§4: Basiswechsel, Eigenvektoren und Determinanten

a) Eigenwerte und Eigenvektoren

Die Matrix einer linearen Abbildung $\varphi\colon V \to V$ bezüglich einer Basis $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$ ist genau dann eine Diagonalmatrix, wenn jeder der Basisvektoren $\vec{b}_i$ von φ auf ein Vielfaches $\lambda_i \vec{b}_i$ von sich selbst abgebildet wird; alsdann ist die Abbildungsmatrix gleich der Diagonalmatrix mit Einträgen $\lambda_1, \ldots, \lambda_n$. Somit hängt es sehr von der Basis ab, ob die Abbildungsmatrix Diagonalgestalt hat oder nicht.

Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 1 & -2 \\ 2 & 1 & -2 & 1 \\ 1 & -2 & 1 & 2 \\ -2 & 1 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

etwa ist ganz sicher keine Diagonalmatrix. Betrachten wir die lineare Abbildung

$$\varphi\colon \mathbb{R}^4 \to \mathbb{R}^4; \quad \vec{v} \mapsto A\vec{v}$$

aber bezüglich der Basis $\mathcal{B} = (\vec{b}_1, \vec{b}_2, \vec{b}_3, \vec{b}_4)$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad \vec{b}_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{b}_4 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix},$$

so rechnet man leicht nach, daß

$$\varphi(\vec{b}_1) = A\vec{b}_1 = 2\vec{b}_1, \quad \varphi(\vec{b}_2) = 2\vec{b}_2, \quad \varphi(\vec{b}_3) = -4\vec{b}_3 \quad \text{und} \quad \varphi(\vec{b}_4) = 4\vec{b}_4$$

ist, bezüglich $\mathcal{B}$ hat φ also die Diagonalmatrix

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

als Abbildungsmatrix. Wenn keine schwerwiegenden anderen Gründe dagegensprechen, wird es bei umfangreichen Rechnungen mit der Matrix A meist eine gute Idee sein, statt mit der Standardbasis von $\mathbb{R}^4$ mit der Basis $\mathcal{B}$ zu rechnen.

Wir werden im Verlauf dieser Vorlesung noch lernen, wie man viele Rechnungen mit der Matrix A auf das Rechnen mit der sehr viel angenehmeren Diagonalmatrix D zurückführen kann und wie man, wo dies möglich ist, eine Basis $\mathcal{B}$ wie im obigen Beispiel finden kann. Im Augenblick wollen wir uns damit begnügen, die dabei auftretenden Begriffe zu definieren:

Definition: Ein Vektor $\vec{v} \in k^n \setminus \{\vec{0}\}$ heißt *Eigenvektor* der Matrix $A \in k^{n\times n}$, wenn es eine Zahl $\lambda \in k$ gibt, so daß $A\vec{v} = \lambda\vec{v}$ ist. Dieses λ bezeichnen wir als einen *Eigenwert* von A.

Der seltsame Name *Eigenwert* läßt sich vielleicht am besten verstehen, wenn man seine Anwendung in der Quantenmechanik betrachtet: Dort werden *Observable,* d.h. physikalische Meßgrößen, durch Matrizen beschrieben und Zustände durch Vektoren. Mögliche Meßergebnisse sind die Eigenwerte; nach der Messung ist der Zustand des Systems ein Eigenvektor zum gemessenen Eigenwert.

Falls es also zu einer Matrix A eine Basis aus Eigenvektoren gibt, können wir sie also bezüglich dieser Basis als Diagonalmatrix darstellen.

Wir werden uns erst im nächsten Semester (im Zusammenhang mit Systemen linearer Differentialgleichungen) genauer mit Eigenwerten und Eigenvektoren befassen und dann auch Kriterien kennenlernen, wann eine Matrix diagonalisierbar ist und welche Möglichkeiten man hat, wenn es sie es nicht ist. In diesem Semester soll es nur darum gehen, die Eigenwerte und Eigenvektoren einer vorgegebenen Matrix zu bestimmen.

Die Gleichung $A\vec{v} = \lambda\vec{v}$ läßt sich umschreiben als

$$(A - \lambda E)\vec{v} = \vec{0},$$

das heißt, als ein homogenes lineares Gleichungssystem für den Vektor $\vec{v}$. Ein solches homogenes Gleichungssystem hat bekanntlich stets den Nullvektor als Lösung, der aber nach Definition genau aus diesem Grund kein Eigenvektor ist. Weitere Lösungen gibt es genau dann, wenn die Matrix $A - \lambda E$ des Gleichungssystems nicht den maximal möglichen Rang hat, wenn ihre Spaltenvektoren also linear abhängig sind. Wir wissen bereits, wie man mit Hilfe von Eliminationsschritten à la GAUSS den Rang einer Matrix bestimmen kann; als alternatives, bei kleinen Dimensionen gelegentlich einfacheres Verfahren, werden wir in diesem Paragraphen noch Determinanten kennenlernen.

Zunächst aber wollen wir uns überlegen, wie man Vektoren und Matrizen von einer Basis in eine andere umrechnet.

b) Beispiel eines Basiswechsels

In den meisten (endlichdimensionalen) Vektorräumen, die wir bislang betrachtet hatten, gab es offensichtliche Basen wie etwa die Standardbasis aus den Koordinateneinheitsvektoren in $\mathbb{R}^n$ oder die reinen Potenzen bei Vektorräumen von Polynomen. Diese Basen sind allerdings für rechnerische Zwecke nicht immer optimal: Sowohl in vielen Anwendungen der Mathematik als auch innerhalb der Mathematik ist es oft günstiger, mit anderen Basen zu rechnen, in denen beispielsweise wichtige Matrizen zu Diagonal- oder Dreiecksmatrizen werden. Das letzte Thema diesen Paragraphen wird sich mit dem Problem befassen, Matrizen und Vektoren bezüglich einer Basis in eine andere Basis umzurechnen.

In diesem Abschnitt betrachten wir zur Einstimmung einen Vektorraum, in dem es keine irgendwie ausgezeichnete Basis gibt.

Wir gehen aus von dem Problem, Farben quantitativ zu charakterisieren: Physikalisch gesehen hängen Farben mit den verschiedenen Wellenlängen des sichtbaren Lichts zusammen, eine Farbe müßte also eigentlich durch eine Funktion auf dem Intervall von etwa 380 bis etwa 780 nm beschrieben werden, wobei keineswegs nur stetige Funktionen in Betracht kommen.

Nun werden Farben aber nur selten mit dem Spektrometer betrachtet; was wirklich interessiert ist der Eindruck auf das menschliche Auge. Dieses hat vier Arten von Photorezeptoren: Die sehr empfindlichen Stäbchen, die nur bei schwachem Licht eine Rolle spielen und die auch nur Helligkeitsinformationen liefern können, sowie drei Arten von weniger lichtempfindlichen Zäpfchen, k, ℓ und m, die bei gutem Licht ein Bild mit Farbinformation liefern, da jede Art ihre eigene spektrale Empfindlichkeitskurve hat. Klassische wie auch digitale Photographie sowie Farbdarstellungen auf Fernseh- und Computermonitoren beruhen allesamt darauf, daß dem Auge etwas vorgesetzt wird, was die k, ℓ und m-Zäpfchen zur gleichen Reaktion veranlaßt wie das „echte“ Farbsignal.

Damit reicht es aus, Farben als Elemente eines dreidimensionalen Raums zu beschreiben. Da sich die Empfindlichkeitsbereiche insbesondere der ℓ- und der m-Zäpfchen stark überlappen, wäre es allerdings weder sinnvoll noch sonderlich praktikabel, eine Basis über die Ausgabewerte der drei Arten von Zäpfchen zu definieren. Stattdessen werden je nach Hauptanwendungsziel mehrere Farbmodelle betrachtet, die ausgehend von den unterschiedlichsten Ansätzen allesamt denselben dreidimensionalen $\mathbb{R}$-Vektorraum beschreiben. Wir wollen uns die beiden einfachsten etwas genauer anschauen.

Am bekanntesten ist wohl das RGB-Modell, das Farben aus den drei Grundfarben Rot, Grün und Blau kombiniert. Basis des Vektorraums sind hierbei also drei Vektoren $\vec{r}, \vec{g}$ und $\vec{b}$, die einem Rot, Grün *bzw.* Blau einer vorgegebenen Frequenz und Intensität entsprechen, und alle anderen Farben werden als Linearkombinationen

$$R\vec{r} + G\vec{g} + B\vec{b}$$

dargestellt. Diese Darstellung ist insbesondere gut geeignet für die Farbdarstellung auf einem Monitor oder auch Fernsehschirm, wo Farben in genau dieser Weise erzeugt werden. Bei digitalen Bildern betrachtet man i.a. nur Koeffizienten R, G, B aus dem Intervall [0, 1], so daß gewisse Grenzhelligkeiten nicht überschritten werden können. Bei der Darstellung mit einer Genauigkeit von acht Bit betrachtet man oft auch das 255-fache der entsprechenden Werte, gerundet zur nächsten ganzen Zahl.

Bei anderen Farbmodellen sieht man nicht so sehr auf die *Erzeugung* der Farben am Bildschirm, sondern auf deren Eigenschaften, wie sie ein menschlicher Betrachter wahrnimmt: Die Ausgabeimpulse der Zäpfchen werden noch im Sehapparat sofort weiterverarbeitet, und was wir bewußt zur Kenntnis nehmen, sind definitiv nicht die R-, G- und B-Anteile der Farben, sondern eher Helligkeiten, Farbsättigungen und ähnliche Eigenschaften.

Eine für die digitale Speicherung und Übermittlung interessante Tatsache ist, daß wir Helligkeiten viel feiner unterscheiden und viel höher auflösen können als Farbnuancen. Es liegt daher nahe, eine Basis zu wählen, in der auch die Gesamthelligkeit als Komponente auftritt, wobei

diese Komponente entweder mit höherer Geauigkeit digitalisiert wird als die beiden anderen oder (häufiger) nur diese Komponente für *jedes* Pixel gespeichert wird, die beiden anderen aber nur für jedes zweite Pixel oder gar nur einmal pro Quadrat aus 2×2 Pixel. Die Wahl der Helligkeit als Basiskomponente hat auch den Vorteil, daß man dann in einfacher Weise diese Komponente für Schwarz-Weiß-Versionen der Bilder verwenden kann, was zum Beispiel beim Fernsehen eine wichtige Rolle spielt.

Wenn wir die Helligkeit als eine Basiskomponente wählen, stehen für die eigentliche chromatische Information nur noch zwei Basisvektoren zur Verfügung; diese können entweder aus zwei Farbanteilen des RGB-Systems abgeleitet werden oder (was für die Gestaltung photorealistischer Bilder gern angewandt wird) aus einer weiteren achromatischen Komponente wie etwa der Farbsättigung und nur *einer* chromatischen Komponente.

Bei den drei derzeit gebräuchlichen Fernsehstandards PAL, SECAM und NTSC sowie auch beim HDTV und beim JPEG-Format für digitale Bilder geht man den ersten Weg: Der erste Basisvektor $\vec{y}$ beschreibt die Helligkeit oder *Luminanz*, die beiden weiteren Vektoren sind gewichtete Differenzen zwischen dieser Helligkeit und dem Rot- oder Blauanteil des Farbvektors. Die Gewichte unterscheiden sich dabei in den einzelnen Fällen; da für die Hörer dieser Vorlesung das JPEG-Format interessanter sein dürfte als Fernsehstandards, betrachten wir das dort verwendete YCbCr-Modell.

Die Helligkeit ist, wie in allen diesen Systemem, als gewichtetes Mittel der drei Farbanteile definiert; die Gewichte L_R, L_G und $L_B \in (0,\ 1)$ bezeichnet man als *Lumared, Lumagreen* und *Lumablue.* Mit diesen Bezeichnungen ist die Helligkeit

$$Y = L_R R + L_G G + L_B B \quad \text{mit} \quad L_R + L_G + L_B = 1\,;$$

Standardwerte sind

$$L_R = \frac{299}{1000}, \quad L_G = \frac{587}{1000} \quad \text{und} \quad L_B = \frac{114}{1000}\,.$$

Die beiden Chrominanzen sind festgelegt durch

$$C_b = \frac{B-Y}{2-2L_B} \quad \text{und} \quad C_r = \frac{R-Y}{2-2L_R}\,.$$

Damit haben wir eine neue Basis $(\vec{y}, \vec{c}_b, \vec{c}_r)$, mittels derer wir dieselben Farben beschreiben wie bezüglich der Basis $(\vec{r}, \vec{g}, \vec{b})$.

Ganz offensichtlich ist es wichtig, zwischen den beiden Basen hin- und herrechnen zu können, denn schließlich werden auch JPEG-Bilder auf RGB-Monitoren betrachtet, und CCD Chips in Digitalkameras messen zunächst einmal RGB-Komponenten, aus denen dann oft ein JPEG-Bild erzeugt wird.

Im vorliegenden Fall ist es einfach, konkrete Formeln zu finden:

$$\begin{aligned}
Y &= L_R R + L_G G + L_B B\,,\\
C_b &= \frac{B-Y}{2-2L_B} = -\frac{L_R}{2-2L_B}\cdot R - \frac{L_G}{2-2L_B}\cdot G + \frac{1}{2}\cdot B\,,\\
C_r &= \frac{R-Y}{2-2L_R} = \frac{1}{2}\cdot R - \frac{L_G}{2-2L_R}\cdot G - \frac{L_B}{2-2L_R}\cdot B\,.
\end{aligned}$$

Auch die Umkehrung läßt sich leicht ausrechnen:

$$\begin{aligned}
R &= Y + (2-2L_R)C_r\,,\\
B &= Y + (2-2L_B)C_b \qquad \text{und}\\
G &= \frac{Y - L_R R - L_B B}{L_G}\\
&= \frac{(1-L_R-L_B)Y - L_B(2-2L_B)C_b - L_R(2-2L_R)C_r}{L_G}\\
&= Y - \frac{L_B(2-2L_B)}{L_G}C_b - \frac{L_R(2-2L_R)}{L_G}C_r\,,
\end{aligned}$$

denn $1 - L_R - L_B = L_G$. Dies können wir auch mit Matrizen formulieren:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} L_R & L_G & L_B \\ \frac{-L_R}{2-2L_B} & \frac{-L_G}{2-2L_B} & \frac{1}{2} \\ \frac{1}{2} & \frac{-L_G}{2-2L_R} & \frac{-L_B}{2-2L_R} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

und

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2-2L_R \\ 1 & -\frac{L_B(2-2L_B)}{L_G} & -\frac{L_R(2-2L_R)}{L_G} \\ 1 & 2-2L_B & 0 \end{pmatrix} \begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} .$$

Der Übergang zwischen den beiden Basen kann also jeweils als Multiplikation mit einer Matrix interpretiert werden, und natürlich sind die beiden zugehörigen Matrizen invers zueinander.

Wenn wir die erste der beiden Matrizen mit A bezeichnen, so sind die Spalten von A die Bilder der Einheitsvektoren des RGB-Systems, umgerechnet ins YCbCr-System; wir erhalten also die Beziehungen

$$\begin{aligned} \vec{r} &= L_R\vec{y} - \frac{L_R}{2-2L_B}\vec{c}_b + \frac{1}{2}\vec{c}_r \\ \vec{g} &= L_G\vec{y} - \frac{L_G}{2-2L_B}\vec{c}_b - \frac{L_G}{2-2L_R}\vec{c}_r \quad \text{und} \\ \vec{b} &= L_B\vec{y} + \frac{1}{2}\vec{c}_b - \frac{L_B}{2-2L_R}\vec{c}_r . \end{aligned}$$

Entsprechend sind die Spalten von A^{-1} die Bilder der Einheitsvektoren des YCbCr-Systems, umgerechnet ins RGB-System, d.h.

$$\begin{aligned} \vec{y} &= \vec{r} + \vec{g} + \vec{b}, \\ \vec{c}_b &= -\frac{2-2L_B}{L_G}\vec{g} + (2-2L_B)\vec{b} \quad \textit{und} \\ \vec{c}_r &= (2-2L_R)\vec{r} - \frac{2-2L_B}{L_G}\vec{g} . \end{aligned}$$

Man beachte die Unterschiede zwischen der Darstellung der Basisvektoren in der jeweils anderen Basis und den Umrechnungsformeln für die Koeffizienten: Beim Umrechnen der RGB-Werte in YCbCr-Werte haben wir als Koeffizienten der einzelnen Gleichungen die *Zeilen* von A; die Basisvektoren $\vec{r}, \vec{g}, \vec{b}$ selbst sind aber im YCbCr-System ausgedrückt durch die *Spalten* der *inversen* Matrix A^{-1}.

Zur Verdeutlichung seien die Gleichungen nochmals angegeben mit numerischen Koeffizienten, näherungsweise berechnet für die Standard-

werte von L_R, L_G und L_B: Damit ist

$$A = \begin{pmatrix} 0{,}2990 & 0{,}5870 & 0{,}1140 \\ -0{,}1687 & -0{,}3313 & 0{,}5000 \\ 0{,}5000 & -0{,}4187 & -0{,}0813 \end{pmatrix}$$

und

$$A^{-1} = \begin{pmatrix} 1 & 0 & 1{,}4020 \\ 1 & -0{,}3441 & -0{,}7141 \\ 1 & 1{,}7720 & 0 \end{pmatrix} ,$$

also erhalten wir die Gleichungen

$$\begin{aligned} Y &= 0{,}2990\,R + 0{,}5870\,G + 0{,}1140\,B \\ Cb &= -0{,}1687\,R - 0{,}3313\,G + 0{,}5000\,B \\ Cr &= 0{,}5000\,R - 0{,}4187\,G - 0{,}08131\,B \end{aligned}$$

und

$$\begin{aligned} \vec{r} &= 0{,}2990\,\vec{y} - 0{,}1687\,\vec{c}_b + 0{,}5000\,\vec{c}_r \\ \vec{g} &= 0{,}5870\,\vec{y} - 0{,}3313\,\vec{c}_b - 0{,}4187\,\vec{c}_r \\ \vec{b} &= 0{,}1140\,\vec{y} + 0{,}5000\,\vec{c}_b - 0{,}08131\,\vec{c}_r . \end{aligned}$$

Entsprechend liefern Zeilen und Spalten von A^{-1} die Beziehungen

$$\begin{aligned} R &= Y + 1{,}402\,Cr \\ G &= Y - 0{,}3441\,Cb - 0{,}7141\,Cr \\ B &= Y + 1{,}772\,Cb \end{aligned}$$

und

$$\begin{aligned} \vec{y} &= \vec{r} + \vec{g} + \vec{b} \\ \vec{c}_b &= -0{,}3441\,\vec{g} + 10{,}772\,\vec{b} \\ \vec{c}_r &= 1{,}402\,\vec{r} - 0{,}7141\,\vec{g} . \end{aligned}$$

c) Basiswechsel im allgemeinen Fall

Wir gehen aus von einen n-dimensionalen k-Vektorraum V und betrachten darin die beiden Basen $\mathcal{B} = \left(\vec{b}_1, \ldots, \vec{b}_n\right)$ und $\mathcal{C} = \left(\vec{c}_1, \ldots, \vec{c}_n\right)$. Da

$\mathcal{B}$ eine Basis ist, lassen sich die Vektoren $\vec{c}_i$ als Linearkombinationen

$$\vec{c}_i = a_{i1}\vec{b}_1 + \cdots + a_{in}\vec{b}_n = \sum_{j=1}^{n} a_{ij}\vec{b}_j$$

der $\vec{b}_j$ schreiben; entsprechen lassen sich natürlich auch die $\vec{b}_j$ als Linearkombinationen

$$\vec{b}_j = m_{j1}\vec{c}_1 + \cdots + m_{jn}\vec{c}_n = \sum_{\ell=1}^{n} m_{j\ell}\vec{c}_\ell$$

schreiben. Dies können wir in die Darstellung von $\vec{c}_i$ einsetzen und erhalten

$$\vec{c}_i = \sum_{j=1}^{n} a_{ij}\vec{b}_j = \sum_{j=1}^{n} a_{ij} \sum_{\ell=1}^{n} m_{j\ell}\vec{c}_\ell = \sum_{\ell=1}^{n} \left(\sum_{j=1}^{n} a_{ij} m_{j\ell} \right) \vec{c}_\ell \,.$$

Da $\mathcal{C}$ eine Basis ist, bezüglich derer $\vec{c}_i$ die eindeutige Basisdarstellung $\vec{c}_i = 1 \cdot \vec{c}_i$ hat, folgt

$$\sum_{j=1}^{n} a_{ij} m_{j\ell} = \begin{cases} 1 & \text{falls } i = \ell \\ 0 & \text{falls } i \neq \ell \end{cases} .$$

Fassen wir die a_{ij} zu einer Matrix A zusammen und die m_{ij} zu einer Matrix M, ist also $AM = E$, d.h. die beiden Matrizen sind invers zueinander (was eigentlich niemanden erstaunen sollte). Formal können wir dies schreiben als

$$\begin{pmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{pmatrix} = A \begin{pmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_n \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_n \end{pmatrix} = A^{-1} \begin{pmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{pmatrix} ,$$

wobei die $\vec{b}_j$ und $\vec{c}_i$ bei der Auswertung dieser Formel nur als Symbole zu betrachten sind: Wir wollen nicht wirklich einen Vektor von Vektoren definieren, sondern diese Formel einfach als kompakte Merkregel für den Zusammenhang zwischen den beiden Basen betrachten.

Beim konkreten Rechnen in Vektorräumen geht es nicht so sehr um die Basisvektoren selbst, sondern um die Koeffizienten der Basisdarstellung. Sei also der Vektor $\vec{v} = v_1\vec{b}_1 + \cdots + v_n\vec{b}_n = w_1\vec{c}_1 + \cdots + w_n\vec{c}_n$ bezüglich

beider Basis dargestellt; wir suchen einen Zusammenhang zwischen den v_j und den w_i. Mit der obigen Matrix $M = (m_{ij})$ ausgedrückt ist

$$\vec{v} = \sum_{j=1}^{n} v_j\vec{b}_j = \sum_{j=1}^{n} v_j \sum_{i=1}^{n} m_{ji}\vec{c}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} v_j m_{ji} \right) \vec{c}_i \,,$$

also folgt wegen der Eindeutigkeit der Basisdarstellung

$$w_i = \sum_{j=1}^{n} v_j m_{ji} = \sum_{j=1}^{n} m_{ji} v_j \,.$$

Letzteres läßt sich leider nicht als Produkt von M mit dem Spaltenvektor der v_j schreiben: Die Indizes von m_{ji} haben die falsche Reihenfolge. Da die Formel trotzdem richtig und nützlich ist, definieren wir eine neue Matrix, die aus der alten durch Vertauschung der Indizes, d.h. also von Zeilen und Spalten, entsteht:

Definition: Die *transponierte* Matrix $N = {}^tM$ zur einer Matrix $M = (m_{ij})_{\substack{i=1\ldots n \\ j=1\ldots m}} \in k^{n\times m}$ ist jene Matrix $N = (n_{ij})_{\substack{i=1\ldots m \\ j=1\ldots n}} \in k^{m\times n}$ mit $n_{ij} = m_{ji}$ für alle i, j.

Somit braucht man zur Umrechnung der Koeffizienten ineinander also die transponierte Matrix zu $M = A^{-1}$. Diese Matrix bezeichnen wir als die *Matrix des Basiswechsels;* sie wird auch gelegentlich als die zu A *kontragrediente* Matrix bezeichnet. Für sie gilt:

Satz: Sind $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$ und $\mathcal{C} = (\vec{c}_1, \ldots, \vec{c}_n)$ zwei Basen eines n-dimensionalen k-Vektorraums und ist

$$\vec{c}_i = \sum_{j=1}^{n} a_{ij}\vec{b}_j \quad \text{mit} \quad A = (a_{ij}) \in k^{n\times n} ,$$

so berechnen sich für $\vec{v} = v_1\vec{b}_1 + \cdots + v_n\vec{b}_n = w_1\vec{c}_1 + \cdots + w_n\vec{c}_n$ die Koeffizienten w_i der Basisdarstellung bezüglich $\mathcal{C}$ aus den v_j nach der Formel

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = {}^t\!\left(A^{-1}\right) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} ;$$

die Matrix des Basiswechsels ist also ${}^t\big(A^{-1}\big)$. ■

Betrachten wir dazu ein Beispiel: Im Vektorraum $\mathcal{C}^0(\mathbb{R},\ \mathbb{R})$ aller stetiger Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ sei der Untervektorraum U erzeugt von den beiden Funktionen e^x und e^{-x}, die wir zur Basis $\mathcal{B}$ von U zusammenfassen. Eine weitere Basis $\mathcal{C}$ von U bestehe aus den Funktionen $\sinh x$ und $\cosh x$. Dann ist

$$\sinh x = \frac{1}{2}e^x - \frac{1}{2}e^{-x} \quad \text{und} \quad \cosh x = \frac{1}{2}e^x + \frac{1}{2}e^{-x}\,,$$

die Matrix A, mittels derer wir $\mathcal{C}$ durch $\mathcal{B}$ ausdrücken, ist also

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\,.$$

Die Berechnung der inversen Matrix ist hier nicht schwierig; man überzeugt sich leicht, daß gilt

$$A^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{und somit} \quad {}^t\big(A^{-1}\big) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\,.$$

Damit kennen wir die Matrix des Basiswechsels; ihr Produkt mit einem Koeffizientenvektor $\binom{a}{b} \in \mathbb{R}^2$ ist

$${}^t\big(A^{-1}\big)\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a-b \\ a+b \end{pmatrix}\,.$$

Also ist nach obigem Satz $ae^x + be^{-x} = (a-b)\sinh x + (a+b)\cosh x$, was man in der Tat leicht nachrechnet.

Als nächstes wollen wir uns überlegen, daß oben in ${}^t\big(A^{-1}\big)$ die Klammern überflüssig waren; genauer gilt

Lemma: Für $A \in k^{n\times m}$ und $B \in k^{m\times p}$ ist ${}^t(AB) = {}^tB\,{}^tA$ und ${}^t\big(A^{-1}\big) = \big({}^tA\big)^{-1}$.

Beweis: Mit $A = (a_{ij})$ und $B = (b_{j\ell})$ hat AB an der Stelle $i\ell$ den Eintrag $\sum_{j=1}^m a_{ij}b_{j\ell}$, und ihre Transponierte hat denselben Eintrag an der Stelle ℓi.

${}^tB$ hat an der Stelle ℓj den Eintrag $b_{j\ell}$ und ${}^tA$ hat an der Stelle ji den Eintrag a_{ij}, das Produkt ${}^tB\,{}^tA$ hat somit an der Stelle ℓi den Eintrag $\sum_{j=1}^m b_{j\ell}a_{ij}$, was wegen der Kommutativität der Multiplikation in k mit dem oben berechneten Ausdruck übereinstimmt.

Damit ist die erste Formel bewiesen. Wenden wir sie an auf $B = A^{-1}$, so folgt die Beziehung

$${}^t\big(A^{-1}\big)\,{}^tA = {}^t\big(AA^{-1}\big) = {}^tE = E\,,$$

die beiden Matrizen sind also invers zueinander. ■

Als letztes Thema im Zusammenhang mit Basiswechseln wollen wir uns überlegen, wie sich die Abbildungsmatrix einer linearen Abbildung unter Basiswechseln verhält.

Wir betrachten also eine lineare Abbildung $\varphi\colon V \to W$; dabei sei V ein n-dimensionaler k-Vektorraum, und W sei ein m-dimensionaler. In jedem der beiden Vektorräume seien zwei Basen gegeben; in V seien dies

$$\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n) \quad \text{und} \quad \mathcal{C} = (\vec{c}_1, \ldots, \vec{c}_n)\,,$$

in W seien es

$$\mathcal{D} = (\vec{d}_1, \ldots, \vec{d}_m) \quad \text{und} \quad \mathcal{E} = (\vec{e}_1, \ldots, \vec{e}_m)\,.$$

Um nicht ganz die Übersicht zu verlieren, ordnen wir jedem Vektor

$$\vec{v} = v_1\vec{b}_1 + \cdots + v_n\vec{b}_n = w_1\vec{c}_1 + \cdots + w_n\vec{c}_n$$

aus V die beiden Spaltenvektoren

$$\vec{v}_{\mathcal{B}} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{und} \quad \vec{v}_{\mathcal{C}} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

aus k^n zu; entsprechend haben wir für jeden Vektor $\vec{w} \in W$ zwei Vektoren $\vec{w}_{\mathcal{D}}$ und $\vec{w}_{\mathcal{E}}$ aus k^m.

Die Abbildungsmatrix $M_{\mathcal{B},\mathcal{D}}$ von φ bezüglich der Basen $\mathcal{B}$ von V und $\mathcal{D}$ von W hat dann die Eigenschaft, daß

$$\varphi(\vec{v})_{\mathcal{D}} = M_{\mathcal{B},\mathcal{D}}\vec{v}_{\mathcal{B}}$$

ist; für die Abbildungsmatrix $M_{\mathcal{C},\mathcal{E}}$ von φ bezüglich der Basen $\mathcal{C}$ von V und $\mathcal{E}$ von W gilt entsprechend

$$\varphi(\vec{v})_{\mathcal{E}} = M_{\mathcal{C},\mathcal{E}}\vec{v}_{\mathcal{C}} \,.$$

Für den Übergang von einer Basis zur anderen rechnen wir der Einfachheit halber nicht mit der zu Beginn dieses Abschnitts betrachteten Matrix, die die Basisvektoren durcheinander ausdrückt, sondern gleich mit der Matrix des Basiswechsels; die Matrizen $A \in k^{n\times n}$ und $B \in k^{m\times m}$ seien also so gewählt, daß für Vektoren $\vec{v} \in V$ und $\vec{w} \in W$ gilt

$$\vec{v}_{\mathcal{C}} = A\vec{v}_{\mathcal{B}} \quad \text{und} \quad \vec{w}_{\mathcal{E}} = B\vec{w}_{\mathcal{D}} \,.$$

Dann ist $\quad \varphi(\vec{v})_{\mathcal{E}} = B\varphi(\vec{v})_{\mathcal{D}} = BM_{\mathcal{B},\mathcal{D}}\vec{v}_{\mathcal{B}} = BM_{\mathcal{B},\mathcal{D}}A^{-1}\vec{v}_{\mathcal{C}} \,, \quad$ d.h.

$$M_{\mathcal{C},\mathcal{E}} = BM_{\mathcal{B},\mathcal{D}}A^{-1} \,.$$

Im nächsten Semester werden wir vor allen den Fall $V = W$ oft benötigen; hier wählt man natürlich fast immer $\mathcal{D} = \mathcal{B}$ und $\mathcal{C} = \mathcal{E}$, so daß jede Abbildungsmatrix von nur *einer* Basis abhängt. Dann ist auch $A = B$, und die obige Formel vereinfacht sich zu $M_{\mathcal{C}} = AM_{\mathcal{B}}A^{-1}$, wobei wir hier in $M_{\mathcal{B}} = M_{\mathcal{B},\mathcal{B}}$ den zweiten Index natürlich weglassen.

Betrachten wir als Beispiel die Differentiation im von e^x und e^{-x} erzeugten Vektorraum; ihre Abbildungsmatrix bezüglich $\mathcal{B}$ ist

$$M_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \,.$$

Die Matrix des Basiswechsels Basis $\mathcal{C}$ aus $\sinh x$ und $\cosh x$ ist, wie wir oben gesehen haben,

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad AM_{\mathcal{B}}A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \,,$$

ganz in Übereinstimmung mit der Tatsache, daß die Differentation $\sinh x$ zu $\cosh x$ macht und umgekehrt.

d) Forderungen an eine Determinante

Nächstes Ziel dieses Paragraphen ist die Entwicklung eines Kriteriums dafür, daß n Vektoren in $\mathbb{R}^n$ linear unabhängig sind. Um zu sehen, was wir da machen können, betrachten wir zunächst den Fall $n = 2$.

Hier sind zwei Vektoren $\vec{v} = \dfrac{v_1}{v_2}$ und $\vec{w} = \dfrac{w_1}{w_2}$ genau dann linear unabhängig, wenn $v_1w_2 - v_2w_1 \neq 0$ ist, denn wenn dieser Ausdruck verschwindet, ist $v_1w_2 = v_2w_1$, so daß im Fall $w_1 \neq 0$ gilt $\vec{v} = \dfrac{v_1}{w_1}\vec{w}$, und andernfalls verschwindet mit w_1 auch v_1 oder w_2, was ebenfalls in beiden Fällen die lineare Abhängigkeit von $\vec{v}$ und $\vec{w}$ impliziert.

Definieren wir die *Determinante* der beiden Vektoren $\vec{v}, \vec{w}$ als

$$\det(\vec{v}, \vec{w}) \underset{\text{def}}{=} v_1w_2 - v_2w_1 \,,$$

so gilt offensichtlich:

(D1) $\det(\vec{w}, \vec{v}) = -\det(\vec{v}, \vec{w})$

(D2) $\det(\vec{v}, \vec{w})$ ist linear in jedem ihrer beiden Argumente, d.h.

$$\det(\lambda\vec{v}_1 + \mu\vec{v}_2, \vec{w}) = \lambda\det(\vec{v}_1, \vec{w}) + \mu\det(\vec{v}_2, \vec{w})$$

und

$$\det(\vec{v}, \lambda\vec{w}_1 + \mu\vec{w}_2) = \lambda\det(\vec{v}, \vec{w}_1) + \mu\det(\vec{v}, \vec{w}_2) \,.$$

(D3) $\det(\vec{v}, \vec{w}) = 0$ genau dann, wenn $\vec{v}$ und $\vec{w}$ linear abhängig sind.

Wir möchten gerne für beliebige endlichdimensionale Vektorräume V eine entsprechende definieren, d.h. wir suchen für einen n-dimensionalen Vektorraum V über einem Körper k eine Abbildung

$$\det\colon \underbrace{V \times \cdots \times V}_{n \text{ mal}} \to k; \quad (\vec{v}_1, \ldots, \vec{v}_n) \mapsto \det(\vec{v}_1, \ldots, \vec{v}_n)$$

mit den Eigenschaften

(D1) $\det(\vec{v}_1, \ldots, \vec{v}_n)$ ändert ihr Vorzeichen, nicht aber ihren Betrag, wenn irgendwelche zwei ihrer Argumente miteinander vertauscht werden.

(D2) $\det(\vec{v}_1, \ldots, \vec{v}_n)$ ist linear in jedem ihrer n Argumente.

(D3) $\det(\vec{v}_1, \ldots, \vec{v}_n) = 0$ genau dann, wenn $\vec{v}_1, \ldots, \vec{v}_n$ linear abhängig sind.

Um zu zeigen, daß es eine solche Abbildung auch tatsächlich gibt, folgen wir einer auch sonst bei Existenzbeiweisen oftmals nützlichen Strategie: Wir nehmen an, wir *hätten* bereits eine Funktion det mit den Eigenschaften (D1) bis (D3), und versuchen dann, diese Funktion

aufgrund dieser Eigenschaften möglichst explizit auszurechnen. Dies wird auf eine Formel für det führen, die wir dann als *Definition* nehmen können, wenn wir nachweisen, daß sie (D1) bis (D3) erfüllt.

Ausgangspunkt ist eine Basis $(\vec{e}_1, \ldots, \vec{e}_n)$ von V; für diese ist wegen Eigenschaft (D3) dann $\det(\vec{e}_1, \ldots, \vec{e}_n) \neq 0$. Über den genauen Wert wissen wir nichts, denn mit det erfüllt offensichtlich auch jedes skalare Vielfache $\lambda \cdot \det$ mit $\lambda \neq 0$ die Forderungen (D1) bis (D3).

Wir betrachten nun n beliebige Vektoren $\vec{v}_1, \ldots, \vec{v}_n$ und schreiben diese bezüglich der Basis $(\vec{e}_1, \ldots, \vec{e}_n)$ als Linearkombinationen

$$\vec{v}_i = \sum_{j=1}^{n} a_{ij} \vec{e}_j$$

der gewählten Basisvektoren.

Sukzessive Anwendung der Linearitätsregel (D2) auf die n Argumente von det führt auf

$$\begin{aligned}
\det(\vec{v}_1, \ldots, \vec{v}_n) &= \det\Big(\sum_{j_1=1}^{n} a_{1j_1} \vec{e}_{j_1}, \vec{v}_2, \ldots, \vec{v}_n\Big) \\
&\underset{(D2)}{=} \sum_{j_1=1}^{n} a_{1j_1} \det(\vec{e}_{j_1}, \vec{v}_2, \ldots, \vec{v}_n) \\
&= \sum_{j_1=1}^{n} a_{1j_1} \det\Big(\vec{e}_{j_1}, \sum_{j_2=1}^{n} a_{2j_2} \vec{e}_{j_2}, \vec{v}_3, \ldots, \vec{v}_n\Big) \\
&\underset{(D2)}{=} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{1j_1} a_{2j_2} \det(\vec{e}_{j_1}, \vec{e}_{j_2}, \vec{v}_3, \ldots, \vec{v}_n) \\
&\quad \vdots \\
&= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} a_{2j_2} \cdot \cdots \cdot a_{nj_n} \det(\vec{e}_{j_1}, \cdots, \vec{e}_{j_n}) \,.
\end{aligned}$$

Um dies noch weiter ausrechnen zu können, müssen wir die Vektoren $\vec{e}_{j_\nu}$ in jedem der Summanden in die natürliche Reihenfolge bringen; dabei

hilft die Regel (D1). Eine ganze Reihe von Summanden können wir allerdings gleich von vornherein außer Acht lassen, denn nach (D3) ist

$$\det(\vec{e}_{j_1}, \cdots, \vec{e}_{j_n}) = 0 \,,$$

wenn (mindestens) zwei der $\vec{e}_{j_\nu}$ übereinstimmen: Falls $\vec{e}_{j_\nu}$ und $\vec{e}_{j_\mu}$ übereinstimmen, ist $\vec{e}_{j_\nu} - \vec{e}_{j_\mu} = \vec{0}$ eine nichttriviale Darstellung des Nullvektors. Ein von Null verschiedener Wert ist also nur möglich, wenn $\{j_1, \ldots, j_n\} = \{1, \ldots, n\}$ ist, wenn also die Abbildung

$$\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}; \quad \nu \mapsto j_\nu$$

bijektiv ist.

Unsere nächste Aufgabe wird daher sein, solche Abbildungen zu studieren und sie auf Vertauschungen zweier Elemente zurückzuführen, so daß wir schließlich die Summanden mittels (D1) auf $\det(\vec{e}_1, \ldots, \vec{e}_n)$ zurückführen können.

e) Gerade und ungerade Permutationen

Erinnern wir uns an §3*j4):* Dort hatten wir bijektive Abbildungen

$$\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$$

als Permutationen bezeichnet; unsere nächste Aufgabe besteht also darin für eine Permutation π nachzurechnen, für welches Vorzeichen

$$\det(\vec{e}_{\pi(1)}, \ldots, \vec{e}_{\pi(n)}) = \pm \det(\vec{e}_1, \ldots, \vec{e}_n)$$

ist. In einem Spezialfall wissen wir das schon: Ist nämlich die Permutation π eine *Transposition,* vertauscht also genau zwei Elemente und bildet die restlichen Elemente auf sich selbst ab, so ist das Vorzeichen nach Regel (D1) ein Minuszeichen.

Der folgende Satz erlaubt es uns, den allgemeinen Fall darauf zurückzuführen:

Satz: Jede Permutation π kann als Produkt

$$\pi = \tau_1 \circ \cdots \circ \tau_r$$

von Transpositionen geschrieben werden.

Beweis: Da es sich hier um eine Mathematikvorlesung für Informatiker handelt, möchte ich sowohl einen mathematischen als auch einen informatischen Beweis geben.

Der einfachste mathematische Beweis benutzt vollständige Induktion nach der Elementanzahl n der zu permutierenden Menge. Für $n = 1$ gibt es keine Permutation außer der Identität, die man als leeres Produkt von Transpositionen betrachten kann; da aber nicht jeder diese Art von logischen Taschenspielertricks liebt, betrachten wir zur Vorsicht auch noch den Fall $n = 2$ als weitere Verankerung der Induktion.

Hier gibt es genau zwei Permutationen: Die Identität, die alles festläßt, und die Vertauschung der beiden Elemente. Letzere ist eine Transposition, erstere kann wieder als leeres Produkt von Transpositionen betrachtet werden oder, falls man das nicht möchte, als Quadrat der Vertauschung, denn nach zweimaligem Vertauschen ist wieder alles beim alten.

Beim *Induktionsschritt* nehmen wir an, wir hätten den Satz für Permutationen $(n-1)$-elementiger Mengen bewiesen und betrachten nun eine Permutation

$$\pi\colon \{1,\dots,n\} \to \{1,\dots,n\}\,.$$

Für diese gibt es ein Element $i \in \{1,\dots,n\}$, so daß $\pi(i) = n$ ist, denn π ist ja bijektiv. Die Permutation

$$\pi' = \pi \circ (i\ n)$$

bildet n ab auf

$$\pi'(n) = \big(\pi \circ (i\ n)\big)(n) = \pi\big((i\ n)(n)\big) = \pi(i) = n\,,$$

d.h. n bleibt fest. Da π' bijektiv ist, muß es daher auch die Menge $\{1,\dots,n-1\}$ auf sich selbst abbilden; wenn wir n ignorieren, können wir π' daher auch als eine Permutation von $\{1,\dots,n-1\}$ auffassen. Von der wissen wir, daß sie als Produkt

$$\pi' = \tau_1 \circ \cdots \circ \tau_r$$

von Transpositionen darstellbar ist. Wegen $\pi' = \pi \circ (i\ n)$ und weil Transpositionen zu sich selbst invers sind, ist damit auch

$$\pi = \pi' \circ (i\ n) = \tau_1 \circ \cdots \circ \tau_r \circ (i\ n)$$

als Produkt von Transpositionen darstellbar, und damit ist der Satz einmal bewiesen.

Zum zweiten Beweis, auf Grundlage der Informatik, führt die Beobachtung, daß jeder gängige Sortieralgorithmus, der auf Vertauschungen von Elementen beruht (und das tut fast jeder) ein konstruktives Verfahren liefert, um Permutationen als Produkte von Transpositionen zu schreiben: Soll π zerlegt werden, so sortiere man durch fortgesetztes Vertauschungen zweier Zahlen die Folge

$$\pi(1),\ \pi(2),\ \pi(3),\ \dots,\pi(n-1),\ \pi(n)$$

der Größe nach; die erste angewandte Transposition sei τ_1, die letzte τ_n.

Die sortierte Folge der $\pi(i)$ ist natürlich $1,\dots,n$ ist; daher für jedes i

$$(\tau_r \circ \cdots \circ \tau_1 \circ \pi)(i) = i\,,$$

d.h.

$$\tau_r \circ \cdots \circ \tau_1 \circ \pi = \text{Identität}\,.$$

Multipliziert man beide Seiten von links mit $\tau_1 \circ \cdots \circ \tau_r$, so folgt

$$\pi = \tau_1 \circ \cdots \circ \tau_r\,,$$

denn

$$\begin{aligned}
&(\tau_1 \circ \cdots \circ \tau_r) \circ (\tau_r \circ \cdots \circ \tau_1)\\
&= \tau_1 \circ \cdots \circ \tau_{r-1} \circ (\tau_r \circ \tau_r) \circ \tau_{r-1} \circ \cdots \circ \tau_1\\
&= \tau_1 \circ \cdots \circ \tau_{r-2} \circ (\tau_{r-1} \circ \tau_{r-1}) \circ \tau_{r-2} \circ \cdots \circ \tau_1\\
&= \tau_1 \circ \cdots \circ \tau_{r-3} \circ (\tau_{r-2} \circ \tau_{r-2}) \circ \tau_{r-3} \cdots \circ \tau_1\\
&= \cdots = \tau_1 \circ (\tau_2 \circ \tau_2) \circ \tau_1 = \tau_1 \circ \tau_1 = \text{Identität}\,.
\end{aligned}$$

Also ist π als Produkt der Transpositionen τ_1 bis τ_r darstellbar. ■

Ein Leser, der mit den gebräuchlichen Sortierverfahren vertraut ist, wird unschwer erkennen, daß der „mathematische“ Beweis ein Spezialfall des informatischen ist, in dem als Sortierverfahren ein relativ einfaches $O(n^2)$-Verfahren benutzt wurde (das allerdings zumindest für einstellige n den meisten „besseren“ Sortierverfahren überlegen sein dürfte). Die beiden Beweise unterscheiden sich also nicht sonderlich.

Die Anzahl n der benötigten Vertauschungen hängt natürlich von der Vorgehensweise und der Effiziens des gewählten Sortierverfahrens ab; die Darstellung von π als Produkt von Transpositionen ist also alles andere als eindeutig. Da wir für die Determinantenberechnung vor allem wissen müssen, ob die Abzahl der Vertauschungen gerade oder ungerade ist, wollen wir uns als nächstes davon überzeugen, daß wenigstens dies, d.h. also die Parität der Anzahl r der benötigten Transpositionen, eindeutig bestimmt ist: r ist für gegebenes π entweder *immer* gerade oder *immer* ungerade.

Dazu definieren wir für jede Permutation $\pi \in \mathfrak{S}_n$ ein Vorzeichen:

Definition: Das *Vorzeichen* der Permutation $\pi \in \mathfrak{S}_n$ ist die Zahl

$$\varepsilon(\pi) \underset{\text{def}}{=} \prod_{\{i,j\}\subset\{1,\dots,n\}} \frac{\pi(j)-\pi(i)}{j-i}\,,$$

wobei das Produkt über die sämtlichen zweielementigen Teilmengen von $\{1,\dots,n\}$ genommen wird.

$\varepsilon(\pi)$ ist wirklich ein Vorzeichen, das heißt gleich ± 1, denn

$$|\varepsilon(\pi)| = \frac{\prod\limits_{\{i,j\}\subset\{1,\dots,n\}} |\pi(j)-\pi(i)|}{\prod\limits_{\{i,j\}\subset\{1,\dots,n\}} |j-i|} = 1\,,$$

weil mit $\{i,j\}$ wegen der Bijektivität von π auch $\{\pi(i),\pi(j)\}$ die sämtlichen zweielementigen Teilmengen von $\{1,\dots,n\}$ durchläuft, so daß im Zähler und Nenner bis auf Reihenfolge genau dieselben Faktoren stehen.

Für eine Transposition $\tau = (k\ \ell)$ ist

$$\frac{\tau(j)-\tau(i)}{j-i} = \frac{j-i}{j-i} = 1 \quad \text{falls} \quad \{i,j\}\cap\{k,l\} = \emptyset\,,$$

da $(k\ \ell)$ alle Zahlen außer k und ℓ festläßt. Die Zahlen k und ℓ werden vertauscht, daher ist

$$\frac{\tau(k)-\tau(\ell)}{k-\ell} = \frac{\ell-k}{k-\ell} = -1\,.$$

Bleibt noch der Fall, daß eine der beiden Zahlen i oder j gleich k oder ℓ ist; da wir zweielementige *Mengen* betrachten und $\{i,j\} = \{j,i\}$ ist, können wir vereinbaren, daß diese Zahl i sein soll.

Ist $i = k$, so ist das Produkt der Terme für $\{k,j\}$ und $\{\ell,j\}$ gleich

$$\frac{\tau(k)-\tau(j)}{k-j}\cdot\frac{\tau(\ell)-\tau(j)}{\ell-j} = \frac{\ell-j}{k-j}\cdot\frac{k-j}{\ell-j} = +1\,;$$

und genau entsprechend kann man auch für $i = \ell$ argumentieren.

Bildet man daher das Produkt über alle zweielementigen Teilmengen von $\{1,\dots,n\}$, so erhält man $\varepsilon(\tau) = -1$ und damit das

Lemma: Das Vorzeichen einer Transposition ist -1. ∎

Der Zusammenhang zwischen dem Vorzeichen und der Transpositionsdarstellung einer Permutation ergibt sich aus

Lemma: Für zwei Permutationen $\pi, \pi' \in \mathfrak{S}_n$ ist

$$\varepsilon(\pi\circ\pi') = \varepsilon(\pi)\cdot\varepsilon(\pi')\,.$$

Beweis: Nach Definition ist

$$\begin{aligned}\varepsilon(\pi\circ\pi') &= \prod_{\{i,j\}\subset\{1,\dots,n\}} \frac{(\pi\circ\pi')(j)-(\pi\circ\pi')(i)}{j-i}\\ &= \prod_{\{i,j\}\subset\{1,\dots,n\}} \frac{\pi\big(\pi'(j)\big)-\pi\big(\pi'(i)\big)}{j-i}\,.\end{aligned}$$

Nach Erweiterung mit

$$\prod_{\{i,j\}\subset\{1,\dots,n\}} \frac{\pi'(j)-\pi'(i)}{\pi'(j)-\pi'(i)}$$

und Umordnung wird dies zu

$$\varepsilon(\pi\circ\pi') = \prod_{\{i,j\}} \frac{\pi\big(\pi'(j)\big)-\pi\big(\pi'(i)\big)}{\pi'(j)-\pi'(i)}\cdot\prod_{\{i,j\}}\frac{\pi'(j)-\pi'(i)}{j-i}\,.$$

Das zweite dieser Produkte ist natürlich $\varepsilon(\pi')$, und das erste ist $\varepsilon(\pi)$, da mit $\{i, j\}$ auch $\{\pi'(i), \pi'(j)\}$ die sämtlichen zweielementigen Teilmengen von $\{1, \ldots, n\}$ durchläuft. ■

Sind nun

$$\pi = \tau_1 \circ \cdots \circ \tau_r = \sigma_1 \circ \cdots \circ \sigma_s$$

zwei Darstellungen einer Permutation π als Produkt von Transpositionen, so ist nach den gerade bewiesenen Lemmata einerseits

$$\varepsilon(\pi) = \prod_{i=1}^{r} \varepsilon(\tau_i) = (-1)^r,$$

andererseits aber auch

$$\varepsilon(\pi) = \prod_{i=1}^{s} \varepsilon(\sigma_i) = (-1)^s,$$

also muß $(-1)^r = (-1)^s$ sein und damit r, s entweder beide gerade oder beide ungerade.

Definition: Eine Permutation $\pi \in \mathfrak{S}_n$ heißt *gerade,* wenn sie als Produkt einer geraden Anzahl von Transpositionen geschrieben werden kann, ansonsten heißt sie *ungerade.*

Das Vorzeichen $\varepsilon(\pi)$ von π ist also $+1$ für gerades π und -1 für ungerades.

f) Existenz von Determinanten

Nach diesem Einschub über Permutationen und Transpositionen haben wir das nötige Rüstzeug, um die Rechnung aus Abschnitt *c)* zu Ende zu führen:

Da die Determinante bei jeder Vertauschung zweier Elemente ihr Vorzeichen wechselt, folgt sofort aus den obigen Betrachtungen, daß für jede Permutation $\pi \in \mathfrak{S}_n$ gilt

$$\det\left(\vec{e}_{\pi(1)}, \ldots, \vec{e}_{\pi(n)}\right) = \begin{cases} +\det(\vec{e}_1, \ldots, \vec{e}_n) & \text{falls } \pi \text{ gerade} \\ -\det(\vec{e}_1, \ldots, \vec{e}_n) & \text{falls } \pi \text{ ungerade} \end{cases},$$

d.h. $\det\left(\vec{e}_{\pi(1)}, \ldots, \vec{e}_{\pi(n)}\right) = \varepsilon(\pi) \cdot \det(\vec{e}_1, \ldots, \vec{e}_n)$. Wenn wir die Rechnung aus Abschnitt *b)* fortführen, erhalten wir daher

$$\begin{aligned} \det\left(\vec{v}_1, \ldots, \vec{v}_n\right) &= \sum_{\pi \in \mathfrak{S}_n} a_{1\pi(1)} \cdots a_{n\pi(n)} \cdot \det\left(\vec{e}_{\pi(1)}, \ldots, \vec{e}_{\pi(n)}\right) \\ &= \left(\sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)} \right) \cdot \det\left(\vec{e}_1, \ldots, \vec{e}_n\right). \end{aligned}$$

Somit ist $\det\left(\vec{v}_1, \ldots, \vec{v}_n\right)$ eindeutig bestimmt bis auf den Wert von $\det\left(\vec{e}_1, \ldots, \vec{e}_n\right)$; falls es überhaupt Determinanten gibt, müssen sie also so aussehen. Wir definieren daher

Definition: Die Determinante von n Vektoren

$$\vec{v}_1 = a_{11}\vec{e}_1 + \cdots + a_{1n}\vec{e}_n, \quad \ldots, \quad \vec{v}_n = a_{n1}\vec{e}_1 + \cdots + a_{nn}\vec{e}_n$$

des n-dimensionalen Vektorraums V bezüglich der Basis $\vec{e}_1, \ldots, \vec{e}_n$ ist

$$\det\left(\vec{v}_1, \ldots, \vec{v}_n\right) = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)}.$$

Insbesondere ist also $\det(\vec{e}_1, \ldots, \vec{e}_n) = 1$.

Satz: Die so definierte Determinante hat die Eigenschaften (D1) bis (D3).

Beweis: Beginnen wir mit (D1), d.h. die Determinante wechselt ihr Vorzeichen, wenn zwei der Argumente vertauscht werden. Vertauscht man etwa die Argumente $\vec{v}_i$ und $\vec{v}_j$, so werden anstelle der Summanden

$$\varepsilon(\pi)\, a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{j\pi(j)} \cdots a_{n\pi(n)}$$

die Summanden

$$\varepsilon(\pi)\, b_{1\pi(1)} \cdots b_{j\pi(j)} \cdots b_{i\pi(i)} \cdots b_{n\pi(n)}$$

aufaddiert, wobei

$$b_{k\ell} = \begin{cases} a_{k\ell} & \text{falls } k \neq i, j \\ a_{j\ell} & \text{falls } k = i \\ a_{i\ell} & \text{falls } k = j \end{cases}$$

ist. Die Summanden sind also

$$\begin{aligned}
&\varepsilon(\pi)\, a_{1\pi(1)} \cdots a_{j\pi(i)} \cdots a_{i\pi(j)} \cdots a_{n\pi(n)} \\
&= \varepsilon(\pi)\, a_{1\pi(1)} \cdots a_{i\pi(j)} \cdots a_{j\pi(i)} \cdots a_{n\pi(n)} \\
&= \varepsilon(\pi)\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)}
\end{aligned}$$

mit

$$\pi'(k) = \begin{cases} \pi(k) & \text{falls } k \neq i, j \\ \pi(j) & \text{falls } k = i \\ \pi(i) & \text{falls } k = j\,, \end{cases}$$

d.h. $\pi' = \pi \circ (i\ j)$. Als Produkt von Transpositionen geschrieben hat π' daher einen Faktor mehr als π; falls π ungerade war, ist also π' gerade, und umgekehrt. Somit ist

$$\varepsilon(\pi') = -\varepsilon(\pi)\,,$$

und damit lassen sich die Summanden auch schreiben als

$$-\varepsilon(\pi')\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)}\,.$$

Summiert man über alle $\pi' \in \mathfrak{S}_n$, ist also

$$\begin{aligned}
&\det(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_i, \ldots, \vec{v}_n) \\
&= \sum_{\pi' \in \mathfrak{S}_n} \big(-\varepsilon(\pi')\big)\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)} \\
&= -\sum_{\pi' \in \mathfrak{S}_n} \varepsilon(\pi')\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)} \\
&= -\det(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_j, \ldots, \vec{v}_n)\,,
\end{aligned}$$

womit (D1) bewiesen wäre.

Eigenschaft (D2), die Linearität in jedem der Argumente, ist klar, denn jedes der Monome $a_{1\pi(1)} \cdots a_{n\pi(n)}$ enthält *genau eine* Komponente des Vektors $\vec{v}_i$ und ist damit linear in $\vec{v}_i$.

Bleibt noch (D3), d.h. $\det\big(\vec{v}_1, \ldots, \vec{v}_n\big)$ soll genau dann verschwinden, wenn die $\vec{v}_i$ linear abhängig sind.

Sind zunächst die Vektoren $\vec{v}_1, \ldots, \vec{v}_n$ linear abhängig, so läßt sich einer von ihnen als Linearkombination der anderen schreiben; durch

Umordnen (wobei die Determinante nach der schon bewiesenen Eigenschaft (D1) höchstens ihr Vorzeichen ändert) können wir annehmen, daß dies etwa $\vec{v}_1$ sei; wir setzen

$$\vec{v}_1 = \sum_{i=2}^{n} \lambda_i \vec{v}_i\,.$$

Dann ist

$$\begin{aligned}
\det\big(\vec{v}_1, \ldots, \vec{v}_n\big) &= \det\left(\sum_{i=2}^{n} \lambda_i \vec{v}_i, \vec{v}_2, \ldots, \vec{v}_n\right) \\
&\underset{(D2)}{=} \sum_{i=2}^{n} \lambda_i \det\big(\vec{v}_i, \vec{v}_2, \ldots, \vec{v}_n\big)\,.
\end{aligned}$$

Im i-ten Summanden der letzten Summe tritt der Vektor $\vec{v}_i$ zweimal als Argument von det auf: Einmal an der ersten, und dann noch an der i-ten. Vertauscht man diese beiden (gleichen) Argumente, ändert sich natürlich nichts; andererseits haben wir aber bereits (D1) bewiesen, wonach sich beim Vertauschen irgend zweier Argumente das Vorzeichen ändert. Der i-te Summand hat also einen Wert s_i, für den $s_i = -s_i$ ist. Falls wir mit reellen Zahlen arbeiten, folgt hieraus natürlich sofort, daß alle s_i verschwinden und somit auch ihre Summe.

Für beliebige Körper gilt das aber leider nicht: Aus $s_i = -s_i$ folgt, wenn wir auf beiden Seiten s_i addieren, zunächst nur, daß

$$s_i + s_i = (1+1)s_i = 0$$

ist, und daraus folgt genau dann, daß auch $s_i = 0$ ist, wenn $1 + 1 \neq 0$ ist. Dies ist aber beispielweise in dem für die digitale Informationsverarbeitung wichtigen Körper $\mathbb{F}_2 = \{0, 1\}$ nicht der Fall; wir müssen uns also mit Rücksicht auf diesen (und eine ganze Reihe anderer) Körper noch überlegen, daß wirklich immer $\det(\vec{v}_1, \ldots, \vec{v}_n) = 0$ ist, wenn zwei der Argumente $\vec{v}_i$ übereinstimmen.

Dazu beachten wir, daß im Falle $\vec{v}_i = \vec{v}_j$ auch für alle ℓ die Koeffizienten $a_{i\ell}$ und $a_{j\ell}$ übereinstimmen, d.h. in der Summe

$$\det A = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{j\pi(j)} \cdots a_{n\pi(n)}$$

ist stets

$$a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{j\pi(j)} \cdots a_{n\pi(n)} = a_{1\pi(1)} \cdots a_{i\pi(j)} \cdots a_{j\pi(i)} \cdots a_{n\pi(n)} .$$

Hier ist die rechte Seite gerade der Summand zur Permutation $\pi \circ (i\ j)$, die genau dann gerade ist, wenn π ungerade ist, und umgekehrt. Also haben beide Terme verschiedenes Vorzeichen und heben sich gegenseitig weg; die Summe über alle Permutationen ist daher Null.

Somit verschwindet die Determinante unabhängig vom Grundkörper immer dann, wenn die Vektoren linear abhängig sind.

Nun seien $\vec{v}_1, \ldots, \vec{v}_n$ linear unabhängig; für dieses Fall müssen wir zeigen, daß die Determinante *nicht* verschwindet. Wir wissen, daß für die Basisvektoren $\vec{e}_1, \ldots, \vec{e}_n$

$$\det\big(\vec{e}_1, \ldots, \vec{e}_n\big) = 1 \neq 0$$

ist; außerdem wissen wir, daß n beliebige linear unabhängige Vektoren eines n-dimensionalen Vektorraums eine Basis bilden.

Daher können wir die $\vec{e}_i$ als Linearkombinationen der $\vec{v}_j$ schreiben, etwa als

$$\vec{e}_i = \sum_{j=1}^{n} b_{ij} \vec{v}_j .$$

Nun können wir die Rechnungen der letzten Paragraphen wörtlich wiederholen, nur daß dieses Mal $\vec{v}_1, \ldots, \vec{v}_n$ die Basisvektoren sind und $\vec{e}_1, \ldots, \vec{e}_n$ die Vektoren, für die wir det berechnen möchten. Wir wissen zwar noch nicht, daß $\det\big(\vec{v}_1, \ldots, \vec{v}_n\big)$ nicht verschwindet, aber das haben wir in der dortigen Rechnung auch nie verwendet: Alles hing nur ab von den bereits bewiesenen Forderungen (D1) und (D2) und der ebenfalls schon bewiesenen Tatsache, daß det verschwindet, wenn zwei der Argumente gleich sind. Somit erhalten wir die Formel

$$\det\big(\vec{e}_1, \ldots, \vec{e}_n\big) = \left(\sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) b_{1\pi(1)} \cdots b_{n\pi(n)} \right) \cdot \det\big(\vec{v}_1, \ldots, \vec{v}_n\big) .$$

In dieser Gleichung steht links die Zahl eins, also kann keiner der beiden Faktoren auf der rechten Seite verschwinden; insbesondere ist $\det\big(\vec{v}_1, \ldots, \vec{v}_n\big) \neq 0$.

Damit ist der Satz vollständig bewiesen. ■

Fassen wir zusammen, was wir bislang in diesem Paragraphen gemacht haben:

- Wir haben anhand des Beispiels von Determinanten im $\mathbb{R}^2$ und $\mathbb{R}^3$ drei zentrale Forderungen an eine allgemeine Determinante aufgestellt.
- Unter der Annahme, daß es Funktionen gibt, die diese drei Forderungen erfüllen, haben wir ausgerechnet, wie solche Funktionen aussehen müssen; inbsbesondere haben wir gesehen, daß sie durch ihren Wert auf den Vektoren einer Basis eindeutig bestimmt sind.
- Dann haben wir die so erhaltene explizite Formel als *Definition* einer allgemeinen Determinanten genommen und gezeigt, daß die so definierte Funktion tatsächlich die drei Forderungen erfüllt.

Insgesamt haben wir also gesehen, daß es Funktionen gibt, die den drei Forderungen (D1) bis (D3) genügen und daß solche Funktionen bis auf eine multiplikative Konstante eindeutig bestimmt sind.

g) Die Determinante einer Matrix

Definition: Unter der Determinante einer Matrix $A \in k^{n \times n}$ verstehen wir die Determinante ihrer Spaltenvektoren; ist also

$$A = (a_{ij}) \quad \text{und} \quad \vec{v}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} ,$$

so ist

$$\det A \underset{\text{def}}{=} \det\big(\vec{v}_1, \ldots, \vec{v}_n\big) .$$

Wir schreiben auch kurz

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} .$$

Um etwas Übung im Umgang mit Determinanten zu bekommen, wollen wir dies für $n = 1$ und die beiden bereits bekannten Fälle $n = 2$ und $n = 3$ noch einmal explizit ausrechnen – auch wenn wir bereits wissen, was herauskommen muß.

Der Fall $n = 1$ einer 1×1-Matrix ist völlig uninteressant: Die einzige Permutation einer einelementigen Menge ist die Identität, die für jedes n gerade ist; für $A = (a) \in k^{1\times 1}$ ist also $\det A = a$. Die Schreibweise $\det A = |a|$ ist hier irreführend, da es zumindest für $k = \mathbb{R}$ oder $\mathbb{C}$ zu Verwechslungen mit der Betragsfunktion kommen kann. Da Determinanten von 1×1-Matrizen völlig uninteressant sind, ist dies jedoch nicht weiter schlimm.

Für $n = 2$ gibt es zwei Permutationen: Die (gerade) Identität und die (ungerade) Transposition $(1\ 2)$. Nach Definition der (allgemeinen) Determinanten ist

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}\,,$$

wie erwartet.

Für $n = 3$ gibt es bereits $3! = 6$ Permutationen, darunter natürlich die (gerade) Identität.

Jede weitere Permutation kann höchstens ein Element festlassen, denn würde sie zwei festlassen, müßte wegen der Bijektivität der Abbildung auch das dritte auf sich selbst abgebildet werden, wir hätten also die Identität.

Falls ein Element festgehalten wird, bleibt einer nichtidentischen Permutation daher nichts anderes übrig, als die beiden anderen zu vertauschen; dies führt auf die drei (ungeraden) Transpositionen $(2\ 3)$, $(1\ 3)$ und $(1\ 2)$.

Falls kein Element auf sich selbst abgebildet wird, muß die Eins entweder auf 2 oder auf 3 abgebildet werden. Im ersten Fall geht dann 2 auf 3, denn sonst müßte zwei entweder festbleiben oder wir hätten die Transposition $(1\ 2)$, die drei auf sich selbst abbildet. Wegen der Bijektivität der Abbildung gibt es dann keine andere Möglichkeit als daß drei auf eins abgebildet wird; wir haben also die zyklische Vertauschung

$$1 \mapsto 2 \mapsto 3 \mapsto 1\,.$$

Diese Permutation ist gerade, denn sie läßt sich beispielsweise als Produkt $(1\ 3)(1\ 2)$ zweier Transpositionen schreiben: Wenden wir dieses Produkt an auf 1, so wird die Eins zunächst von $(1\ 2)$ auf zwei abgebildet, und zwei bleibt fest unter $(1\ 3)$, so daß insgesamt eins auf zwei abgebildet wird, wie verlangt.

Zwei wird von $(1\ 2)$ auf die Eins abgebildet, und die wiederum von $(1\ 3)$ auf drei, also haben wir auch hier insgesamt das richtige Ergebnis. Drei schließlich kann dann wegen der Bijektivität von $(1\ 3)(1\ 2)$ nur auf eins abgebildet werden, was man auch schnell direkt sieht, denn $(1\ 2)$ läßt die Drei unverändert, während $(1\ 3)$ sie auf eins abbildet.

Die noch verbleibende Permutation bildet eins auf drei und daher drei auf zwei ab, sie ist also die zyklische Vertauschung

$$1 \mapsto 3 \mapsto 2 \mapsto 1 \quad \text{oder} \quad 3 \mapsto 2 \mapsto 1 \mapsto 3\,,$$

d.h. die Umkehrabbildung zur gerade betrachteten Permutation. Daher ist sie als Produkt von Transpositionen gleich

$$\big((1\ 2)(1\ 3)\big)^{-1} = (1\ 3)^{-1}(1\ 2)^{-1} = (1\ 3)(1\ 2)$$

und somit auch gerade.

Zur Berechnung der Determinanten einer 3×3-Matrix (a_{ij}) beginnen wir mit den geraden Permutationen: Die Identität liefert den Term $a_{11}a_{22}a_{33}$, d.h. das Produkt der drei Diagonalelemente, die zyklische Vertauschung $1 \to 2 \to 3 \to 1$ liefert das Produkt $a_{12}a_{23}a_{31}$, und ihr Inverses $a_{13}a_{21}a_{32}$.

Die drei Transpositionen führen auf die negativ zu nehmenden Produkte $a_{11}a_{23}a_{32}$, $a_{13}a_{22}a_{31}$ und $a_{12}a_{21}a_{33}$, insgesamt ist also

$$\begin{aligned} \det A &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\qquad - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}\,, \end{aligned}$$

genau wie wir es in Abschnitt *b)* über das Spatprodukt berechnet hatten und uns aufgrund der SARRUSschen Regel leicht merken können.

Für die Determinanten von 4×4-Matrizen gibt es keine entsprechende Regel mehr. Da nun bereits $4! = 24$ Summanden berücksichtigt werden müssen, eignet sich die Formel aus der Definition der Determinanten nicht mehr gut zur Berechnung, von noch größeren Matrizen ganz zu schweigen: Wie wir im nächsten Kapitel sehen werden, ist $\log n! \approx n \log n$, d.h. die Anzahl $n!$ der Summanden wächst mit größer werdendem n noch schneller als die Exponentialfunktion.

Wir werden im folgenden daher auf Verfahren hinarbeiten, die es erlauben, auch größere Determinanten mit vertretbarem Aufwand zu berechnen.

Wir kennen schon einige Rechenregeln für Determinanten als Funktionen von n Vektoren, beispielsweise die Forderungen (D1) bis (D3), aus denen wir die Determinantendefinition hergeleitet haben oder auch die gerade angewandte Regel

$$\begin{aligned}
&\det\left(\vec{v}_1, \ldots, \vec{v}_i + \lambda \vec{v}_j, \ldots, \vec{v}_n\right) \\
&= \det\left(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_n\right) + \lambda \det\left(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_n\right) \\
&= \det\left(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_n\right) .
\end{aligned}$$

In diesem Abschnitt soll es nun um Rechenregeln für Determinanten von Matritzen gehen.

Am wichtigsten ist der *Multiplikationssatz:*

Satz: *a)* Für zwei Matrizen $A, B \in k^{n \times n}$ ist $\det(AB) = \det A \cdot \det B$ *b)* Für eine invertierbare Matrix $A \in k^{n \times n}$ ist $\det(A^{-1}) = (\det A)^{-1}$.

Beweis: a) Die Determinante von AB ist die Determinante der Spaltenvektoren $\vec{v}_1, \ldots, \vec{v}_n$ von AB, die von B ist gleich der Determinante der Spaltenvektoren $\vec{b}_1, \ldots, \vec{b}_n$ von B, und nach Definition der Matrizenmultiplikation ist mit $A = (a_{ij})$

$$\vec{v}_i = \sum_{j=1}^{n} a_{ij} \vec{b}_j .$$

Damit sind wir in der Situation der Abschnitte *b)* und *d)* (wobei hier die $\vec{b}_j$ die Rolle der dortigen $\vec{e}_j$ übernehmen) und können aus den Eigenschaften (D1) bis (D3) folgern, daß

$$\det\left(\vec{v}_1, \ldots, \vec{v}_n\right) = \left(\sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)} \right) \cdot \det\left(\vec{b}_1, \ldots, \vec{b}_n\right)$$

ist. Das ist aber genau die Formel

$$\det(AB) = (\det A) \cdot (\det B) .$$

b) Für eine invertierbare Matrix A ist $A \cdot A^{-1} = E$ die Einheitsmatrix, und deren Determinante ist gleich der Determinante der Basisvektoren in ihrer natürlichen Reihenfolge, also 1. Daher ist nach Teil *a)*

$$\det A \cdot \det\left(A^{-1}\right) = \det E = 1 ,$$

woraus die Behauptung folgt. ■

Der Multiplikationssatz läßt sich in ein Berechnungsverfahren für Determinanten übersetzen: Falls wir die LR-Zerlegung $A = LR$ der Matrix A kennen, sagt uns der Satz, daß

$$\det A = \det L \cdot \det R$$

ist. Die Determinanten von L und R lassen sich leicht ausrechnen nach dem folgenden

Lemma: Die Determinante einer (unteren oder oberen) Dreiecksmatrix ist gleich dem Produkt ihrer Diagonalelemente.

Beweis: $A = (a_{ij})$ sei eine Dreiecksmatrix. Wie üblich ist

$$\det A = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)} .$$

Nun ist aber für eine untere *bzw.* obere Dreiecksmatrix $a_{ij} = 0$ falls $i < j$ *bzw.* $i > j$ ist. Da Permutationen bijektive Abbildungen sind, muß es für jede Permutation außer der Identität mindestens ein i geben mit $\pi(i) < i$ und mindestens ein j mit $\pi(j) > j$, d.h. das Produkt $a_{1\pi(1)} \cdots a_{n\pi(n)}$ enthält für jede nichtidentische Permutation mindestens einen Faktor

Null. Damit kann höchstens die Identität einen von Null verschiedenen Beitrag zur Summe liefern; da sie Vorzeichen +1 hat ist also

$$\det A = a_{11} \cdots a_{nn}$$

das Produkt der Diagonalelemente von A.

■

So wie wir die LR-Zerlegung definiert haben, ist L eine untere Dreiecksmatrix mit lauter Einsen in der Hauptdiagonale und R eine beliebige obere Dreiecksmatrix. Also ist $\det L = 1$, die Determinante von R ist gleich dem Produkt der Diagonalelemente von R, und damit ist auch

$$\det A = \text{Produkt der Diagonalelemente von } R\,.$$

Zumindest für große n ermöglicht dies eine erheblich effizientere Berechnung von Determinanten als die definierende Formel: Der Aufwand für eine LR-Zerlegung ist ungefähr proportional zu n^3, was für große n weitaus günstiger ist als der überexponentiell steigende Aufwand proportional $n \cdot n!$ für die Summe aus der definierenden Formel. Für die Praxis wichtig ist, daß wir die Matrix L nicht wirklich kennen müssen: Ihre Determinante ist auf jeden Fall eins. Daher reicht es, den Algorithmus für die LR-Zerlegung nur auf die Matrix A selbst anzuwenden; wir müssen sie nicht auch noch die Einheitsmatrix mitschleppen. Im wesentlichen geht es also einfach um GAUSS-Elimination. Dies wird auch klar anhand der folgenden Regeln, die man üblicherweise an Stelle der LR-Zerlegung anwendet:

Eine Determinante ändert ihren Wert nicht, wenn man ein Vielfaches einer Spalte zu einer anderen Spalte addiert: Die Determinante einer Matrix ist nach Definition die Determinante der Spaltenvektoren, und wenn man das λ-fache der j-ten Spalte zur i-ten Spalte addiert, ist

$$\begin{aligned}
&\det\big(\vec{v}_1, \ldots, \vec{v}_i + \lambda \vec{v}_j, \ldots, \vec{v}_n\big) \\
&= \det\big(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_n\big) + \lambda \det\big(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_n\big) \\
&= \det\big(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_n\big)\,,
\end{aligned}$$

da der zweite Summand den Vektor $\vec{v}_j$ doppelt enthält und somit verschwindet.

Etwas entsprechendes gilt auch für Zeilen; das könnten wir zwar direkt nachrechnen, aber wenn wir uns daran erinnern, daß die Zeilen einer Matrix gleich den Spalten der transponierten Matrix sind, folgt das viel bequemer aus dem folgenden

Lemma: Für jede $n \times n$-Matrix A ist $\det A = \det {}^tA$.

Beweis: Nach Definition ist für $A = (a_{ij})$

$$\det A = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)}$$

und

$$\det {}^tA = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n}\,.$$

Ordnet man die Faktoren des Produkts $a_{\pi(1)1} \cdots a_{\pi(n)n}$ nach dem ersten Index, so erhält man

$$a_{\pi(1)1} \cdots a_{\pi(n)n} = a_{1\pi^{-1}(1)} \cdots a_{1\pi^{-1}(n)}$$

und damit

$$\det {}^tA = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi^{-1}(1)} \cdots a_{1\pi^{-1}(n)}\,.$$

Da mit π auch π^{-1} die gesamte Gruppe $\mathfrak{S}_n$ durchläuft und $\varepsilon(\pi) = \varepsilon(\pi^{-1})$ ist, stimmt die Summe dieser Terme mit der aus der Formel für $\det A$ überein, die beiden Determinanten sind also gleich.

■

Für Leser, denen nicht klar ist, warum $\varepsilon(\pi)$ und $\varepsilon(\pi^{-1})$ gleich sind, seien zur Übung des Umgangs mit Permutationen hier kurz zwei Beweise angegeben: Am einfachsten sieht man diese Formel aufgrund des obigen Lemmas, wonach für zwei beliebige Permutationen π und π' gilt: $\varepsilon(\pi \circ \pi') = \varepsilon(\pi) \cdot \varepsilon(\pi')$. Setzt man hier $\pi' = \pi^{-1}$, so folgt

$$1 = \varepsilon(\text{Identität}) = \varepsilon(\pi) \cdot \varepsilon(\pi^{-1})\,,$$

so daß die beiden Vorzeichen gleich sein müssen.

Alternativ können wir uns auch explizit überlegen, daß sich π^{-1} als Produkt von r Transpositionen schreiben läßt, falls dies für π der Fall

ist; genauer gilt: Ist $\pi = \tau_1 \circ \cdots \circ \tau_r$, so ist $\pi^{-1} = \tau_r \circ \cdots \circ \tau_1$, denn wie wir bereits im Abschnitt über Permutationen gesehen haben, ist

$$\pi \circ (\tau_r \circ \cdots \circ \tau_1) = (\tau_1 \circ \cdots \circ \tau_r) \circ (\tau_r \circ \cdots \circ \tau_1) = \text{Identität},$$

da jede Transposition zu sich selbst invers ist.

Aus obigem Lemma folgt insbesondere, daß man die Determinante einer Matrix auch als Determinante der Zeilenvektoren definieren kann und daß die Spaltenvektoren einer Matrix genau dann linear unabhängig sind, wenn es die Zeilenvektoren sind. Mit nur wenig mehr Aufwand könnte man so auch zeigen, daß der im Zusammenhang mit der Lösung linearer Gleichungssysteme definierte *Rang* einer Matrix genausogut über Zeilen- wie über Spaltenoperationen berechnet werden kann.

Das obige Lemma zeigt auch die folgende Rechenregel:

Eine Determinante ändert ihren Wert nicht, wenn man ein Vielfaches einer Zeile zu einer anderen Zeile addiert, denn die Zeilen einer Matrix sind die Spalten der transponierten Matrix.

Ebenso zeigt man: *Eine Determinante wird mit -1 multipliziert, wenn man entweder zwei Zeilen oder zwei Spalten miteinander vertauscht,* denn für Spalten ist das gerade die definierende Eigenschaft (D1) der Determinanten.

Als letzte Regel sei noch aufgeführt: *Eine Determinante wird mit λ multipliziert, wenn* eine *ihrer Zeilen oder* eine *ihrer Spalten mit λ multipliziert werden.* Für Spalten folgt das sofort aus der Linearität in jedem Argument, für Zeilen folgt es daraus nach Transponieren.

h) Der Entwicklungssatz von Laplace

Hier geht es um ein Berechnungsverfahren, das die definierenden Eigenschaften der Determinante ausnutzt, die „Entwicklung nach einer Zeile oder Spalte".

Für die Entwicklung einer Matrix beispielsweise nach der j-ten Spalte

schreiben wir den j-ten Spaltenvektor

$$\vec{v}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

als Linearkombination der Einheitsvektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

d.h. als

$$\vec{v}_j = a_{1j}\vec{e}_1 + \cdots + a_{nj}\vec{e}_n .$$

Dann ist wegen der Linearität der Determinante in ihrem j-ten Argument

$$\begin{aligned} \det A &= \det\bigl(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_n\bigr) = \det\bigl(\vec{v}_1, \ldots, \sum_{i=1}^{n} a_{ij}\vec{e}_i, \ldots, \vec{v}_n\bigr) \\ &= \sum_{i=1}^{n} a_{ij} \det\bigl(\vec{v}_1, \ldots, \vec{e}_i, \ldots, \vec{v}_n\bigr) . \end{aligned}$$

Um daraus die Determinante von A zu berechnen, müssen wir die Determinanten

$$D_{ij} \underset{\text{def}}{=} \det\bigl(\vec{v}_1, \ldots, \vec{e}_i, \ldots, \vec{v}_n\bigr)$$

kennen, wobei $\vec{e}_i$ an der j-ten Stelle steht.

Dazu wenden wir die definierende Formel an: In den Produkten $a_{1\pi(1)} \cdots a_{n\pi(n)}$ ist für die gewählte Stelle j, an der wir den Vektor $\vec{e}_i$ eingefügt haben, $a_{j\ell} = 0$ für $\ell \neq i$ und $a_{ji} = 1$. Also verschwindet $a_{1\pi(1)} \cdots a_{n\pi(n)}$ für jede Permutation π mit $\pi(j) \neq i$.

Am einfachsten läßt sich dies im Fall $i = j = n$ weiter ausrechnen: Dann müssen wir nur Permutationen mit $\pi(n) = n$ betrachten, und die entsprechen eineindeutig den Permutationen auf der Menge aller Zahlen von 1 bis $n-1$; was übrigbleibt ist also (wegen $a_{nn} = 1$) gerade die Determinante jener $(n-1) \times (n-1)$-Matrix, die durch Streichung der letzten Spalte und der letzten Zeile aus A entsteht.

Der allgemeine Fall läßt sich durch Vertauschungen darauf zurückführen: Wir können zwar nicht einfach die j-te Spalte mit der n-ten vertauschen, denn dann steht ja die n-te Spalte an j-ter Stelle, aber wir können die j-te Spalte durch sukzessive Vertauschung mit ihrem rechten Nachbarn immer weiter nach außen bringen: Wir vertauschen zunächst die j-te Spalte mit der $(j+1)$-ten, dann die neue $(j+1)$-te (=alte j-te) Spalte mit der $(j+2)$-ten *usw.,* bis schließlich die alte j-te Spalte an n-ter Stelle steht, ohne daß sich vor oder nach der Stelle j irgendetwas an der relativen Reihenfolge der Spalten geändert hätte. Die Anzahl der dazu notwendigen Vertauschungen ist $(n-j)$, die Determinante wurde bei der gesamten Prozedur also mit $(-1)^{n-j}$ multipliziert.

Als nächstes müssen wir auch noch die i-te Zeile nach unten bringen; dazu verwenden wir die gerade gezeigte Formel $\det A = \det {}^tA$. Die i-te Zeile von A ist gleich der i-ten Spalte von ${}^tA$; wie wir oben gesehen haben, läßt sie sich durch $(n-i)$ Spaltenvertauschungen zur letzen Spalte machen, wobei die Determinante mit $(-1)^{n-i}$ multipliziert wird, und durch nochmaliges Transponieren erhalten wir schließlich jene Matrix, in der die i-te Zeile zur n-ten geworden ist, ohne daß sich vor oder nach der Stelle i irgendetwas an der relativen Reihenfolge der Zeilen geändert hätte.

Die Determinante der so entstehenden Matrix ist, wie wir oben beim Spezialfall $i = j = n$ gesehen haben, gleich der Determinante jener $(n-1)\times(n-1)$-Matrix A_{ij}, die durch Streichen der letzten Zeile und Spalte entsteht; bezogen auf die ursprüngliche Matrix A entsteht sie durch Streichen der i-ten Zeile und der j-ten Spalte. Also ist, wenn wir noch die Vertauschungen berücksichtigen,

$$D_{ij} = (-1)^{n-i}\cdot(-1)^{n-j}\cdot\det A_{ij} = (-1)^{i+j}\det A_{ij}\,,$$

denn $(n-i)+(n-j) = 2n-(i+j)$ hat dieselbe Parität wie $(i+j)$. Fassen wir alles zusammen, erhalten wir die Formel

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

Die Entwicklung nach einer Zeile geht ganz entsprechend: Da die i-te Zeile von A gleich der i-ten Spalte von ${}^tA$ ist, führt obige Rechnung für

die transponierte Matrix auf die Formel

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

Damit haben wir den folgenden Satz bewiesen, der trotz seines Namens bereits LEIBNIZ bekannt war:

Entwicklungssatz von Laplace: A sei eine $n\times n$-Matrix und A_{ij} sei jene $(n-1)\times(n-1)$-Matrix, die durch Streichung der i-ten Zeile und der j-ten Spalte von A entsteht. Dann ist für jedes feste j

$$\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij}$$

und für jedes feste i

$$\det A = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij}\,,$$

■

Die erste Formel aus diesem Satz bezeichnet man als die *Entwicklung der Determinante nach ihrer j-ten Spalte,* die zweite (die durch Anwendung der ersten auf die transponierte Matrix entsteht) entsprechend als die *Entwicklung nach der i-ten Zeile.*

PIERRE-SIMON DE LAPLACE (1749–1827) war einer der bedeutendsten französischen Wissenschaftler seiner Zeit; berühmt sind vor allem seine Anwendungen der Analysis auf die Wahrscheinlichkeitstheorie, die Himmelsmechanik, die Potentialtheorie und sein Vergleich des Universums mit einem Uhrwerk. Bekannt wurde er auch durch die nach ihm und KANT benannte Theorie zur Entstehung des Universums. Als politischer Opportunist kam er gut durch die Wirren seiner Zeit; er saß unter anderem in dem Kommitee, das Maßeinheiten nach dem Dezimalsystem einführte, war kurze Zeit Innenminister und wurde schließlich sogar Graf und Marquis.

BARON GOTTFRIED WILHELM VON LEIBNIZ (1646–1716) gilt als der letzte Universalgelehrte, der das gesamte Wissen seiner Zeit überblickte. In der Mathematik ist er vor allem berühmt durch die Entwicklung der Infinitesimalrechnung (bezüglich derer es einen langen Prioritätsstreit mit NEWTON gab); Bezeichnungen wie $\frac{dy}{dx}$ und $\int f(x)\,dx$ gehen auf ihn zurück. Durch seine Begründung der symbolischen Logik legte er auch einen wesentlichen Grundstein der späteren Informatik. Weitere Arbeiten befassen sich mit den Naturwissenschaften und der Technik, der Philosophie, Theologie und der Geschichte.

Als erste Anwendung wollen wir noch einmal die SARRUSsche Regel für Determinanten von 3×3-Matrizen beweisen: Entwicklung nach der ersten Spalte zeigt, daß

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) \\ &\quad + a_{31}(a_{12}a_{23} - a_{13}a_{22})\end{aligned}$$

ist, was ausmultipliziert natürlich genau die alte Formel ergibt.

Rein theoretisch ist es gleichgültig, nach welcher Zeile oder Spalte man eine Determinante entwickelt, aber man wird natürlich versuchen, eine zu finden, in der möglichst viele Nullen stehen, so daß in obiger Summe möglichst wenige Summanden übrigbleiben. In der Tat überlegt man sich leicht, daß die Anwendung des LAPLACEschen Entwicklungssatzes auf eine vollbesetzte Determinante wie bei der definierenden Formel rund $n \cdot n!$ Recheenoperationen erfordert und man somit gegenüber der direkten Anwendung dieser Formel nur den Vorteil hat, daß die Rechenoperationen klarer gegliedert werden.

Interessant zur praktischen Berechnung von Determinanten, vorzugsweise für nicht allzu große n, ist die Formel daher nur, wenn man sie mit den Rechenregeln aus dem vorigen Abschnitt kombiniert, um sich Nullen und einfache Einträge zu verschaffen. Da dieselbe Art von Operationen auch zur LR-Zerlegung führen, wird dann der Übergang zwischen

LAPLACEschem Entwicklungssatz und Berechnung via LR-Zerlegung fließend.

Betrachten wir als Beispiel die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 2 & -2 & 1 \\ -1 & -3 & 2 & -2 \\ -4 & -2 & -1 & 0 \\ 3 & 2 & -1 & 1 \end{pmatrix}.$$

Sie hat zwar einen Eintrag Null, aber da in der dritten und vierten Spalte gleich in zwei Zeilen die Einträge entgegengesetzt gleich sind, ist es besser, zunächst die dritte Spalte durch die Summe von dritter und vierter Spalte zu ersetzen:

$$\begin{vmatrix} 1 & 2 & -2 & 1 \\ -1 & -3 & 2 & -2 \\ -4 & -2 & -1 & 0 \\ 3 & 2 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 1 \\ -1 & -3 & 0 & -2 \\ -4 & -2 & -1 & 0 \\ 3 & 2 & 0 & 1 \end{vmatrix}.$$

Wenn wir jetzt nach der dritten Spalte entwickeln, kommen nur zwei der Summanden wirklich vor: Der erste ($i = 1$ und $j = 3$) und der dritte ($i = j = 3$). In beiden Fällen ist $i + j$ gerade, der Vorfaktor $(-1)^{i+j}$ ist also jeweils +1. Die 3×3-Determinanten, die durch Streichung der dritten Spalte und der ersten *bzw.* dritten Zeile entstehen, können nach der Regel von SARRUS berechnet werden:

$$\begin{vmatrix} -1 & -3 & -2 \\ -4 & -2 & 0 \\ 3 & 2 & 1 \end{vmatrix} = (2 + 0 + 16) - (12 + 0 + 12) = -6$$

und

$$\begin{vmatrix} 1 & 2 & 1 \\ -1 & -3 & -2 \\ 3 & 2 & 1 \end{vmatrix} = (-3 - 12 - 2) - (-9 - 4 - 2) = -2\,.$$

Also ist $\det A = (-1)(-6) + (-1)(-2) = 8$.

Auch als Hilfmittel zur Berechnung von abstrakt gegebenen Determinanten ist der LAPLACEsche Entwicklungssatz oft nützlich, da er es bei geschickter Anwendung gelegentlich erlaubt, eine Rekursionsformel zu finden und damit eine allgemeine Formel für einen gewisse Klasse von Determinanten.

Als Beispiel betrachten wir die unter anderem in der Numerik wichtige VANDERMONDEsche Determinante

$$V(a_1,\dots,a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix} .$$

Der Franzose ALEXANDRE THÉOPHILE VANDERMONDE (1735–1796) war zunächst Musiker; erst im Alter von 35 Jahren begann er sich für Mathematik zu interessieren und publizierte in den Jahren 1771 und 1772 vier Arbeiten über Gleichungen, Determinanten und über das Problem, einen Springer so über ein Schachbrett zu bewegen, daß er jedes Feld genau einmal betritt. Die VANDERMONDEsche Determinante ist nirgends in seinem publizierten Werk zu finden; sie wurde erst um 1935 von HENRI LEBESGUE (1875–1941) nach ihm benannt.

Für die Anwendung des LAPLACEschen Entwicklungssatzes auf diese Determinante bietet sich an, nach der ersten Spalte zu entwickeln, denn diese besteht aus lauter Einsen, und die $(n-1)\times(n-1)$-Determinanten, die nach dem Entwicklungssatz auftreten, sind im wesentlichen wieder VANDERMONDEsche Determinanten. Allerdings entstehen dabei n Summanden, und für jeden von diesen muß die ganz Prozedur wiederholt werden *usw.*– die Berechnung der Determinante auf diese Weise ist also zumindest ziemlich aufwendig.

Ein besserer Ansatz ergibt sich, wenn wir die Einsen in der ersten Spalte dadurch ausnutzen, daß wir etwa die erste Zeile von jeder anderen Zeile subtrahieren. Der Wert der Determinante ändert sich dadurch nicht, d.h.

$$V(a_1,\dots,a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 0 & a_2-a_1 & a_2^2-a_1^2 & \dots & a_2^{n-1}-a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_n-a_1 & a_n^2-a_1^2 & \dots & a_n^{n-1}-a_1^{n-1} \end{vmatrix} .$$

Wenn wir hier nach der ersten Spalte entwickeln, muß nur eine einzige $(n-1)\times(n-1)$-Determinante berücksichtigt werden, alle anderen

haben den Vorfaktor null. Also ist

$$V(a_1,\dots,a_n) = \begin{vmatrix} a_2-a_1 & a_2^2-a_1^2 & \dots & a_2^{n-1}-a_1^{n-1} \\ a_3-a_1 & a_3^2-a_1^2 & \dots & a_3^{n-1}-a_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n-a_1 & a_n^2-a_1^2 & \dots & a_n^{n-1}-a_1^{n-1} \end{vmatrix}$$

gleich jeder Determinanten, die durch Streichung der ersten Spalte und der ersten Zeile entsteht.

Hier können wir in jeder Zeile die jeweils vorne stehende Differenz ausklammern, denn genau wie

$$x^k - 1 = (x-1)(x^{k-1} + x^{k-2} + \dots + x + 1)$$

durch $(x-1)$ teilbar ist, ist auch

$$a_i^k - a_1^k = (a_i - a_1)(a_i^{k-1} + a_i^{k-2}a_1 + a_i^{k-3}a_1^2 + \dots + a_i a_1^{k-2} + a_1^{k-1})$$

durch $(a_i - a_1)$ teilbar; den Quotienten schreiben wir kurz als $q_{i,k-1}$:

$$q_{i,k-1} \underset{\text{def}}{=} a_i^{k-1} + a_i^{k-2}a_1 + a_i^{k-3}a_1^2 + \dots + a_i a_1^{k-2} + a_1^{k-1} .$$

Wegen der Linearität der Determinante können wir jeden Faktor, den wir aus einer Zeile (oder Spalte) ausklammern, vor die Determinante ziehen und erhalten für $V(a_1,\dots,a_n)$ somit den Wert

$$(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1) \begin{vmatrix} 1 & q_{21} & q_{22} & \dots & q_{2,n-2} \\ 1 & q_{31} & q_{32} & \dots & q_{3,n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & q_{n1} & q_{n2} & \dots & q_{n,n-2} \end{vmatrix} .$$

Die Nützlichkeit dieser Formel steht und fällt damit, daß wir die q_{ij} gut miteinander in Verbindung bringen können. Für verschiedene Indizes i haben die entsprechenden Ausdrücke offensichtlich wenig miteinander zu tun; sie enthalten nicht einmal dieselben Variablen. Schreiben wir allerdings

$$\begin{aligned} q_{ij} &= a_i^j + a_i^{j-1}a_1 + a_i^{j-2}a_1^2 + \cdots a_i a_1^{j-1} + a_1^j \\ &= a_i^j + a_1(a_i^{j-1} + a_i^{j-2}a_1 + \cdots a_i a_1^{j-2} + a_1^{j-1}) , \end{aligned}$$

so sehen wir, daß

$$q_{ij} = a_i^j + a_1 q_{i,j-1} \quad \text{oder} \quad q_{ij} - a_1 q_{i,j-1} = a_i^j$$

ist. Subtrahieren wir also zuerst a_1 mal die vorletzte Spalte von der letzten, so werden die Einträge der letzten Spalte zu a_i^{n-2}. Entsprechend subtrahieren wir a_1 mal die $(n-2)$-te Zeile von der $(n-1)$-ten und erhalten lauter Einträge a_i^{n-3} und so weiter, bis schließlich die Subtraktion des a_1-fachen der ersten Spalte von der zweiten die Einträge der letzteren zu

$$q_{i1} - a_1 = (a_i + a_1) - a_1 = a_i$$

macht. Somit ist $V(a_1, \ldots .a_n)$ gleich

$$(a_2 - a_1)(a_3 - a_1) \cdots (a_n - a_1) \begin{vmatrix} 1 & a_2 & a_2^2 & \ldots & a_2^{n-2} \\ 1 & a_3 & a_3^2 & \ldots & a_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-2} \end{vmatrix} .$$

Die Determinante rechts ist offensichtlich wieder eine VANDERMONDEsche Determinante, allerdings mit um eins verminderter Zeilen- und Spaltenzahl und mit einer Variablen weniger.

Damit haben wir die Rekursionsformel

$$V(a_1, \ldots, a_n) = (a_2 - a_1)(a_3 - a_1) \cdots (a_n - a_1)\, V(a_2, \ldots, a_n) ,$$

die es erlaubt, die Berechnung von $V(a_1, \ldots, a_n)$ auf eine einzige VANDERMONDEsche Determinante der Größe $(n-1) \times (n-1)$ zurückzuführen.

Zur vollständigen Berechnung von $V(a_1, \ldots, a_n)$ fehlt uns jetzt nur noch ein Induktionsanfang; direktes Nachrechnen zeigt sofort, daß

$$V(a_n) = \det(1) \quad \text{und} \quad V(a_{n-1}, a_n) = \begin{vmatrix} 1 & a_{n-1} \\ 1 & a_n \end{vmatrix} = a_n - a_{n-1}$$

ist, also folgt induktiv

$$V(a_1, \ldots, a_n) = \prod_{j<i} (a_i - a_j) .$$

i) Determinanten und Eigenwerte

In Abschnitt *a)* haben wir gesehen, daß ein Skalar $\lambda \in k$ genau dann Eigenwert einer Matrix $A \in k^{n \times n}$ ist, wenn die Matrix $A - \lambda E$ nicht den maximalen Rang hat, wenn ihre Spaltenvektoren also linear abhängig sind. Nach der Diskussion der letzten Abschnitte ist klar, daß dies genau dann der Fall ist, wenn $\det(A - \lambda E)$ verschwindet, und daß dann die zugehörigen Eigenvektoren gerade die nichttrivialen Lösungen des homogenen linearen Gleichungssystems $(A - \lambda E)\vec{v} = \vec{0}$ ist.

Dies liefert eine Methode zur Berechnung von Eigenwerten und Eigenvektoren: Man löse die Gleichung $\det(A - \lambda E) = 0$ und dann für jede Nullstelle λ_i dieser Gleichung das lineare Gleichungssystem $(A - \lambda E)\vec{v} = \vec{0}$. Dieses homogene lineare Gleichungssystem hat *nie* maximalen Rang, da es nach Definition eines Eigenwerts nichttriviale Lösungen geben muß; kommt man also auf ein eindeutig lösbares Gleichungssystem (und damit auf den Nullvektor als einzige Lösung), ist das immer ein Zeichen für einen Rechenfehler.

Dies zeigt auch, daß die numerische Bestimmung von Eigenvektoren nicht nach obigem Schema vorgehen kann: Falls man die Nullstellen der charakterischen Gleichung nur näherungsweise kennt, kann die Matrix des linearen Gleichungssystems eine leicht von null verschiedene Determinante haben, so daß das Gleichungssystem nur die triviale Lösung hat. In der Numerik werden Eigenwerte und Eigenvektoren daher simultan berechnet – sofern man beides braucht. Es gibt auch numerische Algorithmen, die nur die (angenäherten) Eigenwerte bestimmen ohne $\det(A - \lambda E)$ zu berechnen; dies ist vor allem nützlich für große n, wo die Berechnung einer Determinanten sehr aufwendig wäre. Für Genaueres sei auf die Numerikvorlesung verwiesen.

Als Beispiel für die Berechnung von Eigenwerten und Eigenvektoren nach obiger Methode betrachten wir die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{pmatrix} .$$