Lineare Algebra für Informatiker und Statistiker

G. Kraus

Wintersemester 2007/08

1 Reelle lineare Gleichungssysteme, Matrizen und Vektoren

1.1 Lineare Gleichungssysteme

Beispiele
Betrachten wir folgende lineare Gleichungssysteme für drei Unbekannte x_1, x_2, x_3:

1.

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 1 \\ x_1 & - & x_2 & + & x_3 & = & 1 \\ x_1 & + & x_2 & - & x_3 & = & 1 \end{array}$$

Ganzzahlige Lösung.
Lösung erfordert Addition/Subtraktion und „aufgehende“ Division von ganzen Zahlen.
Lösung eindeutig bestimmt.

2.

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 1 \\ x_1 & - & x_2 & + & x_3 & = & 0 \\ x_1 & + & x_2 & - & x_3 & = & 0 \end{array}$$

Rationale Lösung.
Lösung erfordert Addition/Subtraktion und „nicht aufgehende“ Division von ganzen Zahlen.
Lösung eindeutig bestimmt.

3.

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 1 \\ x_1 & - & x_2 & + & x_3 & = & 0 \\ 2x_1 & + & & & 2x_3 & = & 1 \end{array}$$

Rationale Lösungen.
Lösung erfordert Addition/Subtraktion und „nicht aufgehende“ Division von ganzen Zahlen.
Nicht eindeutig bestimmte Lösung, sondern einparametrige Schar von Lösungen.

4.

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

Gleichungssystem nicht lösbar.
Rechnung erfordert Addition/Subtraktion Division von ganzen Zahlen.

Wir sehen:

1. Für ähnlich aussehende Gleichungssysteme können völlig unterschiedliche Aussagen bzgl. der Lösbarkeit gelten.

2. Die Lösbarkeit muß unterschiedlich beurteilt werden, je nachdem, in welchem Zahlenbereich man die Lösungen sucht.

3. Das Lösungsverfahren ist nicht auf ganze oder rationale Zahlen usw. beschränkt, sondern funktioniert stets, wenn man addieren, subtrahieren, multiplizieren und dividieren kann.

1.2 Matrix-Schreibweise

Es sei $\mathbb{R}$ die Menge der reellen Zahlen.

Definition 1.2.1 (Matrizen und Vektoren) *Es seien $m, n \geq 1$ natürliche Zahlen.*

1. *Ein rechteckiges Schema reeller Zahlen*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

   heißt Matrix mit m Zeilen und n Spalten und Koeffizienten in $\mathbb{R}$.

   Die Menge solcher Matrizen wird mit $\mathbb{R}^{m\times n}$ bezeichnet. Schreibt man $A \in \mathbb{R}^{m\times n}$, so bedeutet das, daß A eine Matrix mit m Zeilen und n Spalten und Koeffizienten in $\mathbb{R}$ ist.

2. $m = 1$: $\mathbb{R}_n := \mathbb{R}^{1\times n}$ *Zeilenvektoren.*

3. $n = 1$: $\mathbb{R}^m := \mathbb{R}^{m\times 1}$ *Spaltenvektoren.*

4. *Speziell sind die Zeilen- und Spaltenvektoren einer Matrix definiert.*

Definition 1.2.2 (Multiplikation einer Matrix mit einem Vektor) *Es seien $m, n \geq 1$ natürliche Zahlen.*

Ferner sei $A \in \mathbb{R}^{m\times n}$ eine $m \times n$-Matrix und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ ein Vektor.

Unter dem **Produkt** *Ax versteht man dann den Vektor*

$$Ax = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} \in \mathbb{R}^m$$

Bemerkung 1.2.3 (Matrix-Schreibweise eines linearen Gleichungssystems) *Es seien $m, n \geq 1$ natürliche Zahlen.*
Ferner sei

— $A \in \mathbb{R}^{m\times n}$ *eine $m \times n$-Matrix,*

— $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ *ein Vektor,*

— $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in I\!R^m$ *ein weiterer Vektor.*

Dann ist

$$Ax = b$$

eine abkürzende Schreibweise für

$$\begin{array}{rcl} a_{11}x_1 + \ldots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \ldots + a_{mn}x_n & = & b_m \end{array}$$

1.3 Spezielle Matrizen und die entsprechenden linearen Gleichungssysteme.

Beispiel 1.3.1 *Die Einheitsmatrix E_n.*
Beispiel 1.3.2 *Nichtsinguläre Diagonalmatrizen.*
Beispiel 1.3.3 *Allgemeine Diagonalmatrizen, Fallunterscheidung bzgl. der rechten Seiten.*
Beispiel 1.3.4 *Obere Dreiecksmatrizen, Fallunterscheidung bzgl. der rechten Seiten.*
Beispiel 1.3.5 *Untere Dreiecksmatrizen, Fallunterscheidung bzgl. der rechten Seiten.*
Beispiel 1.3.6 *Zeilenstufenform, Fallunterscheidung bzgl. der rechten Seiten.*

1.4 Lösungsmenge eines linearen Gleichungssystems

Definition 1.4.1 *Seien $A \in I\!R^{m\times n}$ und $b \in I\!R^m$ gegeben.*

1. $L(A, b) := \{x \in I\!R^n \mid Ax = b\}$ *heißt* **Lösungsmenge des inhomogenen linearen Gleichungssystems** $Ax = b$.

2. $L(A) := \{x \in I\!R^n \mid Ax = 0\}$ *heißt* **Lösungsmenge des homogenen linearen Gleichungssystems** $Ax = 0$.

2 Der Gauß-Algorithmus

2.1 Stufenmatrizen

Definition 2.1.1 *Eine Matrix $B = (b_{\mu\nu}) \in I\!R^{m\times n}$ heißt* **in Zeilenstufenform** *(Stufenmatrix), wenn es ein $r \leq \min(m, n)$ und Indices $\nu_1 < \ldots < \nu_r \in \{1, \ldots, n\}$ gibt, so daß gilt:*
$\forall \mu \in \{1, \ldots, r\}\ [b_{\mu\nu_\mu} \neq 0]$
$\forall \mu \in \{1, \ldots, m\}\ [\mu > r \vee \nu < \nu_\mu \Longrightarrow b_{\mu\nu} = 0$
Die Zahl r heißt **Rang** *der Stufenmatrix, Bezeichnung $r := \mathrm{Rg}(B)$.*

Die von Null verschiedenen Elemente $b_{\mu\nu_\mu}$, $\mu \in \{1, \dots, r\}$ heißen die **Pivot-Elemente** *der Stufenmatrix.*

Bemerkung 2.1.2 *Ist B in Zeilenstufenform, gilt:*

1. *In jeder Zeile ist $b_{\mu\nu_\mu}$ der erste nichtverschwindende Koeffizient der Matrix.*
2. *Jede Zeile beginnt mit Nullen (die erste Zeile kann mit einem von Null verschiedenen Element beginnen); die Zahl der Nullen wird mit jeder Zeile größer.*
3. *$\mu > r \Longrightarrow$ alle Zeilen bestehen nur aus Nullen.*

2.2 Elementare Zeilenumformungen

Bemerkung 2.2.1
1. *Die Gültigkeit eines linearen Gleichungssystems $Ax = b$ wird durch folgende Operationen nicht geändert:*
   (a) *Multiplikation einer Gleichung mit einem Faktor $\neq 0$.*
   (b) *Addition einer Gleichung zu einer anderen Gleichung.*
   (c) *Addition eines Vielfachen einer Gleichung zu einer anderen Gleichung.*
   (d) *Vertauschung von einzelnen Gleichungen.*
2. *Umformulierung in elementare Zeilenumformungen, angewandt auf die erweiterte Matrix $(A|b)$, das sind folgende Umformungen:*
   (a) *Typ I: Multiplikation einer Zeile mit einem Faktor $\lambda \neq 0$.*
   (b) *Typ II: Addition der j-ten Zeile zur i-ten Zeile.*
   (c) *Typ III: Addition des λ-fachen der j-ten Zeile zur i-ten Zeile.*
   (d) *Typ IV: Vertauschung der i-ten und der j-ten Zeile.*
3. *Die Lösungen bleiben invariant, wenn das Gleichungssystem durch elementare Zeilenumformungen auf Zeilenstufenform gebracht wird.*

Lemma 2.2.2 *Jede Matrix kann durch elementare Zeilenumformungen in Zeilenstufenform gebracht werden.*

Algorithmus 2.2.3 (Gauß-Algorithmus) *.*

1. *(A, b) durch elementare Zeilenumformungen in Zeilenstufenform $(\bar{A}, \bar{b})$ bringen. Nach Bemerkung oben gilt: $L(A, b) = L(\bar{A}, \bar{b})$.*

   Sei $r := \mathrm{Rg}(\bar{A})$.
2. *Spezialfall: In den ersten r Zeilen stehen die jeweils ersten von Null verschiedenen Elemente $\bar{a}_{\mu\nu(\mu)}$ in der Hauptdiagonalen (d.h.: $\nu(\mu) = \mu$ für $\mu = 1, \dots, r$). Man sieht unmittelbar:*
   (a) *$\bar{b}_{r+1} \neq 0 \Longrightarrow L(\bar{A}, \bar{b}) = \emptyset$ (leere Menge).*

(b) $\bar{b}_{r+1} = 0 \Longrightarrow L(\bar{A}, \bar{b}) \neq \emptyset$, *genauer:*

$$\forall \begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix} \in I\!R^{n-r} \exists_1 \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \in I\!R^r \; [\bar{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \bar{b}_1 \\ \vdots \\ \bar{b}_m \end{pmatrix}].$$

Zur Berechnung der $x_1, \ldots, x_r$ *werden die vorgegebenen Werte von* $x_{r+1}, \ldots, x_n$ *in die ersten* r *Gleichungen von* $\bar{A}x = \bar{b}$ *eingesetzt und dann nacheinander die Werte von* x_r, x_{r-1} *bis* x_1 *berechnet.*

3. *Der allgemeine Fall (nicht notwendig* $\nu(\mu) = \mu$*) kann durch Umnumerierung der Variablen (d.h., durch Spaltenvertauschungen der Koeffizientenmatrix) auf den Spezialfall zurückgeführt werden. Nach der Lösung muß die ursprüngliche Numerierung wieder hergestellt werden.*

4. *Einfacher ist es, den allgemeinen Fall in direkter Entsprechung zum Spezialfall zu behandeln. Gibt man die Werte der* x_j *mit* $j \neq \nu(\mu), \mu \in \{1, \ldots, r\}$ *vor — das sind die nicht zu den Pivot-Elementen gehörenden Variablen, sie bilden ebenfalls einen* $(n-r)$*-Vektor —, kann man nacheinander* $x_{\nu(r)}, x_{\nu(r-1)}, \ldots, x_{\nu(1)}$ *berechnen, indem man die ersten* r *Zeilen von hinten her abarbeitet.*

5. *Man erhält eine* **Parametrisierungsabbildung** $\Phi : I\!R^{n-r} \longrightarrow L(A, b)$. *Diese Abbildung ist* **eineindeutig und liefert alle Lösungen** *(die Abbildung ist bijektiv), d.h., jedem Parametervektor*

$$t = \begin{pmatrix} t_1 \\ \vdots \\ t_{n-r} \end{pmatrix} \in I\!R^{n-r}$$

ist genau ein Lösungsvektor

$$x = \Phi(t) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in L(A, b)$$

zugeordnet, zwei verschiedenen Parametervektoren ergeben verschiedene Lösungsvektoren, und jeder Lösungsvektor kommt als solcher parametrisierter Vektor vor.

Bemerkung 2.2.4 *Das Gleichungssystem ist nicht immer lösbar. Es kommt auf die rechten Seiten an. Die Lösungsmenge kann also die leere Menge sein.*

Satz 2.2.5 *Sei* $Ax = b$ *ein inhomogenes lineares Gleichungssystem. Die erweiterte Koeffizientenmatrix* $(A|b)$ *sei (o.E.) in Zeilenstufenform. Dann gilt:*

1. $$L(A, b) \neq \emptyset \Longleftrightarrow Rg(A) = Rg(A|b)$$

2. *Gilt* $Rg(A) = Rg(A|b) =: r$ *und ist* $A \in I\!R^{m \times n}$, *so wird die Lösungsmenge* $L(A, b)$ *durch* $n - r$ *Parameter parametrisiert.*

2.3 Struktur der Lösungsmengen homogener linearer Gleichungssysteme

Bemerkung 2.3.1

1. *Stets lösbar, Nullvektor immer Lösung.*
2. *Lösung i.a. nicht eindeutig. Lösungsmenge. Beispiele.*
3. *Produkt einer Lösung mit einem Faktor ergibt wieder Lösung.*
4. *Summe von Lösungen ist wieder Lösung.*

2.4 Struktur der Lösungsmengen inhomogener linearer Gleichungssysteme

Bemerkung 2.4.1

1. *Nicht immer lösbar. Lösungsmenge dann die* **leere Menge** $\emptyset$.
2. *Lösung i.a. nicht eindeutig. Lösungsmenge. Beispiele.*
3. *Produkt einer Lösung mit einem Faktor ergibt i.a. nicht wieder eine Lösung.*
4. *Summe von Lösungen ist i.a. nicht wieder Lösung.*
5. x, y *Lösungen* $\Longrightarrow x - y$ *Lösung des homogenen Systems.*
6. *Lösungsmenge = partikuläre Lösung + Lösungsmenge des homogenen Systems.*

3 Rechenoperationen für Matrizen und Vektoren

3.1 Addition von Matrizen

Definition 3.1.1 *Seien $m, n \geq 1$ natürliche Zahlen.*
Die Abbildung

$$\begin{array}{rccc} + : & \mathbb{R}^{(m\times n)} \times \mathbb{R}^{(m\times n)} & \longrightarrow & \mathbb{R}^{(m\times n)} \\ & \left(\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \right) & \longmapsto & \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix} \end{array}$$

heißt **Addition (Summe) von Matrizen***.*
In den Fällen $m = 1$ und $n = 1$ spricht man von der Addition von Zeilen- bzw. Spaltenvektoren.
Aussage 3.1.2 (Eigenschaften der Matrizenaddition) .

1. *Assoziativgesetz.*

2. *Kommutativgesetz.*

3. *Neutrales Element (Nullmatrix bzw. Nullvektor (genauer: Nullzeilenvektor, Nullspaltenvektor)).*

   Es gibt nicht einen *Nullvektor bzw.* eine *Nullmatrix schlechthin; es ist jeweils die Zeilen- und Spaltenanzahl mitanzugeben.*

4. *Zu jedem Element existiert ein inverses Element bzgl. der Addition.*

3.2 Die transponierte Matrix

Definition 3.2.1 *Ist $A = (a_{\mu\nu}) \in \mathbb{R}^{(m\times n)}$ eine Matrix, so heißt die Matrix ${}^tA := ({}^ta_{\nu\mu}) \in \mathbb{R}^{(n\times m)}$ mit ${}^ta_{\nu\mu} := a_{\mu\nu}$ die* **transponierte Matrix***.*
Bemerkung 3.2.2 *Für $x \in \mathbb{R}^m$ ist ${}^tx \in \mathbb{R}_m$.*

3.3 Das Produkt von Matrizen mit Skalaren

Definition 3.3.1 *Seien $m, n \geq 1$ natürliche Zahlen.*
Die Abbildung

$$\begin{array}{rccc} \cdot : & \mathbb{R} \times \mathbb{R}^{(m\times n)} & \longrightarrow & \mathbb{R}^{(m\times n)} \\ & \left(\lambda, \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \right) & \longmapsto & \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix} \end{array}$$

heißt Multiplikation der Matrix mit dem Skalar λ.
In den Fällen $m = 1$ und $n = 1$ spricht man von der Multiplikation von Zeilen- bzw. Spaltenvektoren mit Skalaren.
Aussage 3.3.2 *Seien $m, n \geq 1$. Es gilt:*

1. *Assoziativgesetz*

   $\forall_{\lambda,\mu\in\mathbb{R},A\in\mathbb{R}^{(m\times n)}}(\lambda\mu)A = \lambda(\mu A)$

2. *Distributivgesetze:*

   (a) $\forall_{\lambda,\mu\in\mathbb{R},A\in\mathbb{R}^{(m\times n)}}(\lambda+\mu)A = \lambda A + \mu A$

   (b) $\forall_{\lambda\in\mathbb{R},A,B\in\mathbb{R}^{(m\times n)}}\lambda(A+B) = \lambda A + \lambda B$

3.4 Das Matrizenprodukt

Definition 3.4.1 *Seien $k, m, n \in \mathbb{N}^*$ natürliche Zahlen. Die Abbildung*

$$\begin{array}{cccc} \cdot: & \mathbb{R}^{(k\times m)} \times \mathbb{R}^{(m\times n)} & \longrightarrow & \mathbb{R}^{(k\times n)} \\ & \left(\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{km} \end{pmatrix}, \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \right) & \longmapsto & \begin{pmatrix} \sum_{\mu=1}^m a_{1\mu}b_{\mu 1} & \dots & \sum_{\mu=1}^m a_{1\mu}b_{\mu n} \\ \vdots & & \vdots \\ \sum_{\mu=1}^m a_{k\mu}b_{\mu 1} & \dots & \sum_{\mu=1}^m a_{k\mu}b_{\mu n} \end{pmatrix} \end{array}$$

(sog **Zeilen-Spalten-Faltung***) heißt* **Matrizenprodukt***.*
Andere Schreibweise: $(a_{\lambda\mu})(b_{\mu\nu}) = (c_{\lambda\nu})$ mit $c_{\lambda\nu} = \sum_\mu a_{\lambda\mu}b_{\mu\nu}$.
Satz 3.4.2 (Regeln zum Matrizenprodukt) *1. Das Matrizenprodukt ist* **assoziativ***, aber* **nicht kommutativ***.*

2. *Distributivgesetze: $(A+B)C = AB + AC$, $A(B+C) = AB + AC$.*

3. *Für Matrizen A und B und $\lambda \in \mathbb{R}$ gilt $A(\lambda B) = (\lambda A)B = \lambda(AB)$.*

4. *Für Matrizen A und B gilt ${}^t(AB) =^t B\ {}^tA$*

Beweis: Nachrechnen, d.h.:

1. Assoziativgesetz: Seien $A = (a_{\iota\lambda}) \in \mathbb{R}^{(k\times l)}$, $B = (b_{\lambda\mu}) \in \mathbb{R}^{(l\times m)}$, $C = (c_{\mu\nu}) \in \mathbb{R}^{(m\times n)}$.

   Dann ist

   $AB = (d_{\iota\mu}) \in \mathbb{R}^{(k\times m)}$ mit $d_{\iota\mu} = \sum_{\lambda=1}^{l}\ a_{\iota\lambda}b_{\lambda\mu}$

   und

   $(AB)C = (e_{\iota\nu}) \in \mathbb{R}^{(k\times n)}$ mit

   $e_{\iota\nu} = \sum_{\mu=1}^{m}\ d_{\iota\mu}c_{\mu\nu} =$

   $\sum_{\mu=1}^{m}\ (\sum_{\lambda=1}^{l}\ a_{\iota\lambda}b_{\lambda\mu})\ c_{\mu\nu} =$

   $\sum_{\mu=1}^{m}\ \sum_{\lambda=1}^{l}\ a_{\iota\lambda}b_{\lambda\mu}c_{\mu\nu}$,

   andererseits

   $BC = (f_{\lambda\nu}) \in \mathbb{R}^{(l,n)}$ mit $f_{\lambda\nu} = \sum_{\mu=1}^{m}\ b_{\lambda\mu}c_{\mu\nu}$

   und

   $A(BC) = (g_{\iota\nu}) \in \mathbb{R}^{(k\times n)}$ mit

   $g_{\iota\nu} = \sum_{\lambda=1}^{l}\ a_{\iota\lambda}f_{\lambda\nu} =$

$\sum_{\lambda=1}^{l} a_{\iota\lambda} \left(\sum_{\mu=1}^{m} b_{\lambda\mu}c_{\mu\nu}\right) =$
$\sum_{\lambda=1}^{l} \sum_{\mu=1}^{m} a_{\iota\lambda}b_{\lambda\mu}c_{\mu\nu}.$

Die Koeffizienten der beiden Matrizenprodukte sind also Summen der gleichen Summanden und stimmen überein.

Nicht kommutativ, Beweis durch ein Gegenbeispiel, etwa

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

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

2. Distributivgesetze: entsprechende Rechnung.
3. Assoziativität bzgl. der Multiplikation mit Skalaren: elementare Rechnung.
4. Transposition eines Matrizenprodukts: elementare Rechnung.

3.5 Das Kronecker-Symbol

Definition 3.5.1 Kronecker-Symbol

$$\delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Bemerkung 3.5.2 *Die n-reihige Einheitsmatrix ist dann*

$$E = (\delta_{ij})_{i,j=1,\ldots,n}$$

3.6 Elementarmatrizen

Definition 3.6.1 *Sei K ein Körper und $n \in \mathbb{N}$. Die folgenden Matrizen in $\mathbb{R}^{n\times n}$ heißen n-reihige* **Elementarmatrizen***:*

1. *Typ I:*

$$S_i(\lambda) := (\delta_{\mu\nu} + \delta_{i\mu}\delta_{i\nu}(\lambda - 1))_{\mu,\nu=1,\ldots,n} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \lambda & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

($i \in \{1, \ldots, n\}, \lambda \in K \setminus \{0\}$).

Quadratische Matrix wie die Einheitsmatrix E, jedoch in der i-ten Zeile und Spalte λ statt 1.

2. *Typ II:*

$$Q_i^j := (\delta_{\mu\nu} + \delta_{\mu i}\delta_{\nu j})_{\mu,\nu=1,\dots,n} = \dots$$

($i, j \in \{1, \dots, n\}$).

Quadratische Matrix wie die Einheitsmatrix E, jedoch in der i-ten Zeile und j-ten Spalte eine zusätzliche 1.

3. *Typ III:*

$$Q_i^j(\lambda) := (\delta_{\mu\nu} + \delta_{\mu i}\delta_{\nu j}\lambda)_{\mu,\nu=1,\dots,n} = \dots$$

($i, j \in \{1, \dots, n\}, \lambda \in K \setminus \{0\}$).

Wie Typ II, jedoch in der i-ten Zeile und j-ten Spalte $\lambda \neq 0$.

4. *Typ IV:*

$$P_i^j := (\delta_{\mu\nu} + \delta_{\mu i}\delta_{\nu j} + \delta_{\mu j}\delta_{\nu i} - \delta_{\mu i}\delta_{\nu i} - \delta_{\mu j}\delta_{\nu j})_{\mu,\nu=1,\dots,n} = \dots$$

($i, j \in \{1, \dots, n\}$).

Quadratische Matrix wie die Einheitsmatrix E, jedoch in der i-ten Zeile die 1 *nicht in der Hauptdiagonalen, sondern in der j-ten Spalte, in der j-ten Zeile die* 1 *nicht in der Hauptdiagonalen, sondern in der i-ten Spalte.*

Die folgende Aussage ist unmittelbar einzusehen:

Satz 3.6.2 *Ist $A \in \mathbb{R}^{m\times n}$ eine Matrix und $S \in \mathbb{R}^{m\times m}$ eine m-reihige Elementarmatrix, so ist SA die Matrix, die aus aus A durch folgende elementare Zeilenumformung entsteht:*

1. *Typ I, $S = S_i(\lambda)$: Multiplikation der i-ten Zeile mit λ.*
2. *Typ II, $S = Q_i^j$: Addition der j-ten Zeile zur i-ten Zeile.*
3. *Typ III, $S = Q_i^j(\lambda)$: Addition des λ-fachen der j-ten Zeile zur i-ten Zeile.*
4. *Typ IV, $S = P_i^j$: Vertauschung der i-ten und der j-ten Zeile.*

Elementare Spaltenumformungen sind analog zu elementaren Zeilenumformungen definiert. Es gilt entsprechend:

Satz 3.6.3 *Ist $A \in \mathbb{R}^{m\times n}$ eine Matrix und $T \in \mathbb{R}^{n\times n}$ eine n-reihige Elementarmatrix, so ist AT die Matrix, die aus aus A durch folgende elementare Spaltenumformung entsteht:*

1. *Typ I, $T = S_i(\lambda)$: Multiplikation der i-ten Spalte mit λ.*
2. *Typ II, $T = Q_i^j$: Addition der i-ten Spalte zur j-ten Spalte.*
3. *Typ III, $T = Q_i^j(\lambda)$: Addition des λ-fachen der i-ten Spalte zur j-ten Spalte.*
4. *Typ IV, $T = P_i^j$: Vertauschung der i-ten und der j-ten Spalte.*

3.7 Die inverse Matrix

Definition 3.7.1 *Eine quadratische Matrix $A \in I\!R^{n\times n}$ heißt* **invertierbar** *oder* **regulär** *oder* **nichtsingulär**, *wenn es eine Matrix $B \in I\!R^{n\times n}$ gibt mit $AB = BA = E$; dabei sei $E = E_n = (\delta_{i,j})_{i,j=1,\dots,n}$ die n-reihige Einheitsmatrix.*

Bemerkung 3.7.2 *Ist $A \in I\!R^{n\times n}$ invertierbar und sind $B, C \in I\!R^{n\times n}$ Matrizen mit $AB = BA = E$ und $AC = CA = E$, so gilt: $B = C$.*

Beweis: $B = (CA)B = C(AB) = C$.

Definition 3.7.3 *Sei $A \in I\!R^{n\times n}$ invertierbar. Die nach vorheriger Bemerkung eindeutig bestimmte Matrix B mit $AB = BA = E$ heißt die zu A* **inverse Matrix**, *Bezeichnung: $B = A^{-1}$.*

Bemerkung 3.7.4 .

1. *Die inverse Matrix ist wieder invertierbar, und es gilt $(A^{-1})^{-1} = A$.*
2. *Für zwei n-reihige invertierbare Matrizen A, B gilt: AB ist invertierbar, und $(AB)^{-1} = B^{-1}A^{-1}$.*
3. *Elementarmatrizen sind invertierbar, und es gilt:*
   (a) $(S_i(\lambda))^{-1} = S_i(\frac{1}{\lambda})$
   (b) $(Q_i^j)^{-1} = Q_i^j(-1)$
   (c) $(Q_i^j(\lambda))^{-1} = Q_i^j(-\lambda)$
   (d) $(P_i^j)^{-1} = P_i^j$

Satz 3.7.5 1. *Jede Matrix kann durch elementare Zeilen- und Spaltenumformungen auf folgende Normalform gebracht werden:*

$$\begin{pmatrix} E_r & 0 \\ & \\ 0 & 0 \end{pmatrix} = \dots$$

2. *Zu jeder Matrix $A \in I\!R^{m\times n}$ gibt es invertierbare Matrizen $S \in I\!R^{m\times m}$ und $T \in I\!R^{n\times n}$, so daß gilt*

$$SAT = \begin{pmatrix} E_r & 0 \\ & \\ 0 & 0 \end{pmatrix} = \dots$$

Beweis:

1. A durch elementare Zeilenumformungen auf Zeilenstufenform bringen mit $b_{\mu\nu_\mu} = 1$ für $\mu < r$.
2. Spaltenvertauschungen, so daß $\nu_\mu = \mu$ wird.
3. Durch elementare Spaltenumformungen wird erreicht: $b_{\mu\nu} = 0$ für ($\mu > r$ und $\nu > \mu$).

4. S und T seien die Produkte der Elementarmatrizen, die zu den elementaren Zeilen- und Spaltenumformungen der o.e. Schritte gehören.

Für invertierbare (also insbesondere quadratische) Matrizen gilt sogar:

Satz 3.7.6 *Sei $A \in \mathbb{R}^{n\times n}$ eine invertierbare Matrix.*

1. *Man kann A (nur) durch elementare Zeilenumformungen in die Einheitsmatrix $E = E_n$ überführen.*
2. *Es gibt endlich viele Elementarmatrizen $S_1, \dots, S_k$ mit $S_k \cdot \ldots \cdot S_1 \cdot A = E$.*
3. *Ist $S := S_k \cdot \ldots \cdot S_1$ das Produkt dieser Elementarmatrizen (in der angegebenen Reihenfolge), so gilt:*
   (a) $A^{-1} = S$.
   (b) $A = S^{-1} = S_1^{-1} \cdot \ldots \cdot S_k^{-1}$.
4. *A ist endliches Produkt von Elementarmatrizen.*

Zum Beweis: Siehe die Überlegungen zu dem folgenden Algorithmus, der die inverse Matrix A^{-1} von A bestimmt bzw. zeigt, daß A nicht invertierbar ist.
Algorithmus 3.7.7 (Bestimmung der inversen Matrix durch elem. ZUen)
.

1. *A durch elementare Zeilenumformungen auf Zeilenstufenform $\hat{A}$ bringen.*
2. *Im Fall $\mathrm{Rg}(\hat{A}) < n$ ist A nicht invertierbar (denn wäre A invertierbar, so wäre auch die $n \times n$-Matrix $\begin{pmatrix} E_r & 0 \\ & \\ 0 & 0 \end{pmatrix}$ (mit $r < n$) invertierbar, was aber nicht stimmen kann, weil für jede $n \times n$-Matrix B gilt: $\begin{pmatrix} E_r & 0 \\ & \\ 0 & 0 \end{pmatrix} \cdot B$ hat höchstens r von Null verschiedene Zeilen, kann also nicht E_n sein.*
3. *Im Fall $\mathrm{Rg}(\hat{A}) = n$ ist $\hat{A}$ eine obere Dreiecksmatrix, und man kann $\hat{A}$ durch weitere elementare Zeilenumformungen in die Einheitsmatrix E_n überführen.*
4. *Die gleichen elementaren Zeilenumformungen auf die Einheitsmatrix E anwenden (sinnvollerweise simultan zu den Umformungen von A) liefert die inverse Matrix A^{-1}.*

Denn: Die Ausführung der elementaren Umformungen ist äquivalent einer Multiplikation mit Elementarmatrizen $S_1, \dots, S_k$. Im Fall $\mathrm{Rg}(A) = n$ ergibt sich also: $S_k S_{k-1} \dots S_1 A = E_n$.
Sei $S := S_k S_{k-1} \dots S_1$.
Dann ist S als Produkt von Elementarmatrizen (die alle invertierbar sind) invertierbar; ferner gilt $SA = E$.

Es wird behauptet: $S = A^{-1}$.
Da $SA = E$ bereits gezeigt ist, ist hierzu noch zu zeigen: $AS = E$. Das sieht man wie folgt:
$AS = EAS = (S^{-1}S)AS = S^{-1}(SA)S = S^{-1}ES = S^{-1}S = E$.

Beispiel 3.7.8

4 Reelle Vektorräume und lineare Abbildungen

4.1 Vektorräume

Definition 4.1.1 *Ein* **$I\!R$-Vektorraum** *(kurz: VR) ist ein Tripel* $(V, +, \cdot)$ *(kurz meist nur* V *geschrieben) wie folgt:*

1. V *ist eine Menge. Die Elemente von* V *heißen* **Vektoren**.
2. *Es ist eine Vektoraddition*

$$\begin{array}{rccc} + : & V \times V & \longrightarrow & V \\ & (v,w) & \longmapsto & v+w \end{array}$$

   *definiert (***Addition (Summe) von Vektoren** *mit folgenden Eigenschaften:*

   (a) *Assoziativgesetz.*
   (b) *Kommutativgesetz.*
   (c) *Neutrales Element.*
   (d) *Zu jedem Element existiert ein inverses Element bzgl. der Addition.*

   (Dies wird kurz zusammengefaßt als „ $(V, +)$ *ist eine abelsche Gruppe“).*

   Das neutrale Element in V *bzgl.* $+$ *wird ebenfalls mit* 0 *bezeichnet und heißt* **Nullvektor**.

3. *Es gibt eine Abbildung* $\cdot : I\!R \times V \longrightarrow V$ *(***Multiplikation mit Skalaren***), so daß gilt*

   (a) $\forall \alpha, \beta \in I\!R, v \in V\ [\alpha(\beta v) = (\alpha\beta)v]$ *(Assoziativgesetz für* $\cdot$*).*
   (b) $\forall \alpha, \beta \in I\!R, v, w \in V$
      i. $\alpha(v + w) = \alpha v + \alpha w$
      ii. $(\alpha + \beta)v = \alpha v + \beta v$

      (Distributivgesetze)
   (c) $\forall v \in V\ [1v = v]$ *(Unitarität)*

Beispiel 4.1.2 1. $(I\!R, +, \cdot)$ $I\!R$-*VR.*

2. $m, n \in I\!N \Longrightarrow I\!R_n, I\!R^m, I\!R^{m\times n}$ *(mit komponentenweiser Addition und Multiplikation mit Skalaren) sind* $I\!R$-*VR.*
3. $a < b \in I\!R \Longrightarrow \mathcal{C}([a,b]) := \{f : [a,b] \longrightarrow I\!R\ :\ f \text{ stetig}\}$ *ist* $I\!R$-*VR.*

4.2 Lineare Abbildungen

Definition 4.2.1 *Seien V, W $\mathbb{R}$-Vektorräume, $f : V \longrightarrow W$ Abbildung.*
f **linear** $:\Longleftrightarrow$
$\forall \alpha, \beta \in \mathbb{R}, v, w \in V \ [f(\alpha v + \beta w) = \alpha f(v) + \beta f(w)]$
Satz 4.2.2 *Seien $m, n \geq 1$ natürliche Zahlen und $A = (aij) \in \mathbb{R}^{m \times n}$. Die Abbildung*

$$\begin{array}{rccc} \hat{A}: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{array}$$

ist linear.
Beweis: Folgt aus 3.4.2
Beispiel 4.2.3 *Beispiele linearer Abbildungen sind:*

1. *Die Nullabbildung $\forall x \in V \ [f(x) = 0]$.*

   Diese Abbildung kann im Fall $V = \mathbb{R}^m$ auch als Produkt mit der Nullmatrix geschrieben werden.

2. *Die identische Abbildung. Im Fall $V = \mathbb{R}^m$ Produkt mit der Einheitsmatrix E_m.*

3. *Streckung $\hat{\alpha} : V \longrightarrow V$, $v \longmapsto \alpha v$ (mit $\alpha \in \mathbb{R}$).*

   Im Fall $V = \mathbb{R}^m$ Produkt mit αE_m.

4. *Die Projektionsabbildungen $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto x_1.$*

   Produkt mit ${}^t e_i$.

5. *Die Auswertungsabbildungen $Abb(M, V) \longrightarrow V$, $f \longmapsto f(a)$ (mit $a \in M$).*

Satz 4.2.4 1. *Die Komposition von zwei linearen Abbildungen ist wieder eine lineare Abbildung.*

2. *Wird die lineare Abbildung $f : \mathbb{R}^k \to \mathbb{R}^m$ durch die Matrix $A \in \mathbb{R}^{k \times m}$ gegeben, die lineare Abbildung $g : \mathbb{R}^m \to \mathbb{R}^n$ durch die Matrix $B \in \mathbb{R}^{m \times n}$, so wird die Komposition $g \circ f : \mathbb{R}^k \to \mathbb{R}^n$ gegeben durch die Matrix $BA \in \mathbb{R}^{k \times n}$.*

4.3 Untervektorräume, Kern und Bild

Definition 4.3.1 *V VR, $U \subset V$, $U \neq \emptyset$.*
U **Untervektorraum** $:\Longleftrightarrow$

1. $\forall v, w \in U \ [v + w \in U]$

2. $\forall \alpha \in \mathbb{R}, v \in U \ [\alpha v \in U]$

Bemerkung 4.3.2 *$U \subset V$ UVR $\Longrightarrow U$ ist Vektorraum.*
Satz 4.3.3 *$f : V \to W$ lineare Abbildung $\Longrightarrow$*

1. **Ke**$f := f^{-1}(0) \subset V$ *Untervektorraum.*

2. **Im**$f := f(V) \subset W$ *Untervektorraum.*

Beispiel 4.3.4 *1.*

$$\begin{array}{rrcl} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \longmapsto & Ax = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 \end{pmatrix} \end{array}$$

(a) $\mathbf{Ke}f = \{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = 0\}$

(b) $\mathbf{Im}f = \{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = x_2\}$

2. *Ist $A \in \mathbb{R}^{m \times n}$, so ist* $\mathbf{Ke}\hat{A} \subset \mathbb{R}^n$ *die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$.*

4.4 Linearkombinationen, lineare Abhängigkeit und Unabhängigkeit

Definition 4.4.1 *V VR. Seien $v_1, \ldots, v_n \in V, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Der Vektor $\alpha_1 v_1 + \ldots + \alpha_n v_n$ heißt eine* **Linearkombination** *(L.K.) der Vektoren $v_1, \ldots, v_n$, die α_i die Koeffizienten in dieser Linearkombination.*
Beispiel 4.4.2 *1. Linearkombinationen von Vektoren e_i im $\mathbb{R}^n$.*

2. *Verschiedene Fälle von Linearkombinationen von Vektoren im $\mathbb{R}^2$ und $\mathbb{R}^3$.*

Definition 4.4.3 *V VR.*

1. *Seien $v_1, \ldots, v_n \in V$.*

   $\langle v_1, \ldots, v_n \rangle :=$ Menge der LK der Vektoren $v_1, \ldots, v_n$

2. *Analog für eine Teilmenge $M \subset V$:*

   $\langle M \rangle = span(M) := \{v \in V \ : \ \exists n \in \mathbb{N}, \alpha_1, \ldots, \alpha_n \in \mathbb{R}, v_1, \ldots, v_n \in M \ [v = \alpha_1 v_1 + \ldots + \alpha_n v_n]\}$

Bemerkungen und Beispiele 4.4.4 *1. Im Fall $M = \{a\}$ ist $\langle M \rangle = \mathbb{R}a = \{\alpha a \ : \ \alpha \in \mathbb{R}\}$.*

2. *$M = \{e_1, \ldots, e_n\} \in \mathbb{R}^n \Longrightarrow \langle M \rangle = \mathbb{R}^n$.*

3. *$M \subset V \Longrightarrow \langle M \rangle$ ist ein UVR von V.*

Definition 4.4.5 *$M \subset V$ Teilmenge. M* **Erzeugendensystem** *von $V :\Longleftrightarrow \langle M \rangle = V$*
Definition 4.4.6 *V VR.*

1. $\forall v_1, \ldots, v_n \in V$ $[v_1, \ldots, v_n$ **linear unabhängig** *(kurz: l.u.)* $:\Longleftrightarrow$
   $\forall \alpha_1, \ldots, \alpha_n \in I\!R$ $[\alpha_1 v_1 + \ldots + \alpha_n v_n = 0 \Longrightarrow \alpha_1 = \ldots = \alpha_n = 0]]$

2. $\forall M \subset V$ $[M$ *l.u.* $:\Longleftrightarrow$

   $\forall_{n \in I\!N, n \geq 1} \forall_{v_1, \ldots, v_n \in M}$ *pw. verschieden* $v_1, \ldots, v_n$ *l.u.*

   (d.h., je endlich viele Vektoren aus M *sind l.u.)*

3. $\forall M \subset V$ $[M$ **linear abhängig** *(l.a.)* $:\Longleftrightarrow M$ *nicht l.u.]*

Bemerkung 4.4.7 1. $v_1, \ldots, v_n$ *l.u.* $\Longrightarrow$ *Der Nullvektor läßt sich nur auf genau eine Weise (alle Koeffizienten* $= 0$*) als LK von* $v_1, \ldots, v_n$ *darstellen.*

2. *Seien* $v_1, \ldots, v_n \in I\!R^m$ *und* $A \in I\!R^{m \times n}$ *die Matrix mit den Spaltenvektoren* $v_1, \ldots, v_n$*. Dann gilt:*

   $v_1, \ldots, v_n$ *l.u.* $\Longleftrightarrow$ *Das homogene lineare Gleichungssystem* $Ax = 0$ *besitzt nur die triviale Lösung* $x = 0$.

3. *Sind unter den Vektoren* $v_1, \ldots, v_n$ *zwei gleich oder ist der Nullvektor einer dieser Vektoren, so sind* $v_1, \ldots, v_n$ *l.a.*

4. *Zwei Vektoren* v_1, v_2 *sind l.a.* $\Longleftrightarrow$ v_2 *ist Vielfaches von* v_1 *oder* v_1 *ist Vielfaches von* v_2 *(gleichgerichtete Vektoren). Lin. Unabh. von zwei Vektoren bedeutet also: verschiedene Richtung.*

5. $\forall v_1, \ldots, v_n \in V$ $[v_1, \ldots, v_n$ *l.a.* $\Longleftrightarrow \exists \alpha_1, \ldots, \alpha_n \in I\!R$

   (a) $\exists i \in \{1, \ldots, n\}$ $[\alpha_i \neq 0]$

   (b) $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0]$

6. $M \subset V$ *l.u.*, $M' \subset M \Longrightarrow M'$ *l.u.*

4.5 Basis, Dimension

Definition 4.5.1 V *VR.*

1. $\forall B \subset V$ $[B$ **Basis** *(von* V*)* $:\Longleftrightarrow$

   (a) B *l.u.*

   (b) $\langle B \rangle = V$ *(B Erzeugendensystem, Erzeugendenmenge)]*

2. $\dim V := \begin{cases} n & \text{Es gibt eine Basis von } V \text{ mit } n \text{ Elementen} \\ \infty & V \text{ besitzt keine endliche Basis} \end{cases}$

   heißt Dimension von V.

Bemerkungen und Beispiele 4.5.2 1. *Es wird gezeigt werden: Hat eine Basis die Länge* n*, dann jede. Die Dimension ist also wohldefiniert.*

2. $e_1, \ldots, e_n$ *Basis von* $I\!R^n$ *(sog. kanonische Basis). Also* $\dim I\!R^n = n$.

3. *Einen VR oder UVR kann man i.a. nicht explizit angeben derart, daß man z.B. alle Elemente aufzählt. Die vollständige Angabe eines endlichdimensionalen VRs bzw. UVRs ist stattdessen möglich durch die Angabe einer Basis.*

4. $\{v_1, \ldots, v_n\}$ *Basis von* $V \iff$ *Jeder Vektor* $v \in V$ *besitzt eine eindeutige Darstellung* $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$.

5. $v_1, \ldots, v_n$ *Basis von* $V \iff$

   (a) $v_1, \ldots, v_n$ *Erzeugendensystem von* V.

   (b) $\forall_{i \in \{1,\ldots,n\}} v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ *kein Erzeugendensystem von* V.

   D.h.: Eine Basis ist ein minimales Erzeugendensystem, d.h., daß jede echte Teilmenge einer Menge von Basisvektoren kein Erzeugendensystem mehr ist.

   Beweis:

   (a) *"$\Longrightarrow$" Sei* $v_1, \ldots, v_n$ *Basis von* V *und* $i \in \{1, \ldots, n\}$.
   Annahme, $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ *Erzeugendensystem von* V. *Dann gibt es* $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n \in I\!R$ *mit*
   $v_i = \alpha_1 v_1 + \ldots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \ldots + \alpha_n v_n$.
   Daraus folgt:
   $\alpha_1 v_1 + \ldots + \alpha_{i-1} v_{i-1} + (-1) v_1 + \alpha_{i+1} v_{i+1} + \ldots + \alpha_n v_n = 0$.
   Widerspruch zur linearen Unabhängigkeit der $v_1, \ldots, v_n$.

   (b) *„$\Longleftarrow$"*
   Sei $v_1, \ldots, v_n$ *Erzeugendensystem von* V, *und* $\forall_{i \in \{1,\ldots,n\}}$ *sei*
   $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ *kein Erzeugendensystem von* V.
   Zu zeigen: $v_1, \ldots, v_n$ *l.u.*
   Annahme, $v_1, \ldots, v_n$ *l.a.*
   Dann gibt es $\alpha_1, \ldots, \alpha_n \in I\!R$, *die nicht alle* 0 *sind, mit* $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$.
   Sei $i \in \{1, \ldots, n\}$ *mit* $\alpha_i \neq 0$. *Dann gilt:*
   $v_i = \frac{1}{\alpha_i}(-\alpha_1 v_1 - \ldots - \alpha_{i-1} v_{i-1} - \alpha_{i+1} v_{i+1} - \ldots - \alpha_n v_n)$.
   Daraus folgt, daß bereits $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ *Erzeugendensystem von* V *ist, im Widerspruch zur Annahme.*

6. *(sog. Basisauswahlsatz) Ist* V *ein Vektorraum, und bilden* $v_1, \ldots, v_n \in V$ *ein Erzeugendensystem von* V, *so kann man unter den Vektoren* $v_1, \ldots, v_n$ *eine Basis von* V *auswählen.*

   Beweis:

   (a) *Ist* $v_1, \ldots, v_n$ *bereits eine Basis, so ist nichts zu zeigen. Fertig.*

   (b) *Andernfalls ist* $v_1, \ldots, v_n$ *kein minimales Erzeugendensystem, also kann man einen der Vektor weglassen und erhält weiterhin ein Erzeugendensystem. Zurück zu 6a.*

Somit hat man nach endlich vielen Schritte eine Basis ausgewählt.

7. $v_1, \ldots, v_n$ *Basis von* $V \iff$

   (a) $v_1, \ldots, v_n$ *l.u.*

   (b) $\forall_{v \in V} v_1, \ldots, v_n, v$ *l.a.*

   D.h.: Eine Basis ist eine maximale Menge von l.u. Vektoren, d.h., eine, die nicht als solche vergrößert werden kann.

   Beweis:

   (a) „$\Longrightarrow$“
   Es ist zu zeigen: $\forall_{v \in V} v_1, \ldots, v_n, v$ *l.a.*
   Sei $v \in V$*. Wegen* $V = \langle v_1, \ldots, v_n \rangle$ *gibt es* $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ *mit* $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$*. Daraus folgt*
   $(-1)v + \alpha_1 v_1 + \ldots + \alpha_n v_n = 0$*, also sind diese* $n+1$ *Vektoren l.a.*

   (b) „$\Longleftarrow$“
   Es ist zu zeigen: $V = \langle v_1, \ldots, v_n \rangle$.
   Sei $v \in V$*. Dann sind* $v_1, \ldots, v_n, v$ *l.a., also gibt es* $\alpha_1, \ldots, \alpha_n, \alpha \in \mathbb{R}$*, die nicht alle* 0 *sind, mit* $\alpha_1 v_1 + \ldots + \alpha_n v_n + \alpha v = 0$.
   In disem Fall muß gelten: $\alpha \neq 0$*. Denn aus* $\alpha = 0$ *würde folgen:* $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$*, also, da* $v_1, \ldots, v_n$ *l.u.:* $\forall_i \; \alpha_i = 0$*, ein Widerspruch.*
   Damit ist $v = \frac{1}{\alpha}(-\alpha_1 v_1 - \ldots - \alpha_n v_n)$*, also* $v \in \langle v_1, \ldots, v_n \rangle$.

Satz 4.5.3 (Steinitzscher Austauschsatz für endlich-dimensionale Vektorräume)
V VR, $\{b_1, \ldots, b_n\}$ *Basis.*
$v_1, \ldots, v_k \in V$ *l.u.* $\Longrightarrow$

1. $k \leq n$
2. *Man kann die Basis* $\{b_1, \ldots, b_n\}$ *so numerieren, daß auch* $\{v_1, \ldots, v_k, b_{k+1}, \ldots, b_n\}$ *eine Basis ist.*

Beweis: Vollständige Induktion nach k.

1. $k = 0$: Nichts zu zeigen.
2. Sei $k \in \mathbb{N}$ und die Behauptung vorausgesetzt für l.u. Mengen von k Elementen.

   Seien $v_1, \ldots, v_{k+1} \in V$ l.u.

   (a) Dann sind auch $v_1, \ldots, v_k$ l.u. $\Longrightarrow$ (nach Induktionsvoraussetzung)
   $k \leq n$

   (b) Es wird gezeigt, daß sogar $k < n$ gilt.
   Annahme $k = n \Longrightarrow$ (nach Induktionsvoraussetzung)
   $v_1, \ldots, v_k$ Basis von $V \Longrightarrow$
   $v_{k+1} \in \langle \{v_1, \ldots, v_k\} \rangle \Longrightarrow$
   $v_1, \ldots, v_{k+1}$ nicht l.u.

(c) Nach Induktionsvoraussetzung gilt bei geeigneter Numerierung der b_i:
$\{v_1, \ldots, v_k, b_{k+1}, \ldots, b_n\}$ Basis. $\Longrightarrow$
$\exists \alpha_1, \ldots, \alpha_n \in \mathbb{R}\ v_{k+1} = \sum_{i=1}^{k} \alpha_i v_i + \sum_{i=k+1}^{n} \alpha_i b_i$.
Da $v_1, \ldots, v_{k+1}$ l.u. sind, muß mindestens einer der Koeffizienten $\alpha_{k+1}, \ldots, \alpha_n$ von Null verschieden sein (sonst wäre v_{k+1} eine L.K. von $v_1, \ldots, v_k$)
Nach evtl. Umnumerierung der $b_{k+1}, \ldots, b_n$ kann man annehmen: $\alpha_{k+1} \neq 0$

(d) $\{v_1, \ldots, v_{k+1}, b_{k+2}, \ldots, b_n\}$ l.u.
Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\sum_{i=1}^{k+1} \lambda_i v_i + \sum_{i=k+2}^{n} \lambda_i b_i = 0$. $\Longrightarrow$
$0 = \sum_{i=1}^{k} \lambda_i v_i + \lambda_{k+1} v_{k+1} + \sum_{i=k+2}^{n} \lambda_i b_i =$
$\sum_{i=1}^{k} \lambda_i v_i + \lambda_{k+1}(\sum_{i=1}^{k} \alpha_i v_i + \sum_{i=k+1}^{n} \alpha_i b_i) + \sum_{i=k+2}^{n} \lambda_i b_i =$
$\sum_{i=1}^{k}(\lambda_i + \lambda_{k+1}\alpha_i) v_i + \lambda_{k+1}\alpha_{k+1} b_{k+1} + \sum_{i=k+2}^{n}(\lambda_{k+1}\alpha_i + \lambda_i) b_i$
Da $\{v_1, \ldots, v_k, b_{k+1}, \ldots, b_n\}$ Basis, also l.u. ist, verschwinden alle Koeffizienten, insbesondere gilt
$\lambda_{k+1}\alpha_{k+1} = 0 \Longrightarrow$ wegen $\alpha_{k+1} \neq 0$
$\lambda_{k+1} = 0$
Damit folgt $\forall i = 1, \ldots, n\ [\lambda_i = 0]$.

(e) $\langle \{v_1, \ldots, v_{k+1}, b_{k+2}, \ldots, b_n\} \rangle = V$.
Sei $v \in V$.
$\{v_1, \ldots, v_k, b_{k+1}, \ldots, b_n\}$ Basis $\Longrightarrow$
$\exists \beta_1, \ldots, \beta_n \in \mathbb{R}\ [v = \sum_{i=1}^{k} \beta_i v_i + \sum_{i=k+1}^{n} \beta_i b_i]$.
Wegen $b_{k+1} = \frac{1}{\alpha_{k+1}}(v_{k+1} - \sum_{i=1}^{k} \alpha_i v_i - \sum_{i=k+2}^{n} \alpha_i b_i)$
folgt $b_{k+1} \in \langle \{v_1, \ldots, v_{k+1}, b_{k+2}, \ldots, b_n\} \rangle \Longrightarrow$
$v \in \langle \{v_1, \ldots, v_{k+1}, b_{k+2}, \ldots, b_n\} \rangle$.

Korollar 4.5.4 *V VR; in V gebe es eine endliche Basis B der Länge n.*

1. *Jede Basis von V ist endlich.*
2. *Jede Basis von V hat n Elemente.*
3. *Sind $v_1, \ldots, v_r \in V$ l.u., so gibt es Vektoren $v_{r+1}, \ldots, v_n \in B$, so daß $v_1, \ldots, v_n$ eine Basis von V ist (sog. Basisergänzungssatz)).*

Beweis:

1. B_1 Basis von $V \Longrightarrow B_1$ l.u. $\Longrightarrow$ (nach 4.5.3) $|B_1| \leq n$.
2. B_1 Basis von V, $k := |B_1|$. Dann $k \leq n$ (voriger Punkt).
   Durch Vertauschen der Rollen von B und B_1 erhält man: $n \leq k$.

Bemerkung 4.5.5 *Sei V ein endlich-dimensionaler VR , $U \subset V$ UVR. Dann:*

1. *U ist endlich erzeugt, d.h.: Es gibt endlich viele Vektoren $u_1, \ldots, u_m \in U$ mit $U = \langle u_1, \ldots, u_m \rangle$.*

   Beweis: Annahme, U wäre nicht endlich erzeugt. Sind dann $m \in I\!N$ und l.u. Vektoren $u_1, \ldots, u_m \in U$ gegeben, so gilt $\langle u_1, \ldots, u_m \rangle \underset{\neq}{\subset} U$. Es gibt daher ein $u_{m+1} \in U \setminus \langle u_1, \ldots, u_m \rangle$; insbesondere sind $u_1, \ldots, u_m, u_{m+1}$ l.u.

   Auf diese Weise kann man — unter der Annahme, U ist nicht endlich erzeugt — eine unendliche Menge von l.u. Vektoren in $U \subset V$ konstruieren. Dies steht im Widerspruch zum Steinitzschen Austauschsatz.

2. $\dim U \leq \dim V$

3. $\dim U = \dim V \Longleftrightarrow U = V$

Beispiel 4.5.6 *Basis des Lösungsraums eines homogenen linearen Gleichungssystems.*
Sei A eine $m \times n$-Stufenmatrix von Rang r; das Pivot-Element in der μ-ten Zeile sei $a_{\mu\nu_\mu}$, $\mu = 1, \ldots, r$. Sei $\Phi : I\!R^{n-r} \to L(A)$ die Parametrisierungsabbildung für die Lösungsmenge $L(A)$.
Behauptung: Eine Basis von $L(A)$ ist gegeben durch $l_1, \ldots, l_{n-r}$ mit $l_i := \Phi(e_i)$.

Denn man zeigt auf einfache Weise, daß die l_i l.u. sind und den Lösungsraum erzeugen.

Auf entsprechende Weise kann die Lösungsmenge $L(A, b)$ eines inhomogenen linearen Gleichungssystems $Ax = b$ dargestellt werden; denn es ist $L(A, b) = y + L(A)$, wenn y irgendeine (feste) Lösung des inhomogenen Systems ist.

4.6 Isomorphismen, Dimensionsformel für lineare Abbildungen.

Definition 4.6.1 *Es seien V, W Vektorräume und $f : V \to W$ eine lineare Abbildung. f heißt*

1. *Monomorphismus* $:\Longleftrightarrow \mathbf{Ke} f = 0$

2. *Epimorphismus* $:\Longleftrightarrow \mathbf{Im} f = W$

3. *Isomorphismus* $:\Longleftrightarrow$ *f Monomorphismus und Epimorphismus,*

Bemerkung 4.6.2 *Bezeichnungen wie oben.*

1. *f Monomorphismus* $\Longleftrightarrow$ f injektiv, *d.h* $\forall_{x,y \in V}\ f(x) = f(y) \Longrightarrow u = v$.

   Beweis:

   (a) „$\Longrightarrow$“
   Seien $x, y \in V$ mit $f(x) = f(y)$. Zu zeigen: $x = y$.
   Beweis: $f(x - y) = f(x) - f(y) = 0 \Longrightarrow x - y \in \mathbf{Ke} f \Longrightarrow x - y = 0 \Longrightarrow x = y$.

(b) „$\Longleftarrow$“

Klar.

2. *Ein Epimorphismus f wird auch eine surjektive Abbildung oder „Abbildung auf“ genannt.*

3. *f Isomorphismus $\Longleftrightarrow$ f besitzt eine lineare Umkehrabbildung $g : W \to V$, d.h., eine lineare Abbildung $g : W \to V$ mit folgenden Eigenschaften: $g \circ f = \mathrm{id}_V, f \circ g = \mathrm{id}_W$.*

   Zum Beweis: Es ist insbesondere zu zeigen: Besitzt eine lineare Abbildung $f : V \to W$ eine Umkehrabbildung $g : W \to V$, so ist g auch linear.

   Seien $w_1, w_2 \in W$ und $\lambda_1, \lambda_2 \in I\!R$; zu zeigen: $g(\lambda_1 w_1 + \lambda_2 g_2) = \lambda_1 g(w_1) + \lambda_2 g(w_2)$.

   Sei $v_1 := g(w_1), v_2 := g(w_2)$. Dann gilt:

   $\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 f(v_1) + \lambda_2 f(v_2) =$ (da f linear) $= f(\lambda_1 v_1 + \lambda_2 v_2)$.

   Anwendung von g liefert

   $g(\lambda_1 w_1 + \lambda_2 w_2) = g(f(\lambda_1 v_1 + \lambda_2 v_2)) = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 g(w_1) + \lambda_2 g(w_2)$.

4. *$f : V \to W$ Monomorphismus, $v_1, \ldots, v_n \in V$ l.u. $\Longrightarrow$ $f(v_1), \ldots, f(v_n) \in W$ l.u.*

   Beweis. Seien $\alpha_1, \ldots, \alpha_n \in I\!R$ mit $\alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) = 0$. Zu zeigen: $\alpha_1 = \ldots = \alpha_n = 0$.

   Es gilt $f(\alpha_1 v_1 + \ldots + \alpha_n v_n) = \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) = 0$; da f injektiv ist, folgt hieraus: $(\alpha_1 v_1 + \ldots + \alpha_n v_n$.

   Die Behauptung folgt, da $v_1, \ldots, v_n \in V$ l.u. sind.

5. *$f : V \to W$ lineare Abbildung, $v_1, \ldots, v_n \in V$ Erzeugendensystem von V $\Longrightarrow f(v_1), \ldots, f(v_n) \in W$ Erzeugendensystem von $\mathbf{Im} f$.*

   Beweis: $w \in \mathbf{Im} f \Longrightarrow \exists_{v \in V}\ f(v) = w$.

   $v_1, \ldots, v_n \in V$ Erzeugendensystem von $V \Longrightarrow \exists_{\alpha_1, \ldots, \alpha_n \in I\!R}\ v = \alpha_1 v_1 + \ldots + \alpha_n v_n \Longrightarrow$

   $w = f(v) = f(\alpha_1 v_1 + \ldots + \alpha_n v_n) = \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) \Longrightarrow w \in \langle f(v_1), \ldots, f(v_n) \rangle$.

6. *$f : V \to W$ Epimorphismus, $v_1, \ldots, v_n \in V$ Erzeugendensystem von $V \Longrightarrow f(v_1), \ldots, f(v_n) \in W$ Erzeugendensystem von W.*

   Spezialfall der vorherigen Aussage.

7. *Die vorherigen Aussagen gelten auch für ∞e l.u. Mengen bzw. Erzeugendensysteme.*

8. *$f : V \to W$ Isomorphismus $\Longrightarrow$ Das Bild einer Basis von V ist eine Basis von W, insbesondere: $\dim V = \dim W$.*

9. *Ein Monomorphismus $f : V \to W$ definiert einen Isomorphismus auf das Bild, Bezeichnung $f|V \to \mathbf{Im}f$.*

Satz 4.6.3 *Sei A eine $m \times n$-Stufenmatrix von Rang r. Dann gilt:*

1. *Die Parametrisierungsabbildung $\Phi : \mathbb{R}^{n-r} \to L(A)$ ist ein Isomorphismus.*
2. *Eine Basis von $L(A)$ ist gegeben durch $l_1, \ldots, l_{n-r}$ mit $l_i := \Phi(e_i)$.*
3. $\dim L(A) = n - r$

Der Beweis wurde eigentlich schon erbracht, genauer:

1. Es wurde bereits gezeigt:
   (a) Φ ist nach Konstruktion linear und injektiv, also ein Monomorphismus.
   (b) $\Phi : \mathbb{R}^{n-r} \to L(A)$ ist surjektiv.
2. Nach Bemerkung 8 oben sind die l_i eine Basis von $L(A)$. Insbesondere gilt:
3. $\dim L(A) = n - r$.

Satz 4.6.4 (Dimensionsformel für lineare Abbildungen) *V, W VR, $\dim V < \infty$, $f : V \longrightarrow W$ lineare Abbildung $\Longrightarrow$*
$\dim \mathbf{Im}f = \dim V - \dim \ker f$

Bei einer linearen Abbildung kann die Dimension von $\mathbf{Im}f$ also nie größer sein als $\dim V$; sie ist vielmehr um $\dim \mathbf{Ke}f$ kleiner als $\dim V$. Im Falle eines Monomorphismus oder Isomorphismus gilt $\dim Kef = 0$, also (in Übereinstimmung mit Bemerkung oben) $\dim \mathbf{Im}f = \dim V$.
Beweis: Sei $r := \dim \mathbf{Ke}f$, und sei $v_1, \ldots, v_r$ eine Basis von $\mathbf{Ke}f$. Dann kann man diese Basis von $\mathbf{Ke}f$ zu einer Basis $v_1, \ldots, v_n$ von V ergänzen.
O.E. ist $r < n$. Sei $U := \langle v_{r+1}, \ldots, v_n \rangle$. Dann $\dim U = n - r$.
Es reicht zu zeigen:
Die Abbildung $g := f|U \to W$, definiert also durch $g(u) := f(u)$ für $u \in U$, ist ein Isomorphismus.

1. g Monomorphismus, also $\mathbf{Ke}g = 0$.

   Seu $u \in \mathbf{Ke}g$. Dann ist $u \in U = \langle v_{r+1}, \ldots, v_n \rangle$ und $f(u) = 0$, also $u \in \mathbf{Ke}f = \langle v_1, \ldots, v_r \rangle$, also

   $u \in \langle v_1, \ldots, v_r \rangle \cap \langle v_{r+1}, \ldots, v_n \rangle$.

   Es gibt also $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit

   $u = \alpha_1 v_1 + \ldots + \alpha_r v_r = \alpha_{r+1} v_{r+1} + \ldots + \alpha_n v_n$. Daras folgt $\alpha_1 v_1 + \ldots + \alpha_r v_r - \alpha_{r+1} v_{r+1} - \ldots - \alpha_n v_n = 0$. Da $v_1, \ldots, v_n$ l.u. sind, verschwinden alle Koeffizienten α_i. Deshalb ist $u = 0$.

2. g Epimorphismus, also $\mathbf{Im}g = \mathbf{Im}f$.

   Wegen $U \subset V$ gilt $\mathbf{Im}g \subset \mathbf{Im}f$; zu zeigen ist: $\mathbf{Im}f \subset \mathbf{Im}g$.

   Sei $w \in \mathbf{Im}f$. Dann gibt es ein $v \in V$ mit $f(v) = w$. Da $v_1, \ldots, v_n$ eine Basis von V ist, gibt es $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit $u = \alpha_1 v_1 + \ldots + \alpha_r v_r + \alpha_{r+1} v_{r+1} + \ldots + \alpha_n v_n$.

   Wegen $\alpha_1 v_1 + \ldots + \alpha_r v_r \in \mathbf{Ke}f$ folgt

   $f(u) = f(\alpha_{r+1} v_{r+1} + \ldots + \alpha_n v_n \in \mathbf{Im}g$.

Beispiel 4.6.5 *Sei A eine $m \times n$-Stufenmatrix mit r von Null verschiedenen Zeilen und $\hat{A} : \mathbb{R}^n \to \mathbb{R}^m,\ x \mapsto Ax$.*
Dann gilt nach 4.6.3: $\dim L(A) = n - r$. *Daraus folgt:* $\dim \mathbf{Im}\hat{A} = r$. *Im nächsten Abschnitt wird man diese Aussage verallgemeinern und besser verstehen.*

Bemerkung 4.6.6 *Sei $A \in \mathbb{R}^{n \times n}$ eine invertierbare Matrix. Dann ist die durch A definierte lineare Abbildung $\hat{A} : \mathbb{R}^n \to \mathbb{R}^n$ ein Isomorphismus.*

Beweis: $\widehat{A^{-1}}$ ist Umkehrabbildung von $\hat{A}$.

4.7 Der Rang einer Matrix

Definition 4.7.1 *Sei sei $A \in \mathbb{R}^{m \times n}$ eine $m \times n$-Matrix. Seien $a_i := (a_{i1}, \ldots, a_{in}) \in K_n$, $i = 1, \ldots, m$ die Zeilenvektoren von A und $a^j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m,\ j = 1, \ldots, n$ die Spaltenvektoren von A.*

1. *$Z(A) := \langle a_1, \ldots, a_m \rangle \subset \mathbb{R}_n$ heißt der Zeilenraum von A,*

   $\mathrm{ZRg}(A) := \dim Z(A)$ *heißt Zeilenrang von A.*

2. *$S(A) := \langle a^1, \ldots, a^n \rangle \subset \mathbb{R}^m$ heißt der Spaltenraum von A,*

   $\mathrm{SRg}(A) := \dim S(A)$ *heißt Spaltenrang von A.*

Bemerkung 4.7.2

1. *Elementare Zeilenumformungen ändern den Zeilenraum nicht, also auch nicht den Zeilenrang.*

2. *Elementare Spaltenumformungen ändern den Spaltenraum nicht, also auch nicht den Spaltenrang.*

3. *Für eine Stufenmatrix A gilt* $\mathrm{ZRg}(A) = \mathrm{Rg}(A) = \mathrm{SRg}(A)$.

   Vgl. dazu auch 3.7.5

4. *Für $i = 1, \ldots, n$ gilt $Ae_i = a^i$.*

5. *Für die lineare Abbildung $\hat{A} : \mathbb{R}^n \to \mathbb{R}^m,\ x \mapsto Ax$ gilt:*

$$S(A) = \mathbf{Im}\ \hat{A}$$

Beweis: Sei $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m$. *Dann:*

$y \in \mathbf{Im}\hat{A} \iff$

$\exists_{x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n} \quad y = Ax.$

Setzt man hier ein: $x = x_1e_1 + \ldots + x_ne_n$, *so wird die letzte Gleichung*

$y = A(x_1e_1 + \ldots + x_ne_n) = x_1Ae_1 + \ldots + x_nAe_n = x_1a^1 + \ldots + x_na^n$. *Damit:*

$y \in \mathbf{Im}\hat{A} \iff$

$\exists_{x_1,\ldots,x_n \in K}\ y = x_1a^1 + \ldots + x_na^n \iff y \in \langle a^1, \ldots, a^n \rangle = S(A)$.

Es folgt:

6. $\mathrm{SRg}(A) = \dim \mathbf{Im}\ \hat{A}$.

7. *Elementare Zeilenumformungen ändern den Spaltenrang nicht.*

   Beweis: $\bar{A}$ *entstehe aus* A *durch elementare Zeilenumformungen. Dann gibt es Elementarmatrizen* $S_1, \ldots, S_k$ *mit* $\bar{A} = S_k \cdot \ldots \cdot S_1 \cdot A$. *Sei* $S := S_k \cdot \ldots \cdot S_1$.

   Dann ist S *invertierbar und* $\mathbf{Im}\bar{A} = \hat{S}(\mathbf{Im}A)$. *Da* $\hat{S}$ *ein Isomorphismus ist, folgt* $\dim \mathbf{Im}\bar{A} = \dim \mathbf{Im}A$.

Insgesamt ergibt sich:

Satz 4.7.3 *Für jede Matrix* A *gilt* $\mathrm{ZRg}(A) = \mathrm{SRg}(A)$

Definition 4.7.4 *Für eine Matrix* A *sei* $\mathrm{Rg}(A) := \mathrm{ZRg}(A) = \mathrm{SRg}(A)$. *(Rang von* A*)*

Satz 4.7.5 $A \in \mathbb{R}^{n \times n}$ *invertierbar* $\iff \mathrm{Rg}(A) = n$

Satz 4.7.6 *Sei* $A \in \mathbb{R}^{m \times n}$. *Dann*
$\hat{A} : \mathbb{R}^n \to \mathbb{R}^m$ *Isomorphismus* $\iff m = n$ *und* $\mathrm{Rg}(A) = n$

4.8 Die einer linearen Abbildung zugeordnete Matrix

Lemma 4.8.1 (Prinzip der linearen Fortsetzung) *Es seien* V *und* W *Vektorräume und* $v_1, \ldots, v_n$ *eine Basis von* V.
Zu vorgegebenen Vektoren $w_1, \ldots, w_n \in W$ *gibt es genau eine lineare Abbildung* $f : V \to W$ *mit* $f(v_i) = w_i, i = 1, \ldots, n$.
Kurz: Ist $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine Basis von V, so läßt sich jede Abbildung $\mathcal{V} \to W$ eindeutig zu einer linearen Abbildung $V \to W$ fortsetzen.
Beweis:
Seien $w_1, \ldots, w_n \in W$ gegeben.

1. Existenz einer linearen Abbildung $f : V \to W$ mit $f(v_i) = w_i, i = 1, \ldots, n$:

   (a) Definition der Abbildung
   Sei $v \in V$; es ist $f(v) \in W$ anzugeben.
   v hat eine Darstellung $v = x_1 v_1 + \ldots + x_n v_n$ mit eindeutig bestimmten Koeffizienten $x_1, \ldots, x_n \in \mathbb{R}$. Setze $f(v) := x_1 w_1 + \ldots + x_n w_n$.

   (b) Linearität der so definierten Abbildung f:
   Sind $v = x_1 v_1 + \ldots + x_n v_n$ und $v' = x'_1 v_1 + \ldots + x'_n v_n$ sowie $\alpha, \alpha' \in \mathbb{R}$, so gilt
   $f(\alpha v + \alpha' v') =$
   $f((\alpha x_1 + \alpha' x'_1)v_1 + \ldots + (\alpha x_n + \alpha' x'_n)v_n) =$
   $(\alpha x_1 + \alpha' x'_1)w_1 + \ldots + (\alpha x_n + \alpha' x'_n)w_n =$
   $\alpha f(v) + \alpha' f(v')$

2. f ist eindeutig bestimmt:

   Seien $f, g : V \to W$ lineare Abbildungen mit $f(v_i) = w_i, i = 1, \ldots, n$, $g(v_i) = w_i, i = 1, \ldots, n$.

   Dann gilt für $v = x_1 v_1 + \ldots + x_n v_n \in V$:

   $f(v) = f(x_1 v_1 + \ldots + x_n v_n) = x_1 w_1 + \ldots + x_n w_n = g(x_1 v_1 + \ldots + x_n v_n) = g(v)$.

Bemerkungen und Beispiele 4.8.2 *1. Sei V ein Vektorraum und $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine Basis von V. Dann gibt es genau eine lineare Abbildung $\iota_{\mathcal{V}} : \mathbb{R}^n \longrightarrow V$ mit $\iota_{\mathcal{V}}(e_i) = v_1, i = 1, \ldots, n$, nämlich*

$$\iota_{\mathcal{V}} : \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & V \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & \sum_{i=1}^n x_i v_i \end{array}$$

2. Seien $m, n \in \mathbb{N}$ und $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ eine lineare Abbildung. Dann gibt es genau eine Matrix $A \in \mathbb{R}^{(m,n)}$ mit $f = \hat{A}$.

Beweis: Ist $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ gegeben, so sei $A \in \mathbb{R}^{(m,n)}$ die Matrix mit den Spaltenvektoren $a^i = f(e_i), i = 1, \ldots, n$. Die Aussage folgt dann aus 4.8.1.

3. Drehmatrix: Bei Drehung der Ebene $\mathbb{R}^2$ um den Nullpunkt um den Winkel φ gehen die kanonischen Basisvektoren e_1, e_2 über in die Vektoren

$$\begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} \text{und} \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$$

Die Drehung wird also bzgl. der kanonischen Basis durch die Matrix

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

beschrieben.

Satz 4.8.3 *1. Es seien*

(a) V ein n-dimensionaler Vektorraum und $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine Basis von V,

(b) W ein m-dimensionaler Vektorraum und $\mathcal{W} = \{w_1, \ldots, w_m\}$ eine Basis von w,

(c) $f : V \longrightarrow W$ eine lineare Abbildung.

Dann gibt es genau eine $m \times n$-Matrix $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$, so daß das folgende Diagramm kommutiert:

$$\begin{array}{ccccc} & V & \xrightarrow{f} & W & \\ \iota_{\mathcal{V}} & \uparrow & & \uparrow & \iota_{\mathcal{W}} \\ & \mathbb{R}^n & \xrightarrow{\hat{A}} & \mathbb{R}^m & \end{array}$$

Die Matrix A heißt die der linearen Abbildung f bei Wahl der Basen $\mathcal{V}$ und $\mathcal{W}$ **zugeordnete Matrix**, *Bezeichnung: $A = \mathcal{M}^{\mathcal{V}}_{\mathcal{W}}(f)$.*

2. Der Komposition von linearen Abbildungen entspricht das Produkt der zugeordneten Matrizen.

3. Im Fall $m = n$ gilt: Ist f ein Isomorphismus, so ist die zugeordnete Matrix regulär und der Umkehrabbildung ist die inverse Matrix zugeordnet.

Beweis:
Anwendung der Bemerkung oben auf die lineare Abbildung $\iota_{\mathcal{W}}^{-1} \circ f \circ \iota_{\mathcal{V}}$

Bemerkung 4.8.4 *Die Spaltenvektoren von A sind also die den Bildern $f(v_i)$ bzgl. der Basis $\mathcal{W}$ zugeordneten Spalten, d.h.:*
Für $\nu = 1, \ldots, n$: $f(v_\nu)) = \sum_{\mu=1}^{m} a_{\mu\nu} w_\mu$.
Dies folgt aus $\hat{A}(e_\nu) = Ae_\nu = \sum_{\mu=1}^{m} a_{\mu\nu} e_\mu$.

4.9 Basiswechsel (Koordinatentransformationen)

Der folgende Spezialfall einer zugeordneten Matrix ist von besonderem Interesse:

Definition 4.9.1 *Sei V ein n-dimensionaler VR, und seien $\mathcal{V} = \{v_1, \ldots, v_n\}$ und $\mathcal{W} = \{w_1, \ldots, w_m\}$ Basen von V. Die Matrix*

$$\mathcal{T}^{\mathcal{V}}_{\mathcal{W}} := \mathcal{M}^{\mathcal{V}}_{\mathcal{W}}(\mathrm{id}_V)$$

heißt die zum **Basiswechsel** *$\mathcal{V}$ zu $\mathcal{W}$ gehörige* **Transformationsmatrix**.

Bemerkungen und Beispiele 4.9.2 *1. Die Transformationsmatrix ist also durch folgendes kommutatives Diagramm definiert:*

$$\begin{array}{ccccc} & V & \xrightarrow{id_V} & V & \\ \iota_{\mathcal{V}} & \uparrow & & \uparrow & \iota_{\mathcal{W}} \\ & \mathbb{R}^n & \xrightarrow{\hat{\mathcal{T}}^{\mathcal{V}}_{\mathcal{W}}} & \mathbb{R}^m & \end{array}$$

2. *Die Spaltenvektoren von $\mathcal{T}_{\mathcal{W}}^{\mathcal{V}}$ enthalten die Koeffizienten bei der Darstellung der Basisvektoren v_i bzgl. der neuen Basis.*

3. *$V = \mathbb{R}^n$, $\mathcal{V} = \{v_1, \ldots, v_n\}$ und $\mathcal{W} = \{w_1, \ldots, w_m\}$ Basen. Dann ist $\iota_{\mathcal{V}}$ bzgl. der kanonischen Basis durch eine Matrix A gegeben, die die Vektoren $v_1, \ldots, v_n$ als Spaltenvektoren hat, und $\iota_{\mathcal{W}}$ ist durch eine Matrix W mit den Vektoren $w_1, \ldots, w_n$ als Spaltenvektoren gegeben.*

   Die Transformationsmatrix $\mathcal{T}_{\mathcal{W}}^{\mathcal{V}}$ ist dann $B^{-1}A$.

   Denn das kommutative Diagramm

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "id"] & \mathbb{R}^n \\ \mathbb{R}^n \arrow[u, "\iota_{\mathcal{V}}"] \arrow[r, "\hat{\mathcal{T}}_{\mathcal{W}}^{\mathcal{V}}"] & \mathbb{R}^m \arrow[u, "\iota_{\mathcal{W}}"'] \end{tikzcd}$$

   ist

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "id"] & \mathbb{R}^n \\ \mathbb{R}^n \arrow[u, "\hat{A}"] \arrow[r, "\hat{\mathcal{T}}_{\mathcal{W}}^{\mathcal{V}}"] & \mathbb{R}^m \arrow[u, "\hat{B}"'] \end{tikzcd}$$

   Daraus liest man ab: $B\mathcal{T}_{\mathcal{W}}^{\mathcal{V}} = A$; daraus folgt die Behauptung.

4. *Ist $V = \mathbb{R}^n$ und speziell $\mathcal{E} = \{e_1, \ldots, e_n\}$ die kanonische Basis, $\mathcal{W} = \{w_1, \ldots, w_n\}$ eine weitere Basis. Dann ist $\iota_{\mathcal{E}}$ die identische Abbildung, während $\iota_{\mathcal{W}}$ durch eine Matrix B gegeben ist, die die Vektoren $w_1, \ldots, w_n$ als Spaltenvektoren hat.*

   Die Transformationsmatrix $\mathcal{T}_{\mathcal{W}}^{\mathcal{E}}$ ist dann die zu B inverse Matrix B^{-1}.

5. *Es gilt stets $\mathcal{T}_{\mathcal{W}}^{\mathcal{V}} = (\mathcal{T}_{\mathcal{V}}^{\mathcal{W}})^{-1}$*

6. *Wie hatten bereits die Matrix einer Drehung im $\mathbb{R}^2$ bestimmt; in entsprechender Weise erhält man die Matrix für eine Drehung im $\mathbb{R}^3$, wenn die Drehachse eine Achse des kanonischen Koordinatensystems ist. Z.B. wird eine Drehung um die x_3-Achse, also um die durch den Vektor e_3 aufgespannte Gerade, durch folgende Matrix beschrieben:*

$$\begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

   Zur Berechnung der Matrix einer Drehung um eine beliebige Achse empfiehlt es sich, zunächst eine Koordinatentransformation vorzunehmen:

7. *Sei* $f : I\!R^3 \to I\!R^3$ *die Drehung um den Winkel* φ *um die durch den Vektor* $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \in I\!R^3$ *gegebene Gerade (Winkelhalbierende der* x_2, x_3-*Ebene). Wir wollen die Matrix* $M = \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(f)$ *von* f *bzgl. der kanonischen Basis* $\mathcal{E} = \{e_1, e_2, e_3\}$ *berechnen.*

   Eine für die Behandlung dieser Abbildungsaufgabe geeignete Basis $\mathcal{W} = \{w_1, w_2, w_3\}$ *erhält man, wenn man die kanonische Basis* $\{e_1, e_2, e_3\}$ *um die* x_1-*Achse um den Winkel* $\pi/4 = 45°$ *dreht; man erhält die Basisvektoren*

   $$w_1 = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

   $$w_2 = \begin{pmatrix} 0 \\ \cos\frac{\pi}{4} \\ \sin\frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix},$$

   $$w_3 = \begin{pmatrix} 0 \\ -\sin\frac{\pi}{4} \\ \cos\frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}.$$

   Sei B *die Matrix mit den Spaltenvektoren* w_1, w_2 *und* w_3*, also*

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

   Die Transformationsmatrix des Basiswechsels ist dann

   $$\mathcal{T}_{\mathcal{W}}^{\mathcal{E}} = B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

   Bemerkung: Da die Matrix B *in diesem Fall eine Drehmatrix ist, erhält man die inverse Matrix* B^{-1} *in einfacher Weise als die Drehmatrix zum Drehwinkel* $-\pi/4 = -45°$ *um die* x_1-*Achse.*

   In Bezug auf diese neue Basis $\mathcal{W}$ *haben wir nun eine Drehung um die durch* w_2 *bestimmte Achse; bzgl. dieser Basis wird die Drehung um den Winkel* φ *daher durch die Matrix*

   $$N = \mathcal{M}_{\mathcal{W}}^{\mathcal{W}}(f) = \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix}$$

   beschrieben.

Um die die Drehung beschreibende Matrix $M = \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(f)$ bzgl. der kanonischen Basis $\mathcal{E} = \{e_1, e_2, e_3\}$ zu erhalten, muß zuerst die Transformationsmatrix für den Basiswechsel, also die Matrix B^{-1} angewandt werden, dann die Drehmatrix N und schließlich wieder die Transformationsmatrix des umgekehrten Basiswechsels, also die Matrix B.

Insgesamt ergibt sich als Drehmatrix bzgl. der kanonischen Basis:

$$M = BNB^{-1} = \begin{pmatrix} \cos(\varphi) & 1/2\,\sqrt{2}\sin(\varphi) & -1/2\,\sqrt{2}\sin(\varphi) \\ -1/2\,\sqrt{2}\sin(\varphi) & 1/2 + 1/2\,\cos(\varphi) & 1/2 - 1/2\,\cos(\varphi) \\ 1/2\,\sqrt{2}\sin(\varphi) & 1/2 - 1/2\,\cos(\varphi) & 1/2 + 1/2\,\cos(\varphi) \end{pmatrix}$$

(Berechnung der Matrizenmultiplikation mit Maple.)

Man kann dies auch so ausdrücken:

$$\mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(f) = \mathcal{T}_{\mathcal{E}}^{\mathcal{W}} \mathcal{M}_{\mathcal{W}}^{\mathcal{W}}(f) \mathcal{T}_{\mathcal{W}}^{\mathcal{E}}$$

Der folgende Satz verallgemeinert die Situation des letzten Beispiels:
Satz 4.9.3 *Es seien*
Es seien V ein n-dimensionaler Vektorraum und $\mathcal{V} = \{v_1, \ldots, v_n\}$, $\mathcal{W} = \{w_1, \ldots, w_n\}$ Basen von V.
Sei ferner $f : V \longrightarrow V$ eine lineare Abbildung. Dann gilt

$$\mathcal{M}_{\mathcal{W}}^{\mathcal{W}}(f) = \mathcal{T}_{\mathcal{W}}^{\mathcal{V}} \mathcal{M}_{\mathcal{V}}^{\mathcal{V}}(f) \mathcal{T}_{\mathcal{V}}^{\mathcal{W}}$$

Beweis: aus dem entsprechenden Diagramm liest man ab: $\mathcal{M}_{\mathcal{W}}^{\mathcal{W}}(f)\mathcal{T}_{\mathcal{W}}^{\mathcal{V}} = \mathcal{T}_{\mathcal{W}}^{\mathcal{V}}\mathcal{M}_{\mathcal{V}}^{\mathcal{V}}(f)$.
Daraus folgt $\mathcal{M}_{\mathcal{W}}^{\mathcal{W}}(f) = \mathcal{T}_{\mathcal{W}}^{\mathcal{V}}\mathcal{M}_{\mathcal{V}}^{\mathcal{V}}(f)(\mathcal{T}_{\mathcal{W}}^{\mathcal{V}})^{-1} = \mathcal{T}_{\mathcal{W}}^{\mathcal{V}}\mathcal{M}_{\mathcal{V}}^{\mathcal{V}}(f)\mathcal{T}_{\mathcal{V}}^{\mathcal{W}}$

Allgemeiner gilt:
Satz 4.9.4 *Es seien*
Es seien

— *V ein n-dimensionaler, W ein m-dimensionaler Vektorraum,*

— *$\mathcal{V} = \{v_1, \ldots, v_n\}$ und $\mathcal{V}' = \{v'_1, \ldots, v'_n\}$ Basen von V,*

— *$\mathcal{W} = \{w_1, \ldots, w_m\}$ und $\mathcal{W}' = \{w'_1, \ldots, w'_m\}$ Basen von W.*

Sei ferner $f : V \longrightarrow W$ eine lineare Abbildung. Dann gilt

$$\mathcal{M}_{\mathcal{W}'}^{\mathcal{V}'}(f) = \mathcal{T}_{\mathcal{W}'}^{\mathcal{W}} \mathcal{M}_{\mathcal{W}}^{\mathcal{V}}(f) \mathcal{T}_{\mathcal{V}}^{\mathcal{V}'}$$

4.10 Ähnliche und äquivalente Matrizen

Definition 4.10.1 *Zwei Matrizen $A, B \in I\!R^{(m.n)}$ heißen*

1. **äquivalent**, *wenn es invertierbare Matrizen $S \in I\!R^{(m,m)}$ und $T \in I\!R^{(n,n)}$ gibt mit $B = SAT^{-1}$.*

2. **ähnlich**, *wenn* $m = n$ *gilt und wenn es eine invertierbare Matriz* $S \in \mathbb{R}^{(n,n)}$ *gibt mit* $B = SAS^{-1}$.

Bemerkung 4.10.2 *1. Zwei Matrizen* A *und* B *sind äquivalent genau dann, wenn sie bei geigneten Koordinatentransformationen im* $\mathbb{R}^m$ *bzw.* $\mathbb{R}^n$ *die gleiche lineare Abbildung erzeugen.*

Dies ergibt sich aus dem kommutativen Diagramm

$$\begin{array}{cccccc} & \mathbb{R}^n & \xrightarrow{\hat{A}} & \mathbb{R}^m & \\ \hat{T} & \downarrow & & \downarrow & \hat{S} \ , \\ & \mathbb{R}^n & \xrightarrow{\hat{B}} & \mathbb{R}^m & \end{array}$$

Vgl. auch die letzten Sätze 4.9.3, 4.9.4 des vorigen Abschnitts.

2. $\mathrm{Rg}(A) = r \Longrightarrow A$ *äquivalent zu Normalform*

$$\begin{pmatrix} E_r & 0 \\ & \\ 0 & 0 \end{pmatrix} = \ldots$$

3. A, B *äquivalent* $\Longleftrightarrow \mathrm{Rg}(A) = \mathrm{Rg}(B)$.

4. A, B *ähnlich* $\Longleftrightarrow B$ *geht aus* A *durch Anwendung derselben Koordinatentransformation im Urbild- und im Bildraum hervor.*

5. *Eine der wichtigsten Fragestellungen der linearen Algebra ist es, eine Problemstellung durch geeignete Basistransformationen zu vereinfachen (Eigenwerttheorie, Hauptachsentransformation, Jordansche Normalform).*

   Solche Fragestellungen werden für einfache Fälle später in dieser Vorlesung behandelt.

5 Determinanten

5.1 Permutationen

Definition 5.1.1 *Sei $n > 0$ eine natürliche Zahl. $S_n :=$ Menge der bijektiven Abbildungen $\{1, \dots, n\} \longrightarrow \{1, \dots, n\}$. Die Elemente von S_n heißen* **Permutationen** *der Elemente von $\{1, \dots, n\}$.*

Bemerkung 5.1.2

1. *Ist M eine beliebige endliche Menge, so wird die Menge $Aut(M)$ der bijektiven Abbildungen von M auf sich ebenfalls als Menge der Permutationen bezeichnet.*

2. *Sei M eine beliebige endliche Menge und $f : \{1, \dots, n\} \longrightarrow M$ eine bijektive Abbildung. Dann ist die Abbildung*

   $$\begin{array}{ccc} S_n & \longrightarrow & Aut(M) \\ \sigma & \longmapsto & f \circ \sigma \circ f^{-1} \end{array}$$

   bijektiv.

3. *Tabellenschreibweise von Permutationen:*

   $$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

4. *Zyklenschreibweise von Permutationen:*

   $$(1 \quad \sigma(1) \quad \sigma \circ \sigma(1) \quad \dots)$$

   Z.B. gilt für $n = 5$:

   $$(123)(45) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix}$$

Satz 5.1.3 *Sei $n > 0$. S_n hat $n!$ Elemente. Kurz: $\#(S_n) = n!$.*

Beweis: Induktion nach n.

1. $n = 1$: $\mathrm{id}_{\{1\}} : \{1\} \longrightarrow \{1\}$ ist die einzige Abbildung; sie ist bijektiv.

2. Sei $n > 0$, und sei vorausgesetzt: $\#(S_n) = n!$. Für eine bijektive Abbildung

   $$\sigma : \{1, \dots, n+1\} \longrightarrow \{1, \dots, n+1\}$$

   gibt es für das Bild von $n + 1$ die $n + 1$ Möglichkeiten $1, 2, \dots, n + 1$. Gilt $\sigma(n + 1) = j \in \{1, \dots, n + 1\}$, so ist

   $$\sigma|\{1, \dots, n\} \longrightarrow \{1, \dots, n + 1\} \setminus \{j\}$$

   eine bijektive Abbildung zwischen zwei $n-$elementigen Mengen; dafür gibt es nach Induktionsvoraussetzung $n!$ Möglichkeiten.

   Daraus folgt: $\#(S_{n+1}) = (n + 1)n! = (n + 1)!$.

$\square$

Bemerkung 5.1.4 *S_n ist (mit der Komposition als Verknüpfung) eine Gruppe; sie heißt die* **symmetrische Gruppe**.

5.2 Das Signum einer Permutation

Definition 5.2.1 *Sei $n \in I\!N$.*
Eine Permutation $\tau \in S_n$ heißt **Transposition***, wenn gilt:*
$\exists i \in \{1, \ldots, n\}$ [

1. $\tau(i) \neq i$
2. $\forall j \in \{1, \ldots, n\} \setminus \{i, \tau(i)\}$ $[\tau(j) = j]]$

Bemerkung 5.2.2
1. *In Zyklenschreibweise schreibt sich eine Transposition als (ij).*
2. $i \neq j \in \{1, \ldots, n\} \Longrightarrow \exists \sigma \in S_n$ $[(ij) = \sigma(12)\sigma^{-1}$

   Beweis: Im Fall $n = 2$ gilt $(ij) = (12)$, und man kann $\sigma = (1)$ wählen.

   Ist $n > 2$ und $\{i, j\} \neq \{1, 2\}$, so sei $\sigma \in S_n$ eine Permutation, die 1 auf i und 2 auf j abbildet und die übrigen Elemente von $\{1, \ldots, n\}$ festhält. Dann rechnet man die Behauptung direkt nach.

   Genauer: Man hat folgende Fälle zu betrachten:

   Ist $\{i, j\} \cap \{1, 2\} = \emptyset$, so hat man $\sigma := (1i)(2j)$ zu wählen.

   Ist $i \in \{1, 2\}$, etwa $i = 1$, und $j \notin \{1, 2\}$, so erhält man $\sigma := (2j)$.

   Die übrigen Fälle werden entsprechend behandelt.

Satz 5.2.3 *Sei $n \in I\!N, n \geq 2$. Jede Permutation $\sigma \in S_n$ ist Produkt (d.h. Komposition) von endlich vielen Transpositionen.*
Beweis: Sei $\sigma \in S_n$.

1. $\sigma = (1) \Longrightarrow \sigma = (12)(12)$.
2. Sei $\sigma \neq (1)$, und sei $i := \min\{j \in \{1, \ldots, n\} \ : \ \sigma(j) \neq j\}$; dann gilt $\sigma(i) > i$.

   Für $\sigma_1 := (i\sigma(i))\sigma$ gilt dann: $\min\{j \in \{1, \ldots, n\} \ : \ \sigma_1(j) \neq j\} > i$;

   Definiert man in entsprechender Weise $\sigma_2 := (j\sigma(j))(i\sigma(i))\sigma$ und fährt so fort, muß das Verfahren abbrechen, also die identische Abbildung ergeben. Daraus folgt die Behauptung.

Beispiel: $(1352) = (13)(15)(12)$

Bemerkung 5.2.4 *Reihenfolge und Anzahl der Transpositionen sind nicht eindeutig bestimmt.*
Definition 5.2.5 *Sei $n \in I\!N, n \geq 2$ und $\sigma \in S_n$.*

$$\operatorname{sign}(\sigma) := \varepsilon(\sigma) := \prod_{0<i<j<n+1} \frac{\sigma(i) - \sigma(j)}{i - j} := \prod_{j=2}^{n} \prod_{i=1}^{i-1} \frac{\sigma(i) - \sigma(j)}{i - j}$$

heißt **Signum, Signatur** *oder* **Vorzeichen** *von σ.*
Bemerkung 5.2.6 1. *Nach Definition ist $\varepsilon(\sigma) \in Q\!\!\!\!Q$.*

2. *Es wird gezeigt, daß das Signum nur die Werte* $+1$ *und* -1 *annehmen kann (also insbesondere ganzzahlig ist).*

3. $\varepsilon((1)) = 1$ *(Klar)*

4. $\varepsilon((12)) = -1$

   Beweis: Sei $\sigma := (12)$.

   $\prod_{0<i<j<n+1} \frac{\sigma(i)-\sigma(j)}{i-j} =$

   $\frac{2-1}{1-2}$ *(Faktor mit* $i = 1, j = 2$*)*

   $\prod_{j>2} \frac{2-j}{1-j}$ *(Faktoren mit* $i = 1, j > 2$*)*

   $\prod_{j>2} \frac{1-j}{2-j}$ *(Faktoren mit* $i = 2, j > 2$*)*

   $\prod_{j>i>2} \frac{i-j}{i-j}$

   Das letzte Produkt hat den Wert 1*, der erste Faktor den Wert* -1*, das zweite und das dritte Produkt haben zusammen den Wert* 1.

Satz 5.2.7 *Sei* $n \in \mathbb{N}, n \geq 2$ *und* $\sigma, \tau \in S_n$*. Dann gilt:* $\varepsilon(\tau\sigma) = \varepsilon(\tau)\varepsilon(\sigma)$.

Beweis: Es ist

$\varepsilon(\tau\sigma) =$

$\prod_{0<i<j<n+1} \frac{\tau\sigma(i)-\tau\sigma(j)}{i-j} =$

$\prod_{0<i<j<n+1} \frac{\tau\sigma(i)-\tau\sigma(j)}{\sigma(i)-\sigma(j)} \prod_{0<i<j<n+1} \frac{\sigma(i)-\sigma(j)}{i-j}$

Der zweite Faktor in diesem Produkt ist $\varepsilon(\sigma)$; der erste Faktor ist gleich

$\prod_{0<i<j<n+1,\sigma(i)<\sigma(j)} \frac{\tau\sigma(i)-\tau\sigma(j)}{\sigma(i)-\sigma(j)} \prod_{0<i<j<n+1,\sigma(i)>\sigma(j)} \frac{\tau\sigma(i)-\tau\sigma(j)}{\sigma(i)-\sigma(j)} =$

und damit (nach Umbenennung von i und j im zweiten Faktor:

$\prod_{0<i<j<n+1,\sigma(i)<\sigma(j)} \frac{\tau\sigma(i)-\tau\sigma(j)}{\sigma(i)-\sigma(j)} \prod_{0<j<i<n+1,\sigma(j)>\sigma(i)} \frac{\tau\sigma(i)-\tau\sigma(j)}{\sigma(i)-\sigma(j)} =$

$\prod_{0<i,j<n+1,\sigma(i)<\sigma(j)} \frac{\tau\sigma(i)-\tau\sigma(j)}{\sigma(i)-\sigma(j)} =$

$\prod_{0<i<j<n+1} \frac{\tau(i)-\tau(j)}{i-j} =$

$\varepsilon(\tau)$

Korollar 5.2.8 *Sei* $n \in \mathbb{N}, n \geq 2$ *und* $\sigma \in S_n$*. Dann gilt:*

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

   Beweis: $\varepsilon(\sigma^{-1})\varepsilon(\sigma) = \varepsilon(\sigma^{-1}\sigma) = \varepsilon((1)) = 1$

2. σ *Transposition* $\Longrightarrow \varepsilon(\sigma) = -1$

   Beweis. Sei $\sigma = (ij)$ *Transposition. Nach Bemerkung oben gibt es eine Permutation* $\tau \in S_n$ *mit* $(ij) = \tau(12)\tau^{-1}$*. Daraus folgt:*

   $\varepsilon((ij)) = \varepsilon(\tau(12)\tau^{-1}) = \ldots = -1$

3. $\varepsilon(\sigma) \in \{-1, 1\}$.

   Beweis: Sind $\tau_1, \ldots, \tau_k$ *Transpositionen mit* $\sigma = \tau_1 \ldots \tau_k$*, so folgt:* $\varepsilon(\sigma) = \varepsilon(\tau_1) \ldots \varepsilon(\tau_k)$

5.3 Determinanten

In diesem Abschnitt wird folgende Bezeichnung verwendet: Für eine $(n \times n)$-Matrix $A \in \mathbb{R}^{n\times n}$ seien $a_1, \ldots, a_n \in \mathbb{R}_n = \mathbb{R}^{1\times n}$ die Zeilenvektoren von A. Ist also $A = (a_{\mu\nu})$, so sei für $\mu = 1, \ldots, n$ $a_\mu := (a_{\mu 1}, \ldots, a_{\mu n})$ und somit $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$.

Definition 5.3.1 (Determinante) *Eine Abbildung* $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ *heißt* **Determinante**, *wenn gilt:*

1. $\det$ *hängt linear von jedem der Zeilenvektoren ab, d.h., ist* $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^{n\times n}$, $\mu \in \{1, \ldots, n\}$ *und ist* $a_\mu = \beta b_\mu + \gamma c_\mu$ *mit Zeilenvektoren* $b_\mu, c_\mu \in \mathbb{R}_n$ *und* $\beta, \gamma \in \mathbb{R}$, *so ist*

$$\det(A) = \beta \det(\begin{pmatrix} a_1 \\ \vdots \\ b_\mu \\ \vdots \\ a_n \end{pmatrix}) + \gamma \det(\begin{pmatrix} a_1 \\ \vdots \\ c_\mu \\ \vdots \\ a_n \end{pmatrix})$$

   Man sagt auch, $\det$ *ist eine* n**-lineare** *Abbildung in den Zeilenvektoren von* A.

2. $\det$ *ist eine* **alternierende** *Abbildung in den Zeilenvektoren von* A, *d.h.: Besitzt* A *zwei gleiche Zeilen* $a_{\mu_1}, a_{\mu_2}, \mu_1 \neq \mu_2$, *so ist* $\det(A) = 0$.

3. **Normierung:** $\det(E_n) = 1$

Schreibweise: Statt $\det(A)$ wird auch $\det A$ oder $\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{vmatrix}$ geschrieben.

Bemerkung 5.3.2 (Eigenschaften von Determinanten) *Sei* $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^{n\times n}$

1. *Ist eine Zeile* a_i *der Nullvektor, gilt* $\det A = 0$

   Folgerung aus der Linearität der Abbildung $\det$ *an der* i*-ten Stelle.*

2. *Entsteht* $\bar{A}$ *aus* A *durch eine elementare Zeilentransformation vom Typ I, d.h., Multiplikation einer Zeile mit einem Faktor* $\lambda \in \mathbb{R}, \lambda \neq 0$, *so gilt* $\det \bar{A} = \lambda \det A$.

   Beweis: Folgt aus der n*-Linearität von* $\det$.

3. $\forall \lambda \in I\!R\ [\det(\lambda A) = \lambda^n \det A$.

   Beweis: Klar, falls $\lambda = 0$. Im Fall $\lambda \neq 0$ n-malige Anwendung der vorherigen Bemerkung.

4. *Entsteht $\bar{A}$ aus A durch eine elementare Zeilentransformation vom Typ IV, d.h. Vertauschung von zwei Zeilen, so gilt* $\det\bar{A} = -\det A$.

   Beweis: Seien $\mu_1 < \mu_2 \in \{1, \ldots, n\}$ und $\bar{A} := \begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_n \end{pmatrix}$. *Dann gilt:*

   $\det A + \det(\bar{A}) =$

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_n \end{pmatrix} =$$

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_{\mu_1} + a_{\mu_2} \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_{\mu_1} + a_{\mu_2} \\ \vdots \\ a_n \end{pmatrix} =$$

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_1} + a_{\mu_2} \\ \vdots \\ a_{\mu_1} + a_{\mu_2} \\ \vdots \\ a_n \end{pmatrix} = 0$$

5. *Die in der Definition angegebene Eigenschaft „alternierend“ ist äquivalent zu der Eigenschaft in der letzten Bemerkung.*

   Beweis: Es sei vorausgesetzt:

$$\forall A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in I\!R^{n\times n}\ \forall \mu_1 < \mu_2 \in \{1,\ldots,n\}\ [\det(A) = -\det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_n \end{pmatrix}))].$$

Es gelte nun für $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$*:* $a_{\mu_1} = a_{\mu_2}$ *für zwei Zeilenindices* $\mu_1 < \mu_2$*.*
Dann folgt

$$0 = \det A + \det\begin{pmatrix} a_1 \\ \vdots \\ a_{\mu_2} \\ \vdots \\ a_{\mu_1} \\ \vdots \\ a_n \end{pmatrix} =$$

$\det A + \det A =$

$(1+1)\det A$

und daraus $\det A = 0$*.*

6. *Entsteht* $\bar{A}$ *aus* A *durch eine elementare Zeilentransformation vom Typ II oder III, d.h., Addition des* λ*-fachen der* j*-ten Zeile zur* i*-ten Zeile (* $\lambda \in I\!R, \lambda \neq 0$*), so gilt* $\det\bar{A} = \det A$*.*

   Beweis: $\det\bar{A} =$

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} =$$

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \lambda\det\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} =$$

$\det A$

7. *Ist* $A = (a_{\mu\nu})$ *eine obere Dreiecksmatrix, so folgt:* $\det A = a_{11} \dots a_{nn}$*; entsprechend für untere Dreiecksmatrizen.*

   Beweis:

   (a) *Sind alle Hauptdiagonalelemente* $\neq 0$*, folgt:*
   A kann durch elementare Zeilenumformungen vom Typ III in eine Diagonalmatrix $\bar{A}$ *übergeführt werden, wobei die Elemente in der Hauptdiagonalen ungeändert bleiben; nach der vorherigen Bemerkung gilt*
   $\det A = \det \bar{A} = a_{11} \dots a_{nn} \det E_n = a_{11} \dots a_{nn}$.

   (b) *Sind nicht alle Hauptdiagonalelemente* $\neq 0$*, so sei i der größte Index mit* $a_{ii} = 0$*; es gelte also* $a_{i+1,i+1}, \dots, a_{nn} \neq 0$.
   A kann dann durch elementare Zeilenumformungen vom Typ III in eine Matrix $\bar{A}$ *mit i-tem Zeilenvektor* $a_i = 0$ *übergeführt werden; es folgt:*
   $\det A = \det \bar{A} = 0$

8. $\det A = 0 \Longleftrightarrow \mathrm{Rg} A < n$

   Beweis: A kann durch elementare Zeilenumformungen vom Typ II – IV in Zeilenstufenform $\bar{A}$ *gebracht werden; dann gilt* $\det \bar{A} = \pm \det A$*. Die Behauptung folgt nun aus der vorherigen Bemerkung.*

9. *Ist A von der Gestalt* $\begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix}$ *mit quadratischen Matrizen* A_1, A_2*, so gilt*

   $\det A = \det A_1 \det A_2$.

   Analog für Matrizen der Form $\begin{pmatrix} A_1 & 0 \\ B & A_2 \end{pmatrix}$

   Beweis: A_2 *kann durch elementare Zeilenumformungen vom Typ II – IV von A, die die Matrizen* A_1 *und C ungeändert lassen, in Zeilenstufenform* $\bar{A}_2$ *überführt werden; es sei* k_2 *die Anzahl der dabei auftretenden Zeilenvertauschungen.*

   Ebenso kann A_1 *durch elementare Zeilenumformungen vom Typ II – IV von A, die* A_2 *ungeändert lassen, in Zeilenstufenform* $\bar{A}_1$ *überführt werden; aus C wird dabei* $\bar{C}$*. Es sei* k_1 *die Anzahl der jetzt auftretenden Zeilenvertauschungen.*

   Für die Matrix $\bar{A} = \begin{pmatrix} \bar{A}_1 & \bar{C} \\ 0 & \bar{A}_2 \end{pmatrix}$ *folgt die Behauptung aus der vorvorigen Bemerkung; andererseits gilt:*

   $\det \bar{A}_i = (-1)^{k_i} \det A_i, i = 1, 2,$

   $\det \bar{A} = (-1)^{k_1 + k_2} \det A,$

   woraus die Behauptung für A folgt.

10. **Determinantenmultiplikationssatz:** *Ist auch* $B \in \mathbb{R}^{n \times n}$*, so folgt:*

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

Beweis: Ist $\mathrm{Rg}A < n$, *also* $\dim \mathrm{Im}\hat{A} < n$, *so folgt nach der Dimensionsformel für lineare Abbildungen:* $\mathrm{Rg}(AB) = \dim \mathrm{Im}\hat{A}\hat{B} \leq \dim \mathrm{Im}\hat{A} < n$ *und damit die Behauptung.*

Ist $\mathrm{Rg}A = n$, *also* A *invertierbar, so ist* $A = S_1 \ldots S_k$ *mit Elementarmatrizen* $S_1, \ldots, S_k \in \mathbb{R}^{n\times n}$.

Damit (d.h.: vollständige Induktion nach k) ist zu zeigen: Ist A *eine Elementarmatrix und* $B \in \mathbb{R}^{n\times n}$, *so folgt:* $\det(AB) = \det A \, \det B$.

Zum Beweis dieser Behauptung betrachtet man die einzelnen Fälle, und zwar müssen nur die Fälle Typ I und Typ II betrachtet werden, da die Typen III und IV auf diese Fälle zurückgeführt werden können.

(a) *Typ I:* $A = S_i(\lambda) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \lambda & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$. *Dann*

$\det A = \lambda$

$\det(AB) = \det\bar{B}$,

wobei $\bar{B}$ *die Matrix ist, die aus* B *hervorgeht, wenn man die* i*-te Zeile von* B *mit* λ *multipliziert. Es gilt also*

$\det\bar{B} = \lambda\det B$.

Daraus folgt die Behauptung.

(b) *Typ II:* $A = Q_i^j := (\delta_{\mu\nu} + \delta_{\mu i}\delta_{\nu j})_{\mu,\nu=1,\ldots,n}$

Da dies eine Dreiecksmatrix ist, in der jedes Hauptdiagonalelement 1 *ist, gilt* $\det A = 1$.

Andererseits ist $\det(AB) = \det\bar{B}$,

wobei $\bar{B}$ *die Matrix ist, die aus* B *hervorgeht, wenn man die* j*-te Zeile von* B *zur* i*-ten Zeile addiert; es gilt also*

$\det\bar{B} = \det B$.

11. A *invertierbar* $\Longrightarrow \det(A^{-1}) = (\det A)^{-1}$

Beweis: $1 = \det E_n = \det(AA^{-1}) = \det(A)\det(A^{-1})$.

12. $\sigma \in S_n \Longrightarrow \det \begin{pmatrix} {}^t e_{\sigma(1)} \\ \vdots \\ {}^t e_{\sigma(n)} \end{pmatrix} = \varepsilon(\sigma)$.

Beweis: Klar, denn ist σ *Produkt von* k *Transpositionen, kann die Matrix durch die entsprechenden* k *Zeilenvertauschungen in die Einheitsmatrix überführt werden; bei jeder Zeilenvertauschung ändert sich das Vorzeichen.*

Man kann auch so argumentieren: $\sigma = \tau_1 \ldots \tau_k$ *die Darstellung von* σ *als Produkt von Transpositionen, so ist* $\begin{pmatrix} {}^t e_{\sigma(1)} \\ \vdots \\ {}^t e_{\sigma(n)} \end{pmatrix}$ *Produkt der den Transpositionen* *zugeordneten Elementarmatrizen* $P^i_{\tau(i)}, i = 1, \ldots, k$. *Die Behauptung folgt dann aus dem Determinantenmultiplikationssatz.*

13. **Leibniz-Formel:** $A = (a_{\mu\nu}) \in I\!R^{n\times n} \Longrightarrow$

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$$

Beweis: Nutzt man die Linearität von det *nacheinander für die 1. Zeile* a_1, *die 2. Zeile* a_2 *usw. bis zur* n*-ten Zeile* a_n *aus, folgt man aus* $a_i = \sum_{\nu_i=1}^n \ a_{i\nu_i} \ {}^t e_{\nu_i}$ $(i = 1, \ldots, n)$:

$\det A =$

$$\sum_{\nu_1=1}^n a_{1\nu_1} \det \begin{pmatrix} {}^t e_{\nu_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix} =$$

$$\sum_{\nu_1=1}^n a_{1\nu_1} \sum_{\nu_2=1}^n a_{2\nu_2} \det \begin{pmatrix} {}^t e_{\nu_1} \\ {}^t e_{\nu_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix} =$$

$$\sum_{\nu_1=1}^n \sum_{\nu_2=1}^n a_{1\nu_1} a_{2\nu_2} \det \begin{pmatrix} {}^t e_{\nu_1} \\ {}^t e_{\nu_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \ldots =$$

$$\sum_{\nu_1=1}^n \sum_{\nu_2=1}^n \ldots \sum_{\nu_n=1}^n a_{1\nu_1} a_{2\nu_2} \ldots a_{n\nu_n} \det \begin{pmatrix} {}^t e_{\nu_1} \\ \vdots \\ {}^t e_{\nu_n} \end{pmatrix}$$

In dieser Summe von n^n *Summanden sind nur die Summanden von Null verschieden, bei denen die Zahlen* $\nu_1, \ldots, \nu_n$ *eine Permutation von* $1, \ldots, n$ *sind; bei den übrigen Summanden kommen zwei gleiche Zeilenvektoren vor, so daß die in diesem Summanden enthaltene Determinante verschwindet. Damit ist:*

$\det A =$

$$\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \det \begin{pmatrix} {}^t e_{\sigma(1)} \\ \vdots \\ {}^t e_{\sigma(n)} \end{pmatrix} =$$

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$$

14. Die Determinante ist eindeutig bestimmt.

15. *Die Determinante ist ein Polynom in den n^2 Koeffizienten der Matrix.*

    Der Totalgrad dieses Polynoms ist n; der Grad bzgl. jedes einzelnen Koeffizienten (wenn man die übrigen Koeffizienten festhält) ist 1.

    In jedem Summanden kommt aus jeder Zeile genau ein Koeffizient vor; ebenso kommt in jedem Summanden aus jeder Spalte genau eine Koeffizient vor.

16. $\det A$ *hängt differenzierbar, insbesondere stetig von den Koeffizienten ab.*

    (Polynome sind differenzierbare Funktionen).

17. $\det {}^tA = \det A$

    Folgt unmittelbar aus der Leibniz-Formel.

18. *Die Determinante hängt n-linear und alternierend von den Spaltenvektoren der Matrix ab.*

19. *Bei der Herleitung der Leibniz-Formel muß nicht dividiert werden. Man kann daher auf diese Weise Determinanten auch über Ringen, also z.B. für Matrizen mit Koeffizienten in $\mathbb{Z}$, definieren.*

20. *Spezialfälle:*

    (a) $n = 2:$

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

    (b) $n = 3:$

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

    $$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} -$$

    $$a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

    Merkregel: Regel von Sarrus ...

21. *Die Summe in der Leibniz-Formel hat $n!$ Summanden. Für die praktische Berechnung von Determinanten eignet sich die Leibniz-Formel daher nur für kleinere Werte von n, da $n!$ schnell sehr groß wird. So ist z.B.*

    10! = 3628800

    100! =

    93326215443944152681699238856266700490715968264381621468592963895217599993229915

    608941463976156518286253697920827223758251185210916864000000000000000000000000

    $\simeq .9332621544 \cdot 10^{158}$

    1000! =

    402387260077093773543702433923003985719374864210714632543799910429938512398629020

    059204420848696940480047998861019719605863166687299480855890132382966994459099742

    450408707375991882362772718873251977950595099527612087497546249704360141827809464

    649629105639388743788648733711918104582578364784997701247663288983595573543251

31853239584630755574091142624174743493475534286465766116677973966688202912073791
43853719588249808126867838374559731746136085379534524221586593201928090878297308
43139284440328123155861103697680135730421616874760967587134831202547858932076716
91324484262361314125087802080002616831510273418279777047846358681701643650241536
91398281264810213092761244896359928705114964975419909342221566832572080821333186
11681155361583654698404670897560290095053761647584772842188967964624494516076535
34081989013854424879849599533191017233555566021394503997362807501378376153071277
61926849034352625200015888535147331611702103968175921510907788019393178114194545
25722386554146106289218796022383897147608850627686296714667469756291123408243920
81601537808898939645182632436716167621791689097799119037540312746222899880051954
44414282012187361745992642956581746628302955570299024324153181617210465832036786
90611726015878352075151628422554026517048330422614397428693306169089796848259012
54583271682264580665267699586526822728070757813918581788896522081643483448259932
66043367660176999612831860788386150279465955131156552036093988180612138558600301
43569452722420634463179746059468257310379008402443243846565724501440282188525247
09351906209290231364932734975655139587205596542287497740114133469627154228458623
77387538230483865688976461927383814900140767310446640259899490222221765904339901
88601856652648506179970235619389701786004081188972991831102117122984590164192106
88843871218556461249607987229085192968193723886426148396573822911231250241866493
53143970137428531926649875337218940694281434118520158014123344828015051399694290
15348307764456909907315243327828826986460278986432113908350621709500259738986355
42771967428222487575867657523442202075736305694988250879689281627538488633969099
59826280956121450994871701244516461260379029309120889086942028510640182154399457
15680594187274899809425474217358240106367740459574178516082923013535808184009699
63725242305608559037006242712434169090041536901059339838357779394109700277534720
00000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000

$\simeq .4023872601 \cdot 10^{2568}$

Satz 5.3.3 *Sei* $n \in \mathbb{N}, n \geq 1$. *Die Determinante* $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ *existiert.*
Beweis. Im Fall $n = 1$ sei $\det(a_{1,1}) := a_{1,1}$; dann sind offensichtlich die definierenden Eigenschaften einer Determinante erfüllt.
Sei $n > 1$. Man definiert $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ durch $\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$; es ist zu zeigen:

1. det hängt **linear von jedem der Zeilenvektoren** ab.

   Beweis: Direktes Nachrechnen!

2. det ist eine **alternierende Abbildung** in den Zeilenvektoren von A.

   Beweis: A besitze zwei gleiche Zeilen $a_{\mu_1}, a_{\mu_2}, \mu_1 \neq \mu_2$. Sei $\tau := (\mu_1 \mu_2)$.

   Es sei $A_n := \{\sigma \in S_n \mid \varepsilon(\sigma) = 1\}$ ("gerade Permutationen"). Ist σ eine ungerade Permutation (d.h. $\varepsilon(\sigma) = -1$, so ist $\sigma \circ \tau$ gerade und $\sigma \circ \tau \circ \tau = \sigma$. Daraus folgt:

   $S_n = A_n \cup A_n \tau$ und daher

   $\det A =$

   $\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} =$

   $\sum_{\sigma \in A_n} \ a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} - \sum_{\sigma \in A_n} \ a_{1\sigma\circ\tau(1)} \cdot \ldots \cdot a_{n\sigma\circ\tau(n)}.$

Nun gilt für jedes $\sigma \in A_n$:

$a_{1\sigma\circ\tau(1)} \cdot \ldots a_{\mu_1\sigma\circ\tau(\mu_1)} \cdot \ldots a_{\mu_2\sigma\circ\tau(\mu_2)} \cdot \ldots \cdot a_{n\sigma\circ\tau(n)} =$

$a_{1\sigma(1)} \cdot \ldots a_{\mu_1\sigma(\mu_2)} \cdot \ldots a_{\mu_2\sigma(\mu_1)} \cdot \ldots \cdot a_{n\sigma(n)} =$

$a_{1\sigma(1)} \cdot \ldots a_{\mu_2\sigma(\mu_2)} \cdot \ldots a_{\mu_1\sigma(\mu_1)} \cdot \ldots \cdot a_{n\sigma(n)} =$

$a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$

Daher heben sich die Summanden der letzten Summe weg, so daß die Behauptung folgt.

3. **Normierung:** $\det(E_n) = 1$

   Beweis: $\det(E_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma)\delta_{1\sigma(1)} \cdot \ldots \cdot \delta_{n\sigma(n)}$.

   In dieser Summe ist nur der Summand mit $\sigma = (1)$ von Null verschieden.

Lemma 5.3.4 *Es seien*

1. *V ein n-dimensionaler Vektorraum,*
2. *$\mathcal{V} = \{v_1, \ldots, v_n\}$, $\tilde{\mathcal{V}} = \{\tilde{v}_1, \ldots, \tilde{v}_n\}$ Basen von V,*
3. *$f : V \longrightarrow V$ Endomorphismus.*

Dann gilt

$$\det(M_{\mathcal{V}}^{\mathcal{V}}(f) = \det(M_{\tilde{\mathcal{V}}}^{\tilde{\mathcal{V}}}(f)$$

Beweis: Die Matrix $A := M_{\mathcal{V}}^{\mathcal{V}}(f)$ ist durch das kommutative Diagramm

$$\begin{array}{cccccc} & V & \xrightarrow{f} & V & \\ \iota_{\mathcal{V}} & \uparrow & & \uparrow & \iota_{\mathcal{V}} \\ & \mathbb{R}^n & \xrightarrow{\hat{A}} & \mathbb{R}^n & \end{array}$$

definert, die Matrix $\tilde{A} := M_{\tilde{\mathcal{V}}}^{\tilde{\mathcal{V}}}(f)$ durch das kommutative Diagramm

$$\begin{array}{cccccc} & V & \xrightarrow{f} & V & \\ \iota_{\tilde{\mathcal{V}}} & \uparrow & & \uparrow & \iota_{\tilde{\mathcal{V}}} \\ & \mathbb{R}^n & \xrightarrow{\hat{\tilde{A}}} & \mathbb{R}^n & \end{array}$$

Sei $B := M_{\tilde{\mathcal{V}}}^{\mathcal{V}}(\mathrm{id}_V) = \mathcal{T}_{\tilde{\mathcal{V}}}^{\mathcal{V}}$ die Transformationsmatrix zur Koordinatentransformation $\mathcal{V} \mapsto \tilde{\mathcal{V}}$, definiert durch das kommutative Diagramm

$$\begin{array}{cccccc} & V & \xrightarrow{\mathrm{id}_V} & V & \\ \iota_{\mathcal{V}} & \uparrow & & \uparrow & \iota_{\tilde{\mathcal{V}}} \\ & \mathbb{R}^n & \xrightarrow{\hat{B}} & \mathbb{R}^n & \end{array}$$

In dem kommutativen Diagramm

$$\begin{array}{ccccccccc} & V & = & & V & \xrightarrow{f} & V & & = & V & \\ \iota_{\tilde{\mathcal{V}}} & \uparrow & & \iota_{\mathcal{V}} & \uparrow & & \uparrow & \iota_{\mathcal{V}} & & \uparrow & \iota_{\tilde{\mathcal{V}}} \\ & \mathbb{R}^n & \xrightarrow{\hat{B^{-1}}} & & \mathbb{R}^n & \xrightarrow{\hat{A}} & \mathbb{R}^n & & \xrightarrow{\hat{B}} & \mathbb{R}^n & \end{array}$$

ist dann die unterste Zeile von ganz links nach ganz rechts die Abbildung $\hat{\tilde{A}}$, d.h., es gilt:
$\tilde{A} = BAB^{-1}$.
Aus dem Determinantenmultiplikationssatz folgt die Behauptung.
Definition 5.3.5 *Es sei V ein n-dimensionaler Vektorraum, $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine Basis von V und $f : V \longrightarrow V$ ein Endomorphismus.*

$$\det(f) := \det(M^{\mathcal{V}}_{\mathcal{V}}(f)$$

Bemerkung 5.3.6

1. *Nach dem Lemma ist* $\det(f)$ *unabhängig von der Auswahl der Basis.*
2. $\det(f)$ *hat eine einfache geometrische Bedeutung, und zwar ist* $|\det(f)|$ *der Verzerrungsfaktor für Volumina bei Anwendung der Abbildung f, d.h.: Hat eine Teilmenge, z.B. ein Quader $Q \subset V$ das Volumen 1, so hat die Menge $f(Q)$ das Volumen* $\det(f)$.
3. *Das Vorzeichen von* $\det(f)$ *bestimmt, ob die Abbildung f die Orientierung des Raumes erhält (z.B. Drehungen) oder umdreht (z.B. Spiegelungen)*

Beispiel 5.3.7 *Sei*

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

Bei einer 5×5-Matrix ist die Berechnung nach der Leibniz-Formel noch gut möglich; sie erfordert die Berechnung von 5! = 120 Summanden, die jeweils aus 5 Faktoren bestehen, es müssen also 5!(5-1) = 480 Multiplikationen durchgeführt werden.*
In diesem Fall hat man in jeder Zeile eine 0, so daß man mit 4!(5-1)= 96 Multiplikationen auskommt.*
Aber bereits hier ist es günstiger, die Matrix nach dem Gauß-Verfahren in Zeilenstufenform zu bringen (bei nichtverschwindender Matrix also in obere Dreiecksform). Für eine Dreiecksmatrix ist die Determinate das Produkt der Elemente in der Hauptdiagonalen. Für größere n ist das Verfahren nach Gauß grundsätzlich vorzuziehen. Nur bei dünn besetzten Matrizen (sparse matrices) ist die Leibniz-Formel (bzw. die daraus hergeleitete Laplace-Entwicklung) oft mehr angeraten.

Die Rechnung kann z.B. wie folgt durchgeführt werden:

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

$$-4\det \begin{pmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & 2 & -2 & 0 & 0 \\ 0 & 2 & 1 & -3 & 0 \\ 0 & 2 & 1 & 1 & -1 \end{pmatrix} = -8\det \begin{pmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 2 & 1 & -3 & 0 \\ 0 & 2 & 1 & 1 & -1 \end{pmatrix} =$$

$$-8\det \begin{pmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & -3 & -3 & -1 \\ 0 & 0 & -3 & -9 & -2 \\ 0 & 0 & -3 & -5 & -3 \end{pmatrix} = 8\det \begin{pmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & -6 & -1 \\ 0 & 0 & 0 & -2 & -2 \end{pmatrix} =$$

$$-16\det \begin{pmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 6 & 1 \end{pmatrix} = -16\det \begin{pmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix} = 240$$

6 Das kanonische euklidische Skalarprodukt, Längen und Winkel

6.1 Das kanonische euklidische Skalarprodukt

Definition 6.1.1 *Sei $n \geq 1$ eine natürliche Zahl. Die Abbildung*

$$\begin{array}{rlcl} \sigma: & \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ & (x,y) & \mapsto & \langle x,y\rangle := {}^t xy = x_1y_1 + x_2y_2 + \ldots + x_ny_n \end{array}$$

heißt das kanonische euklidische Skalarprodukt auf dem $\mathbb{R}^n$.

Aussage 6.1.2 *Eigenschaften*

1. *Die Abbildung $\sigma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist eine* **Bilinearform**, *d.h., es gilt:*
   (a) *$\forall x \in \mathbb{R}^n$ ist die Abbildung $\sigma(x, \cdot) : \mathbb{R}^n \to \mathbb{R}, y \mapsto \sigma(x,y)$ linear.*
   (b) *$\forall y \in \mathbb{R}$ ist die Abbildung $\sigma(\cdot, y) : \mathbb{R}^n \to V\mathbb{R}, x \mapsto \sigma(x,y)$ linear.*
2. *Die Bilinearform $\sigma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist*
   (a) **symmetrisch** $\Longleftrightarrow \forall x,y \in \mathbb{R}^n\ [\ \sigma(x,y) = \sigma(y,x)\ ]$
   (b) **positiv definit** $\Longleftrightarrow \forall x \in \mathbb{R}^n\ [\ x \neq 0 \Longrightarrow \sigma(x,x) > 0\ ]$

Bemerkungen und Beispiele 6.1.3 *1. Die in 6.1.2 aufgeführten Eigenschaften sind die definierenden Eigenschaften eines (allgemeinen) Skalarprodukts für reelle Vektorräume.*

2. $x = 0 \Longrightarrow \sigma(x,x) = 0$
3. *Kanonisches Skalarprodukt auf dem $\mathbb{R}^2$:*
   $\langle x,y\rangle = \langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \rangle := {}^t xy = x_1y_1 + x_2y_2.$
   Setzt man die Geometrie des euklidischen Anschaungsraums inkl. Längenmessung (Länge $||x|| = \sqrt{x^2+y^2}$ eines Vektors $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$) und Trigonometrie voraus, ergibt sich: $\langle x,y\rangle = ||x|| \cdot ||y|| \cdot \cos\angle(x,y)$. In dem im Rahmen dieser Vorlesung gegebenen Aufbau werden aber Längen- und Winkelmessung mit Hilfe des Skalarprodukts eingeführt.
4. *Die analoge Aussage gilt für das kanonische Skalarprodukt im $\mathbb{R}^3$: $\langle x,y\rangle := {}^t xy = x_1y_1 + x_2y_2 + x_3y_3$.*

Satz 6.1.4 (Ungleichung von Cauchy-Schwarz) *Seien $x, y \in \mathbb{R}^n$*

1. $\langle x,y\rangle^2 \leq \langle x,y\rangle \cdot \langle y,y\rangle$.
2. $\langle x,y\rangle^2 = \langle x,x\rangle \cdot \langle y,y\rangle \Longleftrightarrow x,y$ *l.a.*

Beweis:

1. Ist $y = 0$, folgt aus der Linearität des Skalarprodukts die Gleichheit der beiden Seiten.

   Sei also $y \neq 0$.

   $\forall \alpha \in \mathbb{R}$ [$0 \leq \langle x - \alpha y, x - \alpha y \rangle =$

   $\langle x, x \rangle - 2\alpha \langle x, y \rangle + \alpha^2 \langle y, y \rangle$

   Setzt man in dieser Ungleichung $\alpha := \frac{\langle x,y \rangle}{\langle y,y \rangle}$ und multipliziert man mit $\langle y, y \rangle$, so bleibt die Ungleichung wegen $\langle y, y \rangle > 0$ erhalten, und man bekommt:

   $0 \leq \langle x, x \rangle \langle y, y \rangle - 2\langle x, y \rangle \langle x, y \rangle + \langle x, y \rangle \langle x, y \rangle =$

   $\langle x, x \rangle \langle y, y \rangle - (\langle x, y \rangle)^2,$

   woraus die Behauptung folgt.

2. „ $\Longrightarrow$ “

   Seien $x, y \in \mathbb{R}^n$ mit $\langle x, y \rangle^2 = \langle x, x \rangle \cdot \langle y, y \rangle$

   Zu zeigen: x, y l.a.

   O.E. $x \neq 0, y \neq 0$. Dann $0 < \langle x, x \rangle \cdot \langle y, y \rangle = \langle x, y \rangle^2$, also $\langle x, y \rangle \neq 0$.

   Man verwendet einen ähnlichen Ansatz wie im Beweis des 1. Teil des Satzes:

   $\langle x - \frac{\langle x,x \rangle}{\langle x,y \rangle} y, x - \frac{\langle x,x \rangle}{\langle x,y \rangle} y \rangle =$

   $\langle x, x \rangle - 2\frac{\langle x,x \rangle}{\langle x,y \rangle} \langle x, y \rangle + \frac{\langle x,x \rangle \langle x,x \rangle}{\langle x,y \rangle \langle x,y \rangle} \langle y, y \rangle =$ (wegen $\langle x, y \rangle^2 = \langle x, x \rangle \cdot \langle y, y \rangle$)

   $\langle x, x \rangle - 2\langle x, x \rangle + \frac{\langle x,x \rangle \langle x,x \rangle}{\langle x,x \rangle \langle y,y \rangle} \langle y, y \rangle =$

   $\langle x, x \rangle - 2\langle x, x \rangle + \langle x, x \rangle =$

   $0.$

   Wegen der positiven Definitheit des Skalarprodukts folgt hieraus:

   $x = \frac{\langle x,x \rangle}{\langle x,y \rangle} y,$

   also die lineare Abhängigkeit.

   „ $\Longleftarrow$ “

   Sind x, y l.a., also o.E. $x = \alpha y$ mit $\alpha \in \mathbb{R}$, so folgt:

   $\langle x, y \rangle \langle x, y \rangle = \alpha^2 (\langle x, y \rangle)^2 = \langle x, x \rangle \cdot \langle y, y \rangle$

Korollar 6.1.5 *Für $x, y \in \mathbb{R}^n$ gilt $|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle}$*

Beweis: Da die Wurzelfunktion $x \mapsto \sqrt{x}$ für $x > 0$ streng monoton wachsend ist, folgt die Behauptung aus der Ungleichung von Cauchy-Schwarz.

6.2 Längen- und Winkelmessung

Bemerkung 6.2.1 *1. In der Analysis definiert man eine reelle Funktion* $\cos :$ $I\!R \to I\!R$ *durch* $\cos\varphi := \mathrm{Re}(e^{i\varphi}) = \mathrm{Re}(\sum_{k=0}^{\infty} \frac{(i\varphi)^k}{k!})$ *und eine reelle Zahl* π*, so daß gilt:*

(a) $\cos 0 = 1.$

(b) $\cos\frac{\pi}{2} = 0$

(c) $\cos\pi = -1.$

(d) Die Funktion $\cos$ *ist stetig und im Intervall* $[0, \pi]$ *streng monoton fallend; sie bildet also das Intervall* $[0, \pi]$ *bijektiv auf das Intervall* $[-1, 1]$ *ab.*

$\forall x, y \in I\!R^n \setminus \{0\}\ \exists_1 \varphi \in [0, \pi]\ [\frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle\langle y,y\rangle}} = \cos\varphi\ ]$

Beweis: $\forall x, y \in I\!R^n \setminus \{0\}\ [-1 \le \frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle\langle y,y\rangle}}\rangle \le 1\ ]$

2. Für dieses φ *gilt:* v, w *l.a.* $\Longleftrightarrow \varphi \in \{0, \pi\}.$

Definition 6.2.2 *Seien* $x, y \in I\!R^n$

1. *Seien* $x, y \neq 0$*. Die eindeutig bestimmte reelle Zahl* $\varphi \in [0, \pi]$ *mit* $\frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle\langle y,y\rangle}} = \cos\varphi$ *heißt der* **Winkel** *zwischen* x *und* y*, Bezeichnung:* $\varphi = \angle(x, y)$.
2. x *und* y *heißen* **orthogonal oder senkrecht** *zueinander — in Zeichen:* $x \perp y$ *—, wenn gilt:* $\langle x, y\rangle = 0$ *(im Fall* $x, y \neq 0$ *heißt das:* $\angle(x, y) = \cos\frac{\pi}{2}$*).*

Definition 6.2.3 *Eine Abbildung* $||\cdot|| : I\!R^n \to I\!R, x \mapsto ||x||$ *heißt* **Norm** *auf* $I\!R^n$*, wenn gilt:*

1. $\forall x \in I\!R^n, \alpha \in I\!R\ [\ ||\alpha x|| = |\alpha| \cdot ||x||\ ].$
2. $\forall x, y \in I\!R^n\ [\ ||x + y|| \le ||x|| + ||y||\ ]$ *(***Dreiecksungleichung***).*
3. $\forall x \in I\!R^n\ [\ ||x|| \ge 0\ ].$
4. $\forall x \in I\!R^n\ [\ ||x|| = 0 \Longleftrightarrow x = 0\ ].$

Bemerkung 6.2.4 *Eine Abbildung* $||\cdot|| : I\!R^n \to I\!R, x \mapsto ||x||$ *ist bereits dann eine* **Norm** *auf* $I\!R^n$*, wenn gilt:*

1. $\forall x \in I\!R^n, \alpha \in I\!R\ [\ ||\alpha x|| = |\alpha| \cdot ||x||\ ].$
2. $\forall x, y \in I\!R^n\ [\ ||x + y|| \le ||x|| + ||y||\ ]$ *(***Dreiecksungleichung***).*
3. $\forall x \in I\!R^n\ [\ ||x|| = 0 \Longleftrightarrow x = 0\ ].$

Beweis: Zu zeigen ist: $\forall x \in I\!R^n\ [\ ||x|| \ge 0\ ].$
Sei $x \in I\!R^n$*. Dann:*
$0 = |0| \cdot ||x|| = ||0x|| = ||0|| = ||x + (-x)|| \le ||x|| + ||-x|| = ||x|| + |-1| \cdot ||x|| = 2||x||.$

Satz 6.2.5 *$x \mapsto ||x|| := \sqrt{\langle x, x\rangle}$ ist eine Norm auf $\mathbb{R}^n$, sie heißt die* **euklidische Norm** *auf $\mathbb{R}^n$.*
Beweis:

1. Seien $x \in \mathbb{R}^n, \alpha \in \mathbb{R}$.
   $||\alpha x|| =$
   $\sqrt{\langle \alpha x, \alpha x\rangle} =$
   $\sqrt{\alpha^2}\sqrt{\langle x, x\rangle} =$
   $|\alpha| \cdot ||x||$.
2. Seien $x, y \in \mathbb{R}^n$.
   $||x+y||^2 =$
   $\langle x+y, x+y\rangle =$
   $\langle x, x\rangle + 2\langle xy\rangle + \langle y, y\rangle \le$
   $\langle x, x\rangle + 2\sqrt{\langle x, y\rangle^2} + \langle yy\rangle \le$ (Ungleichung von Cauchy-Schwarz)
   $\langle x, y\rangle + 2\sqrt{\langle x, x\rangle\langle y, y\rangle} + \langle y, y\rangle =$
   $(||x|| + ||y||)^2$
   $\Longrightarrow$ (wegen der Monotonie der Wurzelfunktion)
   $||x+y|| \le ||y|| + ||y||$
3. Sei $x \in V$.
   $||x|| = 0 \Longrightarrow$
   $\langle x, x\rangle = 0 \Longrightarrow$ (wegen der positiven Definitheit)
   $x = 0$.

Bemerkung 6.2.6 *Für alle $x, y \in \mathbb{R}^n \setminus \{0\}$ gilt:*
$\langle x, y\rangle = ||x|| \cdot ||y|| \cdot \cos(\angle(x, y))$

Beweis: $||x|| \cdot ||y|| \cdot \cos(\angle(x, y)) =$
$||x|| \cdot ||y|| \cdot \frac{\langle x,y\rangle}{||x||\cdot||y||} =$
$\langle x, y\rangle$.

Satz 6.2.7 *Für alle $x, y \in \mathbb{R}^n$ gilt:*

1. $||x+y||^2 = ||x||^2 + ||y||^2 + 2\langle x, y\rangle$.
2. $x \perp y \Longrightarrow ||x+y||^2 = ||x||^2 + ||y||^2$ *(***Satz von Pythagoras***)*
3. *Allgemeiner gilt für paarweise orthogonale Vektoren $v_1, \ldots, v_k \in \mathbb{R}^n$:*
   $||v_1 + \ldots + v_k||^2 = ||v_1||^2 + \ldots + ||v_k||^2$
4. $x, y \neq 0 \Longrightarrow ||x-y||^2 = ||x||^2 + ||y||^2 - 2||x|| \cdot ||y|| \cdot \cos(\angle(x, y))$ *(***Cosinussatz***).*

Beweis:

1. $||x+y||^2 =$
   $\langle x+y, x+y\rangle =$
   $\langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle =$
   $||x||^2 + ||y||^2 + 2\langle x,y\rangle.$

2. $y \perp y \Longrightarrow \langle x,y\rangle = 0 \Longrightarrow ||x+y||^2 = ||x||^2 + ||y||^2$

3. $||v_1 + \ldots + v_k||^2 =$
   $\langle v_1 + \ldots + v_k, v_1 + \ldots + v_k\rangle =$
   $\sum_{i,j=1,\ldots,k} \langle v_i, v_j\rangle =$ (da $\langle v_i, v_j\rangle = 0$ für $j \neq j$)
   $\sum_{i=1,\ldots,k} \langle v_i, v_i\rangle = ||v_1||^2 + \ldots + ||v_k||^2$

4. $x, y \neq 0 \Longrightarrow$
   $||x-y||^2 =$
   $\langle x-y, x-y\rangle =$
   $||x||^2 + ||y||^2 - 2\langle x,y\rangle.$
   Die Behauptung folgt hieraus nach der letzten Bemerkung.

6.3 Orthogonale Projektion

Lemma 6.3.1 *Seien $u, v \in I\!R^n$.*
$u \neq 0 \Longrightarrow$ *Es gibt genau einen Vektor $p \in I\!Ru$, so daß gilt: $(v-p) \perp u$.*

Beweis: Mit dem Ansatz $p = \lambda u$ erhält man
$(v-p) \perp u \Longleftrightarrow$
$0 = \langle u, v - \lambda u\rangle = \langle u,v\rangle - \lambda\langle u,u\rangle \Longleftrightarrow$
$\lambda = \frac{\langle v,u\rangle}{\langle u,u\rangle} = \frac{\langle v,u\rangle}{||u||^2}$

Definition 6.3.2 *Situation wie oben.*

1. *p heißt* **orthogonale Projektion** *von v auf u.*

2. *$h := v - p$ heißt* **Lot** *von v auf u.*

Definition 6.3.3 *Sei $U \subset I\!R^n$ ein UVR und $h \in V$.*
$h \perp U :\Longleftrightarrow \forall_{u\in U}\ h \perp u$

Lemma 6.3.4 *Seien $u_1, \ldots, u_k \in I\!R^n$ Vektoren und $U := span(u_1, \ldots, u_k)$.*
Zu jedem Vektor $v \in I\!R^n$ gibt es genau einen Vektor $p \in U$, so daß gilt: $(v-p) \perp U$.
p heißt **orthogonale Projektion** *von v auf U, $h := v - p$ heißt* **Lot** *von v auf U.*

Beweis. Der Beweis wird nur für den Fall geführt, daß $u_1, \ldots, u_k$ l.u. sind; der allgemeine Fall kann auf diesen Fall zurückgeführt werden.
Der Ansatz $p = \lambda_1 u_1 + \ldots + \lambda_k u_k$ liefert
$(v - p) \perp U \Longleftrightarrow$
$\forall_{i=1,\ldots,k}\ (v - p) \perp u_i \Longleftrightarrow$
$\forall_{i=1,\ldots,k}\ 0 = \langle u_i, v - \sum_{j=1}^{k} \lambda_j u_j \rangle = \langle u_i, v \rangle - \sum_{j=1}^{k} \lambda_j \langle u_i, u_j \rangle \Longleftrightarrow$

$$\begin{pmatrix} \langle u_1, u_1 \rangle & \ldots & \langle u_1, u_k \rangle \\ \vdots & & \vdots \\ \langle u_k, u_1 \rangle & \ldots & \langle u_k, u_k \rangle \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = \begin{pmatrix} \langle u_1, v \rangle \\ \vdots \\ \langle u_k, v \rangle \end{pmatrix}$$

(**Normalgleichung**).
In Übungsaufgabe 34 wird gezeigt: Da die Vektoren $u_1, \ldots, u_k$ l. u. sind, ist die Matrix

$$A := \begin{pmatrix} \langle u_1, u_1 \rangle & \ldots & \langle u_1, u_k \rangle \\ \vdots & & \vdots \\ \langle u_k, u_1 \rangle & \ldots & \langle u_k, u_k \rangle \end{pmatrix}$$

positiv definit.
Nun gilt allgemein: Ist eine $k \times k$-Matrix A positiv definit, so ist A invertierbar. Denn hierzu ist zu zeigen, daß das homogene lineare Gleichungssystem $Ax = 0$ nur die triviale Lösung $x = 0$ hat. Sei also $x \in \mathbb{R}^k$ mit $Ax = 0$. Dann folgt ${}^t x A x = 0$, also, da A positiv definit ist, $x = 0$.

Also ist A invertierbar. Die Normalgleichungen haben demnach eine eindeutige Lösung $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix}$.

Bemerkung 6.3.5 *Die orthogonale Projektion erfüllt folgende Minimalitätseigenschaft (Bezeichnungen wie oben):*

$$\forall_{u \in U}\ ||v - p|| \leq ||v - u||$$

D.h.: p ist die **beste Approximation von v in U.**
Beweis: Sei $u \in U$. Dann ist $p - u \in U$ und deshalb $(v - p) \perp (p - u)$. Außerdem gilt $v - u = (v - p) + (p - u)$. Nach Pythagoras folgt
$||v - u||^2 = ||v - p||^2 + ||p - u||^2 \geq ||v - p||^2$

Die folgende Anwendung zeigt eine Methode zur näherungsweisen Lösung von linearen Gleichungssystemen, z.B. bei der Auswertung von Meßreihen:

Korollar 6.3.6 (Methode der kleinsten Quadrate) *(Gauß) Sei $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Gesucht ist eine bestmögliche näherungsweise Lösung der inhomogenen linearen Gleichung $Ax - b = 0$. Es gilt:*
[$x \in \mathbb{R}^n$ ist beste Näherungslösung, d.h.
$\forall_{y \in \mathbb{R}^n}\ ||Ax - b|| \leq ||Ay - b||$] $\Longleftrightarrow$
${}^t A A x = {}^t A b$ (Normalgleichung)

Beweis: Wende 6.3.4 und die nachfolgende Bemerkung an auf die Spaltenvektoren $a^1, \ldots, a^n$ von A, den dadurch aufgespannten Spaltenraum (ein UVR des euklidischen Vektorraums $\mathbb{R}^m$ (mit dem kanonischen euklidischen Skalarprodukt)) und den Vektor $v = b$ an.

Gesucht ist die orthogonale Projektion $p = \sum_{i=1}^n x_i a^i = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Die Spalte $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ist charakterisiert durch die Normalgleichung, welche wegen

$$(\langle a^i, a^j \rangle)_{i,j=1,\ldots,n} = {}^tAA$$

und

$$\begin{pmatrix} \langle a^1, b \rangle \\ \vdots \\ \langle a^n, b \rangle \end{pmatrix} = {}^tAb$$

die o.a. Gleichung ist.

Beispiel 6.3.7 *Für* $\alpha = \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^{n+1}$ *sei* $P_\alpha(t)$ *das Polynom in* t *mit den Koeffizienten* $\alpha_0, \ldots, \alpha_n$*, also*

$$P_\alpha(t) = \sum_{i=0}^n \alpha_i t^i = (1 \ t \ \ldots \ t^n) \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Ist $\alpha_n \neq 0$*, hat* $P_\alpha(t)$ *den Grad* n*.*
Um $P_\alpha(t)$ *durch Messungen zu ermitteln, wählt man Meßpunkte* $t_1, \ldots, t_m$ *und bestimmt die Meßwerte* $b_1 = P_\alpha(t_1), \ldots, b_m = P_\alpha(t_m)$*.*
$P_\alpha(t)$ *ist durch die Werte an* $n+1$ *Stützstellen bestimmt. Wegen der Ungenauigkeiten von Messungen erhebt man jedoch mehr Meßwerte als diese Mindestanzahl und minimiert die Fehlerquadrate. Dazu wendet man 6.3.6 an und löst die Normalgleichung*

$${}^tAA \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = {}^tA \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

mit der Matrix

$$A := \begin{pmatrix} 1 & t_1 & \ldots & t_1^n \\ \vdots & & & \vdots \\ 1 & t_m & \ldots & t_m^n \end{pmatrix}$$

Sind die Meßpunkte pw. verschieden, sind die Spaltenvektoren dieser Matrix l.u. Die Lösung ist dann eindeutig bestimmt.

6.4 Euklidische Vektorräume

Definition 6.4.1 *Sei V ein Vektorraum über $\mathbb{R}$.*

1. *Ein* **Skalarprodukt** *auf V ist eine symmetrische, positiv definite Bilinearform $\sigma : V \times V \to \mathbb{R}$.*

2. *Ist ein Skalarprodukt $\sigma : V \times V \to \mathbb{R}$ auf V gegeben, so heißt V (mit diesem Skalarprodukt) ein* **euklidischer Vektorraum**.

Beispiel 6.4.2

1. *$\mathbb{R}^n$ ist mit dem kanonischen euklidischen Skalarprodukt ein euklidischer Vektorraum.*

2. *Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt* **positiv definit**, *wenn für alle $x \in \mathbb{R}^n$ gilt: $x \neq 0 \Longrightarrow^t x\ Ax > 0$.*

   Ist A eine symmetrische, positiv definite Matrix, so ist durch

   $$\forall x, y \in \mathbb{R}^n\ [\ \langle x, y\rangle =\ {}^t xAy\ ]$$

   ein Skalarprodukt auf $\mathbb{R}^n$ definiert und somit $\mathbb{R}^n$ auch mit diesem Skalarprodukt ein euklidischer Vektorraum.

3. *Sei V ein n-dimensionaler reeller Vektorraum, $\mathcal{V} := \{v_1, \ldots, v_n\}$ eine Basis auf V und $\iota_{\mathcal{V}} : \mathbb{R}^n \to V$ die lineare Abbildung, die die Basisvektoren $e_1, \ldots, e_n$ des $\mathbb{R}^n$ auf die Basisvektoren $v_1, \ldots, v_n$ von V abbildet. Jedem Vektor $v \in V$ entspricht dann genau eine Spalte $x \in \mathbb{R}^n$ mit $\iota_{\mathcal{V}}(x) = v$.*

   Sei σ ein Skalarprodukt auf $\mathbb{R}^n$. Setzt man für $v, w \in V$ $\langle v, w\rangle := \sigma(x, y)$ (wobei x und y die den Vektoren v und w zugeordneten Spalten sind), so erhält man ein Skalarprodukt auf V; V wird dadurch zu einem euklidischen Vektorraum.

4. *Ein Beispiel aus der Analysis: $V := \mathcal{C}([0,1]) = \{f : [0,1] \to \mathbb{R} : f \text{ stetig }\}$ ist mit dem Skalarprodukt $\langle f, g, \rangle := \int_0^1 f(x)g(x)dx$ ein (unendlich-dimensionaler) euklidischer Vektorraum.*

Bemerkung 6.4.3 *Die Cauchy-Schwarzsche Ungleichung und die Folgerungen im letzten Abschnitt gelten in entsprechender Weise für beliebige euklidische Vektorräume.*

6.5 Orthonormalbasen

Definition 6.5.1 *Sei V ein euklidischer Vektorraum*

1. *Eine Teilmenge $\mathcal{V}$ von V heißt* **Orthonormalsystem (ONS)**, *wenn gilt:*

   (a) $\forall v_i \neq v_j \in \mathcal{V}\ [\ v_i \perp v_j\ ]$

   (b) $\forall v_i \in \mathcal{V}\ [\ ||v_i|| = 1\ ]$

2. *Eine Basis $\mathcal{V}$ von V heißt* **Orthonormalbasis (ONB)**, *wenn sie ein ONS ist.*

Bemerkungen und Beispiele 6.5.2 *V euklidischer Vektorraum, $\mathcal{V} \subset V$.*

1. *$\mathcal{V}$ ONS $\Longrightarrow \mathcal{V}$ l.u.*

   Beweis: Seien $v_1, \ldots, v_n \in \mathcal{V}$ und $\alpha_1, \ldots, \alpha_n \in I\!R$ mit $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$. Dann gilt für $i = 1, \ldots, n$:

   $0 = \langle \alpha_1 v_1 + \ldots + \alpha_n v_n, v_i \rangle = \alpha_i.$

2. $\dim V = n, \{v_1, \ldots, v_n\} \subset V$ *ONS* $\Longrightarrow \{v_1, \ldots, v_n\}$ *ONB.*

3. *Die kanonische Basis des $I\!R^n$ ist eine ONB.*

Satz 6.5.3 *Jeder endlich-dimensionale euklidische Vektorraum besitzt eine ONB.*

Ein konstruktiver Beweis ist durch das folgende **Orthonormalisierungsverfahren nach Erhard Schmidt** (1876 – 1959) gegeben:
Sei $w_1, \ldots, w_n$ eine Basis von V.
Zu zeigen ist: Es gibt in V ein ONS der Länge n.
Es wird gezeigt: $\forall k \in \{1, \ldots, n\}$ gibt es in V ein ONS $v_1, \ldots, v_k$ der Länge k mit $< v_1, \ldots, v_k > = < w_1, \ldots, w_k >$ (aufgespannte Untervektorräume).
Der Beweis erfolgt durch durch vollständige Induktion nach k; das gewünschte ONS $v_1, \ldots, v_n$ wird also *rekursiv definiert.*

1. Induktionsanfang, $k = 1$:

   $v_1 := \frac{w_1}{||w_1||}$

   v_1 entsteht also aus w_1 durch Normierung; wegen $w_1 \neq 0$, also auch $||w_1|| \neq 0$ ist v_1 wohldefiniert und hat die Länge 1.

2. Induktionsschluß.

   Sei $1 \leq k < n$, und sei $v_1, \ldots, v_k$ ein ONS in V. Zu zeigen: Es gibt einen Vektor $v_{k+1} \in V$, so daß $v_1, \ldots, v_{k+1}$ ein ONS in V ist.

   (a) Sei p_{k+1} die orthogonale Projektion von w_{k+1} auf den UVR $span(v_1, \ldots, v_k)$; da $v_1, \ldots, v_k$ ein ONS in V ist, berechnet sich diese Projektion einfacher als im allgemeinen Fall zu
   $p_{k+1} = \sum_{i=1}^{k} \langle v_i, w_{k+1} \rangle v_i$.
   Definiere zunächst
   $\tilde{v}_{k+1} := w_{k+1} - p_{k+1} = w_{k+1} - \sum_{i=1}^{k} \langle v_i, w_{k+1} \rangle v_i$.
   Dann gilt:

      i. $\tilde{v}_{k+1} \neq 0$
      Beweis: $w_{k+1} \notin < w_1, \ldots, w_k > = < v_1, \ldots, v_k >$

ii. $\forall j \in \{1, \ldots, k\}$ $[\tilde{v}_{k+1} \perp v_j\ ]$
Beweis: Sei $j \in \{1, \ldots, k\}$.
$\langle \tilde{v}_{k+1}, v_j \rangle =$
$\langle w_{k+1} - \sum_{i=1}^{k} \langle v_i, w_{k+1} \rangle v_i, v_j \rangle =$ (Bilinearität des Skalarprodukts)
$\langle w_{k+1}, v_j \rangle - \sum_{i=1}^{k} \langle v_i, w_{k+1} \rangle \langle v_i, v_j \rangle =$ (da $v_1, \ldots, v_k$ ein ONS in V ist)
$\langle w_{k+1}, v_j \rangle - \langle v_j, w_{k+1} \rangle =$
0

iii. $< v_1, \ldots, v_k, \tilde{v}_{k+1} >=< v_1, \ldots, v_k, w_{k+1} >=< w_1, \ldots, w_{k+1} >$

(b) Setze nun
$v_{k+1} := \frac{\tilde{v}_{k+1}}{||\tilde{v}_{k+1}||}$
Dann gilt:

i. $||v_{k+1}|| = 1$.
(Klar)

ii. $\forall j \in \{1, \ldots, k\}$ $[v_{k+1} \perp v_j\ ]$
(Klar)

iii. $< v_1, \ldots, v_{k+1} >=< v_1, \ldots, v_k, \tilde{v}_{k+1} >=< w_1, \ldots, w_{k+1} >$

$\Longrightarrow v_1, \ldots, v_{k+1}$ ONS in V mit $< v_1, \ldots, v_{k+1} >=< w_1, \ldots, w_{k+1} >$

Satz 6.5.4 *Ist V ein endlich-dimensionaler euklidischer VR und $\mathcal{V}_1 = \{v_1, \ldots, v_k\}$ ein ONS in V, so gibt es ein ONS $\mathcal{V}_2 = \{v_{k+1}, \ldots, v_n\}$ in V, so daß $\mathcal{V} := \mathcal{V}_1 \cup \mathcal{V}_2$ eine ONB in V ist.*

Beweis: Sei $\mathcal{V}_1 = \{v_1, \ldots, v_k\}$ ONS in V.
$\mathcal{V}_1$ l.u. $\Longrightarrow$ (nach)
Es gibt eine Basis $\{v_1, \ldots, v_k, \tilde{v}_{k+1}, \ldots, \tilde{v}_n\}$ von V $\Longrightarrow$ (nach dem Schmidtschen Orthonormalisierungsverfahren)
Es gibt eine ONB $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ von V

Satz 6.5.5 *V euklidischer VR, $n := \dim V, \mathcal{V} = \{v_1, \ldots, v_n\}$ ONB von V und $x, y \in V$. Dann:*

1. $x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$
2. $x \neq 0 \Longrightarrow x = \sum_{i=1}^{n}\ ||x|| \cos \angle(x, v_i) v_i$
3. $\langle x, y \rangle = \sum_{i=1}^{n} \langle x, v_i \rangle \langle y, v_i \rangle$
4. $||x||^2 = \sum_{i=1}^{n} \langle x, v_i \rangle^2$ *(Parsevalsche Gleichung)*

Beweis:

1. $\forall j \in \{1, \ldots, n\}$ [$\langle \sum_{i=1}^{n} \langle x, v_i \rangle v_i, v_j \rangle =$ (Bilinearität)
$\sum_{i=1}^{n} \langle x, v_i \rangle \langle v_i, v_j \rangle =$ ($\mathcal{V}$ ONS)
$\langle x, v_j \rangle \Longrightarrow$

$\forall j \in \{1, \ldots, n\}$ [$\langle \sum_{i=1}^n \langle x, v_i \rangle v_i - x, v_j \rangle = 0 \Longrightarrow$ ($\mathcal{V}$ Basis, also $\forall y \in V \exists_1\, y_1, \ldots, y_n \in \mathbb{R}$ [$y = \sum_{j=1}^n y_j v_j$])
$\langle \sum_{i=1}^n \langle x, v_i \rangle v_i - x, y \rangle =$ (Bilinearität)
$\sum_{i=1}^n y_j \langle \sum_{i=1}^n \langle x, v_i \rangle v_i - x, v_j \rangle =$
$0.$

Setzt man hier insbesondere ein: $y = \sum_{i=1}^n \langle x, v_i \rangle v_i - x$, so folgt:
$\langle \sum_{i=1}^n \langle x, v_i \rangle v_i - x, \sum_{i=1}^n \langle x, v_i \rangle v_i - x \rangle = 0 \Longrightarrow$ (positive Definitheit)
$\sum_{i=1}^n \langle x, v_i \rangle v_i = x.$

2. Umformulierung

3. $\langle x, y \rangle =$
$\langle \sum_{i=1}^n \langle x, v_i \rangle v_i, \sum_{j=1}^n \langle y, v_j \rangle v_j \rangle =$
$\sum_{i,j=1}^n \langle x, v_i \rangle \langle y, v_j \rangle \langle v_i, v_j \rangle =$ ($\mathcal{V}$ ONS)
$\sum_{i=1}^n \langle x, v_i \rangle \langle y, v_i \rangle.$

4. $||x||^2 = \langle x, x \rangle$; die Behauptung ist also ein Spezialfall der vorherigen Aussage.

Definition 6.5.6 *V euklidischer VR, $U \subset V$ UVR.*

$$U^\perp := \{ v \in V \ : \ \forall u \in U \ [\ u \perp v \ \}$$

heißt **orthogonales Komplement** *von U.*

Bemerkung 6.5.7 *1. $U^\perp$ ist ein UVR von V.*

Beweis: Folgt sofort aus der Bilinearität des Skalarprodukts.

2. $U, U^\perp$ sind mit dem von V geerbten Skalarprodukt euklidische Vektorräume.

Bemerkung 6.5.8 *V euklidischer VR, $\dim V < \infty$, $U \subset V$ UVR $\Longrightarrow$*

1. (a) $V = U + U^\perp$, d.h., jedes $v \in V$ hat eine Summendarstellung $v = u + u^\perp$ mit $u \in U, u^\perp \in U^\perp$.

(b) $U \cap U^\perp = 0 (= \{0\})$ (daraus folgt, daß die Summendarstellung $v = u + u^\perp$ eindeutig ist).

2. Gilt $V = U + U'$ und $U \cap U' = 0$, so heißt V **direkte Summe** *von U und U'; jedes $v \in V$ ist dann eindeutig in der Form $v = u + u', u \in U, u' \in U'$ darstellbar. Bezeichnung: $V = U \oplus U'$*

3. Insbesondere: $V = U \oplus U^\perp$

4. $\dim V = \dim U + \dim U^\perp$

5. $U^{\perp\perp} = U$

Beweis: U euklid. VR $\Longrightarrow U$ besitzt eine ONB $\mathcal{U} = \{u_1, \ldots, u_k\}$.
6.5.4 $\Longrightarrow V$ besitzt ONB $\{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\}$. Dann gilt:
$U^\perp = < v_{k+1}, \ldots, v_n >$
Beweis:

1. „$\subset$“: $v \in U^\perp \Longrightarrow (\{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\}$ Basis)

   v hat Darstellung $\alpha_1 u_1 + \ldots + \alpha_k u_k + \beta_{k+1} v_{k+1} + \ldots + \beta_n v_n$

   Wegen $v \in U^\perp$ gilt $\alpha_i = 0$ für alle i, also $v \in < v_{k+1}, \ldots, v_n >$.

2. „$\supset$“: Klar

Die Behauptungen sind aus dieser Darstellung abzulesen.
Definition 6.5.9 *Sei V ein euklidischer Vektorraum, und seien $U_1, \ldots, U_n \subset V$ UVR.*
V heißt **orthogonale Summe** *der U_i $:\Longleftrightarrow$*

1. $V = U_1 + \ldots + U_n$

2. $\forall i \neq j \in \{1, \ldots, n\}[U_i \perp U_j]$, *d.h.:* $\forall u_i \in U_i, u_j \in U_j[u_i \perp u_j]$

Bemerkung 6.5.10 1. *Ein Vektorraum V heißt* **direkte Summe** *von Untervektorräumen $U_1, \ldots, U_k$ — in Zeichen: $V = U_1 \oplus \ldots \oplus U_k$ —, wenn gilt:*

   (a) $V = U_1 + \ldots + U_k$,

   (b) Sind $u_1 \in U_1, \ldots, u_k \in U_k$ gegeben, so ist jede Auswahl von $\{u_1, \ldots, u_k\}$, bei der nur von Null verschiedene Vektoren vorkommen, l.u.

2. *Ist $V = U_1 \oplus \ldots \oplus U_k$, so hat jedes $v \in V$ eine eindeutige Darstellung $v = u_1 + \ldots + u_k$, $u_1 \in U_1, \ldots, u_k \in U_k$.*

3. *Jede orthogonale Summe ist eine direkte Summe.*

4. *Jeder endlich-dimensionale euklidische VR ist orthogonale Summe von 1-dimensionalen Unterverktorräumen.*

Bemerkung 6.5.11 *Sei V ein endlich-dimensionaler euklidischer VR, $U \subset V$ ein UVR und $a \in V$.*

1. $a + U = \{v \in V : \forall u^\perp \in U^\perp[\langle u^\perp, v - a\rangle = 0]\}$

   Beweis: $x \in a + U \Longleftrightarrow x - a \in U = U^{\perp\perp}$

2. *Ist $\{u_1^\perp, \ldots, u_k^\perp\}$ eine ONB von $U^\perp$, so gilt:*

   $a + U = \{v \in V : \forall i \in \{1, \ldots, k\}[\langle u_i^\perp, v - a\rangle = 0]\}$

   Beweis: Folgt aus der vorigen Bemerkung.

3. *Sei $\{u_1^\perp, \ldots, u_k^\perp\}$ eine ONB von $U^\perp$.*

   Die Darstellung

   $a + U = \{v \in V : \forall i \in \{1, \ldots, k\}[\langle u_i^\perp, v - a\rangle = 0]\}$

   heißt **Hessesche Normalform (HNF)** *des affinen Unterraums $a + U$.*

Beispiel 6.5.12 *Ist in einem Vektorraum V ein affiner Unterraum $a+U$ gegeben, der in HNF dargestellt werden soll, ist zunächst eine ONB $\{u_1,\dots,u_k,u_{k+1}^\perp,\dots,u_n^\perp\}$ von V mit $u_1,\dots,u_k \in U$ und $u_{k+1}^\perp,\dots,u_n^\perp \in U^\perp$ zu bestimmen. Vgl. hierzu die Übungsaufgaben. Zur Abkürzung dieses Beispiels nehmen wir an, daß diese Arbeit bereits getan ist, und wir rechnen bzgl. dieser neuen Basis. Im $I\!R^3$ hat man dann i.w. die folgenden typischen Fälle:*

1. $U = I\!Re_1$, $a = e_3$.

   Dann $U^\perp = I\!Re_2 + I\!Re_3$.

   HNF: $a+U = \{x \in I\!R^3 \mid \langle x-e_3, e_2\rangle = 0 \ \wedge\ \langle x-e_3,e_3\rangle = 0\} =$

   $= \{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_2 = 0 \ \wedge\ x_3 - 1 = 0 \ \}$

2. $U = I\!Re_1 + I\!Re_2$, $a = e_3$.

   Dann $U^\perp = I\!Re_3$.

   HNF: $a+U = \{x \in I\!R^3 \mid \langle x-e_3,e_3\rangle = 0\} = \{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_3 - 1 = 0 \ \}$

7 Endomorphismen euklidischer Vektorräume

7.1 Orthogonale Endomorphismen

Lemma 7.1.1 *Sei V ein euklidischer Vektorraum und $f \in \mathrm{End}(V)$. Die folgenden Aussagen sind äquivalent:*

1. $\forall v \in V[||f(v)|| = ||v||]$. *($f$ erhält die Norm.)*

2. $\forall v,w,\in V[\langle f(v),f(w)\rangle = \langle v,w\rangle]$.*($f$ erhält das Skalarprodukt.)*

Beweis:

1. 1. $\Longrightarrow$ 2:

   Seien $v,w \in V$. Dann:

   $2\langle v,w\rangle = ||v+w||^2 - ||v||^2 - ||w||^2$ (sog. **Polarisierung**);

   entsprechend

   $2\langle f(v),f(w)\rangle = ||f(v)+f(w)||^2 - ||f(v)||^2 - ||f(w)||^2$;

   daraus folgt die Behauptung.

2. 2. $\Longrightarrow$ 1: Klar.

Definition 7.1.2 *Sei V ein euklidischer Vektorraum und $f \in \mathrm{End}(V)$.*
f heißt **orthogonal** *oder eine* **Isometrie***, wenn die äquivalenten Eigenschaften des Lemmas erfüllt sind.*
Lemma 7.1.3 *Sei V ein euklidischer Vektorraum.*

1. *$f : V \to V$ Isometrie $\Longrightarrow f$ erhält Winkel, insbesondere: $v \perp w \Longrightarrow f(v) \perp f(w)$.*

   Klar.

2. *$f : V \to V$ Isometrie $\Longrightarrow f$ Monomorphismus.*

   Beweis: $v \in \mathrm{Ker}(f) \Longrightarrow$

   $f(v) = 0 \Longrightarrow$

   $||v|| = ||f(v)|| = 0 \Longrightarrow$

   $v = 0$

3. *$f : V \to V$ Isometrie, $\dim V < \infty \Longrightarrow f$ Isomorphismus.*

   Beweis: Nach der Dimensionsformel ist f auch surjektiv.

4. *$f : V \to V$ Isometrie, $\dim V < \infty \Longrightarrow f^{-1}$ Isometrie.*

   Klar.

5. *Sei $f \in \mathrm{End}(V)$, $\dim V < \infty$.*

   f Isometrie $\Longleftrightarrow \forall \mathcal{V}$ ONB $[f(\mathcal{V})$ ONB $]$.

   Beweis:

   (a) *„$\Longrightarrow$:“ klar.*

   (b) *„$\Longleftarrow$:“ Seien $v, w \in V$; zu zeigen: $\langle f(v), f(w)\rangle = \langle v, w\rangle$.*
   Sei $\mathcal{V} = \{v_1, \dots, v_n\}$ ONB, $v = \sum_{\mu=1}^n \alpha_\mu v_\mu$, $w = \sum_{\nu=1}^n \beta_\nu v_\nu$. Dann:
   $\langle f(v), f(w)\rangle =$
   $\langle f(\sum_{\mu=1}^n \alpha_\mu v_\mu), f(\sum_{\nu=1}^n \beta_\nu v_\nu)\rangle =$ *(Bilinearität des Skalarprodukts)*
   $\sum_{\mu=1,\nu=1}^n \alpha_\mu \beta_\nu \langle f(v_\mu), f(v_\nu)\rangle =$ *(ONB)*
   $\sum_{\mu=1}^n \alpha_\mu \beta_\mu =$ *(6.5.5)*
   $\langle v, w\rangle$.

Definition 7.1.4 *Eine reelle $n \times n$-Matrix $A \in I\!R^{(n,n)}$ heißt* **orthogonal**, *wenn A invertierbar ist und wenn gilt: $A^{-1} = {}^tA$.*

Lemma 7.1.5 *Sei $A \in I\!R^{(n,n)}$. Folgende Aussagen sind äquivalent:*

1. *A orthogonal.*
2. *Die Spaltenvektoren von A bilden eine ONB bzgl. des kanonischen Skalarprodukts.*
3. *Die Zeilenvektoren von A bilden eine ONB bzgl. des kanonischen Skalarprodukts.*

Beweis: Seien $a^1, \dots, a^n$ die Spaltenvektoren von A. Dann ist
$\langle a^\mu, a^\nu \rangle = {}^t a^\mu a^\nu$
das (μ, ν)-te Element von
${}^t AA = E$
Für die Zeilenvektoren analoge Überlegung.

Satz 7.1.6 *V euklidischer Vektorraum, $\mathcal{V}$ ONB, $f \in \mathrm{End}(V)$. Dann:*
f orthogonal $\iff M_\mathcal{V}(f)$ orthogonal.

Beweis: O.E. $V = \mathbb{R}^n$ und $\mathcal{V}$ die kanonische Basis.
$M_\mathcal{V}(f)$ orthogonal $\iff$ (7.1.5)
Die Spaltenvektoren von $M_\mathcal{V}(f)$ bilden ONB $\iff$ (7.1.3, Punkt 5)
f orthogonal.

Bemerkung 7.1.7 *Sei $n \in \mathbb{N}$.*

1. *$A \in \mathbb{R}^{n \times n}$ orthogonal $\Longrightarrow A^{-1}$ orthogonal.*

   Beweis: $A^{-1} = {}^t A \Longrightarrow$

   $(A^{-1})^{-1} = A = {}^t({}^t A) = {}^t(A^{-1})$

2. *$A, B \in \mathbb{R}^{n \times n}$ orthogonal $\Longrightarrow BA$ orthogonal.*

   Beweis:

   $(BA)^{-1} = A^{-1}B^{-1} = {}^t A\, {}^t B = {}^t(BA)$

3. *$\{A \in \mathbb{R}^{n \times n} \mid A$ orthogonal $\}$ ist eine Untergruppe der Gruppe $GL(n, \mathbb{R})$ der invertierbaren $x \times n$-Matrizen. Sie heißt* **orthogonale Gruppe**, *Bezeichnung: $O(n)$.*

Beispiel 7.1.8 1. **Fall $n = 1$:**

$O(1) = \{(-1), (1)\}$

$A = (1)$ beschreibt die Identität.

$A = (-1)$ beschreibt die Spiegelung am Nullpunkt.

2. **Fall $n = 2$:**

   $O(2) =$

   $\{\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} : \varphi \in [0, 2\pi[\} \cup$ *(Drehungen um den Ursprung um den Winkel φ)*

   $\{\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} : \varphi \in [0, 2\pi[\}$ *(Spiegelungen an der Geraden durch den Ursprung mit Winkel $\frac{\varphi}{2}$ zur x-Achse)*

   Beweis: Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{(2,2)}$ mit $A^{-1} = {}^t A$. Dann:

   $${}^t AA = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = E \Longrightarrow$$

(a) $a_{11}^2 + a_{21}^2 = 1 \Longrightarrow \exists\varphi \in [0, 2\pi[\ \ [a_{11} = \cos\varphi, a_{21} = \sin\varphi]$

(b) $a_{11}a_{12} + a_{21}a_{22} = 0$

(c) $a_{12}a_{11} + a_{22}a_{21} = 0$ *(identische Gleichung wie die vorherige)*

(d) $a_{12}^2 + a_{22}^2 = 1 \Longrightarrow \exists\psi \in [0, 2\pi[\ \ [a_{12} = \sin\psi, a_{22} = \cos\psi]$

Es gilt also: $A = \begin{pmatrix} \cos\varphi & \sin\psi \\ \sin\varphi & \cos\psi \end{pmatrix}$

Die 2. Gleichung bedeutet dann:

$0 = \cos\varphi\sin\psi + \sin\varphi\cos\psi =$ *(Additionstheorem für* $\sin$*)*

$\sin(\varphi + \psi)$

$\Longrightarrow$

$\exists k \in \{0, 1, 2, 3\}[\varphi + \psi = k\pi]$

(a) $\varphi + \psi = 0 \Longrightarrow \psi = -\varphi \Longrightarrow A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$

(b) $\varphi + \psi = \pi \Longrightarrow \psi = \pi - \varphi \Longrightarrow A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$

Die gleichen Ergebnisse erhält man in den Fällen $\varphi + \psi = 2\pi$ *bzw.* $\varphi + \psi = 3\pi$*.*

8 Eigenwert-Theorie und Diagonalisierung

8.1 Aussagen über Polynome

Es werden hier Polynome mit reellen oder komplexen Koeffizienten betrachtet. K steht entsprechend für den Körper $\mathbb{R}$ der reellen Zahlen bzw. den Körper $\mathbb{C}$ der komplexen Zahlen.

Bemerkung 8.1.1

1. *Die Polynome $f \in K[X]$ können in der Form $f = \sum_{\nu=1}^{n} a_\nu X^\nu$ geschrieben werden; die a_ν heißen die* **Koeffizienten** *von f. Man verwendet auch die Schreibweise $f(X) = \sum_{\nu=1}^{n} a_\nu X^\nu$. Die Menge der Polynome ist ein Ring (man kann beliebig addieren, substrahieren und multiplizieren, aber i.a. nicht dividieren), und zwar ein kommuativer Ring mit Einselement.*

2. *Das Nullelement des Rings $K[X]$ ist das* **Nullpolynom**, *das mit* 0 *bezeichnet wird; es hat keine von Null verschiedenen Koeffizienten.*

3. *Ist $f(X) = \sum_{\nu=1}^{n} a_\nu X^\nu$ und $a_n \neq 0$, heißt n der* **Grad** *von f, Bezeichnung:* $\deg f$ *oder* $\operatorname{grad} f$. *a_n heißt dann der* **höchste Koeffizient** *von f.*

4. *Als Grad des Nullpolynoms wird definiert:* $\deg 0 := -\infty$.

5. $f, g \in K[X] \Longrightarrow \deg(fg) = \deg f + \deg g$.

   Beweis: Betrachte das Produkt der höchsten Koeffizienten von f und g.

Satz 8.1.2 (Division mit Rest in Polynomringen) *Seien $f, g \in K[X], g \neq 0$. Dann gibt es eindeutig bestimmte Polynome $q, r \in K[X]$ mit*

1. $\deg r < \deg g$
2. $f = qg + r$

Beweis:

1. Eindeutigkeit: Seien $q_1, r_1, q_2, r_2, \in K[X]$ mit $\deg r_1, \deg r_2 < \deg g$ und $f = q_1 g + r_1 = q_2 g + r_2$.

   Dann gilt: $0 = (q_1 - q_2)g + r_1 - r_2$

   Es reicht zu zeigen: $q_1 = q_2$.

   Aus der Annahme $(q_1 - q_2) \neq 0$ folgt dann $\deg(q_1 - q_2) \geq 0$, also (betrachte jeweils die höchsten Koeffizienten):

   $\deg(r_1 - r_2) = \deg((q_1 - q_2)g) = \deg(q_1 - q_2) + \deg g \geq \deg g$,

   also insbesondere $\deg r_1 \geq \deg g$ oder $\deg r_2 \geq \deg g$, Widerspruch.

2. Existenz: Fallunterscheidung:

   (a) $\exists q \in K[X]\ [qg = f]$.
   Setze $r := 0$, und es folgt die Behauptung.

(b) $\forall q \in K[X]\ [qg \neq f]$.

$M := \{\deg(f - qg) : q \in K[X]\} \subset \mathbb{N}$.

Die Menge M ist nicht leer wegen $\deg f \in M$. Es wird nun benutzt, daß jede nichtleere Teilmenge von $\mathbb{N}$ ein minimales Element besitzt (diese einfach nachzuweisende Eigenschaft von $\mathbb{N}$ heißt auch: $\mathbb{N}$ ist wohlgeordnet.)

Sei also $m := \min M$.

Dann gibt es ein $q \in K[X]$ mit $\deg(f - qg) = m$.

Setze $r := f - qg$, also $\deg r = m$. Zu zeigen: $\deg r < \deg g$.

Beweis durch Widerspruch. Annahme, $\deg r \geq \deg g$.

Sei $g := \sum_{\nu=1}^{n} a_\nu X^\nu$, $r := \sum_{\mu=1}^{m} b_\mu X^\mu$, wobei $a_n \neq 0, b_m \neq 0$ und $m \geq n$ gilt. Dann ist

$\deg r > \deg(r - g\frac{b_m}{a_n}X^{m-n})$ (betrachte wiederum die höchsten Koeffizienten)

$= \deg(f - qg - g\frac{b_m}{a_n}X^{m-n})$

$= \deg(f - (q + \frac{b_m}{a_n}X^{m-n})g)$

$\geq m$ nach Definition von m.

Also gilt $m = \deg r > m$, ein Widerspruch.

Bemerkungen und Beispiele 8.1.3 *Explizite Durchführung der Division mit Rest von Polynomen:*

Definition 8.1.4 *$a \in K$ heißt* **Nullstelle** *von $f :\Longleftrightarrow f(a) = 0$.*

Bemerkung 8.1.5 *Sei $f \in K[X]$.*

1. *$a \in K$ Nullstelle von $f :\Longrightarrow f$ ist Vielfaches von $X - a$, und zwar gilt $f = q \cdot (X - a)$ mit $\deg q = \deg f - 1$.*

   Beweis: Division mit Rest:

   $f = q \cdot (X - a) + r$ mit $\deg r < \deg(X - a) = 1$, also $r \in K$.

   Wegen $f(a) = 0$ und $(X - a)(a) = 0$ folgt: $r = 0$.

2. *$f \neq 0,\ \deg f = n \Longrightarrow f$ hat höchstens n Nullstellen in K.*

   Annahme, f hat Nullstellen $a_1, \ldots, a_{n+1}$. Durch Division erhält man Quotienten $q_1, \ldots, q_{n+1}$ mit $\deg q_i = \deg f - i$, Widerspruch.

3. **sog. Fundamentalsatz der Algebra***:Im Fall $K = \mathbb{C}$ gilt: $f \neq 0,\ \deg f = n \Longrightarrow f$ hat genau n Nullstellen in $\mathbb{C}$.*

   Dieser Satz wurde erstmals 1799 vom damals 22-jährigen Gauß bewiesen. Im Gegensatz zu seinem Namen ist er kein Satz der Algebra, sondern gehört in die Analysis. Äquivalente Formulierungen sind:

   (a) Jedes nichtkonstante Polynom über $\mathbb{C}$ hat in $\mathbb{C}$ eine Nullstelle.

   (b) Jedes Polynom $\in \mathbb{C}[X]$ zerfällt in Linearfaktoren.

4. *Über $\mathbb{R}$ gibt es nichtkonstante Polynome ohne Nullstellen, z.B.* $X^2 + 1$.

5. *Man kann zeigen: Jedes Polynom* $f \in \mathbb{R}[X]$ *von ungeradem Grad hat eine Nullstelle in* $\mathbb{R}$.

6. *Sei* $a \in K$. $\mu(f, a) := \max\{k \in \mathbb{N}_0 \;:\; f$ *ist Vielfaches von* $(X-a)^k\}$ *heißt* **Vielfachheit der Nullstelle** a *von* f.

7. *Sei* $a \in K$. $f(a) = 0 \Longleftrightarrow \mu(f, a) > 0$.

8.2 Eigenwerte und Eigenvektoren

Definition 8.2.1 1. *Sei* V *ein* $\mathbb{R}$*-Vektorraum und* $f : V \to V$ *ein Endomorphismus.*

$\lambda \in \mathbb{R}$ *heißt* **Eigenwert** *von* f, *wenn es einen Vektor* $v \neq 0$ *in* V *gibt mit* $f(v) = \lambda v$. *Der Vektor* $v \in V$ *heißt dann ein zum Eigenwert* λ *gehöriger* **Eigenvektor**.

2. *Sei* $A \in \mathbb{R}^{n\times n}$.

$\lambda \in \mathbb{R}$ *heißt* **Eigenwert** *von* A, *wenn es einen Vektor* $v \neq 0$ *in* $\mathbb{R}^n$ *gibt mit* $Av = \lambda v$. *Der Vektor* $v \in \mathbb{R}^n$ *heißt dann ein zum Eigenwert* λ *gehöriger* **Eigenvektor**.

Bemerkungen und Beispiele 8.2.2 1. *Ist* $v \in V$ *Eigenvektor zum Eigenwert* λ, *so ist auch* αv *Eigenvektor zu* λ *für jedes* $\alpha \neq 0$.

2. *Eigenvektoren sind stets* $\neq 0$ *vorausgesetzt (für* $v = 0$ *gilt* $f(v) = \lambda v$ *für alle* λ*).*

Eigenwerte können aber den Wert 0 *annehmen.*

3. *Ist* $\dim V < \infty$, *so ist* λ *genau dann Eigenwert eines Endomorphismus* $f : V \to V$, *wenn* λ *Eigenwert einer zu* f *bzgl. einer Basis* $\mathcal{V} \subset V$ *zugeordneten Matrix ist.*

Beweis: Betrachte den Isomorphismus $\iota_{\mathcal{V}} : \mathbb{R}^n \to V$.

4. *Es seien* $V = \mathbb{R}^n$, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *und* $A = (\delta_{\mu\nu}\lambda_\mu)_{\mu,\nu=1,\ldots,n}$ *die Diagonalmatrix mit den Elementen* $\lambda_1, \ldots, \lambda_n$ *in der Hauptdiagonalen.*

Dann sind $\lambda_1, \ldots, \lambda_n$ *Eigenwerte von* A *(bzw. von* $\hat{A}$*), und* $e_1, \ldots, e_n$ *sind Eigenvektoren zu* $\lambda_1, \ldots, \lambda_n$.

Die Koordinatenachsen werden also durch $\hat{A}$ *jeweils auf sich abgebildet.*

5. *Sei* V *ein VR,* $\dim V = n$, $f \in \mathrm{End}(V)$ *ein Endomorphismus von* V *und* $\mathcal{V} = \{v_1, \ldots, v_n\}$ *eine Basis von* V, *so daß jeder Basisvektor Eigenvektor von* f *ist. Seien* $\lambda_1, \ldots, \lambda_n$ *die zugehörigen Eigenwerte. Dann wird* f *bzgl. dieser Basis* $\mathcal{V}$ *durch die Matrix* $(\delta_{\mu\nu}\lambda_\mu)_{\mu,\nu=1,\ldots,n}$ *dargestellt.*

Die durch die Basisvektoren festgelegten Koordinatenachsen werden also durch f *jeweils auf sich abgebildet.*

Beweis: Betrachte das die zugeordnete Matrix definierende Diagramm.

6. *Die Matrix* $A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ *beschreibt eine Drehung des* $\mathbb{R}^2$ *um den Winkel* φ*. Ist* φ *kein Vielfaches von* π*, wird keine Gerade auf sich abgebildet; die Matrix kann also in diesem Fall keine Eigenwerte besitzen.*

7. *Die Matrix* $A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$ *beschreibt eine Spiegelung an der durch den Vektor* $\begin{pmatrix} \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} \end{pmatrix}$ *bestimmten Achse. Die folgenden Achsen werden auf sich abgebildet:*

   — *Die Achse durch den Vektor* $\begin{pmatrix} \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} \end{pmatrix}$.

   — *Die darauf senkrecht stehende Achse, also die Achse durch den Vektor* $\begin{pmatrix} \cos\frac{\varphi+\pi}{2} \\ \sin\frac{\varphi+\pi}{2} \end{pmatrix}$.

   Diese beiden Vektoren sind also Eigenvektoren; die zugehörigen Eigenwerte sind 1 *und* -1.

Definition 8.2.3 *Sei* $n \in \mathbb{N}, n \geq 1$.

1. *Sei* V *ein VR,* $\dim V = n$*. Ein Endomorphismus* $f \in \mathrm{End}(V)$ *heißt* **diagonalisierbar***, wenn* f *bzgl. einer geeigneten Basis von* V *durch eine Diagonalmatrix dargestellt wird.*

2. *Eine Matrix* $A \in \mathbb{R}^{n\times n}$ *heißt* **diagonalisierbar***, wenn* A *ähnlich zu einer Diagonalmatrix ist, d.h.:* $\exists S \in GL(n, \mathbb{R})$ $[SAS^{-1}$ *Diagonalmatrix* $]$

Bemerkung 8.2.4

1. f *diagonalisierbar* $\Longleftrightarrow$ *die (bzgl. einer festen Basis) zugeordnete Matrix ist diagonalisierbar.*

2. f *diagonalisierbar* $\Longleftrightarrow$ V *besitzt eine Basis, die nur aus Eigenvektoren von* f *besteht.*

Lemma 8.2.5 *Sei* $n \in \mathbb{N}$, V *ein* n*-dimensionaler Vektorraum und* $f \in \mathrm{End}(V)$ *ein Endomorphismus.*
Es seien $m \in \mathbb{N}, \lambda_1, \ldots, \lambda_m$ *paarweise verschiedene Eigenwerte und* $v_1, \ldots, v_m$ *Eigenvektoren zu den Eigenwerten* $\lambda_1, \ldots, \lambda_m$.
Dann sind $v_1, \ldots, v_m$ *l.u., insbesondere gilt* $m \leq n$
Beweis: Induktion nach m:

1. $m = 1$. Dann ist $v_1 \neq 0$, also l.u.

2. Sei $m > 1$ und die Aussage für je $m - 1$ pw. verschiedene Eigenwerte und zugehörige Eigenvektoren bewiesen.

   Seien $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ mit $\alpha_1 v_1 + \ldots + \alpha_m v_m = 0$. Dann gilt:

   $0 = f(\alpha_1 v_1 + \ldots + \alpha_m v_m) =$

   $\alpha_1 f(v_1) + \ldots + \alpha_m f(v_m) =$

   $\alpha_1\lambda_1 v_1 + \ldots + \alpha_m\lambda_m v_m.$

Andererseits gilt:

$0 = \lambda_m(\alpha_1 v_1 + \ldots + \alpha_m v_m) =$

$\alpha_1\lambda_m v_1 + \ldots + \alpha_m\lambda_m v_m.$

Daraus folgt:

$0 = (\alpha_1\lambda_1 v_1 + \ldots + \alpha_m\lambda_m v_m) - (\alpha_1\lambda_m v_1 + \ldots + \alpha_m\lambda_m v_m) =$

$\alpha_1(\lambda_1 - \lambda_m)v_1 + \ldots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)v_m.$

Nach Induktionsvoraussetzung sind $v_1, \ldots, v_{m-1}$ l.u.; daraus folgt:

$\alpha_1(\lambda_1 - \lambda_m) = 0$

$\vdots$

$\alpha_{m-1}(\lambda_{m-1} - \lambda_m) = 0$

und daraus, da die EW pw. verschieden sind, $\alpha_1 = \ldots = \alpha_{m-1} = 0$ und somit, da auch $\alpha_m v_m = 0$, wegen $v_m \neq 0$ somit $\alpha_m = 0$.

Korollar 8.2.6 *Sei $n \in \mathbb{N}$, V ein n-dimensionaler Vektorraum und $f \in \mathrm{End}(V)$ ein Endomorphismus.*
Voraussetzung: f besitze n paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.
Dann ist f diagonalisierbar.
Der Beweis folgt aus dem Lemma und der letzten Bemerkung, da man auf die Existenz einer Basis aus Eigenvektoren von f schließen kann.

8.3 Das charakteristische Polynom, Diagonalisierung

Lemma 8.3.1 *1. Sei V ein Vektorraum, $f \in \mathrm{End}(V)$ ein Endomorphismus und $\lambda \in \mathbb{R}$. Die folgenden Aussagen sind äquivalent:*

(a) λ *ist EW von* f.

(b) $\ker(f - \lambda\mathrm{id}_V) \neq 0$

Sei $A \in \mathbb{R}^{n\times n}$ und $\lambda \in \mathbb{R}$. Die folgenden Aussagen sind äquivalent:

(a) λ *ist EW von* A.

(b) $\ker(A - \lambda E) \neq 0$

Der Beweis ist klar.

Lemma 8.3.2 *Sei $n \in \mathbb{N}$.*

1. *Seien V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$ ein Endomorphismus und $\lambda \in \mathbb{R}$. Die folgenden Aussagen sind äquivalent:*

   (a) λ *ist EW von* f.

   (b) $\det(f - \lambda\mathrm{id}_V) = 0.$

2. *Seien $A \in K^{n\times n}$, $E = E_n$ und $\lambda \in \mathbb{R}$. Die folgenden Aussagen sind äquivalent:*

   (a) λ *ist EW von* A.

(b) $\det(A - \lambda E) = 0$.

Beweis: Man kann o.E. annehmen: $V = \mathbb{R}^n, f = \hat{A}$. Dann gilt:
$\ker(f - \lambda \mathrm{id}_V) \neq 0 \Longleftrightarrow$
$0 < \dim \ker(f - \lambda \mathrm{id}_V) =$ (Dimensionsformel)
$n - \dim \mathrm{Im}(f - \lambda \mathrm{id}_V) =$
$n - \mathrm{Rg}(A - \lambda E) \Longleftrightarrow$
$\mathrm{Rg}(A - \lambda E) < n \Longleftrightarrow$
$\det(A - \lambda E) = 0$.
$\det(f - \lambda \mathrm{id}_V) = 0$.

Definition 8.3.3 (Charakteristisches Polynom) *Sei $n \in \mathbb{N}$.*

1. *Sei $A = (a_{\mu\nu}) \in \mathbb{R}^{n \times n}$. Das Polynom*

$$\chi_A := \sum_{\sigma \in S_n} \varepsilon(\sigma)(a_{1\sigma(1)} - \delta_{1\sigma(1)}X) \cdot \ldots \cdot (a_{n\sigma(n)} - \delta_{n\sigma(n)}X)$$

   heißt **charakteristisches Polynom** *von A.*

2. *Es seien V ein n-dimensionaler K-Vektorraum, $\mathcal{V}$ eine Basis von V und $f \in \mathrm{End}(V)$. Das Polynom*

$$\chi_f := \chi_{M^{\mathcal{V}}_{\mathcal{V}}(f)}$$

   heißt **charakteristisches Polynom** *von f.*

Bemerkung 8.3.4 *Sei $n \in \mathbb{N}$.*

1. $A = (a_{\mu\nu}) \in \mathbb{R}^{n \times n} \Longrightarrow \chi_A$ *ist ein Polynom vom Grad n.*

   Beweis: Der Summand mit $\sigma = (1)$ liefert beim Ausmultiplizieren einen Term X^n; alle anderen Summanden enthalten nur Terme von Grad $< n$.

2. $A = (a_{\mu\nu}) \in \mathbb{R}^{n \times n} \Longrightarrow \forall \lambda \in \mathbb{R}\ [\chi_A(\lambda) = \det(A - \lambda E)]$

   Beweis: Leibniz-Formel.

3. $A = (a_{\mu\nu}) \in \mathbb{R}^{(n,n)} \Longrightarrow [\chi_A = \det(A - XE)]$

   Beweis: Folgt in entsprechender Weise aus der Leibniz-Formel für Matrizen mit Koeffizienten im Ring $\mathbb{R}[X]$.

4. $f \in \mathrm{End}(V), \dim V = n < \infty \Longrightarrow \chi_f = \det(f - X\mathrm{id}_V)$.

   Beweis: Ähnliche Überlegung.

   Insbesondere gilt:

5. $f \in \mathrm{End}(V), \dim V = n < \infty \Longrightarrow \chi_f$ *ist unabhängig von der Auswahl der Basis $\mathcal{V}$ definiert.*

Satz 8.3.5 *Sei $n \in \mathbb{N}$ und $\lambda \in \mathbb{R}$.*

1. *Sei* $A = (a_{\mu\nu}) \in I\!R^{n\times n}$. *Dann gilt:*

   λ *EW von* $A \Longleftrightarrow \chi_A(\lambda) = 0.$

2. *Sei* $f \in \mathrm{End}(V), \dim V = n < \infty$. *Dann gilt:*

   λ *EW von* $f \Longleftrightarrow \chi_f(\lambda) = 0.$

Beweis: Folgt aus den obigen Bemerkungen und aus dem vorangegangenen Lemma.

Definition 8.3.6 *Sei* V *ein Vektorraum,* $f \in \mathrm{End}(V)$ *ein Endomorphismus und* $\lambda \in I\!R$. *Der UVR*

$$\mathrm{Eig}(f;\lambda) := \ker(f - \lambda \mathrm{id}_V) = \{v \in V : f(v) = \lambda v\}$$

von V *heißt* Eigenraum *von* f *bezüglich* λ.

Bemerkung 8.3.7 1. $\lambda_1 \neq \lambda_2 \Longrightarrow \mathrm{Eig}(f;\lambda_1) \cap \mathrm{Eig}(f;\lambda_2) = 0$

Beweis: Wäre $v \neq 0$ *in* $\mathrm{Eig}(f;\lambda_1) \cap \mathrm{Eig}(f;\lambda_2)$, *so wäre* v *EV zu* λ_1 *und zu* λ_2*; zwei EV zu zwei verschiedenen EWen* λ_1, λ_2 *müssen aber l.u. sein.*

2. *Zur Bestimmung des Eigenraums zu einem gegebenen Eigenwert* λ *muß man das homogene Gleichungssystem*

   $(A - \lambda E)x = 0$

   lösen; dabei sei A *die dem Endomorphismus* f *bei fester Wahl einer Basis zugeordnete Matrix.*

Beispiel 8.3.8 1. $A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \Longrightarrow$

$\chi_A(\lambda) = \det \begin{pmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{pmatrix} =$

$(\cos\varphi - \lambda)^2 + \sin^2\varphi =$

$\cos^2\varphi - 2\lambda\cos\varphi + \lambda^2 + \sin^2\varphi =$

$\lambda^2 - 2\lambda\cos\varphi + 1.$

Dieses Polynom hat nur im Fall $\cos^2\varphi - 1 =$ *reelle Lösungen, und zwar die Lösungen* 1 *bzw.* -1*; der Eigenraum ist jeweils* $I\!R^2$.

Das ist klar aus der geometrischen Situation heraus.

Für einen Beweis für den Eigenwert $\varphi = \pi, \lambda = -1$ *muß man das Gleichungssystem* $(A - \lambda E)x = 0$. *für diesen Fall lösen. Es ist*

$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E$, *also*

$A - \lambda E = 0;$

daraus folgt die Behauptung.

2. *Die komplexen Eigenwerte von* $A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ *sind* $\lambda_1 := \cos(\varphi) + \sqrt{(\cos(\varphi))^2 - 1}$ *und* $\lambda_2 := \cos(\varphi) - \sqrt{(\cos(\varphi))^2 - 1}$

   Für $\varphi = \frac{\pi}{3}$ *ergibt sich die Matrix*

   $$\begin{bmatrix} 1/2 & -1/2\sqrt{3} \\ 1/2\sqrt{3} & 1/2 \end{bmatrix}$$

   und als Eigenwerte $\lambda_1 = 1/2 - 1/2\sqrt{-1}\sqrt{3}$*,* $\lambda_2 = 1/2 + 1/2\sqrt{-1}\sqrt{3}$*, der Eigenraum ist jeweils eindimensional.*

   Zur genauen Bestimmung des Eigenraums z.B. für λ_1 *hat man das Gleichungssystem* $(A - \lambda_1 E)x = 0$ *zu lösen, also (für* $\varphi = \frac{\pi}{3}$ *):*

   $$\begin{bmatrix} 1/2\sqrt{-1}\sqrt{3} & -1/2\sqrt{3} \\ 1/2\sqrt{3} & 1/2\sqrt{-1}\sqrt{3} \end{bmatrix} x = 0$$

   Durch Addition des i-fachen der 1. Zeile zur 2. Zeile wird dieses Gleichungssystem:

   $$\begin{bmatrix} 1/2\sqrt{-1}\sqrt{3} & -1/2\sqrt{3} \\ 0 & 0 \end{bmatrix} x = 0$$

   Durch Division der ersten Zeile durch $1/2\sqrt{-1}\sqrt{3}$ *erhält man (als Gauß-Jordan-Form) das Gleichungssystem*

   $$\begin{bmatrix} 1 & \sqrt{-1} \\ 0 & 0 \end{bmatrix} x = 0$$

   Setzt man $x_2 = i$ *ein, berechnet man* $x_1 = 1$*; der Eigenraum zu* λ_1 *wird also aufgespannt durch den Vektor* $\begin{pmatrix} 1 \\ i \end{pmatrix}$.

Bemerkung 8.3.9 *Sei* $n \in \mathbb{N}$.

1. *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

   Beweis: Seien $A, B \in \mathbb{R}^{n \times n}$ *und* $S \in GL(n, \mathbb{R})$ *mit* $B = SAS^{-1}$*. Für die Matrizen* $A - XE, B - XE$ *(Matrizen mit Koeffizienten in* $\mathbb{R}[X]$*) gilt dann:*

   $S(A - XE)S^{-1} = SAS^{-1} - SXES^{-1} = SAS^{-1} - XSES^{-1} = B - XE,$

   daraus folgt:

   $\det(B - XE) = \det S \cdot \det(A - XE) \cdot \det S^{-1} = \det(A - XE)$

2. *Aus der letzten Bemerkung folgt ebenfalls die Unabhängigkeit des charakteristischen Polynoms* χ_f *von der Auswahl der Basis.*

Lemma 8.3.10 *Sei* $n \in \mathbb{N}$.

1. V *ein n-dimensionaler VR und* $f \in \mathrm{End}(V)$ *ein Endomorphismus.*

   f *diagonalisierbar* $\Longrightarrow \chi_f$ *zerfällt in Linearfaktoren* $\lambda_i - X$*; die Nullstellen* λ_i *sind EW von* f*.*

2. *Sei $A \in I\!R^{n\times n}$.*

   A diagonalisierbar $\Longrightarrow \chi_A$ zerfällt in Linearfaktoren $\lambda_i - X$; die Nullstellen λ_i sind EW von A.

Beweis: Es genügt, die erste der beiden Aussagen zu beweisen.
Aus der Voraussetzung folgt, daß es n l.u. EV $v_1, \ldots, v_n$ gibt; bei Wahl dieser EV als Basis ist die f darstellende Matrix A eine Diagonalmatrix mit den zugehörigen (nicht notwendig verschiedenen) Eigenwerten $\lambda_1, \ldots, \lambda_n$ als Hauptdiagonalelementen.
Damit ist auch $A - XE$ eine Diagonalmatrix; die Diagonalelemente sind $\lambda_1 - X, \ldots, \lambda_n - X$.
Es folgt:
$\chi_f = \det(A - XE) = (\lambda_1 - X) \cdot \ldots \cdot (\lambda_n - X) =$
$(-1)^n (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_n)$.
Die EW $\lambda_1, \ldots, \lambda_n$ müssen nicht pw. verschieden sein. Faßt man gleiche EW zusammen, erhält man:
$\chi_f = (-1)^n (X - \lambda_1)^{r_1} \cdot \ldots \cdot (X - \lambda_k)^{r_k} =$
$(-1)^n (X - \lambda_1)^{\mu(\chi_f, \lambda_1)} \cdot \ldots \cdot (X - \lambda_k)^{\mu(\chi_f, \lambda_k}$,
wobei $\lambda_1, \ldots, \lambda_k$ die pw. verschiedenen EW und $\mu(\chi_f, \lambda_1), \ldots, \mu(\chi_f, \lambda_1)$ ihre Vielfachheiten als Nullstellen von χ_f sind.

Lemma 8.3.11 *Sei $n \in I\!N, V$ ein n-dimensionaler VR und $f \in \mathrm{End}(V)$ ein Endomorphismus. Das charakteristische Polynom zerfalle in Linearfaktoren:*
$\chi_f = (-1)^n (X - \lambda_1)^{\mu(\chi_f, \lambda_1)} \cdot \ldots \cdot (X - \lambda_k)^{\mu(\chi_f, \lambda_k}$.
Dann gilt für $i = i, \ldots, k$:
$1 \le \dim \mathrm{Eig}(f, \lambda_i) \le \mu(\chi_f, \lambda_i)$.

Definition: $\mu(\chi_f, \lambda_i)$ wird in diesem Zusammenhang die **algebraische Vielfachheit** von λ_i genannt, $\dim \mathrm{Eig}(f, \lambda_i)$ die **geometrische Vielfachheit**.

Beweis: Sei $i = i, \ldots, k$, $\lambda := \lambda_i$.
λ EW von $f \Longrightarrow \mathrm{Eig}(f, \lambda) \ne 0 \Longrightarrow \dim \mathrm{Eig}(f, \lambda) > 0$.
Zu zeigen: $\dim \mathrm{Eig}(f, \lambda) \le \mu(\chi_f, \lambda)$.
Sei $\{v_1, \ldots, v_s\}$ eine Basis von $\mathrm{Eig}(f, \lambda)$.
Dann sind $v_1, \ldots, v_s$ l.u.; nach dem Steinitzschen Austauschsatz gibt es daher eine Basis $\{v_1, \ldots, v_s, v_{s+1}, \ldots, v_n\}$ von V

Bzgl. dieser Basis wird f dargestellt durch eine Matrix der Form $\begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix}$ mit

$A_1 = \begin{pmatrix} \lambda & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & \lambda \end{pmatrix}$

Daraus folgt nach 5.3.2, 9:
$\chi_f = \det(A - XE) =$
$\det \begin{pmatrix} A_1 - XE_s & C \\ 0 & A_2 - XE_{n-s} \end{pmatrix} =$
$(\lambda - X)^s \cdot \chi_{A_2}$.
Hieraus ergibt sich die Behauptung.

Beispiel 8.3.12 1. $A = \begin{pmatrix} 0 & 1 \\ -1 & -4 \end{pmatrix}$

Das charakteristische Polynom ist $\chi_A(X) = \det \begin{pmatrix} -X & 1 \\ -1 & -4-X \end{pmatrix} = X^2 + 4X + 1$

Es gibt also zwei verschiedene Eigenwerte $\lambda_i = -2 \overset{+}{-} \sqrt{3}$.

Die zugehörigen Eigenvektoren v_i *sind l.u.; sie bilden also eine Basis; damit ist* A *diagonalisierbar.*

Es ist $v_1 = \begin{pmatrix} 2-\sqrt{3} \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2-\sqrt{3} \\ -1 \end{pmatrix}$.

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

Das charakteristische Polynom ist $\chi_A(X) = \det \begin{pmatrix} -X & 1 \\ -1 & -2-X \end{pmatrix} = X^2 + 2X + 1 = (X+1)^2$

Es gibt also nur einen EW $\lambda = -1$ *mit* $\mu(\chi_a; -1) = 2$.

Es ist $\operatorname{Eig}(A; -1) = L(A+E) = L(\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix})$.

Wegen $\operatorname{rg}(A+E) = 1$ *gilt* $\dim \operatorname{Eig}(A; -1) = 1 < 2$. *Eine Basis von* $\operatorname{Eig}(A; -1)$ *ist gegeben z.B. durch den Eigenvektor* $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Es kann also keine aus Eigenvektoren von A *bestehende Basis des* $\mathbb{R}^2$ *geben* $\Longrightarrow A$ *ist nicht diagonalisierbar.*

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

Es ist $\chi_A(X) =$

$$\det \begin{pmatrix} 3-X & 4 & 3 \\ -1 & -X & -1 \\ 1 & 2 & 3-X \end{pmatrix} =$$

$$-\det \begin{pmatrix} 1 & X & 1 \\ 1 & 2 & 3-X \\ 3-X & 4 & 3 \end{pmatrix} =$$

$$-\det \begin{pmatrix} 1 & X & 1 \\ 0 & 2-X & 2-X \\ 3-X & 4 & 3 \end{pmatrix} =$$

$$(2-X)\det \begin{pmatrix} 1 & X & 1 \\ 0 & 1 & 1 \\ 0 & 4-3X+X^2 & X \end{pmatrix} =$$

$(2-X)\det\begin{pmatrix} 1 & 1 & X \\ 0 & 1 & 1 \\ 0 & X & 4-3X+X^2 \end{pmatrix} =$

$(2-X)\det\begin{pmatrix} 1 & 1 & X \\ 0 & 1 & 1 \\ 0 & 0 & 4-4X+X^2 \end{pmatrix} =$

$= (2-X)^3.$

Die Matrix A hat also den einzigen Eigenwert 2, und es gilt: $\mu(\chi_a, 2) = 3$.

Betrachte nun $\mathrm{Eig}(A; 2) = L(A-2E) = L(\begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix})$.

Ohne Rechnung sieht man:

$\mathrm{rg}(A-2E) = 2$, *also* $\dim L(A-2E) = 1 < 3$. *Es gibt damit — bis auf Vielfache* $\neq 0$ *— nur einen einzigen Eigenvektor, nämlich* $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$*; damit ist A nicht diagonalisierbar.*

Beispiel 8.3.13 (Federpendel) *Eine Masse m sei an einer Feder aufgehängt; in Ruhestellung habe sie die Position $x = 0$.*
Bei Abweichung von dieser Ruheposition um den Wert x wirkt auf die Masse die Kraft $-Dx$; dabei ist D die sog. Federkonstante.
Bei Bewegung der Masse wird die Geschwindigkeit durch die Ableitung $u = x'$ von $x(t)$ nach der Zeit t beschrieben; die Beschleunigung durch die zweite Ableitung $x'' = u'$.
Die Federkraft $-Dx$ erzeugt nach Newton eine Beschleunigung x'', für die gilt: $mx'' = -Dx$.
Ferner wirkt bei der Geschwindigkeit x' ene Reibungskraft $-kx'$.
Wenn keine weitere äußere Kraft wirkt, hat man daher die Gleichung $mx'' = -Dx - kx'$. Setzt man, wie üblich, $\frac{D}{m} = \omega^2$, $\frac{k}{m} = 2\mu$, erhält man die Differentialgleichung der gedämpften Schwingung

$$x'' + 2\mu x' + \omega^2 x = 0$$

Diese Gleichung ist gleichbedeutend mit den beiden Gleichungen

$$x' = u$$

$$u' = -\omega^2 x - 2\mu u$$

,
also mit der Vektorgleichung

$$\begin{pmatrix} x' \\ u' \end{pmatrix} = A \begin{pmatrix} x \\ u \end{pmatrix}$$

mit

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}$$

Im letzten Beispiel hatten wir diese Matrix betrachtet mit den Werten $\omega = 1$ *und* $\mu = 2$ *bzw.* $\mu = 1$.
Im ersten Beispiel ($\mu = 2$*, also* $\mu > |\omega|$*) ist diese Matrix* A *über* $\mathbb{R}$ *diagonalisierbar — das gilt stets für* $\mu > \omega$.
Im zweiten Beispiel ($\mu = 1$*, also* $\mu = |\omega|$*) ist die Matrix* A *nicht diagonalisierbar.*
Man kann zeigen (Übungsaufgabe!), daß im Fall $0 \leq \mu < |\omega|$ *die Matrix* A *über* $\mathbb{C}$ *diagonalisierbar ist.*
Die physikalische Bedeutung ist:

$0 \leq \mu < \omega$	*schwache Dämpfung*	*komplex diagonalisierbar*	*periodische Schwingung*
$\mu = \omega$	*aperiodischer Grenzfall*	*nicht diagonalisierbar*	
$\mu > \omega$	*starke Dämpfung*	*reell diagonalisierbar*	*langsames Abklingen*

Auf eine weitere Ausführung der Zusammenhänge wird hier verzichtet.
Bemerkung: *Im aperiodischen Grenzfall ist* A *trigonalisierbar, d.h.: ähnlich zu einer Dreiecksmatrix.*

Satz 8.3.14 *Sei* $n \in \mathbb{N}, V$ *ein* n*-dimensionaler VR und* $f \in \operatorname{End}(V)$ *ein Endomorphismus.*
Die folgenden Bedingungen sind äquivalent:

1. f *ist diagonalisierbar.*
2. *Das charakteristische Polynom* χ_f *zerfällt in Linearfaktoren*

   $\chi_f = (-1)^n (X - \lambda_1)^{\mu(\chi_f, \lambda_1)} \cdot \ldots \cdot (X - \lambda_k)^{\mu(\chi_f, \lambda_k}$,

   und es gilt für $i = i, \ldots, k$:

   $\dim \operatorname{Eig}(f, \lambda_i) = \mu(\chi_f, \lambda_i)$.
3. V *ist die direkte Summe der Untervektorräume* $\operatorname{Eig}(f, \lambda_1), \ldots, \operatorname{Eig}(f, \lambda_k)$*, d.h.*

   $V = \operatorname{Eig}(f, \lambda_1) \oplus \ldots \oplus \operatorname{Eig}(f, \lambda_k)$. *d.h.:*

   (a) $V = \operatorname{Eig}(f, \lambda_1) + \ldots + \operatorname{Eig}(f, \lambda_k)$.

   (b) *Sind* $w_1 \in \operatorname{Eig}(f, \lambda_1) \setminus \{0\}, \ldots, w_k \in \operatorname{Eig}(f, \lambda_k) \setminus \{0\}$*, so sind* $w_1, \ldots, w_k$ *l.u.*

Beweis:

1. Sei f diagonalisierbar.

   Nach dem vorangegangenen Lemma zerfällt χ_f in Linearfaktoren, d.h. $\chi_f = (-1)^n (X - \lambda_1)^{\mu(\chi_f, \lambda_1)} \cdot \ldots \cdot (X - \lambda_k)^{\mu(\chi_f, \lambda_k}$, und die Nullstellen $\lambda_1, \ldots, \lambda_k$ sind EW von f.

   Zu zeigen: $\forall i = i, \ldots, k$ [$s_1 = \dim \operatorname{Eig}(f, \lambda_i) = \mu(\chi_f, \lambda_i) = r_i$].

   Es gibt eine Basis $\{v_1, \ldots, v_n\}$ von V aus EV von f, und zwar kann man o.E. annehmen:

$v_1, \ldots, v_{s_1}$ sind EV zum EW λ_1,

$v_{s_1+1}, \ldots, v_{s_1+s_2}$ sind EV zum EW λ_2,

$\vdots$

$v_{s_1+\ldots+s_{k-1}+1}, \ldots, v_n$ sind EV zum EW λ_k.

Also gilt $s_1 + \ldots + s_k = n = r_1 + \ldots + r_k$; zusammen mit $s_i \leq r_i$ für alle i folgt die Behauptung.

2. Sei nun vorausgesetzt: Das charakteristische Polynom χ_f zerfällt in Linearfaktoren

   $\chi_f = (-1)^n (X - \lambda_1)^{\mu(\chi_f, \lambda_1)} \cdot \ldots \cdot (X - \lambda_k)^{\mu(\chi_f, \lambda_k}$,

   und es gilt für $i = i, \ldots, k$:

   $\dim \mathrm{Eig}(f, \lambda_i) = \mu_i := \mu(\chi_f, \lambda_i)$.

   Dann gibt Basen der Eigenräume $\mathrm{Eig}(f, \lambda_i)$ der Länge $\mu_i, i = 1, \ldots, k$; es seien:

   — $v_1, \ldots, v_{\mu_1}$ l.u. EV zum EW λ_1,

   $v_{\mu_1+1}, \ldots, v_{\mu_1+\mu_2}$ l.u. EV zum EW λ_2,

   $\vdots$

   $v_{\mu_1+\ldots+\mu_{k-1}+1}, \ldots, v_{\mu_1+\ldots+\mu_k} = v_n$ l.u. EV zum EW λ_k.

   Dann sind $v_1, \ldots, v_n$ l.u.

   Denn: Seien $\alpha_1, \ldots \alpha_n \in \mathbb{R}$ mit $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$, und sei angenommen: nicht alle α_i sind 0. Sei etwa $\alpha_1 \neq 0$. Dann ist $w_1 := \alpha_1 v_1 + \ldots + \alpha_{\mu_1} v_{\mu_1}$ EV zum EW λ_1. Wegen $w_1 + \alpha_{\mu_1+1} v_{\mu_1+1} + \ldots + \alpha_n v_n = 0$ gibt es dann weitere, analog aufgebaute Vektoren $w_i = \alpha_{\mu_1+\ldots+\mu_{i-1}+1} v_{\mu_1+\ldots+\mu_{i-1}+1} + \ldots + \alpha_{\mu_1+\ldots\mu_i} v_{\mu_1+\ldots+\mu_i}$, $i = 2, \ldots, n$ mit $w_1 + \ldots + w_n = 0$. Das kann aber nicht sein, da die w_i entweder 0 oder EV zum EW λ_i sind und da EV zu paarweise verschiedenen Eigenwerten nach Lemma 8.2.5 l.u. sind.

   Also bilden $v_1, \ldots, v_n$ eine Basis von V. Daraus folgt unmittelbar die Darstellung

   $V = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_k)$.

3. Sei nun vorausgesetzt: $V = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_k)$. Dann hat man wiederum die Basis $\{v_1, \ldots, v_n\}$ von V aus EV von f, sodaß gilt:

   $v_1, \ldots, v_{s_1}$ sind EV zum EW λ_1,

   $v_{s_1+1}, \ldots, v_{s_1+s_2}$ sind EV zum EW λ_2,

   $\vdots$

   $v_{s_{k-1}+1}, \ldots, v_n$ sind EV zum EW λ_k.

   Bezüglich dieser Basis wird f dargestellt durch die Diagonalmatrix mit den Diagonalelementen λ_1 (s_1−mal), …, λ_k (s_k−mal).

Algorithmus 8.3.15 *Bestimmung aller Eigenwerte eines Endomorphismus bzw. einer Matrix und Diagonalisierung.*
Es reicht, den Fall einer Matrix $A \in \mathbb{R}^{n \times n}$ zu betrachten.

1. *Berechne* $\chi_A = \det(A - XE)$.

   I.a. wird man dazu $A - XE$ *durch elementare Zeilenumformungen in Zeilenstufenform bringen. Bei Zeilenvertauschungen erhält man einen Faktor* -1*; ggflls sind weitere Faktoren zu berücksichtigen. Man erhält also eine obere Dreiecksmatrix, deren Determinante das Produkt der Elemente in der Hauptdiagonalen ist.*

   Bei kleineren Matrizen ist auch die direkte Anwendung der Leibniz-Formel sinnvoll. Häufig verwendet man zur Berechnung die sog. Laplace-Entwicklung, ein rekursives Verfahren zur Leibniz-Formel.

2. *Zerlege* χ_A *in Linearfaktoren:*

   $\chi_A = (-1)^n (X - \lambda_1)^{r_1} \cdot \ldots \cdot (X - \lambda_k)^{r_k}$

   Bemerkung: Dieser Teil des Algorithmus ist nicht direkt Teil der Linearen Algebra, sondern — je nach Aufgabenstellung – der Algebra bzw. der Analysis. Denn für $n > 4$ *gibt es keine Lösungsformel zur Bestimmung der Nullstellen. Die Nullstellen können häufig nur näherungsweise durch numerische Verfahren bestimmt werden.*

3. *Bestimme für jeden Eigenwert* λ_1 *den Eigenraum* $\mathrm{Eig}(f, \lambda_i)$ *durch Lösung des homogenen linearen Gleichungssystems* $(A - \lambda_i E)x = 0$.

4. *Stimmen algebraische und geometrische Vielfachheiten aller EW überein, kann man direkt die zu* A *ähnliche Diagonalmatrix hinschreiben, d.h., die Diagonalmatrix mit den Diagonalelementen* λ_1 *(*s_1*−mal), ...,* λ_k *(*s_k*−mal).*

   Im anderen Fall ist die Matrix (bzw. der Endomorphismus) nicht diagonalisierbar.

Bemerkung 8.3.16 *Ist* f *ein Endomorphismus eines endlich-dimensionalen Vektorraums, dessen charakteristisches Polynom in Linearfaktoren zerfällt, stimmen aber algebraische und geometrische Vielfachheiten der Eigenwerte nicht überein, so ist* f *nach dem Satz nicht diagonalisierbar. Jedoch kann* f *durch eine (obere oder untere) Dreiecksmatrix dargestellt werden (Trigonalisierung). Insbesondere ist jede Matrix* $\in \mathbb{C}^{n\times n}$ *über* $\mathbb{C}$ *trigonalisierbar.*

9 Symmetrische reelle Matrizen

9.1 Eigenwerte symmetrischer reeller Matrizen

Satz 9.1.1 *$A \in \mathbb{R}^{n\times n}$ symmetrische $n \times n$- Matrix $\Longrightarrow$ Das charakteristische Polynom χ_A zerfällt über $\mathbb{R}$ in Linearfaktoren.*

Beweis: Nach dem Fundamentalsatz der Algebra gibt es komplexe Zahlen $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ mit $\chi_A = (\lambda_1 - X) \cdot \ldots \cdot (\lambda_n - X)$. Zu zeigen: Die λ_ν sind reell.
Sei $\nu \in \{1, \dots, n\}$ und $\lambda := \lambda_\nu$. Zu zeigen: $\lambda \in \mathbb{R}$.
Wegen $\chi_A(\lambda) = 0$ gibt es einen Vektor $v \in \mathbb{C}^n$ mit $(A - \lambda E)v = 0$, d.h. $Av = \lambda v$.

Das Matrizenprodukt ${}^t\bar{v}v$ ist reell und > 0 ; denn ist $v = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} x_1 + iy_1 \\ \vdots \\ x_n + iy_n \end{pmatrix}$,
so ist

$$ {}^t\bar{v}v = (x_1 - iy_1, \dots, x_n - iy_n) \begin{pmatrix} x_1 + iy_1 \\ \vdots \\ x_n + iy_n \end{pmatrix} = \sum_{\nu=1}^{n} (x_\nu^2 + y_\nu^2) $$

Es gilt:

$\lambda({}^t\bar{v}v) =$
${}^t\bar{v}(\lambda v) =$
${}^t\bar{v}(Av) =$ wegen ${}^t\bar{A} = A$
$({}^t\bar{v}\, {}^t\bar{A})v =$
${}^t(\overline{Av})v =$
${}^t(\overline{\lambda v})v =$
$\bar{\lambda}\, {}^t\bar{v}v$.

Daraus folgt: $\lambda = \bar{\lambda}$, also $\lambda \in \mathbb{R}$.

Beispiel 9.1.2 *Die symmetrische Matrix*

$$ A := \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{13}{72} & \frac{-5}{72} \\ 0 & \frac{-5}{72} & \frac{13}{72} \end{pmatrix} $$

hat das charakteristische Polynom $\chi_A = -(x+1) * (x^2 - 13/36 * x + 1/36)$, also die Eigenwerte $-1,\ 1/9,\ 1/4$.

Bemerkung 9.1.3 *I.a. muß eine Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt, nicht diagonalisierbar sein. Wir werden aber sehen, daß dies bei symmetrischen reellen Matrizen sehr wohl der Fall ist.*

9.2 Selbstadjungierte Endomorphismen euklidischer Vektorräume

Sei $n \in \mathbb{N}$ und V ein n-dimensionaler euklidischer Vektorraum mit dem Skalarprodukt $\langle \cdot, \cdot \rangle$.

Definition 9.2.1 *Ein Endomorphismus $f : V \to V$ heißt* **selbstadjungiert**, *wenn gilt:*

$$\forall_{v,w \in V} \ \langle (f(v), w\rangle = \langle v, f(w)\rangle$$

Lemma 9.2.2 *Sei $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V und $f \in \mathrm{End}(V)$. Dann gilt:*

$$f \quad \text{selbstadjungiert} \Longleftrightarrow \mathcal{M}_{\mathcal{V}}(f) \quad \text{symmetrisch}$$

Beweis: Für $v, w \in V$ und $i \in \{1, \ldots, n\}$ setzen wir $x_i := \langle v, v_i\rangle, y_i := \langle w, v_i\rangle$, also $v = x_1 v_1 + \ldots + x_n v_n, w = y_1 v_1 + \ldots + y_n v_n$. Sei $A := \mathcal{M}_{\mathcal{V}}(f)$. Dann wird v bzgl $\mathcal{V}$ durch die Spalte $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dargestellt, w durch die Spalte $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, ferner $f(v)$ durch die Spalte Ax und $f(w)$ durch die Spalte Ay. Es folgt:
$\langle v, f(w)\rangle = {}^t x A y$
$\langle f(v), w\rangle = {}^t(Ax)y = {}^t x \, {}^t A y$

Gilt $A = {}^t A$, so folgt: f selbstadjungiert.
Zum Beweis der Umkehrung wendet man die obigen Gleichungen an auf die Spezialfälle $v = v_i, w = v_j$ $(i, j \in \{1, \ldots, n\}$ und erhält (mit $A = (a_{\mu\nu})$: $\quad a_{ij} = a_{ji}$.

Bemerkung 9.2.3 *Es sei V ein euklidischer Vektorraum und $f : V \to V$ ein selbstadjungierter Endomorphismus. Sind $\lambda \neq \mu$ verschiedene Eigenwerte von f, so sind die zugehörigen Eigenvektoren orthogonal.*

Beweis: Es sei $v \in V$ Eigenvektor zum Eigenwert λ und $w \in W$ Eigenvektor zu μ. Dann gilt:
$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle f(v), w\rangle = \langle v, f(w)\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle$.
Daraus folgt $(\lambda - \mu)\langle v, w\rangle = 0$. Wegen $\lambda \neq \mu$ folgt $\langle v, w\rangle = 0$

Der Satz im folgenden Abschnitt liefert eine Basis aus Eigenvektoren, wobei ebenfalls die zu verschiedenen Eigenwerten gehörigen Eigenvektoren zueinander orthogonal sind. Das ergibt sich, ohne daß die obige Bemerkung benutzt werden muß.

9.3 Diagonalisierung symmetrischer reeller Matrizen

Satz 9.3.1 (Spektralsatz für selbstadjungierte Endomorphismen) *Es sei V ein endlich-dimensionaler euklidischer Vektorraum und $f : V \to V$ ein selbstadjungierter Endomorphismus.*
Dann gibt es eine Orthonormalbasis $\mathcal{V} = \{v_1, \ldots, v_n\}$ von V aus Eigenvektoren von f.

Beweis: Sei $\mathcal{W}$ eine beliebige Orthonormalbasis von V und $A := \mathcal{M}_{\mathcal{W}}(f)$. Nach 9.2.2 ist A symmetrisch; nach 9.1.1 zerfällt χ_A über $\mathbb{R}$ in Linearfaktoren. f hat also reelle Eigenwerte $\lambda_1, \ldots, \lambda_n$.
Konstruktion von $\mathcal{V} = \{v_1, \ldots, v_n\}$ durch vollständige Induktion nach n.
Ist $n = 1$, wählt man als v_1 einen beliebigen Eigenvektor zum Eigenwert λ_1 mit $||v_1|| = 1$ und ist fertig.
Sei nun $n > 1$ und sei angenommen, daß die Behauptung für jeden euklidischen Vektorraum der Dimension $n - 1$ richtig ist.
Wähle wiederum v_1 als einen beliebigen Eigenvektor zum Eigenwert λ_1 mit $||v_1|| = 1$ und setze $U :=< v_1 >^{\perp} = \{u \in V \mid \langle u, v_1\rangle = 0\}$.
Dann gilt $\dim U = n - 1$.
Ferner ist $f(U) \subset U$, denn für $u \in U$ gilt: $\langle f(u), v_1\rangle = \langle u, f(v_1)\rangle = \langle u, \lambda_1 v_1\rangle = \lambda_1\langle u, v_1\rangle = 0$. Somit ist $f|U \to U$ ein selbstadjungierter Endomorphismus von U.
Nach Induktionsvoraussetzung gibt es eine Orthonormalbasis $v_2, \ldots, v_n$ von U aus Eigenvektoren von $f|U$, also auch von f.
$\mathcal{V} = \{v_1, \ldots, v_n\}$ ist dann die gesuchte ONB von V.

Korollar 9.3.2 *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch.*
Dann gibt es eine orthogonale Matrix $S \in O(n)$, so daß

$$S^{-1}AS = {}^tSAS$$

eine Diagonalmatrix ist; die Hauptdiagonalelemente sind die Eigenwerte von A.

Bemerkung 9.3.3 *Es sei V ein endlich-dimensionaler euklidischer Vektorraum und $f : V \to V$ ein selbstadjungierter Endomorphismus.*
Seien $\lambda_1, \ldots, \lambda_k$ die Eigenwert von f und $V_i := \mathrm{Eig}(f; \lambda)$. Dann ist V die orthogonale direkte Summe von $V_1, \ldots, V_k$; jeder Vektor $v \in V$ hat also eine eindeutige Darstellung $v = v_1 + \ldots + v_k$ mit $v_i \in V_i$. Die Vektoren v_i sind pw. zueinander orthogonal, und jedes v_i ist die orthogonale Projektion von v auf V_i, vgl. 6.3.2.
Für jedes $i = 1, \ldots, k$ sei $\pi_i : V \to V_i$ die Abbildung, die einem Vektor v seine orthogonale Projektion $v_i \in \mathrm{Eig}(f; \lambda)$ zuordnet. Dann gilt für jedes $v \in V$:

$f(v) = f(v_1+\ldots+v_k) = f(v_1)+\ldots+f(v_k) = \lambda_1\cdot v_1+\ldots+\lambda_k\cdot v_k = \lambda_1\cdot\pi_1(v)+\cdot\lambda_k\cdot\pi_k(v)$,
also

$$f = \lambda_1 \cdot \pi_1 + \cdot\lambda_k \cdot \pi_k$$

Diese Darstellung heißt **Spektralzerlegung des selbstadjungierten Endomorphismus** f.

Die Menge der Eigenwerte eines Endomorphismus f wird auch das Spektrum von f genannt.

Beispiel 9.3.4 *Betrachte $I\!R^3$ mit dem kanonischen euklidischen Skalarprodukt und den durch die Matrix*

$$A := \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{13}{72} & \frac{-5}{72} \\ 0 & \frac{-5}{72} & \frac{13}{72} \end{pmatrix}$$

definierten selbstadjungierten Endomorphismus $f = \hat{A} : I\!R^3 \to I\!R^3$ (vgl. Beispiel 9.1.2). Wie oben festgestellt, ist das charakteristische Polynom $\chi_A = (x+1) * (x^2 - 13/36{*}x{+}1/36)$; die Eigenwerte sind $\lambda_1 = -1,\ \lambda_2 = 1/9,\ \lambda_3 = 1/4$ und die zugehörigen normierten Eigenvektoren $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$.
Da die Eigenvektoren zu paarweise verschiedenen Eigenwerten gehören, sind sie nach 9.2.3 orthogonal, was man auch unmittelbar nachprüft.
Dies ist also eine Orthonormalbasis von Eigenvektoren von A.

Bemerkung 9.3.5 *Ist $f : V \to V$ ein selbstadjungierter Endomorphismus eines euklidischen Vektorraums mit Eigenräumen höherer Dimension als 1, bestimmt man eine Orthonormalbasis aus Eigenwerten von f, indem man für jeden Eigenraum eine Orthonormalbasis berechnet.*

9.4 Hauptachsentransformation

Bemerkung 9.4.1 *Es sei $n \in I\!N$ und $s : I\!R^n \times I\!R^n \to I\!R$ eine Bilinearform (vgl. Definition 2.1.1.1).*

1. *Für eine Basis $\mathcal{V} = \{v_1, \ldots, v_n\}$ des $I\!R^n$ definiert man (wie bereits für Skalarprodukte, vgl. Definition 2.1.1.4) die Matrix*

$$A := \mathcal{M}_{\mathcal{V}}(s) := (s(v_i, v_j))_{i,j=1,\ldots,n} \in I\!R^{n \times n}$$

2. *Sind $x, y \in I\!R^n$ und $v := x_1v_1 + \ldots + x_nv_n, w := y_1v_1 + \ldots + x_nv_n$, so gilt $s(v, w) = {}^t xAy$.*

3. *Ist s symmetrisch, so auch A. In diesem Fall ist also $\hat{A} : I\!R^n \to I\!R^n$ selbstadjungiert.*

4. *Ist s positiv definit, so auch A.*

5. *(Transformationsformel, vgl. 2.1.1.6)*
   Seien $\mathcal{V} = \{v_1, \ldots, v_n\}$, $\tilde{\mathcal{V}} = \{\tilde{v_1}, \ldots, \tilde{v_n}\}$ Basen von V.

Sei $T := \mathcal{T}_{\mathcal{V}}^{\tilde{\mathcal{V}}} = \mathcal{M}_{\mathcal{V}}^{\tilde{\mathcal{V}}}(\mathrm{id}_{I\!R^n})$ die zum Basiswechsel $\tilde{\mathcal{V}} \mapsto \mathcal{V}$ gehörige Transformationsmatrix.

Dann gilt:

$$\mathcal{M}_{\tilde{\mathcal{V}}}(s) = {}^tT\mathcal{M}_{\mathcal{V}}(s)T = {}^t\mathcal{T}_{\mathcal{V}}^{\tilde{\mathcal{V}}}\mathcal{M}_{\mathcal{V}}(s)\mathcal{T}_{\mathcal{V}}^{\tilde{\mathcal{V}}}$$

6. *Sind $\mathcal{V} = \{v_1, \ldots, v_n\}$ ′ und $\tilde{\mathcal{V}} = \{\tilde{v_1}, \ldots, \tilde{v_n}\}$ ONB, so ist die Transformationsmatrix $T := \mathcal{T}_{\mathcal{V}}^{\tilde{\mathcal{V}}}$ orthogonal.*

Satz 9.4.2 (Hauptachsentransformation) *Es sei $n \in I\!N$ und $s : I\!R^n \times I\!R^n \to I\!R$ eine symmetrische Bilinearform. Sei $\mathcal{E} = \{e_1, \ldots, e_n\}$ die kanonische Basis des $I\!R^n$ und $A := \mathcal{M}_{\mathcal{E}}(s)$.*

1. *Ist $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine aus Eigenvektoren von A bestehende Orthonormalbasis (bzgl. des kanonischen euklidischen Skalarprodukts) des $I\!R^n$ (eine solche existiert stets nach 9.3.1), so gilt:*

$$\mathcal{M}_{\mathcal{V}}(s) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

   Dabei sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A zu den Eigenvektoren $v_1, \ldots, v_n$.

2. *Hat A k positive und l negative Eigenwerte, so gilt:*

   Es gibt eine Orthogonalbasis $\mathcal{W} = \{w_1, \ldots, w_n\}$ des $I\!R^n$ mit

$$\mathcal{M}_{\mathcal{W}}(s) = \begin{pmatrix} E_k & & 0 \\ & -E_l & \\ 0 & & 0 \end{pmatrix} = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& 0 && \\ &&&&&&& \ddots & \\ &&&&&&&& 0 \end{pmatrix}$$

Beweis:

1. Sei $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine aus Eigenvektoren von A bestehende Orthonormalbasis, und seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A zu den Eigenvektoren $v_1, \ldots, v_n$, also $Av_i = \lambda_i v_i$, $i = 1, \ldots, n$.

   Für $i, j = 1, \ldots, n$ gilt dann: $s(v_i, v_j) = {}^tv_iAv_j = {}^tv_j\lambda_jv_j = \delta_{ij}v_j$.

2. O.E. seien $\lambda_1, \ldots, \lambda_k > 0$, $\lambda_{k+1}, \ldots, \lambda_{k+l} < 0$ und $\lambda_{k+l+1}, \ldots, \lambda_n = 0$. Setzt man

$$w_i := \begin{cases} |\lambda_i|^{-\frac{1}{2}} v_i & \text{für } i = 1, \ldots, k+l, \\ v_i & \text{für } i = k+l+1, \ldots, n, \end{cases}$$

so erhält man eine Orthogonalbasis von $\mathbb{R}$, und es gilt:

$$s(w_i, w_j) := \begin{cases} 1 & \text{für } 1 \leq i = j \leq k, \\ -1 & \text{für } k+1 \leq i = j \leq k+l, \\ 0 & \text{sonst.} \end{cases}$$

Korollar 9.4.3 *Voraussetzungen wie im Satz.*

1. *Es gibt eine orthogonale $n \times n$-Matrix $S \in O(n)$ mit*

$$S^{-1}AS = {}^tSAS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

   Dabei sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A.

2. *Hat A k positive und l negative Eigenwerte, so gilt:*

   Es gibt eine invertierbare $n \times n$-Matrix $T \in GL(n; \mathbb{R})$ mit

$${}^tTAT = \begin{pmatrix} E_k & & 0 \\ & -E_l & \\ 0 & & 0 \end{pmatrix} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}$$

9.5 Positive Definitheit und Eigenwerte

Satz 9.5.1 *Eine symmetrische reelle Matrix $A \in \mathbb{R}^{n \times n}$ ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.*

Beweis: A ist genau dann positiv definit, wenn die symmetrische Bilinearform

$$s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \quad (x,x) \mapsto {}^t x A y$$

positiv definit ist.
Nach dem Satz über die Hauptachsentransformation wird s bzgl. einer aus Eigenvektoren von A bestehende Orthonormalbasis (bzgl. des kanonischen euklidischen Skalarprodukts auf $\mathbb{R}^n$) $\mathcal{V} = \{v_1, \ldots, n_n\}$ von $\mathbb{R}^n$ dargestellt durch eine Diagonalmatrix

$$B = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix};$$

dabei sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A zu den Eigenvektoren $v_1, \ldots, v_n$.
Es folgt für jedes $v \in \mathbb{R}^n$: Ist v bzgl $\mathcal{V}$ dargestellt durch den Spaltenvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, d.h., $v = x_1 v_1 + \ldots + x_n v_n$, so ist

$$s(v,v) = {}^t x B x = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$$

Sind alle Eigenwerte $\lambda_i > 0$ folgt $s(v,v) > 0$ für jedes $v \neq 0$, d.h.: s ist positiv definit.
Sind nicht alle Eigenwerte positiv, so findet man ein $v \neq 0$ mit $s(v,v) \leq 0$; in diesem Fall ist s nicht positiv definit.

Korollar 9.5.2 *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch.*
A positiv definit $\Longrightarrow$ $\det A > 0$

Beweis: Nach 9.4.3 gibt es eine orthogonale Matrix $S \in O(n)$ mit

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

Daraus folgt: $\det A = \lambda_1 \cdot \ldots \cdot \lambda_n$. Ist A positiv definit, so sind die Eigenwerte λ_i und damit auch ihr Produkt positiv.

9.6 Das Hauptminorenkriterium für positive Definitheit

Definition 9.6.1 *Sei K ein Körper, $n \in \mathbb{N}$ und $A = (a_{ij})_{i,j=1,\ldots,n} \in K^{n\times n}$ eine quadratische Matrix. Sei $k \in \{1, \ldots, n\}$.*
Dann heißt

$$\det(a_{ij})_{i,j=1,\ldots,k}$$

der k-te **Hauptminor** *von A; dies ist also die Determinante der k-reihigen quadratischen Untermatrix von A links oben.*

Für $k \in \{1, \ldots, n\}$ sei $A_k := (a_{ij})_{i,j=1,\ldots,k}$

Satz 9.6.2 (Hauptminorenkriterium von Jacobi) *Eine symmetrische reelle Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn alle Hauptminoren* $\det A_k$*, $k = 1, \ldots, n$, positiv sind.*

Beweis.

1. Sei A positiv definit. Dann ist auch A_k positiv definit für jedes $k \in \{1, \ldots, n\}$; denn für $\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$ ist

$$(x_1, \ldots, x_k) A_k \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = (x_1, \ldots, x_k, 0, \ldots, 0) A \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

   Die Behauptung folgt aus 9.5.2.

2. Es wird nun durch vollständige Induktion nach n gezeigt: Sind alle Hauptminoren einer n-reihigen symmetrischen reellen Matrix A positiv, so ist A positiv definit.

   (a) $n = 1$: Klar

   (b) Sei nun $n > 1$ und die Behauptung für $n-1$ vorausgesetzt. Sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix, und alle Hauptminoren $\det A_k$, $k = 1, \ldots, n$, seien positiv. Nach Induktionsvoraussetzung ist dann $A' := A_{n-1}$ positiv definit.

   Nach dem Satz über die Hauptachsentransformation bzw. seinem Korollar gibt es eine orthogonale Matrix $S \in O(n)$ mit

$$^tSAS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} =: B,$$

   wobei B die gleiche symmetrische Bilinearform s darstellt wie A, jedoch in Bezug auf eine andere Orthonormalbasis.

   Setze

$$V_{>0} := \{\, ^tSx \mid x \in \mathbb{R}^n, x_n = 0\}$$

   .

   Dann gilt $\dim V_{>0} = n - 1$, und für alle $y = \, ^tSx \in V_{>0} \setminus \{0\}$ gilt $^tyBy =$

$$
\begin{aligned}
&{}^t y({}^t SAS)y = \\
&{}^t(Sy)A(Sy) = {}^t(S\,{}^tSx)A(S\,{}^tSx) = \\
&{}^t xAx = \\
&(x_1, \ldots, x_{n-1})A_{n-1} \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} > 0.
\end{aligned}
$$

Unter den Diagonaleinträgen $\lambda_1, \ldots, \lambda_n$ kann es nicht zwei nichtpositive Werte geben.
Denn nimmt man etwa an: $\lambda_i \leq 0$, $\lambda_j \leq 0$, $i \neq j$, so würde folgen:

$$ {}^t e_i B e_i \leq 0, \ {}^t e_j B e_i \leq 0 $$

Setzt man dann

$$ V_\square :=< e_i, e_j >, $$

so folgt $\dim V_\square = 2$ und $\mathbb{R}^n \supset V_{>0} \oplus V_\square$; daraus folgt dann
$n = \dim \mathbb{R}^n >= \dim V_{>0} + \dim V_\square = 2 + (n-1)$, ein Widerspruch.

Andererseits ist $\lambda_1 \cdot \ldots \cdot \lambda_n = \det B = \det A = \det A_n > 0$, also sind alle $\lambda_i > 0$.
Damit ist B positiv definit, also auch s und damit auch A.

10 Normale Endomorphismen euklidischer Vektorräume

10.1 Adjungierte lineare Abbildungen

Definition 10.1.1 *Es seien V, W endlich-dimensionale euklidische Vektorräume und $\mathcal{V} = \{v_1, \ldots, v_n\}$, $\mathcal{W} = \{w_1, \ldots, m_m\}$ Orthonormalbasen von V bzw. von W. Sei $f : V \to W$ eine lineare Abbildung.*
Die lineare Abbildung $f^{ad} : W \to V$, die definiert ist durch

$$ \mathcal{M}^{\mathcal{W}}_{\mathcal{V}}(f^{ad}) := {}^t(\mathcal{M}^{\mathcal{V}}_{\mathcal{W}}(f)) $$

heißt die **zu f adjungierte lineare Abbildung**.

Bemerkungen und Beispiele 10.1.2 *Bezeichnungen wie in der Definition.*

1. *Ist speziell $V = \mathbb{R}^n, W = \mathbb{R}^m$ mit dem kanonischen euklidischen Skalarprodukt und sind $\mathcal{V}, \mathcal{W}$ die Standard-Orthonormalbasen, so gilt:*

   Ist $f = \hat{A}$ mit $A \in \mathbb{R}^{m \times n}$, so ist $f^{ad} = \hat{{}^tA}$

2. *Für $v \in V, w \in W$ gilt:*

$$\langle f(v), w \rangle = \langle v, f^{ad}(w) \rangle$$

Beweis für den o.g. Spezialfall: Seien $x \in I\!R^n, y \in I\!R^m$.
$\langle f(x), y \rangle = {}^t(Ax)y = {}^tx\, {}^tAy = \langle x, f^{ad}(y) \rangle$

3. *Im Fall $V = W$ gilt:*

$$f \quad \text{selbstadjungiert} \Longleftrightarrow f = f^{ad}$$

4. *Speziell im Fall $V = W = I\!R^n, f = \hat{A}$ gilt:*

$$f \quad \text{selbstadjungiert} \Longleftrightarrow f = f^{ad} \Longleftrightarrow A = {}^tA$$

5. *Man kann leicht sehen (vgl. z.B. Fischer, Lineare Algebra, 6.2.4):*

$$\mathrm{Im} f^{ad} = (\mathrm{Ker} f)^{\perp}$$

$$\mathrm{Ker} f^{ad} = (\mathrm{Im} f)^{\perp}$$

10.2 Normale Matrizen und Endomorphismen

Definition 10.2.1 1. *Sei V ein endlich-dimensionaler euklidischer Vektorraum Und $f : V \to V$ ein Endomorphismus.*

f heißt **normal**, *wenn gilt:*

$$f \circ f^{ad} = f^{ad} \circ f$$

2. *Sei $n \in I\!N$ und $A \in I\!R^{n \times n}$.*

A heißt **normal**, *wenn gilt:*

$$A\, {}^tA = {}^tAA$$

Bemerkungen und Beispiele 10.2.2 1. *f selbstadjungiert $\Longrightarrow f$ normal.*

2. *A symmetrisch $\Longrightarrow A$ normal.*

3. *A orthogonal $\Longrightarrow A$ normal.*

4. *Gilt für eine Matrix A: $\quad A = -\,{}^tA$, so ist A normal.*

5. *A normal, $\alpha \in I\!R \Longrightarrow \alpha A$ normal.*

6. *Die Matrix* $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ *ist normal.*

Lemma 10.2.3 *$f : V \to V$ normal $\Longrightarrow$ $\mathbf{Ker} f^{ad} = \mathbf{Ker} f$ und $\mathbf{Im} f^{ad} = \mathbf{Im} f$*
Ferner ist V die orthogonale direkte Summe von $\mathbf{Ker} f$ und $\mathbf{Im} f$.
Beweis: $v \in \mathbf{Ker} f \Longleftrightarrow$
$f(v) = 0 \Longleftrightarrow$
$0 = \langle f(v), f(v)\rangle = \langle v, f^{ad}(f(v))\rangle = \langle v, f(f^{ad}(v))\rangle = \langle f(f^{ad}(v)), v\rangle == \langle f^{ad}(v), f^{ad}(v)\rangle \Longleftrightarrow$
$f^{ad}(v) = 0 \Longleftrightarrow$
$v \in \mathbf{Ker} f^{ad}$

Die zweite Behauptung folgt nun mit 10.1.2, 5.

Korollar 10.2.4 *f und f^{ad} haben die gleichen Eigenwerte und Eigenvektoren.*

Beweis: Betrachte die zu f gehörige Matrix A.
Sei $\lambda \in \mathbb{R}$.
A normal $\Longrightarrow A - \lambda E$ normal $\Longrightarrow \mathbf{Ker}(A - \lambda E) = \mathbf{Ker}\ {}^t(A - \lambda E)$.

Satz 10.2.5 (Spektralsatz für normale Endomorphismen) *Sei V ein endlichdimensionaler euklidischer Vektorraum und $f : V \to V$ ein Endomorphismus.*
Das charakteristische Polynom χ_f zerfalle über $\mathbb{R}$ in Linearfaktoren.
Dann sind folgende Eigenschaften äquivalent:

1. *f ist normal.*
2. *Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f.*

Bemerkung: Die Voraussetzung „Das charakteristische Polynom χ_f zerfalle über $\mathbb{R}$ in Linearfaktoren.“ ist hier deswegen notwendig, weil wir keine systematische Theorie von Vektorräumen über $\mathbb{C}$ entwickelt haben, insbesondere nicht die Theorie der unitären Vektorräume, dem komplexen Analogon zu den euklidischen Vektorräumen. Für den allgemeinen Satz und Beweis siehe z.B. Fischer: Lineare Algebra, 6.2.7.

Zum Beweis.

1. $2 \Longrightarrow 1$:

   Sei $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V und seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte zu den Eigenvektoren $v_1, \ldots, v_n$.

   Für $i = 1, \ldots, n$ gilt dann: $f(v_i) = \lambda_i v_i$ und $f^{ad}(v_i) = \lambda_i v_i$.

   Daraus folgt $f \circ f^{ad}(v_i) = \lambda^2 v_i = f^{ad} \circ f(v_i)$.

   Da somit die linearen Abbildungen $f \circ f^{ad}$ und $f^{ad} \circ f$ auf den Basiselementen übereinstimmen, stimmen sie überall überein.

   Ein zweiter Beweis ergibt sich durch die Betrachtung der definierenden Matrizen:

f wird bzgl. der Basis $\mathcal{V}$ dargestellt durch die Diagonalmatrix mit $\lambda_1, \ldots, \lambda_n$ in der Hauptdiagonalen; nennen wir diese Matrix B.

Bzgl. der kanonischen Basis werde f dargestellt durch die Matrix A. Dann sind A und B ähnlich mit einer orthogonalen Transformationsmatrix S, d.h.

$B = {}^tSAS$, also $A = BB\,{}^tS$ und ${}^tA = S\,{}^tB\,{}^tS$.

Zusammen mit ${}^tBB = B\,{}^tB$ (trivial für eine Diagonalmatrix) rechnet man hieraus leicht nach: ${}^tAA = A\,{}^tA$.

2. $1 \Longrightarrow 2$ (der eigentliche Spektralsatz):

   Der Beweis geht analog zum Beweis des Spektralsatzes für selbstadjungierte Endomorphismen (9.3.1):

   Konstruktion von $\mathcal{V} = \{v_1, \ldots, v_n\}$ durch vollständige Induktion nach n.

   Ist $n = 1$, wählt man als v_1 einen beliebigen Eigenvektor zum Eigenwert λ_1 mit $||v_1|| = 1$ und ist fertig.

   Sei nun $n > 1$ und sei angenommen, daß die Behauptung für jeden euklidischen Vektorraum der Dimension $n-1$ richtig ist.

   Wähle wiederum v_1 als einen beliebigen Eigenvektor zum Eigenwert λ_1 mit $||v_1|| = 1$ und setze $U := < v_1 >^{\perp} = \{u \in V \mid \langle u, v_1 \rangle = 0\}$.

   Dann gilt $\dim U = n - 1$, und V die orthogonale direkte Summe von $\langle v_1 \rangle$ und U.

   Ferner ist $f(U) \subset U$, denn für $u \in U$ gilt:

   $\langle f(u), v_1 \rangle = \langle u, f^{ad}(v_1) \rangle = \langle u, \lambda_1 v_1 \rangle = \lambda_1 \langle u, v_1 \rangle = 0$. Somit ist $f|U \to U$ ein Endomorphismus von U.

   Man zeigt ebenso: $f^{ad}(U) \subset U$.

   Somit ist $f|U \to U$ ein normaler Endomorphismus

   Nach Induktionsvoraussetzung gibt es eine Orthonormalbasis $v_2, \ldots, v_n$ von U aus Eigenvektoren von $f|U$, also auch von f.

   $\mathcal{V} = \{v_1, \ldots, v_n\}$ ist dann die gesuchte ONB von V.

11 Verallgemeinerte Inverse von Matrizen

Literatur:
David A. Harville: Matrix Algebra from a Statistician's Perspective. Springer 1997.
Vgl. auch:
James R. Schott: Matrix Analysis for Statistics. Wiley 1997.

11.1 Rechts- und Linksinverse

Definition 11.1.1 *Seien $m, n \in \mathbb{N}$. Sei $A \in \mathbb{R}^{m\times n}$.*

1. $R \in \mathbb{R}^{n\times m}$ **Rechtsinverse** *zu* $A \Longleftrightarrow AR = E_m$.
2. $L \in \mathbb{R}^{n\times m}$ **Linksinverse** *zu* $A \Longleftrightarrow LA = E_n$.

Bemerkung 11.1.2 $m = n, A \in \mathbb{R}^{n\times n}$ *regulär* $\Longrightarrow A^{-1}$ *Rechts- und Linksinverse zu* A.

Lemma 11.1.3 *Seien $m, n \in \mathbb{N}$. Sei $A \in \mathbb{R}^{m\times n}$.*

1. *A besitzt eine Rechtsinverse R* $\Longleftrightarrow \mathrm{rg}A = m$
2. *A besitzt eine Linksinverse L* $\Longleftrightarrow \mathrm{rg}A = n$

Beweis:

1. (a) $'' \Longrightarrow''$:
   $AR = E_m \Longrightarrow$
   $m = \dim\mathrm{im}(E_m)^\wedge = \dim\mathrm{im}(AR)^\wedge = \dim\mathrm{im}(A^\wedge \circ R^\wedge) = \dim\mathrm{im}A^\wedge = \mathrm{rg}A$

   (b) $'' \Longleftarrow''$:
   $\mathrm{rg}A = m \Longrightarrow$
   $\forall \mu \in \{1, \ldots, m\} \ : \ \mathrm{rg}(A, e_\mu) = \mathrm{rg}A$ (dabei sei e_μ der μ-te kanonische Basisvektor des $\mathbb{R}^m$) $\Longrightarrow$ (die Rangbedingung für die Lösbarkeit des inhomogenen Gleichungssystems $Ax = e_\mu$ ist erfüllt)
   $\forall \mu \in \{1, \ldots, m\} \exists r^\mu \in \mathbb{R}^n \ : \ Ar^\mu = e_\mu \Longrightarrow$ (mit $R := (r^1, \ldots, r^m) \in \mathbb{R}^{n\times m}$)
   $AR = E_m$

2. Wegen ${}^t(LA) = {}^tA\,{}^tL$ gilt:

   A besitzt Linksinverse $L \Longleftrightarrow$

   ${}^tA$ besitzt Rechtsinverse (nämlich ${}^tL$) $\Longleftrightarrow$ (nach dem 1. Teil des Lemmas)

   $\mathrm{rg}{}^tA = n \Longleftrightarrow$ (wegen $\mathrm{rg}{}^tA = \mathrm{rg}A$)

   $\mathrm{rg}A = n$

Korollar 11.1.4 *Eine Matrix A besitzt genau dann sowohl eine Rechtsinverse wie eine Linksinverse, wenn A quadratisch und regulär ist.*

Lemma 11.1.5 *Sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix.*

1. *Besitzt A eine Rechtsinverse R, so ist A regulär, und es gilt $R = A^{-1}$.*
2. *Besitzt A eine Linksinverse L, so ist A regulär, und es gilt $L = A^{-1}$.*

Beweis:

1. Besitzt A eine Rechtsinverse R, so folgt aus dem vorherigen Lemma: $\mathrm{rg}A = n$. Also ist A regulär.

   Ebenso folgt aus dem vorherigen Lemma: A besitzt Linksinverse L.

   Wegen $L = LE = LAR = ER = R$ folgt, daß R auch Linksinverse zu A ist, also $R = A^{-1}$.
2. Ebenso.

Korollar 11.1.6 *Sei $A \in \mathbb{R}^{n\times n}$ eine reguläre quadratische Matrix. Dann besitzt A keine anderen Rechts- oder Linksinversen als A^{-1}.*

Bemerkung 11.1.7 *Sei $A \in \mathbb{R}^{m\times n}$ mit $\mathrm{rg}A = n$. Dann besitzt A eine Linksinverse L.*
Für jeden Vektor $b \in \mathbb{R}^m$ mit $\mathrm{rg}A = \mathrm{rg}(A, b)$ ist dann das Gleichungssystem $Ax = b$ eindeutig lösbar, und zwar findet man als Lösung: $x = Lb$.

11.2 Verallgemeinerte Inverse

Definition 11.2.1 *Seien $m, n \in \mathbb{N}$. Sei $A \in \mathbb{R}^{m\times n}$.*
$G \in \mathbb{R}^{n\times m}$ **verallgemeinerte Inverse** *(Pseudoinverse) zu $A \Longleftrightarrow AGA = A$.*

Bemerkung 11.2.2 *Jede Rechts- oder Linksinverse ist verallgemeinerte Inverse. Insbesondere ist die inverse Matrix einer regulären Matrix eine verallgemeinerte Inverse.*

Lemma 11.2.3 *Sei $A \in \mathbb{R}^{n\times n}$ eine reguläre quadratische Matrix. Dann besitzt A keine anderen verallgemeinerten Inversen als A^{-1}.*
Beweis: Sei G verallgemeinerte Inverse zu A. Dann:
$G = A^{-1}AGAA^{-1} = A^{-1}AA^{-1} = A^{-1}$.

Satz 11.2.4 *Sei $A \in \mathbb{R}^{m\times n}$ und $G \in \mathbb{R}^{n\times m}$. Die folgenden Aussagen sind äquivalent:*

1. *G ist verallgemeinerte Inverse zu A.*
2. *Für jede rechte Seite $b \in \mathbb{R}^m$ mit $L(A, b) \neq \emptyset$ (d.h. $\mathrm{rg}(A, b) = \mathrm{rg}A$) gilt $Gb \in L(A, b)$.*

Beweis:

1. Sei G verallgemeinerte Inverse zu A. Sei $b \in \mathbb{R}^m$ mit $L(A,b) \neq \emptyset$ und $y \in L(A,b)$. Dann:

   $A(Gb) = (AG)b = (AG)(Ay) = AGAy = Ay = b.$

2. Sei nun vorausgesetzt: Für jede rechte Seite $b \in \mathbb{R}^m$ mit $L(A,b) \neq \emptyset$ (d.h. $\text{rg}(A,b) = \text{rg}A$) gilt $Gb \in L(A,b)$.

   Sei $j \in \{1,\ldots,n\}$, und sei $a^j \in \mathbb{R}^m$ der j-te Spaltenvektor von A.

   Dann ist $L(A,a^j) \neq \emptyset$, und zwar ist $e_j \in L(A,a_j)$, dabei sei e_j der j-te kanonische Basisvektor des $\mathbb{R}^n$.

   Nach Voraussetzung folgt:

   $AGa^j = a^j$ für alle $j = 1,\ldots,n$.

   Daraus folgt:

   $AGA = A$.

Satz 11.2.5 *Seien $m,n,r \in \mathbb{N}$.*
Es sei $B \in \mathbb{R}^{m\times r}$ mit $\text{rg}B = r$ (B hat vollen Spaltenrang) und $T \in \mathbb{R}^{r\times n}$ mit $\text{rg}T = r$ (T hat vollen Zeilenrang).
Seien (vgl. 11.1.3) L eine Linksinverse von B und R eine Rechtsinverse von B.
Dann ist RL eine verallgemeinerte Inverse von BT.
Beweis: $(BT)(RL)(BT) = B(TR)(LB)T = BEET = BT$.

Satz 11.2.6 *Jede Matrix besitzt eine verallgemeinerte Inverse.*
Beweis: Sei $A \in \mathbb{R}^{m\times n}$.
Ist $\text{rg}A = 0$, d.h. jeder Matrixkoeffizient von A ist 0, so ist jede beliebige $n \times m$-Matrix verallgemeinerte Inverse zu A.
Sei nun $r := \text{rg}A > 0$. Nach dem vorigen Satz reicht es zu zeigen:

Lemma 11.2.7 $r := \text{rg}A > 0 \Longrightarrow$ *Es gibt Matrizen $B \in \mathbb{R}^{m\times r}$ mit $\text{rg}B = r$ und $T \in \mathbb{R}^{r\times n}$ mit $\text{rg}T = r$, so daß gilt: $A = BT$.*
Seien $a^1,\ldots,a^n \in \mathbb{R}^m$ die Spaltenvektoren von A.
Dann ist $r = \dim \text{im}A^\wedge = \dim\langle a^1,\ldots,a^n\rangle \subset \mathbb{R}^m$.
Sei $\mathcal{B} = \{b_1,\ldots,b_r\}$ eine Basis von $\text{im}A^\wedge = \langle a^1,\ldots,a^n\rangle$ und $B := (b_1,\ldots,b_r) \in \mathbb{R}^{m\times r}$, dann gilt $\text{rg}B = r$.
Betrachte $A^\wedge$ als lineare Abbildung von $\mathbb{R}^n$ in $\text{im}A^\wedge$; diese lineare Abbildung wird bzgl. der Basen $\mathcal{E} = \{e_1,\ldots,e_n\}$ (kanonische Basis des $\mathbb{R}^n$) und $\mathcal{B} = \{b_1,\ldots,b_r\}$ durch dieMatrix $T := M^{\mathcal{E}}_{\mathcal{B}}(A^\wedge)$ gegeben. Dann gilt $\text{rg}T = r$ und $A = BT$.

Bemerkung 11.2.8 *Man kann zeigen: Ist A nicht eine reguläre quadratische Matrix, so gibt es unendlich viele verallgemeinerte Inversen zu A.*

11.3 Moore-Penrose-Inverse

Definition 11.3.1 *Seien $m, n \in \mathbb{N}$. Sei $A \in \mathbb{R}^{m\times n}$.*
$G \in \mathbb{R}^{n\times m}$ **Moore-Penrose-Inverse** *zu $A \Longleftrightarrow$*

1. *$AGA = A$.*
2. *$GAG = G$.*
3. *AG und GA sind symmetrisch, d.h. ${}^t(AG) = AG$ und ${}^t(GA) = GA$*

(Moore 1920, Penrose 1955).

Bemerkung 11.3.2 1. *Eine Moore-Penrose-Inverse ist i.a. keine Inverse, sondern eine verallgemeinerte Inverse.*

2. *A quadratisch und regulär $\Longrightarrow$ Die inverse Matrix A^{-1} ist Moore-Penrose-Inverse.*

Lemma 11.3.3 *Jede Matrix kann höchstens eine Moore-Penrose-Inverse besitzen.*
Beweis: Sei $A \in \mathrm{IR}^{m\times n}$, und seien $G_1, G_2 \in \mathrm{IR}^{n\times m}$ Moore-Penrose-Inverse zu A. Dann:

1. $AG_2 =^t (AG_2) =^t (AG_1AG_2) =^t (AG_2)^t(AG_1) = (AG_2)(AG_1) = AG_2AG_1 = AG_1$
2. Analog: $G_2A = G_1A$.
3. $G_1 = B_1AG_1 = (G_1A)G_1 = (G_2A)G_1 = G_2(AG_1) = G_2(AG_2) = G_2AG_2 = G_2$

Satz 11.3.4 *Jede Matrix besitzt eine Moore-Penrose-Inverse.*

Der Beweis kann mit Hilfe der Eigenwerttheorie geführt werden.
Hier Skizze eines Beweises ohne Eigenwerttheorie. Zunächst einige Vorbereitungen:

Lemma 11.3.5 *Sei $A =\in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^n$.*
Das inhomogene lineare Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn für jeden n-Zeilenvektor ${}^t\lambda = (\lambda_1, \ldots, \lambda_n)$ gilt:

$$ {}^t\lambda A = 0 \Longrightarrow {}^t\lambda b = 0 $$

.
Beweisskizze: Die Multiplikation mit dem Zeilenvektor ${}^t\lambda$ bedeutet die Bildung einer Linearkombination. Die Behauptung folgt damit aus der Rangbedingung für die Lösbarkeit des inhomogenen Gleichungssystems.
Die Aussage ist leicht zu sehen, wenn die erweiterte Matrix (A, b) in Zeilenstufenform vorliegt. Die allgemeine Aussage ist aber nicht ohne weiteres auf diesen Spezialfall zurückzuführen; die vollständige Durchführung ist etwas aufwendiger.

Lemma 11.3.6 *Sei* $A = (a_{ij})_{i=1,\dots,m,j=1,\dots,n} \in \mathbb{R}^{m\times n}$. *Dann:*

$A = 0 \Longleftrightarrow \ {}^tAA = 0$

Beweis: $'' \Longrightarrow''$: klar.
$'' \Longleftarrow''$:
${}^tAA = 0 \Longrightarrow$
$0 = \mathrm{Sp}({}^tAA) = \sum_{i=1}^m \ \sum_{j=1}^n \ a_{ij}^2 \Longrightarrow$
$\forall i,j \ : \ a_{ij} = 0$

Korollar 11.3.7 *Sei* $A = \in \mathbb{R}^{m\times n}$ *und*

1. *Seien* $B, C = \in \mathbb{R}^{n\times p}$. *Dann:*

   $AB = AC \Longleftrightarrow \ {}^tAAB = \ {}^tAAC$

2. *Seien* $B, C = \in \mathbb{R}^{p\times n}$. *Dann:*

   $B\ {}^tA = C\ {}^tA \Longleftrightarrow B\ {}^tAA = C\ {}^tAA$

Die zweite Behauptung kann durch Transposition auf die erste zurückgeführt werden.
Beweis der ersten Behauptung:
$'' \Longrightarrow''$: klar.
$'' \Longleftarrow''$:
Es gelte ${}^tAAB =^t AAC$. Dann:
${}^t(AB - AC)(AB - AC) = ({}^tB -^t C)({}^tAAB -^t AAC) = 0 \Longrightarrow$
$AB - AC = 0$.

Lemma 11.3.8 *Sei* $A = (a_{ij})_{i=1,\dots,m,j=1,\dots,n} = \in \mathbb{R}^{m\times n}$ *eine Matrix,* $j \in \{1,\dots,m\}$ *und* ${}^ta_i = (a_{i1},\dots,a_{in})$ *der* i*-te Zeilenvektor von* A*. Dann ist das inhomogene lineare Gleichungssystem*

$$ {}^tAAx = a_i $$

lösbar.
Beweis: Zu zeigen ist: Für jeden n-Zeilenvektor ${}^t\lambda = (\lambda_1,\dots,\lambda_n)$ gilt:

$$ {}^t\lambda\ {}^tAA = 0 \Longrightarrow \ {}^t\lambda a_i = 0 $$

.
Beweis: ${}^t\lambda\ {}^tAA = 0 \Longrightarrow$ (Anwendung des letzten Lemmas mit $B = \ {}^t\lambda, C = 0$ (Zeilenvektor mit n Komponenten)):
${}^t\lambda\ {}^tA = 0 \Longrightarrow$ (da a_i der i-te Spaltenvektor von ${}^tA$ ist):
${}^t\lambda a_i = 0$.

Lemma 11.3.9 *Sei* $A = \in \mathbb{R}^{m\times n}$. *Dann gilt:*

$$ \mathrm{rg}({}^tAA) = \mathrm{rg}A $$

Beweis: Allgemein gilt: $\mathrm{rg}(AB) \leq \min(\mathrm{rg}A, \mathrm{rg}B)$ und $\mathrm{rg}{}^tA = \mathrm{rg}A$; daraus folgt: $\mathrm{rg}({}^tAA) \leq \mathrm{rg}A$.
Andererseits folgt aus dem letzten Lemma: $\dim \mathrm{im}({}^tAA)^\wedge \geq \dim\langle a_1, \ldots, a_n\rangle = \mathrm{rg}A$

Beweis des Satzes: Ist $A = 0$, so ist $G := 0$ eine Moore-Penrose-Inverse. Sei also $A \neq 0$.

1. Nach dem letzten Lemma des vorigen Abschnitts gilt $A = BT$. mit Matrizen $B \in \mathbb{R}^{m\times r}$ mit $\mathrm{rg}B = r$ und $T \in \mathbb{R}^{r\times n}$ mit $\mathrm{rg}T = r$.
2. ${}^tBB \in \mathbb{R}^{r\times r}$ ist eine quadratische $r \times r$-Matrix. Für ihren Rang gilt nach dem letzten Lemma:

   $\mathrm{rg}({}^tBB) = \mathrm{rg}B = r$.

   Damit ist ${}^tBB$ invertierbar.
3. Ebenso zeigt man: $T\ {}^tT$ ist eine invertierbare quadratische $r \times r$-Matrix.
4. Daher ist auch ${}^tBA\ {}^tT = {}^tBBT\ {}^tT$ eine invertierbare quadratische $r \times r$-Matrix, ihre Inverse ist $(T\ {}^tT)^{-1}({}^tBB)^{-1}$
5. Behauptung: $G :={}^t T({}^tBA\ {}^tT)^{-1}\ {}^tB = {}^tT(T\ {}^tT)^{-1}({}^tBB)^{-1}\ {}^tB$ ist Moore-Penrose-Inverse zu A.

   Beweis

   (a) $AGA = BT\ {}^tT(T\ {}^tT)^{-1}({}^tBB)^{-1}\ {}^tBBT = BT = A$
   (b) $GAG = {}^tT(T\,{}^tT)^{-1}({}^tBB)^{-1}\,{}^tBBT\,{}^tT(T\,{}^tT)^{-1}({}^tBB)^{-1}\,{}^tB = {}^tT(T\,{}^tT)^{-1}({}^tBB)^{-1}\,{}^tB = G$
   (c) $AG = BT\ {}^tT(T\ {}^tT)^{-1}({}^tBB)^{-1}\ {}^tB = B({}^tBB)^{-1}\ {}^tB$; diese Matrix ist, wie man leicht nachprüft, symmetrisch.
   (d) $GA = {}^tT(T\ {}^tT)^{-1}({}^tBB)^{-1}\ {}^tBBT = {}^tT(T\ {}^tT)^{-1}B$; dies ist ebenfalls eine symmetrische Matrix.

Inhaltsverzeichnis

1 Reelle lineare Gleichungssysteme, Matrizen und Vektoren 2
1.1 Lineare Gleichungssysteme 2
1.2 Matrix-Schreibweise 3
1.3 Spezielle Matrizen und die entsprechenden linearen Gleichungssysteme. 4
1.4 Lösungsmenge eines linearen Gleichungssystems 4

2 Der Gauß-Algorithmus 4
2.1 Stufenmatrizen 4
2.2 Elementare Zeilenumformungen 5
2.3 Struktur der Lösungsmengen homogener linearer Gleichungssysteme . 7
2.4 Struktur der Lösungsmengen inhomogener linearer Gleichungssysteme 7

3 Rechenoperationen für Matrizen und Vektoren 8
3.1 Addition von Matrizen 8
3.2 Die transponierte Matrix 8
3.3 Das Produkt von Matrizen mit Skalaren 8
3.4 Das Matrizenprodukt 9
3.5 Das Kronecker-Symbol 10
3.6 Elementarmatrizen 10
3.7 Die inverse Matrix 12

4 Reelle Vektorräume und lineare Abbildungen 14
4.1 Vektorräume 14
4.2 Lineare Abbildungen 15
4.3 Untervektorräume, Kern und Bild 15
4.4 Linearkombinationen, lineare Abhängigkeit und Unabhängigkeit . . . 16
4.5 Basis, Dimension 17
4.6 Isomorphismen, Dimensionsformel für lineare Abbildungen. 21
4.7 Der Rang einer Matrix 24
4.8 Die einer linearen Abbildung zugeordnete Matrix 25
4.9 Basiswechsel (Koordinatentransformationen) 27
4.10 Ähnliche und äquivalente Matrizen 30

5 Determinanten 32
5.1 Permutationen 32
5.2 Das Signum einer Permutation 33
5.3 Determinanten 35

6 Das kanonische euklidische Skalarprodukt, Längen und Winkel 46
6.1 Das kanonische euklidische Skalarprodukt 46
6.2 Längen- und Winkelmessung 48
6.3 Orthogonale Projektion 50
6.4 Euklidische Vektorräume 53
6.5 Orthonormalbasen 53

7 Endomorphismen euklidischer Vektorräume **58**
7.1 Orthogonale Endomorphismen 58

8 Eigenwert-Theorie und Diagonalisierung **62**
8.1 Aussagen über Polynome 62
8.2 Eigenwerte und Eigenvektoren 64
8.3 Das charakteristische Polynom, Diagonalisierung 66

9 Symmetrische reelle Matrizen **76**
9.1 Eigenwerte symmetrischer reeller Matrizen 76
9.2 Selbstadjungierte Endomorphismen euklidischer Vektorräume 77
9.3 Diagonalisierung symmetrischer reeller Matrizen 78
9.4 Hauptachsentransformation 79
9.5 Positive Definitheit und Eigenwerte 81
9.6 Das Hauptminorenkriterium für positive Definitheit 82

10 Normale Endomorphismen euklidischer Vektorräume **84**
10.1 Adjungierte lineare Abbildungen 84
10.2 Normale Matrizen und Endomorphismen 85

11 Verallgemeinerte Inverse von Matrizen **88**
11.1 Rechts- und Linksinverse 88
11.2 Verallgemeinerte Inverse 89
11.3 Moore-Penrose-Inverse 91