

25.10.00

1. Lineare Gleichungssysteme1.1.Bem.: Wir starten mit einem Beispiel :

$$\begin{aligned}
 &x+y+2z=9 \\
 &2x+4y-3z=1 \\
 (*) &3x+6y-5z=0 \\
 &-x+3y+z=8
 \end{aligned}$$

- Fragen: (I) Existieren Lösungen, oder ist das System unlösbar ?  
 (II) Falls (\*) lösbar ist, gibt es genau eine oder mehrere Lösungen ?  
 (III) Wie bestimmt man alle Lösungen ?  
 (IV) Was für eine Struktur bildet die Lösungsmenge ?  
 (V) Wie kann man sich die Lösungen geometrisch veranschaulichen ?

Einsetzen ergibt:  $x=1, y=2, z=3$  also Lösung für (\*)

Allgemein hat ein lineares Gleichungssystem von  $m$  Gleichungen in  $n$  Unbekannten( Variablen)  $x_1, \dots, x_n$  die Form

$$(*) \quad a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

Dabei sind die Zahlen  $a_{ij}$  ( $i=1; \dots; m, j=1; \dots; n$ ) und die Zahlen  $b_1; \dots; b_m$  vorgegeben. Man nennt die Zahlen  $a_{ij}$  die Koeffizienten und die Zahlen  $b_1; \dots; b_m$  die Bekannten von (\*). Eine Lösung von (\*) sind Zahlen  $x_1; \dots; x_n$  die die Gleichungen in (\*) simultan erfüllen. Jede solche Lösung faßt man zu einem  $n$ -Tupel  $(x_1; \dots; x_n)$  zusammen. Dabei bezeichnet man zwei  $n$ -Tupel  $(x_1; \dots; x_n)$  und  $(y_1; \dots; y_n)$  genau dann als verschieden, wenn  $x_i \neq y_i$  ist für ein  $i$  zwischen 1 und  $n$ . Z.B ist  $(1; 2; 3) \neq (1; 3; 2)$ .

Man nennt 2-Tupel auch Paare, 3-Tupel auch Tripel, 4-Tupel Quadrupel, usw. Als Lösungsmenge von  $x$  bezeichnet man die Menge

$$L = \{ (x_1; \dots; x_n) \mid (x_1; \dots; x_n) \text{ ist Lösung von } (*) \}.$$

1.2. Bem.: Wir betrachten ein Gleichungssystem

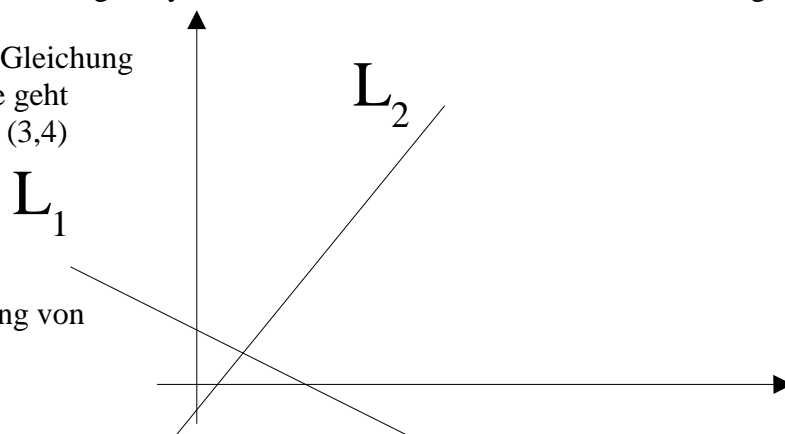
$$(*) \quad \begin{aligned} x+2y &= 1 \\ 3x-2y &= 1 \end{aligned}$$

Die Lösungen der ersten Gleichung  $x+2y=1$  bilden eine Gerade  $l_1$  in der Ebene, diese geht durch die Punkte  $(1,0)$  und  $(-1,1)$

Die Lösungen der zweiten Gleichung bilden eine Gerade  $l_2$ , diese geht durch die Punkte  $(1,1)$  und  $(3,4)$

Beide Geraden schneiden sich genau in einem Punkt, nämlich  $(\frac{1}{2}, \frac{1}{4})$ .

Dieser ist die einzige Lösung von (\*).



Allgemeiner: Eine einzelne Gleichung in zwei Unbekannten hat als Lösungsmenge i.d.R. Eine Gerade in der Ebene. Ausnahmen sind Gleichungen der folgenden Form  $0 \cdot x + 0 \cdot y = b$  die als Lösungsmenge entweder die ganze Ebene ( $b=0$ ) oder die leere Menge  $\emptyset$  (im Fall  $b \neq 0$ ) haben.

Ein lineares Gleichungssystem in zwei Unbekannten hat als Lösungsmenge den Durchschnitt der einzelnen Lösungsmengen der einzelnen Gleichungen. Bei Gleichungssystemen in mehr als zwei

Unbekannten hat man ein ähnliches Verhalten. So ist die Lösungsmenge einer einzelnen Gleichung in drei Variablen i.d.R. eine Ebene im Raum

1.3. Bem.: Gegeben sei ein lineares Gleichungssystem (\*)

*Beispiel*

Unter einer elementaren Umformung von (\*) versteht man folgendes:

Typ I: Die Addition eines Vielfachen einer Gleichung zu einer anderen

Typ II: Die Multiplikation einer Gleichung mit einer von 0 verschiedenen Zahl

Für das folgende ist wichtig, daß sich bei einer elementaren Umformung die Lösungsmenge nicht ändert; aus

(I)  $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$  und

(II)  $a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = b_j$  folgt für jede Zahl r

(II')  $(a_{j1} + ra_{i1})x_1 + (a_{j2} + ra_{i2})x_2 + \dots + (a_{jn} + ra_{in})x_n = b_j + rb_i$

Umgekehrt folgt aus den Gleichungen (I) und (II') die Gleichung (II), indem man das (-r)-fache von (I) zu (II') addiert.

Entsprechend kann man bei elementaren Umformungen vom Typ II argumentieren.

1.4. Bem.: Die Tatsache, daß sich bei einer elementaren Umformung die Lösungsmenge nicht ändert, verwendet man dazu, um ein Verfahren zur Lösung von linearen Gleichungssystemen zu entwickeln. Dabei geht man von einem vorgegeben linearen Gleichungssystem durch mehrfache elementare Umformung systematisch zu einem einfacheren linearen Gleichungssystem über, dessen Lösung man unmittelbar ablesen kann.

1.5. Bem.: Man kann bei der Behandlung von linearen Gleichungssystemen Schreibarbeit sparen, indem man sogenannte Matrizen verwendet. Eine  $m \times n$ -Matrix ist ein rechteckiges Schema von Zahlen der Form

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & \dots & b_{1l} \\ b_{21} & \dots & b_{2l} \\ \vdots & & \vdots \\ b_{k1} & \dots & b_{kl} \end{pmatrix}$$

Zwei Matrizen der Form A wie oben und B betrachtet man als genau dann gleich, wenn  $k=m$  und  $l=n$  und  $a_{ij}=b_{ij}$  für alle  $i,j$  ist.

2. Vorlesung

26.10.00

Jedem linearen Gleichungssystem (\*) ordnet man eine Matrix zu, die Koeffizientenmatrix wie unten( Mitte) und die erweiterte Matrix wie unten rechtsstehend.

$$(*) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{pmatrix} \quad \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

Die elementaren Umformungen von (\*) entsprechen elementaren Zeilenumformungen der erweiterten Matrix.

$$(*) \begin{cases} x+y+z=3 \\ x+2y-z=0 \\ 2x+5y-4z=-3 \end{cases} \quad \text{Erweiterte Matrix} \quad \begin{pmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & -1 & 0 \\ 2 & 5 & -4 & -3 \end{pmatrix} \begin{matrix} \bullet \\ \bullet \\ \bullet \end{matrix} \begin{matrix} \downarrow -1 \\ \downarrow -2 \\ \downarrow -3 \end{matrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & -2 & -3 \\ 0 & 3 & -6 & -9 \end{pmatrix} \begin{matrix} \bullet \\ \bullet \\ \bullet \end{matrix} \begin{matrix} \uparrow -1 \\ \downarrow -3 \\ \rightarrow \end{matrix} \begin{pmatrix} 1 & 0 & 3 & 6 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{matrix} \text{Entsprechendes Gleichungssystem} \\ x+ & & & 3z=6 \\ & y- & & 2z=-3 \end{matrix}$$

Lösungsmenge:  $L = \{(-3z+6; 2z-3; z) \mid z \in \mathbb{R}\}$ . L ist eine Gerade im Raum( Lösungsvektor).

1.6. Bem.: Oft ist es zweckmäßig zwei Zeilen einer Matrix zu vertauschen. Dies kann man durch mehrfache elementare Zeilenumformung erreichen.

$$\begin{pmatrix} \cdot \\ \cdot \\ a_{i1} \dots \\ \cdot \\ \cdot \\ a_{j1} \dots \end{pmatrix} \begin{matrix} \bullet \\ \downarrow +1 \\ \bullet \end{matrix} \begin{pmatrix} \cdot \\ \cdot \\ a_{i1} \dots \\ \cdot \\ \cdot \\ a_{j1} + a_{i1} \dots \end{pmatrix} \begin{matrix} \bullet \\ \uparrow -1 \\ \bullet \end{matrix} \begin{pmatrix} \cdot \\ \cdot \\ -a_{j1} \dots \\ \cdot \\ \cdot \\ a_{j1} + a_{i1} \dots \end{pmatrix} \begin{matrix} \bullet \\ \downarrow +1 \\ \bullet \end{matrix} \begin{pmatrix} \cdot \\ \cdot \\ -a_{j1} \dots \\ \cdot \\ \cdot \\ a_{i1} \dots \end{pmatrix} \mid \cdot (-1) \begin{pmatrix} \cdot \\ \cdot \\ a_{j1} \dots \\ \cdot \\ \cdot \\ a_{i1} \dots \end{pmatrix}$$

( Beweis für die Vertauschbarkeit)

1.7. Bem.: Wie im obigen Beispiel kann man eine beliebige Matrix durch mehrfache elementare Zeilenumformungen in eine Matrix der folgenden Form überführen:

$$Z = \begin{pmatrix} 0 & 0 & 1 & * & \dots & * & 0 & * & \dots & * \\ 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & * & \dots & * \end{pmatrix} \quad \text{Teilweise sind die Abstände etwas verzerrt, um die entstehende Matrix möglichst genau wiederzugeben.}$$

- Dabei steht \* für eine beliebige Zahl( nicht unbedingt immer die gleiche).  
 Man sagt dazu, daß Z reduzierte Zeilenstufenform hat. Diese wird durch folgende Eigenschaften charakterisiert:
1. Die Zeilen, die nur aus 0 bestehen, stehen ganz unten.
  2. Enthält eine Zeile nicht nur 0, so ist die erste von 0 verschiedene Zahl eine 1( führende 1).
  3. Eine Spalte, die eine führende 1 enthält, enthält sonst lauter 0.
  4. In zwei aufeinanderfolgenden Zeilen, die nicht nur 0 enthalten, steht die führende 1 in der oberen Zeile links von der führenden 1 in der unteren Zeile.

Man kann also jede beliebige Matrix A durch mehrfache elementare Zeilenumformungen in eine Matrix Z in reduzierter Zeilenstufenform überführen. Später werden wir das am Anfang etwas überraschende Ergebnis zeigen, daß Z durch A eindeutig bestimmt ist, also unabhängig davon, in welcher Weise man die elementaren Zeilenumformungen vornimmt.  
 An einer Matrix in reduzierter Zeilenstufenform kann man die Lösungsmenge des entsprechenden linearen Gleichungssystems unmittelbar ablesen.  
 Die Variablen, die zu den führenden Einsen( gebundene Variablen) gehören, sind eindeutig bestimmt, die frei wählbaren nennt man auch „freie Variablen“.

1.8. Bem.: Ein lineares Gleichungssystem der Form

$$(*) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0 \end{pmatrix}$$

nennt man homogen. Gegebenfalls gilt:

- (I) (\*) hat stets die triviale Lösung (0;...;0)
- (II) Sind  $s=(s_1; \dots; s_n)$  und  $t=(t_1; \dots; t_n)$  Lösungen von (\*), so ist auch die Summe  $s+t:=(s_1+t_1; \dots; s_n+t_n)$  eine Lösung von (\*).
- (III) Ist  $s=(s_1; \dots; s_n)$  eine Lösung von (\*) und  $r$  eine beliebige Zahl, so ist auch das Produkt  $r \cdot s=(r \cdot s_1; \dots; r \cdot s_n)$  Lösung von (\*).

Satz: Ein homogenes Gleichungssystem mit mehr Unbekannten als Gleichungen hat stets eine nichttriviale Lösung.

Bew.: Wir führen den Gauß-Algorithmus durch. Dabei ändert sich die Anzahl der Gleichungen und Unbekannten nicht. Das entstehende lineare Gleichungssystem in reduzierter Zeilenstufenform hat auch immer noch mehr Unbekannte als Gleichungen( und ist immer noch homogen). Daher muß es freie Variablen geben, und für diese kann man von Null verschiedene Zahlen einsetzen( nichttriviale Lösung).

1.9 Bem: Einem beliebigen linearen Gleichungssystem (I) ordnet man das entsprechende homogene lineare Gleichungssystem (H) zu.

$$(I) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{pmatrix}$$

$$(H) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0 \end{pmatrix}$$

Es kann natürlich vorkommen, daß das inhomogene lineare Gleichungssystem nicht lösbar ist. Hat aber (I) eine Lösung  $s=(s_1; \dots; s_n)$ , so gilt:

- (I) Für jede Lösung  $t=(t_1; \dots; t_n)$  von (H) ist  $u:=s+t$  eine Lösung von (I).
- (II) Umgekehrt existiert zu jeder Lösung  $u$  von (I) genau eine Lösung von (H) mit der Eigenschaft  $u:=s+t$ .
- (III) Diese Beziehung der Lösungsmengen  $L(I)$  und  $L(H)$  drückt man auch so aus:  $L(I)=s+L(H)$ .

2. Matrizen:

2.1 Def.: Die Addition von Matrizen gleichen Formats wird folgendermaßen definiert.

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & \dots & a_{2n}+b_{2n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

Satz: Für die Addition von Matrizen gelten die folgenden Rechenregeln:

- (I)  $A+B=B+A$  (Kommutativgesetz)
- (II)  $(A+B)+C=A+(B+C)$  (Assoziativgesetz)

Bew.: folgt unmittelbar aus den entsprechenden Rechenregeln für reelle Zahlen.

Bem.: Wegen der Assoziativität und der Kommutativität kann man auf die Klammern auch verzichten und kurz  $A+B+C$  schreiben.

2.2.Def.: Die  $m \times n$ -Matrix  $0_{m,n}$  bezeichnet man als Nullmatrix.

$0_{mn} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$  Sind keine Verwechslungen mit der reellen Zahl 0 zu befürchten, so schreibt man 0 statt  $0_{m,n}$ .

**Bem.:** Für eine beliebige  $m \times n$ -Matrix  $A$  gilt dann  $A+0=A=0+A$

**2.3.Def.:** Subtraktion zu einer Matrix.

$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$  Die negative Matrix wird definiert durch:  $-A = \begin{pmatrix} -a_{11} & \dots & -a_{1n} \\ -a_{21} & \dots & -a_{2n} \\ \vdots & & \vdots \\ -a_{m1} & \dots & -a_{mn} \end{pmatrix}$

**Bem.:** Offenbar gilt stets:  $A+(-A)=0$  und  $-(-A)=A$  und  $-(A+B)=(-A)+(-B)$ . Statt  $A+(-B)$  schreiben wir  $A-B$ . Es ist also

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} - \begin{pmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11}-b_{11} & \dots & a_{1n}-b_{1n} \\ a_{21}-b_{21} & \dots & a_{2n}-b_{2n} \\ \vdots & & \vdots \\ a_{m1}-b_{m1} & \dots & a_{mn}-b_{mn} \end{pmatrix}$$

Auf diese Weise erhält man eine Subtraktion für Matrizen.

**2.4.Def.:** Die Multiplikation einer Matrix mit der Zahl  $r$  (Skalar) wird definiert durch:

$$r \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} := \begin{pmatrix} r \cdot a_{11} & \dots & r \cdot a_{1n} \\ r \cdot a_{21} & \dots & r \cdot a_{2n} \\ \vdots & & \vdots \\ r \cdot a_{m1} & \dots & r \cdot a_{mn} \end{pmatrix}$$

**Satz:** Für die Multiplikation von Matrizen mit Zahlen gelten die folgenden Rechenregeln:

- (I)  $r(sA)=(rs)A$  (Assoziativgesetz)
- (II)  $r(A+B)=rA+rB$  (Distributivgesetz)
- (III)  $(r+s)A=rA+sA$  (Distributivgesetz)

**Bew.:** Der Beweis folgt unmittelbar aus den entsprechenden Rechenregeln für Zahlen.

**Bem.:** Wegen (I) kann man Klammern sparen und kurz  $rsA$  schreiben. In (II) und (III) haben wir stillschweigend vorausgesetzt, daß die Multiplikation stärker bindet als die Addition. Ausführlich hätten wir  $r(A+B)=(rA)+(rB)$  schreiben müssen. Offenbar gilt stets das folgende:

$1 \cdot A=A, 0 \cdot A=0_{m,n}, r \cdot 0_{m,n}=0_{m,n}, (-1) \cdot A=-A.$

Man kann Matrizen mit Zahlen nicht nur von links, sondern auch von rechts multiplizieren. Dabei gilt das Kommutativgesetz:  $rA=Ar$ .

**2.5.Def.:** Das Produkt einer  $m \times n$ -Matrix  $A=(a_{ij})$  und einer  $n \times p$ -Matrix  $B=(b_{ij})$  ist die folgendermaßen definierte  $m \times p$ -Matrix  $C=(c_{rs})$ :

$c_{rs} := a_{r1}b_{1s} + a_{r2}b_{2s} + \dots + a_{rn}b_{ns}$

**Bsp.:** (I)  $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{pmatrix}$   $B = \begin{pmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{pmatrix}$   $AB = \begin{pmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{pmatrix}$

Der Koeffizient in  $A \cdot B$  an der Position 2,3 ist gerade  $2 \cdot 4 + 6 \cdot 3 + 0 \cdot 5 = 26$ . Entsprechend wird bei den anderen Positionen verfahren.

$$(II) A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad A \cdot B = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Bem.: (I)  $A \cdot B$  ist nur definiert, wenn die Anzahl der Spalten von A die Anzahl der Zeilen von B ist (Im Beispiel (I) ist  $B \cdot A$  eben nicht definiert!). Die Anzahl der Zeilen von AB ist gleich der Anzahl der Zeilen von A, die Anzahl der Spalten von AB entspricht der von B.

(II) Warum definiert man das Produkt von Matrizen so und nicht anders?

Wir stellen uns vor, daß die Matrizen in Bsp.(II) mit folgenden Gleichungssystemen zu tun haben:

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 &= z_1 & b_{11}x_1 + b_{12}x_2 &= y_1 \\ a_{21}y_1 + a_{22}y_2 &= z_2 & b_{21}x_1 + b_{22}x_2 &= y_2 \end{aligned}$$

Einsetzen ergibt:  $z_1 = a_{11}(b_{11}x_1 + b_{12}x_2) + a_{12}(b_{21}x_1 + b_{22}x_2) = (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2$

Analog ist  $z_2 = (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2$

Man sieht, daß das Produkt AB genau dem Einsetzen der linearen Gleichungssysteme ineinander entspricht.

(III) Multipliziert man die  $m \times n$ -Matrix A mit der  $n \times 1$ -Matrix ergibt sich die  $m \times 1$ -Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad AX = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

Man kann also jedes lineare Gleichungssystem auch als Matrixgleichung( rechts) schreiben:

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{pmatrix} \Leftrightarrow \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Für das Produkt von Matrizen gilt das Kommutativgesetz i.A. nicht!

Satz: Für das Produkt von Matrizen gelten die folgenden Rechenregeln:

(I)  $(A \cdot B) \cdot C = A \cdot (B \cdot C)$  (Assoziativgesetz)

(II)  $A(B+C) = AB+AC$  (Distributivgesetz)

(III)  $(A+B)C = AC+BC$  ( „ „ )

(IV)  $r(AB) = (rA)B = A(rB)$

(V) Dabei setzen wir voraus, daß die Formate von A,B und C so beschaffen sind, daß die entsprechenden Produkte Sinn machen( definiert sind), und r ist eine beliebige reelle Zahl.

Bew.: (I) Sei  $A=(a_{ij})$  eine  $m \times n$ -Matrix,  $B=(b_{ij})$  eine  $n \times p$ -Matrix und  $C=(c_{ij})$  eine  $p \times q$ -Matrix. Dann ist AB eine  $m \times p$ -Matrix,  $(AB)C$  eine  $m \times q$ -Matrix, BC eine  $n \times q$ -Matrix und  $A(BC)$  eine  $m \times q$ -Matrix:

Das Element an der Position (i,j) in AB ist  $a_{i1}b_{1j} + \dots + a_{in}b_{nj}$ .

Das Element an der Position (r,s) in  $(AB)C$  ist also  $(a_{r1}b_{11} + \dots + a_{rn}b_{n1})c_{1s} + \dots + (a_{r1}b_{1p} + \dots + a_{rn}b_{np})c_{ps}$ .

Das Element an der Position (k,l) in BC ist  $b_{k1}c_{1l} + \dots + b_{kp}c_{pl}$ .

Das Element an der Position (r,s) in  $A(BC)$  ist also  $a_{r1}(b_{11}c_{1s} + \dots + b_{1p}c_{ps}) + \dots + a_{rn}(b_{n1}c_{1s} + \dots + b_{np}c_{ps})$ .

Daher stimmen die Elemente an der Position (r,s) in  $(AB)C$  und  $A(BC)$  überein, also ist  $(AB)C = A(BC)$ .

(II) Sei  $A=(a_{ij})$  eine  $m \times n$ -Matrix, und seien  $B=(b_{ij})$  und  $C=(c_{ij})$  zwei  $n \times p$ -Matrizen. Dann ist  $B+C$  eine  $n \times p$ -Matrix und  $A(B+C)$ ,  $AB$ ,  $AC$ ,  $AB+AC$  sind  $m \times p$ -Matrizen.

Das Element an der Position  $(i,j)$  von  $AB$  bzw.  $AC$  ist  $a_{i1}b_{1j}+\dots+a_{in}b_{nj}$  bzw.  $a_{i1}c_{1j}+\dots+a_{in}c_{nj}$ .

Das Element an der Position  $(i,j)$  in  $AB+AC$  ist also die Summe aus den beiden Elementen von  $AB$  und  $AC$  in der Position  $(i,j)$ :  $a_{i1}b_{1j}+\dots+a_{in}b_{nj}+a_{i1}c_{1j}+\dots+a_{in}c_{nj}=a_{i1}(b_{1j}+c_{1j})+\dots+a_{in}(b_{nj}+c_{nj})$ .

Das ist genau das Element an der Position  $(i,j)$  von  $A(B+C)$ . Folglich gilt  $AB+AC=A(B+C)$ .

(III) Analog zu (II)

(IV) Seien  $A$  und  $B$  wie in (II). Das Element an der Position  $(i,j)$  in  $r(AB)$  bzw.  $(rA)B$  bzw.  $A(rB)$  ist dann:  $r(a_{i1}b_{1j}+\dots+a_{in}b_{nj})$  bzw.  $(ra_{i1})b_{1j}+\dots+(ra_{in})b_{nj}$  bzw.  $a_{i1}(rb_{1j})+\dots+a_{in}(rb_{nj})$ . Alle drei Terme sind aber gleich, daraus folgt die Richtigkeit der Behauptung.

2.6Bem.: Da für das Produkt von Matrizen das Assoziativgesetz gilt, können wir wieder Klammern sparen und kurz  $ABC$  bzw.  $rAB$  schreiben.

Def.: Die  $n \times n$ -Matrix

$$1_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & \dots & 0 & & 1 \end{pmatrix}$$

bezeichnet man als Einheitsmatrix des Grades  $n$ .

Satz: Für jede  $n \times n$ -Matrix  $A$  gilt:  $1_n A = A = A 1_n$ .

Bew.: Folgt unmittelbar aus der Definition des Produkts von Matrizen.

2.7.Bem.: Eine Matrix nennt man quadratisch, falls ihre Zeilenanzahl mit ihrer Spaltenanzahl übereinstimmt. Für jede quadratische Matrix bezeichnet man das  $n$ -Tupel  $(a_{11}; a_{22}; \dots; a_{nn})$  als Hauptdiagonale von  $A$ .

2.11.00

Wenn  $A=(a_{ij})$   $m \times n$ -Matrix;  $B=(b_{kl})$   $n \times p$ -Matrix, dann ist  $AB=(c_{rs})$   $m \times p$ -Matrix:

$$c_{rs} = a_{r1}b_{1s} + \dots + a_{rn}b_{ns}$$

Def.: Eine  $m \times n$ -Matrix nennt man:

(I) linksinvertierbar, falls eine zu  $A$  linksinverse  $n \times m$ -Matrix  $B$  existiert mit  $BA=1_n$ .

(II) rechtsinvertierbar, falls eine zu  $A$  rechtsinverse  $n \times m$ -Matrix  $C$  existiert mit  $AC=1_m$ .

(III) invertierbar, falls  $A$  linksinvertierbar und rechtsinvertierbar ist.

Satz: Zu jeder invertierbaren  $m \times n$ -Matrix  $A$  existiert genau eine  $m \times n$ -Matrix  $D$ , die die beiden Eigenschaften mit  $DA=1_n$  und  $AD=1_m$  erfüllt. Man nennt  $D$  die zu  $A$  inverse Matrix und schreibt:  $D=A^{-1}$ .

Bew.: Sei  $A$  eine rechtsinvertierbare  $m \times n$ -Matrix und seien  $B$  und  $C$  wie in (I) und (II):

$B=B \cdot 1_m = B \cdot A \cdot C = 1_n \cdot C = C$ ; also  $BA=1_n$  und  $AB=1_m$ . Ist  $E$  eine beliebige  $m \times n$ -Matrix mit  $EA=1_n$  und  $AE=1_m$ , so zeigt die gleiche Rechnung:  $B=B \cdot 1_m = B \cdot A \cdot E = 1_n \cdot E = E$ .

Bsp.: (I)  $A$  ist invertierbar, denn für die Matrix  $B$  gilt:

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} \quad AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = BA$$

Es ist logisch, daß jede invertierbare Matrix quadratisch ist.

(II)  $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$  ist linksinvertierbar, denn für  $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$  gilt  $BA=1_2$ . Daher ist  $B$  linksinvers zu  $A$ .

Für  $B_1 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$  gilt auch  $BA=1_2$ . Linksinverse Matrizen sind i.A. nicht eindeutig. Aus dem obigen Satz folgt, daß A nicht invertierbar sein kann, denn sonst wäre  $B=B_1_3=BAA^{-1}=1_2A^{-1}=A^{-1}$  und analog  $B_1=A^{-1}$ , also  $B=B_1$ , was nicht der Fall ist. Daß A nicht invertierbar ist, kann man auch direkt sehen, denn für jede  $2 \times 3$ -Matrix  $E=(e_{ij})$  gilt:

$$AE = \begin{pmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \underline{1} \end{pmatrix}$$

**2.8.Bem.:** (I) Für jede invertierbare  $m \times n$ -Matrix A ist  $AA^{-1}=1_m$  und  $A^{-1}A=1_n$ . Daher ist auch  $A^{-1}$  invertierbar mit der inversen Matrix A:  $(A^{-1})^{-1}=A$ .

(II) Für eine beliebige invertierbare  $m \times n$ -Matrix A und eine beliebige invertierbare  $n \times p$ -Matrix B gilt:  $AB \cdot B^{-1} \cdot A^{-1} = A \cdot 1_n \cdot A^{-1} = A \cdot A^{-1} = 1_m$  und  $B^{-1} \cdot A^{-1} \cdot A \cdot B = B^{-1} \cdot 1_m \cdot B = B^{-1} \cdot B = 1_p$ .

Daher ist auch AB invers mit inverser Matrix  $B^{-1}A^{-1}$ :  $(AB)^{-1} = (B^{-1}) \cdot (A^{-1})$  (Reihenfolge beachten!)

(III) Wegen  $1_n \cdot 1_n = 1_n$  ist die Einheitsmatrix stets invertierbar und  $1_n^{-1} = 1_n$ .

(IV) Wie kann man feststellen, ob eine vorg  $m \times n$ -Matrix A rechtsinvertierbar ist, und wie kann man gegebenenfalls eine zu A rechtsinverse Matrix berechnen? Dazu macht man den Ansatz

$X = \begin{pmatrix} x_{11} & \dots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nm} \end{pmatrix}$  mit den Unbekannten  $x_{ij}$  und sieht sich die Matrixgleichung  $AX=1_m$  genauer an. Ein Vergleich der ersten Spalte liefert das lineare Gleichungssystem(I) mit der Koeffizientenmatrix A. Darauf können wir den Gauß-Algorithmus anwenden (Gibt es keine Lösung, so ist die Matrix A nicht rechtsinvertierbar).

Ein Vergleich der zweiten Spalte von AX und  $1_m$  liefert(II):

$$(I) \begin{cases} a_{11}x_{11} + \dots + a_{1n}x_{n1} = 1 \\ a_{21}x_{11} + \dots + a_{2n}x_{n1} = 0 \\ \vdots \\ a_{m1}x_{11} + \dots + a_{mn}x_{n1} = 0 \end{cases} \quad (II) \begin{cases} a_{11}x_{11} + \dots + a_{1n}x_{n1} = 0 \\ a_{21}x_{11} + \dots + a_{2n}x_{n1} = 1 \\ \vdots \\ a_{m1}x_{11} + \dots + a_{mn}x_{n1} = 0 \end{cases}$$

Entsprechend kann man für die übrigen Spalten vorgehen (Gibt es dann irgendein unlösbares Gleichungssystem, so ist die Matrix nicht rechtsinvertierbar).

Wir sehen, daß die Berechnung einer Matrix X mit  $AX=1_m$  auf die Lösung von linearen Gleichungssystemen mit Koeffizientenmatrix A und rechten Seiten hinausläuft. Es ist ökonomischer, diese m linearen Gleichungssysteme simultan zu lösen. Dazu faßt man die m rechten Seiten wieder zur Einheitsmatrix zusammen. Man schreibt die Matrix A und  $1_m$  nebeneinander und wendet auf A und  $1_m$  die gleichen elementaren Zeilenumformungen an.

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

(V) Ähnlich kann man vorgehen, um festzustellen, ob eine vorgegebene Matrix A linksinvertierbar ist. Um festzustellen, ob eine Matrix A invertierbar ist, kann man wie folgt vorgehen:

Wie oben prüft man, ob eine  $m \times n$ -Matrix X mit  $AX=1_m$  existiert (Wenn es keine gibt, ist A nicht mal rechtsinvertierbar). Existieren verschiedene Matrizen X,Y mit  $AX=1_m=AY$ , so ist A nicht invertierbar (Denn sonst wäre  $X=1_n X=A^{-1}AX=A^{-1}1_m=A^{-1}$ , aber analog  $Y=A^{-1}$ ).

Es bleibt der Fall, daß genau eine  $m \times n$ -Matrix X existiert mit  $AX=1_m$ . Man braucht also nur noch nachzuprüfen, ob auch  $XA=1_n$ . Falls nicht, so ist A nicht invertierbar, falls ja, so ist A invertierbar mit  $X=A^{-1}$ .

(VI) Wir werden bald sehen, daß für jede rechtsinvertierbare  $m \times n$ -Matrix  $m \leq n$  gilt.

Analog gilt für jede linksinvertierbare  $m \times n$ -Matrix  $n \leq m$ .

Daraus folgt: jede invertierbare Matrix ist quadratisch; da  $n \leq m$  und  $m \leq n \Rightarrow m=n$ .

Ferner werden wir sehen, daß eine quadratische Matrix genau dann links( rechts)invertierbar ist, wenn sie auch rechts( links) invertierbar ist. Das vereinfacht den Nachweis der Invertierbarkeit.

**2.9.Def.:** Für jede  $m \times n$ -Matrix A und  $r \in \mathbb{N}$  definiert man die r-te Potenz von A durch  $A^r := AA \dots A$  (r Faktoren). Zusätzlich definiert man  $A^0 = 1_n$  ( $0 \notin \mathbb{N}$  in der Vorlesung).

**Bem.:** Dann ist jeweils  $A^r A^s = A^{r+s}$  und  $(A^r)^s = A^{r \cdot s}$ .

Dagegen ist i.A.  $(AB)^r \neq A^r B^r$  !

Sind aber A und B vertauschbar, das heißt  $AB=BA$ , so gilt auch dies.

**2.10 Def.:** Für eine  $m \times n$ -Matrix A definiert man die zu A transponierte Matrix  $A^T$  (die Indizes werden vertauscht).

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

**Satz: (I)** Für jede Matrix A ist  $(A^T)^T = A$

**(II)** Für beliebige  $m \times n$ -Matrizen A,B ist  $(A+B)^T = A^T + B^T$ .

**(III)** Für jede Zahl r und jede Matrix A ist  $(rA)^T = rA^T$ .

**(IV)** Für jede  $m \times n$ -Matrix A und jede  $n \times p$ -Matrix B ist  $(AB)^T = B^T A^T$ .

**(V)** Für jede rechtsinvertierbare Matrix A( mit rechtsinverser Matrix B) ist  $A^T$  linksinvertierbar (mit linksinverser Matrix  $B^T$ ).

**(VI)** Für jede invertierbare Matrix A ist auch  $A^T$  invertierbar mit  $(A^T)^{-1} = (A^{-1})^T$ .

**Bew.:** (I-III) sind offensichtlich.

8.11.00

**(IV)** Da  $AB$  eine  $m \times p$ -Matrix ist, ist  $(AB)^T$  eine  $p \times m$ -Matrix. Da  $A^T$  eine  $m \times n$ -Matrix ist und  $B^T$  eine  $p \times n$ -Matrix ist, ist  $B^T A^T$  eine  $p \times m$ -Matrix. Das Element an der Position (i,j) von  $(AB)^T$  stimmt mit dem Element an der Position (j,i) in  $AB$  überein, und dieses ist:

$$a_{j1}b_{1i} + \dots + a_{jn}b_{ni} = b_{1j}a_{1i} + \dots + b_{nj}a_{in}. \text{ Letzteres ist genau das Element an der Position (i,j) in } B^T A^T.$$

**(V)** Aus  $AB = 1_m$  ergibt sich mit **(IV)**:  $1_m = 1_m^T = (AB)^T = B^T A^T$ .

**(VI)** folgt aus **(V)** und der Aussage, die man aus **(V)** erhält, indem man rechts und links tauscht.

**2.11. Def.:** Eine quadratische Matrix A mit  $A^T = -A$  ( bzw.  $A^T = A$ ) nennt man anti-(symmetrisch).

Bsp.:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} \text{ Symmetrisch} \quad \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{pmatrix} \text{ Antisymmetrisch}$$

**Bem.:** Aus Satz 2.10 folgt

**(I)** Für Zahlen r,s und symmetrische( antisymmetrische)  $m \times n$ -Matrizen A,B ist auch  $rA+sB$  symmetrisch( antisymmetrisch). Ist A invertierbar, so ist auch  $A^{-1}$  symmetrisch( antisymmetrisch). I.A. Ist aber  $AB$  nicht symmetrisch( antisymmetrisch).

Dagegen ist stets  $AB+BA$  symmetrisch(  $AB-BA$  antisymmetrisch).

**(II)** Für eine beliebige quadratische Matrix A ist  $\frac{1}{2}(A+A^T)$  und  $\frac{1}{2}(A-A^T)$  antisymmetrisch. Man kann also jede beliebige quadratische Matrix als Summe einer symmetrischen und einer antisymmetrischen Matrix schreiben.

### 3. Der Rang

**3.1. Def.:** Seien  $m,n \in \mathbb{N}$ . Für  $i=1; \dots; m$  und  $j=1; \dots; n$  sei  $E_{ij}$  die  $m \times n$ -Matrix, die 1 an der Position (i,j) enthält und 0 sonst. Man nennt  $E_{ij}$  eine Matrixeinheit.

**Bem.:** Für jede  $m \times n$ -Matrix  $A=(a_{ij})$  gilt offenbar:  $A=(a_{11}E_{11} + \dots + a_{1n}E_{1n}) + \dots + (a_{m1}E_{m1} + \dots + a_{mn}E_{mn})$ .

**(II)** Es gelten die Regeln 
$$E_{ij} E_{kl} = \begin{cases} E_{il} & j=k \\ 0 & \text{sonst} \end{cases}$$

**3.2. Def.:** Eine  $m \times n$ -Matrix der Form  $U_{ij}(r) = 1_n + rE_{ij}$  ( $r \in \mathbb{R}, i \neq j$ )

nennt man elementar vom Typ I. Eine  $m \times m$ -Matrix der Form  $(0 \neq r \in \mathbb{R})$  nennt man elementar vom Typ II.

$$D_i(r) = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & r & \\ 0 & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

Bsp.: elementar vom Typ I    elementar vom Typ II

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

Bem.: (I) Für alle  $r, s \in \mathbb{R}$  ist  $U_{ij}(r)U_{ij}(s) = (1_m + rE_{ij})(1_m + sE_{ij}) = 1_m + rE_{ij} + sE_{ij} + rsE_{ij}E_{ij} = 1_m + (r+s)E_{ij} = U_{ij}(r+s)$  ( $E_{ij}E_{ij} = 0$ ), das heißt  $U_{ij}(r)U_{ij}(s) = U_{ij}(r+s)$ , insbesondere ist  $U_{ij}(r)U_{ij}(-r) = 1_m$ . Analog ist  $U_{ij}(-r)U_{ij}(r) = 1_m$ . Daher ist  $U_{ij}(r)$  invertierbar und es gilt:  $U_{ij}(r)^{-1} = U_{ij}(-r)$ .

(II) Für  $r, s \in \mathbb{R}$  mit  $r \neq 0 \neq s$  ist analog  $D_i(r)D_i(s) = D_i(rs)$ . Insbesondere ist  $D_i(r)D_i(r^{-1}) = 1_m$ . Daher ist  $D_i(r)$  invertierbar und  $D_i(r)^{-1} = D_i(r^{-1})$ .

(III) Sei  $A$  eine beliebige  $m \times n$ -Matrix. Dann ist  $U_{ij}(r)A$  die Matrix, die aus  $A$  durch Addition des  $r$ -fachen der  $j$ -ten Zeile zur  $i$ -ten Zeile entsteht. Analog ist  $D_i(r)A$  die Matrix, die aus  $A$  durch Multiplikation der  $i$ -ten Zeile mit  $r$  entsteht. Eine elementare Zeilenumformung von  $A$  entspricht also der Multiplikation von  $A$  mit einer elementaren Matrix von links (Daher die Bezeichnung elementare Matrix).

Geht also eine  $m \times n$ -Matrix  $C$  durch  $k$  elementare Zeilenumformungen aus der Matrix  $A$  hervor, so existieren elementare Matrizen  $E_1, \dots, E_k$  mit  $C = E_1 \dots E_k A$ . Daher ist die Matrix  $F := E_1 \dots E_k$  invertierbar mit  $C = FA$ .

Satz: Für eine  $m \times n$ -Matrix  $A = (a_{ij})$  sind die folgenden Aussagen äquivalent:

(1)  $A$  ist linksinvertierbar.

(2) Für jede Wahl von  $y_1, \dots, y_m$  hat das lineare Gleichungssystem  $\left. \begin{matrix} a_{11}x_1 + \dots + a_{1n}x_n = y_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = y_m \end{matrix} \right\} (*)$  höchstens eine Lösung.

(3) Für irgendeine Wahl von  $y_1, \dots, y_m$  hat das lineare Gleichungssystem (\*) genau eine Lösung.

(4) Das homogene lineare Gleichungssystem mit der Koeffizientenmatrix  $A$  hat nur die triviale Lösung.

(5) Der Gauß-Algorithmus macht aus  $A$  eine Matrix der folgenden Form:  $\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ & & \ddots & \\ 0 & 0 & & 1 \\ 0 & \dots & \dots & 0 \\ \vdots & & & \\ 0 & \dots & \dots & 0 \end{pmatrix}$

Bew.: Wir zeigen  $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (1)$

$(1) \Rightarrow (2)$  Sei  $A$  linksinvertierbar. Wir wählen eine feste zu  $A$  linksinverse Matrix  $B$ . Sei  $(x_1, \dots, x_n)$  eine Lösung von (\*). Dann ist

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}, \text{ d.h. } \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 1_m \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = BA \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = B \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

Folglich ist  $(x_1, \dots, x_n)$  eindeutig bestimmt.

$(2) \Rightarrow (3)$  Setze  $y_1 = y_2 = \dots = 0$ .

$(3) \Rightarrow (4)$  folgt aus Bemerkung 1.9

$(4) \Rightarrow (5)$  Sei (4) erfüllt. Wir wenden auf die Matrix  $A$  den Gauß-Algorithmus an und müssen sehen, daß die am Ende entstehende Matrix  $C$  die gewünschte Form hat. Das homogene lineare Gleichungssystem mit Koeffizientenmatrix  $C$  hat ebenfalls nur die triviale Lösung. Man kann also keine der Unbekannten  $x_1, \dots, x_n$  frei wählen. Daher müssen alle Unbekannten zu führenden Einsen gehören. Da  $C$  eine  $m \times n$ -Matrix ist, muß sie die angegebene Form haben.

$(5) \Rightarrow (1)$  Sei  $C$  die angegebene Matrix. Nach Bemerkung 3.2(II) ist  $C = FA$  für eine invertierbare Matrix  $F$ . Setzt man  $D = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & 0 & \dots & \dots & 0 \end{pmatrix}$  dann ist  $DC = 1_n$ . Daher ist  $DF A = DC = 1_n$ . Folglich ist  $A$  linksinvertierbar mit linksinverser Matrix  $DF$ .

3.3. Bsp.: Ist die folgende Matrix linksinvertierbar?  $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \\ 3 & 5 & 6 \end{pmatrix}$  Wir machen den folgenden Ansatz:

$$X = \begin{pmatrix} x_{11} & \dots & x_{14} \\ \vdots & & \vdots \\ x_{31} & \dots & x_{34} \end{pmatrix}$$

und betrachten die Gleichung  $XA=1_3$ . Oder einfacher die dazu äquivalente Gleichung  $A^T X^T = 1_3$ .

$A^T = \begin{pmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 1 & 2 & 3 & 6 \end{pmatrix} \xrightarrow{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $= \begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} \xrightarrow{-1} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$ $= \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{-1} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$ $= \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$	<p>Wir erhalten zum Beispiel die Lösung <math>X^T = \begin{pmatrix} 2 &amp; -1 &amp; 0 \\ -1 &amp; 2 &amp; -1 \\ 0 &amp; -1 &amp; 1 \end{pmatrix}</math>, das heißt <math>X = \begin{pmatrix} 2 &amp; -1 &amp; 0 &amp; 0 \\ -1 &amp; 2 &amp; -1 &amp; 0 \\ 0 &amp; -1 &amp; 1 &amp; 0 \end{pmatrix}</math> (Probe!)</p> <p>Wir betrachten noch das lineare Gleichungssystem</p> $A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 4 \end{pmatrix}$ <p>Ist <math>(x_1; x_2; x_3)</math> eine Lösung, so gilt:</p> $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 1_3 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = XA \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = X \begin{pmatrix} 1 \\ 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$
---	---

Man rechnet nach, daß tatsächlich  $x_1=1, x_2=-1, x_3=1$  eine Lösung ist. Das trifft zu.

9.11.00

**Bem.:** Aus (5) folgt, daß für jede linksinvertierbare  $m \times n$ -Matrix gilt:  $n \leq m$ . Durch Transponieren folgt, daß für jede rechtsinvertierbare  $m \times n$ -Matrix gilt  $n \geq m$ . Daher ist jede invertierbare Matrix quadratisch.

**Satz:** Für eine  $m \times n$ -Matrix  $A=(a_{ij})$  sind die folgenden Aussagen äquivalent:

- (1) A ist rechtsinvertierbar.
- (2) Der Gauß-Algorithmus macht aus A eine Matrix ohne Nullzeilen.
- (3) Für jede Wahl von  $y_1; \dots; y_m$  hat das lineare Gleichungssystem (\*) mindestens eine Lösung:

$$(*) \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = y_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = y_m \end{cases}$$

**Bew.:**

(1) $\Rightarrow$ (2) Sei A rechtsinvertierbar und B eine zu A rechtsinverse Matrix. Der Gauß-Algorithmus macht aus A eine Matrix C. Nach Bemerkung 3.2(III) existiert eine invertierbare  $m \times m$ -Matrix F mit  $C=FA$ . Folglich ist doch  $CBF^{-1}=FABF^{-1}=F1_mF^{-1}=FF^{-1}=1_m$ . Daher kann C keine Nullzeilen enthalten( wegen der Definition der Multiplikation von Matrizen und der Einheitsmatrix ).

(2) $\Rightarrow$ (3) Wir wenden auf unsere Matrix A den Gauß-Algorithmus an und bezeichnen die am Ende entstehende Matrix ohne Nullzeilen mit  $C=(c_{ij})$ . Sind  $y_1; \dots; y_m$  beliebig vorgegeben, so ist das lineare Gleichungssystem (\*) äquivalent zu einem linearen Gleichungssystem der Form:

$$(**) \begin{cases} c_{11}x_1 + \dots + c_{1n}x_n = y_1 \\ \vdots \\ c_{m1}x_1 + \dots + c_{mn}x_n = y_m \end{cases}$$

Da C reduzierte Zeilenstufenform hat und keine Nullzeilen enthält, ist (\*\*) lösbar. Daher ist auch (\*) lösbar.

(3) $\Rightarrow$ (1) Sei (3) erfüllt. Wir setzen zunächst  $y_1:=1, y_2:=\dots=y_n:=0$ . Dann existieren  $x_{11}; \dots; x_{n1}$  mit( s. S.12):

$$\begin{aligned} a_{11}x_{11} + \dots + a_{1n}x_{n1} &= 1 \\ a_{11}x_{11} + \dots + a_{2n}x_{n1} &= 0 \\ &\vdots \\ a_{m1}x_{11} + \dots + a_{mn}x_{n1} &= 0 \end{aligned}$$

Als nächstes setzen wir  $y_1:=0, y_2=1, y_3=\dots=y_n=0$ . Dann existieren analog Zahlen mit  $x_{12}; \dots; x_{n2}$  mit

Auf diese Weise fahren wir fort und erhalten so Zahlen  $x_{ij}(i=1; \dots; n ;j=1; \dots; m)$  mit

$$A \cdot \begin{pmatrix} x_{12} \\ \vdots \\ x_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad A \begin{pmatrix} x_{11} \dots x_{1m} \\ \vdots \\ x_{n1} \dots x_{nm} \end{pmatrix} = 1_m$$

Damit ist A rechtsinvertierbar.

**3.4 Satz:** Für eine  $n \times n$ -Matrix  $A=(a_{ij})$  sind folgende Aussagen äquivalent:

- (1) A ist invertierbar.
- (2) A ist linksinvertierbar.
- (3) A ist rechtsinvertierbar.
- (4) Für jede Wahl von Zahlen  $y_1; \dots; y_n$  hat das folgende lineare Gleichungssystem genau eine Lösung.

$$(*) \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = y_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = y_n \end{cases}$$

- (5) Der Gauß-Algorithmus macht aus A die Einheitsmatrix.
- (6) A läßt sich als Produkt elementarer Matrizen schreiben.

**Bew.:** (1) $\Leftrightarrow$ (4) $\Leftrightarrow$ (5) folgt aus 3.2 und 3.3.

(1) $\Rightarrow$ (2) und (1) $\Rightarrow$ (3) sind trivial.

(2) $\Rightarrow$ (5) Sei A linksinvertierbar. Nach Satz 3.2 macht der Gauß-Algorithmus aus der Matrix A eine Matrix der Form  $C = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ 0 & 0 & 1 & \\ \vdots & & & \ddots \\ 0 & \dots & \dots & 0 \end{pmatrix}$ . Da C außerdem quadratisch ist, ist  $C=1_n$ .

(5) $\Rightarrow$ (6) Sei (5) erfüllt. Nach Bemerkung 3.2(III) existieren elementare Matrizen  $E_1; \dots; E_k$  mit  $1_n = E_1 \dots E_k A$ . Folglich ist die Matrix  $A = E_k^{-1} \dots E_1^{-1}$  mit elementaren Matrizen  $E_1^{-1} \dots E_k^{-1}$ .

(6) $\Rightarrow$ (1) Seien  $E_1; \dots; E_k$  elementare Matrizen mit  $A = E_1 \dots E_k$ . Mit  $E_1 \dots E_k$  ist also auch A invertierbar.

(3) $\Rightarrow$ (1) Sei A rechtsinvertierbar, dann ist  $A^T$  linksinvertierbar. Nach dem, was wir bereits bewiesen haben, ist  $A^T$  invertierbar. Folglich ist auch A invertierbar.

**3.5. Def.:** Man nennt zwei  $m \times n$ -Matrizen zeilenäquivalent und schreibt  $A \sim_z B$ , falls man A durch mehrere elementare Zeilenumformungen in B überführen kann.

**Bem.:** Nach Satz 3.4 und Bemerkung 3.2(III) gilt genau dann  $A \sim_z B$ , wenn eine invertierbare Matrix F mit  $FA=B$  existiert. Ferner zeigt der Gauß-Algorithmus, daß jede Matrix einer Matrix in reduzierter Zeilenstufenform zeilenäquivalent ist.

**Satz:** Für  $m \times n$ -Matrizen A,B,C gilt stets:

- (I)  $A \sim_z A$  (Reflexivität)
- (II) Aus  $A \sim_z B$  folgt  $B \sim_z A$  (Symmetrie)
- (III) Aus  $A \sim_z B$  und  $B \sim_z C$  folgt  $A \sim_z C$  (Transitivität)

**Bew.:** (I) Stets ist  $1_m A = A$ .

(II) Sei  $A \sim_z B$ , dann existiert eine invertierbare Matrix F mit  $FA=B$ . Daher ist  $F^{-1}$  invertierbar mit  $A = F^{-1}B$ , d.h.  $B \sim_z A$ .

(III) Sei  $A \sim_z B$  und  $B \sim_z C$ . Dann existieren zwei invertierbare Matrizen F und G mit  $FA=B$  und

GB=C. Folglich ist das Produkt GF invertierbar und GFA=GB=C, d.h.  $A \sim_z C$ .

**3.6. Satz:** Gegeben seien  $m \times n$ -Matrizen  $Y=(y_{ij})$  und  $Z=(z_{ij})$  in reduzierter Zeilenstufenform mit  $Y \sim_z Z$ . Dann ist  $Y=Z$ .

**Bem.:** (I) Aus dem Satz folgt insbesondere, daß die Matrix, die aus einer vorgegebenen Matrix A durch Anwenden des Gauß-Algorithmus entsteht, durch A eindeutig bestimmt ist, also nicht von den Wahlmöglichkeiten innerhalb des Gauß-Algorithmus abhängt. [Wir wenden zwei Varianten des Gauß-Algorithmus an und erhalten am Ende Y und Z in reduzierter Zeilenstufenform. Dann ist  $A \sim_z Y$  und  $A \sim_z Z$ , also auch  $Y \sim_z Z$ . Aus 3.6 folgt also  $Y=Z$ ].

(II) Will man entscheiden, ob vorgegebene  $m \times n$ -Matrizen A und B zeilenäquivalent sind, kann man folgendermaßen vorgehen. Wir wenden auf A und B den Gauß-Algorithmus an und bezeichnen die am Ende entstehende Matrix mit Y bzw. Z (Dann ist  $A \sim_z Y$  und  $B \sim_z Z$ ). Im Fall  $Y=Z$  ist  $A \sim_z Y=Z \sim_z B$ , also auch  $A \sim_z B$ .

Im Fall  $Y \neq Z$  ist A nicht zeilenäquivalent zu B (Denn sonst wäre  $A \sim_z Y$  und  $A \sim_z B \sim_z Z$ , also auch  $Y \sim_z Z$ ; Widerspruch zu 3.6).

Vorgehensweise: Wir wollen feststellen, ob zwei Matrizen zeilenäquivalent sind und gegebenenfalls eine invertierbare Matrix F mit  $FA=B$  konstruieren. Dazu wenden wir zunächst auf A den Gauß-Algorithmus an und nehmen dabei gleichzeitig dieselben elementaren Zeilenumformungen an der Einheitsmatrix vor.

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad M_1 = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \quad S_1 = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Im Schritt i des Verfahrens gilt dabei jeweils  $M_i A = N_i$  ( $N_i[M_i]$  ist [Einheits-]Matrix von A,  $R_i[S_i]$  entsprechend für B). Dabei gilt jeweils  $R_i B = S_i$ . Insgesamt ist also  $A \sim_z N_x = S_y \sim_z B$ . D.h.  $A \sim_z B$ .

Ferner ist  $R_y B = S_y = N_x = M_x A$ . Wir berechnen als nächstes  $R_y^{-1}$ . Dann ist  $B = R_y^{-1} M_x A =: FA$ .

**3.7. Def.:** Gegeben sei eine  $m \times n$ -Matrix A. Die durch A eindeutig bestimmte Matrix in reduzierter Zeilenstufenform, die zu A zeilenäquivalent ist bezeichnen wir mit Z. Die Anzahl r der von Null verschiedenen Zeilen der Matrix Z nennt man den Zeilenrang von A. Man schreibt:  $r = zr(A)$ .

Hier ist z.B  $M_1 A = N_1 = A$ . Analog für B.

15.11.00

**Bew.(von 3.6.):** Wegen  $Y \sim_z Z$  existiert eine invertierbare  $m \times m$ -Matrix  $U=(u_{ij})$  mit  $Z=UY$ . Wir schreiben:

$$Y = \begin{pmatrix} 0 \dots 0 y_{1s_1} & y_{1n} \\ 0 \dots 0 0 0 \dots 0 y_{2s_2} & y_{2n} \\ \vdots & \vdots \\ 0 \dots 0 0 0 \dots 0 0 \dots 0 y_{rsr} & y_{rn} \\ 0 \dots & 0 \\ \vdots & \vdots \\ 0 \dots & 0 \end{pmatrix} \quad \left. \begin{array}{l} (s_i = \text{Nummer der Spalte in} \\ \text{der die führende 1 stehen} \\ \text{soll}) \\ \} \text{ Weitere Zeilen sind} \\ \} \text{ Nullzeilen} \end{array} \right\}$$

mit  $s_1 < s_2 < \dots < s_r$  und  $y_{1s_1} = \dots = y_{rsr} = 1$  und  $y_{isj} = 0$  für  $i \neq j$ . Mit X bezeichnen wir die Matrix, die aus den ersten r Zeilen von Y besteht, d.h.  $Y = \begin{pmatrix} X \\ 0 \end{pmatrix}$ . Wir werden zeigen, daß U die Form  $U = \begin{pmatrix} 1_r & * \\ 0 & * \end{pmatrix}$  hat.

[Dann ist  $Z = UY = \begin{pmatrix} 1_r & * \\ 0 & * \end{pmatrix} \begin{pmatrix} X \\ 0 \end{pmatrix} = \begin{pmatrix} X \\ 0 \end{pmatrix} = \underline{Y}$ ]

Beim Beweis argumentieren wir mit Induktion nach r!

Im Fall  $r=1$  hat  $Y$  die Form  $Y = \begin{pmatrix} 0 \dots 0 & 1 & * \dots * \\ 0 \dots & & 0 \\ \vdots & & \vdots \\ 0 \dots & & 0 \end{pmatrix}$ . Daher hat  $Z=UY$  die Form  $Z = \begin{pmatrix} 0 \dots 0 & u_{11} & * \dots * \\ 0 \dots 0 & u_{21} & * \dots * \\ \vdots & \vdots & \vdots \\ 0 \dots 0 & u_{m1} & * \dots * \end{pmatrix}$ .

Wäre  $u_{11}=0$ , so wäre auch  $u_{21}=\dots=u_{m1}=0$ , da  $Z$  reduzierte Zeilenstufenform hat.

Dann wäre die erste Spalte von  $U$  gleich 0, also auch die erste Spalte  $U^{-1} \cdot U = 1_m$ , was nicht so ist. Also ist  $u_{11} \neq 0$ . Da  $Z$  reduzierte Zeilenstufenform hat, folgt  $u_{11}=1$  und  $u_{21}=\dots=u_{m1}=0$ . Damit hat  $U$  die gewünschte Form.

Wir haben also den Induktionsanfang erledigt. Sei jetzt  $r > 1$  und die Behauptung für  $r-1$  bewiesen. Wir bezeichnen mit  $Y_{\text{quer}}$  und  $Z_{\text{quer}}$  die Matrizen, die aus den ersten  $s_r-1$  Spalten von  $Y$  bzw.  $Z$  bestehen. Damit haben  $Y_{\text{quer}}$  und  $Z_{\text{quer}}$  ebenfalls reduzierte Zeilenstufenform, und es ist  $Z_{\text{quer}} = U Y_{\text{quer}}$ . Da  $Y_{\text{quer}}$  nur  $r-1$  von 0 verschiedene Zeilen hat, folgt aus der Induktionsvoraussetzung, daß  $U$  die Form  $U = \begin{pmatrix} 1 & & * \\ & \dots & \\ 0 & & * \end{pmatrix}$  hat. Daher ist  $Z_{\text{quer}} = U Y_{\text{quer}} = Y_{\text{quer}}$ , d.h. die ersten  $s_r-1$  Spalten von  $Y$  und  $Z$  stimmen überein.

Wegen  $Z=UY$  und  $y_{rsr}=1$  hat die  $s_r$ -te Spalte von  $Z$  die Form  $\begin{pmatrix} z_{1sr} \\ \vdots \\ z_{msr} \end{pmatrix} = \begin{pmatrix} u_{1r} \\ \vdots \\ u_{mr} \end{pmatrix}$ . Wir nehmen zunächst  $z_{rsr}=0=u_{rr}$  an. Da  $Z$  reduzierte Zeilenstufenform hat, ist dann auch  $u_{rr}=u_{r+1,r}=\dots=u_{mr}=0$ . Damit hat  $U$  die folgende Form:

$$U = \begin{pmatrix} 1 & * & * \\ 0 & 0 & * \\ 0 & 0 & * \end{pmatrix} \leftarrow r.$$

Der Gauß-Algorithmus kann also aus  $U$  nicht die Einheitsmatrix machen. Da  $U$  aber invertierbar ist, widerspricht dies Satz 3.4 (5). Also ist  $z_{rsr}=u_{rr} \neq 0$ . Da  $Z$  reduzierte Zeilenstufenform hat, folgt  $z_{rsr}=u_{rr}=1$  und die  $s_r$ -te Spalte von  $Z$  hat die Form  $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow r$ . Damit hat auch die  $r$ -te Spalte von  $U$  diese Form und ganz  $U$  hat die gewünschte Gestalt.

Bsp.3.7.: (Zeilenrang)

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \xrightarrow{\text{red.}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \text{ Also ist } \text{zr}(A)=2 \text{ ( von 0 verschiedene Zeilen).}$$

Zeilenstufenform

Bem.: Offenbar ist  $\text{zr}(A)$  die Anzahl der gebundenen Variablen im homogenen linearen Gleichungssystem (\*) mit Koeffizientenmatrix  $A$ .  $n-\text{zr}(A)$  ist die Anzahl der freien Variablen in (\*).

3.8. Bem.: (I) Analog zu den elementaren Zeilenumformungen einer Matrix, der reduzierten Zeilenstufenform, der Zeilenäquivalenz  $\sim_z$  und dem Zeilenrang  $\text{zr}(A)$  definiert man elementare Spaltenumformungen, die reduzierte Spaltenstufenform  $\sim_s$  und den Spaltenrang  $\text{sr}(A)$ . Offenbar gehen die entsprechenden Begriffe durch Transponieren ineinander über. Daher gelten analoge Sätze. Diese werden wir im folgenden benutzen:

Z.B.: entsprechen elementare Spaltenumformungen einer Matrix  $A$  der Multiplikation mit elementaren Matrizen von rechts.

(II) Gegeben seien  $2 m \times n$ -Matrizen  $A, B$ . Man sagt, daß  $A$  und  $B$  äquivalent sind, falls man  $A$  durch mehrfache elementare Zeilen- und Spaltenumformungen in  $B$  überführen kann. Dies ist offenbar genau dann der Fall, wenn zwei invertierbare Matrizen  $U, V$  mit  $B=UAV$  existieren. Man schreibt  $A \sim B$ . Man zeigt leicht, daß für Äquivalenz  $\sim$  ein zu Satz 3.5 analoges Ergebnis gilt.

Satz: Für eine natürliche Zahl  $r$  und eine  $m \times n$ -Matrix  $A$  sind gleichwertig:

- (1)  $r = \text{zr}(A)$ .
- (2)  $A$  ist zu einer Matrix der Form  $\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$  äquivalent.
- (3)  $r = \text{sr}(A)$ .

Bew.: (1) $\Rightarrow$ (2) Sei  $r = \text{zr}(A)$ . Der Gauß-Algorithmus macht aus  $A$  eine zu  $A$  zeilenäquivalente Matrix  $Z$  in reduzierter Zeilenstufenform. Durch mehrfache elementare Spaltenumformung( genauer

Spaltenvertauschung) wird aus Z eine Matrix der Form  $\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$ . Durch weitere elementare Spaltenumformungen wird aus T eine Matrix der gewünschten Form.

(2)⇒(1) Sei (2) erfüllt. Dann existieren invertierbare Matrizen U, V mit  $UAV = \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix} =: N$ . In N und  $NV^{-1} = UA$  verschwinden die letzten m-r Zeilen. Der Gauß-

Algorithmus macht also aus UA eine Matrix Z in reduzierter Zeilenstufenform mit mindestens m-r Nullzeilen. Wegen  $Z \sim_z UA \sim_z A$  folgt  $zr(A) \leq r$ . Andererseits existiert eine invertierbare Matrix R mit  $Z = RUA = RNV^{-1}$ . Daher haben Z und  $ZV = RN$  mindestens m-zr(A) Nullzeilen. Der Gauß-Algorithmus macht aus RN eine Matrix Y in reduzierter Zeilenstufenform mit mindestens m-zr(A) Nullzeilen. Da N selbst reduzierte Zeilenstufenform hat und zu RN zeilenäquivalent ist, folgt  $N = Y$ . Insbesondere ist  $m-r \geq m-zr(A)$ , d.h.  $r \leq zr(A)$ . Insgesamt ist also  $r = zr(A)$ .

(2)⇔(3) analog.

**3.9. Bem.:** Nach Satz 3.8 ist  $zr(A) = sr(A)$  für jede Matrix A. Man spricht daher auch kurz von dem Rang von A und schreibt  $rg(A) := zr(A) := sr(A)$ .

Satz: Zwei  $m \times n$ -Matrizen sind genau dann äquivalent, wenn sie den gleichen Rang haben.

Bew.: Sei  $r := rg(A)$  und  $B \sim A$ . Nach Satz 3.8. ist  $A \sim \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$ , also auch  $B \sim \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$ . Aus Satz 3.8 folgt also  $rg(B) = r$ . Sei umgekehrt  $rg(A) = rg(B) =: r$ . Nach Satz 3.8 ist dann:

$$A \sim \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix} \sim B, \text{ also } A \sim B.$$

Nach Satz 3.8 existieren invertierbare Matrizen U und V mit der Eigenschaft, daß  $UAV$  die Gestalt  $UAV = \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$  hat. Wie findet man U und V?

Für eine beliebige  $m \times n$ -Matrix A schreibt man die drei Matrizen nebeneinander. Auf A wendet man elementare Zeilen- und Spaltenumformungen an, bis man A auf die Form  $\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$  gebracht hat. Dabei wendet man die Zeilenumformungen auch auf die Matrix  $1_m$  und jede Spaltenumformung auf  $1_n$  an.

$U = 1_m$ -Matrix,  $V = 1_n$ -Matrix. Jeweils ist  $U_i A V_i = A_i$ ,  $i = i$ -ter Schritt des Verfahrens.

Insbesondere ist dann  $U_5 A V_5 = A_5 = \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$ .

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ \vdots & \vdots & \vdots \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \end{pmatrix}$$

$$U_5(U_i) \quad A_5(A_i) \quad V_5(V_i)$$

**3.10 (Rangkriterium für lineare Gleichungssysteme)**

Ein lineares Gleichungssystem (\*) ist genau dann lösbar, wenn der Rang der Koeffizientenmatrix mit dem Rang der erweiterten Matrix übereinstimmt.

Bew.: Da sich bei elementaren Zeilenumformungen weder an der Lösbarkeit von (\*) noch an den Rängen der beteiligten Matrizen etwas ändert, können wir annehmen, daß die erweiterte Matrix reduzierte Zeilenstufenform hat( und damit auch die Koeffizientenmatrix):

$$\begin{aligned} a_{1s_1}x_{s_1} + \dots + a_{1n}x_n &= b_1 \\ \dots & \dots \\ a_{rs_r}x_{s_r} + \dots + a_{rn}x_n &= b_r \\ & 0 = b_{r+1} \\ & \vdots \\ & 0 = b_m. \end{aligned}$$

Dieses lineare Gleichungssystem ist genau dann lösbar, wenn  $b_{r+1} = \dots = b_m = 0$  ist. Das bedeutet aber

gerade, daß die Koeffizientenmatrix und die erweiterte Matrix den gleichen Rang haben.

**4. Determinanten**

4.1. Bem.: Gegeben sei ein lineares Gleichungssystem mit quadratischer Koeffizientenmatrix:

$$(*) \begin{cases} a_{11}x_{11} + \dots + a_{1n}x_{n1} = b_1 \\ \vdots \\ a_{n1}x_{11} + \dots + a_{nn}x_{n1} = b_n \end{cases}$$

Wir betrachten den Fall, daß (\*) genau eine Lösung hat. Gibt es eine Formel, die diese Lösung  $x_1; \dots; x_n$  durch die Koeffizienten  $a_{ij}$  und die Bekannten  $b_k$  ausdrückt? Ja, sie wird durch „Determinanten“ geliefert.

Analog: Gibt es eine Formel, die die Koeffizienten der inversen Matrix  $A^{-1}$  aus denen der invertierbaren Matrix  $A$  berechnet? Ja, sie enthält „Determinanten“.

Def.: Die Determinante  $\det(A) := |A| := \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$  einer  $n \times n$ -Matrix  $A = (a_{ij})$  wird folgendermaßen definiert:

Ist  $n=1$ , also  $A=(a)$  setzt man  $\det(A) := a$ . Ist  $n > 1$  und die Determinante einer  $(n-1) \times (n-1)$ -Matrix bereits definiert, so setzt man  $|A| = a_{11}|A_{11}| - a_{21}|A_{21}| + \dots + (-1)^{n+1}a_{n1}|A_{n1}|$ ; dabei ist  $A_{ij}$  jeweils die Matrix, die aus der Matrix  $A$  durch Streichen der  $i$ -ten Zeile und der  $j$ -ten Spalte entsteht:

$$A_{ij} = \begin{pmatrix} a_{11} \dots a_{1,j-1} & a_{1j} & a_{1,j+1} \dots a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i-1,1} & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & a_{i-1,n} \\ \hline a_{i1} & a_{ij} & a_{i+1,j} & a_{i+1,j+1} & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n,j-1} & a_{nj} & a_{n,j+1} & a_{nn} \end{pmatrix} \quad \text{Bsp.: } n=2 \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb = ad - bc$$

$$n=3 \quad \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix} = aei - afh - dbi + dch + gbfc - gec.$$

Allgemein ist  $|A|$  eine Summe von  $n!$  Termen. Schon für kleine  $n$  ist die Definition für die Berechnung von Determinanten unbrauchbar( da Rechenaufwand faktoriell wächst; NP-hard).

Wir werden daher die allgemeinen Eigenschaften von Determinanten herleiten, die die Berechnung erleichtern. Wir beginnen mit Spezialfällen.

Eine Matrix  $A$  der Form wie rechts nennt man obere Dreiecksmatrix. Für

jede solche Matrix  $A$  ist  $|A| = a_{11} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ 0 & \dots & 0 & a_{nn} \end{vmatrix} = a_{11} a_{22} \begin{vmatrix} a_{33} & \dots & a_{3n} \\ \vdots & & \vdots \\ 0 & \dots & 0 & a_{nn} \end{vmatrix} = a_{11} \dots a_{nn}$ .  $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$

Insbesondere ist  $\det 0_{n,n} = 0$  und  $\det 1_n = 1$ .

Eine Matrix  $A$  der Form wie rechts nennt man untere Dreiecksmatrix. Für

jedes solche  $A$  ist  $|A| = a_{11} \begin{vmatrix} a_{22} & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} - a_{21} \begin{vmatrix} 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} + \dots + (-1)^{n+1} a_{n1} \begin{vmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ a_{n-1,2} & \dots & a_{n-1,n-1} & 0 \end{vmatrix}$   $A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$

Wir wollen uns überlegen, daß  $|A|$  auch in diesem Fall das Produkt der Elemente auf der Hauptdiagonalen ist. Dazu verwenden wir Induktion nach der Größe  $n$  der Matrix. Im Fall  $n=1$  ist das sicher richtig. Für  $n=2$  wäre das  $ac - b \cdot 0 = ac$ . Sei jetzt  $n > 2$  und die Behauptung für  $(n-1) \times (n-1)$ -Matrizen bereits bewiesen. Dann ergibt die obige Gleichung  $|A| = a_{11}(a_{22} \dots a_{nn}) - a_{21}(0 \cdot a_{33} \dots a_{nn}) + \dots + (-1)^{n+1} a_{n1}(0 \dots 0) = a_{11}(a_{22} \dots a_{nn}) = a_{11} \dots a_{nn}$ . Damit ist die Behauptung für untere Dreiecksmatrizen gezeigt.

Als Spezialfälle halten wir folgendes fest:

- (1) Die Determinante einer elementaren Matrix  $U_{ij}(a)$  vom Typ I ist stets gleich 1.
- (2) Die Determinante einer elementaren Matrix  $D_i(a)$  vom Typ II ist stets gleich  $a$ .

**Satz:** (I) Für  $k=1;\dots;n$  und  $r,s \in \mathbb{R}$  gilt:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,n} \\ rb_1+sc_1 & \dots & rb_n+sc_n \\ a_{k+1,1} & \dots & a_{k+1,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = r \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,n} \\ b_1 & \dots & b_n \\ a_{k+1,1} & \dots & a_{k+1,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + s \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ c_1 & \dots & c_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix} \quad (, * = \text{wie links})$$

(II) Für  $k=1;\dots;n-1$  gilt:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,n} \\ b_1 & \dots & b_n \\ b_1 & \dots & b_n \\ a_{k+2,1} & \dots & a_{k+2,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = 0$$

**Bew.:** Wir argumentieren jeweils mit Induktion nach  $n$ .

(I) Im Fall  $n=1$  ist die Sache einfach  $\det(rb+sc)=rb+sc=r \cdot \det(b)+s \cdot \det(c)$ . Sei also  $n>1$  und die Behauptung für  $(n-1) \times (n-1)$ -Matrizen bewiesen. Wir bezeichnen die drei Matrizen mit  $A=(a_{ij})$ ,  $B=(b_{ij})$  und  $C=(c_{ij})$ .

Zu zeigen ist  $|A|=r|B|+s|C|$ . Nach Definition ist  $|A|=a_{11}|A_{11}|-a_{21}|A_{21}|+\dots+(-1)^{n+1}a_{n1}|A_{n1}|$ . Für  $i \neq k$  ist  $A_{i1}$  eine  $(n-1) \times (n-1)$ -Matrix der gleichen Bauart, also  $|A_{i1}|=r|B_{i1}|+s|C_{i1}|$  nach Induktionsvoraussetzung. Außerdem ist  $a_{i1}=b_{i1}=c_{i1}$ . Für den entsprechenden Summanden gilt also:  $(-1)^{n+1}a_{i1}|A_{i1}|=(-1)^{n+1}b_{i1}r|B_{i1}|+(-1)^{n+1}c_{i1}s|C_{i1}|$ .

Dagegen ist  $A_{k1}=B_{k1}=C_{k1}$  und  $a_{k1}=rb_{k1}+sc_{k1}$ . Der  $k$ -te Summand hat also die Form  $(-1)^{k+1}a_{k1}|A_{k1}|=(-1)^{k+1}b_{k1}r|B_{k1}|+(-1)^{k+1}c_{k1}s|C_{k1}|$ . Insgesamt ergibt sich:  $|A|=rb_{11}|B_{11}|-rb_{21}|B_{21}|+\dots+(1)^{n+1}rb_{n1}|B_{n1}|+sc_{11}|C_{11}|-sc_{21}|C_{21}|+\dots+(-1)^{n+1}sc_{n1}|C_{n1}|=r|B|+s|C|$ .

22.11.00

**Bew.:** (II) Zu zeigen ist, daß die Determinante einer  $n \times n$ -Matrix mit zwei gleichen aufeinanderfolgenden Zeilen Null ist.

Im Fall  $n=2$  ist  $\begin{vmatrix} a & b \\ a & b \end{vmatrix} = ab-ab=0$ . Sei also  $n>2$  und die Behauptung für  $(n-1) \times (n-1)$ -Matrizen bewiesen.

Die Matrix aus dem Satz sei  $A=(a_{ij})$ . Die Definition ergibt:  $|A|=a_{11}|A_{11}|-a_{21}|A_{21}|+\dots+(-1)^{n+1}a_{n1}|A_{n1}|$ . Für  $k \neq i \neq k+1$  enthält auch  $A_{i1}$  zwei gleiche Zeilen, nach Induktionsvoraussetzung ist also  $|A_{i1}|=0$ . Außerdem ist  $A_{k1}=A_{k+1,1}$  und  $a_{k1}=b_1=a_{k+1,1}$ . Folglich ist  $|A|=(-1)^{k+1}a_{k1}|A_{k1}|+(-1)^{k+2}a_{k+1,1}|A_{k+1,1}|=0$ .

**4.2 Satz:** (I) Vertauscht man in einer  $n \times n$ -Matrix zwei Zeilen, so wird ihre Determinante mit  $(-1)$  multipliziert.

(II) Die Determinante einer Matrix mit zwei gleichen Zeilen verschwindet.

(III) Die Determinante einer  $n \times n$ -Matrix mit einer Nullzeile verschwindet.

**Bew.:**

(I)

$$c_1 \dots c_n \quad 0 = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,n} \\ b_1+c_1 & \dots & b_n+c_n \\ b_1+c_1 & \dots & b_n+c_n \\ a_{k+2,1} & \dots & a_{k+2,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ b_1 & \dots & b_n \\ b_1+c_1 & \dots & b_n+c_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix} + \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ c_1 & \dots & c_n \\ b_1+c_1 & \dots & b_n+c_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix} = \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ b_1 & \dots & b_n \\ b_1+c_1 & \dots & b_n+c_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix} + \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ b_1 & \dots & b_n \\ c_1 & \dots & c_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix} + \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ b_1 & \dots & b_n \\ b_1 & \dots & b_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix} + \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ c_1 & \dots & c_n \\ b_1 & \dots & b_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix} + \begin{vmatrix} * & & * \\ \vdots & & \vdots \\ * & & * \\ b_1 & \dots & b_n \\ c_1 & \dots & c_n \\ * & & * \\ \vdots & & \vdots \\ * & & * \end{vmatrix}$$

$= 0 \qquad \qquad \qquad = 0$

(, \* = wie in der ersten Determinante)

Damit ist die Behauptung für den Fall bewiesen, daß die zu vertauschenden Zeilen benachbart

sind. Den allgemeinen Fall führen wir darauf zurück. Wir wollen also Zeile  $i$  und Zeile  $j$  vertauschen; dabei können wir  $i < j$  annehmen.

Zunächst können wir durch  $(j-i)$ -fache Vertauschung von jeweils zwei benachbarten Zeilen die Konfiguration  $i \Rightarrow i+1$  und  $j \Rightarrow i$  erreichen. Dann führen wir noch  $j-i-1$  Vertauschungen von jeweils zwei benachbarten Zeilen durch, um die endgültige Konfiguration  $j_{\text{ende}} = i_{\text{anfang}}$  und  $i_{\text{ende}} = j_{\text{anfang}}$  zu erhalten. Die Determinante ändert sich also insgesamt um den Faktor  $(-1)^{j-1}(-1)^{i-1} = (-1)^{2j-2i-1} = -1$ .

(II) Nach (I) kann man annehmen, daß die beiden ersten Zeilen von  $A$  übereinstimmen. Die Behauptung folgt dann aus Satz 4.1(II).

(III) folgt aus Satz 4.1(I).

**4.3 Satz:** (I) Bei einer elementaren Umformung vom Typ I ändert sich die Determinante garnicht.

(II) Multipliziert man eine Zeile einer Matrix mit einer Zahl  $r$ , so wird ihre Determinante auch mit  $r$  multipliziert.

(III) Multipliziert man eine  $n \times n$ -Matrix mit einer Zahl  $r$ , so wird ihre Determinante mit  $r^n$  multipliziert.

Bew.: (I) Für  $i \neq j$  gilt:

$a_{11} \dots a_{1n}$	$a_{11} \dots a_{1n}$	$a_{11} \dots a_{1n}$	[Die Matrix hat zwei gleiche Zeilen, die $i$ -te Zeile ist gleich der $j$ -ten].
$\vdots$	$\vdots$	$\vdots$	
$a_{i-1,1} \dots a_{i-1,n}$	$a_{i-1,1} \dots a_{i-1,n}$	$a_{i-1,1} \dots a_{i-1,n}$	
$a_{i1} + ra_{j1} \dots a_{in} + ra_{jn}$	$a_{i1} \dots a_{in}$	$+r a_{j1} \dots a_{jn}$	
$\vdots$	$\vdots$	$\vdots$	
$a_{i+1,1} \dots a_{i+1,n}$	$a_{i+1,1} \dots a_{i+1,n}$	$a_{i+1,1} \dots a_{i+1,n}$	
$\vdots$	$\vdots$	$\vdots$	
$a_{n1} \dots a_{nn}$	$a_{n1} \dots a_{nn}$	$a_{n1} \dots a_{nn}$	

(II) folgt unmittelbar aus Satz 4.1

(III) folgt aus (II).

Bem.: Es ergibt sich also die gute Methode, um Determinanten mit dem Gauß-Algorithmus zu berechnen:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 3 & 1 \\ 1 & 2 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 \quad !$$

**4.4 Satz:** Eine  $n \times n$ -Matrix ist genau dann invertierbar, wenn  $\det(A) \neq 0$  ist.

Bew.: Ob  $A$  invertierbar ist, oder nicht, wird von elementaren Zeilenumformungen von  $A$  nach 3.4 nicht beeinflusst. Ob  $\det(A) \neq 0$  ist, wird nach Satz 4.3 auch nicht beeinflusst. Daher kann man annehmen, daß  $A$  reduzierte Zeilenstufenform hat.

Ist  $A$  invertierbar, so ist  $A = I_n$  nach 3.4. Folglich ist  $|A| = 1 \neq 0$ .

Ist  $A$  nicht invertierbar, so ist  $A$  auch nicht rechtsinvertierbar nach 3.4. Nach Satz 3.3 enthält  $A$  Nullzeilen, also ist  $|A| = 0$  nach Satz 4.2(III).

Bem.: Nach Satz 3.4 ist  $A$  auch genau dann invertierbar, wenn  $\text{rg}(A) = n$  ist.

**4.5 Satz:** (Produktregel für Determinanten; *wichtig!*)

Für beliebige  $n \times n$ -Matrizen  $A, B$  ist  $|AB| = |A| \cdot |B|$ .

Bew.: Ist  $A$  nicht invertierbar, so ist auch  $AB$  nicht invertierbar. Nach Satz 4.4 sind in diesem Fall beide Seiten der Gleichung Null. Daher kann man annehmen, daß  $A$  invertierbar ist. Nach Satz 3.4 ist dann  $A = E_1 \dots E_k$  mit elementaren Matrizen  $E_1, \dots, E_k$ . Für jede elementare  $n \times n$ -Matrix  $E$  und jede beliebige  $n \times n$ -Matrix  $G$  gilt nach Satz 4.3  $|EG| = |E| \cdot |G|$ .

Daher ist  $|AB| = |E_1 \dots E_k B| = |E_1| \dots |E_k| |B| = |E_1| \dots |E_{k-2}| |E_{k-1} E_k| |B| = \dots = |E_1| \dots |E_k| |B| = |A| |B|$ .

Bem.: Ist  $A$  invertierbar, so ist  $AA^{-1} = I_n$ . Folglich ist  $1 = |I_n| = |AA^{-1}| = |A| |A^{-1}|$ , also  $|A^{-1}| = |A|^{-1}$ .

**4.6 Satz:** Für jede  $n \times n$ -Matrix  $A$  ist  $|A^T| = |A|$ .

Bew.: Ist  $A$  nicht invertierbar, so ist auch  $A^T$  nicht invertierbar. Daher sind in diesem Fall beide Determinanten Null. Sei also  $A$  invertierbar. Dann ist  $A = E_1 \dots E_k$  mit elementaren Matrizen  $E_1 \dots E_k$ .

also  $A^T = E_k^T \dots E_1^T$ . Nach Satz 4.5 ist also  $|A| = |E_1| \dots |E_k|$  und  $|A^T| = |E_k^T| \dots |E_1^T|$ . Daher genügt zu zeigen, daß  $|E^T| = |E|$  für jede elementare Matrix E ist. Im Fall  $E = U_{ij}(a)$  ist  $E^T = U_{ji}(a)$ , also  $|E| = 1 = |E^T|$  und im Fall  $E = D_i(a)$  ist  $E^T = E$ , also sicher  $|E^T| = |E|$ .

**Bem.:** Aus dem Satz folgt, daß alles was wir für die Zeilen von Determinanten bewiesen haben, entsprechend auch für die Spalten gilt.

**4.7 Satz:** (Entwicklungssatz von Laplace)

Für jede  $n \times n$ -Matrix  $A = (a_{ij})$  und  $k = 1; \dots; n$  gilt:

$$|A| = (-1)^{k+1} a_{1k} |A_{1k}| + \dots + (-1)^{k+n} a_{nk} |A_{nk}|.$$

**Bew.:** Man vertauscht zunächst die Spalten von A, so daß die k-te Spalte zur ersten wird und die Reihenfolge der übrigen unverändert bleibt. Dabei ändert sich die Determinante um den Faktor  $(-1)^{k-1} = (-1)^{k+1}$ . Dann folgt die Behauptung aus der Definition.

**Bem.:** Genauer spricht man bei der obigen Formel von einer Entwicklung nach der k-ten Spalte. Nach 4.6 hat man eine analoge Formel für die Entwicklung nach der k-ten Zeile. In der Praxis entwickelt man in der Regel nach der Zeile oder Spalte mit vielen Nullen.

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 0 \\ 3 & 6 & 7 & 1 \\ 2 & 0 & 3 & 1 \end{vmatrix} = \textcircled{1} \begin{vmatrix} 1 & 3 & 4 \\ 3 & 7 & 1 \\ 2 & 3 & 1 \end{vmatrix} - \textcircled{2} \begin{vmatrix} 1 & 2 & 4 \\ 3 & 6 & 1 \\ 2 & 0 & 1 \end{vmatrix} = \dots$$

Die Vorzeichen der Formel verteilen sich schachbrettartig:

+	-	+	-	+
-	+	-	+	-
+	-	+	-	+

23.11.00

**4.8 Def.:** Für eine  $n \times n$ -Matrix  $A = (a_{ij})$  definiert man die Adjunkte  $\tilde{A} = (\tilde{a}_{ij})$  von A durch  $a_{(adj)i} := (-1)^{i+j} |A_{ji}|$  (Achtung bei der Reihenfolge der Indizes).

Mit den obigen Beziehungen ist stets  $AA_{(adj)} = \det(A) 1_n = A_{(adj)}A$ .

**Bew.:** Für  $i \neq k$  ist das Element an der Position (i,k) von  $AA_{(adj)}$   $a_{i1}a_{(adj)1k} + \dots + a_{in}a_{(adj)nk} = (-1)^{k+1}a_{i1}|A_{k1}| + \dots + (-1)^{k+n}a_{in}|A_{kn}| = |B|$ , wobei die  $n \times n$ -Matrix B aus A gerade dadurch entsteht, daß man die k-te Zeile durch die i-te ersetzt. Da B zwei gleiche Zeilen hat, ist  $|B| = 0$ . Dagegen ist das Element an der Position (k,k) von  $A_{(adj)}A$  gleich  $a_{k1}a_{(adj)1k} + \dots + a_{kn}a_{(adj)nk} = (-1)^{k+1}a_{k1}|A_{k1}| + \dots + (-1)^{k+n}a_{kn}|A_{kn}| = |A|$ . Damit ist  $AA_{(adj)} = \det(A) 1_n$ , und analog beweist man  $A_{(adj)}A = \det(A) 1_n$ .

**Bem.:** Ist A invertierbar, so ist  $|A| \neq 0$ , und aus dem Satz folgt:

Dies ist die früher angekündigte Formel für inverse Matrizen.

In der Praxis ist diese Formel jedoch nur für kleine n brauchbar.

$$A^{-1} = \frac{1}{|A|} \tilde{A}$$

Sei  $n=2$   $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  mit  $0 \neq |A| = ad - bc$ . Dann ist A invertierbar und  $A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ .

Für größere n geht das nicht so gut.

**4.9 Satz:** (Cramersche Regel)

Ein lineares Gleichungssystem mit quadratischer Koeffizientenmatrix (\*) der Form (\*) ist genau dann eindeutig lösbar, wenn gilt:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

$$d = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \neq 0$$

Gegebenfalls gilt für  $i=1; \dots; n$

$$x_i = \frac{1}{d} \begin{vmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{vmatrix}$$

**Bew.:** Aus Satz 3.4 folgt, das (\*) genau dann eindeutig lösbar ist, wenn die Koeffizientenmatrix  $A = (a_{ij})$  invertierbar ist. Nach 4.4 ist dies wiederum äquivalent zu  $d \neq 0$ . Nach der vorigen Bemerkung ist dann  $A^{-1} = \frac{1}{d} A_{(adj)}$ . Außerdem ist  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$  das heißt für  $i=1; \dots; n$  gilt:

( siehe S.20)

$$x_i = \frac{1}{d} (a_{(adj)i1}b_1 + \dots + a_{(adj)in}b_n) = \frac{1}{d} [(-1)^{i+1}b_1|A_{1i}| + \dots + (-1)^{i+n}b_n|A_{ni}|] = \frac{1}{d} \begin{vmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{vmatrix}$$

**Bsp:** Gegeben sei das lineare Gleichungssystem  $\left. \begin{matrix} 2x+3y=1 \\ x+2y=1 \end{matrix} \right\} (*)$

Es ist  $d = \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 1$ , also  $x = \begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} = -1$ ;  $y = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = +1$ . (Probe!)

**Bem.:** Der obige Satz liefert die früher angekündigte Formel zur Lösung linearer Gleichungssysteme mit invertierbarer Koeffizientenmatrix. (Der Rechenaufwand steigt bei größeren n deutlich an). In der Praxis zieht man jedoch meist den Gauß-Algorithmus vor.

**4.10 Def.:** Gegeben sei eine  $m \times n$ -Matrix A. Eine Matrix, die durch Streichen von Zeilen und/oder Spalten von A entsteht, nennt man eine Untermatrix von A.

**Satz:** Für jede  $m \times n$ -Matrix A ist  $rg(A)$  die maximale Größe einer quadratischen Untermatrix von A mit nicht verschwindender Determinante.

**Bsp:** Die Matrix  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$  hat  $|A|=0$ , wie man leicht ausrechnet. Aber es ist  $\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = 5-8 \neq 0$ . Also ist  $rg(A)=2$ . (Vergleiche mit Beispiel 3.7)

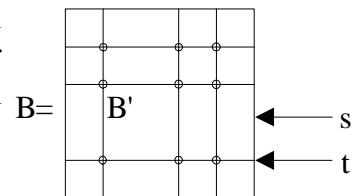
**Bew.:** Wir bezeichnen mit  $m(A)$  die maximale Größe einer quadratischen Untermatrix von A mit nicht verschwindender Determinante. Nach Satz 3.8 genügt es folglich, zu zeigen:

(I) Bei einer elementaren Zeilen- oder Spaltenumformung ändern sich weder der Rang noch  $m(A)$ .

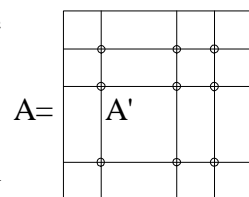
(II) Hat A die Form  $A = \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$ , so ist  $m(A)=r$

zu (I): Daß sich  $rg(A)$  bei elementaren Umformungen nicht ändert, folgt aus Satz 3.9. Daß sich  $m(A)$  bei elementaren Umformungen vom Typ II nicht ändert, ist klar. Daher betrachten wir Umformungen vom Typ I.

Die Matrix B entstehe aus der Matrix A durch Addition des r-fachen der s-ten Zeile zur t-ten ( $s \neq t$ ). Wir müssen die Determinanten der quadratischen Untermatrizen von A und B vergleichen. Dazu setzen wir  $k := m(B)$  und wählen eine  $k \times k$ -Untermatrix B' von B mit  $|B'| \neq 0$ . Gehört die t-te Zeile von A nicht zu A', so ist  $A' = B'$ , also  $|A'| = |B'| \neq 0$ . Wir können also annehmen, daß die t-te Zeile von A zu A' gehört. Gehört auch die s-te Zeile zu A', so entsteht die Matrix B' durch eine elementare Zeilenumformung vom Typ I aus der Untermatrix A'. Daher ist in diesem Fall  $|A'| = |B'| \neq 0$ .



Es bleibt nur der Fall übrig, daß die s-te Zeile von A nicht zu A' gehört. Mit A'' bezeichnen wir die  $k \times k$ -Untermatrix von A, die sich von A' nur dadurch unterscheidet, daß sie statt der t-ten Zeile die s-te enthält. Dann ist  $0 \neq |B'| = |A' \pm rA''|$ , also  $|A'| \neq 0$  oder  $|A''| \neq 0$ .



In jedem Fall ist also  $m(B) \leq m(A)$ . Da man die Rollen von B und A vertauschen kann, folgt  $m(B) = m(A)$ .

Analog ist die Sache bei elementaren Spaltenumformungen.

zu (II): Sei  $A = \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$ . Offenbar existiert eine  $r \times r$ -Untermatrix A' von A mit  $|A'| \neq 0$ . Für  $s > r$  enthält jede  $s \times s$ -Untermatrix A'' eine Nullzeile, d. h.  $|A''| = 0$ . Folglich ist  $m(A) = r$ .

**Bem.:** Für die Berechnung von  $rg(A)$  ist dieser Satz im Allgemeinen ungeeignet.

29.11.00

## 5. Reelle Vektorräume

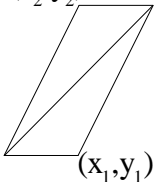
**5.1 Def.:** Ein reeller Vektorraum ist eine Menge mit Addition + und einer skalaren Multiplikat-

ion . Die Addition ordnet je zwei Elementen( Vektoren)  $v, w \in V$  ein Element  $v+w \in V$  zu, die Summe von  $v$  und  $w$ , und die skalare Multiplikation ordnet jedem( Skalar)  $r \in \mathbb{R}$  und jedem  $v \in V$  ein Element  $rv \in V$  zu, das skalare Vielfache.

Dabei sollen die folgenden Rechenregeln gelten:

- (I) Für  $v, w \in V$  ist  $v+w=w+v$  (*Kommutativgesetz*)
- (II) Für  $u, v, w \in V$  ist  $(u+v)+w=u+(v+w)$  (*Assoziativgesetz*)
- (III) Es existiert ein Element  $0 \in V$  (Nullvektor) mit  $v+0=v$  für  $v \in V$
- (IV) Zu jedem  $v \in V$  existiert ein( negatives) Element  $-v \in V$  mit  $v+(-v)=0$
- (V) Für  $r \in \mathbb{R}$  und  $v, w \in V$  ist  $r(v+w)=rv+rw$  (*Distributivgesetz*)
- (VI) Für  $r, s \in \mathbb{R}$ ,  $v \in V$  ist  $(r+s)v=rv+sv$  (*Distributivgesetz*)
- (VII) Für  $r, s \in \mathbb{R}$ ,  $v \in V$  ist  $(rs)v=r(sv)$  (*Assoziativgesetz*)
- (VIII) Für  $v \in V$  ist  $1v=v$ .

Bsp.: (a) Die Menge  $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$  ist ein reeller Vektorraum, dabei ist  $(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$  und  $r(x, y) := (rx, ry)$ .

$(x_2, y_2)$   $(x_1 + x_2, y_1 + y_2)$  Bei der Addition zweier Vektoren wird also ein Parallelogramm konstruiert.  

 Die Multiplikation entspricht einem Strecken des Vektors(  $r$  kann auch negativ sein).

b) Analog ist  $\mathbb{R}^3 = \{x, y, z \mid x, y, z \in \mathbb{R}\}$  ein reeller Vektorraum mit  $(x_1, y_1, z_1) + (x_2, y_2, z_2) := (x_1 + x_2, y_1 + y_2, z_1 + z_2)$  und  $r(x, y, z) := (rx, ry, rz)$ .

c) Allgemeiner ist  $\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{R}\}$  ein reeller Vektorraum mit  $(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$  und  $r(x_1, \dots, x_n) := (rx_1, \dots, rx_n)$ .

Man identifiziert  $\mathbb{R}^1$  mit  $\mathbb{R}$  und setzt zusätzlich  $\mathbb{R}^0 := \{0\}$ , dies ist ein Vektorraum, der nur ein Element enthält.

d) Auch die Menge  $\mathbb{R}^\infty := \{(x_1; x_2; \dots) \mid x_1; x_2; \dots \in \mathbb{R}\}$  aller unendlichen Folgen ist ein reeller Vektorraum mit  $(x_1; x_2; \dots) + (y_1; y_2; \dots) := (x_1 + y_1; x_2 + y_2; \dots)$  und  $r(x_1; x_2; \dots) := (rx_1; rx_2; \dots)$ .

e) Nach §2 ist für  $m, n \in \mathbb{N}$  auch die Menge  $\mathbb{R}^{m \times n}$  aller  $m \times n$ -Matrizen  $A = (a_{ij})$  ein reeller Vektorraum.

f) Für jede Menge  $D$  ist die Menge  $V$  aller Funktionen  $f: D \rightarrow \mathbb{R}$  ein reeller Vektorraum. Für Funktionen  $f, g: D \rightarrow \mathbb{R}$  ist dabei die Summe  $f+g: D \rightarrow \mathbb{R}$  durch  $(f+g)(x) := f(x) + g(x)$  für  $x \in D$  definiert. Für  $r \in \mathbb{R}$  ist  $rf: D \rightarrow \mathbb{R}$  durch  $(rf)(x) := rf(x)$  für  $x \in D$  definiert.

Bem.:(I) Wir sehen also, daß die Elemente eines Vektorraumes verschiedene mathematische Objekte sein können:  $n$ -Tupel, Folgen, Matrizen, Funktionen, ... .

Die Natur der Objekte ist dabei egal, es kommt nur auf die zwischen ihnen bestehenden Rechenregeln an.

(II) Ein Vektorraum kann keine zwei verschiedene Nullvektoren  $0_1$  und  $0_2$  enthalten, denn aus den Axiomen folgt:  $0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2$  .

Es ist üblich den Nullvektor in  $V$  mit dem gleichen Symbol wie die reelle Zahl  $0$  zu schreiben, obwohl es sich um zwei verschiedene Elemente handelt.

(III) Zu jedem  $v \in V$  existiert genau ein negatives  $-v \in V$ : Hätte man nämlich zwei Elemente  $v_1, v_2 \in V$  mit  $v + v_1 = 0 = v + v_2$ , so könnte man mit den Axiomen schließen:  $v_1 = v_1 + 0 = v_1 + (v + v_2) = (v_1 + v) + v_2 = 0 + v_2 = v_2 + 0 = v_2$  .

(IV) Statt  $v+(-w)$  schreibt man meist  $v-w$ .

**Satz:** Für Elemente  $v,w$  eines Vektorraumes  $V$  und alle  $r \in \mathbb{R}$  gilt stets:

(I)  $rv=0 \Leftrightarrow (r=0 \text{ oder } v=0)$

(II)  $-(rv)=(-r)v=r(-v)$

(III)  $-(-v)=v$

(IV)  $-(v+w)=-v-w$

**Bew.:** (I) „ $\Leftarrow$ “ Sei  $v \in V$  und  $w:=0v$ . Dann ist  $w=0v=(0+0)v=w+w$ , also  $0=w+(-w)=w+w+(-w)=w+0=\underline{w}$ .

Sei jetzt  $r \in \mathbb{R}$  und  $u:=r0$ . Dann ist  $u=r0=r(0+0)=r0+r0=u+u$ , also analog  $0=\dots=u$ .

„ $\Rightarrow$ “ Sei  $rv=0$ . Wir müssen  $r=0$  oder  $v=0$  zeigen. Im Fall  $r=0$  sind wir fertig. Sei also  $r \neq 0$ . Nach

dem ersten Teil des Beweises gilt dann:  $0 = \frac{1}{r} 0 = \frac{1}{r} (rv) = (\frac{1}{r} r)v = 1v = \underline{v}$ .

(II) Wegen  $rv+(-r)v=0v=0$  ist  $(-r)v$  das negative Element zu  $rv$ , d.h.  $(-r)v=-rv$ .

Analog ist  $r(-v)=-rv$ .

(III) Aus (II) folgt  $-v=-1v=(-1)v$ . Daher ist  $-(-v)=(-1)[(-1)v]=(-1)(-1)v=1v=\underline{v}$ .

(IV) Aus dem Beweis von (III) folgt:  $-(v+w)=(-1)v+(-1)w=(-v)+(-w)=-v-w$ .

**5.2 Bem.:** Wegen der Assoziativgesetze kann man oft auf Klammern verzichten und  $rsv$  oder  $u+v+w$  schreiben. Beim Rechnen mit Vektoren müssen wir darauf achten, daß wir nur die Axiome und daraus bereits abgeleitete Tatsachen verwenden.

**Def.:** Für Vektoren  $v_1, \dots, v_n$  in einem Vektorraum  $V$  nennt man  $\text{Span}(v_1, \dots, v_n) := \{r_1v_1 + \dots + r_nv_n : r_1, \dots, r_n \in \mathbb{R}\}$  den Aufspann von  $v_1, \dots, v_n$ . Die Elemente im Spann [engl. „Span“] ( $v_1, \dots, v_n$ ) nennt man Linearkombination von  $v_1, \dots, v_n$ .

**Satz:** Für Vektoren  $v_1, \dots, v_n$  in einem Vektorraum  $V$  gilt stets:

(I)  $0 \in \text{Span}(v_1, \dots, v_n)$ .

(II)  $r, s \in \mathbb{R}$  und  $x, y \in \text{Span}(v_1, \dots, v_n) \Rightarrow rx + sy \in \text{Span}(v_1, \dots, v_n)$ .

**Bew.:** (I)  $0=0+0+0+\dots+0=0v_1+\dots+0v_n \in \text{Span}(v_1, \dots, v_n)$  ( $0v_1+\dots+0v_n$  ist Linearkombination)

(II) Wir schreiben  $x=r_1v_1+\dots+r_nv_n$  und  $y=s_1v_1+\dots+s_nv_n$  mit  $r_1, s_1, \dots, r_n, s_n \in \mathbb{R}$ . Dann ist  $rx+sy=r(r_1v_1+\dots+r_nv_n)+s(s_1v_1+\dots+s_nv_n)=rr_1v_1+\dots+rr_nv_n+ss_1v_1+\dots+ss_nv_n=rr_1v_1+ss_1v_1+\dots+rr_nv_n+ss_nv_n=(rr_1+ss_1)v_1+\dots+(rr_n+ss_n)v_n \in \text{Span}(v_1, \dots, v_n)$ .  $[(rr_1+ss_1), \dots, (rr_n+ss_n)] \in \mathbb{R}$ .

**Bsp.:** (I) Für  $v \in V$  ist  $\text{Span}(v) = \{rv | r \in \mathbb{R}\} =: \mathbb{R}v$ . (Eine Gerade)

(II) Sei  $V = \mathbb{R}^3$   $v_1=(1;2;3)$  und  $v_2=(4;5;6)$ .

Dann ist  $\text{Span}(v_1, v_2)$  die Ebene, die durch  $v_1, v_2$  und den Nullpunkt geht. Wie stellt man fest, ob z.B. der Vektor  $(7;8;9)$  in dieser Ebene liegt?

Dazu macht man folgenden Ansatz  $(7;8;9)=r_1v_1+r_2v_2=r_1(1;2;3)+r_2(4;5;6)=\underline{(r_1+4r_2, 2r_1+5r_2, 3r_1+6r_2)}$ .

Man wird auf das folgende lineare Gleichungssystem geführt:

$$\begin{aligned} r_1+4r_2 &= 7 \\ 2r_1+5r_2 &= 8 \\ 3r_1+6r_2 &= 9 \end{aligned}$$

Als Lösung ergibt sich  $r_1=-1, r_2=2$

Probe:  $(-1)(1;2;3)+2(4;5;6)=(7;8;9)$  Daher ist  $(7;8;9) \in \text{Span}(v_1, v_2)$ .

**5.3 Def.:** Man nennt Vektoren  $v_1, \dots, v_n$  in einem Vektorraum  $V$  linear unabhängig, falls aus  $r_1v_1+\dots+r_nv_n=0$  mit  $r_1, \dots, r_n \in \mathbb{R}$  stets  $r_1=\dots=r_n=0$  folgt. Andernfalls nennt man  $v_1, \dots, v_n$  linear abhängig.

**Bsp.:** Wir wollen untersuchen, ob die folgenden Vektoren  $v_1=(1;1;0)$ ,  $v_2=(0;1;1)$  und  $v_3=(1;0;1)$  in  $\mathbb{R}^3$  linear unabhängig sind. Dazu betrachten wir folgende Gleichung:

$$0=r_1v_1+r_2v_2+r_3v_3=(r_1+r_3, r_1+r_2, r_2+r_3).$$

Wir werden so auf das folgende homogene lineare Gleichungssystem geführt:

$$r_1 + r_3 = 0$$

$$r_1 + r_2 = 0$$

$$r_2 + r_3 = 0$$

Der Gauß-Algorithmus liefert nur die triviale Lösung  $r_1=r_2=r_3=0$ . Daher sind  $v_1, v_2, v_3$  linear unabhängig.

30.11.00

**5.3 Satz:** Für Vektoren  $v_1, \dots, v_n$  in einem Vektorraum  $V$  ist äquivalent:

(1)  $v_1, \dots, v_n$  sind linear unabhängig.

(2) Jedes Element in  $\text{Span}(v_1, \dots, v_n)$  läßt sich als Linearkombination  $r_1v_1 + \dots + r_nv_n$  mit eindeutig bestimmten  $r_1, \dots, r_n \in \mathbb{R}$  schreiben.

**Bew.:** (1) $\Rightarrow$ (2) Sei (1) erfüllt und sei  $x=r_1v_1+\dots+r_nv_n$  und  $x=s_1v_1+\dots+s_nv_n$  mit eindeutig bestimmten  $r_1, \dots, r_n, s_1, \dots, s_n \in \mathbb{R}$ .

$$\text{Dann ist } 0=x-x=r_1v_1+\dots+r_nv_n-s_1v_1-\dots-s_nv_n=(r_1-s_1)v_1+\dots+(r_n-s_n)v_n.$$

Da  $v_1, \dots, v_n$  linear unabhängig sind, muß  $r_1-s_1=\dots=r_n-s_n=0$ , d.h.  $r_1=s_1, \dots, r_n=s_n$ .

(2) $\Rightarrow$ (1) Sei (2) erfüllt und seien  $r_1, \dots, r_n \in \mathbb{R}$  mit  $r_1v_1+\dots+r_nv_n=0$ , außerdem ist  $0v_1+\dots+0v_n=0$ .

Wegen (2) folgt  $r_1=\dots=r_n=0$ , also sind  $v_1, \dots, v_n$  linear unabhängig.

**5.4 Satz:** Gegeben seien linear abhängige Vektoren  $v_1, \dots, v_n$  in einem Vektorraum  $V$ . Dann läßt sich einer der Vektoren  $v_1, \dots, v_n$  als Linearkombination der übrigen schreiben.

**Bew.:** Nach Voraussetzung existieren reelle Zahlen  $r_1, \dots, r_n$ , die nicht alle gleich Null sind, und  $r_1v_1+\dots+r_nv_n=0$ .

$$\text{Wir wählen ein } i \text{ mit } r_i \neq 0, \text{ dann ist } v_i = \frac{-r_1}{r_i}v_1 + \dots + \frac{-r_{i-1}}{r_i}v_{i-1} + \dots + \frac{r_{i+1}}{r_i}v_{i+1} + \dots + \frac{r_n}{r_i}v_n.$$

**5.5 Def.:** Ein Untervektorraum eines Vektorraumes ist eine nichtleere Teilmenge  $U$  von  $V$  mit folgenden Eigenschaften:

(I) Für  $u, u' \in U$  ist  $u+u' \in U$ .

(II) Für  $r \in \mathbb{R}$  und  $u \in U$  ist  $ru \in U$ .

**Bem.:** Gegebenfalls enthält  $U$  wegen  $U \neq \emptyset$  ein Element  $u$ . Wegen (II) ist dann auch  $0=0u \in U$ . Jeder Untervektorraum von  $V$  enthält also auch den Nullvektor von  $V$ . Ferner besagen (I) und (II), daß  $U$  gegenüber der Addition und der skalaren Multiplikation „abgeschlossen“ ist. Es ist klar, daß  $U$  dann selbst ein Vektorraum ist.

**Satz:** Eine nichtleere Teilmenge  $U$  eines Vektorraumes  $V$  ist genau dann ein Untervektorraum von  $V$ , wenn für alle  $r, r' \in \mathbb{R}$ ,  $u, u' \in U$  gilt:  $ru+r'u \in U$ .

**Bew.:** Sei zunächst  $U$  ein Untervektorraum von  $V$ . Für  $r, r' \in \mathbb{R}$ ,  $u, u' \in U$  gilt dann nach (II) auch  $ru \in U$ ,  $r'u \in U$ , also nach (I)  $ru+r'u \in U$ .

Sei jetzt umgekehrt die Bedingung erfüllt. Für  $u, u' \in U$  ist dann  $u+u'=1u+1u' \in U$  und für  $r, r' \in \mathbb{R}$  ist  $ru+0u \in U$ .

**Bsp.:** (I) Für Vektoren  $v_1, \dots, v_n$  in einem Vektorraum  $V$  ist  $\text{Span}(v_1, \dots, v_n)$  ein Untervektorraum von  $V$  (Satz 5.2). Im Fall  $\text{Span}(v_1, \dots, v_n)=V$  sagt man auch: „ $v_1, \dots, v_n$  spannen  $V$  auf“ oder „ $v_1, \dots, v_n$  erzeugen  $V$ “.

(II)  $U=\{(x, y, 0) \mid x, y \in \mathbb{R}\}$  ist ein Untervektorraum von  $\mathbb{R}^3$ , denn wegen  $(0; 0; 0) \in U$  ist  $U \neq \emptyset$  und für  $r, r', x, x', y, y' \in \mathbb{R}$  gilt:  $r(x, y, 0)+r'(x', y', 0)=(rx+r'x', ry+r'y', 0) \in U$ .

(III)  $U=\{(x, y, z) \in \mathbb{R}^3 \mid x+y+z=0\}$  ist ein Untervektorraum von  $\mathbb{R}^3$ , denn wegen  $0+0+0=0$  ist  $(0; 0; 0) \in U$ .

$\in U$  unf für  $r,r' \in \mathbb{R}$  und  $(x,y,z), (x',y',z') \in U$  ist  $r(x,y,z)+r'(x',y',z')=(rx+r'x',ry+r'y',rz+r'z') \in U$  wegen  $rx+r'x'+ry+r'y'+rz+r'z'=r(x+y+z)+r'(x'+y'+z')=0$ .

(IV)  $U=\{(x,y,1)|x,y \in \mathbb{R}\}$  ist kein Untervektorraum von  $\mathbb{R}^3$  wegen  $(0;0;0) \notin U$ .

**5.6 Satz:** Für Untervektorräume  $U,U'$  eines Vektorraumes  $V$  ist auch  $U \cap U'$  ein Untervektorraum von  $V$ .

**Bew.:** Nach Bem. 5.5 enthalten  $U$  und  $U'$  den Nullvektor von  $V$ . Daher ist  $U \cap U' \neq \emptyset$ . Für  $r,r' \in \mathbb{R}$  und  $w,w' \in U \cap U'$  ist  $w,w' \in U$  und  $w,w' \in U'$ , also  $rw+r'w' \in U$  und  $rw+r'w' \in U'$ , d.h.  $rw+r'w' \in U \cap U'$ .

**Bem.:** I.A. ist  $U \cup U'$  kein Untervektorraum vom Vektorraum  $V$ .

**Bsp.:**  $V=\mathbb{R}^3, U=\mathbb{R}(1;2), U'=\mathbb{R}(3;4)$   
 Dann ist  $(1;2) \in U \subseteq U \cup U'$  und  $(3;4) \in U' \subseteq U \cup U'$ , aber  $(1;2)+(3;4)=(4;6) \notin U \cup U'$ , denn im Fall  $(4;6) \in U$  wäre  $(4;6)=r(1;2)$  für ein  $r \in \mathbb{R}$  und im Fall  $(4;6) \in U'$  wäre  $(4;6)=s(3;4)$  für ein  $s \in \mathbb{R}$ ; dies ist nicht möglich.

**5.7 Satz:** Für Untervektorräume  $U$  und  $U'$  eines Vektorraumes  $V$  ist auch die Summe  $U+U'=\{u+u' | u \in U, u' \in U'\}$  ein Untervektorraum von  $V$  mit  $U \subseteq U+U', U' \subseteq U+U'$ .

**Bew.:** Wegen  $0 \in U$  und  $0 \in U'$  ist  $0+0=0 \in U+U'$ , d.h.  $U+U' \neq \emptyset$ . Seien  $r,s \in \mathbb{R}$  und  $x,y \in U+U'$ . Wir schreiben  $x=u+u', y=v+v'$  mit  $u,v \in U, u',v' \in U'$ . Dann ist  $rx+sy=\underbrace{ru+sv}_{\in U}+\underbrace{ru'+sv'}_{\in U'} \in U+U'$

**Bem.:** Man sieht leicht, daß für Untervektorräume  $U_1,U_2,U_3$  von  $V$  stets das Assoziativgesetz gilt. Man kann daher auf Klammern verzichten und kurz  $U_1+U_2+U_3$  schreiben.

**5.8 Satz:** Für Untervektorräume  $U_1,\dots,U_n$  eines Vektorraumes  $V$  sind äquivalent:

- (1) Für  $i=1,\dots,n$  ist  $U_i \cap (U_1+\dots+U_{i-1}+U_{i+1}+\dots+U_n)=\{0\}$
- (2) Für  $i=1,\dots,n$  ist  $U_i \cap (U_1+\dots+U_{i-1})=\{0\}$
- (3) Sind  $u_1 \in U_1, \dots, u_n \in U_n$  mit  $u_1+\dots+u_n=0$ , so ist  $u_1=\dots=u_n=0$ .
- (4) Jedes Element in  $U_1+\dots+U_n$  läßt sich in der Form  $u_1+\dots+u_n$  mit eindeutig bestimmten  $u_1 \in U_1, \dots, u_n \in U_n$  schreiben.

**Bew.:** (1) $\Rightarrow$ (2) klar wegen  $U_1+\dots+U_{i-1} \subseteq U_1+\dots+U_{i-1}+U_{i+1}+\dots+U_n$ .  
 (2) $\Rightarrow$ (3) Sei (2) erfüllt und  $u_1+\dots+u_n=0$  mit  $u_1 \in U_1, \dots, u_n \in U_n$ . Wir nehmen an, daß  $u_i \neq 0$  ist für ein  $i \in \{1,\dots,n\}$  und wählen  $i$  maximal mit dieser Eigenschaft.

Dann ist  $-u_i = u_1+\dots+u_{i-1} \in U_i \cap (U_1+\dots+U_{i-1}) = \{0\}$  Widerspruch!!!

$$\underbrace{-u_i}_{\in U_i} = \underbrace{u_1+\dots+u_{i-1}}_{\in U_1+\dots+U_{i-1}}$$

(3) $\Rightarrow$ (4) Sei (3) erfüllt und  $u_1+\dots+u_n=v_1+\dots+v_n$  mit  $u_i,v_i \in U_i$  für  $i \in \{1,\dots,n\}$ . Dann ist  $0=u_1+\dots+u_n-(v_1+\dots+v_n)=\underbrace{(u_1-v_1)}_{\in U_1}+\dots+\underbrace{(u_n-v_n)}_{\in U_n}$ . Wegen (3) ist  $u_1-v_1=\dots=u_n-v_n=0$ , d.h.  $u_1=v_1,\dots,u_n=v_n$ .

(4) $\Rightarrow$ (1) Sei (4) erfüllt,  $i \in \{1,\dots,n\}$  und  $u_i \in U_i \cap (U_1+\dots+U_{i-1}+U_{i+1}+\dots+U_n)$ . Wir schreiben  $u_i = u_1+\dots+u_{i-1}+u_{i+1}+\dots+u_n$  mit  $u_j \in U_j$  für  $j \neq i$ . Dann ist  $\underbrace{0}_{\in U_i} + \dots + \underbrace{0}_{\in U_i} + \underbrace{0}_{\in U_n} + \dots + \underbrace{0}_{\in U_n} = \underbrace{u_1+\dots+u_{i-1}}_{\in U_1} + \underbrace{u_{i+1}+\dots+u_n}_{\in U_{i+1} \dots U_n}$ .

Wegen (4) folgt  $0=u_1=\dots=u_{i-1}, u_i=0, u_{i+1}=\dots=u_n=0$ .

Def.: Gegebenfalls schreibt man  $U_1 \oplus \dots \oplus U_n$  statt  $U_1 + \dots + U_n$  und spricht von direkter Summe von  $U_1, \dots, U_n$ .

Bem.: Im Fall  $n=2$  gilt also  $U_1 + U_2 = U_1 \oplus U_2 \Leftrightarrow U_1 \cap U_2 = \{0\}$ .

Bsp.: Für  $n \in \mathbb{N}$  ist  $\mathbb{R}^{n \times n} = S \oplus T$  mit  $S = \{A \mid A \in \mathbb{R}^{n \times n} \text{ ist symmetrisch}\}$  und  $T = \{A \mid A \in \mathbb{R}^{n \times n} \text{ ist schiefsymmetrisch}\}$ , denn nach Bem. 2.11 sind  $S, T$  Untervektorräume von  $\mathbb{R}^{n \times n}$  mit  $\mathbb{R}^{n \times n} = S + T$  und  $0_{n,n}$  ist die einzige Matrix, die gleichzeitig symmetrisch und schief- (antisymmetrisch) ist, d.h.  $S \cap T = \{0\}$ .

5.9 Satz: Die Lösungsmenge eines homogenen linearen Gleichungssystem (\*) mit  $n$  Unbekannten ist stets ein Untervektorraum von  $\mathbb{R}^n$ .

Bew.: nach Bem. 1.8

Bem.: Man spricht daher auch von dem Lösungsraum von (\*). Die Lösungsmenge eines inhomogenen linearen Gleichungssystem ist dagegen i.A. kein Untervektorraum, da sie den Nullvektor nicht enthält.

6.12.00

## 6. Basis und Dimension

6.1 Def.: Vektoren  $b_1, \dots, b_n$  in einem Vektorraum, die linear unabhängig sind und  $V$  aufspannen, nennt man eine Basis von  $V$ .

Bsp.: (I) Für  $n \in \mathbb{N}$  bilden die Elemente  $e_1 = (1; 0; \dots; 0)$ ,  $e_2 = (0; 1; 0; \dots; 0)$ , ...,  $e_n = (0; \dots; 0; 1)$  eine Basis des  $\mathbb{R}^n$ ; diese nennt man Standardbasis oder kanonische Basis des  $\mathbb{R}^n$ . Insbesondere ist  $1$  eine Basis von  $\mathbb{R}^1 = \mathbb{R}$ . Zusätzlich betrachtet man  $\emptyset$  als Basis von  $\mathbb{R}^0 = \{0\}$ .

(II) Für  $m, n \in \mathbb{N}$  bilden die Matrixeinheiten  $E_{ij} = (i=1; \dots; m, j=1; \dots; n)$  eine Basis des  $\mathbb{R}^{m \times n}$ , diese nennt man Standardbasis oder kanonische Basis des  $\mathbb{R}^{m \times n}$ .

Satz: Für Vektoren  $b_1, \dots, b_n$  in einem Vektorraum  $V$  sind die folgenden Aussagen äquivalent:

(1)  $b_1, \dots, b_n$  bilden eine Basis von  $V$ .

(2)  $b_1, \dots, b_n$  spannen  $V$  auf; läßt man aber irgendeinen der Vektoren  $b_1, \dots, b_n$  weg, so spannen die übrigen  $V$  nicht mehr auf.

(3)  $b_1, \dots, b_n$  sind linear unabhängig, nimmt aber einen Vektor  $b_{n+1} \in V$  hinzu, so sind  $b_1, \dots, b_n, b_{n+1}$  nicht mehr linear unabhängig.

Bew.: (1)  $\Rightarrow$  (2) Sei (1) erfüllt. Wir nehmen an:  $V = \text{Span}(b_2, \dots, b_n)$ . Dann existieren  $r_2, \dots, r_n \in \mathbb{R}$  mit  $b_1 = r_2 b_2 + \dots + r_n b_n$ . Folglich ist  $(-1)b_1 + r_2 b_2 + \dots + r_n b_n = 0$  im Widerspruch zur linearen Unabhängigkeit von  $b_1, \dots, b_n$ .

(2)  $\Rightarrow$  (3) Sei (2) erfüllt. Wir nehmen zunächst an, daß  $b_1, \dots, b_n$  linear abhängig sind. Nach Satz 5.4 können wir einen der Vektoren  $b_1, \dots, b_n$  als Linearkombination der übrigen schreiben. Nach Umm Nummerierung können wir also annehmen, daß  $b_1 = r_2 b_2 + \dots + r_n b_n$  mit  $r_2, \dots, r_n \in \mathbb{R}$  ist.

Zu jedem  $v \in V$  existieren  $s_1, \dots, s_n \in \mathbb{R}$  mit  $v = s_1 b_1 + \dots + s_n b_n = (s_1 \cdot r_2 + s_2) b_2 + \dots + (s_1 \cdot r_n + s_n) b_n \in \text{Span}(b_2, \dots, b_n)$ . Dieser Widerspruch zeigt, daß  $b_1, \dots, b_n$  linear unabhängig sind. Sei jetzt  $b_{n+1} \in V$  beliebig. Dann existieren  $t_1, \dots, t_n \in \mathbb{R}$  mit  $b_{n+1} = t_1 b_1 + \dots + t_n b_n$ , d.h.  $0 = (-1)b_{n+1} + t_1 b_1 + \dots + t_n b_n$ . Folglich sind die Vektoren linear abhängig.

(1)  $\Rightarrow$  (3) Sei (3) erfüllt und  $v \in V$ . Da  $b_1, \dots, b_n, v$  linear abhängig sind, existieren  $r_1, \dots, r_n, s \in \mathbb{R}$ , die nicht alle gleich 0 sind, mit  $r_1 b_1 + \dots + r_n b_n + s v = 0$ .

Im Fall  $s=0$  wäre  $r_1 b_1 + \dots + r_n b_n = 0$ , was im Widerspruch zur linearen Unabhängigkeit von  $b_1, \dots, b_n$  stünde. Also ist  $s \neq 0$  und  $v = -\frac{1}{s}(r_1 b_1 + \dots + r_n b_n) \in \text{Span}(b_1, \dots, b_n)$ . [ $b_1, \dots, b_n$  spannen ganz  $V$  auf]

6.2 Satz: (Austauschsatz von Steinitz)

Gegeben sei ein Vektorraum  $V$ , der von endlich vielen Elementen  $e_1, \dots, e_k$  aufgespannt wird. Sind  $u_1, \dots, u_l \in V$  linear unabhängig, so existieren Indizes  $i_1, \dots, i_m \in \{1, \dots, k\}$  mit der Eigenschaft, daß  $u_1, \dots, u_l, e_{i_1}, \dots, e_{i_m}$  eine Basis von  $V$  bilden.

Bew.: Wir wählen möglichst viele Vektoren  $e_{i_1}, \dots, e_{i_m}$  mit der Eigenschaft, daß  $u_1, \dots, u_l, e_{i_1}, \dots, e_{i_m}$  linear unabhängig sind (evtl.  $m=0$ ). Für  $j \in \{1, \dots, k\} \setminus \{i_1, \dots, i_m\}$  sind  $u_1, \dots, u_l, e_{i_1}, \dots, e_{i_m}, e_j$  linear abhängig. Daher existieren Zahlen  $r_1, \dots, r_l, s_1, \dots, s_m, t \in \mathbb{R}$ , die nicht alle gleich 0 sind, mit  $0 = r_1 u_1 + \dots + r_l u_l + s_1 e_{i_1} + \dots + s_m e_{i_m} + t e_j$ . Im Fall  $t=0$  wäre  $0 = r_1 u_1 + \dots + r_l u_l + s_1 e_{i_1} + \dots + s_m e_{i_m}$ , was im Widerspruch zur linearen Unabhängigkeit von  $u_1, \dots, u_l, e_{i_1}, \dots, e_{i_m}$  steht. Daher ist  $t \neq 0$  und  $e_j = \frac{-1}{t}(r_1 u_1 + \dots + r_l u_l + s_1 e_{i_1} + \dots + s_m e_{i_m}) \in \text{Span}(u_1, \dots, u_l, e_{i_1}, \dots, e_{i_m})$ . Also liegen  $e_1, \dots, e_k$  im  $\text{Span}(u_1, \dots, u_l, e_{i_1}, \dots, e_{i_m})$ . Folglich gehört auch jede Linearkombination von  $e_1, \dots, e_k$  zum  $\text{Span}(u_1, \dots, u_l, e_{i_1}, \dots, e_{i_m})$ .

Bem.: Insbesondere kann man also einige der Vektoren  $e_1, \dots, e_k$  weglassen, um eine Basis von  $V$  zu erhalten (Dies ist der Fall, wenn  $l=0$ ).

6.3 Satz: Gegeben sei eine Basis  $b_1, \dots, b_n$  eines Vektorraumes  $V$ . Für  $r_1, \dots, r_n \in \mathbb{R}$  mit  $r_1 \neq 0$  bilden dann auch  $b := r_1 b_1 + \dots + r_n b_n, b_2, \dots, b_n$  eine Basis von  $V$ .

Bew.: Zum Beweis der linearen Unabhängigkeit seien  $s_1, \dots, s_n \in \mathbb{R}$  mit  $0 = s_1 b + s_2 b_2 + \dots + s_n b_n = s_1 r_1 b_1 + (s_1 r_2 + s_2) b_2 + \dots + (s_1 r_n + s_n) b_n$ . Da  $b_1, \dots, b_n$  linear unabhängig sind, folgt  $s_1 r_1 = 0$ . Wegen  $r_1 \neq 0$  folgt  $s_1 = 0$ . Also ist  $s_2 b_2 + \dots + s_n b_n = 0$ . Da  $b_2, \dots, b_n$  linear unabhängig sind, folgt  $s_2 = \dots = s_n = 0$ . Damit ist gezeigt, daß  $b, b_2, \dots, b_n$  linear unabhängig sind. Sei jetzt  $v \in V$ .

Dann existieren  $t_1, \dots, t_n \in \mathbb{R}$  mit  $v = t_1 b_1 + \dots + t_n b_n = \frac{t_1 b}{r_1} + \left(t_2 - r_2 \frac{t_1}{r_1}\right) b_2 + \dots + \left(t_n - r_n \frac{t_1}{r_1}\right) b_n \in \text{Span}(b, b_2, \dots, b_n)$ . Also ist  $V = \text{Span}(b, b_2, \dots, b_n)$ .

6.4 Satz: Gegeben seien zwei Basen  $a_1, \dots, a_m$  und  $b_1, \dots, b_n$  von  $V$ . Dann ist  $m=n$ .

Bew.: Sei  $k$  die Anzahl der Vektoren, die in beiden Basen gleichzeitig vorkommen (evtl.  $k=0$ ). Nach Ummummerierung können wir  $a_1 = b_1, \dots, a_k = b_k$  annehmen. Sei zunächst  $k < n$ . Wir schreiben  $b_{k+1} = r_1 a_1 + \dots + r_m a_m$  mit  $r_1, \dots, r_m \in \mathbb{R}$ . Im Fall  $r_{k+1} = \dots = r_m = 0$  wäre  $b_{k+1} = r_1 a_1 + \dots + r_k a_k = r_1 b_1 + \dots + r_k b_k$ . Dies steht im Widerspruch zur linearen Unabhängigkeit von  $b_1, \dots, b_n$ .

Also ist  $r_i \neq 0$  für  $i \in \{k+1, \dots, m\}$ . Indem wir  $a_{k+1}, \dots, a_m$  notfalls umnummerieren können  $r_{k+1} \neq 0$  annehmen. Nach Satz 6.3 bilden auch  $a_1, \dots, a_k, b_{k+1}, a_{k+2}, \dots, a_m$  eine Basis von  $V$ .

Diese enthält auch  $m$  Vektoren, hat aber mit der Basis  $b_1, \dots, b_n$  genau  $k+1$  Vektoren gemeinsam. Wir ersetzen  $a_1, \dots, a_m$  durch  $a_1, \dots, a_k, b_{k+1}, a_{k+2}, \dots, a_m$  und wiederholen das Verfahren. Dies können wir solange machen, bis  $k=n$  ist. In diesem Fall sind aber  $b_1, \dots, b_n$  und  $b_1 = a_1, \dots, b_n = a_n, a_{n+1}, \dots, a_m$  Basen von  $V$ . Aus Satz 6.1 folgt also  $m=n$ .

Def.: Man nennt  $n$  die Dimension von  $V$  und schreibt  $n = \dim V$ . Besitzt ein Vektorraum keine endliche Basis, so nennt man  $V$  unendlich dimensional und schreibt  $\dim V = \infty$ .

Bsp.:  $\dim \mathbb{R}^n = n$ ,  $\dim \mathbb{R}^{n \times m} = n \cdot m$ .

6.5 Satz: Für jeden Vektorraum der Dimension  $n < \infty$  gilt:

- (I) Sind  $u_1, \dots, u_n$  in  $V$  linear unabhängig, so bilden  $u_1, \dots, u_n$  eine Basis von  $V$ .
- (II) Ist  $V = \text{Span}(a_1, \dots, a_n)$ , so bilden  $a_1, \dots, a_n$  eine Basis von  $V$ .

Bew.: (I) Wir wählen eine Basis  $b_1, \dots, b_n$  von  $V$ . Nach Steinitz kann man  $u_1, \dots, u_n$  durch einige der Vektoren  $b_1, \dots, b_n$  zu einer Basis  $u_1, \dots, u_n, b_{i_1}, \dots, b_{i_m}$  von  $V$  ergänzen. Nach Satz 6.4 ist  $m=0$ , d.h.  $u_1, \dots, u_n$  bilden selbst schon eine Basis von  $V$ .

(II) Nach Steinitz bilden einige der Vektoren  $a_1, \dots, a_n$  eine Basis  $a_{i_1}, \dots, a_{i_m}$  von  $V$ . Nach Satz 6.4 ist  $m=n$ , d.h.  $a_1, \dots, a_n$  bilden eine Basis von  $V$ .

**6.6 Satz:** Für jeden Untervektorraum  $U$  eines endlich-dimensionalen Vektorraumes  $V$  ist  $\dim U \leq \dim V$ . Im Fall  $\dim U = \dim V$  folgt  $U=V$ .

**Bew.:** Wir wählen eine Basis  $e_1, \dots, e_n$  von  $V$ . Sind  $u_1, \dots, u_m$  beliebige linear unabhängige Vektoren in dem Vektorraum  $U$ , so kann man diese nach Steinitz durch einige der Vektoren zu einer Basis von  $V$  ergänzen. Nach 6.4 ist  $m \leq n$ . Wählt man  $m$  möglichst groß, so bilden  $u_1, \dots, u_m$  nach 6.1 eine Basis von  $U$ . Dann ist aber  $\dim U = m \leq n = \dim V$ .

Im Fall  $m=n$  folgt aus 6.5(I), daß  $u_1, \dots, u_m$  auch eine Basis von  $V$  bilden. Also ist  $U=V$ .

**Def.:** Untervektorräume der Dimension 1 nennt man auch Geraden( durch 0), solche der Dimension 2 Ebenen( durch 0). Im Fall  $\dim V = n < \infty$  nennt man Untervektorräume der Dimension  $n-1$  Hyperebenen( durch 0).

13.12.00

**6.7 Satz:** Für Untervektorräume eines endlich-dimensionalen Vektorraumes  $V$  gilt stets:

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2.$$

**Bew.:** Nach Satz 6.6 sind alle aufgeführten Untervektorräume endlich-dimensional. Wir wählen eine Basis  $a_1, \dots, a_k$  von  $U_1 \cap U_2$  und ergänzen diese zu Basen  $a_1, \dots, a_k, b_1, \dots, b_l$  von  $U_1$  und  $a_1, \dots, a_k, c_1, \dots, c_m$  von  $U_2$ .

Wir wollen zeigen, daß  $a_1, \dots, a_k, b_1, \dots, b_l, c_1, \dots, c_m$  eine Basis von  $U_1 + U_2$  bilden. Sei  $u \in U_1 + U_2$ . Wir schreiben  $u = u_1 + u_2$  mit  $u_1 \in U_1$  und  $u_2 \in U_2$ . Dann existieren  $q_1, r_1, \dots, q_k, r_k, s_1, \dots, s_l, t_1, \dots, t_m \in \mathbb{R}$  mit  $u_1 = q_1 a_1 + \dots + q_k a_k + s_1 b_1 + \dots + s_l b_l$  und  $u_2 = r_1 a_1 + \dots + r_k a_k + t_1 c_1 + \dots + t_m c_m$ .

Folglich ist  $u = (q_1 + r_1)a_1 + \dots + (q_k + r_k)a_k + s_1 b_1 + \dots + s_l b_l + t_1 c_1 + \dots + t_m c_m$  eine Linearkombination von  $a_1, \dots, a_k, b_1, \dots, b_l, c_1, \dots, c_m$ .

Seien jetzt  $x_1, \dots, x_k, y_1, \dots, y_l, z_1, \dots, z_m \in \mathbb{R}$  mit  $x_1 a_1 + \dots + x_k a_k + y_1 b_1 + \dots + y_l b_l + z_1 c_1 + \dots + z_m c_m = 0$ . Dann ist  $x_1 a_1 + \dots + x_k a_k + y_1 b_1 + \dots + y_l b_l = -z_1 c_1 - \dots - z_m c_m \in U_1 \cap U_2$ .

Da  $a_1, \dots, a_k$  eine Basis von  $U_1 \cap U_2$  bilden, folgt  $y_1 = \dots = y_l = 0$ . Folglich ist  $x_1 a_1 + \dots + x_k a_k + z_1 c_1 + \dots + z_m c_m = 0$ . Da  $a_1, \dots, a_k$  und  $c_1, \dots, c_m$  linear unabhängig sind, folgt, daß  $x_1 = \dots = x_k = z_1 = \dots = z_m = 0$ . Damit ist gezeigt, daß  $a_1, \dots, a_k, b_1, \dots, b_l, c_1, \dots, c_m$  eine Basis von  $U_1 + U_2$  bilden.

Wir erhalten:  $\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = (k+l+m) + k = (k+l) + (k+m) = \dim U_1 + \dim U_2$ .

**6.8 Satz:** Zu jedem Untervektorraum  $U$  eines endlich-dimensionalen Vektorraumes  $V$  existiert ein Untervektorraum  $U'$  von  $V$  mit  $V = U \oplus U'$ .

**Bew.:** Wir wählen eine Basis  $b_1, \dots, b_m$  von  $U$ , ergänzen diese zu einer Basis  $b_1, \dots, b_m, c_1, \dots, c_n$  von  $V$  und setzen  $U' := \text{Span}(c_1, \dots, c_n)$ . Für  $v \in V$  existieren Skalare  $r_1, \dots, r_m, s_1, \dots, s_n \in \mathbb{R}$  mit  $v = \underbrace{r_1 b_1 + \dots + r_m b_m}_{\in U} + \underbrace{s_1 c_1 + \dots + s_n c_n}_{\in U'}$ .

Ist  $u \in U \cap U'$ , so existieren  $x_1, \dots, x_m, y_1, \dots, y_n \in \mathbb{R}$  mit  $u = x_1 b_1 + \dots + x_m b_m = y_1 c_1 + \dots + y_n c_n$ , d.h.  $0 = x_1 b_1 + \dots + x_m b_m - y_1 c_1 - \dots - y_n c_n$ . Da  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$  linear unabhängig sind, folgt  $x_1 = \dots = x_m = y_1 = \dots = y_n = 0$ . Daher ist  $u = 0$ . Damit ist gezeigt, daß  $U \cap U' = \{0\}$ , d.h.  $V = U \oplus U'$ .

**Def.:** Man nennt  $U'$  ein Komplement von  $U$  in  $V$ .

**Bem.:** I. A. ist  $U'$  durch  $U$  nicht eindeutig bestimmt.

**6.9 Def.:** Für  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  bezeichnet man  $a_1 = (a_{11}, \dots, a_{1n}), \dots, a_m = (a_{m1}, \dots, a_{mn}) \in \mathbb{R}^n$  als Zeilen von  $A$ .

und den Untervektorraum  $ZR(A)=\text{Span}(a_1, \dots, a_m)$  von  $\mathbb{R}^n$  als Zeilenraum von A.  
 Analog definiert man die Spalten von A und den Spaltenraum  $SR(A) \subseteq \mathbb{R}^m$ .

- Satz:** (I) Elementare Zeilenumformungen ändern  $ZR(A)$  nicht.  
 (II) Elementare Spaltenumformungen ändern  $SR(A)$  nicht.  
 (III) Für  $A \in \mathbb{R}^{m \times n}$  ist  $\text{rg}(A) = \dim ZR(A) = \dim SR(A)$ .

**Bew.:** (I) Mit den obigen Bezeichnungen gilt für  $r \in \mathbb{R}$  und  $i, j \in \{1; \dots; m\}$  mit  $i \neq j$ :  $a_j + ra_i \in \text{Span}(a_1, \dots, a_m)$ . Folglich ist der  $\text{Span}(a_1, \dots, a_{j-1}, a_j + ra_i, a_{j+1}, \dots, a_m) \subseteq ZR(A)$ .  
 Die umgekehrte Inklusion erhält man analog. In ähnlicher Weise kann man bei elementaren Umformungen vom Typ II vorgehen.

(II) Analog.

(III) Nach dem ersten Teil können wir annehmen, daß die Matrix A reduzierte Zeilenstufenform hat:

$$\left( \begin{array}{cccccccc} 0 & \dots & 0 & a_{1s_1} & & & & \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & a_{2s_2} \\ \vdots & & & & & & & \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & a_{rs_r} \\ 0 & \dots & & & & & & & & & \\ \vdots & & & & & & & & & & \\ 0 & \dots & & & & & & & & & \end{array} \right) \quad \text{mit } s_1 < \dots < s_r \text{ und } a_{1s_1} = a_{2s_2} = \dots = a_{rs_r} = 1.$$

} Weitere Zeilen sind  
} Nullzeilen

Offenbar ist  $ZR(A) = \text{Span}(a_1, \dots, a_r)$ . Daher ist zu zeigen, daß die Zeilen  $a_1, \dots, a_r$  linear unabhängig sind. Seien also  $A_1, \dots, A_r \in \mathbb{R}$  mit  $A_1 a_1 + \dots + A_r a_r = 0$ . Für  $j=1; \dots; r$  ist dann  $0 = A_1 a_{1s_j} + \dots + A_r a_{rs_j} = \underline{A_j}$ . Damit ist die Behauptung gezeigt.

**Bsp.:** (a) Wie bestimmt man eine Basis des von  $u_1=(1;-2;1;2)$ ,  $u_2=(2;3;0;-1)$ ,  $u_3=(1;-3;3;7)$  aufgespannten Untervektorraumes U von  $V := \mathbb{R}^4$  ?

Man faßt  $u_1, u_2, u_3$  als Zeilen einer Matrix auf und wendet hierauf den Gauß-Algorithmus an.

$$\left( \begin{array}{cccc} 1 & -2 & 1 & 2 \\ 2 & 3 & 0 & -1 \\ 1 & -3 & 3 & 7 \end{array} \right) \xrightarrow{\text{G.-Algorithmus}} \left( \begin{array}{cccc} 1 & 0 & -3 & -8 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Daher bilden  $u_1=(1;0;-3;-8)$  und  $u_2=(0;1;-2;-5)$  eine Basis von U.

(b) Wie ergänzt man die linear unabhängigen Vektoren  $u_1, u_2$  in (a) zu einer Basis von V ?

Die Rechnung in (a) zeigt  $U = \text{Span}(u_1, u_2) = \text{Span}(v_1, v_2)$ . Wegen

$$\text{rg} \left( \begin{array}{cccc} 1 & 0 & -3 & -8 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) = 4 \quad \left. \begin{array}{l} \} v_3 \\ \} v_4 \end{array} \right\} \text{ ist } V = \text{Span}(v_1, v_2, v_3, v_4) = \text{Span}(u_1, u_2) \text{ mit } v_3=(0;0;1;0), v_4=(0;0;0;1). \text{ Daher bilden } u_1, u_2 \text{ eine Basis von } V.$$

(c) Wie bestimmt man eine Basis von  $U \cap V$ , wobei  $U = \text{Span}(u_1, u_2, u_3)$ ,  $V = \text{Span}(v_1, v_2, v_3) \subseteq \mathbb{R}^4$  und  $u_1=(1;-2;1;1)$ ,  $u_2=(0;1;-2;-5)$ ,  $u_3=(0;0;1;1)$ ,  $v_1=(1;2;3;4)$ ,  $v_2=(0;1;2;3)$ ,  $v_3=(0;0;1;2)$  ist ?

Ist  $w \in U \cap V$ , so existieren  $x_1, x_2, x_3, y_1, y_2, y_3 \in \mathbb{R}$  mit  $w = x_1 u_1 + x_2 u_2 + x_3 u_3 = y_1 v_1 + y_2 v_2 + y_3 v_3$ , d.h.  $0 = x_1 u_1 + x_2 u_2 + x_3 u_3 - y_1 v_1 - y_2 v_2 - y_3 v_3 = (x_1 - y_1, -2x_1 + x_2 - 2y_1 - y_2, x_1 - 2x_2 + x_3 - 3y_1 - 2y_2 - y_3, 2x_1 - 5x_2 + x_3 - 4y_1 - 3y_2 - 2y_3)$ .

Man erhält ein homogenes lineares Gleichungssystem mit Koeffizientenmatrix:

$$\left( \begin{array}{cccccc} 1 & 0 & 0 & -1 & 0 & 0 \\ -2 & 1 & 0 & -2 & -1 & 0 \\ 1 & -2 & 1 & -3 & -2 & -1 \\ -2 & -5 & 1 & -4 & -3 & -2 \end{array} \right) \xrightarrow{\text{G.-Algorithmus}} \left( \begin{array}{cccccc} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -4 & -1 & 0 \\ 0 & 0 & 1 & -10 & -4 & -1 \\ 0 & 0 & 0 & -12 & -4 & -1 \end{array} \right)$$

Lösung:  $-12y_1 - 4y_2 - y_3 = 0$ , d.h.  $y_3 = -12y_1 - 4y_2$ .

Folglich ist  $w = y_1 v_1 + y_2 v_2 + (-12y_1 - 4y_2) v_3 = y_1 (v_1 - 12v_3) + y_2 (v_2 - 4v_3) = y_1 (1;2;-9;-20) + y_2 (0;1;-2;-5)$ .

Daher bilden  $(1;2;-9;-20)$ ,  $(0;1;-2;-5)$  eine Basis von  $U \cap V$ . (Probe)

7. Lineare Abbildungen

7. Def.: Gegeben seien zwei Vektorräume  $V$  und  $W$ . Eine Abbildung  $f:V \rightarrow W$  mit  $f(av+a'v')=af(v)+a'f(v')$  ( $a,a' \in \mathbb{R}; v,v' \in V$ ) nennt man linear oder (Vektorraum-) Homomorphismus. Man setzt  $\text{Hom}(V,W)=\{f:V \rightarrow W \mid f \text{ linear}\}$ .

Bem.: (I) Die obige Linearitätsbedingung läßt sich folgendermaßen aufspalten:

$$\left. \begin{array}{l} (1) f(v+v')=f(v)+f(v') \\ (2) f(av)=af(v) \end{array} \right\} a \in \mathbb{R}; v,v' \in V$$

(II) Für  $f \in \text{Hom}(V,W)$  gilt nach (2):  $f(0)=f(0 \cdot 0)=0f(0)=0$ .

Eine lineare Abbildung bildet stets den Nullvektor in  $V$  auf den Nullvektor in  $W$  ab.

Bsp.: Für  $A \in \mathbb{R}^{m \times n}$  ist die Abbildung  $f:\mathbb{R}^{n \times 1} \rightarrow \mathbb{R}^{m \times 1}, x \rightarrow Ax$  linear; dies folgt sofort aus den Rechenregeln für Matrizen. Wir werden später sehen, daß jede lineare Abbildung vom  $\mathbb{R}^{n \times 1}$  in den  $\mathbb{R}^{m \times 1}$  auf diese Weise gegeben wird( Multiplikation mit einer Matrix von links). Lineare Abbildungen stehen also in enger Verwandtschaft zu Matrizen.

Satz: Für Vektorräume  $V,W$  und  $f \in \text{Hom}(V,W)$  gilt:

(I) Für jeden Untervektorraum  $V'$  von  $V$  ist  $f(V'):=\{f(v') \mid v' \in V'\}$  ein Untervektorraum von  $W$ . Insbesondere ist das  $\text{Bld}(f):=f(V)$  ein Untervektorraum von  $W$ .

(II) Für jeden Untervektorraum  $W'$  von  $W$  ist  $f^{-1}(W'):=\{v \in V \mid f(v) \in W'\}$  ein Untervektorraum von  $V$ , insbesondere ist  $\text{Ker}(f):=f^{-1}\{0\}=\{v \in V \mid f(v)=0\}$  ein Untervektorraum von  $V$ .

*( $f^{-1}$  steht hier für das Urbild; nicht für die Umkehrabbildung!!!)*

(III) Jeweils ist  $f^{-1}(f(V'))=V'+\text{Ker}(f)$  und  $f(f^{-1}(W'))=W' \cap \text{Bld}(f)$ .

(IV) Ist  $V$  endlich-dimensional, so auch  $\text{Bld}(f)$  und  $\text{Ker}(f)$  und es gilt:

$$\boxed{\dim V = \dim \text{Ker}(f) + \dim \text{Bld}(f)}. \text{ ?}$$

Bew.: (I) Wegen  $V' \neq \emptyset$  ist auch  $f(V') \neq \emptyset$ . Seien  $a,a' \in \mathbb{R}$  und  $w,w' \in f(V')$ . Dann existieren  $v,v' \in V$  mit  $w=f(v); w'=f(v')$ . Folglich ist  $aw+a'w'=af(v)+a'f(v')=f(\underbrace{av+a'v'}_{\in V'}) \in f(V')$

(II) Wegen  $f(0)=0 \in W'$  ist  $0 \in f^{-1}(W')$ , d.h.  $f^{-1}(W') \neq \emptyset$ .

Seien  $a,a' \in \mathbb{R}$  und  $v,v' \in f^{-1}(W')$ , d.h.  $f(v), f(v') \in W'$ . Dann ist  $f(\underbrace{av+a'v'}_{\in W'})=af(v)+a'f(v') \in W'$ , d.h.  $av+a'v' \in f^{-1}(W')$ .

14.12.00

Bew.: (III) Sei  $v \in f^{-1}(f(V'))$ , d.f.  $f(v) \in f(V')$ . Dann existiert  $v' \in V'$  mit  $f(v)=f(v')$ . Folglich ist  $f(v-v')=f(v)-f(v')=0$ , d.h.  $v-v' \in \text{Ker } f$  und  $v=v'+(v-v') \in V'+\text{Ker } f$ . Damit ist gezeigt:  $f^{-1}(f(V')) \subseteq V'+\text{Ker } f$ . Für  $v' \in V'$  und  $x \in \text{Ker } f$  ist umgekehrt  $f(v'+x)=f(v')+\underbrace{f(x)}_0=f(v') \in f(V')$ , d.h.  $v'+x \in f^{-1}(f(V'))$ .

Sei jetzt  $w \in f(f^{-1}(W'))$ . Dann existiert  $v \in f^{-1}(W')$  mit  $w=f(v)$ . Folglich ist  $w=f(v) \in W' \cap \text{Bld}(f)$ . Damit ist gezeigt, daß  $f(f^{-1}(W')) \subseteq W' \cap \text{Bld}(f)$ .

Sei umgekehrt  $w' \in W' \cap \text{Bld}(f)$ . Dann existiert  $v \in V$  mit  $f(v)=w' \in W'$ . Folglich ist  $v \in f^{-1}(W')$  und  $w'=f(v) \in f(f^{-1}(W'))$ . Damit ist gezeigt:  $f(f^{-1}(W'))=W' \cap \text{Bld}(f)$ .

(IV) Sei  $V$  endlich-dimensional. Als Untervektorraum von  $V$  ist dann auch  $\text{Ker}(f)$  endlich dimensional. Wir wählen jetzt eine Basis  $b_1, \dots, b_m$  von  $\text{Ker}(f)$  und ergänzen diese zu einer Basis  $b_1, \dots, b_m, b_{m+1}, \dots, b_n$  von  $V$ . Es genügt zu zeigen, daß  $f(b_{m+1}), \dots, f(b_n)$  eine Basis von  $\text{Bld}(f)$  bilden, dann ist  $\text{Bld}(f)$  endlich-dimensional und  $\dim V = n = m + (n - m) = \dim \text{Ker}(f) + \dim \text{Bld}(f)$ .

Sei  $w \in \text{Bld}(f)$ . Dann existiert  $v \in V$  mit  $w=f(v)$ . Wir schreiben  $v=r_1b_1+\dots+r_nb_n$  mit  $r_1, \dots, r_n \in \mathbb{R}$ . Dann ist  $w=f(v)=r_1\underbrace{f(b_1)}_0+\dots+r_m\underbrace{f(b_m)}_0+r_{m+1}f(b_{m+1})+\dots+r_nf(b_n)$ .

Damit ist gezeigt:  $\text{Bld}(f) \subseteq \text{Span}(f(b_{m+1}), \dots, f(b_n))$ . Die umgekehrte Inklusion ist klar. Also ist  $\text{Bld}(f) = \text{Span}(f(b_{m+1}), \dots, f(b_n))$ .

Seien jetzt  $r_{m+1}, \dots, r_n \in \mathbb{R}$  mit  $0 = r_{m+1}f(b_{m+1}) + \dots + r_n f(b_n) = f(r_{m+1}b_{m+1} + \dots + r_n b_n)$ . Dann ist  $r_{m+1}b_{m+1} + \dots + r_n b_n \in \text{Ker}(f)$ . Da  $b_1, \dots, b_m$  eine Basis von  $\text{Ker}(f)$  bilden, folgt  $r_{m+1} = \dots = r_n = 0$ . Also sind tatsächlich  $f(b_{m+1}), \dots, f(b_n)$  linear unabhängig, d.h. sie bilden eine Basis von  $\text{Ker}(f)$ .

Def.: Für Vektorräume  $V, W$  nennt man  $\text{Ker } f := \{v \in V \mid f(v) = 0\}$  den Kern und  $\text{Bld } f := \{f(v) \mid v \in V\}$  das Bild von  $f$ . Ferner nennt man  $\text{def}(f) := \dim \text{Ker}(f)$  den Defekt von  $f$  und  $\text{rg}(f) := \dim \text{Bld}(f)$  den Rang von  $f$ .

Bsp.: Die Abbildung  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2, (x, y, z) \rightarrow (x+y+z, x-y)$  ist linear, denn für alle  $r, r', x, x', y, y', z, z' \in \mathbb{R}$  gilt:  $f(r(x, y, z) + r'(x', y', z')) = f(rx + r'x', ry + r'y', rz + r'z') = (rx + r'x' + ry + r'y' + rz + r'z', rx + r'x' - ry - r'y') = r(x+y+z, x-y) + r'(x'+y'+z', x'-y') = rf(x, y, z) + r'f(x', y', z')$ .

Ferner gilt:  $(x, y, z) \in \text{Ker } f \Leftrightarrow (x+y+z, x-y) = 0 \Leftrightarrow \begin{cases} x+y+z=0 \\ x-y=0 \end{cases}$

Die Bestimmung von  $\text{Ker } f$  führt also auf die Lösung eines homogenen linearen Gleichungssystems.

Als Lösung ergibt sich:  $x=y$  und  $z=-2x$ .

Daher ist  $\text{Ker } f = \mathbb{R}(1; 1; -2)$ . Insbesondere ist  $\text{def}(f) = 1$ . Aus dem Satz 7.1 folgt:  $\text{rg}(f) = 3 - \text{def}(f) = 2$ .

Also ist  $\text{Bld } f = \mathbb{R}^2$ , d.h.  $f$  ist surjektiv.

Satz: Für Vektorräume  $V, W$  und  $f \in \text{Hom}(V, W)$  gilt:  $f$  ist injektiv  $\Leftrightarrow \text{Ker } f = \{0\}$

Bew.: (Wir wissen:  $f$  ist injektiv  $\Leftrightarrow$  aus  $f(x) = f(y)$  folgt  $x=y$ ).

Sei zunächst  $f$  injektiv. Für  $v \in \text{Ker } f$  ist  $f(v) = 0 = f(0)$ . Wegen der Injektivität von  $f$  folgt  $v=0$ .

Daher ist  $\text{Ker } f = \{0\}$ .

Sei umgekehrt  $\text{Ker } f = \{0\}$ . Sind  $x, y \in V$  mit  $f(x) = f(y)$ , so ist  $f(x-y) = f(x) - f(y) = 0$ , d.h.  $x-y \in \text{Ker } f = \{0\}$ , also  $x-y=0$ ; d.h.  $x=y$ . Folglich ist  $f$  injektiv.

Bem.: Injektive lineare Abbildungen nennt man auch Monomorphismen, surjektive lineare Abbildungen nennt man Epimorphismen.

[Erinnerung: Eine (lineare) Abbildung  $f$  nennt man surjektiv, falls  $\text{Bld } f = W$  ist].

7.3 Satz: Für Vektorräume  $V, W$  mit  $\dim V = \dim W < \infty$  und  $f \in \text{Hom}(V, W)$  gilt:  
 $f$  injektiv  $\Leftrightarrow f$  surjektiv

Bew.: Wegen  $\dim \text{Ker } f + \dim \text{Bld } f = \dim V = \dim W < \infty$  gilt:

$f$  injektiv  $\Leftrightarrow \text{Ker } f = \{0\} \Leftrightarrow \dim \text{Ker } f = 0 \Leftrightarrow \dim \text{Bld } f = \dim W \Leftrightarrow \text{Bld } f = W \Leftrightarrow f$  surjektiv.

Bem.: Gegebenfalls ist  $f$  bijektiv. Eine bijektive lineare Abbildung nennt man Isomorphismus.

7.4 Satz: Sind  $V, W$  zwei Vektorräume und ist  $f \in \text{Hom}(V, W)$  bijektiv, so ist auch die Umkehrabbildung  $f^{-1}: W \rightarrow V$  linear.

Bew.: Für  $r, s \in \mathbb{R}, x, y \in W$  gilt:

$$f^{-1}(rx + sy) = f^{-1}(r f(f^{-1}(x)) + s f(f^{-1}(y))) = f^{-1}(f(r f^{-1}(x) + s f^{-1}(y))) = r f^{-1}(x) + s f^{-1}(y).$$

Bem.: Für Vektorräume  $U, V, W$  und  $f \in \text{Hom}(U, V), g \in \text{Hom}(V, W)$  ist die zusammengesetzte Abbildung  $g \circ f: U \rightarrow W; u \rightarrow g(f(u))$  linear, denn für  $r, s \in \mathbb{R}, x, y \in U$  gilt:

$$(g \circ f)(rx + sy) = g(f(rx + sy)) = g(r f(x) + s f(y)) = r g(f(x)) + s g(f(y)) = r(g \circ f)(x) + s(g \circ f)(y).$$

Def.: Man nennt zwei Vektorräume isomorph und schreibt  $V \cong W$ , falls ein Isomorphismus  $f: V \rightarrow W$  existiert.

7.5 Satz: Für Vektorräume  $U, V, W$  gilt stets:

- (I)  $V \cong W$  (Reflexivität)
- (II)  $V \cong W \Leftrightarrow W \cong V$  (Symmetrie)
- (III)  $U \cong V$  und  $V \cong W \Rightarrow U \cong W$  (Transitivität)

Bem.: Der Satz besagt also, daß Isomorphie von Vektorräumen eine Äquivalenzrelation ist.

Bew.: (I) Die Identitätsabbildung  $\text{id}_V: V \rightarrow V, v \rightarrow v$  ist linear und bijektiv. Daher ist  $V \cong V$ .

(II) Sei  $V \cong W$ . Dann existiert ein Isomorphismus  $f: V \rightarrow W$ . Nach Satz 7.4 ist auch  $f^{-1}: W \rightarrow V$  ein Isomorphismus. Folglich ist  $W \cong V$ .

(III) Sei  $U \cong V$  und  $V \cong W$ . Dann existieren Isomorphismen  $f: U \rightarrow V$  und  $g: V \rightarrow W$ . Nach Bemerkung 7.4 ist auch  $g \circ f: U \rightarrow W$  ein Isomorphismus. Also ist  $U \cong W$ .

7.6 Satz: Gegeben seien Vektorräume  $V, W$  und eine Basis  $b_1, \dots, b_n$  von  $V$  und beliebige Elemente  $c_1, \dots, c_n \in W$ . Dann existiert genau ein  $f \in \text{Hom}(V, W)$  mit  $f(b_1) = c_1, \dots, f(b_n) = c_n$ .

Dabei ist:

- (I)  $\text{Bld } f = \text{Span}(c_1, \dots, c_n)$ .
- (II)  $\text{Ker } f = \{a_1 b_1 + \dots + a_n b_n \mid a_1, \dots, a_n \in \mathbb{R}, a_1 c_1 + \dots + a_n c_n = 0\}$ .
- (III)  $f$  ist surjektiv  $\Leftrightarrow W = \text{Span}(c_1, \dots, c_n)$ .
- (IV)  $f$  ist injektiv  $\Leftrightarrow c_1, \dots, c_n$  sind linear unabhängig.
- (V)  $f$  ist bijektiv  $\Leftrightarrow c_1, \dots, c_n$  bilden eine Basis von  $W$  (!)

Bew.: Existenz Da  $b_1, \dots, b_n$  eine Basis von  $V$  bilden, läßt sich jedes Element in  $V$  in der Form  $a_1 b_1 + \dots + a_n b_n$  mit eindeutigen  $a_1, \dots, a_n \in \mathbb{R}$  schreiben.

Wir setzen  $f(a_1 b_1 + \dots + a_n b_n) = a_1 c_1 + \dots + a_n c_n$  und erhalten eine Abbildung  $f: V \rightarrow W$  mit  $f(b_1) = c_1, \dots, f(b_n) = c_n$ . Diese Abbildung ist tatsächlich linear, denn für  $r, r', a'_1, \dots, a'_n \in \mathbb{R}$  gilt:

$$f(r(a_1 b_1 + \dots + a_n b_n) + r'(a'_1 b_1 + \dots + a'_n b_n)) = f((ra_1 b_1 + \dots + ra_n b_n + r'a'_1 b_1 + \dots + r'a'_n b_n)) = f((ra_1 + r'a'_1)b_1 + \dots + (ra_n + r'a'_n)b_n) = (ra_1 + r'a'_1)c_1 + \dots + (ra_n + r'a'_n)c_n = (ra_1 c_1 + r'a'_1 c_1) + \dots + (ra_n c_n + r'a'_n c_n) = r(a_1 c_1 + \dots + a_n c_n) + r'(a'_1 c_1 + \dots + a'_n c_n) = rf(a_1 b_1 + \dots + a_n b_n) + r'f(a'_1 b_1 + \dots + a'_n b_n).$$

15.12.00

7.6 Bew.: Eindeutigkeit Seien  $f, g \in \text{Hom}(V, W)$  mit  $f(b_i) = c_i = g(b_i)$  für  $1; \dots; n$ . Sei  $v \in V$  beliebig. Wir schreiben  $v = a_1 b_1 + \dots + a_n b_n$  mit  $a_1, \dots, a_n \in \mathbb{R}$ . Dann ist  $f(v) = a_1 f(b_1) + \dots + a_n f(b_n) = a_1 c_1 + \dots + a_n c_n = a_1 g(b_1) + \dots + a_n g(b_n) = g(v)$ . Dies zeigt:  $f = g$ .

(I)  $\text{Bld } f = \{a_1 c_1 + \dots + a_n c_n \mid a_1, \dots, a_n \in \mathbb{R}\} = \text{Span}(c_1, \dots, c_n)$ .

(II) klar. (III) folgt aus (I).

(IV) „ $\Rightarrow$ “ Sei  $f$  injektiv und seien  $a_1, \dots, a_n \in \mathbb{R}$  mit  $a_1 c_1 + \dots + a_n c_n = 0$ .

Nach (II) ist dann  $a_1 b_1 + \dots + a_n b_n \in \text{Ker } f = \{0\}$ . Da  $b_1, \dots, b_n$  linear sind, unabhängig folgt  $a_1 = \dots = a_n = 0$ .

Daher sind  $c_1, \dots, c_n$  linear unabhängig.

„ $\Leftarrow$ “ Seien  $c_1, \dots, c_n$  linear unabhängig und sei  $v \in \text{Ker } f$ . Wir schreiben  $v = a_1 b_1 + \dots + a_n b_n$  mit  $a_1, \dots, a_n \in \mathbb{R}$ . Nach (II) ist  $a_1 c_1 + \dots + a_n c_n = 0$ . Folglich ist  $a_1 = \dots = a_n = 0$ , da  $c_1, \dots, c_n$  linear unabhängig sind, d.h.  $v = 0$ . Dies zeigt  $\text{Ker } f = \{0\}$ . Daher ist  $f$  injektiv.

(V) folgt aus (III) und (IV).

7.7 Satz: Für endlich-dimensionale Vektorräume  $V, W$  gilt stets:

$$V \cong W \Leftrightarrow \dim V = \dim W$$

Bem.: Ein reeller Vektorraum der Dimension  $n < \infty$  ist also stets zum  $\mathbb{R}^n$  isomorph.

Bew.: „ $\Rightarrow$ “ Es sei  $V \cong W$ . Dann existiert ein Isomorphismus  $f: V \rightarrow W$ . Aus Satz 7.1 folgt:  $\dim V = \dim \underbrace{\text{Ker } f}_{=0} + \dim \underbrace{\text{Bld } f}_{=W} = \dim W$ .

„ $\Leftarrow$ “ Sei  $\dim V = \dim W$ . Wir wählen Basen  $b_1, \dots, b_n$  von  $V$  und  $c_1, \dots, c_n$  von  $W$ . Nach Satz 7.6 existiert genau ein  $f \in \text{Hom}(V, W)$  mit  $f(b_1) = c_1, \dots, f(b_n) = c_n$  und  $f$  ist Isomorphismus, also  $V \cong W$ .

7.8 Def.: Für Vektorräume  $V, W$  und  $f, g \in \text{Hom}(V, W)$  und  $a \in \mathbb{R}$  definiert man Abbildungen  $f+g: V \rightarrow W$  und  $af: V \rightarrow W$  durch  $(f+g)(v) := f(v) + g(v)$  und  $(af)(v) := af(v)$  für  $v \in V$ .

Bem.: Dann sind  $f+g$  und  $af$  wieder linear, denn für  $r, s \in \mathbb{R}$  und  $x, y \in V$  gilt:

$$(f+g)(rx+sy) = f(rx+sy) + g(rx+sy) = rf(x) + sf(y) + rg(x) + sg(y) = r(f(x) + g(x)) + s(f(y) + g(y)) = r(f+g)(x) + s(f+g)(y).$$

$$(af)(rx+sy) = af(rx+sy) = a(rf(x) + sf(y)) = arf(x) + asf(y) = raf(x) + saf(y) = r(af)(x) + s(af)(y).$$

Satz: Auf diese Weise wird die Menge  $\text{Hom}(V, W)$  zu einem reellen Vektorraum.

Bew.: Man sieht sofort, daß die Addition kommutativ und assoziativ ist. Nullvektor im  $\text{Hom}(V, W)$  ist die Nullabbildung, die jedes  $v \in V$  auf den Nullvektor in  $W$  abbildet. Negativ zu  $f \in \text{Hom}(V, W)$  ist die Abbildung  $-f: V \rightarrow W$ ,  $v \rightarrow -f(v)$ . Man sieht außerdem, daß die Multiplikation mit Skalaren assoziativ ist. Ferner gilt:

$$(a(f+g))(v) = a(f+g)(v) = a(f(v) + g(v)) = af(v) + ag(v) = (af)(v) + (ag)(v) = (af+ag)(v).$$

$$((a+b)f)(v) = (a+b)f(v) = af(v) + bf(v) = (af)(v) + (bf)(v) = (af+bf)(v) \text{ für } v \in V, \text{ d.h. es gelten die Distributivgesetze } a(f+g) = af+ag \text{ und } (a+b)f = af+bf \text{ (} a, b \in \mathbb{R}, f, g \in \text{Hom}(V, W)\text{)}.$$

Schließlich ist  $1f = f$  für  $f \in \text{Hom}(V, W)$ .

7.9 Satz: Für Vektorräume  $U, V, W$  und  $a \in \mathbb{R}$ ,  $f, f' \in \text{Hom}(U, V)$ ,  $g, g' \in \text{Hom}(V, W)$  gilt:

$$(I) \quad a(gof) = (ag)of = go(af)$$

$$(II) \quad (g+g')of = gof + g'of$$

$$(III) \quad go(f+f') = gof + gof'$$

Bew.: Für  $u \in U$  gilt: (I)  $(a(gof))(u) = a(gof)(u) = ag(f(u)) = (ag)(f(u)) = ((ag)of)(u)$ .

$$(a(gof))(u) = ag(f(u)) = g(af(u)) = g((af)(u)) = (go(af))(u).$$

$$(II) \quad ((g+g')of)(u) = (g+g')(f(u)) = g(f(u)) + g'(f(u)) = (gof)(u) + (g'of)(u) = (gof + g'of)(u).$$

$$(III) \quad (go(f+f'))(u) = g((f+f')(u)) = g(f(u) + f'(u)) = g(f(u)) + g(f'(u)) = (gof)(u) + (gof')(u).$$

## 8. Lineare Abbildungen und Matrizen

8.1 Def.: Gegeben seien endlich-dimensionale Vektorräume  $V, W$  mit Basen  $b_1, \dots, b_m$  bzw.  $c_1, \dots, c_n$  und eine lineare Abbildung  $f: V \rightarrow W$ .

Wir schreiben  $f(b_j) = a_{1j}c_1 + \dots + a_{nj}c_n$ , ( $a_{ij} \in \mathbb{R}$ )

Dann nennt man  $A = (a_{ij}) \in \mathbb{R}^{n \times m}$  die Matrix von  $f$  bezüglich der Basen  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$ .

Bem.: (I) Die Zahlen  $a_{ij}$  sind durch  $f$  und  $b_1, \dots, b_m, c_1, \dots, c_n$  eindeutig bestimmt.

(II) Merkregel: Die Matrix  $A$  enthält in der  $j$ -ten Spalte die „Koeffizienten“ (oder „Koordinaten“) von  $f(b_j)$  bezüglich  $c_1, \dots, c_n$ .

(III) Nach Satz 7.6 existiert umgekehrt zu jeder Matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times m}$  genau eine lineare Abbildung  $f: V \rightarrow W$  mit  $f(b_j) = a_{1j}c_1 + \dots + a_{nj}c_n$ . ( $j = 1; \dots; m$ )

(IV) Auf diese Weise kann man Fragen über (abstrakte) lineare Abbildungen in Fragen über (konkrete) Matrizen übersetzen, für deren Lösung wir bereits Verfahren kennen.

Bsp.: a) Nach Bsp. 7.1 ist für  $A = (a_{ij}) \in \mathbb{R}^{n \times m}$  die Abbildung  $f: \mathbb{R}^{m \times 1} \rightarrow \mathbb{R}^{n \times 1}$ ,  $x \rightarrow Ax$  linear. Wir be-

zeichnen mit  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$  die Standardbasen von  $\mathbb{R}^{m \times 1}$  bzw.  $\mathbb{R}^{n \times 1}$ . Für  $j=1, \dots, m$  ist dann

$$f(b_j) = Ab_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = a_{1j}c_1 + \dots + a_{nj}c_n. \text{ Daher ist } A \text{ selbst die Matrix von } f \text{ bezüglich } b_1, \dots, b_m \text{ und } c_1, \dots, c_n.$$

b) Man rechnet leicht nach, daß  $V := \{(a, b, c) \in \mathbb{R}^3 \mid a+b+c=0\}$ ,  $W := \{(r, s, t, u) \in \mathbb{R}^4 \mid r+s+t+u=0\}$  Untervektorräume von  $\mathbb{R}^3$  sind, mit Basen  $v_1 := (1; 0; -1)$ ,  $v_2 := (0; 1; -1)$  bzw.  $w_1 := (1; 0; 0; -1)$ ,  $w_2 := (0; 1; 0; -1)$  und  $w_3 := (0; 0; 1; -1)$ .

Durch  $f(a, b, c) := (a-2b-c, 2a-b-c, -a-b, -6a-2c)$  wird eine lineare Abbildung  $f: V \rightarrow W$  definiert; denn aus  $a+b+c=0$  folgt  $(a-2b-c) + (2a-b-c) + (-a-b) + (-6a-2c) = -4(a+b+c) = 0$  und die Linearität von  $f$  sieht man unmittelbar. Was ist die Matrix von  $f$  bezüglich  $v_1, v_2$  und  $w_1, w_2, w_3$  ?

Wegen  $f(v_1) = (2; 3; -1; -4) = 2(1; 0; 0; -1) + 3(0; 1; 0; -1) - (0; 0; 1; -1)$  und  $f(v_2) = (-1; 0; -1; 2) = -(1; 0; 0; -1) - (0; 0; 1; -1)$  hat die Matrix von  $f$  bezüglich  $v_1, v_2$  und  $w_1, w_2, w_3$  die Form  $A = \begin{pmatrix} 2 & -1 \\ 3 & 0 \\ -1 & -1 \end{pmatrix}$

**Satz:** Für endlich-dimensionale Vektorräume  $U, V, W$  mit Basen  $a_1, \dots, a_l$  bzw.  $b_1, \dots, b_m$  bzw.  $c_1, \dots, c_n$  gilt stets:

- (I) Sind  $s, s' \in \mathbb{R}$  und  $f, f' \in \text{Hom}(V, W)$  mit Matrizen  $R = (r_{ij})$ ,  $R' = (r'_{ij}) \in \mathbb{R}^{n \times m}$  bezüglich  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$ , so ist  $sR + s'R$  die Matrix von  $sf + s'f'$  bezüglich  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$ .
- (II) Sind  $f \in \text{Hom}(U, V)$  und  $g \in \text{Hom}(V, W)$  mit Matrizen  $R = (r_{ij}) \in \mathbb{R}^{m \times l}$  bezüglich  $a_1, \dots, a_l$  und  $b_1, \dots, b_m$  bzw.  $S = (s_{ij}) \in \mathbb{R}^{n \times m}$  bezüglich  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$ , so ist  $SR \in \mathbb{R}^{n \times l}$  die Matrix von  $g \circ f$  bezüglich  $a_1, \dots, a_l, b_1, \dots, b_m$  und  $c_1, \dots, c_n$ .

**Bew.:** (I) Aus  $f(b_j) = r_{1j}c_1 + \dots + r_{nj}c_n$  und  $f'(b_j) = r'_{1j}c_1 + \dots + r'_{nj}c_n$  folgt:

$$(sf + s'f')(b_j) = sf(b_j) + s'f'(b_j) = s(r_{1j}c_1 + \dots + r_{nj}c_n) + s'(r'_{1j}c_1 + \dots + r'_{nj}c_n) = (sr_{1j} + s'r'_{1j})c_1 + \dots + (sr_{nj} + s'r'_{nj})c_n.$$

(II) Aus  $f(a_j) = r_{1j}b_1 + \dots + r_{mj}b_m$  und  $g(b_k) = s_{1k}c_1 + \dots + s_{nk}c_n$  folgt:

$$(g \circ f)(a_j) = g(r_{1j}b_1 + \dots + r_{mj}b_m) = r_{1j}g(b_1) + \dots + r_{mj}g(b_m) = r_{1j}(s_{11}c_1 + \dots + s_{n1}c_n) + \dots + r_{mj}(s_{1m}c_1 + \dots + s_{nm}c_n) = (s_{11}r_{1j} + \dots + s_{1m}r_{mj})c_1 + \dots + (s_{n1}r_{1j} + \dots + s_{nm}r_{mj})c_n.$$

20.12.00

**8.2 Satz:** Für endlich-dimensionale Vektorräume  $V, W$  mit Basen  $b_1, \dots, b_m$  bzw.  $c_1, \dots, c_n$  ist die Abbildung  $F: \text{Hom}(V, W) \rightarrow \mathbb{R}^{n \times m}$ , die jeder linearen Abbildung  $f: V \rightarrow W$  ihre Matrix  $A$  bezüglich  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$  zuordnet, ein Isomorphismus.

**Bew.:** siehe 8.1

**Bem.:** (I) Insbesondere ist  $\dim \text{Hom}(V, W) = \dim \mathbb{R}^{n \times m} = n \cdot m = (\dim V)(\dim W)$ .

Im Fall  $W = \mathbb{R}$  nennt man  $V^* := \text{Hom}(V, \mathbb{R})$  den Dualraum von  $V$  und seine Elemente Linearformen auf  $V$ . Es ist also  $\dim V^* = \dim V$  (falls  $\dim V < \infty$ ).

(II) Im Fall  $V = W$  ist  $m = n$ . Man wählt dann in der Regel  $c_1 = b_1, \dots, c_n = b_n$  und spricht dann auch von der Matrix  $f$  bezüglich  $b_1, \dots, b_n$ .

Nach Satz 8.1 gilt in diesem Fall  $F(g \circ f) = F(g)F(f)$  für  $f, g \in \text{Hom}(V, W)$ .

**8.3 Satz:** Gegeben seien endlich-dimensionale Vektorräume  $V, W$  und eine lineare Abbildung  $f: V \rightarrow W$  mit Matrix  $A$  bezüglich der Basen  $b_1, \dots, b_m$  von  $V$  und  $c_1, \dots, c_n$  von  $W$ . Dann gilt:

$$\boxed{f \text{ ist bijektiv} \Leftrightarrow A \text{ ist invertierbar}}.$$

Gegebenfalls ist  $m = n$  und  $A^{-1}$  die Matrix von  $f^{-1}$  bezüglich  $c_1, \dots, c_n$  und  $b_1, \dots, b_m$ .

**Bew.:** „ $\Rightarrow$ “ Sei  $f$  bijektiv. Nach Satz 7.4 ist dann die Umkehrabbildung  $f^{-1}: W \rightarrow V$  auch linear. Wir bezeichnen mit  $A' = (a'_{ij}) \in \mathbb{R}^{m \times n}$  die Matrix von  $f^{-1}$  bezüglich  $c_1, \dots, c_n$  und  $b_1, \dots, b_m$ . Nach Satz 8.1 ist  $A'A$  die Matrix der linearen Abbildung  $f^{-1} \circ f = \text{id}_V$  bezüglich  $b_1, \dots, b_m$ .

Wegen  $\text{id}_V(b_j) = b_j$  ( $j = 1, \dots, m$ ) ist andererseits  $1_m$  die Matrix von  $\text{id}_V$  bezüglich  $b_1, \dots, b_m$ . Daher ist  $A'A = 1_m$ . Analog ist  $AA' = 1_n$ . Folglich ist  $A$  invertierbar,  $m = n$  und  $A' = A^{-1}$ .

„⇐“ Sei  $A$  invertierbar und  $g:W \rightarrow V$  die nach Satz 8.2 eindeutig bestimmte lineare Abbildung mit Matrix  $A^{-1}$  bezüglich  $c_1, \dots, c_n$  und  $b_1, \dots, b_m$ . Nach Satz 8.1 ist  $A^{-1}A = 1_m$  die Matrix von  $g \circ f$  bezüglich  $b_1, \dots, b_m$ , d.h.  $(g \circ f)(b_j) = b_j = \text{id}_V(b_j)$  für  $j=1; \dots; m$ . Aus Satz 7.6 folgt also  $g \circ f = \text{id}_V$ . Analog ist  $f \circ g = \text{id}_W$ , d.h.  $f$  ist bijektiv und  $g = f^{-1}$ .

**8.4 Satz:** Gegeben seien ein endlich-dimensionaler Vektorraum  $V$  mit Basis  $b_1, \dots, b_n$  und eine Matrix  $S = (s_{ij}) \in \mathbb{R}^{n \times n}$ . Wir setzen  $c_j = s_{1j}b_1 + \dots + s_{nj}b_n$  ( $j=1; \dots; n$ ). Dann gilt:

$S$  invertierbar  $\Leftrightarrow c_1, \dots, c_n$  ist eine Basis von  $V$ .

**Bew.:** Nach Satz 7.6 existiert genau eine lineare Abbildung  $f \in \text{Hom}(V, V)$  mit  $f(b_j) = c_j$  ( $j=1; \dots; n$ ). Offenbar ist  $S$  gerade die Matrix von  $f$  bezüglich  $b_1, \dots, b_n$ . Daher gilt nach 8.3 und 7.6 :  
 $S$  ist invertierbar  $\Leftrightarrow f$  ist bijektiv  $\Leftrightarrow c_1, \dots, c_n$  bilden eine Basis von  $V$ .

**Bem.:** Gegebenfalls ist  $b_j = s'_{1j}c_1 + \dots + s'_{nj}c_n$  für  $j=1; \dots; n$  mit  $S^{-1} = (s'_{ij})$ .

**8.5 Bem.:** Gegeben seien zwei endlich-dimensionale Vektorräume  $V, W$  und eine lineare Abbildung  $f:V \rightarrow W$ . Mit  $A = (a_{ij}) \in \mathbb{R}^{n \times m}$  bezeichnen wir die Matrix von  $f$  bezüglich einer Basis  $b_1, \dots, b_m$  von  $V$  und einer Basis  $c_1, \dots, c_n$  von  $W$ . Hat man weitere Basen  $b'_1, \dots, b'_m$  von  $V$  und  $c'_1, \dots, c'_n$  von  $W$ , so kann man auch die Matrix  $A' = (a'_{ij}) \in \mathbb{R}^{n \times m}$  von  $f$  bezüglich  $b'_1, \dots, b'_m$  und  $c'_1, \dots, c'_n$  betrachten. Es ist wichtig, sich zu überlegen, wie  $A$  und  $A'$  zusammenhängen.

**Satz:** Gegeben seien die obigen Voraussetzungen. Wir schreiben  $b'_j = s_{1j}b_1 + \dots + s_{mj}b_m$  mit  $s_{ij} \in \mathbb{R}$ ,  $i, j = 1; \dots; m$   $c_l = t_{1l}c'_1 + \dots + t_{nl}c'_n$  mit  $t_{kl} \in \mathbb{R}$ ,  $k, l = 1; \dots; n$ . Dann sind  $S = (s_{ij})$  und  $T = (t_{kl})$  invertierbar und es gilt:  $A' = T A S$ , insbesondere sind  $A$  und  $A'$  äquivalent.  $\wp$

**Bew.:** Die Invertierbarkeit von  $S$  und  $T$  folgt aus dem Satz 8.4 und außerdem gilt:

$$a'_{ij}c'_1 + \dots + a'_{nj}c'_n = f(b'_j) = f(s_{1j}b_1 + \dots + s_{mj}b_m) = s_{1j}f(b_1) + \dots + s_{mj}f(b_m) = s_{1j}(a_{11}c_1 + \dots + a_{n1}c_n) + \dots + s_{mj}(a_{1m}c_1 + \dots + a_{nm}c_n) = \underbrace{(a_{11}s_{1j} + \dots + a_{1m}s_{mj})}_{r_{1j}}c_1 + \dots + \underbrace{(a_{n1}s_{1j} + \dots + a_{nm}s_{mj})}_{r_{nj}}c_n = r_{1j}(t_{11}c'_1 + \dots + t_{n1}c'_n) + \dots + r_{nj}(t_{1n}c'_1 + \dots + t_{nn}c'_n) = (t_{11}r_{1j} + \dots + t_{1n}r_{nj})c'_1 + \dots + (t_{n1}r_{1j} + \dots + t_{nn}r_{nj})c'_n.$$

Wegen der linearen Unabhängigkeit von  $c'_1, \dots, c'_n$  folgt daraus  $a'_{ij} = t_{i1}r_{1j} + \dots + t_{in}r_{nj}$  ( $i=1; \dots; n, j=1; \dots; m$ ).

Mit der Matrix  $R := (r_{ij})$  gilt also:  $A' = TR$ , also wegen  $R = AS$  folgt  $A' = T A S$ .

**8.6 Bem.:** Gegeben seien endlich-dimensionale Vektorräume  $V, W$  und eine lineare Abbildung  $f:V \rightarrow W$ . Wir wollen Basen von  $V$  und  $W$  so wählen, daß die Matrix von  $f$  bezüglich dieser Basen eine „möglichst einfache“ Gestalt hat.

**Satz:** In der obigen Situation kann man Basen von  $V$  und  $W$  so wählen, daß die Matrix von  $f$  bezüglich dieser Basen die Form  $\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$  hat. Dabei ist  $r = \text{rg}(f)$ .

**Bew.:** Wir wählen zunächst eine Basis  $b_1, \dots, b_m$  von  $\text{Ker } f$  und ergänzen diese zu einer Basis  $b_1, \dots, b_m, b_{m+1}, \dots, b_n$  von  $V$ . Im Beweis von Satz 7.1 hatten wir gezeigt, daß dann  $f(b_{m+1}), \dots, f(b_n)$  eine Basis von  $\text{Bild}(f)$  ist.

Wir ergänzen diese zu einer Basis  $c_1 = f(b_{m+1}), \dots, c_{n-m} = f(b_n), c_{n-m+1}, \dots, c_k$  von  $W$ . Dann hat die Matrix von  $f$  bezüglich  $b_{m+1}, \dots, b_n, b_1, \dots, b_m$  und  $c_1, \dots, c_{n-m}, c_{n-m+1}, \dots, c_k$  gerade die gewünschte Form.  $\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$   
 Offenbar ist  $r = n - m$ .

**Bsp.:** In Bsp. 8.1 hatten wir zwei Untervektorräume  $V$  von  $\mathbb{R}^3$  und  $W$  von  $\mathbb{R}^4$  mit Basen  $v_1 := (1; 0; -1)$ ,  $v_2 := (0; 1; -1)$  bzw.  $w_1 := (1; 0; 0; -1)$ ,  $w_2 := (0; 1; 0; -1)$  und  $w_3 := (0; 0; 1; -1)$ . Wir hatten gesehen, daß die lineare Abbildung  $f:V \rightarrow W$  mit  $f(a, b, c) := (a - 2b - c, 2a - b - c, -a - b, 6a - 2c)$ ,

(a,b,c ∈ ℝ) die Matrix A ( s.rechts) bezüglich v<sub>1</sub>,v<sub>2</sub> und w<sub>1</sub>,w<sub>2</sub>,w<sub>3</sub> hat.

Wir konstruieren zunächst invertierbare Matrizen S ∈ ℝ<sup>2×2</sup> und T ∈ ℝ<sup>3×3</sup>

mit der Eigenschaft, daß T<sup>-1</sup>AS die folgende Form hat( s.rechts).

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 0 \\ -1 & 1 \end{pmatrix} \quad T^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

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

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

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

$$\begin{pmatrix} 1 & -2/3 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -2/3 & 0 \\ 0 & 1/3 & 0 \\ -1 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

T<sub>i</sub>

A<sub>i</sub>

S<sub>i</sub>

Stets ist dabei T<sub>i</sub>AS<sub>i</sub>=A<sub>i</sub>. Wir setzen jetzt S =  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

und T = T<sub>5</sub><sup>-1</sup> =  $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 1 & -1 & 1 \end{pmatrix}$

Dann ist T<sup>-1</sup>AS=A<sub>5</sub>.

Folglich bilden v'<sub>1</sub>=0v<sub>1</sub>+(-1)v<sub>2</sub>=(0;-1;1) und v'<sub>2</sub>=1v<sub>1</sub>+0v<sub>2</sub>=(1;0;-1) eine Basis von V und w'<sub>1</sub>=1w<sub>1</sub>+0w<sub>2</sub>+1w<sub>3</sub>=(1;0;1;-2), w'<sub>2</sub>=2w<sub>1</sub>+3w<sub>2</sub>+(-1)w<sub>3</sub>=(2;3;-1;-4) und w'<sub>3</sub>=0w<sub>1</sub>+0w<sub>2</sub>+1w<sub>3</sub>=(0;0;1;-1) eine Basis von W mit f(v'<sub>1</sub>)=w'<sub>1</sub>, f(v'<sub>2</sub>)=w'<sub>2</sub> (Probe !)

Die Matrix von f bezüglich der neuen Basis v'<sub>1</sub>,v'<sub>2</sub> und w'<sub>1</sub>,w'<sub>2</sub>,w'<sub>3</sub> hat also die Form  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$

21.12.00

**8.7 Satz:** Gegeben seien endlich-dimensionale Vektorräume V,W und eine lineare Abbildung f:V→W bezüglich einer Basis b<sub>1</sub>,...,b<sub>m</sub> von V und einer Basis c<sub>1</sub>,...,c<sub>n</sub> von W. Dann ist rg f=rg(A).

**Bew.:** Nach 8.6 existieren Basen b'<sub>1</sub>,...,b'<sub>m</sub> von V und c'<sub>1</sub>,...,c'<sub>n</sub> von W mit der Eigenschaft, daß die Matrix von f bezüglich dieser Basen die folgende( s.rechts ) Form hat. Nach Satz 8.5 existieren invertierbare Matrizen S ∈ ℝ<sup>m×m</sup> und T ∈ ℝ<sup>n×n</sup> mit A'=TAS. mit r=rg(f). Daher ist rg f=r=rg A'=rg A.

**8.8 Satz:** (Im folgenden betrachten wir den Fall V=W)

Gegeben seien ein endlich-dimensionaler Vektorraum V und eine lineare Abbildung f:V→V mit Matrix A bezüglich einer Basis b<sub>1</sub>,...,b<sub>n</sub> von V und Matrix A' bezüglich einer Basis b'<sub>1</sub>,...,b'<sub>n</sub> von V. Wir schreiben b'<sub>j</sub>=s<sub>1j</sub>b<sub>1</sub>+...+s<sub>nj</sub>b<sub>n</sub> mit s<sub>ij</sub> ∈ ℝ. Dann ist S=(s<sub>ij</sub>) invertierbar mit A'=S<sup>-1</sup>AS.

**Bew.:** Schreibt man S<sup>-1</sup>=(s'<sub>ij</sub>), so ist b<sub>j</sub>=s'<sub>1j</sub>b'<sub>1</sub>+...+s'<sub>nj</sub>b'<sub>n</sub> (j=1;...;n)( vgl. Bem. 8.4). Die Behauptung folgt dann aus Satz 8.5.

**Bem.:** (I) Es ist daher det A'=(det S)<sup>-1</sup>(det A)(det S)=det A. Also hängt det A nicht von der Wahl der Basis ab. Man bezeichnet det A auch als Determinante von f und schreibt: det f:=det A.

(II) Rechenregeln für Determinanten von Matrizen übertragen sich leicht in Rechenregeln für Determinanten von linearen Abbildungen f:V→V. So ist z.B. det (gof)=(det g)(det f), det id<sub>v</sub>=1, det (rf)=r<sup>n</sup> det f (n=dim V) für r ∈ ℝ. Ferner gilt: f ist bijektiv ⇔ det f≠0.

(III) Es ist nicht immer möglich eine Basis von  $V$  so zu wählen, daß die Matrix von  $f$  bezüglich dieser Basis eine Diagonalmatrix ist, d.h. folgende Form hat:  $\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$   
 Ist eine solche Wahl möglich, so nennt man  $f$  diagonalisierbar.  
 Die Frage, ob eine vorgegebene lineare Abbildung diagonalisierbar ist oder nicht, werden wir ausführlich im folgenden Kapitel behandeln.

8.9 Def.: Man nennt Matrizen  $A, B \in \mathbb{R}^{n \times n}$  ähnlich und schreibt  $A \approx B$ , falls eine invertierbare Matrix  $S \in \mathbb{R}^{n \times n}$  mit  $B = SAS^{-1}$  existiert.

- Bem.: (I) Man zeigt leicht, daß  $\approx$  eine Äquivalenzrelation auf  $\mathbb{R}^{n \times n}$  ist.  
 (II) Nach Satz 8.8 sind Matrizen, die die gleiche lineare Abbildung  $f: V \rightarrow V$  bezüglich verschiedener Basen beschreiben, ähnlich.  
 (III) Wie stellt man fest, ob vorgegebene Matrizen  $A, B \in \mathbb{R}^{n \times n}$  ähnlich sind, oder nicht ?  
 Dafür gibt es Algorithmen, diese gehen aber über den Stoff einer einsemstrigen Vorlesung Lineare Algebra hinaus. (*Syntax in maple: issimilar(A,B)?*)

8.10 Def.: Man nennt eine Matrix  $A \in \mathbb{R}^{n \times n}$  diagonalisierbar, wenn  $A$  zu einer Diagonalmatrix ähnlich ist.

Bem.: (I) Gegeben seien ein endlich-dimensionaler Vektorraum  $V$  und eine lineare Abbildung  $f: V \rightarrow V$  mit Matrix  $A$  bezüglich einer Basis  $b_1, \dots, b_n$  von  $V$ . Nach Satz 8.8 gilt dann:  
 $f$  ist diagonalisierbar  $\Leftrightarrow A$  ist diagonalisierbar.

(II) Im nächsten Kapitel werden wir uns damit beschäftigen, wann eine vorgegebene Matrix diagonalisierbar ist.

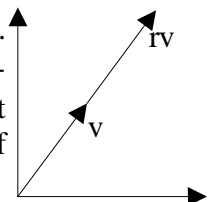
(III) Ist  $A$  zu einer Diagonalmatrix  $\begin{pmatrix} d_1 & 0 \\ 0 & d_n \end{pmatrix}$  ähnlich, so existiert eine invertierbare Matrix  $S$  mit  $A = SDS^{-1}$ .  
 Dann ist  $A^2 = SD(S^{-1}S)DS^{-1} = SD^2S^{-1}$  mit  $D^2 = \begin{pmatrix} d_1^2 & 0 \\ 0 & d_n^2 \end{pmatrix}$ , für  $A^3$  ergibt sich dann also  $A^3 = SD^3(S^{-1}S)DS^{-1} = SD^3S^{-1}$ .

Für  $k \in \mathbb{N}$  ist also  $A^k = SD^kS^{-1}$  mit  $D^k = \begin{pmatrix} d_1^k & 0 \\ 0 & d_n^k \end{pmatrix}$

Auf diese Weise kann man die Potenzen von  $A$  schnell berechnen.

## 9. Eigenwerte und Eigenvektoren

9.1 Def.: Gegeben seien ein Vektorraum  $V$  und eine lineare Abbildung  $f: V \rightarrow V$ . Man nennt  $r \in \mathbb{R}$  Eigenwert von  $f$ , falls ein  $v \in V \setminus \{0\}$  mit  $f(v) = rv$  existiert. Gegebenfalls nennt man  $v$  einen Eigenvektor von  $f$  zum Eigenwert  $r$ . Man bezeichnet  $E_r(f) = \{v \in V \mid f(v) = rv\} = \{v \in V \mid (f - r \cdot \text{id}_V)(v) = 0\} = \text{Ker}(f - r \cdot \text{id}_V)$  als Eigenraum von  $f$  zum Eigenwert  $r$ .



Bem.:  $E_r(f)$  ist also ein Untervektorraum von  $V$ , insbesondere ist  $0 \in E_r(f)$ . Man beachte aber, daß  $0$  nicht als Eigenvektor zählt.

Bsp.: (I) Sei  $V = \mathbb{R}^2$  und  $f(a,b) = (b,a)$  für  $a, b \in \mathbb{R}$ . Dann ist  $(1;1)$  ein Eigenvektor von  $f$  zum Eigenwert  $1$ , und  $(1;-1)$  ist Eigenvektor von  $f$  zum Eigenwert  $-1$ .

(II) Sei  $V = \mathbb{R}^2$  und  $f(a,b) = (b,-a)$  für  $a, b \in \mathbb{R}$ .

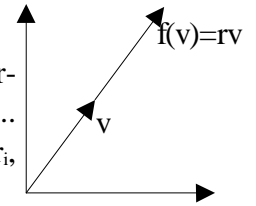
Ist  $(a,b) \in V$  Eigenvektor von  $f$  zum Eigenwert  $r \in \mathbb{R}$ , so ist  $f(f(a,b)) =$

$f(r(a,b)) = rf(a,b) = r^2(a,b) = (r^2a, r^2b)$  und  $f(f(a,b)) = f(b,-a) = (-a,-b)$ ;

also ist  $r^2 = -1$  wegen  $a \neq 0$  oder  $b \neq 0$ . Dies ist unmöglich. Daher besitzt  $f$  weder Eigenwerte noch Eigenvektoren.

10.01.01

Eigenraum  $E_r(f) = \{v \in V \mid f(v) = rv\}$  ist Untervektorraum von  $V$ .



Satz: Gegeben seien ein Vektorraum  $V$ , eine lineare Abbildung  $f: V \rightarrow V$  und paarweise verschiedene Eigenwerte  $r_1, \dots, r_k$  von  $f$ . Dann ist  $E_{r_1}(f) + \dots + E_{r_k}(f) = E_{r_1}(f) \oplus \dots \oplus E_{r_k}(f)$ . Hat man also für  $i=1, \dots, k$  einen Eigenvektor,  $u_i$  von  $f$  zum Eigenwert  $r_i$ , so sind  $u_1, \dots, u_k$  linear unabhängig.

Bew.: Zum Beweis der ersten Aussage argumentieren wir mit Induktion nach  $k$ .

Im Fall  $k=1$  ist nichts zu zeigen. Sei jetzt  $k > 1$  und bereits gezeigt, daß  $E_{r_1}(f) + \dots + E_{r_{k-1}}(f) = E_{r_1}(f) \oplus \dots \oplus E_{r_{k-1}}(f)$  ist. Dazu sei  $v \in E_{r_k}(f) \cap (E_{r_1}(f) \oplus \dots \oplus E_{r_{k-1}}(f))$ . Wir schreiben  $v = v_1 + \dots + v_{k-1}$  mit  $v_1 \in E_{r_1}(f), \dots, v_{k-1} \in E_{r_{k-1}}(f)$ . Dann ist  $f(v) = f(v_1 + \dots + v_{k-1}) = f(v_1) + \dots + f(v_{k-1}) = r_1 v_1 + \dots + r_{k-1} v_{k-1}$  und  $f(v) = r_k v = r_k v_1 + \dots + r_k v_{k-1}$ .

Aus der Induktionsvoraussetzung folgt:  $r_1 v_1 = r_k v_1, \dots, r_{k-1} v_{k-1} = r_k v_{k-1}$ ; d.h.  $0 = \underbrace{(r_1 - r_k)}_{\neq 0} v_1 = \dots = \underbrace{(r_{k-1} - r_k)}_{\neq 0} v_{k-1}$ .

Also ist  $v_1 = \dots = v_{k-1} = 0$ . Damit ist gezeigt:  $E_{r_k}(f) \cap (E_{r_1}(f) \oplus \dots \oplus E_{r_{k-1}}(f)) = \{0\}$ . Aus Satz 5.8 folgt,  $E_{r_1}(f) + \dots + E_{r_{k-1}}(f) + E_{r_k}(f) = E_{r_1}(f) \oplus \dots \oplus E_{r_{k-1}}(f) \oplus E_{r_k}(f)$ . Seien  $a_1, \dots, a_k \in \mathbb{R}$  mit  $\underbrace{a_1 u_1 + \dots + a_k u_k}_{\in E_{r_1}(f) \oplus \dots \oplus E_{r_k}(f)} = 0$ . Nach dem ersten Teil des Beweises ist  $a_i u_i = 0$  für  $i=1, \dots, k$ . Also ist  $a_1 = \dots = a_k = 0$ . (da  $u_1, \dots, u_k \neq 0$ )

9.2 Bem.: Aus dem Satz 9.1 folgt:

$$\dim V \geq \dim [E_{r_1}(f) \oplus \dots \oplus E_{r_k}(f)] = \underbrace{\dim E_{r_1}(f)}_{\geq 1} + \dots + \underbrace{\dim E_{r_k}(f)}_{\geq 1} \geq k.$$

Daher kann  $f$  im Fall  $n = \dim V \leq \infty$  höchstens  $n$  verschiedene Eigenwerte haben.

Satz: Gegeben seien ein endlich-dimensionaler Vektorraum  $V$  mit Basis  $b_1, \dots, b_n$  und eine lineare Abbildung  $f: V \rightarrow V$  mit Matrix  $A$  bezüglich dieser Basis. Für  $r \in \mathbb{R}$  gilt dann:

$$\boxed{r \text{ ist Eigenwert von } f \Leftrightarrow \det(r \cdot 1_n - A) = 0.}$$

Bew.:  $r$  ist Eigenwert von  $f \Leftrightarrow E_r(f) \neq \{0\} \Leftrightarrow \text{Ker}(\text{rid}_V - f) \neq \{0\} \Leftrightarrow^{7.2} r \cdot \text{id}_V - f$  ist nicht injektiv  $\Leftrightarrow r \cdot \text{id}_V - f$  ist nicht bijektiv  $\Leftrightarrow^{7.3} r \cdot 1_n - A$  ist nicht invertierbar  $\Leftrightarrow \det(r \cdot 1_n - A) = 0$ .

Bsp.: Sei  $V = \mathbb{R}^2$  und  $f(a,b) := (3a+b, a+3b)$  für  $a, b \in \mathbb{R}$ . Für  $x \in \mathbb{R}$  ist  $\det(x \cdot 1_n - A) = \begin{vmatrix} x-3 & -1 \\ -1 & x-3 \end{vmatrix} = (x-3)^2 - 1 = x^2 - 6x + 8 \Rightarrow$  Polynom, Nullstellen ausrechnen  $\Rightarrow (x-4)(x-2) \Rightarrow x=4 \vee x=2$ .

Dabei gilt  $x^2 - 6x + 8 = 0 \Leftrightarrow x \in \{2, 4\}$ . Daher sind 2 und 4 die einzigen Eigenwerte von  $f$ . Außerdem gilt:  $(a,b) \in E_2(f) \Leftrightarrow (3a+b, a+3b) = 2(a,b) \Leftrightarrow a+b=2 \Leftrightarrow a=-b$

$(a,b) \in E_4(f) \Leftrightarrow (3a+b, a+3b) = 4(a,b) \Leftrightarrow a-b=0 \Leftrightarrow a=b$

Daher ist  $E_2(f) = \mathbb{R}(1; -1)$  und  $E_4(f) = \mathbb{R}(1; 1)$ .

Nach Satz 9.1 (oder direkt) sind  $(1; -1)$  und  $(1; 1)$  linear unabhängig, bilden also eine Basis von  $V$ . Die Matrix von  $f$  bezüglich  $(1; -1), (1; 1)$  ist die Diagonalmatrix  $\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$ .

9.3 Bem.: Für  $A \in \mathbb{R}^{n \times n}$  nennt man die Abbildung  $\mathbb{R} \rightarrow \mathbb{R}, x \rightarrow \det(x 1_n - A)$  das charakteristische Polynom von  $A$ . Nach dem Entwicklungssatz ist es von der Form  $x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$  ( $a_0, \dots, a_{n-1} \in \mathbb{R}$ ).

Es hat also den Grad  $n$  und ist normiert, d.h. der Koeffizient  $a_n$  von  $x^n$  ist gleich 1. Die Nullstellen des charakteristischen Polynoms von  $A$  nennt man auch die Eigenwerte von  $A$ .

Ist also  $V$  ein endlich-dimensionaler Vektorraum und  $f: V \rightarrow V$  eine lineare Abbildung mit Matrix  $A$  bezüglich einer Basis von  $V$ , so sind die Eigenwerte von  $f$  genau die Eigenwerte von  $A$ .

Bei größeren Matrizen hat das charakteristische Polynom großen Grad und es ist i.d.R. unmöglich seine Nullstellen exakt zu bestimmen. Hat man aber die Eigenwerte erst einmal gefunden, so ist die Bestimmung der Eigenräume vergleichsweise einfach. Wegen  $E_r(f) = \text{Ker}(\text{rid}_V - f)$  führt dies

nun wieder auf die Lösung eines homogenen linearen Gleichungssystems.

**Satz:** Ähnliche Matrizen besitzen immer das gleiche charakteristische Polynom( und damit auch die gleichen Eigenwerte)

**Bew.:** Es seien  $A, B \in \mathbb{R}^{n \times n}$  ähnlich. Es existiert eine invertierbare  $n \times n$ -Matrix  $S$  mit  $B = SAS^{-1}$ . Für  $x \in \mathbb{R}$  ist  $|x1_n - B| = |xSS^{-1} - SAS^{-1}| = |S(x1_n - A)S^{-1}| = |S| \cdot |x1_n - A| \cdot |S^{-1}| = |x1_n - A|$ .

**9.4 Bem.:** Sei  $V$  ein endlich-dimensionaler Vektorraum und  $f: V \rightarrow V$  eine lineare Abbildung mit Matrix  $A$  bezüglich einer Basis  $b_1, \dots, b_n$  von  $V$ . Dann hängt das charakteristische Polynom von  $A$  nur von  $f$  ab, aber nicht von der Wahl der Basis  $b_1, \dots, b_n$ . (Wählt man eine andere Basis  $c_1, \dots, c_n$  von  $V$ , so ist die Matrix  $B$  von  $f$  bezüglich  $c_1, \dots, c_n$  zu  $A$  ähnlich, besitzt also nach Satz 9.3 das gleiche charakteristische Polynom). Man nennt dieses Polynom dann auch das charakteristische Polynom von  $f$ .

**Satz:** Für einen endlich-dimensionalen Vektorraum  $V$  und eine lineare Abbildung  $f: V \rightarrow V$  sind äquivalent:

- (1)  $f$  ist diagonalisierbar.
- (2)  $V$  besitzt eine Basis, die aus Eigenvektoren von  $f$  besteht.
- (3)  $V$  ist die Summe der Eigenräume von  $f$ .
- (4)  $V$  ist die direkte Summe der Eigenräume von  $f$ .
- (5)  $\dim V$  ist die Summe der Dimensionen der Eigenräume von  $f$ .

**Bew.:** (1) $\Rightarrow$ (2) Sei  $f$  diagonalisierbar. Dann existiert eine Basis  $b_1, \dots, b_n$  von  $V$  mit der Eigenschaft, daß die Matrix von  $f$  bezüglich  $b_1, \dots, b_n$  die Form wie rechts hat. Für  $i=1; \dots; n$  ist  $f(b_i) = a_i b_i$ , d.h.  $b_i$  ist Eigenvektor von  $f$  zum Eigenwert  $a_i$ .

$$A = \begin{pmatrix} a_1 & 0 \\ 0 & a_n \end{pmatrix}$$

(2) $\Rightarrow$ (3) Sei  $b_1, \dots, b_n$  eine Basis von  $V$  und jedes  $b_i$  sei Eigenvektor von  $f$  zum Eigenwert  $a_i$ . Ist  $v \in V$  beliebig, so existieren  $r_1, \dots, r_n \in \mathbb{R}$  mit  $v = r_1 b_1 + \dots + r_n b_n$ . Also ist  $V = E_{a_1}(f) + \dots + E_{a_n}(f)$ .

(3) $\Rightarrow$ (4) Siehe Satz 9.1  $E_{a_1}(f) \in E_{a_n}(f)$

(4) $\Rightarrow$ (5) Aus  $V = E_{a_1}(f) \oplus \dots \oplus E_{a_n}(f)$  folgt:  $\dim V = \dim E_{a_1}(f) + \dots + \dim E_{a_n}(f)$ .

(5) $\Rightarrow$ (1) Seien  $r_1, \dots, r_n$  die paarweise verschiedenen Eigenwerte von  $f$ , und sei  $\dim V = \dim E_{r_1}(f) + \dots + \dim E_{r_n}(f) = \sum_{i=1}^n \dim(E_{r_i}(f) \oplus \dots \oplus E_{r_n}(f))$ . Dann ist  $V = E_{r_1}(f) \oplus \dots \oplus E_{r_n}(f)$ . Wir wählen eine Basis  $b_1, \dots, b_{s_1}$  von  $E_{r_1}(f)$ , eine Basis  $b_{s_1+1}, \dots, b_{s_2}$  von  $E_{r_2}(f)$ , usw. Dann bilden  $b_1, \dots, b_{s_1}, b_{s_1+1}, \dots, b_{s_2}, \dots$  eine Basis von  $V$  und bezüglich dieser Basis hat die Matrix von  $f$  die Form

$$A = \begin{pmatrix} r_1 & & & \\ & r_1 & & \\ & & r_2 & \\ & & & r_2 & \\ & & & & \dots & \dots \end{pmatrix}$$

**Bsp.:** (I) Sei  $V = \mathbb{R}^2$  und  $f(a,b) = (a, a+b)$  für  $a, b \in \mathbb{R}$ . Die Matrix von  $f$  bezüglich der Standardbasis von  $V$  ist  $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Das charakteristische Polynom von  $A$  ist  $|x-1 \ 0; -1 \ x-1| = (x-1)^2$ .

Daher ist 1 der einzige Eigenwert von  $A$  und  $f$ . Ferner gilt:

$$(a,b) \in E_1(f) \Leftrightarrow 0 = (0, a) \Leftrightarrow a = 0.$$

Folglich ist  $E_1(f) = \mathbb{R}(0;1)$ , insbesondere ist  $\dim E_1(f) = 1 < 2 = \dim V$ , d.h.  $f$  und  $A$  sind nicht diagonalisierbar.

11.01.01

(II)  $V = \mathbb{R}^4$ ,  $f: V \rightarrow V$   $f(a,b,c,d) = (a-b+d, b, -a-b+2c+d, -b+2d)$ ,  $a, b, c, d \in \mathbb{R}$ . Die Matrix von  $f$  bezüglich der Standardbasis von  $V$  ist( s. nächste Seite). Für  $x \in \mathbb{R}$  ist  $\det(x1_n - A) = \dots = (x-1)^2(x-2)^2$ . Daher sind 1 und 2 die einzigen Eigenwerte von  $f$ .

$A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ -1 & -1 & 2 & 1 \\ 0 & -1 & 0 & 2 \end{pmatrix}$  Wir berechnen zunächst  $E_1(f) = \text{Ker}(\text{id}_V - f)$ :  $1_n - A = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$  Gauß-Algorithmus  $\rightarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Daher bilden  $b_1 = (0; 1; 0; 1)$  und  $b_2 = (1; 0; 1; 0)$  eine Basis von  $E_1(f)$ .

$2 \cdot 1_n - A = \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$  Gauß-Algorithmus  $\rightarrow \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Daher bilden  $b_3 = (1; 0; 0; 1)$  und  $b_4 = (0; 0; 1; 0)$  eine Basis von  $E_2(f)$ . Folglich bilden  $b_1, b_2, b_3, b_4$  eine Basis von  $V$  und die Matrix von  $f$  bezüglich dieser Basis  $b_1, b_2, b_3, b_4$  ist

$\begin{pmatrix} 1 & & & \\ & 1 & & \\ 0 & & 2 & \\ & & & 2 \end{pmatrix}$ . Insbesondere ist  $f$  diagonalisierbar.

**9.5 Satz:** Gegeben sei ein Vektorraum  $V$  der Dimension  $n < \infty$  und eine lineare Abbildung  $f: V \rightarrow V$ . Besitzt die Abbildung  $f$   $n$  verschiedene Eigenwerte, so ist  $f$  diagonalisierbar.

**Bew.:** Offenbar ist  $\dim V \geq \dim E_{r_1}(f) \oplus \dots \oplus \dim E_{r_n}(f) = \underbrace{\dim E_{r_1}(f)}_{\geq 1} + \dots + \underbrace{\dim E_{r_n}(f)}_{\geq 1} \geq n$ . Nach Satz 9.4 ist  $f$  diagonalisierbar.

**Bsp.:** Wie in Aufgabe 6 definiert man für  $n \in \mathbb{N}_0$  die  $n$ -te Fibonacci-Zahl  $F_n$  durch  $F_{n+1} = F_n + F_{n-1}$ ,  $F_0 = 0, F_1 = 1, n \in \mathbb{N}$ . Die ersten Glieder dieser Kette sind demnach  $0; 1; 1; 2; 3; 5; 8; 13; 21; \dots$

Wie findet man die geschlossene Formel für  $F_n$  aus Teil (III) von Aufgabe 6?

Zunächst beobachtet man  $\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = \dots = A^n \begin{pmatrix} F_0 \\ F_1 \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Um die Potenzen von  $A$  zu berechnen, werden wir  $A$  diagonalisieren. Das charakteristische Polynom von  $A$  ist  $\begin{vmatrix} x & -1 \\ 1 & x-1 \end{vmatrix} = x(x+1) - 1$ .

Die Nullstellen sind  $r = \frac{1+\sqrt{5}}{2}$  und  $s = \frac{1-\sqrt{5}}{2}$ . Nach Satz 9.5 ist  $A$  diagonalisierbar. Wir berechnen die entsprechenden Eigenräume:  $\begin{pmatrix} r & -1 \\ -1 & r+1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow rx + y = 0 \wedge -x + (r-1)y = 0$ , daraus folgt  $y = rx$ . Daher ist  $E_r(f) = \mathbb{R} \begin{pmatrix} 1 \\ r \end{pmatrix}$ . Analog ist  $E_s(f) = \mathbb{R} \begin{pmatrix} 1 \\ s \end{pmatrix}$ . Wir erhalten:  $A = \begin{pmatrix} 1 & 1 \\ r & s \end{pmatrix} = \begin{pmatrix} rI & sI \\ rr & ss \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ r & s \end{pmatrix} \begin{pmatrix} r & 0 \\ 0 & s \end{pmatrix} \begin{pmatrix} 1 & 1 \\ r & s \end{pmatrix}^{-1}$  mit  $\begin{pmatrix} 1 & 1 \\ r & s \end{pmatrix}^{-1} = \frac{1}{s-r} \begin{pmatrix} s & -1 \\ -r & 1 \end{pmatrix}$ . Für  $n \in \mathbb{N}$  ist also  $A^n = \begin{pmatrix} 1 & 1 \\ r & s \end{pmatrix} \begin{pmatrix} r & 0 \\ 0 & s \end{pmatrix}^n \begin{pmatrix} 1 & 1 \\ r & s \end{pmatrix}^{-1} = \begin{pmatrix} r^n & s^n \\ r^{n+1} & s^{n+1} \end{pmatrix} \frac{1}{s-r} \begin{pmatrix} s & -1 \\ -r & 1 \end{pmatrix}$   
 $= \frac{1}{\sqrt{5}} \begin{pmatrix} r^n s - s^n r & s^n - r^n \\ r^{n+1} s - s^{n+1} r & s^{n+1} - r^{n+1} \end{pmatrix}$  und  $\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} r^n - s^n \\ r^{n+1} - s^{n+1} \end{pmatrix}$ . Also ist stets  $F_n = \frac{1}{\sqrt{5}} \left[ \frac{1+\sqrt{5}}{2}^n - \frac{1-\sqrt{5}}{2}^n \right]$ .

In ähnlicher Weise kann man andere rekursiv definierte Folgen behandeln.

**9.6 Bem.:** (vergleiche Aufgabe 7)

Eine Maus befindet in einem Labyrinth (s. Aufgabe zur Form des Labyrinths). Für  $i = 1; \dots; 4$  sei  $a_{ij}$  die Übergangswahrscheinlichkeit von Zelle  $i$  nach Zelle  $j$  nach einer Minute. Die Matrix der

Übergangswahrscheinlichkeit habe die Form  $A = \begin{pmatrix} 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$ . Für  $n \in \mathbb{N}$  gibt  $A^n$  die Übergangswahrscheinlichkeit nach  $n$  Minuten an. Wie ist das Verhalten der Maus nach langer Zeit (d.h. für  $n \rightarrow \infty$ )? Um die Potenzen von  $A$  zu berechnen diagonalisieren wir  $A$  zunächst: Das charakteristische Polynom von  $A$  ist  $x^4 - 3/4x^3 + 3/4x^2 + 1/16x - 1/16 = (x-1)(x+1/4)(x-1/2)^2$ . Die entspre-

chenden Eigenräume sind  $E_1=\mathbb{R}(1;1;1;1)$ ,  $E_{-1/4}=\mathbb{R}(2;2;2;-3)^T$ ,  $E_{1/2}=\mathbb{R}(1;-1;0;0)^T+\mathbb{R}(-1;0;1;0)$ .

Daher ist  $A \begin{pmatrix} 1 & 2 & -1 & -1 \\ 1 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & -3 & 0 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -1/2 & -1/2 & -1/2 \\ 1 & -1/2 & 1/2 & 0 \\ 1 & -1/2 & 0 & 1/2 \\ 1 & 3/4 & 0 & 0 \end{pmatrix}}_{=S} = \underbrace{\begin{pmatrix} 1 & 2 & -1 & -1 \\ 1 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & -3 & 0 & 0 \end{pmatrix}}_{=D} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1/4 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \end{pmatrix}$ , d.h.  $A=SDS^{-1}$

und  $A^n=SD^nS^{-1}$  für  $n \in \mathbb{N}$ . Jetzt verwenden wir ein wenig Analysis und erhalten:

$\lim A^n=S \cdot \lim D^nS^{-1}$  mit  $\lim D^n= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ . Daher ist  $\lim A^n= \begin{pmatrix} 1 & 2 & -1 & -1 \\ 1 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & -3 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

$\begin{pmatrix} 1/5 & 1/5 & 1/5 & 2/5 \\ 1/15 & 1/15 & 1/15 & -1/5 \\ -1/3 & 2/3 & -1/3 & 0 \\ -1/3 & -1/3 & 2/3 & 0 \end{pmatrix} = \begin{pmatrix} 1/5 & 1/5 & 1/5 & 2/5 \\ 1/5 & 1/5 & 1/5 & 2/5 \\ 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 2/5 \end{pmatrix}$ . Asymptotisch hält sich die Maus mit der Wahrscheinlichkeit  $1/5$  in jeder der Zellen 1;2;3 und mit der Wahrscheinlichkeit  $2/5$  in Zelle 4 auf und zwar unabhängig von der Ausgangszelle.

In ähnlicher Weise kann man ökonomische Prozesse oder auch Wachstumsprozesse in der Biologie behandeln.

### 10. Euklidische Vektorräume

**10.1 Def.:** Ein Skalarprodukt auf einem reellen Vektorraum  $V$  ist eine Abbildung  $V \times V \rightarrow \mathbb{R}$ ,  $(v,w) \rightarrow (v|w)$  mit folgenden Eigenschaften:

- (I)  $(v+v'|w)=(v|w)+(v'|w)$  ( $v,v',w \in V$ )
- (II)  $(av|w)=a(v|w)$  ( $v,w \in V, a \in \mathbb{R}$ )
- (III)  $(v|w)=(w|v)$  ( $v,w \in V$ )
- (IV)  $(v|v) > 0$  für  $v \in V \setminus \{0\}$

**Bem.:** Aus (I), (II), (III) folgt sofort:

- (V)  $(v|w+w')=(v|w)+(v|w')$  ( $v,w',w \in V$ )
- (VI)  $(v|aw)=a(v|w)$  ( $v,w \in V, a \in \mathbb{R}$ )

Wegen (I) und (III) ist für  $w \in V$  die Abbildung  $f_w:V \rightarrow \mathbb{R}$ ,  $v \rightarrow (v|w)$  linear. Daher ist  $f_w(0)=0$ , d.h.

- (VII)  $(0|w)=0$  ( $w \in V$ )
- Aus (III) folgt (VIII)  $(v|0)=0$  ( $v \in V$ )

**Bsp.:** Für  $n \in \mathbb{N}$  ist die Abbildung  $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $(x,y) \rightarrow (x|y)$  mit  $(x|y)=x_1y_1+\dots+x_ny_n$  für  $x=(x_1,\dots,x_n)$ ,  $y=(y_1,\dots,y_n) \in \mathbb{R}^n$  ein Skalarprodukt, man nennt es das Standardskalarprodukt auf  $\mathbb{R}^n$ .

(II) Seien  $a,b \in \mathbb{R}$  mit  $a < b$  und sei  $\mathcal{C}[a,b]$  der reelle Vektorraum aller stetigen( continuous) Funktionen auf dem abgeschlossenen Intervall  $[a,b]=\{x \in \mathbb{R} \mid a \leq x \leq b\}$ . Dann wird durch  $(f|g)=\int_a^b f(t)g(t)dt$  ein Skalarprodukt auf  $\mathcal{C}[a,b]$  definiert.

Man nennt es das Standardskalarprodukt auf  $\mathcal{C}[a,b]$ ( Mehr darüber in der Analysis).

17.01.01

**Satz:** (Ungleichung von Cauchy und Schwarz)

Gegeben sei ein Skalarprodukt auf einem reellen Vektorraum  $V$ . Für alle  $x,y \in V$  ist dann  $(x|y)^2 \leq (x|x)(y|y)$ . Das Gleichheitszeichen tritt genau dann auf, wenn  $x$  und  $y$  linear abhängig sind.

**Bew.:** Im Fall  $y=0$  steht auf beiden Seiten 0. Sei  $y \neq 0$  und  $c = \frac{(x|y)}{(y|y)}$  Dann gilt:

$0 \leq (x-cy|x-cy)=(x|x)-c(x|y)-c(y|x)+c^2(y|y)=(x|x)-2c(x|y)+c^2(y|y)=(x|x)-2 \frac{(x|y)^2}{(y|y)} + \frac{(x|y)^2}{(y|y)} (y|y)$   
 $= (x|x) - \frac{(x|y)^2}{(y|y)}$

Multiplikation mit  $(y|y)$  ergibt:  $0 \leq (x|x)(y|y) - (x|y)^2$ .

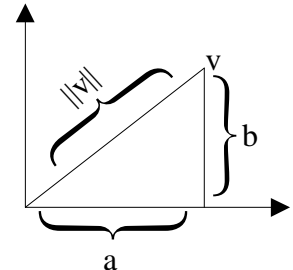
Damit ist die Ungleichung bewiesen.

Im Fall  $(x|x)(y|y) = (x|y)^2$  folgt  $0 = (x - cy|x - cy)$ , d.h.  $x - cy = 0$ . Also sind  $x$  und  $y$  linear abhängig. Sei umgekehrt  $x = dy$  für ein  $d \in \mathbb{R}$ . Dann ist  $(x|y)^2 = d^2(y|y)^2 = (x|x)(y|y)$ .

**10.2 Def.:** Gegeben sei ein reeller Vektorraum  $V$  mit Skalarprodukt  $V \times V \rightarrow \mathbb{R}, (v, w) \rightarrow (v|w)$ . Man spricht dann auch von einem euklidischen Vektorraum. Für  $v \in V$  nennt man  $\|v\| = \sqrt{(v|v)}$  die Länge oder den Betrag oder die Norm von  $v$ .

**Bsp.:** In  $\mathbb{R}^2$  mit dem Standardskalarprodukt gilt:

$$v = (a, b) \Rightarrow \|v\| = \sqrt{a^2 + b^2}, \quad \|v\|^2 = a^2 + b^2 \text{ (siehe rechts).}$$



**Satz:** In jedem euklidischen Vektorraum  $V$  gilt:

(I)  $\|v\| \geq 0$  für alle  $v \in V$

(II)  $\|v\| = 0 \Leftrightarrow v = 0$

(III)  $\|av\| = |a| \cdot \|v\|$  ( $a \in \mathbb{R}, v \in V$ )

(IV)  $\|v+w\| \leq \|v\| + \|w\|$  ( $v, w \in V$ )

(Dreiecksungleichung)

**Bew.:** (I) folgt aus der Definition.

(II)  $\|v\| = 0 \Leftrightarrow (v|v) = 0 \Leftrightarrow v = 0$

(III)  $\|av\| = \sqrt{(av|av)} = \sqrt{a^2(v|v)} = \sqrt{a^2} \sqrt{(v|v)} = |a| \cdot \|v\|$

(IV)  $\|v+w\|^2 = (v+w|v+w) = (v|v) + (v|w) + (w|v) + (w|w) = (v|v) + 2(v|w) + (w|w) \leq (v|v) + 2|(v|w)| + (w|w) \leq_{CSU} \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$ .

**Bem.:** Im Fall  $\|v+w\| = \|v\| + \|w\|$  folgt aus obigem Beweis  $|(v|w)| = \|v\| \cdot \|w\|$ . Aus der CSU folgt also, daß  $v$  und  $w$  linear abhängig sind. Im Fall  $w \neq 0$  ist also  $v = cw$  für ein  $c \in \mathbb{R}$ . Folglich ist  $c(w|w) = (v|w) = |(v|w)| = |c|(w|w)$ , d.h.  $c \geq 0$ . Es gilt also:  $\|v+w\| = \|v\| + \|w\| \Leftrightarrow w = 0$  oder  $v = cw$  für ein  $c \geq 0$ .

**10.3 Satz:** Für alle Elemente  $v, w$  eines euklidischen Vektorraumes  $V$  gilt:

(I)  $\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2(v|w)$  (Satz des Pythagoras)

(II)  $\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2$  (Parallelogrammgleichung)

**Bew.:** (I)  $\|v+w\|^2 = (v+w|v+w) = (v|v) + (v|w) + (w|v) + (w|w) = \|v\|^2 + 2(v|w) + \|w\|^2$ .

(II)  $\|v+w\|^2 + \|v-w\|^2 = (v|v) + 2(v|w) + (w|w) + (v|v) - 2(v|w) + (w|w) = 2\|v\|^2 + 2\|w\|^2$ .

**10.4 Def.:** Für Elemente  $v, w$  eines euklidischen Vektorraumes  $V$  nennt man  $d(v, w) := \|v - w\|$  den Abstand oder die Distanz zwischen  $v$  und  $w$ . Die auf diese Weise definierte Abbildung  $V \times V \rightarrow \mathbb{R}$  nennt man die Metrik von  $V$ .

**Satz:** Für Elemente  $x, y, z$  eines euklidischen Vektorraumes  $V$  gilt stets:

(I)  $d(x, y) = d(y, x)$

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

(III)  $d(x, y) \geq 0$

(IV)  $d(x, y) = 0 \Leftrightarrow x = y$

**Bew.:** (I)  $d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1| \cdot \|y - x\| = d(y, x)$ .

(II)  $d(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$ .

(III) klar.

(IV)  $d(x, y) = 0 \Leftrightarrow \|x - y\| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$ .

10.5 Def.: Elemente  $v, w$  eines euklidischen Vektorraumes  $V$  mit  $(v|w)=0$  nennt man senkrecht oder orthogonal. Gegebenenfalls schreibt man  $v \perp w$ . Man nennt von 0 verschiedene Vektoren  $v_1, \dots, v_n$  in  $V$

(I) ein Orthogonalsystem, falls  $v_i \perp v_j$  für alle  $i \neq j$  gilt.

(II) ein Orthonormalsystem, falls zusätzlich  $\|v_1\| = \dots = \|v_n\| = 1$  gilt.

(III) eine Orthonormalbasis, falls  $v_1, \dots, v_n$  ein Orthonormalsystem und eine Basis von  $V$  bilden.

Satz: Jedes Orthogonalsystem  $v_1, \dots, v_n$  in einem euklidischen Vektorraum  $V$  ist linear unabhängig.

Bew.: Seien  $a_1, \dots, a_n \in \mathbb{R}$  mit  $a_1 v_1 + \dots + a_n v_n = 0$ . Für  $k=1; \dots; n$  ist dann:

$$0 = (0|v_k) = a_1(v_1|v_k) + \dots + a_n(v_n|v_k) = a_k \underbrace{(v_k|v_k)}_{\neq 0}, \text{ d.h. } a_k = 0.$$

Bem.: Für jede Orthonormalbasis  $b_1, \dots, b_n$  in einem euklidischen Vektorraum  $V$  und beliebige Zahlen  $r_1, s_1, \dots, r_n, s_n \in \mathbb{R}$  gilt:

$$(r_1 b_1 + \dots + r_n b_n | s_1 b_1 + \dots + s_n b_n) = r_1 s_1 (b_1|b_1) + \dots + r_n s_n (b_n|b_n) = r_1 s_1 + \dots + r_n s_n.$$

10.6 Satz: ( Orthonormalisierungsverfahren von Gram-Schmidt ) !

Gegeben sei eine Basis  $a_1, \dots, a_n$  eines endlich-dimensionalen euklidischen Vektorraumes  $V$ . Dann existiert eine Orthonormalbasis  $b_1, \dots, b_n$  von  $V$  mit der Eigenschaft, daß jedes  $b_i$  eine Linearkombination von  $a_1, \dots, a_i$  ist.

Bew.: Wir setzen zunächst  $b_1 := \frac{a_1}{\|a_1\|}$ . Dann ist  $\|b_1\|=1$ .

Sind bereits Vektoren  $b_1, \dots, b_k$  für ein  $k < n$  konstruiert, so setzen wir  $c_{k+1} := a_{k+1} - (a_{k+1}|b_1)b_1 - \dots - (a_{k+1}|b_k)b_k$ .

Für  $j=1; \dots; k$  gilt dann:  $(c_{k+1}|b_j) = (a_{k+1}|b_j) - (a_{k+1}|b_1)(b_1|b_j) - \dots - (a_{k+1}|b_k)(b_k|b_j) = (a_{k+1}|b_j) - (a_{k+1}|b_j) = 0$ .

Daher bilden  $b_1, \dots, b_k, c_{k+1}$  eine Orthogonalsystem. Setzt man  $b_{k+1} := \frac{c_{k+1}}{\|c_{k+1}\|}$ , so bilden  $b_1, \dots, b_k, b_{k+1}$  ein

Orthonormalsystem. Nach Satz 10.5 bilden  $b_1, \dots, b_n$  eine Orthonormalbasis von  $V$ .

18.01.01

10.7 Satz: Jedes Orthogonalsystem  $a_1, \dots, a_m$  in einem endlich-dimensionalen Vektorraum  $V$  kann man zu einer Orthonormalbasis von  $V$  ergänzen.

Bew.: Nach Satz 10.5 sind  $a_1, \dots, a_m$  linear unabhängig. Wir ergänzen  $a_1, \dots, a_m$  zu einer Basis  $a_1, \dots, a_m, a_{m+1}, \dots, a_n$  von  $V$  und wenden darauf das Orthonormalisierungsverfahren von Gram-Schmidt an. Wir erhalten eine Orthonormalbasis  $b_1, \dots, b_n$  von  $V$ . Dabei gilt nach Konstruktion:

$b_1 = a_1, \dots, b_m = a_m$  (bereits orthogonale Vektoren bleiben erhalten).

10.8 Def.: Für jede Teilmenge  $M$  eines euklidischen Vektorraumes  $V$  nennt man  $M^\perp = \{v \in V | v \perp m \text{ für alle } m \in M\} = \{v \in V | (v|m)=0 \text{ für alle } m \in M\}$  das orthogonale Komplement von  $M$  in  $V$ .

Bem.: Man rechnet leicht nach, daß  $M^\perp$  ein Untervektorraum von  $V$  ist.

Satz: Für jeden Untervektorraum  $U$  eines endlich-dimensionalen Vektorraumes  $V$  gilt:

(I)  $V = U \oplus U^\perp$ , insbesondere ist  $\dim V = \dim U + \dim U^\perp$ .

(II)  $U = (U^\perp)^\perp = U^{\perp\perp}$ .

Bew.: Wir wählen eine Basis  $a_1, \dots, a_m$  von  $U$  und orthonormalisieren diese mit dem Gram-Schmidt -Verfahren. Wir erhalten so eine Orthonormalbasis  $b_1, \dots, b_m$  von  $U$ .

Diese ergänzen wir unter Verwendung von Satz 10.7 zu einer Orthonormalbasis  $b_1, \dots, b_m, b_{m+1}, \dots, b_n$  zu  $V$ . Dann sind  $b_{m+1}, \dots, b_n \in U^\perp$ . Daher ist  $\dim U^\perp \geq n - m$ .

Für jedes  $u \in U \cap U^\perp$  ist  $(u|u)=0$ , d.h.  $u=0$ . Daher ist  $U \cap U^\perp = \{0\}$ , also  $U \cap U^\perp = U \oplus U^\perp$  und  $\dim U + \dim U^\perp = \dim U + \dim U^\perp \geq m+n-m=n = \dim V$ .

Folglich ist  $V = U + U^\perp = U \oplus U^\perp$ .

Für  $u \in U$  und  $v \in U^\perp$  ist  $0 = (v|u) = (u|v)$ . Daher ist  $U \subseteq U^{\perp\perp}$ .

Andererseits gilt nach (I):  $\dim U^{\perp\perp} = \dim V - \dim U^\perp = \dim V - (\dim V - \dim U) = \dim U$ .

Folglich ist  $U = U^{\perp\perp}$ .

**10.9 Satz:** Für Untervektorräume  $U_1$  und  $U_2$  eines endlich-dimensionalen euklidischen Vektorraumes  $V$  gilt: (I)  $U_1 \subseteq U_2 \Rightarrow U_2^\perp \subseteq U_1^\perp$ .

(II)  $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$ .

(III)  $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$ .

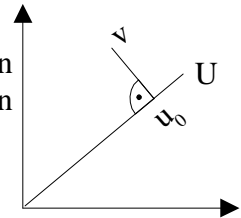
**Bew.:** (I) „ $\Rightarrow$ “ folgt unmittelbar aus der Definition.

„ $\Leftarrow$ “  $U_2^\perp \subseteq U_1^\perp \Rightarrow U_1^{\perp\perp} \subseteq U_2 \Rightarrow U_1 \subseteq U_2$ .

(II) Wegen  $U_1 \subseteq U_1 + U_2$  gilt nach (I):  $(U_1 + U_2)^\perp \subseteq U_1^\perp$ . Analog ist  $(U_1 + U_2)^\perp \subseteq U_2^\perp$ . Daher ist  $(U_1 + U_2)^\perp \subseteq U_1^\perp \cap U_2^\perp$ .

Umgekehrt gilt für  $u_1 \in U_1, u_2 \in U_2, v \in U_1^\perp \cap U_2^\perp, (v|u_1 + u_2) = \underbrace{(v|u_1)}_{=0} + \underbrace{(v|u_2)}_{=0} = 0$ . Also ist  $v \in (U_1 + U_2)^\perp$ .

(III)  $(U_1 \cap U_2)^\perp \stackrel{10.8(II)}{=} (U_1^{\perp\perp} \cap U_2^{\perp\perp})^\perp \stackrel{10.9(II)}{=} (U_1^\perp + U_2^\perp)^\perp = U_1 + U_2$ .



**10.10 Satz:** Gegeben seien ein Untervektorraum  $U$  eines endlich-dimensionalen euklidischen Vektorraumes  $V$  und ein Vektor  $v \in V$ . Dann existiert genau ein Vektor  $u_0 \in U$  mit

$$d(v, u_0) = \min\{d(v, u) \mid u \in U\}.$$

Für jede Orthonormalbasis  $u_1, \dots, u_m$  von  $U$  gilt:  $u_0 = (v|u_1)u_1 + \dots + (v|u_m)u_m$ .

**Bew.:** Seien  $V, U, v$  und  $u_1, \dots, u_m$  gegeben. Wir setzen  $u_0 := (v|u_1)u_1 + \dots + (v|u_m)u_m$ .

Für  $i=1; \dots; m$  gilt dann:  $(v - u_0|u_i) = (v|u_i) - (v|u_1)(u_1|u_i) - \dots - (v|u_m)(u_m|u_i) = (v|u_i) - (v|u_i) = 0$ . Daher ist  $v - u_0 \in U^\perp$ . Für beliebige Elemente  $a_1, \dots, a_m \in \mathbb{R}$  und  $u = a_1u_1 + \dots + a_mu_m \in U$  gilt:  $v - u = v - u_0 + [(v|u_1) - a_1]u_1 + \dots + [(v|u_m) - a_m]u_m$ . Aus dem Satz des Pythagoras folgt also:

$$\|v - u\|^2 = \|v - u_0\|^2 + \underbrace{[(v|u_1) - a_1]^2 + \dots + [(v|u_m) - a_m]^2}_{\geq 0} \geq \|v - u_0\|^2, \text{ und Gleichheit hat man nur im Fall } a_1 = (v|u_1), \dots, a_m = (v|u_m).$$

**Bsp.:** (Methode der kleinsten Quadrate)

Eine Versuchsreihe hat Meßwerte  $(r_1, s_1), \dots, (r_n, s_n) \in \mathbb{R}^2$  ergeben. Wir suchen eine Gerade  $y = ax + b$ , die diese Meßwerte „approximiert“, d.h. es soll gelten:  $s_i \sim ar_i + b$ . Genauer soll  $(ar_1 + b - s_1)^2 + \dots + (ar_n + b - s_n)^2$  möglichst klein werden. Setzt man also  $v_1 = (r_1, \dots, r_n), v_2 = (1; \dots; 1), v = (s_1, \dots, s_n) \in \mathbb{R}^n$ , so sind Zahlen  $a, b \in \mathbb{R}$  gesucht mit  $\|av_1 + bv_2 - v\| = \text{minimal}$ .

Wir suchen also im Untervektorraum  $U = \text{Span}(v_1, v_2)$  von  $\mathbb{R}^n$  einen Vektor  $u_0 = av_1 + bv_2$ , der von  $v$  minimalen Abstand hat. Nach dem Satz ist  $u_0 = (v|u_1)u_1 + (v|u_2)u_2$ , wobei  $u_1, u_2$  eine Orthonormalbasis von  $U$  ist.

Als konkretes Beispiel nehmen wir Meßwerte  $(2; 8), (3; 10), (-5; -3)$ . Dann ist  $v_1 = (2; 3; -5), v_2 = (1; 1; 1), v = (8; 10; -3)$ .

Wendet man das Gram-Schmidt-Verfahren auf  $v_1$  und  $v_2$  an, so erhält man die Orthonormalbasis  $u_1 = \frac{1}{\sqrt{38}} (2; 3; -5), u_2 = \frac{1}{\sqrt{3}} (1; 1; 1)$ . Es ist also  $u_0 = \frac{1}{\sqrt{38}} (2; 3; -5) + \frac{1}{\sqrt{3}} (1; 1; 1) = \frac{1}{38} (312; 373; -115)$ .

Wegen  $u_0 = av_1 + bv_2$  ist  $\frac{312}{38} = 2a + b$  und  $\frac{373}{38} = 3a + b$ , also  $\left. \begin{matrix} a = \frac{61}{38} \\ b = \frac{190}{38} \end{matrix} \right\}$  gesuchte Gerade.

Als „Probe“ berechnen wir:  $(2a + b - 8)^2 + (3a + b - 10)^2 + (-5a + b + 3)^2 = [8^2 + (-7)^2 + (-1)^2] / 38^2 = \frac{114}{1444} \sim 8\%$ .

Bem.: Gegeben seien von 0 verschiedene Vektoren  $a, b$  in einem euklidischen Vektorraum  $V$ . Nach der CSU gilt dann  $-1 \leq \frac{(a|b)}{\|a\| \cdot \|b\|} \leq +1$ .

In der Analysis lernt man, daß genau ein  $\vartheta = \vartheta_{a,b} \in \mathbb{R}$  existiert mit  $0 \leq \vartheta \leq \pi$  und  $\cos \vartheta = \frac{(a|b)}{\|a\| \cdot \|b\|}$

Man nennt  $\vartheta$  den Winkel zwischen  $a$  und  $b$ . Es gilt also  $(a|b) = \|a\| \cdot \|b\| \cdot \cos \vartheta$ . Daher kann man den Satz des Pythagoras auch in der folgenden Form schreiben.  $\|a-b\|^2 = \|a\|^2 + \|b\|^2 - 2\|a\| \cdot \|b\| \cdot \cos \vartheta_{a,b}$ . Diese Gleichung bezeichnet man auch als Cosinus-Satz.

Offenbar gilt stets:

- (I)  $\vartheta_{a,b} = \vartheta_{b,a}$ .
- (II)  $\vartheta_{a,-b} = \pi - \vartheta_{a,b}$ .
- (III)  $\vartheta_{ra,rb} = \vartheta_{a,b}$  für positive  $r, s \in \mathbb{R}$ .
- (IV)  $a, b$  linear abhängig  $\Leftrightarrow \vartheta_{a,b} \in \{0; \pi\}$ .
- (V)  $a \perp b \Leftrightarrow \vartheta_{a,b} = \pi/2$ .

24.01.01

## 11. Isometrien und adjungierte Abbildungen

11.1 Def.: Gegeben seien euklidische Vektorräume  $V, W$ . Eine Abbildung  $f: V \rightarrow W$  mit  $(f(x)|f(y)) = (x|y)$  für alle  $x, y \in V$  nennt man Isometrie.

Bem.: (I) Gegebenenfalls erhält  $f$  Längen von und Winkel zwischen Vektoren. Insbesondere ist  $f$  stets injektiv.

(II) Ist  $f$  sogar bijektiv, dann ist auch die Umkehrabbildung  $f^{-1}: W \rightarrow V$  eine Isometrie; für  $w, w' \in W$  gilt:  $(f^{-1}(w)|f^{-1}(w')) = (f(f^{-1}(w))|f(f^{-1}(w')))) = (w|w')$ .

(III) Hat man euklidische Vektorräume  $U, V, W$  und Isometrien  $f: U \rightarrow V, g: V \rightarrow W$ , so ist auch  $g \circ f: U \rightarrow W$  eine Isometrie. Für  $u, u' \in U$  gilt nämlich:  $(g(f(u))|g(f(u'))) = (f(u)|f(u')) = (u|u')$ .

(IV) Stets ist  $\text{id}_V: V \rightarrow V, v \rightarrow v$  eine Isometrie. Allgemeiner nennt man Isometrien  $f: V \rightarrow V$  orthogonale Transformationen.

11.2 Def.: Zwei euklidische Vektorräume  $V, W$  nennt man isometrisch isomorph, falls eine bijektive Isometrie  $f: V \rightarrow W$  existiert.

Bem.: Die isometrische Isomorphie von euklidischen Vektorräumen ist eine Äquivalenzrelation, dies beweist man in der üblichen Weise.

Satz: Zwei endlich-dimensionale euklidische Vektorräume sind genau dann isometrisch isomorph, wenn sie die gleiche Dimension haben.

Bew.: Wir wissen schon, daß isomorphe Vektorräume stets die gleiche Dimension haben. Umgekehrt zeigen wir, daß jeder euklidische Vektorraum  $V$  der Dimension  $n < \infty$  zum  $\mathbb{R}^n$  mit dem Standardskalarprodukt isometrisch isomorph ist. Dazu wählen wir eine Orthonormalbasis  $b_1, \dots, b_n$  von  $V$ . Betrachten wir die Abbildung  $f: \mathbb{R}^n \rightarrow V, (x_1, \dots, x_n) \rightarrow x_1 b_1 + \dots + x_n b_n$ .

Offenbar ist  $f$  linear und bijektiv.

Für  $x = (x_1, \dots, x_n)$  und  $y = (y_1, \dots, y_n)$  gilt nach Bemerkung 10.5:

$(f(x)|f(y)) = (x_1 b_1 + \dots + x_n b_n | y_1 b_1 + \dots + y_n b_n) \stackrel{10.5}{=} x_1 y_1 + \dots + x_n y_n = (x|y)$ . Daher ist  $f$  Isometrie.

11.3 Satz: Gegeben seien endlich-dimensionale euklidische Vektorräume  $V, W$  mit Orthonormalbasis  $b_1, \dots, b_m$  bzw.  $c_1, \dots, c_n$ . Eine lineare Abbildung  $f: V \rightarrow W$  mit Matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times m}$  bezüglich  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$  ist genau dann Isometrie, wenn gilt:  $\boxed{A^T A = I_m}$ .

Bew.: „ $\Rightarrow$ “ Sei  $f$  Isometrie. Für  $i, j=1; \dots; m$  setzen wir  $\delta_{ij} := \begin{cases} 1 & \text{falls } i=j \\ 0 & \text{sonst} \end{cases}$  (Kronecker-Delta).

Dann ist  $\delta_{ij} = (b_i | b_j) = (f(b_i) | f(b_j)) = (a_{i1}c_1 + \dots + a_{im}c_m | a_{j1}c_1 + \dots + a_{jm}c_m) = \sum_{k=1}^m a_{ik}a_{jk}$ . Folglich ist  $1_m = A^T A$ .

„ $\Leftarrow$ “ Sei umgekehrt  $A^T A = 1_m$ . Für  $i, j=1; \dots; m$  gilt dann:

$$(f(b_i) | f(b_j)) = (a_{i1}c_1 + \dots + a_{im}c_m | a_{j1}c_1 + \dots + a_{jm}c_m) = \sum_{k=1}^m a_{ik}a_{jk} = \delta_{ij}.$$

Für beliebige  $x_1, y_1, \dots, x_m, y_m \in \mathbb{R}$  gilt also:

$$\begin{aligned} (f(x_1 b_1 + \dots + x_m b_m) | f(y_1 b_1 + \dots + y_m b_m)) &= (x_1 f(b_1) + \dots + x_m f(b_m) | y_1 f(b_1) + \dots + y_m f(b_m)) \\ &= x_1 y_1 (f(b_1) | f(b_1)) + \dots + x_1 y_m (f(b_1) | f(b_m)) + \dots + x_m y_1 (f(b_m) | f(b_1)) + \dots + x_m y_m (f(b_m) | f(b_m)) \\ &= x_1 y_1 + \dots + x_m y_m = (x_1 b_1 + \dots + x_m b_m | y_1 b_1 + \dots + y_m b_m). \end{aligned}$$

Also ist  $f$  Isometrie.

Def.: Eine Matrix  $A \in \mathbb{R}^{n \times n}$  mit  $A^T A = 1_n$  nennt man orthogonal.

Bem.: (I) Orthogonale Matrizen entsprechen also orthogonalen Transformationen nach Wahl einer Orthonormalbasis.

(II) Für jede orthogonale Matrix  $A \in \mathbb{R}^{n \times n}$  gilt:

$$1 = \det 1_n = \det (A^T A) = \det A^T \cdot \det A = (\det A)^2, \text{ also } \det A = 1 \text{ oder } \det A = -1.$$

(III) Für jede orthogonale Transformation  $f$  auf einem euklidischen Vektorraum  $V$  gilt also:

$$\det f \in \{1; -1\}$$

Bsp.: Sei  $V$  ein euklidischer Vektorraum,  $\dim V=2$  mit Orthonormalbasis  $e_1, e_2$  und  $f: V \rightarrow V$  eine orthogonale Transformation.

Wir schreiben  $f(e_1) = ae_1 + be_2$  und  $f(e_2) = ce_1 + de_2$  mit  $a, b, c, d \in \mathbb{R}$ . Dann ist

$$1 = \|e_1\|^2 = \|f(e_1)\|^2 = \|ae_1 + be_2\|^2 = a^2 + b^2.$$

Daher existiert  $\alpha \in \mathbb{R}$  mit  $0 \leq \alpha < 2\pi$  und  $a = \cos \alpha$ ,  $b = \sin \alpha$ . Analog ist  $c^2 + d^2 = 1$ . Wegen  $0 = (e_1 | e_2) = (f(e_1) | f(e_2)) = (ae_1 + be_2 | ce_1 + de_2) = ac + bd$  folgt leicht:  $f(e_2) = \pm(-be_1 + ae_2)$

Fall 1:  $f(e_2) = -be_1 + ae_2$ . Dann ist die Matrix von  $f$  bezüglich  $e_1, e_2$  gleich  $A = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$ ,  $x$  sei  $\alpha$ .

Insbesondere ist  $\det f = \det A = 1$ . Geometrisch ist  $f$  eine Drehung um den Winkel  $\alpha$ .

Fall 2:  $f(e_2) = be_1 - ae_2$ . Dann ist die Matrix von  $f$  bezüglich  $e_1, e_2$  gleich  $B = \begin{pmatrix} \cos x & \sin x \\ \sin x & -\cos x \end{pmatrix}$ ,  $x$  sei  $\alpha$ .

Insbesondere ist  $\det f = \det B = -1$ .

Wir benutzen die Additionstheoreme aus der Analysis:

$$\cos(\alpha + \beta) = \cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta.$$

$$\sin(\alpha + \beta) = \sin \alpha \cdot \cos \beta + \cos \alpha \cdot \sin \beta.$$

Dann sieht man, daß  $b_1 = (\cos \alpha/2)e_1 + (\sin \alpha/2)e_2$  ein Eigenvektor von  $f$  zum Eigenwert 1 und  $b_2 = (-\sin \alpha/2)e_1 + (\cos \alpha/2)e_2$  ein Eigenvektor von  $f$  zum Eigenwert -1 ist.

Die Vektoren  $b_1, b_2$  bilden eine Orthonormalbasis von  $V$  und bezüglich  $b_1, b_2$  hat die Matrix von  $f$  die Form  $C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . Geometrisch ist  $f$  Spiegelung an der Geraden  $\mathbb{R}b_1$ . In Dimension 2 ist also jede orthogonale Transformation entweder eine Drehung oder eine Spiegelung.

11.4 Satz: Gegeben sei ein euklidischer Vektorraum  $V$  mit Orthonormalbasis  $b_1, \dots, b_n$  und eine Matrix  $S = (s_{ij}) \in \mathbb{R}^{n \times n}$ . Wir setzen  $b'_j := s_{j1}b_1 + \dots + s_{jn}b_n$  ( $j=1; \dots; n$ ). Dann gilt:

$b'_1, \dots, b'_n$  ist Orthonormalbasis von  $V \Leftrightarrow S$  ist orthogonal.

Bew.: Für  $i, j=1; \dots; n$  gilt  $(b'_i | b'_j) = s_{i1}s_{j1} + \dots + s_{in}s_{jn}$ . Daher gilt:

$$b'_1, \dots, b'_n \text{ ist Orthonormalbasis von } V \Leftrightarrow (b'_i | b'_j) = \delta_{ij} \text{ (} i, j=1; \dots; n) \Leftrightarrow s_{i1}s_{j1} + \dots + s_{in}s_{jn} = \delta_{ij} \text{ (} i, j=1; \dots; n) \Leftrightarrow S^T S = 1_n \Leftrightarrow S \text{ ist orthogonal.}$$

25.01.01

11.5 Satz: Gegeben seien endlich-dimensionale euklidische Vektorräume  $V, W$  und eine lineare Abbildung  $f: V \rightarrow W$ . Zu jedem  $w \in W$  ex. dann genau ein Element  $f^*(w) \in V$  mit  $(f^*(w) | v) = (w | f(v))$

(!) für alle  $v \in V$ . Die so definierte Abbildung  $f^*: W \rightarrow V$  ist linear.

Bew.: Wir wählen eine Orthonormalbasis  $b_1, \dots, b_n$  von  $V$ . Sei  $w \in W$  und  $f^*(w) \in V$  mit  $(f^*(w)|v) = (w|f(v))$  für alle  $v \in V$ . Wir schreiben  $f^*(w) = a_1 b_1 + \dots + a_n b_n$  mit  $a_1, \dots, a_n \in \mathbb{R}$ . Für  $i=1; \dots; n$  gilt dann:  $a_i = (f^*(w)|b_i) = (w|f(b_i))$ . Daher ist  $f^*(w)$  eindeutig bestimmt (da jedes  $a_i$  eindeutig bestimmt ist). Zum Beweis der Existenz setzen wir  $f^*(w) := (w|f(b_1))b_1 + \dots + (w|f(b_n))b_n$ .

Für  $i=1; \dots; n$  ist dann  $(f^*(w)|b_i) = (w|f(b_i))$ . Sei jetzt  $v \in V$  beliebig und  $v = r_1 b_1 + \dots + r_n b_n$  mit  $r_1, \dots, r_n \in \mathbb{R}$ . Dann gilt:

$$(f^*(w)|v) = r_1 (f^*(w)|b_1) + \dots + r_n (f^*(w)|b_n) = r_1 (w|f(b_1)) + \dots + r_n (w|f(b_n)) = (w|r_1 f(b_1) + \dots + r_n f(b_n)) = (w|f(r_1 b_1 + \dots + r_n b_n)) = (w|f(v)).$$

Damit sind die Existenz und Eindeutigkeit von  $f$  bewiesen. Zum Beweis der Linearität von  $f^*$  seien  $w, w' \in V$ ,  $a, a' \in \mathbb{R}$ . Dann gilt:

$$f^*(aw + a'w') = (aw + a'w'|f(b_1))b_1 + \dots + (aw + a'w'|f(b_n))b_n = a(w|f(b_1))b_1 + a'(w|f(b_1))b_1 + \dots + a(w|f(b_n))b_n + a'(w|f(b_n))b_n = a((w|f(b_1))b_1 + \dots + (w|f(b_n))b_n) + a'((w|f(b_1))b_1 + \dots + (w|f(b_n))b_n) = af^*(w) + a'f^*(w').$$

Damit ist  $f^*$  linear.

Def.: Man nennt  $f^*$  die zu  $f$  adjungierte (lineare) Abbildung.

Bem.: Wir wählen zwei Orthonormalbasen  $b_1, \dots, b_m$  von  $V$  und  $c_1, \dots, c_n$  von  $W$  und bezeichnen mit  $A = (a_{ij})$  und  $A^* = (a^*_{ij})$  die Matrix von  $f$  bzw.  $f^*$  bezüglich dieser Basen. Es ist also  $f(b_j) = a_{1j}c_1 + \dots + a_{nj}c_n$  ( $j=1; \dots; m$ ) und  $f^*(c_l) = a^*_{1l}b_1 + \dots + a^*_{ml}b_m$  ( $l=1; \dots; n$ ).

Daher gilt für  $k=1; \dots; m$ ,  $l=1; \dots; n$   $a^*_{lk} = (f^*(c_l)|b_k) = (c_l|f(b_k)) = a_{lk}$ . Folglich ist  $\overline{A^*} = A^T$ .

Bsp.:  $f$  Isometrie  $\Leftrightarrow (v|v') = (f(v)|f(v'))$  für alle  $v \in V \Leftrightarrow f^*(f(v)) = v$  für alle  $v \in V \Leftrightarrow f^* \circ f = \text{id}_V$ .

11.6 Satz: Für endlich-dimensionale euklidische Vektorräume  $U, V, W$  gilt stets:

- (I)  $(f^*)^* = f$  für  $f \in \text{Hom}(U, V)$ .
- (II)  $(af + bg)^* = af^* + bg^*$  für  $a, b \in \mathbb{R}$ ,  $f, g \in \text{Hom}(U, V)$ .
- (III)  $(g \circ f)^* = f^* \circ g^*$  für  $f \in \text{Hom}(U, V)$  und  $g \in \text{Hom}(V, W)$ .
- (IV)  $\det f^* = \det f$  für  $f \in \text{Hom}(U, U)$ .

Bew.: Dies folgt mit Bem. 11.5 aus den entsprechenden Rechenregeln für Matrizen( s. dort):

- (I)  $(A^T)^T = A$ .
- (II)  $(aA + bB)^T = aA^T + bB^T$ .
- (III)  $(BA)^T = A^T \cdot B^T$ .
- (IV)  $\det A^T = \det A$ .

Def.: Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum. Eine lineare Abbildung  $f: V \rightarrow V$  heißt:

- (I) selbstadjungiert, falls  $f^* = f$ .
- (II) normal, falls  $f^* \circ f = f \circ f^*$ .

Bem.: Ist  $A$  die Matrix von  $f$  bezüglich einer Orthonormalbasis von  $V$ , so gilt also:

- (I)  $f$  selbstadjungiert  $\Leftrightarrow A^T = A$ , d.h.  $A$  ist symmetrisch.
- (II)  $f$  normal  $\Leftrightarrow A^T A = A A^T$ . Matrizen mit dieser Eigenschaft nennt man ebenfalls normal.

- Bsp.: (I)  $f$  selbstadjungiert  $\Rightarrow f$  normal
- (II)  $f$  Isometrie  $\Rightarrow f^* = f^{-1} \Rightarrow f$  normal

11.7 Satz: Für jeden endlich-dimensionalen euklidischen Vektorraum  $V$  und jede normale lineare Abbildung  $f: V \rightarrow V$  gilt: (I)  $\text{Ker } f^* = \text{Ker } f$ .

- (II)  $f$  und  $f^*$  besitzen die gleichen Eigenwerte.
- (III) Für jeden Eigenwert  $r$  von  $f$  ist  $E_r(f^*) = E_r(f)$ .

Bew.: (I) Für  $v \in V$  gilt:  $f^*(v)=0 \Leftrightarrow (f^*(v)|f^*(v))=0 \Leftrightarrow (v|f(f^*(v)))=0 \Leftrightarrow (f^*(f(v))|v)=0 \Leftrightarrow (f(v)|f(v))=0 \Leftrightarrow f(v)=0$ .

(II)-(III) Wegen  $1_n^T=1_n$  ist  $(id_v)^*=id_v$ . Für  $r \in \mathbb{R}$  gilt also:  $(f-r \cdot id_v)^*o(f-r \cdot id_v)=(f^*-r(id_v)^*)o(f-r \cdot id_v)=(f^*-r \cdot id_v)o(f-r \cdot id_v)=f^*of-rf-rf^*+r^2 \cdot id_v=f^*of-rf-rf^*+r^2 \cdot id_v=(f-r \cdot id_v)o(f^*-r \cdot id_v)=(f-r \cdot id_v)o(f-r \cdot id_v)^*$ . Daher ist  $f-r \cdot id_v$  auch normal. Nach (I) gilt also  $\underline{E_r(f)}=Ker(f^*-r \cdot id_v)=Ker(f-r \cdot id_v)^*=Ker(f-r \cdot id_v)=\underline{E_r(f)}$ .

Bem.: Im folgenden werden wir beweisen, daß selbstadjungierte lineare Abbildungen stets diagonalisierbar sind. Äquivalent dazu ist die Aussage, daß symmetrische Matrizen stets diagonalisierbar sind. Der Beweis nutzt komplexe Zahlen, diese werden wir jetzt einführen.

12. Die Hauptachsentransformation

12.1 Def.: Ein Körper ist eine Menge  $K$  mit zwei Abbildungen  $K \times K \rightarrow K, (a,b) \rightarrow (a+b)$  (Addition) und  $K \times K \rightarrow K, (a,b) \rightarrow (a \cdot b)$  (Multiplikation) und folgenden Eigenschaften:

- (I)  $a+b=b+a$  (Kommutativgesetz)
- (II)  $(a+b)+c=a+(b+c)$  (Assoziativgesetz)
- (III) Es existiert ein Element  $0 \in K$  mit  $0+a=a$  für  $a \in K$  (Existenz des Nullelements)
- (IV) Zu jedem  $a \in K$  existiert ein  $-a \in K$  mit  $a+(-a)=0$  (negatives Element)
- (V)  $ab=ba$  (Kommutativgesetz)
- (VI)  $(ab)c=a(bc)$  (Assoziativgesetz)
- (VII) Es existiert ein Element  $1 \in K$  mit  $1 \neq 0 \wedge 1a=a$  für  $a \in K$  (Existenz des Einselements)
- (VIII) Zu jedem  $a \in K \setminus \{0\}$  existiert ein  $a^{-1} \in K \setminus \{0\}$  mit  $aa^{-1}=1$  (inverses Element)
- (IX)  $a(b+c)=ab+ac$  (Distributivgesetz)

Für (I)-(IX) gilt  $a,b,c \in K$ .

Bsp.:  $\mathbb{R}$  ist Körper,  $\mathbb{Q}$  ist Körper,  $\mathbb{Z}$  dagegen nicht( kein inverses Element)

Den Körper  $\mathbb{C}$  der komplexen Zahlen werden wir bald kennenlernen. Für die Informatik wichtig ist der Körper mit zwei Elementen( nämlich „0“ und „1“).

Bem.: (I) In jedem Körper ist das Nullelement eindeutig bestimmt. Sind nämlich  $0,0' \in K$  mit  $0+a=a=0'+a$  für  $a \in K$ , so gilt  $\underline{0}=0+0'=0'+0=\underline{0}$ .

31.01.01

(II) Zu jedem Element  $a \in K$  ist das negative Element  $-a \in K$  eindeutig bestimmt, ist nämlich  $a+b=0=a+c$ , so folgt  $\underline{b}=0+b=(a+c)+b=(c+a)+b=c+(a+b)=c+0=0+c=\underline{c}$ .

Statt  $a+(-b)$  schreibt man  $a-b$ .

(III) Analog zeigt man, daß  $K$  genau ein Einselement besitzt und daß zu jedem  $a \in K \setminus \{0\}$  genau ein Inverses  $a^{-1} \in K \setminus \{0\}$  existiert.

(IV) Für  $a \in K$  ist  $0a=(0+0)a=0a+0a$ . Addition von  $(-0a)$  auf beiden Seiten ergibt:  $0=0a$ . Daher ist auch  $a \cdot 0=0$ .

(V) Aus  $ab=0$  folgt  $a=0 \vee b=0$ , denn im Fall  $a \neq 0$  ergibt die Multiplikation mit  $a^{-1}$ :  $\underline{0}=a^{-1}0=a^{-1}(ab)=(a^{-1}a)b=1b=\underline{b}$ .

(VI) Wie bei Vektorräumen beweist man Rechenregeln wie:  $-(-a)=a, -(a+b)=-a-b, -(ab)=(-a)b, (-a)(-b)=ab$ , usw.

Analog beweist man:  $(a^{-1})^{-1}=a, (ab)^{-1}=a^{-1}b^{-1}, \dots$

(VII) Wie bei Vektorräumen vereinbart man, daß Punktrechnung vor Strichrechnung gilt. Wegen der Assoziativgesetze läßt man häufig Klammern weg.

12.2 Satz:  $\mathbb{R}^2=\{(a,b) \mid a,b \in \mathbb{R}\}$  wird zu einem Körper, wenn man Addition und Multiplikation folgendermaßen definiert:  $(a,b)+(c,d)=(a+c,b+d)$  und  $(a,b) \cdot (c,d)=(ac-bd,ad+bc)$

Bew.: Die Rechenregeln für die Addition kennen wir bereits( entspricht der Addition in Vektorräumen)

$$(a,b)(c,d)=(ac-bd,ad+bc)=(ca-db,cb+da)\equiv(c,d)(a,b) \quad \text{Kommutativität}$$

$$[(a,b)(c,d)](e,f)=(ac-bd,ad+bc)(e,f)=(ace-bde-adf-bcf,acf-bdf+ade+bce)=(a,b)(ce-df,cf+de)\equiv(a,b)[(c,d)(e,f)] \quad \text{Assoziativität}$$

$$(a,b)[(c,d)+(e,f)]=(a,b)(c+e,d+f)=(ac+ae-bd-bf,ad+af+bc+be)=(ac-bd,ad+bc)+(ae-bf,af+be)\equiv(a,b)(c,d)+(a,b)(e,f) \quad \text{Distributivität}$$

Einselement ist (1,0), denn für  $(a,b) \in \mathbb{R}^2$  gilt:  $(1,0)(a,b)=(1a-0b,1b+0a)=(a,b)$

$$\text{Ist } (a,b) \in \mathbb{R} \wedge (a,b) \neq (0,0), \text{ so ist } a^2+b^2 \neq 0 \text{ und } (a,b) \left( \frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right) = \left( \frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2}, \frac{-ab}{a^2+b^2} + \frac{ab}{a^2+b^2} \right) = \left( \frac{a^2+b^2}{a^2+b^2}, \frac{ab-ab}{a^2+b^2} \right) = (1,0).$$

Def.: Diesen Körper bezeichnet man mit  $\mathbb{C}$  und seine Elemente als komplexe Zahlen.

Bem.: (I) Man kann sich  $\mathbb{C}$  als (Gaußsche) Zahlenebene veranschaulichen.

(II) Für  $a,b \in \mathbb{R}$  gilt:  $(a,0)+(b,0)=(a+b,0)$

$$(a,0) \cdot (b,0)=(a \cdot b,0)$$

Man rechnet also mit den Zahlen  $(a,0) \in \mathbb{C}$  wie mit den Zahlen  $a \in \mathbb{R}$ . Zur Vereinfachung identifiziert man daher eine Zahl  $a \in \mathbb{R}$  mit der Zahl  $(a,0) \in \mathbb{C}$ . Auf diese Weise kann man  $\mathbb{R}$  auch als Teilmenge( Teilkörper) von  $\mathbb{C}$  auffassen.

(III) Wir setzen  $i:=(1,0) \in \mathbb{C}$ . Für  $a,b \in \mathbb{R}$  gilt:  $a+bi=(a,0)+(b,0)(0,1)=(a,0)+(0,b)\equiv(a,b)$ .

Daher ist  $\mathbb{C}=\{a+bi \mid a,b \in \mathbb{R}\}$  und  $i^2=(0,1)(0,1)=(-1,0)\equiv-1$ . !

Wir schreiben in Zukunft komplexe Zahlen in der Form  $z=a+bi$  mit  $a,b \in \mathbb{R}$  und rechnen folgendermaßen:

$$(a+bi)+(c+di)=a+c+(b+d)i$$

$$(a+bi)(c+di)=(ac-bd)+(ad+bc)i$$

Man a den Real- und b den Imaginärteil von der komplexen Zahl  $z=a+bi$  und schreibt:  $a=\text{Re}(z)$ ,  $b=\text{Im}(z)$ .

Bsp.: Wie schreibt man  $\frac{1+3i}{2-i}$  in der Form  $x+yi$  mit  $x,y \in \mathbb{R}$ ?

$$\frac{1+3i}{2-i} = \frac{1+3i}{2-i} \cdot \frac{2+i}{2+i} = \frac{(1+3i)(2+i)}{4-i^2} = \frac{(1+3i)(2+i)}{5} = \frac{-1+7i}{5} = \frac{-1}{5} + \frac{7}{5}i \quad \left( \frac{-1}{5} =x, \frac{7}{5} =y \right)$$

12.3 Def.: Ist  $z=a+bi \in \mathbb{C}$  mit  $a,b \in \mathbb{R}$ , so nennt man  $\bar{z} :=a-bi$  die zu  $z$  konjugiert komplexe Zahl.

Satz: Für  $z,w \in \mathbb{C}$  gilt: (I)  $\overline{z+w} = \bar{z}+\bar{w}$

(II)  $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$

(III)  $\overline{\bar{z}} =z$

(IV)  $z+\bar{z} =2 \text{Re}(z)$ ,  $z-\bar{z} =2 \text{Im}(z)i$

(V)  $z \in \mathbb{R} \Leftrightarrow \bar{z}=z$

(VI)  $z=a+bi$  mit  $a,b \in \mathbb{R} \Rightarrow z \cdot \bar{z} =a^2+b^2 \geq 0$  ( $a^2,b^2 \in \mathbb{R}$ )

Bew.: Wir schreiben  $z=a+bi$ ,  $w=c+di$ ,  $a,b,c,d \in \mathbb{R}$

$$(I) \quad \overline{z+w} = \overline{a+c+(b+d)i} =a+c-(b+d)i=a-bi+c-di= \bar{z}+\bar{w}$$

$$(II) \quad \overline{z \cdot w} = \overline{(ac-bd)+(ad+bc)i} =ac-bd-(ad+bc)i=(a-bi)(c-di)= \bar{z} \cdot \bar{w}$$

$$(III) \quad \overline{\bar{z}} = \overline{a-bi} = a-b(-i) =a+bi=z$$

$$(IV) \quad z+\bar{z} =(a+bi)+(a-bi)=2a=2 \text{Re}(z), \quad z-\bar{z} =(a+bi)-(a-bi)=2bi=2 \text{Im}(z)i$$

$$(V) \quad z \in \mathbb{R} \Leftrightarrow b=0 \Leftrightarrow (a+bi)=(a-bi) \Leftrightarrow \bar{z}=z$$

$$(VI) \quad z \cdot \bar{z} =a^2-b^2i^2=a^2+b^2 \geq 0 \wedge a^2+b^2 \in \mathbb{R}$$

12.4 Def.: Ist  $z=a+bi \in \mathbb{C}$  mit  $a,b \in \mathbb{R}$ , so nennt man  $|z|:= \sqrt{z \cdot \bar{z}} = \sqrt{a^2+b^2}$  den Betrag von  $z$ .

**Bem.:** Der Betrag von  $z$  ist also gleich der Norm von  $z$  im euklidischen Vektorraum  $\mathbb{R}^2$ , ausgestattet mit dem Standardskalarprodukt. Für alle  $z, w \in \mathbb{C}$  gilt:

- (I)  $|z| \geq 0$
- (II)  $|z|=0 \Leftrightarrow z=0$
- (III)  $|\bar{z}|=|z|$
- (IV)  $|zw|=|z|*|w|$
- (V)  $|z+w| \leq |z|+|w|$

**12.5 Bem.:** Geg. seien Elemente  $a_0, \dots, a_n \in \mathbb{C}$  mit  $a_n \neq 0$ . Dann nennt man die Abbildung  $p: \mathbb{C} \rightarrow \mathbb{C}$ ,  $z \rightarrow a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$  ein komplexes Polynom vom Grad  $n$ .

Ist  $z \in \mathbb{C}$  mit  $p(z)=0$ , so nennt man  $z$  eine Nullstelle von  $p$ . Der Fundamentalsatz der Algebra besagt, daß jedes komplexe Polynom, dessen Grad von 0 verschieden ist, auch eine Nullstelle in  $\mathbb{C}$  hat.

Er wurde von Carl-Friedrich-Gauß( 1777-1855) bewiesen. Sein Beweis geht über den Beweis der linearen Algebra hinaus. Elegante Beweise lernt man in Vorlesungen über Funktionentheorie oder Algebra. Wir werden ihn im folgenden verwenden( aber nicht beweisen).

**12.6 Bem.:** Statt Vektorräumen über reellen Zahlen kann man auch Vektorräume über einem beliebigen Körper  $K$  betrachten. Man spricht dann von K-Vektorräumen( Im Fall  $K=\mathbb{C}$  spricht man auch von komplexen Vektorräumen). Die in den Kapiteln 1-9 bewiesenen Tatsachen übertragen sich ohne Schwierigkeiten auf Vektorräume über beliebigen Körper. Wir werden diese Tatsachen auch im folgenden verwenden. Kapitel 10 und 11 lassen sich nicht ohne weiteres auf beliebige Körper verallgemeinern und zwar, da dort spezielle Eigenschaften von  $\mathbb{R}$  benutzt werden, die beliebige Körper nicht haben( z.B. eine Einteilung in pos. und neg. Elemente, Wurzel, Cosinus,...).

01.02.01

**12.7 Satz:** Jede symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  besitzt einen reellen Eigenwert.

**Bew.:** Das charakteristische Polynom  $p$  von  $A$  ist ein reelles Polynom vom Grad  $n$ . Dieses kann man auch als komplexes Polynom auffassen. Nach dem FS d. Algebra hat  $p$  eine Nullstelle  $t \in \mathbb{C}$ . Faßt man  $A$  als komplexe Matrix auf, so ist  $t$  auch ein Eigenwert von  $A$ . Wir werden zeigen, daß  $t \in \mathbb{R}$  gilt. Dazu betrachten wir die lineare Abbildung  $f: \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{n \times 1}$ ,  $z \rightarrow Az$ .

Diese hat die Matrix  $A$  bezüglich der Standardbasis  $e_1, \dots, e_n$  von  $\mathbb{C}^{n \times 1}$ . Daher ist  $t$  Eigenwert von  $f$ . Wir bezeichnen mit  $z=(z_1, \dots, z_n)^T \in \mathbb{C}^{n \times 1}$  einen entsprechenden Eigenvektor. Wir schreiben  $t=r+si$  mit  $r, s \in \mathbb{R}$  und  $z_j=x_j+y_j i$  mit  $x_j, y_j \in \mathbb{R}$  ( $j=1, \dots, n$ ). Wir setzen ferner  $x:=(x_1, \dots, x_n)^T$ ,  $y:=(y_1, \dots, y_n)^T \in \mathbb{R}^{n \times 1}$ . Aus  $f(z)=tz$  folgt  $f(x)=rx-sy(=Ax)$  und  $f(y)=ry+sx(=Ay)$ .

Folglich ist  $y^T Ax=y^T(rx-sy)=ry^T x-sy^T y=r(y_1 x_1 + \dots + y_n x_n)-s(y_1^2 + \dots + y_n^2) \in \mathbb{R}$  und  $y^T Ax=(y^T Ax)^T=x^T A^T y=-A^T y$  symm.  $x^T Ay=x^T(ry+sx)=rx^T y+sx^T x=r(y_1 x_1 + \dots + y_n x_n)+s(x_1^2 + \dots + x_n^2)$ .

Damit ist  $0=s(x_1^2 + \dots + x_n^2 + y_1^2 + \dots + y_n^2)$  und  $(x_1^2 + \dots + x_n^2 + y_1^2 + \dots + y_n^2) \neq 0$ , da  $z \neq 0$  ( $z$  ist Eigenvektor), also ist  $s=0$ , d.h.  $t=r \in \mathbb{R}$ .

**12.8 Satz:** Gegeben seien ein endlich-dimensionaler euklidischer Vektorraum  $V$  und eine lineare Abbildung  $f: V \rightarrow V$ .  $f$  ist genau dann selbstadjungiert, wenn eine Orthonormalbasis von  $V$  existiert, die aus Eigenvektoren von  $f$  besteht.

**Bew.:** „ $\Rightarrow$ “ Sei  $f$  selbstadjungiert. Wir wählen eine Orthonormalbasis  $b_1, \dots, b_n$  von  $V$  und wir bezeichnen mit  $A$  die Matrix von  $f$  bezüglich  $b_1, \dots, b_n$ . Dann ist  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Nach Satz 12.7 besitzt  $A$  einen Eigenwert  $r_1 \in \mathbb{R}$ . Dieser ist auch Eigenwert von  $f$ . Wir bezeichnen mit  $v_1 \in V$  einen entsprechenden Eigenvektor, o.B.d.A. gilt  $\|v_1\|=1$ .

Dann ist  $V=\mathbb{R}v_1 \oplus U$  mit  $U:=(\mathbb{R}v_1)^\perp$ . Für  $u \in U$  gilt dann:  $(v_1|f(u))=(f(v_1)|u)=(r_1 v_1|u)=0$ , d.h.  $f(u)$

$\in (\mathbb{R}V_1)^\perp = U$ . Daher kann man  $f$  zu einer linearen Abbildung  $g: U \rightarrow U$  einschränken. Offenbar ist  $U$  auch ein euklidischer Vektorraum, und für alle  $u, u' \in U$  gilt:  $(g(u)|u') = (f(u)|u') = (u|f(u')) = (u|g(u'))$ . Daher ist  $g$  auch selbstadjungiert. Wir argumentieren jetzt mit Induktion nach  $\dim V$ . Wir können also annehmen, daß eine Orthonormalbasis  $v_2, \dots, v_n$  von  $U$  existiert, die aus Eigenvektoren von  $g$  besteht. Dann ist  $v_1, \dots, v_n$  eine Orthonormalbasis von  $V$ , die aus Eigenvektoren von  $f$  besteht.

„ $\Leftarrow$ “ Sei  $b_1, \dots, b_n$  eine Orthonormalbasis von  $V$ , die aus Eigenvektoren von  $f$  besteht. Die Matrix von  $f$  bezüglich  $b_1, \dots, b_n$  ist eine Diagonalmatrix, d.h. insbesondere symmetrisch. Daher ist  $f$  selbstadjungiert.

Bsp.: Wir betrachten den euklidischen Vektorraum  $\mathbb{R}^3$  mit dem Standardskalarprodukt und die lineare Abbildung  $f: V \rightarrow V$  mit Matrix  $A = \begin{pmatrix} 6 & -1 & -1 \\ -1 & 6 & -1 \\ -1 & -1 & 6 \end{pmatrix}$  bezüglich der Standardbasis  $e_1, e_2, e_3$  von  $V$ . Wegen  $A^T = A$  ist  $f$  selbstadjungiert. Wir berechnen zunächst die Eigenwerte von  $f$ :  $\begin{vmatrix} r-6 & -1 & -1 \\ -1 & r-6 & -1 \\ -1 & -1 & r-6 \end{vmatrix} = \dots = (r-4)(r-7)^2$ . Daher sind 4 und 7 die beiden einzigen Eigenwerte von  $f$  und  $A$ .

Als nächstes berechnen wir  $E_4(f)$ :  $\begin{pmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix} \xrightarrow{\text{Gauß-Algo.}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$  Eigenvektor ist z.B.  $(1; 1; 1)$ . Normierung ergibt:  $b_1 = \frac{1}{\sqrt{3}} (1; 1; 1)$ .

Dann berechnen wir entsprechend  $E_7(f)$ . Eine Basis von  $E_7(f)$  ist z.B.  $(1; -1; 0)$ ,  $(1; 0; -1)$ . Auf diese wenden wir das GS-Verfahren an. Wir erhalten die Vektoren  $\frac{1}{\sqrt{2}} (1; -1; 0)$  und  $(1; 0; -1) - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} (1; -1; 0) = \left(\frac{1}{2}; \frac{1}{2}; -1\right)$ . Den letzten Vektor müssen wir noch normalisieren und erhalten dann die Orthonormalbasis  $b_1 = \frac{1}{\sqrt{3}} (1; 1; 1)$ ,  $b_2 = \frac{1}{\sqrt{2}} (1; -1; 0)$  und  $b_3 = \frac{1}{\sqrt{6}} (1; 1; -2)$ .

Bezüglich dieser Basis hat die Matrix von  $f$  die folgende Form:  $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}$ .

12.9 Satz: Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist genau dann symmetrisch, wenn eine orthogonale Matrix  $S$  existiert mit der Eigenschaft, daß  $SAS^{-1}$  eine Diagonalmatrix ist.

Bew.: „ $\Rightarrow$ “ Sei  $A$  symmetrisch. Wir betrachten den euklidischen Vektorraum  $V = \mathbb{R}^n$  mit dem Standardskalarprodukt und die (selbstadjungierte) lineare Abbildung  $f: V \rightarrow V$  mit Matrix  $A$  bezüglich der Standardbasis  $e_1, \dots, e_n$  von  $V$ . Nach Satz 12.8 existiert eine Orthonormalbasis  $b_1, \dots, b_n$  von  $V$ , die aus Eigenvektoren von  $f$  besteht. Die Matrix  $B$  bezüglich  $b_1, \dots, b_n$  ist also eine Diagonalmatrix. Wir schreiben:  $e_j = s_{1j}b_1 + \dots + s_{nj}b_n$  mit  $s_{ij} \in \mathbb{R}$  für  $i, j = 1, \dots, n$ . Nach Satz 8.8 ist  $B = SAS^{-1}$ , und nach Satz 11.4 ist  $S$  orthogonal.

„ $\Leftarrow$ “ Sei  $A \in \mathbb{R}^{n \times n}$ ,  $S \in \mathbb{R}^{n \times n}$  orthogonal und  $SAS^{-1}$  eine Diagonalmatrix. Dann ist  $\underline{SAS^{-1}} = (SAS^{-1})^T = (S^{-1})^T A^T S^T = \underline{SA^T S^{-1}}$ , also  $A = A^T ((S^{-1})^T = S$  und  $S^T = S^{-1})$ .

Bem.: Die Sätze 12.8 und 12.9 bezeichnet man manchmal als Hauptachsentransformation.

Bsp.: Gegeben sei die symmetrische Matrix  $A = \begin{pmatrix} 2 & -1 & 2 \\ -1 & 2 & 2 \\ 2 & 2 & -1 \end{pmatrix}$ . Wir suchen eine orthogonale Matrix  $S$  mit der Eigenschaft, daß  $SAS^{-1}$  eine Diagonalmatrix ist. Wir berechnen zunächst die Eigenwerte.  $\begin{vmatrix} r-2 & 1 & -2 \\ 1 & r-2 & -2 \\ -2 & -2 & r+1 \end{vmatrix} = \dots = (r-3)^2(r+3)$ . Daher sind 3 und -3 die einzigen Eigenwerte von  $A$ .

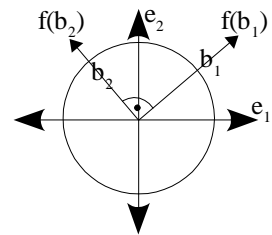
Wir betrachten den euklidischen Vektorraum  $V = \mathbb{R}^n$  mit dem Standardskalarprodukt und die lineare Abbildung  $f: V \rightarrow V$  mit Matrix  $A$  bezüglich der Standardbasis  $e_1, e_2, e_3$ . Wegen  $A^T = A$  ist  $f$  selbstadjungiert. Wir berechnen  $E_{-3}(f)$ . Die Berechnung der Basisvektoren erfolgt wie im vorigen Beispiel. Der Vektor  $(-1; -1; 2)$  bildet also eine Basis von  $E_{-3}(f)$ .

Normierung ergibt:  $b_1 = \frac{1}{\sqrt{6}} (-1; -1; 2)$ . Analog berechnen wir  $E_3(f)$ . Eine Basis von  $E_3(f)$  bilden z.B. die Vektoren  $(2; 0; 1)$  und  $(-1; 1; 0)$ . Auf diese wenden wir das GS-Verfahren an und erhalten  $b_2 = \frac{1}{\sqrt{5}} (2; 0; 1)$  und  $b_3 = \frac{1}{\sqrt{30}} (-1; 5; 2)$ . Die Vektoren  $b_1, b_2, b_3$  bilden also eine Orthonormalbasis von  $V$  (Probe!).

Daher ist die Matrix  $T = \begin{pmatrix} \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{30}} \\ \frac{-1}{\sqrt{6}} & 0 & \frac{5}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \end{pmatrix}$  orthogonal mit  $AT = T \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ , d.h.  $T^{-1}AT = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ . Die gesuchte Matrix erhält man also durch  $S = T^{-1} = T^T = \begin{pmatrix} \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ \frac{-1}{\sqrt{30}} & \frac{5}{\sqrt{30}} & \frac{2}{\sqrt{30}} \end{pmatrix}$ .

07.02.01

Die Hauptachsentransformation kann man sich bildlich wie rechts gezeigt vorstellen( bildlicher Nachtrag).



12.10 Satz: (QR-Zerlegung)

Zu jeder invertierbaren Matrix  $A \in \mathbb{R}^{n \times n}$  existiert eine orthogonale Matrix  $Q \in \mathbb{R}^{n \times n}$  und eine obere Dreiecksmatrix  $R \in \mathbb{R}^{n \times n}$  ( $A, Q$  und  $R$  besitzen die gleiche Größe) mit lauter positiven Elementen auf der Hauptdiagonalen und  $A = QR$ , dabei sind  $Q$  und  $R$  eindeutig bestimmt.

Bew.: Wir versehen den  $\mathbb{R}^n$  (Spaltenvektoren) mit dem Standardskalarprodukt und wenden auf die Spalten  $a_1, \dots, a_n$  von  $A$  das GS-Verfahren an. Wir erhalten eine Orthonormalbasis  $b_1, \dots, b_n$  von

$\mathbb{R}^n$  mit  $\text{Span}(b_1, \dots, b_i) = \text{Span}(a_1, \dots, a_i)$  für  $i = 1; \dots; n$ . Daher können wir schreiben:

$$\begin{aligned} a_1 &= r_{11} b_1 \\ a_2 &= r_{12} b_1 + r_{22} b_2 \\ &\vdots \\ a_n &= r_{1n} b_1 + \dots + r_{nn} b_n \end{aligned}$$

Das GS-Verfahren zeigt auch, daß  $r_{11}, r_{22}, \dots, r_{nn}$  tatsächlich positiv sind. Wir bezeichnen mit  $Q$  die Matrix mit den Spalten  $b_1, \dots, b_n$ . Dann ist  $Q$  orthogonal und  $A = QR$  mit  $R$  wie rechts.

Zum Beweis der Eindeutigkeit sei auch  $A = Q_1 R_1$  mit einer orthogonalen Matrix  $Q_1$  und einer oberen Dreiecksmatrix  $R_1$ , die auf der Hauptdiagonalen lauter positive Zahlen hat. Dann ist  $Q^{-1}Q_1 = R R_1^{-1}$  orthogonal und eine obere Dreiecksmatrix mit lauter positiven Zahlen auf der Hauptdiagonalen.

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & r_{nn} \end{pmatrix}$$

$$\begin{pmatrix} * & \dots & * \\ 0 & \dots & \\ \vdots & & \\ 0 & \dots & 0 & * \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & 1 \end{pmatrix} \text{ also ist } R R_1^{-1} = 1_n, \text{ d.h. } R_1 = R \text{ und } Q_1 = Q.$$

Bsp.:  $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$  Spalten:  $a_1 = (1; 1; 1)$ ,  $a_2 = (1; 2; 1)$ ,  $a_3 = (2; 3; 1)$ .

Das GS-Verfahren liefert  $b_1 = \frac{1}{\sqrt{3}} (1; 1; 1)$ ,  $b_2 = \frac{1}{\sqrt{6}} (-1; 2; -1)$ ,  $b_3 = \frac{1}{\sqrt{2}} (1; 0; -1)$ .

Schreibe:  $a_1 = r_{11} b_1$ ,  $a_2 = r_{12} b_1 + r_{22} b_2$ ,  $a_3 = r_{13} b_1 + r_{23} b_2 + r_{33} b_3$ .

$r_{11} = (a_1 | b_1) = \frac{1}{\sqrt{3}}$ ,  $r_{12} = (a_2 | b_1) = \frac{4}{\sqrt{3}}$ ,  $r_{13} = (a_3 | b_1) = 2\sqrt{3}$ ,  $r_{22} = (a_2 | b_2) = \frac{\sqrt{6}}{3}$ ,  $r_{23} = (a_3 | b_2) = \frac{\sqrt{6}}{2}$ ,  $r_{33} = (a_3 | b_3) = \frac{\sqrt{2}}{2}$ .

QR-Zerlegung:  $\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & \frac{4}{\sqrt{3}} & 2\sqrt{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{2} \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$  (Probe !)

**12.11 Bem.:** Als Nullstellen von Polynomen sind Eigenwerte oft nicht exakt zu berechnen. Der folgende Satz gibt Auskunft über die ungefähre Lage der Eigenwerte.

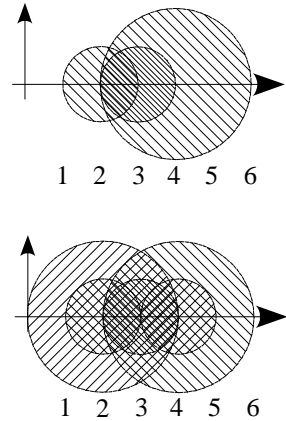
**Satz:** (Gerschgorin)

Gegeben sei eine komplexe  $n \times n$ -Matrix  $A=(a_{ij})$ , für  $i=1;\dots;n$  setzen wir  $r_i:=|a_{i1}|+\dots+|a_{i,i-1}|+|a_{i,i+1}|+\dots+|a_{in}|$  und  $D_i:=\{z \in \mathbb{C} \mid |z-a_{ii}|\leq r_i\}$ . Dann gehört jeder Eigenwert von  $A$  zu  $D_1 \cup \dots \cup D_n$ .

**Bsp.:**  $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 1 & 0 & 3 \end{pmatrix}$   $a_{11}=2, r_1=1+0=1; a_{22}=4, r_2=1+1=2; a_{33}=3, r_3=1+0=1.$

Jeder Eigenwert von  $A$  liegt also in einem der drei „Gerschgorin-Kreise“. Man sieht insbesondere, daß 0 kein Eigenwert von  $A$  ist. Folglich ist  $A$  invertierbar. Bekanntlich haben  $A$  und  $A^T$  die gleichen Eigenwerte. Man kann also das gleiche Spiel mit  $A^T$  machen:  $a_{11}=2, r_1'=2; a_{22}=4, r_2'=1; a_{33}=3, r_3'=1.$

Jeder Eigenwert von  $A$  liegt also auch in einem dieser drei „Gerschgorin-Kreise“. Durch Betrachtung von  $A^T$  erhält man also eventuell Zusatzinformationen. Insgesamt liegt jeder Eigenwert von  $A$  also in dem Durchschnitt der Vereinigungen der Gerschgorin-Kreise von  $A$  und  $A^T$  (Hier nur die Bereiche, wo die Schraffur ein Netz ergibt).



**Bew.:** Sei  $c \in \mathbb{C}$  ein Eigenwert von  $A$ . Wir bezeichnen mit  $f:\mathbb{C}^n \rightarrow \mathbb{C}^n$  die lineare Abbildung mit Matrix  $A$  bezüglich der Standardbasis. Dann ist  $c$  auch Eigenwert von  $f$ , und wir bezeichnen mit  $v=(v_1,\dots,v_n) \in \mathbb{C}^n$  einen entsprechenden Eigenvektor. Wegen  $cv=f(v)=A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$  ist

$cv_i=a_{i1}v_1+\dots+a_{in}v_n$ , also  $(c-a_{ii})v_i=a_{i1}v_1+\dots+a_{i,i-1}v_{i-1}+a_{i,i+1}v_{i+1}+\dots+a_{in}v_n$  und  $|c-a_{ii}| \cdot |v_i| \leq |a_{i1}| \cdot |v_1|+\dots+|a_{i,i-1}| \cdot |v_{i-1}|+|a_{i,i+1}| \cdot |v_{i+1}|+\dots+|a_{in}| \cdot |v_n|$  ( $i=1;\dots;n$ ).

Wir wählen  $i$  so, daß  $|v_i|$  maximal wird. Dann ist  $|c-a_{ii}| \cdot |v_i| \leq (|a_{i1}|+\dots+|a_{i,i-1}|+|a_{i,i+1}|+\dots+|a_{in}|) \cdot |v_i| = r_i \cdot |v_i|$ , d.h.  $|c-a_{ii}| \leq r_i$  und  $c \in D_i$ .

08.02.01

**13. Matrixzerlegungen**

**13.1 Satz:** (Polarzerlegung)

Sei  $V$  ein endlich-dimensionaler euklidischer Vektorraum und sei  $f:V \rightarrow V$  linear und bijektiv. Dann ist  $f=goh$  mit einer orthogonalen Transformation  $g:V \rightarrow V$  und einer selbstadjungierten linearen Abbildung  $h:V \rightarrow V$  mit lauter positiven Eigenwerten. Dabei sind  $g$  und  $h$  eindeutig bestimmt.

**Bew.:** Nach Satz 11.6(IV) ist mit  $f$  auch die adjungierte Abbildung  $f^*:V \rightarrow V$  bijektiv. Daher ist  $F=f^*of:V \rightarrow V$  eine bijektive lineare Abbildung mit  $F^*=(f^*of)^*=f^*of^{**}=f^*of=F$ , d.h.  $F$  ist selbstadjungiert. Nach der Hauptachsentransformation existiert eine Orthonormalbasis  $b_1,\dots,b_n$  von  $V$ , die aus Eigenvektoren von  $f$  besteht. Für  $i=1;\dots;n$  ist also  $F(b_i)=r_i b_i$  mit  $r_i \in \mathbb{R}$ . Da  $f$  bijektiv ist, ist  $f(b_i) \neq 0$ , also  $0 < (f(b_i)|f(b_i))=(f^*(f(b_i))|b_i)=(F(b_i)|b_i)=r_i(b_i|b_i)=r_i$ , d.h.  $r_i > 0$ . Nach Satz 7.6 existiert genau eine lineare Abbildung  $h:V \rightarrow V$  mit  $h(b_i)=\sqrt{r_i} b_i$  ( $i=1;\dots;n$ ). Die Matrix von  $h$  bezüglich

$b_1,\dots,b_n$  ist  $H= \begin{pmatrix} \sqrt{r_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{r_n} \end{pmatrix}$ . Da  $H$  symmetrisch ist, ist  $h$  selbstadjungiert( und bijektiv). Für

$i=1;\dots;n$  ist  $h^2(b_i)=h(h(b_i))=h(\sqrt{r_i} b_i)=\sqrt{r_i} h(b_i)=\sqrt{r_i^2} b_i=r_i b_i=F(b_i)$ . Also ist tatsächlich  $h^2=F$ . Ferner ist  $g=f \circ h^{-1}:V \rightarrow V$  linear und bijektiv mit:

$$g^{-1}=(f \circ h^{-1})^{-1}=(h^{-1})^{-1} \circ f^{-1}=h \circ f^{-1}=h^{-1} \circ h^2 \circ f^{-1}=h^{-1} \circ F \circ f^{-1}=h^{-1} \circ f^* \circ f \circ f^{-1}=(h^*)^{-1} \circ f^*=(h^{-1})^* \circ f^*=(f \circ h^{-1})^*=g^*.$$

Daher ist  $g$  eine orthogonale Transformation und es ist  $f=g \circ h$ . Damit ist die Existenz bewiesen.

Zum Beweis der Eindeutigkeit sei  $g_1:V \rightarrow V$  eine weitere orthogonale Transformation und  $h_1:V \rightarrow V$  eine weitere selbstadjungierte lineare Abbildung mit lauter positiven Eigenwerten mit  $f=g_1 \circ h_1$ . Dann ist  $h_1^2=h_1^* \circ h_1=(g_1^{-1} \circ f)^* \circ (g_1^{-1} \circ f)=f^* \circ (g_1^{-1})^* \circ g_1 \circ f=f^* \circ g_1^{-1} \circ g_1 \circ f=f$ .

Da  $h_1$  selbstadjungiert ist, ist  $h_1$  diagonalisierbar, d.h.  $V$  ist direkte Summe der Eigenräume von  $h_1$ :  $V=E_{s_1}(h_1) \oplus \dots \oplus E_{s_k}(h_1)$  ( $0 < s_1 < \dots < s_k$ ).

Für  $j=1;\dots;k$  und  $v \in E_{s_j}(h_1)$  ist  $F(v)=h_1^2(v)=h_1(h_1(v))=h_1(s_j \cdot v)=s_j \cdot h_1(v)=s_j^2 \cdot v$ , d.h.  $v \in E_{s_j^2}(F)$ . Also ist  $E_{s_j}(h_1) \subseteq E_{s_j^2}(F)$ . Aus Dimensionsgründen folgt  $E_{s_j}(h_1)=E_{s_j^2}(F)$ . Analog ist  $E_{s_j^2}(F)=E_{s_j}(h)$ . Also gilt für  $v \in E_{s_j}(h_1)$ :  $h_1(v)=s_j \cdot v=h(v)$ . Daher gilt  $h=h_1$ . Folglich ist auch  $g=g_1$ .

**Bem.:** Der Beweis liefert folgendes Verfahren zur Bestimmung der polaren Zerlegung:

1. Setze:  $F=f^* \circ f$ .
2. Finde lineare Abbildung  $h:V \rightarrow V$  mit  $h(E) \subseteq E$  für jeden Eigenraum  $E$  von  $F$  und  $h^2=F$  (Ha-trans.)
3. Setze  $g=f \circ h^{-1}$ .

**13.2 Satz:** (Polarzerlegung für Matrizen)

Zu jeder invertierbaren Matrix  $A \in \mathbb{R}^{n \times n}$  existieren eine orthogonale Matrix  $B \in \mathbb{R}^{n \times n}$  und eine symmetrische Matrix  $C \in \mathbb{R}^{n \times n}$  mit lauter positiven Eigenwerten und der Eigenschaft  $A=BC$ . Dabei sind  $B$  und  $C$  eindeutig bestimmt.

**Bew.:** folgt durch Übersetzung in die Matrzensprache.

**Bem.:** Die Polarzerlegung ist wichtig für die Numerik.

**Bsp.:** Gesucht ist die Polarzerlegung von  $A= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 5 & 5 & -5 & -5 \\ 5 & -5 & 5 & -5 \\ 5 & -5 & -5 & 5 \end{pmatrix}$ . Wir berechnen zuerst die symme-

trische Matrix  $X=A^T A= \begin{pmatrix} 76 & -24 & -24 & -24 \\ -24 & 76 & -24 & -24 \\ -24 & -24 & 76 & -24 \\ -24 & -24 & -24 & 76 \end{pmatrix}$ . Mit  $X$  machen wir jetzt eine Hauptachsentransfor-

mation. Die Eigenwerte von  $X$  sind 100 und 4. Die entsprechenden Eigenräume sind:

$$E_{100}=\mathbb{R} \cdot (1;-1;0;0)^T + \mathbb{R}(1;0;-1;0)^T + \mathbb{R}(1;0;0;-1)^T, E_4=\mathbb{R}(1;1;1;1)^T.$$

Das GS-Verfahren liefert Orthonormalbasen für die Eigenräume:

$$E_{100}: b_1= \frac{1}{\sqrt{2}} (1;-1;0;0)^T, b_2= \frac{1}{\sqrt{6}} (1;1;-2;0)^T, b_3= \frac{1}{\sqrt{12}} (1;1;1;-3)^T \text{ und } E_4: b_4= \frac{1}{2} (1;1;1;1)^T.$$

Dann ist  $X(b_1,b_2,b_3,b_4)=(100b_1,100b_2,100b_3,4b_4)=(b_1,b_2,b_3,b_4) \cdot \begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$  und  $X=Y \cdot$

$$\begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} Y^{-1} \text{ mit } Y=(b_1 b_2 b_3 b_4). \text{ Folglich ist } C=Y \begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} Y^T= \begin{pmatrix} 8 & -2 & -2 & -2 \\ -2 & 8 & -2 & -2 \\ -2 & -2 & 8 & -2 \\ -2 & -2 & -2 & 8 \end{pmatrix}$$

symm. mit lauter positiven Eigenwerten( hier 10 und 2) und  $B=AC^{-1}= \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$  ist

orthogonal. Wir erhalten also die Polarzerlegung  $A=BC$  (Probe !).

**13.3 Satz:** Gegeben seien endlich-dimensionale euklidische Vektorräume  $V,W$  und eine lineare Abbildung  $f:V \rightarrow W$ . Dann hat die Matrix von  $f$  bezüglich geeigneter Orthonormalbasen

von  $V$  und  $W$  die folgende Form  $A= \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$  mit einer invertierbaren Matrix  $B \in \mathbb{R}^{r \times r}$  ( $r=rg f$ ).

Bem.: In Satz 3.6 hatten wir gezeigt, daß man Basen von  $V$  und  $W$  stets so wählen kann, daß die entsprechende Matrix die Form  $\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$  hat.

Bew.: Wir schreiben  $V=(\text{Ker } f)\oplus(\text{Ker } f)^\perp$  und  $W=(\text{Bld } f)\oplus(\text{Bld } f)^\perp$ . Dann ist  $\dim (\text{Ker } f)^\perp = \dim V - \dim (\text{Ker } f) = \dim (\text{Bld } f) = \text{rg } f = r$ .

Wir wählen Orthonormalbasen  $b_1, \dots, b_r$  von  $(\text{Ker } f)^\perp$ ,  
 $b_{r+1}, \dots, b_m$  von  $\text{Ker } f$ ,  
 $c_1, \dots, c_r$  von  $\text{Bld } f$  und  
 $c_{r+1}, \dots, c_n$  von  $(\text{Bld } f)^\perp$ .

Dann sind  $b_1, \dots, b_m$  und  $c_1, \dots, c_n$  Orthonormalbasen von  $V$  bzw.  $W$ . Die Abbildung  $g: (\text{Ker } f)^\perp \rightarrow \text{Bld } f, x \rightarrow f(x)$  ist linear und bijektiv.

Daher ist die Matrix  $B$  von  $g$  bezüglich  $b_1, \dots, b_r$  und  $c_1, \dots, c_r$  invertierbar und die Matrix  $A$  hat dann die angegebene Form.

13.4 Satz: (Singularwertzerlegung)

Zu jeder Matrix  $A \in \mathbb{R}^{m \times n}$  existieren orthogonale Matrizen  $S \in \mathbb{R}^{m \times m}$ ,  $T \in \mathbb{R}^{n \times n}$  und positive  $d_1, \dots, d_r \in \mathbb{R}$  mit  $SAT = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_r & 0 \\ 0 & 0 & 0 \end{pmatrix}$ . Dabei sind  $d_1, \dots, d_r$  bis auf die Reihenfolge eindeutig bestimmt.

Def.: Man nennt  $d_1, \dots, d_r$  die Singularwerte von  $A$ .

Bem.: In Satz 3.8 hatten wir gezeigt, daß stets invertierbare Matrizen  $P, Q$  ex. mit  $PAQ = \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$ .

Bew.: Übersetzt man Satz 13.3 in die Matrixsprache, so erhält man orthogonale Matrizen  $P \in \mathbb{R}^{m \times m}$ ,  $Q \in \mathbb{R}^{n \times n}$  mit  $PAQ = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$ ,  $B \in \mathbb{R}^{r \times r}$  ist invertierbar und  $r = \text{rg}(A)$ .

Die Polarzerlegung von  $B$  liefert eine orthogonale Matrix  $X \in \mathbb{R}^{r \times r}$  und eine symmetrische Matrix  $Y \in \mathbb{R}^{r \times r}$  mit lauter positiven Eigenwerten und mit  $B = XY$ .

Die Hauptachsentransformation von  $Y$  ergibt eine orthogonale Matrix  $Z \in \mathbb{R}^{r \times r}$  und eine Diagonalmatrix  $D \in \mathbb{R}^{r \times r}$  mit  $Y = ZDZ^{-1}$ .

Die Diagonalelemente von  $D$  sind genau die Eigenwerte von  $D$ , also auch die Eigenwerte von  $Y$  und damit positiv.

Ergebnis:  $PAQ = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} XY & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} XZDZ^{-1} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} XZ & 0 \\ 0 & 1_{m-r} \end{pmatrix} \cdot \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} Z^{-1} & 0 \\ 0 & 1_{n-r} \end{pmatrix}$ .

Damit ist die Existenz bewiesen.

14.02.01

Zum Beweis der Eindeutigkeit sei  $SAT = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$  mit orthogonalen Matrizen  $S \in \mathbb{R}^{m \times m}$ ,  $T \in \mathbb{R}^{n \times n}$  und einer Diagonalmatrix  $D \in \mathbb{R}^{r \times r}$  mit lauter positiven Diagonalelementen  $d_1, \dots, d_r$ . Dann ist

$\begin{pmatrix} D^2 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}^T = SAT \cdot T^T A^T S^T = SAA^T S^{-1}$ , dabei ist  $D^2$  die Diagonalmatrix mit Diagonalelementen  $d_1^2, \dots, d_r^2$ . Daher sind  $d_1^2, \dots, d_r^2$  die von 0 verschiedenen Eigenwerte von  $AA^T$ . Damit sind  $d_1, \dots, d_r$  durch  $A$  bis auf die Reihenfolge eindeutig bestimmt.

13.5 Satz.: Gegeben seien endlich-dimensionale euklidische Vektorräume  $V, W$  und eine lineare Abbildung  $f: V \rightarrow W$ . Dann hat die Matrix von  $f$  bezüglich geeignet gewählter Orthonormalbasen

von  $V$  und  $W$  die Form  $A = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_r & 0 \\ 0 & 0 & 0 \end{pmatrix}$  mit positiven  $d_1, \dots, d_r \in \mathbb{R}$ , die durch  $f$  bis auf die Reihenfolge eindeutig bestimmt sind.

Bem.: Dies verbessert Satz 13.3.

Bew.: folgt unmittelbar aus Satz 13.4, indem man von Matrizen zu linearen Abbildungen übergeht.

Bsp.: (I) Gesucht ist die Singulärwertzerlegung von  $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 5 & 5 & -5 & -5 \\ 5 & -5 & 5 & -5 \\ 5 & -5 & -5 & 5 \end{pmatrix}$ . In Beispiel 13.2 hatten wir geschrieben:  $A=BC=BYDY^T$ ; dabei ist

$$BY = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{3}} & 0 \end{pmatrix} =: S \text{ orthogonal, } D = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \text{ diagonal mit positiver Hauptdiagonale und } Y^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-3}{\sqrt{12}} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} =: T \text{ orthogonal.}$$

Damit haben wir die Singulärwertzerlegung  $A=SDT$  (Probe!).

(II) Gesucht ist die Singulärwertzerlegung von  $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$ . Zunächst bestimmen wir die Lösungsmenge  $L$  von  $Ax=0$ . Offenbar ist  $L=\mathbb{R}(1,-1)^T$ . Das GS-Verfahren liefert die Orthonormalbasis  $b_1 = \frac{1}{\sqrt{2}} (1,-1)$  von  $L$ .

Als nächstes berechnen wir  $L^\perp = \{(y_1, y_2)^T \mid y_1 - y_2 = 0\} = \mathbb{R}(1,1)^T$ .

Das GS-Verfahren liefert die Orthonormalbasis  $b_2 = \frac{1}{\sqrt{2}} (1,1)^T$  von  $L^\perp$ . Dann ist  $B := (b_2, b_1) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$  orthogonal mit  $AB = \begin{pmatrix} \sqrt{2} & 0 \\ \sqrt{2} & 0 \\ \sqrt{2} & 0 \end{pmatrix}$ . Die von 0 verschiedenen Spalten werden mit GS

orthonormalisiert. Dies ergibt hier  $c_1 = \frac{1}{\sqrt{3}} (1,1,1)^T$ .

Zu dem so aufgespannten Untervektorraum  $M$  wird der Orthogonal-Raum  $M^\perp$  bestimmt.

$M^\perp = \{(z_1, z_2, z_3)^T \mid z_1 + z_2 + z_3 = 0\} = \mathbb{R}(1,-1,0)^T + \mathbb{R}(1,0,-1)^T$ .

Das GS-Verfahren liefert folgende Orthonormalbasis von  $M^\perp$ :  $c_2 = \frac{1}{\sqrt{2}} (1,-1,0)^T$ ,  $c_3 = \frac{1}{\sqrt{6}}$

$(1,1,-2)^T$ . Wir erhalten  $\begin{pmatrix} \sqrt{2} & 0 \\ \sqrt{2} & 0 \\ \sqrt{2} & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} =_D$ .

Damit haben wir die Singulärwertzerlegung  $A=CDB^{-1}=CDB^T$  (Probe!).

13.6 Def.: Gegeben seien Matrizen  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{n \times m}$  mit folgenden Eigenschaften:

(I)  $ABA=A$  und  $BAB=B$ .

(II)  $AB$  und  $BA$  sind symmetrisch.

Dann nennt man  $B$  eine Moore-Penrose-Inverse von  $A$ .

Bem.: R. Penrose, Physiker und Mathematiker):

- Computerdenken, Spektrum Akad. Verlag 1991

- ( mit Stephen Hawking) Raum und Zeit, Rowohlt 2000

**Satz:** Sei  $A \in \mathbb{R}^{m \times n}$ , und sei  $B \in \mathbb{R}^{n \times m}$  eine Moore-Penrose-Inverse von A. Für orthogonale Matrizen  $S \in \mathbb{R}^{m \times m}$ ,  $T \in \mathbb{R}^{n \times n}$  ist dann  $T^{-1}BS^{-1}$  eine Moore-Penrose-Inverse von SAT.

**Bew.:**  $SAT \cdot T^{-1}BS^{-1} = S(AB)S^T$  und  $T^{-1}BS^{-1} \cdot SAT \cdot T^{-1}BS^{-1} = T^T(BA)T$  sind symmetrisch. Ferner ist  $SAT \cdot T^{-1}BS^{-1} \cdot SAT = S(ABA)T = SAT$  und  $T^{-1}BS^{-1} \cdot SAT \cdot T^{-1}BS^{-1} = T^{-1}(BAB)S^{-1} = T^{-1}BS^{-1}$ .

**Bsp.:** A invertierbar  $\Leftrightarrow A^{-1}$  ist Moore-Penrose-Inverse von A.

**13.7 Satz:** Jede Matrix  $A \in \mathbb{R}^{m \times n}$  besitzt genau eine Moore-Penrose-Inverse B.

**Def.:** Man schreibt  $B =: A^\#$ .

**Bew.:** Nach Satz 13.4 existieren orthogonale Matrizen  $S \in \mathbb{R}^{m \times m}$ ,  $T \in \mathbb{R}^{n \times n}$  mit der Eigenschaft, daß SAT die Form  $SAT = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} =: A'$  hat, wobei D eine Diagonalmatrix mit lauter positiven Eigenwerten ist. Nach Satz 13.6 genügt zu zeigen, daß A' genau eine Moore-Penrose-Inverse B' hat. Dazu macht man den Ansatz  $B' = \begin{pmatrix} U & V \\ W & X \end{pmatrix}$ .

Dann ist  $A'B' = (A'B')^T = B'^T A'^T = \begin{pmatrix} U^T & W^T \\ V^T & X^T \end{pmatrix} \cdot \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} U^T D & 0 \\ V^T D & 0 \end{pmatrix}$  und  $A'B' = \begin{pmatrix} DU & DV \\ 0 & 0 \end{pmatrix}$ , also  $DV=0$ . Da D invertierbar ist, folgt  $V=0$ . Analog erhält man  $W=0$  aus  $B'A' = (B'A')^T$ . Also ist  $A' = A'B'A' = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} U & 0 \\ 0 & X \end{pmatrix} \cdot \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} DUD & 0 \\ 0 & 0 \end{pmatrix}$ , also  $D=DUD$ . Da D invertierbar ist, folgt  $U=D^{-1}$ . Ferner ist  $B' = B'A'B' = \begin{pmatrix} U & 0 \\ 0 & X \end{pmatrix} \cdot \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} U & 0 \\ 0 & X \end{pmatrix} = \begin{pmatrix} UDU & 0 \\ 0 & 0 \end{pmatrix}$ , also  $X=0$ , d.h.  $B' = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ . Damit ist die Eindeutigkeit von B' bewiesen. Umgekehrt zeigt man leicht, daß  $B' = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}$  die gewünschten Eigenschaften hat.

**Bem.:** (I) Der Beweis liefert folgendes Verfahren zur Berechnung von  $A^\#$ :

Schreibe  $A = S \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} T$ , wobei  $S \in \mathbb{R}^{m \times m}$  orthogonal,  $T \in \mathbb{R}^{n \times n}$  orthogonal und D Diagonalmatrix mit lauter positiven Eigenwerten ist. Dann ist  $A^\# = T^{-1} \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} S^{-1}$ .

(II) Offenbar ist stets  $(A^\#)^\# = A$ . Dagegen ist im Allgemeinen  $(A_1 A_2)^\# \neq A_2^\# A_1^\#$ .

**Bsp.:** Gesucht ist die Moore-Penrose-Inverse von  $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$ . In Bsp. 13.5 hatten wir gesehen:

$$A = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix} = C \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = D \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = B \begin{pmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = C^T \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \quad (\text{Probe !}).$$

**13.8 Satz:** Für  $A \in \mathbb{R}^{m \times n}$ , und  $b \in \mathbb{R}^{m \times 1}$  gilt:

- (I) Das lineare Gl.-system (\*)  $Ax=b$  ist genau dann lösbar, wenn  $A^\#b$  eine Lösung von (\*) ist.
- (II) Gegebenenfalls hat (\*) die Lösungsmenge  $L = \{A^\#b + y - A^\#Ay \mid y \in \mathbb{R}^{n \times 1}\}$ .
- (III)  $A^\#b$  ist die Lösung von (\*) mit dem kleinsten Betrag [ Dabei betrachtet man das Standardskalarprodukt auf  $\mathbb{R}^{n \times 1}$  ].

**Bew.:** (I) „ $\Rightarrow$ “ Sei  $x \in \mathbb{R}^{n \times 1}$  mit  $Ax=b$ . Dann ist  $AA^\#b = AA^\#Ax = Ax = b$ , d.h.  $A^\#b$  ist Lösung von (\*). „ $\Leftarrow$ “ trivial.

15.02.01

(II) Für  $y \in \mathbb{R}^{n \times 1}$  gilt nach (I):  $A[A^\#b+y-A^\#Ay]=AA^\#b+Ay-\underbrace{AA^\#Ay}_=A=b$ . Sei umgekehrt  $x$  eine beliebige Lösung von (\*), d.h.  $Ax=b$ . Nach (I) ist also  $A(\underbrace{x-A^\#b}_=y)=Ax-AA^\#b=b-b=0$ . Folglich ist  $x=A^\#b+y=A^\#b+y-A^\#Ay$ .

(III) Für  $y \in \mathbb{R}^{n \times 1}$  ist  $(A^\#b|y-A^\#Ay)=(A^\#b|y)-(A^\#b|A^\#Ay)=(A^\#b|y)-(A^\#AA^\#b|y)=(A^\#b|y)-(A^\#b|y)=0$ . Folglich gilt:  $\|A^\#b+y-A^\#Ay\|^2=\|A^\#b\|^2+\underbrace{\|y-A^\#Ay\|^2}_{\geq 0}$ .

13.9 Satz: (Cholesky-Zerlegung)

Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch mit lauter positiven Eigenwerten. Dann existiert eine obere Dreiecksmatrix  $R \in \mathbb{R}^{n \times n}$  mit lauter positiven Zahlen auf der Hauptdiagonalen und  $A=R^TR$ . Dabei ist  $R$  eindeutig bestimmt.

Bew.: Ha-transformation liefert eine orthogonale Matrix  $S \in \mathbb{R}^{n \times n}$  mit  $S^TAS = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix}$ ; dabei

sind  $d_1, \dots, d_n$  die Eigenwerte von  $A$ , also positiv. Wir setzen  $D := \begin{pmatrix} \sqrt{d_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{d_n} \end{pmatrix}$ .

Dann ist  $DS^T \in \mathbb{R}^{n \times n}$  invertierbar. Die QR-Zerlegung liefert eine orthogonale Matrix  $Q \in \mathbb{R}^{n \times n}$  und eine obere Dreiecksmatrix  $R \in \mathbb{R}^{n \times n}$  mit lauter positiven Zahlen auf der Hauptdiagonalen und  $DS^T=QR$ . Folglich ist  $A=SD^2S^T=(QR)^T(QR)=R^TQ^TQR=R^TR$ .

Damit ist die Existenz der Cholesky-Zerlegung bewiesen. Zum Beweis der Eindeutigkeit sei  $R_1 \in \mathbb{R}^{n \times n}$  eine weitere obere Dreiecksmatrix mit lauter positiven Zahlen auf der Hauptdiagonalen und  $A=R_1^TR_1$ . Dann ist  $\underbrace{RR_1^{-1}}=(\underbrace{R^T}^{-1})R_1^T$  eine Diagonalmatrix.

obere Dreiecksmatrix      untere Dreiecksmatrix  
 Schreibt man  $R=(r_{ij})$ ,  $R_1=(r'_{ij})$ , so gilt für  $i=1; \dots; n$  ferner:  
 $r_{ii}(r'_{ii})^{-1}=r_{ii}^{-1}r'_{ii}$ , d.h.  $r_{ii}^2=(r'_{ii})^2$ .  
 Da  $r_{ii}$  und  $r'_{ii}$  positiv sind, folgt  $r_{ii}=r'_{ii}$ . Also ist  $RR_1^{-1}=1_n$ , d.h.  $R_1=R$ .

Bem.: Schreibt man  $A=(a_{ij})$ ,  $R=(r_{ij})$ , so ist  $a_{ij}=r_{i1}r_{1j}+\dots+r_{ii}r_{ij}$  ( $1 \leq i \leq j \leq n$ ). Damit kann man die Koeffizienten  $r_{ij}$  induktiv berechnen.

Bsp.: Gesucht ist die Cholesky-Zerlegung von  $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$ . Wir machen den Ansatz  $A =$

$$\begin{pmatrix} r_{11} & 0 & 0 \\ r_{12} & r_{22} & 0 \\ r_{13} & r_{23} & r_{33} \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix} = R$$

Vergleich der ersten Zeilen ergibt:  $(1;1;0)=r_{11}(r_{11},r_{12},r_{13})$ . Also ist  $r_{11}=1$ ,  $r_{12}=1$ ,  $r_{13}=0$ .

Vergleich der zweiten Zeilen ergibt:  $(1;2;1)=1(1;1;0)+r_{22}(0,r_{22},r_{23})$ . Also ist  $r_{22}=1$ ,  $r_{23}=1$ .

Vergleich der dritten Zeilen ergibt:  $(0;1;3)=0(1;1;0)+1(0;1;1)+r_{33}(0;0;r_{33})$ . Also ist  $r_{33} = \sqrt{2}$ .

Wir erhalten:  $R = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$ .

13.10 Satz: (Hurwitz-Kriterium)

Sei  $A=(a_{ij}) \in \mathbb{R}^{n \times n}$  symmetrisch und  $A_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$  für  $k=1; \dots; n$ . Es sind genau dann alle Eigenwerte von  $A$  positiv, wenn  $\det(A_k) > 0$  für  $k=1; \dots; n$  ist.

Bew.: „ $\Rightarrow$ “ Seien alle Eigenwerte positiv. Dann besitzt  $A$  eine Cholesky-Zerlegung  $A=R^TR$ . Für  $k=1; \dots; n$  ist also auch  $A_k=R_k^TR_k$ ; insbesondere ist  $\det(A_k)=(\det(R_k))^2 > 0$ .

„ $\Leftarrow$ “ (Induktion nach n).

Im Fall n=1 ist die Sache klar. Sei also n>1. Nach Induktion sind alle Eigenwerte von  $A_{n-1}$  positiv. Hauptachsentransformation liefert also eine orthogonale Matrix  $S \in \mathbb{R}^{(n-1) \times (n-1)}$  mit  $S^{-1}A_{n-1}S =$

$$\begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_{n-1} \end{pmatrix} ; \text{ dabei sind } d_1, \dots, d_{n-1} \text{ positiv. Folglich ist } \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix}^{-1} A \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} S^{-1}A_{n-1}S & * \\ * & * \end{pmatrix}$$

$$=: B = \left( \begin{array}{c|c} d_1 & 0 \\ \hline 0 & d_{n-1} \end{array} \middle| \begin{array}{c} b_1 \\ \vdots \\ b_{n-1} \end{array} \right) \text{ und } \det B = \det A > 0. \text{ Andererseits ist } \det B = d_1 \dots d_{n-1} \left( \sqrt{b_n - \frac{b_1^2}{d_1} - \dots - \frac{b_{n-1}^2}{d_{n-1}}} \right) \text{ wie}$$

man leicht mit vollständiger Induktion zeigt; insbesondere ist  $b_n - \frac{b_1^2}{d_1} - \dots - \frac{b_{n-1}^2}{d_{n-1}} > 0$ . Wir setzen  $s_n := \frac{b_i}{\sqrt{d}}$  für  $i=1; \dots; n-1$  und  $s_n := \sqrt{b_n - \frac{b_1^2}{d_1} - \dots - \frac{b_{n-1}^2}{d_{n-1}}}$ .

Dann ist  $R := \left( \begin{array}{c|c} \sqrt{d_1} & 0 \\ \hline 0 & \sqrt{d_{n-1}} \end{array} \middle| \begin{array}{c} s_1 \\ \vdots \\ s_{n-1} \end{array} \right)$  eine obere Dreiecksmatrix mit  $R^T R = B$ . Ist also  $r \in \mathbb{R}$  ein Eigenwert von B und ist  $v \in \mathbb{R}^{n \times 1}$  mit  $Bv = rv$ , so ist  $rv^T v = v^T Bv = v^T R^T R v = (Rv)^T (Rv) > 0$ , also auch  $r > 0$  (Da  $v^T v > 0$ ). Daher ist jeder Eigenwert von B (und A) positiv.

Bsp.:  $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 4 \end{pmatrix}$ . Wegen  $\det A_1=1, \det A_2=1, \det A_3=2, \det A_4=7$  sind alle Eigenwerte von A positiv.

Nachklausur am Samstag 07.04.01 9<sup>00</sup>-12<sup>00</sup> Uhr im Abbeanum, HS 1.