Kapitel III

Lineare und multilineare Algebra

§1 Lineare Gleichungssysteme

Sei K ein beliebiger Körper.

Sind $V_1, \ldots, V_n$ K-Vektorräume, so ist auch

$$V_1 \times \ldots \times V_n := \{(x_1, \ldots, x_n) \mid x_i \in V_i \text{ für } i = 1, \ldots, n\}$$

ein K-Vektorraum, mit

$$\begin{aligned} (x_1, \ldots, x_n) + (y_1, \ldots, y_n) &:= (x_1 + y_1, \ldots, x_n + y_n) \\ \text{und} \quad \alpha \cdot (x_1, \ldots, x_n) &:= (\alpha \cdot x_1, \ldots, \alpha \cdot x_n). \end{aligned}$$

Ist $V_1 = V_2 = \ldots = V_n = K$, so erhält man den Vektorraum K^n.

Ist z.B. $\{a_1, \ldots, a_r\}$ eine Basis des Vektorraumes V und $\{b_1, \ldots, b_s\}$ eine Basis des Vektorraumes W, so ist

$$\{(a_1, 0), \ldots, (a_r, 0), (0, b_1), \ldots, (0, b_s)\}$$

eine Basis des Vektorraumes $V \times W$. Und allgemein gilt:

> *Sind $V_1, \ldots, V_n$ endlich-dimensionale Vektorräume mit Dimensionen $d_1, \ldots, d_n$, so ist auch $V_1 \times \ldots \times V_n$ endlich-dimensional, mit Dimension $d_1 + \cdots d_n$.*

Definition:

V und W seien zwei K-Vektorräume. Eine Abbildung $b : V \times W \to K$ heißt eine *Bilinearform*, falls gilt:

1. Für festes $x \in V$ ist $y \mapsto b(x, y)$ eine Linearform auf W, d.h. es gilt:
   (a) $b(x, y_1 + y_2) = b(x, y_1) + b(x, y_2)$.
   (b) $b(x, \alpha \cdot y) = \alpha \cdot b(x, y)$.
2. Für festes $y \in W$ ist $x \mapsto b(x, y)$ eine Linearform auf V, d.h. es gilt:
   (a) $b(x_1 + x_2, y) = b(x_1, y) + b(x_2, y)$.
   (b) $b(\alpha \cdot x, y) = \alpha \cdot b(x, y)$.

Eine Bilinearform $b : V \times V \to K$ nennt man auch eine *Bilinearform auf V*.

Bemerkung: Man beachte:

$$\text{Es ist} \quad b(\alpha \cdot x, y) = b(x, \alpha \cdot y), \quad \text{aber} \quad b(\alpha \cdot x, \alpha \cdot y) = \alpha^2 \cdot b(x, y).$$

Außerdem gilt:

$$b(x_1 + x_2, y_1 + y_2) = b(x_1, y_1) + b(x_1, y_2) + b(x_2, y_1) + b(x_2, y_2).$$

Beispiele:

1. Durch $b(\mathbf{x}, \mathbf{y}) := \mathbf{x} \bullet \mathbf{y} := \sum_{i=1}^{n} x_i y_i$ wird eine Bilinearform auf dem $\mathbb{R}^n$ definiert.

2. Das Skalarprodukt läßt sich folgendermaßen verallgemeinern:

   Ist K ein beliebiger Körper, so nennt man $b : K^n \times K^n \to K$ mit

   $$b((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := x_1 y_1 + \cdots x_n y_n$$

   die *kanonische Bilinearform* auf K^n. Sie hat zusätzlich folgende Eigenschaften:

   (a) $b(\mathbf{x}, \mathbf{y}) = b(\mathbf{y}, \mathbf{x})$.

   (b) Ist $b(\mathbf{x}, \mathbf{y}) = 0$ für alle $\mathbf{y} \in K^n$, so ist $\mathbf{x} = \mathbf{0}$.

   Beweis für die zweite Aussage:

   Sei $\mathbf{x} = (x_1, \ldots, x_n)$. Setzt man für $\mathbf{y}$ den Einheitsvektor $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ (mit einer 1 an der i-ten Stelle) ein, so erhält man:

   $$0 = b(\mathbf{x}, \mathbf{e}_i) = x_1 \cdot 0 + \ldots + x_i \cdot 1 + \ldots x_n \cdot 0 = x_i.$$

   Das gilt für jedes i, also ist $\mathbf{x} = \mathbf{0}$. $\square$

   Allerdings ist es i.a. nicht richtig, daß $b(\mathbf{x}, \mathbf{x})$ reell und ≥ 0 ist. Das geht schon im Falle $K = \mathbb{C}$ schief, und bei allgemeineren Körpern macht die Aussage überhaupt keinen Sinn.

3. Sei $a : K^2 \times K^2 \to K$ definiert durch

   $$a((x_1, x_2), (y_1, y_2)) := x_1 y_2 - x_2 y_1.$$

   Dann ist a eine Bilinearform, und es gilt:

   (a) $a(\mathbf{x}, \mathbf{y}) = -a(\mathbf{y}, \mathbf{x})$.

   (b) Ist $a(\mathbf{x}, \mathbf{y}) = 0$ für alle $\mathbf{y} \in K^2$, so ist $\mathbf{x} = \mathbf{0}$.

   Beweis: $a(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1 = -(y_1 x_2 - y_2 x_1) = -a(\mathbf{y}, \mathbf{x})$.

   Ist $a(\mathbf{x}, \mathbf{y}) = 0$ für alle $\mathbf{y} \in K^2$, so ist insbesondere $0 = a((x_1, x_2), (1, 0)) = -x_2$ und $0 = a((x_1, x_2), (0, 1)) = x_1$, also $\mathbf{x} = \mathbf{0}$. $\square$

Definition:

Sei $b : V \times V \to K$ eine Bilinearform auf dem K-Vektorraum V.

1. Ist $b(x,y) = b(y,x)$ für alle $x, y \in V$, so heißt b *symmetrisch.*
2. Ist $b(x,y) = -b(y,x)$ für alle $x, y \in V$, so heißt b *schiefsymmetrisch* oder *alternierend.*
3. Man nennt b *nicht entartet* (oder *nicht ausgeartet*), wenn gilt:

   Ist $b(x,y) = 0$ für alle $y \in V$, so ist $x = 0$.

Die kanonische Bilinearform auf dem K^n (und damit erst recht das euklidische Skalarprodukt auf dem $\mathbb{R}^n$) ist symmetrisch und nicht entartet.

Die durch $b((x_1,x_2),(y_1,y_2)) := x_1y_2 - x_2y_1$ definierte Bilinearform ist schiefsymmetrisch, aber auch nicht entartet.

Die durch $b_2((x_1,x_2,x_3),(y_1,y_2,y_3)) := x_1y_1 + x_2y_2$ definierte Bilinearform auf dem K^3 ist offensichtlich symmetrisch, aber ausgeartet, denn es ist $b_2(\mathbf{e}_3, \mathbf{y}) = 0$ für alle $\mathbf{y} \in K^3$, obwohl $\mathbf{e}_3 = (0,0,1)$ nicht der Nullvektor ist.

Wie im $\mathbb{R}^n$ gilt:

Darstellung von Linearformen

Sei $b : K^n \times K^n \to K$ die kanonische Bilinearform.

Ist $f : K^n \to K$ eine Linearform, so gibt es einen eindeutig bestimmten Vektor $\mathbf{m} \in K^n$, so daß gilt:

$$f(\mathbf{x}) = b(\mathbf{m}, \mathbf{x}), \qquad \text{für alle } \mathbf{x} \in K^n.$$

Beweis: Eine Linearform ist schon durch die Werte auf den Elementen einer Basis festgelegt:

$$f(\mathbf{x}) = f(\sum_{i=1}^{n} x_i \mathbf{e}_i) = \sum_{i=1}^{n} x_i \cdot f(\mathbf{e}_i).$$

Also setzen wir $m_i := f(\mathbf{e}_i)$ und $\mathbf{m} := (m_1, \ldots, m_n)$. Offensichtlich ist dann $f(\mathbf{x}) = b(\mathbf{m}, \mathbf{x})$ für jedes $\mathbf{x} \in K^n$.

Um auch die Eindeutigkeit der Darstellung zu beweisen, nehmen wir an, es gebe zwei Vektoren $\mathbf{m}_1, \mathbf{m}_2$ mit $f(\mathbf{x}) = b(\mathbf{m}_1, \mathbf{x}) = b(\mathbf{m}_2, \mathbf{x})$. Dann folgt:

$$b(\mathbf{m}_1 - \mathbf{m}_2, \mathbf{x}) = 0 \qquad \text{für alle } \mathbf{x} \in K^n.$$

Weil b nicht entartet ist, muß $\mathbf{m}_1 - \mathbf{m}_2 = \mathbf{0}$ sein, also $\mathbf{m}_1 = \mathbf{m}_2$. $\square$

Mit $M_{n,m}(K)$ bezeichnen wir den K-Vektorraum aller Matrizen mit Koeffizienten aus K

und n Zeilen und m Spalten. Das Rechnen mit solchen Matrizen geschieht genauso wie im Falle $K = \mathbb{R}$.

Ist $A = \left(a_{ij} \mid {}^{i=1,\ldots,n}_{j=1,\ldots m}\right)$ eine Matrix aus $M_{n,m}(K)$, so definiert man die *transponierte Matrix*

$$A^\top := \left(a^*_{ij} \mid \begin{matrix} i = 1, \ldots, n \\ j = 1, \ldots m \end{matrix}\right)$$

durch $a^*_{ij} := a_{ji}$. Anschaulich gesehen gewinnt man $A^\top$ aus A durch Spiegelung an der Hauptdiagonalen $(a_{11}, a_{22}, \ldots, a_{kk})$ (mit $k := \min(n, m)$). Zum Beispiel ist

$$\begin{pmatrix} 1 & 5 & 2 \\ 3 & 4 & 9 \end{pmatrix}^\top = \begin{pmatrix} 1 & 3 \\ 5 & 4 \\ 2 & 9 \end{pmatrix}.$$

Und speziell ist $\mathbf{x}^\top = M(\mathbf{x}) = \vec{x}$. Das Transponieren macht aus einem Zeilenvektor einen Spaltenvektor.

Die kanonische Bilinearform b kann nun in der Form

$$b(\mathbf{x}, \mathbf{y}) = \mathbf{x} \circ \mathbf{y}^\top$$

geschrieben werden.

Regeln für das Transponieren

1. $(A + B)^\top = A^\top + B^\top$.
2. $(A \circ B)^\top = B^\top \circ A^\top$.
3. $A^{\top\top} = A$.

Beweis: Ist A eine Matrix, so bezeichnen wir den Eintrag an der Stelle (i, j) der Einfachheit halber mit A_{ij}. Dann gilt:

$$\begin{array}{rrcl}
1) & ((A + B)^\top)_{ij} & = & (A + B)_{ji} \;=\; A_{ji} + B_{ji} \\
 & & = & (A^\top)_{ij} + (B^\top)_{ij} \;=\; (A^\top + B^\top)_{ij}. \\
2) & ((A \circ B)^\top)_{ij} & = & (A \circ B)_{ji} \\
 & & = & \sum_k A_{jk} B_{ki} \\
 & & = & \sum_k (B^\top)_{ik} (A^\top)_{kj} \\
 & & = & (B^\top \circ A^\top)_{ij}. \\
3) & (A^{\top\top})_{ij} & = & (A^\top)_{ji} \;=\; A_{ij}.
\end{array}$$

$\square$

Ist $A \in M_{n,m}(K)$, so bezeichnen wir die i-te Zeile von A mit $\mathbf{z}_i(A)$ und die j-te Spalte von A mit $\vec{s}_j(A)$. Offensichtlich gilt:

$$\vec{s}_j(A) = A \circ \vec{e}_j \quad \text{und} \quad \mathbf{z}_i(A) = \mathbf{e}_i \circ A.$$

Daraus folgt:

$$\mathbf{z}_i(A)^\top = \vec{s}_i(A^\top) \quad \text{und} \quad \vec{s}_j(A)^\top = \mathbf{z}_j(A^\top).$$

Wollen wir die j-te Spalte $\vec{s}_j(A)$ als Zeilenvektor schreiben, so können wir statt $\vec{s}_j(A)^\top$ auch das Symbol $\mathbf{s}_j(A)$ verwenden.

Definition:

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

Mit $\mathrm{rg}_s(A)$ bezeichnet man den *Spaltenrang* von A, die maximale Anzahl linear unabhängiger Spalten von A.

Mit $\mathrm{rg}_z(A)$ bezeichnet man den *Zeilenrang* von A, die maximale Anzahl linear unabhängier Zeilen von A.

Die der Matrix A zugeordnete lineare Abbildung $f_A : K^m \to K^n$ wird wie in Kapitel I im Falle $K = \mathbb{R}$ definiert:

$$f_A(\mathbf{x}) := (A \circ \mathbf{x}^\top)^\top = (\mathbf{z}_1(A) \circ \mathbf{x}^\top, \ldots, \mathbf{z}_n(A) \circ \mathbf{x}^\top).$$

Dann folgt:

Charakterisierung des Spaltenranges

Es ist $\quad \mathrm{rg}_s(A) = \dim \mathrm{Im}\,(f_A).$

BEWEIS: $\mathrm{Im}\,(f_A)$ wird erzeugt von den Vektoren

$$f_A(\mathbf{e}_j) = (A \circ \vec{e}_j)^\top = \mathbf{s}_j(A),$$

also den Spalten von A. Die maximale Anzahl linear unabhängiger Spalten gibt dann die Dimension von $\mathrm{Im}\,(f_A)$ an. $\square$

Von großer Bedeutung ist nun die folgende Aussage:

Zeilenrang = Spaltenrang

Für jede Matrix $A \in M_{n,m}(K)$ *gilt:*

$$\mathrm{rg}_z(A) = \mathrm{rg}_s(A).$$

BEWEIS: Sei $r := \mathrm{rg}_z(A)$, $\mathbf{a}_i = \mathbf{z}_i(A)$ für $i = 1, \ldots, n$. Wir müssen zeigen, daß $r = \dim \mathrm{Im}\,(f_A)$ ist. Wir führen das in drei Schritten durch.

1. Fall: Sei $r = n$. Dann ist zu zeigen, daß $f_A : K^m \to K^n$ surjektiv ist.

Weil nach unserer Voraussetzung die Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_n$ linear unabhängig sind, muß $n \leq m$ sein, und man kann die Vektoren zu einer Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ von K^m ergänzen.

Nun definieren wir eine lineare Abbildung $\widehat{f}: K^m \to K^m$ durch

$$\widehat{f}(\mathbf{x}) := (\mathbf{a}_1 \circ \mathbf{x}^\top, \ldots, \mathbf{a}_m \circ \mathbf{x}^\top).$$

Ist $\widehat{f}(\mathbf{x}) = \mathbf{0}$, so ist $\mathbf{a}_i \circ \mathbf{x}^\top = 0$ für $i = 1, \ldots, m$ und damit auch $\mathbf{y} \circ \mathbf{x}^\top = 0$ für jedes $\mathbf{y} \in K^m$. Weil die kanonische Bilinearform nicht entartet ist, muß dann schon $\mathbf{x} = \mathbf{0}$ sein.

Das bedeutet, daß $\widehat{f}$ injektiv und damit auch surjektiv ist. Ist nun ein $\mathbf{y} = (y_1, \ldots, y_n) \in K^n$ vorgegeben, so setzen wir

$$\mathbf{y}^* := (y_1, \ldots, y_n, 0, \ldots, 0) \in K^m.$$

Es gibt ein $\mathbf{x} \in K^m$ mit $\widehat{f}(\mathbf{x}) = \mathbf{y}^*$. Aber dann ist auch $f_A(\mathbf{x}) = \mathbf{y}$. Die Surjektivität von f_A ist gezeigt.

2. Fall: Sei $r < n$, und genau die ersten r Zeilen $\mathbf{a}_1, \ldots, \mathbf{a}_r$ seien linear unabhängig.

Definieren wir $g : K^m \to K^r$ durch $g(\mathbf{x}) := (\mathbf{a}_1 \circ \mathbf{x}^\top, \ldots, \mathbf{a}_r \circ \mathbf{x})$, so ist g nach Fall 1 surjektiv.

Wir führen nun eine weitere Abbildung $h : K^r \to K^n$ ein. Da $\mathbf{a}_{r+1}, \ldots, \mathbf{a}_n$ von $\mathbf{a}_1, \ldots, \mathbf{a}_r$ linear abhängen, gibt es Koeffizienten $\alpha_{\lambda,i} \in K$, so daß gilt:

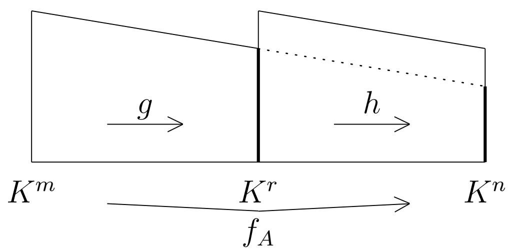
$$\mathbf{a}_{r+\lambda} = \sum_{i=1}^{r} \alpha_{\lambda,i} \mathbf{a}_i, \qquad \text{für } \lambda = 1, \ldots, n-r.$$

Nun setzen wir

$$h(y_1, \ldots, y_r) := (y_1, \ldots, y_r, \sum_{i=1}^{r} \alpha_{1,i} y_i, \ldots, \sum_{i=1}^{r} \alpha_{n-r,i} y_i).$$

Es folgt:

$$\begin{aligned} h \circ g(\mathbf{x}) &= h(\mathbf{a}_1 \circ \mathbf{x}^\top, \ldots, \mathbf{a}_r \circ \mathbf{x}^\top) \\ &= (\mathbf{a}_1 \circ \mathbf{x}^\top, \ldots, \mathbf{a}_r \circ \mathbf{x}^\top, \sum_{i=1}^{r} \alpha_{1,i} \mathbf{a}_i \circ \mathbf{x}^\top, \ldots, \sum_{i=1}^{r} \alpha_{n-r,i} \mathbf{a}_i \circ \mathbf{x}^\top) \\ &= (\mathbf{a}_1 \circ \mathbf{x}^\top, \ldots, \mathbf{a}_r \circ \mathbf{x}^\top, \mathbf{a}_{r+1} \circ \mathbf{x}^\top, \ldots, \mathbf{a}_n \circ \mathbf{x}^\top) \\ &= f_A(\mathbf{x}). \end{aligned}$$



Offensichtlich ist $\mathrm{Im}\,(f_A) = f_A(K^m) = h(g(K^m)) = h(\mathrm{Im}\,(g))$, und da g surjektiv ist, ist $\mathrm{Im}\,(f_A) = h(K^r) = \mathrm{Im}\,(h)$. Aber $\mathrm{Im}\,(h)$ wird von den Vektoren

$$h(\mathbf{e}_1) = (1, 0, \ldots, 0, \alpha_{1,1}, \ldots, \alpha_{n-r,1}), \ldots h(\mathbf{e}_r) = (0, \ldots, 0, 1, \alpha_{1,r}, \ldots, \alpha_{n-r,r})$$

erzeugt. Da diese Vektoren linear unabhängig sind, ist $\dim \mathrm{Im}\,(f_A) = \dim \mathrm{Im}\,(h) = r$.

3. Fall: Sei $r < n$, ohne weitere Bedingungen.

Dann gibt es eine Permutation $\sigma \in S_n$, so daß die Vektoren $\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(r)}$ linear unabhängig sind. Wir betrachten die lineare Abbildung $P_\sigma : K^n \to K^n$ mit

$$P_\sigma(x_1, \ldots, x_n) := (x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

P_σ ist sogar ein Isomorphismus (mit $(P_\sigma)^{-1} = P_{\sigma^{-1}}$), und es ist

$$P_\sigma \circ f_A(\mathbf{x}) = (\mathbf{a}_{\sigma(1)} \circ \mathbf{x}^\top, \ldots, \mathbf{a}_{\sigma(r)} \circ \mathbf{x}^\top, \ldots).$$

Aus (2) folgt, daß $\dim \operatorname{Im}(P_\sigma \circ f_A) = r$ ist. Und da P_σ ein Isomorphismus ist, hat $\operatorname{Im}(P_\sigma \circ f_A) = P_\sigma(\operatorname{Im}(f_A))$ die gleiche Dimension wie $\operatorname{Im}(f_A)$. Damit ist alles gezeigt. $\square$

Definition:

Die Zahl $\operatorname{rg}(A) := \operatorname{rg}_s(A) = \operatorname{rg}_z(A)$ bezeichnet man als den *Rang* von A.

Nun kommen wir endlich zum eigentlichen Thema dieses Paragraphen.

Definition:

Ein Lineares Gleichungssystem für die Unbekannten $x_1, \ldots, x_m$ mit Koeffizienten in K hat die Gestalt

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ \vdots & & & & & & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nm}x_m & = & b_n \end{array}$$

Fassen wir die Koeffizienten $a_{i1}, \ldots, a_{im}$ jeweils zu einem Vektor

$$\mathbf{a}_i := (a_{i1}, \ldots, a_{im}) \in K^m$$

und die Unbekannten x_j zu einem Vektor $\mathbf{x} = (x_1, \ldots, x_m) \in K^m$ zusammen, so können wir das Gleichungssystem auch in der folgenden Form schreiben:

$$\begin{array}{rcl} \mathbf{a}_1 \circ \mathbf{x}^\top & = & b_1, \\ & \vdots & \\ \mathbf{a}_n \circ \mathbf{x}^\top & = & b_n. \end{array}$$

Fassen wir die Vektoren $\mathbf{a}_i$ als Zeilen $\mathbf{z}_i(A)$ einer Matrix $A \in M_{n,m}(K)$ auf, und die b_i als Komponenten eines Vektors $\mathbf{b} = (b_1, \ldots, b_n) \in K^n$, so erhält das Gleichungssystem die Gestalt

$$f_A(\mathbf{x}) = \mathbf{b}.$$

Gerne benutzt man auch eine reine Matrizen-Schreibweise. Dann müssen $\mathbf{x}$ und $\mathbf{b}$ als Spaltenvektoren $\vec{x}$ und $\vec{b}$ geschrieben werden, und aus dem Gleichungssystem wird

$$A \circ \vec{x} = \vec{b}.$$

Wir wissen schon, daß der Lösungsraum des zugehörigen homogenen Gleichungssystems

$$A \circ \vec{x} = \vec{0}$$

der Vektorraum $\text{Ker}\,(f_A)$ ist. Die Lösungsmenge $\text{Lös}(A, \vec{b})$ des inhomogenen Systems kann eventuell leer sein. Existiert jedoch wenigstens eine partikuläre Lösung x_0, so ist $\text{Lös}(A, \vec{b})$ der affine Raum $\mathbf{x}_0 + \text{Ker}\,(f_A)$.

Man kann A und $\vec{b}$ zur *erweiterten Matrix* $(A, \vec{b}) \in M_{n,m+1}(K)$ zusammenfassen. Offensichtlich gilt:

$\text{Lös}(A, \vec{0}) = \text{Ker}\,(f_A)$. Das homogene LGS $A \circ \vec{x} = \vec{0}$ besitzt immer die „triviale Lösung“ $\vec{x} = \vec{0}$. Das ist genau dann die einzige Lösung, wenn f_A injektiv ist.

Das inhomogene LGS $A \circ \vec{x} = \vec{b}$ besitzt genau dann eine Lösung, wenn $\mathbf{b}$ in $\text{Im}\,(f_A)$ liegt.

1. Lösbarkeitskriterium

Ein inhomogenes LGS $A \circ \vec{x} = \vec{b}$ *ist genau dann lösbar, wenn* $\vec{b} \in \, < \vec{s}_1(A), \ldots, \vec{s}_m(A) >$ *ist.*

2. Lösbarkeitskriterium

Das inhomogene LGS $A \circ \vec{x} = \vec{b}$ *ist genau dann lösbar, wenn* $\text{rg}(A, \vec{b}) = \text{rg}(A)$ *ist.*

BEWEIS: Die Rang–Gleichung ist genau dann erfüllt, wenn $\vec{b}$ von den Spalten $\vec{s}_1(A), \ldots, \vec{s}_m(A)$ linear abhängt, wenn also $\vec{b} \in \, < \vec{s}_1(A), \ldots, \vec{s}_m(A) >$ ist. □

Das LGS ist genau dann eindeutig lösbar, wenn das zugehörige homogene System nur die triviale Lösung besitzt.

Beispiel:

Wir betrachten das LGS

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 + x_3 &= -2 \end{aligned}$$

In Matrizenschreibweise sieht das folgendermaßen aus:

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

Wir bestimmen zunächst die Lösungsmenge des zugehörigen homogenen Systems:

$$\begin{aligned}
\text{Lös}(A,\vec{0}) &= \{\vec{x} \mid x_1 + x_2 + x_3 = x_1 + x_3 = 0\} \\
&= \{\vec{x} \mid x_2 = 0 \text{ und } x_3 = -x_1\} \\
&= \{(x,0,-x)^\top \mid x \in \mathbb{R}\} \\
&= \mathbb{R}\cdot(1,0,-1)^\top.
\end{aligned}$$

Nun brauchen wir noch eine spezielle Lösung des inhomogenen Systems: Setzen wir versuchsweise $x_1 = 1$, so bleiben uns die Gleichungen

$$x_2 + x_3 = 0 \text{ und } x_3 = -3.$$

Also ist $\vec{x}_0 := (1,3,-3)^\top$ eine Lösung.

Die Gesamtlösungsmenge ist dann

$$\text{Lös}(A,\vec{b}) = \begin{pmatrix} 1 \\ 3 \\ -3 \end{pmatrix} + \mathbb{R}\cdot\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \{\begin{pmatrix} 1+x \\ 3 \\ -3-x \end{pmatrix} \mid x \in \mathbb{R}\}.$$

Im Allgemeinen geht es leider nicht so einfach. Zur Behandlung von großen Gleichungssystemen muß man systematisch vorgehen.

Wir betrachten eine Matrix $A = \left(a_{ij} \mid \begin{matrix} i = 1,\ldots,n \\ j = 1,\ldots,m \end{matrix}\right)$.

Wir wollen nun die Matrix verändern, indem wir Zeilen oder Spalten gewissen einfachen linearen Transformationen unterwerfen. Wenn wir das geschickt genug machen, bleibt der Lösungsraum erhalten oder wird nur unwesentlich verändert. Wir sprechen dann von *elementaren Umformungen*.

Typ (I):

Für $i \in \{1,\ldots,n\}$ und $\lambda \in K$, $\lambda \neq 0$, sei

$$\varepsilon_{i,\lambda} : K^n \to K^n \text{ definiert durch } \quad \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ \lambda\cdot x_i \\ \vdots \\ x_n \end{pmatrix}.$$

Das ist eine lineare Abbildung, ja sogar ein Isomorphismus mit Umkehrabbildung $\varepsilon_{i,1/\lambda}$.

Ist $A \in M_{n,m}(K)$, so wird bei der Transformation

$$A \mapsto \varepsilon_{i,\lambda}(A) := (\varepsilon_{i,\lambda}(\vec{s}_1(A)),\ldots,\varepsilon_{i,\lambda}(\vec{s}_m(A)))$$

schlicht und ergreifend die i–te Zeile von A mit λ multipliziert.

Typ (II):
Für $i, k \in \{1, \dots, n\}$ mit $i \neq k$ sei

$$\varepsilon_i^{+k} : K^n \to K^n \text{ definiert durch } \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_i + x_k \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix}.$$

Auch das ist eine lineare Abbildung, und auch ein Isomorphismus. Bei der Umkehrabbildung $\varepsilon_i^{-k} = \varepsilon_{k,-1} \circ \varepsilon_i^{+k} \circ \varepsilon_{k,-1}$ wird das k–te Element vom i–ten Element subtrahiert.

Bei der entsprechenden Matrizen–Operation

$$A \mapsto \varepsilon_i^{\pm k}(A) := (\varepsilon_i^{\pm k}(\vec{s}_1(A)), \dots, \varepsilon_i^{\pm k}(\vec{s}_m(A)))$$

wird die k–te Zeile zur i–ten addiert oder von ihr subtrahiert.

Eine *elementare Zeilenumformung* ist eine beliebige Kombination von Transformationen vom Typ (I) oder (II). In diese Kategorie gehört zum Beispiel auch die Addition des λ–fachen der k–ten Zeile zur i–ten Zeile (mit $\lambda \neq 0$ und $i \neq k$). Das nennt man manchmal eine Umformung vom Typ (III).
Sogar Vertauschungen von Zeilen lassen sich so produzieren:

$$\begin{pmatrix} a_i \\ a_k \end{pmatrix} \mapsto \begin{pmatrix} a_i \\ -a_k \end{pmatrix} \mapsto \begin{pmatrix} a_i \\ a_i - a_k \end{pmatrix} \mapsto \begin{pmatrix} a_i - (a_i - a_k) \\ a_i - a_k \end{pmatrix} = \begin{pmatrix} a_k \\ a_i - a_k \end{pmatrix} \mapsto \begin{pmatrix} a_k \\ a_i \end{pmatrix}.$$

Zeilenvertauschungen nennen wir auch Umformungen vom Typ (IV).

Mit Spalten könnte man entsprechende Umformungen vornehmen, aber man darf beide Sorten nicht so ohne weiteres mischen. Deshalb werden wir bei den Spaltenoperationen nur *Vertauschungen von Spalten* zulassen!

Ist $A \in M_{n,m}(K)$ eine Matrix und $\sigma \in S_m$ eine Permutation, so setzen wir $A_\sigma := (\vec{s}_{\sigma(1)}(A), \dots, \vec{s}_{\sigma(m)}(A))$. Das ist diejenige Matrix, die man aus A erhält, wenn man die Spalten gemäß σ permutiert.

Diese ganzen Umformungen sind natürlich nur dann sinnvoll, wenn dabei der Lösungsraum nicht oder nur unwesentlich verändert wird. Das wollen wir nun überprüfen.

Die Wirkung elementarer Umformungen

Wir betrachten ein inhomogenes LGS $A \circ \vec{x} = \vec{b}$ *mit* $A \in M_{n,m}(K)$.

1. *Ist* ε *eine elementare Zeilenumformung, so ist*

$$\text{Lös}(\varepsilon(A), \varepsilon(\vec{b})) = \text{Lös}(A, \vec{b}).$$

2. *Ist* $\sigma \in S_m$ *und* $P_\sigma : K^m \to K^m$ *der Isomorphismus, der die Komponenten gemäß* σ *permutiert, so ist*

$$\text{Lös}(A_\sigma, \vec{b}) = P_\sigma(\text{Lös}(A, \vec{b})).$$

Fazit: Elementare Zeilenoperationen ändern am Lösungsraum überhaupt nichts, wenn man sie auf die erweiterte Matrix $(A, \vec{b})$ anwendet. Spaltenvertauschungen sorgen für eine Permutation der Komponenten der Lösungsvektoren. Darüber muß man sorgfältig Buch führen.

Beweis: Zu den Zeilenumformungen ε :

$$\begin{aligned} A \circ \vec{x} = \vec{b} &\iff x_1 \vec{s}_1(A) + \cdots + x_m \vec{s}_m(A) = \vec{b} \\ &\iff \varepsilon(x_1 \vec{s}_1(A) + \cdots + x_m \vec{s}_m(A)) = \varepsilon(\vec{b}) \\ &\iff x_1 \cdot \varepsilon(\vec{s}_1(A)) + \cdots + x_m \cdot \varepsilon(\vec{s}_m(A)) = \varepsilon(\vec{b}) \\ &\iff \varepsilon(A) \circ \vec{x} = \varepsilon(\vec{b}). \end{aligned}$$

Bei den Spalten–Vertauschungen müssen wir nur beachten, daß (wegen des Kommutativgesetzes) folgendes gilt:

$$x_1 \vec{s}_1(A) + \cdots + x_m \vec{s}_m(A) = x_{\sigma(1)} \vec{s}_{\sigma(1)}(A) + \cdots + x_{\sigma(m)} \vec{s}_{\sigma(m)}(A).$$

Daraus folgt: $A \circ \vec{x} = A_\sigma \circ P_\sigma(\vec{x})$. $\square$

Was wollen wir durch die Umformungen erreichen? Wir wollen die Matrix auf eine besondere Dreiecksgestalt bringen, so daß dann das LGS ganz einfach lösbar wird. Zu dem Zweck betrachten wir die folgende Klasse von Matrizen:

Definition:

Eine Matrix $A \in M_{n,m}(K)$ soll r*–speziell* genannt werden, wenn sie folgende Gestalt besitzt:

$$A = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \\ \hline & 0 & & C \end{array}\right),$$

mit $B \in M_{r,m-r}(K)$, $C \in M_{n-r,m-r}(K)$ und $a_{11}, \ldots, a_{rr} \neq 0$. Im Kästchen links oben stehen unterhalb der Diagonalen nur Nullen.

Beispiele:

1. Jede Matrix A ist 0–speziell! Die Teilmatrix B ist dann gar nicht vorhanden, es ist $A = C$.
2. Ist $r = m = n$, so ist eine r-spezielle Matrix A eine „obere Dreiecks–Matrix“.
3. Es muß natürlich $r \le \min(n, m)$ sein.

   Ist $r = n \le m$, so ist C nicht vorhanden.
   Ist $r = m \le n$, so sind B und C beide nicht vorhanden.

Der Fall $C = 0$

Sei A r–speziell, $C = 0$ oder nicht vorhanden. Dann gilt:

1. *Die ersten r Spalten und die ersten r Zeilen sind jeweils linear unabhängig.*
2. *Es ist $\mathrm{rg}(A) = r$.*
3. *Das Gleichungssystem $A \circ \vec{x} = \vec{b}$ ist genau dann lösbar, wenn $b_{r+1} = b_{r+2} = \ldots = b_n = 0$ ist. Insbesondere ist es immer lösbar, wenn $r = n$ ist.*

BEWEIS:
1) Ist $\sum_{i=1}^{r} \beta_i \mathbf{z}_i(A) = \mathbf{0}$, so ist insbesondere

$$\beta_1 \cdot (a_{11}, \ldots, a_{1r}) + \beta_2 \cdot (0, a_{22}, \ldots, a_{2r}) + \cdots + \beta_r \cdot (0, \ldots, 0, a_{rr}) = (0, \ldots, 0),$$

also

$$\begin{array}{ccccccccc} \beta_1 a_{11} & & & & & & & = & 0, \\ \beta_1 a_{12} & + & \beta_2 a_{22} & & & & & = & 0, \\ \vdots & & & & \ddots & & & & \vdots \\ \beta_1 a_{1r} & + & \beta_2 a_{2r} & + & \cdots & + & \beta_r a_{rr} & = & 0. \end{array}$$

Offensichtlich muß dann $\beta_1 = \beta_2 = \ldots = \beta_r = 0$ sein, die ersten r Zeilen von A sind linear unabhängig.

Ähnlich folgt die Aussage über die ersten r Spalten.

2) Da es nur r Zeilen $\neq \mathbf{0}$ gibt, ist $\mathrm{rg}(A) = r$.

3) Wir müssen das folgende LGS betrachten:

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \\ \hline & 0 & & 0 \end{array}\right) \circ \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ * \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ * \end{pmatrix}.$$

Ist das LGS lösbar, so muß offensichtlich $b_{r+1} = \ldots = b_n = 0$ sein.

Ist umgekehrt diese Bedingung erfüllt, so können wir $x_{r+1} = \ldots = x_m = 0$ setzen und brauchen dann nur noch das kleinere System

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{rr} \end{pmatrix} \circ \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}$$

zu lösen, also

$$\begin{array}{ccccccc} x_1 a_{11} & + & \cdots & + & x_r a_{1r} & = & b_1, \\ & & \ddots & & \vdots & & \vdots \\ & & & & x_r a_{rr} & = & b_r. \end{array}$$

Aber das ist mit „Rückwärtseinsetzen“ ganz einfach zu bewerkstelligen:

Es ist

$$\begin{aligned} x_r &= \frac{1}{a_{rr}} \cdot b_r, \\ x_{r-1} &= \frac{1}{a_{r-1,r-1}} \cdot (b_{r-1} - x_r a_{r-1,r}), \\ x_{r-2} &= \frac{1}{a_{r-2,r-2}} \cdot (b_{r-2} - x_r a_{r-2,r} - x_{r-1} a_{r-2,r-1}) \\ &\text{usw.} \end{aligned}$$

□

Definition:

Wir nennen A eine *Gauß–Matrix*, wenn A r–speziell ist, und $C = 0$ oder nicht vorhanden.

Bevor wir nun versuchen, ein LGS so umzuformen, daß es durch eine Gauß–Matrix beschrieben wird, wollen wir noch untersuchen, wie sich der Rang bei elementaren Umformungen verhält.

Rang-Erhaltungssatz

Bei elementaren Zeilenumformungen oder Vertauschungen von Spalten ändert sich der Rang nicht.

BEWEIS: Sei $A \in M_{n,m}(K)$ gegeben. Es ist klar, daß sich der Spaltenrang bei einer Vertauschung der Spalten nicht ändert. Wir können deshalb o.B.d.A. annehmen, daß die ersten r Spalten linear unabhängig sind, und daß $\text{rg}(A) = r$ ist.

Ist nun $\varepsilon : K^n \to K^n$ eine elementare Zeilen-Umformung, so ist ε insbesondere ein Isomorphismus. Das bedeutet, daß die Vektoren

$$\varepsilon(\vec{s}_1(A)), \ldots, \varepsilon(\vec{s}_r(A))$$

ebenfalls linear unabhängig sind. Wenn es unter den Bildvektoren $\varepsilon(\vec{s}_j(A))$ sogar $r+1$ linear unabhängige gäbe, dann könnte man darauf den Isomorphismus ε^{-1} anwenden und hätte auch mindestens $r+1$ linear unabhängige Vektoren unter den $\vec{s}_j(A)$. Da dem nicht so ist, muß auch $\mathrm{rg}(\varepsilon(A)) = r$ sein. □

Nun gehen wir an die Lösung eines allgemeinen LGS:

Gauß'sches Eliminationsverfahren

Gegeben sei ein LGS $A \circ \vec{x} = \vec{b}$*, mit* $A \in M_{n,m}(K)$ *und* $\mathrm{rg}_s(A) = r$*. Durch endlich viele elementare Zeilenumformungen und Spaltenvertauschungen kann man die Koeffizientenmatrix in eine* r*–spezielle Gauß–Matrix verwandeln.*

Beweis:
1) Jede Matrix, also auch A, ist 0–speziell.

2) Reduktions–Schritt:

Es sei $0 \le k < r$. Wir nehmen an, A sei schon k–speziell:

$$A = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1k} & \\ \vdots & \ddots & \vdots & B_k \\ 0 & \cdots & a_{kk} & \\ \hline & 0 & & C_k \end{array}\right),$$

mit $B_k \in M_{k,m-k}(K)$, $C_k \in M_{n-k,m-k}(K)$ und $a_{11}, \ldots, a_{kk} \neq 0$. Wäre $C_k = 0$, so wäre $\mathrm{rg}_s(A) = k$. Das ist nicht der Fall, also muß $C_k \neq 0$ sein.

Das bedeutet, daß man ein a_{ij} in C_k finden kann, das $\neq 0$ ist. Wir nennen dieses Element das *Pivot–Element*. Es ist nicht eindeutig bestimmt, und es ist Inhalt umfangreicher numerischer Untersuchungen, dieses Pivot–Element möglichst geschickt zu wählen. Für uns hier reicht aber die Existenz eines solchen $a_{ij} \neq 0$. Durch Zeilenvertauschungen und Spaltenvertauschungen kann man erreichen, daß es das Element $a_{k+1,k+1}$ ist.

Nun subtrahiert man Vielfache der $(k+1)$–ten Zeile von den folgenden Zeilen, um die Elemente $a_{k+2,k+1}, \ldots, a_{n,k+2}$ zum Verschwinden zu bringen:

$$A \longrightarrow \left(\begin{array}{ccc|cc} a_{11} & \cdots & a_{1k} & & \\ \vdots & \ddots & \vdots & B_k & \\ 0 & \cdots & a_{kk} & & \\ \hline 0 & \cdots & 0 & a_{k+1,k+1} & \\ \cline{4-4} \vdots & & \vdots & 0 & * \\ \vdots & & \vdots & \vdots & \\ 0 & \cdots & 0 & 0 & \end{array}\right).$$

Das Ergebnis ist eine $(k+1)$–spezielle Matrix.

3) Abschluß des Verfahrens:

Ist schließlich $k = r$ erreicht, so sind die hinteren $m - r$ Spalten von den ersten r Spalten linear abhängig (denn der Spaltenrang bleibt ja immer gleich, und die ersten r Spalten sind linear unabhängig!). Das geht nur, wenn die rechts unten verbliebene Matrix C_r die Null–Matrix ist. Aber das bedeutet, daß wir eine Gauß-Matrix erhalten haben. $\square$

Wir wollen ein Beispiel durchrechnen:

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 3, \\ x_1 & - & x_2 & - & x_3 & = & 4, \\ x_1 & + & 3x_2 & + & 3x_3 & = & 1. \end{array}$$

Wir müssen die folgende erweiterte Matrix betrachten:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -1 & -1 & 4 \\ 1 & 3 & 3 & 1 \end{array}\right)$$

Da $a_{11} = 1 \neq 0$ ist, braucht man beim ersten Schritt nichts zu vertauschen. Man kann bereits a_{11} als Pivot–Element benutzen. Subtraktion der 1. Zeile von der 2. und 3. Zeile ergibt:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 2 & 2 & -2 \end{array}\right)$$

Diese Matrix ist 1–speziell. Als neues Pivot–Element kann man $a_{22} = -2$ benutzen. Addiert man die 2. Zeile zur 3. Zeile, so erhält man:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & -1 \end{array}\right)$$

Der Rang der ursprünglichen Matrix ist also 2, der Rang der erweiterten Matrix aber 3. Das LGS besitzt keine Lösung!

Wie man hier deutlich sehen konnte, dient das Gauß'sche Eliminationsverfahren auch zur Bestimmung des Ranges. Man braucht den Rang nicht schon vorher zu kennen

Ändert man in der Ausgangsgleichung $A \circ \vec{x} = \vec{b}$ den Vektor $\vec{b} = (3,4,1)^\top$ zu $(3,4,2)^\top$, so ergibt sich nach den Umformungen folgende erweiterte Matrix:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Das zugehörige LGS ist lösbar, wie man sofort mit der Rangbedingung erkennt. Um eine spezielle Lösung $\vec{x}_0$ zu erhalten, wählen wir etwa $x_3 = 0$. Dann erhalten wir das folgende vereinfachte und eindeutig lösbare System:

$$\begin{array}{rcrcl} x_1 & + & x_2 & = & 3, \\ & & -2x_2 & = & 1, \end{array} \qquad \text{also } \vec{x}_0 = \frac{1}{2} \cdot \begin{pmatrix} 7 \\ -1 \\ 0 \end{pmatrix}.$$

Wie bekommt man nun die allgemeine Lösung des zugehörigen homogenen Systems? Wir müssen folgendes System lösen:

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 0, \\ & - & 2x_2 & - & 2x_3 & = & 0. \end{array}$$

Setzt man für x_3 einen beliebigen Parameter λ ein, so wird daraus:

$$\begin{array}{rcrcl} x_1 & + & x_2 & = & -\lambda, \\ & - & 2x_2 & = & 2\lambda. \end{array}$$

Das ergibt $\vec{x} = \begin{pmatrix} 0 \\ -\lambda \\ \lambda \end{pmatrix}$. Also ist

$$\text{Lös}(A, \vec{b}) = \{\vec{x} \in K^3 \mid \exists\, \lambda \in \mathbb{R} \text{ mit } \vec{x} = \frac{1}{2} \cdot \begin{pmatrix} 7 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -\lambda \\ \lambda \end{pmatrix} \}.$$

Das hier angewandte Verfahren können wir verallgemeinern:

Wir betrachten die Gleichung $A \circ \vec{x} = \vec{0}$, wobei wir aber schon voraussetzen können, daß $A \in M_{n,m}(K)$ eine Gauß–Matrix vom Rang r ist. Da außerdem für $i > r$ die Zeilen $\mathbf{z}_i(A)$ alle $= \mathbf{0}$ sind und nichts zum Problem beitragen, können wir uns auf den Fall $n = r$ beschränken:

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \end{array}\right) \circ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Die obere Dreiecksmatrix auf der linken Seite von A bezeichnen wir mit D, und den Vektor $\vec{x}$ teilen wir in zwei Vektoren $\vec{x}^{\,*} \in K^r$ und $\vec{x}^{\,**} \in K^{m-r}$ auf, so daß gilt:

$$A = (D, B) \quad \text{und} \quad \vec{x} = \begin{pmatrix} \vec{x}^{\,*} \\ \vec{x}^{\,**} \end{pmatrix}.$$

Dann folgt:

$$\begin{aligned} A \circ \vec{x} = \vec{0} \quad &\Longleftrightarrow \quad D \circ \vec{x}^{\,*} + B \circ \vec{x}^{\,**} = \vec{0} \\ &\Longleftrightarrow \quad D \circ \vec{x}^{\,*} = -B \circ \vec{x}^{\,**}. \end{aligned}$$

Zu jedem Vektor $\vec{x}^{\,**} \in K^{m-r}$ erhalten wir durch Rückwärtseinsetzen einen eindeutig bestimmten Vektor $\vec{x}^{\,*}$, und beide Vektoren zusammen ergeben eine Lösung des Gleichungssystems. Das liefert uns auch einen Algorithmus zur Bestimmung einer Basis des Lösungsraumes:

Ist $\{\vec{e}_1, \ldots, \vec{e}_{m-r}\}$ die Standard–Basis des K^{m-r} und $\vec{u}_j$ jeweils der eindeutig bestimmte Vektor mit $D \circ \vec{u}_j = -B \circ \vec{e}_j$, für $j = 1, \ldots, m-r$, so bilden die Vektoren

$$\vec{x}_j := \begin{pmatrix} \vec{u}_j \\ \vec{e}_j \end{pmatrix}, \quad j = 1, \ldots, m - r\,,$$

eine Basis des Lösungsraumes $\text{Lös}(A, \vec{0}) = \text{Ker}\,(f_A)$.

Daß die gefundenen Vektoren Lösungen sind, ist klar. Ebenso, daß sie linear unabhängig sind. Und daß sie tatsächlich den Lösungsraum erzeugen, ergibt sich aus der Dimensionsformel:

Die Dimension des Lösungsraumes

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

$$\text{rg}(A) + \dim_K(\text{Lös}(A, \vec{0})) = m.$$

Beweis: Es ist $\text{rg}(A) = \dim \text{Im}\,(f_A)$ und $\dim_K(\text{Lös}(A, \vec{0})) = \dim \text{Ker}\,(f_A)$. $\square$

Folgerung

Ist $A \in M_{n,m}(K)$, so besitzt $\text{Lös}(A, \vec{0})$ eine Basis von $m - \text{rg}(A)$ Elementen.

Ist $\text{rg}(A) < m$ (sind also die Spalten von A linear abhängig), so gibt es eine nicht–triviale Lösung von $A \circ \vec{x} = \vec{0}$.

Beispiel:

Wir wollen jetzt ein etwas größeres Gleichungssystem betrachten:

$$\begin{array}{rcrcrcrcr}
x_1 & + & 3x_2 & - & 4x_3 & + & 3x_4 & = & 9, \\
3x_1 & + & 9x_2 & - & 2x_3 & - & 11x_4 & = & -3, \\
4x_1 & + & 12x_2 & - & 6x_3 & - & 8x_4 & = & 6, \\
2x_1 & + & 6x_2 & + & 2x_3 & - & 14x_4 & = & -12.
\end{array}$$

Für die systematische Anwendung des Gauß'schen Eliminationsverfahrens benutzen wir folgendes Schema:

x_1	x_2	x_3	x_4	b_i	
1	3	-4	3	9	
3	9	-2	-11	-3	
4	12	-6	-8	6	
2	6	2	-14	-12	
1	3	-4	3	9	bleibt stehen
0	0	10	-20	-30	$(-3 \times 1.$ Zeile)
0	0	10	-20	-30	$(-4 \times 1.$ Zeile)
0	0	10	-20	-30	$(-2 \times 1.$ Zeile)

Die jetzt erhaltene Matrix ist 1-speziell, und ein Pivot-Element findet sich in der 3. Spalte. Also muß eine Spalten-Vertauschung vorgenommen werden.

x_1	x_2	x_3	x_4	b_i	
1	3	-4	3	9	
0	0	10	-20	-30	
0	0	10	-20	-30	
0	0	10	-20	-30	
x_1	x_3	x_2	x_4		Spaltenvertauschung
1	-4	3	3	9	
0	10	0	-20	-30	
0	10	0	-20	-30	
0	10	0	-20	-30	
1	-4	3	3	9	
0	10	0	-20	-30	bleibt stehen
0	0	0	0	0	$(-\,2.$ Zeile)
0	0	0	0	0	$(-\,2.$ Zeile)

Das ergibt (z.B.) die folgende spezielle Lösung:

$$\begin{aligned} x_4 &= 0 \text{ (willkürlich gewählt)},\\ x_2 &= 0 \text{ (ebenfalls willkürlich)},\\ x_3 &= -3,\\ \text{und } x_1 &= 9 - 12 = -3. \end{aligned}$$

Eine Basis für den Lösungsraum des zugehörigen homogenen Systems erhält man, indem man für $(x_2, x_4)^\top$ die beiden möglichen Einheitsvektoren einsetzt und dann die zugehörigen Werte von x_1 und x_3 bestimmt:

$x_2 = 1$ und $x_4 = 0$ ergibt

$$\begin{aligned} x_1 - 4x_3 + 3 &= 0\\ \text{und} \quad 10x_3 &= 0, \end{aligned}$$

also $x_1 = -3$ und $x_3 = 0$.

$x_2 = 0$ und $x_4 = 1$ ergibt

$$\begin{aligned} x_1 - 4x_3 + 3 &= 0\\ \text{und} \quad 10x_3 - 20 &= 0, \end{aligned}$$

also $x_1 = 5$ und $x_3 = 2$.

Die allgemeine Lösung des inhomogenen Systems ist also gegeben durch

$$\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{pmatrix} = \begin{pmatrix} -3\\ 0\\ -3\\ 0 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -3\\ 1\\ 0\\ 0 \end{pmatrix} + \beta \cdot \begin{pmatrix} 5\\ 0\\ 2\\ 1 \end{pmatrix} = \begin{pmatrix} -3 - 3\alpha + 5\beta\\ \alpha\\ -3 + 2\beta\\ \beta \end{pmatrix}, \text{ mit } \alpha, \beta \in \mathbb{R}.$$

§2 Lineare Koordinaten und Basiswechsel

Sei V ein beliebiger endlich-dimensionaler K-Vektorraum, $A = \{a_1, \ldots, a_n\}$ eine Basis von V.

Zu jedem Vektor $x \in V$ gibt es eindeutig bestimmte Koeffizienten $x_1, \ldots, x_n \in K$, so daß gilt:

$$x = \sum_{i=1}^{n} x_i a_i.$$

Durch $\Phi_A(x) := (x_1, \ldots, x_n)$ wird also eine Abbildung $\Phi_A : V \to K^n$ definiert. Wir nennen Φ_A das durch A bestimmte *lineare Koordinatensystem* für V. Außerdem sei

$$[x]_A := \Phi_A(x)^\top = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Beispiele :

1. Ist V selbst der K^n und $A = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ die Standardbasis, so ist $\Phi_A = \mathrm{id}_V$, und $[\mathbf{x}]_A = \mathbf{x}^\top$.

2. Sei $V = K^n$ und $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ eine beliebige Basis, mit $\mathbf{a}_i = (a_{1i}, \ldots, a_{ni})$ für $i = 1, \ldots, n$, so gilt für $\mathbf{y} = (y_1, \ldots, y_n) \in V$:

$$\Phi_A(\mathbf{y}) = \mathbf{x} \iff \left\{ \begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ & & \vdots & & & & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nn}x_n & = & y_n. \end{array} \right.$$

   Die Koeffizientenmatrix hat den Rang n, weil ihre Spalten eine Basis des K^n bilden. Die erweiterte Matrix hat dann ebenfalls den Rang n, und das Gleichungssystem ist in jedem Fall eindeutig lösbar. Die Abbildung Φ_A ist durch den Lösungsalgorithmus bestimmt.

3. Sei $V = M_{2,2}(K)$. Dann bilden die Matrizen $E_{11} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ und $E_{22} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ eine Basis A von V, und $\Phi_A : M_{2,2}(K) \to K^4$ ist gegeben durch

$$\Phi_A(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = (a, b, c, d).$$

Sehr wichtig ist die Feststellung, daß das lineare Koordinatensystem Φ_A von der Basis A abhängt. Was passiert, wenn man eine zweite Basis $B = \{b_1, \ldots, b_n\}$ von V betrachtet. Welcher Zusammenhang besteht für einen Vektor $x \in V$ zwischen $\Phi_A(x)$ und $\Phi_B(x)$?

Zunächst einmal müssen wir überlegen, welche Eigenschaften Φ_A besitzt. Sehr leicht sieht man, daß Φ_A sogar eine lineare Abbildung ist. Und außerdem ist Φ_A bijektiv, die Umkehrabbildung ist durch

$$\Phi_A^{-1} : (x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} x_i a_i$$

gegeben.

Also ist $\Phi_A : V \to K^n$ und $\Phi_B : V \to K^n$ jeweils ein Isomorphismus, und $\Phi_B \circ \Phi_A^{-1} : K^n \to K^n$ ist ebenfalls ein Isomorphismus. Man kann sich das sehr gut an Hand des folgenden Diagramms veranschaulichen:

$$\begin{tikzcd} K^n \arrow[rr, "\Phi_B \circ \Phi_A^{-1}"] & & K^n \\ & V \arrow[ul, "\Phi_A"] \arrow[ur, "\Phi_B"'] & \end{tikzcd}$$

$\Phi_B \circ \Phi_A^{-1} : K^n \to K^n$ wird durch eine Matrix $W_{B,A} \in M_{n,n}(K)$ beschrieben. Es gilt:

$$\Phi_B \circ \Phi_A^{-1}(\Phi_A(x)) = \Phi_B(x).$$

In der Sprache der Matrizen heißt das:

$$W_{B,A} \circ [x]_A = [x]_B.$$

Man nennt $W_{B,A}$ die *Basiswechsel-Matrix.*

Wir können natürlich die Elemente von A durch die Elemente von B ausdrücken:

$$a_j = \sum_{i=1}^{n} \alpha_{ij} b_i, \quad j = 1, \ldots, n.$$

Das bedeutet:

$$[a_j]_B = (\alpha_{1j}, \ldots, \alpha_{nj})^\top, \text{ für } j = 1, \ldots, n.$$

Für die Spalten der Basiswechselmatrix $W = W_{B,A}$ gilt nun:

$$\vec{s}_j(W) = W \circ \vec{e}_j = W \circ [a_j]_A = [a_j]_B.$$

$$\text{Also ist} \quad W_{B,A} = ([a_1]_B, \ldots, [a_n]_B).$$

Um die Dinge einprägsam zu formulieren, begehen wir jetzt einen Notationsmißbrauch. Normalerweise dürfen wir nur Skalare (also Elemente des Körpers K) zu Vektoren oder Matrizen zusammenfassen. Aber die Gleichung

$$a_j = \alpha_{1j} b_1 + \cdots + \alpha_{nj} b_n$$

könnte man rein formal auch in der Form

$$a_j = (b_1, \ldots, b_n) \circ (\alpha_{1j}, \ldots, \alpha_{nj})^\top$$

schreiben. Tun wir dies, so erhalten wir die

Basiswechsel-Formeln

Sind $A = \{a_1, \ldots, a_n\}$ und $B = \{b_1, \ldots, b_n\}$ zwei Basen in einem Vektorraum V, so gilt für die Basiswechselmatrix $W_{B,A} = ([a_1]_B, \ldots, [a_n]_B)$:

$$\begin{aligned} (b_1, \ldots, b_n) \circ W_{B,A} &= (a_1, \ldots, a_n) \\ &\textit{und} \\ W_{B,A} \circ [x]_A &= [x]_B, \quad \textit{für } x \in V. \end{aligned}$$

Insbesondere folgt:

$$W_{A,B} \circ W_{B,A} \circ [x]_A = W_{A,B} \circ [x]_B = [x]_A$$

und

$$W_{B,A} \circ W_{A,B} \circ [x]_B = W_{B,A} \circ [x]_A = [x]_B.$$

Bezeichnen wir die n-reihige Einheitsmatrix $(\vec{e}_1, \ldots, \vec{e}_n)$ mit E_n, so erhalten wir (weil $[x]_A$ und $[x]_B$ jeweils völlig beliebig gewählt werden können):

$$W_{A,B} \circ W_{B,A} = E_n \quad \text{und} \quad W_{B,A} \circ W_{A,B} = E_n.$$

Definition:

Eine Matrix $A \in M_{n,n}(K)$ heißt *invertierbar*, falls gilt:

$$\exists\, A' \in M_{n,n}(K) \text{ mit } A \circ A' = A' \circ A = E_n.$$

Man nennt dann die (eindeutig bestimmte) Matrix A' die *inverse Matrix* zu A und bezeichnet sie mit A^{-1}.

Eigenschaften invertierbarer Matrizen

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

1. *A ist invertierbar.*
2. *$f_A : K^n \to K^n$ ist ein Isomorphismus (mit $(f_A)^{-1} = f_{A^{-1}}$).*
3. $\operatorname{Ker}(f_A) = \{0\}$.
4. *Das LGS $A \circ \vec{x} = \vec{0}$ ist eindeutig lösbar.*
5. *Das LGS $A \circ \vec{x} = \vec{b}$ ist für jedes $\vec{b} \in K^n$ eindeutig lösbar.*
6. $\operatorname{rg}(A) = n$.

Beweis:

(1) $\Longrightarrow$ (2):

Sei $B := A^{-1}$. Dann ist

$$\mathrm{id}_{K^n} = f_{E_n} = f_{A \circ B} = f_A \circ f_B \text{ und genauso } \mathrm{id}_{K^n} = f_B \circ f_A.$$

Also ist f_A bijektiv (und damit ein Isomorphismus) mit $(f_A)^{-1} = f_B$.

(2) $\Longrightarrow$ (3):

Ist f_A ein Isomorphismus, so ist natürlich $\operatorname{Ker}(f_A) = \{0\}$.

(3) $\Longrightarrow$ (4):

Es ist $\text{Lös}(A, \vec{0}) = \text{Ker}\,(f_A) = \{0\}$. Damit ist das LGS eindeutig (durch $\vec{x} = \vec{0}$) lösbar.

(4) $\Longrightarrow$ (5) :

Daß $A \circ \vec{x} = \vec{0}$ eindeutig lösbar ist, bedeutet, daß $f_A : K^n \to K^n$ injektiv ist. Aber dann ist f_A auch surjektiv, und jedes LGS $A \circ \vec{x} = \vec{b}$ lösbar. Der Lösungsraum hat die Gestalt $\vec{x}_0 + \text{Ker}\,(f_A) = \{\vec{x}_0\}$, also ist die Lösung auch eindeutig bestimmt.

(5) $\Longrightarrow$ (6) :

Nach der Dimensionsformel ist $\text{rg}(A) + \dim_K(\text{Ker}\,(f_A)) = n$. Die eindeutige Lösbarkeit des LGS $A \circ \vec{x} = \vec{0}$ besagt, daß $\text{Ker}\,(f_A) = \{0\}$ sein muß. Also ist $\text{rg}(A) = n$.

(6) $\Longrightarrow$ (1) :

Ist $\text{rg}(A) = n$, so bilden die Spalten $\vec{a}_j := \vec{s}_j(A)$ eine Basis von K^n, und f_A ist gegeben durch $\mathbf{a}_j \mapsto \mathbf{e}_j$, für $j = 1, \ldots, n$. Offensichtlich liefert die Zuordnung $\mathbf{e}_j \mapsto \mathbf{a}_j$ eine Umkehrabbildung zu f_A, und mit f_A ist auch A invertierbar. □

Definition:

Eine Matrix $A \in M_{n,n}(K)$ heißt *regulär*, wenn sie eine der äquivalenten Eigenschaften des obigen Satzes erfüllt.

Die Menge aller n-reihigen regulären Matrizen wird mit $\text{GL}(n, K)$ bezeichnet („General Linear Group“).

„Regulär“ ist also nur ein anderes Wort für „invertierbar“. Zur Rechtfertigung des Namens „Allgemeine Lineare Gruppe“ brauchen wir noch den folgenden Satz:

Gruppeneigenschaft der GL_n

$\text{GL}(n, K) := \{A \in M_{n,n}(K) \mid \text{rg}(A) = n\}$ *bildet mit der Matrizenmultiplikation eine Gruppe.*

BEWEIS:
1) E_n liegt in $\text{GL}(n, K)$ und spielt die Rolle des neutralen Elements.

2) Jedes $A \in \text{GL}(n, K)$ besitzt ein Inverses.

3) Seien $A, B \in \text{GL}(n, K)$. Dann ist $A \circ B \in M_{n,n}(K)$, und es gilt:

$$(A \circ B) \circ (B^{-1} \circ A^{-1}) = A \circ (B \circ B^{-1}) \circ A^{-1} = A \circ A^{-1} = E_n,$$

und genauso

$$(B^{-1} \circ A^{-1}) \circ (A \circ B) = E_n.$$

Also ist $A \circ B$ invertierbar, mit $(A \circ B)^{-1} = B^{-1} \circ A^{-1}$. Man beachte die Reihenfolge!!

4) Offensichtlich ist $A \circ (B \circ C) = (A \circ B) \circ C$. □

Beispiele :

1. $\underline{n = 1}$:

   $\mathrm{GL}(1, K) = \{a \in K \mid ax = 0 \text{ eindeutig lösbar } \} = K^* = K \setminus \{0\}$.

2. $\underline{n = 2}$:

   $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist genau dann regulär, wenn sowohl die beiden Spalten als auch die beiden Zeilen von A linear unabhängig sind.

   Ist etwa $c = 0$, so ist das genau dann der Fall, wenn $a \neq 0$ und $d \neq 0$ ist, also $ad - bc = ad \neq 0$.

   Ist $c \neq 0$ und $ad - bc = 0$, so ist $b = \dfrac{ad}{c}$ und daher

   $$\begin{pmatrix} b \\ d \end{pmatrix} = \frac{d}{c} \cdot \begin{pmatrix} a \\ c \end{pmatrix}.$$

   Dann kann A nicht regulär sein.

   Ist andererseits $c \neq 0$ und A nicht regulär, so gibt es Faktoren $\lambda, \mu \in K$ mit $(\lambda, \mu) \neq (0, 0)$ und

   $$\lambda \cdot \begin{pmatrix} a \\ c \end{pmatrix} + \mu \cdot \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

   Dann kann nicht $\mu = 0$ gelten, und daher ist

   $$b = -\frac{\lambda}{\mu} \cdot a \quad \text{ und } \quad d = -\frac{\lambda}{\mu} \cdot c,$$

   also

   $$ad - bc = -\frac{\lambda}{\mu} \cdot ac + \frac{\lambda}{\mu} \cdot ac = 0.$$

   Die Größe

   $$\det(A) := ad - bc$$

   nennt man die *Determinante* von A. Wir haben gezeigt:

   $$\mathrm{GL}(2, K) = \{A \in M_{2,2}(K) \mid \det(A) \neq 0\}.$$

3. Eine Diagonalmatrix

   $$D = \begin{pmatrix} d_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{nn} \end{pmatrix}$$

   ist genau dann invertierbar, wenn $d_{11}, \ldots, d_{nn} \neq 0$ sind.

   In diesem Spezialfall bezeichnet man die Größe $d_{11} \cdot d_{22} \cdot \ldots \cdot d_{nn}$ als *Determinante* $\det(D)$, und D ist genau dann regulär, wenn $\det(D) \neq 0$ ist.

   Wir werden später jeder Matrix eine Determinante zuordnen, die mißt, ob die Matrix regulär ist.

Satz von der Verkleinerung des Ranges

Sei $A \in M_{n,m}(K)$ und $B \in M_{m,l}(K)$.
Dann ist $\operatorname{rg}(A \circ B) \le \operatorname{rg}(A)$ *und* $\operatorname{rg}(A \circ B) \le \operatorname{rg}(B)$.

BEWEIS: Sei $f := f_A : K^m \to K^n$ und $g := f_B : K^l \to K^m$.

Dann ist $\operatorname{Im}(f \circ g) = f(\operatorname{Im}(g)) \subset f(K^m) = \operatorname{Im}(f)$, also

$$\operatorname{rg}(A \circ B) = \dim \operatorname{Im}(f \circ g) \le \dim \operatorname{Im}(f) = \operatorname{rg}(A).$$

Andererseits gilt: Ist $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ eine Basis von $\operatorname{Im}(g)$, so ist $\{f(\mathbf{x}_1), \ldots, f(\mathbf{x}_k)\}$ ein Erzeugendensystem von $f(\operatorname{Im}(g)) = \operatorname{Im}(f \circ g)$. Also ist

$$\operatorname{rg}(A \circ B) = \dim(f(\operatorname{Im}(g))) \le k = \dim \operatorname{Im}(g) = \operatorname{rg}(B).$$

$\square$

Invarianz des Ranges

Die Multiplikation mit einer regulären Matrix ändert den Rang nicht:

Ist $A \in M_{n,m}(K)$, $P \in \mathrm{GL}(n, K)$ und $Q \in \mathrm{GL}(m, K)$, so ist $\operatorname{rg}(P \circ A \circ Q) = \operatorname{rg}(A)$.

BEWEIS: Nach dem Satz von der Verkleinerung des Ranges ist $\operatorname{rg}(P \circ A \circ Q) \le \operatorname{rg}(A)$. Aber da $A = P^{-1} \circ (P \circ A \circ Q) \circ Q^{-1}$ ist, gilt auch die umgekehrte Ungleichung. $\square$

Wir wollen jetzt ein Verfahren suchen, wie man A^{-1} (für eine reguläre Matrix A) berechnen kann. Dazu benutzen wir die gleichen Methoden wie beim Gauß-Algorithmus.

Ist ε eine Folge von elementaren Zeilenumformungen, so ist $\varepsilon : K^n \to K^n$ ein (auf Spaltenvektoren wirkender) Isomorphismus, der durch eine reguläre Matrix P_ε beschrieben werden kann. Damit gilt:

$$\varepsilon(A) = (\varepsilon(\vec{s}_1(A)), \ldots, \varepsilon(\vec{s}_m(A))) = (P_\varepsilon \circ \vec{s}_1(A), \ldots, P_\varepsilon \circ \vec{s}_m(A)) = P_\varepsilon \circ A.$$

Eine Folge von Spaltenvertauschungen wird durch eine Permutation $\sigma \in S_m$ und den dazu gehörenden (auf Zeilenvektoren wirkenden) Isomorphismus $f_\sigma : K^m \to K^m$ beschrieben, mit

$$f_\sigma(x_1, \ldots, x_m) := (x_{\sigma(1)}, \ldots, x_{\sigma(m)}).$$

Setzen wir $Q_\sigma := (\vec{e}_{\sigma(1)}, \ldots, \vec{e}_{\sigma(m)}) \in M_{m,m}(K)$, so ist Q_σ regulär und $f_\sigma(\mathbf{x}) = \mathbf{x} \circ Q_\sigma$, also

$$A \circ Q_\sigma = \begin{pmatrix} a_{1,\sigma(1)} & \cdots & a_{1,\sigma(m)} \\ \vdots & & \vdots \\ a_{n,\sigma(1)} & \cdots & a_{n,\sigma(m)} \end{pmatrix} = (\vec{s}_{\sigma(1)}(A), \ldots, \vec{s}_{\sigma(m)}(A)) =: A_\sigma.$$

Da man A durch elementare Zeilenumformungen und Spaltenvertauschungen auf Gauß-Form bringen kann, bedeutet das:

Ist $r = \mathrm{rg}(A)$, so gibt es reguläre Matrizen P und Q, so daß $P \circ A \circ Q$ eine Gaußmatrix vom Rang r ist. Dabei kann Q wegfallen, wenn im Gaußverfahren keine Spaltenvertauschungen benötigt werden.

Nun betrachten wir den Fall, daß $A \in \mathrm{GL}(n, K)$ ist. Wir können annehmen, daß A schon r-speziell für ein $r \geq 0$ ist:

$$A = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1r} & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & a_{rr} & \\ \hline & & & \\ & 0 & & C \\ & & & \end{array}\right).$$

Wäre $r < n$ und würde die 1. Spalte von C verschwinden, so wäre die (r+1)-te Spalte von A linear abhängig von den ersten r Spalten. Andererseits ist $\mathrm{rg}(A) = n$, es sind also alle Spalten linear unabhängig.

Also muß es in der ersten Spalte von C ein Element $\neq 0$ geben, und mit Hilfe von Zeilenvertauschungen kann man es an die Stelle $(r+1, r+1)$ bringen. Das bedeutet, daß man keine Spaltenvertauschung braucht, um A auch $(r+1)$-speziell zu machen.

Dieses Verfahren läßt sich fortsetzen, solange $r < n$ ist. Das bedeutet: Es gibt eine Matrix $P \in \mathrm{GL}(n, K)$, die einer Folge von Zeilenumformungen entspricht, so daß $P \circ A$ eine obere Dreiecksmatrix vom Rang n ist. Durch weitere Zeilenumformungen kann man daraus schließlich sogar die Einheitsmatrix machen.

Ist aber $P \circ A = E_n$, so ist $P = A^{-1}$. Um nun bei der Durchführung des Gaußverfahrens gleichzeitig auch die Matrix P zu erhalten, erweitern wir A zur Matrix (A, E_n). Dann gilt:

$$P \circ (A, E_n) = (P \circ A, P \circ E_n) = (E_n, A^{-1}).$$

Zusammengefaßt ergibt das:

Verfahren zur Invertierung von Matrizen

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

1. *Es gibt eine Folge ε von Zeilenumformungen mit $\varepsilon(A) = E_n$.*
2. *Ist $\varepsilon(A, E_n) = (E_n, A^*)$, so ist $A^* = A^{-1}$.*

Beweis: Der erste Teil wurde gerade gezeigt.

Wenn ε durch die Matrix P verwirklicht wird, hat man die Gleichung

$$P \circ (A, E_n) = (E_n, A^*),$$

also $P \circ A = E_n$ und $P = A^*$. Damit ist $A^* = A^{-1}$. ∎

Beispiel:

Sei $B := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

Da die Spalten von B eine Basis des $\mathbb{R}^3$ bilden, ist $\mathrm{rg}(B) = 3$. Wir berechnen nun B^{-1} nach dem beschriebenen Verfahren:

$$
\begin{array}{ccc|cccl}
\hline
1 & 1 & 0 & 1 & 0 & 0 & \\
1 & 0 & 1 & 0 & 1 & 0 & \\
0 & 1 & 1 & 0 & 0 & 1 & \\
\hline
1 & 1 & 0 & 1 & 0 & 0 & \\
0 & -1 & 1 & -1 & 1 & 0 & (\,-\ 1.\ \text{Zeile}) \\
0 & 1 & 1 & 0 & 0 & 1 & \\
\hline
1 & 0 & 1 & 0 & 1 & 0 & (\,+\ 2.\ \text{Zeile}) \\
0 & -1 & 1 & -1 & 1 & 0 & \\
0 & 0 & 2 & -1 & 1 & 1 & (\,+\ 2.\ \text{Zeile}) \\
\hline
1 & 0 & 1 & 0 & 1 & 0 & \\
0 & 1 & -1 & 1 & -1 & 0 & (\,\times(-1)\,) \\
0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & (\,\times\frac{1}{2}\,) \\
\hline
1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & (\,-\ 3.\ \text{Zeile}) \\
0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & (\,+\ 3.\ \text{Zeile}) \\
0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \\
\hline
\end{array}
$$

Also ist $B^{-1} = \dfrac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

Mit Hilfe der inversen Matrix lassen sich manche Dinge einfacher beschreiben.

1. Sei $A \circ \vec{x} = \vec{b}$ ein LGS mit einer regulären Matrix A. Dann erhält man die eindeutig bestimmte Lösung sofort durch
$$\vec{x} = A^{-1} \circ \vec{b}.$$

2. Sind A und B zwei Basen des Vektorraumes V, so besteht für jeden Vektor $x \in V$ die Gleichung
$$W_{B,A} \circ [x]_A = [x]_B.$$

Da die Basiswechselmatrix $W_{B,A}$ auf jeden Fall regulär ist, kann man die Gleichung nach $[x]_A$ auflösen und erhält

$$[x]_A = W_{B,A}^{-1} \circ [x]_B.$$

Natürlich ist $W_{B,A}^{-1} = W_{A,B}$, und der Vorteil ist nicht sofort sichtbar. Betrachten wir daher den Spezialfall $V = K^n$.

Ist $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ eine Basis von K^n, so ist das zugehörige Koordinatensystem $\Phi_A : K^n \to K^n$ gegeben durch

$$\Phi_A(x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n) := (x_1, \ldots, x_n).$$

Diese Beschreibung von Φ_A ist unhandlich. Setzen wir jedoch $E := \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ (Standardbasis), so ist

$$W_{E,A}([\mathbf{x}]_A) = [\mathbf{x}]_E = \vec{x}, \quad \text{also } \Phi_A(\mathbf{x}) = (W_{E,A}^{-1} \circ \vec{x})^\top.$$

Nun ist $W_{E,A} = ([\mathbf{a}_1]_E, \ldots, [\mathbf{a}_n]_E) = (\vec{a}_1, \ldots, \vec{a}_n)$. Das liefert die Gleichung

$$\boxed{[\mathbf{x}]_A = (\vec{a}_1, \ldots, \vec{a}_n)^{-1} \circ \vec{x}.}$$

Beispiel:

Die Vektoren $\mathbf{a}_1 := (1,1,0)$, $\mathbf{a}_2 := (1,0,1)$ und $\mathbf{a}_3 := (0,1,1)$ bilden eine Basis A des $\mathbb{R}^3$. Die Matrix $B := (\vec{a}_1, \vec{a}_2, \vec{a}_3)$ haben wir oben schon einmal betrachtet.

Wir wollen herausfinden, wie die Darstellung des Vektors $\mathbf{x} := (12, 6, 30)$ bezüglich der Basis A aussieht. Nach der obigen Formel ist

$$[\mathbf{x}]_A = B^{-1} \circ \vec{x} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix} = \begin{pmatrix} -6 \\ 18 \\ 12 \end{pmatrix}.$$

Probe: Tatsächlich ist

$$-6\vec{a}_1 + 18\vec{a}_2 + 12\vec{a}_3 = \begin{pmatrix} -6 \\ -6 \\ 0 \end{pmatrix} + \begin{pmatrix} 18 \\ 0 \\ 18 \end{pmatrix} + \begin{pmatrix} 0 \\ 12 \\ 12 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \\ 30 \end{pmatrix} = \vec{x}.$$

Eine lineare Abbildung $f : K^m \to K^n$ wird durch eine eindeutig bestimmte Matrix $M(f) \in M_{n,m}(K)$ beschrieben, mit

$$M(f) \circ \mathbf{x}^\top = f(\mathbf{x})^\top.$$

Die Spalten von $M(f)$ gewinnt man also durch die Formeln

$$\vec{s}_j(M(f)) = f(\mathbf{e}_j)^\top, \quad \text{für } j = 1, \ldots, m.$$

Ist $M(f) = \left(a_{ij} \;\middle|\; \begin{array}{l} i = 1, \ldots, n \\ j = 1, \ldots, m \end{array} \right)$, so ist

$$f(\mathbf{e}_j) = (a_{1j}, \ldots, a_{nj}) = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i,$$

wobei links die Einheitsvektoren im K^m und rechts die Einheitsvektoren im K^n gemeint sind.

Die Beschreibung von f durch die Matrix $M(f)$ benutzt die Darstellung von Vektoren mit Hilfe der Standardbasen. Indem wir nun beliebige Basen benutzen, können wir auch lineare Abbildungen zwischen beliebigen endlich-dimensionalen Vektorräumen durch Matrizen beschreiben. Allerdings hängt die Beschreibung von der Auswahl der Basen ab!

Gegeben seien nun

- Zwei endlich-dimensionale K-Vektorräume V und W,
- zwei Basen $A = \{a_1, \ldots, a_m\}$ und $B = \{b_1, \ldots, b_n\}$ von V bzw. W,
- eine lineare Abbildung $f : V \to W$.

f ist durch die m Werte $f(a_1), \ldots, f(a_m)$ bereits vollständig festgelegt. Und jeder der Vektoren $f(a_j)$ läßt sich auf eindeutige Weise als Linearkombination der Basisvektoren $b_1, \ldots, b_n$ darstellen:

$$f(a_j) = \sum_{i=1}^{n} \alpha_{ij} b_i, \qquad \text{mit } \alpha_{ij} \in K, \qquad \text{für } j = 1, \ldots, m.$$

Es sei $M_{B,A}(f) \in M_{n,m}(K)$ die Matrix der α_{ij}.

Wir wollen jetzt die Situation mit Hilfe von Koordinaten beschreiben. Dazu betrachten wir das folgende Diagramm:

$$\begin{array}{ccccc} & V & \xrightarrow{f} & W & \\ \Phi_A & \downarrow & & \downarrow & \Phi_B \\ & K^m & \xrightarrow{f_{B,A}} & K^n & \end{array}$$

Die Abbildung $f_{B,A} := \Phi_B \circ f \circ \Phi_A^{-1} : K^m \to K^n$ ist ebenfalls linear und kann standardmäßig durch die Matrix $M(f_{B,A})$ beschrieben werden. Wir zeigen, daß $M(f_{B,A}) = M_{B,A}(f)$ ist. Es gilt nämlich:

$$\begin{aligned} \vec{s}_j(M(f_{B,A})) &= f_{B,A}(\mathbf{e}_j)^\top \\ &= (\Phi_B \circ f \circ \Phi_A^{-1}(\mathbf{e}_j))^\top \\ &= (\Phi_B \circ f(a_j))^\top \\ &= (\Phi_B(\alpha_{1j}b_1 + \cdots + \alpha_{nj}b_n))^\top \\ &= (\alpha_{1j}, \ldots, \alpha_{nj})^\top \\ &= \vec{s}_j(M_{B,A}(f)). \end{aligned}$$

Definition:

Die Matrix $M_{B,A}(f) := M(f_{B,A})$ zur linearen Abbildung

$$f_{B,A} = \Phi_B \circ f \circ \Phi_A^{-1} : K^m \to K^n$$

nennt man *die Matrix, die $f : V \to W$ bezüglich der Basen A und B beschreibt.*

Da $f_{B,A}(\Phi_A(x)) = \Phi_B(f(x))$ ist, folgt nun:

$$\boxed{M_{B,A}(f) \circ [x]_A = [f(x)]_B, \quad \text{für } x \in V.}$$

Es ist $[a_j]_A = \vec{e}_j$, für $j = 1, \ldots, m$, also

$$\vec{s}_j(M_{B,A}(f)) = M_{B,A}(f) \circ \vec{e}_j = M_{B,A}(f) \circ [a_j]_A = [f(a_j)]_B.$$

Das ergibt folgende Formel zur Berechnung von $M_{B,A}(f)$:

$$\boxed{M_{B,A}(f) = ([f(a_1)]_B, \ldots, [f(a_m)]_B).}$$

Ein Sonderfall liegt vor, wenn $V = W$ ist und wenn auf beiden Seiten die gleiche Basis A benutzt wird. Dann bezeichnen wir die Matrix zur linearen Abbildung $f_{A,A} = \Phi_A \circ f \circ \Phi_A^{-1} : K^n \to K^n$ mit $M_A(f)$. Es gilt:

$$\boxed{M_A(f) = ([f(a_1)]_A, \ldots, [f(a_n)]_A).}$$

Beispiele:

1. Sei $V := \{p \in \mathbb{R}[x] \; : \; \text{grad}(p) \leq 3\}$ der Vektorraum der Polynome vom Grad ≤ 3. Dann ist $A := \{1, x, x^2, x^3\}$ eine Basis von V. Das Koordinatensystem $\Phi_A : V \to \mathbb{R}^4$ ist dann gegeben durch

$$\Phi_A(a_0 + a_1 x + a_2 x^2 + a_3 x^3) = (a_0, a_1, a_2, a_3).$$

Nun betrachten wir die lineare Abbildung $D : V \to V$ mit $D(p) := p'$ (Ableitung von p). Es ist

$$D(a_0 + a_1 x + a_2 x^2 + a_3 x^3) = a_1 + 2a_2 + 3a_3.$$

Die Matrix $M := M_A(D)$, die D bezüglich der Basis A beschreibt, hat also die Gestalt

$$M = ([D(1)]_A, [D(x)]_A, [D(x^2)]_A, [D(x^3)]_A) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Der Rang dieser Matrix ist 3, also ist $\dim \text{Im}\,(D) = 3$ und $\dim \text{Ker}\,(D) = 4 - 3 = 1$. Tatsächlich ist der Kern der Unterraum der konstanten Polynome, der von 1 erzeugt wird.

Ist etwa $p(x) = 2x - 5x^2 + x^3$, so ist $[p]_A = (0, 2, -5, 1)^\top$, und

$$[p']_A = M \circ [p]_A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \circ \begin{pmatrix} 0 \\ 2 \\ -5 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -10 \\ 3 \\ 0 \end{pmatrix} = [2 - 10x + 3x^2]_A.$$

Tatsächlich ist $p'(x) = 2 - 10x + 3x^2$.

2. Sei $V := \mathbb{C}$, aufgefaßt als Vektorraum über $\mathbb{R}$, mit der Basis $A = \{1, \mathbf{j}\}$.

   Für $z = x + y\mathbf{j}$ ist $\Phi_A(z) = (x, y)$, also $[z]_A = \begin{pmatrix} x \\ y \end{pmatrix}$.

   Sei nun $w := a + b\mathbf{j} \in \mathbb{C}$ fest gewählt und $f : \mathbb{C} \to \mathbb{C}$ definiert durch $f(z) := w \cdot z$. Diese Abbildung ist $\mathbb{R}$-linear, und ihre Matrix bezüglich A ist gegeben durch

   $$M := M_A(f) = ([f(1)]_A, [f(\mathbf{j})]_A) = ([a + b\mathbf{j}]_A, [-b + a\mathbf{j}]_A) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

   Da jeder komplexen Zahl w genau eine solche Matrix zugeordnet ist, kann man $\mathbb{C}$ auch als Menge aller Matrizen der Form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M_{2,2}(\mathbb{R})$ auffassen.

3. Am Anfang dieses Paragraphen haben wir schon den Raum $V = M_{2,2}(K)$ mit der Basis $A = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ betrachtet. Das Koordinaten system $\Phi_A : V \to K^4$ ist gegeben durch

   $$\Phi_A(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = (a, b, c, d).$$

   Wir betrachten nun die Abbildung $f : V \to V$ mit

   $$f(X) := X \circ I - I \circ X, \quad I := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

   Hier ist nicht sofort erkennbar, daß es sich um eine lineare Abbildung handelt, aber man kann es leicht nachrechnen:

   $$\begin{aligned} f(X + Y) &= (X + Y) \circ I - I \circ (X + Y) \\ &= X \circ I + Y \circ I - I \circ X - I \circ Y \\ &= (X \circ I - I \circ X) + (Y \circ I - I \circ Y) \\ &= f(X) + f(Y) \\ \text{und} \qquad f(\alpha \cdot X) &= (\alpha \cdot X) \circ I - I \circ (\alpha \cdot X) \\ &= \alpha \cdot (X \circ I - I \circ X) \\ &= \alpha \cdot f(X). \end{aligned}$$

   Man rechnet nun leicht nach:

   $$f(E_{11}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = E_{12} + E_{21},$$

$$\begin{aligned} f(E_{12}) &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -E_{11} + E_{22}, \\ f(E_{21}) &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -E_{11} + E_{22}, \\ \text{und} \quad f(E_{22}) &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = -E_{12} - E_{21}. \end{aligned}$$

Also wird f beschrieben durch die Matrix

$$M := M_A(f) = \begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Offensichtlich ist $\text{rg}(M) = 2$, also $\dim \text{Im}\,(f) = 2$ und $\dim \text{Ker}\,(f) = 4 - 2 = 2$. Wir wollen noch eine Basis des Kerns von f bestimmen.

Das wäre auf dem üblichen Wege möglich (Gauß-Verfahren), aber da wir die Dimension des Kerns schon kennen, versuchen wir es mit Raten. Offensichtlich ist $E_2 \circ I - I \circ E_2 = I - I = 0$ und $I \circ I - I \circ I = 0$. Also ist $\{E_2, I\}$ eine Basis von $\text{Ker}\,(f)$.

Die Beispiele handelten alle von linearen Abbildungen eines Vektorraumes auf sich. Solche Abbildungen nennt man auch *Endomorphismen*.

Wir wollen nun ganz speziell Endomorphismen von K^n untersuchen. Es sei $A \in M_{n,n}(K)$ eine quadratische Matrix und $f = f_A : K^n \to K^n$ die zugehörige lineare Abbildung. Dann ist $A = M_E(f)$, wenn wir mit E die Standardbasis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ bezeichnen.

Ist nun $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eine beliebige Basis, so suchen wir nach einer Formel für die Matrix $M_B(f_A)$.

Wir wissen schon:

$$W_{E,B} = (\vec{b}_1, \ldots, \vec{b}_n).$$

Wir bezeichnen die reguläre Matrix, deren Spalten die Elemente der Basis B bilden, mit dem Buchstaben $\mathbf{B}$. Dann ist

$$W_{E,B} = \mathbf{B} \quad \text{und} \quad [\mathbf{x}]_B = (W_{E,B})^{-1} \circ [\mathbf{x}]_E = \mathbf{B}^{-1} \circ \vec{x}.$$

Nun ist $M_B(f_A) = ([f_A(\mathbf{b}_1)]_B, \ldots, [f_A(\mathbf{b}_n)]_B)$, mit

$$[f_A(\mathbf{b}_j)]_B = [(A \circ \vec{b}_j)^\top]_B = \mathbf{B}^{-1} \circ A \circ \vec{b}_j, \text{ für } j = 1, \ldots, n.$$

Daraus folgt:

$$\boxed{M_B(f_A) = \mathbf{B}^{-1} \circ A \circ \mathbf{B}.}$$

Definition:

Zwei Matrizen $M, M' \in M_{n,n}(K)$ heißen *ähnlich*, falls es eine reguläre Matrix P mit $M' = P^{-1} \circ M \circ P$ gibt.

Unter dem *Normalformenproblem für Endomorphismen* verstehen wir die Aufgabe, zu einer gegebenen Matrix eine möglichst einfache ähnliche Matrix zu finden. In speziellen Fällen werden wir das später durchführen. Im allgemeinen würde man stets die sogenannte „Jordansche Normalform"erhalten, auf die wir auch noch zu sprechen kommen, aber der Beweis für deren Existenz ist schwierig, und sie ist auch nicht gut für numerische Verfahren geeignet.

Im Falle des $\mathbb{R}^n$ wollen wir nun mit Hilfe des Skalarproduktes eine spezielle Klasse von Basen auszeichnen.

Definition:

Ein *Orthogonalsystem* im $\mathbb{R}^n$ ist eine Menge $\mathcal{O} \subset \mathbb{R}^n$ von paarweise zueinander orthogonalen Vektoren $\neq \mathbf{0}$. (Es ist also $\mathbf{v} \bullet \mathbf{w} = 0$ für $\mathbf{v}, \mathbf{w} \in \mathcal{O}$ mit $\mathbf{v} \neq \mathbf{w}$).

Orthogonalsysteme sind linear unabhängig

Sei $\mathcal{O} \subset \mathbb{R}^n$ ein Orthogonalsystem. Dann ist $\mathcal{O}$ linear unabhängig. Insbesondere enthält $\mathcal{O}$ höchstens n Elemente.

Beweis:
Es seien $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathcal{O}$ beliebig gewählt. Nach Voraussetzung sind alle $\mathbf{a}_i \neq \mathbf{0}$. Ist $\sum_{i=1}^{k} \lambda_i \mathbf{a}_i = \mathbf{0}$, so gilt für $l = 1, \ldots, k$:

$$0 = \mathbf{0} \bullet \mathbf{a}_l = \sum_{i=1}^{k} \lambda_i \mathbf{a}_i \bullet \mathbf{a}_l = \lambda_l \cdot \| \mathbf{a}_l \|^2.$$

Also ist $\lambda_l = 0$ für $l = 1, \ldots, k$. Damit sind die Vektoren linear unabhängig, und da sie beliebig ausgewählt wurden, ist ganz $\mathcal{O}$ linear unabhängig. □

Definition:

Ein *Orthonormalsystem* (kurz: *ON-System*) im $\mathbb{R}^n$ ist ein Orthogonalsystem $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$, in dem die Vektoren $\mathbf{a}_i$ zusätzlich normiert sind, in dem also $\| \mathbf{a}_i \| = 1$ für alle i ist.

Eine *Orthonormalbasis* (kurz: *ON-Basis*) ist eine Basis des $\mathbb{R}^n$, die zugleich ein ON-System ist.

Analog kann man auch Orthonormalbasen von Untervektorräumen des $\mathbb{R}^n$ definieren.

Man beachte: Ein System $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ von Vektoren des $\mathbb{R}^n$ ist genau dann ein ON-System, wenn für alle i und j gilt:

$$\mathbf{a}_i \bullet \mathbf{a}_j = \delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst.} \end{cases}$$

Man nennt δ_{ij} das *Kronecker-Symbol.*

Beispiele :

1. Die Standardbasis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ ist eine ON-Basis des $\mathbb{R}^n$.

2. Die Vektoren $\mathbf{a}_1 := (1, 1)$ und $\mathbf{a}_2 := (1, -1)$ bilden ein Orthogonalsystem im $\mathbb{R}^2$, sie sind aber nicht normiert.

   $\mathbf{b}_1 := \dfrac{1}{\sqrt{2}}(1, 1)$ und $\mathbf{b}_2 := \dfrac{1}{\sqrt{2}}(1, -1)$ bilden ein ON-System und natürlich auch eine ON-Basis.

Die Überprüfung, ob ein System von Vektoren eine Basis bildet, ist in solchen Fällen sehr einfach geworden. Besonders bequem ist auch die Ermittlung der Koordinaten eines Vektors bezüglich einer ON-Basis:

Die Koordinaten bezüglich einer ON-Basis

Sei $U \subset \mathbb{R}^n$ ein Untervektorraum, $A := \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ eine ON-Basis von U. Ist $\mathbf{x} \in U$ ein beliebiger Vektor, so findet man die Koeffizienten x_i in der Darstellung

$$\mathbf{x} = \sum_{i=1}^{k} x_i \mathbf{a}_i$$

durch die Formel $\boxed{x_i = \mathbf{x} \bullet \mathbf{a}_i, \text{ für } i = 1, \ldots, k.}$

Es ist also $[\mathbf{x}]_A = (\mathbf{x} \bullet \mathbf{a}_1, \ldots, \mathbf{x} \bullet \mathbf{a}_k)^\top$.

Der BEWEIS ist trivial, die Aussage sehr nützlich.

Im $\mathbb{R}^n$ kennen wir schon ein Beispiel für eine ON-Basis. Bei einem beliebigen Unterraum des $\mathbb{R}^n$ stellt sich zunächst die Frage, ob es dort überhaupt ON-Basen gibt. Daß das in der Tat der Fall ist, zeigt der folgende Satz:

Schmidt'sches Orthogonalisierungsverfahren

Sei $U \subset \mathbb{R}^n$ ein k-dimensionaler Untervektorraum und $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ eine Basis von U.
Dann gibt es ein ON-System $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ in U, so daß gilt:

$$\langle \mathbf{a}_1, \ldots, \mathbf{a}_l \rangle = \langle \mathbf{x}_1, \ldots, \mathbf{x}_l \rangle \text{ für } l = 1, \ldots, k.$$

BEWEIS: Wir konstruieren die $\mathbf{a}_i$ rekursiv aus den $\mathbf{x}_i$.

Sei $\mathbf{a}_1 := \dfrac{1}{\|\mathbf{x}_1\|} \cdot \mathbf{x}_1$. Dann ist $\|\mathbf{a}_1\| = 1$ und $\langle \mathbf{a}_1 \rangle = \langle \mathbf{x}_1 \rangle$.

Nun nehmen wir an, wir hätten $\mathbf{a}_1, \dots, \mathbf{a}_l$ schon mit den gewünschten Eigenschaften konstruiert. Dann sei

$$\mathbf{b}_{l+1} := \mathbf{x}_{l+1} - \sum_{i=1}^{l} (\mathbf{x}_{l+1} \bullet \mathbf{a}_i)\mathbf{a}_i.$$

Es gilt:

1. $\mathbf{b}_{l+1} \in \langle \mathbf{x}_1, \dots, \mathbf{x}_{l+1} \rangle$.
2. Für $j = 1, \dots, l$ ist $\mathbf{b}_{l+1} \bullet \mathbf{a}_j = \mathbf{x}_{l+1} \bullet \mathbf{a}_j - \mathbf{x}_{l+1} \bullet \mathbf{a}_j = 0$.
3. $\mathbf{a}_1, \dots, \mathbf{a}_l, \mathbf{b}_{l+1}$ sind linear unabhängig (wegen (2)).

Also bilden $\mathbf{a}_1, \dots, \mathbf{a}_l, \mathbf{b}_{l+1}$ ein Orthogonalsystem mit

$$\langle \mathbf{a}_1, \dots, \mathbf{a}_l, \mathbf{b}_{l+1} \rangle = \langle \mathbf{x}_1, \dots, \mathbf{x}_{l+1} \rangle.$$

Schließlich setzen wir $\mathbf{a}_{l+1} := \dfrac{1}{\| \mathbf{b}_{l+1} \|} \cdot \mathbf{b}_{l+1}$. □

Folgerung

Jeder Untervektorraum des $\mathbb{R}^n$ besitzt eine ON-Basis.

BEWEIS: Klar! □

Beispiel:

Die Vektoren $\mathbf{x}_1 = (1,1,0)$ und $\mathbf{x}_2 = (0,1,0)$ bilden eine Basis des Untervektorraumes $U := \{\mathbf{x} \in \mathbb{R}^3 \mid x_3 = 0\}$. Das Schmidt'sche Verfahren liefert:

$$\begin{aligned}
\mathbf{a}_1 &= \frac{1}{\sqrt{2}}(1,1,0), \\
\mathbf{b}_2 &= \mathbf{x}_2 - (\mathbf{x}_2 \bullet \mathbf{a}_1)\mathbf{a}_1 = (-\frac{1}{2}, \frac{1}{2}, 0) \\
\text{Dann ist} \quad \| \mathbf{b}_2 \| &= \frac{1}{\sqrt{2}}, \quad \text{und wir setzen} \\
\mathbf{a}_2 &= \sqrt{2} \cdot \mathbf{b}_2 = \frac{1}{\sqrt{2}}(-1,1,0).
\end{aligned}$$

Definition:

Ist $U \subset \mathbb{R}^n$ ein Untervektorraum, so heißt

$$U^\perp := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \bullet \mathbf{u} = 0 \text{ für alle } \mathbf{u} \in U\}$$

das *orthogonale Komplement* von U.

Eigenschaften des orthogonalen Komplements

Sei $U \subset \mathbb{R}^n$ ein Untervektorraum. Dann ist $U^\perp$ ebenfalls ein Untervektorraum, und es gilt:

1. $\dim(U) + \dim(U^\perp) = n$.
2. *Jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ kann eindeutig zerlegt werden:*

$$\mathbf{x} = \mathbf{u} + \mathbf{v}, \ \textit{mit } \mathbf{u} \in U \textit{ und } \mathbf{v} \in U^\perp.$$

3. *Es ist $U \cap U^\perp = \{\mathbf{0}\}$.*

Beweis: Zunächst zeigen wir, daß $U^\perp$ ein Unterraum ist:

Offensichtlich ist $\mathbf{0} \in U^\perp$.
Sind $\mathbf{x}, \mathbf{y} \in U^\perp$, so ist $(\mathbf{x} + \mathbf{y}) \bullet \mathbf{u} = \mathbf{x} \bullet \mathbf{u} + \mathbf{y} \bullet \mathbf{u} = 0$ für alle $\mathbf{u} \in U$, also $\mathbf{x} + \mathbf{y} \in U^\perp$.
Ist $\mathbf{x} \in U^\perp$ und $\lambda \in \mathbb{R}$, so ist $(\lambda \cdot \mathbf{x}) \bullet \mathbf{u} = \lambda \cdot (\mathbf{x} \bullet \mathbf{u}) = 0$ für alle $\mathbf{u} \in U$, also $\lambda \cdot \mathbf{x} \in U^\perp$.

Nun kommen wir zu den weiteren Eigenschaften:

1) Wir wählen eine ON-Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ von U. Dann definieren wir eine lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^k$ durch

$$f(\mathbf{x}) := (\mathbf{x} \bullet \mathbf{a}_1, \ldots, \mathbf{x} \bullet \mathbf{a}_k).$$

Da $f(\mathbf{a}_i) = \mathbf{e}_i$ jeweils der i-te Einheitsvektor ist, für $i = 1, \ldots, k$, ist f surjektiv.

$$\text{Weiter ist} \quad f(\mathbf{x}) = 0 \iff \mathbf{x} \bullet \mathbf{u} = 0 \text{ für alle } \mathbf{u} \in U$$
$$\iff \mathbf{x} \in U^\perp,$$

also $U^\perp = \operatorname{Ker}(f)$. Nach der Dimensionsformel ist

$$\dim(U) + \dim(U^\perp) = k + \dim(\operatorname{Ker}(f)) = \dim(\operatorname{Im}(f)) + \dim(\operatorname{Ker}(f)) = n.$$

2) Es gibt eine ON-Basis $\{\mathbf{a}_{k+1}, \ldots, \mathbf{a}_n\}$ von $U^\perp$. Da definitionsgemäß $\mathbf{a}_i \bullet \mathbf{a}_j = 0$ für $i = 1, \ldots, k$ und $j = k+1, \ldots, n$ ist, bildet

$$\{\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n\}$$

eine ON-Basis des ganzen $\mathbb{R}^n$.

Ist $\mathbf{x} \in \mathbb{R}^n$ beliebig vorgegeben, so ist $\mathbf{x} = \sum_{\nu=1}^{n} \lambda_\nu \mathbf{a}_\nu = \mathbf{u} + \mathbf{v}$ mit

$$\mathbf{u} := \sum_{i=1}^{k} \lambda_i \mathbf{a}_i \quad \text{und} \quad \mathbf{v} := \sum_{j=k+1}^{n} \lambda_j \mathbf{a}_j.$$

Das ist die gewünschte Zerlegung. Die Eindeutigkeit folgt aus (3).

3) Ist $\mathbf{x} \in U \cap U^\perp$, so muß $\mathbf{u} \bullet \mathbf{u} = 0$ sein, also $\mathbf{u}$ der Nullvektor.

Wir tragen nun den Beweis der Eindeutigkeit in (2) nach:

Sind durch $\mathbf{x} = \mathbf{u} + \mathbf{v} = \mathbf{u}' + \mathbf{v}'$ zwei Zerlegungen gegeben, so ist $\mathbf{u} - \mathbf{u}' = \mathbf{v}' - \mathbf{v} \in U \cap U'$, also $= \mathbf{0}$. Aber das bedeutet, daß $\mathbf{u} = \mathbf{u}'$ und $\mathbf{v} = \mathbf{v}'$ ist. $\square$

Folgerung

Ist $U \subset \mathbb{R}^n$ ein Untervektorraum, so ist $U^{\perp\perp} = U$.

Beweis: Ist $\mathbf{x} \in U$, so ist $\mathbf{x} \bullet \mathbf{v} = 0$ für alle $\mathbf{v} \in U^\perp$. Daher ist $U \subset U^{\perp\perp}$. Wegen der Dimensionsformel ist

$$\dim(U^{\perp\perp}) = n - \dim(U^\perp) = n - (n - \dim(U)) = \dim(U).$$

Also muß $U^{\perp\perp} = U$ sein. $\square$

Ist $U \subset \mathbb{R}^n$ ein Untervektorraum, so kann man eine Projektion von $\mathbb{R}^n$ auf U wie folgt definieren:

Da jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ eine **eindeutige** Zerlegung $\mathbf{x} = \mathbf{u} + \mathbf{v}$ mit $\mathbf{u} \in U$ und $\mathbf{v} \in U^\perp$ besitzt, definiert man einfach

$$\mathrm{pr}_U : \mathbb{R}^n \to U$$

durch

$$\mathrm{pr}_U : \mathbf{u} + \mathbf{v} \mapsto \mathbf{u}.$$

Diese Abbildung ist vernünftig definiert und - wie man leicht nachrechnen kann - auch linear. Wir nennen pr_U die *orthogonale Projektion auf U*. Sie hat folgende Eigenschaften:

1. $\mathrm{pr}_U \circ \mathrm{pr}_U(\mathbf{x}) = \mathrm{pr}_U(\mathbf{x})$ für alle $\mathbf{x} \in \mathbb{R}^n$.
2. $\mathrm{Ker}\,(\mathrm{pr}_U) = U^\perp$.
3. $\mathrm{pr}_U : \mathbb{R}^n \to U$ ist surjektiv.
4. $\|\, \mathbf{x} - \mathrm{pr}_U(\mathbf{x}) \,\| \le \|\, \mathbf{x} - \mathbf{u} \,\|$ für alle $\mathbf{u} \in U$ und $\mathbf{x} \in \mathbb{R}^n$.

Zum Beweis:

1) Ist $\mathbf{x} = \mathbf{u} + \mathbf{v}$, so ist $\mathrm{pr}_U(\mathbf{x}) = \mathbf{u} = \mathbf{u} + \mathbf{0}$, und daher

$$\mathrm{pr}_U \circ \mathrm{pr}_U(\mathbf{x}) = \mathrm{pr}_U(\mathbf{u}) = \mathbf{u} = \mathrm{pr}_U(\mathbf{x}).$$

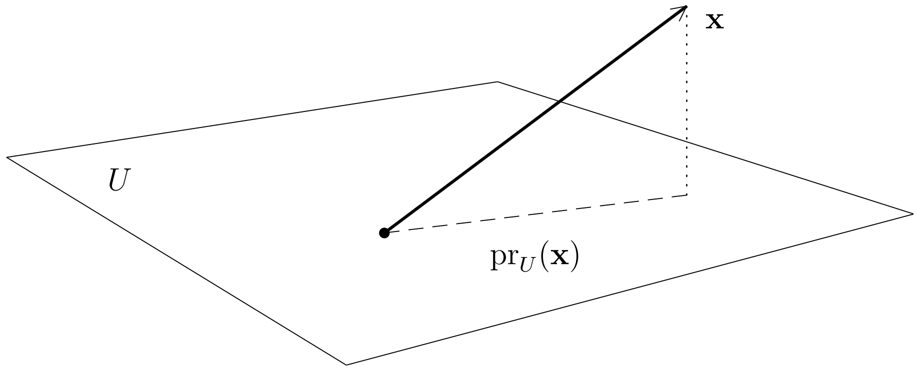
2) Es ist $\mathrm{pr}_U(\mathbf{u} + \mathbf{v}) = \mathbf{0} \iff \mathbf{u} = \mathbf{0} \iff \mathbf{u} + \mathbf{v} = \mathbf{v} \in U^\perp$.

3) ist klar.

4) Sei $\mathbf{x} = \mathbf{u}_0 + \mathbf{v}_0$ mit $\mathbf{u}_0 \in U$ und $\mathbf{v}_0 \in U^\perp$. Dann ist $\mathrm{pr}_U(\mathbf{x}) = \mathbf{u}_0$, und für $\mathbf{u} \in U$ ist $(\mathbf{x} - \mathbf{u}_0) \bullet (\mathbf{u}_0 - \mathbf{u}) = \mathbf{v}_0 \bullet (\mathbf{u}_0 - \mathbf{u}) = 0$. Daraus folgt:

$$\begin{aligned} \|\, \mathbf{x} - \mathbf{u} \,\|^2 &= \|\, (\mathbf{x} - \mathbf{u}_0) + (\mathbf{u}_0 - \mathbf{u}) \,\|^2 \\ &= \|\, \mathbf{x} - \mathbf{u}_0 \,\|^2 + \|\, \mathbf{u}_0 - \mathbf{u} \,\|^2 \\ &\ge \|\, \mathbf{x} - \mathbf{u}_0 \,\|^2. \end{aligned}$$

Die orthogonale Projektion von $\mathbf{x}$ auf U ist derjenige Vektor $\mathbf{u} \in U$, der von $\mathbf{x}$ den kleinsten Abstand hat, der also $\mathbf{x}$ am besten approximiert.



Die Berechnung der orthogonalen Projektion eines Vektors ist ganz einfach:

Wir wählen eine ON-Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ von U und ergänzen sie wie oben zu einer ON-Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ des ganzen $\mathbb{R}^n$. Dann gibt es eine Darstellung

$$\mathbf{x} = \sum_{i=1}^{n} \lambda_i \mathbf{a}_i, \text{ mit } \lambda_i = \mathbf{x} \bullet \mathbf{a}_i.$$

Offensichtlich ist

$$\mathrm{pr}_U(\mathbf{x}) = \sum_{i=1}^{k} (\mathbf{x} \bullet \mathbf{a}_i)\mathbf{a}_i,$$

und dafür braucht man die zusätzlichen Basisvektoren $\mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$ gar nicht zu kennen!

Hier sind zwei Anwendungen der orthogonalen Projektion:

1) Zwei Vektoren $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ im $\mathbb{R}^n$ spannen eine Ebene auf und umschließen dort einen Winkel $\alpha = \angle(\mathbf{x}, \mathbf{y})$. Sei $\mathbf{a} := \dfrac{\mathbf{x}}{\| \mathbf{x} \|}$ und $U := \mathbb{R}\mathbf{a}$ die durch $\mathbf{a}$ aufgespannte Gerade. Aus der anschaulichen Definition der Winkelfunktionen ergibt sich die Gleichung

$$\mathrm{pr}_U(\mathbf{y}) = (\| \mathbf{y} \| \cdot \cos(\alpha)) \cdot \mathbf{a}.$$

Andererseits ist $\mathrm{pr}_U(\mathbf{y}) = (\mathbf{y} \bullet \mathbf{a}) \cdot \mathbf{a}$. Also ist $\mathbf{y} \bullet \mathbf{a} = \| \mathbf{y} \| \cdot \cos(\alpha)$ und

$$\mathbf{x} \bullet \mathbf{y} = (\| \mathbf{x} \| \cdot \mathbf{a}) \bullet \mathbf{y} = \| \mathbf{x} \| \cdot \| \mathbf{y} \| \cdot \cos(\alpha).$$

Damit stimmt der anschauliche Winkelbegriff mit dem in Kapitel I, §6, abstrakt definierten Winkel überein:

$$\alpha = \arccos \frac{\mathbf{x} \bullet \mathbf{y}}{\| \mathbf{x} \| \cdot \| \mathbf{y} \|}.$$

2) Eine weitere Anwendung betrifft die näherungsweise Lösung von linearen Gleichungssystemen. Ist $f = f_A : \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung, so ist die Gleichung

$$f_A(\mathbf{x}) = \mathbf{b}$$

nur dann lösbar, wenn $\mathbf{b} \in U := \mathrm{Im}\,(f_A)$ ist.

Ist diese Bedingung nicht erfüllt, etwa weil die Matrix $A \in M_{n,m}(\mathbb{R})$ experimentell bestimmt und daher nicht exakt bekannt ist, so interessiert man sich wenigstens für eine Näherungslösung.

Bei einer exakten Lösung $\mathbf{x}_0$ der Original-Gleichung wäre $f_A(\mathbf{x}_0) = \mathbf{b}$ und daher $\| f_A(\mathbf{x}_0) - \mathbf{b} \| = 0$. Ist das nicht erreichbar, so sucht man unter allen möglichen $\mathbf{x}$ ein $\mathbf{x}_0$, so daß $\| f_A(\mathbf{x}_0) - \mathbf{b} \|$ wenigstens minimal wird. Und das ist genau dann der Fall, wenn $f_A(\mathbf{x}_0) = \mathrm{pr}_U(\mathbf{b})$ ist.

Wenden wir pr_U auf $f_A(\mathbf{x}_0) - \mathbf{b}$ an, so bleibt $f_A(\mathbf{x}_0)$ als Element von U unverändert. Also gilt:

$$\mathrm{pr}_U(f_A(\mathbf{x}_0) - \mathbf{b}) = \mathrm{pr}_U(\mathbf{b}) - \mathrm{pr}_U(\mathbf{b}) = \mathbf{0}, \text{ d.h. } f_A(\mathbf{x}_0) - \mathbf{b} \in \mathrm{Ker}\,(\mathrm{pr}_U) = U^\perp.$$

Wir betrachten nun auch die Abbildung $f_{A^\top} : \mathbb{R}^n \to \mathbb{R}^m$.

Behauptung: $(\mathrm{Im}\,(f_A))^\perp = \mathrm{Ker}\,(f_{A^\top})$.

BEWEIS: Es gilt

$$\begin{aligned}
\mathbf{x} \in (\mathrm{Im}\,(f_A))^\perp &\iff \mathbf{x} \circ (A \circ \mathbf{y}^\top) = 0 \text{ für alle } \mathbf{y} \\
&\iff (\mathbf{x} \circ A) \circ \mathbf{y}^\top = 0 \text{ für alle } \mathbf{y} \\
&\iff \mathbf{x} \circ A = \mathbf{0} \text{ (weil das Skalarprodukt nicht entartet ist)} \\
&\iff A^\top \circ \mathbf{x}^\top = \mathbf{0}^\top \\
&\iff \mathbf{x} \in \mathrm{Ker}\,(f_{A^\top}).
\end{aligned}$$

Fassen wir das mit dem obigen Ergebnis zusammen, so erhalten wir:

Ist $\mathbf{x}_0$ eine Näherungslösung der Gleichung $A \circ \mathbf{x}^\top = \mathbf{b}^\top$, so ist $A^\top \circ (A \circ \mathbf{x}_0^\top - \mathbf{b}^\top) = \mathbf{0}^\top$, also $\mathbf{x}_0$ eine exakte Lösung der sogenannten „Normal-Gleichung“

$$(A^\top \circ A) \circ \mathbf{x}^\top = A^\top \circ \mathbf{b}.$$

Man überzeugt sich leicht davon, daß die obigen Schlüsse auch umkehrbar sind: erfüllt $\mathbf{x}_0$ die Normalgleichung, so ist $\mathbf{x}_0$ eine Näherungslösung der ursprünglichen Gleichung.

Eine besondere Situation liegt vor, wenn f_A injektiv ist, also $\mathrm{rg}(A) = m \le n$.

Behauptung: Ist $\mathrm{rg}(A) = m$, so ist $A^\top \circ A$ regulär.

BEWEIS: Ist $(A^\top \circ A) \circ \mathbf{x}^\top = \mathbf{0}$, so liegt $A \circ \mathbf{x}$ in $\mathrm{Ker}\,(f_{A^\top}) = (\mathrm{Im}\,(f_A))^\perp$. Also gehört $A \circ \mathbf{x}$ zu $\mathrm{Im}\,(f_A) \cap (\mathrm{Im}\,(f_A))^\perp = \{\mathbf{0}\}$, d.h. $A \circ \mathbf{x}^\top = \mathbf{0}^\top$. Wegen der Injektivität von f_A ist dann $\mathbf{x} = \mathbf{0}$. Da somit die Gleichung $(A^\top \circ A) \circ \mathbf{x}^\top = \mathbf{0}$ eindeutig lösbar ist, muß $A^\top \circ A$ regulär sein! $\square$

In diesem Fall ist die Normalgleichung sogar eindeutig lösbar.

§3 Multilinearformen

Definition:

$V_1, \ldots, V_q$ seien K-Vektorräume. Eine Abbildung

$$\varphi : V_1 \times \ldots \times V_q \to K$$

heißt eine *(q-fache) Multilinearform*, falls für $i = 1, \ldots, q$ gilt:

1. $\varphi(x_1, \ldots, x_i + x_i', \ldots, x_n) = \varphi(x_1, \ldots, x_i, \ldots, x_n) + \varphi(x_1, \ldots, x_i', \ldots, x_n)$.
2. $\varphi(x_1, \ldots, \alpha \cdot x_i, \ldots, x_n) = \alpha \cdot \varphi(x_1, \ldots, x_i, \ldots, x_n)$.

Die Menge aller q-fachen Multilinearformen wird mit $L_q(V_1, \ldots, V_q; K)$ bezeichnet.

φ ist also genau dann multilinear, wenn φ in jedem einzelnen Argument linear ist. Das verallgemeinert den Begriff der Bilinearform (vgl. §1).

Man rechnet leicht nach, daß $L_q(V_1, \ldots, V_q; K)$ auf die übliche Weise zu einem K-Vektorraum gemacht werden kann:

$$\begin{array}{rrcl} & (\varphi + \psi)(x_1, \ldots, x_q) & := & \varphi(x_1, \ldots, x_q) + \psi(x_1, \ldots, x_q) \\ \text{und} & (\alpha \cdot \varphi)(x_1, \ldots, x_q) & := & \alpha \cdot \varphi(x_1, \ldots, x_q). \end{array}$$

Definition:

$$L_0(V; K) := K \quad \text{und} \quad L_q(V; K) := L_q(\underbrace{V, \ldots, V}_{q\text{-mal}}; K) \quad \text{für } q \in \mathbb{N}.$$

Die Elemente von $L_q(V; K)$ nennt man Multilinearformen auf V.

Wir untersuchen nun verschiedene Fälle.

1. Fall: q=1.

$V^* := L_1(V; K)$ ist der Raum der Linearformen auf V. Wir kennen diesen Raum schon, wollen ihn aber noch ein wenig genauer erforschen.

Beispiel :

Sei $V := \mathcal{C}^0(I)$ der Raum der stetigen Funktionen auf dem abgeschlossenen Intervall $I := [a, b]$. V ist nicht endlich-dimensional! Schon die unendlich vielen Funktionen $f_n(x) := x^n$ sind linear unabhängig.

Ist $p : I \to \mathbb{R}$ eine beliebig oft differenzierbare Funktion, so setzen wir

$$\lambda_p[f] := \int_a^b p(t)f(t)\,dt, \qquad \text{für } f \in V.$$

Wegen der Linearität des Integrals ist $\lambda_p : V \to \mathbb{R}$ eine Linearform auf V.

Sei nun $f \in V$, $f \neq 0$. Das bedeutet, daß es ein $t_0 \in I$ mit $f(t_0) \neq 0$ gibt. Aber da f stetig ist, gibt es ein $\varepsilon > 0$, so daß $f(t) \neq 0$ für $t \in I$ und $|\,t - t_0\,| < \varepsilon$ ist. Wir können annehmen, daß $t_0 \in (a, b)$ und auch noch $U_\varepsilon(t_0) \subset (a, b)$ ist. Außerdem nehmen wir an, daß $f(t) > c > 0$ für $|\,t - t_0\,| < \varepsilon$ ist (der Fall $f < 0$ wird analog behandelt).

Man kann jetzt ein $\varepsilon_1 > 0$ mit $0 < \varepsilon_1 < \varepsilon$ und eine $\mathcal{C}^\infty$-Funktion p auf $\mathbb{R}$ finden, so daß gilt:

1. $p(t) \geq 0$ für alle $t \in \mathbb{R}$,
2. $p(t) = 1$ für $|\,t - t_0\,| < \varepsilon_1$.
3. $p(t) = 0$ für $|\,t - t_0\,| \geq \varepsilon$.

Dann gilt:

$$\lambda_p[f] = \int_{t_0-\varepsilon}^{t_0+\varepsilon} p(t)f(t)\,dt \geq \int_{t_0-\varepsilon_1}^{t_0+\varepsilon_1} p(t)f(t)\,dt \geq 2c\varepsilon_1 > 0.$$

Wir können also zu jedem $f \in V$ mit $f \neq 0$ eine Linearform λ_p mit $\lambda_p[f] \neq 0$ finden.

In Wirklichkeit gibt es noch viel mehr Linearformen als nur die von der Gestalt λ_p. Diejenigen Linearformen, die in einem hier nicht näher zu erörterndem Sinne stetig sind, bezeichnet man als *Distributionen.* Ein prominentes Beispiel ist die sogenannte *Dirac'sche δ-Distribution* :

$$\delta_{t_0}[f] := f(t_0).$$

Man kann zeigen: Wäre δ_{t_0} von der Gestalt λ_p, so müßte $p(t_0) = \infty$ und $p(t) = 0$ für $t \neq 0$ sein. Das ist natürlich absurd, aber trotzdem wird manchmal so getan, als gäbe es eine solche Funktion p. Sie wird dann als „Delta-Funktion“ bezeichnet.

Das folgende Ergebnis erscheint nun plausibel:

Existenzsatz für Linearformen

Sei V ein beliebiger K-Vektorraum, $v \in V$ und $v \neq 0$. Dann gibt es eine Linearform $\lambda \in V^*$ mit $\lambda(v) = 1$.

Den Beweis können wir nur andeuten. Als K-Vektorraum besitzt V eine Basis $(a_\iota)_{\iota \in I}$. Man kann es – wegen der Austauschsätze – so einrichten, daß v einer der Basisvektoren ist. Da eine Linearform auf V schon festgelegt ist, wenn man ihre Werte auf den Elementen einer Basis kennt, kann man das gesuchte λ definieren durch

$$\lambda(a_\iota) := \begin{cases} 1 & \text{falls } a_\iota = v, \\ 0 & \text{sonst.} \end{cases}$$

Dieses λ tut's! $\square$

Im Folgenden beschäftigen wir uns vorwiegend mit endlich-dimensionalen Vektorräumen, und wir beginnen mit dem einfachsten Beispiel, dem K^n.

Jedem Vektor $\mathbf{m} \in K^n$ ist eine Linearform $\lambda_{\mathbf{m}}$ zugeordnet, mit

$$\lambda_{\mathbf{m}}(\mathbf{x}) := b(\mathbf{m}, \mathbf{x}) = m_1 x_1 + \cdots m_n x_n = \mathbf{m} \circ \mathbf{x}^\top .$$

Dabei bezeichnet $b : K^n \times K^n \to K$ die kanonische Bilinearform.

Die Zuordnung $\mathbf{m} \mapsto \lambda_{\mathbf{m}}$ definiert eine lineare Abbildung von K^n auf $(K^n)^* = L_1(K^n; K)$, mit $\lambda_{\mathbf{m}}(\mathbf{e}_i) = m_i$. Ist umgekehrt $\lambda \in (K^n)^*$ gegeben, so setzen wir $\mathbf{m} := (\lambda(\mathbf{e}_1), \ldots, \lambda(\mathbf{e}_n))$. Das liefert uns eine lineare Abbildung von $(K^n)^*$ nach K^n, und wegen

$$x_1 \cdot \lambda(\mathbf{e}_1) + \cdots + x_n \lambda(\mathbf{e}_n) = \lambda(x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n) = \lambda(\mathbf{x})$$

sind die beiden Abbildungen invers zueinander.

Nachdem es nun also einen Isomorphismus von K^n nach $(K^n)^*$ gibt, wissen wir, daß $\dim(K^n)^* = n$ ist, und die Linearformen $\varepsilon^i := \lambda_{\mathbf{e}_i}$ bilden eine Basis von $(K^n)^*$. Es ist

$$\varepsilon^i(x_1, \ldots, x_n) = b(\mathbf{e}_i, \mathbf{x}) = \mathbf{e}_i \circ \mathbf{x}^\top = x_i,$$

d.h., ε^i ist einfach die Projektion auf die i-te Komponente. Speziell ist

$$\boxed{\varepsilon^i(\mathbf{e}_j) = \delta_{ij} \quad \text{(Kronecker-Symbol).}}$$

Leider läßt sich die Zuordnung $\mathbf{m} \mapsto \lambda_{\mathbf{m}}$ nicht so ohne weiteres auf beliebige (endlich-dimensionale) Vektorräume übertragen. Deshalb versuchen wir es mit der Umkehrabbildung:

Definition:

Sei V ein K-Vektorraum, $A = \{a_1, \ldots, a_n\}$ eine Basis von V. Dann wird $\varphi^A : V^* \to K^n$ definiert durch

$$\varphi^A(\lambda) := (\lambda(a_1), \ldots, \lambda(a_n)).$$

φ^A hängt von der gewählten Basis A ab, aber immerhin gilt:

Behauptung: $\varphi^A : V^* \to K^n$ ist ein Isomorphismus.

BEWEIS: Den Nachweis der Linearität überlasse ich dem Leser.

Injektivität: Sei $\varphi^A(\lambda) = \mathbf{0}$. Dann ist $\lambda(a_i) = 0$ für alle i, und da λ durch die Werte auf den Elementen der Basis A festgelegt ist, muß $\lambda = 0$ sein.

Die Surjektivität ist etwas schwieriger zu zeigen. Sei $\mathbf{m} \in K^n$ vorgegeben. Dann ist die zugehörige Linearform in $(K^n)^*$ gegeben durch $\mathbf{x} \mapsto \mathbf{m} \circ \mathbf{x}^\top$. Als zugehörige Linearform in V^* wählen wir versuchsweise stattdessen

$$\lambda(x) := \mathbf{m} \circ \Phi_A(x)^\top ,$$

wobei $\Phi_A : V \to K^n$ das Koordinatensystem zur Basis A ist, mit $\Phi_A(x_1a_1+\cdots+x_na_n) := (x_1,\ldots,x_n)$. Tatsächlich ist dann

$$\varphi^A(\lambda) = (\lambda(a_1),\ldots,\lambda(a_n)) = (\mathbf{m}\circ\Phi_A(a_1)^\top,\ldots,\mathbf{m}\circ\Phi_A(a_n)^\top) = (\mathbf{m}\circ\mathbf{e}_1^\top,\ldots,\mathbf{m}\circ\mathbf{e}_n^\top) = \mathbf{m}.$$

Jeder Vektor aus K^n kommt als Wert vor. □

Definition:

Ist V ein K-Vektorraum, so heißt der Raum V^* der Linearformen auf V auch der *Dualraum* von V.

Die Dimension des Dualraumes

Ist $\dim_K(V) = n$, *so ist auch* $\dim_K(V^*) = n$.

Den Beweis haben wir oben gerade geführt. Wir können nun auch eine Basis von V^* angeben, nämlich

$$\alpha^i := (\varphi^A)^{-1}(\mathbf{e}_i), \text{ für } i = 1,\ldots,n.$$

Es folgt:

$$(\alpha^i(a_1),\ldots,\alpha^i(a_n)) = \varphi^A(\alpha^i) = \mathbf{e}_i,$$

also

$$\boxed{\alpha^i(a_j) = \delta_{ij} = \begin{cases} 0 & \text{falls } i = j, \\ 1 & \text{sonst.} \end{cases}}$$

Definition:

Sei V ein K-Vektorraum, $A = \{a_1,\ldots,a_n\}$ eine Basis von V. Die durch $\alpha^i(a_j) = \delta_{ij}$ eindeutig festgelegte Basis $A^* := \{\alpha^1,\ldots,\alpha^n\}$ von V^* nennt man *die zu A duale Basis.*

Bemerkungen :

1. Die Vorgabe der Basis A^* liefert ein Koordinatensystem $\Phi_{A^*} : V^* \to K^n$. Es ist

$$\Phi_{A^*}(x_1\alpha^1 + \cdots + x_n\alpha^n) = (x_1,\ldots,x_n), \qquad \text{also}$$

$\Phi_{A^*}(\alpha^i) = \mathbf{e}_i = \varphi^A(\alpha^i)$. Die Abbildung φ^A ist nichts anderes als das Koordinatensystem Φ_{A^*}.

2. Die Linearformen $\varepsilon^i \in (K^n)^*$ mit $\varepsilon^i(x_1,\ldots,x_n) = x_i$ bilden die duale Basis zur Standardbasis $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ des K^n.

Beispiel :

Wir betrachten im Vektorraum $V := \mathbb{R}^3$ die Basis $A = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ mit

$$\mathbf{a}_1 := (1,1,1),\ \mathbf{a}_2 := (0,2,3) \text{ und } \mathbf{a}_3 := (0,2,-1).$$

Daß es sich tatsächlich um eine Basis handelt, rechnet man leicht nach.

Die zu A duale Basis $A^* = \{\alpha^1, \alpha^2, \alpha^3\}$ wird gegeben durch die Kronecker-Beziehungen $\alpha^i(\mathbf{a}_j) = \delta_{ij}$. Um dieses Gleichungssystem lösen zu können, müssen wir die Elemente von A bzw. A^* als Linearkombinationen der Standard-Basiselemente schreiben:

$$\mathbf{a}_j = \sum_{k=1}^{3} a_{kj}\mathbf{e}_k \quad \text{und} \quad \alpha^i = \sum_{l=1}^{3} a^{il}\varepsilon^l.$$

Die Koeffizienten a_{kj} sind bekannt, die a^{il} suchen wir. Nun gilt:

$$\begin{aligned}
\delta_{ij} &= \alpha^i(\mathbf{a}_j) \\
&= \sum_{l=1}^{3} a^{il}\varepsilon^l(\sum_{k=1}^{3} a_{kj}\mathbf{e}_k) \\
&= \sum_{l,k} a^{il}a_{kj}\varepsilon^l(\mathbf{e}_k) \\
&= \sum_{l,k} a^{il}a_{kj}\delta_{lk} \\
&= \sum_{k=1}^{3} a^{ik}a_{kj}.
\end{aligned}$$

Für die Matrizen

$$\mathbf{A} := \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (\vec{a}_1, \vec{a}_2, \vec{a}_3) \quad \text{und} \quad \mathbf{A}^* := \begin{pmatrix} a^{11} & a^{12} & a^{13} \\ a^{21} & a^{22} & a^{23} \\ a^{31} & a^{32} & a^{33} \end{pmatrix}$$

gilt daher: $\mathbf{A}^* \circ \mathbf{A} = E_n$, also

$$\mathbf{A}^* = (\vec{a}_1, \vec{a}_2, \vec{a}_3)^{-1} = \frac{1}{8}\begin{pmatrix} 8 & 0 & 0 \\ -3 & 1 & 2 \\ -1 & 3 & -2 \end{pmatrix}.$$

Die Koeffizienten a^{il} der Linearform α^i finden sich in der i-ten Zeile von $\mathbf{A}^*$. Daraus ergibt sich:

$$\begin{aligned}
\alpha^1 &= \varepsilon^1, \\
\alpha^2 &= -\frac{3}{8}\varepsilon^1 + \frac{1}{8}\varepsilon^2 + \frac{1}{4}\varepsilon^3 \\
\text{und} \quad \alpha^3 &= -\frac{1}{8}\varepsilon^1 + \frac{3}{8}\varepsilon^2 - \frac{1}{4}\varepsilon^3.
\end{aligned}$$

Das bedeutet:

$$\alpha^1(\mathbf{x}) = x_1, \quad \alpha^2(\mathbf{x}) = -\frac{3}{8}x_1 + \frac{1}{8}x_2 + \frac{1}{4}x_3 \text{ und } \alpha^3(\mathbf{x}) = -\frac{1}{8}x_1 + \frac{3}{8}x_2 - \frac{1}{4}x_3.$$

Wir machen noch eine Probe. Tatsächlich ist

$$\alpha^2(\mathbf{a}_3) = -\frac{3}{8} \cdot 0 + \frac{1}{8} \cdot 2 + \frac{1}{4} \cdot (-1) = 0.$$

Jeder Vektorraum besitzt einen Dualraum, also auch der Raum V^* der Linearformen auf V. Seinen Dualraum V^{**} nennt man auch den *Bidualraum.*

Die Einbettung in den Bidualraum

Sei V ein beliebiger K-Vektorraum, $j : V \to V^{**}$ definiert durch

$$j(v)(\lambda) := \lambda(v).$$

Dann gilt:

1. *j ist linear und injektiv.*
2. *Ist V endlich-dimensional, so ist j ein Isomorphismus.*

BEWEIS: Zunächst einmal gilt es, die Bedeutung des Satzes zu verstehen.

Ein Element von V^{**} ist eine Linearform auf V^*, also eine lineare Abbildung $\Lambda : V^* \to K$. Wir wollen einen Vektor $v \in V$ als eine solche Linearform auffassen. Dazu benutzen wir die bilineare Abbildung

$$b : V \times V^* \to K \quad \text{mit } b(v, \lambda) := \lambda(v)$$

Halten wir λ fest, so ergibt sich eine Linearform in v, halten wir aber den Vektor v fest, so ergibt sich die Linearform $j(v)$ in λ. Wir sprechen auch von der Dualität zwischen V und V^*. Daß $j : V \to V^{**}$ tatsächlich linear ist, rechnet man leicht nach.

Für die Injektivität müssen wir nur zeigen, daß $\text{Ker}\,(j) = \{0\}$ ist. Ist aber $j(v) = 0$, so ist $\lambda(v) = 0$ für alle Linearformen $\lambda \in V^*$. Wegen des Existenzsatzes für Linearformen geht das nur, wenn v selbst schon der Nullvektor ist.

Ist schließlich V n-dimensional, so ist auch V^* und V^{**} n-dimensional. Aber eine injektive Abbildung zwischen zwei n-dimensionalen Vektorräumen ist automatisch ein Isomorphismus. □

Die Abbildung $j : V \to V^{**}$ konnte ohne Benutzung einer Basis definiert werden. Man spricht dann auch von einer *kanonischen* Abbildung. Und da j in jedem Falle injektiv ist, also einen Isomorphismus von V auf $j(V) \subset V^{**}$ definiert, nennt man j eine *Einbettung.* Zusammengefaßt ist $j : V \to V^{**}$ die *kanonische Einbettung des Vektorraumes V in seinen Bidualraum.* Im Falle endlich-dimensionaler Räume spricht man von einem kanonischen Isomorphismus.

2. Fall: $q = 2.$

Beispiele von Bilinearformen kennen wir schon:

- Die kanonische Bilinearform $b : K^n \times K^n \to K$ mit $b(\mathbf{x}, \mathbf{y}) := \mathbf{x} \circ \mathbf{y}^\top$.
- Die Dualität zwischen V und V^*, mit $(v, \lambda) \mapsto \lambda(v)$.
- Das (kanonische) euklidische Skalarprodukt auf dem $\mathbb{R}^n$, mit $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \bullet \mathbf{y}$.

Definition:

Sei V ein $\mathbb{R}$-Vektorraum. Eine symmetrische Bilinearform $s : V \times V \to \mathbb{R}$ heißt *positiv definit*, falls $s(v, v) > 0$ für alle $v \neq 0$ ist.

Ein *(euklidisches) Skalarprodukt* auf V ist eine positiv definite symmetrische Bilinearform auf V.

Dieser Begriff wurde schon in Kapitel I eingeführt. Wir wollen nun auch für komplexe Vektorräume Skalarprodukte einführen.

Definition:

Sei V ein $\mathbb{C}$-Vektorraum. Eine *Sesquilinearform* (d.h. $1\frac{1}{2}$-fache Multilinearform) auf V ist eine Abbildung $s : V \times V \to \mathbb{C}$, für die gilt:

1. $s(v + v', w) = s(v, w) + s(v', w)$ und $s(v, w + w') = s(v, w) + s(v', w)$
2. $s(c \cdot v, w) = c \cdot s(v, w)$ und $s(v, c \cdot w) = \overline{c} \cdot s(v, w)$ für $c \in \mathbb{C}$.

Eine *hermitesche Form auf* V ist eine Sesquilinearform s auf V, für die zusätzlich gilt:

$$s(w, v) = \overline{s(v, w)}.$$

Eine hermitesche Form ist also keine Bilinearform!

Ist $s : V \times V \to \mathbb{C}$ eine hermitesche Form, so gilt für $v \in V$:

$$s(v, v) \in \mathbb{C} \quad \text{und} \quad \overline{s(v, v)} = s(v, v).$$

Das bedeutet, daß $s(v, v)$ stets reell ist. Ist sogar $s(v, v) > 0$ für alle $v \neq 0$, so heißt s positiv-definit.

Definition:

Sei V ein $\mathbb{C}$-Vektorraum. Ein *hermitesches Skalarprodukt* auf V ist eine positiv definite hermitesche Form auf V.

Beispiel :

Der einfachste komplexe Vektorraum ist der $\mathbb{C}^n$. Das kanonische hermitesche Skalarprodukt auf dem $\mathbb{C}^n$ wird definiert durch

$$<\mathbf{z}, \mathbf{w}> := z_1 \overline{w}_1 + \cdots + z_n \overline{w}_n.$$

Ist $\mathbf{z} = (z_1, \ldots, z_n)$ und $z_i = x_i + \mathbf{j}y_i$ für $i = 1, \ldots, n$, so ist

$$<\mathbf{z}, \mathbf{z}> = z_1\overline{z}_1 + \cdots + z_n\overline{z}_n = (x_1)^2 + (y_1)^2 + \cdots + (x_n)^2 + (y_n)^2.$$

Die *Norm* $\| \mathbf{z} \| := \sqrt{<\mathbf{z}, \mathbf{z}>}$ eines Vektors $\mathbf{z} \in \mathbb{C}^n$ stimmt also mit der euklidischen Norm des Vektors überein, wenn man den $\mathbb{C}^n$ mit dem $\mathbb{R}^{2n}$ identifiziert.

Vieles überträgt sich sinngemäß aus dem Reellen, so auch der Begriff der Orthonormalbasis oder die Orthogonalisierung.

Bilinearformen auf dem K^n kann man auch mit Matrizen beschreiben:

Die Matrix zu einer Bilinearform

Sei $b : K^n \times K^n \to K$ eine Bilinearform, $b_{ij} := b(\mathbf{e}_i, \mathbf{e}_j)$ und $B \in M_{n,n}(K)$ die Matrix mit den Einträgen b_{ij}. Dann gilt:

$$b(\mathbf{x}, \mathbf{y}) = \mathbf{x} \circ B \circ \mathbf{y}^\top.$$

B ist durch diese Beziehung eindeutig bestimmt.

Beweis:

$$\begin{aligned} b(\mathbf{x}, \mathbf{y}) &= b(\sum_{i=1}^n x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j b(\mathbf{e}_i, \mathbf{e}_j) \\ &= \sum_{i,j} x_i y_j b_{ij} = \mathbf{x} \circ B \circ \mathbf{y}^\top. \end{aligned}$$

□

Analog kann man eine Sesquilinearform auf dem $\mathbb{C}^n$ beschreiben durch

$$s(\mathbf{z}, \mathbf{w}) = \mathbf{z} \circ S \circ \overline{\mathbf{w}}^\top,$$

wobei die Matrix $S = (s_{ij})$ wieder durch $s_{ij} := s(\mathbf{e}_i, \mathbf{e}_j)$ gegeben ist.

Eigenschaften von Bilinear- und Sesquilinearformen kann man schon an den Matrizen ablesen.

Euklidische Skalarprodukte und Matrizen

Sei $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ gegeben durch $b(\mathbf{x}, \mathbf{y}) = \mathbf{x} \circ B \circ \mathbf{y}^\top$. Dann ist b bilinear, und es gilt:

$$\begin{aligned} b \text{ symmetrisch} &\iff B^\top = B, \\ b \text{ positiv definit} &\iff \mathbf{x} \circ B \circ \mathbf{x}^\top > 0 \text{ für } \mathbf{x} \neq \mathbf{0}. \end{aligned}$$

Hermitesche Skalarprodukte und Matrizen

Sei $s : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ gegeben durch $s(\mathbf{z},\mathbf{w}) = \mathbf{z} \circ S \circ \overline{\mathbf{w}}^\top$. Dann ist s sesquilinear, und es gilt:

$$\begin{aligned} s \text{ hermitesch} &\iff \overline{S}^\top = S, \\ s \text{ positiv definit} &\iff \mathbf{z} \circ S \circ \overline{\mathbf{z}}^\top > 0 \text{ für } \mathbf{z} \neq \mathbf{0}. \end{aligned}$$

Man spricht dann sinngemäß auch von symmetrischen, hermiteschen und positiv definiten Matrizen.

Wir wollen nun Koordinaten für den Vektorraum $L_2(V;K)$ einführen.

Am einfachsten geht das im Falle $V = K^n$. Wir können dann jeder Matrix $B \in M_{n,n}(K)$ eine Bilinearform $\varphi_B : K^n \times K^n \to K$ zuordnen, mit

$$\varphi_B(\mathbf{x},\mathbf{y}) := \mathbf{x} \circ B \circ \mathbf{y}^\top.$$

Ist umgekehrt eine Bilinearform φ gegeben, so gewinnt man die Matrix B durch

$$B = \left(b_{ij} \;\middle|\; \begin{matrix} i = 1,\dots,n \\ j = 1,\dots,n \end{matrix} \right), \qquad \text{mit } b_{ij} := \varphi(\mathbf{e}_i,\mathbf{e}_j).$$

Also ist die Zuordnung $M_{n,n}(K) \to L_2(K^n;K)$ mit $B \mapsto \varphi_B$ ein Isomorphismus, und es ist $\dim L_2(K^n;K) = n^2$.

Eine Basis von $M_{n,n}(K)$ bilden die Matrizen

$$E_{ij} := \left(\delta_{i\nu} \cdot \delta_{j\mu} \;\middle|\; \begin{matrix} \nu = 1,\dots,n \\ \mu = 1,\dots,n \end{matrix} \right) = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix},$$

mit genau einer Eins am Schnittpunkt der i-ten Zeile mit der j-ten Spalte und lauter Nullen sonst.

Die Bilinearformen $\varepsilon^{ij} := \varphi_{E_{ij}}$ müssen dann eine Basis von $L_2(K^n;K)$ bilden. Dabei ist

$$\varepsilon^{ij}(\mathbf{x},\mathbf{y}) = \mathbf{x} \circ E_{ij} \circ \mathbf{y}^\top = x_i y_j, \text{ für } i,j = 1,\dots,n.$$

Definition:

Sind $f, g \in V^*$ zwei Linearformen auf dem Vektorraum V, so wird deren *Tensorprodukt* $f \otimes g \in L_2(V;K)$ definiert durch

$$(f \otimes g)(v,w) := f(v) \cdot g(w).$$

Man rechnet sofort nach, daß $f \otimes g$ tatsächlich eine Bilinearform ist. Und speziell im Falle $V = K^n$ ist $\varepsilon^{ij} = \varepsilon^i \otimes \varepsilon^j$.

Rechenregeln für das Tensorprodukt

Die durch $(f, g) \mapsto f \otimes g$ gegebene Abbildung $V^* \times V^* \to L_2(V; K)$ ist bilinear, d.h. es gilt:

$$\begin{aligned} (f_1 + f_2) \otimes g &= f_1 \otimes g + f_2 \otimes g, \\ f \otimes (g_1 + g_2) &= f \otimes g_1 + f \otimes g_2 \\ \textit{und} \quad \alpha \cdot (f \otimes g) &= (\alpha \cdot f) \otimes g \,=\, f \otimes (\alpha \cdot g), \qquad \textit{für } \alpha \in K. \end{aligned}$$

Beispiel:

Sei $V = \mathbb{R}^2$, und

$$\begin{aligned} \varphi &:= (2\varepsilon^1 + 5\varepsilon^2) \otimes (\varepsilon^1 - \varepsilon^2) \\ &= 2\varepsilon^1 \otimes \varepsilon^1 - 2\varepsilon^1 \otimes \varepsilon^2 + 5\varepsilon^2 \otimes \varepsilon^1 - 5\varepsilon^2 \otimes \varepsilon^2. \end{aligned}$$

Eine weitere Vereinfachung ist nicht möglich.

Ist nun $\mathbf{x} = (1, 1)$ und $\mathbf{y} = (2, -3)$, so ist

$$\begin{aligned} \varphi(\mathbf{x}, \mathbf{y}) &= 2\varepsilon^1(\mathbf{x})\varepsilon^1(\mathbf{y}) - 2\varepsilon^1(\mathbf{x})\varepsilon^2(\mathbf{y}) + 5\varepsilon^2(\mathbf{x})\varepsilon^1(\mathbf{y}) - 5\varepsilon^2(\mathbf{x})\varepsilon^2(\mathbf{y}) \\ &= 2 \cdot 1 \cdot 2 - 2 \cdot 1 \cdot (-3) + 5 \cdot 1 \cdot 2 - 5 \cdot 1 \cdot (-3) \\ &= 4 + 6 + 10 + 15 \,=\, 35. \end{aligned}$$

Tatsächlich ist auch

$$(2\varepsilon^1 + 5\varepsilon^2)(\mathbf{x}) \cdot (\varepsilon^1 - \varepsilon^2)(\mathbf{y}) = (2 \cdot 1 + 5 \cdot 1)(2 - (-3)) = 35.$$

Achtung!! I.a. ist $f \otimes g \neq g \otimes f$. Wählt man $\mathbf{x}$ und $\mathbf{y}$ wie im obigen Beispiel, so ist

$$\begin{aligned} \varepsilon^1 \otimes \varepsilon^2(\mathbf{x}, \mathbf{y}) &= 1 \cdot (-3) \,=\, -3 \\ \text{und} \quad \varepsilon^2 \otimes \varepsilon^1(\mathbf{x}, \mathbf{y}) &= 1 \cdot 2 = 2. \end{aligned}$$

Die Tensorprodukte $\varepsilon^i \otimes \varepsilon^j$ bilden eine Basis von $L_2(K^n; K)$. Analog gilt nun:

Eine Basis für $L_2(V; K)$

Sei V ein K-Vektorraum, $A = \{a_1, \ldots, a_n\}$ eine Basis von V und $A^* = \{\alpha^1, \ldots, \alpha^n\}$ die dazu duale Basis von V^*. Dann bilden die Bilinearformen

$$\alpha^{ij} := \alpha^i \otimes \alpha^j, \quad i, j = 1, \ldots, n,$$

eine Basis von $L_2(V; K)$. Insbesondere ist $\dim L_2(V; K) = n^2$.

Wir verzichten hier auf den Beweis, der im Wesentlichen darauf beruht, daß **jede** Bilinearform φ auf V schon durch die Werte von φ auf den Paaren (a_ν, a_μ) festgelegt ist.

Definition:

Sind $f, g \in V^*$ zwei Linearformen auf V, so wird deren *Dachprodukt* (oder *äußeres Produkt*) $f \wedge g \in L_2(V; K)$ definiert durch

$$f \wedge g := f \otimes g - g \otimes f.$$

Auch hier rechnet man leicht nach, daß $f \wedge g$ eine Bilinearform ist. Sie hat aber noch eine zusätzliche Eigenschaft:

Das Dachprodukt zweier Linearformen ist eine alternierende Bilinearform

Es ist $\quad (f \wedge g)(v, w) = -(f \wedge g)(w, v).$

BEWEIS: Es ist

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

□

Die alternierenden Bilinearformen bilden einen Untervektorraum von $L_2(V; K)$, den wir mit $A^2(V)$ bezeichnen. Man nennt die Elemente von $A^2(V)$ auch *(alternierende) 2-Formen.* Es gilt:

1. $(f + f') \wedge g = f \wedge g + f' \wedge g$ und $f \wedge (g + g') = f \wedge g + f \wedge g'$.
2. $\alpha \cdot (f \wedge g) = (\alpha \cdot f) \wedge g = f \wedge (\alpha \cdot g)$ für $c \in K$.
3. $f \wedge g = -g \wedge f$. („Anti-Kommutativgesetz“)

Wir wollen nun versuchen, Tensor- und Dachprodukte von Linearformen auf dem K^n mit Hilfe von Matrizen zu beschreiben.

Sei $f = \lambda_{\mathbf{a}}$ und $g = \lambda_{\mathbf{b}}$. Dann ist

$$\begin{aligned}(f \otimes g)(\mathbf{e}_i, \mathbf{e}_j) &= \left((\sum_{\nu=1}^n a_\nu \varepsilon^\nu) \otimes (\sum_{\mu=1}^n b_\mu \varepsilon^\mu) \right) (\mathbf{e}_i, \mathbf{e}_j) \\ &= (\sum_{\nu,\mu} a_\nu b_\mu \varepsilon^\nu \otimes \varepsilon^\mu)(\mathbf{e}_i, \mathbf{e}_j)\end{aligned}$$

$$\begin{aligned} &= \sum_{\nu,\mu} a_\nu b_\mu \delta_{\nu i}\delta_{\mu j} \\ &= a_i b_j, \end{aligned}$$

Also ist $f \otimes g = \varphi_B$, mit der Matrix

$$B = \mathbf{a}^\top \circ \mathbf{b} = \begin{pmatrix} a_1b_1 & \cdots & a_1b_n \\ \vdots & & \vdots \\ a_nb_1 & \cdots & a_nb_n \end{pmatrix},$$

und $f \wedge g$ ist durch die Matrix $\mathbf{a}^\top \circ \mathbf{b} - \mathbf{b}^\top \circ \mathbf{a}$ gegeben.

Eine beliebige Bilinearform $\varphi = \varphi_B$ auf dem K^n ist genau dann alternierend oder schiefsymmetrisch, wenn $\varphi(\mathbf{x},\mathbf{y}) = -\varphi(\mathbf{y},\mathbf{x})$ für alle $\mathbf{x},\mathbf{y} \in K^n$ gilt, und das ist genau dann der Fall, wenn $B^\top = -B$ ist. Eine solche Matrix nennt man auch *schiefsymmetrisch.* Für die Einträge b_{ij} in B gilt dann:

1. $b_{ii} = 0$ für $i = 1, \ldots, n$, d.h., alle Diagonalelemente verschwinden.
2. $b_{ji} = -b_{ij}$ für $i < j$, d.h., die Elemente unterhalb der Diagonalen erhält man aus denen oberhalb der Diagonalen durch Spiegelung und den Übergang zum Negativen.

Im Falle $n = 3$ hat eine schiefsymmetrische Matrix also die Gestalt

$$B = \begin{pmatrix} 0 & b_{12} & b_{13} \\ -b_{12} & 0 & b_{23} \\ -b_{13} & -b_{23} & 0 \end{pmatrix}.$$

Im Falle $n = 3$ sind das 3 unabhängige Parameter und im allgemeinen Fall $\frac{n^2-n}{2} = \binom{n}{2}$ unabhängige Parameter.

Tatsächlich bilden die $\binom{n}{2}$ Matrizen $S_{ij} := E_{ij} - E_{ji}$ für $1 \le i < j \le n$ eine Basis des Vektorraumes der n-reihigen schiefsymmetrischen Matrizen. Daraus folgt unmittelbar, daß die Bilinearformen $\varphi_{S_{ij}} = \varepsilon^i \wedge \varepsilon^j$ für $1 \le i < j \le n$ eine Basis des Vektorraumes $A^2(K^n)$ bilden.

Mit Hilfe von Koordinaten kann man dieses Ergebnis sofort verallgemeinern:

Eine Basis für $A^2(V)$

Sei wieder $\{a_1, \ldots, a_n\}$ *eine Basis von* V *und* $\{\alpha_1, \ldots, \alpha^n\}$ *die duale Basis von* V^*.

Dann bilden die Dachprodukte $\alpha^i \wedge \alpha^j$ *mit* $1 \le i < j \le n$ *eine Basis des Raumes* $A^2(V)$ *der alternierenden 2-Formen.*

Insbesondere ist $\dim A^2(V) = \dfrac{n(n-1)}{2} = \dbinom{n}{2}$.

Ist z.B. $n = 3$, so besteht die Basis von $A^2(V)$ aus den Elementen $\alpha^1 \wedge \alpha^2$, $\alpha^1 \wedge \alpha^3$ und $\alpha^2 \wedge \alpha^3$.

Das Ergebnis des Satzes ist so plausibel, daß wir auf einen detaillerten Beweis verzichten können.

3. Fall: q beliebig ≥ 2.

Sind $f_1, \ldots, f_q$ Linearformen auf V, so wird deren *Tensorprodukt* $f_1 \otimes \ldots \otimes f_q$ definiert durch

$$(f_1 \otimes \ldots \otimes f_q)(v_1, \ldots, v_q) := f_1(v_1) \cdot \ldots \cdot f_q(v_q).$$

Ist $\{a_1, \ldots, a_n\}$ eine Basis von V und $\{\alpha^1, \ldots, \alpha^n\}$ die dazu duale Basis, so kann man zeigen, daß die Tensorprodukte $\alpha^{i_1} \otimes \ldots \alpha^{i_q}$ mit $1 \leq i_1, \ldots, i_q \leq n$ eine Basis des Raumes $L_q(V; K)$ bilden, so daß $\dim L_q(V; K) = n^q$ ist. Wir wollen uns hier aber nicht weiter mit allgemeiner Tensorrechnung befassen, sondern uns auf „alternierende q-Formen“ beschränken.

Definition:

Eine Multilinearform $\varphi \in L_q(V; K)$ heißt *alternierend* oder *schiefsymmetrisch*, falls für $i = 1, \ldots, q-1$ gilt:

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

Da man beliebige Permutationen aus Vertauschungen zusammensetzen kann, folgt:

Eigenschaften alternierender Multilinearformen

1. $\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(q)}) = \text{sgn}(\sigma) \cdot \varphi(x_1, \ldots, x_q)$ *für alle Permutationen* $\sigma \in S_q$.
2. $\varphi(x_1, \ldots, x_q) = 0$*, falls zwei Argumente gleich sind.*

BEWEIS: Die 1. Aussage ergibt sich daraus, daß das Signum einer Permutation unabhängig davon ist, wie man diese Permutation aus Vertauschungen zusammensetzt.

Sind zwei der Argumente in $\varphi(x_1, \ldots, x_q)$ gleich, so ergibt sich nach Vertauschung dieser Elemente:

$$\varphi(x_1, \ldots, x_q) = -\varphi(x_1, \ldots, x_q),$$

also $2 \cdot \varphi(x_1, \ldots, x_q) = 0$. Ist $2 \neq 0$ (und wir wollen hier nur solche Körper K betrachten, für die das richtig ist), so folgt nun auch: $\varphi(x_1, \ldots, x_q) = 0$. □

Definition:

Der Untervektorraum $A^q(V) \subset L_q(V; K)$ aller alternierenden q-fachen Multilinearformen wird als Raum der alternierenden q-Formen bezeichnet.

Das Dachprodukt zwischen p- und q-Formen

Es gibt für alle p und q eine bilineare Abbildung $A^p(V) \times A^q(V) \to A^{p+q}(V)$ mit $(\varphi, \omega) \mapsto \varphi \wedge \omega$, so daß gilt:

1. **Antikommutativgesetz:** $\omega \wedge \varphi = (-1)^{pq} \varphi \wedge \omega$.
2. **Assoziativgesetz:** $(\omega \wedge \varphi) \wedge \psi = \omega \wedge (\varphi \wedge \psi) =: \omega \wedge \varphi \wedge \psi$.
3. (a) *Sind $\lambda_1, \ldots, \lambda_q$ Linearformen, so ist*

$$\lambda_1 \wedge \ldots \wedge \lambda_q = \sum_{\sigma \in S_q} \operatorname{sgn}(\sigma) \lambda_{\sigma(1)} \otimes \ldots \otimes \lambda_{\sigma(q)}.$$

   (b) *Die Formen $\alpha^{i_1} \wedge \ldots \wedge \alpha^{i_q}$ mit $1 \le i_1 < \ldots < i_q \le n$ bilden eine Basis von $A^q(V)$.*

Den technisch aufwendigen Beweis dieses Satzes lassen wir weg. Es reicht, wenn man mit dem Kalkül umgehen kann. Die Definition des Dachproduktes wird dabei natürlich auch verheimlicht, aber aus (3) kann man sie sich zusammenreimen. Außerdem folgt, daß für $f, g \in A^1(V) = V^*$ gilt:

$$f \wedge g = f \otimes g - g \otimes f.$$

Damit stimmt das neue Dachprodukt im Falle der 1-Formen mit dem alten überein.

Beispiele :

1. Sei $\varphi := 2\varepsilon^1 - 5\varepsilon^2 + 7\varepsilon^3 - 3\varepsilon^4 \in A^1(\mathbb{R}^4) = (\mathbb{R}^4)^*$ und $\omega := 15\varepsilon^1 \wedge \varepsilon^3 - 12\varepsilon^2 \wedge \varepsilon^4 \in A^2(\mathbb{R}^4)$. Dann ist

$$\begin{aligned} \varphi \wedge \omega &= (2\varepsilon^1 - 5\varepsilon^2 + 7\varepsilon^3 - 3\varepsilon^4) \wedge (15\varepsilon^1 \wedge \varepsilon^3 - 12\varepsilon^2 \wedge \varepsilon^4) \\ &= -24\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^4 - 75\varepsilon^2 \wedge \varepsilon^1 \wedge \varepsilon^3 \\ &\quad - 84\varepsilon^3 \wedge \varepsilon^2 \wedge \varepsilon^4 - 45\varepsilon^4 \wedge \varepsilon^1 \wedge \varepsilon^3 \\ &= 75\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3 - 24\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^4 \\ &\quad - 45\varepsilon^1 \wedge \varepsilon^3 \wedge \varepsilon^4 + 84\varepsilon^2 \wedge \varepsilon^3 \wedge \varepsilon^4. \end{aligned}$$

   Also ist

$$\begin{aligned} (\varepsilon^1 - \varepsilon^4) \wedge \varphi \wedge \omega &= 84\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3 \wedge \varepsilon^4 - 75\varepsilon^4 \wedge \varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3 \\ &= 159\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3 \wedge \varepsilon^4. \end{aligned}$$

2. Sei $\varphi := \varepsilon^1 - 2\varepsilon^2 + 3\varepsilon^3 \in A^1(\mathbb{R}^3)$ und $\omega := 4\varepsilon^1 \wedge \varepsilon^2 + 5\varepsilon^2 \wedge \varepsilon^3 - 6\varepsilon^1 \wedge \varepsilon^3 \in A^2(\mathbb{R}^3)$. Man kann sofort mit diesen Formen rechnen, aber vielleicht ist es ja lehrreich, sich erst einmal zu überlegen, was man vor sich hat.

   φ ist eine Linearform, mit $\varphi(x_1, x_2, x_3) = x_1 - 2x_2 + 3x_3$. ω ist eine alternierende Bilinearform und kann daher in der Form $\omega(\mathbf{x}, \mathbf{y}) = \mathbf{x} \circ \Omega \circ \mathbf{y}^\top$ geschrieben werden, wobei gilt:

$$\Omega = \begin{pmatrix} 0 & 4 & -6 \\ -4 & 0 & 5 \\ 6 & -5 & 0 \end{pmatrix}.$$

Nun berechnen wir $\varphi \wedge \omega \in A^3(\mathbb{R}^3)$:

$$\begin{aligned}\varphi \wedge \omega &= (\varepsilon^1 - 2\varepsilon^2 + 3\varepsilon^3) \wedge (4\varepsilon^1 \wedge \varepsilon^2 + 5\varepsilon^2 \wedge \varepsilon^3 - 6\varepsilon^1 \wedge \varepsilon^3)\\ &= 5\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3 + 12\varepsilon^2 \wedge \varepsilon^1 \wedge \varepsilon^3 + 12\varepsilon^3 \wedge \varepsilon^1 \wedge \varepsilon^2\\ &= (5 - 12 + 12)\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3 \;=\; 5\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3.\end{aligned}$$

Das Ergebnis ist eine alternierende 3-Form. Angewandt auf Vektoren $\mathbf{x}$, $\mathbf{y}$ und $\mathbf{z}$ ergibt sie:

$$\begin{aligned}5\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3(\mathbf{x},\mathbf{y},\mathbf{z}) &= 5 \cdot \sum_{\sigma\in S_3} \operatorname{sgn}(\sigma)\varepsilon^{\sigma(1)} \otimes \varepsilon^{\sigma(2)} \otimes \varepsilon^{\sigma(3)}(\mathbf{x},\mathbf{y},\mathbf{z})\\ &= 5 \cdot \sum_{\sigma\in S_3} \operatorname{sgn}(\sigma)x_{\sigma(1)}y_{\sigma(2)}z_{\sigma(3)}\\ &= 5 \cdot (x_1y_2z_3 - x_2y_1z_3 \pm \ldots).\end{aligned}$$

Vertauschungsregel

Sind $\lambda_1, \ldots, \lambda_q$ Linearformen und ist $\tau \in S_q$, so ist

$$\lambda_{\tau(1)} \wedge \ldots \wedge \lambda_{\tau(q)} = \operatorname{sgn}(\tau) \cdot \lambda_1 \wedge \ldots \wedge \lambda_q.$$

BEWEIS: Es gilt: $\lambda_\nu \wedge \lambda_\mu = -\lambda_\mu \wedge \lambda_\nu$. Vertauscht man also in $\lambda_1 \wedge \ldots \wedge \lambda_q$ zwei aufeinander folgende Faktoren, so ändert sich das Vorzeichen. Jede Permutation kann man durch eine Folge solcher Vertauschungen herstellen, und das Signum der Permutation ist $+1$ oder -1, je nachdem, ob man eine gerade oder ungerade Anzahl von Vertauschungen benötigt hat. □

Die Dimension von $A^q(V)$ ist die Anzahl der q-Tupel $(i_1, \ldots, i_q)$ mit $1 \le i_1 < \ldots < i_q \le n$. Jedes solche q-Tupel bestimmt genau eine q-elementige Teilmenge von $\{1, \ldots, n\}$, und zu jeder der Mengen gibt es nur eine zulässige Anordnung der Elemente. Also folgt:

$$\boxed{\text{Ist } \dim(V) = n, \text{ so ist } \dim A^q(V) = \binom{n}{q}.}$$

Ist $q > n$, so ist $A^q(V) = \{0\}$. Das liegt daran, daß die q-Formen φ schon durch ihre Werte $\varphi(a_{i_1}, \ldots, a_{i_q})$ auf den Basiselementen festgelegt sind, daß aber höchstens n Basisvektoren zur Verfügung stehen.

Betrachten wir nun den Fall $q = n = \dim(V)$.

Es gibt zu jeder Basis $A = \{a_1, \ldots, a_n\}$ von V genau eine n-Form ω_A mit $\omega_A(a_1, \ldots, a_n) = 1$, nämlich die n-Form $\omega_A = \alpha^1 \wedge \ldots \wedge \alpha^n$, wobei $A^* = \{\alpha^1, \ldots, \alpha^n\}$ wie üblich die zu A duale Basis bezeichnet.

Wir wollen das überprüfen: Zunächst ist

$$\omega_A(a_1, \ldots, a_n) \;=\; \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\alpha^{\sigma(1)} \otimes \ldots \otimes \alpha^{\sigma(n)}(a_1, \ldots, a_n)$$

$$\begin{aligned} &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \alpha^{\sigma(1)}(a_1) \cdot \ldots \cdot \alpha^{\sigma(n)}(a_n) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \delta_{\sigma(1),1} \cdot \ldots \cdot \delta_{\sigma(n),n} \\ &= \operatorname{sgn}(\mathrm{id}) = 1. \end{aligned}$$

Jede andere alternierende n-Form hat die Gestalt $\omega = c \cdot \omega_A$, mit einem konstanten Faktor c, und dann ist $\omega(a_1, \ldots, a_n) = c$. Damit ist die Behauptung richtig.

Insbesondere gilt:

Existenz der Determinantenform

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

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

Wir nennen $\Delta = \varepsilon^1 \wedge \ldots \wedge \varepsilon^n$ die *Determinantenform* auf dem K^n.

Setzt man noch $A^0(V) := K$, so sind $A^0(V)$ und $A^n(V)$ beide 1-dimensional und daher isomorph. Weiter sind $A^1(V) = V^*$ und $A^{n-1}(V)$ beide n-dimensional und daher ebenfalls isomorph, so wie allgemein $A^q(V) \cong A^{n-q}(V)$ ist. Wir wollen noch den Raum $A^{n-1}(V)$ untersuchen, und dabei beschränken wir uns auf den Fall $V = K^n$.

Eine Basis von $A^{n-1}(K^n)$ bilden die n (n-1)-Formen

$$\omega^i := \varepsilon^1 \wedge \ldots \wedge \widehat{\varepsilon^i} \wedge \ldots \wedge \varepsilon^n, \quad i = 1, \ldots, n,$$

wobei das Dach über ε^i bedeutet, daß dieser Faktor weggelassen werden soll. Man erhält also die Formen

$$\omega^1 = \varepsilon^2 \wedge \ldots \wedge \varepsilon^n, \quad \omega^2 = \varepsilon^1 \wedge \varepsilon^3 \wedge \ldots \wedge \varepsilon^n, \quad \ldots, \quad \omega^n = \varepsilon^1 \wedge \ldots \wedge \varepsilon^{n-1}.$$

Wir kennen schon den Isomorphismus $\mathbf{a} \mapsto \lambda_{\mathbf{a}}$, der jedem Vektor $\mathbf{a} \in K^n$ eine Linearform $\lambda_{\mathbf{a}} \in (K^n)^*$ zuordnet. Die Umkehrabbildung $\varphi^E : (K^n)^* \to K^n$ ist gegeben durch $\varphi^E(\lambda) := (\lambda(\mathbf{e}_1), \ldots, \lambda(\mathbf{e}_n))$. Nun wollen wir auch noch einen Isomorphismus $K^n \to A^{n-1}(K^n)$ definieren.

Die kanonische $(n-1)$-Form zu einem Vektor

Es gibt zu jedem Vektor $\mathbf{a} \in K^n$ genau eine $(n-1)$-Form $\eta_{\mathbf{a}} \in A^{n-1}(K^n)$, so daß gilt:

$$\lambda \wedge \eta_{\mathbf{a}} = \lambda(\mathbf{a}) \cdot \Delta, \qquad \text{für alle } \lambda \in (K^n)^*.$$

In Koordinaten ist

$$\eta_{\mathbf{a}} = \sum_{i=1}^{n} a_i (-1)^{i+1} \varepsilon^1 \wedge \ldots \wedge \widehat{\varepsilon^i} \wedge \ldots \wedge \varepsilon^n.$$

Beweis: Der Eindeutigkeitsbeweis liefert auch gleich die Formel:

Wenn es eine Form $\eta_{\mathbf{a}} = \sum_{j=1}^{n} \eta_j \omega_j$ mit der geforderten Eigenschaft gibt, so muß gelten:

$$\begin{aligned} a_i \cdot \Delta &= \varepsilon^i(\mathbf{a}) \cdot \Delta \\ &= \varepsilon^i \wedge \eta_{\mathbf{a}} \\ &= \sum_{j=1}^{n} \eta_j \varepsilon^i \wedge \varepsilon^1 \wedge \ldots \wedge \widehat{\varepsilon^j} \wedge \ldots \wedge \varepsilon^n \\ &= \eta^i \varepsilon^i \wedge \varepsilon^1 \wedge \ldots \wedge \widehat{\varepsilon^i} \wedge \ldots \wedge \varepsilon^n \\ &= \eta^i \cdot (-1)^{i+1} \cdot \Delta. \end{aligned}$$

Also ist dann $\eta_{\mathbf{a}} = \sum_{i=1}^{n} a_i (-1)^{i+1} \varepsilon^1 \wedge \ldots \wedge \widehat{\varepsilon^i} \wedge \ldots \wedge \varepsilon^n$.

Da jede Linearform λ eine Linearkombination der ε^i ist, hat die so definierte Form $\eta_{\mathbf{a}}$ die gewünschte Eigenschaft. □

Der Stern-Operator

Durch $*(\lambda_{\mathbf{a}}) := \eta_{\mathbf{a}}$ *wird ein Isomorphismus*

$$* : (K^n)^* = A^1(K^n) \to A^{n-1}(K^n)$$

definiert.

Für $\mathbf{a}, \mathbf{b} \in K^n$ *ist dann stets*

$$\lambda_{\mathbf{a}} \wedge *\lambda_{\mathbf{b}} = (\mathbf{a} \circ \mathbf{b}^\top) \cdot \Delta.$$

BEWEIS: Sei $\psi^E : K^n \to A^{n-1}(K^n)$ die durch $\psi^E(\mathbf{a}) := \eta_{\mathbf{a}}$ gegebene Abbildung. Man sieht sehr leicht, daß ψ^E linear und $\operatorname{Ker}(\psi^E) = \{0\}$ ist. Wegen $\dim(A^{n-1}(K^n)) = n = \dim(K^n)$ ist ψ^E tatsächlich ein Isomorphismus, und der Stern-Operator ist gegeben durch

$$* = \psi^E \circ \varphi^E.$$

Weiter gilt:

$$\lambda_{\mathbf{a}} \wedge *\lambda_{\mathbf{b}} = \lambda_{\mathbf{a}} \wedge \eta_{\mathbf{b}} = \lambda_{\mathbf{a}}(\mathbf{b}) \cdot \Delta = (\mathbf{a} \circ \mathbf{b}^\top) \cdot \Delta.$$

□

Beispiel:

Sei $V = \mathbb{R}^3$. Dann ist $n = 3$ und $n - 1 = 2$. Es gibt also nur die Räume

$$A^0(\mathbb{R}^3) = \mathbb{R}, \quad A^1(\mathbb{R}^3) = (\mathbb{R}^3)^*, \quad A^2(\mathbb{R}^3) \cong \mathbb{R}^3 \text{ und } A^3(\mathbb{R}^3) = \mathbb{R} \cdot \Delta.$$

Der Stern-Operator $* : A^1(\mathbb{R}^3) \to A^2(\mathbb{R}^3)$ ist in diesem Falle gegeben durch

$$*(a_1\varepsilon^1 + a_2\varepsilon^2 + a_3\varepsilon^3) = a_1\varepsilon^2 \wedge \varepsilon^3 + a_2\varepsilon^3 \wedge \varepsilon^1 + a_3\varepsilon^1 \wedge \varepsilon^2.$$

Man beachte die zyklischen Vertauschungen der Indizes bei der Definition.

Weiter ist
$$\lambda_{\mathbf{a}} \wedge *\lambda_{\mathbf{b}} = (a_1b_1 + a_2b_2 + a_3b_3) \cdot \varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^3.$$

Man definiert nun auch einen Stern-Operator in der umgekehrten Richtung.

Definition:

Die Abbildung
$$* : A^{n-1}(K^n) \to A^1(K^n) = (K^n)^*$$
mit $*(\eta_{\mathbf{a}}) := (-1)^{n-1} \cdot \lambda_{\mathbf{a}}$ nennt man auch *Stern-Operator*.

Offensichtlich ist dieser Stern-Operator nicht die Umkehrabbildung des zuvor definierten, vielmehr gilt:
$$* \circ * = (-1)^{n-1} \cdot \mathrm{id}.$$
Der Grund für diese merkwürdige Konvention kann hier nicht im Detail erörtert werden. Immerhin gilt nun:

$$\begin{aligned}
\eta_{\mathbf{a}} \wedge *\eta_{\mathbf{b}} &= (-1)^{n-1} \cdot \eta_{\mathbf{a}} \wedge \lambda_{\mathbf{b}} \\
&= (-1)^{n-1} \cdot (-1)^{n-1} \cdot \lambda_{\mathbf{b}} \wedge \eta_{\mathbf{a}} \\
&= \lambda_{\mathbf{b}} \wedge *\lambda_{\mathbf{a}} \\
&= (\mathbf{b} \circ \mathbf{a}^\top) \cdot \Delta \\
&= (\mathbf{a} \circ \mathbf{b}^\top) \cdot \Delta.
\end{aligned}$$

Im Falle $n = 3$ verschwindet das Vorzeichen gerade, und dann ist $* = *^{-1}$.

Definition:

Für $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ wird das *Vektorprodukt* $\mathbf{a} \times \mathbf{b} \in \mathbb{R}^3$ definiert durch
$$\mathbf{a} \times \mathbf{b} := \varphi^E(*(\lambda_{\mathbf{a}} \wedge \lambda_{\mathbf{b}})).$$

Eigenschaften des Vektorproduktes

1. $\mathbf{v} \times \mathbf{w} = (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1)$.
2. $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$, *insbesondere* $\mathbf{v} \times \mathbf{v} = \mathbf{0}$.
3. $(\mathbf{v} + \mathbf{v}') \times \mathbf{w} = \mathbf{v} \times \mathbf{w} + \mathbf{v}' \times \mathbf{w}$.
4. $c \cdot (\mathbf{v} \times \mathbf{w}) = (c \cdot \mathbf{v}) \times \mathbf{w}$.
5. $\mathbf{u} \bullet (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \bullet (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \bullet (\mathbf{u} \times \mathbf{v})$, *und speziell*
   $\mathbf{v} \bullet (\mathbf{v} \times \mathbf{w}) = \mathbf{w} \bullet (\mathbf{v} \times \mathbf{w}) = 0$.

Beweis: 1) Es ist

$$\begin{aligned}(v_1, v_2, v_3) \times (w_1, w_2, w_3) &= \varphi^E(*(v_1\varepsilon^1 + v_2\varepsilon^2 + v_3\varepsilon^3) \wedge (w_1\varepsilon^1 + w_2\varepsilon^2 + w_3\varepsilon^3)) \\ &= \varphi^E(*((v_1w_2 - v_2w_1)\varepsilon^1 \wedge \varepsilon^2 + (v_1w_3 - v_3w_1)\varepsilon^1 \wedge \varepsilon^3 \\ &\quad + (v_2w_3 - v_3w_2)\varepsilon^2 \wedge \varepsilon^3)) \\ &= \varphi^E((v_1w_2 - v_2w_1)\varepsilon^3 + (v_3w_1 - v_1w_3)\varepsilon^2 + (v_2w_3 - v_3w_2)\varepsilon^1) \\ &= (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1).\end{aligned}$$

2) $\mathbf{v} \times \mathbf{w} = \varphi^E(*(\lambda_{\mathbf{v}} \wedge \lambda_{\mathbf{w}})) = -\varphi^E(*(\lambda_{\mathbf{w}} \wedge \lambda_{\mathbf{v}})) = -\mathbf{w} \times \mathbf{v}$.

3) und 4) folgt aus der Bilinearität des Dach-Produktes.

5) Nach Definition ist $\mathbf{v} \times \mathbf{w} = \varphi^E(*(\lambda_{\mathbf{v}} \wedge \lambda_{\mathbf{w}}))$ und $\varphi^E(\lambda_{\mathbf{v}\times\mathbf{w}}) = \mathbf{v} \times \mathbf{w}$, also

$$*(\lambda_{\mathbf{v}} \wedge \lambda_{\mathbf{w}}) = \lambda_{\mathbf{v}\times\mathbf{w}},$$

und wegen $* \circ * = \mathrm{id}$ dann

$$\mathbf{u} \bullet (\mathbf{v} \times \mathbf{w}) \cdot \Delta = \lambda_{\mathbf{u}} \wedge *\lambda_{\mathbf{v}\times\mathbf{w}} = \lambda_{\mathbf{u}} \wedge \lambda_{\mathbf{v}} \wedge \lambda_{\mathbf{w}}.$$

Wegen $f \wedge g \wedge h = g \wedge h \wedge f = h \wedge f \wedge g$ folgt die erste Behauptung, und wegen $f \wedge f \wedge g = g \wedge f \wedge g = 0$ die zweite. □

Bemerkungen:

1. Man kann auch schreiben:

$$\mathbf{v} \times \mathbf{w} = (\det\begin{pmatrix} v_2 & v_3 \\ w_2 & w_3 \end{pmatrix}, \det\begin{pmatrix} v_3 & v_1 \\ w_3 & w_1 \end{pmatrix}, \det\begin{pmatrix} v_1 & v_2 \\ w_1 & w_2 \end{pmatrix}).$$

2. Der zweite Teil von Eigenschaft (5) bedeutet, daß das Vektorprodukt von $\mathbf{v}$ und $\mathbf{w}$ auf der durch $\mathbf{v}$ und $\mathbf{w}$ aufgespannten Ebene senkrecht steht.

Beispiel:

Es ist

$$\mathbf{e}_1 \times \mathbf{e}_2 = \varphi^E(*(\varepsilon^1 \wedge \varepsilon^2)) = \varphi^E(\varepsilon^3) = \mathbf{e}_3,$$

und analog

$$\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1 \quad \text{und} \quad \mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2.$$

§4 Determinanten

Sei $\Delta = \varepsilon^1 \wedge \varepsilon^2 \wedge \ldots \wedge \varepsilon^n : K^n \times \ldots K^n \to K$ die Determinantenform.

Definition:

Ist $A \in M_{n,n}(K)$, so heißt

$$\det(A) := \Delta(\mathbf{z}_1(A), \ldots, \mathbf{z}_n(A))$$

die *Determinante* von A.

Die Determinante ist also multilinear in den Zeilen der Matrix, eine Vertauschung von Zeilen führt zu einem Vorzeichenwechsel und es ist $\det(E_n) = 1$.

Ist $A = \left(a_{ij} \;\middle|\; \begin{array}{l} i = 1, \ldots, n \\ j = 1, \ldots, n \end{array} \right)$, so folgt aus der Multilinearität:

$$\begin{aligned} \det(A) &= \Delta(\sum_{j_1=1}^{n} a_{1j_1}\mathbf{e}_{j_1}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}\mathbf{e}_{j_n}) \\ &= \sum_{j_1=1}^{n} \ldots \sum_{j_n=1}^{n} a_{1j_1} \cdot \ldots \cdot a_{nj_n} \Delta(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_n}). \end{aligned}$$

Dabei ist

$$\begin{aligned} \Delta(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_n}) &= \varepsilon^1 \wedge \ldots \wedge \varepsilon^n(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_n}) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\varepsilon^{\sigma(1)} \otimes \ldots \otimes \varepsilon^{\sigma(n)}(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_n}) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\varepsilon^{\sigma(1)}(\mathbf{e}_{j_1}) \cdot \ldots \cdot \varepsilon^{\sigma(n)}(\mathbf{e}_{j_n}) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\delta_{\sigma(1),j_1} \cdot \ldots \cdot \delta_{\sigma(n),j_n} \\ &= \begin{cases} \operatorname{sgn}(\sigma) & \text{falls } (j_1, \ldots, j_n) = (\sigma(1), \ldots, \sigma(n)), \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Es fallen also alle Summanden weg, bis auf diejenigen, bei denen

$$(j_1, \ldots, j_n) = (\sigma(1), \ldots, \sigma(n)) \text{ für ein } \sigma \in S_n \text{ ist.}$$

Damit erhalten wir:

$$\boxed{\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} \,.}$$

Beispiel:

Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in M_{2,2}(K)$.

Die Gruppe S_2 besteht nur aus den 2 Permutationen $\binom{12}{12}$ und $\binom{12}{21}$. Also ist $\det(A) = a_{11}a_{22} - a_{12}a_{21}$. Das stimmt mit der früheren Definition überein.

Bemerkung : Man schreibt manchmal auch

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

Wir berechnen nun die Determinante einer (3×3)-Matrix. S_3 besteht aus 6 Permutationen. Davon sind $\binom{123}{123}$, $\binom{123}{231}$ und $\binom{123}{312}$ gerade und $\binom{123}{132}$, $\binom{123}{213}$ und $\binom{123}{321}$ ungerade. Also gilt:

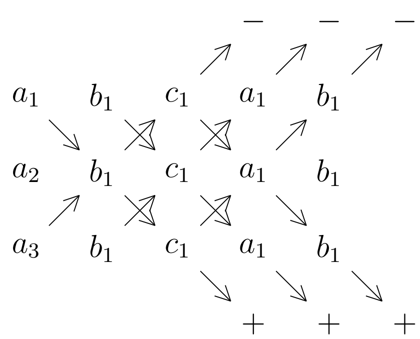
$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}.$$

Man nennt diese Formel auch die *Sarrus'sche Regel*. Es gibt ein Schema, nach dem man sie sich leicht merken kann:

Schreibe die Matrix–Elemente in der folgenden Form auf:

$$\begin{array}{ccc|cc} a_1 & b_1 & c_1 & a_1 & b_1 \\ a_2 & b_2 & c_2 & a_2 & b_2 \\ a_3 & b_3 & c_3 & a_3 & b_3 \end{array}$$

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:



Man beachte aber, daß diese Formel **nur** im Falle $n = 3$ angewandt werden darf.

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.

Wir beginnen mit der

Determinante der transponierten Matrix

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

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

Beweis: Wir sammeln erst einmal ein paar einfache Tatsachen:

1) Die Zuordnung $\sigma \mapsto \sigma^{-1}$ liefert eine bijektive Abbildung $S_n \to S_n$.

2) Für alle $\sigma \in S_n$ ist $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$. Man braucht ja nur alle Vertauschungen rückgängig zu machen.

3) σ und τ seien Elemente von S_n. Dann gilt wegen der Kommutativität der Multiplikation:

$$a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} = a_{\tau(1),\sigma\circ\tau(1)} \cdot \ldots \cdot a_{\tau(n),\sigma\circ\tau(n)}.$$

Unter Verwendung von (1) und (2), sowie (3) im Falle $\tau = \sigma^{-1}$ erhält man (mit $b_{ij} := a_{ji}$):

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1),1} \cdot \ldots \cdot a_{\sigma^{-1}(n),n} \\
&= \sum_{\sigma^{-1}\in S_n} \operatorname{sgn}(\sigma^{-1}) b_{1,\sigma^{-1}(1)} \cdot \ldots \cdot b_{n,\sigma^{-1}(n)} \\
&= \det(A^\top).
\end{aligned}$$

□

Damit ergeben sich folgende Regeln:

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

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 $\operatorname{rg}(A) < n$, so ist $\det(A) = 0$.*

5. *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 $f = \det$.*

Beweis: Das meiste ist schon bekannt, denn es ergibt sich aus den Eigenschaften der Determinantenform und dem Satz über die Determinante der transponierten Matrix.

(4): Ist $\operatorname{rg}(A) < n$, so sind die Spalten $\vec{s}_1(A), \ldots, \vec{s}_n(A)$ linear abhängig. Sei etwa

$$\vec{s}_1(A) = \sum_{j=2}^{n} \lambda_j \vec{s}_j(A).$$

Dann ist

$$\begin{aligned}
\det(A) &= \det(\vec{s}_1(A), \ldots, \vec{s}_n(A)) \\
&= \det(\sum_{j=2}^{n} \lambda_j \vec{s}_j(A), \vec{s}_2(A), \ldots, \vec{s}_n(A)) \\
&= \sum_{j=2}^{n} \lambda_j \det(\vec{s}_j(A), \vec{s}_2(A), \ldots, \vec{s}_n(A)) \\
&= 0,
\end{aligned}$$

denn in der letzten Determinante kommt $\vec{s}_j(A)$ zweimal als Argument vor.

(5): f kann als alternierende Multilinearform auf dem K^n mit $f(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$ aufgefaßt werden und muß daher mit der Determinantenform übereinstimmen. $\square$

Durch elementare Umformungen kann man jede Matrix $A \in M_{n,n}(K)$ auf Dreiecksgestalt bringen. Daher 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 $\text{rg}(A) < n$ 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 aus der Determinantenformel:

$$\det(A) = a_{11} \cdot \ldots \cdot a_{nn}.$$

$\square$

Der nächste Satz hat sowohl praktische als auch theoretische Bedeutung:

Determinanten–Produktsatz

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

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

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

Da aber $\mathrm{rg}(A \circ B) \le \mathrm{rg}(B)$ ist, ist auch $\det(A \circ B) = 0$.

2) Ist $\mathrm{rg}(B) = n$, 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 höchstens um einen Faktor $\neq 0$. Also ist $\det(B) \neq 0$.

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

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

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 $A \circ B$ 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. $\square$

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 $\quad \det(A^{-1}) = \dfrac{1}{\det(A)}\,.$

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:*

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

Beweis: Es ist

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

Die letzte Gleichung ergibt sich aus dem folgenden Satz! □

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 dann 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:

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

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

□

Laplace'scher Entwicklungssatz

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

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

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:
$$
\begin{matrix}
+ & - & + & - & \cdots \\
- & + & - & + & \\
+ & - & + & - & \\
\vdots & & & &
\end{matrix}
$$

Beispiel:

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

Das war eine Entwicklung nach der ersten Zeile. Im allgemeinen wird man eine Zeile oder Spalte suchen, in der möglichst viele Nullen zu finden sind.

Eine Anwendung der Determinantentheorie ergibt sich für die LGS:

Cramersche Regel

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

1. $A \circ \vec{x} = \vec{b}$ *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* $\vec{x} = (x_1, \ldots, x_n)^\top$ *der eindeutig bestimmte Lösungsvektor des LGS* $A \circ \vec{x} = \vec{b}$*, 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 $A \circ \vec{x} = \vec{b}$, so ist

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

Daraus folgt:

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

$\square$

Beispiel:

Sei $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} -13 \\ -7 \end{pmatrix}$. Dann gilt:

$$\det(A) = 2 + 15 = 17, \text{ d.h., das LGS ist eindeutig lösbar,}$$

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

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

Für den Lösungsvektor $\vec{x} = (x_1, x_2)^\top$ 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 $\vec{y}_j := (y_{1j}, \dots, y_{nj})^\top$ die j-te Spalte von A^{-1}. Da $A \circ A^{-1} = E_n$ ist, gilt:

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

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

$$B := \begin{pmatrix} \mathbf{j} & 0 & 2\mathbf{j} \\ 1 & 0 & 2 \\ 0 & \mathbf{j} & \mathbf{j} \\ 1 & 3 & 5 \end{pmatrix} \in M_{4,3}(\mathbb{C}).$$

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 $\mathbf{j} \cdot 5 - 3 \cdot \mathbf{j} = 2\mathbf{j} \neq 0$. 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 $\mathcal{A} = \{a_1, \ldots, a_n\}$ von V wählen. Dann wird f durch die Matrix $M = M_{\mathcal{A}}(f) \in M_{n,n}(K)$ beschrieben, mit

$$M_{\mathcal{A}}(f) := M(\Phi_{\mathcal{A}} \circ f \circ \Phi_{\mathcal{A}}^{-1}) = ([f(a_1)]_{\mathcal{A}}, \ldots, [f(a_n)]_{\mathcal{A}}).$$

Ist nun $\mathcal{A}'$ eine andere Basis von V, $M' = M_{\mathcal{A}'}(f)$ und P die zugehörige Basiswechselmatrix, so ist $M' = P^{-1} \circ M \circ P$, 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, $\mathcal{A}$ eine Basis für V und $f : V \to V$ ein Endomorphismus, so setzt man

$$\det(f) := \det(M_{\mathcal{A}}(f)).$$

Wir wollen jetzt noch eine spezielle Klasse von Endomorphismen betrachten.

Definition:

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

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

Isometrien lassen das Skalarprodukt invariant

Ist $f : \mathbb{R}^n \to \mathbb{R}^n$ eine Isometrie, so ist f ein Isomorphismus und

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

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

2) Es ist

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

Ist f eine Isometrie, so ist $\|f(\mathbf{x})\|^2 = \|\mathbf{x}\|^2$, $\quad \|f(\mathbf{y})\| = \|\mathbf{y}\|^2$ und

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

Also ist auch $f(\mathbf{x}) \bullet f(\mathbf{y}) = \mathbf{x} \bullet \mathbf{y}$. □

Eine Isometrie ist also nicht nur längentreu, sondern sie respektiert auch das Skalarprodukt und ist daher winkeltreu.

Die Matrix zu einer Isometrie

Sei $A \in M_{n,n}(\mathbb{R})$, $f = f_A : \mathbb{R}^n \to \mathbb{R}^n$. f ist genau dann eine Isometrie, wenn $A^\top \circ A = E_n$ ist.

BEWEIS: Durch $\varphi(\mathbf{x}, \mathbf{y}) := f(\mathbf{x}) \bullet f(\mathbf{y})$ wird eine Bilinearform φ auf dem $\mathbb{R}^n$ definiert.

Da $\varphi(\mathbf{x}, \mathbf{y}) = (A \circ \mathbf{x}^\top)^\top \circ (A \circ \mathbf{y}^\top) = \mathbf{x} \circ (A^\top \circ A) \circ \mathbf{y}^\top$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ist, können wir $A^\top \circ A$ als diejenige Matrix identfizieren, die φ beschreibt.

Andererseits ist f eine Isometrie, also $\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{x} \bullet \mathbf{y} = \mathbf{x} \circ \mathbf{y}^\top$. Daher muß $A^\top \circ A = E_n$ sein. □

Definition:

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

Bemerkung: Ist $A^\top \circ A = E_n$, so ist $\det(A)^2 = 1$, also insbesondere $\det(A) \neq 0$ und daher A invertierbar, mit $A^{-1} = A^\top$. Aber dann ist auch $A \circ A^\top = E_n$, d.h. $A^\top$ orthogonal.

Eigenschaften orthogonaler Matrizen

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

1. *A ist orthogonal.*
2. *Die Spalten von A bilden ein ON-System.*
3. *Die Zeilen von A bilden ein ON-System.*
4. *f_A bildet ON-Basen auf ON-Basen ab.*
5. *f_A ist eine Isometrie.*

BEWEIS:
5) $\Longrightarrow$ 4): Ist f_A Isometrie, so ist $\| f_A(\mathbf{x}) \| = \| \mathbf{x} \|$ und $f_A(\mathbf{x}) \bullet f_A(\mathbf{y}) = \mathbf{x} \bullet \mathbf{y}$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Dann ist aber klar, daß f_A ON-Basen auf ON-Basen abbildet.

4) $\Longrightarrow$ 2): f_A bilde ON-Basen auf ON-Basen ab. Da die Standardbasis eine ON-Basis ist und $f_A(\mathbf{e}_j)^\top = \vec{s}_j(A)$ die j-te Spalte von A ergibt, bilden die Spalten von A eine ON-Basis.

2) $\Longrightarrow$ 1): Die Einträge in $A^\top \circ A$ sind gerade die Produkte

$$\mathbf{e}_i \circ (A^\top \circ A) \circ \mathbf{e}_j^\top = (A \circ \vec{e}_i)^\top \circ (A \circ \vec{e}_j) = \vec{s}_i(A)^\top \circ \vec{s}_j(A) = \mathbf{s}_i(A) \bullet \mathbf{s}_j(A) = \delta_{ij}.$$

Also ist $A^\top \circ A = E_n$.

1) $\Longrightarrow$ 5): Wir wissen sogar schon, daß (1) und (5) äquivalent sind.

Es fehlt noch die Äquivalenz zu der Aussage 3): Aber die ist trivial, weil mit A auch $A^\top$ orthogonal ist. □

Notwendiges Kriterium für Orthogonalität

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

BEWEIS: Klar, denn wegen $A^\top \circ A = E_n$ 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 $A \circ S$ orthogonal, aber die neue Matrix besitzt die Determinante $+1$. Also gibt es ein φ, so daß $A = R(\varphi) \circ S^{-1} = R(\varphi) \circ S$ ist. Dabei ist $f_S(x, y) = (x, -y)$ eine 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.

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 $< \mathbf{z}, \mathbf{w} > = \mathbf{z} \circ \overline{\mathbf{w}}^\top$ 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 $\overline{A}^\top \circ A = E_n$ ist.

Der Beweis funktioniert genauso wie beim euklidischen Skalarprodukt.

Definition:

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

Die äquivalenten Bedingungen für die Orthogonalität einer Matrix lassen sich sinngemäß auf die unitären Matrizen übertragen. Weiter folgt:

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

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

Bemerkung : Das Produkt zweier orthogonaler Matrizen ist wieder orthogonal:

$$(A \circ B)^\top \circ (A \circ B) = B^\top \circ A^\top \circ A \circ B = E_n.$$

Und die inverse Matrix $A^{-1} = A^\top$ einer orthogonalen Matrix ist auch wieder orthogonal. Das gleiche gilt für unitäre Matrizen.

Ist schließlich $\det(A) = 1$ und $\det(B) = 1$, so ist auch $\det(A \circ B) = \det(A) \cdot \det(B) = 1$ und $\det(A^{-1}) = (\det(A))^{-1} = 1$.

Definition:

1. $\mathrm{SL}(n, K) := \{A \in \mathrm{GL}(n, K) \mid \det(A) = 1\}$ heißt *spezielle lineare Gruppe.*
2. $\mathrm{O}(n) := \{A \in M_{n,n}(\mathbb{R}) \mid A^\top \circ A = E_n\}$ heißt *orthogonale Gruppe*, $\mathrm{SO}(n) := \{A \in O(n) \mid \det A = 1\}$ heißt *spezielle orthogonale Gruppe.*
3. $\mathrm{U}(n) := \{A \in M_{n,n}(\mathbb{C}) \mid \overline{A}^\top \circ A = E_n\}$ heißt *unitäre Gruppe*, $\mathrm{SU}(n) := \{A \in \mathrm{U}(n) \mid \det(A) = 1\}$ heißt *spezielle unitäre Gruppe.*

Die spezielle orthogonale Gruppe ist die Gruppe der Drehungen. Im Falle $n = 2$ ist die Situation besonders einfach:

$$\mathrm{SO}(2) = \{ \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \mid \varphi \in [0, 2\pi) \}.$$

Man sagt auch, SO(2) ist eine kontinuierliche 1-parametrige Gruppe. Die Gruppe SO(3) der räumlichen Drehungen ist schon komplizierter, sie besitzt 3 kontinuierliche Parameter, die sogenannten *Eulerschen Winkel.*

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

§5 Eigenwerte und Eigenvektoren

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 $f(x) = \lambda \cdot x$ 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 bzw. Eigenwert von A einen Eigenvektor bzw. Eigenwert von $f_A : K^n \to K^n$.

Beispiel:

Sei $A := \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ und $\vec{x}_0 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Dann ist

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

Also ist 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

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

ein Untervektorraum von V.

Beweis:
1) Offensichtlich liegt 0 in $E(\lambda)$. Man beachte allerdings, daß 0 kein Eigenvektor ist!
2) Ist $x \in E(\lambda)$, also $f(x) = \lambda \cdot x$, so ist auch

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

also $\alpha \cdot x \in E(\lambda)$, für $\alpha \in K$.
3) Sind $x, y \in E(\lambda)$, so ist $f(x) = \lambda x$ und $f(y) = \lambda y$. Also ist

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

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 $M_B(f) = \left(d_{ij} \mid \begin{matrix} i = 1, \ldots, n \\ j = 1, \ldots, n \end{matrix} \right)$, 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 $M_B(f_A) = \mathrm{B}^{-1} \circ A \circ \mathrm{B}$ beschrieben, mit der regulären Matrix $\mathrm{B} := (\vec{b}_1, \ldots, \vec{b}_n)$. 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:*

$$D = \mathrm{B}^{-1} \circ A \circ \mathrm{B}$$

Wie findet man Eigenvektoren zu einer Matrix A ?

Es ist einfacher, mit den Eigenwerten zu beginnen.

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

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:

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

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

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

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 $(A - \lambda \cdot E_2) \circ \vec{x} = \vec{0}$ lösen:

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

$\lambda = 4$: $\begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ 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 $\mathrm{B} := (\vec{b}_1, \vec{b}_2) = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. Dann ist $\mathrm{B}^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$ und

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

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 $\lambda_{1,2} = \cos\varphi \pm \mathbf{j}\sin\varphi = e^{\pm\mathbf{j}\varphi}$ komplexe Eigenwerte sind.

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

$$\begin{pmatrix} -\mathbf{j}\sin\varphi & -\sin\varphi \\ \sin\varphi & -\mathbf{j}\sin\varphi \end{pmatrix} \circ \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Man sieht, daß $\begin{pmatrix} 1 \\ -\mathbf{j} \end{pmatrix}$ eine Lösung ist. Genauso ist $\begin{pmatrix} 1 \\ \mathbf{j} \end{pmatrix}$ ein Eigenvektor zum Eigenwert $\lambda_2 = e^{-\mathbf{j}\varphi}$. Offensichtlich sind $\begin{pmatrix} 1 \\ \mathbf{j} \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -\mathbf{j} \end{pmatrix}$ 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.

$\underline{k - 1 \rightarrow k}$: 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! $\square$

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

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

$\square$

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

$$p_{B^{-1}AB}(x) = \det(B^{-1} \circ A \circ B - x \cdot E_n)$$

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

$\square$

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

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

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

$$\begin{array}{rcl} a(\lambda) & := & \text{Vielfachheit der Nullstelle } \lambda \text{ des charakteristischen Polynoms } p_A(x) \\ & & \text{heißt } \textit{algebraische Vielfachheit von } \lambda. \\ g(\lambda) & := & \dim_K(E(\lambda)) \text{ heißt } \textit{geometrische Vielfachheit von } \lambda. \end{array}$$

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 $\mathrm{B} := (\vec{v}_1, \ldots, \vec{v}_n)$ eine reguläre Matrix, und es gilt:

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

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

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

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

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

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. $\square$

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 $\Delta = M_B(f_A) = \mathbf{B}^{-1} \circ A \circ \mathbf{B}$, mit $\mathbf{B} = (\vec{b}_1, \ldots, \vec{b}_n)$. 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. $\square$

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

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

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

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

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

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

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

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:

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

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. $\square$

Eigenwerte von symmetrischen und hermiteschen Matrizen

Sei $A \in M_{n,n}(K)$ symmetrisch (im Falle $K = \mathbb{R}$) oder hermitesch (im Falle $K = \mathbb{C}$). 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 (2) automatisch.

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

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

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 $f(\mathbf{z}) = \lambda \cdot \mathbf{z}$ ist. Dann folgt:

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

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. $< \mathbf{z}, \mathbf{w} > = 0$. $\square$

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 $S^{-1} \circ A \circ S$ eine reelle Diagonalmatrix ist. Die Einträge in der Diagonalmatrix sind die Eigenwerte von 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 $\mathbf{B} := (\vec{v}_1, \ldots, \vec{v}_n)$ eine unitäre Matrix. Sei

$$M := M_B(f_A) = \mathbf{B}^{-1} \circ A \circ \mathbf{B} = \overline{\mathbf{B}}^\top \circ A \circ \mathbf{B}.$$

Dann ist $\overline{M}^\top = \overline{\mathbf{B}}^\top \circ \overline{A}^\top \circ \mathbf{B} = \overline{\mathbf{B}}^\top \circ A \circ \mathbf{B} = M$, also M wieder hermitesch.

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

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

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

$$\Delta_1 := \overline{S}_1^\top \circ A_1 \circ S_1$$

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

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

Das ist die gewünschte Diagonalisierung von A. Analog führt man die Diagonalisierung von A_1 auf die einer $(n-2)$-reihigen hermiteschen Matrix A_2 zurück, usw. $\square$

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 $q_A(\mathbf{x}) := \mathbf{x} \circ A \circ \mathbf{x}^\top$ die *zugehörige quadratische Form*. Ist z.B. $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, so ist

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

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

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

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

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

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

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

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

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

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 halben Längen 2 und 3.

Zum Schluß wollen wir noch kurz auf das allgemeine Normalformenproblem eingehen.

Sei $A \in M_{n,n}(K)$, K beliebig, und es sei vorausgesetzt, daß $p_A(x)$ ganz in Linearfaktoren zerfällt:

$$p_A(x) = (x - \lambda_1)^{m_1} \cdot \ldots \cdot (x - \lambda_r)^{m_r}.$$

Im allgemeinen kann man nicht voraussetzen, daß $a(\lambda_i) = g(\lambda_i)$ ist. Deshalb arbeitet man nicht mit den Eigenräumen

$$E(\lambda_i) := \{\mathbf{x} \in K^n \mid f_A(\mathbf{x}) = \lambda \cdot \mathbf{x}\},$$

sondern mit den „verallgemeinerten Eigenräumen“

$$V(\lambda_i) := \{\mathbf{x} \in K^n \mid \exists\, m \geq 0 \text{ mit } (f_A - \lambda \cdot \mathrm{id})^m(\mathbf{x}) = \mathbf{0}\}.$$

Man kann zeigen, daß $\dim V(\lambda_i) = m_i$ ist, für $i = 1, \dots, r$.

Die Räume $V(\lambda_i)$ können getrennt voneinander behandelt werden, die Schwierigkeit des Verfahrens liegt im Detail.

Für $q \in \mathbb{N}$ sei

$$N_q := \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \in M_{q,q}(K).$$

Dann ist $(N_q)^q = 0$, d.h., N_q ist eine sogenannte „nilpotente Matrix“.

Unter einer *Partition* π einer Zahl $m \in \mathbb{N}$ versteht man ein s-Tupel $\pi = (q_1, \dots, q_s)$ von natürlichen Zahlen mit folgenden Eigenschaften:

1. $m = q_1 + \cdots + q_s$.
2. $q_1 \geq q_2 \geq \ldots \geq q_s$.

Man setzt dann $N_\pi := \begin{pmatrix} N_{q_1} & & 0 \\ & \ddots & \\ 0 & & N_{q_s} \end{pmatrix}$.

Die *Jordan-Matrix* $J_\pi(\lambda)$ zur Zerlegung π und zum Eigenwert λ ist nun die Matrix

$$J_\pi(\lambda) := \lambda \cdot E_m + N_\pi = \begin{pmatrix} \lambda \cdot E_{q_1} + N_{q_1} & & 0 \\ & \ddots & \\ 0 & & \lambda \cdot E_{q_s} + N_{q_s} \end{pmatrix}.$$

Sie hat in der Hauptdiagonalen immer den Wert λ stehen, oberhalb der Hauptdiagonalen Einsen und einige Nullen und sonst überall Nullen.

Satz von der Jordanschen Normalform

Sei $A \in M_{n,n}(K)$, das charakteristische Polynom $p_A(x)$ zerfalle wie oben beschrieben in Linearfaktoren. Dann gibt es Partitionen

$$\pi_i = (q_1^{(i)}, \dots, q_{s_i}^{(i)}) \text{ von } m_i, \text{ für } i = 1, \dots, r,$$

und eine reguläre Matrix $P \in M_{n,n}(K)$, so daß gilt:

$$P^{-1} \circ A \circ P = \begin{pmatrix} J_{\pi_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{\pi_r}(\lambda_r) \end{pmatrix}.$$

Man nennt das die Jordansche Normalform *von A*

Wenn $p_A(x)$ nicht zerfällt, wird es noch schwieriger.