty = \max v_j$$

Für $1 \leq p, q \leq \infty$ existiert eine Konstante c_{pqn}, sodass

$$\forall \vec{v} \in \mathbb{R}^n : \frac{1}{c_{pqn}} \|\vec{v}\|_p \leq \|\vec{v}\|_q \leq c_{pqn} \|\vec{v}\|_p$$

Definition 2.1.20. Für Normen $\|-\|_{\mathbb{R}^n}$ und $\|-\|_{\mathbb{R}^m}$ auf $\mathbb{R}^n$ und $\mathbb{R}^m$ definieren wir die Operatornorm auf $\hom(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{m \times n}$ als

$$\|A\|_{op} = \sup_{x \in \mathbb{R}^n : \|x\|_{\mathbb{R}^n} = 1} \|Ax\|_{\mathbb{R}^m}$$

Lemma 2.1.21. *Die Operatornorm ist eine Norm.*

Beweis.
1. Skalare können aus der inneren Norm und dem Supremum wie nötig herausgezogen werden.
2. Das Supremum ist über einer Menge positiver Zahlen, falls $x \neq \vec{0}$ gibt es mindestens einen Vektor größer 0.
3. Dreiecksungleichung folgt aus der Dreiecksungleichung für $\|-\|_{\mathbb{R}^m}$.

□

Lemma 2.1.22.

$$\|A\|_{op} = \inf\{c > 0 : \forall x \in \mathbb{R}^n \|Ax\| \leq c\|x\|\}$$

Lemma 2.1.23. *Für $A \neq 0$ und $x \in \mathbb{R}^n$ mit $\|x\| \leq 1$ und $\|Ax\| = \|A\|_{op}$ folgt $\|x\| = 1$*

Korollar 2.1.24. *Für alle $x \in \mathbb{R}^n$ gilt $\|Ax\| \leq \|A\|_{op} \|x\|$*

Lemma 2.1.25. *Es gibt Vektoren, sodass die Matrixnorm ihr* inf *und ihren* sup *annimmt.*

Beweis. Es handelt sich um eine stetige Funktion auf einem Kompaktum. □

Beispiel 2.1.26. 1. Die **Spaltensummennorm** $\|-\|_1$ ist eine Operatornorm:

$$\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{m} |a_{ij}|$$

2. Die **Zeilensummennorm** $\|-\|_\infty$ ist eine Operatornorm:

$$\|A\|_\infty = \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{ij}|$$

3. Die **Spektralnorm** $\|-\|_2$ ist eine Operatornorm:

$$\|A\|_2 = \rho(A^\top A) = (\max\{|\lambda| : \lambda \text{ ist Eigenwert von } A^\top A\})^{\frac{1}{2}}$$

Lemma 2.1.27. *Für $A \in \mathbb{R}^{l\times m}$, $B \in \mathbb{R}^{m\times n}$ und eine beliebige Operatornorm $\|-\|$ gilt $\|AB\| \leq \|A\|\|B\|$*

Beweis.

$$\begin{aligned} &\|ABx\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\| \\ \implies &\|AB\| \leq \|A\|\|B\| \end{aligned}$$

□

Lemma 2.1.28. *Falls die Normen in Bild und Urbild gleich sind, gilt*

$$\|E_n\| = 1$$

Lemma 2.1.29. *Falls die Normen in Bild und Urbild gleich sind, gilt für A symmetrisch mit Eigenwert λ*

$$\|A\| \geq |\lambda|$$

Beispiel 2.1.30. Die Frobeniusnorm $\|-\|_{\mathcal{F}}$ einer Matrix $A \in \mathbb{R}^{m\times n}$ ist gegeben durch

$$\|A\|_{\mathcal{F}} = \left(\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}$$

Lemma 2.1.31. *Für $n > 1$ ist die Frobeniusnorm keine Operatornorm!*

Beweis.

$$\|E_n\| = \sqrt{n}$$

Normieren wir die Norm, gilt

$$\frac{1}{\sqrt{2}}\left\|\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right\|_{\mathcal{F}} = \frac{1}{\sqrt{2}} < 1$$

Wobei 1 ein Eigenwert ist, was unseren vorherigen Lemmata widerspricht. □

2.1.4 Konditionszahl

Satz 2.1.32. *Sei $\|-\|$ eine Operatornorm auf $\mathbb{R}^{n\times n}$. Sei $A \in \mathbb{R}^{n\times n}$ regulär und seien $x, x', b, b' \in \mathbb{R}^n$, sodass $Ax = b$, $Ax' = b'$. Dann gilt:*

$$\frac{\|x - x'\|}{\|x\|} \leq \|A\|\left\|A^{-1}\right\|\frac{\|b - b'\|}{\|b\|}$$

Satz 2.1.33.

$$\|x - x'\| = \left\|A^{-1}(b - b')\right\| \leq \left\|A^{-1}\right\|\|b - b'\|$$

und

$$\|b\| = \|Ax\| \leq \|A\|\|x\|$$

Es folgt:

$$\frac{\|x - x'\|}{\|x\|} \leq \frac{\left\|A^{-1}\right\|\|b - b'\|}{\|x\|} \leq \frac{\left\|A^{-1}\right\|\|b - b'\|}{\|A^{-1}\|\|b\|}$$

Definition 2.1.34. Die **Konditionszahl** einer regulären Matrix $A \in \mathbb{R}^{n\times n}$ bezüglich der durch $\|-\|$ auf $\mathbb{R}^n$ induzierten Operatornorm ist gegeben durch:

$$\operatorname{cond}_{\|-\|}(A) = \|A\|\left\|A^{-1}\right\|$$

Wir schreiben oft cond_p statt $\operatorname{cond}_{\|-\|_p}$.

Lemma 2.1.35. $\operatorname{cond}(A) \geq 1$

Lemma 2.1.36. *Für A symmetrisch mit Eigenwerten λ_i gilt*

$$\operatorname{cond}_2(A) = \frac{\max|\lambda_j|}{\min|\lambda_j|}$$

Beispiel 2.1.37.

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{pmatrix}$$

Besitzt die Eigenwerte

$$\lambda_{1,2} = 1 + \frac{\varepsilon}{2} \pm \left(1 + \frac{\varepsilon^2}{4}\right)^{\frac{1}{2}}$$

Also $\lambda_1 \approx 2 + \frac{\varepsilon}{2}$ und $\lambda_2 \approx \frac{\varepsilon}{2}$. Für $\varepsilon \to 0$ geht also die Konditionszahl gegen unendlich.

Satz 2.1.38. *Für A symmetrisch und positiv definit mit Cholesky-Zerlegung $A = LL^\top$ gilt*

$$\operatorname{cond}_2(L) = \operatorname{cond}_2(L^\top) = \sqrt{\operatorname{cond}(A)}$$

Also kann das Problem, welches bei der LU-Zerlegung auftrat, bei der Cholesky-Zerlegung nicht vorkommen.

Beweis. Wir bemerken zunächst, dass $L^\top L$ und $LL^\top$ die selben Eigenwerte haben. Beide Matrizen sind symmetrisch und es gilt

$$\begin{aligned}&\forall x \in \mathbb{R}^n, \lambda \in \mathbb{R} : \\ &L^\top L x = \lambda x \Leftrightarrow LL^\top L x = \lambda(Lx) := \lambda y\end{aligned}$$

Somit folgt

$$\rho(LL^\top) = \rho(L^\top L)$$

und somit auch

$$\|L\|_2 = \left\|L^\top\right\|_2$$

analog gilt

$$\left\|L^{-1}\right\|_2 = \left\|L^{-\top}\right\|_2$$

Also $\operatorname{cond}_2(L) = \operatorname{cond}_2(L^\top)$. Da $LL^\top = A$, und A symmaterisch, folgt

$$\begin{aligned}\|L\|_2^2 &= \left\|L^\top\right\|_2^2 \\ &= \rho(LL^\top) \\ &= \rho(A) \\ &= \|A\|_2\end{aligned}$$

und

$$\begin{aligned}\left\|L^{-1}\right\|_2^2 &= \rho(L^{-\top}L^{-1}) \\ &= \rho(A^{-1}) \\ &= \left\|A^{-1}\right\|_2,\end{aligned}$$

also insgesamt

$$\begin{aligned}\operatorname{cond}_2(L) &= \|L\|_2\left\|L^{-1}\right\|_2 \\ &= \|A\|_2^{1/2}\left\|A^{-1}\right\|_2^{1/2} \\ &= \sqrt{\operatorname{cond}_2(A)}\end{aligned}$$

□

Chapter 3

Eliminationsverfahren

3.1 Gauss-Jordan-Elimination

Definition 3.1.1. **Gauss-Jordan-Elimination** Sei $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$.

1. Setze $A^{(1)} = A$, $b^{(1)} = b$, $k = 1$.

2. Für $A^{(k)}$ gelte für $1 \leq j \leq k-1$ und $i \geq j+1$ $a_{ij}^{k} = 0$, d.h. $A^{(k)}$ habe die Form

$$\begin{pmatrix} a_{11} & \dots & & & \dots & a_{1n} \\ & a_{22} & & & & \vdots \\ & & \ddots & & & \\ & & & a_{kk} & \dots & a_{kn} \\ & & & \vdots & & \vdots \\ & & & a_{nk} & \dots & a_{nn} \end{pmatrix}$$

   mit Nullen im unteren linken Teil.

3. Wir setzen $l_{ik} = \dfrac{a_{ik}}{a_{kk}^{(k)}}$ und definieren $L \in \mathbb{R}^{n \times n}$ als

$$L = \begin{pmatrix} 1 & \dots & & & \dots & 0 \\ & 1 & & & & \vdots \\ & & \ddots & & & \\ & & & 1 & \dots & 0 \\ & & & -l_{k+1,k} & \dots & 0 \\ \vdots & & & \vdots & & \vdots \\ 0 & \dots & & -l_{n,k} & \dots 0 \dots & 1 \end{pmatrix}$$

4. Setze $A^{(k+1)} = L^{(k)} A^{(k)}$, $b^{(k+1)} = L^{(k)} b^{(k)}$

5. Stoppe, falls $k+1 = n$, sonst erhöhe k und gehe zu Schritt 2.

Satz 3.1.2. *Ist $A \in \mathbb{R}^{n\times n}$ regulär, so ist Gauß-Jordan-Elimination genau dann durchführbar, wenn A eine LU-Zerlegung hat. Das Verfahren liefert dann die normierte LU-Zerlegung $U = A^{(n)}$ und $L = (L^{(n-1)} \cdot \ldots \cdot L^{(1)})^{-1}$. Die rechte Seite $y = b^{(n)}$ löst dann $y = L^{-1}b$ und die Lösung des Linearen Gleichungssystems $Ax = b$ ist gegeben durch die Lösung von $Ux = y$.*

3.2 Pivotsuche

Das Gaußverfahren ist für

$$A = \begin{pmatrix} \varepsilon & 1 \\ 1 & 1 \end{pmatrix}$$

zwar durchführbar, aber instabil. Wir führen deswegen eine sogenannte Pivotsuche durch - im k-ten Schritt bestimmen wir $p \in \{k, \ldots, n\}$, sodass

$$\left| a_{pk}^{(k)} \right| = \max a_{ik}^{(k)}$$

und vertauschen dann die Zeilen p und k in $A^{(k)}$ und $b^{(k)}$. Wir müssen diese Vertauschung jedoch nicht im Speicher tatsächlich durchführen, es reicht, einen Permutationsvektor $\pi \in \mathbb{N}^n$ vorzuschalten. Wir initialisieren π als $(1, 2, \ldots, n)^\top$, sollen daraufhin k und p vertauscht werden, vertauschen wir die jeweiligen Komponenten in π. Wollen wir daraufhin im Programm auf a_{ij} Zugreifen, müssen wir stattdessen auf $a_{\pi(i)j}$ zugreifen.

Satz 3.2.1. *Ist $A \in \mathbb{R}^{n\times n}$ regulär, $b \in \mathbb{R}^n$, so ist das Gaußverfahren mit Pivotsuche durchführbar und liefert die normalisierte LU-Zerlegung*

$$PA = LU$$

mit $|l_{ij}| \leq 1$ für alle i, j, sowie die modifizierte rechte Seite $b^{(n)} = L^{-1}Pb$, wobei

$$P = P^{(n-1)} P^{(n-2)} \ldots P^{(1)}$$

Chapter 4

Ausgleichsprobleme

Beispiel 4.0.1. Zu Messdaten $t_i, y_i, i = 1, \ldots, m$ wird die Ausgleichsgerade gesucht, also die Gerade definiert durch $c, b \in \mathbb{R}$, sodass der Least-Squares Abstand

$$\sum_{i=1}^{n} ((c \cdot t_i + b) - y_i)^2$$

zu den Messdaten minimiert wird.

Im Allgemeinen haben solche Probleme die Form

$$\min(x \mapsto \|Ax - b\|_2^2).$$

Ist A eine reguläre $n \times n$-Matrix, so wird das Problem eindeutig durch $x = A^{-1}b$ gelöst. Typischerweise ist dies aber nicht der Fall - in der Praxis sind die meisten Gleichungssysteme überbestimmt.

4.1 Die Gaußsche Normalengleichung

Definition 4.1.1. Durch $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ wird das Ausgleichsproblem

$$\min(x \mapsto \|Ax - b\|_2^2).$$

definiert. Für $x \in \mathbb{R}^n$ heißt

$$r = b - Ax$$

das **Residuum** von x.

Satz 4.1.2. *Die Lösungen des Ausgleichsproblems sind genau die Lösung der* ***Gaußschen Normalengleichung***

$$A^\top Ax = A^\top b.$$

Insbesondere existiert immer eine Lösung $x \in \mathbb{R}^n$. Ist $z \in \mathbb{R}^n$ eine weiter Lösung, so gilt $Ax = Az$ und die dazugehörigen Residuen stimmen überein.

Diese Normalengleichung erhält man durch Ableitung der Residuenfunktion $Ax - b$ nach x.

Beweis. Aus der linearen Algebra ist bekannt, dass

$$\mathbb{R}^m = \mathrm{im}(A) + \ker(A^\top),$$

wobei diese Zerlegung direkt und orthogonal ist. Damit existieren zu $b \in \mathbb{R}^m$ eindeutig bestimmte Vektoren $y \in \mathrm{im}(A)$, $r \in \ker(A^\top)$, sodass $y \cdot r = 0$ und $b = r + y$. Weiter existiert $x \in \mathbb{R}^n$ mit $y = Ax$.

Es folgt

$$A^\top b = A^\top y + A^\top r = A^\top Ax + 0 = A^\top Ax.$$

Somit löst x die Gaußsche Normalengleichung. Es bleibt zu zeigen, dass x auch das Ausgleichsproblem löst. Sei $z \in \mathbb{R}^n$. So rechnen wir

$$\begin{aligned}
\|b - Az\|_2^2 &= \|(b - Ax) + A(x - z)\|_2^2 \\
&= \|b - Ax\|_2^2 + \|A(x - z)\|_2^2 + 2r \cdot (Ax - z) \\
&= \|b - Ax\|_2^2 + \|A(x - z)\|_2^2 + \underbrace{2A^\top r}_{=0} \cdot (x - z) \\
&= \|b - Ax\|_2^2 + \|A(x - z)\|_2^2 \\
&\geq \|b - Ax\|_2^2
\end{aligned}$$

somit gilt insbesondere Gleichheit genau dann, wenn $Ax = Az$. □

Lemma 4.1.3. *Die Matrix $A^\top A$ ist symmetrisch und positiv semidefinit. Weiter ist A genau dann positiv definit, wenn $\ker A = \{0\}$. In diesem Fall sind die Lösungen der Gaußschen Normalengleichung eindeutig.*

Beweis. Symmetrie ist offensichtlich. Positive Semidefinitheit gilt, da

$$x(A^\top A)x = (Ax) \cdot (Ax) = \|Ax\|_2^2 \geq 0.$$

Insbesondere gilt also Gleichheit genau dann, wenn $Ax = 0$, also wenn $x \in \ker A$.

Die Eindeutigkeit der Lösung der Gaußschen Normalengleichung folgt aus der Regularität positiv definiter Matrizen. □

Für $m = n$, $A \in \mathbb{R}^{n\times n}$ gilt

$$\mathrm{cond}_2(A^\top A) = \frac{\lambda_{\max}(A^\top A)}{\lambda_{\min}(A^\top A)} = (\mathrm{cond}_2(A))^2$$

da $\mathrm{cond}_2(A) \geq 1$ ist somit $A^\top A$ immer schlechter konditioniert als A. Die Lösung eines Ausgleichsproblems durch die Gaußsche Normalengleichung ist somit instabil.

4.2 Householder-Matrizen

Sei $Q \in O(n)$ (also eine $n \times n$-Orthogonalmatrix). So gilt $\|Q(Ax-b)\|_2^2 = \|Ax-b\|$. Wir versuchen, eine orthogonale Matrix Q so zu konstruieren, dass QA Dreiecksgestalt hat.

Lemma 4.2.1. *Für alle Orthogonalen Matrizen Q gilt* $\mathrm{cond}_2(Q) = 1$

Beweis.

$$\|Q\|_2 \|Q^\top\|_2 = \|Q\|_2 \|Q^{-1}\|_2 = 1$$

□

Definition 4.2.2. Für $v \in \mathbb{R}^l$ mit $\|v\|_2 = 1$ heißt die Matrix

$$P_v = E_l - 2vv^\top$$

Die Householder-Transformation zu v.

$(vv^\top)x$ entspricht der Projektion von x auf den von v aufgespannten Vektorraum. Insgesamt spiegelt die Householder-Transformation also x an der Ursprungsebene orthogonal zu v.

Lemma 4.2.3. *P_v ist symmetrisch und orthogonal. Außerdem gilt $P_v v = -v$ und*

$$\forall w \in \mathbb{R}^l : wv = 0 \implies P_v w = w$$

Lemma 4.2.4. *Sei $x \in \mathbb{R}^l \neq 0, x \neq \lambda e_1$ und sei*

$$\sigma = \begin{cases} \mathrm{sgn}(x_1) & x_1 \neq 0 \\ 1 & \text{sonst} \end{cases}$$

Setzen wir nun

$$v = \frac{x + \sigma \|x\|_2 e_1}{\|x + \sigma \|x\|_2 e_1\|_2},$$

so gilt

$$P_v x = (E_l x - 2vv^\top)x = -\sigma \|x\|_2 e_1$$

Beweis. Da $x \neq \lambda e_1$ ist v wohldefiniert mit $\|v\|_2 = 1$. Weiter folgt

$$\begin{aligned} \|x + \sigma \|x\|_2 e_1\|_2^2 &= \|x\|^2 + 2\sigma \|x\|_2 x \cdot e_1 + \sigma^2 \|x\|_2^2 \|e_1\|_2^2 \\ &= 2(x + \sigma \|x\|_2 e_1)^\top x \end{aligned}$$

Mit $\tilde{v} = x + \sigma\|x\|_2 e_1$ gilt

$$\begin{aligned} 2\tilde{v}^\top x &= 2(x + \sigma\|x\|_2 e_1)^\top x \\ &= \|x + \sigma\|x\|_2 e_1\|_2^2 \\ &= \|\tilde{v}\|_2^2 \end{aligned}$$

Es gilt $v = \frac{\tilde{v}}{\|v\|_2}$, also

$$\begin{aligned} P_v x &= (E_l - 2vv^\top)x \\ &= x - 2v\frac{\tilde{v}^\top x}{\|\tilde{v}\|_2} \\ &= x - v\frac{\|\tilde{v}\|_2^2}{\|\tilde{v}\|_2} \\ &= x - v\|\tilde{v}\|_2 \\ &= -\sigma\|x\|_2 e_1 \end{aligned}$$

□

σ verhindert hier sogenannte Auslöschungseffekte, also schlechte Konditionierung der Subtraktion zweier fast identischer Zahlen.

4.3 QR-Zerlegung

Satz 4.3.1. *Sei $A \in \mathbb{R}^{m\times n}$, $m \geq n$,* rang *$A = n$. So existiert $Q \in O(m)$ und eine verallgemeinerte obere Dreiecksmatrix $R \in \mathbb{R}^{m\times n}$, sodass*

$$A = QR$$

Außerdem gilt $\forall i : |r|_{ii} > 0$

Beweis. Wir setzen $A_1 = A$, und es sei $x = a_1 \in \mathbb{R}^m$ die erste Spalte von A_1. Falls $x \in \mathbb{R}e_1$, setzen wir $Q_1 = E_m$. Ansonsten sei

$$Q_1 = P_v$$

mit v wie im Lemma. Es folgt

$$Q_1 a_1 = r_{11} e_1$$

mit $|r_{11}| = \|a_1\|_2 \geq 0$. Somit folgt

$$Q_1 A_1 = \begin{pmatrix} r_{11} & r_1^\top \\ \vec{0} & A_2 \end{pmatrix}$$

mit $A_2 \in \mathbb{R}^{(m-1)\times(n-1)}$. Wir setzen nun

$$\tilde{Q}_2 A_2 = \begin{pmatrix} r_{12} & r_2^\top \\ \vec{0} & A_3 \end{pmatrix}$$

und

$$Q_2 = \begin{pmatrix} 1 & \vec{0}^\top \\ \vec{0} & \tilde{Q}_2 \end{pmatrix}$$

Die Matrix Q_2 ist orthogonal, insbesondere ist sie die Householder-Matrix zu $v = \binom{0}{\tilde{v}}$ mit $\tilde{v}$, wobei $\tilde{v}$ der Vektor ist, der $\tilde{Q}_2$ als Householder-Matrix gibt.

Nach n solchen Schritten erhalten wir $QA := (Q_n Q_{n-1} \dots Q_1)A = R$. Da Q ein Produkt orthogonaler Matrizen ist gilt insbesonderee $Q \in O(m)$. Die Einträge erfüllen $r_{ii} = \|a_i\|_2 > 0$, da die Matrix A vollen Rang hat. □

Anmerkung 4.3.2. Im Fall $m = n$ ist die Faktorisierung abgesehen von Vorzeichen der Diagonaleinträge von R eindeutig, denn falls $A = QR = Q'R'$, so folgt mit

$$E := (Q')^{-1}Q = R'R^{-1}$$

Dass E eine orthogonale obere Dreiecksmatrix ist. Die orthogonalen Dreiecksmatrizen sind genau die Diagonalmatrizen mit Einträgen ± 1. Damit folgt aber $Q = Q'E$ und $R = ER'$.

Anmerkung 4.3.3. Für die Anwendung einer Householder-Transformation kann jede Matrixmultiplikation als Householder-Transformation dargestellt werden, was wesentlich schneller als allgemeine Matrixmultiplikation ist:

$$\begin{aligned} P_v A &= (E_m - 2vv^\top)A \\ &= A - 2v(v^\top A) \end{aligned}$$

Anmerkung 4.3.4. Die Vektoren v zu den Householder-Transformationen lassen sich in den frei werdenden Einträgen von A speichern. Weiter gilt

$$Q = \prod_{i=1}^{n} (E_m - 2v_i v_i^\top)$$

Anmerkung 4.3.5. Der Aufwand ist $\mathcal{O}(n^3)$.

4.4 Lösung des Ausgleichsproblems

Mithilfe der QR-Zerlegung bekommen wir ein stabiles Verfahren zur Lösung von Ausgleichsproblemen.

Satz 4.4.1. *Sei* $A \in \mathbb{R}^{m\times n}$*,* $m \geq n$*,* rang $A = n$*. Sei* $A = QR$ *und* $Q^\top b = \binom{c}{d}$*,* $Q^\top A = R = \binom{\hat{R}}{0}$ *mit* $c \in \mathbb{R}^n$*,* $d \in \mathbb{R}^{m-n}$*,* $\hat{R} = \mathbb{R}^{n\times n}$ *obere Dreiecksmatrix. So ist die Lösung des Anfangswertproblems gegeben durch*

$$\hat{R}x = c$$

Anmerkung 4.4.2. Da $\mathrm{cond}_2(Q) = 1$ folgt für reguläre $A \in \mathbb{R}^{n\times n}$ direkt

$$\mathrm{cond}_2(R) = \mathrm{cond}_2(A).$$

Somit ist insbesondere unser Algorithmus stabil.

4.5 Singulärwertzerlegung

Wir betrachten $A^\top A \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times n}$.

Dank Symmetrie ist $A^\top A$ diagonalisierbar, und es existiert eine Orthonormalbasis aus Eigenvektoren $v_1, \dots, v_n$ mit Eigenwerten $\lambda_1 \geq \dots \geq \lambda_p > \lambda_{p+1} = \dots = \lambda_n = 0$. Für $i \in 1, \dots p$ setzten wir

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i$$

Dann gilt für $i, j \in 1, \dots, p$, dass

$$\begin{aligned} u_i^\top u_j &= \frac{1}{\sqrt{\lambda_i \lambda_j}} (A v_i)^\top A v_j \\ &= \frac{1}{\sqrt{\lambda_i \lambda_j}} v_i^\top (A^\top A) v_j \\ &= \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} v_i^\top v_j \\ &= \delta_{ij} \end{aligned}$$

Der letzte Schritt folgt, da für $i \neq j$ alles sowieso 0 ist und für $i = j$ auch $\lambda_i = \lambda_j$ gilt und sich dann die Lambdas kürzen.

Die Vektoren $u_1, \dots, u_p$ bilden also eine Orthonormalbasis von $\operatorname{Im} A$. Wir ergänzen zu einer Orthonormalbasis $u_1, \dots, u_m$ des $\mathbb{R}^m$. Dann gilt

$$A^\top u_i = \frac{1}{\sqrt{\lambda_i}} A^\top A v_i = \sqrt{\lambda_i} v_i$$

Für $1 \leq i \leq p$. Für $i \geq p+1$ müssen die u_i im Kern von A liegen, also gilt die Gleichheit ebenfalls. Wir setzten $\sigma_i = \sqrt{\lambda_i}$ für $i = 1, \dots, p$ und bekommen:

Satz 4.5.1. **<u>Singulärwertzerlegung:</u>** *Sei* $A \in \mathbb{R}^{m\times n}$*. Dann existieren Zahlen* $\sigma_1 \geq \dots \geq \sigma_p$ *und Orthonormalbasen* $(u_i)_{i=1}^m$ *des* $\mathbb{R}^m$ *und* $(v_i)_{i=1}^n$ *des* $\mathbb{R}^n$*, sodass für alle* $1 \leq i \leq p$*:*

$$\begin{aligned} A v_i &= \sigma_i u_i \\ A^\top u_i &= \sigma_i v_i \end{aligned}$$

und für alle $p+1 \leq j \leq n$ *und* $p+1 \leq k \leq m$

$$\begin{aligned} Av_j &= 0 \\ A^\top u_k &= 0. \end{aligned}$$

Die Zahlen σ_i^2 *sind genau die von Null verschiedenen Eigenwerte von* $A^\top A$*. Für*

$$\begin{aligned} U &= (u_1, \dots, u_m) \in \mathbb{R}^{m \times m}, \\ V &= (u_1, \dots, u_n) \in \mathbb{R}^{n \times n} \end{aligned}$$

gilt $U \in O(m)$, $V \in O(n)$*. Ist* Σ *die Diagonalmatrix, die die* σ *in absteigender Reihenfolge enthält, gilt*

$$A = U \Sigma V^\top$$

Beweisskizze. Folgt direkt aus der Konstruktion :) "□"

4.6 Pseudoinverse

Definition 4.6.1. Ist $A = U\Sigma V^\top$ die Singulärwertzerlegung von A und $\Sigma^+ \in \mathbb{R}^{n \times m}$ gegeben durch Inversion der Einträge ungleich Null, dann heißt

$$A^+ = V \Sigma^+ U^\top = \sum_{i=1}^{p} \sigma_i^{-1} v_i u_i^\top$$

die **Pseudoinverse** oder **Moore-Penrose-Inverse** von A.

Anmerkung 4.6.2. Es gilt $\ker A^+ = \ker A^\top$ und $\operatorname{Im} A^+ = \operatorname{Im} A^\top$.

Anmerkung 4.6.3. Die Pseudoinverse ist die eindeutige Lösung des Gleichungssystems

$$\begin{aligned} AXA &= A, \\ XAX &= X, \\ (AX)^\top &= AX \\ (XY)^\top &= XA \end{aligned}$$

Satz 4.6.4. *Der Vektor* A^+b *löst das Ausgleichsproblem* $\min \|Ax - b\|_2^2$ *und ist von allen Lösungen diejenige mit minimaler euklidischer Norm.*

Beweis. Mit $A^+AA^+ = A^+$ und $\ker A^+ = (\operatorname{Im} A)^\perp$ folgt

$$AA^+b - d \in \ker A^+ = (\operatorname{Im} A)^\perp = \ker A^\top,$$

also

$$A^\top A(A^+b) = A^\top b,$$

damit ist aber A^+b eine Lösung der Gaußschen Normalengleichung.

Falls $z \in \mathbb{R}^n$ eine weitere Lösung ist, so gilt $\ker A^\top A = \ker A$, also

$$A^\top A(A^+b) - z = 0 \Leftrightarrow A(A^+b - z) = 0$$

Wir setzen nun $w = A^+b - z \in \ker A$. Es gilt $A^+b \in \operatorname{Im} A^+ = (\ker A)^\perp$, also folgt

$$(A^+b)w = 0$$

Für $z = A^+b - w$ gilt nun

$$\|z\|_2^2 = \left\|A^+b\right\|_2^2 + \|w\|_2^2,$$

da der gemischte Term $\|(A^+b)w\|_2^2$ wegfällt. Somit ist A^+b die Lösung mit minimaler Norm. $\square$

Chapter 5

Eigenwertaufgaben

Im Prinzip gibt es die Möglichkeit, Nullstellen des charakteristischen Polynoms zu suchen, zum Beispiel durch das sog. *Newton-Raphson-Verfahren.* Das ist aber typischerweise nicht praktikabel.

5.1 Abschätzungen

Satz 5.1.1. *Sei $A \in \mathbb{R}^{n \times n}$ und $\lambda \in \mathbb{C}$ ein Eigenwert von A. Dann gilt*

$$\lambda \in \bigcup_{i=1}^{n} K_i,$$

wobei

$$K_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=i, i \neq j}^{n} |a_{ij}| \right\}.$$

Die Mengen K_i heißen ***Gerschgorin-Kreise****.*

Beweis. Es sei $Ax = \lambda x$ für ein $x \in \mathbb{C}^n \setminus \{0\}$. Dann existiert ein maximaler Eintrag $i \in 1, \ldots, n$, also $|x_j| \leq |x_i|$ für alle $j \in 1, \ldots, n$ und $x_i \neq 0$. Dann gilt

$$\lambda x_i = (Ax)_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Wir teilen durch $x_i \neq 0$ und erhalten

$$\lambda - a_{ii} = \sum_{j=1, i \neq j}^{n} a_{ij} \frac{x_j}{x_i}.$$

Durch Dreiecksungleichung und $\frac{|x_j|}{x_i} \leq 1$ folgt $\lambda \in K_i$ und damit die Behauptung.

□

Satz 5.1.2. *Sei $a \in \mathbb{R}^{n\times n}$ symmetrisch, also alle Eigenwerte reell. Für den maximalen und den minimalen Eigenwert von A gilt dann:*

$$\lambda_{max} = \max_{x\in\mathbb{R}^n\setminus\{0\}} \frac{x^\top Ax}{\|x\|_2^2}$$

$$\lambda_{min} = \min_{x\in\mathbb{R}^n\setminus\{0\}} \frac{x^\top Ax}{\|x\|_2^2}$$

*Diese Brüche sind auch als die **Rayleigh-Quotienten** bekannt.*

Beweis. Sei $(v_i) \subset \mathbb{R}^n$ eine Orthonormalbasis des $\mathbb{R}^n$ aus Eigenvektoren zu den Eigenwerten $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \in \mathbb{R}$ der Matrix A. Wir schreiben $x \in \mathbb{R}^n$ als

$$x = \sum_{i=1}^{n} \alpha_i v_i$$

mit $\alpha_i \in \mathbb{R}$. So gilt

$$Ax = \sum_{i=1}^{n} \alpha_i \lambda_i v_i$$

Dank Orthonormalität der v_i folgt

$$x^\top x = \sum_{i,j=1}^{n} \alpha_i\alpha_j \langle v_i, v_j\rangle = \sum_{i=1}^{n} \alpha_i^2$$

und insbesondere

$$x^\top Ax = \sum_{i=1}^{n} \lambda_i \alpha_i^2$$

daraus folgt

$$\begin{aligned} x^\top Ax &\geq \lambda_n \sum_{i=1}^{n} \alpha_i^2 \\ &= \lambda_n \|x\|_2^2 \\ \implies \frac{x^\top Ax}{\|x\|_2^2} &\geq \lambda_n \end{aligned}$$

und

$$\begin{aligned} x^\top Ax &\leq \lambda_1 \sum_{i=1}^{n} \alpha_i^2 \\ &= \lambda_1 \|x\|_2^2 \\ \implies \frac{x^\top Ax}{\|x\|_2^2} &\leq \lambda_1 \end{aligned}$$

□

5.2 Konditionierung des Eigenwertproblems

Satz 5.2.1. *Sei $A \in \mathbb{R}^{n\times n}$ komplex diagonalisierbar mit $A = VDV^{-1}$. Sei $E \in \mathbb{R}^{n\times n}$ eine beliebige "Störungsmatrix" und sei $\bar{\lambda} \in \mathbb{C}$ ein Eigenwert von $A+E$. Dann existiert ein komplexer Eigenwert λ von A, sodass*

$$\left|\bar{\lambda} - \lambda\right| \leq \operatorname{cond}_2(V)\|E\|_2$$

Definition 5.2.2. Die Abschätzung lässt sich auch durch den sog. Hausdorff-Abstand schreiben. Für einen metrischen Raum M mit Metrik d und Teilmengen A, B ist dieser definiert als

$$d_H(A,B) := \max\left\{\sup_{x\in A}\inf_{y\in B} d(x,y), \sup_{y\in A}\inf_{x\in B} d(x,y)\right\}$$

Damit ist

$$d_H(\Sigma, \bar{\Sigma}) \leq \operatorname{cond}_2(V)\|E\|_2$$

wobei Σ die Menge der Eigenwerte (das *Spektrum*) von A ist und $\bar{\Sigma}$ das Spektrum von $A - E$.

Anmerkung 5.2.3. Nicht jede Matrix ist komplex diagonalisierbar.

Korollar 5.2.4. *Jede Matrix A, sodass $A^\top A = AA^\top$ (also jede sog. Normale Matrix) ist komplex diagonalisierbar. In diesem Fall ist V unitär, also insbesondere* $\operatorname{cond}(V)_2 = 1$. *Somit gilt sogar*

$$\left|\lambda - \bar{\lambda}\right| \leq \|E\|_2$$

Proposition 5.2.5. *Sei $p(t) = t^n + \sum_{i=0}^{n-1} a_i t^i$ ein normalisiertes Polynom. So gilt*

$$p(t) = (-1)^n \det(A - tE_n)$$

*mit der sogenannten **Frobenius-Begleitmatrix** A, deren Einträge überall 0 sind, außer auf der Subdiagonalen, wo sie 1 sind, und in der letzten Spalte, in der der i-te Eintrag (wobei wir bei 1 mit dem Zählen anfangen) genau $-a_{i-1}$ ist. Die komplexen Eigenwerte von A sind genau die komplexen Nullstellen von p.*

Korollar 5.2.6. *Da das Finden von Nullstellen eines Polynoms im Allgemeinen schlecht konditioniert ist, ist auch das Finden von Eigenwerten von Matrizen schlecht konditioniert.*

Beispiel 5.2.7. Das Polynom $p_\varepsilon(t) = (t-a)^n - \varepsilon$ besitzt die Nullstellen

$$\chi_k = a - \varepsilon^{\frac{1}{n}} e^{i2\pi\frac{k}{n}}$$

Die Polynome p_0, p_ε Unterscheiden sich nur im konstanten Koeffizienten, und für die dazugehörigen Begleitmatrizen A_0 und A_ε gilt

$$\|A_0 - A_\varepsilon\| = \varepsilon$$

die Nullstellen unterscheiden sich jedoch durch

$$\left|\lambda - \bar{\lambda}\right| = \varepsilon^{\frac{1}{n}},$$

der relative Fehler ist letzendlich

$$\frac{|\lambda - \lambda_k|}{|\lambda|} \sim \frac{\varepsilon^{\frac{1}{n}}}{\varepsilon}$$

und dieser Faktor wächst für $n > 1$ und $\varepsilon \to 0$ unbeschränkt.