Universität Potsdam

Mathematisch-Naturwissenschaftliche Fakultät

Institut für Mathematik

Hauptminoren reeller symmetrischer Matrizen: Anwendung für die Semidefinitheit mithilfe des charakteristischen Polynoms

Bachelorarbeit

von

Lukas Hellwig

betreut von

Prof. Dr. Joachim Gräter

Zweitgutachter:

Prof. Dr. Martin Weese

Potsdam, den 26. April 2018

Zusammenfassung

In dieser Arbeit wird die Beziehung zwischen den Hauptminoren einer Matrix und ihrer Semidefinitheit untersucht. Ausgangspunkt bildet dabei die Betrachtung der Koeffizienten des charakteristischen Polynoms. Ich werde zeigen, dass diese bis auf das Vorzeichen eindeutig durch die Summen von Hauptminoren der gleichen Größe dargestellt werden können. Außerdem werde ich für diese Koeffizienten eine Repräsentation durch Summen von Produkten der Eigenwerte motivieren und nachweisen.

Schließlich wird das Hauptminorenkriterium der positiven Definitheit auf singuläre Matrizen erweitert. Dabei werden alle Hauptminoren betrachtet, um die positive Semidefinitheit zu schlussfolgern. Für die Semidefinitheit werden hinreichende und äquivalente Bedingungen gefunden und deren Implikationen gezeigt. Es findet außerdem eine abgrenzende Diskussion des erweiterten Hauptminorenkriteriums zum Hauptminorenkriterium der positiven Definitheit statt.

Inhaltsverzeichnis

1 Einleitung

Der Ausgangspunkt für die Überlegungen dieser Arbeit geht zurück auf das Jahr 1900, als Hilbert [3] vor den zweiten Internationalen Mathematikerkongress trat und in seinem Vortrag eine Reihe von 23 Problemen lieferte, welche die Mathematik des kommenden 20. Jahrhunderts lösen sollte. Darunter war das 17. Hilbert'sche Problem, welches danach fragte, ob sich reelle Polynome $f \in \mathbb{R}[x_1, \ldots, x_n]$ in n Unbestimmten über $\mathbb{R}$ mit $f \geq 0$ an jeder Stelle als Summen endlicher vieler Quadrate reellwertiger rationaler Funktionen aus $\mathbb{R}(x_1, \ldots, x_n)$ schreiben lassen.

Dieses Problem löste Artin [1] 1923, indem er zeigte, dass eine symmetrische Matrix mit Einträgen aus $\mathbb{R}[x_1, \ldots, x_n]$ dies erfüllt, sofern sie für alle Substitutionen $(x_1, \ldots, x_n)^\tau \in \mathbb{R}^n$ nichtnegativ ist. Diese Eigenschaft bezeichnet man als *Semidefinitheit* einer Matrix.

Im Rahmen dieser Arbeit befasse ich mich eingehend mit der Semidefinitheit einer symmetrischen Matrix $A \in M_n(\mathbb{R})$. Diese Fragestellung ist ebenso untrennbar verwoben mit der Suche nach Nullstellen (reeller) homogener Polynome des Grades 2 über n Unbestimmten $x_1, \ldots, x_n$ in $\mathbb{R}$. Fasst man die Unbestimmten eines Polynoms als Komponenten eines Vektors der Dimension n und die Koeffizenten in $\mathbb{R}$ als Einträge einer Matrix auf, so erhält man

$$p(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i x_j = x^\tau A x$$

für eine symmetrische Matrix $A \in M_n$. Die Wahl eines bestimmten Vektors $x \in \mathbb{R}^n$ kommt nun dem Einsetzen in eben jenes Polynom $p(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]$ gleich. Trifft man also eine Aussage über das Verhalten der sogenannten *quadratischen Form* $x^\tau Ax$ für alle Vektoren $x \in \mathbb{R}^n$, so bildet man das Verhalten des beschriebenen homogenen Polynoms p auf dem gesamten Definitionsbereich $\mathbb{R}$ ab. Die positive Definitheit entspricht in diesem Deutungsmuster einer alleinigen Nullstelle beim Nullpunkt.

Die Aufgabe dieser Arbeit ist die Erweiterung des Hauptminorenkriteriums der positiven Definitheit auf den Fall der positiven Semidefinitheit. Dazu erfolgt eine eingehende Beschreibung von Hauptminoren und ihrem Einfluss auf grundlegende Eigenschaften einer Matrix, wie ihr charakteristisches Polynom, ihre Eigenwerte und ihre Definitheit.

Im Kontext des charakteristischen Polynoms (Kapitel 3) werde ich zeigen, dass die Beträge von dessen Koeffizienten allein durch die Kenntnis der Hauptminoren bzw. deren Summen festgelegt werden können. Diese eindeutige Darstellbarkeit durch Summen von Hauptminoren hat einen bemerkenswerten Effekt auf die Nullstellen des charakteristischen Polynoms. Dies wird beim Beweis des erweiterten Hauptminorenkriteriums in Kapitel 4 eine Rolle spielen.

Eine anderer Fokus wird im Rahmen dieser Arbeit auf die Eigenwerte einer Matrix gelegt. Ich bemühe mich, den Einfluss der Eigenwerte auf die Koeffizienten des charakteristischen Polynoms in Kapitel 3 herauszustellen. Dies mündet in einigen faszinierenden Zusammenhängen zwischen den Eigenwerten und den Hauptminoren einer Matrix, welche in Abschnitt 3.3 dargestellt werden.

Zuletzt skizziere ich den Beweis des Hauptminorenkriteriums der positiven Definitheit für invertierbare Matrizen. Dieser ist maßgeblich durch Meyer [6] geprägt und führt deshalb einige Verbindungen zu numerischen Anwendungen in Abschnitt 4.4 auf.

2 Grundlagen

Im Allgemeinen setze ich beim Leser dieser Arbeit die Kenntnis der Grundlagenmodule der Linearen Algebra voraus. Im Speziellen verstehe ich darunter die Grundlagen der Vektor- und Matrizenrechnung, die Kenntnis symmetrischer und Diagonalmatrizen sowie die Bedeutung der Ähnlichkeit von Matrizen. Außerdem mache man sich die Determinantenfunktion und deren Berechnung bewusst. Bei Interesse studiere man die ersten beiden Kapitel, sowie Kapitel vier im Buch von Knabner u. Barth [5][S. 1-122, 145-308, 383-436].

Zunächst zur Erklärung einiger Notationen. $\mathbb{N}$ bezeichne die Menge der natürlichen Zahlen ohne 0 und $\mathbb{R}$ die Menge der reellen Zahlen. $M_{m,n}$ bezeichne die Menge aller $m \times n$-Matrizen mit Einträgen aus $\mathbb{R}$. Abgekürzt sei M_n die Menge aller quadratischen $n \times n$-Matrizen mit Einträgen aus $\mathbb{R}$. Weiter sei $\mathbb{1}_n$ die $n \times n$-Einheitsmatrix. Für ein $A \in M_n$ sind $A = (a_{i,j})_{i,j}$ die Einträge der Matrix.

Streichungsmatrizen

Im Folgenden möchte ich mir Matrizen unter dem Gesichtspunkt ansehen, inwiefern sich Eigenschaften einer Matrix aus Ausschnitten dieser Matrix vorhersagen lassen. Deshalb will ich zunächst formal fassen, was ich unter dem „Ausschnitt" einer Matrix verstehe. Zu diesem Zweck definiere ich eine Streichungsabbidung ϕ.

Definition 2.1. *Seien $n \in \mathbb{N}$ und $\alpha, \beta \subset \{1, \ldots, n\}$. Eine Abbildung $\phi_{\alpha,\beta}$ der Form*

$$\begin{aligned} \phi_{\alpha,\beta} : \quad & M_n \quad \to \quad M_{n-|\alpha|,n-|\beta|} \\ & A = (a_{i,j})_{i \in \{1,\ldots,n\}, j \in \{1,\ldots,n\}} \mapsto A(\alpha, \beta) = (a_{i,j})_{i \in \{1,\ldots,n\} \setminus \alpha, j \in \{1,\ldots,n\} \setminus \beta} \end{aligned}$$

heißt ***Streichungsabbildung****.*

Nun interessiere ich mich vor allem für das Bild der Streichungsabbildung ϕ.

Definition 2.2. *Eine Matrix $A(\alpha, \beta) \in M_{n-|\alpha|,n-|\beta|}$ mit $\alpha, \beta \subset \{1, \ldots, n\}$ heißt* ***Streichungsmatrix*** *einer Matrix $A \in M_n$, falls gilt*

$$A(\alpha, \beta) = \phi_{\alpha,\beta}(A).$$

Ist $\alpha = \beta$, so notiert man $A(\alpha, \beta)$ kurz als $A(\alpha)$. In diesem Fall nennt man $A(\alpha) \in M_{n-|\alpha|}$ eine ***Hauptstreichungsmatrix*** *von A.*

Einleitende Definitionen

Zur Wiederholung werde ich hier noch weitere Hauptaussagen festhalten, die sich später als nützlich erweisen sollen.

Definition 2.3. *Sei $n \in \mathbb{N}$. Eine Funktion $f(a_1, a_2, \ldots, a_n)$ mehrerer Spaltenvektoren heißt* ***Determinante*** *wenn sie folgende Eigenschaften erfüllt.*

(i) f ist linear in jedem Eingang, d.h. es gilt für $i \in \{1, \ldots, n\}$:

$$f(a_1, \ldots, \lambda_i a_i + b_i, \ldots, a_n) = \lambda_i f(a_1, \ldots, a_i, \ldots, a_n) + f(a_1, \ldots, b_i, \ldots, a_n)$$

für alle $a_k, b_k \in \mathbb{R}^n, \lambda_i \in \mathbb{R}$ *mit* $1 \leq k \leq n$.

(ii) f *ist alternierend, d.h. es gilt für* $i, j \in \{1, \dots, n\}$*:*

$$f(a_1, \dots, a_i, \dots, a_j, \dots a_n) = -f(a_1, \dots, a_j, \dots, a_i, \dots, a_n)$$

für alle $a_k \in \mathbb{R}^n$ *mit* $1 \leq k \leq n$.

(iii) f *ist normiert, d.h. es gilt* $f(e_1, e_2, \dots, e_n) = 1$ *für die kanonischen Einheitsvektoren* $e_i \in \mathbb{R}^n$.

Man kann nun zeigen, dass die in Definition 2.3 gestellten Forderungen von genau einer Funktion erfüllt werden.

Lemma 2.4 (Laplace'scher Entwicklungssatz)**.** *Sei* $A \in M_1$*. Dann ist die Determinante von* $\det(A) = a_{1,1}$*. Für alle* $\mathbb{N} \ni n > 1, k \in \{1, \dots, n\}$ *und* $A \in M_n$ *berechnet sich die Determinante durch*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{i,k} \ \det\left(A(\{i\}, \{k\})\right) = \sum_{j=1}^{n} (-1)^{k+j} a_{k,j} \ \det\left(A(\{k\}, \{j\})\right).$$

Beweis. Knabner u. Barth [5, S. 277 f.] □

Die Determinante von den oben eingeführten Streichungsmatrizen wird eine herausragende Rolle spielen, sodass ich die folgende Definition treffe.

Definition 2.5. *Sei* $n \in \mathbb{N}$*,* $A \in M_n$ *und* $\alpha, \beta \subset \{1, \dots, n\}$*. Ist die Streichungsmatrix* $A(\alpha, \beta)$ *eine quadratische Matrix, so nennt man die Determinante* $\det\left(A(\alpha, \beta)\right)$ *einen* ***Minor*** *der Matrix* A*.*

Die Determinante einer Hauptstreichungsmatrix $\det\left(A(\alpha)\right)$ *nenne ich entsprechend einen* ***Hauptminor*** *der Matrix* A*.*

Die Mächtigkeit der Menge $|\alpha| = k$ *nenne ich* ***Ordnung*** *der Hauptstreichungsmatrix und korrespondierend Ordnung des Hauptminors.*

Neben der Determinante werde ich Aussagen über die Eigenwerte treffen. Dies motiviert die folgende Definition.

Definition 2.6. *Sei* $n \in \mathbb{N}$ *und* $A \in M_n$*. Eine reelle Zahl* $\lambda \in \mathbb{R}$*, welche der Gleichung*

$$Ax = \lambda x$$

für ein $x \in \mathbb{R}^n$ *mit* $x \neq 0$ *genügt, nennt man* ***Eigenwert*** *der Matrix* A*. Der korrespondierende Vektor* x *wird* ***Eigenvektor*** *zum Eigenwert* λ *genannt.*

Die Menge aller Eigenwerte einer Matrix A *nennt man* ***Spektrum*** *einer Matrix* A*. Es wird notiert durch* $\sigma(A) := \{\lambda \in \mathbb{R} \mid Ax = \lambda x \text{ für } x \in \mathbb{R}^n, x \neq 0\}$*.*

Äquivalentes Umformen der Gleichung in Definition 2.6 führt auf

$$\det\left(\lambda \mathbb{1}_n - A\right) = 0. \tag{2.1}$$

Den linken Teil von (2.1) kann man auch als Polynom auffassen.

Definition 2.7. *Sei* $A \in M_n$ *und* x *eine Unbestimmte*[1] *über* $\mathbb{K}$*. Dann ist*

$$\chi_A(x) := \det\left(x\mathbb{1}_n - A\right)$$

[1] Ein Ausflug in das Teilgebiet der Unbestimmtentheorie der algebraischen Polynome würde den Rahmen dieser Arbeit sprengen. Eine erhellende Einführung findet man bei Scheja u. Storch [9, S. 23 ff.].

das ***charakteristische Polynom*** *der Matrix A.*

Der Hauptteil der in dieser Arbeit getroffenen Aussagen wird sich auf eine besondere Gruppe von Matrizen beziehen, die ich im Folgenden einführe.

Definition 2.8. *Sei $A \in M_n$. A heißt* ***symmetrische Matrix****, falls $a_{i,j} = a_{j,i}$ für alle verschiedenen $i, j \in \{1, \ldots, n\}$ gilt.*

Diese symmetrischen Matrizen haben einige spezielle Eigenschaften, da sie diagonalisierbar sind.

Definition 2.9. *Eine Matrix $A \in M_n$ heißt* ***diagonalisierbar****, wenn eine invertierbare Matrix $B \in M_n$ existiert, sodass*

$$A = B^{-1}DB$$

für eine Diagonalmatrix D erfüllt ist.

Lemma 2.10. *Eine symmetrische Matrix $A \in M_n$ ist diagonalisierbar.*

Beweis. Siehe Fischer [2][S. 208 f.]. □

Die angesprochenen Eigenschaften halte ich in folgendem Lemma fest.

Lemma 2.11. *Sei $n \in \mathbb{N}$ und $A \in M_n$ eine diagonalisierbare Matrix mit dem charakteristischen Polynom $\chi_A(x)$. Dann gilt:*

(i) $\chi_A(x)$ zerfällt über $\mathbb{R}$ vollständig in Linearfaktoren.

(ii) A besitzt k paarweise verschiedene Eigenwerte $\lambda_i \in \mathbb{R}$ für ein $1 \leq k \leq n$ mit $i \in \{1, \ldots, k\}$.

(iii) $\det(A) = \prod_{i=1}^{n} \lambda_i$.

Beweis. Siehe Fischer [2][S. 171 ff., 205 ff.]. □

Zuletzt komme ich zur namensgebenden Eigenschaft dieser Arbeit.

Definition 2.12. *Sei $A \in M_n$. Die Matrix A heißt genau dann* ***positiv definit****, wenn für alle Vektoren $x \in \mathbb{R}^n$ gilt*

$$x^\tau Ax > 0.$$

Bemerkung 2.13. Entsprechend definiert man *negative Definitheit* durch $x^\tau Ax < 0$ für alle $x \in \mathbb{R}^n$.

Eine schwächere Aussage, die den Hauptteil dieser Arbeit prägen wird, ist die *positive Semidefinitheit* einer Matrix. Sie ist gegeben, falls $x^\tau Ax \geq 0$ für alle $x \in \mathbb{R}^n$ gilt. Auch die negative Semidefinitheit definiert man entsprechend durch $x^\tau Ax \leq 0$.

3 Das charakteristische Polynom

In diesem Kapitel werde ich studieren, wie sich die Hauptminoren und Eigenwerte einer Matrix auf die Koeffizienten des charakteristischen Polynoms der Matrix auswirken. Dies wird einige spannende Zusammenhänge zwischen den Hauptminoren und Eigenwerten zutage fördern, welche ich am Ende des Kapitels darstelle und in Kapitel 4 benutze und erweitere.

3.1 Allgemeine charakteristische Polynome

In diesem Abschnitt befasse ich mich zunächst mit den charakteristischen Polynomen allgemeiner $n \times n$-Matrizen. Deshalb möchte ich von der formulierten Definition 2.7 ausgehen.

Notiert man die Matrix $x\mathbb{1}_n - A$ in Spaltenvektorenschreibweise, so erhält man mit den Einheitsvektoren des $\mathbb{R}^n$ und den Spaltenvektoren $a_i \in \mathbb{R}^n$ der Matrix A die folgende Darstellung: $x\mathbb{1}_n - A = [xe_1 - a_1, xe_2 - a_2, \ldots, xe_n - a_n]$. Nun kommt die Multilinearität der Determinante (Definition 2.3) zum Tragen, welche die folgende Zerlegung ermöglicht:

$$\begin{aligned} &\det\left([xe_1 - a_1, \ldots, xe_i - a_i, \ldots, xe_n - a_n]\right) = \\ &\quad x \cdot \det\left([xe_1 - a_1, \ldots, e_i, \ldots, xe_n - a_n]\right) - \det\left([xe_1 - a_1, \ldots, a_i, \ldots, xe_n - a_n]\right). \end{aligned} \tag{3.1}$$

Diese Zerlegung gibt ein Verfahren vor, bei dem durch wiederholtes Ausnutzen der Multilinearität der Determinante das charakteristische Polynom entwickelt werden kann. Dessen Koeffizienten werden dann durch Kombinationen von Spaltenvektoren der Matrix A gebildet, welche durch Einheitsspaltenvektoren zu $n \times n$-Matrizen vervollständigt werden. Konsequentes Anwenden dieses Verfahrens soll in diesem Kapitel zu einer neuen Darstellung des charakteristischen Polynoms führen.

Die Hauptidee zum Beweis dieser Darstellung ist eine Induktion über die Anzahl der Hauptdiagonaleneinträge, welche die Gestalt $x - a_{i,i}$ haben. Im Spezialfall, dass alle Hauptdiagonaleneinträge dieser Form sind, erhalte ich dann das charakteristische Polynom. Zunächst will ich jedoch eine neue Notationen einführen, um während des Beweises die Übersicht zu erhalten.

3.1.1 Kurznotation für das charakteristische Polynom

Da das charakteristische Polynom $\chi_A(x)$ nach Definition 2.7 als Determinante der Matrix $x\mathbb{1}_n - A$ zu verstehen ist, führe ich die folgende Benennung ein: $A_x := x\mathbb{1}_n - A$. Betrachtet man genauer die Herleitung des charakteristischen Polynoms, so findet man, dass die Äquivalenzumformungen der Eigenwertgleichung (vgl. Definition 2.6), welche zum charakteristischen Polynom in (2.1) führten, ebensogut mit einer Definition des charakteristischen Polynoms durch die Determinante der Matrix $A - x\mathbb{1}_n$ hätten enden können. Diese Matrix benenne ich entsprechend $\underline{A_x} := A - x\mathbb{1}_n$.

Tatsächlich unterscheiden sich beide Definitionen lediglich durch die Vorzeichen der Summanden. Während das aus $\det\left(\underline{A_x}\right)$ hervorgehende charakteristische Polynom stets 1 als Koeffizient vor dem Absolutglied $\det(A)x^0$ zu stehen hat, besitzt das aus $\det(A_x)$ Entstandene selbigen Term als führenden Koeffizienten. Man nennt diese Eigenschaft *Normiertheit* des Polynoms, weshalb ich diese Definition gewählt habe.

Um das in (3.1) angesprochene Verfahren anwenden und verallgemeinern zu können, ist es nötig, diejenigen Stellen in A_x zu kennen, an denen eine solche Zerlegung vorgenommen werden kann. Dies sind diejenigen Hauptdiagonaleneinträge der Form $x - a_{i,i}$. Der (identische) Zeilen- und Spaltenindex soll im Folgenden in einen Mengenindex α einfließen.

Weiterhin interessiere ich mich dafür, an welchen Stellen - respektive in welchen Spalten von A_x bereits eine solche Aufteilung geschehen ist. Diese Spalten enthalten dann entweder die originalen Spaltenvektoren von A oder den entsprechenden Einheitsvektor e_i. Besonders die Einheitsvektoren sind für mich dabei von Interesse, da Laplace-Entwicklung nach einer Spalte mit einem Einheitsvektor die Determinante direkt in einen Hauptminor überführt. Demzufolge fasse ich alle Spaltenindizes von Einheitsspaltenvektoren folgend im Mengenindex β zusammen.

Die Notation soll dann wie folgt funktionieren[1].

Definition 3.1. *Seien $n \in \mathbb{N}$, $A \in M_n$, x eine Unbestimmte über $\mathbb{R}$ sowie $\alpha, \beta \subset \{1, \dots, n\}$ mit $\alpha \cap \beta = \emptyset$. Definiere dann die Spaltenvektoren $a_{x,i}$ der Matrix $A_x(\alpha, \beta) \in M_n$ wie folgt:*

$$A_x(\alpha, \beta) = [a_{x,i}]_{i \in \{1,\dots,n\}} = \begin{cases} xe_i - a_i & \textit{falls } i \in \alpha \\ e_i & \textit{falls } i \in \beta \\ a_i & \textit{sonst} \end{cases}.$$

Bemerkung 3.2. Die Matrix $\underline{A_x}(\alpha, \beta)$ definiere ich entsprechend durch

$$A_x(\alpha, \beta) = [a_{x,i}]_{i \in \{1,\dots,n\}} = \begin{cases} a_i - xe_i & \text{falls } i \in \alpha \\ -e_i & \text{falls } i \in \beta \\ a_i & \text{sonst} \end{cases}.$$

Zur Illustration betrachte man zunächst zwei einfache Beispiele.

Beispiel 3.3. Sei $\alpha = \{i\}$ und $\beta = \{j\}$. Dann ist

$$A_x(\alpha, \beta) = A_x(\{i\}, \{j\}) = \begin{pmatrix} a_{1,1} & \dots & -a_{1,i} & \dots & 0 & \dots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i,1} & \dots & x - a_{i,i} & \dots & 0 & \dots & a_{i,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1} & \dots & -a_{j,i} & \dots & 1 & \dots & a_{j,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \dots & -a_{n,i} & \dots & 0 & \dots & a_{n,n} \end{pmatrix}.$$

Beispiel 3.4. Im obigen Fall (3.1) des charakteristischen Polynoms sind zunächst alle Spaltenvektoren der Gestalt $xe_i - a_i$. Also ist $\alpha = \{1, \dots, n\}$. Dafür enthält die Matrix $x\mathbb{1}_n - A$ keinen Einheitsspaltenvektor. Demnach ist $\beta = \emptyset$. Das Beispiel in (3.1) wird in dieser neuen Notation zu folgendem Ausdruck:

$$\begin{aligned} \det(A_x(\alpha, \beta)) &= \det(A_x(\{1, \dots, n\}, \emptyset)) \\ &= x \cdot \det(A_x(\{1, \dots, n\} \setminus \{i\}, \{i\})) - \det(A_x(\{1, \dots, n\} \setminus \{i\}, \emptyset)). \end{aligned}$$

[1]Sie steht nicht in Widerspruch zu Definition 2.2, da $\det(A_x(\emptyset, \beta)) = \det(A(\beta))$ gilt. Der Index x der Matrix $A_x(\alpha, \beta)$ zeigt hier und im Folgenden an, um welche Notation es sich handelt.

Das Verfahren endet, falls keine Aufspaltung mehr möglich ist, also $\alpha = \emptyset$ gilt. Dann ist eine bestimmte Anzahl von Spaltenvektoren während der Aufspaltung durch Einheitsvektoren ersetzt worden, also ist $n \geq |\beta| = l \geq 0$. Diese Zahl l entspricht der Anzahl, der bei Laplace-Entwicklung nach den Einheitsvektoren zu streichenden Zeilen bzw. Spalten. Die entstehende Determinante ist die einer $(n-l) \times (n-l)$-Hauptstreichungsmatrix von A. Es handelt sich damit also um einen Hauptminor der Ordnung l.

3.1.2 Verallgemeinerung des charakteristischen Polynoms

Nun verallgemeinere ich den Fall des „klassischen“ charakteristischen Polynoms insofern, dass nicht länger alle Hauptdiagonaleneinträge von $A_x(\alpha, \beta)$ der Form $x - a_{i,i}$ sein sollen. Diesem wird die Einheitsmatrix in $x\mathbb{1}_n - A$ nicht gerecht, sodass ich nun eine angepasste Definition treffe.

Definition 3.5. *Seien $n \in \mathbb{N}$ und $\emptyset \neq \alpha \subseteq \{1, \ldots, n\}$. Dann definiere die Matrix $\mathbb{1}_{n,\alpha}$ durch folgende Vorschrift:*

$$\mathbb{1}_{n,\alpha} = (h_{i,j})_{i,j\in\{1,\ldots,n\}} = \begin{cases} h_{i,j} = 1 & \textit{falls} \;\; i = j \in \alpha \\ h_{i,j} = 0 & \textit{sonst} \end{cases}.$$

Nun verstehe ich unter dem verallgemeinerten charakteristischen Polynom - analog zu Definition 2.7 - die Determinante $\det(x\mathbb{1}_{n,\alpha} - A)$ zur Unbestimmten x und der in Definition 3.5 eingeführten Matrix $\mathbb{1}_{n,\alpha}$.

Beispiel 3.6. Die eingangs eingeführte Motivation (3.1) wird (im Unterschied zu Beispiel 3.4) für den allgemeineren Fall mit $\mathbb{1}_{n,\alpha}$ zu folgender Kurzschreibweise:

$$\det\left(A_x\left(\alpha, \emptyset\right)\right) = x \cdot \det\left(A_x\left(\alpha \setminus \{i_j\}, \{i_j\}\right)\right) - \det\left(A_x\left(\alpha \setminus \{i_j\}, \emptyset\right)\right).$$

Hier ist $\alpha \neq \emptyset$ und damit $i_j \in \alpha$.

Nun komme ich zum eigentlichen Ziel dieses Kapitels. Die Beschreibung eines verallgemeinerten charakteristischen Polynoms folgt als Korollar 3.9 aus folgendem Satz.

Satz 3.7. *Sei $n \in \mathbb{N}$ und $\emptyset \neq \alpha, \beta \subseteq \{1, \ldots, n\}$ mit $\alpha \cap \beta = \emptyset$. Außerdem sei $A \in M_n$ und x eine Unbestimmte über $\mathbb{R}$. Dann gilt*

$$\det\left(\underline{A_x}\left(\alpha, \beta\right)\right) = \sum_{l=0}^{|\alpha|} \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l.$$

Beweis. Diesen Beweis führe ich mittels vollständiger Induktion via $k = |\alpha|$.

<u>Induktionsanfang:</u> Sei $k = 1$.
Dann sei $\alpha = \{i\}$ und es gilt

$$\begin{aligned} \det\left(\underline{A_x}\left(\{i\}, \beta\right)\right) &= \det\left(A_x\left(\emptyset, \beta \cup \emptyset\right)\right) + (-1)^1 \det\left(A_x\left(\emptyset, \beta \cup \{i\}\right)\right) x^1 \\ &= \sum_{l=0}^{1} \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l. \end{aligned}$$

Induktionsvoraussetzung: Es gilt für ein beliebiges $k \in \mathbb{N}$ und $\emptyset \neq \epsilon, \beta \subset \{1, \dots, n\}$ mit $\beta \cap \epsilon = \emptyset$ und $|\epsilon| = k - 1$:

$$\det\left(\underline{A_x}\left(\epsilon, \beta\right)\right) = \sum_{l=0}^{k-1} \sum_{\substack{\gamma \subseteq \epsilon \\ |\gamma| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l.$$

Induktionsschritt: Betrachte $\alpha = \{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ mit $|\alpha| = k \in \mathbb{N}$.
Ohne Beschränkung der Allgemeinheit sei $i_j \in \alpha$ für ein $j \in \{1, \dots, k\}$.

$$\begin{aligned}
\det\left(\underline{A_x}\left(\alpha, \beta\right)\right) &= \det\left(\underline{A_x}\left(\alpha \setminus \{i_j\}, \beta\right)\right) - x \det\left(\underline{A_x}\left(\alpha \setminus \{i_j\}, \beta \cup \{i_j\}\right)\right) \\
&\overset{\text{I.V.}}{=} \sum_{l=0}^{k-1} \sum_{\substack{\gamma \subseteq \alpha \\ |\gamma| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l \\
&+ \sum_{l=0}^{k-1} \sum_{\substack{\gamma \subseteq \alpha \\ |\gamma| = l}} (-1)^{l+1} \det\left(A_x\left(\emptyset, \beta \cup \gamma \cup \{i_j\}\right)\right) x^{l+1} \\
&= \sum_{l=0}^{k-1} \sum_{\substack{\gamma \subseteq \alpha \\ |\gamma| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l \\
&+ \sum_{l=0}^{k-1} \sum_{\substack{\gamma^\star \subseteq \alpha \\ i_j \in \gamma^\star \\ |\gamma^\star| = l+1}} (-1)^{l+1} \det\left(A_x\left(\emptyset, \beta \cup \gamma^\star\right)\right) x^{l+1} \\
&= \sum_{l=0}^{k-1} \sum_{\substack{\gamma \subseteq \alpha \\ |\gamma| = l}} (-1)^l \det\left(A_x\left(\emptyset, \gamma\right)\right) x^l + \sum_{l=1}^{k} \sum_{\substack{\gamma^\star \subseteq \alpha \\ i_j \in \gamma^\star \\ |\gamma^\star| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma^\star\right)\right) x^l \\
&= \det\left(A_x\left(\emptyset, \beta\right)\right) + (-1)^k \det\left(A_x\left(\emptyset, \beta \cup \{i_1, \dots, i_k\}\right)\right) x^k \\
&+ \sum_{l=1}^{k-1} \sum_{\substack{\gamma \subseteq \alpha \\ |\gamma| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l + \sum_{l=1}^{k-1} \sum_{\substack{\gamma^\star \subseteq \alpha \\ i_j \in \gamma^\star \\ |\gamma^\star| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma^\star\right)\right) x^l \\
&= \det\left(A_x\left(\emptyset, \beta\right)\right) + (-1)^k \det\left(A_x\left(\emptyset, \beta \cup \{i_1, \dots, i_k\}\right)\right) x^k \\
&+ \sum_{l=1}^{k-1} \sum_{\substack{\gamma \subseteq \alpha \\ |\gamma| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l \\
&= \sum_{l=0}^{k} \sum_{\substack{\gamma \subseteq \alpha \\ |\gamma| = l}} (-1)^l \det\left(A_x\left(\emptyset, \beta \cup \gamma\right)\right) x^l
\end{aligned}$$

Damit gilt die Behauptung für alle $|\alpha| = k \in \mathbb{N}$. □

Korollar 3.8. *Sei $n \in \mathbb{N}$ und $\emptyset \neq \alpha \subseteq \{1, \dots, n\}$ sowie $\mathbb{1}_{n,\alpha}, A \in M_n$. Ferner sei x eine Unbestimmte über $\mathbb{R}$. Dann gilt:*

$$\det\left(A_x\left(\alpha,\beta\right)\right) = \sum_{l=0}^{|\alpha|} \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=l}} (-1)^{n-l} \det\left(A_x\left(\emptyset, \beta\cup\gamma\right)\right) x^l.$$

Beweis. Aus der Multilinearität der Determinante (Definition 2.3) folgt:

$$\det\left(A_x\left(\alpha,\beta\right)\right) = (-1)^n \det\left(\underline{A_x}\left(\alpha,\beta\right)\right).$$

Nun gilt mit Satz 3.7

$$\begin{aligned}
\det\left(A_x\left(\alpha,\beta\right)\right) &= (-1)^n \cdot \sum_{l=0}^{|\alpha|} \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=l}} (-1)^l \det\left(A_x\left(\emptyset, \beta\cup\gamma\right)\right) x^l \\
&= \sum_{l=0}^{|\alpha|} \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=l}} (-1)^{n+l} \underbrace{\left((-1)^{-l}\right)^2}_{=1} \det\left(A_x\left(\emptyset, \beta\cup\gamma\right)\right) x^l \\
&= \sum_{l=0}^{|\alpha|} \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=l}} (-1)^{n-l} \det\left(A_x\left(\emptyset, \beta\cup\gamma\right)\right) x^l
\end{aligned}$$

und die Behauptung. □

Für den Spezialfall $\alpha = \{1, \dots, n\}, \beta = \emptyset$ ergibt sich damit das charakteristische Polynom.

Korollar 3.9. *Sei $n \in \mathbb{N}$ und $A \in M_n$ eine Matrix mit dem charakteristischen Polynom $\chi_A(x)$.*

Dann ist der Koeffizient des k-ten Glieds des Polynoms bis auf das Vorzeichen eindeutig durch die Summe der Hauptminoren der Ordnung k bestimmt und ferner gilt

$$\chi_A(x) = \sum_{l=0}^{n} \sum_{\substack{\gamma\subseteq\{1,\dots,n\} \\ \gamma=l}} (-1)^{n-l} \det\left(A_x\left(\emptyset,\gamma\right)\right) x^l.$$

Beweis. Nach Definition 2.7 ist das charakteristische Polynom $\chi_A(x) = \det\left(A_x\left(\{1,\dots,n\},\emptyset\right)\right)$. Das ist ein Spezialfall von Korollar 3.8 für $\alpha = \{1, \dots, n\}$ und $\beta = \emptyset$. Daraus folgt die Behauptung. □

3.2 Vollständig zerfallende charakteristische Polynome

Grenzt man die Klasse der betrachteten Matrizen auf symmetrische Matrizen ein, so bekommt man nach Lemma 2.11 ein über $\mathbb{R}$ vollständig in Linearfaktoren zerfallendes Polynom mit den Eigenwerten der Matrix als Nullstellen. Diese mächtige Eigenschaft reell symmetrischer Matrizen, lenkt den Blick zwangsläufig auf eine genauere Beschreibung desselben.

Elementarsymmetrische Polynome

Naiv nähere ich mich dem durch Ausmultiplizieren. Im Folgenden gilt die folgende Bezeichnung $\lambda_{i,j} = \lambda_i \cdot \lambda_j$ für die über $\mathbb{R}$ unabhängigen Unbestimmten $\lambda_1, \dots, \lambda_4, x$.

$$\begin{aligned}
(x-\lambda_1)(x-\lambda_2) &= x^2 - (\lambda_1+\lambda_2)x + \lambda_{1,2} \\
(x-\lambda_1)(x-\lambda_2)(x-\lambda_3) &= x^3 - (\lambda_1+\lambda_2+\lambda_3)x^2 + (\lambda_{1,2}+\lambda_{1,3}+\lambda_{2,3})x - \lambda_{1,2,3} \\
(x-\lambda_1)(x-\lambda_2)(x-\lambda_3)(x-\lambda_4) &= x^4 - (\lambda_1+\lambda_2+\lambda_3+\lambda_4)x^3 \\
&\quad + (\lambda_{1,2}+\lambda_{1,3}+\lambda_{1,4}+\lambda_{2,3}+\lambda_{2,4}+\lambda_{3,4})x^2 \\
&\quad - (\lambda_{1,2,3}+\lambda_{1,3,4}+\lambda_{2,3,4})x + \lambda_{1,2,3,4}
\end{aligned}$$

Es wird deutlich, dass der Koeffizient der $n-k$-ten Potenz der Unbestimmten x alle verschiedenen Kombinationen der Indexmenge $\{1, \dots, n\}$ der entsprechenden Länge k als Summanden aufführt. Dies lässt sich durch folgende Notation codieren.

Definition 3.10. *Sei $n \in \mathbb{N}$ und $i_1, \dots, i_n \in \{1, \dots, n\}$ paarweise verschieden. Weiterhin seien $k \in \mathbb{N}$ und $\lambda_1, \dots, \lambda_n$ unabhängige Unbestimmte über $\mathbb{R}$. Der Term*

$$S_k(\lambda_1, \dots, \lambda_n) := \sum_{1 \le i_1 < \dots < i_k \le n} \prod_{j=1}^{k} \lambda_{i_j}$$

heißt das k-te ***Elementarsymmetrische Polynom*** *in den Unbestimmten $\lambda_1, \dots, \lambda_n$.*

Bemerkung 3.11. Es ist sinnvoll $S_0(\lambda_1, \dots, \lambda_n) = 1$ zu setzen.
Für $k = 1$ gilt:

$$S_1(\lambda_1, \dots, \lambda_n) = \sum_{1 \le i_1 \le n} \prod_{j=1}^{1} \lambda_{i_j} = \sum_{1 \le i_1 \le n} \lambda_{i_1} = \lambda_1 + \dots + \lambda_n.$$

Dagegen gilt für $k = n$:

$$S_n(\lambda_1, \dots, \lambda_n) = \sum_{1 \le i_1 < \dots < i_n \le n} \prod_{j=1}^{n} \lambda_{i_j} = \lambda_1 \cdot \dots \cdot \lambda_n$$

Weiter gelten die folgenden Rechenregeln für die Elementarsymmetrischen Polynome. Ich verwende zur Abkürzung die Schreibweise $S_k(\lambda_1, \dots, \lambda_n) = S_{k,n}$.
Für $k = 1$ gilt:

$$S_{1,n} = S_{1,n-1} + \lambda_n. \tag{3.2}$$

Für $k \in \{2, \dots, n-1\}$ gilt die rekursive Formel:

$$S_{k,n} = S_{k,n-1} + S_{k-1,n-1} \cdot \lambda_n. \tag{3.3}$$

Für $k = n$ gilt:

$$S_{n,n} = S_{n-1,n-1} \cdot \lambda_n. \tag{3.4}$$

Die Regeln (3.2) und (3.4) sieht man schnell ein.

$$\begin{aligned}
S_{1,n} &= \sum_{1 \le i \le n} \lambda_i = \lambda_1 + \dots + \lambda_{n-1} + \lambda_n = \sum_{1 \le i \le n-1} \lambda_i + \lambda_n = S_{1,n-1} + \lambda_n \\
S_{n,n} &= \prod_{1 \le i \le n} \lambda_i = \lambda_1 \cdot \dots \cdot \lambda_{n-1} \cdot \lambda_n = \prod_{1 \le i \le n-1} \lambda_i \cdot \lambda_n \qquad = S_{n-1,n-1} \cdot \lambda_n
\end{aligned}$$

Regel (3.3) zeige ich dagegen mit einer kurzen Induktion.

Beweis. Den Beweis führe ich via vollständiger Induktion nach $k \in \{2, \dots, n-1\} \subset \mathbb{N}$.
Induktionsanfang: Sei $k = 2$.

$$S_{2,n} = \sum_{i=1}^{n-1} \sum_{j>i}^{n} \lambda_i \cdot \lambda_j = \sum_{i=1}^{n-2} \sum_{j>i}^{n-1} \lambda_i \cdot \lambda_j + \sum_{i=1}^{n-1} \lambda_i \cdot \lambda_n = S_{2,n-1} + S_{1,n-1} \cdot \lambda_n$$

Induktionsvoraussetzung: Es gelte $S_{k,n} = S_{k,n-1} + S_{k-1,n-1} \cdot \lambda_n$ für ein $k \in \{2, \dots, n-2\} \subset \mathbb{N}, k < n$.
Induktionsschritt: Betrachte $k+1 < n$.

$$\begin{aligned} S_{k+1,n} = \sum_{1 \leq i_1 < \dots < i_{k+1} \leq n} \prod_{j=1}^{k+1} \lambda_{i_j} &= \lambda_n \cdot \sum_{1 \leq i_1 < \dots < i_k < n} \prod_{j=1}^{k} \lambda_{i_j} + \sum_{1 \leq i_1 < \dots < i_{k+1} < n} \prod_{j=1}^{k+1} \lambda_{i_j} \\ &= \lambda_n \cdot \sum_{1 \leq i_1 < \dots < i_k \leq n-1} \prod_{j=1}^{k} \lambda_{i_j} + \sum_{1 \leq i_1 < \dots < i_{k+1} \leq n-1} \prod_{j=1}^{k+1} \lambda_{i_j} = \lambda_n \cdot S_{k,n-1} + S_{k+1,n-1} \end{aligned}$$

Damit gilt (3.3) für alle $k \in \{2, \dots, n-1\} \subset \mathbb{N}$ für beliebige $n \in \mathbb{N}$. □

Die Anwendbarkeit der vorgestellten Elementarsymmetrischen Polynome auf das vorliegende Problem des charakteristischen Polynoms symmetrischer Matrizen garantiert das folgende Lemma. Für dessen Beweis werden die eben gezeigten Rechenregeln von Bedeutung sein.

Lemma 3.12. *Für beliebige $n \in \mathbb{N}$ und die über $\mathbb{R}$ unabhängigen Unbestimmten $\lambda_1, \dots, \lambda_n, x$ gilt:*

$$\prod_{l=1}^{n} (x - \lambda_l) = \sum_{l=0}^{n} (-1)^l S_l(\lambda_1, \dots, \lambda_n) x^{n-l}. \tag{3.5}$$

Beweis. Den Beweis führe ich via vollständiger Induktion nach $n \in \mathbb{N}$.
Induktionsanfang: Sei $n = 1$.

$$\prod_{l=1}^{1} (x - \lambda_l) = x - \lambda_1 = (-1)^0 S_0(\lambda_1) x^1 + (-1)^1 S_1(\lambda_1) x^0 = \sum_{l=0}^{1} (-1)^l S_l(\lambda_1) x^{1-l}$$

Induktionsvoraussetzung: Es gelte $\prod_{l=1}^{n} (x - \lambda_l) = \sum_{l=0}^{n} (-1)^l S_l(\lambda_1, \dots, \lambda_n) x^{n-l}$ für ein $n \in \mathbb{N}$.

<u>Induktionsschritt:</u> Betrachte $n+1$.

$$\begin{aligned}
\prod_{l=1}^{n+1}(x-\lambda_l) &= (x-\lambda_{n+1})\cdot\prod_{l=1}^{n}(x-\lambda_l) \overset{\text{I.V.}}{=} (x-\lambda_{n+1})\cdot\sum_{l=0}^{n}(-1)^l S_l(\lambda_1,\dots,\lambda_n)x^{n-l} \\
&= \sum_{l=0}^{n}(-1)^l S_{l,n}x^{n-l+1} - \lambda_{n+1}\cdot\sum_{l=0}^{n}(-1)^l S_{l,n}x^{n-l} \\
&= \sum_{l=0}^{n}(-1)^l S_{l,n}x^{n-l+1} + \sum_{l=0}^{n}(-1)^{l+1}\lambda_{n+1}S_{l,n}x^{n-l} \\
&= \sum_{l=0}^{n}(-1)^l S_{l,n}x^{n-l+1} + \sum_{l=2}^{n-1}(-1)^{l+1}\lambda_{n+1}S_{l,n}x^{n-l} \\
&\quad + (-1)^1\lambda_{n+1}S_{0,n}x^n + (-1)^2\lambda_{n+1}S_{1,n}x^{n-1} + (-1)^{n+1}\lambda_{n+1}S_{n,n}x^0 \\
&= \sum_{l=3}^{n}(-1)^l S_{l,n}x^{n+1-l} + (-1)^0 S_{0,n}x^{n+1} + (-1)^1 S_{1,n}x^n + (-1)^2 S_{2,n}x^{n-1} \\
&\quad + \sum_{l=2}^{n-1}(-1)^{l+1}\underbrace{\lambda_{n+1}S_{l,n}}_{\overset{(3.3)}{=}S_{l+1,n+1}-S_{l+1,n}}x^{n-l} + (-1)^1\lambda_{n+1}S_{0,n}x^n \\
&\quad + (-1)^2\lambda_{n+1}S_{1,n}x^{n-1} + (-1)^{n+1}\lambda_{n+1}S_{n,n}x^0 \\
&= \sum_{l=2}^{n-1}(-1)^{l+1}S_{l+1,n}x^{n-l} + \sum_{l=2}^{n-1}(-1)^{l+1}\left(S_{l+1,n+1}-S_{l+1,n}\right)x^{n-l} + (-1)^0 x^{n+1} \\
&\quad + (-1)^1\underbrace{(S_{1,n}+\lambda_{n+1})}_{\overset{(3.2)}{=}S_{1,n+1}}x^n + (-1)^2\underbrace{(\lambda_{n+1}S_{1,n}+S_{2,n})}_{\overset{(3.3)}{=}S_{2,n+1}}x^{n-1} + (-1)^{n+1}\underbrace{\lambda_{n+1}S_{n,n}}_{\overset{(3.4)}{=}S_{n+1,n+1}} \\
&= \sum_{l=2}^{n-1}(-1)^{l+1}S_{l+1,n+1}x^{n-l} + (-1)^0 S_{0,n+1}x^{n+1} + (-1)^1 S_{1,n+1}x^n \\
&\quad + (-1)^2 S_{2,n+1}x^{n-1} + (-1)^{n+1}S_{n+1,n+1}x^0 \\
&= \sum_{l=3}^{n}(-1)^l S_{l,n+1}x^{n+1-l} + (-1)^0 S_{0,n+1}x^{n+1-0} + (-1)^1 S_{1,n+1}x^{n+1-1} \\
&\quad + (-1)^2 S_{2,n+1}x^{n+1-2} + (-1)^{n+1}S_{n+1,n+1}x^{n+1-(n+1)} \\
&= \sum_{l=0}^{n+1}(-1)^l S_{l,n+1}x^{n+1-l}
\end{aligned}$$

Damit gilt $\prod\limits_{l=1}^{n}(x-\lambda_l) = \sum\limits_{l=0}^{n}(-1)^l S_{l,n}x^{n-l}$ für beliebige $n\in\mathbb{N}$. $\square$

Bemerkung 3.13. Lemma 3.12 auf das charakteristische Polynom mit seinen Nullstellen (den Eigenwerten $\lambda_1,\dots,\lambda_n$ der Matrix A) angewandt, erlaubt die folgende Schreibweise.

$$\chi_A(x) = \prod_{l=1}^{n}(x-\lambda_l) = \sum_{l=0}^{n}(-1)^l S_l(\lambda_1,\dots,\lambda_n)x^{n-l} \tag{3.6}$$

3.3 Schlussfolgerungen

Die bemerkenswerte Folgerung aus Korollar 3.9 ist, dass der i-te Koeffizient[2] des charakteristischen Polynoms bis auf das Vorzeichen durch die Summe der Hauptminoren von der Ordnung i eindeutig bestimmt ist. Eine zweite Schlussfolgerung kann man aus der Synthese von Lemma 3.12 und Korollar 3.9 gewinnen.

Korollar 3.14. *Sei $A \in M_n$ eine symmetrische Matrix mit den Eigenwerten $\lambda_1, \ldots, \lambda_n$. $S_k(\lambda_1, \ldots, \lambda_n)$ bezeichne das k-te elementarsymmetrische Polynom der Eigenwerte von A. Dann gilt:*

$$S_{n-k}(\lambda_1, \ldots, \lambda_n) = \sum_{\substack{\alpha \subseteq \{1,\ldots,n\} \\ |\alpha| = k}} \det\left(A_x\left(\emptyset, \alpha\right)\right).$$

Beweis. Sei A eine symmetrische Matrix. Nach Lemma 2.11 zerfällt dann das charakteristische Polynom χ_A vollständig in Linearfaktoren $(x - \lambda_i)$ für $1 \leq i \leq n$. Nun gibt es laut Lemma 3.12 und Korollar 3.9 zwei Möglichkeiten des Umschreibens des charakteristischen Polynoms.

$$\sum_{l=0}^{n} \sum_{\substack{\alpha \subseteq \{1,\ldots,n\} \\ |\alpha| = l}} (-1)^{n-l} \det\left(A_x\left(\emptyset, \alpha\right)\right) x^l = \chi_A(x) = \prod_{k=0}^{n} (x - \lambda_k) = \sum_{k=0}^{n} (-1)^k S_{k,n} x^{n-k}$$

Ein Koeffizientenvergleich für $l = n - k$ liefert die Behauptung. □

Mittels Korollar 3.14 findet man nun, dass die Summe der Hauptminoren einer Größe und die Summe der Produkte von Eigenwerten der entsprechenden Länge übereinstimmen. Dieses Resultat ist insofern bemerkenswert, als dass sich bekannte Identitäten für Determinante und Spur einer symmetrischen Matrix als Spezialfälle aus diesem Korollar ergeben.

Im Spezialfall des 0-ten Summanden liefert Korollar 3.14 die bereits in Lemma 2.11 bewiesene Identität der Determinante als Produkt der Eigenwerte einer symmetrischen Matrix.

Weiterhin bekommt man für den $n-1$-ten Summanden den Spezialfall, dass die Summe der Diagonaleinträge[3] einer symmetrischen Matrix mit der Summe ihrer Eigenwerte übereinstimmt. Dies gilt, obwohl Diagonaleinträge einer symmetrischen Matrix[4] von den Eigenwerten völlig verschieden sein können.

[2]Ich meine den Koeffizienten der i-ten Potenz von x.
[3]Diese Summe nennt man auch die *Spur* $\mathrm{tr}(A)$ einer Matrix A.
[4]Sofern diese nicht Diagonalgestalt hat.

4 Hauptminoren und Definitheit

Dieses Kapitel soll bei der Lösung des Problems des Auffindens von Nullstellen reeller homogener quadratischer Polynome einen Beitrag leisten. Zu diesem Zweck erinnere man sich an die oben gemachte Definition 2.12 der positiven Definitheit. Von dieser ausgehend will ich zunächst einige Charakterisierungen der positiven Semidefinitheit vornehmen, welche die folgenden Beweise vereinfachen werden.

Der Hauptgegenstand dieses Kapitels wird eine Verallgemeinerung des Hauptminorenkriteriums für die positive Semidefinitheit sein. Nach der Einführung einer neuen Schreibweise für die quadratische Form einer Streichungsmatrix gelingt die Beweisführung durch Benutzung einiger Ergebnisse aus Kapitel 3. Vor allem die zuvor ermittelte Beschaffenheit der Koeffizienten des charakteristischen Polynoms wird eine tragende Rolle spielen.

4.1 Charakterisierungen der positiven Semidefinitheit

Zunächst will ich die Beziehung von Eigenwerten und Determinante einer symmetrischen Matrix mit der positiven Semidefinitheit herausarbeiten. Erstere findet sich in folgendem Lemma.

Lemma 4.1. *Sei $A \in M_n$ eine symmetrische Matrix mit den Eigenwerten $\lambda_1, \ldots, \lambda_n$. A ist genau dann positiv semidefinit, das heißt es gilt $x^\tau A x \geq 0$ für alle $x \in \mathbb{R}^n$, wenn jeder Eigenwert λ_i für $1 \leq i \leq n$ der Matrix nichtnegativ ist.*

Beweis. Sei $A \in M_n$ symmetrisch.

„$\Rightarrow$“: Sei A positiv semidefinit, dann gilt für alle $x \in \mathbb{R}^n : x^\tau A x \geq 0$. Da dies für alle x gilt, ist dies insbesondere für die Eigenvektoren $\tilde{x} \neq 0$ der Matrix A erfüllt und es ergibt sich für beliebige Eigenwerte

$$\tilde{x}^\tau A \tilde{x} = \tilde{x}^\tau \lambda \tilde{x} = \lambda \underbrace{|\tilde{x}|^2}_{>0} \Rightarrow \lambda \geq 0.$$

„$\Leftarrow$“: Sei jeder Eigenwert λ_i der Matrix A nichtnegativ. Da A symmetrisch, gibt es eine Diagonalmatrix D, so dass $A = O^\tau D O$. Dann folgt für alle $x \in \mathbb{R}^n$:

$$x^\tau A x = x^\tau O^\tau D O x = (Ox)^\tau D (Ox) = \sum_{i=1}^{n} \underbrace{\lambda_i}_{\geq 0} \underbrace{\left(\sum_{j=1}^{n} o_{i,j} x_j \right)^2}_{\geq 0} \geq 0.$$

Dann ist A positiv semidefinit und es folgt die behauptete Äquivalenz. □

Ich erweitere diese Aussage noch in Bezug auf die Determinante der Matrix.

Lemma 4.2. *Sei $A \in M_n$ eine symmetrische Matrix. Ist A positiv semidefinit, so ist auch* $\det(A) \geq 0$.

Beweis. Aus Lemma 4.1 folgt die Nichtnegativität aller Eigenwerte von A. Wegen Lemma 2.11 ist die Determinante der symmetrischen Matrix gerade das Produkt ihrer Eigenwerte, sodass die Behauptung folgt. □

Die Umkehrung des Lemma 4.2 gilt im Allgemeinen nicht, denn die Nichtnegativität eines Produkts reeller Zahlen impliziert nicht zwangsläufig die Nichtnegativität aller Zahlen. Dies erschwert den Beweis der Rückrichtung des Hauptminorenkriteriums. Andernfalls wäre eine nichtnegative Determinante schon hinreichend für die positive Semidefinitheit einer Matrix.

4.2 Das erweiterte Hauptminorenkriterium

Nun werde ich in diesem Kapitel in der Tradition der Schreibweise in Definition 2.2 eine Notation für die quadratische Form einer Hauptstreichungsmatrix entwickeln. Mit dieser wird die Folgerung nichtnegativer Hauptminoren aus der positiven Semidefinitheit wesentlich vereinfacht.

Nach Definition 2.2 besitzt eine Hauptstreichungsmatrix $A(\alpha)$ Nullspalten-, bzw. Nullzeilenvektoren an den in α indizierten Stellen. Wegen

$$x^\tau A(\alpha)x = \sum_{j\notin\alpha} x_j \left(\sum_{i\notin\alpha} x_i a_{i,j}\right)$$

lässt sich diese Situation ebenso nachbilden, wenn der Vektor x in der quadratischen Form in entsprechenden Komponenten $i, j \in \alpha$ eine 0 enthält. Deshalb treffe ich folgende Definition.

Definition 4.3. *Sei $n \in \mathbb{N}$ und $\alpha \subset \{1, \ldots, n\}$. Dann definiere den Vektor $x_\alpha \in \mathbb{R}^n$ durch*

$$x_\alpha = (x_i)_{i\in\{1,\ldots,n\}} = \begin{cases} x_i = 0 & \textit{falls } i \in \alpha \\ x_i \in \mathbb{R} & \textit{falls } i \notin \alpha \end{cases}.$$

Diese Notation ist insofern gewinnbringend, als dass man nun erhält:

$$x^\tau A(\alpha)x = \sum_{j\notin\alpha} x_j \left(\sum_{i\notin\alpha} x_i a_{i,j}\right) = x_\alpha^\tau A x_\alpha. \tag{4.1}$$

Dabei vergegenwärtige man sich, dass $x_\alpha \in \mathbb{R}^n$. Darauf stützt sich die folgende Argumentation.

Lemma 4.4. *Sei $A \in M_n$ eine symmetrische Matrix und positiv semidefinit. Sodann sind alle Hauptminoren der Matrix A nichtnegativ.*

Beweis. Sei $A \in M_n$ eine positiv semidefinite Matrix. Dann gilt nach Definition

$$\forall x \in \mathbb{R}^n, x \neq 0 : x^\tau A x \geq 0.$$

Da dies für alle $x \in \mathbb{R}^n$ gilt, ist dies insbesondere für alle $x_\alpha \in \mathbb{R}^n$ erfüllt, die $\alpha \subsetneq \{1, \ldots, n\}$ genügen. Man erhält mit (4.1) für beliebige (nicht verschwindende) Hauptstreichungsmatrizen $A(\alpha)$:

$$\forall x \in \mathbb{R}^n, x \neq 0 : x^\tau A(\alpha)x \overset{(4.1)}{=} x_\alpha^\tau A x_\alpha \geq 0,$$

welche somit positiv semidefinit sind. Da jede beliebige Hauptstreichungsmatrix positiv semidefinit und symmetrisch ist, folgt mit Lemma 4.2 die Nichtnegativität jedes Hauptminors und somit die Behauptung. □

Dieses Lemma verkürzt die Beweisführung im folgenden Satz erheblich.

Satz 4.5 (erweitertes Hauptminorenkriterium)**.** *Sei $n \in \mathbb{N}$ und $A \in M_n$ eine symmetrische Matrix. Dann sind äquivalent:*

(i) A ist positiv semidefinit.

(ii) Alle Hauptminoren von A sind nichtnegativ.

(iii) Die Summen aller Hauptminoren der gleichen Ordnung sind nichtnegativ.

(iv) Alle Eigenwerte der Matrix A sind nichtnegativ.

Beweis. Sei $A \in M_n$ symmetrisch.

„(i) $\Rightarrow$ (ii)“: Das ist die Aussage von Lemma 4.4.

„(ii) $\Rightarrow$ (iii)“: Dies folgt direkt, da die Hauptminoren reelle Zahlen sind und die Summation nichtnegativer reeller Zahlen den Wert nicht verringert.

„(iii) $\Rightarrow$ (iv)“: Widerspruchsbeweis. Angenommen es gibt einen Eigenwert λ der Matrix A mit $\lambda < 0$.
Nun sind nach Korollar 3.9 die Koeffizienten des charakteristischen Polynoms (bis auf das Vorzeichen) eindeutig durch die nichtnegativen Summen der Hautpminoren bestimmt. Folglich gilt für ein Polynom mit geradem Grad n:

$$\chi_A(\lambda) = \underbrace{\lambda^n}_{>0} - \underbrace{\lambda^{n-1}}_{<0} \underbrace{\sum_{\substack{\gamma \subset \{1,\ldots,n\} \\ |\gamma| = n-1}} \det(A_x(\emptyset,\gamma))}_{\overset{\text{(iii)}}{\geq} 0} + \underbrace{\lambda^{n-2}}_{>0} \underbrace{\sum_{\substack{\gamma \subset \{1,\ldots,n\} \\ |\gamma| = n-2}} \det(A_x(\emptyset,\gamma))}_{\overset{\text{(iii)}}{\geq} 0} \mp \cdots + \underbrace{\det(A_x(\emptyset,\emptyset))}_{\overset{\text{(iii)}}{\geq} 0}$$
$$> 0.$$

Analog folgt $\chi_A(\lambda) < 0$ für ein Polynom ungeraden Grades.

Dies ist jedoch jeweils ein Widerspruch zur Eigenwerteigenschaft $\chi_A(\lambda) = 0$ von λ. Also muss $\lambda \geq 0$ für alle Eigenwerte gelten.

„(iv) $\Rightarrow$ (i)“: Diese Aussage findet man in Lemma 4.1. □

In diesem Satz 4.5 steckt zunächst bereits in kondensierter Form das erweiterte Hauptminorenkriterium. Im Gegensatz zur positiven Definitheit, ist es nicht ausreichend, nur die führenden Hauptminoren zu betrachten, da die Koeffizienten des charakteristischen Polynoms aus allen Hauptminoren gebildet werden.

Schließlich beweist Satz 4.5 auch eine erstaunliche Implikation für symmetrische Matrizen. Die Richtung „(iii) $\Rightarrow$ (ii)“ ist sehr bemerkenswert, da im Allgemeinen die Summe reeller Zahlen kaum Rückschluss auf ihre Summanden erlaubt.

4.3 Vergleich mit dem Hauptminorenkriterium der positiven Definitheit

In diesem Abschnitt will ich einen Einblick in das Hauptminorenkriterium von Sylvester[1] geben und dessen Beweis vermittels der Cholesky-Zerlegung[2] skizzieren. Das Kriterium von Sylvester greift die positive Definitheit auf, welche in Definition 2.12 eingeführt wurde. Es stützt sich lediglich auf Aussagen über führende Hauptminoren, die ich an dieser Stelle kurz einführen will.

Bemerkung 4.6. Ein Hauptminor $\det(A(\alpha))$ einer Matrix $A \in M_n$ heißt **führend**, wenn die Indexmenge α der Gestalt $\alpha = \{k+1, k+2, \ldots, n-1, n\}$ für ein $k \in \mathbb{N}, k < n$ ist.

Sodann notiere ich im Folgenden $\det(A_k)$, wobei A_k selbst als führende Hauptstreichungsmatrix referenziert wird.

[1]benannt nach James Joseph Sylvester (1814-1897)
[2]benannt nach André-Louis Cholesky (1875-1914)

Dieses Beweisverfahren ist auf invertierbare Matrizen beschränkt, wie ich bei der Diskussion der Existenz der LU-Zerlegung nachweisen werde. Deshalb eignet es sich auch nicht für die Beweisführung in den vorangegangenen Kapiteln, denn in der positiven Semidefinitheit ist der Fall einer verschwindenden Determinante enthalten und solche Matrizen sind nicht invertierbar.

Eine Einführung in die Natur der LU-Zerlegung findet man bei Meyer [6, S. 141 ff.]. Abgekürzt geht diese auf das Gauß'sche Eliminationsverfahren zum Lösen von Gleichungssystemen zurück und betrachtet den Fall, dass während des gesamten Algorithmus niemals Zeilenvertauschungen nötig werden[3]. Dann lassen sich alle elementaren Zeilenumformungen durch sogenannte Elementarmatrizen darstellen, welche der Form $G_k = \mathbb{1}_n - \alpha_k e_j e_i^\tau$ sind. Diese entsprechen bei linksseitiger Multiplikation der Addition der mit $\alpha_k \in \mathbb{R} \setminus \{0\}$ vervielfachten i-ten Zeile zur j-ten Zeile einer Matrix im k-ten Eliminationsschritt. Von diesen werden insgesamt n benötigt, sodass sich diese Elementarmatrizen zusammenfassen lassen.

$$(G_n \cdot \dots \cdot G_1)\, A = U \Leftrightarrow A = \underbrace{\left(G_1^{-1} \cdot \dots \cdot G_n^{-1}\right)}_{=:L} U$$

Darin ist L eine untere Dreiecksmatrix der Gestalt

$$L = (l_{i,j})_{i\in\{1,\dots,n\},j\in\{1,\dots,n\}} = \begin{cases} l_{i,j} \in \mathbb{R} & \text{für } i > j \\ l_{i,j} = 1 & \text{für } i = j \\ l_{i,j} = 0 & \text{sonst} \end{cases} \tag{4.2}$$

und U eine obere Dreiecksmatrix der Form

$$U = (u_{i,j})_{i\in\{1,\dots,n\},j\in\{1,\dots,n\}} = \begin{cases} u_{i,j} \in \mathbb{R} & \text{für } i < j \\ u_{i,j} \neq 0 & \text{für } i = j\ . \\ u_{i,j} = 0 & \text{sonst} \end{cases} \tag{4.3}$$

Zunächst werde ich diskutieren, wann diese Zerlegung sicher existiert.

Lemma 4.7. *Sei $A \in M_n$. A besitzt genau dann eine LU-Zerlegung, d.h. es gilt $A = LU$ für Matrizen $L, U \in M_n$ wie in* (4.2) *und* (4.3)*, wenn alle führenden Hauptminoren der Matrix A nicht* 0 *sind. Weiterhin sind dann L und U durch* (4.2) *und* (4.3) *eindeutig festgelegt.*

Beweis. [4] Sei $A \in M_n$.

<u>*Existenz:*</u> „$\Rightarrow$“:

Es gelte $A = LU$ mit L und U wie oben. Nun teile man die Matrix A wie folgt auf.

$$A = LU = \begin{pmatrix} L_{1,1} & 0 \\ L_{2,1} & L_{2,2} \end{pmatrix} \begin{pmatrix} U_{1,1} & U_{1,2} \\ 0 & U_{2,2} \end{pmatrix} = \begin{pmatrix} L_{1,1}U_{1,1} & \star \\ \star & \star \end{pmatrix}$$

Hier sind $L_{1,1}$ und $U_{1,1}$ gerade $k \times k$-Blockmatrizen, welche ebenso (4.2) und (4.3) genügen. Beide sind regulär, da sie von 0 verschiedene Hauptdiagonaleneinträge haben[5]. Sodann ist die führende Hauptstreichungsmatrix $A_k = L_{1,1}U_{1,1}$ ein Produkt regulärer Matrizen - also selbst regulär und besitzt somit eine von 0 verschiedene Determinante.

[3] Dies ist z.B. der Fall, wenn das Pivot-Element 0 ist.

[4] Ich halte mich hier sehr eng an Meyer [6, S. 144,149 f.].

[5] Um dies einzusehen, entwickle man nach dem Laplace'schen Entwicklungssatz jeweils nach der führenden Zeile bzw. Spalte. Diese hat nur den Hauptdiagonaleneintrag und sonst 0, so dass als Determinante das Produkt der Hauptdiagonaleneinträge verbleibt.

„$\Leftarrow$“ Sei jeder führende Hauptminor von A verschieden von 0. Dann ist jede führende Hauptstreichungsmatrix A_k regulär. Zeige induktiv nach k, dass jede führende Hauptstreichungsmatrix eine LU-Zerlegung besitzt.

Induktionsanfang: Sei $k = 1$ und A_1 regulär.
Die gesuchte Zerlegung von $A_1 = (a_{1,1})$ ist $A_1 = (1)(a_{1,1})$.

Induktionsvoraussetzung: A_k sei regulär und habe eine LU-Zerlegung für ein $k \in \mathbb{N}$.

Induktionsschritt: Sei A_{k+1} regulär.
Es sei $A_k = L_k U_k$ die LU-Zerlegung von A_k und weiter gilt $A_k^{-1} = U_k^{-1} L_k^{-1}$. Nun betrachte die Zerlegung

$$A_{k+1} = \begin{pmatrix} A_k & b \\ c^\tau & \alpha_{k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} L_k & 0 \\ c^\tau U_k^{-1} & 1 \end{pmatrix}}_{=L_{k+1}} \underbrace{\begin{pmatrix} U_k & L_k^{-1} b \\ 0 & \alpha_{k+1} - c^\tau A_k^{-1} b \end{pmatrix}}_{=U_{k+1}}.$$

Hierin sind $b, c \in \mathbb{R}^k$ und $\alpha_{k+1} \in \mathbb{R} \setminus \{0\}$ wie oben.

Nun ist L_{k+1} eine untere Dreiecksmatrix wie in (4.2) und insbesondere regulär. Es gilt außerdem nach Voraussetzung, dass A_{k+1} regulär und somit $U_{k+1} = L_{k+1}^{-1} A_{k+1}$ das Produkt regulärer Matrizen - also ebenso regulär ist. Demnach folgt $a_{k+1} - c^\tau A_k^{-1} b \neq 0$ und U_{k+1} ist wie von (4.3) gefordert.

Demnach besitzt jede führende Hauptstreichungsmatrix A_k eine LU-Zerlegung und da $A_n = A$ auch die Matrix selbst.

Eindeutigkeit: Sei $A \in M_n$ mit den zwei LU-Zerlegungen $L_1 U_1 = A = L_2 U_2$, welche (4.2) und (4.3) genügen.
Da die Faktoren beider Zerlegungen von 0 verschiedene Hauptdiagonaleneinträge haben, sind die L_i, U_i regulär. Demnach kann man umschreiben: $L_2^{-1} L_1 = U_2 U_1^{-1}$.

Das Inverse einer oberen bzw. unteren Dreiecksmatrix ist wieder eine obere bzw. untere Dreiecksmatrix. Auch das Produkt zweier oberer bzw. unterer Dreiecksmatrizen ist wieder eine obere bzw. untere Dreiecksmatrix. Daher ist $U_2 U_1^{-1}$ eine obere und $L_2^{-1} L_1$ eine untere Dreiecksmatrix. Wegen Gleichheit muss $L_2^{-1} L_1 = D = U_2 U_1^{-1}$ für eine Diagonalmatrix $D = (d_{i,j}) = l_{i,i}^2 = u_{i,i}^2$ gelten. Nun sind die Diagonaleinträge $l_{i,i} = 1$, so dass $D = \mathbb{1}_n$ ist. Aus

$$L_2^{-1} L_1 = \mathbb{1}_n = U_2 U_1^{-1}$$

erhält man nun $L_1 = L_2$ und $U_1 = U_2$. Damit sind die beiden Zerlegungen gleich. $\square$

Die vorgestellte LU-Zerlegung weist eine gewisse Asymmetrie der Faktoren L und U auf der Hauptdiagonalen auf. Aus diesem Grund, definiert man $\tilde{U}$ analog zu (4.2) durch

$$\tilde{U} = (\tilde{u}_{i,j})_{i \in \{1,\dots,n\}, j \in \{1,\dots,n\}} = \begin{cases} \tilde{u}_{i,j} = \frac{u_{i,j}}{u_{i,i}} & \text{für } i < j \\ \tilde{u}_{i,j} = 1 & \text{für } i = j \\ \tilde{u}_{i,j} = 0 & \text{sonst} \end{cases}. \tag{4.4}$$

Mit dieser neuen Definition erhält man $U = D\tilde{U}$ für eine Diagonalmatrix

$$D = \operatorname{diag}(u_{1,1}, u_{2,2}, \dots, u_{n,n}). \tag{4.5}$$

Dies führt auf eine neue Zerlegung

$$A = LD\tilde{U}. \tag{4.6}$$

Entsprechen die Matrizen den in (4.2), (4.5) und (4.4) geforderten Bedingungen, so spricht man von der *LDU-Zerlegung* einer Matrix A. Diese ist (wie eben gezeigt) durch die drei

Bedingungen eindeutig bestimmt. Weiter nennt man die Einträge der Diagonalmatrix D auch *Pivotelemente*.

Betrachtet man den Fall symmetrischer Matrizen genauer, so ergibt sich für die LDU-Zerlegung:

$$LD\tilde{U} = A = A^\tau = \left(LD\tilde{U}\right)^\tau = \tilde{U}^\tau D^\tau L^\tau = \tilde{U}^\tau D L^\tau$$

und daraus folgt wegen Eindeutigkeit $\tilde{U} = L^\tau$. Dies bedeutet für die LDU-Zerlegung einer reell symmetrischen, invertierbaren Matrix $A = LD\tilde{U} = LDL^\tau$.

Daraus resultierend komme ich zur *Cholesky-Zerlegung* einer Matrix. Es sind alle Pivotelemente positiv, also existieren deren Wurzeln $\sqrt{u_{i,i}} \in \mathbb{R}$ und man kann die Diagonalmatrix D schreiben als $D = D^{1/2}D^{1/2}$. Setzt man nun $R := D^{1/2}L^\tau$, so hat man eine noch kürzere Zerlegung der Matrix A gefunden, die Cholesky-Zerlegung:

$$A = LD^{1/2}D^{1/2}L^\tau = \left(D^{1/2}L\right)^\tau \left(D^{1/2}L\right) = R^\tau R. \tag{4.7}$$

Lässt man dabei nur die positiven Wurzeln zu, so ergibt sich sogar ein eindeutiger *Cholesky-Faktor R*.

$$R = (r_{i,j})_{i\in\{1,\dots,n\},j\in\{1,\dots,n\}} = \begin{cases} r_{i,j} = \frac{u_{i,j}}{\sqrt{u_{i,i}}} & \text{für } i < j \\ r_{i,j} = \sqrt{u_{i,j}} & \text{für } i = j \\ r_{i,j} = 0 & \text{sonst} \end{cases} . \tag{4.8}$$

Als letzte Vorbereitung betrachte man das folgende Lemma 4.8, welches sicherstellen wird, dass ich die Cholesky-Zerlegung wiederfinde, wenn ich sie benötige. Die vorgestellte Zerlegung bezeichnet man als QR-Zerlegung.

Lemma 4.8. *Sei $A \in M_n$ eine reguläre Matrix. Dann gibt es in eindeutiger Art und Weise eine orthogonale Matrix $Q \in M_n$ und eine obere Dreiecksmatrix $R \in M_n$ mit positiven Hauptdiagonaleneinträgen, sodass gilt*

$$A = QR.$$

Beweis. Siehe Horn u. Johnson [4, S.112 f.]. □

Nach dem Umreißen der verwendeten Begriffe will ich mich nun auf die Hauptaussage des Hauptminorenkriteriums konzentrieren, welche sich in folgendem Satz niederschlägt.

Satz 4.9. *Sei $n \in \mathbb{N}$ und $A \in M_n(\mathbb{R})$ eine symmetrische, reguläre Matrix. Dann sind äquivalent:*

(i) A ist positiv definit, d.h. es gilt für alle $x \in \mathbb{R}^n$ mit $x \neq 0$: $x^\tau Ax > 0$.

(ii) Alle Eigenwerte von A sind positiv.

(iii) Es gibt eine invertierbare Matrix $B \in M_n$, sodass $A = B^\tau B$.

(iv) A besitzt einen eindeutigen Cholesky-Faktor R mit $A = R^\tau R$.

(v) A besitzt eine eindeutige LDU-Zerlegung mit positiven Pivotelementen.

(vi) Alle führenden Hauptminoren von A sind positiv.

(vii) Alle Hauptminoren von A sind positiv.

Beweis. [6] Sei $A \in M_n$ symmetrisch und regulär.

„(i)⇒(ii)“: Gilt $x^\tau Ax > 0$ für alle $x \in \mathbb{R}^n$ mit $x \neq 0$, so gilt dies insbesondere für die Eigenvektoren $y \in \mathbb{R}^n$ von A und es folgt mit der Eigenwertgleichung in Definition 2.6

$$y^\tau Ay = y^\tau \lambda y = \lambda y^\tau y = \lambda \underbrace{|x|^2}_{>0} > 0 \Rightarrow \lambda > 0.$$

„(ii)⇒(iii)“: Da A symmetrisch ist, gibt es eine orthogonale Matrix $O \in M_n$, sodass $A = O^\tau DO$. Dabei ist $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ die Diagonalmatrix mit den Eigenwerten von A auf der Hauptdiagonalen. Sind diese alle positiv, so existiert $D^{1/2}$ und es gilt mit $B = D^{1/2}O$:

$$A = O^\tau DO = O^\tau D^{1/2} D^{1/2} O = B^\tau B.$$

Da alle Eigenwerte von A positiv sind, sind auch deren Wurzeln positiv und damit auch $\det\left(D^{1/2}\right) > 0$. Nach dem Determinantenmultiplikationssatz ist dann auch $\det(B) \neq 0$ und somit B invertierbar.

„(iii)⇒(i)“: Ist $A = B^\tau B$ für eine reguläre Matrix B, so gilt in

$$x^\tau Ax = x^\tau B^\tau Bx = (Bx)^\tau Bx = |Bx|^2 \geq 0$$

die Gleichheit nur, falls $Bx = 0$ und da B invertierbar nur wenn $x = 0$. Demnach ist für alle $x \neq 0$ die Bedingung $x^\tau Ax > 0$ erfüllt.

„(iii)⇒(iv)“: Ist $A = B^\tau B$ für eine invertierbare Matrix $B \in M_n$, so kann man diese Matrix B nach Lemma 4.8 eindeutig in $B = QR$ umschreiben. Nun gilt

$$A = B^\tau B = (QR)^\tau QR = R^\tau \underbrace{Q^\tau Q}_{=\mathbb{1}_n} R = R^\tau R$$

für eine eindeutige obere Dreiecksmatrix R mit positiven Hauptdiagonaleneinträgen. R erfüllt somit die Anforderungen von (4.8) und ist der Cholesky-Faktor von A.

„(iv)⇒(v)“: Gilt $A = R^\tau R$ für eine obere Dreiecksmatrix R mit positiven Einträgen auf der Hauptdiagonalen, kann man diese auch multiplikativ umschreiben in eine obere Dreiecksmatrix $\tilde{U}$[7] mit 1 auf der Hauptdiagonalen und eine Diagonalmatrix D. Es gilt $R = D\tilde{U}$ und es wird

$$A = R^\tau R = \left(D\tilde{U}\right)^\tau D\tilde{U} = \tilde{U}^\tau D^\tau D\tilde{U} = \tilde{U}^\tau D^2 \tilde{U}.$$

Vergleicht man (4.5) mit (4.2) bekommt man $\tilde{U}^\tau = L$ und $A = LD^2\tilde{U}$ ist offensichtlich die LDU-Zerlegung von A. Die Pivotelemente in $D^2 = \operatorname{diag}\left(r_{1,1}^2, \ldots, r_{n,n}^2\right)$ sind allesamt Quadrate in $\mathbb{R}$ und deshalb positiv.

„(v)⇒(vi)“: Besitzt die Matrix A eine LDU-Zerlegung $A = LD\tilde{U}$, kann man sie durch Transformation von $\tilde{U}$[8] in U[9] in eine LU-Zerlegung $A = LU$ umwandeln. U hat dann strikt positive Hauptdiagonaleneinträge. Nun betrachte die führende $k \times k$-Hauptstreichungsmatrix von A.

$$A = LU = \begin{pmatrix} L_k & 0 \\ L_{2,1} & L_{2,2} \end{pmatrix} \begin{pmatrix} U_k & U_{1,2} \\ 0 & U_{2,2} \end{pmatrix} = \begin{pmatrix} L_k U_k & \star \\ \star & \star \end{pmatrix}$$

[6] Die Beweisidee ist dem Buch von Meyer [6, S. 558 f.] entnommen.
[7] Man vergleiche hier mit (4.5).
[8] siehe (4.5)
[9] siehe (4.3)

Hierin sind $L_k, U_k \in M_k$ wie in (4.2) bzw. (4.3) gefordert. Die weiteren Blockmatrizen sind entsprechend angepasst. Nun kann man A_k schreiben als

$$A_k = L_k U_k = \begin{pmatrix} L_{k-1} & 0 \\ d^\tau & 1 \end{pmatrix} \begin{pmatrix} U_{k-1} & c \\ 0 & u_{k,k} \end{pmatrix} = \begin{pmatrix} A_{k-1} & \star \\ \star & \star \end{pmatrix}$$

und es gilt auch $A_{k-1} = L_{k-1}U_{k-1}$, sowie $\det(L_k) = \det(L_{k-1}) = 1$. Laplace-Entwicklung der Determinante nach der letzten Zeile bringt $\det(U_k) = \det(U_{k-1})\, u_{k,k}$. Nun ergibt der Determinantenmultiplikationssatz

$$\det(A_k) = \det(L_k)\det(U_k) = 1 \cdot \det(U_{k-1})\, u_{k,k} = \det(A_{k-1})\, u_{k,k}.$$

Der k-te Hauptdiagonaleneintrag von U wird also durch den Quotienten zweier Hauptminoren, nämlich

$$u_{k,k} = \frac{\det(A_k)}{\det(A_{k-1})} > 0$$

gegeben. Nun beweise per vollständiger Induktion nach k, dass $\det(A_k) > 0$.

Induktionsanfang: Sei $k = 2$. Es ist $\det(A_1) = \det(U_1) = u_{1,1} = 1$ wegen (4.3) und somit folgt mit der Voraussetzung $u_{2,2} > 0$, dass $\det(A_2) > 0$.

Induktionsvoraussetzung: Es ist $\det(A_k) > 0$ für ein $k < n$.

Induktionsschritt: Es gilt nach Voraussetzung $u_{k+1,k+1} = r > 0$. Dann ist

$$\frac{\det(A_{k+1})}{\det(A_k)} = r > 0 \Leftrightarrow \det(A_{k+1}) = \underbrace{\det(A_k)}_{\overset{\text{I.V.}}{>}\, 0} r > 0.$$

Somit ist jeder führende Hauptminor der Matrix A positiv.

„(vi)$\Rightarrow$(vii)“: Man kann jede Hauptstreichungsmatrix einer Matrix A durch eine orthogonale Permutationsmatrix P vermittels Zeilen- bzw. Spaltenvertauschung in eine führende Hauptstreichungsmatrix überführen. Ist $A(\alpha)$ mit $\alpha \subseteq \{1, \dots, n\}$ eine (nichtführende) Hauptstreichungsmatrix von A, so gibt es eine orthogonale Matrix $P \in M_n$, sodass

$$B = P^\tau A P = \begin{pmatrix} A(\alpha) & \star \\ \star & \star \end{pmatrix}.$$

$A(\alpha)$ ist hier ein führender Hauptminor der Matrix B.

Ist A positiv definit, so ist auch B positiv definit, denn es gilt für alle $x \in \mathbb{R}^n, x \neq 0$ und $Px = y \in \mathbb{R}^n, y \neq 0$

$$x^\tau B x = x^\tau P^\tau A P x = (Px)^\tau A (Px) = y^\tau A y > 0.$$

Aus der positiven Definitheit von B folgt dann $\det(A(\alpha)) > 0$. Da $A(\alpha)$ jedoch beliebig gewählt war, ist somit jeder Hauptminor von A positiv.

„(vii)$\Rightarrow$(ii)“: Sind alle Hauptminoren der Matrix A positiv, so sind es insbesondere die Summen der Hauptminoren gleicher Ordnung, da es sich hierbei nur um Addition reeller Zahlen handelt. Nach Korollar 3.9 sind die Koeffizienten des charakteristischen Polynoms bis auf das Vorzeichen eindeutig durch die Summen der Hauptminoren gleicher Ordnung festgelegt.

Ich führe nun einen Widerspruchsbeweis. Angenommen es gibt einen Eigenwert $\lambda \leq 0$ der Matrix A. Dann gilt für das charakteristische Polynom mit geradem Grad n:

$$\begin{aligned}\chi_A(\lambda) &= \underbrace{\lambda^n}_{\geq 0} - \underbrace{\lambda^{n-1}}_{\leq 0} \underbrace{\sum_{\substack{\gamma \subset \{1,\ldots,n\} \\ |\gamma|=n-1}} \det\left(A_x\left(\emptyset, \gamma\right)\right)}_{\overset{\text{Vor.}}{>0}} + \underbrace{\lambda^{n-2}}_{\geq 0} \underbrace{\sum_{\substack{\gamma \subset \{1,\ldots,n\} \\ |\gamma|=n-2}} \det\left(A_x\left(\emptyset, \gamma\right)\right)}_{\overset{\text{Vor.}}{>0}} \mp \ldots \\ &\quad \cdots + \underbrace{\det\left(A_x\left(\emptyset, \emptyset\right)\right)}_{\overset{\text{Vor.}}{>0}} \\ &> 0.\end{aligned}$$

Analog folgt $\chi_A(x) < 0$ für ein Polynom ungeraden Grades.

Dies ist jeweils ein Widerspruch zur Eigenwerteigenschaft $\chi_A(\lambda) = 0$ von λ. Es muss also $\lambda > 0$ für alle Eigenwerte von A gelten.

Damit ist der Ringschluss für alle Aussagen vollendet und die behauptete Äquivalenz gezeigt. □

Das Bemerkenswerte an diesem Hauptminorenkriterium „(vi)⇔(i)“ ist, dass es sich vollständig auf die führenden Hauptminoren einer Matrix verlassen kann. Diese zu berechnen ist wesentlich weniger aufwendig, denn es gibt maximal n zu berechnende Hauptminoren, während das erweiterte Hauptminorenkriterium der Semidefinitheit die Kenntnis aller 2^n Hauptminoren verlangt.

5 Ausblick

Alle in dieser Arbeit getroffenen Aussagen über reell-symmetrische Matrizen sind völlig analog auch auf komplex-hermitesche Matrizen übertragbar. Diese Klasse von Matrizen ist zwar auf dem Körper der komplexen Zahlen $\mathbb{C}$ definiert, besitzt aber lediglich reelle Eigenwerte.

Nur für diese Klassen ist es sinnvoll, Begriffe wie (Semi-)Definitheit einzuführen und die Nichtnegativität von Eigenwerten zu betrachten, da bei der Körpererweiterung von $\mathbb{R}$ auf $\mathbb{C}$ die eindeutige Anordenbarkeit der Zahlen verloren geht. Im Allgemeinen – für $x \in \mathbb{C}^n$ und $A \in M_n(\mathbb{C})$ – wäre schließlich $x^\star Ax = (\bar{x})^\tau Ax = z \in \mathbb{C}$, was nicht eindeutig als positiv oder negativ identifiziert werden kann.

Manchmal behilft man sich dann mit der Forderung, dass $\Re(x^\tau Ax) > 0$ sei. Diese Definition wirkt dann auch auf nicht hermitesche Matrizen und ist dadurch allgemeiner, jedoch nicht so stark wie die hier verwendete.

Die in Kapitel 3 gemachten Beobachtungen zum charakteristischen Polynom sind jedoch in allgemeinerer Form auf komplexe Matrizen übertragbar. Das charakteristische Polynom einer Matrix $A \in M_n(\mathbb{C})$ wäre selbst ein komplexes Polynom $\chi_A(x) \in \mathbb{C}[x]$. Der Fundamentalsatz der Algebra[1] garantiert dann eine Anzahl von Nullstellen, die dem Grad des Polynoms entspricht.

Es zerfällt also jedes charakteristische Polynom $\chi_A(x) \in \mathbb{C}[x]$ vollständig in Linearfaktoren, sodass die Identität der Summe der Hauptminoren der Ordnung k mit dem k-ten elementarsymmetrischen Polynom in den Eigenwerten (Korollar 3.14) für beliebige komplexe Matrizen $A \in M_n(\mathbb{C})$ stimmt.

Richtet man den Blick nicht auf Erweiterungen, sondern auf Teilmengen von $\mathbb{R}$, so sieht man, dass die Bedingung der Anordenbarkeit für alle nichtendlichen Teilkörper von $\mathbb{R}$ erhalten wird. Endliche Teilkörper sind dagegen nicht anordenbar[2] und eignen sich somit nicht als Grundlage für einen Matrizenring, dessen Elemente Definitheit aufweisen sollen.

Die Körpereigenschaften aufzugeben (z.B. bei Zugrundelegung eines Ringes), führt dagegen zu fundamentalen Problemen, da dann auch die (multiplikative) Invertierbarkeit der Matrizen im Allgemeinen nicht mehr gegeben ist. Dies führt sofort zu Schwierigkeiten zum Beispiel bei der Diagonalisierbarkeit und der Identifikation von Eigenwerten mit dem charakteristischen Polynom (2.1).

Allgemein ist also festzuhalten, dass die Betrachtungen zur Definitheit von Matrizen vor allem auf formal reellen Körpern wegen ihrer Anordenbarkeit durchführbar sind. Dass auf solchen nicht zwangsläufig[3] Nullstellen für Polynome existieren und so die Existenz und Ermittlung von Eigenwerten unsicher ist, bleibt dann ein Spannungsfeld, dem man sich aussetzen muss.

Bekannte Anwendungen von Matrizen und ihren Eigenwerten finden sich in der Physik und dort insbesondere in der Quantenmechanik. Durch Anwendung einer linearen Abbildung

[1] Die genaue Formulierung mit Beweis kann bei Remmert [8] nachgelesen werden.
[2] Eine Betrachtung dessen findet sich in Prieß-Crampe [7].
[3] Nämlich sicher nur für Polynome ungeraden Grads.

innerhalb des Hilbertraumes der quadratintegrierbaren Funktionen L^2, entstehen Matrizen, welche dort als Operatoren bezeichnet werden.

Jedes Element $\psi \in L^2$ stellt in dieser Theorie den Zustand eines Systems (i.d.R. endlich viele Teilchen) dar. Die Anwendung eines solchen Operators ist in dieser Deutung eine Messung. Die Messergebnisse entsprechen den Eigenwerten des Operators, da sie $\hat{O}\psi = \lambda\psi$ erfüllen. Beobachtbare Größen wie z.B. Energie, Impuls, Drehimpuls (sog. „Observablen“) können demnach nur durch reelle Eigenwerte repräsentiert werden.

Diese Bedeutung der Eigenwerte macht die Klassen der reell-symmetrischen Matrizen und der komplex-hermiteschen Matrizen so wichtig, da sie „realen“ Messungen der quantisierten Umwelt entsprechen können. In diesem Kontext ist es besonders interessant, über das Kriterium der Definitheit eines Operators generelle Aussagen über die Natur der Eigenwerte und damit der erwartbaren Messergebnisse treffen zu können.

6 Literaturverzeichnis

[1] Artin, Emil: Über die Zerlegung definiter Funktionen in Quadrate. In: *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 5 (1927), S. 100–115

[2] Fischer, Gerd: *Lineare Algebra*. 9. Aufl. Friedrich Vieweg & Sohn, 1989 (1975). – ISBN 3–528–57217–5

[3] Hilbert, David: Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-Physikalische Klasse* 3 (1900), S. 253–297

[4] Horn, Roger A. ; Johnson, Charles R.: *Matrix Analysis*. 21. Aufl. Cambridge University Press, 2007 (1985). – ISBN 978–0–521–38632–6

[5] Knabner, Peter ; Barth, Wolf: *Lineare Algebra – Grundlagen und Anwendungen*. 1. Aufl. Springer-Verlag, 2013. – ISBN 978–3–642–32185–6

[6] Meyer, Carl D.: *Matrix Analysis and Applied Linear Algebra*. 1. Aufl. Siam, 2000. – ISBN 978–0–898714–54–8

[7] *Kapitel* II: Angeordnete Additionen und Multiplikationen. In: Priess-Crampe, Sibylla: *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Bd. 98: *Angeordnete Strukturen – Gruppen, Körper, projektive Ebenen*. 1. Aufl. Springer-Verlag, 1983. – ISBN 3–540–11646–X, S. 19–64

[8] Remmert, Reinhold: Fundamentalsatz der Algebra. In: Ebbinghaus, Heinz-Dieter (Hrsg.) ; Hermes, Hans (Hrsg.) ; Hirzebruch, Friedrich (Hrsg.) ; Koecher, Max (Hrsg.) ; Lamotke, Klaus (Hrsg.) ; Mainzer, Klaus (Hrsg.) ; Neukirch, Jürgen (Hrsg.) ; Prestel, Alexander (Hrsg.) ; Remmert, Reinhold (Hrsg.): *Zahlen*. Springer-Verlag, 1992 (Springer-Lehrbuch). – ISBN 978–3–540–55654–1, Kapitel 5, S. 79–99

[9] *Kapitel* VII: Kommutative Algebra. In: Scheja, Günther ; Storch, Uwe: *Lehrbuch der Algebra – Unter Einschluss der linearen Algebra*. Bd. 2: *Lehrbuch der Algebra – Unter Einschluss der linearen Algebra, Teil 2*. 1. Aufl. Vieweg+Teubner Verlag, 1988. – ISBN 978–3–519–02212–1, S. 7–256

Selbstständigkeitserklärung

Ich versichere, dass ich diese Arbeit selbstständig und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt habe. Alle Stellen der Arbeit, die ich aus diesen Quellen und Hilfsmitteln dem Wortlaut oder dem Sinne nach entnommen habe, sind kenntlich gemacht und im Literaturverzeichnis aufgeführt.

Lukas Hellwig
Potsdam, den 26.04.2018