Kapitel 7 Multilineare Algebra

§ 1 Alternierende Formen

Inhalt:

Alternierende Bilinearformen, äußeres (oder Dach-)produkt, Differentialformen, Zusammenhang zwischen äußerem Produkt und Vektorprodukt, alternierende Multilinearformen, Determinantenformen und Determinanten, Regel von Sarrus, Produkte von 1-Formen und 2-Formen (im $\mathbb{R}^3$).

Im Folgenden sei der Körper K stets $= \mathbb{R}$ oder $= \mathbb{C}$.

Definition:

Sei V ein endlich-dimensionaler K-Vektorraum. Eine Bilinearform $\varphi : V \times V \to K$ heißt *alternierend* oder *schiefsymmetrisch*, falls für alle $v, w \in V$ gilt:

$$\varphi(v, w) = -\varphi(w, v).$$

Beispiele.

1. Sei $V = K^n$, $A \in M_{n,n}(K)$ und $\varphi_A : K^n \times K^n \to K$ definiert durch $\varphi_A(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot A \cdot \mathbf{w}^t$. Dann ist φ_A eine Bilinearform (und umgekehrt sieht jede Bilinearform auf dem K^n so aus). Die Einträge a_{ij} in A sind durch $a_{ij} = \varphi_A(\mathbf{e}_i, \mathbf{e}_j)$ gegeben.

   Es ist $\mathbf{v} \cdot A \cdot \mathbf{w}^t = (\mathbf{v} \cdot A \cdot \mathbf{w}^t)^t = \mathbf{w} \cdot A^t \cdot \mathbf{v}^t$, und φ_A ist genau dann alternierend, wenn $\mathbf{v} \cdot A \cdot \mathbf{w}^t = -\mathbf{w} \cdot A \cdot \mathbf{v}^t = \mathbf{w} \cdot (-A) \cdot \mathbf{v}^t$ ist, also $A^t = -A$. Das ist genau dann der Fall, wenn gilt:

   $$A = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ -a_{12} & 0 & a_{23} & & a_{2n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 & a_{n-1,n} \\ -a_{1n} & -a_{2n} & \cdots & -a_{n-1,n} & 0 \end{pmatrix}.$$

   Solche Matrizen nennt man *schiefsymmetrisch*.

Ist etwa $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, so ist

$$\varphi_A((v_1, v_2), (w_1, w_2)) = (v_1, v_2) \cdot \begin{pmatrix} w_2 \\ -w_1 \end{pmatrix} = v_1 w_2 - v_2 w_1 = \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}.$$

2. Uns interessiert ganz besonders der Fall $n = 3$. Die Matrizen

$$E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; E_{31} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

sind schiefsymmetrisch, und durch

$$S(a_1, a_2, a_3) := a_1 \cdot E_{23} + a_2 \cdot E_{31} + a_3 \cdot E_{12} = \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_2 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix}$$

wird jedem Vektor $\mathbf{a} \in \mathbb{R}^3$ eine schiefsymmetrische Matrix $S(\mathbf{a}) \in M_{3,3}(\mathbb{R})$ zugeordnet. Diese Abbildung ist bijektiv, ihre Umkehrung liefert ein Koordinatensystem für den 3-dimensionalen Vektorraum aller schiefsymmetrischen (3×3)-Matrizen.

Die Menge aller alternierenden Bilinearformen auf einem Vektorraum V bezeichnen wir mit $A^2(V)$. Offensichtlich ist dann $A^2(\mathbb{R}^3)$ ein 3-dimensionaler $\mathbb{R}$-Vektorraum. Die Formen

$$\varphi_{ij}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot E_{ij} \cdot \mathbf{y}^t, \; 1 \le i < j \le 3,$$

bilden eine Basis von $A^2(\mathbb{R}^3)$. Dabei setzen wir $E_{13} := -E_{31}$.

Offensichtlich ist

$$\mathbf{v} \times \mathbf{w} = (\varphi_{23}(\mathbf{v}, \mathbf{w}), \varphi_{31}(\mathbf{v}, \mathbf{w}), \varphi_{12}(\mathbf{v}, \mathbf{w}))\,.$$

Daraus kann man sofort ableiten, daß das Vektorprodukt bilinear und alternierend ist (allerdings mit Werten in $\mathbb{R}^3$, es ist also keine Bilinear**form**).

3. Sind $f, g \in L(K^n, K)$ zwei Linearformen, so ist ihr *äußeres Produkt* (oder *Dachprodukt* oder *Grassmann-Produkt*) $f \wedge g$ eine Bilinearform auf K^n, definiert durch

$$(f \wedge g)(\mathbf{v}, \mathbf{w}) := f(\mathbf{v}) \cdot g(\mathbf{w}) - g(\mathbf{v}) \cdot f(\mathbf{w}).$$

Die Bilinearität folgt sofort aus der Linearität von f und g. Darüber hinaus ist $f \wedge g$ alternierend:

$$\begin{aligned} (f \wedge g)(\mathbf{v}, \mathbf{w}) &= f(\mathbf{v}) \cdot g(\mathbf{w}) - g(\mathbf{v}) \cdot f(\mathbf{w}) \\ &= -\big(f(\mathbf{w}) \cdot g(\mathbf{v}) - g(\mathbf{w}) \cdot f(\mathbf{v})\big) \\ &= -(f \wedge g)(\mathbf{w}, \mathbf{v}). \end{aligned}$$

Für das Rechnen mit dem Dachprodukt notieren wir noch die folgenden Regeln:

(a) $f \wedge g = -g \wedge f$, also insbesondere stets $f \wedge f = 0$.

(b) $f \wedge (g_1 + g_2) = f \wedge g_1 + f \wedge g_2$.

(c) $(f_1 + f_2) \wedge g = f_1 \wedge g + f_2 \wedge g$.

(d) $(r \cdot f) \wedge g = f \wedge (r \cdot g) = r \cdot (f \wedge g)$.

Die einfachen Beweise führt man, indem man die Formen auf beliebige Vektor-Paare anwendet, z.B.

$$\begin{aligned}
(f \wedge (g_1 + g_2))(\mathbf{v}, \mathbf{w}) &= f(\mathbf{v}) \cdot ((g_1 + g_2)(\mathbf{w}) - (g_1 + g_2)(\mathbf{v}) \cdot f(\mathbf{w}) \\
&= f(\mathbf{v}) \cdot (g_1(\mathbf{w}) + g_2(\mathbf{w})) - (g_1(\mathbf{v}) + g_2(\mathbf{v})) \cdot f(\mathbf{w}) \\
&= f(\mathbf{v}) \cdot g_1(\mathbf{w}) + f(\mathbf{v}) \cdot g_2(\mathbf{w}) - g_1(\mathbf{v}) \cdot f(\mathbf{w}) - g_2(\mathbf{v}) \cdot f(\mathbf{w}) \\
&= [f(\mathbf{v}) \cdot g_1(\mathbf{w}) - g_1(\mathbf{v}) \cdot f(\mathbf{w})] + [f(\mathbf{v}) \cdot g_2(\mathbf{w}) - g_2(\mathbf{v}) \cdot f(\mathbf{w}) \\
&= (f \wedge g_1)(\mathbf{v}, \mathbf{w}) + (f \wedge g_2)(\mathbf{v}, \mathbf{w}) \\
&= (f \wedge g_1 + f \wedge g_2)(\mathbf{v}, \mathbf{w}).
\end{aligned}$$

Im Falle $K = \mathbb{R}$ haben wir jedem Vektor $\mathbf{a} \in \mathbb{R}^n$ eine Linearform $\lambda_\mathbf{a} \in L(\mathbb{R}^n, \mathbb{R})$ zugeordnet, durch

$$\lambda_\mathbf{a}(\mathbf{x}) = \mathbf{a} \bullet \mathbf{x} = \mathbf{a} \cdot \mathbf{x}^t \; (= \mathbf{x} \cdot \mathbf{a}^t) = a_1 x_1 + \cdots + a_n x_n.$$

Dann ist

$$\begin{aligned}
\lambda_\mathbf{a} \wedge \lambda_\mathbf{b}(\mathbf{x}, \mathbf{y}) &= (\mathbf{a} \cdot \mathbf{x}^t) \cdot (\mathbf{b} \cdot \mathbf{y}^t) - (\mathbf{b} \cdot \mathbf{x}^t) \cdot (\mathbf{a} \cdot \mathbf{y}^t) \\
&= \mathbf{x} \cdot (\mathbf{a}^t \cdot \mathbf{b} - \mathbf{b}^t \cdot \mathbf{a}) \cdot \mathbf{y}^t.
\end{aligned}$$

Also ist $A = \mathbf{a}^t \cdot \mathbf{b} - \mathbf{b}^t \cdot \mathbf{a}$ die Matrix zu der Bilinearform $\lambda_\mathbf{a} \wedge \lambda_\mathbf{b}$.

Die Linearform $\lambda_{\mathbf{e}_\nu}$ (mit $\lambda_{\mathbf{e}_\nu}(\mathbf{v}) = \mathbf{e}_\nu \bullet \mathbf{v} = v_\nu$) haben wir auch mit dx_ν bezeichnet, für $\nu = 1, \ldots, n$. Jede Linearform auf dem $\mathbb{R}^n$ hat daher die Form

$$\lambda_\mathbf{a} = a_1\, dx_1 + a_2\, dx_2 + \cdots + a_n\, dx_n.$$

Man spricht auch von einer *Differentialform der Dimension* 1 oder kurz von einer 1-*Form*, speziell von der *kanonischen* 1-*Form* zum Vektor $\mathbf{a}$.

Im $\mathbb{R}^3$ folgt nun:

$$\begin{aligned}
(dx_\nu \wedge dx_\mu)(\mathbf{v}, \mathbf{w}) &= v_\nu w_\mu - v_\mu w_\nu \\
&= \mathbf{v} \cdot E_{\nu\mu} \cdot \mathbf{w}^t,
\end{aligned}$$

also $dx_\nu \wedge dx_\mu = \varphi_{\nu\mu}$ für alle $\nu < \mu$ (und $dx_\nu \wedge dx_\nu = 0$).

Jede alternierende Bilinearform auf dem $\mathbb{R}^3$ hat daher die Form

$$\Lambda_\mathbf{a} = a_1\, dx_2 \wedge dx_3 + a_2\, dx_3 \wedge dx_1 + a_3\, dx_1 \wedge dx_2.$$

Man nennt $\Lambda_{\mathbf{a}}$ eine *Differentialform der Dimension* 2 (oder kurz eine 2-*Form*), speziell auch die *kanonische* 2-*Form* zum Vektor **a**. Man beachte, daß dies nur im Falle $n = 3$ geht!

Ist n beliebig, so hat eine alternierende Bilinearform auf dem $\mathbb{R}^n$ die Form

$$\varphi = \sum_{1 \le i < j \le n} a_{ij}\, dx_i \wedge dx_j.$$

Auch in diesem Falle spricht man von einer 2-Form. Die Anzahl der Koeffizienten beträgt

$$\frac{n^2 - n}{2} = \frac{n(n-1)}{2} = \frac{n!}{2!(n-2)!} = \binom{n}{2}.$$

Damit $(n^2 - n)/2 = n$ ist, muß $n^2 - 3n = 0$, also $n = 3$ sein. Das ist der Grund, warum man nur für $n = 3$ einen Isomorphismus zwischen dem $\mathbb{R}^n$ und dem Raum der alternierenden Bilinearformen auf dem $\mathbb{R}^n$ erhält, und warum es nur im Falle $n = 3$ ein Vektorprodukt gibt.

Satz (Zusammenhang zwischen äußerem Produkt und Vektorprodukt)

1. *Ist* $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ *und* $\Lambda_{\mathbf{a}} = a_1\, dx_2 \wedge dx_3 + a_2\, dx_3 \wedge dx_1 + a_3\, dx_1 \wedge dx_2$ *die zugehörige* 2*-Form, so ist*

$$\Lambda_{\mathbf{a}}(\mathbf{v}, \mathbf{w}) = \mathbf{a} \bullet (\mathbf{v} \times \mathbf{w}), \ \textit{für } \mathbf{v}, \mathbf{w} \in \mathbb{R}^3.$$

2. *Sind* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, *so ist* $\quad \lambda_{\mathbf{a}} \wedge \lambda_{\mathbf{b}} = \Lambda_{\mathbf{a} \times \mathbf{b}}$.

Beweis: 1) Es ist $\Lambda_{\mathbf{a}} = a_1\varphi_{23} + a_2\varphi_{31} + a_3\varphi_{12}$, also

$$\begin{aligned}
\Lambda_{\mathbf{a}}(\mathbf{v}, \mathbf{w}) &= a_1\varphi_{23}(\mathbf{v}, \mathbf{w}) + a_2\varphi_{31}(\mathbf{v}, \mathbf{w}) + a_3\varphi_{12}(\mathbf{v}, \mathbf{w}) \\
&= (a_1, a_2, a_3) \bullet (\varphi_{23}(\mathbf{v}, \mathbf{w}), \varphi_{31}(\mathbf{v}, \mathbf{w}), \varphi_{12}(\mathbf{v}, \mathbf{w})) \\
&= \mathbf{a} \bullet (\mathbf{v} \times \mathbf{w}).
\end{aligned}$$

2) Es ist $\varphi_{\nu\mu}(\mathbf{a}, \mathbf{b}) = a_\nu b_\mu - a_\mu b_\nu$, also

$$\begin{aligned}
\lambda_{\mathbf{a}} \wedge \lambda_{\mathbf{b}} &= \sum_{\nu,\mu} a_\nu b_\mu\, dx_\nu \wedge dx_\mu \\
&= \sum_{\nu<\mu} (a_\nu b_\mu - a_\mu b_\nu)\, dx_\nu \wedge dx_\mu \\
&= \sum_{\nu<\mu} \varphi_{\nu\mu}(\mathbf{a}, \mathbf{b})\varphi_{\nu\mu} \\
&= \Lambda_{(\varphi_{23}(\mathbf{a},\mathbf{b}), \varphi_{31}(\mathbf{a},\mathbf{b}) \varphi_{12}(\mathbf{a},\mathbf{b}))} \ = \ \Lambda_{\mathbf{a} \times \mathbf{b}}.
\end{aligned}$$

Damit ist alles gezeigt. ■

Sei nun $\varphi : K^2 \times K^2 \to K$ eine beliebige alternierende Bilinearform. Dann ist

$$\begin{aligned}\varphi(\mathbf{x},\mathbf{y}) &= \varphi(x_1\mathbf{e}_1 + x_2\mathbf{e}_2, y_1\mathbf{e}_1 + y_2\mathbf{e}_2)\\ &= x_1y_2\cdot\varphi(\mathbf{e}_1,\mathbf{e}_2) + x_2y_1\cdot\varphi(\mathbf{e}_2,\mathbf{e}_1)\\ &= (x_1y_2 - x_2y_1)\cdot\varphi(\mathbf{e}_1,\mathbf{e}_2)\\ &= \det(\vec{x},\vec{y})\cdot\varphi(\mathbf{e}_1,\mathbf{e}_2).\end{aligned}$$

Es gibt also genau eine alternierende Bilinearform Δ_2 auf K^2 mit $\Delta_2(\mathbf{e}_1,\mathbf{e}_2) = 1$, nämlich

$$\Delta_2(\mathbf{x},\mathbf{y}) = \det(\vec{x},\vec{y}).$$

Man nennt Δ_2 die *Determinantenform* auf K^2.

Definition:

Eine alternierende n-Form auf einem n-dimensionalen K-Vektorraum V ist eine Funktion $\varphi : V \times \ldots \times V \to K$ mit folgenden Eigenschaften:

1. $\varphi(x_1,\ldots,x_n)$ ist in jedem Argument linear (man nennt φ daher auch n-fach *multilinear*).
2. Für $i = 1,\ldots,n-1$ ist stets

$$\varphi(x_1,\ldots,x_i,x_{i+1},\ldots,x_n) = -\varphi(x_1,\ldots,x_{i+1},x_i,\ldots,x_n).$$

Satz (Eigenschaften alternierender n-Formen)

1. *Für jede Permutation $\sigma \in S_n$ ist*

$$\varphi(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = \operatorname{sign}(\sigma)\cdot\varphi(x_1,\ldots,x_n).$$

2. *Sind $x_1,\ldots,x_n \in V$ linear abhängig, so ist $\varphi(x_1,\ldots,x_n) = 0$.*

Beweis: Die erste Behauptung folgt daraus, daß jede Permutation in Transpositionen zerlegt werden kann. Insbesondere ist $\varphi(x_1,\ldots,x_n) = 0$, falls zwei der Argumente gleich sind. Daraus folgt die zweite Behauptung. ■

Zum Beispiel ist jede alternierende 2-Form auf dem K^2 ein Vielfaches der Determinantenform Δ_2.

Satz

Es gibt genau eine alternierende n-Form Δ_n auf dem K^n mit

$$\Delta_n(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1.$$

Beweis:
Eindeutigkeit: Wir nehmen an, daß Δ_n existiert. Setzen wir irgendwelche Vektoren $\mathbf{a}_i = (a_{1i}, \ldots, a_{ni})$ ein, so erhalten wir:

$$\begin{aligned}
\Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) &= \Delta_n\Big(\sum_{i_1=1}^{n} a_{i_1,1}\mathbf{e}_{i_1}, \ldots, \sum_{i_n=1}^{n} a_{i_n,n}\mathbf{e}_{i_n}\Big) \\
&= \sum_{i_1} \cdots \sum_{i_n} a_{i_1,1} \cdots a_{i_n,n}\, \Delta_n(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}\,.
\end{aligned}$$

Damit ist Δ_n festgelegt.

Zum Beweis der **Existenz** müssen wir nur zeigen, daß durch

$$\Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

eine alternierende n-Form auf dem K^n gegeben ist. Daß Δ_n multilinear ist, ist offensichtlich. Den Rest der Behauptung zeigen wir in mehreren Schritten:

1. Schritt: Ist $\mathbf{a}_1 = \mathbf{a}_2$, so ist $\Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$.

Es sei S_n^+ die Menge aller geraden Permutationen aus S_n, und $\tau := (1, 2)$ die Vertauschung der ersten zwei Zahlen. Durchläuft σ alle Elemente von S_n^+, so durchläuft $\sigma \circ \tau$ alle ungeraden Permutationen.[1] Daher ist

$$\begin{aligned}
\Delta_n(\mathbf{a}_1, \mathbf{a}_1, \mathbf{a}_3, \ldots, \mathbf{a}_n) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} a_{\sigma(2),1} a_{\sigma(3),3} \cdots a_{\sigma(n),n} \\
&= \sum_{\sigma \in S_n^+} a_{\sigma(1),1} a_{\sigma(2),1} a_{\sigma(3),3} \cdots a_{\sigma(n),n} \\
&\quad - \sum_{\sigma \in S_n^+} a_{\sigma(2),1} a_{\sigma(1),1} a_{\sigma(3),3} \cdots a_{\sigma(n),n} \;=\; 0\,.
\end{aligned}$$

Genauso folgt, daß $\Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ schon verschwindet, wenn zwei beliebige Argumente gleich sind.

[1] Offensichtlich ist $\sigma \circ \tau$ ungerade. Ist umgekehrt ϱ eine beliebige ungerade Permutation, so ist $\sigma := \varrho \circ \tau^{-1}$ gerade und $\sigma \circ \tau = \varrho$.

2. Schritt: Für alle i, k und alle $\lambda \in K$ ist

$$\Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_i + \lambda \mathbf{a}_k, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n) = \Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n).$$

Das ergibt sich sofort aus Schritt 1.

3. Schritt: Jetzt folgt, daß Δ_n alternierend ist, denn die Vertauschung zweier Argumente kann man aus elementaren Transformationen, wie sie in Schritt 2 beschrieben wurden, gewinnen. Dabei ergibt sich ein Vorzeichenwechsel:

$$\begin{aligned}
\Delta_n(\ldots, x_i, \ldots, x_k, \ldots) &= -\Delta_n(\ldots, x_i, \ldots, -x_k, \ldots) \\
&= -\Delta_n(\ldots, x_i, \ldots, x_i - x_k, \ldots) \\
&= -\Delta_n(\ldots, x_i - (x_i - x_k), \ldots, x_i - x_k, \ldots) \\
&= -\Delta_n(\ldots, x_k, \ldots, x_i - x_k, \ldots) \\
&= -\Delta_n(\ldots, x_k, \ldots, (x_i - x_k) + x_k, \ldots) \\
&= -\Delta_n(\ldots, x_k, \ldots, x_i, \ldots). \qquad \blacksquare
\end{aligned}$$

Definition:

Die eindeutig bestimmte alternierende n-Form Δ_n auf dem K^n mit

$$\Delta_n(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$$

heißt die *Determinantenform* auf dem K^n. Ist $A = (\vec{a}_1, \ldots, \vec{a}_n) \in M_{n,n}(K)$, so heißt

$$\det(A) := \Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_n)$$

die *Determinante* von A.

Eigenschaften der Determinante:

1. $\det(A)$ *ist multilinear und alternierend in den Spalten von* A.
2. *Ist* $\operatorname{rg}(A) < n$, *so ist* $\det(A) = 0$.
3. *Es ist* $\det(E_n) = 1$.
4. *Ist* $A = (a_{ij})$, *so ist* $\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\, a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$.

Der Beweis der ersten drei Aussagen ergibt sich direkt aus der Definition der Determinante. Zur letzten Aussage stellen wir folgende Überlegungen an:

- Da das gewöhnliche Produkt in K kommutativ ist und die Werte $\sigma(1), \ldots, \sigma(n)$ wieder genau die Zahlen $1, \ldots, n$ durchlaufen, ist
$$a_{\sigma(1),1} \cdots a_{\sigma(n),n} = a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)}.$$
- Weil $\operatorname{sign}(\sigma) \cdot \operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma \circ \sigma^{-1}) = 1$ ist, ist $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$.
- Die Abbildung $I : S_n \to S_n$ mit $I(\sigma) = \sigma^{-1}$ ist bijektiv, denn es ist offensichtlich $I \circ I = \mathrm{id}$. Durchläuft also σ alle Elemente von S_n, so durchläuft gleichzeitig auch σ^{-1} alle Elemente von S_n.

Mit $\tau := \sigma^{-1}$ folgt nun:

$$\begin{aligned} \Delta_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} \\ &= \sum_{\tau \in S_n} \operatorname{sign}(\tau) a_{1,\tau(1)} \cdots a_{n,\tau(n)}. \end{aligned}$$

Jetzt kann man aber wieder σ statt τ schreiben und erhält so die gewünschte Formel.

Beispiel.

Ist $n = 3$ und $\mathbf{a}_i = (a_{1,i}, a_{2,i}, a_{3,i})$ für $i = 1, 2, 3$, so ist

$$\begin{aligned} \det(\vec{a}_1, \vec{a}_2, \vec{a}_3) &= \Delta_n(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \\ &= \sum_{\sigma \in S_3} \operatorname{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot a_{3,\sigma(3)} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{aligned}$$

Die so gewonnene Berechnungsformel heißt die *Sarrus'sche Regel*. Es gibt ein Schema, nach dem man sie sich leicht merken kann. Man schreibe die Matrix–Elemente in der folgenden Form auf:

$$\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Die Produkte der Elemente in den „Hauptdiagonalen“ werden mit „+“ versehen, die Produkte der Elemente in den „Nebendiagonalen“ werden mit „−“ versehen, und schließlich werden alle Produkte aufsummiert:

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

Determinante der transponierten Matrix

Sei $A \in M_{n,n}(K)$. Dann gilt:

$$\det(A^t) = \det(A).$$

BEWEIS: Sei $A = (a_{ij})$. Mit $b_{ij} := a_{ji}$ erhält man:

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n} \\
&= \sum_{\sigma^{-1} \in S_n} \operatorname{sign}(\sigma^{-1}) b_{1,\sigma^{-1}(1)} \cdots b_{n,\sigma^{-1}(n)} \\
&= \det(A^t).
\end{aligned}$$

∎

Also ist die Determinante auch in den Zeilen der Matrix multilinear und alternierend.

Wir kommen jetzt noch einmal auf die spezielle Situation des $\mathbb{R}^3$ zurück. Hier kennen wir schon alle 1-Formen, sie sind von der Gestalt

$$\lambda_{\mathbf{a}} = a_1\,dx_1 + a_2\,dx_2 + a_3\,dx_3.$$

Wir kennen auch alle 2-Formen, sie sind von der Gestalt

$$\Lambda_{\mathbf{a}} = a_1\,dx_2 \wedge dx_3 + a_2\,dx_3 \wedge dx_1 + a_3\,dx_1 \wedge dx_2.$$

Damit ist $\lambda_{\mathbf{a}}(\mathbf{v}) = \mathbf{a} \bullet \mathbf{v}$ und $\Lambda_{\mathbf{a}}(\mathbf{v}, \mathbf{w}) = \mathbf{a} \bullet (\mathbf{v} \times \mathbf{w})$. Da eine alternierende n-Form auf dem $\mathbb{R}^3$ für $n \geq 4$ nur noch den Wert 0 ergibt (denn die 4 Argumente müssen zwangsläufig linear abhängig sein), brauchen wir bloß noch die alternierenden 3-Formen zu untersuchen.

Definition:

Sind f_1, f_2, f_3 Linearformen auf dem $\mathbb{R}^3$, so erhält man eine alternierende 3-Form $f_1 \wedge f_2 \wedge f_3$ auf dem $\mathbb{R}^3$ durch

$$(f_1 \wedge f_2 \wedge f_3)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) := \det\big(f_i(\mathbf{v}_j) \,:\, i, j = 1, \ldots, 3\big).$$

Daß $f_1 \wedge f_2 \wedge f_3$ tatsächlich eine alternierende 3-Form ist, folgt aus den Eigenschaften der Determinante.

Speziell ist

$$(dx_1 \wedge dx_2 \wedge dx_3)(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \det(\vec{a}_1, \vec{a}_2, \vec{a}_3),$$

also $dx_1 \wedge dx_2 \wedge dx_3 = \Delta_3$.

3-Formen und Permutationen

$$f_{\sigma(1)} \wedge f_{\sigma(2)} \wedge f_{\sigma(3)} = \operatorname{sign}(\sigma)\, f_1 \wedge f_2 \wedge f_3, \ \textit{für alle } \sigma \in S_3.$$

Beweis: Weil die Determinante alternierend in den Spalten und $\det(A^t) = \det(A)$ ist, folgt:

$$\begin{aligned}
(f_{\sigma(1)} \wedge f_{\sigma(2)} \wedge f_{\sigma(3)})(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) &= \det\big(f_{\sigma(i)}(\mathbf{v}_j) \,:\, i, j = 1, \ldots, 3\big) \\
&= \det\big(f_{\sigma(j)}(\mathbf{v}_i) \,:\, i, j = 1, \ldots, 3\big) \\
&= \operatorname{sign}(\sigma) \cdot \det\big(f_j(\mathbf{v}_i) \,:\, i, j = 1, \ldots, 3\big) \\
&= \operatorname{sign}(\sigma) \cdot \det\big(f_i(\mathbf{v}_j) \,:\, i, j = 1, \ldots, 3\big) \\
&= \operatorname{sign}(\sigma) \cdot (f_1 \wedge f_2 \wedge f_3)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3).
\end{aligned}$$

■

Unter Verwendung von $\det(A^t) = \det(A)$ folgt auch:

$$\lambda_{\mathbf{a}_1} \wedge \lambda_{\mathbf{a}_2} \wedge \lambda_{\mathbf{a}_3}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \det(\vec{a}_1, \vec{a}_2, \vec{a}_3)\,,$$

also

$$\lambda_{\mathbf{a}_1} \wedge \lambda_{\mathbf{a}_2} \wedge \lambda_{\mathbf{a}_3} = \det(\vec{a}_1, \vec{a}_2, \vec{a}_3)\, dx_1 \wedge dx_2 \wedge dx_3.$$

Generell ist jede alternierende 3-Form auf dem $\mathbb{R}^3$ ein Vielfaches der Determinantenform, hat also die Gestalt $\omega = c\, dx_1 \wedge dx_2 \wedge dx_3$. Wir sprechen dann auch von einer *Differentialform der Dimension* 3.

Zum Schluß wollen wir noch das äußere Produkt einer 1-Form mit einer 2-Form einführen.

Existenz des äußeren Produktes einer 1-Form und einer 2-Form

Einer 1*-Form* f *und einer (alternierenden)* 2*-Form* φ *auf dem* $\mathbb{R}^3$ *kann man äußere Produkte* $f \wedge \varphi$ *und* $\varphi \wedge f$ *zuordnen, so daß gilt:*

1. $f \wedge \varphi$ *und* $\varphi \wedge f$ *sind (alternierende)* 3*-Formen, und es ist stets*

$$f \wedge \varphi = \varphi \wedge f.$$

2. *Es gelten die Formeln*

$$\begin{aligned} (f_1 + f_2) \wedge \varphi &= f_1 \wedge \varphi + f_2 \wedge \varphi, \\ f \wedge (\varphi_1 + \varphi_2) &= f \wedge \varphi_1 + f \wedge \varphi_2 \\ \textit{und} \quad (rf) \wedge \varphi &= f \wedge (r\varphi) \;=\; r(f \wedge \varphi). \end{aligned}$$

3. *Sind* f_1, f_2, f_3 *Linearformen, so ist*

$$f_1 \wedge (f_2 \wedge f_3) = (f_1 \wedge f_2) \wedge f_3 = f_1 \wedge f_2 \wedge f_3.$$

Auf den Beweis müssen wir hier verzichten. Es sei nur erwähnt, daß man z.B. definiert:

$$(f \wedge \varphi)(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) := \frac{1}{2} \sum_{\sigma \in S_3} \operatorname{sign}(\sigma) f(\mathbf{a}_{\sigma(1)}) \cdot \varphi(\mathbf{a}_{\sigma(2)}, \mathbf{a}_{\sigma(3)}).$$

Der Faktor 1/2 findet sich übrigens nicht bei allen Autoren.

Berechnung des Produktes einer 1-Form und einer 2-Form

Sind $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$*, so ist*

$$\lambda_{\mathbf{a}} \wedge \Lambda_{\mathbf{b}} = (\mathbf{a} \bullet \mathbf{b})\, dx_1 \wedge dx_2 \wedge dx_3\,.$$

Beweis: Es ist $\lambda_{\mathbf{a}} = a_1\, dx_1 + a_2\, dx_2 + a_3\, dx_3$ und

$$\Lambda_{\mathbf{b}} = b_1\, dx_2 \wedge dx_3 + b_2\, dx_3 \wedge dx_1 + a_3\, dx_1 \wedge dx_2,$$

also

$$\begin{aligned}\lambda_{\mathbf{a}} \wedge \Lambda_{\mathbf{b}} &= a_1 b_1\, dx_1 \wedge (dx_2 \wedge dx_3) + a_2 b_2\, dx_2 \wedge (dx_3 \wedge dx_1) \\ &\quad + a_3 b_3\, dx_3 \wedge (dx_1 \wedge dx_2) \\ &= (\mathbf{a} \bullet \mathbf{b})\, dx_1 \wedge dx_2 \wedge dx_3.\end{aligned}$$

Alle anderen Terme fallen weg. ∎

§ 2 Determinanten

Inhalt:

Rechenregeln für Determinanten, Dreiecksmatrizen, Determinanten-Produktsatz, Kästchensatz, Adjunkte und Streichungsmatrix, Laplacescher Entwicklungssatz, Cramersche Regel, Formel für die inverse Matrix, Rangbestimmung durch Unterdeterminanten.

Volumen eines Parallelotops, Orientierung von Basen und Vektorräumen.

Isometrien, orthogonale Matrizen, Drehungen und Spiegelungen, orthogonale und spezielle orthogonale Gruppe, unitäre Matrizen, unitäre Gruppen.

Wir wiederholen noch einmal die wichtigsten Eigenschaften der Determinante:

Ist $A = \left(a_{ij} \;\middle|\; \begin{matrix} i = 1, \ldots, n \\ j = 1, \ldots, n \end{matrix} \right)$, so ist $\boxed{\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} \,.}$

Bemerkung. Man schreibt manchmal auch

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \quad \text{statt} \quad \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Die Originalformel für die Determinante ist im Falle $n \geq 4$ zur Berechnung denkbar ungeeignet, denn die Anzahl der Permutationen beträgt ja $n!$. Also suchen wir nach besseren Methoden. Der Weg dorthin führt über einige allgemeine Sätze.

Die folgenden Regeln haben wir schon hergeleitet:

Rechenregeln für Determinanten

Sei $A \in M_{n,n}(K)$.

1. *Multipliziert man in A eine Zeile oder eine Spalte mit $\lambda \in K$, so muß man auch $\det(A)$ mit λ multiplizieren.*

   Insbesondere ist $\det(\lambda \cdot A) = \lambda^n \cdot \det(A)$.

Rechenregeln für Determinanten (Fortsetzung)

2. *Addiert man das Vielfache einer Zeile (bzw. Spalte) zu einer anderen Zeile (bzw. Spalte), so ändert sich der Wert der Determinante nicht.*
3. *Vertauscht man zwei Zeilen oder Spalten, so wechselt das Vorzeichen der Determinante.*
4. *Ist* $\mathrm{rg}(A) < n$*, so ist* $\det(A) = 0$.
5. *Es ist* $\det(A^t) = \det(A)$.
6. *Ist* $f : M_{n,n}(K) \to K$ *eine Funktion, die multilinear und alternierend in den Zeilen von* A *ist, mit* $f(E_n) = 1$*, so ist* $f = \Delta_n$ *die Determinantenform und* $f(A) = \det(A)$.

Durch elementare Umformungen kann man jede Matrix $A \in M_{n,n}(K)$ auf Dreiecksgestalt bringen. Da man dabei die Änderung der Determinante gut kontrollieren kann, ist der folgende Satz sehr nützlich:

Determinante einer Dreiecksmatrix

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{vmatrix} = a_{11} \cdot \ldots \cdot a_{nn}.$$

BEWEIS: Die Matrix sei mit A bezeichnet. Wir unterscheiden zwei Fälle:

1. Fall: Ist $a_{ii} = 0$ für ein i, so verschwinden beide Seiten der Gleichung: Bei der rechten Seite ist es klar, und die Determinante auf der linken Seite verschwindet, weil $\mathrm{rg}(A) < n$ ist.

3. Fall: $a_{11}, \ldots, a_{nn} \neq 0$.

Durch elementare Zeilenumformungen, die nicht die Determinante verändern, kann man alle Elemente oberhalb der Diagonalen zum Verschwinden bringen. Wenn aber $a_{ij} = 0$ für $i \neq j$ ist, dann ergibt sich die Behauptung sofort aus der Determinantenformel. ∎

Der nächste Satz hat sowohl praktische als auch theoretische Bedeutung:

Determinanten–Produktsatz

Es seien $A, B \in M_{n,n}(K)$. Dann gilt:

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

Beweis:
1) Zunächst sei $\text{rg}(B) < n$. Dann ist $\det(A) \cdot \det(B) = 0$.
Da aber $\text{rg}(A \cdot B) \leq \text{rg}(B)$ ist, ist auch $\det(A \cdot B) = 0$.

2) Ist $\text{rg}(B) = n$, so kann man B durch elementare Zeilenumformungen auf obere Dreiecksgestalt bringen, so daß in der Diagonale nur Elemente $\neq 0$ stehen. Die Determinante ändert sich dabei nur um einen Faktor $\neq 0$. Also ist $\det(B) \neq 0$.

Nun sei $\delta : M_{n,n}(K) \to K$ definiert durch

$$\delta(A) := \frac{\det(A \cdot B)}{\det(B)}.$$

Man rechnet leicht nach, daß δ multilinear in den Zeilen von A ist. Und wenn A zwei gleiche Zeilen enthält, dann trifft das auch auf $A \cdot B$ zu, so daß $\delta(A) = 0$ ist. Schließlich ist noch $\delta(E_n) = 1$. Aber dann muß $\delta(A) = \det(A)$ sein, und die Produktformel folgt. ∎

Wir haben implizit mitbewiesen:

Determinante und Regularität

$A \in M_{n,n}(K)$ ist genau dann regulär, wenn $\det(A) \neq 0$ *ist.*
In diesem Falle ist $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

Kästchensatz

Sind $A \in M_{r,r}(K)$, $B \in M_{r,n-r}(K)$ und $C \in M_{n-r,n-r}(K)$, so ist

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det(A) \cdot \det(C).$$

Beweis: 1) Mit Hilfe von Zeilenoperationen im Bereich der ersten r Zeilen und Spaltenvertauschungen im Bereich der ersten r Spalten kann man A in eine obere

Dreiecksmatrix Δ_1 umformen. Es gibt ein $a \neq 0$ und eine Matrix $B^* \in M_{r,n-r}(K)$, so daß gilt:

$$\det(A) = a \cdot \det(\Delta_1) \quad \text{und} \quad \det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = a \cdot \det\begin{pmatrix} \Delta_1 & B^* \\ 0 & C \end{pmatrix}.$$

2) Mit Hilfe von Zeilenoperationen im Bereich der letzten $n-r$ Zeilen und Spaltenvertauschungen im Bereich der letzten $n-r$ Spalten kann man C in eine obere Dreiecksmatrix Δ_2 umformen. Es gibt dann ein $c \neq 0$ und eine Matrix $B^{**} \in M_{r,n-r}(K)$, so daß gilt:

$$\det(C) = c \cdot \det(\Delta_2) \quad \text{und} \quad \det\begin{pmatrix} \Delta_1 & B^* \\ 0 & C \end{pmatrix} = c \cdot \det\begin{pmatrix} \Delta_1 & B^{**} \\ 0 & \Delta_2 \end{pmatrix}.$$

3) Da die Determinante einer oberen Dreiecksmatrix gleich dem Produkt der Diagonalelemente ist, folgt:

$$\begin{aligned} \det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} &= a \cdot c \cdot \det\begin{pmatrix} \Delta_1 & B^{**} \\ 0 & \Delta_2 \end{pmatrix} \\ &= a \cdot c \cdot \det(\Delta_1) \cdot \det(\Delta_2) \\ &= (a \cdot \det(\Delta_1)) \cdot (c \cdot \det(\Delta_2)) \\ &= \det(A) \cdot \det(B). \end{aligned}$$

■

Als nächstes wollen wir den allgemeinen Laplace'schen Entwicklungssatz beweisen, der es erlaubt, die Berechnung einer n-reihigen Determinante auf die von $(n-1)$–reihigen zurückzuführen.

Definition:

Sei $A = (\vec{a}_1, \ldots, \vec{a}_n) \in M_{n,n}(K)$. Dann nennt man

$$A_{ij} := \det(\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{e}_i, \vec{a}_{j+1}, \ldots, \vec{a}_n)$$

den *Cofaktor* (oder das *algebraische Komplement* oder die *Adjunkte*) von A zur i–ten Zeile und j–ten Spalte.

Mit $S_{ij}(A)$ wird diejenige Matrix bezeichnet, die man aus A durch Streichen der i–ten Zeile und der j–ten Spalte gewinnt. Man nennt sie auch *Streichungsmatrix*.

Adjunkte und Streichungsmatrix

Für alle i, j gilt: $\quad A_{ij} = (-1)^{i+j} \det S_{ij}(A)$.

Beweis: Es ist

$$
\begin{aligned}
A_{ij} &= \left|\begin{array}{ccccccc} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ & & & 1 & & & \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{array}\right| \leftarrow i \\
&= (-1)^{(i-1)+(j-1)} \cdot \left|\begin{array}{ccccccc} 1 & a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{in} \\ 0 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & & & & \vdots \\ 0 & a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{array}\right| \\
&= (-1)^{i+j} \cdot \det \left(\begin{array}{c|c} 1 & * \\ \hline 0 & \\ \vdots & S_{ij}(A) \\ 0 & \end{array}\right) \\
&= (-1)^{i+j} \cdot \det S_{ij}(A).
\end{aligned}
$$

Die letzte Gleichung ergibt sich aus dem Kästchensatz! ■

Laplace'scher Entwicklungssatz

Sei $A = (a_{ij}) \in M_{n,n}(K)$. Dann gilt für festes j (bzw. i) :

$$
\begin{aligned}
\det(A) &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det S_{ij}(A) \\
(\textit{bzw.} &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det S_{ij}(A) \quad).
\end{aligned}
$$

Man spricht von der Entwicklung nach der j-ten Spalte (bzw. nach der i-ten Zeile).

Beweis: Der zweite Fall folgt aus dem ersten durch Übergang zur transponierten Matrix.

Sind $\vec{a}_1, \ldots, \vec{a}_n$ die Spalten von A, so gilt:

$$\begin{aligned}
\det(\vec{a}_1,\ldots,\vec{a}_n) &= \det(\vec{a}_1,\ldots,\vec{a}_{j-1},\sum_{i=1}^n a_{ij}\vec{e}_i,\vec{a}_{j+1},\ldots,\vec{a}_n) \\
&= \sum_{i=1}^n a_{ij}\det(\vec{a}_1,\ldots,\vec{a}_{j-1},\vec{e}_i,\vec{a}_{j+1},\ldots,\vec{a}_n) \\
&= \sum_{i=1}^n a_{ij}A_{ij} \;=\; \sum_{i=1}^n a_{ij}(-1)^{i+j}\det S_{ij}(A).
\end{aligned}$$

∎

Die Vorzeichen sind wie bei einem Schachbrett verteilt:

$$\begin{array}{ccccc}
+ & - & + & - & \cdots \\
- & + & - & + & \\
+ & - & + & - & \\
\vdots & & & &
\end{array}$$

Beispiel.

Die Entwicklung nach der ersten Zeile ergibt:

$$\begin{aligned}
\det\begin{pmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} &= 0 - 1\cdot\det\begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} + 2\cdot\det\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \\
&= 0\cdot(-1) - 1\cdot(-1) + 2\cdot(3-2) \;=\; 3.
\end{aligned}$$

Im allgemeinen wird man eine Zeile oder Spalte suchen, in der möglichst viele Nullen zu finden sind.

Eine Anwendung der Determinantentheorie ergibt sich für Lineare Gleichungen:

Cramersche Regel

Sei $A = (\vec{a}_1,\ldots,\vec{a}_n) \in M_{n,n}(K)$. Dann gilt:

1. *$A\cdot\vec{x} = \vec{b}$ ist genau dann für jedes $\vec{b} \in K^n$* ***eindeutig*** *lösbar, wenn $\det(A) \neq 0$ ist.*
2. *Ist $\det(A) \neq 0$ und $\vec{x} = (x_1,\ldots,x_n)^t$ der eindeutig bestimmte Lösungsvektor des LGS $A\cdot\vec{x} = \vec{b}$, so ist*

$$x_i = \frac{1}{\det(A)}\cdot\det(\vec{a}_1,\ldots,\vec{a}_{i-1},\vec{b},\vec{a}_{i+1},\ldots,\vec{a}_n), \text{ für } i = 1,\ldots,n.$$

Beweis:

1) Schon bekannt!

2) Ist $\vec{x}$ der Lösungsvektor des LGS $A \cdot \vec{x} = \vec{b}$, so ist

$$\vec{b} = \sum_{k=1}^{n} x_k \vec{a}_k.$$

Daraus folgt:

$$\begin{aligned}
&\det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, \dots, \vec{a}_n) = \\
&= \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \sum_{k=1}^{n} x_k \vec{a}_k, \vec{a}_{i+1}, \dots, \vec{a}_n) \\
&= \sum_{k=1}^{n} x_k \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_k, \vec{a}_{i+1}, \dots, \vec{a}_n) \\
&= x_i \cdot \det(\vec{a}_1, \dots, \vec{a}_n).
\end{aligned}$$

∎

Beispiel.

Sei $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} -13 \\ -7 \end{pmatrix}$. Dann gilt: $\det(A) = 2 + 15 = 17$, d.h., das LGS ist eindeutig lösbar,

$$\det(\vec{b}, \mathbf{a}_2) = \det \begin{pmatrix} -13 & -3 \\ -7 & 1 \end{pmatrix} = -13 - 21 = -34$$

$$\text{und} \qquad \det(\mathbf{a}_1, \vec{b}) = \det \begin{pmatrix} 2 & -13 \\ 5 & -7 \end{pmatrix} = -14 + 65 = 51.$$

Für den Lösungsvektor $\vec{x} = (x_1, x_2)^t$ gilt dann:

$$x_1 = \frac{-34}{17} = -2 \text{ und } x_2 = \frac{51}{17} = 3.$$

Als weitere Anwendung ergibt sich eine Berechnungsmöglichkeit für die inverse Matrix:

Formel für die inverse Matrix

Sei $A \in \mathrm{GL}(n, K)$. Dann ist $A^{-1} = (y_{ij})$ gegeben durch

$$y_{ij} = \frac{1}{\det(A)} \cdot A_{ji} = \frac{1}{\det(A)} \cdot (-1)^{i+j} \det S_{ji}(A).$$

Man beachte die Reihenfolge der Indizes!

BEWEIS: Sei $\vec{y}_j := (y_{1j}, \dots, y_{nj})^t$ die j-te Spalte von A^{-1}. Da $A \cdot A^{-1} = E_n$ ist, gilt:

$$A \cdot \vec{y}_j = \vec{e}_j \text{ für } j = 1, \dots, n.$$

Aus der Cramerschen Regel folgt dann:

$$\begin{aligned} y_{ij} &= \frac{1}{\det(A)} \cdot \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{e}_j, \vec{a}_{i+1}, \dots, \vec{a}_n) \\ &= \frac{1}{\det(A)} \cdot (-1)^{i+j} \det S_{ji}(A). \end{aligned}$$

∎

Beispiel.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Schließlich kann man auch den Rang einer Matrix mit Hilfe von Determinanten bestimmen:

Rangbestimmung durch Unterdeterminanten

Sei $A \in M_{n,n}(K)$ nicht die Null–Matrix. Dann ist $\mathrm{rg}(A)$ die größte natürliche Zahl r, zu der es eine r-reihige Unterdeterminante $\neq 0$ von A gibt.

BEWEIS:

Sei r die Anzahl der Spalten der größten Unterdeterminante $\neq 0$ von A.

1) Sei A' eine r-reihige Untermatrix von A mit $\det(A') \neq 0$. Dann sind die Spalten von A' und damit auch r Spalten von A linear unabhängig. Also ist $\mathrm{rg}(A) \geq r$.

2) Sei $k = \mathrm{rg}(A)$. Dann gibt es k linear unabhängige Spalten in A. Sie bilden eine Matrix $A' \in M_{n,k}(K)$, die ebenfalls den Rang k hat. Aber dann gibt es in A' k linear unabhängige Zeilen. Die bilden eine k-reihige quadratische Untermatrix A'' mit $\det(A'') \neq 0$. Also ist $r \geq \mathrm{rg}(A)$. ∎

Beispiele.

1. Wir betrachten
$$A := \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix}.$$
Es ist $\det(A) = 2 + 0 + 8 - 6 - 0 - 4 = 0$, also $\mathrm{rg}(A) < 3$. Links oben findet sich die Unterdeterminante $1 \cdot 1 - 2 \cdot 1 = -1 \neq 0$. Also ist $\mathrm{rg}(A) = 2$.

2. Sei
$$B := \begin{pmatrix} \mathrm{j} & 0 & 2\mathrm{j} \\ 1 & 0 & 2 \\ 0 & \mathrm{j} & \mathrm{j} \\ 1 & 3 & 5 \end{pmatrix} \in M_{4,3}(\mathbb{C}).$$
Man kann nachrechnen, daß alle 3-reihigen Unterdeterminanten Null sind. Man kann aber auch leicht sehen, daß $B = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$ mit $2 \cdot \vec{b}_1 + \vec{b}_2 = \vec{b}_3$ ist. Also ist $\mathrm{rg}(B) < 3$. Rechts unten findet sich die Unterdeterminante $\mathrm{j} \cdot 5 - 3 \cdot \mathrm{j} = 2\mathrm{j} \neq 0$. Also ist $\mathrm{rg}(B) = 2$.

Sei nun V ein beliebiger n-dimensionaler K-Vektorraum und $f : V \to V$ linear, also ein Endomorphismus. Wir können eine Basis $A = \{a_1, \ldots, a_n\}$ von V wählen. Dann wird f durch die Matrix $M = M_A(f) \in M_{n,n}(K)$ beschrieben, mit

$$M_A(f) := M(\Phi_A \circ f \circ \Phi_A^{-1}) = ([f(a_1)]_A, \ldots, [f(a_n)]_A).$$

Ist nun A' eine andere Basis von V, $M' = M_{A'}(f)$ und P die zugehörige Basiswechselmatrix, so ist $M' = P^{-1} \cdot M \cdot P$, also

$$\det(M') = \det(P)^{-1} \cdot \det(M) \cdot \det(P) = \det(M).$$

Daher ist die folgende Definition sinnvoll:

Definition:

Sei V ein beliebiger endlich-dimensionaler K-Vektorraum, A eine Basis für V und $f : V \to V$ ein Endomorphismus. Dann setzt man

$$\det(f) := \det(M_A(f)).$$

Wir wollen uns noch ein wenig mit der geometrischen Deutung der Determinante beschäftigen.

Am einfachsten ist es im Falle $n = 2$. Es seien zwei Spaltenvektoren $\vec{a} = \binom{a_1}{a_2}$ und $\vec{b} = \binom{b_1}{b_2}$ gegeben. Dann ist

$$(a_1, a_2, 0) \times (b_1, b_2, 0) = (0, 0, a_1 b_2 - a_2 b_1) = (0, 0, \det(\vec{a}, \vec{b})),$$

also

$$\|(a_1, a_2, 0) \times (b_1, b_2, 0)\| = |\det(\vec{a}, \vec{b})|.$$

Andererseits wissen wir aus dem 1. Semester, daß

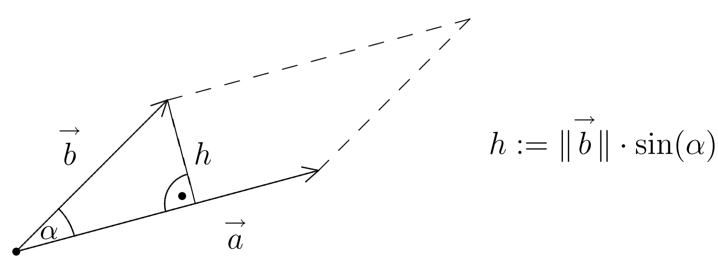
$$\|(a_1, a_2, 0) \times (b_1, b_2, 0)\| = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \sin(\alpha)$$

ist, wobei α den Winkel zwischen $\vec{a}$ und $\vec{b}$ bezeichnet. Er muß so gewählt werden, daß er zwischen 0 und π liegt. Aber dann ist

$$|\det(\vec{a}, \vec{b})| = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \sin(\alpha)$$

der Flächeninhalt des von $\vec{a}$ und $\vec{b}$ aufgespannten Parallelogramms.

Im Bild sieht das folgendermaßen aus:



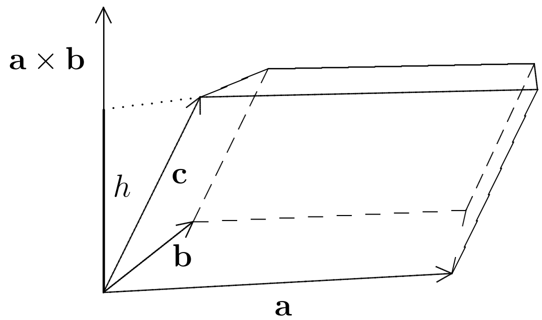
Wir kommen jetzt zum Fall $n = 3$. Gegeben seien drei Vektoren $\vec{a}$, $\vec{b}$ und $\vec{c}$. Wir nehmen an, daß $\vec{a}$ und $\vec{b}$ in der x-y-Ebene liegen. Die Höhe des von $\vec{a}$, $\vec{b}$ und $\vec{c}$ aufgespannten *Parallelotops* (oder *Spats*) ist die Größe

$$h := |\mathrm{pr}_{\mathbf{a}\times\mathbf{b}}(\mathbf{c})|,$$

wobei $\mathrm{pr}_{\mathbf{a}\times\mathbf{b}}(\mathbf{c})$ die orthogonale Projektion von $\mathbf{c}$ auf die von $\mathbf{a} \times \mathbf{b}$ aufgespannte Gerade bedeutet. Da $\|\mathbf{a} \times \mathbf{b}\|$ die Grundfläche des Parallelotops wiedergibt, ist

$$V(\mathbf{a}, \mathbf{b}, \mathbf{c}) := h \cdot \|\mathbf{a} \times \mathbf{b}\|$$

das Volumen des Parallelotops.



Für die orthogonale Projektion gilt:

$$\mathrm{pr}_{\mathbf{a}\times\mathbf{b}}(\mathbf{c}) = \mathbf{c} \bullet \frac{\mathbf{a}\times\mathbf{b}}{\|\mathbf{a}\times\mathbf{b}\|}.$$

Also ist

$$V = |\mathrm{pr}_{\mathbf{a}\times\mathbf{b}}(\mathbf{c})| \cdot \|\mathbf{a}\times\mathbf{b}\| = |\mathbf{c} \bullet (\mathbf{a}\times\mathbf{b})| = |\Lambda_{\mathbf{c}}(\mathbf{a},\mathbf{b})|.$$

Man nennt $[\mathbf{a},\mathbf{b},\mathbf{c}] := \mathbf{c} \bullet (\mathbf{a}\times\mathbf{b})$ auch das *Spatprodukt* der Vektoren $\mathbf{a}$, $\mathbf{b}$ und $\mathbf{c}$. Nach dem Laplaceschen Entwicklungssatz ist aber

$$\begin{aligned}
\det(\vec{a},\vec{b},\vec{c}) &= \det(\vec{c},\vec{a},\vec{b}) \\
&= c_1(a_2b_3 - a_3b_2) - c_2(a_1b_3 - a_3b_1) + c_3(a_1b_2 - a_2b_1) \\
&= \mathbf{c} \bullet (\mathbf{a}\times\mathbf{b}).
\end{aligned}$$

Also ist $V(\mathbf{a},\mathbf{b},\mathbf{c}) = |\det(\vec{a},\vec{b},\vec{c})|$.

In höheren Dimensionen können wir nicht mehr anschaulich argumentieren. Dort setzen wir das „Volumen“ per definitionem fest.

Definition:

Sind $\mathbf{a}_1,\dots\mathbf{a}_n$ linear unabhängig im $\mathbb{R}^n$, so nennt man die Zahl

$$V(\mathbf{a}_1,\dots,\mathbf{a}_n) := |\det(\vec{a}_1,\dots,\vec{a}_n)|$$

das *Volumen* des von $\mathbf{a}_1,\dots,\mathbf{a}_n$ aufgespannten Parallelotops.

Sind die Vektoren linear abhängig, so spannen sie ein niederdimensionales Gebilde auf, dessen Volumen im $\mathbb{R}^n$ natürlich $= 0$ sein sollte.

Als nächstes befassen wir uns mit dem Vorzeichen der Determinante.

Seien $\mathbf{a},\mathbf{b}$ zwei Vektoren im $\mathbb{R}^2$. Ist $\mathbf{a} = (a_1,a_2) \neq (0,0)$, so setzen wir $\mathbf{a}^* := (-a_2,a_1)$. Weil $\mathbf{a} \bullet \mathbf{a}^* = 0$ ist, bildet $\{\mathbf{a},\mathbf{a}^*\}$ eine Basis des $\mathbb{R}^2$. Dabei entsteht $\mathbf{a}^*$ aus $\mathbf{a}$ durch eine Drehung um $90°$ (gegen den Uhrzeigersinn). Es ist

$$\mathbf{b} = \lambda\mathbf{a} + \mu\mathbf{a}^*.$$

Genau dann ist $\mu > 0$, wenn $\mathbf{b}$ in der gleichen Halbebene wie $\mathbf{a}^*$ liegt. Andererseits ist

$$\det(\mathbf{a}, \mathbf{b}) = \det(\mathbf{a}, \lambda\mathbf{a} + \mu\mathbf{a}^*) = \mu \cdot \det(\mathbf{a}, \mathbf{a}^*) = \mu \cdot (a_1^2 + a_2^2).$$

Das Vorzeichen der Determinante kennzeichnet also die „Orientierung" der Vektoren.

Sind jetzt $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ und $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ zwei Basen des $\mathbb{R}^n$, so ist $B \cdot W_{B,A} = A$, also $\det(B) \cdot \det(W_{B,A}) = \det(A)$. Haben $\det(A)$ und $\det(B)$ das gleiche Vorzeichen, so muß $\det(W_{B,A}) > 0$ sein. Wir sagen in diesem Fall, daß A und B *gleich orientiert* sind, und andernfalls, daß sie *entgegengesetzt orientiert* sind.

Definition:

Eine Basis A des $\mathbb{R}^n$ heißt *positiv orientiert*, falls A und die Standardbasis $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ gleich orientiert sind, falls also $\det(A) = \det(W_{E,A}) > 0$ ist. Andernfalls nennt man A *negativ orientiert*.

Ein Isomorphismus $f : \mathbb{R}^n \to \mathbb{R}^n$ heißt *orientierungstreu*, falls $\det(f) > 0$ ist.

Wie weit lassen sich diese Dinge auf beliebige Vektorräume übertragen? Zunächst ist klar, daß wir nur reelle Vektorräume zu betrachten brauchen, denn wir müssen zwischen positiven und negativen Determinanten unterscheiden können. Außerdem brauchen wir die Beschreibung linearer Abbildungen durch Matrizen. Das geht nur in endlich-dimensionalen Räumen. Wir können uns also auf Unterräume $V \subset \mathbb{R}^n$ beschränken. Was ist nun die Schwierigkeit? Was ist anders als beim $\mathbb{R}^n$ selbst?

Nehmen wir als Beispiel den Raum $V = \{(-x, x)\ x \in \mathbb{R}\} \subset \mathbb{R}^2$. Es gibt genau zwei ON-Basen von V, nämlich $\{\frac{1}{\sqrt{2}}(-1, 1)\}$ und $\{\frac{1}{\sqrt{2}}(1, -1)\}$. Welche von beiden ist besser? Welche sollte man auszeichnen? Wir können keine auszuzeichnende Basis finden, und deshalb ist es auch nicht möglich zu sagen, wann eine Basis positiv orientiert ist. Was bleibt?

Zwei Basen A, B von $V \subset \mathbb{R}^n$ heißen *gleich orientiert*, falls $\det(W_{A,B}) > 0$ ist. Wir schreiben dafür kurz $A \sim B$. Dann gilt:

1. Für jede Basis A gilt: $A \sim A$ („Reflexivität").
2. Ist $A \sim B$, so ist auch $B \sim A$ („Symmetrie").
3. Ist $A \sim B$ und $B \sim C$, so ist auch $A \sim C$ („Transitivität").

BEWEIS: 1) $\det(W_{A,A}) = \det(E_n) = 1 > 0$.

2) Ist $\det(W_{B,A}) > 0$, so ist $\det(W_{A,B}) = \det(W_{B,A}^{-1}) = 1/\det(W_{B,A}) > 0$.

3) Ist $\det(W_{B,A}) > 0$ und $\det(W_{C,B}) > 0$, so ist $\det(W_{C,A}) = \det(W_{C,B} \cdot W_{B,A}) = \det(W_{C,B}) \cdot \det(W_{B,A}) > 0$. ■

Eine Relation, die reflexiv, symmetrisch und transitiv ist, nennt man eine *Äquivalenzrelation*. Äquivalenzrelationen sind ein Wundermittel in der Mathematik. Sie schaffen nämlich neue abstrakte Begriffe. Im Grunde kennt das aber jeder aus dem Alltag. Niemand kann Farben definieren (versuchen Sie mal, einem kleinen Kind die Farbe „Rot“ zu erklären), aber man begreift, welche Dinge die gleiche Farbe haben. Das Zusammenfassen solcher Dinge führt erst zum Farbbegriff. Der Mathematiker sagt dazu, daß *Äquivalenzklassen* gebildet werden. Ein Beispiel sind die Restklassen in $\mathbb{Z}$. Die zugehörige Äquivalenzrelation ist die Kongruenz.

Zurück zu unserem Thema. Die Relation „gleich orientiert“ liefert genau zwei Äquivalenzklassen in V. Hält man irgend eine Basis fest, so ist jede andere Basis zu ihr entweder gleich oder entgegengesetzt orientiert. Jede Äquivalenzklasse heißt eine *Orientierung* von V.

Fassen wir zusammen: Auf jedem endlich-dimensionalen Vektorraum gibt es zwei Orientierungen. Aber nur im Falle des $\mathbb{R}^n$ können wir die eine Orientierung positiv und die andere Orientierung negativ nennen.

Definition:

Eine lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ heißt eine *Isometrie*, falls gilt:

$$\|f(\mathbf{x})\| = \|\mathbf{x}\| \text{ für alle } \mathbf{x} \in \mathbb{R}^n.$$

Satz (Charakterisierung von Isometrien)

Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ linear und $A = M(f)$. Dann sind die folgenden Aussagen äquivalent:

1. *f ist eine Isometrie.*
2. *f ist ein Isomorphismus, und es ist*

$$f(\mathbf{x}) \bullet f(\mathbf{y}) = \mathbf{x} \bullet \mathbf{y} \text{ für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

3. *A ist orthogonal, d.h. es ist $A^t \cdot A = A \cdot A^t = E_n$.*
4. *Die Zeilen von A bilden eine ON-Basis.*
5. *Die Spalten von A bilden eine ON-Basis.*

BEWEIS: Vorbemerkung: Ist $A^t \cdot A = E_n$, so ist $\det(A)^2 = \det(A \cdot A^t) = \det(E_n) = 1$, also insbesondere $\det(A) \neq 0$. Daraus folgt, daß A invertierbar ist, und dann ist $A^{-1} = A^t$ und auch $A \cdot A^t = E_n$.

(1) $\implies$ (2): Ist $f(\mathbf{x}) = \mathbf{0}$, so ist $\|\mathbf{x}\| = \|f(\mathbf{x})\| = 0$, also auch $\mathbf{x} = \mathbf{0}$. Damit ist f injektiv, und eine injektive lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ ist automatisch ein Isomorphismus.

Aus der Beziehung

$$\begin{aligned} \|\mathbf{x}-\mathbf{y}\|^2 &= (\mathbf{x}-\mathbf{y}) \bullet (\mathbf{x}-\mathbf{y}) \\ &= \mathbf{x} \bullet \mathbf{x} + \mathbf{y} \bullet \mathbf{y} - 2\mathbf{x} \bullet \mathbf{y} \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x} \bullet \mathbf{y}, \\ \text{folgt:} \qquad \mathbf{x} \bullet \mathbf{y} &= \frac{1}{2} \cdot (\|\mathbf{x}-\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2). \end{aligned}$$

Es ist $\|f(\mathbf{x})\|^2 = \|\mathbf{x}\|^2$, $\|f(\mathbf{y})\| = \|\mathbf{y}\|^2$ und

$$\|f(\mathbf{x}) - f(\mathbf{y})\|^2 = \|f(\mathbf{x}-\mathbf{y})\|^2 = \|\mathbf{x}-\mathbf{y}\|^2.$$

Daraus folgt: $f(\mathbf{x}) \bullet f(\mathbf{y}) = \mathbf{x} \bullet \mathbf{y}$.

(2) $\implies$ (3): Weil $\mathbf{x} \cdot \mathbf{y}^t = \mathbf{x} \bullet \mathbf{y} = f(\mathbf{x}) \bullet f(\mathbf{y}) = \mathbf{x} \cdot A^t \cdot A \cdot \mathbf{y}^t$ ist, folgt: $A^t \cdot A = E_n$

Nach der Vorbemerkung ist dann aber auch $A \cdot A^t = E_n$.

(3) $\implies$ (4): Sind $\mathbf{a}_1, \ldots, \mathbf{a}_n$ die Zeilen von A, so ist

$$E_n = A \cdot A^t = \Big(\mathbf{a}_i \bullet \mathbf{a}_j \;:\; i,j = 1, \ldots, n\Big),$$

also $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ eine ON-Basis.

(4) $\implies$ (5): Bilden die Zeilen von A eine ON-Basis, so ist $A \cdot A^t = E_n$, und nach der Vorbemerkung ist dann auch $A^t \cdot A = E_n$. Daraus folgt wie beim vorigen Schritt, daß die Spalten von A eine ON-Basis bilden.

(5) $\implies$ (1): Nach Voraussetzung ist $\{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)\}$ eine ON-Basis. Ist $\mathbf{x} = (x_1, \ldots, x_n)$, so ist

$$\begin{aligned} \|f(\mathbf{x})\|^2 &= f\Big(\sum_{i=1}^n x_i \mathbf{e}_i\Big) \bullet f\Big(\sum_{j=1}^n x_j \mathbf{e}_j\Big) \\ &= \sum_{i,j} x_i x_j f(\mathbf{e}_i) \bullet f(\mathbf{e}_j) \\ &= \|\mathbf{x}\|^2. \end{aligned}$$

Damit ist f eine Isometrie. ∎

Definition:

$A \in M_{n,n}(\mathbb{R})$ heißt eine *orthogonale Matrix*, falls $A^t \cdot A = E_n$ ist.

Notwendiges Kriterium für Orthogonalität

Ist A orthogonal, so ist $|\det(A)| = 1$.

BEWEIS: Klar, denn wegen $A^t \cdot A = E_n$ ist $\det(A)^2 = 1$. ∎

Im Falle $n = 2$ wollen wir alle orthogonalen Matrizen bestimmen:

Sei also $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ orthogonal. Da dann die Spaltenvektoren ein ON-System bilden, gilt:

$$\|(a, c)\|^2 = 1 \quad \text{und} \quad (a, c) \bullet (b, d) = 0.$$

Also ist $a^2+c^2 = 1$, und weil $(a, c)\bullet(-c, a) = 0$, $(-c, a) \neq (0, 0)$ und das orthogonale Komplement zu $\mathbb{R}(a, c)$ im $\mathbb{R}^2$ 1-dimensional ist, gibt es ein $\lambda \in \mathbb{R}$ mit $(b, d) = \lambda \cdot (-c, a)$.

Aus der ersten Aussage folgt: Es gibt ein $\varphi \in [0, 2\pi)$, so daß $a = \cos(\varphi)$ und $c = \sin(\varphi)$ ist.

Wir müssen nun zwei Fälle unterscheiden:

1) Ist $\det(A) = 1$, so ist $1 = ad - bc = \lambda(a^2 + c^2) = \lambda$, also

$$A = R(\varphi) := \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Man nennt $R(\varphi)$ die *Drehmatrix zum Winkel* φ.

$R(\varphi)$ bildet z.B. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ auf $\begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ auf $\begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix}$ ab.

2) Sei $\det(A) = -1$.

Mit A und $S := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ist auch $A \cdot S$ orthogonal, aber die neue Matrix besitzt die Determinante $+1$. Also gibt es ein φ, so daß $A = R(\varphi) \cdot S^{-1} = R(\varphi) \cdot S$ ist. Dabei ist $f_S(x, y) = (x, -y)$ die Spiegelung an der x–Achse.

Definition:

Eine orthogonale Matrix $A \in M_{n,n}(\mathbb{R})$ heißt *Drehung*, falls $\det(A) = 1$ ist, und *Drehspiegelung*, falls $\det(A) = -1$ ist.

Die Menge aller orthogonalen Matrizen wird mit $O(n)$ bezeichnet, die Menge aller Drehungen mit $SO(n)$.

Satz

$O(n)$ ist eine Gruppe und $SO(n)$ eine Untergruppe.

BEWEIS: Offensichtlich liegt die Einheitsmatrix E_n in beiden Mengen. Sind A, B orthogonal, so ist

$$(A \cdot B) \cdot (A \cdot B)^t = A \cdot B \cdot B^t \cdot A^t = E_n,$$

also auch $A \cdot B$ orthogonal. Außerdem ist

$$A^{-1} \cdot (A^{-1})^t = A^t \cdot A^{tt} = A^t \cdot A = E_n,$$

also auch A^{-1} orthogonal.

Wegen $\det(A \cdot B) = \det(A) \cdot \det(B)$ und $\det(A^{-1}) = \det(A)^{-1}$ folgt sofort, daß $SO(n)$ eine Untergruppe von $O(n)$ ist. ∎

Ist $A \in O(n)$, so gibt es zwei Möglichkeiten: Entweder ist schon $A \in SO(n)$, oder es ist $A \cdot S \in SO(n)$, mit

$$S := \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & & & \\ 0 & 0 & 1 & & \\ \vdots & & & \ddots & 0 \\ 0 & & & 0 & 1 \end{pmatrix}.$$

Das heißt, jede orthogonale Transformation ist eine Drehung oder eine Drehspiegelung.

Wir wollen noch Drehungen im $\mathbb{R}^3$ betrachten, weil sie wichtig für die Computergraphik sind.

Sei also $A \in SO(3)$.

Behauptung: Es gibt ein $\vec{v} \neq \vec{0}$ mit $A \cdot \vec{v} = \vec{v}$.

BEWEIS: $A\cdot\vec{v} = \vec{v}$ ist gleichbedeutend mit $(A-E_3)\cdot\vec{v} = \vec{0}$. Da A eine Drehmatrix ist, ist

$$\begin{aligned}\det(A - E_3) &= \det(A\cdot(E_3 - A^t)) \\ &= (-1)^3\det(A)\cdot\det(A^t - E_3) \\ &= -\det(A)\cdot\det(A - E_3),\end{aligned}$$

also

$$2\cdot\det(A - E_3) = 0.$$

Das Gleichungssystem $(A - E_3)\cdot\vec{x} = \vec{0}$ muß demnach eine Lösung $\vec{v} \neq \vec{0}$ besitzen, und dann ist $A\cdot\vec{v} = \vec{v}$. ∎

Wir können annehmen, daß $\vec{v}$ normiert ist. Dann gibt es eine ON-Basis $V = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ des $\mathbb{R}^3$ mit $\vec{v}_1 = \vec{v}$. Der Vektor $\vec{v}_1$ wird sich als Drehachse erweisen.

Die Matrix $V = W_{E,V}$ ist orthogonal, und $f = f_A$ wird bezüglich V durch die Matrix

$$\widetilde{A} := M_V(f) = V^{-1}\cdot A\cdot V = V^t\cdot A\cdot V$$

beschrieben. Nun ist

$$\widetilde{A}\cdot\widetilde{A}^t = V^t\cdot A^t\cdot V\cdot V^t\cdot A\cdot V = E_3,$$

also $\widetilde{A}$ orthogonal. Außerdem ist $\det(\widetilde{A}) = \det(A) = 1$, also sogar $\widetilde{A} \in SO(3)$. Aber jetzt ist

$$\widetilde{A}\cdot\vec{e}_1 = M_V(f)\cdot[\mathbf{v}_1]_V = [f(\mathbf{v}_1)]_V = [\mathbf{v}_1]_V = \vec{e}_1.$$

Das bedeutet, daß $\widetilde{A}$ folgende Gestalt hat:

$$\widetilde{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$

Die Matrix $A_0 := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist orthogonal, mit $\det(A_0) = 1$. Also gibt es ein $\varphi \in [0, 2\pi)$ mit

$$\widetilde{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}.$$

Das zeigt, daß $\widetilde{A}$ eine Drehung um $\vec{e}_1$ mit Winkel φ beschreibt.

Wegen $V\cdot\widetilde{A} = A\cdot V$ ist

$$A\cdot\vec{v}_j = V\cdot\widetilde{A}\cdot\vec{e}_j = (\vec{v}_1, \vec{v}_2, \vec{v}_3)\cdot\vec{s}_j(\widetilde{A}),$$

also

$$\begin{aligned} f(\vec{v}_1) &= \vec{v}_2. \\ f(\vec{v}_2) &= \cos(\varphi)\vec{v}_2 + \sin(\varphi)\vec{v}_3, \\ f(\vec{v}_3) &= -\sin(\varphi)\vec{v}_2 + \cos(\varphi)\vec{v}_3. \end{aligned}$$

Das ist eine Drehung um $\mathbf{v}_1$ mit Winkel φ.

Bemerkung. Eine andere Beschreibung der Drehungen benutzt die sogenannten *Eulerschen Winkel*.

Im Komplexen sehen die Isometrien etwas anders aus.

Wir nennen $f : \mathbb{C}^n \to \mathbb{C}^n$ eine Isometrie, wenn $<f(\mathbf{z}), f(\mathbf{w})> = <\mathbf{z}, \mathbf{w}>$ für alle $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$ ist. Da $< \mathbf{z}, \mathbf{w} > = \mathbf{z} \cdot \overline{\mathbf{w}}^t$ ist, folgt:

Die Matrix zu einer Isometrie des $\mathbb{C}^n$

Sei $A \in M_{n,n}(\mathbb{C})$, $f = f_A : \mathbb{C}^n \to \mathbb{C}^n$. f ist genau dann eine Isometrie (bezüglich des hermiteschen Skalarproduktes), wenn $\overline{A}^t \cdot A = E_n$ ist.

Der BEWEIS funktioniert genauso wie beim euklidischen Skalarprodukt.

Definition:

$A \in M_{n,n}(\mathbb{C})$ heißt eine *unitäre Matrix*, falls $\overline{A}^t \cdot A = E_n$ ist.

Die Gruppe der unitären Matrizen wird mit $U(n)$ bezeichnet, die Untergruppe aller unitären Matrizen A mit $\det(A) = 1$ wird mit $SU(n)$ bezeichnet.

Es gilt:

$$A \text{ unitär } \implies |\det(A)|^2 = 1.$$

Aber VORSICHT! Eine komplexe Zahl vom Betrag 1 hat die Gestalt $e^{\mathrm{j}\,t}$. Daher kann man die Unterscheidung zwischen Drehungen und Drehspiegelungen **nicht** auf unitäre Matrizen übertragen!

Es ist $SU(1) = \{1\}$ und $U(1) = \{z \in \mathbb{C}^* : |z| = 1\} = \{\cos(\varphi) + \mathrm{j}\,\sin(\varphi)\}$, also $U(1) \cong SO(2)$. Ebenso besteht ein enger Zusammenhang (aber keine Isomorphie) zwischen $SO(3)$ und $SU(2)$, auf den wir hier aber nicht eingehen können.

§ 3 Eigenwerte und Eigenvektoren

Inhalt:

Eigenwerte, Eigenvektoren, Eigenraum, Diagonalisierbarkeit, charakteristisches Polynom, lineare Unabhängigkeit von Eigenvektoren, Invarianz des charakteristischen Polynoms unter Basiswechsel, algebraische und geometrische Vielfachheit, Charakterisierung der Diagonalisierbarkeit, der Satz von Cayley-Hamilton.

Eigenwerte von orthogonalen, unitären, symmetrischen und hermiteschen Matrizen, Hauptachsentransformation.

Definition:

Sei V ein K-Vektorraum, $f : V \to V$ linear. Ein Vektor $x \in V$ heißt *Eigenvektor* von f, falls gilt:

1. x ist nicht der Nullvektor.
2. Es gibt ein $\lambda \in K$, so daß $f(x) = \lambda x$ ist.

Der Skalar λ heißt *Eigenwert* von f zum Eigenvektor x.

Ist $A \in M_{n,n}(K)$ eine quadratische Matrix, so versteht man unter einem Eigenvektor (Eigenwert) von A einen Eigenvektor (Eigenwert) von $f_A : K^n \to K^n$.

Ist z.B. $A := \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ und $\vec{x}_0 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, so ist

$$A \cdot \vec{x}_0 = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 3 \\ 3 \cdot 2 + 2 \cdot 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = 4 \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 4\vec{x}_0,$$

also 4 ein Eigenwert von A zum Eigenvektor $\vec{x}_0$.

Der Raum $E(\lambda)$

Sei V ein K-Vektorraum, $f : V \to V$ linear und $\lambda \in K$. Dann ist

$$E(\lambda) := \{x \in V \mid f(x) = \lambda x\}$$

ein Unterverktorraum von V.

BEWEIS:

1) Offensichtlich liegt der Nullvektor in $E(\lambda)$. Man beachte allerdings, daß 0 kein Eigenvektor ist!

2) Ist $x \in E(\lambda)$, also $f(x) = \lambda x$, so ist auch

$$f(\alpha x) = \alpha \cdot f(x) = \alpha(\lambda x) = \lambda(\alpha x),$$

also $\alpha x \in E(\lambda)$, für $\alpha \in K$.

3) Sind $x, y \in E(\lambda)$, so ist $f(x) = \lambda x$ und $f(y) = \lambda y$. Also ist

$$f(x + y) = f(x) + f(y) = \lambda x + \lambda y = \lambda(x + y),$$

also $x + y \in E(\lambda)$. ∎

Definition:

Ist $f : V \to V$ linear, $\lambda \in K$, so heißt $E(\lambda)$ der *Eigenraum* von f zu λ.

Bemerkung. Die Menge $E(\lambda) \setminus \{0\}$ besteht aus Eigenvektoren zum Eigenwert λ. Im allgemeinen wird sie natürlich leer sein.

Wozu sind Eigenvektoren gut?

Definition:

Sei V ein endlich-dimensionaler K-Vektorraum. Eine lineare Abbildung $f : V \to V$ heißt *diagonalisierbar*, wenn es eine Basis $\{b_1, \ldots, b_n\}$ von V gibt, bzgl. der f durch eine Diagonalmatrix beschrieben wird.

Eine Matrix $A \in M_{n,n}(K)$ heißt diagonalisierbar, wenn $f_A : K^n \to K^n$ diagonalisierbar ist.

Die Matrix $M_B(f) = \big(d_{ij}\big)$, die f bzgl. der Basis $B = \{b_1, \ldots, b_n\}$ beschreibt, ist gegeben durch

$$f(b_j) = \sum_{i=1}^{n} d_{ij} b_i.$$

$M_B(f)$ ist genau dann eine Diagonalmatrix, wenn $d_{ij} = 0$ für $i \neq j$ ist, wenn also $f(b_j) = d_{jj} b_j$ für $j = 1, \ldots, n$ ist. Das ist genau dann der Fall, wenn B nur aus Eigenvektoren besteht.

Nun sei speziell $f = f_A : K^n \to K^n$, mit $A \in M_{n,n}(K)$.

Bezüglich einer beliebigen Basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ wird f durch die Matrix $M_B(f_A) = B^{-1} \cdot A \cdot B$ beschrieben, mit der regulären Matrix $B = (\vec{b}_1, \ldots, \vec{b}_n)$. Damit haben wir:

Diagonalisierbarkeit von Matrizen

Folgende Aussagen über eine Matrix $A \in M_{n,n}(K)$ sind äquivalent:

1. *A ist diagonalisierbar.*
2. *Im K^n gibt es eine Basis von Eigenvektoren von A.*
3. *Es gibt eine Diagonalmatrix D und eine reguläre Matrix B, so daß gilt:*

$$D = B^{-1} \cdot A \cdot B$$

Wie findet man Eigenvektoren zu einer Matrix A ? Es ist einfacher, mit den Eigenwerten zu beginnen:

$$\begin{aligned}
\lambda \text{ Eigenwert von } A \quad &\Longleftrightarrow \quad \exists\, \mathbf{v} \neq \mathbf{0} \text{ mit } f_A(\mathbf{v}) = \lambda \cdot \mathbf{v} \\
&\Longleftrightarrow \quad \exists\, \mathbf{v} \neq \mathbf{0} \text{ mit } f_A(\mathbf{v}) - \lambda \cdot \mathbf{v} = \mathbf{0} \\
&\Longleftrightarrow \quad \exists\, \mathbf{v} \neq \mathbf{0} \text{ mit } (f_A - \lambda \cdot \mathrm{id})(\mathbf{v}) = \mathbf{0} \\
&\Longleftrightarrow \quad \mathrm{Ker}(f_A - \lambda \cdot \mathrm{id}) \neq \{\mathbf{0}\} \\
&\Longleftrightarrow \quad \mathrm{rg}(A - \lambda \cdot E_n) < n \\
&\Longleftrightarrow \quad \det(A - \lambda \cdot E_n) = 0 \\
&\Longleftrightarrow \quad \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.
\end{aligned}$$

Wie sieht diese Determinante aus? Wir setzen $\widetilde{a}_{ij} := \begin{cases} a_{ij} & \text{für } i \neq j \\ a_{ii} - \lambda & \text{für } i = j \end{cases}$.

Außerdem beachten wir folgende Tatsache:

Ist $\sigma \in S_n$ eine Permutation $\neq \mathrm{id}$, so gibt es ein i mit $j := \sigma(i) \neq i$. Weil Permutationen bijektive Abbildungen sind, kann auch nicht $\sigma(j) = j$ gelten, denn dann hätten ja zwei Zahlen das gleiche Bild. Damit folgt:

$$\begin{aligned}
\det(A - \lambda \cdot E_n) &= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \widetilde{a}_{1,\sigma(1)} \cdot \ldots \cdot \widetilde{a}_{n,\sigma(n)} \\
&= (a_{11} - \lambda) \cdot \ldots \cdot (a_{nn} - \lambda) + \text{Terme, bei denen} \\
&\qquad \lambda \text{ in höchstens } n-2 \text{ Faktoren vorkommt} \\
&= (-1)^n \lambda^n + (-1)^{n-1} \Big(\sum_{i=1}^{n} a_{ii}\Big) \lambda^{n-1} + \text{Terme vom Grad} \leq n-2.
\end{aligned}$$

Definition:

$p_A(x) := \det(A - x \cdot E_n)$ heißt *charakteristisches Polynom* von A.

$p_A(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1} + c_n x^n$ ist ein Polynom vom Grad n über K, mit

$$\begin{aligned} c_n &= (-1)^n, \\ c_{n-1} &= (-1)^{n-1} \cdot \mathrm{Spur}(A), \\ &\vdots \\ c_0 &= p_A(0) \ = \ \det(A). \end{aligned}$$

Dabei ist $\mathrm{Spur}(A) = \sum_{i=1}^{n} a_{ii}$.

Das Entscheidende ist nun:

> *Die Eigenwerte von A sind die Nullstellen des charakteristischen Polynoms $p_A(x)$.*

Das liefert eine praktische Berechnungsmöglichkeit.

Beispiele.

1. Wir betrachten noch einmal die Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Es gilt:

$$\det \begin{pmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4.$$

Also ist

$$p_A(\lambda) = 0 \iff \lambda = \frac{3 \pm \sqrt{9+16}}{2} = \begin{cases} 4 \\ -1 \end{cases}$$

Den einen dieser beiden Eigenwerte hatten wir schon kennengelernt. Nun suchen wir die zugehörigen Eigenvektoren. Dazu muß man bei gegebenem λ die Gleichung $(A - \lambda \cdot E_2) \cdot \vec{x} = \vec{0}$ lösen:

$\lambda = -1$: $\begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ hat z.B. $\vec{b}_1 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ als Lösung.

$\lambda = 4$: $\begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ hat $\vec{b}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ als Lösung.

$\vec{b}_1$ und $\vec{b}_2$ sind Eigenvektoren zu den Eigenwerten -1 bzw. 4. Da $\det \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = 5 \neq 0$ ist, bilden sie auch eine Basis.

Damit ist bewiesen, daß A diagonalisierbar ist. Wir machen noch die Probe. Dazu setzen wir $B := (\vec{b}_1, \vec{b}_2) = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. Dann ist $B^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$ und

$$\begin{aligned} B^{-1} \cdot A \cdot B &= \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 8 \\ 1 & 12 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} -5 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Es kommt tatsächlich eine Diagonalmatrix heraus, und die Einträge darin sind gerade die Eigenwerte. Das muß natürlich so sein, weil $\{\vec{b}_1, \vec{b}_2\}$ eine Basis von Eigenvektoren ist.

2. $R(\varphi)$ sei die Drehmatrix $R(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$. Dann ist

$$\begin{aligned} \det(R(\varphi) - \lambda \cdot E_2) &= \det\begin{pmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{pmatrix} \\ &= (\cos\varphi - \lambda)^2 + \sin^2\varphi \\ &= \lambda^2 - 2\lambda\cos\varphi + 1. \end{aligned}$$

Dieser Ausdruck wird genau dann Null, wenn

$$\lambda = \frac{1}{2} \cdot (2\cos\varphi \pm \sqrt{4\cos^2\varphi - 4}) = \cos\varphi \pm \sqrt{-\sin^2\varphi} \quad \text{ist.}$$

Das ist nur möglich, wenn $\varphi = 0$ oder $\varphi = \pi$ ist, also $\sin\varphi = 0$. Das bedeutet, daß unter allen Drehmatrizen nur die speziellen Matrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 \quad \text{und} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E_2$$

Eigenwerte besitzen (und zwar $\lambda = 1$ bzw. $\lambda = -1$). Andere Drehungen besitzen keinen Eigenwert und daher auch keinen Eigenvektor.

Spaßeshalber kann man ja einmal $R(\varphi)$ als Endomorphismus des $\mathbb{C}^n$ auffassen und nach *komplexen* Eigenwerten suchen. Auf Grund der obigen Berechnung ist sofort klar, daß $\lambda_{1,2} = \cos\varphi \pm \mathsf{j}\,\sin\varphi = e^{\pm\mathsf{j}\,\varphi}$ komplexe Eigenwerte sind.

Die Gleichung für einen Eigenvektor $\begin{pmatrix} z \\ w \end{pmatrix} \in \mathbb{C}^2$ zum Eigenwert $\lambda_1 = e^{\mathsf{j}\,\varphi}$ lautet:

$$\begin{pmatrix} -\mathrm{j}\,\sin\varphi & -\sin\varphi \\ \sin\varphi & -\mathrm{j}\,\sin\varphi \end{pmatrix} \cdot \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Man sieht, daß $\begin{pmatrix} 1 \\ -\mathrm{j} \end{pmatrix}$ eine Lösung ist. Genauso ist $\begin{pmatrix} 1 \\ \mathrm{j} \end{pmatrix}$ ein Eigenvektor zum Eigenwert $\lambda_2 = e^{-\mathrm{j}\,\varphi}$. Offensichtlich sind $\begin{pmatrix} 1 \\ \mathrm{j} \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -\mathrm{j} \end{pmatrix}$ linear unabhängig (über $\mathbb{C}$), es gibt also eine Basis von Eigenvektoren. Man sieht, daß eine Matrix über $\mathbb{R}$ nicht diagonalisierbar, aber über $\mathbb{C}$ diagonalisierbar sein kann.

Lineare Unabhängigkeit von Eigenvektoren

Sind $\mathbf{v}_1, \ldots, \mathbf{v}_k \in K^n$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$ der Matrix $A \in M_{n,n}(K)$, so sind sie linear unabhängig.

BEWEIS: Dies ist Gelegenheit für einen Induktionsbeweis. Wir führen Induktion nach der Anzahl k.

$\underline{k=1}$: *Ein* einzelner Eigenvektor $\mathbf{v}_1$ ist immer linear unabhängig, weil er ja $\neq \mathbf{0}$ ist.

$\underline{k-1 \to k}$: Es seien Eigenvektoren $\mathbf{v}_1, \ldots, \mathbf{v}_k$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$ gegeben. Wir nehmen an, sie wären linear abhängig, und versuchen, daraus einen Widerspruch zu konstruieren:

Nach Induktionsvoraussetzung sind $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ linear unabhängig. Wenn sie durch Hinzunahme von $\mathbf{v}_k$ linear abhängig werden, muß es Koeffizienten $\alpha_j \in K$ geben, so daß gilt:

$$\mathbf{v}_k = \sum_{j=1}^{k-1} \alpha_j \mathbf{v}_j.$$

Daraus folgt:

$$\begin{aligned} \mathbf{0} = (f_A - \lambda_k \cdot \mathrm{id})(\mathbf{v}_k) &= (f_A - \lambda_k \cdot \mathrm{id})(\sum_{j=1}^{k-1} \alpha_j \mathbf{v}_j) \\ &= \sum_{j=1}^{k-1} \alpha_j \cdot (f_A - \lambda_k \cdot \mathrm{id})(\mathbf{v}_j) \\ &= \sum_{j=1}^{k-1} \alpha_j \cdot (\lambda_j - \lambda_k) \cdot \mathbf{v}_j. \end{aligned}$$

Da die $\lambda_j \neq \lambda_k$ sind, also $\lambda_j - \lambda_k \neq 0$ für $j = 1, \ldots, k-1$, und da $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ linear unabhängig sind, muß $\alpha_1 = \ldots = \alpha_{k-1} = 0$ sein. Aber dann ist $\mathbf{v}_k = \mathbf{0}$, und das kann bei einem Eigenvektor nicht sein. WS! ∎

Folgerung

Hat $A \in M_{n,n}(K)$ n verschiedene Eigenwerte, so ist A diagonalisierbar.

Eine besondere Rolle spielt der Eigenwert 0:

Der Eigenwert 0

$A \in M_{n,n}(K)$ ist genau dann regulär, wenn 0 ***kein*** *Eigenwert von A ist.*

Beweis: Es ist $p_A(0) = \det(A)$, also

$$\begin{aligned} A \quad \text{regulär} &\iff \det(A) \neq 0 \\ &\iff p_A(0) \neq 0 \\ &\iff 0 \text{ kein Eigenwert.} \end{aligned}$$

■

Das charakteristische Polynom p_A einer Matrix $A \in M_{n,n}(K)$ hängt in Wirklichkeit nur von der linearen Abbildung $f = f_A$ ab, ganz egal, bezüglich welcher Basis man f beschreibt:

Invarianz von $p_A(x)$ unter Basiswechsel

Ist B invertierbar, so ist

$$p_{B^{-1}AB}(x) = p_A(x) \text{ für alle } x \in \mathbb{R}.$$

Beweis: Es ist

$$\begin{aligned} p_{B^{-1}AB}(x) &= \det(B^{-1} \cdot A \cdot B - x \cdot E_n) \\ &= \det(B^{-1} \cdot A \cdot B - B^{-1} \cdot (x \cdot E_n) \cdot B) \\ &= \det(B^{-1} \cdot (A - x \cdot E_n) \cdot B) \\ &= (\det B)^{-1} \cdot \det(A - x \cdot E_n) \cdot \det B \\ &= \det(A - x \cdot E_n) \;=\; p_A(x). \end{aligned}$$

■

Leider kann man nicht immer erwarten, daß das charakteristische Polynom in lauter verschiedene Linearfaktoren zerfällt.

Beispiel.

Sei $A := \begin{pmatrix} 6 & 4 & 4 \\ -1 & 0 & -1 \\ -2 & 0 & 0 \end{pmatrix}$. Dann gilt:

$$\begin{aligned} \det(A - \lambda \cdot E_3) &= \det \begin{pmatrix} 6-\lambda & 4 & 4 \\ -1 & -\lambda & -1 \\ -2 & 0 & -\lambda \end{pmatrix} \\ &= (6-\lambda)\lambda^2 + 8 + 4 \cdot 0 - 8\lambda - 0 - 4\lambda \\ &= -\lambda^3 + 6\lambda^2 - 12\lambda + 8 \\ &= (2-\lambda)^3. \end{aligned}$$

Also ist $\lambda = 2$ die einzige Nullstelle von $p_A(x)$. Sie hat die Vielfachheit 3. Mehr Nullstellen kann es aus Gradgründen selbst im Komplexen nicht geben. Wie steht es nun mit den Eigenvektoren? Die zugehörige Gleichung hat die Form

$$\begin{pmatrix} 4 & 4 & 4 \\ -1 & -2 & -1 \\ -2 & 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Aus der Koeffizientenmatrix gewinnt man durch elementare Umformungen zunächst die Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & -2 & -1 \\ 1 & 0 & 1 \end{pmatrix}, \text{ dann } \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ und schließlich } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Das resultierende Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 &= 0, \\ x_2 &= 0 \end{aligned}$$

liefert als Lösungsraum den Eigenraum

$$E(2) = \{\alpha \cdot (-1, 0, 1) \mid \alpha \in \mathbb{R}\}.$$

Der ist nur 1-dimensional.

Definition:

Sei $\lambda \in K$ ein Eigenwert der Matrix $A \in M_{n,n}(K)$.

$a(\lambda)$:= Vielfachheit der Nullstelle λ des charakteristischen Polynoms $p_A(x)$ heißt *algebraische Vielfachheit von* λ.

$g(\lambda)$:= $\dim_K(E(\lambda))$ heißt *geometrische Vielfachheit von* λ.

Das Verhalten im vorigen Beispiel war kein Zufall:

Vergleich zwischen geometrischer und algebraischer Vielfachheit

Ist λ Eigenwert von $A \in M_{n,n}(K)$, so ist

$$1 \le g(\lambda) \le a(\lambda) \le n.$$

BEWEIS: Sei $k := g(\lambda)$.
Dann gibt es eine Basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ von $E(\lambda)$. Wir ergänzen sie zu einer Basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ des ganzen K^n. Dann ist $B := (\vec{v}_1, \ldots, \vec{v}_n)$ eine reguläre Matrix, und es gilt:

$$\begin{aligned}
B^{-1} \cdot A \cdot B &= B^{-1} \cdot (A \cdot (\vec{v}_1, \ldots, \vec{v}_k), A \cdot (\vec{v}_{k+1}, \ldots, \vec{v}_n)) \\
&= B^{-1} \cdot (\lambda \vec{v}_1, \ldots, \lambda \vec{v}_k, A \cdot (\vec{v}_{k+1}, \ldots, \vec{v}_n)) \\
&= (\lambda \cdot B^{-1} \cdot (\vec{v}_1, \ldots, \vec{v}_k), B^{-1} \cdot A \cdot (\vec{v}_{k+1}, \ldots, \vec{v}_n)) \\
&= \left(\begin{array}{ccc|c} \lambda & \cdots & 0 & \\ \vdots & \ddots & \vdots & X \\ 0 & \cdots & \lambda & \\ \hline & 0 & & Y \end{array}\right),
\end{aligned}$$

mit irgendwelchen Matrizen $X \in M_{k,n-k}(K)$ und $Y \in M_{n-k,n-k}(K)$. Daraus folgt:

$$\begin{aligned}
p_A(x) &= p_{B^{-1}AB}(x) \\
&= \det \left(\begin{array}{ccc|c} \lambda - x & \cdots & 0 & \\ \vdots & \ddots & \vdots & X \\ 0 & \cdots & \lambda - x & \\ \hline & 0 & & Y - x \cdot E_{n-k} \end{array}\right) \\
&= (\lambda - x)^k \cdot \det(Y - x \cdot E_{n-k}).
\end{aligned}$$

Also ist λ eine Nullstelle von mindestens k-ter Ordnung, es ist $a(\lambda) \ge g(\lambda)$.

Da $p_A(x)$ höchstens n Nullstellen haben kann, ist $a(\lambda) \le n$, und da λ nur dann Eigenwert ist, wenn es dazu mindestens einen Eigenvektor gibt, ist $g(\lambda) \ge 1$. ∎

Hilfssatz

$\lambda_1, \ldots, \lambda_r$ seien paarweise verschiedene Eigenwerte von $A \in M_{n,n}(K)$.

Ist $\mathbf{w}_i \in E(\lambda_i)$ für $i = 1, \ldots, r$ und $\mathbf{w}_1 + \cdots + \mathbf{w}_r = \mathbf{0}$, so ist $\mathbf{w}_1 = \ldots = \mathbf{w}_r = \mathbf{0}$.

BEWEIS: Ist $r = 1$, so ist nichts zu zeigen. Sei also $r \geq 2$.

Wir nehmen an, daß es ein k mit $1 \leq k \leq r$ gibt, so daß – nach geeigneter Numerierung – die ersten k Vektoren $\neq \mathbf{0}$ sind, während $\mathbf{w}_{k+1} = \ldots = \mathbf{w}_r = \mathbf{0}$ ist. Dann ist auch $\mathbf{w}_1 + \cdots + \mathbf{w}_k = \mathbf{0}$, also $k \geq 2$ und

$$\mathbf{0} = f_A\bigl(\sum_{i=1}^{k} \mathbf{w}_i\bigr) = \sum_{i=1}^{k} f_A(\mathbf{w}_i) = \sum_{i=1}^{k} \lambda_i \cdot \mathbf{w}_i.$$

Da Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, muß $\lambda_1 = \ldots = \lambda_k = 0$ sein. Aber das ist ein Widerspruch, denn die λ_i sollten paarweise verschieden sein. ■

Charakterisierung der Diagonalisierbarkeit

Sei $A \in M_{n,n}(K)$. A ist genau dann diagonalisierbar, wenn gilt:

1. *$p_A(x)$ zerfällt vollständig in Linearfaktoren.*
2. *Für jeden Eigenwert λ ist $a(\lambda) = g(\lambda)$.*

BEWEIS: 1) Ist A diagonalisierbar, so gibt es eine Basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ von Eigenvektoren. Wir können die Basisvektoren so anordnen, daß gilt:

Die ersten k_1 von ihnen sind Eigenvektoren zum Eigenwert λ_1, die nächsten k_2 sind Eigenvektoren zum Eigenwert λ_2 usw., wobei die Eigenwerte $\lambda_1, \ldots, \lambda_r$ paarweise verschieden sind. Dann wird f_A bezüglich B durch die Diagonalmatrix

$$\Delta := \begin{pmatrix} \lambda_1 \cdot E_{k_1} & & & 0 \\ & \lambda_2 \cdot E_{k_2} & & \\ & & \ddots & \\ 0 & & & \lambda_r \cdot E_{k_r} \end{pmatrix}.$$

beschrieben, d.h. es ist $\Delta = M_B(f_A) = B^{-1} \cdot A \cdot B$, mit $B = (\vec{b}_1, \ldots, \vec{b}_n)$. Daher gilt:

$$p_A(x) = p_\Delta(x) = (\lambda_1 - x)^{k_1} \cdot \ldots \cdot (\lambda_r - x)^{k_r}.$$

Das charakteristische Polynom zerfällt in Linearfaktoren, und es ist $a(\lambda_i) = k_i$ für $i = 1, \ldots, r$.

Wir wissen schon, daß $k_i \geq g(\lambda_i)$ ist. Aber $B_i := B \cap E(\lambda_i)$ besteht aus k_i linear unabhängigen Vektoren, die alle in $E(\lambda_i)$ liegen. Also ist auch $k_i \leq \dim E(\lambda_i) = g(\lambda_i)$.

2) Sei umgekehrt das Kriterium erfüllt, $\lambda_1, \ldots, \lambda_r$ die verschiedenen Eigenwerte mit den Vielfachheiten $k_i = a(\lambda_i) = g(\lambda_i)$ und $k_1 + \cdots + k_r = n$.

Dann wählen wir Basen B_i von $E(\lambda_i)$, für $i = 1, \ldots, r$, und setzen

$$B := B_1 \cup \ldots \cup B_r.$$

Offensichtlich besteht B aus n Elementen. Eine Linearkombination $\mathbf{w}$ von Elementen von B kann in der Form $\mathbf{w} = \mathbf{w}_1 + \cdots + \mathbf{w}_r$ geschrieben werden, wobei $\mathbf{w}_i \in E(\lambda_i)$ jeweils Linearkombination von Elementen von B_i ist. Ist $\mathbf{w} = \mathbf{0}$, so verschwinden nach dem Hilfssatz auch alle $\mathbf{w}_i$, und da die Elemente von B_i linear unabhängig sind, kann $\mathbf{w}$ nur die triviale Linearkombination sein. Also sind die Elemente von B linear unabhängig und bilden somit eine Basis von Eigenvektoren von A, d.h. A ist diagonalisierbar. ■

Bemerkung. Über $\mathbb{C}$ zerfällt jedes Polynom in Linearfaktoren. Da ist die erste Bedingung überflüssig.

Ist λ Eigenwert einer Matrix A, so gibt es einen Vektor $\mathbf{x} \neq \mathbf{0}$ mit $f_A(\mathbf{x}) = \lambda \cdot \mathbf{x}$, und dann ist

$$f_{A^2}(\mathbf{x}) = f_A \circ f_A(\mathbf{x}) = f_A(\lambda\mathbf{x}) = \lambda \cdot f_A(\mathbf{x}) = \lambda \cdot (\lambda\mathbf{x}) = \lambda^2\mathbf{x}.$$

Also ist λ^2 Eigenwert der Matrix $A^2 = A \cdot A$, und allgemeiner ist λ^p Eigenwert der Matrix $A^p := \underbrace{A \cdot \ldots \cdot A}_{p\text{-mal}}$.

Bei diagonalisierbaren Matrizen kann man mehr aussagen und zudem die Matrixpotenz A^p bequem ausrechnen:

Ist $B^{-1}AB = D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ eine Diagonalmatrix, so ist

$$B^{-1}A^pB = \underbrace{(B^{-1}AB) \cdot \ldots \cdot (B^{-1}AB)}_{\text{p-mal}} = D^p.$$

Aber die Potenzen einer Diagonalmatrix kann man sofort hinschreiben, und es ist

$$A^p = B \cdot \begin{pmatrix} \lambda_1^p & & 0 \\ & \ddots & \\ 0 & & \lambda_n^p \end{pmatrix} \cdot B^{-1}.$$

Als Anwendung kann man sogar mit Polynomen von (diagonalisierbaren) Matrizen rechnen. Ist $f(x) = \sum_{i=1}^{n} a_i x^i$, so setzt man

$$f(A) := a_0 \cdot E_n + a_1 \cdot A + a_2 \cdot A^2 + \cdots + a_n \cdot A^n.$$

Ist A wie oben diagonalisierbar, so erhält man:

$$f(A) = B \cdot (a_0 \cdot E_n + a_1 \cdot D + a_2 \cdot D^2 + \cdots + a_n \cdot D^n) \cdot B^{-1} = B \cdot \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} \cdot B^{-1}.$$

Eine besondere Situation liegt vor, wenn $f(x) = p_A(x)$ das charakteristische Polynom von A ist. Dann ist $f(\lambda_i) = 0$ für alle i, also $p_A(A) = 0$. Diese hier nur für diagonalisierbare Matrizen bewiesene Aussage ist ein Spezialfall eines allgemeinen Satzes.

Satz von Cayley-Hamilton

Ist $A \in M_{n,n}(K)$, so gilt: $\quad p_A(A) = 0.$

Der Beweis ist zu schwer für uns. Hier soll nur vor folgendem „Kurzschluß“ gewarnt werden:

$p_A(x) = \det(A - xE_n)$, also $p_A(A) = \det(A - AE_n) = \det(A - A) = \det(0) = 0$. Das ist natürlich völliger Unsinn!! (Warum?)

Eine wichtige Folgerung aus dem Satz von Cayley-Hamilton ist die Tatsache, daß A^n eine Linearkombination der Matrizen E_n, A, $A^2, \ldots, A^{n-1}$ ist.

Wir wollen uns jetzt mit den Eigenwerten spezieller Matrizen befassen:

Eigenwerte von orthogonalen und unitären Matrizen

Sei $A \in M_{n,n}(K)$ orthogonal (im Falle $K = \mathbb{R}$) oder unitär (im Falle $K = \mathbb{C}$). Dann gilt:

1. *Ist $\lambda \in K$ ein Eigenwert von A, so ist $|\lambda| = 1$.*
2. *Eigenvektoren zu verschiedenen Eigenwerten von A sind zueinander orthogonal.*

BEWEIS: Wir betrachten nur den komplexen Fall. Es sei $f := f_A$.

1) Sei $\mathbf{x} \neq \mathbf{0}$, $f(\mathbf{x}) = \lambda \cdot \mathbf{x}$. Da f eine Isometrie ist, gilt:

$$\lambda\overline{\lambda} < \mathbf{x}, \mathbf{x} > = < \lambda\mathbf{x}, \lambda\mathbf{x} > = < f(\mathbf{x}), f(\mathbf{x}) > = < \mathbf{x}, \mathbf{x} > .$$

Daraus folgt, daß $|\lambda|^2 = 1$ ist.

2) Ist $f(\mathbf{x}) = \lambda\mathbf{x}$ und $f(\mathbf{y}) = \mu\mathbf{y}$, mit $\lambda \neq \mu$, so folgt:

$$\lambda\overline{\mu} < \mathbf{x}, \mathbf{y} > = < f(\mathbf{x}), f(\mathbf{y}) > = < \mathbf{x}, \mathbf{y} > .$$

Also ist $(\lambda\overline{\mu} - 1) \cdot < \mathbf{x}, \mathbf{y} > = 0$. Wäre $\lambda\overline{\mu} = 1$, so wäre $\overline{\mu} = \lambda^{-1} = \overline{\lambda}$ (weil $|\lambda| = 1$ ist). Aber dann wäre auch $\lambda = \mu$, im Widerspruch zur Voraussetzung. Also muß $< \mathbf{x}, \mathbf{y} > = 0$ sein. ∎

Eigenwerte von symmetrischen und hermiteschen Matrizen

Sei $A \in M_{n,n}(K)$ symmetrisch (d.h. $K = \mathbb{R}$ und $A^t = A$) oder hermitesch (d.h. $K = \mathbb{C}$ und $\overline{A}^t = A$). Dann gilt:

1. *Das charakteristische Polynom von A zerfällt in Linearfaktoren.*
2. *Alle Eigenwerte von A sind reell.*
3. *Eigenvektoren zu verschiedenen Eigenwerten von A sind zueinander orthogonal.*

Beweis: Eine reelle symmetrische Matrix ist natürlich auch hermitesch, und über $\mathbb{C}$ zerfällt das charakteristische Polynom in Linearfaktoren. Wir brauchen also nur zu zeigen, daß alle Eigenwerte einer hermiteschen Matrix A reell sind, dann folgt (1) und (2) automatisch für $\mathbb{R}$ und $\mathbb{C}$.

Wieder arbeiten wir mit der zugehörigen linearen Abbildung $f = f_A$. Diesmal ist f „selbstadjungiert", d.h. $< f(\mathbf{z}), \mathbf{w} > = < \mathbf{z}, f(\mathbf{w}) >$ für alle $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$, denn es ist

$$\begin{aligned} < f(\mathbf{z}), \mathbf{w} > &= (A \cdot \mathbf{z}^t)^t \cdot \overline{\mathbf{w}}^t \\ &= \mathbf{z} \cdot A^t \cdot \overline{\mathbf{w}}^t \\ &= \mathbf{z} \cdot \overline{A} \cdot \overline{\mathbf{w}}^t \\ &= \mathbf{z} \cdot \overline{A \cdot \mathbf{w}^t} = \mathbf{z} \cdot \overline{f(\mathbf{w})}^t \\ &= < \mathbf{z}, f(\mathbf{w}) > . \end{aligned}$$

Ist $\lambda \in \mathbb{C}$ Eigenwert von f, so gibt es einen Vektor $\mathbf{z} \in \mathbb{C}^n$ mit $\mathbf{z} \neq \mathbf{0}$, so daß $f(\mathbf{z}) = \lambda\mathbf{z}$ ist. Dann folgt:

$$\begin{aligned}
\lambda < \mathbf{z}, \mathbf{z} > &= < \lambda\mathbf{z}, \mathbf{z} > = < f(\mathbf{z}), \mathbf{z} > \\
&= < \mathbf{z}, f(\mathbf{z}) > \\
&= < \mathbf{z}, \lambda\mathbf{z} > = \overline{\lambda} < \mathbf{z}, \mathbf{z} > .
\end{aligned}$$

Da $< \mathbf{z}, \mathbf{z} >$ reell und positiv ist, folgt: $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$.

Zur Orthogonalität der Eigenvektoren:

Seien $\lambda \neq \mu$ zwei Eigenwerte, $\mathbf{z}$ bzw. $\mathbf{w}$ zugehörige Eigenvektoren. Dann gilt:

$$\begin{aligned}
\lambda\cdot < \mathbf{z}, \mathbf{w} > &= < \lambda\mathbf{z}, \mathbf{w} > \\
&= < f(\mathbf{z}), \mathbf{w} > \\
&= < \mathbf{z}, f(\mathbf{w}) > \\
&= < \mathbf{z}, \mu\mathbf{w} > \\
&= \mu\cdot < \mathbf{z}, \mathbf{w} > \quad \text{(weil } \mu \text{ reell).}
\end{aligned}$$

Also ist $(\lambda - \mu)\cdot < \mathbf{z}, \mathbf{w} > = 0$, d.h. $< \mathbf{z}, \mathbf{w} > = 0$. ■

Satz von der Hauptachsentransformation

Ist $A \in M_{n,n}(K)$ eine symmetrische (bzw. hermitesche) Matrix, so gibt es eine orthogonale (bzw. unitäre) Matrix $S \in \mathrm{GL}(n, K)$, so daß $S^{-1} \cdot A \cdot S$ eine reelle Diagonalmatrix ist. Die Einträge in der Diagonalmatrix sind die Eigenwerte der Matrix A.

BEWEIS: Wir betrachten den komplexen Fall, also eine hermitesche Matrix A.

Es gibt mindestens einen Eigenwert λ und dazu einen Eigenvektor $\mathbf{v}_1$. Diesen kann man normieren, so daß $\|\mathbf{v}_1\| = 1$ ist. Schließlich setzen wir $U := \mathbb{C}\mathbf{v}_1$.

Offensichtlich ist $f_A(U) \subset U$, und wir wollen zeigen, daß auch $f_A(U^\perp) \subset U^\perp$ ist. Dazu geben wir uns einen beliebigen Vektor $\mathbf{x} \in U^\perp$ vor. Dann ist

$$<\mathbf{v}_1\,,\, f_A(\mathbf{x})> \; = \; <f_A(\mathbf{v}_1)\,,\, \mathbf{x}> \; = \; <\lambda \cdot \mathbf{v}_1\,,\, \mathbf{x}> \; = \lambda \cdot <\mathbf{v}_1\,,\, \mathbf{x}> \; = 0,$$

also auch $f_A(\mathbf{x}) \in U^\perp$.

Wir wählen nun eine ON-Basis $\{\mathbf{v}_2, \ldots, \mathbf{v}_n\}$ von $U^\perp$. Dann ist $B := \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ eine ON-Basis des $\mathbb{C}^n$ und $B := (\vec{v}_1, \ldots, \vec{v}_n)$ eine unitäre Matrix. Sei

$$M := M_B(f_A) = B^{-1} \cdot A \cdot B = \overline{B}^t \cdot A \cdot B.$$

Dann ist $\overline{M}^t = \overline{B}^t \cdot \overline{A}^t \cdot B = \overline{B}^t \cdot A \cdot B = M$, also M wieder hermitesch.

Wir wollen M genauer berechnen. Wegen $f_A(U^\perp) \subset U^\perp$ gibt es Zahlen α_{ji}, so daß $\vec{w}_i := A \cdot \vec{v}_i = \sum_{j=1}^{n} \alpha_{ji}\vec{v}_j$ ist, für $i = 2, \ldots, n$. Daraus folgt:

$$\begin{aligned}
M &= \begin{pmatrix} \overline{\mathbf{v}}_1 \\ \vdots \\ \overline{\mathbf{v}}_n \end{pmatrix} \cdot A \cdot (\vec{v}_1, \ldots, \vec{v}_n) \\
&= \begin{pmatrix} \overline{\mathbf{v}}_1 \\ \vdots \\ \overline{\mathbf{v}}_n \end{pmatrix} \cdot (\lambda \vec{v}_1, \vec{w}_2, \ldots, \vec{w}_n) \\
&= \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right).
\end{aligned}$$

Weil M hermitesch ist, muß auch A_1 hermitesch sein. Wir werden sehen, daß wir damit das Problem um eine Dimension reduziert haben. Zu diesem Zweck nehmen wir an, es gebe schon eine unitäre Matrix S_1, so daß

$$\Delta_1 := \overline{S}_1^{\,t} \cdot A_1 \cdot S_1$$

eine Diagonalmatrix ist. Anschließend setzen wir $S := B \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array}\right)$. Dann ist auch S unitär, und es ist

$$\begin{aligned}
S^{-1} \cdot A \cdot S &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \overline{S}_1^{\,t} \end{array}\right) \cdot \overline{B}^{\,t} \cdot A \cdot B \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array}\right) \\
&= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \overline{S}_1^{\,t} \end{array}\right) \cdot \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & A_1 \end{array}\right) \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S_1 \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \overline{S}_1^{\,t} \cdot A_1 \cdot S_1 \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & \Delta_1 \end{array}\right).
\end{aligned}$$

Das ist die gewünschte Diagonalisierung von A. Analog führt man die Diagonalisierung von A_1 auf eine $(n-2)$-reihige hermitesche Matrix A_2 zurück, usw. ■

Um den Begriff „Hauptachsentransformation" zu erläutern, sei noch ein Beispiel angegeben:

Beispiel.

Ist $A \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix, so nennt man $q_A : \mathbb{R}^n \to \mathbb{R}$ mit $q_A(\mathbf{x}) := \mathbf{x} \cdot A \cdot \mathbf{x}^t$ die *zugehörige quadratische Form.* Ist z.B. $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, so ist

$$\begin{aligned} q_A(x_1,x_2) &= (x_1,x_2)\cdot\begin{pmatrix} a & b \\ b & c \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1,x_2)\cdot\begin{pmatrix} ax_1+bx_2 \\ bx_1+cx_2 \end{pmatrix} = ax_1^2+2bx_1x_2+cx_2^2. \end{aligned}$$

Für $c \in \mathbb{R}$ bezeichnet man die Menge $Q_A(c) := \{\mathbf{x} : q_A(\mathbf{x}) = c\}$ als *Quadrik* oder *quadratische Hyperfläche*.

Wir betrachten speziell $A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$. Dann ist

$$p_A(x) = \det\begin{pmatrix} 5-x & -2 \\ -2 & 8-x \end{pmatrix} = x^2 - 13x + 36 = (x-4)(x-9).$$

Einen Eigenvektor zum Eigenwert $\lambda_1 = 4$ finden wir durch die Gleichung

$$\begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Eine Lösung ist der Vektor $\mathbf{v}_1 = (2,1)$, mit $\|\mathbf{v}_1\| = \sqrt{5}$.

Einen Eigenvektor zum Eigenwert $\lambda_2 = 9$ erhalten wir aus der Gleichung

$$\begin{pmatrix} -4 & -2 \\ -2 & -1 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Eine Lösung ist $\mathbf{v}_2 = (-1,2)$, mit $\|\mathbf{v}_2\| = \sqrt{5}$.

Dann bilden die Vektoren $\mathbf{w}_1 := \frac{1}{\sqrt{5}}(2,1)$ und $\mathbf{w}_2 := \frac{1}{\sqrt{5}}(1,-2)$ eine ON-Basis von Eigenvektoren von A. Die zugehörige Matrix

$$S = \begin{pmatrix} \dfrac{2}{\sqrt{5}} & -\dfrac{1}{\sqrt{5}} \\ \dfrac{1}{\sqrt{5}} & \dfrac{2}{\sqrt{5}} \end{pmatrix}$$

definiert eine Drehung f_S der Ebene um den Nullpunkt, und es ist

$$S^t\cdot A\cdot S = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}.$$

Nun gilt (mit $\mathbf{y} := (f_S)^{-1}(\mathbf{x}) = (S^{-1}\cdot\mathbf{x}^t)^t = (S^t\cdot\mathbf{x}^t)^t = \mathbf{x}\cdot S$) :

$$\begin{aligned} \mathbf{x}\in Q_A(c) &\iff q_A(\mathbf{x}) = c \\ &\iff \mathbf{x}\cdot A\cdot\mathbf{x}^t = c \\ &\iff \mathbf{x}\cdot(S\cdot\begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}\cdot S^t)\cdot\mathbf{x}^t = c \\ &\iff \mathbf{y}\cdot\begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}\cdot\mathbf{y}^t = c. \end{aligned}$$

Sei etwa $c = 36$. Dann ist

$$\begin{aligned} Q_A(c) &= \{\mathbf{x} \mid 5x_1^2 - 4x_1x_2 + 8x_2^2 = 36\} \\ &= f_S(\{\mathbf{y} : 4y_1^2 + 9y_2^2 = 36\}) \\ &= f_S(\{\mathbf{y} : (\frac{y_1}{3})^2 + (\frac{y_2}{2})^2 = 1\} \end{aligned}$$

das Bild einer achsenparallelen Ellipse unter der Drehung f_S. So haben wir die „Hauptachsen" von $Q_A(c)$ bestimmt, sie zeigen in die durch die Spalten von S vorgegebenen Richtungen und haben die Halbachsen 2 und 3.