

Funktionen symmetrischer Matrizen



Wissenschaftliche Arbeit
am Fachbereich Mathematik
der
Universität des Saarlandes
(Erstes Staatsexamen)

von

Thomas Michael Luxenburger

Matrikelnummer: 2521973

Saarbrücken, Januar 2013

Betreuer und Erstgutachter: Prof. Dr. Bernhard Burgeth
Zweitgutachter: Prof. Dr. Anselm Lambert

Inhaltsverzeichnis

1	Einleitung	3
2	Grundlagen	5
2.1	Grundbegriffe der Matrizen­theorie	5
2.2	Eigenwerte und Eigenvektoren	11
2.3	Diagonalisierbarkeit	14
2.4	Löwner-Ordnung	20
3	Visualisierungen symmetrischer Matrizen	23
3.1	Visualisierung positiv definit­er Matrizen als Ellipsen	23
3.2	Visualisierung als Punkte im \mathbb{R}^3	29
3.2.1	Lage von Matrizen gleicher Spur	29
3.2.2	Lage von Matrizen gleicher Determinante	30
3.2.3	Lage von Matrizen in Abhängigkeit ihrer Definitheit	31
3.2.4	Lage von Matrizen mit gleichen Eigenwerten	32
3.2.5	Lage von Matrizen mit gleichen Eigenvektoren	34
3.3	Visualisierung der Löwner-Ordnung	36
3.3.1	Visualisierung für positiv definite Matrizen über Ellipsen	36
3.3.2	Visualisierung im \mathbb{R}^3	39
4	Matrixfunktionen einer Variablen	42
4.1	Definition und Eigenschaften	43
4.2	Visualisierungen von Matrixfunktionen	47
4.2.1	Visualisierung von Matrixfunktionen in Abhängigkeit der Spur	47
4.2.2	Visualisierung von Matrixfunktionen in Abhängigkeit der Determinante	50
4.3	Polarzerlegung einer symmetrischen Matrix	52
4.3.1	Berechnung und Visualisierung	52
4.3.2	Einfluss von Matrixfunktionen auf die Polarzerlegung	56
4.4	Monotone Matrixfunktionen	57
5	Einfache Matrixfunktionen mehrerer Variablen	62
5.1	Matrixprodukte und parallele Summe	62
5.1.1	Jordan-Produkt	62
5.1.2	P-Produkt	64
5.1.3	Feynman-Produkt	66

5.1.4	Vergleich der Matrixprodukte	68
5.1.5	Parallele Summe	71
5.1.6	Monotoniebetrachtungen	72
5.2	Bestimmung von Supremum und Infimum symmetrischer Matrizen	73
5.2.1	Algorithmus zur Berechnung von Supremum und Infimum	73
5.2.2	Berechnung über Matrixfunktionen	78

Literaturverzeichnis		85
-----------------------------	--	-----------

Kapitel 1

Einleitung

Die Untersuchung von Matrixfunktionen begann mit der von Arthur Cayley im Jahre 1850 veröffentlichten Arbeit *A memoir on the Theory of Matrices* [5]. Cayley führte in dieser Arbeit neben der Addition und Multiplikation quadratischer Matrizen die Quadratwurzeln von 2×2 - und 3×3 -Matrizen ein. Eine erste allgemeine Definition von $f(A)$ wurde 1883 von James Joseph Sylvester in [15] formuliert. In den Folgejahren wurden von weiteren Mathematikern dazu äquivalente Definitionen entwickelt. Mit der Theorie von Matrixfunktionen beschäftigen sich mehrere Bücher, wobei insbesondere die Werke von Horn und Johnson [11] und Higham [9] zu nennen sind, die für das Verfassen dieser Arbeit als Quellen dienten. Die vorliegende Arbeit beschäftigt sich mit Funktionen symmetrischer Matrizen und ist wie folgt aufgebaut:

Zunächst werden wichtige Definitionen und Sätze der Linearen Algebra behandelt, die für diese Arbeit grundlegend sind. Im Vordergrund stehen dabei symmetrische Matrizen, ihre Eigenschaften und damit verwandte Begrifflichkeiten. In diesem Zusammenhang wird auch die Löwner-Ordnung betrachtet, die eine nicht vollständige Ordnungsrelation auf der Menge der symmetrischen Matrizen definiert.

Im darauf folgenden Kapitel werden verschiedene Möglichkeiten zur Visualisierung symmetrischer 2×2 -Matrizen erarbeitet. Für allgemeine symmetrische 2×2 -Matrizen betrachten wir eine Abbildung, welche die Matrizen auf Punkte im \mathbb{R}^3 abbildet. Für positiv definite Matrizen wird zudem eine Visualisierung als Ellipsen vorgestellt, die sich auch auf 3×3 -Matrizen in Form von Ellipsoiden übertragen lässt. Neben den Visualisierungen der Matrizen gehen wir auch der Frage nach, inwiefern die Löwner-Ordnung mit Hilfe dieser Darstellungsformen visuell repräsentiert werden kann. Die in diesem Kapitel behandelten Visualisierungen bilden die Grundlage für Visualisierungen im Hinblick auf Funktionen symmetrischer 2×2 -Matrizen.

Das vierte Kapitel widmet sich Funktionen symmetrischer Matrizen einer Variablen. Neben der Definition und wichtigen Eigenschaften von $f(A)$ werden verschiedene Visualisierungsmöglichkeiten für Funktionen symmetrischer 2×2 -Matrizen vorgestellt. Als Beispiel für eine Zerlegung symmetrischer Matrizen, die mit Hilfe von Matrixfunktionen berechnet wird, untersuchen wir die Polarzerlegung einer nicht singulären symmetrischen Matrix.

Das Kapitel schließt mit der Betrachtung von Monotonie im Hinblick auf Matrixfunktionen.

Im letzten Kapitel wenden wir uns Matrixfunktionen mehrerer Variablen anhand konkreter Beispiele zu. Darunter fallen verschiedene Matrixprodukte, welche die Symmetrie der verknüpften Matrizen erhalten, sowie die parallele Summe zweier positiv definiter Matrizen. Als weiteres Beispiel wird das Auffinden von Supremum und Infimum einer Menge symmetrischer Matrizen betrachtet. Dazu wird zunächst ein algorithmisches Vorgehen zur Bestimmung von Matrixsupremum und Matrixinfimum vorgestellt und anschließend untersucht, inwiefern Matrixfunktionen zur Berechnung von Supremum und Infimum symmetrischer Matrizen herangezogen werden können.

Es wurde versucht, die Arbeit in einer möglichst geschlossenen Form zu halten, sodass sämtliche Aussagen nach Möglichkeit hergeleitet und bewiesen wurden. An einigen Stellen war dies jedoch auf Grund des Umfangs nicht möglich, sodass dort auf externe Quellen verwiesen wird, in denen entsprechende Herleitungen zu finden sind. Zudem werden einige Standardresultate der Linearen Algebra als bekannt vorausgesetzt.

Danksagung

Ich möchte mich an dieser Stelle bei all jenen bedanken, die mich während meines Studiums unterstützt haben. Im Hinblick auf diese Arbeit gilt mein besonderer Dank Herrn Prof. Dr. Bernhard Burgeth für das interessante Thema und die hervorragende Betreuung. Er hat mit seinen Anregungen und Ideen zum Gelingen dieser Arbeit beigetragen und hat sich stets Zeit genommen, um auftretende Probleme zu diskutieren.

Kapitel 2

Grundlagen

In diesem Kapitel werden mathematische Grundlagen behandelt, die für den weiteren Verlauf der Arbeit relevant sind. Es werden wichtige Definitionen und Sätze der Linearen Algebra aufgegriffen, wobei der Schwerpunkt auf symmetrischen Matrizen liegt. Entsprechende Einführungen finden sich auch in [10], [14] oder [16].

2.1 Grundbegriffe der Matrizen­theorie

Definition 2.1.1. Es sei K ein Körper und $m, n \in \mathbb{N}$. Unter einer $m \times n$ -Matrix A über K versteht man ein rechteckiges Schema

$$A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

von Elementen $a_{ij} \in K$ mit m Zeilen und n Spalten. Dabei bezeichnet i den Zeilenindex und j den Spaltenindex. Das Matrixelement a_{ij} befindet sich in der i -ten Zeile und j -ten Spalte.

Wir bezeichnen im Folgenden die Menge aller $m \times n$ -Matrizen über einem Körper K mit $K^{m \times n}$. Für $(a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ schreiben wir $(a_{ij})_{mn}$.

Definition 2.1.2. Es seien $A = (a_{ij})_{mn}$, $B = (b_{ij})_{mn} \in K^{m \times n}$. Man definiert in $K^{m \times n}$

i) eine *Skalarmultiplikation* durch

$$\lambda A = \lambda(a_{ij})_{mn} := (\lambda a_{ij})_{mn} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix},$$

ii) eine *Addition* durch

$$A + B = (a_{ij})_{mn} + (b_{ij})_{mn} := (a_{ij} + b_{ij})_{mn} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Definition 2.1.3. Es sei $A = (a_{ij})_{mn} \in K^{m \times n}$ und $B = (b_{jk})_{nl} \in K^{n \times l}$. Dann ist das *Matrixprodukt* $C = AB = (c_{ik})_{ml}$ von A und B eine $m \times l$ -Matrix definiert durch

$$c_{ik} := \sum_{j=1}^n a_{ij} b_{jk}.$$

Das Matrixprodukt ist nur definiert, wenn die Spaltenanzahl von A mit der Zeilenanzahl von B übereinstimmt. Es ist assoziativ und distributiv, aber im Allgemeinen nicht kommutativ. Dazu betrachten wir die Matrizen

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \text{ und } B = \begin{pmatrix} 2 & 4 \\ 4 & 1 \end{pmatrix}.$$

Dann ist

$$AB = \begin{pmatrix} 10 & 13 \\ 10 & 6 \end{pmatrix} \neq \begin{pmatrix} 10 & 9 \\ 13 & 6 \end{pmatrix} = BA.$$

Definition 2.1.4. i) Es sei $A = (a_{ij})_{mn} \in \mathbb{R}^{m \times n}$. Die zu A *transponierte Matrix* A^\top ist definiert durch

$$A^\top = (a_{ij})_{mn}^\top = (a_{ji})_{nm} \in \mathbb{R}^{n \times m}.$$

Sie entsteht aus der Matrix A durch Vertauschen der Zeilen und Spalten von A .

ii) Es sei $A = (a_{ij})_{mn} \in \mathbb{C}^{m \times n}$. Die zu A *konjugierte Matrix* \bar{A} ist definiert durch

$$\bar{A} = \overline{(a_{ij})_{mn}} = (\bar{a}_{ij})_{mn} \in \mathbb{C}^{m \times n}.$$

Sie entsteht aus der Matrix A , indem alle Elemente von A konjugiert werden.

Ist für zwei Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ das Matrixprodukt definiert, so ist

$$(AB)^\top = B^\top A^\top. \tag{2.1}$$

Zum Beweis der Gültigkeit von (2.1) sei $(AB)^\top = (c_{ij})$ und $B^\top A^\top = (d_{ij})$. Der Eintrag c_{lm} der Matrix $(AB)^\top$ ist das Skalarprodukt der m -ten Zeile von A mit der l -ten Spalte von B . Für die Matrix $B^\top A^\top$ errechnet sich der Eintrag d_{lm} als Skalarprodukt der l -ten

Zeile von B^\top mit der m -ten Spalte von A^\top . Dies entspricht dem Skalarprodukt der l -ten Spalte von B mit der m -ten Zeile von A . Wir haben also

$$c_{lm} = \sum_i a_{mi} b_{il} \quad \text{und} \quad d_{lm} = \sum_i b_{il} a_{mi}.$$

Aus der Gleichheit $c_{lm} = d_{lm}$ folgt die Behauptung.

Definition 2.1.5. Eine Matrix $A \in K^{m \times n}$ heißt *quadratisch*, wenn $n = m$ ist, d.h. wenn die Anzahl der Spalten der Anzahl der Zeilen entspricht. In diesem Fall werden die Matrixelemente $a_{11}, a_{22}, \dots, a_{nn}$ als *Hauptdiagonale* (kurz: *Diagonale*) der Matrix A bezeichnet.

Definition 2.1.6. Die Summe der Hauptdiagonalelemente einer quadratischen Matrix $A = (a_{ij}) \in K^{n \times n}$ heißt *Spur* von A .

$$\text{Spur}(A) := \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}. \quad (2.2)$$

Quadratische Matrizen können nach ihren Eigenschaften in weitere Klassen untergliedert werden.

Definition 2.1.7. i) Eine quadratische Matrix $A \in K^{n \times n}$ heißt *Diagonalmatrix*, falls alle außerhalb der Hauptdiagonalen liegenden Elemente verschwinden:

$$a_{ij} = 0 \quad \text{für} \quad i \neq j \quad (i, j = 1, \dots, n).$$

ii) Eine Diagonalmatrix $A \in K^{n \times n}$ heißt *n-dimensionale Einheitsmatrix* I_n , falls

$$a_{ii} = 1 \quad \text{für} \quad i = 1, \dots, n.$$

iii) Eine quadratische Matrix $A = (a_{ij})_{nn} \in \mathbb{R}^{n \times n}$ heißt *symmetrisch*, falls

$$a_{ij} = a_{ji} \quad \text{für} \quad i, j = 1, \dots, n.$$

Eine symmetrische Matrix ist gleich ihrer Transponierten: $A = A^\top$.

iv) Eine quadratische Matrix $A = (a_{ij})_{nn} \in \mathbb{R}^{n \times n}$ heißt *schiefssymmetrisch*, falls

$$a_{ij} = -a_{ji} \quad \text{und} \quad a_{ii} = 0 \quad \text{für} \quad i, j = 1, \dots, n.$$

Schiefssymmetrische Matrizen sind Matrizen mit der Eigenschaft $A = -A^\top$.

Für die n -dimensionale Einheitsmatrix I_n schreiben wir im Folgenden auch I , wenn die Dimension aus dem Kontext klar erkennbar ist. Mit $\text{Sym}(n)$ bezeichnen wir die Menge aller symmetrischen $n \times n$ -Matrizen über \mathbb{R} .

Das Produkt symmetrischer Matrizen ist im Allgemeinen nicht symmetrisch. Betrachten wir die Matrizen

$$A = \begin{pmatrix} 1 & 5 \\ 5 & 3 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

so sind beide symmetrisch, ihr Produkt

$$AB = \begin{pmatrix} 7 & 6 \\ 13 & 8 \end{pmatrix}$$

ist jedoch nicht symmetrisch. Wir können zeigen, dass das Produkt zweier symmetrischer Matrizen $A, B \in \mathbb{R}^{n \times n}$ genau dann symmetrisch ist, wenn A und B bezüglich des Matrixproduktes kommutieren.

Satz 2.1.8. *Es seien $A, B \in \mathbb{R}^{n \times n}$ symmetrisch. Dann ist $AB = BA$ genau dann, wenn AB symmetrisch ist.*

Beweis. Ist das Produkt zweier symmetrischer Matrizen A und B kommutativ, so ist

$$(AB)^\top = B^\top A^\top = BA = AB$$

und damit ist AB symmetrisch. Ist umgekehrt AB symmetrisch, so zeigt die Betrachtung von

$$AB = (AB)^\top = B^\top A^\top = BA,$$

dass A und B bezüglich der Matrixmultiplikation kommutativ sind. \square

Proposition 2.1.9. *Jede quadratische Matrix $A \in \mathbb{R}^{n \times n}$ lässt sich eindeutig in eine Summe aus einer symmetrischen Matrix $A_s \in \mathbb{R}^{n \times n}$ und einer schiefsymmetrischen Matrix $A_{ss} \in \mathbb{R}^{n \times n}$ zerlegen gemäß*

$$A = A_s + A_{ss} \quad \text{mit} \quad A_s = \frac{1}{2}(A + A^\top), \quad A_{ss} = \frac{1}{2}(A - A^\top).$$

Beweis. Es ist $A = A_s + A_{ss}$. Die Matrix A_s ist symmetrisch, denn

$$A_s^\top = \frac{1}{2}(A^\top + (A^\top)^\top) = \frac{1}{2}(A + A^\top) = A_s.$$

Die Matrix A_{ss} ist schiefsymmetrisch, denn

$$A_{ss}^\top = \frac{1}{2}(A - A^\top)^\top = \frac{1}{2}(A^\top - (A^\top)^\top) = \frac{1}{2}(A^\top - A) = -\frac{1}{2}(A - A^\top) = -A_{ss}.$$

Zum Beweis der Eindeutigkeit betrachten wir eine weitere Zerlegung der Matrix A als $A = A'_s + A'_{ss}$ mit einer symmetrischen Matrix A'_s und einer schiefsymmetrischen Matrix A'_{ss} . Dann gilt

$$A_s = \frac{1}{2}(A + A^\top) = \frac{1}{2}(A'_s + A'_{ss}) + \frac{1}{2}(A'_s + A'_{ss})^\top = \frac{1}{2}(A'_s + A'_{ss}) + \frac{1}{2}(A'_s - A'_{ss}) = A'_s.$$

Ebenso zeigt man $A_{ss} = A'_{ss}$, womit die Eindeutigkeit bewiesen ist. \square

Definition 2.1.10. Eine Matrix $A \in K^{n \times n}$ heißt *invertierbar* oder *regulär*, falls es eine Matrix $B \in K^{n \times n}$ gibt mit

$$A \cdot B = B \cdot A = I_n.$$

B wird als die zu A *inverse Matrix* bezeichnet. Wir schreiben $B = A^{-1}$. Eine nicht invertierbare Matrix heißt *singulär*.

Sind A und B invertierbare Matrizen, deren Produkt definiert ist, so ist ihr Produkt AB ebenfalls invertierbar mit der Inversen

$$(AB)^{-1} = B^{-1}A^{-1},$$

denn unter Ausnutzung der Assoziativität des Matrixproduktes ist

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I.$$

Entsprechend folgt auch

$$(B^{-1}A^{-1})(AB) = I.$$

Definition 2.1.11. Eine quadratische Matrix $Q \in \mathbb{R}^{n \times n}$ heißt *orthogonal*, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

- a) $QQ^\top = Q^\top Q = I_n$,
- b) Q ist invertierbar und es gilt $Q^{-1} = Q^\top$.

Aus der Definition können wir entnehmen, dass die Skalarprodukte je zweier verschiedener Spalten oder Zeilen einer orthogonalen Matrix Null ergeben. Die Skalarprodukte jeder Spalte oder Zeile mit sich selbst ergeben eins. Zeilen- und Spaltenvektoren einer orthogonalen Matrix sind daher orthonormal.

Mit einer orthogonalen Matrix $Q \in \mathbb{R}^{n \times n}$ können wir durch Multiplikation mit einem Vektor $x \in \mathbb{R}^n$ eine orthogonale Koordinatentransformation durchführen. Diese hat die Eigenschaft, Längen, Abstände und Winkel zu erhalten. Es ist

$$|Qx| = |x|, \tag{2.3}$$

denn

$$(Qx)^\top(Qx) = x^\top Q^\top Qx = x^\top Ix = x^\top x.$$

Der Abstand $|Qx - Qy|$ entspricht daher dem Abstand zwischen x und y , denn

$$|Qx - Qy| = |Q(x - y)| = |x - y|. \quad (2.4)$$

Für den Winkel ϕ zwischen zwei Vektoren x und y gilt

$$\cos \phi = \frac{x^\top y}{|x||y|}.$$

Dies ist auch der Winkel zwischen Qx und Qy , denn

$$\frac{(Qx)^\top(Qy)}{|Qx||Qy|} = \frac{x^\top Q^\top Qy}{|x||y|} = \frac{x^\top y}{|x||y|} = \cos \phi. \quad (2.5)$$

Die Determinante $\det(A)$ einer $n \times n$ -Matrix wird induktiv definiert.

Definition 2.1.12. Es sei $A = (a_{ij}) \in K^{n \times n}$.

i) Für $n = 1$ ist

$$\det(A) := a_{11}.$$

ii) Für $n = 2$ ist

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{21}a_{12}.$$

iii) Für $n \geq 2$ ist

$$\det(A) := \sum_{k=1}^n (-1)^{k+1} \cdot a_{1k} \cdot \det(A_{1k}). \quad (2.6)$$

Dabei bezeichnet A_{1k} diejenige Matrix, die aus A durch Streichung der ersten Zeile und k -ten Spalte entsteht.

Aus (2.6) können wir folgern, dass aus einer Zeile oder Spalte ein gemeinsamer Faktor vor die Determinante gezogen werden kann. Damit ist für eine $n \times n$ -Matrix A

$$\det(mA) = m^n \det(A). \quad (2.7)$$

Es gilt der Determinantenmultiplikationssatz:

$$\det(AB) = \det(A) \det(B). \quad (2.8)$$

2.2 Eigenwerte und Eigenvektoren

Eigenwerte und Eigenvektoren gehören zu wichtigen Konzepten der Linearen Algebra und finden in vielen Bereichen Anwendung. Die Bestimmung von Eigenwerten und Eigenvektoren bildet eine Grundlage zur Berechnung von Funktionen symmetrischer Matrizen.

Definition 2.2.1. Es sei $A \in K^{n \times n}$. Ein Skalar $\lambda \in K$ heißt *Eigenwert* von A , falls es ein $v \in K^n$ gibt mit $v \neq 0$ und

$$Av = \lambda v. \quad (2.9)$$

Jedes solche v wird *Eigenvektor* zum Eigenwert λ genannt. Die Menge $\rho(A)$ aller Eigenwerte von A wird als *Spektrum* von A bezeichnet.

Definition 2.2.2. Es sei $A \in K^{n \times n}$. Die Abbildung

$$\chi_A: K \rightarrow K, \lambda \mapsto \det(A - \lambda I_n)$$

heißt *charakteristisches Polynom* von A .

Mit Hilfe des charakteristischen Polynoms lassen sich die Eigenwerte einer Matrix berechnen.

Satz 2.2.3. *Es sei $A \in K^{n \times n}$. Dann ist ein Skalar $\lambda \in K$ genau dann Eigenwert von A , wenn λ Nullstelle des charakteristischen Polynoms ist.*

$$\lambda \text{ ist Eigenwert von } A \iff \det(A - \lambda I_n) = 0. \quad (2.10)$$

Beweis. Wir schreiben die Eigenwertbedingung (2.9) als

$$(A - \lambda I)v = 0.$$

Dies ist ein lineares Gleichungssystem und hat genau dann eine nichttriviale Lösung, wenn $\det(A - \lambda I) = 0$ ist. Die Eigenwerte von A sind daher genau die λ , die $\det(A - \lambda I) = 0$ erfüllen. \square

Beispiel 2.2.4. Wir berechnen die Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}.$$

Die Eigenwerte λ_1 und λ_2 sind die Nullstellen des charakteristischen Polynoms

$$\chi_A(\lambda) = \det(A - \lambda I) = \lambda^2 - 20\lambda + 64 = (\lambda - 4)(\lambda - 16),$$

womit wir als Eigenwerte $\lambda_1 = 4$ und $\lambda_2 = 16$ erhalten. Den Eigenvektor v_1 zum Eigenwert λ_1 berechnen wir durch Lösen des Gleichungssystems

$$(A - \lambda_1 I)v_1 = 0 \iff \begin{pmatrix} 6 & 6 \\ 6 & 6 \end{pmatrix} v_1 = 0 \iff \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} v_1 = 0.$$

Damit ist $v_1 = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ mit $c_1 \in \mathbb{C}$. Durch analoge Rechnung erhalten wir $v_2 = c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ mit $c_2 \in \mathbb{C}$.

Determinante und Spur einer quadratischen Matrix lassen sich auch über ihre Eigenwerte berechnen.

Proposition 2.2.5. *Die Matrix $A \in K^{n \times n}$ habe die Eigenwerte $\lambda_1, \dots, \lambda_n$. Dann ist*

$$\det(A) = \prod_{i=1}^n \lambda_i \quad \text{und} \quad \text{Spur}(A) = \sum_{i=1}^n \lambda_i.$$

Beweis. Es ist

$$\det(A - \lambda I_n) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nn} & \cdots & \cdots & a_{nn} - \lambda \end{vmatrix} \quad (2.11)$$

$$= (a_{11} - \lambda) \begin{vmatrix} a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} + \dots \quad (2.12)$$

$$= (a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda) \cdots (a_{nn} - \lambda) + \dots \quad (2.13)$$

$$= (-\lambda)^n + (-\lambda)^{(n-1)}(a_{11} + a_{22} + \dots + a_{nn}) + \dots + c, \quad (2.14)$$

wobei c eine von λ unabhängig Konstante ist. Setzen wir $\lambda = 0$, ergibt die linke Seite von (2.11) die Determinante von A , in (2.14) bleibt c übrig. Also muss $c = \det(A)$ gelten. Damit haben wir

$$\det(A - \lambda I_n) = (-\lambda)^n + (-\lambda)^{(n-1)} \text{Spur}(A) + \dots + \det(A). \quad (2.15)$$

Da die Eigenwerte $\lambda_1, \dots, \lambda_n$ von A die Lösungen der Gleichung $\det(A - \lambda I) = 0$ sind, können wir $\det(A - \lambda I)$ als Polynom n -ten Grades ausdrücken durch

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) \cdots (\lambda_n - \lambda) \quad (2.16)$$

$$= (-\lambda)^n + (-\lambda)^{n-1}(\lambda_1 + \lambda_2 + \dots + \lambda_n) + \dots + (\lambda_1 \lambda_2 \cdots \lambda_n). \quad (2.17)$$

Der Vergleich von (2.15) mit (2.17) liefert die Behauptung. \square

Für symmetrische Matrizen lassen sich Aussagen über die Eigenwerte und Eigenvektoren treffen.

Satz 2.2.6. *Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann gilt:*

- i) *Alle Eigenwerte von A sind reell.*
- ii) *Eigenvektoren zu verschiedenen Eigenwerten von A sind zueinander orthogonal.*

Beweis. i) Es sei λ ein beliebiger Eigenwert der symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ und v ein Eigenvektor zu λ . Dann ist

$$\bar{\lambda} \bar{v}^\top v = (\overline{Av})^\top v = \bar{v}^\top \bar{A}^\top v = \bar{v}^\top Av = \bar{v}^\top (\lambda v) = \lambda \bar{v}^\top v.$$

Wir verwenden, dass $\bar{v}^\top v = |v|^2$ eine reelle Zahl ist, die als Betrag eines Eigenvektors von Null verschieden sein muss. Damit ist

$$\bar{\lambda} \bar{v}^\top v = \lambda \bar{v}^\top v \iff \bar{\lambda} = \lambda,$$

womit $\lambda \in \mathbb{R}$ gelten muss.

- ii) Es seien λ und μ verschiedene Eigenwerte von A , v ein Eigenvektor zu λ und u ein Eigenvektor zu μ . Dann ist

$$\lambda v^\top u = (Av)^\top u = v^\top A^\top u = v^\top \mu u = \mu v^\top u.$$

Es gilt damit

$$(\lambda - \mu)v^\top u = 0.$$

Da λ und μ verschieden sind, ist $\lambda - \mu \neq 0$, sodass $v^\top u = 0$ gelten muss. Die Vektoren u und v sind daher orthogonal. □

Ein wichtiger Begriff im Zusammenhang mit symmetrischen Matrizen ist die Definitheit. Sie gibt Auskunft über die Vorzeichen der Eigenwerte.

Definition 2.2.7. Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch mit den Eigenwerten $\lambda_1, \dots, \lambda_n$. A heißt

- i) *positiv definit*, wenn $\lambda_i > 0$ für $i = 1, \dots, n$,
- ii) *positiv semidefinit*, wenn $\lambda_i \geq 0$ für $i = 1, \dots, n$,
- iii) *negativ definit*, wenn $\lambda_i < 0$ für $i = 1, \dots, n$,
- iv) *negativ semidefinit*, wenn $\lambda_i \leq 0$ für $i = 1, \dots, n$,
- v) *indefinit*, wenn $j, k \in \{1, \dots, n\}$ existieren mit $\lambda_j < 0$ und $\lambda_k > 0$.

Die Menge der positiv (semi)definiten Matrizen bezeichnen wir mit $\text{Sym}^+(n)$, die Menge der negativ (semi)definiten Matrizen mit $\text{Sym}^-(n)$.

2.3 Diagonalisierbarkeit

Diagonalisierbarkeit von Matrizen kann zur Berechnung von Matrixfunktionen herangezogen werden. Bevor wir zur Definition von Diagonalisierbarkeit kommen, betrachten wir ähnliche Matrizen. Der Begriff der Ähnlichkeit zweier Matrizen steht im Zusammenhang mit dem Diagonalisierbarkeitsbegriff.

Definition 2.3.1. Zwei Matrizen $A, B \in K^{n \times n}$ heißen *ähnlich*, wenn es eine invertierbare Matrix $T \in K^{n \times n}$ gibt mit

$$B = TAT^{-1}.$$

Man schreibt $A \approx B$.

Definition 2.3.2. Es seien M_1 und M_2 Mengen. Eine *Relation* R zwischen M_1 und M_2 ist eine Teilmenge R von $M_1 \times M_2$. Statt $(a, b) \in R$ schreibt man auch aRb .

Definition 2.3.3. Es sei M eine Menge. Eine Relation auf M heißt

- i) *reflexiv*, falls xRx für alle $x \in M$,
- ii) *symmetrisch*, falls für alle $x, y \in M$ mit xRy auch yRx gilt,
- iii) *antisymmetrisch*, falls für alle $x, y \in M$ mit xRy und yRx stets $x = y$ folgt,
- iv) *transitiv*, falls für alle $x, y \in M$ mit xRy und yRz auch xRz gilt,
- v) *konnex*, falls xRy oder yRx für alle $x, y \in M$ gilt.

Eine reflexive, symmetrische und transitive Relation heißt *Äquivalenzrelation*. Unter einer *Ordnungsrelation* versteht man eine Relation, die reflexiv, antisymmetrisch und transitiv ist. Ist eine Ordnungsrelation zudem konnex, heißt sie *vollständige Ordnung* oder *Totalordnung*.

Proposition 2.3.4. Die Ähnlichkeit von zwei Matrizen nach Definition 2.3.1 ist eine Äquivalenzrelation auf der Menge der $n \times n$ -Matrizen über K .

Beweis. i) *Reflexivität:* $A \approx A$, denn $A = IAI^{-1}$. ii) *Symmetrie:* Es sei $A \approx B$, also $B = TAT^{-1}$. Dann ist $T^{-1}BT = A$, also $A = T'BT'^{-1}$ mit $T' := T^{-1}$. Damit gilt $B \approx A$. iii) *Transitivität:* Es gelte $A \approx B$ und $B \approx C$, also $B = TAT^{-1}$ und $C = VBV^{-1}$. Dann ist $C = (VT)A(VT)^{-1}$, also $A \approx C$. \square

Proposition 2.3.5. *Es seien $A, B \in K^{n \times n}$ ähnliche Matrizen. Dann gilt:*

- i) $\rho(A) = \rho(B)$,*
- ii) $\text{Spur}(A) = \text{Spur}(B)$,*
- iii) $\det(A) = \det(B)$.*

Beweis. i) Zwei ähnlich Matrizen haben das gleiche charakteristische Polynom, denn

$$\begin{aligned} \chi_{TAT^{-1}}(\lambda) &= \det(TAT^{-1} - \lambda I) \\ &= \det(T(A - \lambda I)T^{-1}) \\ &= \det(T) \det(A - \lambda I) \det(T^{-1}) \\ &= \det(A - \lambda I) \\ &= \chi_A(\lambda). \end{aligned}$$

Damit haben sie dieselben Eigenwerte. Die Aussagen ii) und iii) folgen aus i) und Proposition 2.2.5, □

Definition 2.3.6. Eine Matrix $A \in K^{n \times n}$ heißt *diagonalisierbar*, wenn sie ähnlich zu einer Diagonalmatrix ist.

Im Bezug auf Diagonalisierbarkeit nehmen symmetrische Matrizen eine besondere Rolle ein. Sie sind stets diagonalisierbar. Da die Diagonalisierung symmetrischer Matrizen immer durch eine orthogonale Matrix möglich ist, bezeichnet man sie auch als *orthogonal diagonalisierbar*.

Satz 2.3.7. *Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann ist A orthogonal diagonalisierbar, d.h. es existiert eine orthogonale Matrix $S \in \mathbb{R}^{n \times n}$ mit*

$$A = S \text{diag}(\lambda_1, \dots, \lambda_n) S^\top,$$

wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A sind. Die Spaltenvektoren von S sind die orthogonalen Eigenvektoren von A .

Beweis. Wir beweisen die orthogonale Diagonalisierbarkeit einer symmetrischen $n \times n$ -Matrix durch vollständige Induktion nach n . Ein entsprechendes Vorgehen zum Beweis findet sich in [7].

Induktionsanfang

Eine Matrix $A \in \mathbb{R}^{1 \times 1}$ hat die Gestalt $A = (\lambda) = I(\lambda)I^\top$ und ist damit orthogonal diagonalisierbar.

Induktionsvoraussetzung

Es sei bewiesen, dass eine symmetrische $n \times n$ -Matrix für ein $n \in \mathbb{N}$ orthogonal diagonalisierbar ist.

Induktionsschluss

Wir zeigen, dass die Induktionsvoraussetzung die orthogonale Diagonalisierbarkeit einer symmetrischen $(n+1) \times (n+1)$ -Matrix impliziert. Sei dazu $A \in \mathbb{R}^{(n+1) \times (n+1)}$ und λ ein Eigenwert von A mit zugehörigem normiertem Eigenvektor u_1 . Wir setzen $U := \mathbb{R}u_1$. Dann ist $AU := \{Ax : x \in U\} \subset U$. Wir zeigen, dass auch $AU^\perp := \{Ax : x \in U^\perp\} \subset U^\perp$ ist. Es sei $x \in U^\perp$. Dann ist

$$u_1^\top Ax = u_1^\top A^\top x = (Au_1)^\top x = (\lambda u_1)^\top x = \lambda \underbrace{u_1^\top x}_{=0} = 0,$$

sodass $AU^\perp \subset U^\perp$. Es sei $\{u_2, \dots, u_n\}$ eine Orthonormalbasis von U^\perp . Dann bildet $\{u_1, u_2, \dots, u_n\}$ eine Orthonormalbasis von \mathbb{R}^n und die Matrix $B := (u_1, u_2, \dots, u_n)$, deren Spalten aus den Vektoren u_1, \dots, u_n bestehen, ist orthogonal. Es sei $M := B^\top AB$. Die Matrix M ist symmetrisch, denn

$$M^\top = B^\top A^\top (B^\top)^\top = B^\top AB = M.$$

Da $AU^\perp \subset U^\perp$, liegen für $i = 2, \dots, n$ die Vektoren $w_i := Au_i$ in U^\perp . Damit haben wir

$$\begin{aligned} M &= \begin{pmatrix} u_1^\top \\ \vdots \\ u_n^\top \end{pmatrix} A(u_1, \dots, u_n) = \begin{pmatrix} u_1^\top \\ \vdots \\ u_n^\top \end{pmatrix} (\lambda u_1, w_2, \dots, w_n) \\ &= \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array} \right), \end{aligned}$$

wobei $A_1 \in \mathbb{R}^{n \times n}$. Mit M ist auch A_1 symmetrisch und als $n \times n$ -Matrix nach Induktionsanfang orthogonal diagonalisierbar. Es existiert eine orthogonale Matrix S_1 , sodass

$$S_1^\top A_1 S_1 = D_1,$$

wobei D_1 diagonal ist. Wir setzen

$$S := B \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array} \right).$$

Da wir aus der Orthogonalität von S_1 die Orthogonalität von S folgern können, erhalten

wir

$$\begin{aligned}
 S^\top AS &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1^\top \end{array} \right) B^\top AB \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array} \right) \\
 &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1^\top \end{array} \right) \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & A_1 \end{array} \right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array} \right) \\
 &= \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & S_1^\top A_1 S_1 \end{array} \right) \\
 &= D,
 \end{aligned}$$

wobei D diagonal ist. Damit ist die orthogonale Diagonalisierbarkeit bewiesen.

Die Diagonalmatrix D besteht aus den Eigenwerten von A , denn A ist ähnlich zu D und Proposition 2.3.5 zu Folge haben ähnliche Matrizen die gleichen Eigenwerte.

Die Spaltenvektoren v_1, \dots, v_n von S sind die Eigenvektoren von A sind. Denn ist

$$A = S \operatorname{diag}(\lambda_1, \dots, \lambda_n) S^\top,$$

können wir A schreiben als

$$A = \lambda_1 v_1 v_1^\top + \dots + \lambda_n v_n v_n^\top.$$

Da die Spaltenvektoren einer orthogonalen Matrix normiert und zueinander orthogonal sind, erhalten wir für den Spaltenvektor v_k der Matrix S

$$Av_k = \sum_{i=1}^n \lambda_i v_i v_i^\top v_k = \lambda_k v_k.$$

Damit ist v_k Eigenvektor zum Eigenwert λ_k ist und die geforderte Eigenwertzerlegung bewiesen. \square

Beispiel 2.3.8. Wir diagonalisieren die Matrix

$$A = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}.$$

Dazu verwenden wir die in Beispiel 2.2.4 errechneten Eigenwerte $\lambda_1 = 4$ und $\lambda_2 = 16$ und normieren die zugehörigen Eigenvektoren zu

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{und} \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Mit der so gefundenen orthogonalen Matrix

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

erhalten wir

$$\begin{aligned} S^T A S &= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 16 \\ -4 & 16 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 8 & 0 \\ 0 & 32 \end{pmatrix} \\ &= \text{diag}(4, 16), \end{aligned}$$

sodass wir mit $S^T A S$ eine orthogonale Diagonalisierung von A gefunden haben. Wir können daher A schreiben als $A = S \text{diag}(4, 16) S^T$.

Mit Hilfe der Diagonalisierbarkeit einer symmetrischen Matrix können wir zeigen, dass die Eigenwerte 1 oder -1 für eine symmetrische Orthogonalmatrix charakteristisch sind.

Proposition 2.3.9. *Eine symmetrische Matrix ist genau dann orthogonal, wenn ihre Eigenwerte 1 oder -1 sind.*

Beweis. Es sei $Q \in \mathbb{R}^{n \times n}$ symmetrisch und orthogonal, v ein Eigenvektor von Q zum Eigenwert λ . Es ist

$$(Qv)^T Qv = v^T Q^T Qv = v^T v$$

und

$$(Qv)^T Qv = (\lambda v)^T \lambda v = \lambda^2 v^T v,$$

sodass

$$\lambda^2 v^T v = v^T v \tag{2.18}$$

gilt. Nun ist $v^T v = \sum_{i=1}^n |v_i|^2$ als Summe reeller Zahlen ebenfalls reell. Da $v \neq 0$ ist, folgt $v^T v = \sum_{i=1}^n |v_i|^2 > 0$. Wir können daher (2.18) kürzen zu $\lambda^2 = 1$. Da Q symmetrisch ist, müssen nach Satz 2.2.6 alle Eigenwerte von Q reell sein. Damit bleiben 1 und -1 als mögliche Eigenwerte.

Ist umgekehrt eine Matrix $U \in \mathbb{R}^{n \times n}$ symmetrisch mit der Eigenwertzerlegung $U = S \text{diag}(\lambda_1, \dots, \lambda_n) S^T$ und hat außer 1 oder -1 keine weiteren Eigenwerte, dann ist $\lambda_i^2 = 1$. Damit ist $U U^T = S \text{diag}(\lambda_1^2, \dots, \lambda_n^2) S^T = I$, sodass U orthogonal ist. \square

Aus Satz 2.3.7 können wir folgern, dass wir aus den Eigenvektoren einer symmetrischen $n \times n$ -Matrix eine Orthonormalbasis des \mathbb{R}^n konstruieren können. Damit lässt sich eine zu Definition 2.2.7 äquivalente Charakterisierung der Definitheit angeben. Dazu betrachten wir quadratische Formen. Ist $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix, so heißt

$$x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j \tag{2.19}$$

quadratische Form.

Satz 2.3.10. *Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und $x \in \mathbb{R}^n$. A ist genau dann*

- i) positiv definit, wenn $x^\top Ax > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$,*
- ii) positiv semidefinit, wenn $x^\top Ax \geq 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$,*
- iii) negativ definit, wenn $x^\top Ax < 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$,*
- iv) negativ semidefinit, wenn $x^\top Ax \leq 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.*

Beweis. Wir beweisen nur den Fall (i). Die Aussagen (ii)-(iv) werden analog bewiesen. Wir nehmen an, es gelte $x^\top Ax > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ und A sei nicht positiv definit. Dann existiert ein Eigenwert μ von A mit $\mu \leq 0$ und zugehörigem Eigenvektor $v \neq 0$. Damit steht aber

$$v^\top Av = v^\top \mu v = \mu v^\top v = \mu |v|^2 \leq 0$$

im Widerspruch zu $x^\top Ax > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$. Also sind alle Eigenwerte von A positiv.

Es sei nun $A \in \mathbb{R}^{n \times n}$ positiv definit mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ und v_1, \dots, v_n die zugehörigen normierten Eigenvektoren, die eine Orthonormalbasis des \mathbb{R}^n bilden. Wir können jeden Vektor $x \in \mathbb{R}^n$ als Linearkombination $x = \sum_{i=1}^n \alpha_i v_i$ schreiben mit $\alpha_i \in \mathbb{R}$. Für $x \neq 0$ ist dann

$$\begin{aligned} x^\top Ax &= \left(\sum_{i=1}^n \alpha_i v_i \right)^\top A \left(\sum_{j=1}^n \alpha_j v_j \right) = \left(\sum_{i=1}^n \alpha_i v_i \right)^\top \left(\sum_{j=1}^n \alpha_j Av_j \right) \\ &= \sum_{i,j=1}^n \alpha_i \alpha_j \lambda_j v_i^\top v_j = \sum_{i,j=1}^n \alpha_i \alpha_j \lambda_j \delta_{ij} \\ &= \sum_{i=1}^n \alpha_i^2 \lambda_i > 0. \end{aligned}$$

□

2.4 Löwner-Ordnung

Auf der Menge der reellen Zahl ist durch " \leq " eine vollständige Ordnungsrelation gegeben. Für $a, b \in \mathbb{R}$ ist $a \leq b$ genau dann, wenn $b - a$ nicht negativ ist. Auch auf der Menge der symmetrischen Matrizen können wir in Analogie dazu eine Ordnungsrelation definieren. Diese geht auf Karl Löwner zurück, der sie 1934 in [12] einführte. Sie wird ihm zu Ehren als *Löwner-Ordnung* bezeichnet. Es handelt sich dabei jedoch nicht um eine vollständige Ordnung. Mit Eigenschaften der Löwner-Ordnung beschäftigt sich unter anderem [10] ausführlich.

Definition 2.4.1 (Löwner-Ordnung). Es seien $A, B \in \mathbb{R}^{n \times n}$ symmetrische Matrizen. Dann ist $A \preceq B$ genau dann, wenn $B - A$ positiv semidefinit ist. Es ist $A \prec B$ genau dann, wenn $B - A$ positiv definit ist.

In Analogie zur Ordnung der reellen Zahlen sind die Schreibweisen $A \succeq B$ und $B \preceq A$ äquivalent. Ebenso können wir für $A \succ B$ auch $B \prec A$ schreiben.

Beispiel 2.4.2. Es seien

$$A = \begin{pmatrix} 8 & 2 \\ 2 & 13 \end{pmatrix} \text{ und } B = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}.$$

Die Matrix

$$A - B = \begin{pmatrix} 4 & 0 \\ 0 & 11 \end{pmatrix}$$

besitzt die Eigenwerte $\lambda_1 = 4$ und $\lambda_2 = 11$. Da $\lambda_1 \geq 0$ und $\lambda_2 \geq 0$, ist $B \preceq A$. Da beide Eigenwerte strikt positiv sind, gilt hier sogar $B \prec A$.

Proposition 2.4.3. *Durch die in Definition 2.4.1 definierte Relation " \preceq " wird eine Ordnungsrelation definiert.*

Beweis. i) *Reflexivität.* Für alle Matrizen $A \in \mathbb{R}^{n \times n}$ ist $A \preceq A$, da $A - A = 0$ und die Nullmatrix positiv semidefinit ist.

ii) *Antisymmetrie.* Es seien $A, B \in \mathbb{R}^{n \times n}$ symmetrisch mit $A \preceq B$ und $B \preceq A$. Nach Definition sind dann die Matrizen $B - A$ und $A - B$ positiv semidefinit, ihre Eigenwerte größer oder gleich Null. Wir zeigen, dass unter der Voraussetzung $A \preceq B$ und $B \preceq A$ alle Eigenwerte von $A - B$ Null sein müssen und folgern daraus, dass die Matrizen A und B identisch sind. Sei dazu λ ein beliebiger Eigenwert von $A - B$. Dann gilt:

$$\det((A - B) - \lambda I) = 0.$$

Unter Verwendung der nach (2.7) gültigen Formel $\det(-A) = (-1)^n \det(A)$ zeigt die Rechnung

$$0 = \det((A - B) - \lambda I) = \det((-1)((B - A) - (-\lambda)I)) = (-1)^n \det((B - A) - (-\lambda)I),$$

dass dann auch $\det((B - A) - (-\lambda)I) = 0$ ist. Daher ist $-\lambda$ Eigenwert von $B - A$. Nach Voraussetzung ist $-\lambda \geq 0$. Als Eigenwert einer positiv semidefiniten Matrix ist aber auch $\lambda \geq 0$, sodass $\lambda = 0$ gelten muss. Die Matrix $A - B$ ist als Differenz zweier symmetrischer Matrizen symmetrisch und lässt sich daher mit einer orthogonalen Matrix S als $A - B = SDS^T$ schreiben, wobei die Diagonalmatrix D aus den Eigenwerten von $A - B$ besteht, die alle Null sind. Damit ist D die Nullmatrix und $A - B = SDS^T = 0$, woraus die Gleichheit $A = B$ folgt.

- iii) *Transitivität.* Es seien $A, B, C \in \mathbb{R}^{n \times n}$ symmetrisch mit $A \preceq B$ und $B \preceq C$. Somit sind die Matrizen $B - A$ und $C - B$ positiv semidefinit, sodass für alle $x \in \mathbb{R}^n \setminus \{0\}$ sowohl $x^T(B - A)x \geq 0$, als auch $x^T(C - B)x \geq 0$ sind. Damit ist auch für alle $x \in \mathbb{R}^n \setminus \{0\}$

$$x^T(C - A)x = x^T((C - B) + (B - A))x = \underbrace{x^T(C - B)x}_{\geq 0} + \underbrace{x^T(B - A)x}_{\geq 0} \geq 0.$$

Dies zeigt, dass $C - A$ positiv semidefinit ist und somit $A \preceq C$ gilt. □

Wie oben erwähnt, ist die Löwner-Ordnung keine vollständige Ordnung, d.h. es existieren Matrizen $A, B \in \mathbb{R}^{n \times n}$, die nicht vergleichbar sind. Für sie gilt weder $A \succeq B$ noch $B \succeq A$.

Beispiel 2.4.4. Betrachte die Matrizen

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}.$$

Die Matrix

$$A - B = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$$

hat die Eigenwerte -1 und 3 und ist damit indefinit. Die Matrizen A und B sind in der Löwner-Ordnung nicht vergleichbar.

Die Löwner-Ordnung ist auch in weiteren Punkten nicht mit der gewöhnlichen Ordnung der reellen Zahlen vergleichbar. So folgt aus $A \succeq B$ und $A \neq B$ im Allgemeinen nicht, dass $A \succ B$ ist.

Beispiel 2.4.5. Für die Matrizen

$$A = \begin{pmatrix} 8 & 2 \\ 2 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

ist $B \preceq A$, denn die Eigenwerte von

$$A - B = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$$

sind 0 und 3 und damit nicht negativ. Die Aussage $B \prec A$ ist jedoch falsch, da einer der Eigenwerte 0 ist. Dennoch sind die Matrizen A und B verschieden.

Lemma 2.4.6. *Es seien $A, B \in \mathbb{R}^{n \times n}$ symmetrisch, A positiv definit und B positiv semidefinit. Dann ist $B \preceq A$ genau dann, wenn alle Eigenwerte von BA^{-1} kleiner oder gleich 1 sind.*

Beweis. Ist A positiv definit und B positiv semidefinit, können wir eine reguläre Matrix $C \in \mathbb{R}^{n \times n}$ finden, sodass $A = C I C^\top$ und $B = C D C^\top$, wobei $D = \text{diag}(d_1, \dots, d_n)$ und I die Einheitsmatrix bezeichnet. Dann ist $B \preceq A$ genau dann, wenn $C D C^\top \preceq C I C^\top$ und damit $0 \preceq C(I - D)C^\top$. Die Eigenwerte von $I - D$ sind genau dann nicht negativ, wenn $d_i \leq 1$ für alle $i = 1, \dots, n$. Da

$$BA^{-1} = C D C^\top (C^\top)^{-1} C^{-1} = C D C^{-1},$$

sind die Eigenwerte von BA^{-1} nach Proposition 2.3.5 genau d_1, \dots, d_n . □

Für reelle positive Zahlen a und b ist $a \geq b$ genau dann, wenn $b^{-1} \geq a^{-1}$. Mit Hilfe von Lemma 2.4.6 können wir beweisen, dass diese Aussage auch im Bezug auf die Löwner-Ordnung für positiv definite Matrizen und ihre Inversen gilt.

Satz 2.4.7. *Es seien $A, B \in \mathbb{R}^{n \times n}$ positiv definite Matrizen. Dann gilt: $B \preceq A$ genau dann, wenn $A^{-1} \preceq B^{-1}$.*

Beweis. Gemäß Lemma 2.4.6 ist $A \succeq B$ genau dann, wenn alle Eigenwerte von BA^{-1} kleiner oder gleich 1 sind. Die Matrizen BA^{-1} und $A^{-1}B$ haben die gleichen Eigenwerte, denn für die charakteristischen Polynome erhalten wir

$$\begin{aligned} \chi_{BA^{-1}}(\lambda) &= \det(BA^{-1} - \lambda I) \\ &= \det(A^{-1}(BA^{-1} - \lambda I)A) \\ &= \det(A^{-1}B - \lambda I) \\ &= \chi_{A^{-1}B}(\lambda). \end{aligned}$$

Damit ist $\rho(BA^{-1}) = \rho(A^{-1}(B^{-1})^{-1})$ und damit sind auch alle Eigenwerte von $A^{-1}(B^{-1})^{-1}$ kleiner oder gleich 1. Lemma 2.4.6 besagt, dass dies genau dann der Fall ist, wenn $B^{-1} \succeq A^{-1}$ ist. □

Kapitel 3

Visualisierungen symmetrischer Matrizen

In diesem Kapitel werden Visualisierungen für symmetrische 2×2 -Matrizen und positiv definite 3×3 -Matrizen behandelt und ihre Eigenschaften untersucht. Die hier vorgestellten bildlichen Darstellungen symmetrischer 2×2 -Matrizen bilden die Grundlage für Visualisierungen im Zusammenhang mit Matrixfunktionen.

3.1 Visualisierung positiv definiter Matrizen als Ellipsen

Unser Ziel ist die Visualisierung positiv definiter Matrizen. Dazu betrachten wir quadratische Formen und Quadriken. Als Quadrik wird die Menge aller Punkte bezeichnet, welche die Gleichung

$$q(x) = x^\top Ax + b^\top x + c = 0$$

lösen, wobei A eine symmetrische $n \times n$ -Matrix ist und $b, c, x \in \mathbb{R}^n$ sind. Zur Klassifikation von Quadriken $\{x \in \mathbb{R}^2: x^\top Ax = 1\}$ in Abhängigkeit der Definitheit von A vergleiche man auch [14]. Quadriken im \mathbb{R}^2 werden zum Beispiel durch die Ellipsengleichung oder die Hyperbelgleichung beschrieben. Die Ellipsengleichung lautet in Normalform, d.h. für zusammenfallende Koordinaten- und Ellipsenachsen

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (3.1)$$

wobei a und b die Längen der Halbachsen sind. Die Hyperbelgleichung in Normalform ist durch

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (3.2)$$

gegeben, wobei a den Abstand der Scheitel vom Ursprung angibt. Wir betrachten die positiv definite Matrix

$$A = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix} \quad (3.3)$$

und die quadratische Gleichung

$$x^\top Ax = 10x_1^2 + 12x_1x_2 + 10x_2^2 = 1. \quad (3.4)$$

Anhand dieser Gleichung können wir nicht direkt ablesen, ob es sich bei der zugehörigen Quadrik um eine Ellipse oder eine Hyperbel handelt. Die graphische Darstellung in Abbildung 3.1 zeigt, dass die Quadrik eine Ellipse ergibt, die um den Ursprung des kartesischen Koordinatensystems zentriert ist.

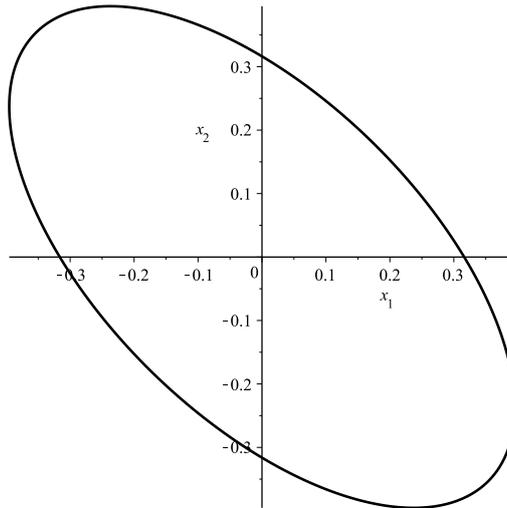


Abbildung 3.1: Die Quadrik (3.4) beschreibt eine Ellipse.

Wir können zeigen, dass eine Quadrik $\{x \in \mathbb{R}^2: x^\top Ax = 1\}$ für eine positiv definite Matrix A stets eine Ellipse beschreibt, deren Hauptachsen entlang der Eigenvektoren von A ausgerichtet sind und dass die Längen der Halbachsen von den Eigenwerten von A bestimmt werden. Dazu führen wir eine orthogonale Koordinatentransformation durch, welche den Ausdruck $x^\top Ax = 1$ auf die Normalform der Ellipsengleichung zurückführt.

Die oben betrachtete Matrix A (3.3) besitzt gemäß Beispiel 2.2.4 die Eigenwerte $\lambda_1 = 4$ und $\lambda_2 = 16$ mit den zugehörigen normierten Eigenvektoren

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{und} \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und lässt sich mit Hilfe der orthogonalen Matrix $S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ ausdrücken als

$$A = SDS^\top = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Dann ist

$$x^\top Ax = (S^\top x)^\top D(S^\top x) = 4 \left(\frac{x_1 - x_2}{\sqrt{2}} \right)^2 + 16 \left(\frac{x_1 + x_2}{\sqrt{2}} \right)^2.$$

Die Koeffizienten 4 und 16 sind die Eigenwerte der Matrix A . In den quadratischen Termen können wir die Eigenvektoren v_1 und v_2 identifizieren. Die Achsen der Ellipse zeigen in Richtung der Eigenvektoren von A . Durch Einführung der neuen Koordinaten

$$X = \frac{x_1 - x_2}{\sqrt{2}} \quad \text{und} \quad Y = \frac{x_1 + x_2}{\sqrt{2}}$$

lässt sich die Ellipse in der Normalform

$$4X^2 + 16Y^2 = 1 \tag{3.5}$$

ausdrücken (vgl. Abbildung 3.2 (a)). Die Längen der Halbachsen sind damit 0,5 und 0,25. Da durch eine orthogonale Koordinatentransformation nach (2.3), (2.4) und (2.5) Längen, Abstände und Winkel erhalten bleiben, sind dies auch die Längen der Halbachsen im alten Koordinatensystem.

Symmetrische Matrizen sind orthogonal diagonalisierbar. Daher können wir eine symmetrische Matrix $A \in \mathbb{R}^{2 \times 2}$ in der Form $A = SDS^T$ schreiben, wobei die Matrix S orthogonal und D eine Diagonalmatrix mit den Eigenwerten λ_1 und λ_2 von A auf der Diagonalen ist. Wir können dann $x^T Ax = 1$ umformen zu

$$x^T Ax = \begin{pmatrix} x_1 & x_2 \end{pmatrix} SDS^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} X & Y \end{pmatrix} D \begin{pmatrix} X \\ Y \end{pmatrix} = \lambda_1 X^2 + \lambda_2 Y^2 = 1.$$

Ist $A = SDS^T = S \operatorname{diag}(\lambda_1, \lambda_2) S^T$ positiv definit, so sind $\lambda_1, \lambda_2 > 0$. Der Graph von $x^T Ax = 1$ ist gemäß (3.1) eine Ellipse mit den Halbachsenlängen

$$\frac{1}{\lambda_1} \quad \text{und} \quad \frac{1}{\lambda_2}.$$

Ihre Achsen zeigen den obigen Überlegungen zufolge in Richtung der Eigenvektoren von A . Ist A indefinit, so ist ein Eigenwert positiv und ein Eigenwert negativ. Die Gleichung $x^T Ax = 1$ beschreibt demnach gemäß (3.2) eine Hyperbel (vgl. Abbildung 3.2 (b)). Für eine negativ definite Matrix A hat $x^T Ax = \lambda_1 X^2 + \lambda_2 Y^2 = 1$ keine Lösung, denn λ_1 und λ_2 sind beide negativ.

Wir nehmen die Quadrik $\{x \in \mathbb{R}^2: x^T Ax = 1\}$ einer positiv definiten Matrix $A \in \mathbb{R}^{2 \times 2}$ als Ausgangspunkt zur Herleitung einer Visualisierung für eine positiv definite 2×2 -Matrix A . Die Visualisierung der Menge $\{x \in \mathbb{R}^2: x^T Ax = 1\}$ selbst ist unzureichend. Die Hauptachsen der entsprechenden Ellipse sind zwar entlang der Eigenvektoren von A ausgerichtet, die Längen der Halbachsen sind allerdings durch

$$\frac{1}{\sqrt{\lambda_1}} \quad \text{und} \quad \frac{1}{\sqrt{\lambda_2}}$$

gegeben. Daher betrachten wir

$$\{x \in \mathbb{R}^2: x^T A^{-2} x = 1\},$$

wobei A^{-2} die Inverse von A^2 bezeichnet. Wir treffen zunächst Aussagen über die Eigenwerte und Eigenvektoren von A^{-2} .

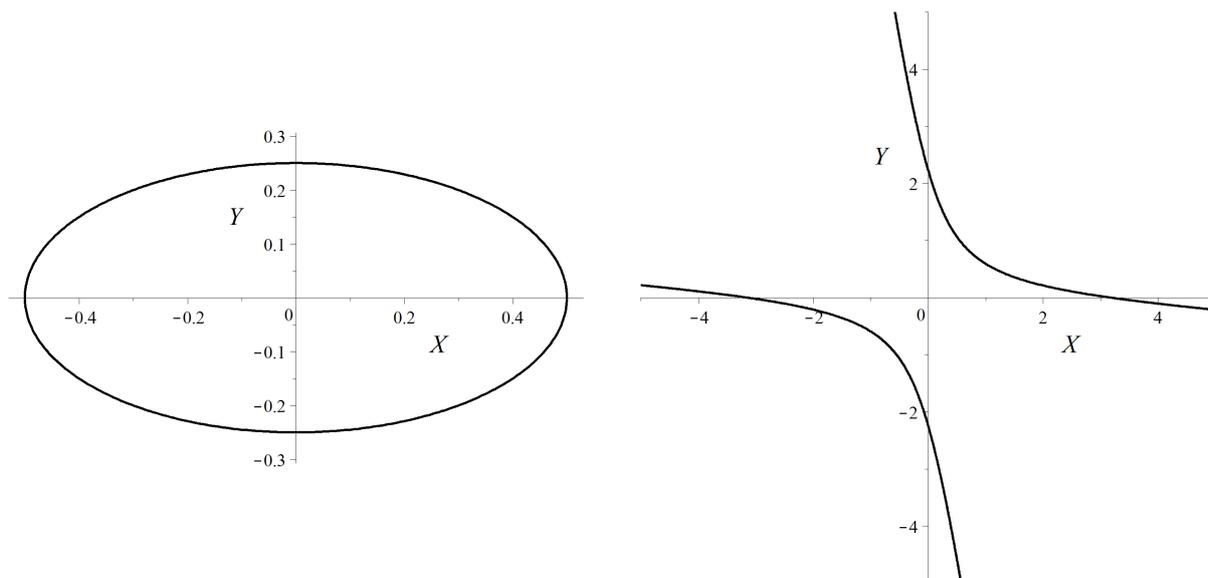


Abbildung 3.2: **(a) Links:** Die Ellipse $\{x \in \mathbb{R}^2: x^\top Ax = 1\}$ zur Matrix A aus (3.3) in den X-Y-Koordinaten gemäß (3.5). Die Hauptachsen fallen mit den Koordinatenachsen zusammen. **(b) Rechts:** Ist A indefinit, beschreibt $\{x \in \mathbb{R}^2: x^\top Ax = 1\}$ eine Hyperbel.

Proposition 3.1.1. *Es sei $A \in \mathbb{R}^{2 \times 2}$ positiv definit. Dann gilt:*

- i) Die Eigenwerte von A^{-2} sind die Quadrate der reziproken Eigenwerte von A .*
- ii) Die Matrizen A und A^{-2} haben die gleichen Eigenvektoren.*

Beweis. i) Mit einer orthogonalen Matrix S können wir A schreiben als

$$A = S \operatorname{diag}(\lambda_1, \lambda_2) S^\top,$$

wobei λ_1 und λ_2 die Eigenwerte von A sind. Damit ist

$$A^2 = S \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \underbrace{S^\top S}_I \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} S^\top = S \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} S^\top,$$

womit gezeigt ist, dass λ_1^2 und λ_2^2 die Eigenwerte von A^2 sind. Die zu A^2 inverse Matrix A^{-2} hat die zu den Eigenwerten von A^2 reziproken Eigenwerte $1/\lambda_1^2$ und $1/\lambda_2^2$, denn

$$A^{-2} = \left(S \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} S^\top \right)^{-1} = S \begin{pmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_2^2 \end{pmatrix} S^\top.$$

- ii) A habe den Eigenvektor v zum Eigenwert λ . Dann ist v auch Eigenvektor von A^2 zum Eigenwert λ^2 , denn

$$A^2 v = A(Av) = A\lambda v = \lambda Av = \lambda\lambda v = \lambda^2 v.$$

Somit ist $(A^2 - \lambda^2 I)v = 0$. Dies können wir durch Multiplikation von $-\lambda^{-2}A^{-2}$ von links umformen:

$$\begin{aligned} (A^2 - \lambda^2 I)v = 0 &\iff (-\lambda^{-2}A^{-2}A^2 + \lambda^{-2}A^{-2}\lambda^2 I)v = 0 \\ &\iff (-\lambda^{-2}I + A^{-2})v = 0 \\ &\iff (A^{-2} - \lambda^{-2}I)v = 0. \end{aligned}$$

Da $1/\lambda^2$ Eigenwert von A^{-2} ist, ist v auch Eigenvektor von A^{-2} . □

Daher lässt sich $x^\top A^{-2}x = 1$ auf die Normalform

$$\frac{X^2}{\lambda_1^2} + \frac{Y^2}{\lambda_2^2} = 1$$

bringen, wobei λ_1 und λ_2 die Eigenwerte von A sind. Die Punktmenge

$$\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$$

beschreibt für eine positiv definite Matrix A somit eine Ellipse, deren Halbachsenlängen durch die beiden Eigenwerte von A gegeben sind. Die Hauptachsen dieser Ellipse sind Proposition 3.1.1 und den vorausgegangenen Überlegungen zufolge entlang der Eigenvektoren von A ausgerichtet. Die Ellipse $\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$ liefert daher eine Visualisierung positiv definiter 2×2 -Matrizen. Die Matrix A aus (3.3) ist dementsprechend in Abbildung 3.3 visualisiert. Positiv definite 3×3 -Matrizen lassen sich entsprechend als Ellipsoide

$$\{x \in \mathbb{R}^3: x^\top A^{-2}x = 1\}$$

darstellen. Eine solche Visualisierung einer positiv definiten 3×3 -Matrix findet sich in Abbildung 3.4. Die entsprechenden Visualisierungen bestimmen die Matrizen eindeutig.

Der Flächeninhalt einer Ellipse mit den Halbachsenlängen a und b beträgt

$$S_{\text{Ellipse}} = \pi ab.$$

Das Volumen eines Ellipsoiden mit den Halbachsenlängen a, b und c ist

$$V_{\text{Ellipsoid}} = \frac{4}{3}\pi abc.$$

Damit ist für positiv definite 2×2 -Matrizen A der Flächeninhalt der zugehörigen Ellipse $\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$ bzw. für positiv definite 3×3 -Matrizen B das Volumen des Ellipsoiden $\{x \in \mathbb{R}^3: x^\top B^{-2}x = 1\}$ proportional zur Determinante der repräsentierten Matrix.

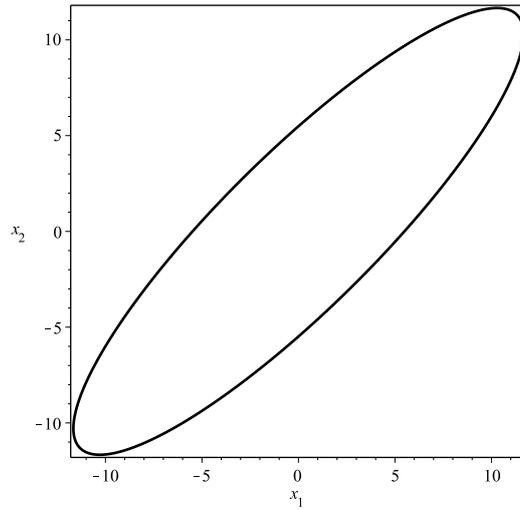


Abbildung 3.3: Visualisierung der Matrix A aus (3.3) als Ellipse $\{x \in \mathbb{R}^2 : x^\top A^{-2}x = 1\}$.

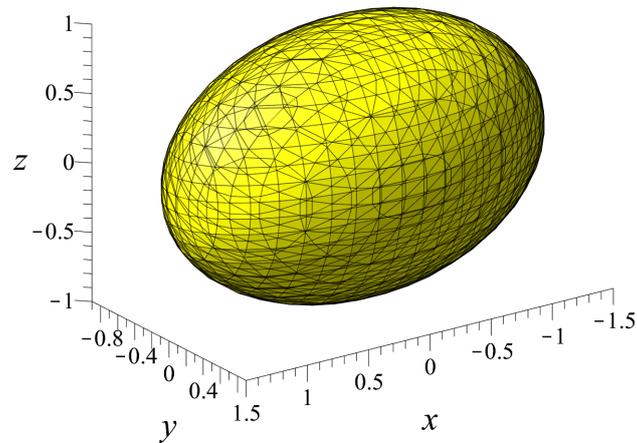


Abbildung 3.4: Visualisierung einer positiv definiten 3×3 -Matrix als Ellipsoid.

3.2 Visualisierung als Punkte im \mathbb{R}^3

Eine weitere Möglichkeit zur graphischen Repräsentation von symmetrischen Matrizen $A \in \mathbb{R}^{2 \times 2}$ ist die Abbildung der Matrizen auf Punkte im \mathbb{R}^3 über die in [3] eingeführte Abbildungsvorschrift

$$\varphi: \text{Sym}(2) \rightarrow \mathbb{R}^3, \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \mapsto \frac{1}{\sqrt{2}}(2\beta, \gamma - \alpha, \alpha + \gamma)^\top, \quad (3.6)$$

die eine bijektive Abbildung zwischen der Menge der symmetrischen Matrizen $\text{Sym}(2)$ und dem \mathbb{R}^3 beschreibt. Die Abbildung ist surjektiv, denn für jeden Punkt $(x, y, z)^\top \in \mathbb{R}^3$ ist ein Urbild in $\text{Sym}(2)$ durch

$$\frac{1}{\sqrt{2}} \begin{pmatrix} z - y & x \\ x & z + y \end{pmatrix}$$

gegeben. Zum Beweis der Injektivität betrachten wir zwei Matrizen $A, B \in \text{Sym}(2)$ mit

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad B = \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{b} & \tilde{c} \end{pmatrix}$$

und den identischen Funktionswerten

$$\varphi(A) = (2b, c - a, a + c)^\top = (2\tilde{b}, \tilde{c} - \tilde{a}, \tilde{a} + \tilde{c})^\top = \varphi(B). \quad (3.7)$$

Aus (3.7) erhalten wir das Gleichungssystem

$$2b = 2\tilde{b}, \quad (3.8)$$

$$c - a = \tilde{c} - \tilde{a}, \quad (3.9)$$

$$a + c = \tilde{a} + \tilde{c}. \quad (3.10)$$

Aus Gleichung (3.8) können wir sofort $b = \tilde{b}$ folgern. Die Addition von (3.9) und (3.10) liefert $c = \tilde{c}$, womit wir auch $a = \tilde{a}$ erhalten. Die Matrizen A und B sind daher identisch. Somit impliziert $\varphi(A) = \varphi(B)$ für alle $A, B \in \text{Sym}(2)$ die Gleichheit von A und B , womit die Injektivität von φ gezeigt ist.

3.2.1 Lage von Matrizen gleicher Spur

Im Rahmen der Visualisierung (3.6) liegen alle symmetrischen 2×2 -Matrizen A mit gleicher Spur auf je einer Ebene parallel zur x-y-Ebene. Die z-Höhe dieser Ebenen ist proportional zur Spur der Matrizen, denn die z-Koordinate liefert

$$\frac{1}{\sqrt{2}}(\alpha + \gamma) = \frac{1}{\sqrt{2}}\text{Spur}(A).$$

Positiv (semi)definite Matrizen besitzen daher eine nicht negative z-Komponente, während die z-Komponente negativ (semi)definiten Matrizen nicht positiv ist.

3.2.2 Lage von Matrizen gleicher Determinante

Wir wollen auch die Fläche beschreiben, auf der Matrizen mit gleicher Determinante d liegen. Im Punkt $(x, y, z)^\top \in \mathbb{R}^3$ liegt die Matrix

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} z - y & x \\ x & z + y \end{pmatrix}$$

mit der Determinante

$$\det(A) = \frac{1}{2}((z - y)(z + y) - x^2) = \frac{1}{2}(z^2 - y^2 - x^2).$$

Die Lage aller symmetrischen 2×2 -Matrizen mit gleicher Determinante d wird daher durch die Fläche

$$F_d = \{(x, y, z)^\top \in \mathbb{R}^3 : z^2 - y^2 - x^2 = 2d\} \quad (3.11)$$

beschrieben. Matrizen mit gleicher positiver Determinante bilden demnach einen zweischaligen Hyperboloiden, Matrizen mit gleicher negativer Determinante einen einschaligen Hyperboloiden und Matrizen mit Determinante Null einen Doppelkegel (vgl. Abbildung 3.5).

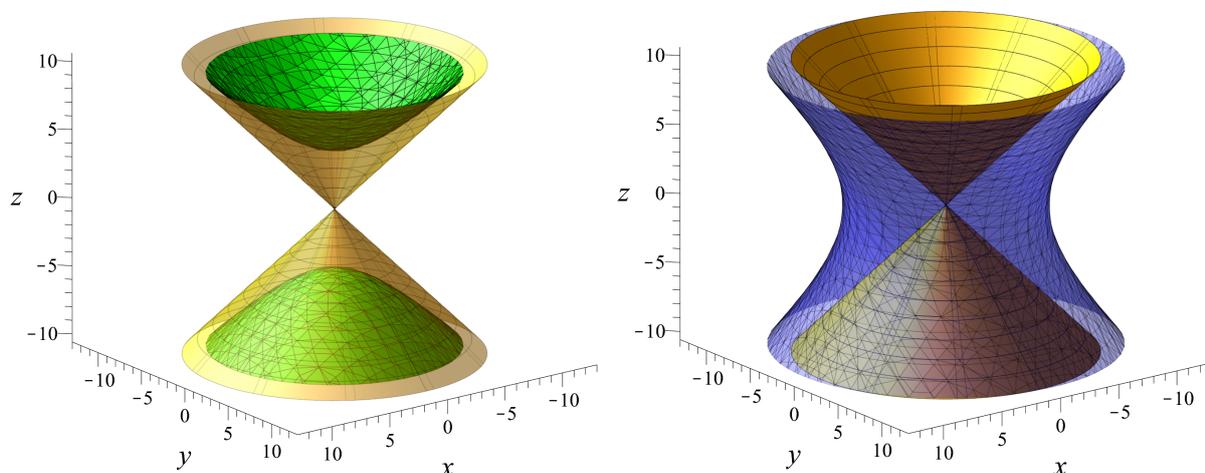


Abbildung 3.5: **(a) Links:** Alle symmetrischen 2×2 -Matrizen mit gleicher positiver Determinante (hier: 10) liegen auf einem zweischaligen Hyperboloiden (grün). **(b) Rechts:** Der einschalige Hyperboloid (blau) beschreibt die Lage aller symmetrischer 2×2 -Matrizen mit gleicher negativer Determinante (hier: -25). Auf den gelben Doppelkegeln liegen die Matrizen, deren Determinante Null ist.

3.2.3 Lage von Matrizen in Abhängigkeit ihrer Definitheit

Wir können das Resultat aus Abschnitt 3.2.2 nutzen, um auf die Lage symmetrischer 2×2 -Matrizen in Abhängigkeit ihrer Definitheit zu schließen. Dazu verwenden wir, dass eine 2×2 -Matrix genau dann semidefinit ist, wenn ihre Determinante Null ist. Zusammen mit der oben getroffenen Aussage über die z-Koordinaten positiv (semi)definiten und negativ (semi)definiten Matrizen ergibt sich, dass die positiv semidefiniten Matrizen den in positiver z-Richtung geöffneten Kegel

$$K^+ = \{(x, y, z)^\top \in \mathbb{R}^3 : z^2 - y^2 - x^2 = 0, z \geq 0\} \quad (3.12)$$

formen und die negativ semidefinite Matrizen den in negativer z-Richtung geöffneten Kegel

$$K^- = \{(x, y, z)^\top \in \mathbb{R}^3 : z^2 - y^2 - x^2 = 0, z \leq 0\}. \quad (3.13)$$

Im Inneren dieser Kegel liegen die Matrizen mit positiver Determinante, dies sind die positiv definiten und negativ definiten Matrizen. Die positiv definiten Matrizen liegen im Inneren von K^+ , die negativ Definiten im Innern von K^- . Die Matrizen außerhalb der Kegel sind indefinit. Sie haben eine negative Determinante. Der Kegel K^+ entspricht der Menge $\text{Sym}^+(2)$, der Kegel K^- der Menge $\text{Sym}^-(2)$. Den von K^- und K^+ gebildeten Doppelkegel bezeichnen wir als Löwner-Doppelkegel $K_L(2)$. Dieser ist in Abbildung 3.6 dargestellt.

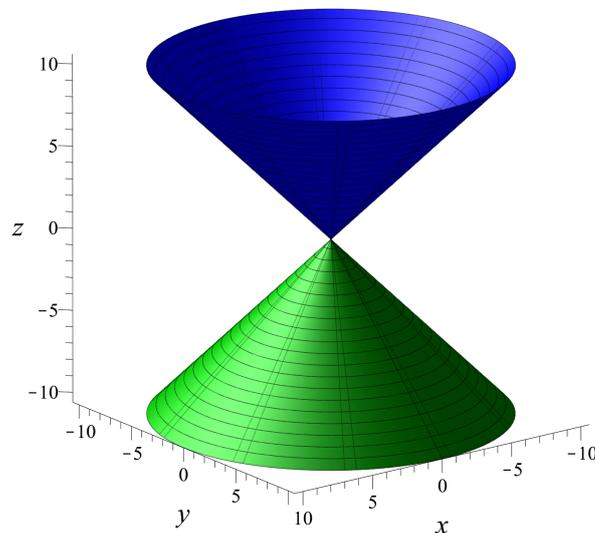


Abbildung 3.6: Der Doppelkegel (Löwner-Doppelkegel) beschreibt die Lage der symmetrischen 2×2 -Matrizen in Abhängigkeit ihrer Definitheit. Die Matrizen auf den Kegelwänden des blauen Kegels sind positiv semidefinit, die Matrizen in seinem Innern positiv definit. Dementsprechend liegen die negativ semidefiniten Matrizen auf dem Rand des grünen Kegels und die Negativ definiten in seinem Innern. Alle Matrizen außerhalb des Doppelkegels sind indefinit.

3.2.4 Lage von Matrizen mit gleichen Eigenwerten

Da die Spur einer Matrix der Summe ihrer Eigenwerte entspricht, liegen auf je einer Ebene parallel zur x-y-Ebene die Matrizen, bei denen die Summen der Eigenwerte konstant sind. Sind λ und μ die Eigenwerte einer Matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

so liegen insbesondere auch alle Matrizen

$$A' = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix},$$

die ebenfalls die Eigenwerte λ und μ besitzen, auf der gleichen Ebene wie A . Um ihre Lage genauer zu beschreiben, betrachten wir das charakteristische Polynom

$$\chi_A(t) = t^2 - (a + c)t - b^2 + ac$$

von A , dessen Nullstellen

$$\lambda = \frac{a + c}{2} + \frac{\sqrt{(c - a)^2 + 4b^2}}{2} \quad (3.14)$$

und

$$\mu = \frac{a + c}{2} - \frac{\sqrt{(c - a)^2 + 4b^2}}{2} \quad (3.15)$$

die Eigenwerte von A sind. Entsprechend sind

$$\lambda' = \frac{\alpha + \gamma}{2} + \frac{\sqrt{(\gamma - \alpha)^2 + 4\beta^2}}{2} \quad (3.16)$$

und

$$\mu' = \frac{\alpha + \gamma}{2} - \frac{\sqrt{(\gamma - \alpha)^2 + 4\beta^2}}{2} \quad (3.17)$$

die Eigenwerte der Matrizen A' . Stimmen die Eigenwerte der Matrizen A' mit denen von A überein, so ist insbesondere

$$\text{Spur}(A') = \alpha + \gamma = a + c = \text{Spur}(A).$$

Die Betrachtung von (3.14) bis (3.17) zeigt, dass die Eigenwerte von A' und A daher genau dann identisch sind, wenn

$$\frac{(\gamma - \alpha)^2}{2} + \frac{4\beta^2}{2} = \frac{(c - a)^2}{2} + \frac{4b^2}{2} = \text{const.} \quad (3.18)$$

Für die x- und y-Koordinaten, auf welche die Matrizen A' durch (3.6) abgebildet werden, gilt

$$x = \frac{2\beta}{\sqrt{2}} \quad \text{und} \quad y = \frac{\gamma - \alpha}{\sqrt{2}},$$

sodass (3.18) die Kreisgleichung

$$x^2 + y^2 = r^2 \text{ mit } r \in \mathbb{R}, r = \text{const.}$$

beschreibt. Somit liegen auf je einem Kreis um die z -Achse die Matrizen mit den gleichen Eigenwerten λ und μ . Die z -Koordinate der Matrizen auf dem Kreis wird durch ihre Spur und damit durch die Summe ihrer Eigenwerte bestimmt, sodass sich für die z -Koordinate

$$z = \frac{(\lambda + \mu)}{\sqrt{2}}$$

ergibt. Die Bestimmung des Kreisradius erfolgt durch die Betrachtung der Diagonalmatrizen auf dem Kreis, die in der x - y -Ebene liegen. Der Radius des Kreises entspricht dem Betrag ihrer y -Koordinate. Damit erhalten wir für den Kreisradius r einen Wert von

$$r = \frac{|\lambda - \mu|}{\sqrt{2}}.$$

Abbildung 3.7 zeigt einen Kreis von Matrizen mit gleichen Eigenwerten.

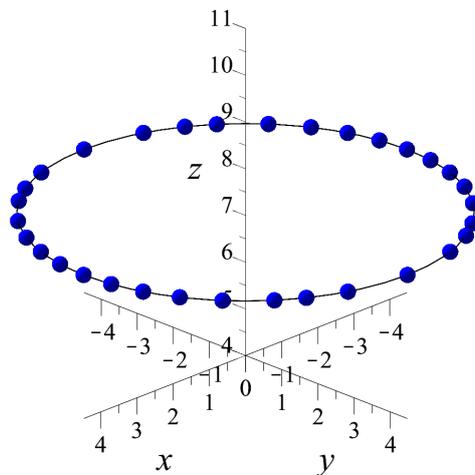


Abbildung 3.7: Kreis von Matrizen gleicher Eigenwerten λ und μ . In blau sind exemplarisch Matrizen mit entsprechenden Eigenwerten eingezeichnet.

3.2.5 Lage von Matrizen mit gleichen Eigenvektoren

Wir betrachten die Lage aller symmetrischen 2×2 -Matrizen, welche die gleichen Eigenvektoren u und v haben. Dazu bezeichnen wir mit $u^0 = (u_1 \ u_2)^\top$ und $v^0 = (v_1 \ v_2)^\top$ die normierten Vektoren von u und v . Sind u und v nicht orthogonal, müssen sie Eigenvektoren zu Vielfachen der Einheitsmatrix sein, die auf der z -Achse liegen. Für alle anderen symmetrischen 2×2 -Matrizen sind die Eigenvektoren u und v orthogonal, da die beiden Eigenwerte stets verschieden sind. Jede solche Matrix hat die Gestalt

$$\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} = \begin{pmatrix} \lambda u_1^2 + \mu v_1^2 & \lambda u_1 u_2 + \mu v_1 v_2 \\ \lambda u_1 u_2 + \mu v_1 v_2 & \lambda u_2^2 + \mu v_2^2 \end{pmatrix},$$

wobei $\lambda, \mu \in \mathbb{R}$. Matrizen dieser Form werden auf die Koordinate

$$\frac{1}{\sqrt{2}}(x, y, z)^\top$$

abgebildet mit

$$\begin{aligned} x &= 2\lambda u_1 u_2 + 2\mu v_1 v_2, \\ y &= \lambda(u_2^2 - u_1^2) + \mu(v_2^2 - v_1^2), \\ z &= \lambda + \mu. \end{aligned}$$

Daran erkennen wir, dass Matrizen mit gleichen orthogonalen Eigenvektoren u und v eine Ebene E_{uv} bilden, deren Parameterform durch

$$E_{uv}: (x, y, z)^\top = \frac{\lambda}{\sqrt{2}} \begin{pmatrix} 2u_1 u_2 \\ u_2^2 - u_1^2 \\ 1 \end{pmatrix} + \frac{\mu}{\sqrt{2}} \begin{pmatrix} 2v_1 v_2 \\ v_2^2 - v_1^2 \\ 1 \end{pmatrix}$$

mit $\lambda, \mu \in \mathbb{R}$ gegeben ist. Unter Ausnutzung der Orthonormalität von u^0 und v^0 erhalten wir durch Bildung des Normalenvektors die Ebenengleichung der Ebene E_{uv} in Komponentenschreibweise:

$$E_{uv}: (u_2^2 + v_1^2 - u_1^2 - v_2^2)x + (v_1 v_2 - u_1 u_2)y = 0. \quad (3.19)$$

Anhand (3.19) ist erkennbar, dass die z -Achse in jeder der Ebenen E_{uv} enthalten ist. Dies ist damit zu erklären, dass jeder beliebige Vektor Eigenvektor zu Vielfachen der Einheitsmatrix ist. Eine solche Ebene, die von Matrizen gleicher Eigenvektoren gebildet wird, ist in Abbildung 3.8 dargestellt.

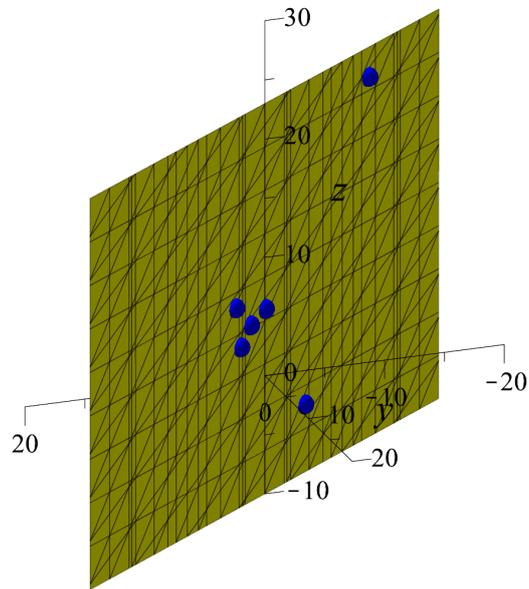


Abbildung 3.8: Beispiel einer Ebene von Matrizen gleicher Eigenvektoren u und v . In blau sind verschiedene Matrizen eingezeichnet, welche diese Eigenvektoren besitzen.

3.3 Visualisierung der Löwner-Ordnung

3.3.1 Visualisierung für positiv definite Matrizen über Ellipsen

Positiv definite Matrizen $A \in \mathbb{R}^{2 \times 2}$ können als Ellipsen

$$\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$$

dargestellt werden. Wir überprüfen, ob sich mit Hilfe dieser Darstellung auch die Löwner-Ordnung visualisieren lässt. Dazu betrachten wir die Matrizen

$$A = \begin{pmatrix} 3 & 2\sqrt{3} \\ 2\sqrt{3} & 7 \end{pmatrix}, B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix}, C = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \text{ und } D = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}.$$

Die Eigenwerte der Matrix

$$A - B = \begin{pmatrix} \frac{5}{2} & 2\sqrt{3} \\ 2\sqrt{3} & 5 \end{pmatrix}$$

sind beide positiv, sodass $A \succeq B$ ist. Die Matrizen C und D sind nach Beispiel 2.4.4 in der Löwner-Ordnung nicht vergleichbar. Abbildung 3.9 (a) zeigt die Visualisierungen der Matrizen A und B als Ellipsen $\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$, bzw. $\{x \in \mathbb{R}^2: x^\top B^{-2}x = 1\}$. Die Matrizen C und D sind dementsprechend in Abbildung 3.9 (b) visualisiert.

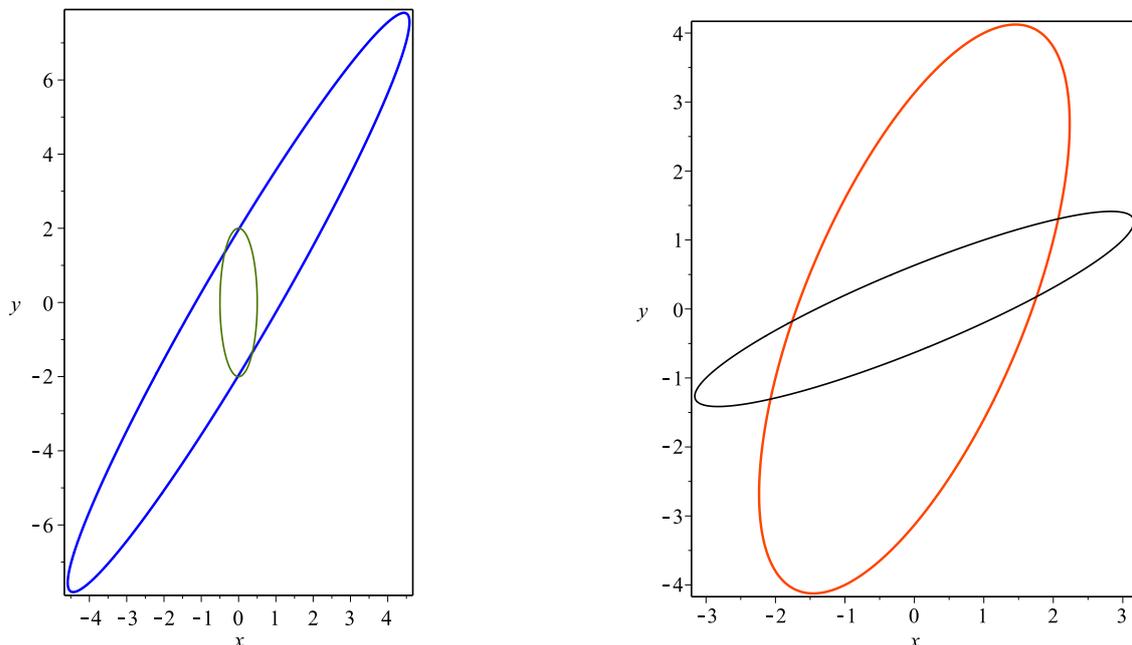


Abbildung 3.9: **(a) Links:** Visualisierungen der in der Löwner-Ordnung vergleichbaren Matrizen A (blau) und B (grün) aus dem einführenden Beispiel. **(b) Rechts:** Visualisierungen der nicht vergleichbaren Matrizen C (rot) und D (schwarz) aus dem einführenden Beispiel.

Wir erkennen, dass sich sowohl die zu den Matrizen A und B , als auch die zu C und D gehörenden Ellipsen schneiden, obwohl A und B in der Löwner-Ordnung vergleichbar sind, C und D aber nicht. Die gewählte Visualisierung lässt daher keine Aussagen über die Vergleichbarkeit der Matrizen zu, womit sie zur graphischen Darstellung der Löwner-Ordnung ungeeignet ist. Als geeignet erweist sich hingegen die Darstellung positiver definiten Matrizen $A \in \mathbb{R}^{2 \times 2}$ als Ellipsen

$$\{x \in \mathbb{R}^2 : x^\top A^{-1} x = 1\},$$

die wir im Folgenden als *Löwner-Visualisierung* bezeichnen. Wir betrachten zwei positiv definite Matrizen $A, B \in \mathbb{R}^{2 \times 2}$ mit $A \succeq B$ und den Löwner-Visualisierungen

$$E_1 = \{x \in \mathbb{R}^2 : x^\top A^{-1} x = 1\} \quad \text{und} \quad E_2 = \{x \in \mathbb{R}^2 : x^\top B^{-1} x = 1\}.$$

Gemäß Satz 2.4.7 ist für positiv definite Matrizen $A \succeq B$ genau dann, wenn $A^{-1} \preceq B^{-1}$, sodass $x^\top A^{-1} x \leq x^\top B^{-1} x$ für alle $x \in \mathbb{R}^2 \setminus \{0\}$ erfüllt ist. Es sei $y \in E_1$, also $y^\top A^{-1} y = 1$. Insbesondere ist

$$y^\top A^{-1} y \leq y^\top B^{-1} y,$$

sodass ein $c \in \mathbb{R}$ existiert mit $c \geq 1$ und

$$y^\top B^{-1} y = c.$$

Dann ist

$$\left(\frac{y}{\sqrt{c}}\right)^\top B^{-1} \frac{y}{\sqrt{c}} = 1, \quad \text{also} \quad \frac{y}{\sqrt{c}} \in E_2.$$

Da $c \geq 1$, ist die Länge des Ortsvektors y größer oder gleich der Länge des Ortsvektors y/\sqrt{c} . Weil $y \in E_1$ beliebig gewählt ist und die Vektoren y und y/\sqrt{c} vom Koordinatenursprung aus in die gleiche Richtung zeigen, sind alle Punkte der Ellipse E_1 gleich weit oder weiter vom Ursprung des kartesischen Koordinatensystems entfernt, als die Punkte der Ellipse E_2 . Im Fall der Gültigkeit der strikten Ungleichung $A \succ B$ liegt die Ellipse E_2 vollständig innerhalb der Ellipse E_1 . Gilt lediglich $A \succeq B$ und die Ungleichung $A \succ B$ ist nicht erfüllt, berühren sich beide Ellipsen. Sind A und B in der Löwner-Ordnung nicht vergleichbar, existieren $u, v \in \mathbb{R}^2$ mit

$$u^\top A^{-1} u < u^\top B^{-1} u$$

und

$$v^\top A^{-1} v > v^\top B^{-1} v.$$

Die entsprechenden Ungleichungen gelten auch für alle Vielfachen der Vektoren u und v . Insbesondere existiert ein Vielfaches \tilde{u} von u und ein Vielfaches \tilde{v} von v mit $\tilde{u}, \tilde{v} \in E_1$, wobei

$$\tilde{u}^\top A^{-1} \tilde{u} < \tilde{u}^\top B^{-1} \tilde{u}$$

und

$$\tilde{v}^\top A^{-1} \tilde{v} > \tilde{v}^\top B^{-1} \tilde{v}.$$

Indem wir wie oben argumentieren, erkennen wir, dass in Richtung des Vektors \tilde{u} der Punkt der Ellipse E_1 näher am Koordinatenursprung liegt, als der Punkt der Ellipse

E_2 , während in Richtung des Vektors \tilde{v} der Punkt der Ellipse E_1 weiter vom Koordinatenursprung entfernt ist, als der Punkt der Ellipse E_2 . Wir erkennen, dass sich die Ellipsen E_1 und E_2 in diesem Fall schneiden müssen. Abbildung 3.10 (a) zeigt die Löwner-Visualisierungen der Matrizen A und B aus dem einführenden Beispiel. Die Ellipsen besitzen keinen Schnittpunkt mehr, die zu B gehörende Ellipse liegt vollständig innerhalb der Ellipse, welche die Matrix A in der Löwner-Visualisierung repräsentiert. In Abbildung 3.10 (b) sind die Löwner-Visualisierungen der Matrizen C und D aus dem einführenden Beispiel dargestellt. Da C und D in der Löwner-Ordnung nicht vergleichbar sind, schneiden sich die Ellipsen.

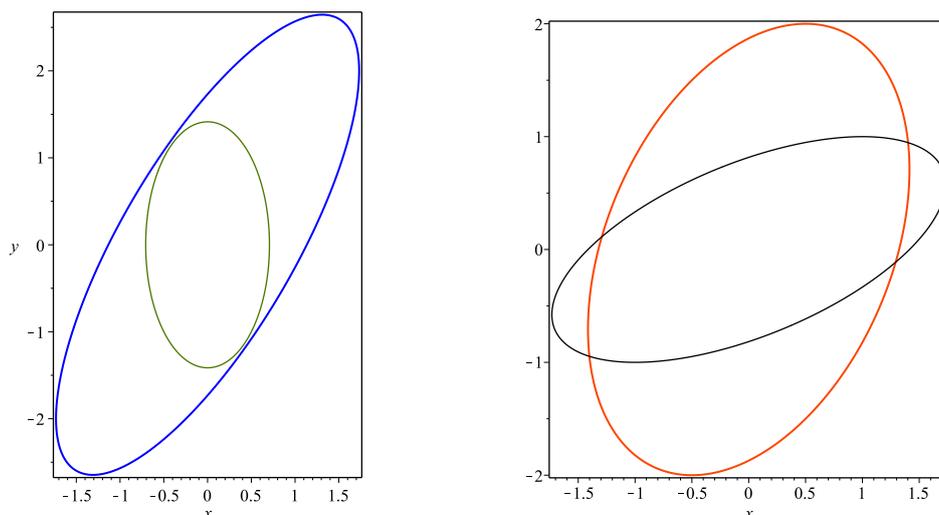


Abbildung 3.10: **(a) Links:** Löwner-Visualisierungen der in der Löwner-Ordnung vergleichbaren Matrizen A (blau) und B (grün) aus dem einführenden Beispiel. **(b) Rechts:** Löwner-Visualisierungen der nicht vergleichbaren Matrizen C (rot) und D (schwarz) aus dem einführenden Beispiel.

Für die Matrizen A und B aus Beispiel 2.4.5 gilt $A \succeq B$, wobei die strikte Ungleichung $A \succ B$ nicht erfüllt ist. Wir sehen in Abbildung 3.11, dass sich die Ellipsen der Löwner-Visualisierung berühren.

Anstelle der Löwner-Visualisierung positiv definitiver Matrizen $A \in \mathbb{R}^{2 \times 2}$ als Ellipsen

$$\{x \in \mathbb{R}^2: x^\top A^{-1}x = 1\},$$

würde auch die Repräsentation als Ellipsen

$$\{x \in \mathbb{R}^2: x^\top Ax = 1\}$$

eine Visualisierung ergeben, die Rückschlüsse über das Verhalten der Matrizen in der Löwner-Ordnung zuließe. Durch Verwendung der Löwner-Visualisierung können wir allerdings erreichen, dass die zu einer Matrix A mit $A \succeq B$ gehörende Ellipse die Ellipse

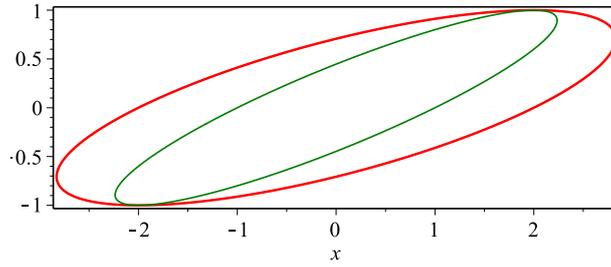


Abbildung 3.11: Die Ellipsen der löwnervisualisierten Matrizen A (grün) und B (rot) aus Beispiel 2.4.5 berühren sich. Für sie ist $A \succeq B$, die strikte Ungleichung $A \succ B$ aber nicht erfüllt.

der Matrix B umschließt, was mit unserer Auffassung von Größenordnungen besser zu vereinbaren ist. Die Verwendung der Visualisierung

$$\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$$

versagt, da durch Quadrieren von Matrizen die Ordnung im Allgemeinen nicht erhalten bleibt. Es existieren Matrizen $A, B \in \mathbb{R}^{n \times n}$, für die $A \succeq B$ gilt, A^2 und B^2 in der Löwner-Ordnung aber nicht vergleichbar sind. Dies ist auch für die Matrizen A und B aus dem einführenden Beispiel der Fall. Daher überschneiden sich in Abbildung 3.9 die Ellipsen $\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$ und $\{x \in \mathbb{R}^2: x^\top B^{-2}x = 1\}$.

3.3.2 Visualisierung im \mathbb{R}^3

Die Abbildung (3.6) symmetrischer Matrizen als Punkte im \mathbb{R}^3 durch

$$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \longleftrightarrow \frac{1}{\sqrt{2}}(2\beta, \gamma - \alpha, \alpha + \gamma)$$

stellt eine weitere Möglichkeit zur Sichtbarmachung der Löwner-Ordnung dar. Wir interessieren uns für die Lage aller Matrizen, die mit einer beliebigen symmetrischen 2×2 -Matrix vergleichbar sind. Für die Nullmatrix ist eine entsprechende Aussage bereits bekannt. Die Lage aller Matrizen, die in der Löwner-Ordnung mit der Nullmatrix vergleichbar sind, wird durch den Löwner-Doppelkegel $K_L(2)$ beschrieben. Die nicht mit der Nullmatrix vergleichbaren Matrizen liegen außerhalb dieses Doppelkegels. Für eine beliebige symmetrische 2×2 -Matrix

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

erhalten wir eine entsprechende Aussage, indem wir den Löwner-Doppelkegel so verschieben, dass M in der Spitze des Doppelkegels liegt und dessen Symmetrieachse parallel zur z -Achse verläuft. Die Verschiebung entspricht der Abbildung

$$\psi_M: K_L(2) \rightarrow \mathbb{R}^3, (x, y, z)^\top \mapsto \frac{1}{\sqrt{2}}(x + 2b, y + c - a, z + a + c)^\top. \quad (3.20)$$

Den so entstehenden in positiver z-Richtung geöffneten Kegel bezeichnen wir als *obere Penumbra* von M , wir schreiben $P^+(M)$. Den nach unten geöffneten Kegel $P^-(M)$ nennen wir *untere Penumbra* von M . Die obere Penumbra $P^+(M)$ wird somit durch die Menge

$$P^+(M) = M + \text{Sym}^+(2)$$

beschrieben, wobei wir die Notation $M + \text{Sym}^+(2) = \{M + S : S \in \text{Sym}^+(2)\}$ verwendet haben. Die untere Penumbra ist dementsprechend

$$P^-(M) = M - \text{Sym}^+(2).$$

Also ist $N \in P^+(M)$ genau dann, wenn $N - M$ positiv semidefinit ist und $N \in P^-(M)$ genau dann, wenn $N - M$ negativ semidefinit ist. Daher ist

$$P^+(M) = \{N \in \text{Sym}(2) : N \succeq M\}$$

und

$$P^-(M) = \{N \in \text{Sym}(2) : N \preceq M\}.$$

Die Matrizen außerhalb von $P^+(M)$ und $P^-(M)$ sind in der Löwnerordnung nicht mit M vergleichbar.

Abschließend betrachten wir, was die Vergleichbarkeit von zwei Matrizen M_1 und M_2 in der Löwner-Ordnung für ihre Penumbras bedeutet, wobei wir die unteren Penumbras betrachten. Es ist $M_1 \preceq M_2$ genau dann, wenn die untere Penumbra von M_2 die untere Penumbra von M_1 überdeckt, d.h. es gilt $P^-(M_1) \subseteq P^-(M_2)$. Sind zwei Matrizen in der Löwner-Ordnung nicht vergleichbar, schneiden sich ihre unteren Penumbras. Abbildung 3.12 zeigt links die unteren Penumbras von zwei in der Löwner-Ordnung vergleichbaren Matrizen und rechts die unteren Penumbras von zwei nicht vergleichbaren Matrizen. In Abbildung 3.13 sind die unteren Penumbras der in der Löwner-Ordnung vergleichbaren Matrizen aus Beispiel 2.4.5 dargestellt.

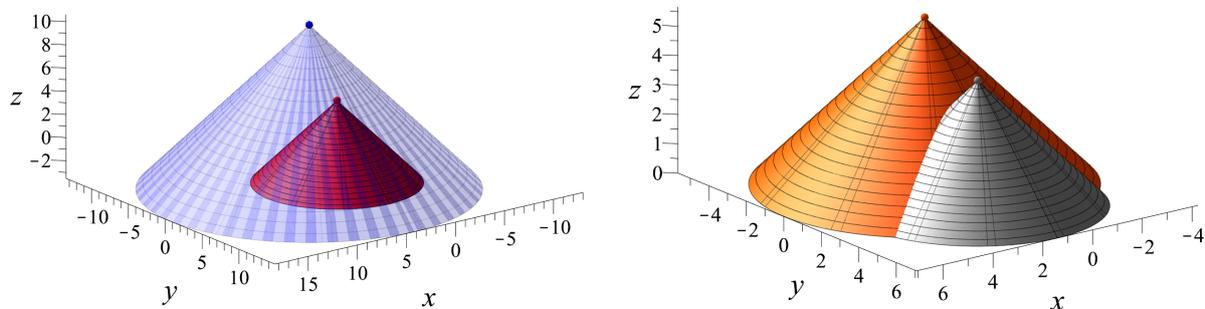


Abbildung 3.12: (a) **Links:** Untere Penumbras zweier in der Löwner-Ordnung vergleichbarer Matrizen. (b) **Rechts:** Untere Penumbras zweier in der Löwner-Ordnung nicht vergleichbarer Matrizen.

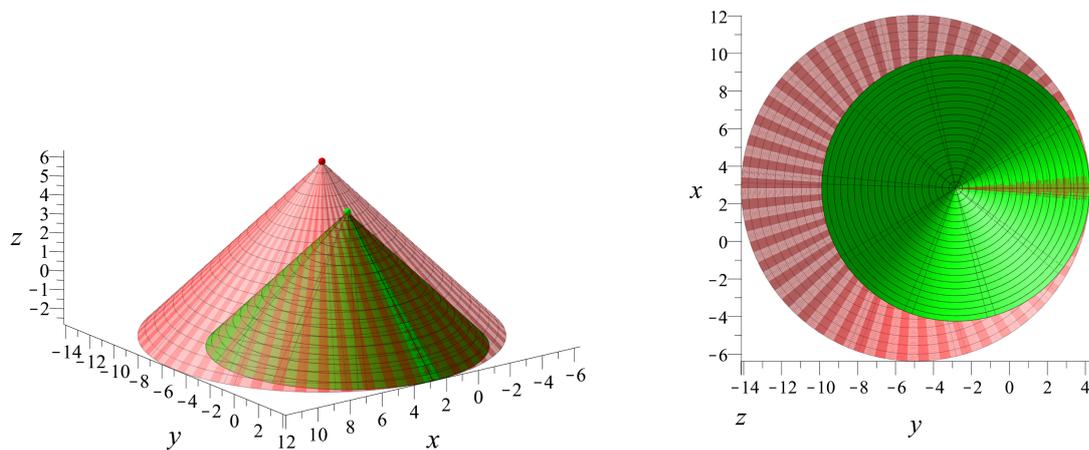


Abbildung 3.13: (a) **Links:** Untere Penumbras der in der Löwner-Ordnung vergleichbaren Matrizen A (grün) und B (rot) aus Beispiel 2.4.5. Ein Eigenwert von $B - A$ ist Null. Die unteren Penumbras berühren sich. (b) **Rechts:** Blick auf die unteren Penumbras in der x - y -Ebene.

Kapitel 4

Matrixfunktionen einer Variablen

Unter Matrixfunktionen können im Allgemeinen Funktionen verstanden werden, bei denen eine Matrix als Argument oder Funktionswert auftritt. In diesem Kapitel liegt das Interesse auf Funktionen, die sowohl als Argument, als auch als Funktionswert eine Matrix gleicher Dimension aufweisen und von skalaren Funktionen abstammen. Eine skalare Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ wird auf eine Matrix $A \in \mathbb{R}^{n \times n}$ angewendet und $f(A) \in \mathbb{R}^{n \times n}$ berechnet. Andere Auffassungen von Matrixfunktionen, die hier nicht näher betrachtet werden, sind unter anderem folgende:

- Funktionen, die ein Skalar auf eine Matrix abbilden, wie $f: \mathbb{R} \rightarrow \mathbb{R}^{n \times n}, t \mapsto tA$ mit $A \in \mathbb{R}^{n \times n}$.
- Funktionen, die eine Matrix auf ein Skalar abbilden. Dies leisten zum Beispiel die Determinantenfunktion oder das charakteristische Polynom einer Matrix.
- Funktionen, die eine Matrix aus $\mathbb{R}^{n \times n}$ auf eine Matrix aus $\mathbb{R}^{m \times m}$ abbilden und nicht von skalaren Funktionen abstammen. Darunter fällt zum Beispiel das Transponieren von Matrizen.

Im Rahmen dieser Arbeit werden Funktionen symmetrischer Matrizen betrachtet. Definitionen für Funktionen nichtsymmetrischer Matrizen können zum Beispiel in [11] oder [9] nachgelesen werden.

4.1 Definition und Eigenschaften

Der Versuch, eine Matrixfunktion $f(A)$ über komponentenweise Anwendung einer skalaren Funktion auf die einzelnen Matrixelemente zu definieren, scheitert. So würde die Funktion $f(t) = t^2$ angewendet auf die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

einen Funktionswert von

$$f(A) = \begin{pmatrix} 1 & 4 \\ 4 & 16 \end{pmatrix}$$

liefern. Eine Matrixfunktion sollte jedoch mit den Operationen der Matrixalgebra verträglich sein, sodass die Funktion $f(t) = t^2$ angewandt auf eine Matrix A das Produkt von A mit sich selbst ergeben sollte. Die Berechnung des Produktes

$$A^2 = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix} \neq \begin{pmatrix} 1 & 4 \\ 4 & 16 \end{pmatrix} = f(A)$$

zeigt jedoch, dass die komponentenweise Anwendung der skalaren Funktion auf die Matrixelemente dieser Forderung nicht gerecht wird und daher keine sinnvolle Definition darstellt, sodass andere Ansätze notwendig sind. Ist

$$p(t) = a_0 + a_1 t + \dots + a_{m-1} t^{m-1} + a_m t^m$$

ein skalares Polynom mit gegebenen Koeffizienten $a_0, a_1, \dots, a_m \in \mathbb{R}$ und skalarem Argument t , können wir für eine Matrix $A \in \mathbb{R}^{n \times n}$ das Polynom $p(A)$ berechnen, indem wir unter Beibehaltung der Koeffizienten das Argument t durch A und eine 1 durch die Einheitsmatrix I ersetzen. Wir können damit für $A \in \mathbb{R}^{n \times n}$ definieren

$$p(A) := a_0 I + a_1 A + \dots + a_{m-1} A^{m-1} + a_m A^m. \quad (4.1)$$

Als unmittelbare Folgerung aus (4.1) erhalten wir für Matrizen $S, T \in \mathbb{R}^{n \times n}$ mit S invertierbar die Gleichheit $p(STS^{-1}) = Sp(T)S^{-1}$. Daher können wir für eine diagonalisierbare Matrix $A \in \mathbb{R}^{n \times n}$ mit der Eigenwertzerlegung $A = SDS^{-1} = S \operatorname{diag}(\lambda_1, \dots, \lambda_n) S^{-1}$ das Polynom $p(A)$ durch

$$p(A) = p(SDS^{-1}) = Sp(D)S^{-1} = S \begin{pmatrix} p(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & p(\lambda_n) \end{pmatrix} S^{-1} \quad (4.2)$$

berechnen. Dieses Vorgehen lässt sich über Potenzreihen auf beliebige analytische Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ ausweiten. Da symmetrische Matrizen nach Satz 2.3.7 stets orthogonal diagonalisierbar sind und gemäß Satz 2.2.6 nur reelle Eigenwerte besitzen, können wir Funktionen symmetrischer Matrizen somit wie folgt definieren:

Definition 4.1.1. Es sei f auf dem Spektrum einer symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ definiert. A habe die Eigenwertzerlegung $A = SDS^\top = S \operatorname{diag}(\lambda_1, \dots, \lambda_n)S^\top$. Dann ist $f(A)$ definiert als

$$f(A) := Sf(D)S^\top = S \operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n))S^\top.$$

Beispiel 4.1.2. Die Matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ mit der Eigenwertzerlegung

$$A = \frac{1}{5} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

hat unter der Funktion $f(t) = t^2$ den Funktionswert

$$f(A) = \frac{1}{5} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} f(5) & 0 \\ 0 & f(0) \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix}.$$

Der Funktionswert erfüllt die Erwartungen an die Matrixalgebra, denn $A^2 = f(A)$.

Der Funktionswert einer symmetrischen Matrix gemäß Definition 4.1.1 ist wohldefiniert. Der Wert von $f(A)$ ist unabhängig von der für A gewählten Eigenwertzerlegung. Um dies zu zeigen, betrachten wir eine symmetrische Matrix $A = SDS^\top = S \operatorname{diag}(\lambda_1, \dots, \lambda_n)S^\top$ und eine weitere Eigenwertzerlegung von A , die durch $A = TMT^\top = T \operatorname{diag}(\mu_1, \dots, \mu_n)T^\top$ gegeben sei. Dann stellen $\lambda_1, \dots, \lambda_n$ eine Permutation von μ_1, \dots, μ_n dar, denn die λ_i und μ_i sind die Eigenwerte von A . Es seien t_1, \dots, t_k die verschiedenen Werte der Eigenwerte $\lambda_1, \dots, \lambda_n$. Über Lagrange-Hermite-Interpolation können wir ein Polynom $r(t)$ finden, sodass $r(t_i) = f(t_i)$ für $i = 1, \dots, k$. Näheres über Lagrange-Hermite-Interpolation kann zum Beispiel in [11] oder [9] nachgelesen werden. Wir haben damit

$$\begin{aligned} f(A) &= S \operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n))S^\top = S \operatorname{diag}(r(\lambda_1), \dots, r(\lambda_n))S^\top \\ &= Sr(D)S^\top = r(SDS^\top) = r(A) = r(TMT^\top) \\ &= Tr(M)T^\top = T \operatorname{diag}(r(\mu_1), \dots, r(\mu_n))T^\top \\ &= T \operatorname{diag}(f(\mu_1), \dots, f(\mu_n))T^\top \end{aligned}$$

und somit gezeigt, dass die für A verwendete Eigenwertzerlegung keinen Einfluss auf den Wert von $f(A)$ nimmt.

Wir betrachten im Folgenden wichtige Eigenschaften von Funktionen symmetrischer Matrizen. Zunächst bemerken wir, dass der Funktionswert einer symmetrischen Matrix wieder symmetrisch ist. Denn für eine symmetrische Matrix A mit der Eigenwertzerlegung $A = SDS^\top$ ist

$$f(A)^\top = (Sf(D)S^\top)^\top = Sf(D)^\top S^\top = Sf(D)S^\top = f(A).$$

Proposition 4.1.3 beschreibt den Einfluss einer Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ auf die Eigenwerte und Eigenvektoren der Matrix, auf welche die Funktion angewendet wird. Die Aussagen folgen unmittelbar aus der Definition von $f(A)$.

Proposition 4.1.3. *Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und $f: \mathbb{R} \rightarrow \mathbb{R}$. Dann sind die Eigenwerte von $f(A)$ die Funktionswerte der Eigenwerte von A . Die Matrizen A und $f(A)$ haben die gleichen Eigenvektoren.*

Die entsprechenden Aussagen haben wir ohne Kenntnisse der Definition von Matrixfunktionen für die Funktion $f(t) = t^{-2}$ bereits in Proposition 3.1.1 bewiesen.

Die Aussagen von Proposition 4.1.3 können wir visuell darstellen, indem wir eine positiv definite Matrix A und eine Funktion f betrachten, sodass $f(A)$ ebenfalls positiv definit ist. Die Visualisierungen von A und $f(A)$ als Ellipsen

$$\{x \in \mathbb{R}^2: x^\top A^{-2}x = 1\}$$

bzw.

$$\{x \in \mathbb{R}^2: x^\top f(A)^{-2}x = 1\}$$

liefern Ellipsen, die entlang der gleichen Hauptachsen ausgerichtet sind, da die Hauptachsen durch die Eigenvektoren bestimmt werden. Die beiden Ellipsen unterscheiden sich lediglich in der Länge der Halbachsen, die den Eigenwerten von A , bzw. von $f(A)$ entsprechen. Solche Visualisierungen positiv definiter Matrizen und ihrer positiv definiten Funktionswerte finden sich in Abbildung 4.1.

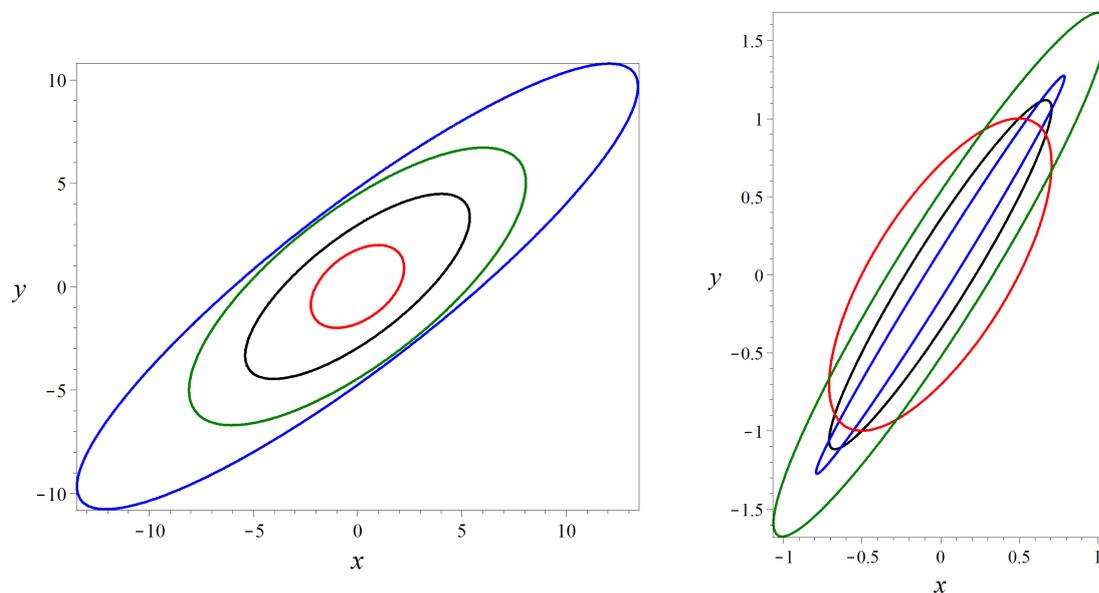


Abbildung 4.1: **(a) Links:** Visualisierung einer positiv definiten Matrix mit Eigenwerten größer 1 (schwarz) und ihrer positiv definiten Funktionswerte unter den Funktionen $f(t) = \sqrt{t}$ (rot), $f(t) = 1,5t$ (grün) und $f(t) = t^{1,5}$ (blau). **(b) Rechts:** Entsprechende Visualisierungen für eine positiv Matrix, bei der ein Eigenwert größer und ein Eigenwert kleiner als 1 ist.

Wie oben erwähnt, können wir mittels Lagrange-Hermite-Interpolation $f(A)$ als Polynom in A ausdrücken. Mit Hilfe dieser Eigenschaft lässt sich die folgende Proposition beweisen.

Proposition 4.1.4. *Es sei f auf dem Spektrum einer symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ definiert. Dann gilt:*

- i) $Af(A) = f(A)A$,
- ii) wenn $XA = AX$ für eine Matrix X , dann ist $Xf(A) = f(A)X$,
- iii) $f(A^\top) = f(A)^\top$.

Beweis. Wir drücken $f(A)$ als Polynom p in A aus und schreiben $f(A) = p(A)$. (i) Dann ist $Af(A) = Ap(A)$ und da A mit $p(A)$ kommutiert, ist $Ap(A) = p(A)A = f(A)A$. Zum Beweis von (ii) können wir ähnlich argumentieren. Kommutiert X mit A , dann auch mit $p(A)$, sodass wir $Xf(A) = Xp(A) = p(A)X = f(A)X$ erhalten. Für (iii) verwenden wir $p(A)^\top = p(A^\top)$ und erhalten $f(A^\top) = p(A^\top) = p(A)^\top = f(A)^\top$. \square

Ungleichungen für reellwertige Funktionen bezüglich der Ordnung " \leq " ergeben auch die entsprechenden Ungleichungen für Funktionen symmetrischer Matrizen bezüglich der Löwner-Ordnung.

Proposition 4.1.5. *Es seien $f: \mathbb{R} \rightarrow \mathbb{R}$ und $g: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) \leq g(x)$ für alle $x \in [a, b] \subseteq \mathbb{R}$. Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch mit $\rho(A) \in [a, b]$. Dann ist $f(A) \preceq g(A)$.*

Beweis. Es sei $A = S \text{diag}(\lambda_1, \dots, \lambda_n) S^\top$ eine Eigenwertzerlegung von A . Die Aussage $f(A) \preceq g(A)$ ist äquivalent zu $f(A) - g(A) \preceq 0$ und damit zu

$$x^\top (f(A) - g(A))x \leq 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

Dies ist äquivalent zu

$$x^\top S \begin{pmatrix} f(\lambda_1) - g(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) - g(\lambda_n) \end{pmatrix} S^\top x \leq 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}. \quad (4.3)$$

Wir setzen $S^\top x = y$ und da S nicht singulär ist, ist (4.3) äquivalent zu

$$y^\top \begin{pmatrix} f(\lambda_1) - g(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) - g(\lambda_n) \end{pmatrix} y \leq 0 \quad \forall y \in \mathbb{R}^n \setminus \{0\}. \quad (4.4)$$

Die linke Seite von (4.4) ist gerade $\sum_{i=1}^n y_i^2 (f(\lambda_i) - g(\lambda_i))$ und damit nicht positiv. \square

4.2 Visualisierungen von Matrixfunktionen

In Abbildung 4.1 sind positiv definite Matrizen und ihre positiv definiten Funktionswerte unter verschiedenen Funktionen als Ellipsen visualisiert. In diesem Abschnitt betrachten wir Visualisierungen für Funktionen symmetrischer 2×2 -Matrizen, die auch nicht positiv definite Matrizen und nicht positiv definite Funktionswerte mit einschließen und zugleich größere Mengen symmetrischer 2×2 -Matrizen erfassen. Dazu betrachten wir Teilmengen der symmetrischen 2×2 -Matrizen und deren Funktionswerte und visualisieren diese über die Abbildungsvorschrift (3.6) als Punkte im \mathbb{R}^3 . Geeignete Teilmengen bilden zum Beispiel symmetrische 2×2 -Matrizen mit gleicher Spur oder mit gleicher Determinante. Diese Möglichkeiten der Visualisierung werden im Folgenden näher untersucht.

4.2.1 Visualisierung von Matrixfunktionen in Abhängigkeit der Spur

Matrizen mit gleicher Spur s liegen auf einer Ebene parallel zur x-y-Ebene, deren z-Koordinate proportional zur Spur s ist. Abbildung 4.2 zeigt links die Ebene der Matrizen mit Spur 3 und rechts die Fläche, die von den Funktionswerten dieser Matrizen unter der Funktion $f(t) = t^2$ gebildet wird.

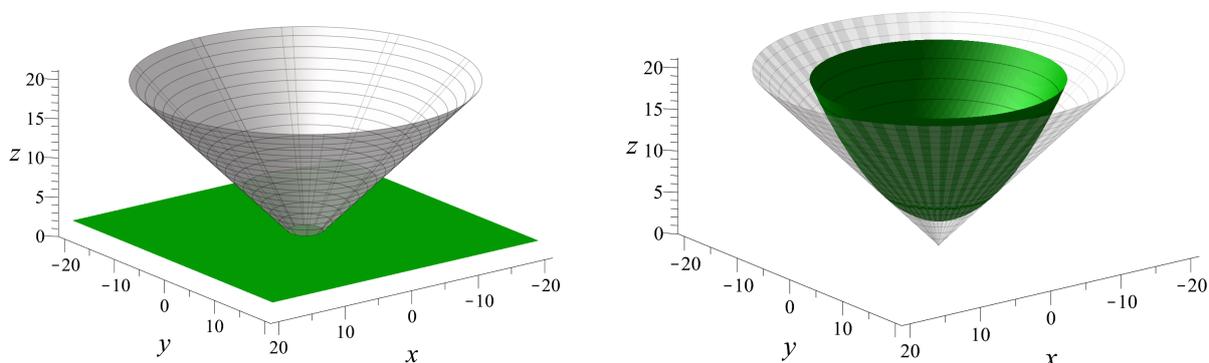


Abbildung 4.2: **(a) Links:** Die Fläche der Matrizen mit Spur 3 (grün) und der Kegel der positiv definiten Matrizen. **(b) Rechts:** Die zur z-Achse rotationssymmetrische grüne Fläche wird von den Bildern der Matrizen mit Spur 3 unter der Funktion $f(t) = t^2$ gebildet. Die Bilder liegen alle im Innern des Kegels der positiv (semi)definiten Matrizen oder auf dessen Kegelwand.

Wie wir in Abbildung 4.2 erkennen können, liegt die Fläche, die von Matrizen mit Spur 3 unter der Quadratfunktion gebildet wird, rotationssymmetrisch zur z-Achse. Eine entsprechende Symmetrie trifft auf alle Bilder von Matrizen gleicher Spur unter einer Funktion f zu. Um dies zu erklären, betrachten wir die in Abschnitt 3.2.4 untersuchte Lage von Matrizen mit den gleichen Eigenwerten λ und μ . Diese liegen rotationssymmetrisch zur

z-Achse. Die Funktionswerte dieser Matrizen haben gemäß Proposition 4.1.3 alle die Eigenwerte $f(\lambda)$ und $f(\mu)$ und liegen daher wieder auf einem Kreis um die z-Achse mit konstanter z-Koordinate. Da sich zudem die Eigenvektoren der Matrizen durch Anwendung einer Funktion nicht ändern, können wir aus der in Abschnitt 3.2.5 hergeleiteten Lage von Matrizen gleicher Eigenvektoren folgern, dass der gesamte Kreis der Matrizen mit den Eigenwerten $f(\lambda)$ und $f(\mu)$ das Bild der Matrizen mit den Eigenwerten λ und μ ist. Damit folgt, dass die Funktionswerte der Matrizen, die zur z-Achse rotationssymmetrische Flächen bilden, selbst wieder rotationssymmetrisch zur z-Achse liegen.

Abbildung 4.3 zeigt links die Bilder der Matrizen mit Spur 3, 8 und 13 unter der Funktion $f(t) = t^2$. Die Darstellung der Bilder von Matrizen verschiedener Spuren wird im \mathbb{R}^3 schnell unübersichtlich. Auf Grund der Rotationssymmetrie ist die Betrachtung der Schnitte der Flächen mit der y-z-Ebene jedoch ausreichend. Die entsprechenden Flächen ergeben sich dann durch Rotation der Schnittkurven um die z-Achse. Auf diese Weise lassen sich Bilder von Matrizen verschiedener Spuren übersichtlich darstellen. So zeigt Abbildung 4.3 rechts die Schnittkurven der Bilder der Matrizen mit Spur $s = 7, 8, \dots, 14$ unter der Quadratfunktion. Abbildung 4.4 zeigt eine Visualisierung der Cosinusfunktion $f(t) = \cos t$ in Abhängigkeit der Spur. In Abbildung 4.5 sind entsprechende Ergebnisse für die Funktion $f(t) = \sqrt{t}$ dargestellt.

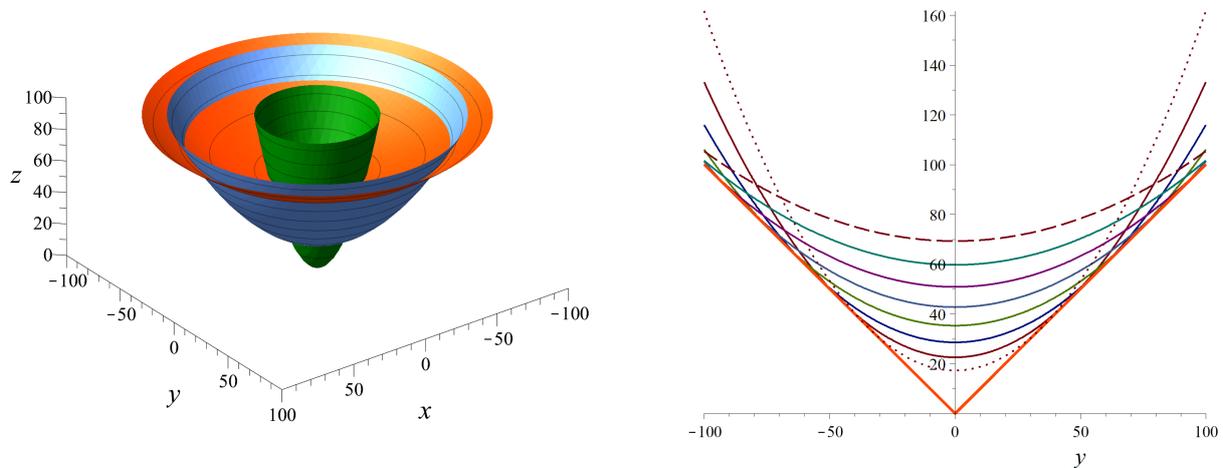


Abbildung 4.3: **(a) Links:** Bilder der Matrizen mit Spur 3 (grün), 8 (blau) und 13 (rot) unter der Funktion $f(t) = t^2$. **(b) Rechts:** Schnittkurven der Quadrate der Matrizen mit Spur $s = 7$ (gepunktet) bis $s = 14$ (gestrichelt) mit der y-z-Ebene. In orange ist der Schnitt des Kegels der positiv (semi)definiten Matrizen mit der y-z-Ebene dargestellt.

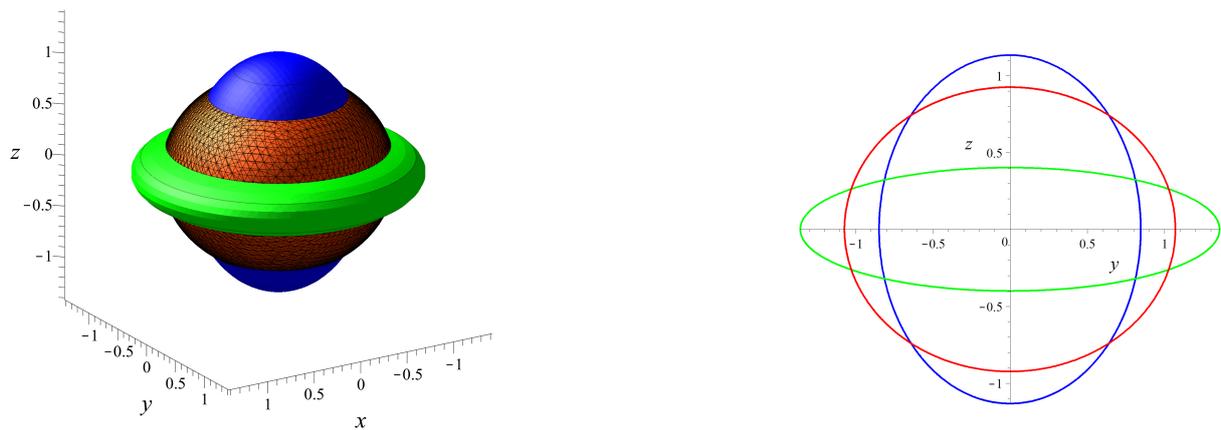


Abbildung 4.4: **(a) Links:** Flächen, die durch die Bilder der Matrizen mit Spur 5 (blau), 8 (orange) und 10 (grün) unter der Cosinusfunktion gebildet werden. **(b) Rechts:** Die Schnitte der Flächen aus (a) mit der y - z -Ebene.

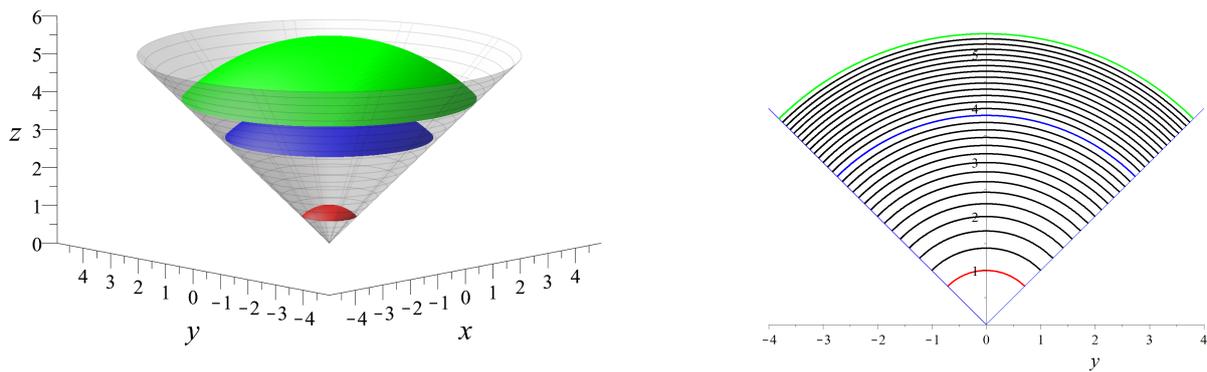


Abbildung 4.5: **(a) Links:** Die Bilder der Matrizen mit Spur 1 (rot), Spur 15 (blau) und Spur 29 (grün) unter der Wurzelfunktion. **(b) Rechts:** Schnitte der y - z -Ebene mit den Bildern der Matrizen mit Spur $1, 2, \dots, 29$ unter der Wurzelfunktion. Die Bilder der Matrizen aus (a) sind entsprechend farblich markiert.

4.2.2 Visualisierung von Matrixfunktionen in Abhängigkeit der Determinante

Auch Matrizen mit gleicher Determinante bilden Flächen, die rotationssymmetrisch zur z-Achse liegen, sodass wir Matrixfunktionen in Abhängigkeit ihrer Determinante in ähnlicher Weise visualisieren können. Auch ihre Bilder liegen symmetrisch zur z-Achse. So zeigt Abbildung 4.6 links die Exponentialfunktion für Matrizen gleicher positiver Determinante und rechts die Funktionswerte von Matrizen gleicher negativer Determinante unter der Funktion $f(t) = 0.25t^3$. Da die Determinante das Produkt der Eigenwerte ist, bilden Funktionen, welche die Funktionalgleichung

$$f(ab) = f(a)f(b)$$

erfüllen, Matrizen mit gleicher Determinante wieder auf Matrizen gleicher Determinante ab. In diesem Fall gilt

$$\det(f(A)) = f(\det(A)).$$

Diese Eigenschaft hat zum Beispiel die Funktion $f(t) = 0.25t^3$ in Abbildung 4.6 (b).

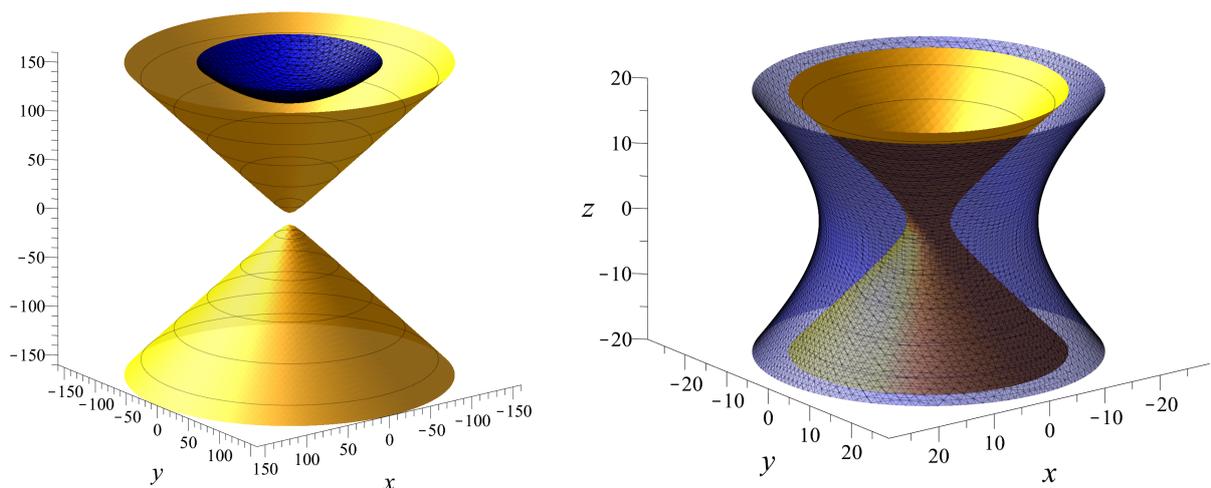


Abbildung 4.6: **(a) Links:** Matrizen mit Determinante 20 (gelb) und ihre Matrixexponentiale (blau). **(b) Rechts:** Matrizen mit Determinante -10 (gelb) und ihre Funktionswerte unter der Funktion $f(t) = 0.25t^3$ (blau).

In Abbildung 4.7 sind die Schnitte der Bilder mit Determinante 10 mit der y-z-Ebene unter der Cosinusfunktion dargestellt. In Abbildung 4.8 finden sich die entsprechenden Ergebnisse für Matrizen mit Determinante 50.

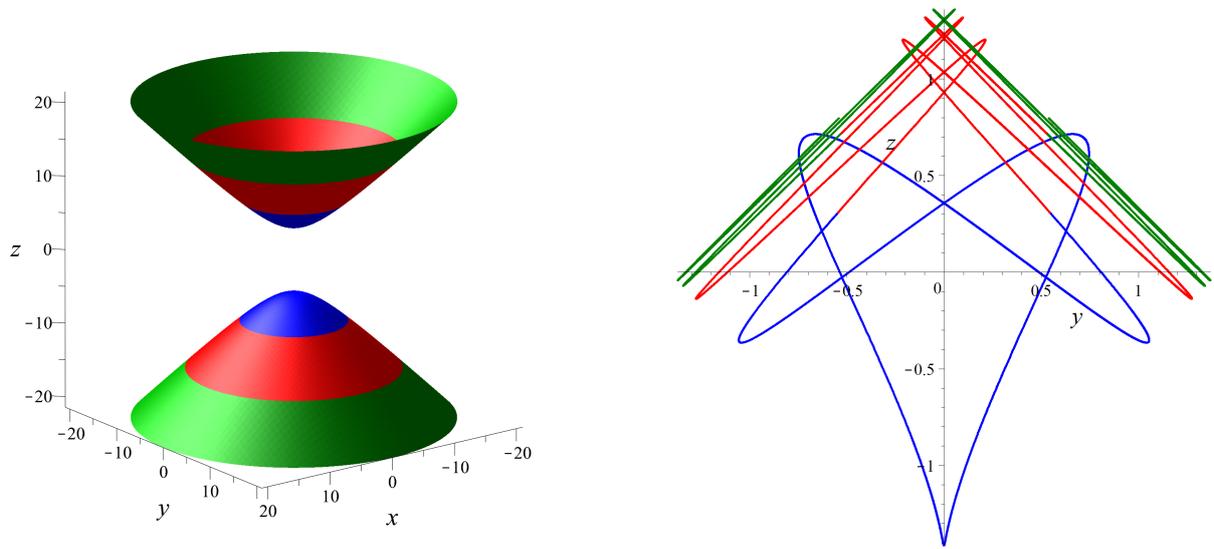


Abbildung 4.7: **(a) Links:** Matrizen mit Determinante 10. **(b) Rechts:** Schnitte der y - z -Ebene mit den Bildern der Matrizen mit Determinante 10 unter der Cosinusfunktion. Der Zusammenhang zwischen Bild und Urbild ist farblich gekennzeichnet.

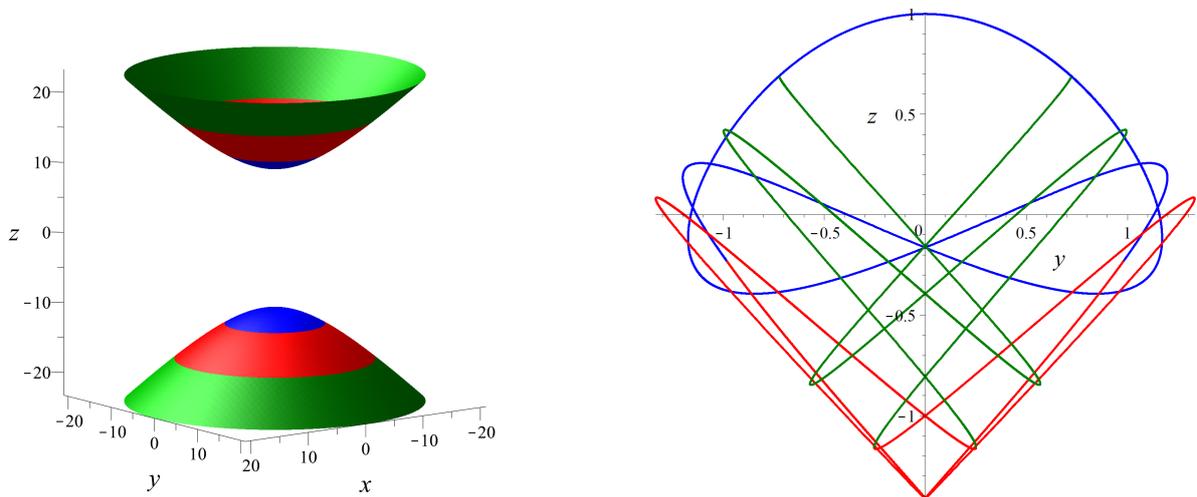


Abbildung 4.8: **(a) Links:** Matrizen mit Determinante 50. **(b) Rechts:** Schnitte der y - z -Ebene mit den Bildern der Matrizen mit Determinante 50 unter der Cosinusfunktion.

4.3 Polarzerlegung einer symmetrischen Matrix

Wir können Matrixfunktionen nutzen, um Zerlegungen symmetrischer Matrizen zu berechnen. Exemplarisch betrachten wir die Polarzerlegung einer nicht singulären symmetrischen Matrix. Diese zerlegt eine reguläre symmetrische Matrix eindeutig in ein Produkt aus einer positiv definiten Matrix und einer symmetrischen Orthogonalmatrix.

4.3.1 Berechnung und Visualisierung

Zur Untersuchung der Polarzerlegung benötigen wir den Begriff der Hauptquadratwurzel $A^{1/2}$ einer positiv semidefiniten Matrix A . Diese ist definiert als der Funktionswert von A unter der Funktion $f: \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$, $t \mapsto \sqrt{t}$ und somit als

$$A^{1/2} := S \operatorname{diag} \left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n} \right) S^\top,$$

wobei $A = S \operatorname{diag}(\lambda_1, \dots, \lambda_n) S^\top$ eine Eigenwertzerlegung von A ist. Der Begriff der Hauptquadratwurzel dient der Abgrenzung zu weiteren Quadratwurzeln einer Matrix. Denn eine Quadratwurzel X von $A \in \mathbb{R}^{n \times n}$ ist jede Lösung der Matrixgleichung $X^2 = A$. So hat zum Beispiel die Matrix

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$$

neben ihrer Hauptquadratwurzel noch die Quadratwurzeln

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \text{ und } \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Satz 4.3.1 (Polarzerlegung). *Es sei $A \in \mathbb{R}^{n \times n}$ nicht singulär. Dann existiert eine eindeutig bestimmte positiv definite Matrix $P \in \mathbb{R}^{n \times n}$ und eine eindeutig bestimmte symmetrische Orthogonalmatrix $Q \in \mathbb{R}^{n \times n}$, sodass*

$$A = PQ,$$

wobei $P = (A^2)^{1/2}$ und $Q = P^{-1}A$.

Beweis. Die Matrix A^2 ist symmetrisch und positiv definit, sodass die Hauptquadratwurzel $(A^2)^{1/2}$ definiert ist. Die Matrix $Q := P^{-1}A$ ist orthogonal, denn sie erfüllt

$$Q^\top Q = (P^{-1}A)^\top P^{-1}A = A^\top P^{-2}A = A(A^{-2})A = I.$$

Da die Matrix P^{-1} gemäß Proposition 4.1.4 mit A kommutiert, ist Q symmetrisch. Dies zeigt die Existenz der geforderten Zerlegung. Zum Beweis der Eindeutigkeit betrachten wir eine weitere passende Zerlegung $A = P'Q'$. Es sei

$$N := P^{-1}P' = Q(Q')^{-1}.$$

Dann ist N eine orthogonale Matrix und sie ist ähnlich zu $(P')^{1/2}P^{-1}(P')^{1/2}$. Dieses Produkt ist offenbar symmetrisch und positiv definit, denn setzen wir $y := (P')^{1/2}x$, dann ist

$$x^\top((P')^{1/2}P^{-1}(P')^{1/2})x = y^\top P^{-1}y > 0$$

für alle $x \in \mathbb{R}^n \setminus \{0\}$. Damit ist N diagonalisierbar und hat, da sie ähnlich zu einer positiv definiten Matrix ist, nur positive Eigenwerte. Als positiv definite Orthogonalmatrix müssen nach Proposition 2.3.9 ihre Eigenwerte alle den Wert 1 haben. Wir erhalten mit einer orthogonalen Matrix $S \in \mathbb{R}^{n \times n}$ die Eigenwertzerlegung $N = SIS^\top = SS^\top = I$. Aus der Definition von N als $N = P^{-1}P' = Q(Q')^{-1}$ folgt daher $P = P'$ und $Q = Q'$ und so die Eindeutigkeit der Polarzerlegung. \square

Beispiel 4.3.2. Wir berechnen die Polarzerlegung von $A = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$. Wir erhalten

$$P = (A^2)^{1/2} = \frac{1}{\sqrt{17}} \begin{pmatrix} 7 & 6 \\ 6 & 10 \end{pmatrix} \quad \text{und} \quad Q = P^{-1}A = \frac{1}{\sqrt{17}} \begin{pmatrix} -1 & 4 \\ 4 & 1 \end{pmatrix}.$$

Damit hat A die Polarzerlegung

$$A = PQ = \frac{1}{\sqrt{17}} \begin{pmatrix} 7 & 6 \\ 6 & 10 \end{pmatrix} \frac{1}{\sqrt{17}} \begin{pmatrix} -1 & 4 \\ 4 & 1 \end{pmatrix}.$$

Satz 4.3.1 über die Polarzerlegung besagt, dass die Menge der positiv definiten Matrizen und die Menge der symmetrischen Orthogonalmatrizen ausreicht, um eine nicht singuläre symmetrische Matrix darzustellen. Dies ist in Abbildung 4.9 (a) für symmetrische 2×2 -Matrizen visualisiert. Abbildung 4.9 (b) zeigt die Polarzerlegung der Matrix aus Beispiel 4.3.2.

In Abbildung 4.9 ist zum einen auffällig, dass A , P und Q in einer Ebene mit der z-Achse liegen. Der Grund hierfür ist, dass bei einer Polarzerlegung $A = PQ$ die Matrizen A , P und Q die gleichen Eigenvektoren haben. Denn P ist der Funktionwert von A unter der Funktion $f(t) = (t^2)^{1/2} = |t|$, womit nach Proposition 4.1.3 die Matrizen P und auch P^{-1} die gleichen Eigenvektoren wie A besitzen. Als Produkt der symmetrischer Matrizen P^{-1} und A hat damit auch Q die gleichen Eigenvektoren wie A .

Eine weitere Besonderheit bei der visuellen Darstellung der Polarzerlegung einer nicht singulären symmetrischen 2×2 -Matrix ist, dass die positiv definite Matrix P dem Spiegelbild von A am Kegel der positiv (semi)definiten Matrizen entspricht, wenn A indefinit ist. Zum Beweis dieser Eigenschaft können wir ohne Einschränkung eine indefinite Diagonalmatrix der Form

$$A = \begin{pmatrix} -\lambda & 0 \\ 0 & \mu \end{pmatrix}$$

mit $\lambda, \mu > 0$ betrachten. Diese liegt im Punkt

$$a = \frac{1}{\sqrt{2}}(0, \mu + \lambda, \mu - \lambda)^\top$$

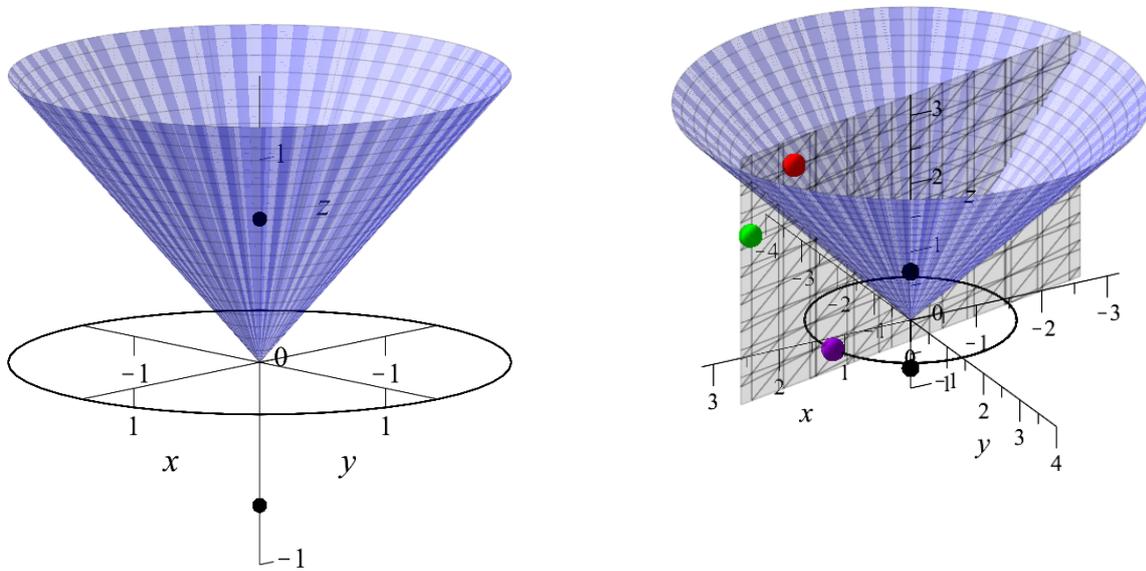


Abbildung 4.9: **(a) Links:** Alle nicht singulären symmetrischen 2×2 -Matrizen lassen sich eindeutig in ein Produkt aus einer positiv definiten Matrix (im Inneren des blauen Kegels) und einer symmetrischen Orthogonalmatrix (Matrix auf dem schwarzen Kreis, bzw. schwarz gekennzeichnete positive oder negative Einheitsmatrix) zerlegen. **(b) Rechts:** Visualisierung der Polarzerlegung der Matrix A aus Beispiel 4.3.2: Die grün eingezeichnete Matrix A lässt sich eindeutig in das Produkt der positiv definiten Matrix P (rot) und der orthogonalen Matrix Q (violett) zerlegen. Die Matrizen A , P und Q liegen in einer Ebene mit der z -Achse. P ist die Spiegelung von A am Kegel der positiv definiten Matrizen.

und die zu ihrer Polarzerlegung gehörende positiv definite Matrix

$$P = (A^2)^{1/2} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

im Punkt

$$p = \frac{1}{\sqrt{2}}(0, \mu - \lambda, \mu + \lambda)^T.$$

Die y - und z -Koordinaten von a und p sind also vertauscht. Der Punkt P entspricht daher der Spiegelung von a an der Geraden $y = |z|$, die den Schnitt des Kegels der positiv definiten Matrizen mit der y - z -Ebene beschreibt. Die Aussage für allgemein indefinite 2×2 -Matrizen folgt aus der Lage von Matrizen gleicher Eigenwerte und aus der Lage von Matrizen gleicher Eigenvektoren. Alle Matrizen A' mit den Eigenwerten $-\lambda$ und μ liegen zusammen mit A auf einem Kreis symmetrisch zur z -Achse. Die zu ihren Polarzerlegungen gehörigen Matrizen P' liegen gemeinsam mit P auf einem symmetrisch zur z -Achse angeordneten Kreis. Den vorausgegangenen Überlegungen zufolge ist dieser Kreis der Matrizen P' das Spiegelbild des Kreises der Matrizen A' am Kegel der positiv definiten Matrizen. Da Matrizen mit gleichen Eigenvektoren auf einer Ebene liegen, die senkrecht auf der

x-y-Ebene steht und durch die z-Achse verläuft und bei der Polarzerlegung $A = PQ$ die beteiligten Matrizen die gleichen Eigenvektoren haben, steht die Verbindungsgerade zwischen A und der positiv definiten Matrix P stets senkrecht auf der Kegelwand. Damit muss P das Spiegelbild von A am Kegel sein.

Ist eine Matrix A positiv definit, so ist die zu ihrer Polarzerlegung gehörende positiv definite Matrix P die Matrix A selbst. Die orthogonale Matrix ist in diesem Fall die Einheitsmatrix. Bei einer negativ definiten Matrix A ist die positiv definite Matrix P das Spiegelbild von A an der x-y-Ebene und die orthogonale Matrix Q die negative Einheitsmatrix. Abbildung 4.10 zeigt die Polarzerlegungen verschiedener symmetrischer 2×2 -Matrizen. Die oben beschriebenen Eigenschaften sind darin erkennbar.

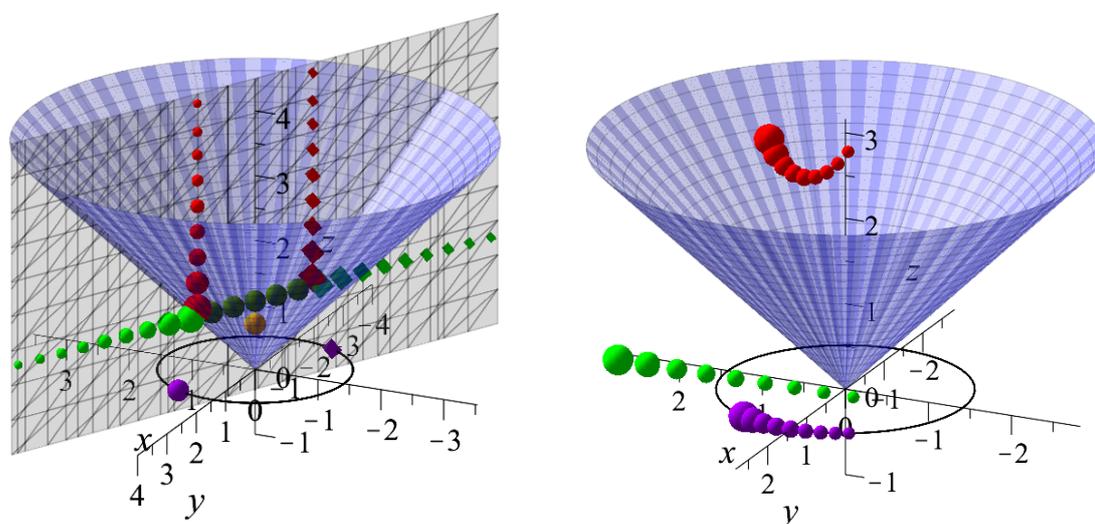


Abbildung 4.10: **(a) Links:** Polarzerlegungen der grün dargestellten Matrizen. Jede der indefiniten Matrizen (hellgrün) lässt sich eindeutig in ein Produkt aus einer rot markierten positiv definiten Matrix und einer der violett eingezeichneten Orthogonalmatrizen zerlegen. Die zu einer Zerlegung gehörenden Matrizen sind dabei an dem entsprechenden Symbol und der entsprechenden Größe erkennbar. Bei der Polarzerlegung der dunkelgrünen positiv definiten Matrizen ist die positiv definite Matrix die Matrix selbst, die orthogonale Matrix ist die Einheitsmatrix (gelb). **(b) Rechts:** Polarzerlegungen der grün dargestellten Matrizen in positiv definite Matrizen (rot) und orthogonale Matrizen (violett).

Die Polarzerlegung lässt sich auch auf singuläre Matrizen und auf nichtsymmetrische Matrizen ausdehnen. Diese Aspekte werden unter anderem in [9] und [13] thematisiert.

4.3.2 Einfluss von Matrixfunktionen auf die Polarzerlegung

Wir untersuchen den Einfluss von Matrixfunktionen auf die Polarzerlegung einer invertierbaren symmetrischen Matrix A . Hat A die Polarzerlegung $A = PQ$ und $f(A)$ die Polarzerlegung $f(A) = P'Q'$ mit P, P' positiv definit und Q, Q' orthogonal, untersuchen wir den Zusammenhang zwischen P und P' sowie zwischen Q und Q' . Wir betrachten dabei nur solche Funktionen, die keinen Eigenwert der Matrix auf Null abbilden. Ist die Eigenwertzerlegung von A gegeben durch $A = S \operatorname{diag}(\lambda_1, \dots, \lambda_n) S^\top = S \operatorname{diag}(\lambda_i) S^\top$, so ist

$$P = (A^2)^{1/2} = S \operatorname{diag}(|\lambda_i|) S^\top$$

und

$$Q = P^{-1}A = S \operatorname{diag}\left(\frac{\lambda_i}{|\lambda_i|}\right) S^\top = S \operatorname{diag}(\operatorname{sgn}(\lambda_i)) S^\top.$$

Für die Matrizen P' und Q' der Polarzerlegung von $f(A) = S \operatorname{diag}(f(\lambda_i)) S^\top$ ergeben sich dementsprechend

$$P' = (f(A)^2)^{1/2} = S \operatorname{diag}(|f(\lambda_i)|) S^\top$$

und

$$Q' = (P')^{-1}f(A) = S \operatorname{diag}\left(\frac{f(\lambda_i)}{|f(\lambda_i)|}\right) S^\top = S \operatorname{diag}(\operatorname{sgn}(f(\lambda_i))) S^\top.$$

Damit erhalten wir

$$P' = S \operatorname{diag}\left(\frac{|f(\lambda_i)|}{|\lambda_i|}\right) S^\top P$$

und

$$Q' = S \operatorname{diag}\left(\frac{\operatorname{sgn}(f(\lambda_i))}{\operatorname{sgn}(\lambda_i)}\right) S^\top Q.$$

Die orthogonalen Matrizen Q und Q' sind daher identisch, wenn die Funktion f die Vorzeichen der Eigenwerte λ_i nicht ändert.

4.4 Monotone Matrixfunktionen

Der Begriff der monotonen Matrixfunktion geht auf Karl Löwner zurück, der sich 1934 in [12] mit der Monotonie von Matrixfunktionen beschäftigte. Mit Hilfe der nach ihm benannten Löwner-Ordnung führte er den Monotoniebegriff für Matrixfunktionen in Analogie zur Monotonie skalarer Funktionen ein.

Definition 4.4.1. Es seien $A, B \in \mathbb{R}^{n \times n}$ symmetrisch mit $A \preceq B$. Dann heißt eine reelle Funktion f *monoton wachsend von n -ter Stufe*, wenn stets auch

$$f(A) \preceq f(B)$$

gilt und *monoton fallend von n -ter Stufe*, wenn stets

$$f(A) \succeq f(B)$$

erfüllt ist.

Die Monotonie erster Stufe entspricht gerade der gewöhnlichen Monotonie einer Funktion. Da eine Funktion f genau dann monoton fallend ist, wenn die Funktion $-f$ monoton wächst, beschränken wir uns bei der Untersuchung der Monotonie von Matrixfunktionen auf monoton wachsende Funktionen. Eine monoton wachsende Funktion von n -ter Stufe wird im Folgenden als *matrixmonotone Funktion n -ter Stufe* bezeichnet. Ist eine Funktion f für alle n matrixmonoton von n -ter Stufe, so heißt f *matrixmonoton*.

Beispiel 4.4.2. 1) Die Funktion $f(t) = \alpha t + \beta$ mit $\alpha, \beta \in \mathbb{R}$ und $\alpha \geq 0$ ist matrixmonoton. Denn sind $A, B \in \mathbb{R}^{n \times n}$ symmetrisch mit $A \preceq B$, so ist $x^\top(A - B)x \leq 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$, womit wir auch $x^\top(f(A) - f(B))x = x^\top((\alpha A + \beta I_n) - (\alpha B + \beta I_n))x = \alpha x^\top(A - B)x \leq 0$ erhalten. Damit ist $f(A) \preceq f(B)$.

2) Die Funktion $f(t) = -\frac{1}{t}$ ist matrixmonoton auf $(0, \infty)$, denn für positiv definite Matrizen $A, B \in \mathbb{R}^{n \times n}$ ist nach Satz 2.4.7 $A \preceq B$ genau dann, wenn $A^{-1} \succeq B^{-1}$ ist. Dies ist äquivalent zu $-A^{-1} \preceq -B^{-1}$ und damit zu $f(A) \preceq f(B)$.

Die Monotonie der zugrundeliegenden reellen Funktion (Matrixmonotonie erster Stufe) ist offenbar eine notwendige Bedingung für die Matrixmonotonie höherer Stufe und damit für die Matrixmonotonie einer Funktion. Wie Beispiel 4.4.3 zeigt, ist die Monotonie erster Stufe jedoch keine hinreichende Bedingung für das Vorliegen von Matrixmonotonie höherer Stufen.

Beispiel 4.4.3. Die Funktion $f(t) = t^2$ ist auf $[0, \infty)$ monoton von erster Stufe, aber nicht matrixmonoton von zweiter Stufe und damit nicht matrixmonoton.

Betrachte die Matrizen

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

deren Spektren in $[0, \infty)$ liegen. Die Differenz der beiden Matrizen ergibt sich zu

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

sodass $A \preceq B$ ist. Wäre die Funktion f matrixmonoton von zweiter Stufe, so müsste $f(A) \preceq f(B)$ gelten. Es ist

$$f(A) = A^2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \quad \text{und} \quad f(B) = B^2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}.$$

Die Matrix

$$f(B) - f(A) = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$$

ist indefinit. Daher sind die Matrizen $f(A)$ und $f(B)$ in der Löwner-Ordnung nicht vergleichbar, womit die Forderung $f(A) \preceq f(B)$ verletzt ist.

Wir visualisieren das Ergebnis aus Beispiel 4.4.3. Dazu betrachten wir die im Beispiel untersuchte Matrix A und ihre obere Penumbra

$$P^+(A) = \{N \in \text{Sym}(2) : N \succeq A\},$$

die in Abbildung 4.11 links dargestellt sind. Wir wenden die Funktion $f(t) = t^2$ auf die Matrix A und ihre obere Penumbra an und zeichnen zudem die obere Penumbra

$$P^+(f(A)) = \{N \in \text{Sym}(2) : N \succeq f(A)\}$$

von $f(A)$ ein (vgl. Abbildung 4.11 (b)). Wir erkennen, dass das Bild von $P^+(A)$ teilweise außerhalb von $P^+(f(A))$ liegt. Dies zeigt die Existenz positiv (semi)definiter Matrizen $B \in \mathbb{R}^{2 \times 2}$ mit $A \preceq B$, für die jedoch $f(A) = A^2 \preceq B^2 = f(B)$ nicht erfüllt ist. Diese Beobachtung machen wir auch bei Funktionen, die bereits im skalaren Fall nicht monoton sind. So zeigt Abbildung 4.12 ein entsprechendes Ergebnis für die Funktion $f(t) = t^2$, angewandt auf eine indefinite Matrix und ihre obere Penumbra und Abbildung 4.13 für die Funktion $f(t) = \cos t$.

Ist hingegen f matrixmonoton, liegen für jede symmetrische Matrix $A \in \mathbb{R}^{2 \times 2}$ im Definitionsbereich von f die Funktionswerte der obere Penumbra von A nie außerhalb von $P^+(f(A))$. Dies ist exemplarisch für die nach [2] matrixmonotone Funktion $f(t) = \sqrt{t}$ mit einer positiv semidefiniten Matrix und ihrer oberen Penumbra in Abbildung 4.14 dargestellt.

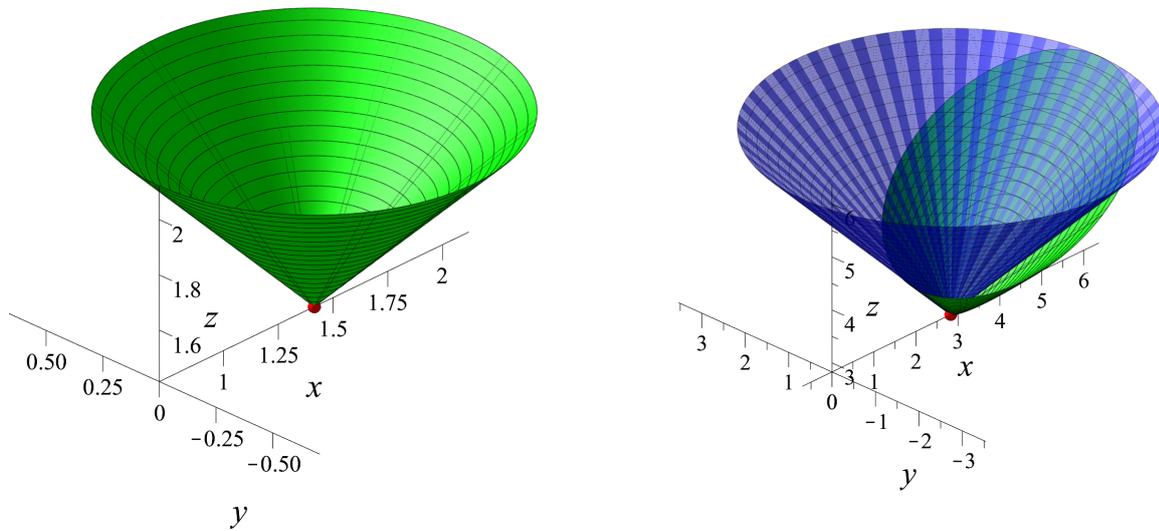


Abbildung 4.11: **(a) Links:** Die positiv semidefinite Matrix A aus Beispiel 4.4.3 (rot) und ihre obere Penumbra $P^+(A)$ (grün). **(b) Rechts:** Funktionswert von A unter $f(t) = t^2$ (rot) und in blau die obere Penumbra von $f(A)$. Die grüne Fläche beschreibt die Funktionswerte der oberen Penumbra von A unter der Quadratfunktion und liegt zum Teil außerhalb des blauen Kegels.

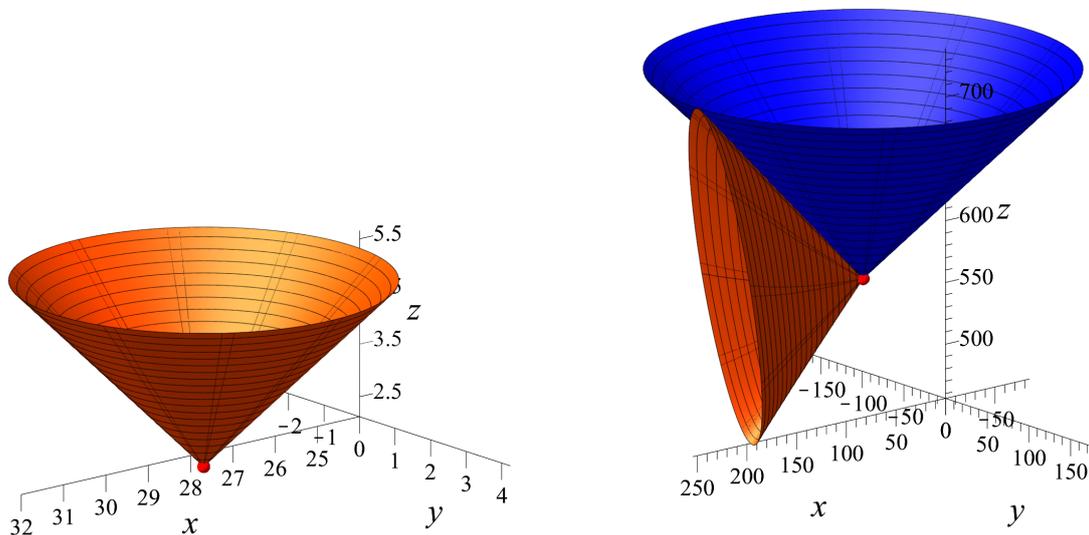


Abbildung 4.12: **(a) Links:** Indefinite Matrix A (rot) und in orange ihre obere Penumbra $P^+(A)$. **(b) Rechts:** Quadrat der Matrix A (rot) und ihrer oberen Penumbra (orange). In blau die obere Penumbra von A^2 .

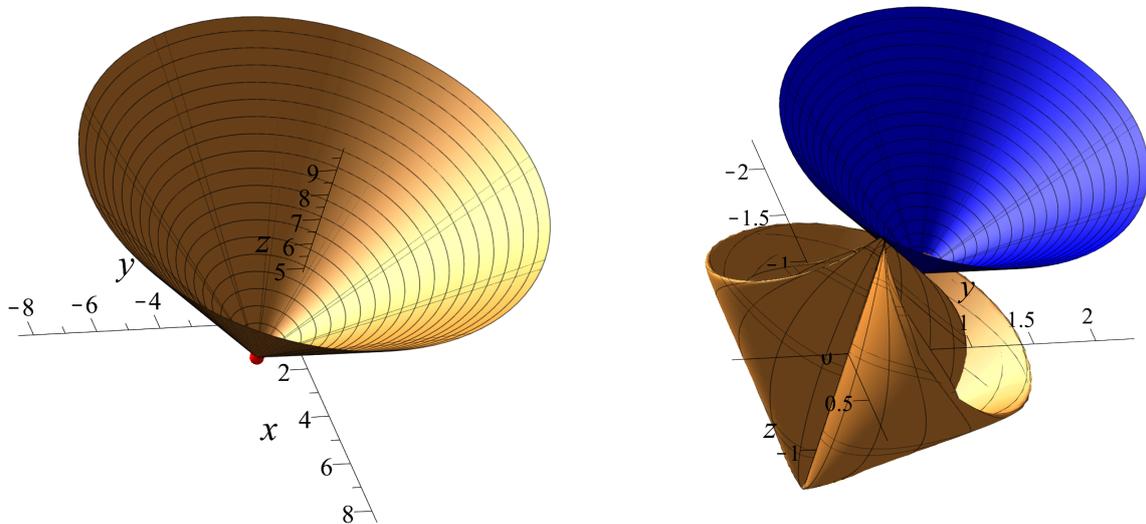


Abbildung 4.13: **(a) Links:** Visualisierung einer Matrix A (rot) mit ihrer oberen Penumbra (gold). **(b) Rechts:** Cosinus der oberen Penumbra von A (gold). In blau die obere Penumbra von $f(A)$.

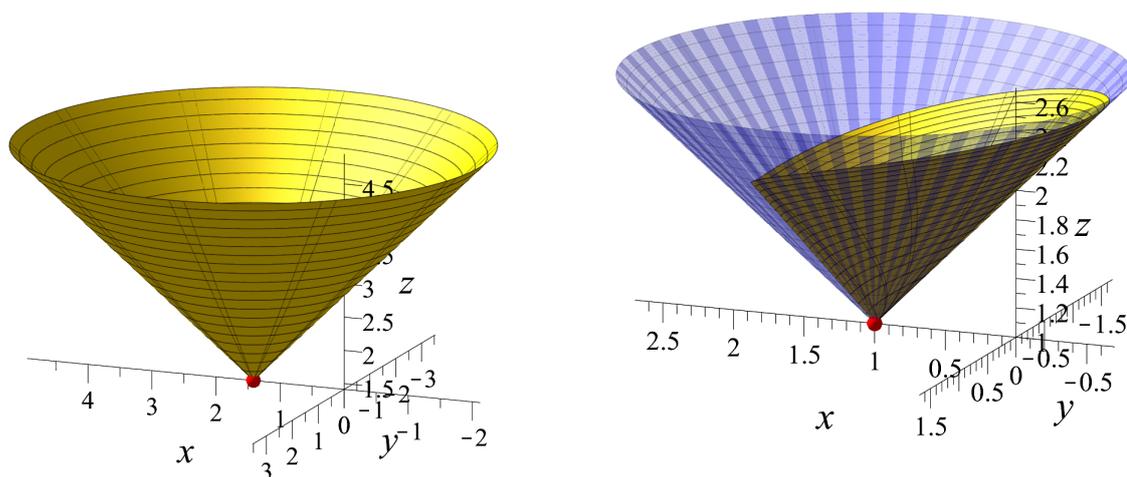


Abbildung 4.14: **(a) Links:** Positiv semidefinite Matrix A (rot) und ihre obere Penumbra (gelb). **(b) Rechts:** Quadratwurzel der positiv semidefiniten Matrix A (rot) und ihrer oberen Penumbra (gelb). Sie liegt nicht außerhalb der oberen Penumbra von $f(A)$ (blau).

Löwner untersuchte in seiner 1934 publizierten Arbeit [12] charakteristische Eigenschaften matrixmonotoner Funktionen. Er fand heraus, dass eine Funktion genau dann matrixmonoton ist, wenn sie zur Klasse der sogenannten Pick-Funktionen gehört. Diese Funktionen besitzen eine für sie charakteristische Integraldarstellung. Damit lässt sich eine sowohl notwendige, als auch hinreichende Bedingung für die Matrixmonotonie einer Funktion formulieren.

Satz 4.4.4. *Eine Funktion $f(t)$ ist genau dann matrixmonoton, wenn sie eine Integraldarstellung der Form*

$$f(t) = \alpha t + \beta + \int_{-\infty}^{\infty} \frac{1 + \lambda t}{\lambda - t} d\mu(\lambda)$$

besitzt, wobei $\alpha, \beta \in \mathbb{R}$ mit $\alpha \geq 0$ und μ ein positives Maß auf der reellen Achse ist.

Entsprechende Herleitungen können auch in [2] oder [6] nachgelesen werden.

Kapitel 5

Einfache Matrixfunktionen mehrerer Variablen

In Kapitel 4 wurden Matrixfunktionen einer Variablen behandelt. Der Funktionswert $f(A)$ wird dabei berechnet, indem die Funktion f auf die Eigenwerte von A angewendet wird. In diesem Kapitel betrachten wir Beispiele für Matrixfunktionen mit zwei Variablen. Die dabei auftretenden Funktionswerte $f(A, B)$ lassen sich nicht durch Anwendung einer Funktion auf das Spektrum von A und B berechnen. Zunächst untersuchen wir die Eigenschaften von Verknüpfungen zweier symmetrischer Matrizen in Form von Matrixprodukten, welche die Symmetrie der verknüpften Matrizen erhalten sowie die parallele Summe von zwei symmetrischen Matrizen. Abschließend erarbeiten wir Möglichkeiten, Supremum und Infimum symmetrischer Matrizen bezüglich der Löwner-Ordnung zu bestimmen.

5.1 Matrixprodukte und parallele Summe

Das gewöhnliche Matrixprodukt nach Definition 2.1.3 kann als Funktion zweier Matrizen aufgefasst werden. Es ist für symmetrische Matrizen $A, B \in \mathbb{R}^{n \times n}$ nach Satz 2.1.8 genau dann symmetrisch, wenn A und B kommutieren. In diesem Abschnitt betrachten wir alternative Produkte symmetrischer Matrizen, welche die Symmetrie der Matrizen auch im Falle einer Nichtkommutativität erhalten und für kommutative Matrizen mit dem gewöhnlichen Matrixprodukt übereinstimmen. Als weitere Verknüpfung untersuchen wir die parallele Summe symmetrischer Matrizen.

5.1.1 Jordan-Produkt

Das *Jordan-Produkt* von zwei symmetrischen Matrizen $A, B \in \mathbb{R}^{n \times n}$ ist definiert als

$$J(A, B) = A \bullet_J B := \frac{1}{2}(AB + BA). \quad (5.1)$$

Wir erkennen, dass es mit dem gewöhnlichen Matrixprodukt übereinstimmt, falls A und B kommutieren. Das Jordan-Produkt erhält die Symmetrie, denn

$$(A \bullet_J B)^\top = \frac{1}{2}(B^\top A^\top + A^\top B^\top) = \frac{1}{2}(BA + AB) = A \bullet_J B.$$

Es ist kommutativ, denn

$$A \bullet_J B = \frac{1}{2}(AB + BA) = \frac{1}{2}(BA + AB) = B \bullet_J A$$

und distributiv in beiden Argumenten, da

$$\begin{aligned} (A + B) \bullet_J C &= \frac{1}{2}((A + B)C + C(A + B)) \\ &= \frac{1}{2}(AC + CA) + \frac{1}{2}(BC + CB) \\ &= A \bullet_J C + B \bullet_J C \end{aligned}$$

und nach analoger Rechnung auch

$$A \bullet_J (B + C) = A \bullet_J B + A \bullet_J C$$

gilt. Das Assoziativgesetz wird vom Jordan-Produkt nicht erfüllt. Betrachten wir

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

dann ist

$$(A \bullet_J B) \bullet_J C = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix},$$

aber

$$A \bullet_J (B \bullet_J C) = \begin{pmatrix} 4 & 3 \\ 3 & 0 \end{pmatrix}.$$

Positive Definitheit bleibt durch das Jordan-Produkt nicht erhalten. So sind zum Beispiel

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

positiv definit, ihr Jordan-Produkt

$$A \bullet_J B = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}$$

indefinit. Entsprechendes erkennen wir auch in Abbildung 5.1. Dort wird das Jordan-Produkt einer positiv definiten Matrix mit der oberen Penumbra einer weiteren positiv definiten Matrix gebildet, sodass nur positiv definite Matrizen verknüpft werden. Die Funktionswerte liegen aber zum Teil im Bereich der indefiniten Matrizen.

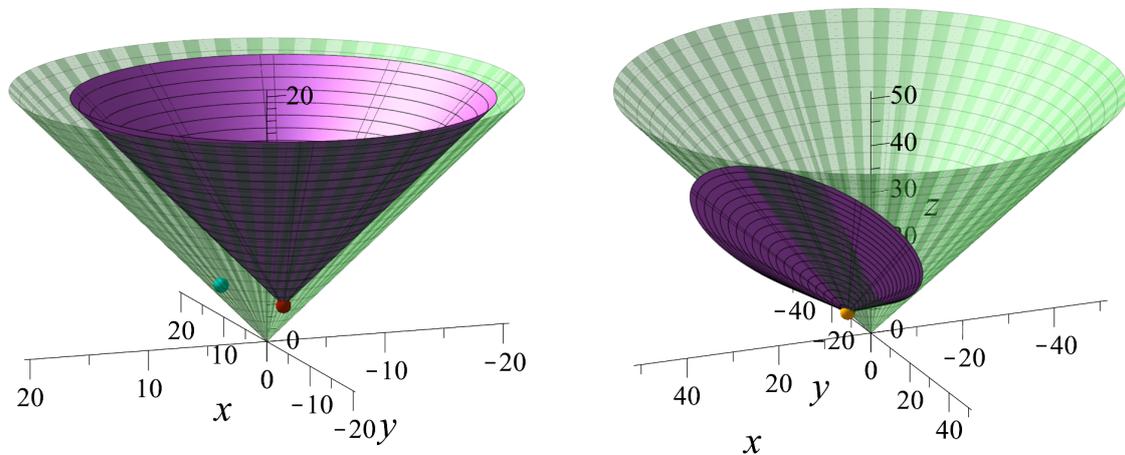


Abbildung 5.1: **(a) Links:** Der Kegel der positiv (semi)definiten Matrizen (grün) mit einer positiv definiten Matrix (rot) und ihrer oberen Penumbra (violet) sowie eine weitere positiv definite Matrix (cyan). **(b) Rechts:** Jordanprodukt der cyanen Matrix mit der roten Matrix (gelb) und ihrer oberen Penumbra (violet). Die Produkte liegen zum Teil außerhalb des Kegels der positiv (semi)definiten Matrizen.

5.1.2 P-Produkt

Für zwei symmetrische Matrizen $A, B \in \mathbb{R}^{n \times n}$ mit A positiv semidefinit ist das *P-Produkt* definiert als

$$P(A, B) = A \bullet_P B := A^{1/2} B A^{1/2}, \quad (5.2)$$

wobei $A^{1/2}$ die Hauptquadratwurzel von A bezeichnet. Das P-Produkt stammt aus der Numerik und wird zum Lösen großer Gleichungssysteme verwendet. Das P steht dabei für *preconditioning* (Vorkonditionieren). Auch das P-Produkt stimmt im Fall $AB = BA$ mit dem gewöhnlichen Matrixprodukt überein. Denn ist $AB = BA$, gilt nach Proposition 4.1.4 auch $A^{1/2} B = B A^{1/2}$, sodass $A^{1/2} B A^{1/2} = A^{1/2} A^{1/2} B = AB$ ist. Das P-Produkt erhält die Symmetrie, denn mit A ist auch $A^{1/2}$ symmetrisch und

$$(A \bullet_P B)^\top = (A^{1/2})^\top B^\top (A^{1/2})^\top = A^{1/2} B A^{1/2} = A \bullet_P B.$$

Im Gegensatz zum Jordanprodukt ist es nicht kommutativ. Dies zeigt das Beispiel der Matrizen

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}. \quad (5.3)$$

Als Hauptquadratwurzeln erhalten wir

$$A^{1/2} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{und} \quad B^{1/2} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

die Präkonditionsprodukte liefern

$$A \bullet_P B = \begin{pmatrix} 8 & 18 \\ 18 & 45 \end{pmatrix} \neq \begin{pmatrix} 13 & 22 \\ 22 & 40 \end{pmatrix} = B \bullet_P A.$$

Die fehlende Kommutativität ist auch in Abbildung 5.2 zu erkennen. Dort wird das P-Produkt einer einzelnen Matrix B mit einer Menge von Matrizen A_i gleicher Eigenwerte λ und μ gebildet, wobei $A_i \bullet_P B$ und $B \bullet_P A_i$ gemäß der Abbildungsvorschrift (3.6) als Punkte im \mathbb{R}^3 visualisiert sind. Neben der fehlenden Kommutativität erkennen wir auch, dass die Werte von $B \bullet_P A_i$ in einer Ebene liegen, was für die entsprechenden Werte von $A_i \bullet_P B$ nicht zutrifft.

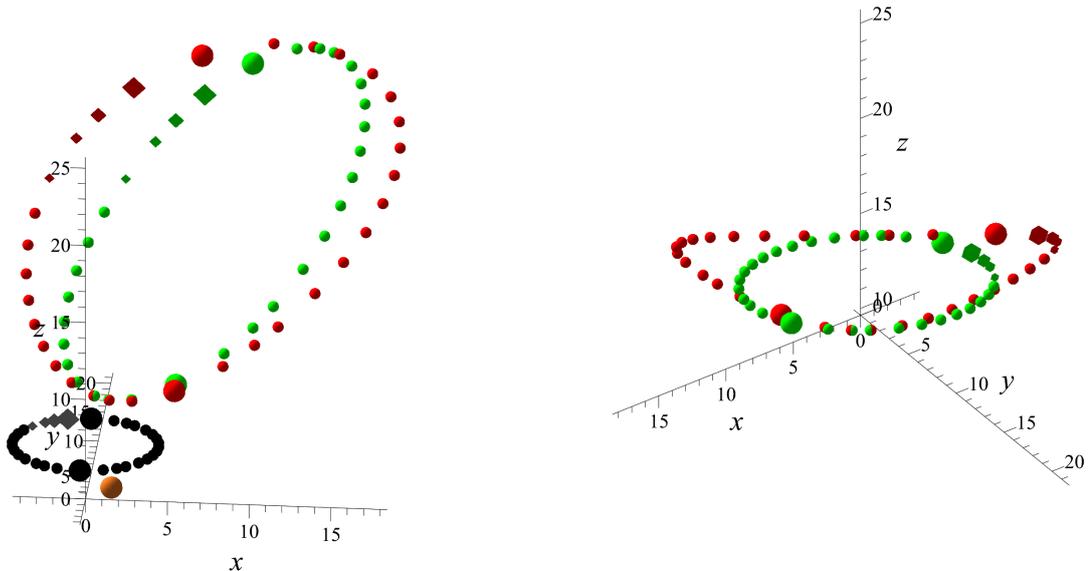


Abbildung 5.2: **(a) Links:** P-Produkt der braun eingezeichneten Matrix B mit den Matrizen A_i auf dem Kreis (schwarz), wobei $A_i \bullet_P B$ (rot) und $B \bullet_P A_i$ (grün) visualisiert sind. **(b) Rechts:** Die Produkte aus (a) aus anderer Perspektive. Die Werte von $B \bullet_P A_i$ (grün) liegen in einer Ebene. Für die Matrizen $A_i \bullet_P B$ (rot) trifft dies nicht zu.

Betrachten wir zu den in (5.3) definierten Matrizen A und B eine weitere Matrix

$$C = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix},$$

zeigt die Rechnung

$$(A \bullet_P B) \bullet_P C = \frac{1}{65} \begin{pmatrix} 4616 & 10002 \\ 10002 & 21969 \end{pmatrix} \neq \begin{pmatrix} 72 & 162 \\ 162 & 369 \end{pmatrix} = A \bullet_P (B \bullet_P C),$$

dass, wie bereits beim Jordan-Produkt, das Assoziativgesetz auch hier nicht gültig ist. Das Distributivgesetz gilt nur im zweiten Argument. Es ist

$$A \bullet_P (B + C) = A^{1/2}(B + C)A^{1/2} = A^{1/2}BA^{1/2} + A^{1/2}CA^{1/2} = A \bullet_P B + A \bullet_P C.$$

Im ersten Argument kann das Distributivgesetz allerdings nicht angewendet werden, wie das explizite Berechnen der Produkte $(A + B) \bullet_P C$ und $A \bullet_P C + B \bullet_P C$ mit den

zuvor betrachteten Matrizen zeigt. Im Gegensatz zum Jordan-Produkt bleibt die positive Definitheit beim P-Produkt erhalten, wenn A nicht singular und B positiv definit sind. Denn ist B positiv definit, so ist $y^\top B y > 0$ für alle $y \in \mathbb{R} \setminus \{0\}$. Auf Grund der positiven Definitheit von A ist $A^{1/2}x \neq 0$ für alle $x \in \mathbb{R} \setminus \{0\}$. Damit ist

$$x^\top (A^{1/2} B A^{1/2}) x = x^\top ((A^{1/2})^\top B A^{1/2}) x = \underbrace{(A^{1/2}x)^\top}_y B \underbrace{(A^{1/2}x)}_y > 0$$

für alle $x \in \mathbb{R} \setminus \{0\}$. Es lässt sich sogar zeigen, dass im Fall der Invertierbarkeit von A die Signatur von B erhalten bleibt. Dies bedeutet, dass B und $A^{1/2} B A^{1/2}$ die gleiche Anzahl positiver und negativer Eigenwerte sowie die gleiche Anzahl von Eigenwerten mit dem Wert Null haben. Dies ist eine Folgerung aus dem *Trägheitssatz von Sylvester*, dessen Herleitung in [10] oder [14] nachgelesen werden kann.

Satz 5.1.1 (Trägheitssatz von Sylvester). *Es sei C eine nichtsinguläre Matrix und $C^\top A C$ definiert. Dann hat die Matrix $C^\top A C$ die gleiche Anzahl positiver und negativer Eigenwerte wie A . Die Anzahl der Eigenwerte mit dem Wert Null stimmt ebenfalls überein.*

Daraus folgt unmittelbar die Behauptung, denn $A^{1/2} B A^{1/2} = (A^{1/2})^\top B A^{1/2}$.

5.1.3 Feynman-Produkt

Für positive reelle Zahlen a und b gilt die Identität

$$\frac{1}{ab} = \int_0^1 \frac{1}{(at + b(1-t))^2} dt,$$

denn

$$\begin{aligned} \int_0^1 \frac{1}{(at + b(1-t))^2} dt &= \int_0^1 \frac{1}{(t(a-b) + b)^2} dt \\ &= \left[\frac{-1}{(a-b)(t(a-b) + b)} \right]_{t=0}^{t=1} \\ &= -\frac{1}{(a-b)a} + \frac{1}{(a-b)b} \\ &= \frac{a-b}{(a-b)ab} = \frac{1}{ab}. \end{aligned}$$

Damit lässt sich das Produkt zweier positiver reeller Zahlen a, b ausdrücken als

$$ab = \int_0^1 \frac{1}{(a^{-1}t + b^{-1}(1-t))^2} dt. \quad (5.4)$$

Wir übertragen (5.4) auf Matrizen und definieren für positiv definite Matrizen $A, B \in \mathbb{R}^{n \times n}$ das *Feynman-Produkt*

$$F(A, B) = A \bullet_F B := \int_0^1 (A^{-1}t + B^{-1}(1-t))^{-2} dt. \quad (5.5)$$

Die Auswertung des Integrals zur Berechnung von Funktionswerten erfolgt mittels numerischer Integration. Dazu wird das Integral (5.4) für den skalaren Fall als Summe approximiert und die in der Summe auftretenden reellen Zahlen a und b durch die Matrizen A und B ersetzt. So liefert zum Beispiel die Simpson-3/8-Regel für 100 Intervalle

$$\begin{aligned}
A \bullet_F B \approx & \frac{1}{800} \sum_{i=0}^{99} \left(3 \left(A^{-1} \left(\frac{1}{100}i + \frac{1}{300} \right) + B^{-1} \left(\frac{299}{300} - \frac{1}{100}i \right) \right)^{-2} \right. \\
& + 3 \left(A^{-1} \left(\frac{1}{100}i + \frac{1}{150} \right) + B^{-1} \left(\frac{149}{150} - \frac{1}{100}i \right) \right)^{-2} \\
& + \left(A^{-1} \left(\frac{1}{100}i + \frac{1}{100} \right) + B^{-1} \left(\frac{99}{100} - \frac{1}{100}i \right) \right)^{-2} \\
& \left. + \left(\frac{1}{100}iA^{-1} + B^{-1} \left(1 - \frac{1}{100}i \right) \right)^{-2} \right).
\end{aligned}$$

Im Fall $AB = BA$ entspricht das Feynman-Produkt dem gewöhnlichen Matrixprodukt. Die Struktur des Integrals lässt erkennen, dass das Feynman-Produkt kommutativ ist. Die Symmetrie und auch die positive Definitheit der Matrizen A und B bleiben erhalten.

5.1.4 Vergleich der Matrixprodukte

Im Fall kommutierender Matrizen $AB = BA$ sind die betrachteten Matrixprodukte identisch. Da eine Matrix A sowohl mit der Einheitsmatrix, als auch im Falle der Existenz ihrer Inversen A^{-1} mit dieser kommutiert, gilt

$$A \circ I = I \circ A = A,$$

sowie

$$A \circ A^{-1} = A^{-1} \circ A = I$$

für jede der entsprechenden Verknüpfungen. Damit ist I das neutrale Element und im Fall der Existenz A^{-1} das inverse Element. In Tabelle 5.1 sind Eigenschaften der untersuchten Matrixprodukte zum Vergleich zusammengefasst. Dabei bedeutet "–", dass die entsprechende Eigenschaft im Allgemeinen nicht erfüllt ist, wohingegen ein "+" für die Allgemeingültigkeit der jeweiligen Eigenschaft steht. Für das Feynman-Produkt liegen keine Ergebnisse zur Assoziativität und zur Distributivität vor. Die Überprüfung anhand von Beispielen zeigt leichte Abweichungen, wobei nicht ausgeschlossen werden kann, dass diese auf die Approximation der Funktionswerte durch numerische Verfahren zurückzuführen sind. In der letzten Zeile von Tabelle 5.1 wird vorausgesetzt, dass die zu A inverse Matrix A^{-1} existiert.

Tabelle 5.1: Eigenschaften der Matrixprodukte im Vergleich

	AB	$A \bullet_J B$	$A \bullet_P B$	$A \bullet_F B$
$A \circ B$ symmetrisch	–	+	+	+
positive Definitheit bleibt erhalten	–	–	+	+
$A \circ B = B \circ A$	–	+	–	+
$(A \circ B) \circ C = A \circ (B \circ C)$	+	–	–	
$A \circ (B + C) = A \circ B + A \circ C$	+	+	+	
$(A + B) \circ C = A \circ C + B \circ C$	+	+	–	
$A \circ I = I \circ A = A$	+	+	+	+
$A \circ A^{-1} = A^{-1} \circ A = I$	+	+	+	+

Die Abbildungen 5.3 und 5.4 zeigen anhand eines ausgewählten Beispiels den Vergleich der Funktionswerte von Jordan-Produkt, P-Produkt und Feynman-Produkt. Dabei wird analog zum Vorgehen in Abbildung 5.2 das jeweilige Produkt einer Menge von Matrizen A_i gleicher Spur und einer einzelnen Matrix B gebildet. In Abbildung 5.3 sind die Ergebnisse für $B \circ A_i$ und in Abbildung 5.4 für $A_i \circ B$ dargestellt.

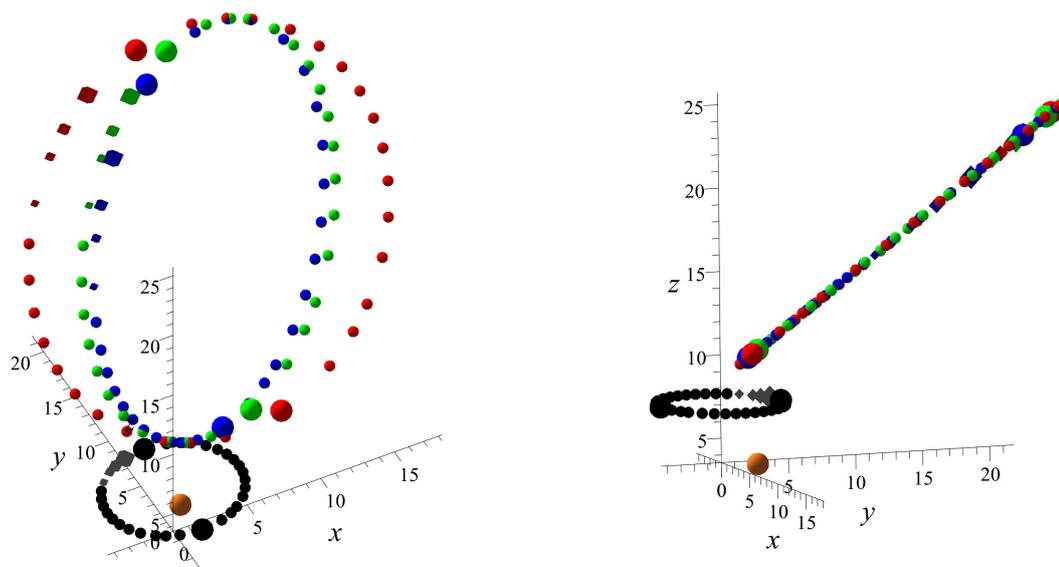


Abbildung 5.3: **(a) Links:** Jordan-Produkt (rot), P-Produkt (grün) und Feynman-Produkt (blau) der braunen Matrix B mit den Matrizen A_i auf dem Kreis (schwarz) als $B \circ A_i$ **(b) Rechts:** Matrizen aus (a) aus anderer Perspektive. Die Funktionswerte der Produkte liegen in einer Ebene.

Für Jordan- und Feynman-Produkt sind die Resultate in Abbildung 5.3 und 5.4 auf Grund der Kommutativität identisch. Eine Abweichung tritt nur beim P-Produkt auf. Hier unterscheiden sich $A_i \bullet_P B$ und $B \bullet_P A_i$. Die Betrachtung der Produkte aus den Perspektiven in Abbildung 5.3 (b) und in Abbildung 5.4 (b) liefert zudem Informationen über die Lage der Matrixprodukte. Während im Fall $B \circ A_i$ alle Funktionswerte der drei Produkte in einer Ebene liegen, liegen die Funktionswerte des P-Produktes für $A_i \bullet_P B$ zum Teil außerhalb dieser Ebene. Daher stellt die Betrachtung von $B \circ A_i$ einen sinnvolleren Ansatz zum visuellen Vergleich der Produkte dar.

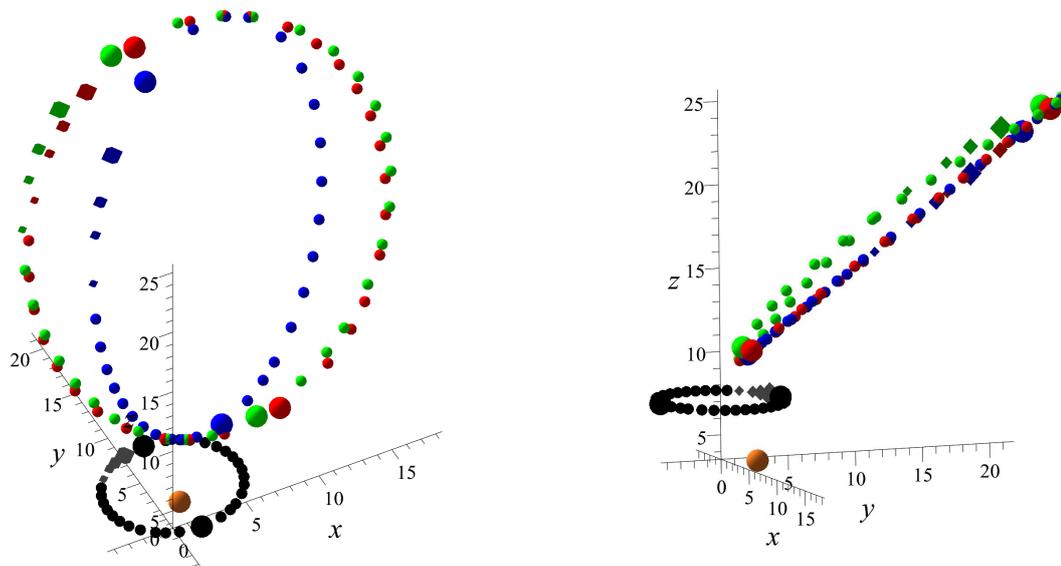


Abbildung 5.4: **(a) Links:** Jordan-Produkt (rot), P-Produkt (grün) und Feynman-Produkt (blau) der braunen Matrix B mit den Matrizen A_i auf dem Kreis (schwarz) als $A_i \circ B$ **(b) Rechts:** Matrizen aus (a) aus anderer Perspektive. Jordan- und Feynman-Produkt liegen in einer Ebene, das P-Produkt fällt zum Teil aus der Ebene heraus.

5.1.5 Parallele Summe

Wir betrachten eine weitere Verknüpfung zweier Matrizen, die parallele Summe. Diese tritt im physikalischen Kontext der Netzwerktheorie auf. Sind zwei Widerstände R_1 und R_2 in Serie geschaltet, ist der Gesamtwiderstand $R_{ges} = R_1 + R_2$, sind die Widerstände parallel geschaltet, erhalten wir $R_{ges} = R_1 R_2 / (R_1 + R_2)$. Im ersten Fall spricht man von einer seriellen Addition, im zweiten Fall von einer parallelen Addition. Anderson und Duffin übertrugen dieses Konzept auf Matrizen und definierten in [1] die *parallele Summe* zweier positiv definiten Matrizen $A, B \in \mathbb{R}^{n \times n}$ als

$$A : B := A(A + B)^{-1}B = (A^{-1} + B^{-1})^{-1}. \quad (5.6)$$

Die *serielle Summe* entspricht der gewöhnlichen Summe zweier Matrizen nach Definition 2.1.2. Aus der Definition der parallelen Summe nach (5.6) folgt unmittelbar die Kommutativität $A : B = B : A$. Die Symmetrie bleibt durch die parallele Addition erhalten, denn

$$(A : B)^T = B^T((A + B)^{-1})^T A^T = B(A + B)^{-1}A = B : A = A : B.$$

Das Gleiche gilt für die positive Definitheit. Denn sind A und B positiv definit, dann auch ihre Inversen A^{-1} und B^{-1} . Damit ist auch $A^{-1} + B^{-1}$ positiv definit, woraus die positive Definitheit von $(A^{-1} + B^{-1})^{-1} = A : B$ folgt. Abbildung 5.5 zeigt die parallele und die serielle (gewöhnliche) Summe einer Menge von Matrizen A_i auf einem Kreis mit einer einzelnen Matrix B . Die parallele Summe unterscheidet sich im Allgemeinen deutlich von der gewöhnlichen Summe.

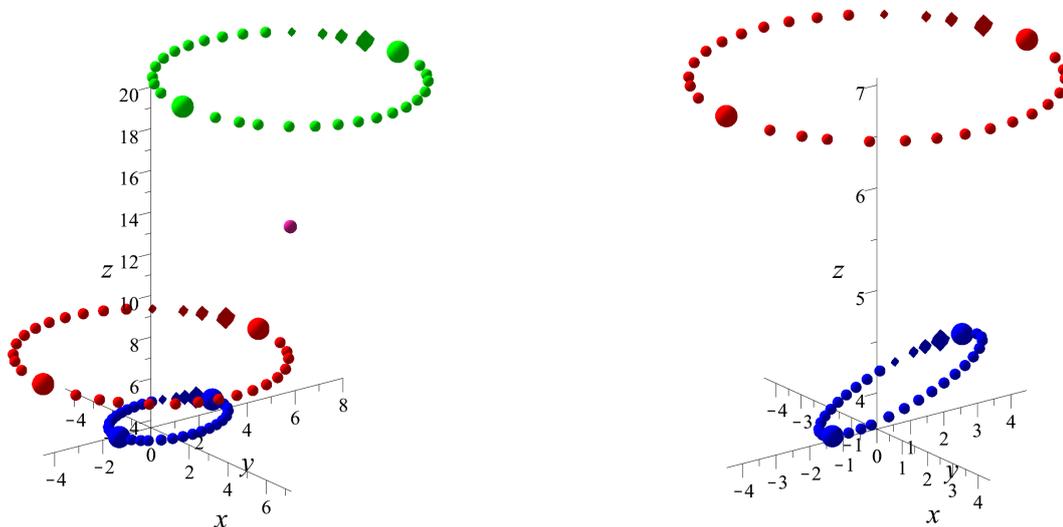


Abbildung 5.5: **(a) Links:** Parallele Summe (blau) und gewöhnliche Summe (grün) der roten Matrizen aus dem Kreis und der violetter Matrix. **(b) Rechts:** Vergrößerte Ansicht der parallelen Summe (blau) aus (a).

5.1.6 Monotoniebetrachtungen

Ist a eine positive reelle Zahl, so ist die Funktion $f(t) = at$ nach Beispiel 4.4.2 matrixmonoton. Wir ersetzen a durch eine positiv definite Matrix A und untersuchen die Monotonie der Funktion

$$f_A(X) = A \circ X,$$

wobei wir für die Verknüpfung " \circ " das Jordan-Produkt, das P-Produkt, das Feynman-Produkt und die parallele Summe einsetzen. Betrachten wir die Matrizen

$$A = \begin{pmatrix} 4 & 1 \\ 1 & 5 \end{pmatrix}, \quad X_1 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{und} \quad X_2 = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix},$$

dann ist A positiv definit und $X_2 \succeq X_1$. Allerdings sind $A \bullet_J X_1$ und $A \bullet_J X_2$ in der Löwner-Ordnung nicht vergleichbar. Gleiches gilt für $A \bullet_F X_1$ und $A \bullet_F X_2$, womit $f_A(X)$ für das Jordan-Produkt und für das Feynman-Produkt nicht matrixmonoton ist.

Für das P-Produkt ist $f_A(X) = A \bullet_P X = A^{1/2} B A^{1/2}$ hingegen matrixmonoton. Denn ist $X_2 \succeq X_1$, dann ist

$$y^\top (X_2 - X_1) y \geq 0$$

für alle $y \in \mathbb{R}^n \setminus \{0\}$. Damit ist auch

$$x^\top (A^{1/2} X_2 A^{1/2} - A^{1/2} X_1 A^{1/2}) x = \underbrace{(A^{1/2} x)^\top}_y (X_2 - X_1) \underbrace{(A^{1/2} x)}_y \geq 0$$

für alle $x \in \mathbb{R}^n \setminus \{0\}$ und somit $f_A(X_2) \succeq f_A(X_1)$.

Für die parallele Summe ist $f_A(X) = A : X = (A^{-1} + B^{-1})^{-1}$ matrixmonoton. Denn für positiv definite Matrizen ist mit $X_2 \succeq X_1$ auch $X_2^{-1} \preceq X_1^{-1}$ und somit

$$(A^{-1} + X_2^{-1}) \preceq (A^{-1} + X_1^{-1}).$$

Damit haben wir

$$(A^{-1} + X_2^{-1})^{-1} \succeq (A^{-1} + X_1^{-1})^{-1},$$

was äquivalent zu $A : X_2 \succeq A : X_1$ ist.

5.2 Bestimmung von Supremum und Infimum symmetrischer Matrizen

In diesem Abschnitt wird ein Algorithmus zur Bestimmung von Supremum und Infimum einer Menge symmetrischer Matrizen vorgestellt, der auf der Ausführung von Burgeth et al. in [3] basiert. Zudem betrachten wir, inwiefern sich Supremum und Infimum symmetrischer Matrizen durch Matrixfunktionen berechnen lassen. Im skalaren Fall beruht die Definition von Supremum und Infimum auf der vollständigen Ordnungsrelation " \leq ". Auf der Menge der symmetrischen Matrizen existiert die in Abschnitt 2.4 vorgestellte Löwner-Ordnung, die allerdings nicht vollständig und keine *lattice order* ist. Dies bedeutet, dass der Begriff eines eindeutigen Supremums nicht existiert.

5.2.1 Algorithmus zur Berechnung von Supremum und Infimum

Bevor wir den allgemeinen Fall symmetrischer $n \times n$ -Matrizen betrachten, demonstrieren wir die Grundidee anhand symmetrischer 2×2 -Matrizen A_1, \dots, A_m .

Supremum und Infimum symmetrischer 2×2 -Matrizen

Wir nehmen zunächst an, dass die Spuren der Matrizen $A_1, \dots, A_m \in \mathbb{R}^{2 \times 2}$ positiv sind. Zum Auffinden des Supremums $\sup(A_1, \dots, A_m)$ betrachten wir für $i = 1, \dots, m$ die unteren Penumbas

$$P^-(A_i) = \{N \in \text{Sym}(2) : N \preceq A_i\}$$

der Matrizen A_i und suchen den Kegel, der diese Penumbas minimal umschließt (vgl. Abbildung 5.6). Dieser Kegel ist die untere Penumbra des Supremums der Matrizen A_i , sodass $\sup(A_1, \dots, A_m)$ in dessen Kegelspitze zu finden ist.

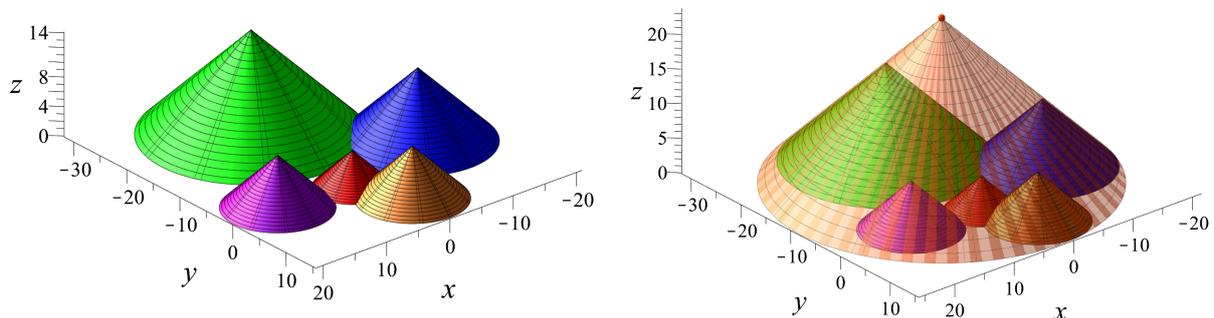


Abbildung 5.6: (a) **Links:** Untere Penumbas verschiedener Matrizen. (b) **Rechts:** Untere Penumbas aus (a) zusammen mit der kleinsten überdeckenden Penumbra (orange). Die Matrix in deren Spitze ist das Supremum der betrachteten Matrizen.

Die Bestimmung dieses minimal umschließenden Kegels erfolgt dabei durch Betrachtung der Schnitte der unteren Penumbra $P^-(A_i)$ mit der x-y-Ebene. Dies sind Kreise K_i , deren Mittelpunkte und Radien durch die Penumbra eindeutig festgelegt sind (vgl. Abbildung 5.7). Der Mittelpunkt des zur unteren Penumbra der Matrix A_i gehörenden Kreises entspricht der Projektion von A_i auf die x-y-Ebene und ist damit durch die Matrix

$$m_{A_i} = A_i - \frac{\text{Spur}(A_i)}{2}I \quad (5.7)$$

gegeben. Die x- und y-Komponenten bleiben so unverändert, während die z-Koordinate Null wird. Der Radius r_{A_i} stimmt mit der z-Komponente von A_i überein, da die Penumbra an ihrer Spitze einen rechten Winkel aufweisen und ihre Symmetrieachsen parallel zur z-Achse verlaufen. Damit ist

$$r_{A_i} = \frac{\text{Spur}(A_i)}{\sqrt{2}}. \quad (5.8)$$

Wir bestimmen den kleinsten Kreis K , der diese Kreise umschließt (vgl. Abbildung 5.7), wozu ein von Gärtner [8] entwickelter Algorithmus verwendet werden kann.

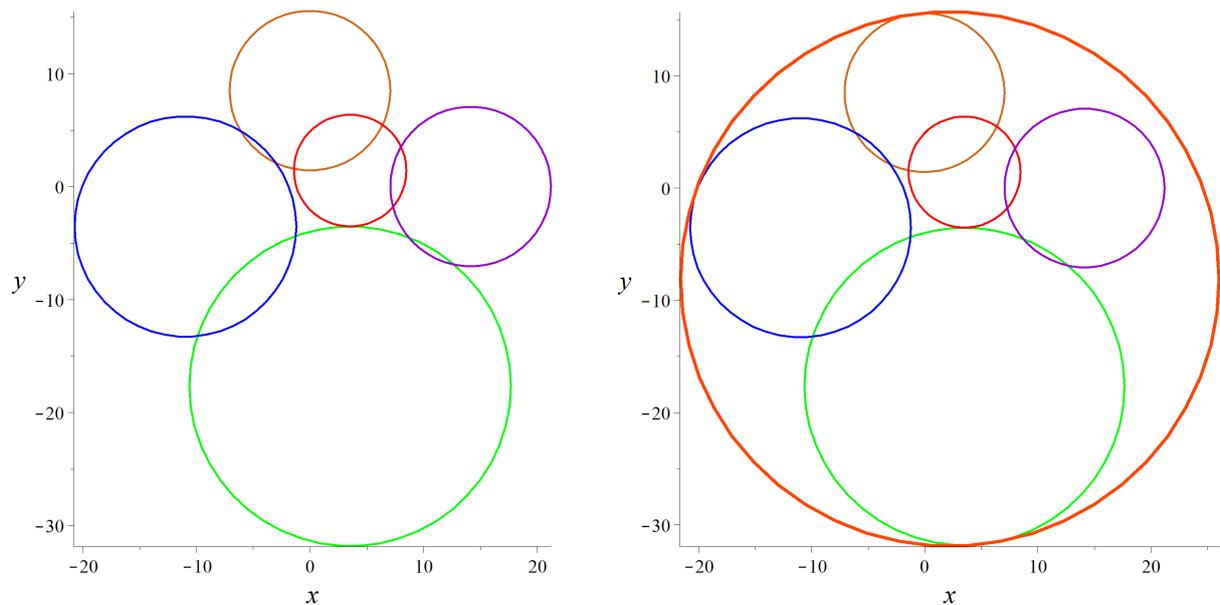


Abbildung 5.7: **(a) Links:** Kreise K_i , die durch die unteren Penumbra aus Abbildung 5.6 mit der x-y-Ebene gebildet werden. **(b) Rechts:** Kleinster Kreis K (orange), der die Kreise K_i minimal umschließt. Er bestimmt die untere Penumbra des Matrixsupremums der betrachteten Matrizen.

Der Kreis K bestimmt die minimal umschließende Penumbra und damit das Matrixsupremum in deren Spitze. Hat K den Mittelpunkt m_K und den Radius r_K , ist das Supremum von A_1, \dots, A_m den vorausgegangenen Überlegungen zufolge

$$\sup(A_1, \dots, A_m) = m_s + \frac{r_s}{\sqrt{2}}I. \quad (5.9)$$

Haben nicht alle betrachteten Matrizen positive Spuren, verwenden wir, dass wir das Supremum von A_1, \dots, A_m über

$$\sup(A_1, \dots, A_m) = f^{-1}(\sup(f(A_1), \dots, f(A_m))) \quad (5.10)$$

berechnen können, wobei f eine monoton steigende Funktion ist. Wir betrachten daher die matrixmonotone Funktion $f(X) = B + X$ und wählen die Matrix B so, dass $\text{Spur}(f(A_i))$ für die zu untersuchenden Matrizen A_i stets positiv ist. Dies leistet zum Beispiel die Matrix

$$B = (\max_i |\text{Spur}(A_i)| + 1)I.$$

Die Berechnung des Infimums einer Menge von symmetrischen Matrizen A_1, \dots, A_m erfolgt durch Betrachtung ihres Supremums. Ist g eine monoton fallende Funktion, erhalten wir

$$\inf(A_1, \dots, A_m) = g^{-1}(\sup(g(A_1), \dots, g(A_m))). \quad (5.11)$$

Dabei bietet sich die monoton fallende Matrixfunktion $g(X) = C - X$ an, wobei wir C so wählen, dass die betrachteten Matrizen alle eine positive Spur haben.

Supremum und Infimum symmetrischer $n \times n$ -Matrizen

Die Idee zum Auffinden von Supremum und Infimum symmetrischer 2×2 -Matrizen kann auf symmetrische Matrizen höherer Dimensionen übertragen werden. Dazu ist es notwendig, eine Verallgemeinerung des Löwner-Kegels der positiv (semi)definiten Matrizen für höhere Dimensionen zu finden. Wir suchen zunächst nach einer Basis des Löwner-Kegels der positiv (semi)definiten Matrizen $\text{Sym}^+(n)$ im $\mathbb{R}^{n \times n}$. Eine entsprechende Herleitung findet sich in [3]. Eine Basis von $\text{Sym}^+(n)$ ist eine Menge $B \subseteq \text{Sym}^+(n)$, sodass alle von der Nullmatrix verschiedenen $Y \in \text{Sym}^+(n)$ eine eindeutige Darstellung als $Y = r \cdot A$ mit $r > 0$ und $A \in B$ haben. Wir verwenden, dass jede konvexe und kompakte Menge S endlicher Dimension als Konvexkombination ihrer Extrempunkte $\text{ext}(S)$ ausgedrückt werden kann gemäß

$$\begin{aligned} S &= \text{convexhull}(\text{ext}(S)) \\ &= \left\{ \sum_{i=1}^N \lambda_i e_i : N \in \mathbb{N}, e_i \in \text{ext}(S), \lambda_i \geq 0, \text{ für } i = 1, \dots, N, \sum_{i=1}^N \lambda_i = 1 \right\}. \end{aligned}$$

Ein Extrempunkt von S ist dabei ein Punkt x , sodass $S \setminus \{x\}$ immer noch konvex ist. Eine konvexe und kompakte Basis von $\text{Sym}^+(n)$ ist durch

$$\{M \in \text{Sym}^+(n) : \text{Spur}(M) = 1\} \quad (5.12)$$

gegeben. Die Extrempunkte dieser Menge sind Matrizen der Form vv^\top , wobei $v \in \mathbb{R}^n$ Einheitsvektoren sind mit $|v| = 1$. Denn die entsprechenden Matrizen vv^\top haben alle Rang 1 und 1 ist zudem ihr einziger von Null verschiedener Eigenwert, da für jeden Einheitsvektor $v \in \mathbb{R}^n$

$$vv^\top v = v \cdot |v|^2 = v$$

ist. Damit ist $\text{Spur}(vv^\top) = 1$, sodass Matrizen vv^\top mit $|v| = 1$ die Extrempunkte der Menge (5.12) bilden. Eine Basis des Löwner-Kegels für n -dimensionale positiv (semi)definite Matrizen ist daher durch

$$\text{convexhull}(\{vv^\top : v \in \mathbb{R}^n, |v| = 1\})$$

gegeben.

Beispiel 5.2.1. Die Matrizen

$$\begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \begin{pmatrix} \cos t & \sin t \end{pmatrix} = \begin{pmatrix} \cos^2 t & \sin t \cos t \\ \sin t \cos t & \sin^2 t \end{pmatrix}$$

mit $t \in [0, 2\pi]$ bilden eine Basis des bislang betrachteten Löwner-Kegels der positiv (semi)definiten 2×2 -Matrizen. Dieser kann daher auch über die Koordinaten

$$\frac{1}{\sqrt{2}}(2r \sin t \cos t, r(\cos^2 t - \sin^2 t), r)^\top$$

mit $t \in [0, 2\pi]$ und $r \geq 0$ im \mathbb{R}^3 visualisiert werden.

Die untere Penumbra einer symmetrischen Matrix $M \in \mathbb{R}^{n \times n}$ ist analog zum 2×2 -Fall gegeben durch

$$P^-(M) = \{N \in \text{Sym}(n) : N \preceq M\} = M - \text{Sym}^+(n).$$

Für symmetrische $n \times n$ -Matrizen A_1, \dots, A_m lässt sich das Supremum $\sup(A_1, \dots, A_m)$ in Analogie zum Supremum symmetrischer 2×2 -Matrizen ermitteln. Auch hierbei können wir uns auf Matrizen mit nichtnegativer Spur beschränken. Ansonsten erfolgt die Berechnung des Supremums über (5.10). Wir betrachten die unteren Penumbras $P^-(A_1), \dots, P^-(A_m)$ und bestimmen die Penumbra, welche diese minimal umschließt. Durch sie wird das Supremum $\sup(A_1, \dots, A_m)$ festgelegt. Für symmetrische 2×2 -Matrizen wird diese minimal umschließende Penumbra bestimmt, indem die zu $P^-(A_i)$ gehörenden Kreise in der x-y-Ebene betrachtet werden. Die x-y-Ebene entspricht dabei gerade der Menge aller symmetrischen Matrizen mit Spur Null. Für symmetrische $n \times n$ -Matrizen betrachten wir dementsprechend die durch die unteren Penumbras $P^-(A_i)$ eindeutig festgelegten Kugeln in

$$\{A \in \text{Sym}(n) : \text{Spur}(A) = 0\}.$$

In dieser Menge bildet

$$\{A_i - \text{Spur}(A_i) \cdot \text{convexhull}\{vv^\top : v \in \mathbb{R}^n, |v| = 1\}\}$$

eine Basis von $P^-(A_i)$. Die orthogonale Projektion von A_i auf die Menge der symmetrischen $n \times n$ -Matrizen mit Spur Null erhalten wir durch

$$m_{A_i} = A_i - \frac{\text{Spur}(A_i)}{n} I, \tag{5.13}$$

wobei wir mit I die n -dimensionale Einheitsmatrix bezeichnen. Die Extrempunkte der Basis von $P^-(A_i)$ liegen dann auf einem Kreis mit Mittelpunkt m_{A_i} , dessen Radius

$$r_{A_i} = \|A_i - \text{Spur}(A_i)vv^\top - m_{A_i}\| = \text{Spur}(A_i)\sqrt{1 - \frac{1}{n}} \quad (5.14)$$

beträgt. Die Bestimmung der kleinsten Kugel, welche die zu A_i gehörenden Kugeln umschließt, kann wie im 2×2 -Fall mit dem Algorithmus von Gärtner [8] erfolgen. Sind der Mittelpunkt m_K und der Radius r_K dieser minimal umschließenden Kugel K bekannt, erhalten wir die zur entsprechenden Penumbra gehörende Matrix S durch

$$S = m_K + \frac{r_K}{n} \frac{1}{\sqrt{1 - \frac{1}{n}}} I. \quad (5.15)$$

Die Matrix S entspricht dabei dem Supremum der Matrizen A_i . Für $n = 2$ stimmen (5.13) bis (5.15) mit den oben betrachteten Formeln (5.7) bis (5.9) überein. Entsprechende Visualisierungen für höhere Dimensionen sind allerdings nicht bekannt. Die Berechnung des Infimums erfolgt auch für symmetrische $n \times n$ -Matrizen gemäß (5.11) über die Betrachtung des Supremums.

5.2.2 Berechnung über Matrixfunktionen

Für reelle Zahlen a und b liefert

$$\max(a, b) = \frac{1}{2}(a + b) + \frac{1}{2}|a - b| \quad (5.16)$$

das Maximum von a und b . Um (5.16) zu verifizieren, betrachten wir zunächst den Fall $a \geq b$. Dann ist

$$\frac{1}{2}(a + b) + \frac{1}{2}|a - b| = \frac{1}{2}(a + b + a - b) = a = \max(a, b).$$

Ist hingegen $b > a$, so ist

$$\frac{1}{2}(a + b) + \frac{1}{2}|a - b| = \frac{1}{2}(a + b - a + b) = b = \max(a, b).$$

Es sei nun V eine symmetrische $n \times n$ -Matrix mit der Eigenwertzerlegung

$$V = S \operatorname{diag}(\lambda_1, \dots, \lambda_n) S^\top.$$

Dann definieren wir

$$|V| := S \operatorname{diag}(|\lambda_1|, \dots, |\lambda_n|) S^\top. \quad (5.17)$$

Mit Hilfe dieser Definition können wir mit (5.16) das Supremum von zwei symmetrischen Matrizen bezüglich der Löwner-Ordnung berechnen. Wir erhalten

$$\sup(A, B) = \frac{1}{2}(A + B) + \frac{1}{2}|A - B|. \quad (5.18)$$

Ist $A \succeq B$ oder $B \succeq A$, liefert diese Formel offenbar das Supremum. Für den Fall nicht vergleichbarer Matrizen verifizieren wir ihre Gültigkeit für symmetrische 2×2 -Matrizen. Wir zeigen, dass das Supremum von zwei symmetrischen Matrizen $A_1, A_2 \in \mathbb{R}^{2 \times 2}$, das wir über die im vorherigen Abschnitt vorgestellte Methode der minimal umschließenden Penumbra erhalten, mit der aus (5.18) errechneten Matrix

$$S = \frac{1}{2}(A_1 + A_2 + |A_1 - A_2|)$$

übereinstimmt. Dazu können wir ohne Einschränkung die Matrizen

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad A_2 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

mit $\operatorname{Spur}(A_2) > 0$ betrachten. Denn zu zwei beliebigen symmetrischen Matrizen A und B können wir für unsere Betrachtung ohne Einschränkung eine beliebige symmetrische Matrix X addieren, da

$$\frac{1}{2}((A + X) + (B + X) + |(A + X) - (B + X)|) = \frac{1}{2}(A + B + |A - B|) + X.$$

Zudem ist (5.18) rotationsinvariant, denn mit einer orthogonalen Matrix V erhalten wir

$$\frac{1}{2}(VAV^\top + VBV^\top + |VAV^\top - VBV^\top|) = \frac{1}{2}V(A + B + |A - B|)V^\top.$$

Des Weiteren können wir uns auf den Fall $\text{Spur}(A) > 0$ und $\text{Spur}(B) > 0$ beschränken. Ansonsten können wir eine Matrix C finden, sodass $\text{Spur}(A+C) > 0$ und $\text{Spur}(B+C) > 0$ gilt. Es genügt, den Fall $\text{Spur}(A) \leq \text{Spur}(B)$ zu betrachten. Ist $\text{Spur}(B) > \text{Spur}(A)$, vertauschen wir einfach die Rollen von A und B .

Wählen wir nun $Y := I - A$, dann ist $Y + A = I$. Setzen wir $B' := Y + B$, ist

$$\text{Spur}(B') = \text{Spur}(I) - \text{Spur}(A) + \text{Spur}(B) > 0.$$

Die Matrix B' ist symmetrisch und habe die Eigenwertzerlegung $B' = Z \text{diag}(\lambda_1, \dots, \lambda_n) Z^\top$. Damit ist

$$\frac{1}{2}(I + B' + |I - B'|) = \frac{1}{2}Z(I + \text{diag}(\lambda_1, \lambda_2) + |I - \text{diag}(\lambda_1, \lambda_2)|)Z^\top,$$

wobei $\lambda_1 + \lambda_2 > 0$ ist.

Wir betrachten also die Matrizen

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad A_2 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

mit $\text{Spur}(A_2) > 0$. Wir nehmen an, dass $a < 1$ und $b > 1$ gilt. Im Fall $a > 1$ und $b < 1$ lässt sich analog argumentieren. Der Mittelpunkt des Kreises in der x-y-Ebene, der zur unteren Penumbra von A_1 gehört, liegt im Ursprung des kartesischen Koordinatensystems. Sein Radius ist

$$R_1 = \frac{2}{\sqrt{2}}. \tag{5.19}$$

Der entsprechende zur unteren Penumbra von A_2 gehörende Kreis hat den auf der y-Achse liegenden Mittelpunkt

$$M_2 = \frac{1}{\sqrt{2}}(0, b - a)^\top \tag{5.20}$$

und den Radius

$$R_2 = \frac{a + b}{\sqrt{2}}. \tag{5.21}$$

Sind in der x-y-Ebene zwei Kreise mit den Radien r_1 und r_2 und den Mittelpunkten m_1 und m_2 gegeben, wobei m_1 im Ursprung und $m_2 = (0, m'_2)^\top$ auf der y-Achse liegt, ist der Radius R_K des minimal umschließenden Kreises

$$R_K = \frac{r_1 + r_2 + |m'_2|}{2} \tag{5.22}$$

und sein Mittelpunkt

$$M_K = \frac{1}{2} \left(0, m'_2 + \frac{m'_2}{|m'_2|} (r_1 - r_2) \right)^\top. \tag{5.23}$$

Für den kleinsten umschließenden Kreis, der zum Supremum von A_1 und A_2 gehört, erhalten wir damit durch Einsetzen von (5.19), (5.20) und (5.21) in (5.23) den Radius

$$\begin{aligned} R_{Sup} &= \frac{1}{2\sqrt{2}}(2 + a + b + |a - b|) \\ &= \frac{1 + b}{\sqrt{2}} \end{aligned}$$

und mit (5.23) den Mittelpunkt

$$\begin{aligned} M_{Sup} &= \frac{1}{\sqrt{2}}(0, (b - a) - (2 - a - b))^\top \\ &= \frac{1}{\sqrt{2}}(0, b - 1)^\top. \end{aligned}$$

Die Verwendung der Formel (5.18) führt auf die Matrix

$$S = \frac{1}{2}((A_1 + A_2) + |A_1 - A_2|) = \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix}.$$

Der zur unteren Penumbra von S gehörende Kreis in der x-y-Ebene hat damit den Radius

$$R_s = \frac{1 + b}{\sqrt{2}}$$

und den Mittelpunkt

$$M_S = \frac{1}{\sqrt{2}}(0, b - 1)^\top.$$

Damit entspricht er genau dem Kreis, welcher die Kreise der unteren Penumbras von A_1 und A_2 in der x-y-Ebene minimal umschließt. Die Formel

$$\sup(A, B) = \frac{1}{2}(A + B + |A - B|)$$

liefert daher das Supremum von zwei symmetrischen Matrizen A und B . Die Berechnung des Infimums erfolgt über das Supremum gemäß

$$\inf(A, B) = f^{-1}(\sup(f(A), f(B))),$$

wobei f eine monoton fallende Matrixfunktion ist.

Im geometrischen Kontext der Löwner-Visualisierung positiv definiter Matrizen $A \in \mathbb{R}^{2 \times 2}$ als Ellipsen

$$\{x \in \mathbb{R}^2: x^\top A^{-1}x = 1\}$$

wird das Infimum zweier positiv definiter Matrizen A und B durch die maximale Ellipse beschrieben, die im Innern der beiden zu A und B gehörenden Ellipsen liegt. Die Ellipse, die in der Löwner-Visualisierung das Supremum repräsentiert, umschließt die Ellipsen zu A und B minimal. Eine entsprechende Veranschaulichung findet sich in Abbildung 5.8.

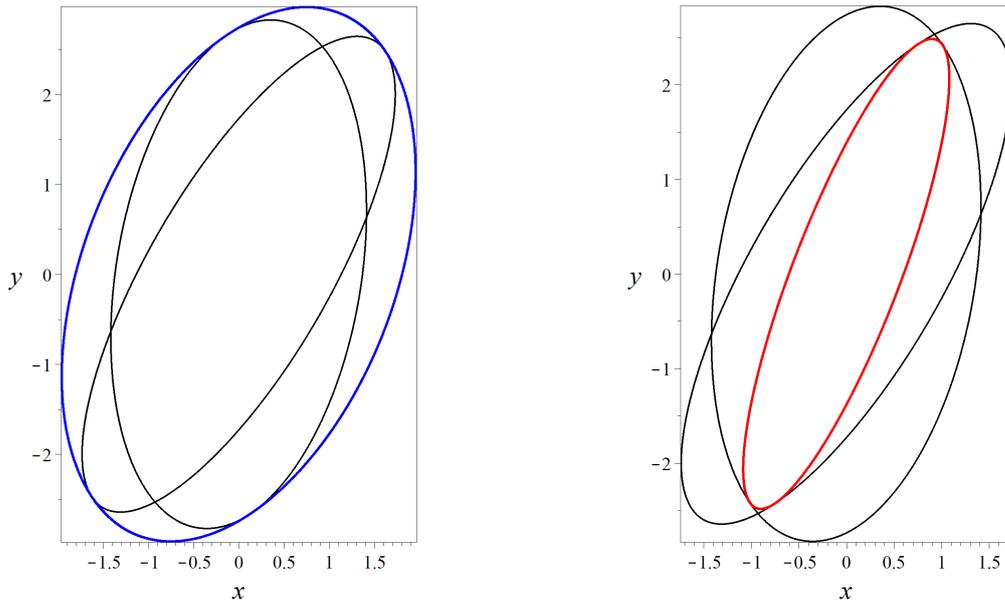


Abbildung 5.8: **(a) Links:** Löwner-Visualisierungen zweier positiv definitiver 2×2 -Matrizen (schwarz) und ihres Supremums (blau). **(b) Rechts:** Löwner-Visualisierungen zweier positiv definitiver 2×2 -Matrizen (schwarz) und ihres Infimums (rot).

Für drei symmetrische Matrizen A_1, A_2 und A_3 erhalten wir durch wiederholte Anwendung von (5.18) die Matrizen

$$S_1 = \sup(A_1, \sup(A_2, A_3)), \quad S_2 = \sup(A_2, \sup(A_1, A_3)), \quad S_3 = \sup(A_3, \sup(A_1, A_2)),$$

die alle obere Schranken von A_1, A_2 und A_3 sind. Die Berechnung anhand konkreter Beispiele zeigt, dass S_1, S_2 und S_3 im Allgemeinen jedoch nicht identisch sind. Entsprechendes erkennen wir auch in Abbildung 5.9 (a). Burgeth et al. schlagen daher in [4] vor, den Mittelwert

$$S_m := \frac{1}{3}(S_1 + S_2 + S_3) \geq A_i, \quad i = 1, 2, 3$$

zu bilden und diesen mit einem $\tau \geq 0$ zu optimieren, sodass für $i = 1, 2, 3$

$$S_m - \tau I \geq A_i$$

erfüllt ist. Bezeichnen $\lambda_{ij} \geq 0$, $j = 1, \dots, n$ die Eigenwerte von $S_m - A_i$ für $i = 1, 2, 3$, ist das optimale τ gegeben durch

$$\tau_{opt} = \min_{\substack{i=1,2,3 \\ j=1,\dots,n}} (\lambda_{ij}).$$

Wir erhalten mit

$$\sup_{opt}(A_1, A_2, A_3) = S_m - \tau_{opt} I \tag{5.24}$$

ein passendes Supremum der drei Matrizen A_1, A_2 und A_3 , welches für ein ausgewähltes Beispiel positiv definitiver Matrizen in Abbildung 5.9 (b) dargestellt ist.

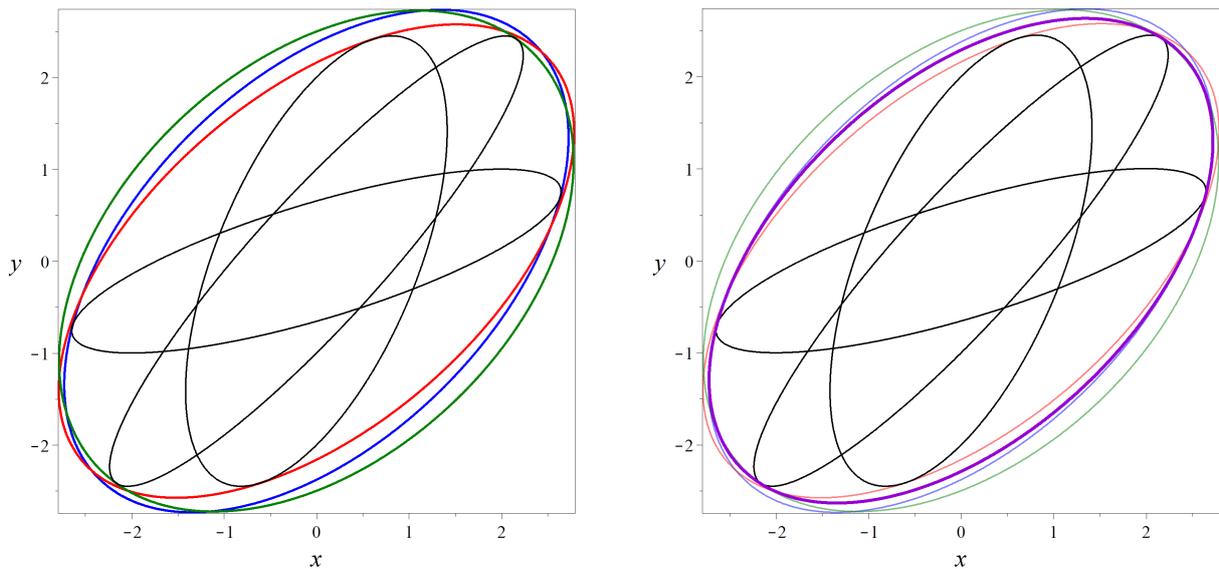


Abbildung 5.9: **(a) Links:** Löwner-Visualisierungen von drei positiv definiten Matrizen A_1 , A_2 und A_3 (schwarz) sowie der oberen Schranken S_1 , S_2 und S_3 aus iterativer Anwendung von (5.18). **(b) Rechts:** Löwner-Visualisierungen der Matrizen aus (a) mit dem optimierten Supremum $\sup_{opt}(A_1, A_2, A_3)$ (violet) gemäß (5.24).

Grenzwertbetrachtungen

Maximum und Minimum von zwei positiven reellen Zahlen lassen sich auch als Grenzwerte angeben.

Proposition 5.2.2. *Es seien a, b positive reelle Zahlen. Dann gilt:*

$$i) \max(a, b) = \lim_{n \rightarrow \infty} \sqrt[n]{a^n + b^n},$$

$$ii) \min(a, b) = \lim_{x \rightarrow 0} \log_x(x^a + x^b).$$

Beweis. i) Ohne Einschränkung sei $a \geq b$, sodass $\max(a, b) = a$. Dann ist

$$a^n \leq a^n \left(1 + \left(\frac{b}{a} \right)^n \right) \leq 2a^n. \quad (5.25)$$

Die Funktion $\sqrt[n]{\cdot}$ ist auf \mathbb{R}^+ monoton wachsend, sodass (5.25) äquivalent zu

$$a \leq \sqrt[n]{a^n + b^n} \leq a \sqrt[n]{2}$$

ist. Die Behauptung folgt dann aus $\lim_{n \rightarrow \infty} a = \lim_{n \rightarrow \infty} a \sqrt[n]{2} = a$.

ii) Ohne Einschränkung sei $a \leq b$ und damit $\min(a, b) = a$. Da $x > 0$, ist $x^a + x^b \geq x^a$ und mit der Monotonie von $\ln(\cdot)$ erhalten wir

$$a = \frac{\ln(x^a)}{\ln(x)} \leq \frac{\ln(x^a + x^b)}{\ln(x)} = \log_x(x^a + x^b). \quad (5.26)$$

Da $x^a + x^b \leq 2x^a$ für $x < 1$, können wir für $x < 1$ den Ausdruck $\log_x(x^a + x^b)$ nach oben abschätzen durch

$$\log_x(x^a + x^b) = \frac{\ln(x^a + x^b)}{\ln(x)} \leq \frac{\ln(2x^a)}{\ln(x)} = a + \frac{\ln(2)}{\ln(x)}. \quad (5.27)$$

Zusammen ergeben (5.26) und (5.27) für $x < 1$

$$a \leq \log_x(x^a + x^b) \leq \frac{\ln(2)}{\ln(x)} + a.$$

Aus

$$\lim_{x \rightarrow 0} a = \lim_{x \rightarrow 0} \left(a + \frac{\ln(2)}{\ln(x)} \right) = a$$

folgt die Behauptung. □

Wir überprüfen, ob die Grenzwertformeln aus Proposition 5.2.2 auch auf Supremum und Infimum positiv definiter Matrizen übertragbar sind, d.h. ob wir Supremum und Infimum zweier positiv definiter Matrizen $A, B \in \mathbb{R}^{n \times n}$ über

$$\sup(A, B) = \lim_{n \rightarrow \infty} \sqrt[n]{A^n + B^n} \quad (5.28)$$

und

$$\inf(A, B) = \lim_{x \rightarrow 0} \log_x(x^A + x^B) \quad (5.29)$$

berechnen können. Dazu betrachten wir die Löwner-Visualisierungen der Matrizen A und B , sowie der Matrizen, die wir über die Grenzwerte (5.28) und (5.29) berechnen. Abbildung 5.10 zeigt dementsprechend die Löwner-Visualisierungen $\{x \in \mathbb{R}^2: x^\top A^{-1}x = 1\}$ und $\{x \in \mathbb{R}^2: x^\top B^{-1}x = 1\}$ zweier positiv definiter Matrizen $A, B \in \mathbb{R}^{2 \times 2}$. Im linken Bild ist zudem die Matrix

$$S' = \lim_{n \rightarrow \infty} \sqrt[n]{A^n + B^n} \quad (5.30)$$

über $\{x \in \mathbb{R}^2: x^\top (S')^{-1}x = 1\}$ visualisiert und rechts dementsprechend die Matrix

$$I' = \lim_{x \rightarrow 0} \log_x(x^A + x^B). \quad (5.31)$$

Anhand der Visualisierungen erkennen wir, dass S' und I' in diesem Fall zwar eine obere bzw. untere Schranke für die Matrizen A und B ergeben, unseren Erwartungen an das Supremum und Infimum aber nicht gerecht werden. Die beiden betrachteten Grenzwerte können somit zur Bestimmung von Matrixsupremum und Matrixinfimum nicht herangezogen werden.

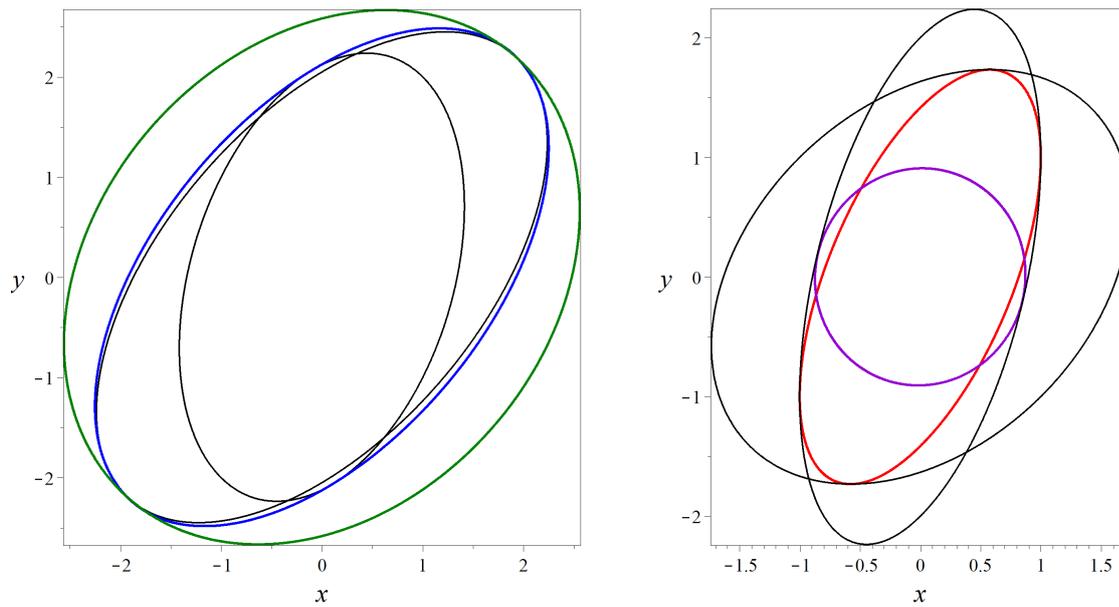


Abbildung 5.10: **(a) Links:** Löwner-Visualisierungen zweier positiv definiter Matrizen (schwarz), ihres Supremums gemäß (5.18) (blau) sowie des Grenzwertes S' aus (5.30) (grün). **(b) Rechts:** Löwner-Visualisierungen zweier positiv definiter Matrizen (schwarz), ihres Infimums gemäß (5.18) (rot) sowie des Grenzwertes I' aus (5.31) (violett).

Literaturverzeichnis

- [1] William N. Anderson, Richard J. Duffin, *Series and Parallel Addition of matrices*. Journal of mathematical analysis and applications **26** (1969), 576-594.
- [2] Rajendra Bhatia, *Matrix Analysis*. Graduate Texts in Mathematics, Vol. 169, Springer, New York; Berlin; Heidelberg, 1997.
- [3] Bernhard Burgeth, Nils Papenberg, Andres Bruhn, Martin Welk, Joachim Weickert, *Mathematical Morphology for Tensor Data Induced by the Loewner Ordering in Higher Dimensions*. Preprint No. 161, Universität des Saarlandes, Fachrichtung 6.1, Saarbrücken, 2005.
- [4] Bernhard Burgeth, Michael Breuß, Stephan Didas, Joachim Weickert, *PDE-based Morphology for Matrix Fields: Numerical Solution Schemes*. Preprint No. 220, Universität des Saarlandes, Fachrichtung 6.1, Saarbrücken, 2007.
- [5] Arthur Cayley, *A memoir on the theory of matrices*. Philosophical Transactions of the Royal Society London **148** (1858), 17-37.
- [6] William F. Donoghue, *Monotone matrix functions and analytic continuation*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Vol. 207. Springer, New York; Berlin; Heidelberg, 1974.
- [7] Klaus Fritzsche, *Kurze Ausarbeitung zur Vorlesung Lineare Algebra 1*. http://www2.math.uni-wuppertal.de/~fritzsche/kf_lss05.htm (Stand: Januar 2013)
- [8] Bernd Gärtner, *Fast and Robust Smallest Enclosing Balls*. Lecture Notes in Computer Science **1643** (1999), 325-338. C-Programm abrufbar unter <http://www.inf.ethz.ch/personal/gaertner/miniball.html> (Stand: Januar 2013)
- [9] Nicholas J. Higham, *Functions of Matrices: Theorie and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, 2008.
- [10] Roger A. Horn, Charles R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.
- [11] Roger A. Horn, Charles R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1999.

- [12] Karl Löwner, *Über monotone Matrixfunktionen*. Mathematische Zeitschrift, Vol. 38 (1934), 177-216.
- [13] Denis Serre, *Matrices: Theorie and Applications*. Springer, New York, 2002.
- [14] Gilbert Strang, *Introduction to Linear Algebra*. Brooks Cole, Belmont, 2007.
- [15] James J. Sylvester, *On the equation to the secular inequalities in the planetary theory*. Philosophical Magazine **16** (1883), 267-269.
- [16] Detlef Wille, *Repetitorium der Linearen Algebra*. Binomi Verlag, Springe, 2006.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides Statt, dass ich die vorliegende Arbeit selbstständig und unter ausschließlicher Verwendung der angegebenen Literatur und Hilfsmittel verfasst habe.

Saarbrücken, 31. Januar 2013