Lineare Algebra 2a/Diskrete Mathematik A
Frühjahrssemester 2011
Mannheim

Claus Hertling

24.03.2011

Inhaltsverzeichnis

Die Vorlesung LA 2a/DMA ist eine Fortsetzung der Vorlesung LA I im Herbstsemester 2010 in Mannheim. Die Notationen und Begriffe, die in der Vorlesung LA I entwickelt wurden, werden hier als bekannt vorausgesetzt. Die Vorlesung LA 2a/DMA dauert nur 7 Wochen. In den zweiten 7 Wochen des FS 2011 gibt es für einige von Ihnen Teil LA 2b.

hertling@math.uni-mannheim.de

.

7 Determinanten

In diesem Kapitel bezeichnet K irgendeinen Körper und R irgendeinen kommutativen Ring mit Eins.
Eine Motivation für Determinanten besteht in der Beziehung zum Volumen eines Parallelotops, siehe Bemerkung 7.20.

Definition 7.1 (Leibniz-Formel) Die *Determinante* $\det A$ einer quadratischen Matrix $A = (a_{ij}) \in M(n \times n, R)$ (mit $n \geq 1$) ist

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

Notation: manchmal schreibt man $|A| = \det A$ (aber nur bei $n \geq 2$).

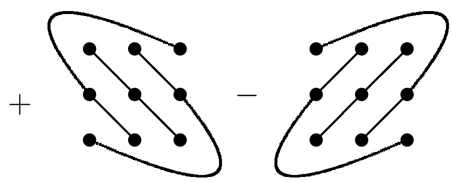
Beispiele 7.2 a) $n = 2$: $\quad S_2 = \{\mathrm{id}, (1\ 2)\}$,

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

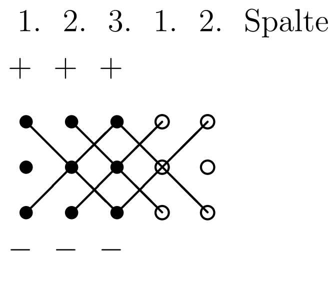
b) $n = 3$: $\quad S_3 = \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2), (1\ 2), (1\ 3), (2\ 3)\}$,

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

Sarrussche Regel:



oder



c) $A = (a_{ij}) \in M(n \times n, R)$ obere Dreiecksmatrix, d.h. $a_{ij} = 0$ für $i > j$:

$$\det A = a_{11} \cdot \ldots \cdot a_{nn},$$

denn für $\sigma \in S_n - \{id\}$ gibt es ein $i \in \{1, ..., n\}$ mit $i > \sigma(i)$.

d) (Verallgemeinerung von c)) Sei $A = (a_{ij}) \in M((n+m) \times (n+m), R)$ eine Matrix der Gestalt

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

mit $B \in M(n \times n, R)$, $C \in M(n \times m, R)$, $D \in M(m \times m, R)$ und 0 die Nullmatrix in $M(m \times n, R)$. Dann ist

$$\det A = \det B \cdot \det D.$$

Beweis: Die Null links unten in A besagt $a_{ij} = 0$, falls $i > n, j \leq n$. Daher verschwinden in der Leibniz-Formel alle Summanden für $\sigma \in S_n$ mit $\sigma(i) \leq n$ für irgendein $i > n$. Es bleiben nur die, für die gilt: σ bildet $\{n+1, ..., n+m\}$ auf sich ab (also bijektiv), und bildet daher auch $\{1, ..., n\}$ auf sich ab. Solche σ lassen sich eindeutig schreiben als $\sigma = \sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$ mit

$$\sigma_1|_{\{n+1,...,n+m\}} = \mathrm{id} \quad \text{ und } \quad \sigma_2|_{\{1,...,n\}} = \mathrm{id}\,.$$

Die Summe in der Leibniz-Formel läßt sich aufspalten in eine Doppelsumme, über die möglichen σ_1 und über die möglichen σ_2. Es ist auch

$$\begin{aligned} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{(n+m)\sigma(n+m)} &= \mathrm{sign}(\sigma_1) \cdot a_{1\sigma_1(1)} \cdot ... \cdot a_{n\sigma_1(n)} \\ &\cdot \ \mathrm{sign}(\sigma_2) \cdot a_{(n+1)\sigma_2(n+1)} \cdot ... \cdot a_{(n+m)\sigma_2(n+m)} \end{aligned}$$

Das zeigt $\det A = \det B \cdot \det D$. □

Definition 7.3 Eine Abbildung $\delta : M(n \times n, R) \to R$ heißt

1) *multilinear* bezüglich der Zeilen, falls gilt:

   a) Entsteht A' aus A durch Multiplikation einer Zeile mit $\lambda \in R$, so ist
   $$\delta(A') = \lambda \cdot \delta(A).$$

   b) Unterscheiden sich A, A', A'' nur in der j-ten Zeile und ist die j-te Zeile von A die Summe der j-ten Zeilen von A' und A'', so ist
   $$\delta(A) = \delta(A') + \delta(A'').$$

2) *alternierend* bezüglich der Zeilen, falls gilt:

   Ist A eine Matrix mit zwei gleichen Zeilen, so ist $\delta(A) = 0$.

3) *schiefsymmetrisch* bezüglich der Zeilen, falls gilt:
   Entsteht A' aus A durch Vertauschen von zwei Zeilen, so ist
   $$\delta(A') = -\delta(A).$$

4) *normiert*, falls gilt:
   $$\delta(E_n) = 1.$$

Lemma 7.4 *Sei $\delta : M(n \times n, R) \to R$ eine Abbildung.*
a) Erfüllt sie 1) und 2), so auch 3).
b) Ist $2 = 1_R + 1_R \in R$ in R invertierbar (im Fall $R = K$ bedeutet das, daß $\mathrm{char}(K) \neq 2$ sein muß) und erfüllt δ 1) und 3), so auch 2).

Beweis: Es sei $1 \leq k < l \leq n$ und $A(x, y)$ die Matrix mit i-ter Zeile $(a_{i1}, ..., a_{in})$ für $i \notin \{k, l\}$, k-ter Zeile $x = (x_1, ..., x_n)$ und l-ter Zeile $y = (y_1, ..., y_n)$.
a) Die folgende Gleichung zeigt a),

$$\begin{aligned} 0 &= \delta(A(x+y, x+y)) \\ &= \delta(A(x,x)) + \delta(A(x,y)) + \delta(A(y,x)) + \delta(A(y,y)) \\ &= \delta(A(x,y)) + \delta(A(y,x)). \end{aligned}$$

b) Die folgende Gleichung zeigt b),

$$\delta(A(x,x)) = -\delta(A(x,x)), \text{ also } \quad 0 = 2 \cdot \delta(A(x,x)).$$

□

Satz 7.5 *Die Abbildung* $\det : M(n \times n, R) \to R$ *ist multilinear bezüglich der Zeilen, alternierend bezüglich der Zeilen und normiert.*

Beweis: 1) a)

$$\begin{aligned} \det A' &= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot (\lambda a_{i\sigma(i)}) \cdot ... \cdot a_{n\sigma(n)} \\ &= \lambda \cdot \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{i\sigma(i)} \cdot ... \cdot a_{n\sigma(n)} \\ &= \lambda \cdot \det A. \end{aligned}$$

1) b)

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \cdot ... \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} ... + \sum_{\sigma \in S_n} ... = \det A' + \det A''. \end{aligned}$$

2) Die k-te und die l-te Zeile von $A = (a_{ij})$ seien gleich (mit $k \neq l$ natürlich), d.h. $a_{ki} = a_{li}$ für alle i. Es sei $\tau := (k\ l) \in S_n$. Es ist $S_n = A_n \cup A_n\tau$.

$$\begin{aligned} \det A &= \sum_{\sigma \in A_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{k\sigma(k)} \cdot ... \cdot a_{l\sigma(l)} \cdot ... \cdot a_{n\sigma(n)} \\ &+ \sum_{\sigma \in A_n} \mathrm{sign}(\sigma \circ \tau) \cdot a_{1\sigma(1)} \cdot ... \cdot a_{k\sigma\circ\tau(k)} \cdot ... \cdot a_{l\sigma\circ\tau(l)} \cdot ... \cdot a_{n\sigma(n)} \\ &= 0, \end{aligned}$$

denn
$$\operatorname{sign}(\sigma) = +1, \qquad \operatorname{sign}(\sigma \circ \tau) = (-1) \cdot \operatorname{sign}(\sigma) = -1,$$
und
$$a_{k\sigma\circ\tau(k))} = a_{k\sigma(l)} = a_{l\sigma(l)} \quad \text{und} \quad a_{l\sigma\circ\tau(l)} = a_{l\sigma(k)} = a_{k\sigma(k)}.$$
4) Klar nach Beispiel 7.2 c). □

Bemerkungen 7.6 i) Aus Satz 7.5 und Lemma 7.4 folgt, daß det in folgender Weise mit den elementaren Zeilenumformungen $Z_I(\lambda; i), Z_{II}(\lambda; i, j)$ und $Z_{III}(i, j) : M(n \times n, R) \to M(n \times n, R)$ von Definition 4.4 verträglich ist:
$$\begin{aligned} \det(Z_I(\lambda; i)(A)) &= \lambda \cdot \det A, \\ \det(Z_{II}(\lambda; i, j)(A)) &= \det A, \\ \det(Z_{III}(i, j)(A)) &= -\det A. \end{aligned}$$

ii) Im Fall $R = K$ (und bei günstigen Koeffizienten auch in anderen Fällen) kann man A mit elementaren Zeilenumformungen der Typen Z_{II} und Z_{III} auf obere Dreiecksgestalt bringen (Gauß-Algorithmus, Satz 7.4). Z_{II} ändert dabei nichts an der Determinante, Z_{III} ändert das Vorzeichen. Danach wird die Berechnung von $\det A$ wegen Beispiel 7.2 c) trivial.

iii) **Methoden zur Berechnung von** det**:**

- **Gauß-Algorithmus:** Die Methode in ii) ist nur bei $R = K$ anwendbar (aber das ist meistens der Fall), oder wenn man Glück mit den Koeffizienten hat. Aber dann ist sie fast immer die schnellste.

- **Leibniz-Formel:** Nur bei $n = 2$ oder $n = 3$ (Beispiel 7.2) schnell. Für großes n ist $n!$ horrend groß.

- **Laplacescher Entwicklungssatz:** Satz 7.15; bei $n = 3$ okay; sonst nur gut, wenn in einer Zeile oder Spalte viele Einträge Null sind.

iv) Die Leibniz-Formel und der Laplacesche Entwicklungssatz sind aber für theoretische Aussagen über det wichtig (die Leibniz-Formel zum Beispiel zur Definition von det).
v) Im folgenden Satz muß man Koeffizienten in einem Körper K betrachten, denn nur dann ist $\operatorname{rang}(A)$ wohldefiniert.

Satz 7.7 *Sei $A \in M(n \times n, K)$ eine Matrix mit Koeffizienten in einem Körper K. Dann gilt:*
$$\det(A) \neq 0 \iff \operatorname{rang}(A) = n \iff A \text{ ist invertierbar.}$$

Beweis: Spezialfall: Für obere Dreiecksmatrizen gelten die Äquivalenzen oben wegen Beispiel 7.2 c), Lemma 4.12 und Satz 4.13.
Allgemeiner Fall: Durch elementare Zeilenumformungen der Typen Z_{II} und Z_{III} läßt sich eine Matrix auf obere Dreiecksgestalt bringen. Ihre Determinante ändert dabei höchstens das Vorzeichen. □

Satz 7.8 *a) Für $A \in M(n \times n, R)$ gilt*

$$\det A = \det A^{tr}.$$

b) Daher ist det *multilinear und alternierend auch bezüglich der Spalten, und Formeln analog zu denen in Bemerkung 7.6 i) gelten für die elementaren Spaltenumformungen der Typen S_I, S_{II} und S_{III} (Bemerkung 4.8 ii)).*

Beweis: a) Die folgende Rechnung benutzt drei Aussagen:

i) Weil σ eine Bijektion auf $\{1, ..., n\}$ ist, ist

$$\{(\sigma(1), 1), ..., (\sigma(n), n)\} = \{(1, \sigma^{-1}(1)), ..., (n, \sigma^{-1}(n))\}.$$

ii) Wenn σ die Menge S_n durchläuft, durchläuft auch σ^{-1} sie, denn die Abbildung $S_n \to S_n, \quad \sigma \mapsto \sigma^{-1}$, ist bijektiv.

iii) $\text{sign}(\sigma) = \text{sign}(\sigma^{-1})$.

$$\begin{aligned} \det A^{tr} &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{\sigma(1),1} \cdot ... \cdot a_{\sigma(n),n} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma^{-1}) \cdot a_{1\sigma^{-1}(1)} \cdot ... \cdot a_{n\sigma^{-1}(n)} = \det A. \end{aligned}$$

b) Klar. □

Beispiel 7.9 Daher kann man Matrizen mit elementaren Zeilen- **und** Spaltenumformungen vereinfachen, wenn man ihre Determinanten ausrechnen will. Im folgenden Beispiel wurden erst untereinanderstehende Zeilen subtrahiert (von oben nach unten); dann wurde die erste Spalte zu allen anderen addiert.

$$\begin{vmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & 5 & 4 & 3 & 2 \\ 3 & 4 & 5 & 4 & 3 \\ 2 & 3 & 4 & 5 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 \\ 1 & 2 & 3 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 1 & 2 & 2 & 0 & 0 \\ 1 & 2 & 2 & 2 & 0 \\ 1 & 3 & 4 & 5 & 6 \end{vmatrix} = 48.$$

Satz 7.10 (Weierstraß-Axiome)
Ist $\delta : M(n \times n, R) \to R$ multilinear und alternierend bezüglich der Zeilen und normiert, so ist $\delta = \det$. Das heißt, det *ist durch diese Eigenschaften (=Weierstraß-Axiome) eindeutig charakterisiert.*

Beweis (nicht in der Vorlesung): Sei $\delta : M(n \times n, R) \to R$ multilinear und alternierend. Mit $e_j = (0, ..., 0, 1, 0, ..., 0) = (\delta_{jk})_{k=1,...,n}$ (für $j = 1, ..., n$) wird der Zeilenvektor mit einer Eins an der j-ten Stelle und Nullen sonst bezeichnet.

1. Teil: Für $\sigma : \{1, ..., n\} \to \{1, ..., n\}$ sei $P_\sigma \in M(n \times n, R)$ die Matrix mit

$$(P_\sigma)_{ij} := \delta_{\sigma(i),j}.$$

Ihre i-te Zeile ist $e_{\sigma(i)}$. Sind $1 \le k < l \le n$ und ist $\tau = (k\ l)$ die Transposition, die k und l vertauscht, so ist

$$P_{\sigma\circ\tau} = Z_{III}(k, l)(P_\sigma).$$

Wegen Lemma 7.4 ist δ schiefsymmetrisch. Daher ist

$$\delta(P_{\sigma\circ\tau}) = -\delta(P_\sigma),$$

Im Fall $\sigma \in S_n$ folgt

$$\delta(P_\sigma) = \mathrm{sign}(\sigma) \cdot \delta(E_n).$$

Im Fall $\sigma \notin S_n$ (d.h. σ ist nicht bijektiv), gibt es zwei Zeilen von P_σ, die gleich sind; weil σ alternierend ist, ist dann

$$\delta(P_\sigma) = 0.$$

2. Teil: Sei $A = (a_{ij}) \in M(n \times n, R)$. Mit $a_i = (a_{ij})_{j=1,...,n} \in M(1 \times n, R)$ wird die i-te Zeile von A bezeichnet. Aus der Multilinearität von δ folgt

$$\begin{aligned}
\delta(A) &= \sum_{j_1=1}^{n} a_{1j_1} \cdot \delta(\begin{pmatrix} e_{j_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix}) = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{1j_1} a_{2j_2} \cdot \delta(\begin{pmatrix} e_{j_1} \\ e_{j_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix}) \\
&= ... = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} ... \sum_{j_n=1}^{n} a_{1j_1} a_{2j_2} ... a_{nj_n} \cdot \delta(\begin{pmatrix} e_{j_1} \\ e_{j_2} \\ \vdots \\ e_{j_n} \end{pmatrix}) \\
&= \sum_{\sigma:\{1,...,n\}\to\{1,...,n\}} a_{1\sigma(1)} a_{2\sigma(2)} ... a_{n\sigma(n)} \cdot \delta(P_\sigma)
\end{aligned}$$

Mit dem 1. Teil und der Leibniz-Formel folgt

$$\delta(A) = \det(A) \cdot \delta(E_n).$$

$\square$

Satz 7.11 *a)* **(Cauchy)** *Seien A und B in $M(n \times n, R)$. Es ist*

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

b) (Achtung, hier ist der Koeffizientenbereich ein Körper K.) Die Einschränkung von det *auf $GL(n,K)$ ist ein Gruppenhomomorphismus*

$$\det : (GL(n,K), \circ) \to (K - \{0\}, \cdot).$$

Insbesondere ist

$$\det(A^{-1}) = (\det A)^{-1}.$$

Beweis (nicht in der Vorlesung): a) Sei $B \in M(n \times n, R)$ fest und $X \in M(n \times n, R)$ variabel. Die Abbildung

$$\delta : M(n \times n, R) \to M(n \times n, R), \quad X \mapsto \det(X \cdot B)$$

ist multilinear und alternierend bezüglich der Zeilen; denn elementare Zeilenumformungen vertauschen mit der Multiplikation von rechts mit B. Nach dem Beweis von Satz 7.10 ist daher

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

b) Das folgt aus a). □

Definition/Lemma 7.12 *Die Menge*

$$SL(n,K) := \{A \in M(n \times n, K) \mid \det A = 1\}$$

ist eine Untergruppe (sogar ein Normalteiler) von $(GL(n,K), \circ)$. Sie heißt spezielle lineare Gruppe.

Beweis: Satz 7.11 b) und Beispiel 1.35 ii) (und Lemma 1.22). □

Definition 7.13 Sei $A \in M(n \times n, R)$ und $i, j \in \{1, ..., n\}$.
Mit $A[i,j]$ wird die $((n-1) \times (n-1))$-Matrix bezeichnet, die man aus A erhält, indem man die i-te Zeile und die j-te Spalte streicht.
Skizze in der Vorlesung
Die Matrix $A^\sharp \in M(n \times n, R)$ hat die Einträge

$$(A^\sharp)_{ij} := (-1)^{i+j} \det A[j,i]$$

(Vorsicht; $[j,i]$, nicht $[i,j]$) und wird *Komplementärmatrix zur Matrix A* genannt (Satz 7.15 c) zeigt, warum).

Die $n \times n$-Matrix mit Eintrag $(-1)^{i+j}$ an der Stelle (i, j) sieht so aus (die Einsen sind weggelassen):

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Beispiele 7.14 i) Im Fall einer (2×2)-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, R)$ ist

$$A^\sharp = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

also

$$A \cdot A^\sharp = A^\sharp \cdot A = (ad - bc) \cdot E_2 = \det A \cdot E_2.$$

ii) (Vgl. Beispiel 4.14 .)

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 9 \end{pmatrix}, \\ \det A &= \begin{vmatrix} 0 & 1 & 2 \\ 3 & 3 & 3 \\ 0 & 0 & 1 \end{vmatrix} = (-1) \cdot \begin{vmatrix} 3 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = -3. \end{aligned}$$

$$A^\sharp = \begin{pmatrix} \begin{vmatrix} 4 & 5 \\ 7 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 7 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \\ -\begin{vmatrix} 3 & 5 \\ 6 & 9 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ 6 & 9 \end{vmatrix} & -\begin{vmatrix} 0 & 2 \\ 3 & 5 \end{vmatrix} \\ \begin{vmatrix} 3 & 4 \\ 6 & 7 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 6 & 7 \end{vmatrix} & \begin{vmatrix} 0 & 1 \\ 3 & 4 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 5 & -3 \\ 3 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}$$

Beispiel 4.14 zeigt

$$A^\sharp = (-3) \cdot A^{-1}, \qquad \text{also } A \cdot A^\sharp = A^\sharp \cdot A = \det A \cdot E_3.$$

Satz 7.15 *Sei $A \in M(n \times n, R)$.*
*a) (***Laplacescher Entwicklungssatz** *für die i-te Zeile) Sei $i \in \{1, ..., n\}$.*

$$\det A = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} \cdot \det A[i, j] = \sum_{j=1}^{n} a_{ij} \cdot (A^\sharp)_{ji}.$$

*b) (***Laplacescher Entwicklungssatz** *für die j-te Spalte) Sei $j \in \{1, ..., n\}$.*

$$\det A = \sum_{i=1}^{n} a_{ij} \cdot (-1)^{i+j} \cdot \det A[i,j] = \sum_{i=1}^{n} a_{ij} \cdot (A^\sharp)_{ji}.$$

c)

$$A \cdot A^\sharp = A^\sharp \cdot A = \det A \cdot E_n.$$

*d) (***Cramersche Regel***, allgemeine Version für R und für* $\det A$ *beliebig)*
Sei $b \in M(n \times 1, R)$. Eine Lösung $y \in M(n \times 1, R)$ des linearen Gleichungssystems

$$A \cdot x = \det A \cdot b$$

ist offenbar

$$y = A^\sharp \cdot b.$$

Die Koeffizienten y_j dieser Lösung $y = (y_1, ..., y_n)^{tr}$ sind gegeben durch

$$y_j = \det B_j \qquad \text{für } j = 1, ..., n,$$

wo B_j aus A entsteht, indem man die j-te Spalte von A durch b ersetzt.

Beweis nach Korollar 7.17.

Beispiel 7.16 (Vgl. Beispiel 7.14 iii)) Laplace-Entwicklung nach der zweiten Zeile,

$$\begin{aligned} \det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 9 \end{pmatrix} &= -3 \cdot \begin{vmatrix} 1 & 2 \\ 7 & 9 \end{vmatrix} + 4 \cdot \begin{vmatrix} 0 & 2 \\ 6 & 9 \end{vmatrix} - 5 \cdot \begin{vmatrix} 0 & 1 \\ 6 & 7 \end{vmatrix} \\ &= -3 \cdot (-5) + 4 \cdot (-12) - 5 \cdot (-6) = -3. \end{aligned}$$

Korollar 7.17 ist eine unmittelbare Folgerung von Satz 7.15.

Korollar 7.17 *Sei $A \in M(n \times n, K)$ mit* $\det A \neq 0$.
a)

$$A^{-1} = (\det A)^{-1} \cdot A^\sharp.$$

*b) (***Cramersche Regel***, klassische Version für K und für* $\det A \neq 0$*)*
Sei $b \in M(n \times 1, K)$. Die eindeutige Lösung $y = (y_1, ..., y_n)^{tr} = A^{-1} \cdot b \in M(n \times 1, K)$ des linearen Gleichungssystems

$$A \cdot x = b$$

ist gegeben durch

$$y_j = \frac{\det B_j}{\det A} \qquad \text{für } j = 1, ..., n,$$

wo B_j aus A entsteht, indem man die j-te Spalte von A durch b ersetzt.

Beweis von Satz 7.15: a) Das zweite Gleichheitszeichen folgt aus der Definition von $A^\sharp$.
Zum ersten Gleichheitszeichen: Für einen Moment wird mit $\widetilde{A}[i,j] \in M(n \times n, R)$ die Matrix bezeichnet, die aus A entsteht, indem man die i-te Zeile von A durch die Zeile $e_j = (\delta_{jk})_{k=1,\dots,n}$ ersetzt. Aus der Multilinearität von det folgt

$$\det A = \sum_{j=1}^{n} a_{ij} \cdot \det \widetilde{A}[i,j].$$

Durch $i-1$ Zeilenvertauschungen und $j-1$ Spaltenvertauschungen erhält man aus $\widetilde{A}[i,j]$ eine Matrix, die so aussieht,

$$\begin{pmatrix} 1 & 0 & \cdots \\ * & & \\ \vdots & & A[i,j] \end{pmatrix}.$$

Nun zeigen Beispiel 7.2 d) und Satz 7.8

$$\det \widetilde{A}[i,j] = (-1)^{i+j} \cdot \det A[i,j].$$

b) Analog zu a).

c) Wegen a) ist der Diagonaleintrag $(A \cdot A^\sharp)_{ii}$ von $A \cdot A^\sharp$

$$(A \cdot A^\sharp)_{ii} = \sum_{j=1}^{n} a_{ij} \cdot A^\sharp_{ji} = \det A = (\det A \cdot E_n)_{ii}.$$

Ersetzt man in a) die Matrix A durch die Matrix, die man aus A erhält, indem man die i-te Zeile durch die k-te Zeile ersetzt (mit $k \neq i$), so hat sie zwei gleiche Zeilen. Also ist ihre Determinante 0, und a) gibt $\overset{a)}{=}$ in

$$(A \cdot A^\sharp)_{ki} = \sum_{j=1}^{n} a_{kj} \cdot A^\sharp_{ji} \overset{a)}{=} 0 = (\det A \cdot E_n)_{ki}.$$

Also ist $A \cdot A^\sharp = \det A \cdot E_n$. Analog zeigt man $A^\sharp \cdot A = \det A \cdot E_n$.

d) Laplace-Entwicklung von B_j nach der j-ten Spalte gibt

$$\det B_j = \sum_{k=1}^{n} b_k \cdot (-1)^{k+j} \cdot \det A[k,j] = \sum_{k=1}^{n} b_k \cdot (A^\sharp)_{jk},$$

also

$$(\det B_1, ..., \det B_n)^{tr} = A^\sharp \cdot b = y.$$

Wegen $A \cdot A^\sharp = \det A \cdot E_n$ ist das eine Lösung des Gleichungssystems $A \cdot x = \det A \cdot b$. $\square$

Bemerkungen 7.18 i) $A^\sharp$ ist mehr aus theoretischen als aus praktischen Gründen wichtig. Ein Pluspunkt ist, daß $A^\sharp$ für beliebige kommutative Ringe R mit Eins definiert ist, ohne Annahmen über $\det A$ und ohne daß man dividieren muß.
Aber meistens ist $R = K$ ein Körper, und dann berechnet man A^{-1} am besten wie in Bemerkung 4.14.

ii) Das gleiche gilt für die Cramersche Regel, beide Versionen. *Praktisch* löst man lineare Gleichungssysteme besser mit dem Gauß-Algorithmus als mit der Cramerschen Regel.
Aber die geschlossene Formel in der Cramerschen Regel zeigt zum Beispiel, daß die Lösung y stetig von den Koeffizienten in A und b abhängt.

Satz/Definition 7.19 *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums.*
a) (Satz) Sind $\mathcal{A}$ und $\mathcal{B}$ zwei Basen von V, so ist

$$\begin{aligned} \det M(\mathcal{A}, f, \mathcal{A}) &= \det\left(M(\mathcal{A}, \mathcal{B}) \cdot M(\mathcal{B}, f, \mathcal{B}) \cdot M(\mathcal{B}, \mathcal{A})\right) \\ &= \det M(\mathcal{B}, \mathcal{A})^{-1} \cdot \det M(\mathcal{B}, f, \mathcal{B}) \cdot \det M(\mathcal{B}, \mathcal{A}) \\ &= \det M(\mathcal{B}, f, \mathcal{B}). \end{aligned}$$

Daher ist $\det M(\mathcal{B}, f, \mathcal{B})$ *unabhängig von der Wahl der Basis $\mathcal{B}$.*
b) (Definition) Daher ist die Determinante von f wohldefiniert durch

$$\det f := \det M(\mathcal{B}, f, \mathcal{B}).$$

Beweis: Klar. □

Bemerkungen 7.20 (Volumen und Orientierung im $\mathbb{R}^n$)
i) Es seien $b_1, ..., b_n$ (Zeilen-)Vektoren im $\mathbb{R}^n$. Das von ihnen erzeugte *Parallelotop* ist

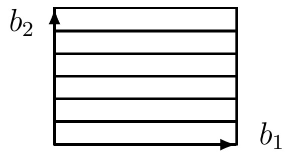
$$P(b_1, ..., b_n) = \{\sum_{j=1}^{n} x_j b_j \mid 0 \le x_j \le 1 \text{ für alle } i = 1, ..., n\}.$$

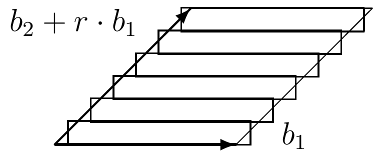
Sein *Volumen* ist

$$\text{Volumen } (P(b_1, ..., b_n)) = |\det(\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix})|.$$

Das wird hier nicht bewiesen, denn zuerst bräuchte man eine Definition von "Volumen", die hier auch nicht gegeben wird.
Immerhin ist klar, daß das Volumen unter "Scherungen" invariant sein soll, und das paßt zur Multilinearität von det.

Ein Parallelotop heißt im Fall $n = 2$ *Parallelogramm*, im Fall $n = 3$ *Spat.*





ii) Beim Volumen oben vergißt man das Vorzeichen von $\det(b_1, ..., b_n)^{tr}$. Was ist seine Rolle?
Jeder Basis $(b_1, ..., b_n)$ des $\mathbb{R}^n$ ist eine *Orientierung* zugeordnet, die Zahl

$$\frac{\det(b_1, ..., b_n)^{tr}}{|\det(b_1, ..., b_n)^{tr}|}.$$

Es gibt also nur zwei Orientierungen, $+1$ oder -1.
Bei einer Permutation $(e_{\sigma(1)}, ..., e_{\sigma(n)})$ $(\sigma \in S_n)$ der Standardbasis ist die Orientierung gerade $\mathrm{sign}(\sigma)$.
Man kann zeigen, daß die Gruppe $GL(n, \mathbb{R})$ zwei "Zusammenhangskomponenten hat", die Teilmengen

$$\{A \in GL(n, \mathbb{R}) \mid \det A > 0\} \text{ und } \{A \in GL(n, \mathbb{R}) \mid \det A < 0\}.$$

Daher lassen sich je zwei Basen mit gleicher Orientierung durch eine "stetige Familie" von Basen verbinden.
Beim $\mathbb{C}^n$ hat man das Phänomen der Orientierung nicht: im Gegensatz zu $\mathbb{R} - \{0\}$ ist $\mathbb{C} - \{0\}$ "zusammenhängend".

8 Eigenvektoren und Eigenwerte

In diesem Kapitel bezeichnet K irgendeinen Körper. Hier geht es um Normalformen für Endomorphismen.

Definition/Lemma 8.1 *a) (Definition) Sei $A \in M(n \times n, K)$. Das* charakteristische Polynom *von A ist das Polynom*

$$\begin{aligned} P_A(t) := \det(t \cdot E_n - A) &= (-1)^n \det(A - t \cdot E_n) \\ &= (-1)^n \cdot \begin{vmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} - t \end{vmatrix}. \end{aligned}$$

(Lemma) Es ist

$$\begin{aligned} P_A(t) &= t^n - (a_{11} + a_{22} + ... + a_{nn})t^{n-1} \\ &\quad +(...)t^{n-2} + ... + (...)t + (-1)^n \det A \in K[t]. \end{aligned}$$

Es ist also ein unitäres (d.h. Leitkoeffizient 1) Polynom vom Grad n. Der Koeffizient

$$\mathrm{Spur}(A) := a_{11} + a_{22} + ... + a_{nn}$$

von $-t^{n-1}$ heißt Spur *von A.*

b) (Lemma) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Sei $\mathcal{B}$ eine Basis von V. Das Polynom

$$P_f(t) := (-1)^n \det(M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n) = P_{M(\mathcal{B},f,\mathcal{B})}(t)$$

ist unabhängig von der Wahl der Basis $\mathcal{B}$.
(Definition) Es heißt charakteristisches Polynom *von f.*
Auch die Zahl

$$\mathrm{Spur}(f) := \mathrm{Spur}(M(\mathcal{B}, f, \mathcal{B}))$$

ist unabhängig von der Wahl von $\mathcal{B}$. Sie heißt Spur *von f.*

Beweis: a) Rechnet man $P_A(t)$ mit der Leibniz-Formel aus, so sieht man, daß nur der Summand $(-1)^n(a_{11} - t)(a_{22} - t)...(a_{nn} - t)$ zu $\mathrm{id} \in S_n$ Beiträge zu t^n und t^{n-1} liefert. Daher sind deren Koeffizienten 1 und $-(a_{11} + ... + a_{nn})$. Der konstante Koeffizient ist natürlich $(-1)^n \det A$.
b) Ist $\widetilde{\mathcal{B}}$ eine zweite Basis, so sei $B := M(\mathcal{B}, \widetilde{\mathcal{B}})$ die Basiswechselmatrix. Es ist

$$\begin{aligned} \det(M(\widetilde{\mathcal{B}}, f, \widetilde{\mathcal{B}}) - t \cdot E_n) &= \det\left[B^{-1} \cdot M(\mathcal{B}, f, \mathcal{B}) \cdot B - t \cdot E_n\right] \\ &= \det\left[B^{-1} \cdot (M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n) \cdot B\right] \\ &= \det(B^{-1}) \cdot \det(M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n) \cdot \det B \\ &= \det(M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n). \end{aligned}$$

Daher ist $P_f(t)$ unabhängig von der Wahl einer Basis $\mathcal{B}$. Das gleiche gilt für Spur(f), da es der Koeffizient von $-t^{n-1}$ in $P_f(t)$ ist. $\square$

Beispiele 8.2 i) (Obere Dreiecksmatrix) Ist $A = (a_{ij}) \in M(n \times n, K)$ eine obere Dreiecksmatrix (also $a_{ij} = 0$ für $i > j$), so ist auch $t \cdot E_n - A$ eine obere Dreiecksmatrix, mit Diagonaleinträgen $t - a_{11}, ..., t - a_{nn}$. Daher ist

$$P_A(t) = \det(t \cdot E_n - A) = \prod_{i=1}^{n} (t - a_{ii}).$$

Ein wichtiger Spezialfall: A eine *Diagonalmatrix*, d.h. $a_{ij} = 0$ für $i \neq j$,

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

ii) (Jordanblock) Eine Matrix $A \in M(n \times n, K)$ ist ein *Jordanblock*, falls sie die Gestalt hat

$$A = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}.$$

Das ist auch ein Spezialfall einer oberen Dreiecksmatrix, mit

$$P_A(t) = (t - \lambda)^n.$$

iii) (Obere Blockdreiecksmatrix) Ist $A \in M(n \times n, K)$ eine obere Blockdreiecksmatrix mit Blöcken $A_i \in M(r_i \times r_i, K)$ in der Diagonalen, so ist auch $t \cdot E_n - A$ eine obere Blockdreiecksmatrix:

$$A = \begin{pmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix}, \ t \cdot E_n - A = \begin{pmatrix} tE_{r_1} - A_1 & & & * \\ & tE_{r_2} - A_2 & & \\ & & \ddots & \\ 0 & & & tE_{r_k} - A_k \end{pmatrix}$$

Daher ist dann

$$P_A(t) = \det(t \cdot E_n - A) = \prod_{i=1}^{k} \det(t \cdot E_{r_i} - A_i) = \prod_{i=1}^{k} P_{A_i}(t).$$

iv)(Drehmatrix) Sei $\alpha \in \mathbb{R}$. Die Matrix

$$A := \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

hat das charakteristische Polynom

$$P_A(t) = (\cos\alpha - t)(\cos\alpha - t) + \sin^2\alpha = t^2 - 2\cos\alpha \cdot t + 1 = (t - e^{i\alpha})(t - e^{-i\alpha}).$$

v) (Typische Klausurmatrix) $A := \begin{pmatrix} 1 & 3 & 1 \\ -1 & 5 & 2 \\ 0 & 0 & 3 \end{pmatrix} \in M(3 \times 3, \mathbb{Q})$.

$$\begin{aligned} P_A(t) &= (-1)^3 \cdot \det \begin{pmatrix} 1-t & 3 & 1 \\ -1 & 5-t & 2 \\ 0 & 0 & 3-t \end{pmatrix} = (t-3) \cdot \det \begin{pmatrix} 1-t & 3 \\ -1 & 5-t \end{pmatrix} \\ &= (t-3) \cdot ((1-t)(5-t) + 3) = (t-3)(t^2 - 6t + 8) \\ &= (t-3)(t-2)(t-4). \end{aligned}$$

vi) (Telefonmatrix) Die Matrix

$$\mathrm{Tel} := \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}$$

hat das charakteristische Polynom

$$P_{\mathrm{Tel}}(t) = ... = t^3 - 12t^2 - 18t = t(t - 6 + 3\sqrt{6})(t - 6 - 3\sqrt{6}).$$

Definition 8.3 a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein Element $v \in V - \{0\}$ heißt *Eigenvektor* von f, falls es ein $\lambda \in K$ gibt mit

$$f(v) = \lambda \cdot v.$$

Ein solches λ heißt *Eigenwert* von f.

b) (Lemma) Für jedes $\lambda \in K$ ist die Menge

$$\mathrm{Eig}(f, \lambda) := \{v \in V \mid f(v) = \lambda \cdot v\} = \ker(f - \lambda \cdot \mathrm{id})$$

offenbar ein Untervektorraum von V. Und offenbar ist λ genau dann ein Eigenwert von f, wenn $\mathrm{Eig}(f, \lambda) \neq \{0\}$ ist.
(Definition) Der Vektorraum $\mathrm{Eig}(f, \lambda)$ heißt *Eigenraum* von f bezüglich λ.

c) Eigenwerte, Eigenvektoren und Eigenräume einer Matrix $A \in M(n \times n, K)$ sind die Eigenwerte, Eigenvektoren und Eigenräume des Endomorphismus

$$l_A : M(n \times 1, K) \to M(n \times 1, K), \quad b \mapsto A \cdot b.$$
$$(l_A = \text{ Linksmultiplikation mit } A)$$

Äquivalent und konkreter: $v \in M(n \times 1, K) - \{0\}$ heißt *Eigenvektor* von A, falls es ein $\lambda \in K$ gibt mit

$$A \cdot v = \lambda \cdot v.$$

Ein solches λ heißt *Eigenwert* von A. Für *jedes* $\lambda \in K$ ist

$$\begin{aligned} \mathrm{Eig}(A, \lambda) &:= \{v \in M(n \times 1, K) \mid A \cdot v = \lambda \cdot v\} \\ &= \ker(l_A - \lambda \cdot \mathrm{id}) = \mathrm{Lös}(A - \lambda \cdot E_n, 0) \end{aligned}$$

der Eigenraum von A zum Wert λ. Es ist offenbar ein Untervektorraum von $M(n \times 1, K)$. Und offenbar ist $\lambda \in K$ genau dann ein Eigenwert von A, wenn $\mathrm{Eig}(A, \lambda) \neq \{0\}$ ist.

Satz 8.4 (Bestimmung von Eigenwerten und Eigenvektoren)
a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Dann gilt folgende Äquivalenz:

$$\lambda \in K \text{ ist ein Eigenwert von } f \iff P_f(\lambda) = 0.$$

b) Sei $A \in M(n \times n, K)$. Dann gilt folgende Äquivalenz:

$$\lambda \in K \text{ ist ein Eigenwert von } A \iff P_A(\lambda) = 0.$$

Beweis: Zuerst b):

$$\begin{aligned} & \lambda \text{ ist ein Eigenwert von } A \\ \iff & \mathrm{Eig}(A, \lambda) \neq \{0\} \\ \iff & \mathrm{Lös}(A - \lambda \cdot E_n, 0) \neq \{0\} \\ \overset{6.3}{\iff} & A - \lambda \cdot E_n \text{ ist nicht invertierbar} \\ \overset{7.7}{\iff} & 0 = \det(\lambda \cdot E_n - A) = P_A(\lambda). \end{aligned}$$

Nun a):

$$\begin{aligned} & \lambda \text{ ist ein Eigenwert von } f \\ \iff & \mathrm{Eig}(f, \lambda) \neq \{0\} \\ \iff & \ker(f - \lambda \cdot \mathrm{id}) \neq \{0\} \\ \iff & f - \lambda \cdot \mathrm{id} \text{ ist nicht invertierbar} \\ \overset{*}{\iff} & 0 = \det(\lambda \cdot \mathrm{id} - f) \overset{!}{=} P_f(\lambda). \end{aligned}$$

$\overset{*}{\iff}$ benutzt: ein Endomorphismus (hier $\lambda \cdot \mathrm{id} - f$) ist invertierbar $\iff$ seine Determinante ist Null. Das folgt mit 7.19 und 5.11 d) aus der analagen Aussage 7.7 für Matrizen. $\square$

Beispiele 8.5 Es werden dieselben Beispiele wie in 8.2 betrachtet.
i) (Obere Dreiecksmatrix) Wegen $P_A(t) = \prod_{i=1}^n (t - a_{ii})$ sind die Eigenwerte $a_{11}, ..., a_{nn}$. Aber die Eigenvektoren sind schwerer zu bestimmen. Sofort sieht man nur den Eigenvektor $e_1 = (1, 0..., 0)^{tr}$ zum Eigenwert a_{11}. Wenn mehrere a_{ii} übereinstimmen, gibt es zu ihnen eventuell nur einen Eigenvektor, siehe ii). Nur im Fall einer Diagonalmatrix sieht man sofort viele Eigenvektoren: e_i ist Eigenvektor zum Eigenwert a_{ii}. Mehr dazu kommt in 8.6.

ii) (Jordanblock) Hier ist $P_A(t) = (t - \lambda)^n$, also hat man nur einen Eigenwert. Tatsächlich ist hier

$$A - \lambda \cdot E_n = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix},$$

also $\operatorname{rang}(A - \lambda \cdot E_n) = n - 1$, also $\dim \operatorname{Eig}(A, \lambda) = 1$. Genauer:

$$\operatorname{Eig}(A, \lambda) = \text{Lös}(A - \lambda \cdot E_n, 0) = K \cdot e_1.$$

iii) (Blockdiagonalmatrix) Wenn A_i einen Eigenwert λ und Eigenvektor $v \in M(r_i \times 1, K)$ hat, so ist $(0, ..., 0, v^{tr}, 0, ...0)^{tr} \in M(n \times 1, K)$ (mit den richtigen Anzahlen von Nullen) ein Eigenvektor von A mit Eigenwert λ.

iv) (Vgl. Beispiel 5.9 ii)) Sei $\alpha \in \mathbb{R}$. Die Linksmultiplikation l_A mit der Matrix

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

beschreibt eine Drehung des $\mathbb{R}$-Vektorraums $M(2 \times 1, \mathbb{R}) \cong \mathbb{R}^2$ um 0 um den Winkel α. Skizze in der Vorlesung.
Für $\alpha \notin \pi \cdot \mathbb{Z}$ hat sie offenbar keinen Eigenvektor und keinen Eigenwert. Das paßt dazu, daß

$$P_A(t) = t^2 - 2\cos\alpha \cdot t + 1 = (t - e^{i\alpha})(t - e^{-i\alpha})$$

keine reelle Nullstelle hat.
Aber die Linksmultiplikation mit derselben Matrix auf dem $\mathbb{C}$-Vektorraum $M(2 \times 1, \mathbb{C}) \cong \mathbb{C}^2$ hat die Eigenwerte $e^{i\alpha}$ und $e^{-i\alpha}$ und die Eigenvektoren $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ und $\begin{pmatrix} 1 \\ i \end{pmatrix}$:

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \mp i \end{pmatrix} = \begin{pmatrix} \cos\alpha \pm i\sin\alpha \\ \sin\alpha \mp i\cos\alpha \end{pmatrix} = e^{\pm i\alpha} \cdot \begin{pmatrix} 1 \\ \mp i \end{pmatrix}.$$

v) (Typische Klausurmatrix) Wegen $P_A(t) = (t-3)(t-2)(t-4)$ sind die Eigenwerte $2, 3$ und 4. Die Eigenräume sind:

$$Eig(A,2) = \text{Lös}(A - 2 \cdot E_3, 0) = \text{Lös}(\begin{pmatrix} -1 & 3 & 1 \\ -1 & 3 & 2 \\ 0 & 0 & 1 \end{pmatrix}, 0) = \text{span}_{\mathbb{Q}}(\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}),$$

$$Eig(A,3) = \text{Lös}(A - 3 \cdot E_3, 0) = \text{Lös}(\begin{pmatrix} -2 & 3 & 1 \\ -1 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}, 0) = \text{span}_{\mathbb{Q}}(\begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix}),$$

$$Eig(A,4) = \text{Lös}(A - 4 \cdot E_3, 0) = \text{Lös}(\begin{pmatrix} -3 & 3 & 1 \\ -1 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}, 0) = \text{span}_{\mathbb{Q}}(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}).$$

vi) (Telefonmatrix) Wegen

$$P_{\text{Tel}}(t) = ... = t^3 - 12t^2 - 18t = t(t - 6 + 3\sqrt{6})(t - 6 - 3\sqrt{6})$$

hat Tel als Matrix in $M(3 \times 3, \mathbb{R})$ die drei Eigenwerte 0, $6 - 3\sqrt{6}$ und $6 + 3\sqrt{6}$. Mit einiger Mühe rechnet man aus: Die Eigenräume sind hier alle eindimensional und werden erzeugt durch die drei Eigenvektoren

$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 5 \\ 8 - 3\sqrt{6} \\ 11 - 3\sqrt{6} \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 5 \\ 8 + 3\sqrt{6} \\ 11 + 3\sqrt{6} \end{pmatrix}.$$

Als Matrix in $M(3 \times 1, \mathbb{Q})$ hat Tel dagegen nur den Eigenwert 0 und den zugehörigen eindimensionalen Eigenraum $\text{Eig}(Tel, 0)$, der vom Eigenvektor $(1, -2, 1)^{tr}$ erzeugt wird.

Definition/Lemma 8.6 *a) (Definition) Ein Endomorphismus $f : V \to V$ eines K-Vektorraums V mit $\dim V = n \in \mathbb{N}$ heißt* diagonalisierbar*, falls es eine Basis von V aus Eigenvektoren von f gibt.*
b) (Lemma) Ist $\mathcal{B} = (b_1, ..., b_n)$ eine solche Basis mit $f(b_i) = \lambda_i$, so ist

$$M(\mathcal{B}, f, \mathcal{B}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n, \end{pmatrix}.$$

c) (Definition) Eine Matrix $A \in M(n \times n, K)$ heißt diagonalisierbar*, falls es es eine Basis von $M(n \times 1, K)$ aus Eigenvektoren von A gibt.*
d) (Lemma) Eine Matrix A ist diagonalierbar genau dann, wenn es eine Matrix $B \in GL(n, K)$ gibt mit

$$B^{-1} \cdot A \cdot B = \text{ Diagonalmatrix.}$$

Beweis: a)+c): Definitionen. b): Klar.
d) Die Spalten von B bilden eine Basis von Eigenvektoren. $\square$

Definition 8.7 a) Eine Matrix $A \in M(n \times n, K)$ ist in *Jordannormalform*, falls sie die Gestalt hat

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix}.$$

Hier sind

$$A_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in M(r_i \times r_i, K)$$

Jordanblöcke; alle anderen Einträge von A sind Null. Der Jordanblock A_i hat die *Größe* $r_i \times r_i$. Natürlich ist $n = r_1 + ... + r_k$. Offenbar ist

$$P_A(t) = \prod_{i=1}^{k} (t - \lambda_i)^{r_i}.$$

b) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Er *läßt sich in Jordannormalform bringen*, falls eine Basis $\mathcal{B}$ von V existiert, so daß $M(\mathcal{B}, f, \mathcal{B})$ in Jordannormalform ist.
Dann heißt die Matrix $M(\mathcal{B}, f, \mathcal{B})$ eine Jordannormalform von f.
c) Er heißt *nilpotent*, falls ein $m \in \mathbb{N}$ existiert mit $f^m := f \circ ... \circ f = 0$.

Satz 8.8 *(In dieser Vorlesung ohne Beweis)*
a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. *Er läßt sich genau dann in Jordannormalform bringen, wenn sein charakteristisches Polynom in Linearfaktoren zerfällt, d.h. wenn es $\lambda_1, ..., \lambda_k \in K$ und $r_1, ..., r_k \in \mathbb{N}$ gibt mit*

$$P_f(t) = \prod_{i=1}^{k} (t - \lambda_i)^{r_i}.$$

b) Je zwei Jordannormalformen eines Endomorphismus haben die gleichen Jordanblöcke; nur in der Reihenfolge der Blöcke können sie sich unterscheiden.

c) Ein nilpotenter Endomorphismus $f : V \to V$ mit $\dim V = n \in \mathbb{N}$ *hat charakteristisches Polynom $P_f(t) = t^n$. (Spezialfall von a):) Er läßt sich in*

Jordannormalform bringen. Die Jordanblöcke A_i haben dann alle in der Diagonalen $\lambda_i = 0$.

d) (Spezialfall von a) wegen des Fundamentalsatzes der Algebra) Im Fall $K = \mathbb{C}$ läßt sich jeder Endomorphismus in Jordannormalform bringen. (Im Fall $K = \mathbb{R}$ gilt das nicht!)

Bemerkungen 8.9 i) Eine Diagonalmatrix ist eine Matrix in Jordannormalform, bei der alle Blöcke die Größe 1×1 haben.

ii) Wenn ein Endomorphismus diagonalisierbar ist, ist das etwas ganz besonderes. Viele Endomorphismen lassen sich nicht diagonalisieren. Bei $K = \mathbb{C}$ hat man aber immerhin stets Jordannormalformen, bei $K = \mathbb{R}$ oder $K = \mathbb{Q}$ oft nicht mal das.
Tatsächlich (aber nicht in dieser Vorlesung) gibt es Verallgemeinerungen der Jordannormalform für Endomorphismen über einem beliebigen Körper.

iii) (Beweis in einer Algebra-Vorlesung) Zu jedem Körper K gibt es einen (bis auf Isomorphie eindeutigen) kleinsten Körper $\overline{K} \supset K$, in dem jedes Polynom in $K[t]$ in Linearfaktoren zerfällt. Er heißt *algebraischer Abschluß* von K. Es ist $\overline{\mathbb{R}} = \mathbb{C}$. Aber es ist $\overline{\mathbb{Q}} \neq \mathbb{C}$.

iv) Eine Matrix $A \in M(n \times n, K)$ liegt auch in $M(n \times n, \overline{K})$. Daher gibt es nach Satz 8.8 a) eine invertierbare Matrix $B \in M(n \times n, \overline{K})$, so daß $B^{-1} \cdot A \cdot B \in M(n \times n, \overline{K})$ in Jordannormalform ist. Das ist auch für das Verständnis der Normalformen von Endomorphismen über K nützlich.

v) Es gibt verschiedene Beweise von Satz 8.8. Konzeptionell gute Beweise benutzen die Theorie der *Moduln über Hauptidealringen.*

Bemerkung 8.10 Ein Endomorphismus $f : V \to V$ eines K-Vektorraums V induziert eine Abbildung

$$\begin{aligned} \Phi_f : K[t] &\to \mathrm{End}(V), \\ \sum_{i=0}^{m} a_i t^i &\mapsto \sum_{i=0}^{m} a_i f^i. \end{aligned}$$

Hier ist $f^0 := \mathrm{id}_V$. Die Abbildung Φ_f ist ein Ringhomomorphismus und ein Vektorraumhomomorphismus. Das Bild $\Phi_f(K[t])$ ist ein *kommutativer* Unterring mit Eins von $\mathrm{End}(V)$.
Sei $\dim V = n \in \mathbb{N}$. Dann ist $\dim \mathrm{End}(V) = n^2$. Es ist $\dim K[t]_{\leq n^2} = n^2 + 1$. Daher ist schon die Einschränkung von Φ_f auf $K[t]_{\leq n^2}$ nicht injektiv. Satz 8.11 sagt, daß $P_f(t)$ im Kern liegt.

Satz 8.11 *(Satz von Cayley und Hamilton) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Es ist*

$$P_f(f) = 0 \in \mathrm{End}(V).$$

Beweis: Sei $\mathcal{B} = (b_1, ..., b_n)$ irgendeine Basis von V und $M(\mathcal{B}, f, \mathcal{B}) =: A = (a_{ij})$. Es ist $f(b_j) = \sum_{i=1}^n a_{ij} b_i$. Wir wenden auf jeden Eintrag der Matrix

$$C(t) := (A - t \cdot E_n)^{tr} \in M(n \times n, K[t])$$

die Abbildung Φ_f an und erhalten die Matrix

$$C(f) = \begin{pmatrix} a_{11} \cdot \mathrm{id}_V - f & a_{21} \cdot \mathrm{id}_V & \cdots & a_{n1} \cdot \mathrm{id}_V \\ a_{12} \cdot \mathrm{id}_V & a_{22} \cdot \mathrm{id}_V - f & & a_{n2} \cdot \mathrm{id}_V \\ \vdots & & \ddots & \\ a_{1n} \cdot \mathrm{id}_V & & & a_{nn} \cdot \mathrm{id}_V - f \end{pmatrix} \in M(n \times n, \Phi_f(K[t])).$$

Es ist $\det C(f) = (-1)^n P_f(f)$. Multipliziert man von rechts den Spaltenvektor $(b_1, ..., b_n)^{tr} \in M(n \times 1, V)$ dran, so erhält man

$$C(f) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} -f(b_1) + \sum_{i=1}^n a_{i1} b_i \\ \vdots \\ -f(b_n) + \sum_{i=1}^n a_{in} b_i \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$C(f)^\sharp$ ist die Komplementärmatrix zu $C(f)$ von Definition 7.13. Mit Satz 7.15 c) erhält man

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = C(f)^\sharp \cdot C(f) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \det C(f) & & 0 \\ & \ddots & \\ 0 & & \det C(f) \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Also ist $P_f(f)(b_i) = 0$ für alle $i = 1, ..., n$; also $P_f(f)(v) = 0$ für alle $v \in V$. Daher ist $P_f(f) = 0 \in \mathrm{End}(V)$. □

Bemerkung 8.12 Bei einem Endomorphismus $f : V \to V$ mit einer Jordannormalform $M(\mathcal{B}, f, \mathcal{B})$ wie in 8.7 a) sieht man Satz 8.11 einfacher. In dem Fall sei $\rho_i := r_1 + ... + r_{i-1} \in \mathbb{N} \cup \{0\}$. Zum Block A_i der Jordannormalform "gehören" die Basisvektoren $b_{\rho_i+1}, ..., b_{\rho_i+r_i}$ der Basis $\mathcal{B} = (b_1, ..., b_n)$. Der Endomorphismus $f - \lambda_i \cdot \mathrm{id}$ operiert auf ihnen so:

$$b_{\rho_i+r_i} \mapsto b_{\rho_i+r_i-1} \mapsto ... \mapsto b_{\rho_i+1} \mapsto 0.$$

Also ist er auf $\sum_{j=1}^{r_i} K \cdot b_{\rho_i+j}$ nilpotent mit $(f - \lambda_i \cdot \mathrm{id})^{r_i}(b_{\rho_i+j}) = 0$. Daher ist

$$P_f(f)(b_{\rho_i+j}) = \left(\prod_{l \neq i} (f - \lambda_l \cdot \mathrm{id})^{r_l} \right) \circ (f - \lambda_i \cdot \mathrm{id})^{r_i}(b_{\rho_i+j}) = 0.$$

Definition/Lemma 8.13 *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$.*

a) (Definition) Der Kern der Abbildung Φ_f von Bemerkung 8.10 enthält ein unitäres Polynom $M_f(t)$ mit kleinstem positiven Grad. Es heißt Minimalpolynom *von f.*
b) (Lemma) Das Minimalpolynom $M_f(t)$ teilt jedes andere Element von $\ker \Phi_f$, insbesondere teilt es das charakteristische Polynom $P_f(t)$.
c) (Lemma) Läßt sich f in Jordannormalform bringen, so ist (mit den Notationen von 8.7 a))

$$M_f(t) = \prod_{\lambda \text{ Eigenwert}} (t-\lambda)^{r(\lambda)}$$
$$\text{mit} \qquad r(\lambda) = \max(r_i \mid \lambda_i = \lambda).$$

Beweis: a) Definition.
b) Würde $M_f(t)$ ein Element $g(t)$ von $\ker \Phi_f$ nicht teilen, so würde man mit Polynomdivision $g(t) = q(t)M_f(t) + r(t)$ ein Polynom $r(t)$ von kleinerem Grad erhalten, das auch in $\ker \Phi_f$ liegen würde. Widerspruch.
c) 8.12 hat gezeigt, daß $f - \lambda_i \cdot \mathrm{id}$ so auf $V_i := \sum_{j=1}^{r_i} K \cdot b_{\rho_i+j}$ operiert:

$$b_{\rho_i+r_i} \mapsto b_{\rho_i+r_i-1} \mapsto \ldots \mapsto b_{\rho_i+1} \mapsto 0.$$

Daher ist $(f-\lambda\cdot\mathrm{id})^{r_i}$ die kleinste Potenz von $f-\lambda\cdot\mathrm{id}$, die V_i auf Null abbildet. Und $(f-\lambda\cdot\mathrm{id})^{r(\lambda)}$ ist die kleinste Potenz, die den Raum $\sum_{i \text{ mit } \lambda_i=\lambda} V_i$ auf Null abbildet. Daher ist $M_f(t)$ wie oben beschrieben. □

Beispiel 8.14 Läßt sich $f : V \to V$ in Jordannormalform bringen mit $\dim V = 15$ und sechs Jordanblöcken mit λ_i und r_i wie in der Tabelle (mit α, β und $\gamma \in K$ paarweise verschieden)

λ_i	α	α	β	β	β	γ
r_i	1	5	2	3	3	1

so ist

$$\begin{aligned} P_f(t) &= (t-\alpha)^6(t-\beta)^8(t-\gamma), \\ M_f(t) &= (t-\alpha)^5(t-\beta)^3(t-\gamma). \end{aligned}$$

Satz/Definition 8.15 *a) (Notation) Seien $V_i \subset V$ $(i = 1, \ldots, k)$ Untervektorräume eines Vektorraums V. Der von ihnen erzeugte Untervektorraum von V ist*

$$\sum_{i=1}^{k} V_i := \{v_1 + \ldots + v_k \mid v_i \in V_i\}.$$

b) (Satz) Folgende drei Bedingungen sind äquivalent:

α) *Jedes Element von $\sum_{i=1}^k V_i$ lässt sich auf eindeutige Weise als Linearkombination von Elementen der V_i schreiben, d.h.*

$$v_1 + ...v_k = \widetilde{v}_1 + ... + \widetilde{v}_k \text{ mit } v_i, \widetilde{v}_i \in V_i \Rightarrow \text{ für alle } i \quad v_i = \widetilde{v}_i.$$

β) *Beliebige Basen der V_i bilden zusammen eine Basis von $\sum_{i=1}^k V_i$.*

γ) *Für alle $i = 1, ..., k$ ist $V_i \cap (\sum_{j\neq i} V_j) = \{0\}$.*

c) (Definition) Wenn die drei äquivalenten Bedingungen in b) erfüllt sind, ist der von den V_i erzeugte Untervektorraum die direkte Summe *der V_i; Notation: $\bigoplus_{i=1}^k V_i$ oder $V_1 \oplus ... \oplus V_k$.*

Beweis von b): $\alpha) \iff \beta)$: Leicht, Übung.
$\alpha) \Rightarrow \gamma)$: Wäre $v \in V_i \cap (\sum_{j\neq i} V_j)$ und $v \neq 0$, so ließe sich v auf zwei Weisen als Linearkombination der Elemente der V_k schreiben, einmal als $v \in V_i$, einmal als $v \in \sum_{j\neq i} V_j$.
$\gamma) \Rightarrow \alpha)$: Indirekt. Sei α) nicht erfüllt; sei $\sum_j v_j = \sum_j \widetilde{v}_j$ mit $v_j, \widetilde{v}_j \in V_j$ und $v_i \neq \widetilde{v}_i$ für irgendein (mindestens ein) i. Dann ist

$$0 \neq v_i - \widetilde{v}_i = \sum_{j\neq i} (\widetilde{v}_j - v_j) \in V_i \cap (\sum_{j\neq i} V_j).$$

Also ist auch γ) nicht erfüllt. □

Beispiel 8.16 Sei $V = K^4$ mit Standardbasis (e_1, e_2, e_3, e_4). Sei

$$V_1 := Ke_1 + Ke_2,\ V_2 := Ke_3,\ V_3 := Ke_1 + Ke_4,\ V_4 := Ke_4.$$

Dann hat man direkte Summen $V_1 \oplus V_2,\ V_1 \oplus V_4,\ V_2 \oplus V_3,\ V_2 \oplus V_4, V_1 \oplus V_2 \oplus V_4$. Aber die Summen $V_1 + V_3,\ V_3 + V_4$ und alle Summen, die diese enthalten $(V_1 + V_2 + V_3,\ V_1 + V_3 + V_4,\ V_2 + V_3 + V_4, V_1 + V_2 + V_3 + V_4)$, sind nicht direkt.

Definition 8.17 a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein Element $v \in V - \{0\}$ heißt *verallgemeinerter Eigenvektor* oder *Hauptvektor*, falls es ein $\lambda \in K$ und ein $m \in \mathbb{N}$ gibt mit

$$(f - \lambda \cdot \mathrm{id})^m(v) = 0.$$

b) Für jedes $\lambda \in K$ ist die Menge

$$\mathrm{Hau}(f, \lambda) := \{v \in V \mid \text{ es gibt ein } m \in \mathbb{N} \text{ mit } (f - \lambda \cdot \mathrm{id})^m(v) = 0\}$$

offenbar ein Untervektorraum von V. Offenbar ist $\mathrm{Hau}(f, \lambda) \supset \mathrm{Eig}(f, \lambda)$. Also gilt

$$\lambda \text{ Eigenwert } \iff \mathrm{Eig}(f, \lambda) \neq \{0\} \iff \mathrm{Hau}(f, \lambda) \neq \{0\}.$$

$\mathrm{Hau}(f, \lambda)$ heißt *Hauptraum* von f bezüglich λ.

Satz 8.18 *(In dieser Vorlesung ohne Beweis) Sei* $f: V \to V$ *ein Endomorphismus eines* K*-Vektorraums mit* $\dim_K V = n \in \mathbb{N}$*. Sein charakteristisches Polynom zerfalle in Linearfaktoren,*

$$P_f(t) = \prod_{\lambda \text{ Eigenwert}} (t-\lambda)^{d(\lambda)} \qquad (\text{mit } d(\lambda) \in \mathbb{N}).$$

Dann gilt:

$$V = \bigoplus_{\lambda \text{ Eigenwert}} \operatorname{Hau}(f,\lambda) \quad \text{und} \quad \dim \operatorname{Hau}(f,\lambda) = d(\lambda).$$

Bemerkungen 8.19 i) Läßt sich f in Jordannormalform bringen, so ist mit den Notationen von 8.7 und 8.12

$$\operatorname{Hau}(f,\lambda) = \bigoplus_{i \text{ mit } \lambda_i = \lambda} V_i.$$

In dem Fall ist der Satz leicht zu sehen.

ii) Tatsächlich ist Satz 8.18 eine Vorstufe für den Satz 8.8 a) über die Jordannormalform. Aus Satz 8.18 und dem Spezialfall Satz 8.8 c) folgt Satz 8.8 a) leicht.

9 Bilinearformen

In diesem Kapitel bezeichnet K irgendeinen Körper.

Definition 9.1 a) Seien V und W K-Vektorräume. Eine Abbildung $\phi : V \times W \to K$ ist eine *Bilinearform*, falls gilt:

i) für jedes $y \in W$ ist folgende Abbildung linear:

$$V \to K, \quad x \mapsto \phi(x, y).$$

ii) für jedes $x \in V$ ist folgende Abbildung linear:

$$W \to K, \quad y \mapsto \phi(x, y).$$

b) Eine *Bilinearform auf einem Vektorraum* V ist eine Bilinearform

$$\phi : V \times V \to K.$$

[Die meisten interessanten Bilinearformen sind von diesem Typ, d.h. $V = W$.]
c) Eine *symmetrische Bilinearform* ist eine Bilinearform $\phi : V \times V \to K$ mit

$$\phi(x, y) = \phi(y, x).$$

d) Eine (*alternierende* oder) *schiefsymmetrische Bilinearform* ist eine Bilinearform $\phi : V \times V \to K$ mit

$$\phi(x, y) = -\phi(y, x).$$

Beispiele 9.2 i) Das *Standard-Skalarprodukt* auf $M(n \times 1, K)$:

$$\begin{aligned} \langle\ ,\ \rangle : M(n \times 1, K) \times M(n \times 1, K) &\to K, \\ (x, y) &\mapsto \langle x, y \rangle := \sum_{i=1}^{n} x_i \cdot y_i = x^{tr} \cdot y. \end{aligned}$$

$\langle\ ,\ \rangle$ ist eine symmetrische Bilinearform. Sie wird in Kapitel 10 studiert.

ii) Auf dem $\mathbb{R}$-Vektorraum $\mathcal{C}^0([0,1], \mathbb{R})$ der stetigen Funktionen auf dem Intervall $[0, 1]$ (vgl. Beispiel 3.3 d)) hat man die symmetrische Bilinearform

$$\begin{aligned} \phi_{int} : \mathcal{C}^0([0,1], \mathbb{R}) \times \mathcal{C}^0([0,1], \mathbb{R}) &\to \mathbb{R}, \\ (f, g) &\mapsto \int_0^1 f(x)g(x)\mathrm{d}x. \end{aligned}$$

iii) Eine *Linearform* auf einem K-Vektorraum V ist ein Vektorraumhomomorphismus $f : V \to K$.

[Das paßt zum Begriff *Bilinearform* in Definition 9.1.]
Der *Dualraum* zu V ist der K-Vektorraum

$$V^* := \mathrm{Hom}(V, K).$$

Die folgende Bilinearform auf $V^* \times V$ ist zugleich natürlich, etwas abstrakt und ziemlich trivial:

$$\phi : V^* \times V \to K, \quad (f, x) \mapsto \phi(f, x) := f(x).$$

iv) (Matrizen und Bilinearformen, 1. Teil)
Eine $(m \times n)$-Matrix $A = (a_{ij}) \in M(m \times n, K)$ definiert (offenbar) eine Bilinearform auf $M(m \times 1, K) \times M(n \times 1, K)$ durch

$$\begin{aligned} \mathrm{Bil}_A : M(m \times 1, K) \times M(n \times 1, K) &\to K \\ (x, y) &\mapsto x^{tr} \cdot A \cdot y = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \cdot a_{ij} \cdot y_j. \end{aligned}$$

Beispiel 9.2 iv) ist ein Spezialfall von Definition/Satz 9.3 a).

Definition/Satz 9.3 *(Matrizen und Bilinearformen, 2. Teil)*
Seien V und W endlich-dimensionale K-Vektorräume; sei $\mathcal{B} = (b_1, ..., b_m)$ eine Basis von V und $\mathcal{C} = (c_1, ..., c_n)$ eine Basis von W.

a) (Definition) Eine $(m \times n)$-Matrix $A = (a_{ij}) \in M(m \times n, K)$ definiert offenbar eine Bilinearform auf $V \times W$ durch

$$\begin{aligned} \mathrm{Bil}_{A,\mathcal{B},\mathcal{C}} : V \times W &\to K \\ \left(\sum_{i=1}^{m} x_i b_i, \sum_{j=1}^{n} y_j c_j\right) &\mapsto \begin{pmatrix} x_1 \cdots x_m \end{pmatrix} \cdot A \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \cdot a_{ij} \cdot y_j. \end{aligned}$$

b) (Definition) Für jede Bilinearform $\phi : V \times W \to K$ ist $M(\mathcal{B}, \phi, \mathcal{C}) \in M(m \times n, K)$ die Matrix mit den Einträgen

$$(M(\mathcal{B}, \phi, \mathcal{C}))_{ij} := \phi(b_i, c_j).$$

c) (Satz) Die Menge

$$\mathrm{Bil}(V, W) := \{\phi : V \times W \to K \text{ Bilinearform}\}$$

ist ein K-Vektorraum, und die Abbildung

$$M(m \times n, K) \to \mathrm{Bil}(V, W), \quad A \mapsto \mathrm{Bil}_{A,\mathcal{B},\mathcal{C}},$$

ist ein Vektorraumisomorphismus. Der inverse Vektorraumisomorphismus ist

$$\mathrm{Bil}(V,W) \to M(m \times n, K), \quad \phi \mapsto M(\mathcal{B}, \phi, \mathcal{C}).$$

d) (Satz) (Transformationsverhalten von Matrizen zu Bilinearformen bezüglich Basiswechsel)
Sei $\mathcal{B}'$ eine weitere Basis von V und $\mathcal{C}'$ eine weitere Basis von W. Dann ist

$$M(\mathcal{B}', \phi, \mathcal{C}') = M(\mathcal{B}, \mathcal{B}')^{tr} \cdot M(\mathcal{B}, \phi, \mathcal{C}) \cdot M(\mathcal{C}, \mathcal{C}').$$

Beweis: a) Definition. $\mathrm{Bil}_{A,\mathcal{B},\mathcal{C}}$ ist bilinear, weil in jedem Summanden der Summe $\sum_{i=1}^m \sum_{j=1}^n x_i \cdot a_{ij} \cdot y_j$ die x_k linear auftreten und die y_l linear auftreten.

b) Definition.

c) [Die Aussagen sind verwandt zu denen von Satz 5.5 b) und Satz 5.11 b).]

1.Teil: $\mathrm{Bil}(V,W)$ *ist ein K-Vektorraum.*
Die Menge $\mathrm{Abb}(V \times W, K)$ aller Abbildungen von $V \times W$ nach K ist ein Vektorraum mit der *punktweisen Addition und Multiplikation* (Beispiel 3.3 c)): Für $\phi_1, \phi_2 \in \mathrm{Abb}(V,W)$ und $\lambda \in K$ ist

$$\begin{aligned} (\phi_1 + \phi_2)(x,y) &:= \phi_1(x,y) + \phi_2(x,y), \\ (\lambda \cdot \phi_1)(x,y) &:= \lambda \cdot \phi_1(x,y). \end{aligned}$$

Es reicht nun zu zeigen, daß die Teilmenge $\mathrm{Bil}(V,W) \subset \mathrm{Abb}(V \times W, K)$ ein Untervektorraum ist, d.h. daß aus $\phi_1, \phi_2 \in \mathrm{Bil}(V,W)$ auch $\phi_1 + \phi_2 \in \mathrm{Bil}(V,W)$ und $\lambda \cdot \phi_1 \in \mathrm{Bil}(V,W)$ folgt. Aber weil ϕ_1 und ϕ_2 bei festem $y \in W$ linear in $x \in V$ sind, gilt das auch für $\phi_1 + \phi_2$ und $\lambda \cdot \phi_1$. Analog folgt, daß $\phi_1 + \phi_2$ und $\lambda \cdot \phi_1$ bei festem $x \in V$ linear in $y \in W$ sind. Also sind sie bilinear und in $\mathrm{Bil}(V,W)$.

2. Teil: *Vektorraumisomorphismen.*
Behauptungen: *Für eine $(m \times n)$-Matrix A ist*

$$M(\mathcal{B}, \mathrm{Bil}_{A,\mathcal{B},\mathcal{C}}, \mathcal{C}) = A.$$

Für $\phi \in \mathrm{Bil}(V,W)$ ist

$$\mathrm{Bil}_{M(\mathcal{B},\phi,\mathcal{C}),\mathcal{B},\mathcal{C}} = \phi.$$

Beweis der Behauptungen:

$$(M(\mathcal{B}, \mathrm{Bil}_{A,\mathcal{B},\mathcal{C}}, \mathcal{C}))_{kl} = \mathrm{Bil}_{A,\mathcal{B},\mathcal{C}}(b_k, c_l) = \sum_{i=1}^m \sum_{j=1}^n \delta_{ik} \cdot a_{ij} \cdot \delta_{jl} = a_{kl}.$$

und

$$\begin{aligned} \mathrm{Bil}_{M(\mathcal{B},\phi,\mathcal{C}),\mathcal{B},\mathcal{C}}\Big(\sum_{i=1}^m x_i b_i, \sum_{j=1}^n y_j c_j\Big) &\overset{\text{Def.}}{=} \sum_{i=1}^m \sum_{j=1}^n x_i \cdot \phi(b_i, c_j) \cdot y_j \\ &\overset{\phi \text{ bilinear}}{=} \phi\Big(\sum_{i=1}^m x_i b_i, \sum_{j=1}^n y_j c_j\Big). \qquad (\Box) \end{aligned}$$

Daher sind die beiden Abbildungen $\phi \mapsto M(\mathcal{B}, \phi, \mathcal{C})$ und $A \mapsto \mathrm{Bil}_{A,\mathcal{B},\mathcal{C}}$ invers zueinander und beide bijektiv.
Die Abbildung $\phi \mapsto M(\mathcal{B}, \phi, \mathcal{C})$ ist ein Vektorraumhomomorphismus wegen der punktweisen Definition von $\phi_1 + \phi_2$ und $\lambda \cdot \phi_1$.

d) Bei

$$b'_j = \sum_{i=1}^{m} \beta_{ij} b_i \quad \text{und} \quad c'_l = \sum_{k=1}^{n} \gamma_{kl} c_k$$

ist

$$M(\mathcal{B}, \mathcal{B}') = (\beta_{ij}), \qquad M(\mathcal{C}, \mathcal{C}') = (\gamma_{kl})$$

und

$$\begin{aligned}
(M(\mathcal{B}', \phi, \mathcal{C}'))_{jl} &= \phi(b'_j, c'_l) \\
&= \sum_{i=1}^{m} \sum_{k=1}^{n} \beta_{ij} \cdot \phi(b_i, c_k) \cdot \gamma_{kl} \\
&= \left(M(\mathcal{B}, \mathcal{B}')^{tr} \cdot M(\mathcal{B}, \phi, \mathcal{C}) \cdot M(\mathcal{C}, \mathcal{C}')\right)_{jl}.
\end{aligned}$$

□

Beispiele 9.4 i) Polynome geben auf $[0,1]$ stetige Abbildungen. Daher ist $\mathbb{R}[t] \subset \mathcal{C}^0([0,1], \mathbb{R})$. Der Untervektorraum $\mathbb{R}[t]_{\leq 3}$ hat die Basis $\mathcal{B} = (1, t, t^2, t^3)$. Wegen $\int_0^1 t^k \mathrm{d}t = \frac{1}{k+1}$ hat ϕ_{int} von Beispiel 9.2 ii) auf $\mathbb{R}[t]_{\leq 3}$ bezüglich dieser Basis die Matrix

$$M(\mathcal{B}, \phi_{int}, \mathcal{B}) = (\int_0^1 t^{i+j-2} \mathrm{d}t)_{i,j=1,\ldots,4} = (\frac{1}{i+j-1})_{i,j=1,\ldots,4} = \begin{pmatrix} \frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

ii) Auf $M(2 \times 1, K)$ hat man die Standardbasis $\mathcal{B} = (e_1, e_2) = (\binom{1}{0}, \binom{0}{1})$ und die Basis $\mathcal{B}' = (\binom{1}{2}, \binom{1}{-1})$. Die Basiswechselmatrix ist natürlich $T := M(\mathcal{B}, \mathcal{B}') = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$. Die symmetrische Bilinearform ϕ_2 mit

$$M(\mathcal{B}, \phi, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

erfüllt

$$\begin{aligned}
M(\mathcal{B}', \phi_2, \mathcal{B}') &= T^{tr} \cdot M(\mathcal{B}, \phi_2, \mathcal{B}) \cdot T = T^{tr} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot T \\
&= \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 3 \\ 3 & 0 \end{pmatrix}.
\end{aligned}$$

iii) Die Bilinearform ϕ_3 auf $M(2 \times 1, K)$ mit

$$M(\mathcal{B}, \phi, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$$

ist nicht symmetrisch, wegen $\phi_3(e_2, e_1) = 2 \neq 0 = \phi_3(e_1, e_2)$. Es ist $\phi_3(v, e_2) = 0$ für alle $v \in M(2 \times 1, K)$.

Bemerkungen 9.5 a) Häufig ist eine Bilinearform $\phi : V \times W \to K$ durch eine Matrix $M(\mathcal{B}, \phi, \mathcal{C})$ gegeben, wobei $\mathcal{B}$ und $\mathcal{C}$ irgendwelche (bekannten) Basen von V und W sind. Dann sagen Satz 9.3 c) und die Formel in 9.3 a) mit $A = M(\mathcal{B}, \phi, \mathcal{C})$, wie man $\phi(X, Y)$ für $X \in V$ und $Y \in W$ ausrechnet: Man bestimmt die Koeffizienten x_i und y_j in $X = \sum_{i=1}^m x_i b_i$ und $Y = \sum_{j=1}^n y_j c_j$ und berechnet das Matrizenprodukt $x^{tr} \cdot A \cdot y$, wobei $x = (x_1 \cdots x_m)^{tr}$ und $y = (y_1 \cdots y_n)^{tr}$ ist.

b) Das Transformationsverhalten in 9.3 d) von Matrizen $M(\mathcal{B}, \phi, \mathcal{C})$ zu Bilinearformen $\phi : V \times W \to K$ bei Basiswechseln ist anders als das von Matrizen $M(\mathcal{B}, f, \mathcal{C})$ zu Homomorphismen $f : W \to V$.
Erinnerung an Bemerkung 5.12 i):

$$M(\mathcal{B}', f, \mathcal{C}') = M(\mathcal{B}, \mathcal{B}')^{-1} \cdot M(\mathcal{B}, f, \mathcal{C}) \cdot M(\mathcal{C}, \mathcal{C}').$$

Bei Bilinearformen hat man links $(..)^{tr}$ statt $(..)^{-1}$.

c) (Matrizen und Bilinearformen, 3. Teil)
Sei $\phi : V \times V \to K$ eine Bilinearform auf V mit Matrix $T := M(\mathcal{B}, \phi, \mathcal{B})$ bezüglich einer Basis $\mathcal{B} = (b_1, ..., b_n)$ von V. Es gilt offenbar:

$$\begin{array}{rcl} \phi \text{ ist symmetrisch} & \iff & T \text{ ist symmetrisch, d.h. } T^{tr} = T; \\ \phi \text{ ist schiefsymmetrisch} & \iff & T \text{ ist schiefsymmetrisch, d.h. } T^{tr} = -T. \end{array}$$

Definition 9.6 Sei $\phi : V \times V \to K$ eine symmetrische oder schiefsymmetrische Bilinearform.

a) (Definition) Zwei Elemente $x, y \in V$ heißen orthogonal, falls $\phi(x, y) = 0$ ist. Notation: $x \perp y$.

b) (Triviales Lemma) Weil ϕ symmetrisch oder schiefsymmetrisch ist, ist $x \perp y \iff y \perp x$. [Sonst müßte man rechts-orthogonal und links-orthogonal unterscheiden.]

c) (Definition) Sei $U \subset V$ ein Untervektorraum. Der orthogonale Untervektorraum $U^\perp \subset V$ ist

$$U^\perp := \{x \in V \mid \text{für alle } \; y \in U \; \text{ ist } x \perp y\}.$$

d) (Definition) Das Radikal von ϕ ist $\mathrm{Rad}(\phi) := V^\perp \subset V$.

e) (Definition) ϕ heißt nichtentartet, falls $\mathrm{Rad}(\phi) = \{0\}$. Falls $\mathrm{Rad}(\phi) \neq \{0\}$ ist, heißt ϕ entartet.
f) (Definition) Ein Element $v \in V$ heißt isotrop, falls es erfüllt:

$$\phi(v, v) = 0 \quad \text{und} \quad v \neq 0.$$

g) (Triviales Lemma) Offenbar ist jedes Element von $\mathrm{Rad}(\phi) - \{0\}$ isotrop.

Beispiele 9.7 i) Es gibt in $\mathcal{C}^0([0,1], \mathbb{R})$ bezüglich ϕ_{int} (Beispiel 9.2 ii)) kein isotropes Element, denn $f^2(x) \geq 0$ für alle $x \in [0,1]$, und $f \in \mathcal{C}^0([0,1], \mathbb{R})$ erfüllt (Beweis: Analysis)

$$\int_0^1 f^2(x)\mathrm{d}x = 0 \iff f = 0.$$

Insbesondere ist $\mathrm{Rad}(\phi_{int}) = \{0\}$. Also ist ϕ_{int} nichtentartet.
ii) Die Bilinearform ϕ_2 von Beispiel 9.4 ii) ist nichtentartet (Übung oder Lemma 9.8 c) und $\det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -1 \neq 0$). Aber es gibt isotrope Elemente:

$$\begin{aligned} x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq 0 \text{ ist isotrop} \quad &\iff \quad \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \\ &\iff \quad x_1^2 - x_2^2 = 0. \end{aligned}$$

Also sind zum Beispiel $\binom{1}{1}$ und $\binom{1}{-1}$ isotrop. [Bei $K = \mathbb{F}_2$ sind sie natürlich gleich.] Bei $U := K \cdot \binom{1}{1}$ ist

$$U^\perp = U.$$

Lemma 9.8 *Sei V ein endlich-dimensionaler K-Vektorraum, $\phi : V \times V \to K$ eine symmetrische oder schiefsymmetrische Bilinearform, $\mathcal{B} = (b_1, ..., b_n)$ eine Basis von V und $A := M(\mathcal{B}, \phi, \mathcal{B})$.*

a) Sei $U \subset V$ ein Unterverktorraum. Sei $\mathcal{C} = (c_1, ..., c_k)$ eine Basis von U und $\Gamma = (\gamma_{ij})$ die $(n \times k)$-Matrix mit $c_j = \sum_{i=1}^n \gamma_{ij} b_i$. Dann ist

$$\begin{aligned} &\operatorname{rang} \Gamma = k \geq \operatorname{rang} \left(\Gamma^{tr} \cdot A\right), \\ &U^\perp = \{\sum_{i=1}^n x_i b_i \mid \Gamma^{tr} \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0\}, \\ &\dim U^\perp = n - \operatorname{rang}(\Gamma^{tr} \cdot A) \geq n - k = n - \dim U, \\ &\dim U^\perp + \dim U \geq n. \end{aligned}$$

b) Insbesondere ist

$$\begin{aligned} &\mathrm{Rad}(\phi) = \{\sum_{i=1}^n x_i b_i \mid A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0\}, \\ &\dim \mathrm{Rad}(\phi) = n - \operatorname{rang} A. \end{aligned}$$

c) ϕ ist nichtentartet $\iff \det A \neq 0$.

d) Wenn ϕ nichtentartet ist, ist

$$\dim U^\perp + \dim U = n.$$

e) Wenn es keine isotropen Vektoren gibt ($\Rightarrow \mathrm{Rad}(\phi) = \{0\} \Rightarrow \phi$ *ist nichtentartet), ist $U \cap U^\perp = \{0\}$ und*

$$U \oplus U^\perp = V.$$

Beweis: a) Die Ungleichung $\operatorname{rang}\Gamma \geq \operatorname{rang}(\Gamma^{tr} \cdot A)$ ist klar. Es ist

$$\phi(c_j, \sum_{i=1}^n x_i b_i) = (\gamma_{1j} \cdots \gamma_{nj}) \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

wegen Bemerkung 9.5 a) bzw. Satz 9.3. Daraus folgt die Beschreibung von $U^\perp$. Der Rest ist klar.
b) Das folgt aus a) und der Definition des Radikals.
c) Das folgt aus b).
d) In dem Fall ist $\operatorname{rang}\Gamma = \operatorname{rang}(\Gamma^{tr} \cdot A)$ und $\dim U^\perp = n - k = n - \dim U$.
e) Jeder Vektor in $U \cap U^\perp - \{0\}$ wäre isotrop. Wenn es keine isotropen Vektoren gibt, ist $U \cap U^\perp = \{0\}$, also $U + U^\perp = U \oplus U^\perp$. Wegen d) ist dann $U \oplus U^\perp = V$. □

Bemerkung 9.9 Ist $A \in M(n \times n, K)$ (schief)symmetrisch und $T \in M(n \times n, K)$ beliebig, so ist auch $T^{tr} \cdot A \cdot T$ (schief)symmetrisch. Denn wegen $(B \cdot C)^{tr} = C^{tr} \cdot B^{tr}$ (Lemma 4.15 b)) ist

$$(T^{tr} \cdot A \cdot T)^{tr} = T^{tr} \cdot A^{tr} \cdot T = (-)T^{tr} \cdot A \cdot T.$$

10 Euklidische Vektorräume

In diesem Kapitel wird hauptsächlich mit dem Körper $\mathbb{R}$ gearbeitet.

Definition 10.1 a) Eine symmetrische Bilinearform $\phi : V \times V \to \mathbb{R}$ auf einem $\mathbb{R}$-Vektorraum V heißt *positiv definit*, falls gilt:

$$\phi(v, v) > 0 \text{ für alle } v \in V - \{0\}.$$

Sie heißt *negativ definit*, falls gilt:

$$\phi(v, v) < 0 \text{ für alle } v \in V - \{0\}.$$

Sie heißt *positiv semidefinit* [bzw. *negativ semidefinit*], falls nur gilt:

$$\phi(v, v) \geq 0 \text{ [bzw } \phi(v, v) \leq 0] \text{ für alle } v \in V.$$

b) Eine symmetrische Matrix A heißt positiv oder negativ (semi)definit, wenn die Bilinearform Bil_A auf $M(n \times 1, \mathbb{R})$ es ist.

Definition 10.2 a) Ein $\mathbb{R}$-Vektorraum V zusammen mit einer positiv definiten symmetrischen Bilinearform $\phi : V \times V \to \mathbb{R}$ heißt *Euklidischer Vektorraum*. Die Bilinearform ϕ ist sein *Skalarprodukt*.
b) Dann ist die *Norm* (oder *Länge*) eines Vektors $v \in V$

$$\|v\| := \sqrt{\phi(v, v)} \geq 0.$$

Bemerkungen 10.3 i) Die Formulierung "zusammen mit" ist nicht präzise. Genaugenommen ist ein Euklidischer Vektorraum das Paar (V, ϕ) Aber der Vektorraum V wird als das primäre Objekt angesehen.
ii) Bei "natürlich" gegebenen Skalarprodukten gibt es in der Literatur neben $\phi(x, y)$ eine ganze Reihe von gebräuchlichen Notationen, zum Beispiel $x \cdot y$, $\langle x, y \rangle$, (x, y).
iii) In einem Euklidischen Vektorraum V gibt es offensichtlich keine isotropen Vektoren. Daher erfüllen ein Untervektorraum $U \subset V$ und der orthogonale Untervektorraum $U^\perp$

$$U + U^\perp = U \oplus U^\perp.$$

Im Fall $\dim V < \infty$ gilt wegen Lemma 9.8 e)

$$U \oplus U^\perp = V.$$

Im Fall $\dim V = \infty$ gilt das im allgemeinen nicht, siehe 10.4 iv) und v). Allerdings gilt es, falls $\dim U < \infty$ ist, siehe Satz 10.17.

iv) Eine symmetrische Matrix $A \in M(n \times n, \mathbb{R})$ ist nach Definition positiv definit, falls

$$(x_1 \cdots x_n) \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} > 0 \quad \text{für alle} \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M(n \times 1, \mathbb{R}) - \{0\} \text{ ist.}$$

Am Ende des Kapitels werden zwei andere Charakterisierungen positiv definiter Matrizen gegeben.

Beispiele 10.4 i) Der wichtigste Euklidische Vektorraum ist der $\mathbb{R}^n$ mit dem Standard-Skalarprodukt ϕ_{st},

$$\phi_{st} : ((x_1, ..., x_n), (y_1, ..., y_n)) \overset{\phi_{st}}{\mapsto} (x_1 \cdots x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i.$$

ii) Die Vektorräume $\mathbb{R}^n$ und $M(1 \times n, \mathbb{R})$ waren in Definition 4.1 c) identifiziert worden, aber nicht $\mathbb{R}^n$ und $M(n \times 1, \mathbb{R})$. Das Standard-Skalarprodukt auf $M(n \times 1, \mathbb{R})$ war in Beispiel 9.2 i) $\langle\ ,\rangle$ genannt worden: $\langle\ ,\rangle = \mathrm{Bil}_{E_n}$,

$$(x, y) \mapsto \langle x, y\rangle = x^{tr} \cdot E_n \cdot y = x^{tr} \cdot y = \sum_{i=1}^{n} x_i y_i.$$

iii) Der $\mathbb{R}$-Vektorraum $\mathcal{C}^0([0,1], \mathbb{R})$ mit der Bilinearform ϕ_{int},

$$\phi_{int} : (f, g) \mapsto \int_0^1 f(x)g(x)\mathrm{d}x,$$

ist ein Euklidischer Vektorraum wegen (Beweis: Analysis)

$$\int_0^1 f^2(x)\mathrm{d}x > 0, \qquad \text{falls } f \neq 0.$$

iv) Sei $V := C^0([0,1], \mathbb{R})$, $U := \mathbb{R}[t]$. Dann ist $U \subsetneq V$.

Behauptung: $U^\perp = \{0\}$. Also ist $U + U^\perp = U \oplus U^\perp = U \subsetneq V$.

(Zitat aus der Analysis:) *Approximationssatz von Weierstrass:*

$$\forall\ f \in V \quad \forall\ \varepsilon > 0 \quad \exists\ g \in U \quad \forall\ t \in [0,1] \quad |f(t) - g(t)| < \varepsilon.$$

Indirekter Beweis der Behauptung: Sei $U^\perp \supsetneq \{0\}$ und $f \in U^\perp - \{0\}$. Sei $a := \phi_{int}(f, f) > 0$. Sei $b := \max(|f(t)|\,|\, t \in [0,1])$. Sei $\varepsilon := \frac{a}{2b}$. Wähle ein

$g \in U$ mit $\forall\ t \in [0,1] \quad |f(t) - g(t)| < \varepsilon$. Dann ist

$$\begin{aligned}
\phi_{int}(f,g) &= \int_0^1 f(t)g(t)\mathrm{d}t = \int_0^1 f(t)\left(f(t) - (f(t) - g(t))\right)\mathrm{d}t \\
&= \int_0^1 f^2(t)\mathrm{d}t - \int_0^1 f(t)(f(t) - g(t))\mathrm{d}t \\
&\geq \phi_{int}(f,f) - \int_0^1 |f(t)| \cdot |f(t) - g(t)|\mathrm{d}t \\
&\geq a - b \cdot \varepsilon = a - \frac{a}{2} = \frac{a}{2},
\end{aligned}$$

also $f \notin U^\perp$, Widerspruch. □

v) Aber falls V ein *Hilbertraum* und $U \subset V$ ein *abgeschlossener Unterraum* ist, gilt $U \oplus U^\perp = V$ (Definitionen und Beweis in der Funktionalanalysis).

vi) Längen und Winkel im $\mathbb{R}^2$:

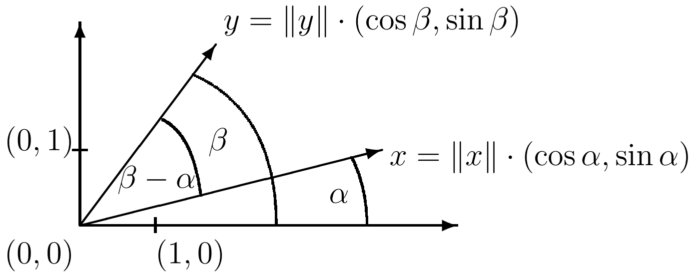
Nach dem Satz von Pythagoras ist die Länge des Vektors $(x_1, x_2) \in \mathbb{R}^2$ tatsächlich

$$\|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2}.$$

Sind x und $y \in \mathbb{R}^2 - \{0\}$, so haben $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ Länge 1, und es gibt eindeutige α und $\beta \in [0, 2\pi)$ mit

$$x = \|x\| \cdot (\cos\alpha, \sin\alpha) \text{ und } y = \|y\| \cdot (\cos\beta, \sin\beta).$$

Skizze:



Es ist

$$\begin{aligned}
\phi_{st}(x,y) &= \|x\| \cdot \|y\| \cdot (\cos\alpha\cos\beta + \sin\alpha\sin\beta) \\
&= \|x\| \cdot \|y\| \cdot \cos(\beta - \alpha) \text{ (nach einem Additionstheorem).}
\end{aligned}$$

Es gibt ein eindeutiges $\gamma \in [0, \pi]$ mit $\cos(\beta - \alpha) = \cos\gamma$. Es ist $\gamma \equiv \pm(\beta - \alpha)$ mod 2π. Man kann γ als *den Winkel* zwischen x und y auffassen. Er wird eindeutig bestimmt durch $\gamma \in [0, \pi]$ und

$$\phi_{st}(x,y) = \|x\| \cdot \|y\| \cdot \cos(\gamma).$$

v) Aufgrund der Cauchy-Schwarzschen Ungleichung (Satz 10.5) wird es möglich sein, diesen Begriff eines Winkels zwischen zwei Vektoren auf beliebige Euklidische Vektorräume zu verallgemeinern (Definition 10.6). Danach wird auch der Begriff der Länge im allgemeinen diskutiert werden (Definition 10.7 und Korollar 10.8).

Satz 10.5 *Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ. Dann gilt die Cauchy-Schwarzsche Ungleichung:*

$$|\phi(x,y)| \leq \|x\| \cdot \|y\| \text{ für } x, y \in V,$$

und Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Beweis: Falls $x = 0$ oder $y = 0$ ist, sind beide Seiten gleich Null, und x und y sind linear abhängig.
Seien $x \neq 0$ und $y \neq 0$. Dann haben die Vektoren $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ Länge 1. Es ist (mit $\pm$ = plus oder minus)

$$\begin{aligned} 0 &\leq \phi(\frac{x}{\|x\|} \pm \frac{y}{\|y\|}, \frac{x}{\|x\|} \pm \frac{y}{\|y\|}) \\ &= \phi(\frac{x}{\|x\|}, \frac{x}{\|x\|}) \pm 2\phi(\frac{x}{\|x\|}, \frac{y}{\|y\|}) + \phi(\frac{y}{\|y\|}, \frac{y}{\|y\|}) \\ &= 1 \pm 2\frac{\phi(x,y)}{\|x\| \cdot \|y\|} + 1, \end{aligned}$$

also

$$-(\pm 1)\phi(x,y) \leq \|x\| \cdot \|y\|,$$

also

$$|\phi(x,y)| \leq \|x\| \cdot \|y\|$$

Bei Gleichheit ist $-(\pm 1)\phi(x,y) = \|x\| \cdot \|y\|$, also

$$0 = \phi(\frac{x}{\|x\|} \pm \frac{y}{\|y\|}, \frac{x}{\|x\|} \pm \frac{y}{\|y\|}),$$

also

$$\frac{x}{\|x\|} \pm \frac{y}{\|y\|} = 0.$$

Bei $x \neq 0$ und $y \neq 0$ ist das äquivalent dazu, daß x und y linear abhängig sind. □

Definition 10.6 Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ. Wegen der Cauchy-Schwarzschen Ungleichung gibt es zu je zwei Vektoren x und $y \in V - \{0\}$ ein eindeutiges $\gamma \in [0, \pi]$ mit

$$\phi(x,y) = \|x\| \cdot \|y\| \cdot \cos\gamma.$$

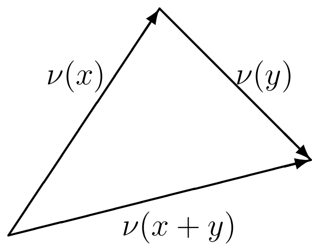
Dieses γ heißt der *Winkel* zwischen x und y.
Bemerkung: Nun sind x und y orthogonal ($x \perp y$) genau dann, wenn der Winkel zwischen ihnen $\pi/2$ ist. Sie sind linear abhängig genau dann, wenn der Winkel zwischen ihnen 0 oder π ist.

Definition 10.7 Sei V ein $\mathbb{R}$-Vektorraum. Eine *Norm* auf V ist eine Abbildung $\nu : V \to \mathbb{R}_0^+$ mit den Eigenschaften:

(N1) $\nu(x) = 0 \iff x = 0$.

(N2) $\nu(\lambda \cdot x) = |\lambda| \cdot \nu(x)$ für $\lambda \in \mathbb{R}$, $x \in V$.

(N3) $\nu(x + y) \leq \nu(x) + \nu(y)$ (Dreiecksungleichung).



Korollar 10.8 *(zur Cauchy-Schwarzschen Ungleichung, Satz 10.5) Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ. Die Längen-Abbildung*

$$\|.\| : V \to \mathbb{R}_0^+, \ x \mapsto \sqrt{\phi(x,x)}$$

ist eine Norm.

Beweis: (N1) und (N2) sind klar. (N3) folgt aus

$$\begin{aligned} \|x+y\|^2 &= \phi(x+y, x+y) = \phi(x,x) + 2\phi(x,y) + \phi(y,y) \\ &= \|x\|^2 + 2\phi(x,y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

□

Definition/Lemma 10.9 *a) (Definition) Den Begriff des Abstandes erfaßt man abstrakt mit dem Begriff einer Metrik:*
Sei X irgendeine nichtleere Menge. Eine Abbildung $d : X \times X \to \mathbb{R}_0^+$ heißt Metrik, falls sie erfüllt:

(M1) $d(x,y) = 0 \iff x = y$.

(M2) $d(y,x) = d(x,y)$.

(M3) $d(x,z) \leq d(x,y) + d(y,z)$ *(Dreiecksungleichung)*

b) Sei V *ein* $\mathbb{R}$*-Vektorraum mit einer Norm* $\|.\| : V \to \mathbb{R}_0^+$*. Dann ist die Abbildung* $d : V \times V \to \mathbb{R}_0^+$ *mit* $d(x,y) := \|x - y\|$ *eine Metrik auf* V*.*

Beweis: a) Definition. b) Klar. □

Bemerkung 10.10 Also hat man auf einem Euklidischen Vektorraum den Begriff eines Winkels nach 10.6 und den Begriff des Abstandes nach 10.9 und 10.8.

Definition 10.11 Sei V ein endlich-dimensionaler Euklidischer Vektorraum mit Skalarprodukt $\phi : V \times V \to \mathbb{R}$.
Eine Basis $(b_1, ..., b_n)$ heißt *Orthogonalbasis*, falls $b_i \perp b_j$ für $i \neq j$ ist.
Eine Basis $(b_1, ..., b_n)$ heißt *Orthonormalbasis* oder kürzer *ON-Basis*, falls $b_i \perp b_j$ für $i \neq j$ und $\|b_i\| = 1$ für alle i ist. Die gleiche Bedingung kürzer geschrieben:

$$\phi(b_i, b_j) = \delta_{ij}.$$

Satz 10.12 *Sei* V *ein endlich-dimensionaler Euklidischer Vektorraum mit Skalarprodukt* $\phi : V \times V \to \mathbb{R}$*.*
a) Er besitzt eine Orthonormalbasis.
b) (Gram-Schmidtsches Orthogonalisierungsverfahren)
Aus einer gegebenen Basis $(a_1, ..., a_n)$ *erhält man in folgender Weise eine Orthogonalbasis* $(b_1, ..., b_n)$*.*

$$\begin{aligned}
b_1 &:= a_1, \\
b_2 &:= a_2 - \frac{\phi(a_2, b_1)}{\phi(b_1, b_1)} \cdot b_1, \\
b_3 &:= a_3 - \sum_{i=1}^{2} \frac{\phi(a_3, b_i)}{\phi(b_i, b_i)} \cdot b_i, \\
&\vdots \\
b_n &:= a_n - \sum_{i=1}^{n-1} \frac{\phi(a_n, b_i)}{\phi(b_i, b_i)} \cdot b_i.
\end{aligned}$$

c) Aus einer Orthogonalbasis $(b_1, ..., b_n)$ *erhält man eine Orthonormalbasis* $(c_1, ..., c_n)$ *durch Normieren:*

$$c_i := \frac{b_i}{\|b_i\|}.$$

d) Ist $(b_1, ..., b_n)$ *eine Orthogonalbasis von* V *und* $x \in V$ *beliebig, so ist*

$$x = \sum_{i=1}^{n} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot b_i.$$

Im Spezialfall einer Orthonormalbasis $(c_1, ..., c_n)$ *ist dann natürlich*

$$x = \sum_{i=1}^{n} \phi(x, c_i) \cdot c_i.$$

Beweis: a) folgt aus b) und c).
b) Induktiv.
Induktionsanfang: $b_1 = a_1$.
Induktionsannahme: für ein $j \in \{2, ..., n\}$ ist $(b_1, ..., b_{j-1})$ eine Orthogonalbasis des von $(a_1, ..., a_{j-1})$ erzeugten Untervektorraums.
Induktionsschritt: Die Definition von b_j für $j = 2, ..., n$ zeigt, daß b_j orthogonal zu allen b_k mit $k < j$ ist:

$$\begin{aligned} b_j &:= a_j - \sum_{i=1}^{j-1} \frac{\phi(a_j, b_i)}{\phi(b_i, b_i)} \cdot b_i, \\ \phi(b_j, b_k) &= \phi(a_j, b_k) - \sum_{i=1}^{j-1} \frac{\phi(a_j, b_i)}{\phi(b_i, b_i)} \cdot \phi(b_i, b_k) \\ &= \phi(a_j, b_k) - \frac{\phi(a_j, b_k)}{\phi(b_k, b_k)} \cdot \phi(b_k, b_k) = 0. \end{aligned}$$

Daher ist $(b_1, ..., b_j)$ eine Orthogonalbasis des von $(a_1, ..., a_j)$ erzeugten Untervektorraums.
c) Klar.
d) Weil $(b_1, ..., b_n)$ eine Basis von V ist, gibt es $x_i \in \mathbb{R}$ mit $x = \sum_{i=1}^{n} x_i \cdot b_i$. Es ist

$$\phi(x, b_k) = \sum_{i=1}^{n} x_i \cdot \phi(b_i, b_k) = x_k \cdot \phi(b_k, b_k).$$

□

Beispiel 10.13 (Legendre-Polynome) Der $\mathbb{R}$-Vektorraum $\mathbb{R}[x]$ mit der Bilinearform

$$\phi_{int,2} : (f, g) \mapsto \int_{-1}^{1} f(x)g(x)\mathrm{d}x$$

ist ein Euklidischer Vektorraum. Denn die Bilinearform $\phi_{int,2}$ ist offenbar symmetrisch und positiv definit:

$$\int_{-1}^{1} f^2(x)\mathrm{d}x > 0 \qquad \text{falls } f \neq 0.$$

Eine Basis von ihm ist $(a_i)_{i\in\mathbb{N}}$ mit $a_i = x^{i-1}$.
Eine Basis des Untervektorraums $\mathbb{R}[x]_{\leq n}$ ist $(a_1, a_2, ..., a_{n+1}) = (1, x, ..., x^n)$.
Das Orthogonalisierungsverfahren oben liefert eine Familie $(b_i)_{i\in\mathbb{N}}$ und eine Familie $(c_i)_{i\in\mathbb{N}}$, so daß $(b_1, ..., b_{n+1})$ die Orthogonalbasis von $\mathbb{R}[x]_{\leq n}$ ist,

die man aus $(a_1, ..., a_{n+1})$ konstruiert, und $(c_1, ..., c_{n+1})$ die zugehörige ON-Basis. Die Polynome c_i heißen *Legendre-Polynome*. Die ersten drei werden hier ausgerechnet.

$$\begin{aligned}
b_1 &= a_1 = 1, \\
\|b_1\|^2 &= \|1\|^2 = \phi_{int,2}(1,1) = \int_{-1}^{1} 1 \cdot 1 \mathrm{d}x = 2, \\
c_1 &= \frac{b_1}{\|b_1\|} = \frac{1}{\sqrt{2}}, \\
b_2 &= a_2 - \frac{\phi_{int,2}(a_2, b_1)}{\phi_{int,2}(b_1, b_1)} \cdot b_1 \\
&= x - \left(\int_{-1}^{1} x \cdot 1 \mathrm{d}x\right) \cdot \frac{1}{2} \cdot 1 = x, \\
\|b_2\|^2 &= \phi_{int,2}(x, x) = \int_{-1}^{1} x^2 \mathrm{d}x = \frac{2}{3}, \\
c_2 &= \frac{b_2}{\|b_2\|} = \sqrt{\frac{3}{2}} \cdot x, \\
b_3 &= a_3 - \sum_{i=1}^{2} \frac{\phi_{int,2}(a_3, b_i)}{\phi_{int,2}(b_i, b_i)} \cdot b_i \\
&= x^2 - \left(\int_{-1}^{1} x^2 \cdot 1 \mathrm{d}x\right) \cdot \frac{1}{2} \cdot 1 - \left(\int_{-1}^{1} x^2 \cdot x \mathrm{d}x\right) \cdot \frac{3}{2} \cdot x \\
&= x^2 - \frac{1}{3}, \\
\|b_3\|^2 &= \int_{-1}^{1} (x^2 - \frac{1}{3})^2 \mathrm{d}x = ... = \frac{8}{45}, \\
c_3 &= \frac{b_3}{\|b_3\|} = \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3}) = \frac{3}{2}\sqrt{\frac{5}{2}}x^2 - \frac{1}{2}\sqrt{\frac{5}{2}}.
\end{aligned}$$

Ohne Beweis: man kann für das Legendre-Polynom c_{k+1} $(k \in \mathbb{N} \cup \{0\})$ die allgemeine Formel zeigen:

$$c_{k+1} = \frac{\sqrt{2k+1}}{k!\sqrt{2^{2k+1}}} \cdot \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k \left((x^2-1)^k\right).$$

Korollar 10.14 *(QR-Zerlegung, Korollar zum Gram-Schmidtschen Orthogonalisierungsverfahren)*
Sei $k \leq n$ und sei $A \in M(n \times k, \mathbb{R})$ eine Matrix mit linear unabhängigen Spalten.
Dann existieren Matrizen $Q \in M(n \times k, \mathbb{R})$ und $R \in M(k \times k, \mathbb{R})$ mit den Eigenschaften:

i)
$$A = Q \cdot R;$$

ii) die Spalten von Q bilden eine ON-Basis des von ihnen erzeugten Untervektorraums von $M(n \times 1, \mathbb{R})$ bezüglich des Standard-Skalarproduktes;

iii) R ist eine invertierbare obere Dreiecksmatrix.

Beweis: Die Spalten von A werden mit $a_1, ..., a_k$ bezeichnet. Der von ihnen erzeugte Untervektorraum $V := \text{span}\langle a_1, ..., a_k \rangle$ von $M(n \times 1, \mathbb{R})$ ist zusammen mit der Einschränkung des Standard-Skalarproduktes $\langle \ , \rangle$ ein Euklidischer Vektorraum. Das Gram-Schmidtsche Orthogonalisierungsverfahren liefert eine ON-Basis $c_1, ...c_k$.
Die Basiswechselmatrix $M((a_1, ..., a_k), (c_1, ..., c_k))$ und die inverse Matrix $M((c_1, ..., c_k), (a_1, ..., a_k)) =: R$ sind invertierbare obere Dreiecksmatrizen. Die Matrix mit den Spalten $c_1, ..., c_k$ wird Q genannt. Dann gelten $A = Q \cdot R$ und die anderen Eigenschaften. □

Beispiel 10.15 A ist die 4×3-Matrix

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

mit den Spalten a_1, a_2, a_3. Durch Hingucken findet man die Matrix B, deren Spalten die Orthogonalbasis von $\text{span}\langle a_1, a_2, a_3 \rangle$ bilden, die man mit Gram-Schmidt bekommt:

$$B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wieder im Kopf findet man die Matrix Q, deren Spalten c_1, c_2, c_3 man aus denen von B durch Normieren erhält,

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Ebenfalls im Kopf findet man die obere Dreiecksmatrix R mit $(a_1, a_2, a_3) = (c_1, c_2, c_3) \cdot R$, d.h. mit $A = Q \cdot R$,

$$R = \begin{pmatrix} 2 & 1 & 1 \\ 0 & \sqrt{2} & \sqrt{2} \\ 0 & 0 & \sqrt{3} \end{pmatrix}.$$

Bemerkungen 10.16 i) $x = (x_1 \cdots x_n)^{tr} \in M(n \times 1, \mathbb{R})$ hat Norm 1 bezüglich des Standard-Skalarproduktes genau dann, wenn $\sum_{i=1}^n x_i^2 = 1$ ist. Dann gilt $|x_i| \leq 1$ für alle Einträge von x.
ii) Die QR-Zerlegung ist interessant für numerische Verfahren, weil die Matrix Q robust gegen kleine Störungen ist (unter anderem wegen i)) und weil man Störungen bei oberen Dreiecksmatrizen gut kontrollieren kann.

Definition/Satz 10.17 *Sei V ein Euklidischer Vektorraum mit Skalarprodukt ϕ. Sei $U \subset V$ ein endlich-dimensionaler Untervektorraum.*
a)

$$V = U \oplus U^\perp.$$

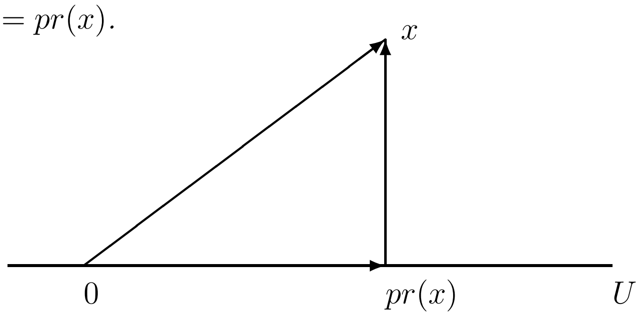
b) (Definition) Für jedes $x \in V$ gibt es eindeutige $y \in U$ und $z \in U^\perp$ mit $x = y + z$. Die Abbildung

$$pr_U : V \to U, \qquad x \mapsto y$$

heißt Projektion *von V auf U. Sie wird pr genannt, wenn klar ist, welcher Unterraum U gemeint ist.*
c) Man kann $pr(x)$ folgendermaßen ausrechnen. Man wählt eine Orthogonalbasis $(b_1, ..., b_k)$ von U. Dann ist für jedes $x \in V$

$$pr(x) = \sum_{i=1}^{k} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot b_i.$$

d) Für jedes $x \in V$ gibt es genau ein $\widetilde{y} \in U$, so daß $\|x - \widetilde{y}\|$ minimal ist. Es ist $\widetilde{y} = pr(x)$.



Beweis: a) Aus ϕ positiv definit folgt $U \cap U^\perp = \{0\}$, also $U + U^\perp = U \oplus U^\perp$. Daher reicht es, $U + U^\perp = V$ zu zeigen.

Sei $(b_1, ..., b_k)$ eine Orthogonalbasis von U. Man definiert eine Abbildung $\widetilde{pr} : V \to U$ durch

$$\widetilde{pr}(x) := \sum_{i=1}^{k} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot b_i.$$

Dann ist

$$\begin{aligned} \phi(x - \widetilde{pr}(x), b_j) &= \phi(x, b_j) - \sum_{i=1}^{k} \frac{\phi(x, b_i)}{\phi(b_i, b_i)} \cdot \phi(b_i, b_j) \\ &= \phi(x, b_j) - \phi(x, b_j) = 0, \end{aligned}$$

also $x - \widetilde{pr}(x) \perp b_j$, also $x - \widetilde{pr}(x) \perp U$, also $x - \widetilde{pr}(x) \in U^\perp$. Mit

$$x = \widetilde{pr}(x) + (x - \widetilde{pr}(x))$$

und $\widetilde{pr}(x) \in U$ folgt $V = U + U^\perp$.
b) Definition.
c) Der Beweis von a) zeigt $pr = \widetilde{pr}$.
d) Sei $x \in V$ und $y \in U$. Es ist

$$\begin{aligned} \|x - y\|^2 &= \|(x - pr(x)) + (pr(x) - y)\|^2 \\ &= \|x - pr(x)\|^2 + 2\phi(x - pr(x), pr(x) - y) + \|pr(x) - y\|^2 \\ &= \|x - pr(x)\|^2 + \|pr(x) - y\|^2 \end{aligned}$$

denn $x - pr(x) \in U^\perp$, $pr(x) - y \in U$. Der Abstand $\|x - y\|$ ist also genau dann minimal, wenn $pr(x) = y$ ist. □

Bemerkungen 10.18 i) Satz 10.17 c)+d) hat viele Anwendungen. Normalerweise sucht man innerhalb eines endlich-dimensionalen Unterraums $U \subset V$ das Element $\widetilde{y}$, das ein Element $x \in V$ am besten approximiert.
Das ist offensichtlich relevant in praktischen Problemen aus der Wirtschaft. Innerhalb der Mathematik ist V häufig ein Vektorraum von gewissen Funktionen und U ist ein Unterraum von spezielleren Funktionen, z.B.

$$V = \mathcal{C}^0([-1, 1], \mathbb{R}) \quad \text{und} \quad U = \mathbb{R}[t]_{\leq n} \text{ für ein } n \in \mathbb{N},$$

oder

$$\begin{aligned} V &= \{f : \mathbb{R} \to \mathbb{R} \text{ stetig } \mid f(x + 2\pi) = f(x)\} \quad \text{und} \\ U &= \{f = a_0 + \sum_{p=1}^{n} (a_p \cos(px) + b_p \sin(px)) \mid a_p, b_p \in \mathbb{R}\}. \end{aligned}$$

Das erste Beispiel führt zu den Legendre-Polynomen, das zweite zu trigonometrischen Polynomen und Fourier-Reihen.

ii) Satz 10.19 passt eigentlich besser in Kapitel 9. Er betrifft nicht nur Skalarprodukte, sondern beliebige symmetrische Bilinearformen und symmetrische Matrizen über fast beliebigen Körpern K (Einschränkung: $\operatorname{char} K \neq 2$).
Er bereitet die letzten drei Sätze 10.22, 10.23 und 10.24 vor. Sie behandeln reelle symmetrische Matrizen, die nicht notwendig positiv definit sind. Man kann ihre Aussagen auch übersetzen in Aussagen über reelle symmetrische Bilinearformen. Die Beweise dieser drei Sätze werden in dieser Vorlesung nicht gegeben (aber in der Vorlesung LA 2b in den hinteren 7 Wochen des FS 2011).

Satz 10.19 *(Ein allgemeines Orthogonalisierungsverfahren)*
Sei K ein Körper mit $\operatorname{char}(K) \neq 2$.
a) (Matrix-Version) Zu einer symmetrischen Matrix $A \in M(n \times n, K)$ gibt es eine invertierbare Matrix $T \in GL(n, K)$ und Skalare $\alpha_1, ..., \alpha_n \in K$ mit

$$T^{tr} \cdot A \cdot T = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

Zur Berechnung gibt es einen Algorithmus (Details im Beweis und in Beispiel 10.20.)

b) (Abstrakte Version) Sei V ein K-Vektorraum mit $\dim V = n < \infty$. Zu einer symmetrischen Bilinearform $\phi : V \times V \to K$ gibt es eine Basis $\mathcal{B}$ von V und Skalare $\alpha_1, ..., \alpha_n \in K$ mit

$$M(\mathcal{B}, \phi, \mathcal{B}) = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

Beweis: nach Beispiel 10.21.

Beispiel 10.20 Sei K ein Körper mit $\operatorname{char}(K) \neq 2$. Dann ist $2 \neq 0$, und $\frac{1}{2}$ existiert in K. Die Einheitsmatrix E_n wird rechts neben eine symmetrische Matrix A geschrieben. An A werden gewisse Spaltenumformungen und *die gleichen* Zeilenumformungen vorgenommen, an E_n nur die Spaltenumformungen. Nach Beispiel 4.10 (vii) lassen sich die Spaltenumformungen als Multiplikation *von rechts* mit gewissen Matrizen $S_1, ..., S_k$ deuten. Die *gleichen* Zeilenumformungen sind dann tatsächlich (!) Multiplikationen *von links* mit den transponierten Matrizen $S_1^{tr}, ..., S_k^{tr}$. Man erhält

$$E_n \cdot S_1 \cdot ... \cdot S_k =: T \quad \text{und}$$

$$\begin{aligned} \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} &= S_k^{tr} \cdot ... \cdot S_1^{tr} \cdot A \cdot S_1 \cdot ... \cdot S_k \\ &= T^{tr} \cdot A \cdot T \end{aligned}$$

Ob man erst Spaltenumformungen und dann die zugehörigen Zeilenumformungen durchführt oder abwechselnd oder umgekehrt, ist egal; denn die Matrizenmultiplikation ist assoziativ.
Ein Beispiel:

$$\begin{array}{rlll}
A = & \begin{pmatrix}1&-2&3\\-2&4&-5\\3&-5&9\end{pmatrix} & & \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = E_3 \\
 & & S_{II}(2;1,2) & \\
 & & S_{II}(-3;1,3) & \\
 & \begin{pmatrix}1&0&0\\-2&0&1\\3&1&0\end{pmatrix} & & \begin{pmatrix}1&2&-3\\0&1&0\\0&0&1\end{pmatrix} \\
Z_{II}(2;1,2) & & & \\
Z_{II}(-3;1,3) & \begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix} & & \\
 & & S_{II}(\frac{1}{2};3,2) & \\
 & \begin{pmatrix}1&0&0\\0&\frac{1}{2}&1\\0&1&0\end{pmatrix} & & \begin{pmatrix}1&\frac{1}{2}&-3\\0&1&0\\0&\frac{1}{2}&1\end{pmatrix} \\
Z_{II}(\frac{1}{2};3,2) & & & \\
 & \begin{pmatrix}1&0&0\\0&1&1\\0&1&0\end{pmatrix} & & \\
 & & S_{II}(-1;2,3) & \\
 & \begin{pmatrix}1&0&0\\0&1&0\\0&1&-1\end{pmatrix} & & \begin{pmatrix}1&\frac{1}{2}&-\frac{7}{2}\\0&1&-1\\0&\frac{1}{2}&\frac{1}{2}\end{pmatrix} = T \\
Z_{II}(-1;2,3) & & & \\
T^{tr}\cdot A\cdot T = & \begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix} & &
\end{array}$$

Anstelle von $S_{II}(\frac{1}{2};3,2)$ und $Z_{II}(\frac{1}{2};3,2)$ könnte man auf die dritte Matrix links $S_{II}(1;3,2)$ und $Z_{II}(1;3,2)$ anwenden. Das gäbe

$$\begin{pmatrix}1&0&0\\0&2&1\\0&1&0\end{pmatrix}.$$

Nun wäre man doch gezwungen, $\frac{1}{2}$ zu benutzen. $S_{II}(-\frac{1}{2};2,3)$ und $Z_{II}(-\frac{1}{2};2,3)$ gäben

$$\begin{pmatrix}1&0&0\\0&2&0\\0&0&-\frac{1}{2}\end{pmatrix}.$$

Falls $\sqrt{2}\in K$ ist (z.B bei $K=\mathbb{R}$), könnte man mit $S_I(\frac{1}{\sqrt{2}};2)$, $Z_I(\frac{1}{\sqrt{2}};2)$, $S_I(\sqrt{2};3)$ und $Z_I(\sqrt{2};3)$ die Diagonaleinträge auf ±1 normieren. Aber $\sqrt{2}$ zu benutzen ist nicht so schön (es geht nicht in $\mathbb{Q}$). Bei $K=\mathbb{R}$ kann man das Vorzeichen in -1 nicht wegbekommen. Da braucht man $\sqrt{-1}\in K$, z.B. $K=\mathbb{C}$.

Unser Standpunkt ist erstmal, daß es wichtig ist, Diagonalgestalt zu erreichen, aber nicht, die Einträge auf ± 1 zu normieren.

Beispiel 10.21 Beim vorigen Beispiel wurde in der Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

mit $S_{II}(\frac{1}{2},;3,2)$ und $Z_{II}(\frac{1}{2},;3,2)$ ein Diagonaleintrag $\neq 0$ an der Position $(2,2)$ erzeugt. Wenn man weiter unten rechts schon einen Diagonaleintrag $\neq 0$ hat, ist es besser, den mit Umformungen vom Typ III nach oben zu vertauschen. Denn bei Umformungen vom Typ II könnte er stören. Ein Beispiel:

$$\begin{array}{rcl} & \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} & S_{III}(1,2) \\ Z_{III}(1,2) & \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} & \\ & \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix} & \end{array}$$

ist gut.

$$\begin{array}{rcl} & \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} & S_{II}(1;2,1) \\ Z_{II}(1;2,1) & \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} & \\ & \begin{pmatrix} 0 & -1 \\ -1 & -2 \end{pmatrix} & \end{array}$$

ist schlecht.

Beweis von Satz 10.19: a) Das allgemeine Prinzip mit Spaltenumformungen und gleichen Zeilenumformungen wurde in Beispiel 10.20 beschrieben. Hier ist ein präziser Algorithmus:
1. Schritt:
1. Fall: $a_{11} \neq 0$.
Dann löscht man mit den Spaltenumformungen $S_{II}(-\frac{a_{1i}}{a_{11}};1,i)$ für $i = 2, ..., n$ und den entsprechenden Zeilenumformungen die anderen Einträge in der ersten Spalte und der ersten Zeile. Der 1. Schritt ist fertig.
2. Fall: $a_{11} = 0$.
Unterfall (a): Es gibt ein $a_{ii} \neq 0$; man betrachtet das minimale i mit $a_{ii} \neq 0$. Durch Vertauschen mit $S_{III}(1,i)$ und $Z_{III}(1,i)$ erhält man an der Position $(1,1)$ den Eintrag $a_{ii} \neq 0$. Nun weiter wie im 1. Fall.
Unterfall (b): Alle $a_{ii} = 0$. Wenn $A = 0$ ist, ist man trivialerweise fertig. Sonst gibt es ein $a_{ij} \neq 0$ mit i minimal und dann (für dieses i) j minimal.

Mit $S_{II}(1; j, i)$ und $Z_{II}(1; j, i)$ erhält man an der Position (i, i) den Eintrag $2 \cdot a_{ij}$. Hier ist $a_{jj} = 0$ wichtig. Wegen $\mathrm{char}(K) \neq 2$ ist $2 \cdot a_{ij} \neq 0$. Nun weiter wie im Unterfall (a).
2. Schritt: Wie der erste Schritt für die $(n-1) \times (n-1)$-Matrix, die man aus A durch Weglassen der ersten Zeile und Spalte erhält.
Weitere Schritte: analog.
Es ist klar, daß der Algorithmus funktioniert.

b) Man wählt irgendeine Basis $\mathcal{C} = (c_1, ..., c_n)$ von V. Die Matrix $A := M(\mathcal{C}, \phi, \mathcal{C})$ ist symmetrisch (Bemerkung 9.5 c)). Man konstruiert mit dem Algorithmus in a) eine invertierbare Matrix T mit

$$T^{tr} \cdot A \cdot T = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

für geeignete $\alpha_1, ..., \alpha_n \in K$. Die Basis $\mathcal{B} := \mathcal{C} \cdot T = (c_1, ..., c_n) \cdot T$ erfüllt $M(\mathcal{C}, \mathcal{B}) = T$ und

$$M(\mathcal{B}, \phi, \mathcal{B}) = M(\mathcal{C}, \mathcal{B})^{tr} \cdot M(\mathcal{C}, \phi, \mathcal{C}) \cdot M(\mathcal{C}, \mathcal{B}) = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

□

Satz 10.22 *(Spektralsatz für reelle symmetrische Matrizen, hier ohne Beweis)*
Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Es gilt:

(i) $P_A(t) = \prod_{i=1}^n (t - \lambda_i)$ mit $\lambda_i \in \mathbb{R}$. Also sind alle Eigenwerte von A (als komplexer Matrix) reell.

ii) Die Matrix A ist diagonalisierbar. Mit Satz 8.18 folgt: $M(n \times 1, \mathbb{R})$ ist direkte Summe aller Eigenräume.

iii) Die Eigenräume zu verschiedenen Eigenwerten sind orthogonal bezüglich des Standardskalarproduktes auf $M(n \times 1, \mathbb{R})$.

iv) Daher gibt es eine ON-Basis von $M(n \times 1, \mathbb{R})$ (bezüglich des Standardskalarproduktes) aus Eigenvektoren von A. Ist T eine Matrix, deren Spalten eine solche ON-Basis bilden, so ist $T^{-1} = T^{tr}$ (dann heißt T orthogonal), und es ist

$$T^{tr} \cdot A \cdot T = T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Satz/Definition 10.23 *(Satz von Sylvester, hier ohne Beweis)*
a) (Matrix-Version) Sei $A \in M(n \times n, \mathbb{R})$ *symmetrisch. Seien* $\lambda_1, ..., \lambda_n \in \mathbb{R}$ *die Eigenwerte von* A*. Jedes* $T \in GL(n, \mathbb{R})$ *mit*

$$T^{tr} \cdot A \cdot T = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

für gewisse $\alpha_1, ..., \alpha_n \in \mathbb{R}$ *erfüllt*

$$\begin{aligned} \text{(die Anzahl der positiven } \alpha_i) &= \text{(die Anzahl der positiven } \lambda_i) =: r_+, \\ \text{(die Anzahl der } \alpha_i = 0) &= \text{(die Anzahl der } \lambda_i = 0) =: r_0, \\ \text{(die Anzahl der negativen } \alpha_i) &= \text{(die Anzahl der negativen } \lambda_i) =: r_-. \end{aligned}$$

Das Tripel (r_+, r_0, r_-) *heißt* Signatur *von* A*. Die Zahlen erfüllen auch*

$$r_+ + r_0 + r_- = n \quad \text{und} \quad r_0 = n - \operatorname{rang} A.$$

b) (Abstrakte Version) Sei V *ein endlich-dimensionaler* $\mathbb{R}$*-Vektorraum und* $\phi : V \times V \to \mathbb{R}$ *eine symmetrische Bilinearform. Sei*

$$\begin{aligned} r_+ &:= \max(\dim U \mid U \text{ ist ein Untervektorraum von } V, \\ &\qquad\qquad \text{auf dem } \phi \text{ positiv definit ist}, \\ r_0 &:= \dim \operatorname{Rad}(\phi), \\ r_- &:= \max(\dim U \mid U \text{ ist ein Untervektorraum von } V, \\ &\qquad\qquad \text{auf dem } \phi \text{ negativ definit ist}. \end{aligned}$$

Es gibt Unterräume $U_1, U_2 \subset V$ *mit den Eigenschaften*

$$\begin{aligned} &V = U_1 \oplus \operatorname{Rad}(\phi) \oplus U_2, \\ &U_1 \perp U_2 \text{ (d.h. } u_1 \perp u_2 \text{ für alle } u_1 \in U_1, u_2 \in U_2), \\ &\phi \text{ ist positiv definit auf } U_1 \text{ und negativ definit auf } U_2. \end{aligned}$$

Es gilt immer:

$$\dim U_1 = r_+, \quad \dim U_2 = r_-.$$

Das Tripel (r_+, r_0, r_-) *heißt* Signatur *von* ϕ*.*

Satz 10.24 *(Hauptminorenkriterium für Definitheit, hier ohne Beweis)*
Sei $A = (a_{ij}) \in M(n \times n, \mathbb{R})$ *eine symmetrische Matrix.*
a) (Definition) Ihre Hauptminoren sind die Determinanten $\det A_k$ *der Untermatrizen*

$$A_k := (a_{ij})_{i,j=1,...,k} \in M(k \times k, \mathbb{R}) \text{ für } k = 1, 2, ..., n.$$

b) (Satz) A *ist genau dann positiv definit, wenn alle Hauptminoren positiv sind.*

Beispiel 10.25 Sei $r > 1$. Die Hauptminoren der Matrix

$$\begin{pmatrix} \frac{1}{r-1} & \frac{1}{r} \\ \frac{1}{r} & \frac{1}{r+1} \end{pmatrix}$$

sind $\frac{1}{r-1} > 0$ und

$$\frac{1}{r-1} \cdot \frac{1}{r+1} - \frac{1}{r} \cdot \frac{1}{r} = \frac{1}{r^2-1} - \frac{1}{r^2} > 0.$$

Daher ist die Matrix positiv definit.