“Lineare Algebra und analytische Geometrie II”

Meinolf Geck, Universität Stuttgart, SoSe 2019

Literaturverzeichnis

[An] T. ANDREESCU, Essential linear algebra with applications. A problem-solving approach, Birkhäuser/Springer, New York, 2014.

[Ar] M. ARTIN, Algebra. Aus dem Englischen übersetzt von Annette A'Campo. Birkhäuser Verlag, 1993.

[Ax] S. AXLER, Linear Algebra done right. Undergraduate texts in mathematics, Springer-Verlag, 2015.

[Bo] N. BOURBAKI, Éléments de Mathématiques. Algèbre. Chap. 1 à 3, Masson, Paris, 1970; Chap. 4 à 7, Masson, Paris, 1981.

[Fi] G. FISCHER, Lineare Algebra: Eine Einführung für Studienanfänger, Vieweg + Teubner Verlag; 17. Auflage 2010.

[FIS] S. H. FRIEDBERG, A. J. INSEL UND L. E. SPENCE, Linear Algebra, 4th ed., Pearson, 2002.

[GAP] THE GAP GROUP, GAP - Groups, Algorithms, and Programming, Version 4.10.0, 2018. Frei verfügbares Computer-Algebra-System, siehe `http://www.gap-system.org`.

[HW] B. HUPPERT UND W. WILLEMS, Lineare Algebra: Mit zahlreichen Anwendungen in Kryptographie, Codierungstheorie, Mathematischer Physik und Stochastischen Prozessen, Vieweg + Teubner Verlag, 2. Auflage 2010.

[Ko] M. KOECHER, Lineare Algebra und analytische Geometrie, Grundwissen Mathematik, Springer-Verlag, 4. Auflage, 2002.

[Po] D. POOLE, Linear Algebra: A Modern Introduction. Brooks Cole Pub Co., 3. Auflage, 2010.

[Pr] V. V. PRASOLOV, Problems and Theorems in Linear Algebra. Translations of Mathematical Monographs, Amer. Math. Soc., 1994.

[SAGE] SAGEMATH, version 8.6, 2018. Frei verfügbares Computer-Algebra-System, siehe `https://www.sagemath.org/`.

[Se] D. SERRE, Matrices: Theory and Applications. Graduate Texts in Mathematics 216, Springer-Verlag, 2. Auflage, 2010.

KAPITEL 4

Determinanten (6 Doppelstunden)

Ziel: Ordne einer Matrix $A \in M_n(K)$ ein Element $\det(A) \in K$ zu. Es soll zum Beispiel gelten: $\det(A) \neq 0 \Leftrightarrow A$ invertierbar. Kann man in diesem Fall eine allgemeine Formel für die Einträge von A^{-1} finden? Für $n = 2$ haben wir dies bereits in Kapitel 1, §5, gemacht:

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) := ad - bc, \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \text{falls } ad - bc \neq 0.$$

Für beliebiges $n \geqslant 2$ benötigen wir zunächst einige Vorbereitungen über Permutationen.

4.1. *Permutationen und die Leibniz-Formel*

Sei X eine nicht-leere Menge. Eine bijektive Abbildung $\pi\colon X \to X$ heißt dann auch eine ***Permutation*** von X. Sei

$$S_X := \{\pi\colon X \to X \mid \ \pi \text{ Permutation}\}.$$

Dann ist S_X eine Gruppe mit der Hintereinanderausführung "$\circ$" als Verknüpfung, genannt die ***symmetrische Gruppe*** auf X. (Die Assoziativität von $\circ$ wird in Aufgabe 0.3 gezeigt.) Das neutrale Element ist die identische Abbildung id, die jedes $x \in X$ auf sich abbildet. Das Inverse von $\pi \in S_X$ ist die Umkehrabbildung $\pi^{-1}\colon X \to X$.

Beispiel 4.1.1. Sei X eine endliche Menge mit $|X| = n \in \mathbb{N}$. Dann ist auch S_X endlich und es gilt $|S_X| = n!$ (Fakultät von n), siehe Kapitel 0, Lemma 5.17. Sei nun speziell $X = \{1, 2, \ldots, n\}$. Dann schreiben wir kurz $S_n := S_X$; diese Gruppe heißt symmetrische Gruppe vom Grad n. Die Elemente $\pi \in S_n$ schreiben wir in der Form

$$\pi = \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ \pi(1) & \pi(2) & \ldots & \pi(n-1) & \pi(n) \end{pmatrix}.$$

Für $n = 1$ gilt $S_1 = \{\text{id}\}$; für $n = 2$ gilt $S_2 = \{\text{id}, \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)\}$. Beispiele für $n = 3$:

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Also $\pi(1) = 3$, $\pi(2) = 1$, $\pi(3) = 2$ und $\sigma(1) = 2$, $\sigma(2) = 1$, $\sigma(3) = 3$. Damit erhalten wir

$$\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma \circ \pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \qquad \pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Wir sehen also hier $\sigma \circ \pi \neq \pi \circ \sigma$, d.h., S_3 ist nicht-abelsch.

Im Folgenden fassen wir S_n als eine Teilmenge von S_{n+1} auf: jede Permutation $\pi \in S_n$ können wir auch als eine Permutation $\pi \in S_{n+1}$ betrachten, indem wir $\pi(n+1) := n+1$ setzen. Wegen $S_3 \subseteq S_n$ für alle $n \geqslant 3$ folgt dann, dass S_n nicht abelsch ist für alle $n \geqslant 3$.

Einige spezielle Permutationen in S_n für $n \geqslant 2$:
Seien $1 \leqslant i, j \leqslant n$ mit $i \neq j$. Dann definiere $\tau_{ij} \in S_n$ durch $\tau_{ij}(k) := \begin{cases} j & \text{falls } k = i, \\ i & \text{falls } k = j, \\ k & \text{falls } k \neq i, j. \end{cases}$

D.h., τ_{ij} vertauscht die Ziffern i und j, lässt aber alle anderen Ziffern fest. Eine solche Permutation heißt ***Transposition***. Es gilt $\tau_{ij} \circ \tau_{ij} = \mathrm{id}$ und $\tau_{ij}^{-1} = \tau_{ij}$; außerdem $\tau_{ij} = \tau_{ji}$.

SATZ 4.1.2. *Jedes $\pi \in S_n$ lässt sich als Produkt von höchstens $n-1$ Transpositionen schreiben.*

BEWEIS. Vollständige Induktion nach n. Ist $n = 1$, so ist $\pi = \mathrm{id}$, also Produkt von 0 Transpositionen. Sei nun $n \geqslant 2$ und die Aussage bereits für alle Permutationen in S_{n-1} gezeigt. Ist $\pi(n) = n$, so können wir π als Permutation in S_{n-1} auffassen und die Aussage gilt nach Induktion. Sei nun $k := \pi(n) < n$ mit $1 \leqslant k \leqslant n-1$. Betrachte die Transposition $\tau_{kn} \in S_n$ und setze $\pi' := \tau_{kn} \circ \pi \in S_n$. Dann gilt $\pi'(n) = (\tau_{kn} \circ \pi)(n) = \tau_{kn}\big(\pi(n)\big) = \tau_{kn}(k) = n$, also $\pi' \in S_{n-1}$. Nach Induktion können wir π' als Produkt von $n-2$ Transpositionen schreiben, also ist $\pi = \tau_{kn}^{-1} \circ \pi' = \tau_{kn} \circ \pi'$ Produkt von höchstens $n-1$ Transpositionen. $\square$

Sei zum Beispiel $n = 4$ und $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$. Es gilt $\pi(4) = 2$, also betrachte $\pi' := \tau_{24} \circ \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in S_3$. Es gilt $\pi'(3) = 1$, also betrachte anschließend $\pi'' := \tau_{13} \circ \pi' = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \tau_{12} \in S_2$. Damit folgt $\pi = \tau_{24} \circ \tau_{13} \circ \tau_{12}$.

DEFINITION 4.1.3. Sei $\pi \in S_n$. Dann heißt

$$N(\pi) := \big\{(i, j) \in \mathbb{N} \times \mathbb{N} \mid 1 \leqslant i < j \leqslant n \text{ und } \pi(i) > \pi(j)\big\}$$

die Menge der ***Fehlstände*** von π und $\mathrm{sgn}(\pi) := (-1)^{|N(\pi)|} \in \mathbb{Z}$ das ***Signum*** von π. Es gilt also $\mathrm{sgn}(\pi) = 1$ falls π eine gerade Anzahl von Fehlständen hat, und $\mathrm{sgn}(\pi) = -1$ falls π eine ungerade Anzahl von Fehlständen hat; entsprechend nennen wir π auch eine ***gerade Permutation*** bzw. eine ***ungerade Permutation***. Seien

$$A_n := \{\pi \in S_n \mid \mathrm{sgn}(\pi) = 1\} \qquad \text{und} \qquad A_n' := \{\pi \in S_n \mid \mathrm{sgn}(\pi) = -1\}.$$

Dann ist $S_n = A_n \cup A_n'$ (disjunkte Vereinigung). Es gilt stets $N(\mathrm{id}) = \varnothing$, also $\mathrm{id} \in A_n$.

Im obigen Beispiel $n = 4$ ist $N(\pi) = \{(1,2),(1,3),(1,4),(2,3),(2,4)\}$ also $\operatorname{sgn}(\pi) = (-1)^5 = -1$, d.h., π ist eine ungerade Permutation. Außerdem gilt

$$\operatorname{sgn}(\tau_{12}) = -1 \qquad \text{für } \tau_{12} \in S_n \text{ und } n \geqslant 2.$$

Denn sei $1 \leqslant i < j \leqslant n$; ist $j \geqslant 3$, so $\tau_{12}(j) = j$ und $\tau_{12}(1) = 2$, $\tau_{12}(2) = 1$, $\tau_{12}(i) = i$ für $i \geqslant 3$, in jedem Fall also $\tau_{12}(i) < j = \tau_{12}(j)$ und damit $(i,j) \notin N(\tau_{12})$. Dann bleibt nur noch $i = 1$, $j = 2$ übrig, und hier gilt $\tau_{12}(1) = 2 > 1 = \tau_{12}(2)$; also folgt $N(\tau_{12}) = \{(1,2)\}$.

Satz 4.1.4. *Für alle* $\pi, \sigma \in S_n$ *gilt* $\operatorname{sgn}(\pi \circ \sigma) = \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)$ *und* $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$. *Für jede Transposition* $\tau_{ij} \in S_n$ *gilt* $\operatorname{sgn}(\tau_{ij}) = -1$. *Lässt sich* π *als Produkt von* r *Transpositionen schreiben (wobei* $0 \leqslant r \leqslant n$*, siehe Satz 4.1.2), so gilt* $\operatorname{sgn}(\pi) = (-1)^r$.

Beweis. Wie in Kapitel 0, Satz 5.15 (Formel von Pascal) sei $T(n,2)$ die Menge der Teilmengen von $\{1,2,\ldots,n\}$ mit genau 2 Elementen; es gilt also $|T(n,2)| = \binom{n}{2} = \frac{1}{2}n(n-1)$. Ist $S \in T(n,2)$, so schreibe $S = \{i,j\}$ mit $i < j$ und setze $\pi(S) := \{\pi(i),\pi(j)\}$. Weil π injektiv ist, folgt $\pi(S) \in T(n,2)$; es gilt also entweder $\pi(i) < \pi(j)$ oder $\pi(i) > \pi(j)$. Dann definere

$$\varepsilon_\pi(S) := \begin{cases} 1 & \text{falls } \pi(i) < \pi(j), \\ -1 & \text{falls } \pi(i) > \pi(j). \end{cases}$$

Also gilt $\varepsilon_\pi(S) = -1$ genau dann, wenn $(i,j) \in N(\pi)$. Damit folgt $\operatorname{sgn}(\pi) = \prod_{S \in T(n,2)} \varepsilon_\pi(S) = \pm 1$. Nun betrachte die folgenden vier möglichen Fälle für $S = \{i,j\}$ mit $i < j$:

Fall	$\varepsilon_\sigma(S)$	$\varepsilon_\pi(\sigma(S))$	$\varepsilon_{\sigma\circ\pi}(S)$
$\sigma(i) < \sigma(j),\ \pi(\sigma(i)) < \pi(\sigma(j))$	1	1	1
$\sigma(i) < \sigma(j),\ \pi(\sigma(i)) > \pi(\sigma(j))$	1	-1	-1
$\sigma(i) > \sigma(j),\ \pi(\sigma(i)) < \pi(\sigma(j))$	-1	-1	1
$\sigma(i) > \sigma(j),\ \pi(\sigma(i)) > \pi(\sigma(j))$	-1	1	-1

In allen vier Fällen gilt $\varepsilon_{\pi\circ\sigma}(S) = \varepsilon_\sigma(S)\varepsilon_\pi(\sigma(S))$. Damit folgt

$$\operatorname{sgn}(\pi \circ \sigma) = \prod_{S \in T(n,2)} \varepsilon_{\pi\circ\sigma}(S) = \Big(\prod_{S \in T(n,2)} \varepsilon_\sigma(S)\Big)\Big(\prod_{S \in T(n,2)} \varepsilon_\pi(\sigma(S))\Big).$$

Nun beachte: Weil σ injektiv ist, folgt nicht nur $\sigma(S) \in T(n,2)$ sondern auch, dass die Abbildung $T(n,2) \to T(n,2)$, $S \mapsto \sigma(S)$, bjiektiv ist. (Die Umkehrabbildung ist gegeben durch $S \mapsto \sigma^{-1}(S)$.) Also folgt $\prod_{S \in T(n,2)} \varepsilon_\pi(\sigma(S)) = \prod_{S \in T(n,2)} \varepsilon_\pi(S) = \operatorname{sgn}(\pi)$ und damit schließlich $\operatorname{sgn}(\pi \circ \sigma) = \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)$. Damit erhalten wir auch $1 = \operatorname{sgn}(\mathrm{id}) = \operatorname{sgn}(\pi \circ \pi^{-1}) = \operatorname{sgn}(\pi)\operatorname{sgn}(\pi^{-1})$, also $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$ (weil beide Werte gleich ± 1 sind).

Nun zu Transpositionen. Wir haben oben gesehen, dass $\operatorname{sgn}(\tau_{12}) = -1$ gilt. Sei nun $n \geqslant 3$ und $1 < i < j \leqslant n$. Dann rechnet man sofort nach: $\quad \tau_{2j} \circ \tau_{12} \circ \tau_{2j} = \tau_{1j}$ und $\tau_{1i} \circ \tau_{1j} \circ \tau_{1i} = \tau_{ij}$.

Mit der gerade bewiesenen Regel für das Signum von Produkten von Permutationen folgt $\mathrm{sgn}(\tau_{1j}) = \mathrm{sgn}(\tau_{12})\mathrm{sgn}(\tau_{2j})^2 = \mathrm{sgn}(\tau_{12}) = -1$ und $\mathrm{sgn}(\tau_{ij}) = \mathrm{sgn}(\tau_{1j})\mathrm{sgn}(\tau_{1i})^2 = \mathrm{sgn}(\tau_{1j}) = -1$. Damit erhalten wir auch sofort die Formel für das Produkt von r Transpositionen. $\square$

FOLGERUNG 4.1.5. *Sei* $n \geqslant 2$. *Dann gilt* $|S_n| = 2|A_n|$, *d.h., je eine Hälfte der Permutationen in* S_n *ist gerade und die andere Hälfte ist ungerade. Ist* $\tau \in S_n$ *eine beliebige Permutation mit* $\mathrm{sgn}(\tau) = -1$ *(zum Beispiel eine Transposition), so ist* $A_n \to A_n'$, $\pi \mapsto \pi \circ \tau$, *eine bijektive Abbildung, also* $S_n = A_n \cup \{\pi \circ \tau \mid \pi \in A_n\}$.

BEWEIS. Wie bereits bemerkt gilt $S_n = A_n \cup A_n'$ und $A_n \cap A_n' = \varnothing$; insbesondere also $|S_n| = |A_n| + |A_n'|$. Sei nun $\tau \in S_n$ fest mit $\mathrm{sgn}(\tau) = -1$. Für $\pi \in A_n$ gilt dann $\mathrm{sgn}(\pi \circ \tau) = \mathrm{sgn}(\pi)\mathrm{sgn}(\tau) = -1$, also erhalten wir eine Abbildung $f\colon A_n \to A_n'$, $\pi \mapsto \pi \circ \tau$. Umgekehrt: Für $\sigma \in A_n'$ ist $\mathrm{sgn}(\sigma \circ \tau^{-1}) = \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau^{-1}) = \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) = (-1)(-1) = 1$, also erhalten wir auch eine Abbildung $g\colon A_n' \to A_n$, $\sigma \mapsto \sigma \circ \tau^{-1}$. Offenbar ist $f \circ g = \mathrm{id}_{A_n'}$ und $g \circ f = \mathrm{id}_{A_n}$, also sind f, g bijektiv und $g = f^{-1}$. Damit folgt $|A_n| = |A_n'|$ und $|S_n| = |A_n| + |A_n'| = 2|A_n|$. $\square$

DEFINITION 4.1.6 (Leibniz-Formel). Für $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in M_n(R)$ definieren wir

$$\det(A) := \sum_{\pi \in S_n} \mathrm{sgn}(\pi) a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} \in R \qquad \text{“\textbf{\textit{Determinante}} von } A\text{”}.$$

Ist $A = [a] \in M_1(R)$ mit $a \in R$, so gilt $\det(A) = a$. Ist $n = 2$, so ist $S_2 = \{\mathrm{id}, \tau_{12}\}$ und

$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11}a_{22} + \mathrm{sgn}(\tau_{12}) a_{1\tau_{12}(1)} a_{2\tau_{12}(2)} = a_{11}a_{22} - a_{12}a_{21}$$

wie in Kapitel 1, §4, definiert. Im Allgemeinen erhalten wir einen Ausdruck mit $n!$ Summanden $a_{1\pi(1)} \cdots a_{n\pi(n)}$. (Für grosse Werte von n ist diese Definition also praktisch unbrauchbar !)

LEMMA 4.1.7. *Sei* $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in M_n(R)$ *eine untere Dreiecksmatrix, d.h.,* $a_{ij} = 0$ *für* $1 \leqslant i < j \leqslant n$. *Dann gilt* $\det(A) = a_{11}a_{22} \cdots a_{nn}$. *Inbesondere ist* $\det(I_n) = 1$.

BEWEIS. Sei $\pi \in S_n$ so, dass der entsprechende Summand $\mathrm{sgn}(\pi) a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}$ in der Formel für $\det(A)$ ungleich 0 ist. Dann ist $a_{k\pi(k)} \neq 0$ für alle k, also muss gelten

$$\pi(1) \leqslant 1, \qquad \pi(2) \leqslant 2, \qquad \ldots, \qquad \pi(n) \leqslant n.$$

Aus der ersten Bedingung folgt $\pi(1) = 1$, dann aus der zweiten $\pi(2) = 2$ und so weiter bis $\pi(n) = n$. Damit liefert nur der Term für $\pi = \mathrm{id}$ einen Beitrag ungleich 0 zu $\det(A)$, also gilt die genannte Formel. Angewandt auf $A = I_n$ erhalten wir $\det(I_n) = 1$. $\square$

LEMMA 4.1.8. *Es gilt* $\det(A) = \det(A^{\mathrm{tr}})$ *für alle* $A \in M_n(R)$. *Damit gilt eine analoge Aussage wie in Lemma 4.1.7 auch für obere Dreiecksmatrizen.*

Beweis. Sei $A = [a_{ij}]_{1\leqslant i,j\leqslant n}$; dann ist $A^{tr} = [a_{ji}]_{1\leqslant i,j\leqslant n}$, also gilt

$$\det(A^{tr}) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{\pi(1)1} a_{\pi(2)2}\cdots a_{\pi(n)n} = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{k=1}^{n} a_{\pi(k)k}.$$

Sei $\pi \in S_n$ fest und setze $l := \pi(k)$ für $1 \leqslant k \leqslant n$. Mit k durchläuft auch l alle Ziffern von 1 bis n, nur in einer anderen Reihenfolge. Mit $k := \pi^{-1}(l)$ erhalten wir also

$$\prod_{k=1}^{n} a_{\pi(k)k} = \prod_{l=1}^{n} a_{l\pi^{-1}(l)}$$

und damit $\det(A^{tr}) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{1\pi^{-1}(1)} a_{2\pi^{-1}(2)} \cdots a_{n\pi^{-1}(n)}$. Da S_n eine Gruppe ist, ist die Abbildung $S_n \to S_n$, $\pi \mapsto \pi^{-1}$ bijektiv. Also können wir in der letzteren Formel für $\det(A^{tr})$ auch überall π^{-1} durch π ersetzen und sehen damit $\det(A^{tr}) = \det(A)$. □

Die Aussagen in Folgerung 4.1.5 haben folgende Konsequenz für Determinanten.

Lemma 4.1.9. *Sei $A \in M_n(R)$. Sind zwei Zeilen von A gleich, so gilt* $\det(A) = 0$.

Beweis. Sei $A = [a_{ij}]_{1\leqslant i,j\leqslant n} \in M_n(R)$. Seien $1 \leqslant k < l \leqslant n$ so, dass die k-te und l-te Zeile von A gleich sind, also $a_{kj} = a_{lj}$ für $1 \leqslant j \leqslant n$. Betrachte nun die Transposition $\tau_{kl} \in S_n$. Nach Folgerung 4.1.5 ist $S_n = A_n \cup \{\pi \circ \tau_{kl} \mid \pi \in A_n\}$. Damit erhalten wir

$$\begin{aligned}
\det(A) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i\pi(i)} = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{k\pi(k)} a_{l\pi(l)} \prod_{1\leqslant i\leqslant n, i\neq k,l} a_{i\pi(i)} \\
&= \sum_{\pi\in A_n} \underbrace{\operatorname{sgn}(\pi)}_{=1} a_{k\pi(k)} a_{l\pi(l)} \prod_{1\leqslant i\leqslant n, i\neq k,l} a_{i\pi(i)} \\
&\quad + \sum_{\pi\in A_n} \underbrace{\operatorname{sgn}(\pi\circ\tau_{kl})}_{=-\operatorname{sgn}(\pi)=-1} \underbrace{a_{k,(\pi\circ\tau_{kl})(k)}}_{=a_{k\pi(l)}} \underbrace{a_{l,(\pi\circ\tau_{kl})(l)}}_{=a_{l\pi(k)}} \prod_{1\leqslant i\leqslant n, i\neq k,l} \underbrace{a_{i,(\pi\circ\tau_{kl})(i)}}_{=a_{i\pi(i)}}.
\end{aligned}$$

Nach Voraussetzung ist $a_{k\pi(l)} = a_{l\pi(l)}$ und $a_{l\pi(k)} = a_{k\pi(k)}$. Also ist die zweite Summe gleich

$$\sum_{\pi\in A_n} (-1) a_{l\pi(l)} a_{k\pi(k)} \prod_{1\leqslant i\leqslant n, i\neq k,l} a_{i\pi(i)},$$

d.h., genau gleich dem Negativen der ersten Summe. Also ist $\det(A) = 0$. □

Bemerkung 4.1.10. In LAAG1, Aufgabe 8.6 wurde folgende geometrische Interpretation der Determinante im Fall $n = 2$ behandelt. Gegeben seien zwei Punkte im $\mathbb{R}^2$:

$$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2 \qquad \text{und} \qquad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^2 \qquad \text{mit } a_1, a_2, b_1, b_2 \geqslant 0.$$

Dann ist $\left|\det\left(\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}\right)\right|$ der Flächeninhalt des von a und b aufgespannten Parallelogramms. Analog gilt eine solche Interpretation auch im $\mathbb{R}^n$. ???mehr Literaturhinweis

4.2. *Determinantenfunktionen*

Sei R ein kommutativer Ring mit 1 und $M_n(R)$ der Ring der $n \times n$ Matrizen mit Einträgen in R. Eine Abbildung $\Delta\colon M_n(R) \to R$ heißt ***Determinantenfunktion***, wenn folgende beiden Eigenschaften gelten:

(D1) $\Delta(A)$ ist linear in jeder Zeile von $A \in M_n(R)$.
(D2) Es gilt $\Delta(A) = 0$ wenn zwei Zeilen von A gleich sind.

Genauer ist mit (D1) gemeint: Sei $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in M_n(R)$. Für festes $1 \leqslant k \leqslant n$ seien $s, t \in R$ sowie $a'_{k1}, \ldots, a'_{kn} \in R$ und $a''_{k1}, \ldots, a''_{kn} \in R$ gegeben mit $a_{kj} = sa'_{kj} + ta''_{kj}$ für $1 \leqslant j \leqslant n$; dann gilt

$$\Delta\left(\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}\right) = s\Delta\left(\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a'_{k1} & a'_{k2} & \ldots & a'_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}\right) + t\Delta\left(\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a''_{k1} & a''_{k2} & \ldots & a''_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}\right)$$

wobei alle Zeilen außer der k-ten in den beiden Matrizen auf der rechten Seiten gleich denen von A auf der linken Seite sind. Als Spezialfall von (D1) halten wir fest:

(D1') Entsteht die Matrix $A' \in M_n(R)$ aus $A \in M_n(R)$ durch Multiplikation einer Zeile mit $c \in R$, so gilt $\Delta(A') = c\Delta(A)$.

Aus den obigen Regeln folgen sofort:

(D3) Sind alle Einträge in einer Zeile von A gleich 0, so gilt $\Delta(A) = 0$.
(D4) Seien $1 \leqslant k, l \leqslant n$, $k \neq l$, und $c \in R$. Entsteht $A' \in M_n(R)$ dadurch, dass man das c-Fache der k-ten Zeile von A zur l-ten Zeile addiert, so gilt $\Delta(A') = \Delta(A)$.
(D5) Entsteht $A' \in M_n(R)$ durch Vertauschen von zwei Zeilen in $A \in M_n(R)$, so gilt $\Delta(A') = -\Delta(A)$.

Zu (D3): Ist $1 \leqslant k \leqslant n$ so, dass alle Einträge in der k-ten Zeilen von A gleich sind, dann können wir die k-te Zeile mit $c = 0$ multiplizieren, ohne dass sich A ändert. Also ergibt (D1') sofort $\Delta(A) = 0\Delta(A) = 0$.

Zu (D4): Im Folgenden schreiben wir nur die Einträge in den Zeilen k und l aus, alle anderen Zeilen sind gleich denen in A selbst. Sei zuerst $k < l$. Wegen (D1) ist $\Delta(A')$ gleich

$$\Delta\left(\begin{bmatrix} \vdots & & \vdots \\ a_{k1} & \ldots & a_{kn} \\ \vdots & & \vdots \\ a_{l1} + ca_{k1} & \ldots & a_{ln} + ca_{kn} \\ \vdots & & \vdots \end{bmatrix}\right) = \Delta\left(\begin{bmatrix} \vdots & & \vdots \\ a_{k1} & \ldots & a_{kn} \\ \vdots & & \vdots \\ a_{l1} & \ldots & a_{ln} \\ \vdots & & \vdots \end{bmatrix}\right) + c\Delta\left(\begin{bmatrix} \vdots & & \vdots \\ a_{k1} & \ldots & a_{kn} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kn} \\ \vdots & & \vdots \end{bmatrix}\right).$$

Mit (D2) folgt nun, dass der zweite Summand gleich 0, also das Ergebnis gleich $\Delta(A)$ ist. Das Argument ist völlig analog im Fall $l < k$.

Zu (D5): Seien $1 \leqslant k < l \leqslant n$ so, dass A' durch Vertauschen der Zeilen k und l in A entsteht. Diese Operation erhält man auch wie folgt (Beweis analog zu Aufgabe 6.4):

(i) Zuerst addiere das (-1)-Fache der k-ten Zeile zur l-ten Zeile.
(ii) Dann addiere die l-te Zeile zur k-ten Zeile.
(iii) Anschließend addiere das (-1)-Fache der k-ten Zeile zur l-ten Zeile.
(iv) Zum Schluss multipliziere die l-te Zeile mit -1.

Am Ende sind genau die k-te und l-te Zeile in A vertauscht. Bei (i)–(iii) ändert sich nach (D4) der Wert von Δ nicht. Mit (D1') im Schritt (iv) folgt also $\Delta(A') = -\Delta(A)$.

Die Regel (D5) lässt sich wie folgt verallgemeinern.

LEMMA 4.2.1. *Sei* $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in M_n(R)$. *Für* $\pi \in S_n$ *setzen wir*

$$A^\pi := \begin{bmatrix} a_{\pi(1)1} & a_{\pi(1)2} & \dots & a_{\pi(1)n} \\ a_{\pi(2)1} & a_{\pi(2)2} & \dots & a_{\pi(2)n} \\ \vdots & \vdots & & \vdots \\ a_{\pi(n)1} & a_{\pi(n)2} & \dots & a_{\pi(n)n} \end{bmatrix} \in M_n(R).$$

Dann gilt $\Delta(A^\pi) = \operatorname{sgn}(\pi)\Delta(A)$.

BEWEIS. Nach (D5) gilt die Aussage, wenn $\pi = \tau_{ij} \in S_n$ eine Transposition ist; außerdem gilt die Aussage natürlich für $\pi = \mathrm{id}$. Für den allgemeinen Fall benötigen wir die Regel

$$(*) \qquad A^{\pi \circ \sigma} = (A^\pi)^\sigma \qquad \text{für alle } \pi, \sigma \in S_n.$$

Dazu seien $A_1, \dots, A_n \in R^{1 \times n}$ die Zeilen von A und $B_1, \dots, B_n \in R^{1 \times n}$ die Zeilen von A^π, d.h., es gilt $B_k = A_{\pi(k)}$ für $1 \leqslant k \leqslant n$. Dann sind $B_{\sigma(1)}, \dots, B_{\sigma(n)}$ die Zeilen von $(A^\pi)^\sigma$. Wegen $B_{\sigma(k)} = A_{\pi(\sigma(k))} = A_{(\pi \circ \sigma)(k)}$ gilt also $(*)$. Sei nun $\mathrm{id} \neq \pi \in S_n$. Nach Satz 4.1.2 gibt es Transpositionen $\tau_{i_1 j_1}, \dots, \tau_{i_r j_r} \in S_n$ mit $1 \leqslant r \leqslant n$ und $\pi = \tau_{i_1 j_1} \circ \dots \circ \tau_{i_r j_r}$. Setzen wir $\pi' := \tau_{i_1 j_1} \circ \dots \circ \tau_{i_{r-1} j_{r-1}} \in S_n$ und $B := A^{\pi'}$, so folgt mit $(*)$ und (D5):

$$\Delta(A^\pi) = \Delta(A^{\pi' \circ \tau_{i_r j_r}}) = \Delta\big((A^{\pi'})^{\tau_{i_r j_r}}\big) = \Delta\big(B^{\tau_{i_r j_r}}\big) = -\Delta(B) = -\Delta(A^{\pi'}).$$

Danach wenden wir ein analoges Argument auf $A^{\pi'}$ an; nach insgesamt r Schritten folgt $\Delta(A^\pi) = (-1)^r \Delta(A)$, wobei $(-1)^r = \operatorname{sgn}(\pi)$, siehe Satz 4.1.4. $\square$

HAUPTSATZ 4.2.2. *Die Abbildung* $\det\colon M_n(R) \to R$ *in Definition 4.1.6 ist eine Determinantenfunktion. Für eine beliebige Determinantenfunktion* $\Delta\colon M_n(R) \to R$ *gilt*

$$\Delta(A) = \det(A)\Delta(I_n) \qquad \textit{für alle } A \in M_n(R).$$

BEWEIS. Ist $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in M_n(R)$, so schreiben wir nun

$$\Delta(A) = \Delta(A_1, \dots, A_n) \qquad \text{und} \qquad \det(A) = \det(A_1, \dots, A_n),$$

wobei $A_1, \dots, A_n \in R^{1\times n}$ die Zeilen von A sind. Sei $1 \leqslant k \leqslant n$ fest und seien $s, t \in R$ sowie Zeilenvektoren $A'_k, A''_k \in R^{1\times n}$ gegeben mit $A_k = sA'_k + tA''_k$. Dann gilt

$$\begin{aligned}
\det(A_1, \dots, A_n) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i\pi(i)} = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{k\pi(k)} \prod_{1\leqslant i\leqslant n, i\neq k} a_{i\pi(i)} \\
&= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)(sa'_{k\pi(k)} + ta''_{k\pi(k)}) \prod_{1\leqslant i\leqslant n, i\neq k} a_{i\pi(i)} \\
&= s\underbrace{\Big(\sum_{\pi\in S_n} \operatorname{sgn}(\pi) a'_{k\pi(k)} \prod_{1\leqslant i\leqslant n, i\neq k} a_{i\pi(i)}\Big)}_{=\det(A_1,\dots,A_{k-1},A'_k,A_{k+1},\dots,A_n)} + t\underbrace{\Big(\sum_{\pi\in S_n} \operatorname{sgn}(\pi) a''_{k\pi(k)} \prod_{1\leqslant i\leqslant n, i\neq k} a_{i\pi(i)}\Big)}_{=\det(A_1,\dots,A_{k-1},A''_k,A_{k+1},\dots,A_n)},
\end{aligned}$$

also ist $\det(A_1, \dots, A_n)$ linear in jeder Zeile A_k, d.h., es gilt (D1) für det. Da (D2) bereits in Lemma 4.1.9 gezeigt wurde, ist damit det eine Determinantenfunktion. Sei nun umgekehrt $\Delta\colon M_n(R) \to R$ eine beliebige Determinantenfunktion. Seien außerdem $e_1, \dots, e_n \in R^{1\times n}$ die Einheits-Zeilenvektoren, d.h., e_k hat Eintrag 1 an der Stelle k und 0 überall sonst. Damit ist

$$A_k = a_{k1}e_1 + a_{k2}e_2 + \ldots + a_{kn}e_n = \sum_{i=1}^{n} a_{ki}e_i \qquad \text{für } 1 \leqslant k \leqslant n.$$

Durch wiederholte Anwendung von (D1) folgt

$$\begin{aligned}
\Delta(A) = \Delta(A_1, A_2, \dots, A_n) &= \det\Big(\sum_{i_1=1}^{n} a_{1i_1} e_{i_1}, A_2, \dots, A_n\Big) = \sum_{i_1=1}^{n} a_{1i_1} \Delta(e_{i_1}, A_2, \dots, A_n) \\
&= \sum_{i_1=1}^{n} a_{1i_1} \Delta\Big(e_{i_1}, \sum_{i_2=1}^{n} a_{2i_2} e_{i_2}, A_3, \dots, A_n\Big) = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} a_{1i_1} a_{2i_2} \Delta(e_{i_1}, e_{i_2}, A_3, \dots, A)_n) \\
&= \ldots = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \ldots \sum_{i_n=1}^{n} a_{1i_1} a_{2i_2} \cdots a_{ni_n} \Delta(e_{i_1}, e_{i_2}, \dots, e_{i_n}) \\
&= \sum_{1\leqslant i_1, i_2, \dots, i_n \leqslant n} a_{1i_1} a_{2i_2} \cdots a_{ni_n} \Delta(e_{i_1}, e_{i_2}, \dots, e_{i_n}).
\end{aligned}$$

Wegen (D2) ist $\Delta(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = 0$ wenn zwei der Indizes $i_1, i_2, \dots, i_n$ gleich sind. Also brauchen wir nur die Tupel $(i_1, i_2, \dots, i_n)$ zu betrachten, bei denen alle Komponenten verschieden sind; dies ist aber genau dann der Fall, wenn $\left(\begin{smallmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{smallmatrix}\right) \in S_n$ eine Permutation ist. Also können wir den obigen Ausdruck auch schreiben als

$$\Delta(A) = \sum_{\pi\in S_n} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} \Delta(e_{\pi(1)}, e_{\pi(2)}, \dots, e_{\pi(n)}).$$

Nun ist aber die Matrix mit Zeilen $e_{\pi(1)}, e_{\pi(2)}, \dots, e_{\pi(n)}$ genau die Matrix I_n^π; mit Lemma 4.2.1 folgt also $\Delta(e_{\pi(1)}, e_{\pi(2)}, \dots, e_{\pi(n)}) = \Delta(I_n^\pi) = \operatorname{sgn}(\pi)\Delta(I_n)$. $\square$

Bemerkung 4.2.3. Sei $R = K$ ein Körper und $A \in M_n(K)$. Mit elementaren Zeilenumformungen (Gauß-Algorithmus) erhalten wir $A \to A'$ wobei $A' \in M_n(K)$ in Zeilenstufenform ist, mit r Stufen und Pivots $1 \leqslant j_1 < \ldots < j_r \leqslant n$. Da wir nun wissen, dass det eine Determinantenfunktion ist, gelten die obigen Regeln (D1), (D1'), (D2), (D3), (D4), (D5). Aus diesen folgt insbesondere, dass sich bei einer elementaren Zeilenumformung die Determinante einer Matrix entweder gar nicht ändert (siehe (D4)), oder das Vorzeichen wechselt (siehe (D5)), oder mit einer Konstante ungleich 0 multipliziert wird (siehe (D1')). Damit folgt sofort:

(a) Es gibt ein $0 \neq c \in K$ mit $\det(A) = c\det(A')$ (und dieses c kann berechnet werden, wenn man sich die Umformungen merkt, die für $A \to A'$ nötig sind).

Nach Kapitel 3, §1, ist $r = \mathrm{Rang}(A)$, und A ist invertierbar genau dann, wenn $r = n$ gilt. In diesem Fall ist $A' = I_n$ (siehe Kapitel 1, §3), also $\det(A) = c\det(I_n) = c \neq 0$. Ist dagegen $r < n$, so besteht die letzte Zeile in A' nur aus Nullen; mit (D3) folgt, dass $\det(A) = c\det(A') = 0$ gilt. Damit haben wir gezeigt:

(b) Es gilt $\det(A) = c \neq 0$, falls $r = n$; und $\det(A) = 0$, falls $r < n$. Insbesondere ist A invertierbar genau dann, wenn $\det(A) \neq 0$ gilt.

Die durch (a) und (b) gegebene Methode zur Berechnung von $\det(A)$ ist praktisch viel effizienter als die Leibniz-Formel.

Satz 4.2.4 (Cramersche Regel, 1750). *Sei $R = K$ ein Körper, $b \in K^n$ und $A \in M_n(K)$ invertierbar. Dann ist $\det(A) \neq 0$ und das lineare Gleichungssystem mit erweiterter Matrix $[A|b]$ hat genau eine Lösung $x \in K^n$ (nämlich $x := A^{-1}b$). Seien $b_1, \ldots, b_n \in K$ die Komponenten von b. Für $1 \leqslant k \leqslant n$ sei $x_k \in K$ die k-te Komponente von x. Dann gilt*

$$x_k = \det(A)^{-1}\det\left(\begin{bmatrix} a_{11} & \ldots & a_{1,k-1} & b_1 & a_{1,k+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,k-1} & b_2 & a_{2,k+1} & \ldots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,k-1} & b_n & a_{n,k+1} & \ldots & a_{nn} \end{bmatrix}\right).$$

Beweis. Nach Bemerkung 4.2.3 gilt $\det(A) \neq 0$. Seien nun $S_1, \ldots, S_n \in K^n$ die Spalten von A; wir schreiben dann auch $\det(A) = \det(S_1, \ldots, S_n)$. Für $1 \leqslant i \leqslant n$ sei $e_i \in K^n$ der i-te Einheitsvektor. Dann gilt $x = x_1e_1 + \ldots + x_ne_n$ und damit

$$b = Ax = A(x_1e_1 + \ldots + x_ne_n) = x_1Ae_1 + \ldots + x_nAe_n = x_1S_1 + \ldots + x_nS_n.$$

Nach Lemma 4.1.8 gelten die Regeln (D1), (D2) analog auch für Spalten. Damit folgt:

$$\det(S_1,\ldots,S_{k-1},b,S_{k+1},\ldots,S_n) = \det(S_1,\ \ldots,\ S_{k-1},\ \sum_{i=1}^n x_i S_i, S_{k+1},\ \ldots,\ S_n)$$

$$= \sum_{i=1}^n x_i \underbrace{\det(S_1,\ldots,S_{k-1},S_i,S_{k+1},\ldots,S_n)}_{=0 \text{ wenn } i\neq k, \text{ wegen (D2)}} = x_k \underbrace{\det(S_1,\ldots,S_{k-1},S_k,S_{k+1},\ldots,S_n)}_{=\det(A)}.$$

Wegen $\det(A) \neq 0$ ergibt dies $x_k = \det(A)^{-1}\det(S_1,\ldots,S_{k-1},b,S_{k+1},\ldots,S_n)$. □

4.3. *Produktregel und Laplace-Entwicklung*

Mit Hilfe des Hauptsatzes 4.2.2 können wir nun weitere Eigenschaften der Determinante zeigen (ohne noch einmal explizit mit der Leibniz-Formel argumentieren zu müssen). Sei wieder zunächst R ein beliebiger kommutativer Ring mit 1.

Satz 4.3.1 (Produktregel). *Für alle* $A, B \in M_n(R)$ *gilt* $\det(AB) = \det(A)\det(B)$.

Beweis. Sei $B \in M_n(R)$ fest. Dann definieren wir eine Abbildung $\Delta\colon M_n(R) \to R$ durch $\Delta(A) := \det(AB)$ für alle $A \in M_n(R)$. Wenn wir zeigen können, dass Δ eine Determinantenfunktion ist, so folgt aus Hauptsatz 4.2.2 sofort

$$\det(AB) = \Delta(A) = \det(A)\Delta(I_n) = \det(A)\det(I_nB) = \det(A)\det(B).$$

Sei nun $A \in M_n(R)$ und seien $A_1,\ldots,A_n \in R^{1\times n}$ die Zeilen von A. Dann sind (nach Definition des Matrixproduktes) die Zeilen von AB gegeben durch $A_1B,\ldots,A_nB$. Damit ist

$$\Delta(A) = \Delta(A_1,\ldots,A_n) = \det(A_1B,\ldots,A_nB).$$

Gibt es $1 \leqslant k < l \leqslant n$ mit $A_k = A_l$, so gilt auch $A_kB = A_lB$ und damit ist die rechte Seite oben gleich 0 (weil det eine Determinantenfunktion ist). Also gilt (D2) für Δ. Nun zu (D1). Sei $1 \leqslant k \leqslant n$ fest und seien $s, t \in R$ sowie Zeilenvektoren $A_k', A_k'' \in R^{1\times n}$ gegeben mit $A_k = sA_k' + tA_k''$. Dann gilt auch $A_kB = (sA_k' + tA_k'')B = sA_k'B + tA_k''B$ und damit

$$\begin{aligned}\Delta(A) &= \det(A_1B,\ldots,A_{k-1}B, sA_k'B + tA_k''B, A_{k+1}B,\ldots,A_nB)\\ &= s\det(A_1B,\ldots,A_{k-1}B,A_k'B,A_{k+1}B,\ldots,A_nB)\\ &\qquad + t\det(A_1B,\ldots,A_{k-1}B,A_k''B,A_{k+1}B,\ldots,A_nB)\\ &= s\Delta(A_1,\ldots,A_{k-1},A_k',A_{k+1},\ldots,A_n) + t\Delta(A_1,\ldots,A_{k-1},A_k'',A_{k+1},\ldots,A_n),\end{aligned}$$

d.h., Δ ist in jeder Zeile linear. Damit gilt auch (D1) für Δ. □

Satz 4.3.2. *Sei* $1 \leqslant d < n$ *und* $A \in M_n(R)$ *eine untere Blockdreiecksmatrix der Form*

$$A = \left[\begin{array}{c|c} B & 0 \\ \hline D & C \end{array}\right] \quad \textit{mit} \quad B \in M_d(R), \quad C \in M_{n-d}(R), \quad D \in R^{(n-d)\times d}.$$

Dann gilt $\det(A) = \det(B)\det(C)$*. Eine analoge Aussage gilt auch für obere Blockdreiecksmatrizen (mit Hilfe von Lemma 4.1.8).*

Beweis. Zunächst rechnen wir nach, dass folgendes gilt:

$$A = \left[\begin{array}{c|c} B & 0 \\ \hline D & C \end{array}\right] = \left[\begin{array}{c|c} B & 0 \\ \hline D & I_{n-d} \end{array}\right] \cdot \left[\begin{array}{c|c} I_d & 0 \\ \hline 0 & C \end{array}\right].$$

Definieren wir nun $\Delta_1 \colon M_d(R) \to R$ und $\Delta_2 \colon M_{n-d}(R) \to R$ durch

$$\Delta_1(B) := \det\left(\left[\begin{array}{c|c} B & 0 \\ \hline D & I_{n-d} \end{array}\right]\right) \qquad \text{und} \qquad \Delta_2(C) := \det\left(\left[\begin{array}{c|c} I_d & 0 \\ \hline 0 & C \end{array}\right]\right)$$

für alle $B \in M_d(R)$ und alle $C \in M_{n-d}(R)$. Nach Satz 4.3.1 ist $\det(A) = \Delta_1(B)\Delta_2(C)$, also genügt es zu zeigen, dass $\Delta_1(B) = \det(B)$ und $\Delta_2(C) = \det(C)$ gilt.

Behauptung: Δ_1 ist eine Determinantenfunktion. Nun, schreiben wir eine Zeile von B als Linearkomination von zwei Zeilenvektoren in $R^{1\times d}$, so ist auch die entsprechende Zeile von

$$\left[\begin{array}{c|c} B & 0 \\ \hline D & I_{n-d} \end{array}\right]$$

eine solche Linearkombination (weil der rechte obere $d \times (n-d)$ Block nur aus 0 besteht), also gilt (D1). Analog: Sind zwei Zeilen von B gleich, so sind die entsprechenden Zeilen auch in obiger Matrix gleich (wiederum weil der rechte obere $d \times (n-d)$ Block nur aus 0 besteht), also gilt (D2). Nach Hauptsatz 4.2.2 folgt also $\Delta_1(B) = \det(B)\Delta_1(I_d)$; und mit Lemma 4.1.7 (Determinante einer unteren Dreiecksmatrix) folgt, dass $\Delta_1(I_d) = 1$ gilt. Also ist $\Delta_1(B) = \det(B)$. Völlig analog sieht man, dass auch $\Delta_2(C) = \det(C)$ gilt. □

Zum Beispiel ist

$$\det\left(\begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 4 & -4 & -4 \\ 0 & 0 & -4 & -8 \\ 0 & 0 & 4 & 0 \end{bmatrix}\right) = \det\left(\left[\begin{array}{cc|cc} 1 & -1 & 2 & 3 \\ 0 & 4 & -4 & -4 \\ \hline 0 & 0 & -4 & -8 \\ 0 & 0 & 4 & 0 \end{array}\right]\right)$$

$$= \det\left(\begin{bmatrix} 1 & -1 \\ 0 & 4 \end{bmatrix}\right) \cdot \det\left(\begin{bmatrix} -4 & -8 \\ 4 & 0 \end{bmatrix}\right) = (1 \cdot 4) \cdot ((-4) \cdot 0 - (-8) \cdot 4)) = 4 \cdot 32 = 128.$$

Beispiel 4.3.3. Seien $x_1, \dots, x_n \in R$. Dann betrachte die ***Vandermonde-Matrix***

$$V(x_1, \dots, x_n) := \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix} \in M_n(R).$$

In Kapitel 2, §3, haben wir gesehen, dass $V(x_1,\ldots,x_n)$ invertierbar ist, wenn $R = K$ ein Körper ist und die x_i paarweise verschieden sind. Wir behaupten nun:

$$\det\bigl(V(x_1,\ldots,x_n)\bigr) = \prod_{1\leqslant i<j\leqslant n}(x_j - x_i).$$

Dazu: Nach Lemma 4.1.8 können wir auch die tranponierte Matrix betrachten; wir wenden nun geschickt elementare Zeilenumformungen auf $V(x_1,\ldots,x_n)^{\mathrm{tr}}$ an, nämlich: Ziehe zuerst das x_1-Fache der $(n-1)$-Zeile von der n-ten Zeile ab, dann das x_1-Fache der $(n-2)$-ten Zeile von der $(n-1)$-ten Zeile und so fort. Mit (D4) erhalten wir

$$\det\bigl(V(x_1,\ldots,x_n)\bigr) = \det\bigl(V(x_1,\ldots,x_n)^{\mathrm{tr}}\bigr) = \det\left(\begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-2} & x_2^{n-2} & \ldots & x_n^{n-2} \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{bmatrix}\right)$$

$$= \det\left(\left[\begin{array}{c|ccc} 1 & 1 & \ldots & 1 \\ \hline 0 & x_2-x_1 & \ldots & x_n-x_1 \\ 0 & x_2(x_2-x_1) & \ldots & x_n(x_n-x_1) \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2-x_1) & \ldots & x_n^{n-2}(x_n-x_1) \end{array}\right]\right) = \det\left(\begin{bmatrix} x_2-x_1 & \ldots & x_n-x_1 \\ x_2(x_2-x_1) & \ldots & x_n(x_n-x_1) \\ \vdots & & \vdots \\ x_2^{n-2}(x_2-x_1) & \ldots & x_n^{n-2}(x_n-x_1) \end{bmatrix}\right),$$

wobei im letzten Schritt Satz 4.3.2 angewandt wurde. In der letzten Matrix (die nur noch $n-1$ Zeilen und Spalten hat) können wir den Faktor $x_2 - x_1$ aus der ersten Spalte herausziehen, dann den Faktor $x_3 - x_1$ aus der zweiten Spalte und so fort bis zum Faktor $x_n - x_1$ in der letzten Spalte. Mit Lemma 4.1.8 und (D1') folgt also:

$$\det\bigl(V(x_1,\ldots,x_n)\bigr) = (x_2-x_1)(x_3-x_1)\cdots(x_n-x_1)\det\left(\begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_2 & x_3 & \ldots & x_n \\ \vdots & \vdots & & \vdots \\ x_2^{n-2} & x_3^{n-2} & \ldots & x_n^{n-2} \end{bmatrix}\right).$$

Die Matrix auf der rechten Seite ist nun eine (transponierte) Vandermonde-Matrix der Größe $(n-1)\times(n-1)$. Also folgt die Behauptung mit vollständiger Induktion nach n (wobei der Induktionsanfang $n = 1$ einfach durch $\det(V(x_1)) = 1$ gegeben ist).

Schließlich kommen wir noch zu einer rekursiven Formel zur Berechnung von Determinanten. Sei dazu $n \geqslant 2$ und $A = [a_{ij}]_{1\leqslant i,j\leqslant n} \in M_n(R)$. Seien $1 \leqslant k,l \leqslant n$ fest. Dann sei $A^{kl} \in$

$M_{n-1}(R)$ die Matrix, die aus A entsteht, wenn man die k-Zeile und l-te Spalte streicht, also

$$A^{kl} = \begin{bmatrix} a_{11} & \dots & a_{1,l-1} & a_{1,l+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,l-1} & a_{k-1,l+1} & \dots & a_{k-1,n} \\ a_{k+1,1} & \dots & a_{k+1,l-1} & a_{k+1,l+1} & \dots & a_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,l-1} & a_{n,l+1} & \dots & a_{nn} \end{bmatrix} \in M_{n-1}(R).$$

Wir definieren dann eine neue Matrix $\tilde{A} = [\tilde{a}_{kl}]_{1 \leqslant k,l \leqslant n} \in M_n(R)$ durch

$$\tilde{a}_{kl} := (-1)^{k+l} \det(A^{lk}) \qquad \text{für } 1 \leqslant k, l \leqslant n.$$

Diese Matrix $\tilde{A}$ heißt auch die ***Adjunkte*** zu A (englisch: "adjugate" oder "adjunct") und wird manchmal mit $\operatorname{adj}(A)$ bezeichnet. Der folgende Satz zeigt, dass man $\det(A)$ rekursiv mit Hilfe der Einträge der Adjunkten berechnen kann.

Hauptsatz 4.3.4 (Laplace-Entwicklung). *Mit obigen Bezeichnungen gilt* $A \cdot \tilde{A} = \tilde{A} \cdot A = \det(A) I_n$*, wobei* $\tilde{A} = \operatorname{adj}(A)$ *die Adjunkte zu* A *ist. Insbesondere erhält man damit:*

$$\det(A) = \sum_{l=1}^{n} (-1)^{k+l} a_{kl} \det(A^{kl}) \qquad \textit{(Entwicklung nach der } k\textit{-ten Zeile)},$$

$$\det(A) = \sum_{k=1}^{n} (-1)^{k+l} a_{kl} \det(A^{kl}) \qquad \textit{(Entwicklung nach der } l\textit{-ten Spalte)}.$$

Beweis. Seien $A_1, \dots, A_n \in R^{1 \times n}$ die Zeilen von A. Für $1 \leqslant l \leqslant n$ sei $e_l \in R^{1 \times n}$ der l-te Zeileneinheitsvektor; es gilt dann also $A_j = \sum_{l=1}^{n} a_{jl} e_l$ für $1 \leqslant j \leqslant n$. Für $1 \leqslant k, l \leqslant n$ sei $B^{kl} \in M_n(R)$ die Matrix mit Zeilen $A_1, \dots, A_{k-1}, e_l, A_{k+1}, \dots, A_n$, also

$$B^{kl} = \begin{bmatrix} a_{1,1} & \dots & a_{1,l-1} & a_{1,l} & a_{1,l+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,l-1} & a_{k-1,l} & a_{k-1,l+1} & \dots & a_{k-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{k+1,1} & \dots & a_{k+1,l-1} & a_{k+1,l} & a_{k+1,l+1} & \dots & a_{k+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,l-1} & a_{n,l} & a_{n,l+1} & \dots & a_{n,n} \end{bmatrix} \in M_n(R).$$

Wenden wir nacheinander die $k-1$ Transpositionen $\tau_{k-1,k}, \tau_{k-2,k-1}, \dots, \tau_{1,2}$ auf die Zeilen von B^{kl} an, so wird einfach die k-te Zeile (die aus e_l besteht) in die erste Zeile nach oben versetzt. Wendet man anschließend die $l-1$ Transpositionen $\tau_{l-1,l}, \tau_{l-2,l-1}, \dots, \tau_{12}$ auf die

Spalten an, so erhält man am Ende die Matrix

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ a_{1,l} & a_{1,1} & \dots & a_{1,l-1} & a_{1,l+1} & \dots & a_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{k-1,l} & a_{k-1,1} & \dots & a_{k-1,l-1} & a_{k-1,l+1} & \dots & a_{k-1,n} \\ a_{k+1,l} & a_{k+1,1} & \dots & a_{k+1,l-1} & a_{k+1,l+1} & \dots & a_{k+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,l} & a_{n,1} & \dots & a_{n,l-1} & a_{n,l+1} & \dots & a_{nn} \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline * & & & \\ \vdots & & A^{kl} & \\ * & & & \end{array}\right] \in M_n(R).$$

Mit den Regeln (D5) (sowie Lemma 4.1.8) und Satz 4.3.2 folgt also

$$\det(B^{kl}) = (-1)^{k-1}(-1)^{l-1}\det(A_{kl}) = (-1)^{k+l}\det(A^{kl}) = \tilde{a}_{lk}.$$

Damit können wir nun wie folgt argumentieren. Sei $1 \leqslant j \leqslant n$ und $C \in M_n(R)$ die Matrix mit Zeilen $A_1, \dots, A_{k-1}, A_j, A_{k+1}, \dots, A_n$. Weil det linear in der k-ten Zeile ist, gilt

$$\begin{aligned} \det(C) &= \det\bigl(A_1, \dots, A_{k-1}, A_j, A_{k+1}, \dots, A_n) \\ &= \det\bigl(A_1, \dots, A_{k-1}, \sum_{l=1}^{n} a_{jl}e_l, A_{k+1}, \dots, A_n) \\ &= \sum_{l=1}^{n} a_{jl}\det(A_1, \dots, A_{k-1}, e_l, A_{k+1}, \dots, A_n) \\ &= \sum_{l=1}^{n} a_{jl}\det(B^{kl}) = \sum_{l=1}^{n}(-1)^{k+l}a_{jl}\det(A^{kl}) = \sum_{l=1}^{n} a_{jl}\tilde{a}_{lk} = (A \cdot \tilde{A})_{jk}. \end{aligned}$$

Ist $j \neq k$, so sind zwei Zeilen von C gleich, also $\det(C) = 0$. Ist $j = k$, so ist $C = A$. Also:

$$\sum_{l=1}^{n} a_{jl}\tilde{a}_{lk} = \begin{cases} \det(A) & \text{falls } j = k, \\ 0 & \text{falls } j \neq k, \end{cases}$$

d.h., es gilt $A \cdot \tilde{A} = \det(A)I_n$. Insbesondere gilt damit die Formel zur Entwicklung nach der k-ten Zeile. Durch Transponieren erhalten wir

$$\tilde{A}^{\mathrm{tr}} \cdot A^{\mathrm{tr}} = (A \cdot \tilde{A})^{\mathrm{tr}} = \bigl(\det(A)I_n\bigr)^{\mathrm{tr}} = \det(A)I_n = \det(A^{\mathrm{tr}})I_n.$$

Ersetzen wir A durch A^{tr} und beachten $\operatorname{adj}\bigl(A^{\mathrm{tr}}\bigr) = \bigl(\operatorname{adj}(A)\bigr)^{\mathrm{tr}}$, so folgt $\tilde{A} \cdot A = \det(A)I_n$, und dies ergibt dann auch die Formel zur Entwicklung nach der k-ten Spalte von A. □

FOLGERUNG 4.3.5. *Sei* $R = K$ *ein Körper und* $A \in M_n(K)$ *mit* $\det(A) \neq 0$*; nach Bemerkung 4.2.3 ist* A *invertierbar. Dann gilt* $A^{-1} = \det(A)^{-1}\tilde{A}$ *und* $\det(A^{-1}) = \det(A)^{-1}$.

BEWEIS. Setze $B := \det(A)^{-1}\tilde{A}$. Nach Hauptsatz 4.3.4 gilt $A \cdot B = I_n$, also $B = A^{-1}$. Mit der Produktregel folgt $1 = \det(AA^{-1}) = \det(A)\det(A^{-1})$, also $\det(A^{-1}) = \det(A)^{-1}$. □

Beispiel Entwicklung nach der 1. Spalte bzw. 3. Zeile:

$$\begin{aligned}
\det\left(\begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 5 \\ 1 & 6 & 9 \end{bmatrix}\right) &= 2 \cdot \det\left(\begin{bmatrix} 2 & 5 \\ 6 & 9 \end{bmatrix}\right) - 4 \cdot \det\left(\begin{bmatrix} 1 & 3 \\ 6 & 9 \end{bmatrix}\right) + 1 \cdot \det\left(\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}\right) \\
&= 2 \cdot (18 - 30) - 4 \cdot (9 - 18) + 1 \cdot (5 - 6) = 11, \\
\text{oder (3. Zeile)} \quad &= 1 \cdot \det\left(\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}\right) - 6 \cdot \det\left(\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}\right) + 9 \cdot \det\left(\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}\right) \\
&= 1 \cdot (5 - 6) - 6 \cdot (10 - 12) + 9 \cdot (4 - 4) = 11.
\end{aligned}$$

In manchen Büchern wird die Formel aus der Laplace-Entwicklung als *Definition* der Determinante benutzt. (Dann muss man am Ende zeigen, dass auch die Leibniz-Formel gilt.)

4.4. *Der Satz von Cayley–Hamilton*

Sei K ein Körper. Wir betrachten nun auch den Polynomring $R = K[X]$ über K in der Unbestimmten X. Die Elemente von R sind also (abstrakte) Polynome

$$f = a_0 + a_1 + \ldots + a_d X^d \qquad \text{mit } d \geqslant 0,\ a_i \in K.$$

Ist $f \neq 0$ und $a_d \neq 0$, so heißt $\mathrm{Grad}(f) = d$ der Grad von f. Eine wichtige Regel: Sind $f, g \in K[X]$ mit $f \neq 0$ und $g \neq 0$, so gilt auch $f * g \neq 0$ und $\mathrm{Grad}(f * g) = \mathrm{Grad}(f) + \mathrm{Grad}(g)$. (Für all dies siehe Kapitel 3, §4.) Da $R = K[X]$ ein kommutativer Ring mit 1 ist, können wir auch Matrizen mit Einträgen in $K[X]$ bilden und davon die Determinante; es gilt also

$$\det(F) \in K[X] \qquad \text{für alle} \qquad F \in M_n(K[X]).$$

Zu einer gegebenen Matrix $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in M_n(K)$ bilden wir nun

$$F := A - XI_n \in M_n(K[X]) \qquad \text{und} \qquad \chi_A := \det(F) = \det(A - XI_n) \in K[X];$$

dann heißt χ_A das ***charakteristische Polynom*** von A.

Beispiel 4.4.1. (a) Sei $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in M_2(K)$. Dann ist $F = A - XI_2 = \begin{bmatrix} -X & 1 \\ 1 & 1 - X \end{bmatrix}$ und

$$\chi_A = \det(A - XI_2) = \det\left(\begin{bmatrix} -X & 1 \\ 1 & 1 - X \end{bmatrix}\right) = (-X) * (1 - X) - 1 \cdot 1 = X^2 - X - 1.$$

(b) Sei $A \in M_n(K)$ eine obere Dreiecksmatrix mit $c_1, \ldots, c_n \in K$ auf der Diagonalen. Dann ist $F = A - XI_n$ eine obere Dreiecksmtarix mit $c_1 - X, \ldots, c_n - X$ auf der Diagonalen. Mit Lemmas 4.1.7 und 4.1.8 folgt also

$$\chi_A = \det(A - XI_n) = (c_1 - X) * (c_2 - X) * \ldots * (c_n - X).$$

(c) Sei $K = \mathbb{Q}$ und $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix} \in M_3(\mathbb{Q})$ die Matrix aus Kapitel 3, Beispiel 3.8, mit Minimalpolynom $\tilde{\mu}_A = X * (X+2)$. Mit Entwicklung nach der 2. Spalte erhalten wir

$$\chi_A = \det\left(\begin{bmatrix} -1-X & 0 & 1 \\ 3 & -X & -3 \\ 1 & 0 & -1-X \end{bmatrix}\right) = (-1)^{2+2}(-X) * \det\left(\begin{bmatrix} -1-X & 1 \\ 1 & -1-X \end{bmatrix}\right)$$

$$(-X) * \big((-1-X) * (-1-X) - 1 \cdot 1\big) = (-X) * (X^2 + 2X) = -X^2 * (X+2).$$

Das charakteristische Polynom ist also "fast" gleich dem Minimalpolynom. Nach Kapitel 1, §4, hat die Matrix in (a) das Minimalpolynom $\tilde{\mu}_A = X^2 - X - 1$; hier ist also sogar $\chi_A = \tilde{\mu}_A$.

Lemma 4.4.2. *Es gilt* $\chi_A \neq 0$ *und* $\mathrm{Grad}(\chi_A) = n$. *Schreiben wir* $\chi_A = a_0 + a_1 X + \ldots + a_n X^n$, *so gilt* $a_n = (-1)^n$ *und* $a_{n-1} = (-1)^{n-1}\,\mathrm{Spur}(A)$, *wobei* $\mathrm{Spur}(A) := \sum_{i=1}^n a_{ii}$ *(Summe der Diagonaleinträge) als die* ***Spur*** *von* A *bezeichnet wird.*

Beweis. Sei $F = [f_{ij}]_{1 \leqslant i,j \leqslant n} = A - XI_n \in M_n(K[X])$ und $\chi_A = \det(F) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) f_\pi$ wobei $f_\pi := f_{1\pi(1)} * \ldots * f_{n\pi(n)}$. Sei $\pi \in S_n$ mit $f_\pi \neq 0$, also $f_{i\pi(i)} \neq 0$ für alle i. Für $i = \pi(i)$ ist $f_{i\pi(i)} = a_{ii} - X$, hat also Grad 1; für $i \neq \pi(i)$ ist $f_{i\pi(i)} = a_{ij}$, hat also Grad 0. Damit ist $\mathrm{Grad}(f_\sigma) \leqslant n$. If $\pi \neq \mathrm{id}$, so gibt es mindestens zwei Ziffern $i \neq j$ mit $\pi(i) \neq i$ und $\pi(j) \neq j$, also ist dann $\mathrm{Grad}(f_\pi) \leqslant n-2$. Für $\pi = \mathrm{id}$ ist $f_{\mathrm{id}} = (a_{11} - X) * \ldots * (a_{nn} - X)$. Damit folgt

$$\chi_A = (a_{11} - X) * \ldots * (a_{nn} - X) + \text{Terme mit Grad} \leqslant n-2.$$

Sind irgendwelche $c_1, \ldots, c_n \in K$ gegeben und multipliziert man $(c_1 - X) * \ldots * (c_n - X)$ aus, so erhält man $(-X)^n + (-X)^{n-1}(c_1 + \ldots + c_n) +$ Terme vom Grad $\leqslant n-2$ (einfacher Beweis mit vollständiger Induktion nach n). □

Lemma 4.4.3. *Sei* $\lambda \in K$. *Dann gilt* $\chi_A(\lambda) = \det(A - \lambda I_n) \in K$. *Insbesondere ist* $\det(A) = \chi_A(0)$ *der konstante Term von* χ_A.

Beweis. Wie im obigen Beweis schreibe $\chi_A = \det(F) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) f_\pi$. Da Einsetzen ein Homomorphismus ist, folgt $\chi_A(\lambda) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) f_\pi(\lambda)$ und $f_\pi(\lambda) = f_{1\pi(1)}(\lambda) \cdots f_{n\pi(n)}(\lambda)$. Ist $i = \pi(i)$, so ist $f_{i\pi(i)} = a_{ii} - X$ also $f_{i\pi(i)}(\lambda) = a_{ii} - \lambda$; ist $i \neq \pi(i)$, so ist $f_{i\pi(i)} = a_{i\pi(i)}$ also $f_{i\pi(i)}(\lambda) = a_{i\pi(i)}$. Damit ist $\sum_{\pi \in S_n} \mathrm{sgn}(\pi) f_\pi(\lambda) = \det(B)$, wobei $B = [b_{ij}]_{1 \leqslant i,j \in n} \in M_n(K)$ gegeben ist durch $b_{ij} = a_{ij}$ falls $i \neq j$, und $b_{ii} = a_{ii} - \lambda$, d.h., $B = A - \lambda I_n$. □

Lemma 4.4.4. *Sind* $A, B \in M_n(K)$ *ähnlich (d.h., es gibt eine invertiebare Matrix* $T \in M_n(K)$ *mit* $B = T^{-1}AT$*), so gilt* $\chi_A = \chi_B$, *also auch* $\det(A) = \det(B)$ *und* $\mathrm{Spur}(A) = \mathrm{Spur}(B)$.

BEWEIS. Es gilt $T^{-1}(A - XI_n)T = T^{-1}AT^{-1} - XT^{-1}T = B - XI_n$ und damit $\chi_B = \det\big(T^{-1}(A - XI_n)T\big) = \det(T^{-1})\chi_A \det(T) = \det(T^{-1}T)\chi_A = \det(I_n)\chi_A = \chi_A$. Die Aussagen zu $\det(A)$ und $\mathrm{Spur}(A)$ folgen dann aus Lemmas 4.4.2 und 4.4.4. $\square$

FOLGERUNG 4.4.5. *Sei $\lambda \in K$. Genau dann ist λ ein Eigenwert von A, wenn $\chi_A(\lambda) = 0$ gilt. Damit haben also χ_A und das Minimalpolynom $\tilde{\mu}_A$ die gleichen Nullstellen in K (nämlich genau die Eigenwerte von A).*

Damit haben wir eine neue Methode, um Eigenwerte von A zu bestimmen!

BEWEIS. Es gelten die Äquivalenzen: λ ist ein Eigenwert von A

$\Leftrightarrow$ es gibt ein $0_n \neq v \in K^n$ (Spaltenvektor) mit $Av = \lambda v$

$\Leftrightarrow$ das lineare Gleichungssystem $(A - \lambda I_n)x = 0_n$ hat eine Lösung $0_n \neq x \in K^n$

$\Leftrightarrow$ die Matrix $A - \lambda I_n$ ist nicht invertierbar $\Leftrightarrow$ $\det(A - \lambda I_n) = 0$,

wobei die letzte Äquivalenz nach Bemerkung 4.2.3 gilt. Schließlich zur Erinnerung: λ ist ein Eigenwert von A genau dann, wenn $\tilde{\mu}_A(\lambda) = 0$ gilt (siehe Kapitel 1, §4). $\square$

HAUPTSATZ 4.4.6 (Cayley–Hamilton, 1853/58; allgemeiner Beweis Frobenius 1878). *Für $A \in M_n(K)$ gilt stets $\chi_A(A) = 0_{n\times n}$; damit folgt insbesondere* $\mathrm{Grad}(\tilde{\mu}_A) \leqslant n$.

In Kapitel 1, §5, haben wir dies für 2×2 Matrizen nachgerechnet; für $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ gilt

$$\chi_A = \det\left(\begin{bmatrix} a - X & b \\ c & d - X \end{bmatrix}\right) = X^2 - (a + b)X + ad - bc = X^2 - \mathrm{Spur}(A)X + \det(A)$$

und damit $\chi_A(A) = A^2 - \mathrm{Spur}(A)A + \det(A)I_2 = 0_{2\times 2}$. Für den allgemeinen Fall braucht man also einen "Trick".

BEWEIS. Sei $F = [f_{ij}]_{1\leqslant i,j\leqslant n} = A - XI_n \in M_n(K[X])$ und $\chi_A = \det(F)$. Mit Laplace-Entwicklung gilt dann $F \cdot \mathrm{adj}(F) = \det(F)I_n$; mit $\mathrm{adj}(F) = [\tilde{f}_{ij}] \in M_n(K[X])$ also:

$$\sum_{l=1}^{n} \tilde{f}_{lj} * f_{il} = \sum_{l=1}^{n} f_{il} * \tilde{f}_{lj} = \big(F \cdot \mathrm{adj}(F)\big)_{ij} = \begin{cases} \chi_A & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

(Für die erste Gleichheit haben wir nur die Reihenfolge der Faktoren $\tilde{f}_{lj}$ und f_{il} getauscht.) Dies sind Gleichungen zwischen Polynomen in $K[X]$. Wir setzen A ein und erhalten

$$\sum_{l=1}^{n} \tilde{f}_{lj}(A) f_{il}(A) = \begin{cases} \chi_A(A) & \text{falls } i = j, \\ 0_{n\times n} & \text{falls } i \neq j. \end{cases}$$

Für $1 \leqslant i \leqslant n$ sei $e_i \in K^n$ der i-te Einheits-Spaltenvektor. Für festes $1 \leqslant j \leqslant n$ gilt dann

$$\begin{aligned}\chi_A(A)e_j &= \sum_{i=1}^{n} \underbrace{\Big(\sum_{l=1}^{n} \tilde{f}_{lj}(A)f_{il}(A)\Big)}_{=0_{n\times n} \text{ falls } i\neq j} e_i = \sum_{i=1}^{n}\sum_{l=1}^{n} \tilde{f}_{lj}(A)f_{il}(A)e_i \\ &= \sum_{l=1}^{n}\sum_{i=1}^{n} \tilde{f}_{lj}(A)f_{il}(A)e_i = \sum_{l=1}^{n} \tilde{f}_{lj}(A)\Big(\sum_{i=1}^{n} f_{il}(A)e_i\Big).\end{aligned}$$

Betrachten wir die letzte Summe. Für $i = l$ ist $f_{ll} = a_{ll} - X$ also $f_{ll}(A)e_l = a_{ll}e_l - Ae_l$. Für $i \neq l$ ist $f_{il} = a_{il}$ also $f_{il}(A)e_i = a_{il}e_i$. Damit erhalten wir

$$\sum_{i=1}^{n} f_{il}(A)e_i = (a_{ll}e_l - Ae_l) + \sum_{1\leqslant i\leqslant n, i\neq l} a_{il}e_i = \Big(\sum_{i=1}^{n} a_{il}e_i\Big) - Ae_l = 0_n$$

(weil Ae_l genau die l-te Spalte von A ist). Also ist $\chi_A(A)e_j = 0_n$ für $1 \leqslant j \leqslant n$. Damit sind alle Spalten von $\chi_A(A)$ gleich 0_n, also $\chi_A(A) = 0_{n\times n}$. □

Wir werden später (im Kapitel über Normalformen von Matrizen) noch einen völlig anderen Beweis des obigen Satzes sehen.

Beispiel 4.4.7. Sei $n \geqslant 1$ und $f = a_0 + a_1X + \ldots + a_nX^n \in K[X]$ beliebig mit $a_n = 1$. Dann ist die zugehörige (Frobenius-)***Begleitmatrix*** definiert als

$$A_f := \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix} \in M_n(K).$$

Ist $\{e_1, \ldots, e_n\}$ die Basis der Einheits-Spaltenvektoren in K^n, so gilt also

$$(*) \quad A_fe_n = -(a_0e_1 + a_1e_2 + \ldots + a_{n-1}e_n) \qquad \text{und} \qquad A_fe_i = e_{i+1} \quad \text{für } 1 \leqslant i \leqslant n-1.$$

Behauptung: $\tilde{\mu}_{A_f} = f$ und $\chi_A = (-1)^n f = \pm\tilde{\mu}_A$.
D.h., jedes Polynom mit höchstem Koeffizienten 1 ist Minimalpolynom einer Matrix!

Beweis. Zuerst bestimmen wir $\tilde{\mu}_A$. Aus $(*)$ folgt $A_fe_1 = e_2$, $A_f^2e_1 = A_f(A_fe_1) = A_fe_2 = e_3$, $A_f^3 = A_f(A_f^2e_1) = A_fe_3 = e_4$, usw., d.h., $A_f^ie_1 = e_{i+1}$ für $1 \leqslant i \leqslant n-1$. Sei nun

$$\tilde{\mu}_{A_f} = b_0 + b_1X + \ldots + b_{d-1}X^{d-1} + X^d \quad \text{mit } d \geqslant 1.$$

Wäre $d < n$, so erhielten wir

$$\begin{aligned}0_n = \tilde{\mu}_{A_f}(A_f)e_1 &= b_0e_1 + b_1A_fe_1 + \ldots + b_{d-1}A_f^{d-1}e_1 + A_f^de_1 \\ &= b_0e_1 + b_1e_2 + \ldots + b_{d-1}e_d + e_{d+1},\end{aligned}$$

Widerspruch dazu, dass $(e_1, e_2, \ldots, e_{d+1})$ linear unabhängig sind. Also ist $d \geqslant n$.

Jetzt genügt es zu zeigen, dass $f(A_f) = 0_{n\times n}$ gilt. Dazu müssen wir zeigen, dass $f(A_f)e_i = 0$ gilt für $1 \leqslant i \leqslant n$. Für $i = 1$ ist

$$\begin{aligned} f(A_f)e_1 &= a_0 I_n e_1 + a_1 A_f e_1 + \ldots + a_{n-1} A_f^{n-1} e_1 + A_f^n e_1 \\ &= a_0 e_1 + a_1 e_2 + \ldots + a_{n-1} e_n + A_f e_n = 0_n \quad \text{(siehe } (*)\text{)}. \end{aligned}$$

Für $2 \leqslant i \leqslant n$ ist $e_i = A_f^{i-1} e_1$, und damit

$$f(A_f)e_i = f(A_f)A_f^{i-1}e_1 = (fX^{i-1})(A_f)e_1 = (X^{i-1}f)(A_f)e_1 = A_f^{i-1}f(A_f)e_1 = 0_n.$$

Also ist $\tilde{\mu}_A = f$. Andererseits folgt aus dem Satz von Cayley–Hamilton auch $\chi_{A_f}(A_f) = 0_{n\times n}$. Wegen $\mathrm{Grad}(\chi_{A_f}) = n = \mathrm{Grad}(\tilde{\mu}_{A_f})$ folgt nun $\chi_{A_f} = c\tilde{\mu}_{A_f}$ mit $0 \neq c \in K$. Da der höchste Koeffizient von $\tilde{\mu}_{A_f}$ gleich 1 und derjenige von χ_{A_f} gleich $(-1)^n$ ist, folgt $c = (-1)^n$. □

Es ist eine gute Übung, direkt $\det(A_f - XI_n)$ mit Hilfe von geschickten Zeilen- und Spaltenumformungen zu berechnen (so wie im Fall der Determinante der Vandermonde-Matrix).