TECHNISCHE HOCHSCHULE NÜRNBERG
GEORG SIMON OHM

Lineare Algebra, Teil II

Determinanten

©Edgar M. E. Wermuth
TH Nürnberg
V/2000 - I / 2023

Inhalt

Vorbemerkungen

Determinanten sind viel älter als die Matrizenrechnung, traten im Zusammenhang mit der **Lösung von linearen Gleichungssystemen** schon auf bei Gottfried Wilhelm Leibniz (1646-1716) und bei Seki Kowa (1642-1708) in Japan (unabhängig voneinander). Wesentliche Mitbegründer der Determinantenrechnung sind Gabriel Cramer (1704-1752) und später Bézout, Vandermonde, Laplace, Gauß, Cauchy, Jacobi. Das Wort „Determinante" prägte 1801 Carl Friedrich Gauß (1777-1855) in seinen *Disquisitiones Arithmeticae*.

Die Determinante einer quadratischen Matrix

- ordnet einem quadratischen Schema von n^2 Zahlen eine einzige Zahl als Kenngröße zu;
- „entscheidet", ob Gleichungssysteme mit dieser Koeffizientenmatrix *eindeutig lösbar* sind (det $\neq 0$) und ob die der Matrix entsprechende lineare Abbildung die *Orientierung* erhält (det > 0);
- ist der orientierte Volumenänderungsfaktor der mit der Matrix verbundenen linearen Abbildung (also so etwas wie eine „Matrix-Dicke mit Vorzeichen").

2×2-Determinanten

Sei

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (\mathbf{a}_1, \mathbf{a}_2),\ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Die Lösung $\mathbf{x}$ des Gleichungssystems

$$A\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$$

findet man unmittelbar durch skalare Multiplikation der Linearkombinations-Form mit $\mathbf{a}_2^{\vdash} := \begin{pmatrix} a_{22} \\ -a_{12} \end{pmatrix}$ und $\mathbf{a}_1^{\dashv} := \begin{pmatrix} -a_{21} \\ a_{11} \end{pmatrix}$.
Da $\mathbf{a}_1^{\dashv}$ durch $\frac{\pi}{2}$-Drehung im Gegenuhrzeigersinn aus $\mathbf{a}_1$ und $\mathbf{a}_2^{\vdash}$ durch $\frac{\pi}{2}$-Drehung im Uhrzeigersinn aus $\mathbf{a}_2$ entsteht, ergibt sich

$$\boxed{x_1 = \frac{\mathbf{b} \cdot \mathbf{a}_2^{\vdash}}{\mathbf{a}_1 \cdot \mathbf{a}_2^{\vdash}} = \frac{\det(\mathbf{b}, \mathbf{a}_2)}{\det(\mathbf{a}_1, \mathbf{a}_2)},\ x_2 = \frac{\mathbf{b} \cdot \mathbf{a}_1^{\dashv}}{\mathbf{a}_2 \cdot \mathbf{a}_1^{\dashv}} = \frac{\det(\mathbf{a}_1, \mathbf{b})}{\det(\mathbf{a}_1, \mathbf{a}_2)}}$$

mit der sog. **Determinante** einer 2×2-Matrix:

$$\boxed{\begin{aligned} \det(A) &= \det(\mathbf{a}_1, \mathbf{a}_2) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &:= \mathbf{a}_1 \cdot \mathbf{a}_2^{\vdash} = \mathbf{a}_2 \cdot \mathbf{a}_1^{\dashv} = a_{11}a_{22} - a_{21}a_{12} \end{aligned}}$$

$\det(A)$ ist das *orientierte Volumen* des von den Spalten $\mathbf{a}_1$ und $\mathbf{a}_2$ von A aufgespannten *Parallelogramms*:

$$\det(\mathbf{a}_2, \mathbf{a}_1) = -\det(\mathbf{a}_1, \mathbf{a}_2).$$

3×3-Determinanten

Sei

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3),\ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},\ \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Die Lösung $\mathbf{x}$ des Gleichungssystems

$$A\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}$$

ergibt sich – analog zum vorherigen Fall – durch Skalarmultiplikation mit $\mathbf{a}_2 \times \mathbf{a}_3$, mit $\mathbf{a}_3 \times \mathbf{a}_1$ und mit $\mathbf{a}_1 \times \mathbf{a}_2$:

$$\boxed{\begin{aligned} x_1 &= \frac{\mathbf{b} \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} = \frac{\det(\mathbf{b}, \mathbf{a}_2, \mathbf{a}_3)}{\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)},\ x_2 = \frac{\mathbf{b} \cdot (\mathbf{a}_3 \times \mathbf{a}_1)}{\mathbf{a}_2 \cdot (\mathbf{a}_3 \times \mathbf{a}_1)} \\ &= \frac{\det(\mathbf{a}_1, \mathbf{b}, \mathbf{a}_3)}{\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)},\ x_3 = \frac{\mathbf{b} \cdot (\mathbf{a}_1 \times \mathbf{a}_2)}{\mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2)} = \frac{\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{b})}{\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)} \end{aligned}}$$

mit der **Determinante**

$$\boxed{\begin{aligned} \det(A) = \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ :&= \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = \mathbf{a}_2 \cdot (\mathbf{a}_3 \times \mathbf{a}_1) \\ &= \mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2) \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{aligned}}$$

Die erhaltenen Formeln für die Lösung des Gleichungssystems, $x_1 = \frac{\det(\mathbf{b},\mathbf{a}_2,\mathbf{a}_3)}{\det(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3)}$ etc., heißen *Cramersche Regel*; ebenso die analogen Formeln im 2×2-Fall.

Die Determinante ist *das orientierte Volumen des* von den Vektoren $\mathbf{a}_1$, $\mathbf{a}_2$ und $\mathbf{a}_3$ aufgespannten *Spats (Parallelepipeds)*; sie wird auch (*nur* im 3×3-Fall!) **Spatprodukt** genannt.

Die Berechnungsformel lässt sich anschaulich gemäß der Regel von Sarrus darstellen:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ \underbrace{a_{31} \quad a_{32} \quad a_{33}}_{-} & & & \underbrace{a_{31} \quad a_{32}}_{+} & \end{array}$$

Man ergänzt die Matrix in Gedanken rechts *zyklisch* um zwei Spalten und *addiert* dann die Hauptdiagonal-Produkte und *subtrahiert* die Nebendiagonal-Produkte. Beim Berechnen konkreter 3×3-Determinanten braucht man so nicht an Indizes zu denken.

Wie geht es nun weiter im 4×4-Fall?

Um nach x_1 aufzulösen, braucht man einen Vektor, der auf $\mathbf{a}_2$, $\mathbf{a}_3$ und $\mathbf{a}_4$ senkrecht steht, um mit diesem skalar zu multiplizieren, usw. Das wird aber schon unübersichtlicher, weil man ja kein Kreuzprodukt mehr zur Verfügung hat.

Eine *axiomatische* Betrachtungsweise löst den gordischen Knoten des $n \times n$-Falles.

Determinanten-Axiome

Nur *zwei algebraische Eigenschaften der Determinante* waren nötig zur Auflösung der Gleichungssysteme im Falle $n = 2, 3$; diese wählen wir als *Axiome* für den Allgemeinfall:

- Die Determinante ist *linear* in jeder Spalte; z.B. $\det(\lambda\mathbf{a} + \mu\mathbf{b}, \mathbf{a}_2, \dots, \mathbf{a}_n) = \lambda \det(\mathbf{a}, \mathbf{a}_2, \dots, \mathbf{a}_n) + \mu \det(\mathbf{b}, \mathbf{a}_2, \dots, \mathbf{a}_n)$
- $\det(\mathbf{a}_1, \dots, \mathbf{a}_n)$ *verschwindet*, wenn zwei der Spaltenvektoren übereinstimmen. (Es genügt, dies nur für *benachbarte* Spalten vorauszusetzen.)

In der Tat folgt damit aus $\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$:

$$\begin{aligned}\det(\mathbf{b}, \mathbf{a}_2) &= \det(x_1\mathbf{a}_1 + x_2\mathbf{a}_2, \mathbf{a}_2)\\ &= x_1 \det(\mathbf{a}_1, \mathbf{a}_2) + x_2 \det(\mathbf{a}_2, \mathbf{a}_2)\\ &= x_1 \det(\mathbf{a}_1, \mathbf{a}_2)\end{aligned}$$

und analog $\det(\mathbf{a}_1, \mathbf{b}) = x_2 \det(\mathbf{a}_1, \mathbf{a}_2)$.

Sofort klar: *Genau so* funktioniert's auch im 3×3- und $n \times n$-Fall; es folgt die *allgemeine* Cramer-Regel!

Umgekehrt impliziert die Cramersche Regel die Determinanten-Axiome: Aus $x_1 = \frac{\det(\mathbf{b}, \mathbf{a}_2, \dots, \mathbf{a}_n)}{\det(\mathbf{a}_1, \dots, \mathbf{a}_n)}$ und den analogen Formeln für $x_2, \dots, x_n$ folgt, dass det linear in allen Spalten ist (Lin.-Komb. zweier rechter Seiten beim GL-System); die rechten Seiten $\mathbf{a}_1, \dots, \mathbf{a}_n$ zeigen, dass det verschwindet, wenn zwei Spalten gleich sind.

Aus beiden Eigenschaften zusammen folgt eine Verallgemeinerung der zweiten:

- Vertauscht man zwei Spalten, wechselt die Determinante ihr Vorzeichen.

Beweis:
Aus den Axiomen folgt, dass man beliebige Vielfache einer Spalte zu einer *anderen* Spalte addieren darf, ohne die Determinante zu ändern. Also gilt:

$$\begin{aligned}\det(\mathbf{a}_1, \mathbf{a}_2, \dots) &= \det(\mathbf{a}_1 - \mathbf{a}_2, \mathbf{a}_2, \dots) = \det(\mathbf{a}_1 - \mathbf{a}_2, \mathbf{a}_1, \dots)\\ &= \det(-\mathbf{a}_2, \mathbf{a}_1, \dots) = -\det(\mathbf{a}_2, \mathbf{a}_1, \dots)\,.\end{aligned}$$

Eine Funktion $d(\mathbf{a}_1, \dots, \mathbf{a}_n)$, die lineare Funktion jedes ihrer Argumente ist und bei Vertauschung zweier das Vorzeichen wechselt, nennt man alternierende n-Linearform.

Der springende Punkt: Bis auf einen Normierungsfaktor gibt es *genau eine* alternierende n-Linearform! Das zeigt sich, wenn jedes Argument als Linearkombination der Standardbasis dargestellt und dann entsprechend den beiden Axiomen ausmultipliziert wird.

Fall $n = 2$: $\det(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2, a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2) = a_{11}a_{12} \det(\mathbf{e}_1, \mathbf{e}_1) + a_{11}a_{22} \det(\mathbf{e}_1, \mathbf{e}_2) + a_{21}a_{12} \det(\mathbf{e}_2, \mathbf{e}_1) + a_{21}a_{22} \det(\mathbf{e}_2, \mathbf{e}_2) = (a_{11}a_{22} - a_{21}a_{12}) \det(\mathbf{e}_1, \mathbf{e}_2)\,.$

Der Normierungsfaktor $\det(\mathbf{e}_1, \mathbf{e}_2)$ kürzt sich bei der Berechnung der Lösungen x_1 und x_2 als *Quotienten* zweier Determinanten heraus. Die natürliche Wahl ist $\det(\mathbf{e}_1, \mathbf{e}_2) = 1$.

Fall $n = 3$ ist völlig analog; Ausmultiplizieren ergibt:

$$\begin{aligned}\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) &= \textstyle\sum_{i,j,k=1}^{3} a_{i1}a_{j2}a_{k3} \det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)\\ &= \textstyle\sum_{\pi \in S_3} \sigma_\pi\, a_{\pi_1 1}a_{\pi_2 2}a_{\pi_3 3} \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)\end{aligned}$$

Dabei durchläuft $\pi = (\pi_1, \pi_2, \pi_3)$ die „symmetrische Gruppe“ S_3 der sechs Permutationen von (1,2,3), und $\det(\mathbf{e}_{\pi_1}, \mathbf{e}_{\pi_2}, \mathbf{e}_{\pi_3}) = \sigma_\pi \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, $\sigma_\pi \in \{1, -1\}$.

Eine *zweite* axiomatisch-algebraische Charakterisierung der Determinante ergibt sich aus der *geometrischen* Bedeutung in den betrachteten Fällen $n=2$ und $n=3$ als *orientiertes Volumen* des von den Spaltenvektoren aufgespannten Parallelepipeds:

Die Determinante ist *homogen und scherungsinvariant*.

- $\det(\lambda\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \cdots = \det(\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}, \lambda\mathbf{a}_n)$
  $= \lambda \cdot \det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ (Homogenität)
- $\det(\mathbf{a}_1, \ldots, \mathbf{a}_k + \mathbf{a}_l, \ldots, \mathbf{a}_n) = \det(\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n)$
  für $l \neq k$ (Scherungsinvarianz)

Klar: Jede alternierende n-Linearform ist homogen und scherungsinvariant. Es gilt aber auch die Umkehrung:

Alternation: $d(\ldots, \mathbf{a}_k, \ldots, \mathbf{a}_l, \ldots) = d(\ldots, \mathbf{a}_k + \mathbf{a}_l, \ldots, \mathbf{a}_l, \ldots)$
$= -d(\ldots, \mathbf{a}_k + \mathbf{a}_l, \ldots, -\mathbf{a}_l, \ldots) = -d(\ldots, \mathbf{a}_k + \mathbf{a}_l, \ldots, \mathbf{a}_k, \ldots)$
$= d(\ldots, \mathbf{a}_l, \ldots, -\mathbf{a}_k, \ldots) = -d(\ldots, \mathbf{a}_l, \ldots, \mathbf{a}_k, \ldots)$.

Linearität: $d(\ldots, \mathbf{a}_k, \ldots, \mathbf{a}_l, \ldots) = \frac{1}{\lambda} d(\ldots, \mathbf{a}_k, \ldots, \lambda\mathbf{a}_l, \ldots)$
$= \frac{1}{\lambda} d(\ldots, \mathbf{a}_k + \lambda\mathbf{a}_l, \ldots, \lambda\mathbf{a}_l, \ldots) = d(\ldots, \mathbf{a}_k + \lambda\mathbf{a}_l, \ldots, \mathbf{a}_l, \ldots)$.

Sind $\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$ linear unabhängig (l.u.), werden sie durch $\mathbf{b}$ zu einer Basis ergänzt, und mit $\mathbf{a}_k = \lambda\mathbf{b}$+LK der anderen, $\mathbf{a}_k^* = \lambda^*\mathbf{b}$+LK der anderen folgt $d(\ldots, \mathbf{a}_k + \mathbf{a}_k^*, \ldots) = d(\ldots, (\lambda + \lambda^*)\mathbf{b}, \ldots) =$
$= (\lambda + \lambda^*) d(\ldots, \mathbf{b}, \ldots) = d(\ldots, \lambda\mathbf{b}, \ldots) + d(\ldots, \lambda^*\mathbf{b}, \ldots)$
$= d(\ldots, \mathbf{a}_k, \ldots) + d(\ldots, \mathbf{a}_k^*, \ldots)$.

Sind $\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$ *nicht* l.u., gilt
$d(\ldots, \mathbf{a}_k + \mathbf{a}_k^*, \ldots) = d(\ldots, \mathbf{a}_k, \ldots) = d(\ldots, \mathbf{a}_k^*, \ldots) = 0$.

Permutationen

Für die präzise Formulierung des allgemeinen Determinantenbegriffs brauchen wir ein paar Fakten über Permutationen.
Eine Permutation $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ ist eine Umordnung von $(1, 2, \ldots, n)$.
Anders gesagt: Permutationen sind Bijektionen $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Statt π_1, π_2, etc. schreibt man auch $\pi(1)$, $\pi(2)$, etc.
Die *symmetrische Gruppe* $\boldsymbol{S_n}$ ist die Gesamtheit aller Permutationen von $(1, 2, \ldots, n)$.

Wird ein n-Tupel $(x_1, x_2, \ldots, x_n)$ *verschiedener* Zahlen gemäß einer Permutation π vertauscht, entsteht das n-Tupel $(x_{\pi_1}, x_{\pi_2}, \ldots, x_{\pi_n})$.
Wird dieses neue n-Tupel nun gemäß einer Permutation π^* vertauscht, entsteht

$$(x_{\pi_{\pi_1^*}}, x_{\pi_{\pi_2^*}}, \ldots, x_{\pi_{\pi_n^*}}) = (x_{\pi\pi^*{}_1}, x_{\pi\pi^*{}_2}, \ldots, x_{\pi\pi^*{}_n}) .$$

Die der zweifachen Vertauschung entsprechende Permutation $\pi\pi^* = \pi \circ \pi^*$ entsteht durch Verkettung der Permutationen *in umgekehrter Reihenfolge*.
Permutationen kann man auch repräsentieren durch **Permutationsmatrizen**

$$P = (\mathbf{e}_{\pi_1}, \mathbf{e}_{\pi_2}, \ldots, \mathbf{e}_{\pi_n}).$$

AP entsteht aus A durch *Spalten*vertauschung gemäß π (und $P^T A$ durch *Zeilen*vertauschung).

Wie sich für $n = 2$ und $n = 3$ zeigte, kommt's bei der Determinantenberechnung darauf an, wie viele Spaltenvertauschungen nötig sind, um die Matrix P in die Einheitsmatrix $E = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ umzuwandeln. Dem dienen die folgenden Überlegungen.

Bei einer Permutation $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ heißt ein Paar (π_i, π_k) mit $i < k$ und $\pi_i > \pi_k$ eine *Inversion* oder *Fehlstellung*. Ist $\nu(\pi)$ die Gesamtzahl der Inversionen der Permutation, heißt

$$\sigma(\pi) := (-1)^{\nu(\pi)} = \prod_{i<k} \frac{\pi_k - \pi_i}{k - i}$$

das Signum der Permutation. $(\prod_{i<k} = \prod_{i=1}^{n-1} \prod_{k=i+1}^{n})$
Es gilt auch für *irgendwelche* untereinander verschiedenen Zahlen $x_1, x_2, \ldots, x_n$:

$$\boldsymbol{\sigma(\pi) = \prod_{i<k} \frac{x_{\pi_k} - x_{\pi_i}}{x_k - x_i}}$$

Denn vom *Differenzenprodukt* $\prod_{i<k}(x_k - x_i)$ der Zahlen $x_1, x_2, \ldots, x_n$ unterscheidet sich $\prod_{i<k}(x_{\pi_k} - x_{\pi_i})$ nur dadurch, dass genau alle die Differenzen $x_{\pi_k} - x_{\pi_i}$, bei denen (π_i, π_k) eine Inversion ist, in umgekehrter Reihenfolge, also mit anderem Vorzeichen vorkommen.

Eine Permutation, die nur zwei benachbarte der Zahlen 1 bis n vertauscht, hat offenbar genau eine Inversion. Eine Permutation, die nur irgend zwei Zahlen i und k vertauscht, hat genau $2|i - k| - 1$ Inversionen, weil nur die zwischen ihnen stehenden zu den beiden Zahlen und diese zueinander nicht mehr in natürlicher Reihenfolge stehen. Auf jeden Fall gilt $\sigma(\pi) = -1$ für eine Paarvertauschung π; Paarvertauschungen (*Transpositionen*) sind, wie man sagt, *ungerade* Permutationen. Entsprechend sind die Permutationen mit positivem Signum die *geraden* Permutationen. Klar:

Jede Permutation ist durch Aufeinanderfolge von $\leq n - 1$ Transpositionen erzeugbar.

Die fundamentale Formel für's Signum:

Es gilt

$$\boldsymbol{\sigma(\pi \circ \pi^*) = \sigma(\pi) \cdot \sigma(\pi^*)}$$

für je zwei Permutationen $\pi, \pi^* \in S_n$.

Beweis:

$$\sigma(\pi \circ \pi^*) = \prod_{i<k} \frac{\pi_{\pi^*_k} - \pi_{\pi^*_i}}{k - i} = \prod_{i<k} \frac{\pi_k - \pi_i}{k - i} \cdot \prod_{i<k} \frac{\pi_{\pi^*_k} - \pi_{\pi^*_i}}{\pi_k - \pi_i}$$
$$= \sigma(\pi) \cdot \sigma(\pi^*) . \quad \blacksquare$$

Insbesondere ist das Produkt zweier gerader Permutationen gerade, genauso das Produkt zweier ungerader; und durch eine Paarvertauschung wird das Signum einer Permutation geändert. Ferner: π gerade $\Leftrightarrow \pi^{-1}$ gerade.

Allgemeine Determinanten

Definition Determinantenfunktion

Eine Funktion $d : \mathbb{R}^{n\times n} \to \mathbb{R}$, Schreibweise $d(A) = d(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ für $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n) \in \mathbb{R}^{n\times n}$, heißt *Determinantenfunktion*, wenn sie *homogen und scherungsinvariant* ist.

Wir haben schon bewiesen, dass eine Determinantenfunktion nichts anderes ist als eine alternierende n-Linearform:

- $d(\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_l, \ldots, \mathbf{a}_n)$
  $= -d(\mathbf{a}_1, \ldots, \mathbf{a}_l, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n)$
  (Alternation)
- $d(\mathbf{a}_1, \ldots, \lambda\mathbf{a}_k + \hat{\lambda}\hat{\mathbf{a}}_k, \ldots, \mathbf{a}_n)$
  $= \lambda d(\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n) + \hat{\lambda} d(\mathbf{a}_1, \ldots, \hat{\mathbf{a}}_k, \ldots, \mathbf{a}_n)$
  (n-Linearität)

Durch Ausmultiplizieren gemäß diesen Rechenregeln, wie es in den Fällen $n = 2$ und $n = 3$ schon exemplarisch diskutiert wurde, ergibt sich – bis auf einen Normierungsfaktor $d(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) = d(E)$ – die *Eindeutigkeit* der Determinantenfunktionen und eine explizite Darstellungsformel mithilfe des Signums von Permutationen.

Satz und Definition Determinante

Für jede Determinantenfunktion zur Reihenzahl n gilt

$$d(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \\ = d(E) \cdot \sum_{\pi\in S_n} \sigma(\pi) \cdot a_{\pi_1 1} \ldots a_{\pi_n n}.$$

Setzt man $d(E) := 1$, erhält man die Determinante:

$$\boldsymbol{\det(A) := \sum_{\pi\in S_n} \sigma(\pi) \cdot a_{\pi_1 1} \ldots a_{\pi_n n}}$$

(Leibnizsche Entwicklungsformel).

Die Determinante ist also im wesentlichen *die einzige* alternierende n-Linearform bzw. homogene und scherungsinvariante Abbildung von $\mathbb{R}^{n\times n}$ nach $\mathbb{R}$.

Zur tatsächlichen *Berechnung* von Determinanten ist die Leibnizsche Entwicklungsformel fast immer ungeeignet! Die Anzahl $n!$ der Summanden wächst mit n extrem schnell.

Alle bisherigen und so gut wie alle folgenden Überlegungen lassen sich genauso auf Matrizen $A \in \mathbb{C}^{n\times n}$ anwenden, auch wenn dies meist nicht explizit erwähnt wird.

Eine kleine Lücke hat die bisherige Argumentation. Wir haben nur gezeigt:
Wenn es eine alternierende n-Linearform gibt, ist sie konstantes Vielfaches der Determinante. Es ist noch nicht ganz offensichtlich – im Allgemeinfall –, dass die so definierte Determinante eine alternierende n-Linearform *ist*.

Klar ist, dass die Leibnizsche Entwicklungsformel linear in allen Spalten $\mathbf{a}_k$ ist, also eine n-Linearform definiert. Dass sie auch *alternierend* ist, bleibt zu zeigen.

Die Vertauschung zweier Spalten der Matrix sei ausgedrückt durch eine Permutation π^*. Wir beachten, dass $\sigma(\pi) = -\sigma(\pi \circ \pi^*)$, dass $\pi^* = (\pi^*)^{-1}$ und dass mit π auch $\pi \circ \pi^*$ alle Permutationen aus S_n durchläuft und erhalten mit $\tau = \pi \circ \pi^*$

$$\begin{aligned}
\det(\mathbf{a}_{\pi_1^*}, \ldots, \mathbf{a}_{\pi_n^*}) &= \sum_{\pi \in S_n} \sigma(\pi) \cdot a_{\pi_1\, \pi_1^*} \cdots a_{\pi_n\, \pi_n^*} \\
&= \sum_{\pi \in S_n} \sigma(\pi) \cdot a_{\pi_{\pi_1^*}\, 1} \cdots a_{\pi_{\pi_n^*}\, n} \\
&= \sum_{\tau \in S_n} (-\sigma(\tau)) \cdot a_{\tau_1\, 1} \cdots a_{\tau_n\, n} \\
&= -\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)\,.
\end{aligned}$$

Beispiel*

Die Enwicklungsformel taugt nur in Ausnahmefällen zur konkreten Berechnung von Determinanten,

Einen solchen Ausnahmefall diskutieren wir jetzt, die *Vandermondesche Determinante*

$$V_n := \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} = \sum_{\pi \in S_n} \sigma(\pi) \cdot x_{\pi_1}^0 x_{\pi_2}^1 \cdots x_{\pi_n}^{n-1}\,.$$

Beim Ausmultiplizieren des *Differenzenprodukts*

$$\begin{aligned}
\prod_{i<k} (x_k - x_i) = {} & (x_n - x_{n-1})(x_n - x_{n-2}) \cdots (x_n - x_2)(x_n - x_1) \\
& \cdot (x_{n-1} - x_{n-2}) \cdots (x_{n-1} - x_2)(x_{n-1} - x_1) \\
& \vdots \\
& \cdot (x_3 - x_2)(x_3 - x_1) \\
& \cdot (x_2 - x_1)
\end{aligned}$$

tritt der Summand $x_n^{n-1} x_{n-1}^{n-2} \cdots x_3^2\, x_2^1\, x_1^0$ mit *positivem* Vorzeichen auf, folglich auch $x_{\pi_n}^{n-1} x_{\pi_{n-1}}^{n-2} \cdots x_{\pi_3}^2\, x_{\pi_2}^1\, x_{\pi_1}^0$ beim Ausmultiplizieren von $\prod_{i<k}(x_{\pi_k} - x_{\pi_i})$.
Wegen $\prod_{i<k}(x_{\pi_k} - x_{\pi_i}) = \sigma(\pi) \cdot \prod_{i<k}(x_k - x_i)$ ist also $\sigma(\pi)$ das Vorzeichen von $x_{\pi_1}^0 x_{\pi_2}^1 \cdots x_{\pi_n}^{n-1}$ bei $\prod_{i<k}(x_k - x_i)$.
Da Summanden mit x_i und x_k in *gleicher* Potenz (mit $i \neq k$) mit *beiden* Vorzeichen vorkommen müssten im Produkt $\prod_{i<k}(x_{\pi_k} - x_{\pi_i})$ (π die nur x_i und x_k vertauschende Permutation, $\sigma(\pi) = -1$), heben solche Terme bei $\prod_{i<k}(x_k - x_i)$ sich weg. Das bedeutet:

$$\boldsymbol{V_n = \sum_{\pi \in S_n} \sigma(\pi) \cdot x_{\pi_1}^0 x_{\pi_2}^1 \cdots x_{\pi_n}^{n-1} = \prod_{i<k} (x_k - x_i).}$$

Rechenregeln

Bei der praktischen Ausrechnung von Determinanten hilft – neben Alternation, Scherungsinvarianz und n-Linearität – die folgende Zusammenstellung weiterer Rechenregeln.
Ein halbwegs effizienter genereller Rechenweg: Die Determinante durch Zeilenumformungen (Elimination) auf obere Dreiecksform bringen und dann die Diagonalelemente miteinander multiplizieren (Aussage c) des Satzes).

Satz Determinanten-Eigenschaften

a) $\det(AB) = \det(A)\det(B)$
für $A, B \in \mathbb{R}^{n\times n}$
(Determinanten-Multiplikationssatz)

b) $\det(A^T) = \det(A)$

c) $\det \begin{pmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} = \alpha_1 \dots \alpha_n$

d) $\det \begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix} = \det(A)$

e) $\det \begin{pmatrix} A & * \\ 0 & B \end{pmatrix} = \det(A)\det(B)$
für $A \in \mathbb{R}^{m\times m}, B \in \mathbb{R}^{n\times n}$

f) Laplacescher Entwicklungssatz
Für $1 \le k \le n$ gilt („Entwicklung nach der k-ten Spalte“)

$$\det(A) = \sum_{i=1}^{n} a_{ik} \cdot (-1)^{i+k} D_{ik},$$

wobei die D_{ik} die Determinanten der durch Streichen der i-ten Zeile und k-ten Spalte aus A entstehenden Untermatrizen sind, die sogenannten *Cofaktoren* der a_{ik}. Die Vorzeichen $(-1)^{i+k}$ sind *schachbrettartig* verteilt.
Nach b) klar: Man kann Determinanten auch nach jeder *Zeile* entwickeln.

g) Eine Matrix A ist invertierbar genau dann, wenn $\det(A) \neq 0$; es gilt dann

$$A^{-1} = \frac{1}{\det(A)} \left((-1)^{i+k} D_{ik}\right)^T_{\substack{1\le i\le n\\ 1\le k\le n}}$$

h) Falls $A\mathbf{x} = \mathbf{b}$ und $\det(A) \neq 0$, gilt

$$x_k = \frac{\det(\mathbf{a}_1, \dots, \mathbf{a}_{k-1}, \mathbf{b}, \mathbf{a}_{k+1}, \dots, \mathbf{a}_n)}{\det(A)}$$

für $(1 \le k \le n)$.
(Cramersche Regel)
Allenfalls im 2×2-Fall praktisch brauchbar; von eher theoretischer Bedeutung.

i) Binet/Cauchy-Formel

Für $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times m}$ mit $n \geq m$, $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ und $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)^T$ gilt

$$\det(AB) = \sum_{1\leq k_1<k_2<\ldots<k_m\leq n} \det(\mathbf{a}_{k_1}, \ldots, \mathbf{a}_{k_m}) \cdot \det(\mathbf{b}_{k_1}, \ldots, \mathbf{b}_{k_m})$$

Der Determinanten-Multiplikationssatz, also Aussage a), ist der Spezialfall $m = n$.

j) Vandermonde-Determinante

$$V_n = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1\leq i<k\leq n} (x_k - x_i)$$

k) Cauchy-Determinante

$$\begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & & \vdots \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix} = \frac{\prod_{i<k}(a_k-a_i)\cdot\prod_{i<k}(b_k-b_i)}{\prod_{i,k}(a_i+b_k)}$$

Die letzten beiden Formeln sind wichtige Beispiele *spezieller Determinanten*.

Beweis:

a) Durch

$$d(\mathbf{b}_1, \ldots, \mathbf{b}_n) := \det(A\mathbf{b}_1, \ldots, A\mathbf{b}_n) = \det(AB)$$

mit $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ wird offenbar eine alternierende n-Linearform definiert, so dass also $d(B) = d(E)\det(B)$. Und $d(E) = \det(A)$.

b) Mit $A = (a_{i\,k})_{1\leq i,k\leq n}$ und $A^T = (a^T_{i\,k})_{1\leq i,k\leq n}$ gilt $a^T_{i\,k} = a_{k\,i}$ und damit gemäß Leibniz-Formel

$$\begin{aligned}\det(A^T) &= \sum_{\pi\in S_n} \sigma(\pi)\, a^T_{\pi(1)\,1}\cdots a^T_{\pi(n)\,n} \\ &= \sum_{\pi\in S_n} \sigma(\pi)\, a_{1\,\pi(1)}\cdots a_{n\,\pi(n)} \\ &\underset{(1)}{=} \sum_{\pi\in S_n} \sigma(\pi)\, a_{\pi^{-1}(1)\,1}\cdots a_{\pi^{-1}(n)\,n} \\ &\underset{(2)}{=} \sum_{\pi\in S_n} \sigma(\pi^{-1})\, a_{\pi^{-1}(1)\,1}\cdots a_{\pi^{-1}(n)\,n} \\ &\underset{(3)}{=} \det(A)\,.\end{aligned}$$

(1): $a_{1\,\tau(1)}\cdots a_{n\,\tau(n)} = a_{\pi(1)\,\tau\circ\pi(1)}\cdots a_{\pi(n)\,\tau\circ\pi(n)}$.
(2): $\sigma(\pi^{-1}) = \sigma(\pi)$.
(3): Mit π durchläuft auch π^{-1} ganz S_n.

c) Bei dieser oberen Dreiecksmatrix hat die Leibniz-Entwicklung nur *einen* (evtl.) von 0 verschiedenen Summanden, nämlich $\alpha_1\cdots\alpha_n$; denn bei allen von der Identität verschiedenen Permutationen π gilt $\pi(k) > k$ für mindestens ein k.

d) Sei $S'_n := \{ \pi \in S_n \mid \pi_1 = 1 \}$ und

$$\pi' := (\pi_2 - 1, \pi_3 - 1, \ldots, \pi_n - 1) \quad (\pi \in S'_n),$$

so dass $\pi' \in S_{n-1}$ und

$$\pi = (1, \pi'_1 + 1, \ldots, \pi'_{n-1} + 1).$$

Dann ist offenbar $\pi \mapsto \pi' \quad (\pi \in S'_n)$ eine Bijektion von S'_n auf S_{n-1} mit $\sigma(\pi) = \sigma(\pi')$ für alle $\pi \in S'_n$. (Dieselben Inversionen!)
Daher folgt (Entwicklungsformel):

$$\begin{vmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{vmatrix} = \sum_{\pi \in S'_n} \sigma(\pi) a_{\pi_2 - 1, 1} \cdots a_{\pi_n - 1, n-1}$$

$$= \sum_{\pi \in S'_n} \sigma(\pi') a_{\pi'_1\, 1} \cdots a_{\pi'_{n-1}\, n-1}$$

$$= \det(A)$$

mit $A = (a_{i\,k})_{1 \le i,k \le n-1}$.

e) Zunächst beachten wir, dass
$\begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}$ und $\begin{pmatrix} & & & * \\ & A & & \vdots \\ & & & * \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ *beide* die Determinante $\det(A)$ besitzen, wie $n-1$ Zeilen- und $n-1$ Spaltenvertauschungen sowie d) zeigen.

Damit folgt (Mehrfachanwendung)
$\begin{vmatrix} E & * \\ 0 & A \end{vmatrix} = \begin{vmatrix} A & * \\ 0 & E \end{vmatrix} = \det(A)$ für Einheitsmatrizen E und quadratische Matrizen A.
Da $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} E_m & C \\ 0 & B \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & E_n \end{pmatrix}$ mit $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{m \times n}$ und $m \times m$- bzw. $n \times n$-Einheitsmatrizen E_m und E_n, folgt mit a)
$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det(A) \det(B)$.

f) Mit $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ und $\mathbf{a}_k = (a_{1\,k}, \ldots, a_{n\,k})^T$ $= a_{1\,k}\mathbf{e}_1 + \cdots + a_{n\,k}\mathbf{e}_n$ ergibt die n-Linearität der Determinante

$$\det(\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, \sum_{i=1}^{n} a_{i\,k}\mathbf{e}_i, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n)$$

$$= \sum_{i=1}^{n} a_{i\,k} \det(\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, \mathbf{e}_i, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n)$$

$$\underset{(*)}{=} \sum_{i=1}^{n} a_{i\,k} (-1)^{i+k} \cdot \begin{vmatrix} 1 & a_{i\,1} & \cdots & a_{i\,k-1} & a_{i\,k+1} & \cdots & a_{i\,n} \\ 0 & a_{1\,1} & \cdots & a_{1\,k-1} & a_{1\,k+1} & \cdots & a_{1\,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{i-1\,1} & \cdots & a_{i-1\,k-1} & a_{i-1\,k+1} & \cdots & a_{i-1\,n} \\ 0 & a_{i+1\,1} & \cdots & a_{i+1\,k-1} & a_{i+1\,k+1} & \cdots & a_{i+1\,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n\,1} & \cdots & a_{n\,k-1} & a_{n\,k+1} & \cdots & a_{n\,n} \end{vmatrix}$$

Bei $(*)$ wurden bei den Summanden mit den Faktoren $a_{i\,k}$ jeweils $k-1$ Spalten- und $i-1$ Zeilenvertauschungen vorgenommen. Mit d) folgt der Laplacesche Entwicklungssatz.

g) Nach e) gilt

$$\sum_{i=1}^{n}(-1)^{i+k}D_{i\,k}a_{i\,l} = \text{det}(\mathbf{a}_1,\ldots,\mathbf{a}_{k-1},\mathbf{a}_l,\mathbf{a}_{k+1},\ldots,\mathbf{a}_n)$$
$$= \begin{cases}\text{det } A, & l=k,\\ 0, & \text{sonst.}\end{cases}$$

Also

$$\left((-1)^{i+k}D_{i\,k}\right)^T_{\substack{1\le i\le n\\1\le k\le n}} A = \begin{pmatrix}\text{det } A & & 0\\ & \ddots & \\ 0 & & \text{det } A\end{pmatrix}$$

Also ist A im Falle det $A \neq 0$ invertierbar und besitzt die angegebene Inverse.
Falls hingegen det $A = 0$, kann es keine Inverse A^{-1} geben, da ja anderenfalls nach a) det A det $A^{-1} = \text{det}(A\,A^{-1}) = \text{det } E = 1$.

h) Mit $\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$ folgt

$$\begin{aligned}&\text{det}(\mathbf{a}_1,\ldots,\mathbf{a}_{k-1},\mathbf{b},\mathbf{a}_{k+1},\ldots,\mathbf{a}_n)\\ &= \text{det}(\mathbf{a}_1,\ldots,\mathbf{a}_{k-1},x_k\mathbf{a}_k,\mathbf{a}_{k+1},\ldots,\mathbf{a}_n)\\ &= x_k\cdot\text{det}(\mathbf{a}_1,\ldots,\mathbf{a}_{k-1},\mathbf{a}_k,\mathbf{a}_{k+1},\ldots,\mathbf{a}_n)\,.\end{aligned}$$

(Dieser Gedankengang war der Ausgangspunkt der axiomatischen Determinanten-Definition.)

i) Mit $A = (a_{i\,k})_{\substack{1\le i\le m\\1\le k\le n}}$ und $B = (b_{i\,k})_{\substack{1\le i\le n\\1\le k\le m}}$ gilt

$$C = AB = \begin{pmatrix}\sum_{l=1}^n a_{1\,l}\,b_{l\,1} & \cdots & \sum_{l=1}^n a_{1\,l}\,b_{l\,m}\\ \vdots & \ddots & \vdots\\ \sum_{l=1}^n a_{m\,l}\,b_{l\,1} & \cdots & \sum_{l=1}^n a_{m\,l}\,b_{l\,m}\end{pmatrix},$$

also

$$\begin{aligned}\text{det}\,C &= \sum_{\pi\in S_m}\sigma(\pi)\left(\sum_{l_1=1}^n a_{\pi_1\,l_1}\,b_{l_1\,1}\right)\cdots\left(\sum_{l_m=1}^n a_{\pi_m\,l_m}\,b_{l_m\,m}\right)\\
&= \sum_{l_1=1}^n\cdots\sum_{l_m=1}^n\sum_{\pi\in S_m}\sigma(\pi)a_{\pi_1\,l_1}\cdots a_{\pi_m\,l_m}\,b_{l_1\,1}\cdots b_{l_m\,m}\\
&\underset{1)}{=} \sum_{1\le l_1<\cdots<l_m\le n}\sum_{\tau\in S_m}\sum_{\pi\in S_m}\sigma(\pi)a_{\pi_1\,l_{\tau_1}}\cdots a_{\pi_m\,l_{\tau_m}}\,b_{l_{\tau_1}\,1}\cdots b_{l_{\tau_m}\,m}\\
&= \sum_{1\le l_1<\cdots<l_m\le n}\sum_{\tau\in S_m}\sum_{\pi\in S_m}\sigma(\pi)a_{\pi\tau^{-1}{}_1\,l_1}\cdots a_{\pi\tau^{-1}{}_m\,l_m}b_{l_{\tau_1}\,1}\cdots b_{l_{\tau_m}\,m}\\
&\underset{2)}{=} \sum_{1\le l_1<\cdots<l_m\le n}\sum_{\tau\in S_m}\sigma(\tau)b_{l_{\tau_1}\,1}\cdots b_{l_{\tau_m}\,m}\,\cdot\\
&\qquad\cdot\sum_{\pi\in S_m}\sigma(\pi\tau^{-1})a_{\pi\tau^{-1}{}_1\,l_1}\cdots a_{\pi\tau^{-1}{}_m\,l_m}\\
&= \sum_{1\le l_1<\cdots<l_m\le n}\text{det}(\mathbf{a}_{l_1},\ldots,\mathbf{a}_{l_m})\cdot\text{det}\begin{pmatrix}\mathbf{b}^T_{l_1}\\ \vdots\\ \mathbf{b}^T_{l_m}\end{pmatrix}.\end{aligned}$$

Bei 1) wird benutzt, dass $\sum_{\pi\in S_m}\sigma(\pi)a_{\pi_1\,l_1}\cdots a_{\pi_m\,l_m}$ verschwindet, wenn nicht alle l_i verschieden sind, bei 2) nutzen wir die Identität $\sigma(\tau)\cdot\sigma(\pi\tau^{-1}) = \sigma(\pi)$.

j) Die Vandermonde-Determinante wurde schon als – etwas kniffliges – Beispiel zur Leibnizschen Entwicklungsformel behandelt.

An dieser Stelle zwei einfache *Induktions*beweise; der erste mithilfe des Laplaceschen Entwicklungssatzes, der zweite durch unmittelbare Umformung.

Sei dazu

$$V(x) := \begin{vmatrix} 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \\ 1 & x & \cdots & x^n \end{vmatrix}.$$

Entwicklung nach der letzten Zeile zeigt, dass $V(x)$ ein Polynom höchstens n-ten Grades ist. Klar ist, dass dieses Polynom an den Stellen $x_1, \ldots, x_n$ verschwindet. Sind alle diese x_i verschieden, gilt also

$$V(x) = c(x - x_1) \cdots (x - x_n)$$

mit einer gewissen Konstanten c; für diese, den Koeffizienten von x^n beim Polynom, gilt nach Laplace

$$c = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix} =: V_n\,,$$

also $V(x_{n+1}) = V_{n+1} = (x_{n+1} - x_1) \cdots (x_{n+1} - x_n) V_n$.
Per Induktion folgt die Behauptung.
(Im Falle nicht sämtlich verschiedener x_i ist die Vandermonde-Formel trivial.)

Alternativer Induktionsschritt durch direkte Umformung:

$$\begin{aligned}
\begin{vmatrix} 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \\ 1 & x_{n+1} & \cdots & x_{n+1}^n \end{vmatrix}
&= \begin{vmatrix} 1 & x_1 - x_{n+1} & \cdots & x_1^{n-1}(x_1 - x_{n+1}) \\ \vdots & \vdots & & \vdots \\ 1 & x_n - x_{n+1} & \cdots & x_n^{n-1}(x_n - x_{n+1}) \\ 1 & 0 & \cdots & 0 \end{vmatrix} \\
&= (-1)^n \begin{vmatrix} x_1 - x_{n+1} & \cdots & x_1^{n-1}(x_1 - x_{n+1}) \\ \vdots & & \vdots \\ x_n - x_{n+1} & \cdots & x_n^{n-1}(x_n - x_{n+1}) \end{vmatrix} \\
&= (-1)^n (x_1 - x_{n+1}) \cdots (x_n - x_{n+1}) V_n \\
&= (x_{n+1} - x_1) \cdots (x_{n+1} - x_n) V_n\,.
\end{aligned}$$

k) Im Fall $n = 1$ gilt $\prod_{i<k}(a_k - a_i) = \prod_{i<k}(b_k - b_i) = 1$ und $\prod_{i,k}(a_i + b_k) = a_1 + b_1$; daher ist die angegebene Formel für $n = 1$ richtig. Nun der Induktionsschritt:

$$\begin{aligned}
&\begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} & \frac{1}{a_1+b_{n+1}} \\ \vdots & & \vdots & \vdots \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n} & \frac{1}{a_n+b_{n+1}} \\ \frac{1}{a_{n+1}+b_1} & \cdots & \frac{1}{a_{n+1}+b_n} & \frac{1}{a_{n+1}+b_{n+1}} \end{vmatrix} \\
\underset{1)}{=}\ & \frac{1}{a_{n+1}+b_{n+1}} \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} & 1 + \frac{a_{n+1}-a_1}{a_1+b_{n+1}} \\ \vdots & & \vdots & \vdots \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n} & 1 + \frac{a_{n+1}-a_n}{a_n+b_{n+1}} \\ \frac{1}{a_{n+1}+b_1} & \cdots & \frac{1}{a_{n+1}+b_n} & 1 \end{vmatrix} \\
\underset{2)}{=}\ & \frac{1}{a_{n+1}+b_{n+1}} \begin{vmatrix} \frac{a_{n+1}-a_1}{(a_1+b_1)(a_{n+1}+b_1)} & \cdots & \frac{a_{n+1}-a_1}{(a_1+b_n)(a_{n+1}+b_n)} & \frac{a_{n+1}-a_1}{a_1+b_{n+1}} \\ \vdots & & \vdots & \vdots \\ \frac{a_{n+1}-a_n}{(a_n+b_1)(a_{n+1}+b_1)} & \cdots & \frac{a_{n+1}-a_n}{(a_n+b_n)(a_{n+1}+b_n)} & \frac{a_{n+1}-a_n}{a_n+b_{n+1}} \\ \frac{1}{a_{n+1}+b_1} & \cdots & \frac{1}{a_{n+1}+b_n} & 1 \end{vmatrix} \\
\underset{3)}{=}\ & \frac{\prod_{i=1}^n (a_{n+1} - a_i)}{\prod_{k=1}^n (a_{n+1} + b_k)} \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} & \frac{1}{a_1+b_{n+1}} \\ \vdots & & \vdots & \vdots \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n} & \frac{1}{a_n+b_{n+1}} \\ 1 & \cdots & 1 & 1 \end{vmatrix} \\
\underset{4)}{=}\ & \frac{\prod_{i=1}^n (a_{n+1} - a_i)}{\prod_{k=1}^n (a_{n+1} + b_k)} \begin{vmatrix} \frac{b_{n+1}-b_1}{(a_1+b_1)(a_1+b_{n+1})} & \cdots & \frac{b_{n+1}-b_n}{(a_1+b_n)(a_1+b_{n+1})} & \frac{1}{a_1+b_{n+1}} \\ \vdots & & \vdots & \vdots \\ \frac{b_{n+1}-b_1}{(a_n+b_1)(a_n+b_{n+1})} & \cdots & \frac{b_{n+1}-b_n}{(a_n+b_n)(a_n+b_{n+1})} & \frac{1}{a_n+b_{n+1}} \\ 0 & \cdots & 0 & 1 \end{vmatrix} \\
\underset{5)}{=}\ & \frac{\prod_{i=1}^n (a_{n+1} - a_i) \prod_{k=1}^n (b_{n+1} - b_k)}{\prod_{i=1}^n (a_i + b_{n+1}) \prod_{k=1}^n (a_{n+1} + b_k)} \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & & \vdots \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix}
\end{aligned}$$

1): Faktor aus der letzten Spalte herausziehen; 2): letzte Zeile von den anderen abziehen; 3): aus Zeilen Zählerfaktoren, aus Spalten Nennerfaktoren herausziehen; 4): letzte Spalte von den anderen abziehen; 5): nach letzter Zeile entwickeln, dann Faktoren herausziehen. Per Induktion folgt die Behauptung. ■

Satz Determinanten-Abschätzungen

a) Die Hadamardsche Ungleichung
Für $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{K}^n$ gilt

$$|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)| \leq |\mathbf{a}_1| \cdots |\mathbf{a}_n|.$$

Das Gleichheitszeichen gilt bei linear unabhängigen $\mathbf{a}_i$ genau dann, wenn diese untereinander *orthogonal* sind.

b) Eine Stetigkeitsabschätzung
Für $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{K}^n$ mit $|\mathbf{a}_i - \mathbf{b}_i| \leq \varepsilon,\ |\mathbf{a}_i|, |\mathbf{b}_i| \leq M$ $(1 \leq i \leq n)$ folgt

$$|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n) - \det(\mathbf{b}_1, \ldots, \mathbf{b}_n)| \leq n\varepsilon M^{n-1}.$$

Beim Beweis setzen wir einfachste Orthogonalitäts-Tatsachen im $\mathbb{K}^n$ als bekannt voraus, die hier erst im folgenden Teil III behandelt werden.

Beweis:
Seien zunächst die $\mathbf{a}_i$ als *untereinander orthogonal* angenommen, $\mathbf{a}_i = c_i \mathbf{b}_i$ $(1 \leq i \leq n)$ mit einer ONB $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ des $\mathbb{K}^n$, so dass also $B^*B = BB^* = E$ für $B := (\mathbf{b}_1, \ldots, \mathbf{b}_n)$.

Mit $A := (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ folgt $A^*A = \begin{pmatrix} |c_1|^2 & & 0 \\ & \ddots & \\ 0 & & |c_n|^2 \end{pmatrix}$

und damit $\det(A^*A) = |c_1|^2 \ldots |c_n|^2$, andererseits $\det(A^*) = \det\left(\overline{A}^T\right) = \overline{\det(A)}$, also nach dem Determinanten-Multiplikationssatz

$$\det(A^*A) = |\det(A)|^2.$$

Wegen $|\mathbf{a}_i| = |c_i|$ folglich $|\det(A)| = |\mathbf{a}_1| \cdots |\mathbf{a}_n|$.

Nun seien die $\mathbf{a}_i$ linear unabhängig, aber ansonsten beliebig. (Im Falle linearer Abhängigkeit *verschwindet* die Determinante, und es bleibt nichts zu beweisen.)
Wir zerlegen $\mathbf{a}_1$ in einen auf $\langle \mathbf{a}_2, \ldots, \mathbf{a}_n \rangle$, dem von den restlichen Vektoren aufgespannten Teilraum, senkrechten und einen darin enthaltenen Anteil: $\mathbf{a}_1 = \mathbf{a}_1^{(0)} + \mathbf{a}_1^{(1)}$,

$$\mathbf{a}_1^{(0)} \perp \langle \mathbf{a}_2, \ldots, \mathbf{a}_n \rangle, \quad \mathbf{a}_1^{(1)} \in \langle \mathbf{a}_2, \ldots, \mathbf{a}_n \rangle.$$

Also $|\mathbf{a}_1|^2 = |\mathbf{a}_1^{(0)}|^2 + |\mathbf{a}_1^{(1)}|^2$ sowie
$\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \det(\mathbf{a}_1^{(0)}, \mathbf{a}_2, \ldots, \mathbf{a}_n)$.
Analoge Zerlegung: $\mathbf{a}_2 = \mathbf{a}_2^{(0)} + \mathbf{a}_2^{(1)}$ mit

$$\mathbf{a}_2^{(0)} \perp \langle \mathbf{a}_3, \ldots, \mathbf{a}_n \rangle, \quad \mathbf{a}_2^{(1)} \in \langle \mathbf{a}_3, \ldots, \mathbf{a}_n \rangle.$$

Usw. Am Ende hat man $\mathbf{a}_1^{(0)}, \ldots, \mathbf{a}_n^{(0)}$, die alle untereinander orthogonal sind, und es gilt
$\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \det(\mathbf{a}_1^{(0)}, \mathbf{a}_2^{(0)}, \ldots, \mathbf{a}_n^{(0)})$;
ferner $|\mathbf{a}_1^{(0)}| \cdots |\mathbf{a}_n^{(0)}| < |\mathbf{a}_1| \cdots |\mathbf{a}_n|$, wenn nicht *alle* $\mathbf{a}_i$ untereinander orthogonal sind. Mit dem zu Anfang Gezeigten folgt a).

Nun zu b). Es gilt

$$
\begin{aligned}
&|\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) - \det(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)| \le \\
&\le |\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) - \det(\mathbf{b}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)| \\
&+ |\det(\mathbf{b}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) - \det(\mathbf{b}_1, \mathbf{b}_2, \mathbf{a}_3 \ldots, \mathbf{a}_n)| \\
&+ \cdots \\
&+ |\det(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}, \mathbf{a}_n) - \det(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)| \\
&= \sum_{k=1}^{n} |\det(\mathbf{b}_1, \ldots, \mathbf{b}_{k-1}, \mathbf{a}_k - \mathbf{b}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n)| \\
&\underset{\text{a)}}{\le} \sum_{k=1}^{n} |\mathbf{b}_1| \cdots |\mathbf{b}_{k-1}|\, |\mathbf{a}_k - \mathbf{b}_k|\, |\mathbf{a}_{k+1}| \cdots |\mathbf{a}_n| \\
&\le \sum_{k=1}^{n} M^{k-1} \cdot \varepsilon \cdot M^{n-k} = n\,\varepsilon \cdot M^{n-1}.
\end{aligned}
$$

■

Zu a und b) analoge Abschätzungen gelten natürlich bzgl. der *Zeilen* der Determinanten.

Wesentlich verschärfen kann man die Abschätzung b) im allgemeinen *nicht*; dies zeigt schon die skalare Ungleichung (Diagonalmatrix!)

$$n\,\varepsilon\, M^{n-1} \le (M+\varepsilon)^n - M^n \le n\,\varepsilon\,(M+\varepsilon)^{n-1}\,.$$

Die Aussage b) besagt insbesondere, dass die Determinante eine *stetige* Abbildung vom $\mathbb{K}^{n\times n}$ nach $\mathbb{K}$ ist.

Determinanten-Beispiele

1.

$$
\begin{vmatrix} 2 & 3 & 4 \\ 3 & 4 & 2 \\ 4 & 2 & 3 \end{vmatrix} = 2 \cdot \begin{vmatrix} 4 & 2 \\ 2 & 3 \end{vmatrix} - 3 \cdot \begin{vmatrix} 3 & 2 \\ 4 & 3 \end{vmatrix} + 4 \cdot \begin{vmatrix} 3 & 4 \\ 4 & 2 \end{vmatrix} = \\
= 16 - 3 + 4 \cdot (-10) = \mathbf{-27}
$$

Ein alternativer Rechenweg:

$$
\begin{vmatrix} 2 & 3 & 4 \\ 3 & 4 & 2 \\ 4 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 4 \\ 1 & 1 & -2 \\ 4 & 2 & 3 \end{vmatrix} = -\begin{vmatrix} 1 & 1 & -2 \\ 2 & 3 & 4 \\ 4 & 2 & 3 \end{vmatrix} = \\
= -\begin{vmatrix} 1 & 1 & -2 \\ 0 & 1 & 8 \\ 0 & -2 & 11 \end{vmatrix} = -\begin{vmatrix} 1 & 1 & -2 \\ 0 & 1 & 8 \\ 0 & 0 & 27 \end{vmatrix} = \mathbf{-27}
$$

2.

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

$$= -3\left(-3\begin{vmatrix}1 & -3\\ 3 & 2\end{vmatrix} + (-2)\begin{vmatrix}1 & -1\\ 3 & 0\end{vmatrix}\right) =$$
$$= -3(-33-6) = \mathbf{117}$$

3.
$$\begin{vmatrix}1 & 2 & 3 & 4\\ -1 & 2 & -3 & 4\\ -2 & 0 & 1 & 1\\ 1 & 1 & 1 & 0\end{vmatrix} = \begin{vmatrix}1 & 2 & 3 & 4\\ 0 & 4 & 0 & 8\\ -2 & 0 & 1 & 1\\ 1 & 1 & 1 & 0\end{vmatrix} =$$
$$= \begin{vmatrix}1 & 2 & 3 & 0\\ 0 & 4 & 0 & 0\\ -2 & 0 & 1 & 1\\ 1 & 1 & 1 & -2\end{vmatrix} = 4\begin{vmatrix}1 & 3 & 0\\ -2 & 1 & 1\\ 1 & 1 & -2\end{vmatrix} =$$
$$= 4\left(\begin{vmatrix}1 & 1\\ 1 & -2\end{vmatrix} - 3\begin{vmatrix}-2 & 1\\ 1 & -2\end{vmatrix}\right) =$$
$$= 4(-3-9) = -\mathbf{48}$$

4.
$$\begin{vmatrix}2 & 1 & 1 & 3 & 7 & 9 & 8\\ -1 & 2 & -4 & 2 & 8 & -9 & 17\\ 0 & 0 & 3 & 0 & 1 & 11 & -7\\ 0 & 0 & 4 & 2 & 2 & -1 & 9\\ 0 & 0 & 1 & 0 & 0 & 12 & 3\\ 0 & 0 & 0 & 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 & 0 & 1 & 2\end{vmatrix} =$$

$$= \begin{vmatrix}2 & 1\\ -1 & 2\end{vmatrix} \cdot \begin{vmatrix}3 & 0 & 1\\ 4 & 2 & 2\\ 1 & 0 & 0\end{vmatrix} \cdot \begin{vmatrix}1 & 1\\ 1 & 2\end{vmatrix} = 5 \cdot \begin{vmatrix}0 & 1\\ 2 & 2\end{vmatrix} \cdot 1 =$$
$$= -\mathbf{10}$$

5.
$$\begin{pmatrix}2 & 3 & 4\\ 3 & 4 & 2\\ 4 & 2 & 3\end{pmatrix}^{-1} =$$
$$= \frac{1}{-27}\begin{pmatrix}\begin{vmatrix}4 & 2\\ 2 & 3\end{vmatrix} & -\begin{vmatrix}3 & 2\\ 4 & 3\end{vmatrix} & \begin{vmatrix}3 & 4\\ 4 & 2\end{vmatrix}\\ -\begin{vmatrix}3 & 4\\ 2 & 3\end{vmatrix} & \begin{vmatrix}2 & 4\\ 4 & 3\end{vmatrix} & -\begin{vmatrix}2 & 3\\ 4 & 2\end{vmatrix}\\ \begin{vmatrix}3 & 4\\ 4 & 2\end{vmatrix} & -\begin{vmatrix}2 & 4\\ 3 & 2\end{vmatrix} & \begin{vmatrix}2 & 3\\ 3 & 4\end{vmatrix}\end{pmatrix}^T$$
$$= \frac{1}{-27}\begin{pmatrix}8 & -1 & -10\\ -1 & -10 & 8\\ -10 & 8 & -1\end{pmatrix}^T = \frac{1}{27}\begin{pmatrix}-8 & 1 & 10\\ 1 & 10 & -8\\ 10 & -8 & 1\end{pmatrix}$$

6.
$$\begin{pmatrix}1 & 0 & 3\\ 0 & 1 & -2\\ 1 & 0 & 1\end{pmatrix}^{-1} = ?$$

$$\begin{vmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & -2 \end{vmatrix} = \mathbf{-2},$$

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{pmatrix}^{-1} =$$

$$= \frac{1}{-2} \begin{pmatrix} \begin{vmatrix} 1 & -2 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 0 & -2 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} 0 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} 0 & 3 \\ 1 & -2 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 0 & -2 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \end{pmatrix}^T$$

$$= \frac{1}{2} \begin{pmatrix} -1 & 2 & 1 \\ 0 & 2 & 0 \\ 3 & -2 & -1 \end{pmatrix}^T = \begin{pmatrix} -0.5 & 0 & 1.5 \\ 1 & 1 & -1 \\ 0.5 & 0 & -0.5 \end{pmatrix}$$

7. Das Gleichungssystem

$$\begin{aligned} 3x + 4y + 5z &= 1 \\ 2x + 3y + z &= 0 \\ 2x + y + 2z &= 1 \end{aligned}$$

besitzt, da die Koeffizientendeterminante nicht verschwindet, nach *Cramer* die Lösung

$$x = \begin{vmatrix} 1 & 4 & 5 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{vmatrix} \Big/ \begin{vmatrix} 3 & 4 & 5 \\ 2 & 3 & 1 \\ 2 & 1 & 2 \end{vmatrix} = \frac{-6}{-13},$$

$$y = \begin{vmatrix} 3 & 1 & 5 \\ 2 & 0 & 1 \\ 2 & 1 & 2 \end{vmatrix} \Big/ \begin{vmatrix} 3 & 4 & 5 \\ 2 & 3 & 1 \\ 2 & 1 & 2 \end{vmatrix} = \frac{5}{-13},$$

$$z = \begin{vmatrix} 3 & 4 & 1 \\ 2 & 3 & 0 \\ 2 & 1 & 1 \end{vmatrix} \Big/ \begin{vmatrix} 3 & 4 & 5 \\ 2 & 3 & 1 \\ 2 & 1 & 2 \end{vmatrix} = \frac{-3}{-13}.$$

Generell ist die Cramersche Regel *kein praktisch brauchbares Verfahren* zur Lösung linearer Gleichungssysteme! Nur im 2×2-Fall und allenfalls noch im 3×3-Fall ist es effizient. Ansonsten ist die Berechnung der $n+1$ Determinanten viel rechenaufwendiger als Gauß (bzw. LR-Zerlegung). Die Cramersche Regel ist eher eine *theoretisch* nützliche Beziehung.

8. Die Vandermonde-Determinante und Polynom-Interpolation

Gesucht eine Parabel, also eine Funktion

$$p : x \mapsto c_2 x^2 + c_1 x + c_0,$$

die durch drei gegebene Punkte (x_0, y_0), (x_1, y_1) und (x_2, y_2) verläuft. D. h.: die Koeffizienten c_0, c_1 und c_2 so wählen, dass sie das Gleichungssystem

$$p(x_k) = y_k \quad (0 \le k \le 2)$$

erfüllen. Ein *Interpolationsproblem*. In Matrixform lautet das Gleichungssystem

$$\begin{pmatrix} 1 & x_0 & {x_0}^2 \\ 1 & x_1 & {x_1}^2 \\ 1 & x_2 & {x_2}^2 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}.$$

Man kann eine Lösung leicht explizit angeben:

$$p(x) = \\ y_0 \cdot \frac{x-x_1}{x_0-x_1} \cdot \frac{x-x_2}{x_0-x_2} \\ + y_1 \cdot \frac{x-x_0}{x_1-x_0} \cdot \frac{x-x_2}{x_1-x_2} + y_2 \cdot \frac{x-x_0}{x_2-x_0} \cdot \frac{x-x_1}{x_2-x_1}$$

Die *Lagrangesche Interpolationsformel* im Fall der Parabel.
Da zwei verschiedene Polynome zweiten Grades höchstens an zwei Stellen übereinstimmen (andernfalls: ihre Differenz ein Polynom $\le$ 2-ten Grades, das mehr als zwei Nullstellen besitzt, unmöglich!), besitzt das Parabel- Interpolationsproblem *höchstens* eine Lösung, also (Lagrange liefert eine) *genau* eine.

Im Falle $y_0 = y_1 = 0$, $y_2 = 1$ liefern die Lagrange-Formel und die Cramersche Regel

$$c_2 = \frac{1}{(x_2 - x_0)(x_2 - x_1)} \text{ bzw.}$$

$$c_2 = \begin{vmatrix} 1 & x_0 & 0 \\ 1 & x_1 & 0 \\ 1 & x_2 & 1 \end{vmatrix} \Big/ \begin{vmatrix} 1 & x_0 & {x_0}^2 \\ 1 & x_1 & {x_1}^2 \\ 1 & x_2 & {x_2}^2 \end{vmatrix}.$$

Vergleicht man beides, folgt:

$$\begin{vmatrix} 1 & x_0 & {x_0}^2 \\ 1 & x_1 & {x_1}^2 \\ 1 & x_2 & {x_2}^2 \end{vmatrix} = (x_2 - x_0)(x_2 - x_1)(x_1 - x_0) \\ = \prod_{i<k} (x_k - x_i)$$

Diese Überlegungen lassen sich praktisch ohne Änderung auf den Fall des Polynoms n-ten Grades verallgemeinern.
Für die Lösung des allgemeinen Polynom-Interpolationsproblems

$$p(x_k) = y_k \quad (0 \le k \le n),$$

wobei passende Koeffizienten c_k von

$$p(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n$$

zu vorgegebenen (x_k, y_k) („Stützstellen") gesucht sind, ergibt sich einerseits die allgemeine Lagrange-Formel

$$p(x) = \sum_{k=0}^{n} y_k \cdot \prod_{\substack{0\le i\le n\\ i\neq k}} \frac{x - x_i}{x_k - x_i}.$$

Andererseits folgt mit

$$V(x_0, \dots, x_{n-1}) := \begin{vmatrix} 1 & x_0 & \cdots & x_0^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n-1} & \cdots & x_{n-1}^{n-1} \end{vmatrix}$$

nach Cramer für die Lösung im Falle
$y_0 = \ldots = y_{n-1} = 0,\ y_n = 1$:
$c_n = V(x_0, \dots, x_{n-1})/V(x_0, \dots, x_{n-1}, x_n)$
Die Lösung ist wiederum eindeutig bestimmt, da Polynome n-ten Grades höchstens n verschiedene Nullstellen haben. Mit der Lagrange-Formel also im Spezialfall: $c_n = 1/\prod_{0\le i\le n-1}(x_n - x_i)$.
Vergleich ergibt

$$\frac{V(x_0, \dots, x_n)}{V(x_0, \dots, x_{n-1})} = \prod_{0\le i\le n-1} (x_n - x_i),$$

woraus die Formel für die Vandermonde-Determinante per Induktion folgt.

9. Die Inverse der Vandermonde-Matrix

Mit $p(x_i) = y_i\ (0\le i\le n) \Leftrightarrow p(x) = \sum_{k=0}^{n} y_k \cdot \prod_{i\neq k} \frac{x-x_i}{x_k-x_i}$,

$$\begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}^{-1} \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix},\ p(x) = \sum_{k=0}^{n} c_k x^k$$

besteht die k-te Spalte der inversen Vandermonde-Matrix gerade aus den Polynom-Koeffizienten c_i im Spezialfall $y_k = 1,\ y_i = 0\,(i \neq k)$.
Das sind aber gerade die Koeffizienten des jeweiligen Lagrange-Grundpolynoms $L_k(x) = \prod_{i\neq k} \frac{x-x_i}{x_k-x_i}$.
Um diese kompakt ausdrücken zu können, benutzen wir die sogenannten *elementarsymmetrischen Funktionen* σ_k der Größen $x_0, x_1, \dots, x_n$. Es gilt

$$\begin{aligned}\prod_{k=0}^{n}(x - x_k) = \sigma_0\, x^{n+1} &- \sigma_1(x_0, \dots, x_n)\, x^n \\ &+ \sigma_2(x_0, \dots, x_n)\, x^{n-1} \\ &- + \cdots \\ &+ (-1)^{n+1}\sigma_{n+1}(x_0, \dots, x_n).\end{aligned}$$

Also ist σ_k jeweils die Summe aller k-fachen Produkte verschiedener x_i. Im Hinblick auf die Lagrange-Polynome L_k brauchen wir noch die modifizierten Summen (stets $\sigma_0 = 1$, also auch $\sigma_0^{(k)} = 1$)
$\sigma_j^{(k)} := \sigma_j(x_0, \dots, x_{k-1}, x_{k+1}, \dots, x_n)\ (0 \le j \le n,\ 0 \le k \le n)$.

Damit $$\begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}^{-1} = \left(\frac{(-1)^{n-i}\sigma_{n-i}^{(k)}}{\prod_{j\neq k}(x_k - x_j)}\right)_{\substack{0\le i\le n\\ 0\le k\le n}}$$

Das Berechnen inverser Vandermonde-Matrizen von Hand ist recht mühsam. Daher sei hier gezeigt, wie man entsprechende MAXIMA–Befehle formulieren kann.

```
VANDERM(n):=genmatrix(lambda([i,k],x[i-1]^(k-1)),n+1,n+1);
```

erzeugt eine $n+1$-reihige Vandermonde-Matrix in den Variablen $x_0, x_1, \dots, x_n$; entsprechende Variablen-Liste:

`VANDVAR(n):=makelist(x[i],i,0,n);` Mit

```
SIG(k,L):=if k=0 then 1
          else expand(sum(SIG(k-1,rest(L,-i))
          *L[length(L)+1-i],i,1,length(L)-k+1));

ELSYM(k,L):=if k>length(L) then print("Inkompatible Daten!")
            else SIG(k,L);
```

werden **elementarsymmetrische Funktionen** definiert.

`DEL(k,L):=delete(L[k+1],L);` streicht ein Element einer Liste;

`pro(k,u,n):=product(u-x[i],i,0,n)/(u-x[k]);` und

`PRO(k,n):=subst(u=x[k],pro(k,u,n));`

erzeugen die Nennerprodukte. Die **inverse Vandermonde-Matrix**:

```
INVVANDERM(n):=block([L,M],L:VANDVAR(n),
               M:genmatrix(lambda([i,k],((-1)^(n+1-i))
               *ELSYM(n+1-i,DEL(k-1,L))/PRO(k-1,n)),n+1,n+1),M);
```

Die Probe:

```
matrixmap(ratsimp,matrixmap(expand,VANDERM(7).INVVANDERM(7)));
```

ergibt – nach einer Weile –, wenn alles richtig ist, eine 8×8-Einheitsmatrix.

10. Kreis durch drei Punkte

Dies ist eine andere Interpolationsaufgabe: Gegeben drei Punkte (x_0, y_0), (x_1, y_1) und (x_2, y_2), gibt es genau einen Kreis in der Ebene, der diese Punkte trifft. Gleichung des Kreises mit dem Mittelpunkt (a, b) und dem Radius r:

$$(x-a)^2 + (y-b)^2 = r^2$$

Ausmultipliziert und sortiert:

$$x^2 + y^2 + c_1 x + c_2 y + d = 0$$

Dabei

$$4r^2 = c_1^2 + c_2^2 - 4d > 0$$

Entwicklung nach der letzten Spalte zeigt:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ x_0 & x_1 & x_2 & x \\ y_0 & y_1 & y_2 & y \\ x_0^2+y_0^2 & x_1^2+y_1^2 & x_2^2+y_2^2 & x^2+y^2 \end{vmatrix} = 0$$

ist vom Typ einer Kreisgleichung; außerdem: setzt man für x, y die Werte x_k, y_k mit $k = 0, 1, 2$ ein, *ist* die Gleichung erfüllt, weil die Determinante wegen zweier gleicher Spalten verschwindet. Also ist es *die* Gleichung des Kreises durch die drei gegebenen Punkte.

Die Determinantenbeziehung ergibt *keine* Kreisgleichung, wenn

$$\begin{vmatrix} 1 & 1 & 1 \\ x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \end{vmatrix} = 0,$$

d. h. genau dann, wenn $(x_1 - x_0, y_1 - y_0)^T$ und $(x_2 - x_0, y_2 - y_0)^T$ Vielfache voneinander sind, also die drei Punkte auf einer Geraden liegen.

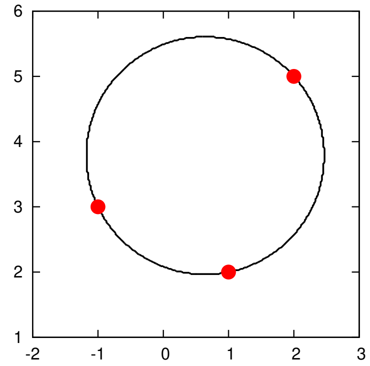
Mittels MAXIMA erzeugen wir ein konkretes Zahlenbeispiel; dabei empfielt es sich, zur grafischen Darstellung das Paket *draw* zu nutzen, eine MAXIMA-Schnittstelle zu GNUPLOT.

```
load(draw)$

Kreisgleich(x0,y0,x1,y1,x2,y2):=
expand(determinant(matrix([1,1,1,1],[x0,x1,x2,x],
[y0,y1,y2,y],[x0^2+y0^2,x1^2+y1^2,x2^2+y2^2,
x^2+y^2]))/determinant(matrix([1,1,1],[x0,x1,x2],
[y0,y1,y2])))=0;

Kreisgleich(1,2,-1,3,2,5);
(Ergebnis: y^2-(53*y)/7+x^2-(9*x)/7+80/7=0)

Kreis(x0,y0,x1,y1,x2,y2,xa,xb,ya,yb):=
draw2d(
[ color=black,line_width=3,implicit(determinant(
matrix( [1,1,1,1],[x0,x1,x2,x],[y0,y1,y2,y],
[x0^2+y0^2,x1^2+y1^2,x2^2+y2^2,x^2+y^2]))=0,
x,xa,xb,y,ya,yb),
color=red,point_type=filled_circle,point_size=4,
points([[x0,y0],[x1,y1],[x2,y2]]) ],
nticks=200,dimensions=[600,600], user_preamble=
"set size ratio -1");
```

Der Kreis durch die Punkte $\binom{1}{2}, \binom{-1}{3}, \binom{2}{5}$ ist gegeben durch die Gleichung

$$y^2 - \frac{53y}{7} + x^2 - \frac{9x}{7} + \frac{80}{7} = 0$$

(so gibt MAXIMA es aus). Siehe nebenstehende MAXIMA-Grafik.

Analog sind weitere Gleichungstypen erfassbar: *Parabeln* durch 3 Punkte in der Ebene, *Ebenen* durch 3 Punkte im Raum, *Kugeln* durch 4 Punkte, allgemeine *Kegelschnitte* durch 5, allgemeine *Quadriken* durch 9 Punkte, usw.

11. Die Determinanten $\begin{vmatrix} a_1 + b_1 & \cdots & a_1 + b_n \\ \vdots & & \vdots \\ a_n + b_1 & \cdots & a_n + b_n \end{vmatrix}$

Im Fall $\mathbf{n = 2}$ ergibt sich:

$$\begin{vmatrix} a_1 + b_1 & a_1 + b_2 \\ a_2 + b_1 & a_2 + b_2 \end{vmatrix} = \begin{vmatrix} a_1 + b_1 & a_1 + b_2 \\ a_2 - a_1 & a_2 - a_1 \end{vmatrix}$$

$$= (a_2 - a_1) \begin{vmatrix} a_1 + b_1 & a_1 + b_2 \\ 1 & 1 \end{vmatrix}$$

$$= (a_2 - a_1) \begin{vmatrix} b_1 - b_2 & a_1 + b_2 \\ 0 & 1 \end{vmatrix}$$

$= (a_2 - a_1)(b_1 - b_2)$, und für $\mathbf{n > 2}$ folgt:

$$\begin{vmatrix} a_1+b_1 & \cdots & a_1+b_n \\ \vdots & & \vdots \\ a_n+b_1 & \cdots & a_n+b_n \end{vmatrix} =$$

$$\begin{vmatrix} b_1-b_3 & b_2-b_3 & a_1+b_3 & \cdots & a_1+b_n \\ \vdots & \vdots & \vdots & & \vdots \\ b_1-b_3 & b_2-b_3 & a_n+b_3 & \cdots & a_n+b_n \end{vmatrix} = 0$$

12. Eine einfache Cauchy-Determinante:

$$\begin{vmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{vmatrix} = \begin{vmatrix} \frac{1}{1+0} & \frac{1}{1+1} & \frac{1}{1+2} \\ \frac{1}{2+0} & \frac{1}{2+1} & \frac{1}{2+2} \\ \frac{1}{3+0} & \frac{1}{3+1} & \frac{1}{3+2} \end{vmatrix}$$

$$= \frac{(3-2)(3-1)(2-1)\cdot(2-1)(2-0)(1-0)}{(1+0)(2+0)(3+0)(1+1)(2+1)(3+1)(1+2)(2+2)(3+2)}$$

$= \frac{1}{2160}$ (ein erstaunlich kleiner Wert!)

Dabei wurde die Formel aus dem Satz angewandt für $a_k = k$, $b_k = k-1$ mit $n = 3$ (Spezialfall **Hilbert-Matrix** H_3).

Einige weitere Determinanten-Werte: $\det(H_4) = 1/6048000$, $\det(H_5) = 1/266716800000$ und $\det(H_{10}) \approx 2.164 \cdot 10^{-53}$. Diese *fast singulären* Matrizen H_n dienen oft als Test-Matrizen in der Numerik.

13. Inverse Cauchy- und Hilbert-Matrizen

Die Cofaktoren D_{ik} einer Cauchy-Matrix sind selbst wieder Cauchy-Determinanten; nur a_i und b_k müssen weggelassen werden. Bei der Gesamt-Cauchy-Determinante ist daher folgendes wegzulassen:

$$\frac{(a_n-a_i)\cdots(a_{i+1}-a_i)(a_i-a_{i-1})\cdots(a_i-a_1)(b_n-b_k)\cdots(b_{k+1}-b_k)(b_k-b_{k-1})\cdots(b_k-b_1)}{(a_i+b_1)(a_i+b_2)\cdots(a_i+b_n)(a_1+b_k)(a_2+b_k)\cdots(a_n+b_k)/(a_i+b_k)}$$

Bildet man die inversen Cauchy-Matrizen gemäß der früher angegebenen Formel mit Cofaktoren, ergibt sich also für diese die Darstellung (*falls* alle Nenner $\neq 0$ sind)

$$\left(\frac{(-1)^{i+k}(a_i+b_1)(a_i+b_2)\cdots(a_i+b_n)(a_1+b_k)(a_2+b_k)\cdots(a_n+b_k)/(a_i+b_k)}{(a_n-a_i)\cdots(a_{i+1}-a_i)(a_i-a_{i-1})\cdots(a_i-a_1)(b_n-b_k)\cdots(b_{k+1}-b_k)(b_k-b_{k-1})\cdots(b_k-b_1)}\right)^T_{\substack{1\le i\le n\\1\le k\le n}}$$

Im Falle einer Hilbert-Matrix H_n, bei der also $a_i = i$ $(1 \le i \le n)$ und $b_k = k-1$ $(1 \le k \le n)$ gilt, ergibt sich die (symmetrische!) inverse Matrix

$$H_n^{-1} = \left(\frac{(-1)^{i+k}\frac{(n-1+i)!}{(i-1)!}\cdot\frac{(n-1+k)!}{(k-1)!}\cdot\frac{1}{i+k-1}}{(n-i)!\cdot(i-1)!\cdot(n-k)!\cdot(k-1)!}\right)^T_{\substack{1\le i\le n\\1\le k\le n}}$$

Bemerkenswerterweise sind alle Elemente dieser Inversen *ganzzahlig*.

Ein Binomialkoeffizient $(k+l)!/(k!\,l!)$, also k aufeinanderfolgende Zahlen geteilt durch $k!$, ist stets ganzzahlig. Die Aufteilung $(1,\ldots,i-1),(i,\ldots,i+k-2),i+k-1,(i+k,\ldots,i+n-1)$ bei $(n-1+i)!$, die Aufteilung $(1,\ldots,i-1),(i,\ldots,i+k-2)$, $i+k-1,(i+k,\ldots,k+n-1)$ bei $(n-1+k)!$ machen dies auf die Elemente der inversen Hilbert-Matrix anwendbar.

14. Verallgemeinerte Kreuzprodukte*

Das Kreuzprodukt zweier Vektoren im $\mathbb{R}^3$ hat auf den ersten Blick keine Verallgemeinerung auf den $\mathbb{R}^n$.

Welche Eigenschaften sollte es haben?

Es sollte *linear* in seinen Faktoren sein, und das Ergebnis sollte ein zu den Faktoren *senkrecht* stehender Vektor sein.

Diese Bedingungen sind leicht erfüllbar, wenn man dabei ein Produkt nicht von 2, sondern von $n-1$ (!) Faktoren bildet:

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} a_1 & b_1 & \mathbf{e}_1 \\ a_2 & b_2 & \mathbf{e}_2 \\ a_3 & b_3 & \mathbf{e}_3 \end{vmatrix} \qquad (\mathbf{a},\mathbf{b}\in\mathbb{R}^3)$$

(Entwicklung nach der letzten Spalte!)
lässt sich verallgemeinern zu

$$\boxed{\mathbf{a}_1\times\cdots\times\mathbf{a}_{n-1} = \begin{vmatrix} a_{1\,1} & \cdots & a_{1\,n-1} & \mathbf{e}_1 \\ \vdots & & \vdots & \vdots \\ a_{n\,1} & \cdots & a_{n\,n-1} & \mathbf{e}_n \end{vmatrix}}$$

für beliebige $\mathbf{a}_1 = [a_{1\,1}, a_{2\,1}\ldots, a_{n\,1}]^T$, ..., $\mathbf{a}_{n-1} = [a_{1\,n-1}, a_{2\,n-1}, \ldots, a_{n\,n-1}]^T \in \mathbb{R}^n$. Bezeichnet A'_k die aus $A := (\mathbf{a}_1,\ldots,\mathbf{a}_{n-1})$ durch Streichen der k-ten Zeile entstehende Matrix, gilt

$$\boxed{\mathbf{a}_1\times\cdots\times\mathbf{a}_{n-1} := \sum_{k=1}^{n}(-1)^{n+k}\det(A'_k)\mathbf{e}_k}$$

Mit dem Laplace-Entwicklungssatz folgt $(\mathbf{a}_1\times\cdots\times\mathbf{a}_{n-1})\cdot\mathbf{a}_k = 0$ $(1\le k\le n-1)$, ferner $\mathbf{e}_1\times\cdots\times\mathbf{e}_{n-1} = \mathbf{e}_n$. Und generell ist, wenn $\mathbf{a}_1,\ldots,\mathbf{a}_{n-1}$ linear unabhängig sind, $\{\mathbf{a}_1,\ldots,\mathbf{a}_{n-1},\mathbf{a}_n\}$ eine *positiv orientierte* Basis im Falle $\mathbf{a}_n = \mathbf{a}_1\times\cdots\times\mathbf{a}_{n-1}$:

$$\det(\mathbf{a}_1,\ldots,\mathbf{a}_{n-1},\mathbf{a}_n) = \sum_{k=1}^{n}\det(A'_k)^2\,.$$

Mit diesem verallgemeinerten Kreuzprodukt kann man die Determinante als *verallgemeinertes Spatprodukt* schreiben:

$$\det A = (\mathbf{a}_1\times\cdots\times\mathbf{a}_{n-1})\cdot\mathbf{a}_n = \sum_{k=1}^{n}(-1)^{n+k}\det(A'_k)a_{k\,n}$$

$$\text{für } A = (\mathbf{a}_1,\ldots,\mathbf{a}_n) = \begin{pmatrix} a_{1\,1} & \cdots & a_{1\,n} \\ \vdots & & \vdots \\ a_{n\,1} & \cdots & a_{n\,n} \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Dies entspricht einer *rekursiven Definition* der Determinante, da ja die $n\times n$-Determinante $\det A$ zurückgeführt wird auf die $(n-1)\times(n-1)$-Determinanten $\det(A'_k)$.

Man kann die ganze Determinantentheorie auch auf Basis dieser Definition entwickeln, ausgehend von den 1×1-Determinanten

$$\det A = a \quad \text{für} \quad A = (a)\,.$$

Dass es eine n-Linearform ist, ist offensichtlich (Induktion).

Etwas aufwendiger: der Nachweis per Induktion, dass bei zwei übereinstimmenden Spalten die Determinante verschwindet. Dazu bezeichne $A''_{k\,l}$ die aus A'_k durch Streichen der l-ten Zeile $(l\ne k)$ und letzten Spalte entstehende $(n-2)\times(n-2)$-Matrix – mit A'_k wie vorher. Man beachte: $A''_{k\,l} = A''_{l\,k}$.

Beim Induktionsschluss von $n-1$ auf n muss man nur den Fall $\mathbf{a}_{n-1} = \mathbf{a}_n$ behandeln, da ja für die $\det A'_k$ die Eigenschaften schon als Induktionsannahme vorausgesetzt werden.
Mit $\det(\mathbf{a}_1,\ldots,\mathbf{a}_{n-1},\mathbf{a}_{n-1}) = \sum_{k=1}^{n}(-1)^{n+k}a_{k\,n-1}\det A'_k$ und

$$\det A'_k = \sum_{l=1}^{k-1}(-1)^{n-1+l}a_{l\,n-1}\det A''_{k\,l} + \sum_{l=k+1}^{n}(-1)^{n-1+l-1}a_{l\,n-1}\det A''_{k\,l}$$

sieht man, dass jeder Ausdruck $a_{k\,n-1}a_{l\,n-1}\det A''_{k\,l}$ $(k\ne l)$ genau zweimal vorkommt, mit unterschiedlichem Vorzeichen. Also folgt $\det(\mathbf{a}_1,\ldots,\mathbf{a}_{n-1},\mathbf{a}_{n-1}) = 0$.
Fazit: Die Determinanten-Axiome gelten. ($\det E = 1$ ist klar.)

(*Permutationen* vermeidet diese Determinanten-Definition, was aber angesichts der Wichtigkeit von Permutationen kein Vorzug ist; eher ist die *Unmotiviertheit* dieses Ansatzes ein Nachteil.)

15. Hauptminoren, Definitheit und Rang

Bei einer $m \times n$-Matrix $A = \left(a_{i\,j}\right)_{\substack{1\le i\le m\\ 1\le j\le n}}$ nennt man die Unterdeterminanten

$$H_k = \begin{vmatrix} a_{1\,1} & \cdots & a_{1\,k} \\ \vdots & & \vdots \\ a_{k\,1} & \cdots & a_{k\,k} \end{vmatrix} \qquad (1 \le k \le \min(m,n)\,)$$

die *Hauptminoren* der Matrix.

Führt man bei A Gauß-Eliminationsschritte *ohne Zeilenvertauschungen* durch, ändern sich die Hauptminoren nicht.

Angenommen, bei *k-1* solchen Eliminationsschritten seien von 0 verschiedene Hauptdiagonalelemente $a_1^{(0)}, \ldots, a_{k-1,k-1}^{(k-2)}$ entstanden, also am Ende folgende Konfiguration erreicht:

$$\begin{pmatrix} a_{1\,1}^{(0)} & & & & * & \\ 0 & a_{2\,2}^{(1)} & & & & \\ \vdots & 0 & \ddots & & & \\ \vdots & \vdots & 0 & a_{k\,k}^{(k-1)} & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots\ 0 & * & \cdots & * \end{pmatrix}$$

Dann gilt

$$H_j = a_{1\,1}^{(0)} \cdot a_{2\,2}^{(1)} \cdots a_{j\,j}^{(j-1)} \qquad (1 \le j \le k)$$

und damit

$$a_{1\,1}^{(0)} = H_1, \quad a_{j\,j}^{(j-1)} = \frac{H_j}{H_{j-1}} \qquad (1 < j \le k)\,.$$

Folglich ergibt sich:

Satz Hauptminoren und Elimination

a) Genau dann, wenn die Hauptminoren H_1, H_2, …, H_{k-1} von 0 verschieden sind, sind die ersten $k-1$ Gaußeliminations-Schritte ohne Zeilenvertauschung durchführbar; dabei entstehen die Pivotelemente
$H_1,\ H_2/H_1,\ \ldots,\ H_k/H_{k-1} = a_{k\,k}^{(k-1)}$.

b) Der Gaußalgorithmus ist bei einer invertierbaren $n \times n$-Matrix genau dann insgesamt ohne Zeilenvertauschungen durchführbar, wenn alle Hauptminoren H_1, …, H_n von 0 verschieden sind.

Quadratische Formen $Q(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x}$ (mit einer Matrix $A \in \mathbb{R}^{n\times n}$) spielen u.a. eine wichtige Rolle bei der *Diskussion stationärer Stellen* von Funktionen mehrerer Variablen.

O.B.d.A. kann man davon ausgehen, dass A *symmetrisch* ist: Da $Q(\mathbf{x})$ ein Skalar ist, gilt $Q(\mathbf{x}) = Q(\mathbf{x})^T = \mathbf{x}^T A^T \mathbf{x}$, also letztlich

$$\boldsymbol{Q(\mathbf{x}) = \mathbf{x}^T \frac{A+A^T}{2}\,\mathbf{x}}.$$

Also kann man A immer durch die *symmetrische* Matrix $(A+A^T)/2$ ersetzen, weshalb wir von nun an $A = A^T$ voraussetzen.
Die quadratische Form und die zugehörige symmetrische Matrix A heißen *positiv definit*, wenn

$$\boxed{\boldsymbol{Q}(\mathbf{x}) = \mathbf{x}^{\boldsymbol{T}} \boldsymbol{A}\, \mathbf{x} > \mathbf{0} \text{ für alle } \mathbf{x} \neq \mathbf{0}}$$

Analog sind A und Q *negativ definit* im Falle $\mathbf{x}^T A\,\mathbf{x} < 0$ $(\mathbf{x} \neq \mathbf{0})$; sie heißen *indefinit*, falls $Q(\mathbf{x})$ für $\mathbf{x} \neq \mathbf{0}$ sowohl positive als auch negative Werte annimmt.
Klar: A ist positiv definit genau dann, wenn $-A$ negativ definit ist.
(Bei stationären Stellen ist A die *Hesse-Matrix* der Funktion, und die drei Begriffe korrespondieren *Minima*, *Maxima* und *Sattelpunkten* der Funktion.)

Satz Definitheitskriterien
Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist positiv definit
a) genau dann, wenn alle ihre Hauptminoren positiv sind;
b) genau dann, wenn Gaußelimination *ohne Zeilenvertauschungen* eine obere Dreiecksmatrix mit positiver Hauptdiagonale ergibt.

Beweis:
Aufgrund der vorherigen Überlegungen über Hauptminoren muss nur b) bewiesen werden.

Gegeben eine symmetrische Matrix A, ist von vornherein klar: $a_{11} > 0$ ist *notwendig* für positive Definitheit. $(\mathbf{e}_k{}^T A\,\mathbf{e}_k = a_{kk} \Rightarrow$ *Alle* Diagonalelemente sind > 0.)

Wir denken uns den ersten Schritt der Gauß-Elimination durchgeführt durch Addition des l_2-Fachen, ..., l_n-Fachen der ersten Zeile zur zweiten, ..., n-ten Zeile. Also

$$L\,A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix} \text{ mit } L := \begin{pmatrix} 1 & & & \\ l_2 & 1 & & \\ \vdots & & \ddots & \\ l_n & & & 1 \end{pmatrix}.$$

Dabei gilt wegen $A^T = A$ auch

$$LAL^T = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix} \quad \text{und} \quad A'^T = A'.$$

Also ist A positiv definit genau dann, wenn $a_{11} > 0$ und A' positiv definit; per Induktion folgt die Behauptung. ∎
(Der Beweis von b) benutzt *nichts* außer der simplen Gauß-Elimination.)

Der *Rang* der Koeffizientenmatrix eines Gleichungssystems wurde in Teil I zunächst definiert als die Anzahl der (von 0 verschiedenen) Pivots am Ende der Gauß-Elimination und erwies sich später als identisch mit der Maximalzahl linear unabhängiger Zeilen und Spalten („Zeilenrang = Spaltenrang") der Matrix. Wir zeigen nun:

Satz Rang und Unterdeterminanten

Der Rang einer Matrix $A \in \mathbb{K}^{m\times n}$ ist genau gleich der maximalen Reihenzahl nichtverschwindender Unterdeterminanten der Matrix.

Beweis:

Gibt es eine Unterdeterminante $\neq 0$ der Reihenzahl r, sind die Spalten der zugehörigen Untermatrix linear unabhängig und damit auch die entsprechenden Spalten der Gesamt-Matrix. Was zeigt, dass der Rang der Matrix dann *mindestens* gleich r ist.

Gehen wir umgekehrt davon aus, der Rang sei gleich r. Wir denken uns die im Laufe der Gauß-Elimination erforderlichen Zeilenvertauschungen *vorab* durchgeführt. Dann ergibt anschließende Gauß-Elimination ohne Zeilenvertauschungen am Ende r Pivotzeilen in Stufenform. Die Pivotzeilen und -spalten bilden eine $r\times r$-Untermatrix in oberer Dreiecksform mit Determinante $\neq 0$. Diese Determinante ist aber *dieselbe* wie *vor* Beginn der Elimination. ■

16.** Kombinationen von Einheitspotenzen

Die folgende analytische Aussage über Linearkombinationen von Potenzen des Betrags 1 ist bemerkenswert und von Bedeutung für lineare Differenzengleichungen. Vorab einige

Beispiele:

(i) Sei $x_n := 1 + \mathrm{e}^{ni\alpha}$ $(n \in \mathbb{N})$ mit α/π irrational.

Dann hat die Folge (x_n) den *gesamten Kreis* $|z-1| = 1$ als Menge aller Häufungswerte, da nach einem bekannten (einfach beweisbaren) Satz von Dirichlet die Zahlen $n\alpha$ mod 2π dicht im gesamten Intervall $[0, 2\pi]$ liegen.

Nach einem berühmten Satz von Hermann Weyl „Über die Gleichverteilung von Zahlen mod. Eins" (1916) ist die Folge sogar *gleichmäßig* auf dem Rand des Kreises verteilt.

(ii) Sei $x_n := \varepsilon_1^n + \varepsilon_2^n + \cdots + \varepsilon_k^n$ $(n \in \mathbb{N})$ mit $\varepsilon_j := \mathrm{e}^{j\cdot 2\pi \mathrm{i}/k}$ $(1 \le j \le k)$.

Dann gilt, mit $\varepsilon := \mathrm{e}^{2\pi\mathrm{i}/k}$, also $\varepsilon_j = \varepsilon^j$,

$$x_n = \varepsilon^n(1 + \varepsilon^n + \varepsilon^{2n} + \cdots + \varepsilon^{(k-1)n}) = \begin{cases} \varepsilon^n \cdot \frac{1-\varepsilon^{nk}}{1-\varepsilon^n} = 0, & n \notin k\mathbb{Z}, \\ k, & n \in k\mathbb{Z}. \end{cases}$$

Die Folge nimmt also *nur zwei* Werte an:

$$(x_n) = (\underbrace{0,\ldots,0}_{k-1 \text{ Glieder}}, k, \underbrace{0,\ldots,0}_{k-1 \text{ Glieder}}, k, \ldots).$$

(iii) Sei $x_n := 1 + \mathrm{e}^{\mathrm{i}\alpha n} + \mathrm{e}^{\mathrm{i}\beta n}$, wobei 1, $\alpha/2\pi$ und $\beta/2\pi$ linear unabhängig seien bzgl. rationaler Koeffizienten. (Etwa $\frac{\alpha}{2\pi} = \sqrt{2}$, $\frac{\beta}{2\pi} = \sqrt{3}$.)

Die Punkte (αn mod 1, βn mod 1) liegen gleichmäßig dicht im Quadrat $[0,1] \times [0,1]$, wenn 1, α, β linear unabhängig bzgl. rationaler Koeffizienten sind; siehe etwa L.Kuipers/H.Niederreiter, *Uniform Distribution of Sequences*, Dover 2006 (Nachdruck), p. 48. Also liegt die oben definierte Folge (x_n) *dicht in der gesamten Kreisscheibe* $|z-1| \leq 2$.

Nun ein Satz über *konvergente* Folgen des in den Beispielen betrachteten Typs.

Satz Konvergente EK-Potenz-Linearkombinationen

Seien $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k$ verschiedene komplexe Zahlen vom Betrage 1 und $c_1, c_2, \ldots, c_k \in \mathbb{C}$ beliebige Konstanten. Dann gilt

$$c_1\,\varepsilon_1^n + c_2\,\varepsilon_2^n + \cdots + c_k\,\varepsilon_k^n \to 0 \ (n \to \infty) \qquad (*)$$

nur im Trivialfall $c_1 = c_2 = \cdots = c_k = 0$.

Äquivalent zur Formulierung des Satzes:

Ist eine Folge $x_n = c_1\,\mathrm{e}^{n\,\mathrm{i}\varphi_1} + \cdots + c_k\,\mathrm{e}^{n\,\mathrm{i}\varphi_k}$ mit $0 \leq \varphi_1 < \varphi_2 < \cdots < \varphi_k < 2\pi$ *konvergent*, gilt $c_i = 0$ für alle $\varphi_i > 0$.

Denn durch einfache Umformung kann man die Bedingung der Konvergenz in die Gestalt $(*)$ bringen.

Beweis des Satzes:

Wir setzen $\mathbf{a}_n := (\varepsilon_1^n, \ldots, \varepsilon_k^n)^T$ $(n \in \mathbb{N})$. Dann gilt

$$\det(\mathbf{a}_n, \mathbf{a}_{n+1}, \ldots, \mathbf{a}_{n+k-1}) = \varepsilon_1^n \varepsilon_2^n \cdots \varepsilon_k^n \begin{vmatrix} 1 & \varepsilon_1 & \cdots & \varepsilon_1^{k-1} \\ 1 & \varepsilon_2 & \cdots & \varepsilon_2^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \varepsilon_k & \cdots & \varepsilon_k^{k-1} \end{vmatrix}$$

$= (\varepsilon_1 \cdots \varepsilon_k)^n \cdot \prod_{j>i}(\varepsilon_j - \varepsilon_i)$ (Vandermonde), also

$$|\det(\mathbf{a}_n, \mathbf{a}_{n+1}, \ldots, \mathbf{a}_{n+k-1})| = \prod_{j>i} |\varepsilon_j - \varepsilon_i|. \qquad (1)$$

Angenommen, *nicht* alle c_i seien $= 0$. Wir ergänzen $\mathbf{b}_1 := \mathbf{c}/|\mathbf{c}|$ mit $\mathbf{c} := (\bar{c}_1, \ldots, \bar{c}_k)^T$ zur ONB $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ des Raumes $\mathbb{C}^k$. Dann gilt

$$\mathbf{a}_n = (\mathbf{a}_n, \mathbf{b}_1)\,\mathbf{b}_1 + \cdots + (\mathbf{a}_n, \mathbf{b}_k)\,\mathbf{b}_k.$$

Wir setzen

$$\mathbf{a}_n^* = (\mathbf{a}_n, \mathbf{b}_2)\,\mathbf{b}_2 + \cdots + (\mathbf{a}_n, \mathbf{b}_k)\,\mathbf{b}_k.$$

Wegen $(*)$ folgt $(\mathbf{a}_n, \mathbf{b}_1) \to 0$, also $|\mathbf{a}_n - \mathbf{a}_n^*| \to 0$ für $n \to \infty$.

Daher ergibt die *Stetigkeitsungleichung* für Determinanten, $|\mathbf{a}_n^*| \leq |\mathbf{a}_n| \leq \sqrt{k}$ berücksichtigend,

$$|\det(\mathbf{a}_n, \ldots, \mathbf{a}_{n+k-1}) - \det(\mathbf{a}_n^*, \ldots, \mathbf{a}_{n+k-1}^*)| \to 0 \ (n \to \infty). \qquad (2)$$

Da offenbar $\det(\mathbf{a}_n^*, \mathbf{a}_{n+1}^*, \ldots, \mathbf{a}_{n+k-1}^*) = 0$, steht (2) im Widerspruch zu (1).

Daher muss die Annahme $\mathbf{c} = (\bar{c}_1, \ldots, \bar{c}_k)^T \neq \mathbf{0}$ *falsch* sein, und die Behauptung ist bewiesen. ■

Mit diesem Satz lässt sich relativ leicht zeigen, dass die Folgen $(\lambda_i^n), (n\,\lambda_i^n), \ldots, (n^{l_i-1}\lambda_i^n)$ $(1 \leq i \leq k)$ für lauter verschiedene $\lambda_i \in \mathbb{C}$ ein insgesamt linear unabhängiges System von Folgen bilden, also einen $(l_1 + \cdots + l_k)$-dimensionalen Unterraum des Raumes aller Folgen aufspannen.

17.* Die Determinante von e^A

Als Umkehrung der Mercatorschen Logarithmusreihe wurde die *Exponentialreihe* gefunden (Isaac Newton, 1669), später ihr Bezug zur *stetigen Verzinsung* erkannt (Jakob Bernoulli, 1696). Zu letzterem passt gut die *Grenzwertdefinition* $e^x = \lim_{k\to\infty}(1+\frac{x}{k})^k$ (Leonhard Euler, 1743), die leicht auf *quadratische Matrizen* zu übertragen ist (Grenzwerte *elementweise*):

$$\mathbf{e}^{\boldsymbol{A}} := \lim_{k\to\infty}\left(\boldsymbol{E}+\tfrac{1}{k}\boldsymbol{A}\right)^{\boldsymbol{k}} \quad (\boldsymbol{A}\in\mathbb{K}^{n\times n})\,.$$

Dass dies gleichwertig ist mit der Reihe

$$\mathbf{e}^{\boldsymbol{A}} = \sum_{k=0}^{\infty}\tfrac{1}{k!}\boldsymbol{A}^{\boldsymbol{k}}\,,$$

ist für Matrizen praktisch mit denselben Argumenten nachweisbar wie im skalaren Fall. Die Matrizen-Exponentialfunktion spielt vor allem eine Rolle bei *linearen Differenzialgleichungssystemen erster Ordnung mit konstanten Koeffizienten*:

$$\mathbf{y}'(t) = A\,\mathbf{y}(t)+\mathbf{f}(t), \quad \mathbf{y}(t_0)=\mathbf{y}_0,$$

hat die Lösung (*komponentenweises* Integral)

$$\mathbf{y}(t) = \mathbf{y}_0\,e^{A(t-t_0)} + \int_{t_0}^{t} e^{A(t-\tau)}\mathbf{f}(\tau)d\tau\,.$$

Zum anderen ergibt sie eine interessante Darstellung *dreidimensionaler Drehungen*:

Ist $\mathbf{a}$ die Drehachse, $|\mathbf{a}|=1$, und ε ein sehr kleiner Drehwinkel, wird jeder Punkt $\mathbf{x}$ in erster Ordnung *senkrecht zur Drehachse* bewegt; die Verschiebung ist $\varepsilon\,\mathbf{a}\times\mathbf{x}+\mathcal{O}(\varepsilon^2)$; also

$$D_{\mathbf{a},\varepsilon}\,\mathbf{x} = \mathbf{x}+\varepsilon\,\mathbf{a}\times\mathbf{x}+\mathcal{O}(\varepsilon^2) = (E+\varepsilon\,A_\times+\mathcal{O}(\varepsilon^2))\mathbf{x}$$

mit $\boldsymbol{A}_\times := \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$. Mit $\varepsilon=\varphi/k$, $k\to\infty$ folgt $D_{\mathbf{a},\varphi}\,\mathbf{x} = D^k_{\mathbf{a},\varphi/k}\mathbf{x} = (E+\frac{\varphi A_\times}{k}+\mathcal{O}(k^{-2}))^k\mathbf{x} \to e^{\varphi A_\times}\mathbf{x}$; denn

$$\left(\mathbf{1}+\tfrac{\boldsymbol{x_k}}{\boldsymbol{k}}\right)^{\boldsymbol{k}} \to \mathbf{e}^{\boldsymbol{x}}, \text{ falls } \boldsymbol{x_k}\to\boldsymbol{x}. \qquad (*)$$

Da $A_\times=(\mathbf{a}\times\mathbf{e}_1,\mathbf{a}\times\mathbf{e}_2,\mathbf{a}\times\mathbf{e}_3)$ und $\mathbf{a}\times(\mathbf{a}\times\mathbf{e}_i)=a_i\,\mathbf{a}-\mathbf{e}_i$, gilt $A_\times^2=\mathbf{a}\,\mathbf{a}^T-E$, $A_\times^3=-A_\times=A_\times^T$. Also ist die spezielle Exponentialreihe $\sum_{k=0}^{\infty}\frac{1}{k!}(\varphi\,A_\times)^k$ leicht auszurechnen:

$$D_{\mathbf{a},\varphi}=e^{\varphi A_\times}=\sum_{k=0}^{\infty}\frac{\varphi^k}{k!}A_\times^k=E+\sin\varphi\,A_\times+(1-\cos\varphi)A_\times^2.$$

Nun zur Determinante von e^A:

$$\det\left(E+\frac{A}{k}\right)^k = \begin{vmatrix} 1+\frac{a_{11}}{k} & \frac{a_{12}}{k} & \cdots & \frac{a_{1n}}{k} \\ \frac{a_{21}}{k} & 1+\frac{a_{22}}{k} & \cdots & \frac{a_{2n}}{k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{k} & \frac{a_{n2}}{k} & \cdots & 1+\frac{a_{nn}}{k} \end{vmatrix}^k,$$

$$\begin{vmatrix} 1+\frac{a_{11}}{k} & \frac{a_{12}}{k} & \cdots & \frac{a_{1n}}{k} \\ \frac{a_{21}}{k} & 1+\frac{a_{22}}{k} & \cdots & \frac{a_{2n}}{k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{k} & \frac{a_{n2}}{k} & \cdots & 1+\frac{a_{nn}}{k} \end{vmatrix} = 1+\frac{a_{11}+a_{22}+\cdots+a_{nn}}{k}+\mathcal{O}(k^{-2})$$

und damit (man beachte $(*)$)

$$\det(e^A) = \lim_{k\to\infty}\det\left(E+\frac{A}{k}\right)^k = \lim_{k\to\infty}\left(1+\frac{a_{11}+\cdots+a_{nn}}{k}\right)^k,$$

also:

$$\boxed{\det(\mathbf{e}^{\boldsymbol{A}}) = \mathbf{e}^{a_{11}+a_{22}+\cdots+a_{nn}} \quad \text{für} \quad \boldsymbol{A}\in\mathbb{K}^{n\times n}}$$

Insbesondere: $\det(e^{At}) = (\det(e^A))^t$ $(A\in\mathbb{R}^{n\times n},\ t\in\mathbb{R})$.

18. Determinante und lineare Abbildungen

Hier werden Determinanten vorrangig im Zusammenhang mit konkret gegebenen *Matrizen* definiert und betrachtet. Bedeutende Mathematiker betonten die Vorzüge des *koordinatenfreien* Umgangs mit Vektorräumen. Dem entspricht eine Definition der Determinante durch Bezugnahme auf lineare Abbildungen, wie sie z.B. Jean Dieudonné der konventionellen Herangehensweise vorzog und daher in seiner knappen Einführung in die lineare Algebra im Anhang der zweiten Auflage seines einflussreichen Buchs *Foundations of Modern Analysis* zugrundelegte.

Ausgehend von der leicht nachweisbaren Tatsache, dass es bei gegebenem n-dimensionalen Vektorraum V über $\mathbb{K}$ bis auf einen konstanten Faktor genau eine nichtverschwindende alternierende n-Linearform $d : V \times \cdots \times V \to \mathbb{K}$ gibt, wird die *Determinante einer linearen Abbildung* $f : V \to V$ definiert durch

$$d(f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)) = \det(f) \cdot d(\mathbf{x}_1, \ldots, \mathbf{x}_n)$$

für beliebige $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$; $\det(f \circ g) = \det(f) \det(g)$ folgt unmittelbar. Legt man eine Basis von V zugrunde, entspricht f eine Abbildungsmatrix A, und $\det(f) =: \det(A)$. Der Determinanten-Multiplikationssatz zeigt, dass beim Übergang zu einer anderen Basis und damit Abbildungsmatrix $B^{-1}AB$ die Determinante dieselbe bleibt – wie es sein muss.

Der weitgehende Verzicht auf den expliziten Gebrauch von Matrizen sei von Vorteil, gibt Dieudonné in seiner Darstellung zu verstehen. Viele halten dennoch Matrizen für sehr nützlich.

19.*** Orientierung: Rechts und Links

Ein Körper, der *genau eine* Symmetrie-Ebene besitzt, hat zwei spiegelbildliche Hälften, eine „rechte“ und eine „linke“ Hälfte. Solche Körper sind z.B. (näherungsweise) ein Mensch, ein Stuhl, ein Löffel, eine Tasse, aber kein (perfekter) Ziegelstein, kein Weinglas, kein Baum und i.a. kein Felsbrocken. Fast jedes Tier hat genau eine Symmetrieebene, wohl weil Erdanziehung und Vorwärtsdrang als orientierte Achsen eine Ebene aufspannen, welche die Umwelt des Tieres in zwei gleich gefährliche Hälften zerlegt. (Da das Erdreich die Gravitation praktisch wegnimmt, hat der Regenwurm nur den Vorwärtsdrang, also bloß eine Symmetrie*achse*. Bäume wiederum haben nur die Erdanziehung und daher auch bloß eine Symmetrieachse – es sei denn, sie stehen ziemlich oft im selben Wind.)

Zu jedem Detail auf der „rechten“ Seite eines einfach-symmetrischen Körpers gibt es ein spiegelbildliches Gegenstück auf der „linken“ Seite: Ein rechtes und ein linkes Ohr, eine rechte und eine linke Hand des Menschen, ein rechtes und ein linkes Hinterbein des Stuhls, etc. Was als „rechts“ und was als „links“ bezeichnet wird, ist natürlich willkürlich.

Aber ist an der „rechten“ Seite eines Löffelgriffs ein Markierungspunkt angebracht, so wird dieser Punkt bei einer Bewegung immer auf der rechten Seite bleiben. Denn jede Bewegung lässt sich zusammengesetzt denken aus sehr vielen fast unmerklich kleinen Bewegungen, bei denen der Löffel jeweils kaum seine Lage ändert: Jede endliche Bewegung ist Resultat einer stetigen Folge „unendlich kleiner“ Bewegungen. Kurz: Durch eine Bewegung wird aus einer rechten Hand niemals eine linke.

Anders die Spiegelung: Spiegeln wir den Löffel an irgendeiner Raumebene, so können wir uns diese Transformation ersetzt denken durch eine Spiegelung an der Symmetrieebene des Löffels, gefolgt von einer Bewegung. Der Punkt ist also nach einer Spiegelung nicht mehr auf der rechten, sondern auf der linken Seite des Löffelgriffs. Bei jedem Rechts-Links-Körper geschieht notwendigerweise dasselbe: Jede Spiegelung an irgendeiner Ebene vertauscht die rechte und linke Seite dieses Körpers. Durch eine Spiegelung wird jede rechte Hand zu einer linken und umgekehrt. Man sagt: Bei Spiegelungen ändert sich die „Orientierung“ des Körpers. Demgemäß können Spiegelungen auch niemals ganz allmählich aus unendlich kleinen Bewegungen aufgebaut werden. Eine Spiegelung ist eine diskontinuierliche Änderung der Konfiguration, bei der die Orientierung umschlägt.

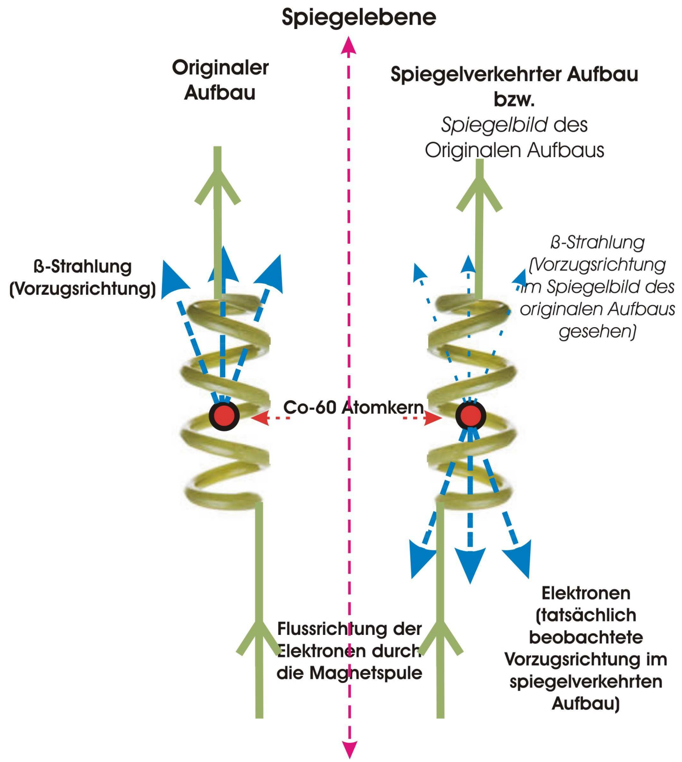
Was ändert sich überhaupt bei einer Spiegelung durch das Umschlagen der „Orientierung“? Gewissermaßen *gar nichts*: alle gespiegelten Punkte haben untereinander dieselbe Beziehung wie vorher. Erst aus der Perspektive eines von der Raumtransformation verschonten „absoluten“ Beobachters wird es relevant: Der Orientierungsumschlag ist eine Änderung im Verhältnis zu einem Beobachter mit orientierungsfester Raumwahrnehmung! Mit einem wirklichen Spiegel können wir dies realisieren und unsere rechte Hand abwechselnd diesseits und – zur linken umgeschlagen – jenseits der Spiegelebene anschauen.

Die absolute Qualifizierung einer Hand als rechte oder linke Hand, einer Schraube als Rechts- oder Linksschraube ist nicht möglich anhand ihrer mathematischen Beschreibung als Punktmenge im $\mathbb{R}^3$, sondern nur, wenn vorher durch Blick auf das zugrundeliegende Koordinatensystem absolut entschieden ist, dass es sich um ein positiv orientiertes handelt.

Dieser auf das Koordinatensystem gerichtete Blick des Beobachters ist ein besonders leicht zu erkennendes Beispiel der *unvermeidlichen Subjektverankerung* „objektiver“ Naturbeschreibung. Wie unterscheiden wir sicher zwischen „Rechts“ und „Links“? Man denke daran, dass es Kindern im Kindergarten-Alter durchaus nicht immer leicht fällt. Vermutlich hat es mit der Unsymmetrie unserer Selbstwahrnehmung zu tun: Trotz einer ungefähren Rechts-Links-Symmetrie fühlen sich beide Körperhälften *nicht genau gleich* an; was auch, aber nicht nur, aus der Händigkeit resultiert. Die Rechts-Links-Symmetrie gilt ja bekanntlich auch nicht für die Anordnung aller *inneren* Organe. So haben wir eine unmittelbar intuitive Unterscheidungsmöglichkeit. Hinzu kommen Orientierungsmerkmale in unserer sichtbaren Umwelt, von der Schrift-Ausrichtung bis zur Lenkrad-Anordnung im Auto.

Man könnte den Begriff „rechte Hand“ *empirisch* mittels der Händigkeit definieren, also als diejenige Hand, die bei der deutlichen Mehrheit der Menschen die geschicktere Hand ist. Diese sozusagen objektive Definition ist es auch, die den Zusammenhang zwischen den Begriffen „rechts“ und „richtig“ verdeutlicht. Die wissenschaftlich-mathematische Formulierung dieser Händigkeits-Definition benutzt das Urteil des Beobachters und damit ein *existierendes Subjekt*: Als bloße Punktmenge in einem Koordinatensystem lässt sich eine rechte Hand nicht beschreiben, ohne dass dieses Koordinatensystem als rechtshändig auf die Wahrnehmungsrealität bezogen wird.

Will man also einem intelligenten Wesen, mit dem nur mittels elektromagnetischer Strahlung über große Distanz kommuniziert werden kann, die Tatsache der überwiegenden Rechtshändigkeit bei uns mitteilen, braucht man irgendeinen Bezug zur Existenz in einer gemeinsamen Realität. Diesen liefern z.B. die Paritätsverletzungen beim π^+- und beim β^--Zerfall (Yang und Lee 1956, Wu 1957).



Prinzip des Wu-Experiments, Abbildung aus Wikipedia

Dadurch lässt sich die Koordinatensystem-Orientierung mit dem entfernten Beobachter synchronisieren; ein Low-Level-Ersatz für die Zählung von Rechtshändern, aber *nichts grundsätzlich anderes*. In beiden Fällen wird die Bedeutung der Aussagen erst durch den Bezug zu den beobachtenden Subjekten festgelegt.

(Kommt zur Spiegelung der *Übergang zu Antiteilchen* hinzu, wird's wieder ununterscheidbar: C-P-Invarianz des Beta-Zerfalls.)

Es scheint, dass genaueres Nachdenken erweisen wird, dass dies letztlich für die Bedeutung *aller* Begriffe und Aussagen gilt: Objektivität ist Intersubjektivität. Wissenschaft ist bedeutungsunsicher ohne *letztendliche Verankerung aller Begriffe im unmittelbaren Existenzerleben* der sie nutzenden Subjekte. Selbst die rein mathematischen Fakten lassen sich kaum subjektfrei mit Sinn füllen. Was sind *Zahlen* ohne die Interaktion von Subjekten mit ihren Wahrnehmungs-Strukturen, die Wiederkehr desselben (oder *fast* desselben) in Raum und Zeit?

All unsere abstrakten Theorien, wie hoch aufgetürmt auch immer sie über unseren Alltagsbegriffen aufragen, haben nur insofern Sinn, als sie letzten Endes durch eine Kette von Bezugnahmen in unserem unmittelbaren anschaulichen Erleben verankert sind.

Zurück zum *innermathematischen* Orientierungsbegriff. Indem die Standardbasis des $\mathbb{R}^3$ und allgemeiner des $\mathbb{R}^n$ *per definitionem* als positiv orientiert vorausgesetzt wird, liefert die Determinante relativ dazu die Orientierung jedes n-Tupels $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ von Vektoren:

Sie sind *positiv* orientiert genau dann, wenn $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n) > 0$.

Wie schon früher erwähnt, kann man die Determinante selbst *definieren* als orientiertes Volumen des von ihren Spalten aufgespannten n-dimensionalen Spats (Parallelepipeds) und aus der daraus folgenden *Homogenität und Scherungsinvarianz* alle Determinanteneigenschaften rein algebraisch ableiten. Dies setzt jedoch voraus, dass man, neben dem n-dimensionalen Volumen, auch die Orientierung schon definiert hat. Wie aber kann man dies *determinantenfrei* machen?

Es liegt nahe, die positive Orientierung im Vektorraum $\mathbb{R}^n$ durch Bezugnahme auf die Standardbasis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ zu definieren:

> Jedes durch eine *stetige, durchgehend linear unabhängige* Schar $\big(\mathbf{b}_1(t), \ldots, \mathbf{b}_n(t)\big)_{0 \le t \le 1}$ in das Standardbasis-Tupel $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ überführbare Vektor-n-Tupel $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ ist *positiv* orientiert.

Dabei gilt natürlich $\mathbf{b}_i(0) = \mathbf{b}_i$ $(1 \le i \le n)$. Man kann eine solche Basis-Schar etwa folgendermaßen konstruieren.

(Bei den weiteren Erörterungen werden Skalarprodukt, Orthogonalität und die damit zusammenhängenden elementaren Begriffe und Konstruktionen als bekannt vorausgesetzt, obwohl sie erst im Teil III dieser Vorlesung systematisch behandelt werden.)

Im ersten Schritt wird eine „Homotopie-Variante" des Gram/-Schmidt-Orthogonalisierungsverfahrens genutzt, um das gegebene Vektor-n-Tupel stetig zu einer Orthonormalbasis zu machen:

$$\mathbf{b}_k(t) = \left(\mathbf{b}_k - t\sum_{j=1}^{k-1}(\mathbf{b}_k\cdot\mathbf{b}_j(t))\mathbf{b}_j(t)\right)\left(1 - t + t\left|\mathbf{b}_k - t\sum_{j=1}^{k-1}(\mathbf{b}_k\cdot\mathbf{b}_j(t))\mathbf{b}_j(t)\right|\right)^{-1}$$

für $1 \le k \le n$ und $0 \le t \le 1$. Da jedes $\mathbf{b}_k(t)$ immer eine Linearkombination von $\mathbf{b}_1, \ldots, \mathbf{b}_k$ mit Koeffizient $\neq 0$ bei $\mathbf{b}_k$ ist, ist diese Schar durchgehend linear unabhängig. Und mit $\mathbf{b}_k(1) =: \mathbf{b}_k^*$ gilt für $1 \le k \le n$:

$$\mathbf{b}_k^* = \left(\mathbf{b}_k - \sum_{j=1}^{k-1}(\mathbf{b}_k\cdot\mathbf{b}_j^*)\mathbf{b}_j^*\right)\left|\mathbf{b}_k - \sum_{j=1}^{k-1}(\mathbf{b}_k\cdot\mathbf{b}_j^*)\mathbf{b}_j^*\right|^{-1}.$$

Genau die ONB, die nach dem Gram/Schmidt-Verfahren aus der ursprünglich gegebenen Basis entsteht.

Im zweiten Schritt wird nun die erhaltene ONB $(\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*)$ nach und nach durch stetige *ebene* Rotationen in die Standard-Position gebracht.

Als erstes betrachten wir die von $\mathbf{b}_1^*$ und $\mathbf{e}_1$ aufgespannte Ebene; im Falle $\mathbf{b}_1^* = \pm\mathbf{e}_1$ gehen wir gleich zu $\mathbf{b}_2^*$ und $\mathbf{e}_2$ weiter.

$\mathbf{b}_1^+ := (\mathbf{e}_1 - (\mathbf{e}_1\cdot\mathbf{b}_1^*)\mathbf{b}_1^*)/|\mathbf{e}_1 - (\mathbf{e}_1\cdot\mathbf{b}_1^*)\mathbf{b}_1^*|$ ist ein zu $\mathbf{b}_1^*$ senkrechter Einheitsvektor in dieser Ebene, und $\cos t\,\mathbf{b}_1^* + \sin t\,\mathbf{b}_1^+$ nimmt für $t = t_0 := \arctan\frac{|\mathbf{e}_1 - (\mathbf{e}_1\cdot\mathbf{b}_1^*)\mathbf{b}_1^*|}{\mathbf{e}_1\cdot\mathbf{b}_1^*}$ den Wert $\mathbf{e}_1$ an, mit Ausnahme des Sonderfalls $\mathbf{e}_1 \perp \mathbf{b}_1^*$, also $\mathbf{b}_1^+ = \mathbf{e}_1$, für den stattdessen $t_0 = \pi/2$ zu wählen ist. Die Drehung der $\mathbf{b}_1^*$-$\mathbf{e}_1$-Ebene wird durch

$\mathbf{b}_i^* \mapsto \mathbf{b}_i^* - (\mathbf{b}_i^*\cdot\mathbf{b}_1^+)\mathbf{b}_1^+ + (\mathbf{b}_i^*\cdot\mathbf{b}_1^+)\big(\cos t\,\mathbf{b}_1^+ - \sin t\,\mathbf{b}_1^*\big) \quad (2 \le i \le n)$

zu einer n-dimensionalen Rotation vervollständigt, da $\mathbf{b}_i^*\cdot\mathbf{b}_1^* = 0$.

Sind $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$ schon in die Positionen $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_k$ gedreht, betrachtet man als nächstes die $\mathbf{b}_{k+1}^*$-$\mathbf{e}_{k+1}$-Ebene, wobei der inzwischen durch Drehungen *veränderte* Vektor $\mathbf{b}_{k+1}^*$ gemeint ist. Daher stehen diese Ebene sowie $\mathbf{b}_{k+2}^*, \ldots, \mathbf{b}_n^*$ senkrecht auf $\mathbf{e}_1, \ldots, \mathbf{e}_k$ und die $\mathbf{b}_{k+1}^*$ in die Position $\pm\mathbf{e}_{k+1}$ überführende ebene Rotation kann genau wie die bei $\mathbf{b}_1^*$ angesetzt werden, ausschließlich die *späteren* (natürlich ebenfalls veränderten) $\mathbf{b}_i^*$ beachtend.

Hat man alle $\mathbf{b}_i^*$ in die Positionen $\pm\mathbf{e}_i$ gedreht, ist es nun leicht, durch eventuelle Halbkreisdrehungen erst in der $\mathbf{e}_1$-$\mathbf{e}_2$-Ebene, dann in der $\mathbf{e}_2$-$\mathbf{e}_3$-Ebene, usw., sicherzustellen, dass die Endkonfiguration $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \pm\mathbf{e}_n)$ erreicht wird.

Die insgesamt durchlaufenen Deformationen und Rotationen lassen sich zu *einer einzigen* stetigen Basis-Schar zusammenfassen.

Es bleibt aber noch eine *logische Lücke* dieser determinantenfreien Definition der positiven Orientierung im $\mathbb{R}^n$:

Wir müssen nachweisen, dass eine stetige Basis-Schar nicht von $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{e}_n)$ zu $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, -\mathbf{e}_n)$ führen kann, es also wirklich *zwei* Orientierungen gibt und die Definition *eindeutig* ist.

Wir nennen zwei Basen des $\mathbb{R}^n$ *deformationsäquivalent*, wenn eine stetige Schar von Basen von der einen zur anderen führt. Wir haben also bisher nur gesehen: Jede Basis ist deformantionsäquivalent zu $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{e}_n)$ *oder* zu $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, -\mathbf{e}_n)$.

Prinzipiell könnte man sich *vor* Entwicklung der Determinantentheorie mit der anschaulichen Definition der beiden Orientierungen durch Deformationsäquivalenz begnügen und sagen:

Die Stetigkeit der Determinanten wird *nachträglich* zeigen, dass es wirklich *zwei verschiedene* Orientierungen gibt.

Aber wir wollen hier dennoch *unmittelbar elementar und determinantenfrei* nachweisen:

Keine stetige Basis-Schar $(\mathbf{b}_1(t), \ldots, \mathbf{b}_n(t))_{0 \le t \le 1}$ *führt von der Standardbasis zu* $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, -\mathbf{e}_n)$.

Vorbemerkungen vor dem eigentlichen Beweis dieser Ausage:

1) Gegeben irgendeine *Spiegelung* $\mathbf{x} \mapsto (E - 2\mathbf{n}\mathbf{n}^T)\mathbf{x}$ $(\mathbf{x} \in \mathbb{R}^n)$, ist die Bildbasis $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n) := E - 2\mathbf{n}\mathbf{n}^T$ deformationsäquivalent zu $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, -\mathbf{e}_n)$. Dazu muss nur die Einheitsnormale $\mathbf{n}$ durch Drehung stetig in $\mathbf{e}_n$ überführt werden.

2) Eine stetige Deformationsschar lässt sich bei Bedarf lokal ein bisschen abändern. Denn jedes $\mathbf{b}_i$ liegt immer außerhalb der von den anderen aufgespannten Hyperebene. Also gibt's einen *senkrechten Mindestabstand* $d > 0$ zur jeweilige Hyerebene der anderen, den kein Basiselement je unterschreitet. Das erlaubt praktisch beliebige hinreichend kleine *lokale Verlaufsabänderungen*.

3) Stehen am Anfang und Ende einer stetigen Basis-Schar ONBs, kann man o.B.d.A. voraussetzen, dass die Schar insgesamt aus ONBs besteht (Gram/Schmidt!). Nach 2) dürfen die ONBs lokal

bei Bedarf ein bisschen verdreht werden.

Nun zum eigentlichen **Beweis:**

Wir gehen aus von einer stetigen Deformationsschar, welche von der Basis $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{e}_n)$ zur Basis $(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, -\mathbf{e}_n)$ führt und werden per Induktion über die Anzahl der Dimensionen zeigen, dass dies einen *Widerspruch* beinhaltet, also eine solche Schar *nicht existiert*.

Die geometrische Grundidee des Beweises ist einfach:

Der n-te Vektor $\mathbf{b}_n(t)$ der Deformationsschar muss irgendwie stetig vom Startwert $\mathbf{e}_n$ zum entgegengesetzten Endwert $-\mathbf{e}_n$ übergehen. O.B.d.A. kann man von einer *ONB*-Schar $(\mathbf{b}_1(t), \ldots, \mathbf{b}_n(t))$ ausgehen. Nach Induktionsannahme sind im $\mathbb{R}^{n-1}$ zwei durch Spiegelung auseinander hervorgehende Basen *nicht* deformationsäquivalent. Wir betrachten die *Nulldurchgänge* von $\mathbf{b}_n(t) \cdot \mathbf{e}_n$ und zeigen, dass dabei gleichzeitig die Projektionen der anderen $\mathbf{b}_i$ in die von $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$ aufgespannte Hyperebene einer Spiegelung unterworfen werden. Da die Anzahl der Nulldurchgänge *ungerade* ist, widerspricht dies der Induktionsannahme.

Induktionsanfang:

Im Falle $n = 1$ ist nichts zu zeigen. Wir betrachten illustrationshalber auch noch $n = 2$. Ist $(\mathbf{b}_1(t), \mathbf{b}_2(t))$ eine ONB-Schar, die stetig von $(\mathbf{e}_1, \mathbf{e}_2)$ $(t = 0)$ zu $(\mathbf{e}_1, -\mathbf{e}_2)$ $(t = 1)$ führt, kann man durch Abänderung erreichen, dass $\mathbf{b}_2(t)$ nur endlich oft (aber mindestens einmal!) die $\mathbf{e}_1$-Achse durchtritt $(\mathbf{b}_2(t) \cdot \mathbf{e}_2 = 0)$, und zwar mit Vorzeichen-Umschlag bei $\mathbf{b}_2(t) \cdot \mathbf{e}_2$, und dabei $\mathbf{b}_1(t)$ orthogonal zu $\mathbf{b}_2(t)$ bleibt. (Man kann ja den Gesamt-Verlauf von $\mathbf{b}_2(t)$ so abändern, dass er aus gleichmäßig durchlaufenen Kreisbogenstücken zusammengesetzt ist.)

Also hat $\mathbf{b}_1 \cdot \mathbf{e}_1$ beim betrachteten Nulldurchgang einen Vorzeichenwechsel und damit die Projektion $(\mathbf{b}_1 \cdot \mathbf{e}_1)\mathbf{e}_1$ von $\mathbf{b}_1$ in die „Hyperebene" $\mathbf{e}_2^\perp$ eine *Orientierungsumkehr*. Da es insgesamt eine ungerade Anzahl von Nulldurchgängen geben muss, aber die Projektion am Anfang und am Ende gleich $\mathbf{e}_1$ ist, ergibt sich ein *Widerspruch*.

Induktionsschluss:

Nun sei eine stetige Basis-Schar $(\mathbf{b}_1(t), \dots, \mathbf{b}_n(t))$ gegeben mit $\mathbf{b}_i(0) = \mathbf{b}_i(1) = \mathbf{e}_i$ $(1 \le i \le n-1)$, $\mathbf{b}_n(0) = \mathbf{e}_n = -\mathbf{b}_n(1)$. O.B.d.A. setzen wir voraus, dass die $\mathbf{b}_1, \dots, \mathbf{b}_n$ immer eine ONB sind.

Wir verfolgen die $\mathbf{b}_n$-Komponente in $\mathbf{e}_n$-Richtung, also $\mathbf{b}_n(t) \cdot \mathbf{e}_n$, und zugleich die orthogonalen Projektionen der $\mathbf{b}_i(t)$ in die Hyperebene $\mathbf{e}_n^\perp$, also $\mathbf{b}_i^*(t) = \mathbf{b}_i(t) - (\mathbf{b}_i(t) \cdot \mathbf{e}_n)\mathbf{e}_n$ $(1 \le i \le n-1)$.

So lange $\mathbf{b}_n(t) \cdot \mathbf{e}_n \ne 0$, bilden die $\mathbf{b}_i^*$ eine Basis des von $\mathbf{e}_1, \dots, \mathbf{e}_{n-1}$ aufgespannten (mit $\mathbb{R}^{n-1}$ identifizierbaren) Raumes $\mathbf{e}_n^\perp$; denn gilt $\sum_{1 \le i < n} \lambda_i \mathbf{b}_i^* = \mathbf{0}$, folgt durch Skalarmultiplikation mit $\mathbf{b}_n$:
$\left(\sum_{1 \le i < n} \lambda_i (\mathbf{b}_i \cdot \mathbf{e}_n)\right) \underbrace{(\mathbf{e}_n \cdot \mathbf{b}_n)}_{\ne 0} = 0$, also $\sum_{1 \le i < n} \lambda_i \mathbf{b}_i = \mathbf{0}$.

Wir können davon ausgehen, dass es nur *endlich viele* Stellen t_0 mit $\mathbf{b}_n(t_0) \cdot \mathbf{e}_n = 0$ gibt: Die Menge dieser Stellen ist beschränkt und abgeschlossen; endlich viele ε-Kugeln ($\varepsilon > 0$ beliebig) um die entsprechenden $\mathbf{b}_n(t)$ überdecken alle diese Punkte; durch lokales Verdrehen von $\mathbf{b}_n(t)$ innerhalb dieser Kugeln ändern wir die Bahn so ab, dass in jeder ausgewählten Kugel maximal eine Stelle mit $\mathbf{b}_n(t_0) \cdot \mathbf{e}_n = 0$ vorkommt.

Wir können voraussetzen, dass in einer kleinen Umgebung jedes solchen Parameterwertes t_0 die Basis-Deformation eine *Rotation in der* $\mathbf{e}_n$-$\mathbf{b}_n(t_0)$-*Ebene* ist. Also $\mathbf{b}_n(t_0 + \tau) = \cos\tau\, \mathbf{b}_n^0 \pm \sin\tau\, \mathbf{e}_n$ mit $\mathbf{b}_n^0 = \mathbf{b}_n(t_0)$ ($\pm$ je nach Durchlaufsinn). Die anderen $\mathbf{b}_i$ zerlegen wir orthogonal gemäß $\mathbf{b}_i = \mathbf{b}_i^0 + (\mathbf{b}_i \cdot \mathbf{b}_n^0)\mathbf{b}_n^0 + (\mathbf{b}_i \cdot \mathbf{e}_n)\mathbf{e}_n$, so dass also $\mathbf{b}_i^* = \mathbf{b}_i^0 + (\mathbf{b}_i \cdot \mathbf{b}_n^0)\mathbf{b}_n^0$.

Rotation in der $\mathbf{e}_n$-$\mathbf{b}_n^0$-Ebene bedeutet:

$$\begin{aligned}\mathbf{b}_i(t_0 + \tau) &= \mathbf{b}_i^0 + \left(\cos\tau(\mathbf{b}_i(t_0) \cdot \mathbf{b}_n^0) \mp \sin\tau(\mathbf{b}_i(t_0) \cdot \mathbf{e}_n)\right)\mathbf{b}_n^0 + \\ &\quad + \left(\pm\sin\tau(\mathbf{b}_i(t_0) \cdot \mathbf{b}_n^0) + \cos\tau(\mathbf{b}_i(t_0) \cdot \mathbf{e}_n)\right)\mathbf{e}_n \\ &= \mathbf{b}_i^0 \mp \sin\tau(\mathbf{b}_i(t_0) \cdot \mathbf{e}_n)\mathbf{b}_n^0 + \cos\tau(\mathbf{b}_i(t_0) \cdot \mathbf{e}_n)\mathbf{e}_n\end{aligned}$$

mit *konstantem* $\mathbf{b}_i^0 = \mathbf{b}_i^0(t_0) \perp \mathbf{b}_n^0, \mathbf{e}_n$. Somit $\mathbf{b}_i^*(t_0 + \tau) = \mathbf{b}_i^0 \mp \sin\tau(\mathbf{b}_i(t_0) \cdot \mathbf{e}_n)\mathbf{b}_n^0$ $(1 \le i < n)$. D.h.: Die $\mathbf{b}_i^*$ werden an der Hyperebene $\mathbf{b}_n^{0\perp}$ (*innerhalb* der Hyperebene $\mathbf{e}_n^\perp$) exakt gespiegelt durch diese ebene Rotation. Laut Induktionsannahme ändert sich die ($n-1$-dimensionale) Orientierung der $\mathbf{b}_i^*$. Da es eine *ungerade Anzahl* von Nulldurchgängen bei $\mathbf{b}_n(t) \cdot \mathbf{e}_n$ gibt, folgt ein *Widerspruch* zu $\mathbf{b}_i^*(0) = \mathbf{b}_i^*(1) = \mathbf{e}_i$ $(1 \le i < n)$. ∎

Eleganter und einfacher wird alles, wenn wir *Determinanten* nutzen: Bei einer stetigen Basis-Schar wechselt die Determinante nie ihr Vorzeichen. Also führt keine Schar von der Standard-Basis zu $(\mathbf{e}_1, \dots, \mathbf{e}_{n-1}, -\mathbf{e}_n)$.

Plausibel, je nach Vorzeichen der Determinante von *positiv* bzw. *negativ* orientierten Basen zu sprechen.

Dennoch zusätzlich wichtig: Alle positiv (bzw. alle negativ) orientierten Basen sind durch stetige Deformation (ONBs durch stetige *Rotation*) ineinander überführbar.

Ergebnis der „erfolgreichen" Bemühungen, ohne Determinanten auszukommen:

Orientierung wird am besten erfasst durch die Denkökonomie des Determinantenbegriffs, eines etwas verwickelten, aber schon altehrwürdigen Begleiters aller Überlegungen zur Lösung linearer Gleichungssysteme.