Lineare Algebra

Lucas Westermann & Florian Scheibner

1. August 2011

Inhaltsverzeichnis

Literatur

Mathematik für Informatiker: Teschl, Hackenberger
Lineare Algebra: Beutelspacher, Fischer, Lang (auf Englisch), Stammbach.

1 Grundlegendes

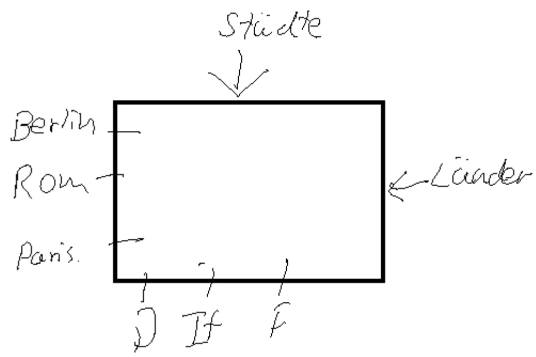
1.1 Mengen

1.1.1 Definition (Relation)

Gegeben seien zwei Mengen X und Y. Eine Teilmenge R des kartesisches Produkts $X \times Y = \{(x, y) : x \in X, y \in Y\}$ heißt Relation (R) zwischen X und Y; im Fall $X = Y$ spricht man von einer Relation auf X. Ferner: $R_1^{-1} = \{(y, x) \in Y \times X : (x, y) \in R\}$ heißt Umkehrrelation.

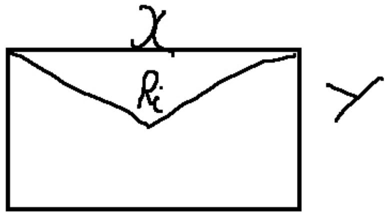
1.1.2 Beispiel

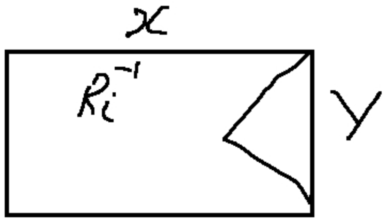
Die Menge $R_0 = \{(x, y) \in X \in Y : y$ ist Hauptstadt von x ist eine Relation zwischen der Menge X aller Länder und Y aller Städte.



1.1.3 Beispiel

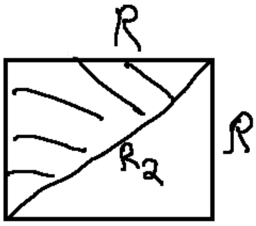
Mit den Mengen $X = \mathbb{R}$ $Y = [0, \infty)$ ist $R_1 = \{(x, |x|) \in X \times Y, X \in X\}$ ist eine Relation mit der Umkehrrelation $R^{-1} = \{(|x|, x) : x \in X\}$.

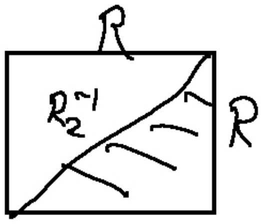




1.1.4 Beispiel

Mit den Mengen $X = Y = \mathbb{R}$ ist $R_2 = \{(x, y) \in X \times Y : x \leq y\}$ eine Relation $R_2^{-1} = (y, x) : x \leq y$



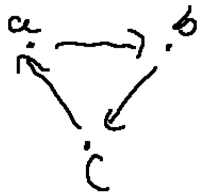


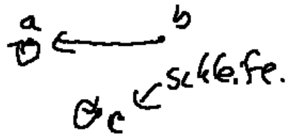
1.1.5 Beispiel

Die Menge $R_3 = \{(x, y) \in C \times C : x$ und y haben gleichen Hersteller$\}$ ist eine Relation auf der Menge aller Computer C.

1.1.6 Beschreibung (Gerichtete Graphen)

Relation R auf endlichen Mengen X können alternative wie folgt dargestellt werden. Man repräsentiert die Elemente von X als Punkte in der Ebene (Knoten) und verbindet $x, y \in X$ genau dann durch einen Pfeil (gerichtete Kante), wenn $(x, y) \in R$. Das paar (X, R) heißt gerichteter Graph oder Digraph, z.B. $X = \{a, b, c\}$ $R = \{(a, b), (b, c), (c, d)\}$.





$X = \{a, b, c\}$ $R = \{(b, a), (a, a), (c, c)\}$.
Eine Relation R auf X heißt
reflexiv $\Leftrightarrow (x, x) \in R$ für alle $x \in X$
transitiv $\Leftrightarrow (x, y) \in R \Rightarrow (x, z) \in R$ für alle $x, y, z \in X$
symmetrisch $\Leftrightarrow (x, y) \in R$ für alle $x, y \in X$

1.1.7 Beispiel

Die Relation R_2 aus Beispiel 1.1.4 ist reflexiv, transitiv, aber nicht Symmetrisch. Die Relation R_3 aus Beispiel 1.1.5 ist reflexiv, transitiv und symmetrisch.

1.1.8 Definition (Äquivalenzrelation)

Eine Relation A auf eine Menge X heißt eine Äquivalenzrelation, falls sie reflexiv, transitiv und symmetrisch ist. Für ein Paar $(x, y) \in A$ Schreiben wir $x{\sim}y$ und nennen x und y äquivalent.

1.1.9 Beispiel

1. Sei X eine beliebige Menge. Dann ist $\{(x, y) \in X \times X : x = y\}$ eine Äquivalenzrelation (Identitätsrelation).

2. Ebenso ist das ganze Produkt $X \times X$ eine Äquivalenzrelation (Allrelation).

3. Die Relation R_3 aus Beispiel 1.1.5 ist eine Äquivalenzrelation. Mit ihr lassen sich Computer nach ihrem Hersteller klassifizieren. Für jedes $[x] := \{y \in X : x{\sim}y\}$ die von X erzeugte Äquivalenzklasse und ein Element $y \in [x]$ heißt Repräsentant von $[x]$.

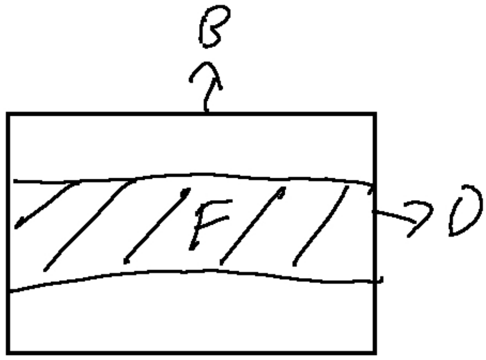
1.1.10 Beispiel

1. Für die Identitätsrelation ist $[x] = \{x\}$ für alle $x \in X$. Die Allrelation besitzt genau eine Äquivalenzklasse $[x] = X$.

2. Im Beispiel 1.1.5 sind die Äquivalenzklassen die Menge aller Hersteller.

1.2 Abbildungen

$F \subseteq D \times B$.

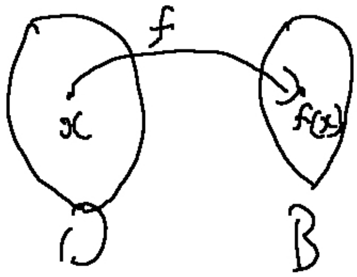


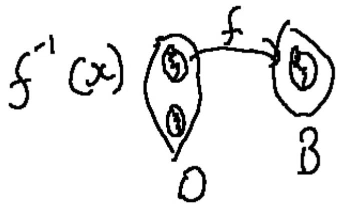
1.2.1 Definition (Abbildungen, Funktion)

Eine Relation F zwischen zwei nichtleeren Mengen D und B heißt Abbildung oder Funktion von D nach B, falls für alle $x \in D$ gilt.
1) Es existiert ein $y \in B$ mit $(x, y) \in F$
2) Mit $y_1, y_2 \in B$ folgt aus $(x, y_1) \in F$ und $(x, y_2) \in F$, dass $y_1 = y_2$.
Die Menge D heißt Definitionsbereich und B Bildbereich von F. Im Fall $D = B$ spricht man von einer Abbildung auf D oder um einer Selbstabbildung auf D.

1.2.2 Bemerkung

Veranschaulicht man Funktionen auf (endlichen) Mengen D als gerichtete Graphen (Beschreibung 1.1.7), so geht von jedem Knoten genau eine Kante ab. Anstelle der Notation $F \subseteq D \times B$, $(x, y) \in F$ schreibt man auch $f : D \to B, x \mapsto f(x)$ oder $y := f(x)$ Mit einer weiteren nichtleeren Menge C und einer Abbildung $g : B \to C$ ist die Verknüpfung (Komposition) von g und f definiert als $g \circ f : D \to C, (g \circ f)(x) := g(f(x))$. Im Fall von Abbildungen f, g auf D gilt i.A. $f \circ g \neq g \circ f$. Statt einzelner Punkte $x \in D$ kann man auch Mengen $X \subseteq D$ abbilden: $f(X) := \{y \in B :$ es gibt ein $x \in X$ mit $y = f(x)\}$. $f(X)$ heißt Bild von X unter f. Das Urbild einer Menge $Y \subseteq B$ ist definiert durch $f^{-1}(Y) := \{x \in D \; f(x) \in Y\}$. Eine Abbildung $f : D \to B$ heißt
injektiv $\Leftrightarrow f^{-1}(\{y\}$ enthält für alle $y \in B$ höchstens ein Element
surjektiv $\Leftrightarrow f^{-1}(\{y\})$ enthält für alle $y \in B$ mindestens ein Element.
bijektiv $\Leftrightarrow f^{-1}(\{y\})$ enthält für alle $y \in B$ genau ein Element.
Eine Abbildung $f : D \to B$ ist genau dann bijektiv, wenn sie injektiv und surjektiv ist.





1.2.3 Beispiel (identische Abbildung)

Die identische Abbildung auf eine Menge $D \neq \emptyset$ ist $id_D : D \to D, id_D(x) := x$. Sie ist bijektiv.

Beispiel

Die Relation R_0 aus Beispiel 1.1.2 zwischen $X = \{Land\}$ und $Y = \{Stadt\}$ ist eine Funktion $r_o : X \to Y$ $r_0(Land) :=$ Hauptstadt vom Land. Ihr Bild ist $r_0(X) = \{\text{Hauptstädte}\}$ und die Urbilder lauten:

$$r_0^{-1}(\{s\}) = \begin{cases} \emptyset & \text{falls } s \text{ keine Hauptstadt,} \\ \{l\} & \text{falls } s \text{ Hauptstadt von } l. \end{cases}$$

Folglich ist r_0 injektiv, aber nicht surjektiv. Betrachtet man die Menge aller Haupstädte als Bildbereich von r_0, so ist diese Abbildung auch surjektiv.

1.2.4 Beispiel

Die Relation R_1 zwischen $\mathbb{R}$ und $[0, \infty)$ aus Beispiel 1.1.3 ist eine Abbildung und lässt sich schreiben als $r_1 : \mathbb{R} \to [0, \infty),\ r_1(x) := |x|$ Für sie gilt $r_1(\mathbb{R}) := [0, \infty)$ und $r_1^{-1}(\{y\}) = \{-y, y\}$ für alle $y \in [0, \infty)$. Also ist $r_1 : \mathbb{R} \to [0, \infty)$ surjektiv, aber nicht injektiv. Betrachten wir r_1 mit ganz $\mathbb{R}$ als Bildbereich, so gilt $r_1^{-1}(\{y\}) = \emptyset$ für $y < 0$ und dann ist r_1 nicht mehr surjektiv.

1.2.5 Beispiel (ASCII-Code)

Der ASCII-Code zur Codierung alpha-numerischer Zeichen ist gegeben durch eine bijektive Abbildung $f : \{0, 1, \cdots, 255 \text{ bzw. } 127\} \to \{\text{Zeichen}\}$.
Einfache Beispiele (etwa Beispiel 1.2.5) zeigen, dass die Umkehrrelation F^{-1} einer Abbildung $F \subseteq D \times B$ bzw. $f : D \to B$ nicht unbedingt eine Abbildung ist.

1.2.6 Definition (Umkehrabbildung)

Eine Abbildung $f : D \to B$ heißt umkehrbar, falls ihre Umkehrrelation F^{-1} wieder eine Abbildung ist. Für letztere schreibt man $f^{-1} : B \to D$ und nennt sie Umkehrabbildung von f.

1.2.7 Bemerkung

Mit einer umkehrbaren Abbildung $f : D \to B$ ist auch ihre Umkehrfunktion $f^{-1} : B \to D$ umkehrbar mit $f^{-1} \circ f = id_D$ und $f \circ f^{-1} = id_B$.

1.2.8 Korollar

Eine Abbildung $f : D \to B$ ist genau dann umkehrbar, wenn f bijektiv ist. Für umgekehrtes f existiert die Umkehrfunktion nur auf $f(D)$.
Beweis: Hausaufgabe

1.3 Matrizen

Wir führen kurz die komplexen Zahlen $\mathbb{C}$ ein. Darunter versteht man alle Paare $z = (x, y)$ reeller Zahlen $x, y \in \mathbb{R}$ mit der Addition:

$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2)$$

und der Multiplikation:

$$z_1 \cdot z_2 := z_1 z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1)$$

wobei $z_1 = (x_1, y_1), z_2 = (x_2, y_2)$
Differenz und Quotient ergeben sich zu:

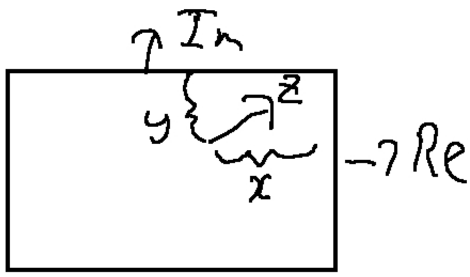
$$z_1 - z_2 = (x_1 - x_2, y_1 - y_2)$$

$$\frac{z_1}{z_2} = \left(\frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2}, \frac{x_2 y_1 - x_1 y_2}{x_2^2 + y_2^2} \right) \text{ falls } x_2^2 + y_2^2 \neq 0$$

Alternative Darstellung:

$$z = (x, y) = x + iy \text{ mit der Konvention } i^2 = -1$$

Wobei x der Realteil ($Rez = x$) ist und y der Imaginärteil ($Imz = y$).



Im Folgenden stehe $\mathbb{K}$ für eine der drei Mengen $\mathbb{Q}$ (rationalen Zahlen), $\mathbb{R}$ (reelle Zahlen) oder $\mathbb{C}$.

1.3.1 Definition (Matrix)

Eine $m \times m$-Matrize ist ein rechteckiges Schema von Zahlen $a_{ij} \in \mathbb{K}$ der Form

$$A = (a_{i,j})1 \leq i \leq m, 1 \leq j \leq m = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,m} \end{pmatrix}$$

Der erste Index $i \in \{1, \cdots, m\}$ nummeriert die m Zeilen, der zweite Index $j \in \{1, \cdots, m\}$ die m Spalten der Matrix A, das Element $a_{ij} \in \mathbb{K}$ steht daher in der i-ten Zeile und der j-ten Spalte. Für die Menge aller solchen Matrizen schreiben wir $\mathbb{K}^{m \times m}$. Für eine quadratische Matrix A gilt $m = n$ und die $a_{i,i}$ heißen Diagonalelement.

$$A' = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

1.3.2 Beispiel (n Tupeln, m-Spalten)

Ein n-Tupel $x = (x_1, \cdots, x_n)$ von Zahlen X aus $\mathbb{K}$ wird als $1 \times m$-Matrix interpretiert. Eine m-Spalte $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ wird als $m \times 1$-Matrix verstanden, identifiziert durch $\mathbb{K}^m = k^{m \times 1}$.

1.3.3 Kronecker-Symbol, Einheits- und Nullmatrixe)

Wir definieren das Kronecker-Symbol $S_{i,j} := \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$ und $I_m := (S_{i,j})_{1 \leq i,j \leq m}$ ist die Einheitsmatrix. Bei der Nullmatrix $0 = (0)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ sind alle Elemente gleich $0 \in \mathbb{N}$.

1.3.4 Beispiel (Diagonal- und Dreieckmatrizen)

Man nennt eine quadratische Matrix $A = (a_{i,j})_{1\leq i,j\leq n}$ diagonal falls $a_{i,j} = 0$ für $i \neq j$. Wir schreiben dann $A = \begin{pmatrix} a_{1,1} & 0 & \cdot & 0 \\ 0 & a_{2,2} & \cdot & 0 \\ 0 & \cdot & \cdot & a_{n,n} \end{pmatrix} = \operatorname{diag}(a_{1,1}, \cdot, a_{n,n})$. Eine obere Dreiecksmatrix ist quadratisch und erfüllt $a_{i,j} = 0$ für $i > j$, wogegen eine untere Dreiecksmatrix $a_{i,j} = 0$ für $i < j$ erfüllt. Sie sind von der Form: $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{pmatrix}$ bzw. $A = \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$.

Mathematische Operationen für Matrizen:

- Skalare Multiplikation: $\mathbb{K} \times \mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n}, \alpha \cdot A = \alpha A = (\alpha a_{i,j})_{\substack{1\leq i\leq m \\ i\leq j\leq n}}$. Wir schreiben $-A := (-1)\cdot A$
- Addition: $+ : \mathbb{K}^{m\times n} \times \mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n}, A + B = (a_{i,j} + b_{i,j})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$. Die Subtraktion lautet $A - B = A + (-B)$.
- Genau für $m \times n$-Matrizen A und $n \times p$-Matrizen B lässt sich eine Multiplikation erklären. $\cdot : \mathbb{K}^{m\times n} \times \mathbb{K}^{n\times p} \to \mathbb{K}^{m\times p}$. $A \cdot B = AB := (\sum_{k=1}^{n} a_{i,k}b_{k,j})_{\substack{1\leq i\leq m \\ 1\leq j\leq p}}$. das Produkt ist also eine $m \times p$-Matrix.

Merke: Das Produkt macht nur Sinn, falls die Spaltenzahl der ersten mit der Zeilenzahl der zweiten Matrix übereinstimmt.

1.3.5 Bemerkung

(1) Um Produkte von Matrizen $A \in \mathbb{K}^{m\times n}$ und $B \in \mathbb{K}^{n\times p}$ zum berechnen ergibt sich das Schema
$\begin{array}{c|c} & B \\ A & C \end{array}$ $C = (\sum_{k=1}^{n} a_{i,k}b_{k,j})_{\substack{1\leq i\leq m \\ 1\leq j\leq p}}$

(2) Spezialfall: $A \in \mathbb{K}^{m\times m}, x \in \mathbb{K}^m$ $Ax = \sum_{k=1}^{m} \begin{pmatrix} a_{1,k} & x_1 \\ \vdots & \vdots \\ a_{m,k} & x_k \end{pmatrix}$.

1.3.6 Beispiel

Das Produkt von $A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$ und $B = \begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix}$ lautet:

$$\begin{array}{c|cc} & 4 & 5 \\ & 6 & 7 \\ \hline & & \\ 0\ 1 & 0 \cdot 4 + 1 \cdot 6 & 5 \cdot 0 + 1 \cdot 7 \\ 2\ 3 & 2 \cdot 4 + 6 \cdot 3 & 2 \cdot 5 + 7 \cdot 3 \end{array}$$

also $C = AB = \begin{pmatrix} 6 & 7 \\ 26 & 31 \end{pmatrix}$.

Im Umgekehrter Reihenfolge gilt $BA = \begin{pmatrix} 10 & 19 \\ 14 & 27 \end{pmatrix}$. Daher ist das Produkt von Matrizen nicht kommutativ $AB \neq BA$.

1.3.7 Beispiel

(1) Für $A \in \mathbb{K}^{m \times n}$ gilt $I_m A = A = A I_m$

(2) Für $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ gilt $AB = 0$, womit das Produkt von Matrizen nicht <u>nullteilerfrei</u> ist, d.h. $AB = 0$ kann gelten, ohne dass ein Faktor Null ist.

(3) Das Produkt von $\begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix}$ und $\begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix}$ ist nicht definiert, $\begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 21 & 34 & 47 \\ 27 & 44 & 61 \end{pmatrix}$ dagegen schon.

1.3.8 Beispiel (RGB - Raum)

Im RGB-Farbmodell werden Farben durch Tupel (r, g, b) reeller Zahlen $r, g, b \in \mathbb{R}$ beschreiben: $(1, 0, 0) =$ rot, $(0, 0, 1)$ blau, $(1, 1, 0)$ gelb. Alternativ: YIQ-Modell (y, i, q).

Umrechnung $\begin{pmatrix} y \\ i \\ q \end{pmatrix} = \begin{pmatrix} 0.3 & 0.6 & 0.1 \\ 0.6 & -0.3 & -0.3 \\ 0.2 & -0.5 & 0.3 \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix}$.

1.3.9 Beispiel (Inzedenzmatrix)

Gerichtete Graphen ohne Schleifen (kein Knoten wird durch eine Kante mit sich selbst verbunden, siehe Beschreibung 1.1.7) mit den Knoten $\hat{1}, \cdots, \hat{m}$ mit den Knoten $1, \cdots, m$ lassen sich durch eine sogenannte Inzedenzmatrix $A \in \mathbb{K}^{m \times n}$ beschreiben mit

$$a_{i,j} = \begin{cases} 1, & \text{Von Knoten } \hat{1} \text{ geht die Kante } j \text{ aus.} \\ -1, & \text{ein Knoten } \hat{1} \text{ mündet die Kante } j \\ 0, & \text{Knoten } \hat{1} \text{ und Kante } j \text{ berühren sich nicht.} \end{cases}$$

$$\begin{tikzcd} \hat{1} \arrow[r, "1"] & \hat{2} \arrow[d, "3"] \\ & \hat{3} \arrow[ul, "2"] \end{tikzcd}$$

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

$$\begin{tikzcd} \hat{1} \arrow[r, "1"] \arrow[dr, "2"] & \hat{2} \\ & \hat{3} \end{tikzcd}$$

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

$$\begin{tikzcd} \hat{1} \arrow[r, "1"] \arrow[d, shift left, "2"] & \hat{2} \\ \hat{3} \arrow[u, shift left, "3"] \arrow[loop below, "4"] & \end{tikzcd}$$

Nicht schleifen frei!

1.3.10 Satz (Rechenregeln für Matrizen)

Für Zahlen $\alpha \in \mathbb{K}$ und Matrizen $A \in \mathbb{K}^{m \times n}, B \in \mathbb{K}^{m \times p}$ gilt das Distributiv-Gesetz. $A(B + C) = AB + AC$ für alle $C \in \mathbb{K}^{m \times p}$ und die Assoziativ-Gesetze $(\alpha A)B = A(\alpha B), A(BC) = (AB)C$ für alle $C \in \mathbb{K}^{p \times q}$. Beweis: Übung.

1.4 Lineare Gleichungen

1.4.1 Definition (lineare Gleichung)

Es seien $A \in \mathbb{K}^{m \times n}$ und $b \in \mathbb{K}^m$. Dann bezeichnet man (L_b) $Ax = b$ als lineares Gleichungssystem mit m Gleichungen für die n unbekannten $x_m \in \mathbb{K}$ oder kurz also lineare Gleichung in $\mathbb{K}^m$. A heißt Koeffizientenmatrix und b Inhomogenität von (L_b). Im Fall $b \neq 0$ nennt man (L_b) inhomogen und erhält andernfalls die homogene Gleichung: (L_0) $Ax = 0$. Eine Lösung von (L_b) ist ein Element $x \in \mathbb{K}^m$ mit $Ax = b$ und $L_b := \{x \in \mathbb{K}^m : Ax = b\}$ steht für die Lösungsmenge von (L_b).

1.4.2 Bemerkung

(1) Ausgeschrieben lautet (L_b):
$a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1$
$a2, 1x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2$
$\cdots$
$am, 1x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = b_m$.
Oder noch unübersichtlicher $\sum_{j=1}^n a_{i,j}x_j = b_i$ für $1 \leq i \leq m$.
(2) (L_b) hat stehts die triviale Lösung $0 \in \mathbb{K}^m$. Inhomogene Gleichungen müssen nicht unbedingt lösbar sein: $0x = 1$.

1.4.3 Satz (Superpositionsprinzip)

Es seien $x, y \in \mathbb{K}^n$ Lösungen von (L_0). Dann ist auch $\alpha x + \beta y$ eine Lösung von (L_0), d.h. $\alpha x + \beta y \in L_0$ für alle $\alpha, \beta \in \mathbb{K}$.
Beweis: Übung.

1.4.4 Satz

Ist $\hat{x} \in \mathbb{K}^n$ eine Lösung von (L_b) so gilt $L_b = \hat{x} + L_0$. Hierbei: Für gegebene $x \in \mathbb{K}^n, A \subseteq \mathbb{K}^n$ ist $x + A := \{y \in \mathbb{K}^n : \text{ es gibt ein } a \in A \text{ mit } y = x + a\}$
Beweis: Übung. Nun: Explizite Lösung von (L_b)!
Besonders einfach, falls $A \in \mathbb{K}^{m \times n}$ diagonal ist gilt nämlich $a_{i,i} \neq 0, 1 \leq i \leq n$, so besitzt (L_b) die eindeutige Lösung $x \in \mathbb{K}^n$ mit Elementen $X_1 = \dfrac{b_i}{a_{i,i}}$ für $1 \leq 1 \leq n$ ist dagegen $d_{i,i} = 0$ für ein $1 \leq i \leq n$, so besitzt (L_b) unendlich viele Lösungen für $b_i = 0$ und anderenfalls keine Lösung.
Allgemeinere Klasse: Ein $A \in \mathbb{K}^{m \times n}$ ist in Zeilen-Stufen-Form (ZSF) falls in jeder Zeile gilt:

(1) Beginnt sie mit k Nullen, so stehen unter diesen Nullen lediglich weitere Nullen.

(2) Unter dem ersten Element $\neq 0$ stehen nur Nullen.

Bei strenger ZSF muss zusätzlich gelten:

(3) Über jedem ersten Element $\neq 0$ stehen nur Nullen

1.4.5 Beispiel

(1) Obere Dreiecksmatrizen sind in ZSF, Diagonalmatrizen sogar in strenger ZSF.

(2) Bezeichnet $*$ ein Element $\neq 0$, so gilt:

- $\begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{pmatrix}, \begin{pmatrix} 0 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}, \begin{pmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{pmatrix}$ sind nicht in ZSF.
- $\begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & 0 & * \end{pmatrix}$ ist in ZSF (aber nicht strenger ZSF).
- $\begin{pmatrix} * & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ ist in strenger ZSF

1.4.6 Beispiel (Rückwärts-Substitution)

Die inhomogene lineare Gleichung $(1.4b)\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1, \\ x_2 + 2x_3 + 3x_4 = 1, \\ x_3 + 2x_4 = 1 \end{cases}$ hat die Koeffizientenmatrix bzw. Inhomogenität $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

<u>Rückwärtssubstitution</u>: Aus der letzten Gleichung $x_3 + 2x_4 = 1$ sieht man, dass $x_4 = t$ frei gewählt wenden kann, $t \in \mathbb{K}$. Dies liefert $x_3 = 1 - 2t$. Die bekannten variablen x_3, x_4 können in die zweite Gleichung von $(1.4b)$ eingesetzt werden, also $x_2 = 1 - 2x_3 - 3x_4 = t - 1$ und analog liefert die erste Gleichung $x_1 = 1 - 2x_2 - 3x_3 - 4x_4 = 0$. Die Lösungsmenge von $(1.4b)$ ist also:

$$L_b = \left\{ \begin{pmatrix} 0 \\ t-1 \\ 1-2t \\ t \end{pmatrix} \in \mathbb{K}^4 : t \in \mathbb{K} \right\} = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \mathbb{K} \begin{pmatrix} 0 \\ 1 \\ -2 \\ 1 \end{pmatrix}$$

Die Lösungsmenge L_b von (L_b) ändert sich nicht, wenn folgende Operationen auf $(1.4b)$ angewandt werden:

- Vertauschen von Gleichungen
- Multiplikation von Gleichungen mit $\alpha \in \mathbb{K}$
  $\{0\}$

- Addition des α-fachen der k-ten Gleichung zur j-ten.
  Diese sind elementare Zeilentransformationen.

ZIEL: Transformiere A bzw. (L_b) auf ZSF mittels elementarer Zeilentransformationen. Systematisch: Gauß Algorithmus.
Zu seiner Beschreibung gehen wir davon aus, dass die erste Spalte von A von 0 verschieden ist (anderenfalls sind $x_1, \cdots, x_n$ umzunummerieren). Ohne Sonderfälle zu berücksichtigen gilt:

1. Ordne die Gleichungen in $(1.4a)$ so an, dass $a_m \neq 0$. In der gängigen Notation schreibt man nun $(1.4b)$ als $\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}$

2. Subtrahiere von der i-ten Gleichung, $2 \leq i \leq m$ in $(1.4a)$ das $\frac{a_{i,1}}{a_{1,1}}$-fache der ersten Gleichung:
$$\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ 0 & a_{2,2}^{(1)} & \cdots & a_{2,n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{m,2}^{(2)} & \cdots & a_{m,n}^{(1)} & b_m^{(1)} \end{array} \quad \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ A^{(1)}x^{(1)} = b^{(1)} \end{cases}$$
mit $A^{(1)} \in \mathbb{K}^{(m-1)\times(n-1)}, b \in \mathbb{K}^{m-1}$.

3. Transformiere $A^{(1)}x^{(1)} = b^{(1)}$ entsprechend und fahre sukzessive fort, bis (idealerweise) eine Dreiecks- oder ZSF entstanden ist.

4. Löse das resultierende System durch Rückwärts-Substitution.

1.4.7 Beispiel

Als Kurzschreibweise für
$$(1.4d) \begin{cases} x_1 + 2x_2 + 3x_3 = 0 \\ 4x_1 + 5x_2 + 6x_3 = 0 \\ 7x_1 + 8x_2 + 9x_3 = 0 \end{cases}$$
$$\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & \\ 4 & 5 & 6 & 0 & II-4I \\ 7 & 8 & 9 & 0 & II-7I \end{array} \Leftrightarrow \begin{array}{ccc|cc} 1 & 2 & 3 & 0 & \\ 0 & -3 & -6 & 0 & (-\frac{1}{3}) \\ 0 & -6 & -12 & 0 & III-2II \end{array} \Leftrightarrow \begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$
Damit ist $(1.4d)$ äquivalent zu $\begin{cases} x_1 + 2x_2 + 3x_3 = 0 \\ x_2 + 2x_3 = 0 \end{cases}$
Rückwärts-Substitution: Wähle $x_3 = t$ mit $t \in \mathbb{K}$ und es folgt $x_2 = -2x_3 = -2t, x_1 = -2x_2 + 3x_3 =$

t. Die Lösungsmenge von $(1.4d)$ ergibt sich zu:

$$L_0 = \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \in \mathbb{K}^3 : t \in \mathbb{K} \right\} = \mathbb{K} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

1.4.8 Satz

Hat (L_0) weniger Gleichungen als Unbekannte, also $m < n$, so besitzt sie unendlich viele Lösungen.
Beweis:

I. Man zeigt (*) (L_0) hat eine nichttriviale Lösung.

II. Da (L_0) nach Schritt (I) eine Lösung $x \neq 0$ besitzt ist nach dem superpositionsprinzip aus Satz 1.4.3 auch jeder $tx, t \in \mathbb{K}$, eine Lösung #.

1.4.9 Satz

Besitzt (L_b) genauso viele Gleichungen wie Unbekannte, also $m = n$, so gilt:

(a) Ist $L_0 = \{0\}$, so besitzt (L_b) genau eine Lösung.

(b) Besitzt (L_0) eine nichttriviale Lösung, so existieren entweder keine oder unendlich viele verschiedene Lösungen von (L_b)

Beweis:

(a) Wie gehen mittels vollständiger Induktion vor.
Für $n = 1$ gilt die Behauptung offenbar. Im Induktionsschritt gelte (a) für $n - 1$. Da (L_0) nur die triviale Lösung hat gilt $A \neq 0$. Durch Umnummerieren erreichen wir $a_{1,1} \neq 0$. Dann wird zur i-ten Gleichung, $2 \leq i$, in $(1.4a)$ das $-\frac{a_{i,1}}{a_{1,1}}$-fache der ersten Gleichung addiert:

$$(1.4f) \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ A^* = \begin{bmatrix} x_2 \\ \cdots \\ x_n \end{bmatrix} = b^* \end{cases} \qquad \text{mit } A^* \in \mathbb{K}^{(n-1)\times(m-1)}, b^* \in \mathbb{K}^{n-1}$$

Beweis:
Wir wissen:

(a) Die homogene Gleichung $A^*x^* = 0$ hat nur die triviale Lösung, denn sonst hätte (L_0) eine nicht triviale Lösung. Das Teilsystem $A^*x^* = b^*$ besitzt nach Induktionsannahme genau eine Lösung x^* mit Elementen $x_2, \cdots, x_m$. Durch Einsetzten in die erste Gleichung in $(1.4f)$ folgt ein eindeutiger Wert x_1 und die Lösung von (L_b) in eindeutiger Weise.

(b) Es sei $\hat{x}$ eine Lösung von (L_b) und x eine nichttriviale Lösung von (L_o). Dann liefern die Sätze 1.4.3 und 1.4.4, dass $\hat{x} + \alpha x$ die Gleichung löst für jedes $\alpha \in \mathbb{K}$. In diesem Fall hat (L_b) unendlich viele Lösungen. Die einzige verbleibende Möglichkeit ist, dass (L_b) keine Lösung besitzt.

2 Lineare Räume

2.1 Algebraische Strukturen

Bezeichnet $M \neq \emptyset$ eine Menge und $F(M)$ die Menge aller Selbstabbildungen auf M, so kann die Komposition $\circ$ als Abbildung $\circ : F(M) \times F(M) \to F(M)$ interpretiert werden - man spricht von einer Verknüpfung.

2.1.1 Definition (Gruppe)

Eine Gruppe $(G, \cdot)$ ist eine nichtleere Menge $\mathbb{G}$ mit einer Veknüpfung $\cdot : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$ mit den Eigenschaften:

(G_1) $\cdot$ ist Assoziativ, d.h. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für $a, b, c \in \mathbb{G}$

(G_2) es existiert ein neutrales Element $e \in \mathbb{G}$ mit $a \cdot e = a = e \cdot a$ für $a \in \mathbb{G}$

(G_3) zu jedem $a \in \mathbb{G}$ existiert ein inverses Element $a^{-1} \in \mathbb{G}$ mit $a \cdot a^{-1} = a^{-1} \cdot a = e$ für $a \in \mathbb{G}$

Bei einer kommutativen oder Abel schen Gruppe gilt ferner

(G_4) $a \cdot b = b \cdot a$ für alle $a, b \in \mathbb{G}$.

Für eine Halbgruppe müssen nur (G_1) und (G_2) gelten.

2.1.2 Bemerkung

(1) Das neutrale Element $e \in \mathbb{G}$ ist eindeutig: In der Tat, bezeichnen $e_1, e_2 \in \mathbb{G}$ zwei neutrale Elemente, so folgt nach (G_2) ist: $e_2 = e_1 \cdot e_2$ und $e_1 \cdot e_2 = e_1$, also $e_1 = e_2$

(2) Zu gegebenem $a \in \mathbb{G}$ ist auch das inverse Element $a^{-1} \in \mathbb{G}$ eindeutig. Für inverse Elemente a_1^{-1}, a_2^{-1} von a gilt nämlich

$$a_1^{-1} \overset{(G_2)}{=} a_1^{-1} \cdot e \overset{(G_3)}{=} a_1^{-1} \cdot (a \cdot a_2^{-1}) \overset{(G_1)}{=} (a_1^{-1} \cdot a) \cdot a_2^{-1} \overset{(G_3)}{=} e \cdot a_2^{-1} \overset{(G_2)}{=} a_2^{-1}$$

(3) Entsprechend $e = e^{-1}$, $a = (a^{-1})^{-1}$

2.1.3 Bemerkung (Potenzen)

Die Potenzen $a^n \in \mathbb{G}$ eines $a \in \mathbb{G}$ (G ist eine multiplikative Halbgruppe) sind rekursiv erklärt durch $a^0 := e, a^{n+1} := a \cdot a^n$ für alle $n \in \mathbb{N}_0$. In einer Gruppe setzen wir $a^n := (a^{-n})^{-1}$ für $n < 0$.

2.1.4 Beispiel

(1) $(\mathbb{Z}, +)$ ist eine kommutative additive Gruppe mit neutralen Element 0 und dem zu $a \in \mathbb{Z}$ inverses Element $-a$. Dagegen ist $(\mathbb{Z}, \cdot)$ keine Gruppe, denn das multiplikative Inverses lässt sich innerhalb von $\mathbb{Z}$ nicht erklären. Ebenso ist $(\mathbb{N}, +)$ keine (additive) Gruppe.

(2) Es sei $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Dann ist $(\mathbb{K}, +)$ eine kommutative additive Gruppe mit neutralem Element 0 und $-a$ als zu a Inversen. Auch $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine kommutative multiplikative Gruppe mit neutralem Element 1 und dem zu a inversen Element $\frac{1}{a}$.

(3) Mit $\mathbb{K} \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{C}\}$ bilden die Matrizen $(\mathbb{K}^{m \times n}, +)$ eine kommutative additive Gruppe mit neutralem Element 0 und den Inversen $-A$ zu A. Die quadratischen reellen rationalen oder komplexen Matrizen $(\mathbb{K}^{m\times n} \setminus \{0\}, \cdot)$ bilden keine Gruppe, da etwa $\operatorname{diag}(1, 0) \neq 0$ kein Inverses besitzt.

2.1.6 Beispiel (modulo)

Es sei $p \geq 2$ eine ganze Zahl und $\mathbb{Z}_p := \{0, \cdots, p-1\}$. Für beliebige $a, b \in \mathbb{Z}$ gibt es vermöge der Division mit Rest eindeutige $m \in \mathbb{Z}$ und $k \in \mathbb{Z}_p$ mit $a + b = mp + k$ wir schreiben dann $k = a + b \mod p$ oder $k =: a +_p b$. Dann ist $(\mathbb{Z}_p, +_p)$ eine kommutative Gruppe mit dem neutralem Element 0.

2.1.7 Beispiel (symmetrische Gruppe)

Es sei M eine nichtleere Menge und $S(M)$ bezeichnet alle bijektiven Selbstabbildungen $f : M \to M$. Dann ist die <u>symmetrischen Gruppe</u> $(S(M), \circ)$ eine i.A. nicht-kommutative Gruppe mit id_m als neutralem Element und $f^{-1} : M \to M$ als inversen Element zu f. Im Fall $M = \{1, \cdots, n\}$ schreiben wir $S_n := S(\{1, \cdots, n\})$. Die Menge aller nicht-notwendig bijektiven Selbstabbildungen $F(M)$ ist dagegen eine Halbgruppe bezüglich $\circ$.

2.1.8 Korollar (Rechnen in Gruppen)

Für alle $a, b, c \in \mathbb{G}$ gilt $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$, wie auch $a \cdot b = a \cdot c \Rightarrow b = c, a \cdot b = e \Rightarrow a = b^{-1}$.
Beweis:
Es seien $a, b, c \in \mathbb{G}$. Wir zeigen zunächst, dass $b^{-1} \cdot a^{-1}$ das inverse Element von $a \cdot b$ ist. Dazu

$$(b^{-1} \cdot a^{-1}) \cdot (a \cdot b) \overset{(G_1)}{=} b^{-1} \cdot (a^{-1} \cdot (a \cdot b \cdot)) \overset{(G_1)}{=} b^{-1} \cdot ((a^{-1} \cdot a) \cdot b) \overset{(G_3)}{=} b^{-1} \cdot (e \cdot b) \overset{(G_2)}{=} b^{-1} \cdot b \overset{(G_3)}{=} e$$

und entsprechend $(a \cdot b) \cdot (b^{-1} \cdot a^{-1}) = e$. Die erste Implikation ergibt sich nach Voraussetzung durch

$$b \overset{(G_2)}{=} e \cdot b \overset{(G_3)}{=} (a^{-1} \cdot a) \cdot b \overset{(G_1)}{=} a^{-1} + (a \cdot b) = a^{-1} \cdot (a \cdot c) \overset{(G_1)}{=} (a^{-1} \cdot a) \cdot c \overset{(G_3)}{=} e \cdot c \overset{(G_2)}{=} c$$

Die verbleibende Implikation sei den Leser überlassen.

2.1.9 Definition (Körper)

Ein Körper $(\mathbb{K}, +, \cdot)$ ist eine Menge $\mathbb{K}$ mit mindestens zwei Elementen versehen, mit den arithmetischen Op
$+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ (Addition) und $\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$(Multiplikation).
$(\mathbb{K}_1)(\mathbb{K}, +)$ ist eine kommutative Gruppe mit neutralem Element 0 und dem zu $\alpha \in \mathbb{K}$ inversen Element $-\alpha$, d.h. für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt:

$$(\mathbb{K}_1^1)\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$$
$$(\mathbb{K}_1^2)\alpha + 0 = 0 + \alpha = \alpha$$
$$(\mathbb{K}_1^3)\alpha \cdot -\alpha = -\alpha \cdot \alpha = 0$$
$$(\mathbb{K}_1^4)\alpha + \beta = \beta + \alpha$$

$(\mathbb{K}_2)$ $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe mit neutralem Element 1 und zu $\alpha \in \mathbb{K}$ Inversem $\frac{1}{\alpha}$, d.h. es gilt für $\alpha, \beta, \gamma \in \mathbb{K} \setminus \{0\}$.

$$(\mathbb{K}_2^1)\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$$
$$(\mathbb{K}_2^2)\alpha \cdot 1 = 1 \cdot \alpha = \alpha$$
$$(\mathbb{K}_2^3)\alpha \cdot \frac{1}{\alpha} = \frac{1}{\alpha} \cdot \alpha = 1$$
$$(\mathbb{K}_2^4)\alpha \cdot \beta = \beta \cdot \alpha$$

$(\mathbb{K}_3)$ es gelten die Distributivgesetze $\alpha(\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$, $(\alpha + \beta) \cdot \gamma = \alpha\gamma + \beta\gamma$ für alle $\alpha, \beta, \gamma \in \mathbb{K}$. Üblich $\alpha\beta := \alpha \cdot \gamma$. Subtraktion als $\alpha - \beta := \alpha + (-\beta)$. Division $\frac{\alpha}{\beta} := \alpha \cdot \frac{1}{\beta}$.

2.1.10 Beispiel

$\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sind Körper bzgl. $+, \cdot$

2.1.11 Beispiel (Restklassenkörper modulo p)

Mit einer gegebenen Primzahl $p \in \mathbb{N}$ definieren wir die Mengen $\mathbb{Z}_p := \{0, \cdots, p\}$. Dann gibt es für beliebige $\alpha, \beta \in \mathbb{Z}_p$ eindeutige Zahlen $m, n \in \mathbb{Z}$ und $k, l \in \mathbb{Z}_p$ derart, dass

$$\alpha + \beta = m \cdot p + k$$

$$\alpha \cdot \beta = np + l \text{ Divison mit Rest.}$$

$$\text{Addition: } \alpha +_p \beta := k$$

$$\text{Multiplikation: } \alpha \cdot_p \beta := l \text{ (2.1a)}$$

$(\mathbb{Z}_p, +_p, \cdot_p)$ ist Körper, der sogenannten Restklassenkörper modulo p.

$$\mathbb{Z}_2 : \begin{array}{c|cc} +_2 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

$$\begin{array}{c|cc} \cdot_2 & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

$$\mathbb{Z}_3 : \begin{array}{c|ccc} +_3 & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array}$$

$$\begin{array}{c|ccc} \cdot_3 & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{array}$$

2.1.12 Korollar

Ist $(\mathbb{K}, +, \cdot)$ ein Körper, so gilt für alle $\alpha, \beta, \gamma \in \mathbb{K}$, dass

$$\begin{aligned} 0 \cdot \alpha = \alpha \cdot 0 = 0, \qquad & \beta \cdot (-\alpha) = -(\beta \cdot \alpha) = (-\beta) \cdot \alpha (2.1b) \\ (-1) \cdot \alpha = -\alpha, \qquad & (-\alpha) \cdot (-\beta) = \alpha \cdot \beta (2.1c) \end{aligned}$$

Und ferner die Implikation $\alpha \cdot \beta = 0 \to \alpha = 0$ oder $\beta = 0$.

2.1.13 Bemerkung

Es gilt $1 \neq 0$, da die Annahme $1 = 0$ folgenden Widerspruch impliziert: Da $\mathbb{K}$ mindestens 2 Elemente enthält, gibt es ein $\alpha \in \mathbb{K}, \alpha \neq 0$ mit:

$$\alpha \overset{(\mathbb{K}_2^2)}{=} \alpha \cdot 1 = \alpha \cdot 0 \overset{(2.1b)}{=} 0$$

Daher ist der Restklassenkörper modulo 2 $\mathbb{Z}_2$ der kleinste Körper.

2.1.14 Beweis

Wähle ein $\alpha, \beta, \gamma \in \mathbb{K}$. Es gilt $0 \cdot \alpha \overset{(\mathbb{K}_1^2)}{=} (0+0) \cdot \alpha \overset{(\mathbb{K})}{=} 0\alpha + 0\alpha$ mittels Korollar 2.1.8 ($+, a = b = 0$ und $c = 0$) folgt $0 \cdot \alpha = 0$, kommutativ liefert $\alpha 0 = 0$. Aus dieser Behauptung resultiert

$$(-\beta)\alpha + \beta\alpha \overset{(\mathbb{K}_3)}{=} (-\beta + \beta)\alpha = 0 \cdot \alpha = 0$$

mit Korollar 2.1.8 ($+, a = (-\beta)\alpha, b = \beta\alpha$). Dies liefert $-(\beta\alpha) = (-\beta)\alpha$ und $\beta(-\alpha) = -(\beta\alpha)$. Die Beziehung $(-1)\alpha = -\alpha$ resultiert aus dem eben gezeigten $\beta = 1$ und

$$(-1)\alpha = 1 \cdot (-\alpha) \overset{\mathbb{K}_2^2)}{=} -\alpha.$$

2.1c ergibt sich mit Bemerkung 2.1.2(3) aus

$$(-\alpha)(-\beta) \overset{2.1b}{=} -(\alpha(-\beta)) \overset{2.1b}{=} -(-(\alpha\beta)) = \alpha\beta = 0$$

Annahme: $\alpha \neq 0$ und $\beta \neq 0$ dann $1 \overset{\mathbb{K}_2^3)}{=} \frac{1}{\beta} \cdot \frac{1}{\alpha} \cdot \alpha \cdot \beta \overset{2.1b}{=} 0$

2.2 Vektorräume

2.2.1 Definition (linearer Raum, Vektorraum)

Es sei $\mathbb{K}$ ein Körper. Ein Vektorraum oder linearer Raum $(X, +, \cdot)$ (über $\mathbb{K}$) ist eine nichtleere Menge X mit arithmetische Operationen:

(1) Addition $+ : \ X \times X \to X$ derart, dass $(X, +)$ eine kommutative Gruppe mit neutralem Element 0 oder Nullvektor.

(2) Skalare Multiplikation $\cdot : \mathbb{K} \times X \to X$ derart, dass für alle $\alpha, \beta \in \mathbb{K}$ und $x, y \in X$ gilt:

(V_1) $\alpha(x + y) = \alpha x + \alpha y$ Distributiv Gesetz
(V_2) $(\alpha + \beta) \cdot x = \alpha x + \beta x$ Distributiv Gesetz
(V_3) $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$ Assoziativ Gesetz
(V_4) $1 \cdot x = x$

Die Elemente aus $\mathbb{K}$ heißen Skalare und X heißen Vektoren.

Konventionen: $\alpha x := \alpha \cdot x \quad x - y := x + (-y)$

2.2.2 Beispiel

Es sei $(\mathbb{K}, +, \cdot)$ ein Körper.

(0) Der triviale Raum $\{0\}$ der nur die 0 enthält.

(1) Weiter ist $\mathbb{K}$ ein Vektorraum über sich selbst.

(2) Die Menge aller $m \times n$-Matrizen $\mathbb{K}^{m\times n}$ ist ein linearer Raum über $\mathbb{K}$ bezüglich

(1.3b) $\alpha A := \alpha A = (\alpha a_{i,j})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$

(1.3c) $A + B := (\alpha_{i,j} + \beta_{i,j})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$

Ein n-Tupel $(x_1, \cdots, x_n) \in \mathbb{K}^{1\times n}$ bezeichnen wir als Zeilenvektor und eine m-Spalte (1.3a) als Spaltenvektor.

2.2.3 Beispiel

Es sei $p \in \mathbb{N}$ eine Primzahl und $n \in \mathbb{N}$. Dann sind die n-Spalten $\mathbb{Z}_p^n$ in $\mathbb{Z}_p$ mit den komponentenweisen Addition $+_p$ und skalaren Multiplikation $\cdot_p$ ein linearer Raum über $\mathbb{Z}_p$.
Insbesondere für $\mathbb{Z}_2^2$

$+_2$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$
$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$
$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$
$\begin{pmatrix}0\\1\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$
$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$

$\cdot_2$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$
0	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$
1	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$

2.2.4 Beispiel (Lösungsmengen)

Mit Satz 1.4.3 ist L_0 einer homogenen Gleichung ein Vektorraum über $\mathbb{K}$. Die Lösungsmenge L_b inhomogener Systeme ist kein linearer Raum über $\mathbb{K}$.

2.2.5 Beispiel (Funktionsräume)

Es sei $\omega \neq \emptyset$ und X ein linearer Raum über $\mathbb{K}$. Dann ist $F(\omega, X) := \{u : \omega \to X\}$ ein Vektorraum über $\mathbb{K}$ mit punktweise definierten arithmetischen Operationen $(a+v)(t) := u(t)+v(t)$, $(\alpha u)(t) := \alpha u(t)$ für alle $t \in \omega, \alpha \in \mathbb{K}$.
Die Menge $F(\omega, X)$ wird als Funktionenraum bezeichnet. $\omega \in \mathbb{N}$, $\omega \in \mathbb{Z}$, dann bezeichnen wir $F(\omega, X)$ als Folgenraum.

2.2.6 Korollar

Ist $(X, +, \cdot)$ ein linearer Raum über $\mathbb{K}$ so gilt für alle Skalare $\alpha, \beta \in \mathbb{K}$ und Vektoren $x, y \in X$:

$$\begin{aligned} &(a) && 0_{\mathbb{K}} \cdot x = \alpha \cdot 0_x = 0_x \\ &(b) && \text{Falls } \alpha x = 0_x, \text{ so folgt } \alpha = 0 \in \mathbb{K} \text{ oder } x \in 0 \in X \\ &(c) && (-\alpha)x = \alpha(-\alpha) = -(\alpha x) \\ &(d) && \alpha(x - y) = \alpha x - \alpha y \text{ und } (\alpha - \beta)x = \alpha x - \beta x \end{aligned}$$

Beweis: Es sei $\alpha \in \mathbb{K}$ und $x \in X$:

(a) Es gilt $0_{\mathbb{K}}x = (0_{\mathbb{K}} + 0_{\mathbb{K}})x = 0_{\mathbb{K}}x + 0_{\mathbb{K}}x$ wegen V_2. Nach Definition 2.2.1 (a) existiert zum Vektor $z := 0_{\mathbb{K}}x$ ein Vektor $-z$ mit $0 \cdot x + (-z) = 0_X$ und wir erhalten $0_X = 0 \cdot x + (-z) = (0 \cdot x + 0 \cdot x) + (-z) = 0 \cdot x + (0 \cdot x + (-z)) = 0 \cdot x + 0_x = 0 + x$ und die Beziehung $\alpha \cdot 0 = 0$ folge analog.

(b) (b) Es gelte $\alpha x = 0$ mit $\alpha \neq 0$ und wir zeigen $x = 0_x$ $\alpha \neq 0$ existiert $\frac{1}{\alpha}$. Nach (a) folgt $\frac{1}{\alpha}(\alpha \cdot x) = \frac{1}{\alpha} \cdot 0 = 0$ und andererseits $\frac{1}{\alpha}(\alpha x) = (\frac{1}{\alpha} \cdot \alpha) \cdot x = 1 \cdot x = x$

(c) , (d)

2.2.7 Definition (Unterraum)

Eine nicht leere Teilmenge $Y \subseteq X$ eines linearen Raumes $(X, +, \cdot)$ über $\mathbb{K}$ heißt Unterraum von X, falls gilt $\alpha_1 y_1 + \alpha_2 y_2 \in Y$ für alle $\alpha_1, \alpha_2 \in \mathbb{K}$ und $y_1, y_2 \in Y$

2.2.8 Bemerkung

Jeder lineare Raum x hat die trivialen Unterräume $\{0\}$ und X.

2.2.9 Beispiel (Stetige und stetig-differenzierbare Funktion)

Es sei $I \subseteq \mathbb{R}$ ein Intervall. Die Menge der stetigen Funktionen $C(I, \mathbb{R}^n)$ auf I mit Bildern in $\mathbb{R}^n$ ist ein Unterraum von $F(I, \mathbb{R})$. Ebenso sind stetig differenzierbare Funktionen $C^1(I, \mathbb{R})$ ein Unterraum von $C(I, \mathbb{R})$ und $F(I, \mathbb{R}^n)$

2.2.10 Beispiel (Polynome)

Mit gegebenem Körper $\mathbb{K}$ definieren wir den Raum der Polynome (über $\mathbb{K}$) durch
$P(\mathbb{K}) := \{p \in F(\mathbb{K}, \mathbb{K}) \exists n \in \mathbb{N}_0 : \exists a_0, ..., a_n \in \mathbb{K} : p(t) = \sum_{l=0}^{n} a_l \cdot t^l\}$;
seine Elemente heißen Polynome und die a_k deren Koeffizienten. Dann ist $P(\mathbb{K})$ ein Unterraum von $F(\mathbb{K}, \mathbb{K})$.
Der Grad $deg\ p$ eines Polynoms $p \in P(\mathbb{K})$ ist der maximale Index $k \in \mathbb{N}_0$ für den $a_k = 0$ ist. Für $m \in \mathbb{N}_0$ sind die Mengen $P_m(\mathbb{K}) := \{p \in P(\mathbb{K}) : deg\ p \leq m\}$

Unterräume von $P(\mathbb{K})$, wogegen $\{p \in P(\mathbb{K}) : deg\ p = m\}$ für $m \neq 0$ kein Unterraum ist. Ferner ist jedes $P_n(\mathbb{K})$ Unterraum von $P_m(\mathbb{K})$ für $0 \leq n \leq m$.

2.2.11 Satz (Schnitte und Summen von Unterräumen)

Ist I eine nichtleere Indexmenge und $(Y_i)_{i \in I}$ eine Familie von Unterräumen von X.

(a) Der Durchschnitt $\bigcap\limits_{i \in I} Y_i$ ist ein Unterraum von X.

(b) Für endliche I ist die Summe $\sum\limits_{i \in I} Y_i := \{\sum\limits_{i \in I} y_i \in X : y_i \in Y_i \text{ mit } i \in I\}$ der kleinste Unterraum von X, der jedes y_i enthält.

Für $I = \{1, ..., n\}$ schreibt man auch $Y_1 + ... + Y_m = \sum\limits_{i \in J} Y_i$.

Beweis:

(a) Es seien $\alpha, p \in \mathbb{R}$ und $x, y \in \cap_{i \in I} Y_i$. Dann gilt $x, y \in Y_i$ für alle $i \in I$ und da jedes Y_i ein Unterraum von X ist, folgt $\alpha \cdot x + \beta \cdot y \in Y_i$ für jedes $i \in I$. Dies impliziert, dass $\alpha \cdot x + \beta \cdot y \in \cap_{i \in I} Y_i$

(b) Wir zeigen $Y := \sum\limits_{i \in I} Y_i$ ist ein Unterraum von X. Dazu sei $x = \sum\limits_{i \in I} x_i$ und $y = \sum\limits_{i \in I} y_i$ mit $x_i, y_i \in Y_i$ und wir erhalten für alle $\alpha, \beta \in \mathbb{R}$:
$\alpha \cdot x + \beta \cdot y = \alpha \sum\limits_{i \in I} x_i + \beta \sum\limits_{i \in I} y_i = \sum\limits_{i \in I} (\underbrace{\alpha x_i + \beta y_i}_{\in Y_i})$

Zu zeigen y ist kleinster Unterraum der alle Y_i enthält.
Dazu sei $z \subseteq X$ ein weiterer Unterraum von X der alle Y_i enthält. Für $x_i \in Y_i$ ist dann auch

$x_i \in Z$ für alle $i \in I$, da Y_i in Z enthalten sind.
Aus der Unterraumeigenschaft von Z resultiert $\sum\limits_{i \in I} x_i \in Z$ und folglich ist $Y \subseteq Z$

2.3 Lineare Abhängigkeiten

Gegeben sei eine nichtleere Menge S von Vektoren aus einem linearen Raum X über dem Körper $\mathbb{K}$. Existieren zu einem gegebenem $x \in X$ dann endlich viele Koeffizienten $a_i \in \mathbb{R}$ und $x_i \in \mathcal{S}$, $1 \leq i \leq n$, mit $x = \sum\limits_{i=1}^{n} a_i \cdot x_i$ so bezeichnen wir x als Linearkombination der Vektoren aus S.

2.3.1 Definition (Spann)

Es sei $\mathcal{S} \leq X$. Der Spann oder die lineare Hülle $\operatorname{span} \mathcal{S}$ von $\mathcal{S}$ ist die Menge aller Linearkombinationen. Ferner setzt man $\operatorname{span}\{0\} = \{0\}$.

2.3.2 Beispiel

Für endliche $\mathcal{S} = \{x_0, ..., x_n\}$ ist der $\operatorname{span} \mathcal{S} = \{\sum\limits_{i=1}^{n} \alpha_i \cdot x_i \in X : \alpha_i \in \mathbb{K}\}$
$\mathbb{K} = \mathbb{R}$: $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ gilt $\operatorname{span}\{e_i, e_2\} = \mathbb{R}^2$
$\operatorname{span}\{x_1, x_2\}$ wenn $x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $x_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ aber
$y_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $y_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ dann $\operatorname{span}\{y_1, y_2\} = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$

2.3.3 Beispiel (Monome)

Polynome $m_n(l) := t^n$, $n \in \mathbb{N}_0$ heißen Monome. Dann lassen sich die Polynome als lineare Hülle der Monome darstellen, d.h. $\operatorname{span}\{m_n\}_{n \in \mathbb{N}_0} = P(\mathbb{K})$ insbesondere ist
$\operatorname{span}\{m_0, ..., m_n\} = P_n(\mathbb{K}^n)$
$\operatorname{span}\{m_{2n}\}_{n \in \mathbb{N}_0} = \{p \in P(\mathbb{K}) : p(t) = p(-t) \text{ auf } \mathbb{K}\}$
$\operatorname{span}\{m_{2n-1}\}_{n \in \mathbb{N}_0} = \{p \in P(\mathbb{K}) : p(t) = -p(-t) \text{ auf } \mathbb{K}\}$

2.3.4 Proposition

Es sei $\mathcal{S} \in X$ nicht leer. Dann ist die lineare Hülle der kleinste $\mathcal{S}$ umfassende Unterraum von X

Beweis: $x, y \in \mathcal{S}$ ist $\alpha x + \beta y$, $\alpha, \beta \in \mathbb{K}$ in $\operatorname{span} \mathcal{S}$. Also ist $\operatorname{span} \mathcal{S}$ Unterraum von X. $\operatorname{span} \mathcal{S}$ enthält die Vektoren aus $\mathcal{S}$ und damit ist $\mathcal{S} \subseteq \operatorname{span} \mathcal{S}$, $Y \subseteq X$ ein Unterraum von X mit $x \in Y$ für sämtliche $x \in \mathcal{S}$. Dann liegen sämtliche Linearkombinationen von Vektoren aus $\mathcal{S}$ in Y. Also ist $\operatorname{span} \mathcal{S}$ in Y enthalten.

2.3.5 Korollar

Ist x eine Linearkombination von Vektoren aus $\mathcal{S} \subseteq X$, so gilt $\operatorname{span} \mathcal{S} = \operatorname{span}(\mathcal{S} \cup \{x\})$.
Beweis: Wir zeigen die Behauptung durch zwei Inklusionen:
($\subseteq$) Es ist klar dass $\operatorname{span} \mathcal{S} \subseteq \operatorname{span}(\mathcal{S} \cup \{x\})$
($\supseteq$) Also Linearkombination von Vektoren aus $\mathcal{S}$ liegt x auch in $\operatorname{span} \mathcal{S}$.
Demnach ist $\operatorname{span} \mathcal{S}$ derjenige Unterraum welcher $\mathcal{S}$ und $\{x\}$ enthält.
Damit folgt aus Prop 2.3.4, dass $\operatorname{span}(\mathcal{S} \cup \{x\}) \operatorname{span} \mathcal{S}$.

2.3.6 Definition (lineare Unabhängigkeit)

Eine endliche Menge $\{x_1, \cdots, x_n\}$ von Vektoren aus X heißt linear unabhängig falls gilt:

$$\sum_{k=1}^{n} \xi_k x_k = 0 \Rightarrow \xi_k = 0 \forall n = 1, n$$

Griechische Buchstaben:

$$\eta \ - \ \text{eta}$$
$$\xi \ - \ \text{xi}$$
$$\zeta \ - \ \text{zeta}$$

Für beliebige Mengen $\mathcal{S} \subseteq X$ nennt man $\mathcal{S}$ linear unabhängig, wenn jede endliche Teilmenge von $\mathcal{S}$ linear unabhängig ist, die leere Menge $\emptyset$ wird als lineare unabhängig betrachtet. Eine Teilmenge von X heißt linear abhängig, falls sie nicht linear unabhängig ist.
Man nennt Vektoren $x_1, x_2, \cdots$ linear unabhängig, wenn $\{x_1, x_2, \cdots\}$ diese Eigenschaft hat.

2.3.7 Bemerkung

(1) lineare Abhängigkeit einer endlichen Menge $\{x_1, \cdots x_n\}$ bedeutet, dass eine nichttriviale Darstellung der Null aus Vektoren x_u existiert:
Man kann also

$$(2.3a) \sum_{k=1}^{n} \xi_k x_k = 0$$

schreiben, ohne dass alle ξ_k verschwinden.

(2) Jede Obermenge einer linear abhängigen Menge ist linear abhängig. Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.

2.3.8 Beispiel

Die Menge $\{0\}$ ist linear abhängig, dagegen ist $\{x\}, x \neq 0$, linear unabhängig.

2.3.9 Proposition

Es sei $\mathcal{S} \subseteq X$ nichtleer und $x, x_1, \cdots, x_n \in X$

(a) Ist $\mathcal{S} = \{x_1, \cdots, x_n\}$ linear abhängig, so lässt sich mindestens ein Vektor aus $\mathcal{S}$ als Linearkombination der weiteren Elementen von $\mathcal{S}$ darstellen.

(b) Für jede Linearkombination x aus $\mathcal{S}$ ist $\mathcal{S} \cup \{x\}$ linear abhängig.

Beweis:

(a) Weil $\{x_1, \cdots, x_n\}$ linear abhängig ist, besitzt 0 die Darstellung $(2.3a)$ in welcher nicht alle ξ_k verschwinden. Also existiert ein Index $1 \leq k^* \leq n$ mit $\xi_{k^*} \neq 0$ und damit

$$X_{k^*} = -\xi_k^{-1} \sum_{\substack{k=1 \\ k \neq k^*}}^{n} \xi_k x_k = \sum_{\substack{k=1 \\ k \neq k^*}}^{n} (-\xi_k^{-1} \xi_k) x_k$$

(b) Mit $x = \sum_{k=1}^n \xi_k x_k$ ist $x - \sum_{k=1}^n \xi_k x_k$ eine nichttriviale Darstellung der 0

In $X = \mathbb{K}^m$ gilt: Es sei $\mathcal{S} = \{a_1, \cdots, a_n\} \subseteq \mathbb{K}^m$. Mit der $m \times n$-Matrize $A := (a_1, \cdots, a_n)$ ist die Beziehung $\sum_{k=1}^n \xi_k a_k = 0$ (vgl. $(2.3a)$) äquivalent zu:

$$(2.3b) Ax = 0, x \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

Demzufolge ist $\mathcal{S}$ genau dann linear unabhängig, wenn $Ax = 0$ nur die triviale Lösung hat. Aus Satz 1.4.8 (in Verbindung mit Blatt 5, Aufg. 1) erhalten wir daher, dass mehr als m Vektoren stets linear abhängig sind.

2.3.10 Beispiel

(1) Für die kanonischen Einheitsvektoren in $\mathbb{K}^m$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots e_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

gilt in obiger Terminologie $A = I_m$. Also besitzt $Ax = 0$ nur die triviale Lösung und $\{e_1, \cdots, e_m\}$ ist linear unabhängig.

(2) Es sei $\lambda \in \mathbb{R}$ um die lineare Unabhängigkeit von

$$x_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, x_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, x_3 = \begin{pmatrix} 7 \\ 8 \\ \lambda \end{pmatrix}$$

in $\mathbb{R}^3$ zu untersuchen, betrachten wir die Gleichung (2.3b) mit

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & \lambda \end{pmatrix}$$

und lösen sie mit dem in Beispiel 1.4.6 beschriebenen Schema:

$$\begin{array}{ccc|cl} 1 & 4 & 7 & 0 & \\ 2 & 5 & 6 & 0 & 2I - II \\ 3 & 6 & \lambda & 0 & 3I - II \end{array} \Rightarrow \begin{array}{ccc|cl} 1 & 4 & 7 & 0 & \\ 0 & 3 & 6 & 0 & :3 \\ 0 & 6 & 21-\lambda & 0 & III - 2II \end{array} \Leftrightarrow \begin{array}{ccc|cl} 1 & 4 & 7 & 0 & \\ 0 & 1 & 2 & 0 & :3 \\ 0 & 0 & 9-\lambda & 0 & III - 2II \end{array}$$

Also hat $Ax = 0$ für $\lambda \neq 9$ nur die triviale Lösung (lineare Unabhängigkeit von $\{x_1, x_2, x_3\}$ und für $\lambda = 9$ nichttriviale Lösungen (lineare Abhängigkeit).

2.3.11 Satz

Eine Menge $\mathcal{S} \subseteq X$ ist genau dann linear unabhängig, wenn jedes $x \in \mathcal{S}$ auf nur eine Art (bis auf Glieder mit Null-Koeffizienten) als Linearkombinationen von Vektoren aus $\mathcal{S}$ dargestellt werden kann.

2.4 Basis und Dimensionen

Es sei X ein linearer Raum über dem Körper $\mathbb{K}$.

2.4.1 Definition (Basis)

Eine Menge $\mathcal{X} \subseteq X$ heißt Basis von X, falls $\mathcal{X}$ linear unabhängig mit $X = \operatorname{span} \mathcal{X}$ ist:
Eine Menge $\mathcal{X}$ mit $X = \operatorname{span} \mathcal{X}$ wird Erzeugendessystem (EZS) von X genannt. Man nennt X endlich erzeugt, falls er ein endliches EZS hat.

2.4.2 Beispiel

Die Basis von $\{0\}$ ist die leere Menge.

2.4.3 Beispiel (Standardbasis)

Die mittels der kanonischen Einheitsvektoren aus Beispiel 2.3.10 (1) gebildete Menge $\mathcal{E}_m := \{e_1, \cdots, e_m\}$ ist eine Basis von $\mathbb{K}^m$, die sogenannte Standardbasis, damit ist $\mathbb{K}^m$ endlich erzeugt.

2.4.4 Beispiel (Polynome)

Für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ sind $\mathcal{M}_n := \{m_0, \cdots m_n\}$ aus Beispiel 2.3.3 eine Basis der Polynome $\mathcal{P}_n(\mathbb{K})$ von maximalem Grad n. Ebenso ist $\{m_n\}_{n\in\mathbb{N}_0}$ eine Basis von $\mathcal{P}(\mathbb{K})$. Somit ist jedes $\mathcal{P}_n(\mathbb{K})$ endlich erzeugt, $\mathcal{P}(\mathbb{K})$ dagegen nicht.

2.4.5 Lemma

Es sei $\mathcal{S} \subseteq X$ linear unabhängig. Gilt dann $x \notin \operatorname{span}\mathcal{S}$, so ist auch $\mathcal{S} \cup \{x\}$ linear unabhängig.
Beweis: Es ist nachzuweisen, dass jede endliche Teilmenge von $\mathcal{S} \cup \{x\}$ linear unabhängig ist. Dazu sei $\{x_1, \cdots x_n, x\}$ eine solche Menge und $\sum_{k=1}^n \xi_k x_k + \eta x = 0$ eine Darstellung der Null. Wäre $\eta \neq 0$, so könnte man x als Linearkombination der $x_1, \cdots, x_n$ darstellen, dies widerspricht $x \notin \operatorname{span}\mathcal{S}$. Also gilt $\eta = 0$. Da aber $\{x_1, \cdots, x_n\}$ linear unabhängig ist, folgt $\xi_1 = \cdots = \xi_n = 0$.
In trivialer Weise: ist X ein EZS von X. Unser Interesse besteht aber gerade in „kleinen" EZSen.

2.4.6 Satz

Mit nicht leerem $\mathcal{X} \subseteq X$ sind äquivalent:

(a) $\mathcal{X}$ ist eine Basis von X

(b) Jeder Vektor $x \in X$ lässt sich eindeutig als Linearkombination von Vektoren aus $\mathcal{X}$ darstellen.

(c) $\mathcal{X}$ ist maximal linear unabhängig, d.h. $\mathcal{X}$ ist linear unabhängig und für jedes $x \in X \setminus \mathcal{X}$ ist $\mathcal{X} \cup \{x\}$ linear abhängig.

(d) $\mathcal{X}$ ist ein minimales EZS, d.h. keine echte Teilmenge von $\mathcal{X}$ ist ein EZS.

2.4.7 Bemerkung (Koordinaten)

Besitzt $x \in X$ bzgl. der Basis $\mathcal{X} := \{x_1, \cdots, x_n\}$ die nach Satz 2.4.6 (b) eindeutige Darstellung $x = \sum_{k=1}^n \xi_k x_k$ mit Koeffizienten $\xi_k \in \mathbb{K}$ so bezeichnet man das n-tupel $(\xi_1, \cdots, \xi_n)$ also Koordinaten von x (bzgl. $\mathcal{X}$). Von nun an sei X endlich erzeugt.

2.4.8 Satz

Jedes endliche EZS eines Vektorraumes enthält eine Basis. Insbesondere hat jeder endlich erzeugte lineare Raum eine Basis.
Beweis: Es sei $\mathcal{X}$ ein endliches EZS. Ist $\mathcal{X}$ keine Basis, so kann $\mathcal{X}$ nicht minimal sein und es existiert eine echte Teilmenge $\mathcal{X}^1 \overset{\subset}{\neq} \mathcal{X}$, die ebenfalls ein EZS ist. Ist wiederum $\mathcal{X}^1$ keine Basis, so existiert erneut eine echte Teilmenge $\mathcal{X}^2 \overset{\subset}{\neq} \mathcal{X}^1$, die X erzeugt. Durch Iteration erhält man eine echt absteigende Folge von Teilmengen $\cdots \subsetneqq \mathcal{X}^2 \subsetneqq \mathcal{X}^1 \subsetneqq \mathcal{X}$. Diese Folge bricht nach endlich vielen

schritten ab, da $\mathcal{X}$ endlich ist, d.h. es gibt ein minimales $\mathcal{X}^k$. Dieses $\mathcal{X}^k$ ist nach Satz Satz 2.4.6 eine Basis von X.

2.4.9 Proposition

Ist X endlich erzeugbar und $\mathcal{S} \subseteq X$ linear unabhängig, so existiert eine Basis von X, welche $\mathcal{S}$ als Teilmenge enthält.

2.4.10 Lemma (Austauschsatz von Steinitz)

Ist $\{x_1, \cdots x_p\}$ linear unabhängig und $\{y_1, \cdots, y_n\}$ ein EZS von X, so gilt $p \leq n$ und nach einer Umnummerierung der y_k ist $\{x_1, \cdots, x_p, y_{p+1}, \cdots, y_n\}$ ein EZS von X.

2.4.11 Satz (Dimension)

Falls X eine Basis von n Elementen besitzt, enthält jede Basis von X genau n Elemente. Wir bezeichnen n als Dimension von X und schreiben $n = \dim X$.

2.4.12 Bemerkung

Ein linearer Raum X heißt unendlich-dimensional (symbolisch $\dim X = \infty$) falls er kein endlichen EZS besitzt, anderenfalls heißt er endlich-dimensional.
Beweis: Es seien $\{x_1, \cdots, x_n\}$ und auch $\{y_1, \cdots, y_m\}$ Basen von X. Mit Lemma 2.4.10 folgt dann $n \leq m$, wie auch $m \leq n$, und somit $m = n$.

2.4.13 Beispiel

Für die bislang betrachteten Räume ist $\dim \mathbb{K}^n = n$, $\dim \mathbb{K}^{m \times n} = m \cdot n$, $\dim \mathcal{P}_n(\mathbb{R}) = n + 1$ und $\dim \mathcal{P}(\mathbb{R}) = \dim C^1(\mathbb{R}, \mathbb{R}) = \dim C(\mathbb{R}, \mathbb{R}) = \dim F(\mathbb{R}, \mathbb{R}) = \infty$.

2.4.14 Beispiel

Die komplexen Zahlen $\mathbb{C}$ sind ein 2-dimensionaler Vektorraum über $\mathbb{R}$ und ein 1-dimensionaler Raum über $\mathbb{C}$.

2.4.15 Korollar

In linearen Räumen X mit $n := \dim X$ gilt:

(a) Weniger als n Vektoren aus X sind kein EZS.

(b) Mehr als n Vektoren aus X sind linear abhängig.

(c) Jedes EZS mit n Elementen ist eine Basis.

(d) Jede linear unabhängige Menge mit n Elementen ist eine Basis.

<u>Beweis</u>:

(a) jedes EZS enthält laut Satz 2.4.8 eine Basis. Für jedes aus weniger als n Vektoren bestehenden EZS gäbe es dann auch eine Basis mit Weniger als n Elementen. Dies widerspricht Satz 2.4.11.

(b) Laut Proposition 2.4.9 ist jede linear unabhängige Menge Teil einer Basis. Somit hätte man mit einer linear unabhängigen Familie von mehr als n Vektoren auch eine Basis mit mehr als n Elementen - im Widerspruch zu Satz 2.4.11.

(c) Ein EZS enthält wegen Satz 2.4.8 eine Basis und ist wegen Satz 2.4.11 bereits eine solche.

(d) Mit Proposition 2.4.9 ist eine linear unabhängige Familie Teilmenge einer Basis und mit Satz 2.4.11 eine Basis.

2.4.16 Korollar

Für jeden Unterraum Y eines endlich-dimensionalen Raumes X ist $\dim Y \leq \dim X$, Gleichheit gilt genau für $X = Y$.

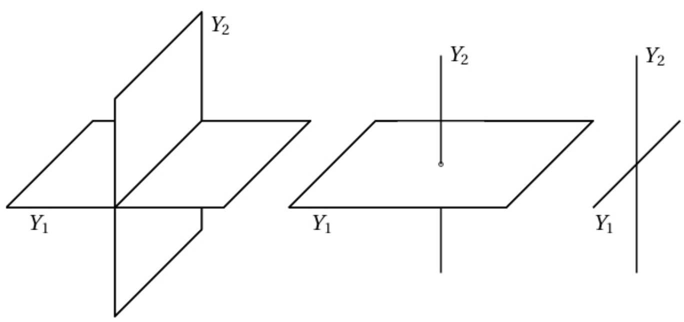
2.5 Komplemente und direkte Summen

Wieder sei X ein linearer Raum über $\mathbb{K}$.

2.5.1 Definition (direkte Summen)

Es seien $Y_1, Y_2 \subseteq X$ Unterräume. Dann heißt Y_2 Komplement von Y_1 in X, falls gilt $Y_1 + Y_2 = X$ $Y_1 \cap Y_2 = \{0\}$, man schreibt $X = y_1 \oplus y_2$ und nennt X direkte Summe von Y_1, Y_2.
<u>Beispiele</u>:

$Y_1 \cap Y_2 \neq 0$ $\qquad$ $Y_1 \oplus Y_2 = \mathbb{R}^3$ $\qquad$ $Y_1 \oplus Y_2 \subsetneq \mathbb{R}^3$



2.5.2 Beispiel

Im Raum $X = \mathbb{R}^3$ ist die Gerade $Y_2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in X : x_1 = x_2 = x_3 \right\}$ ein Komplement zur Ebene $Y_1 := \{x \in X : x_1 - x_2 + x_3 = 0\}$.
In der Tat liegt $x \in Y_1 \cap Y_2$, so erfüllen die Elemente des Durchschnitts

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

und folglich $x = 0$, dies bedeutet $Y_1 \cap Y_2 = \{0\}$. Andererseits liegen $y_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $y_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ in Y_1 und $y_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ in Y_2. Da $\{y_1, y_2, y_3\}$ eine Basis von $\mathbb{R}^3 = X$ bilden, gilt auch $Y_1 + Y_2 = X$.

2.5.3 Beispiel

Wir betrachten den Unterraum $Y_1 := \{p \in P(\mathbb{K}) : p(0) = 0\}$ von $X = P(\mathbb{K})$. Dann gilt $P(\mathbb{K}) = Y_1 \oplus P_0(\mathbb{K})$, d.h. der lineare Raum aller konstanten Polynome ist ein Komplement von Y_1.

2.5.4 Satz

Es seien $Y_1, Y_2 \subseteq X$ Unterräume. Es ist $X = Y_1 \oplus Y_2$ genau dann, wenn es zu jedem $x \in X$ eindeutige $y_1 \in Y_1, y_2 \in Y_2$ mit $x = y_1 + y_2$ gilt.
Beweis:
($\Rightarrow$) Es sei Y_2 ein Komplement von Y_1 in X. Wegen $Y_1 + Y_2 = X$ lässt sich jedes $x \in X$ darstellen als $x = y_1 + y_2$ mit $y_1 \in Y_1, y_2 \in Y_2$. Um deren Eindeutigkeit zu verifizieren, sein $\hat{y}_1 \in Y_1, \hat{y}_2 \in X_2$ zwei weitere Vektoren mit $x = \hat{y}_1 + \hat{y}_2$. Dies impliziert $y_1 - \hat{y}_2 = \hat{y}_2 - y_2$ und $y_1 - \hat{y}_1 \in Y_1$ für $i = 1, 2$ und folglich $y_i - \hat{y}_i \in Y_1 \cap Y_2$ für $i = 1, 2$. Wegen $Y_1 \cap Y_2 = \{0\}$ folgt $y_1 = \hat{y}_1$ und $y_2 = \hat{y}_2$.

($\Leftarrow$) Umgekehrt seien $Y_1, Y_2 \subseteq X$ Unterräume derart, dass sich jedes $x \in X$ eindeutig als Summe $X = y_1 + y_2$ mit $y_i \in Y_i, i = 1, 2$, darstellen lässt. Dann gilt sicherlich $Y_1 + Y_2 = X$. Ist nun $x \in Y_1 \cap Y_2$, so gilt $x = x + 0 = 0 + x$ und da die Darstellung eindeutig sein muss, resultiert $x = 0$; also $Y_1 \cap Y_2 = \{0\}$.

2.5.5 Satz

Jeder Unterraum eines endlich dimensionalen linearen Raumes hat ein Komplement.
Beweis(-skizze):
Ergänze eine Basis von Y_1 zu einer Basis von X gemäß Proposition 2.4.9.

2.6 Anwendung: Matrizen und lineare Gleichungen

Es sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{m\times n}$ mit den n Spalten und den m Zeilen. Die n Spalten seien $a_1, \cdots, a_n \in \mathbb{K}^m$ und $a^1, \cdots, a^m \in \mathbb{K}^{1\times n}$ die Zeilen von A. Man bezeichnet den Unterraum $\text{span}\{a_k\}_{1\leq k\leq n} \subseteq \mathbb{K}^m$ als Spaltenraum und $\text{span}\{a^1, \cdots, a^m\} \subseteq \mathbb{K}^{1\times n}$ als Zeilenraum von A.

$$A = \begin{pmatrix} a^1 \\ \vdots \\ a^m \end{pmatrix} = (a_1, \cdots, a_n)$$

2.6.1 Definition (Rang einer Matrix)

Der Rang (rk$A$) einer Matrix $A \in \mathbb{K}^{m\times n}$ ist die Dimension ihres Zeilenraumes.

2.6.2 Bemerkung

$0 \leq \text{rk}A \leq m$

$$(L_0) Ax = 0$$

2.6.3 Proposition

Der Lösungsraum $L_0 \subseteq \mathbb{K}^n$ von (L_0) erfüllt $\dim L_0 = n - \text{rk}A$.
Beweis:
Wir können o.B.d.A (ohne Beschränkung der Allgemeinheit) annehmen, dass $A \in \mathbb{K}^{m\times n}$ in strenger Zeilen-Stufen-Form ist. Es sei r die Anzahl der Zeilen von A, welche mindestens ein Element $\neq 0$ besitzen - dies ist der Rang von A. Für $1 \leq i \leq r$ sei j_i derjenige Spaltenindex, in welcher das erste Element $\neq 0$ der i-ten Zeile steht. Weiter seien $k_1, \cdots, k_{n-r}$ diejenigen Element von $\{1, \cdots, n\}$, welche nicht in $\{j_1, \cdots, j_r\}$ sind. Dann gilt

$$L_0 = \left\{ x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n : \xi_1, \cdots, \xi_{k_{n-r}} \in \mathbb{K} \text{ und } \xi_{j_i} = -\frac{1}{a_{i,j_i}} \sum_{j=1}^{n-r} a_{i,k_j}\xi_{k_j} \text{ für } 1 \leq i \leq r \right\}$$

und $x_1, \cdots x_{n-r} \in \mathbb{K}^n$ bezeichne die Vektoren in L_0 mit $\xi_{k_j} = 1$ und $\xi_{k_i} = 0$ für $i \neq j$. Man überlegt sich nun, dass $\{x_1, \cdots x_{n-r}\}$ eine Basis von L_0 und die Behauptung folgt.

3 Lineare Abbildungen

Es seien X, Y lineare Räume über dem selben Körper $\mathbb{K}$.

3.1 Grundlagen

3.1.1 Definition (lineare Abbildung)

Eine lineare Abbildung $T : X \to Y$ erfüllt die Eigenschaft

$$(3.1a)\ T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T x_1 + \alpha_2 T x_2$$

für alle $\alpha_1, \alpha_2 \in \mathbb{K}, x_1, x_2 \in X$. Für die Menge aller solchen linearen Abbildungen schreiben wir $L(X, Y)$.
Für lineare Abbildungen schreibt man $Tx := T(x)$.

3.1.2 Bemerkung

(1) Für $T \in L(X, Y)$ ist $T0 = 0$

(2) Die Menge $L(X, Y)$ ist ein Unterraum von $F(X, Y)$; wir kürzen ferner ab $L(X) := L(X, X)$. Ist Z ein weiterer linearer Raum und $T \in L(X, Y), S \in L(Y, Z)$, so ist auch die Komposition $S \circ T : X \to Z$ linear. $(L(X), \circ)$ ist eine Halbgruppe mit neutralem Element id_x.

3.1.3 Beispiel

Die Nullabbildung $0 : X \to Y,\ 0x := 0 \in Y$ ist linear, wie auch die identische Abbildung $id_x : X \to X$ aus Beispiel 1.2.3

3.1.4 Beispiel (affine Abbildungen)

Eine Abbildung von $S : X \to Y$ heißt affin, falls es $T \in L(X, Y)$ und $y \in Y$ derart gibt, dass $S(x) = Tx + y$. S ist genau dann linear , falls $y = 0$.

3.1.5 Beispiel (die Abbildung T_A)

Die wichtigsten linearen Abbildungen dieser Vorlesung sind von der Form $T_A : \mathbb{K}^n \to \mathbb{K}^m,\ T_A x = Ax$ mit $A \in \mathbb{K}^{m \times n}$. Auch die Abbildung $\mathbb{K}^{m \times n} \to L(\mathbb{K}^n, \mathbb{K}^m),\ A \mapsto T_A$ ist linear.

3.1.6 Beispiel

(1) Es sei $\Omega \neq \emptyset$ eine Menge und $x \in \Omega$. Dann ist die Auswertung $ev_x : F(\Omega, X) \to X,\ ev_x(u) := u(x)$ linear.

(2) Es sei $I \subseteq \mathbb{R}$ ein Intervall. Dann ist die Differenziation $D : C^1(I, \mathbb{R}) \to C(I, \mathbb{R}),\ Du := u'$ linear.

(3) Mit einem Intervall $I \subseteq \mathbb{R}$, den fixen $t_0 \in I$ und reellen Zahlen $a < b$ definieren auch nachfolgende Integrale lineare Abbildungen:

$$T_1 : C([a,b], \mathbb{R}) \to \mathbb{R}\ T_{1,u} := \int_a^b u(s)ds$$

$$T_2 : C(I, \mathbb{R}) \to C^1(I, \mathbb{R}),\ (T_{2,u})(t) := \int_{t_0}^t u(s)ds$$

3.1.7 Beispiel (Vorwärts-Shift)

Es sei X ein linearer Raum und $\mathbb{I} \in \{\mathbb{N}_0, \mathbb{Z}\}$. Bezeichnet dann $l(\mathbb{I})$ den linearen Raum aller Folgen $F(\mathbb{I}, X)$, so ist der durch $(S\phi)_k := \phi_{k+1}$ definierte Vorwärts-Shift eine Abbildung $S \in L(l(\mathbb{I}))$.

3.1.8 Definition (Kern, Bild, Rang)

Ist $T \in L(X, Y)$, so bezeichnet $N(T) := \{x \in X : Tx = 0\}$ den Kern, $R(T) := TX$ das Bild und $\mathrm{rk}T := \dim R(T)$ den Rang von T. Eine Verbindung des Begriffes „Rang einer Matrix“ (Definition 2.6.1) und Definition 3.1.8 wird in Satz 3.3.8 hergestellt.

3.1.9 Proposition

Für jedes $T \in L(X, Y)$ ist der Kern $N(T)$ ein Unterraum von X.
Beweis: Übungsaufgabe.

3.1.10 Satz

Für jedes $T \in L(X, Y)$ gilt:

(a) T ist genau dann injektiv, wenn $N(T) = \{0\}$

(b) T ist genau dann surjektiv, wenn $R(T) = Y$

Beweis:

(a) Die Abbildung T ist genau dann nicht injektiv, wenn es $y \in Y$ und $x_1, x_2 \in X$ derart gibt, dass $x_1 \neq x_2$ und $Tx_1 = y = Tx_2$. Dies ist äquivalent zu $T(x_1 - x_2) = 0$, also $0 \neq x_1 - x_1 \in N(T)$

(b) ist genau die Definition von Surjektivität.

3.1.11 Beispiel

(1) Die Auswertung $ev_x : F(\Omega, X) \to X$ aus Beispiel 3.1.6 (1) hat den Kern $N(ev_x) := \{m \in F(\Omega, X) : u(x) = 0\}$ und das Bild $R(ev_x) = X$, ein Urbild zu einem beliebigen $y \in X$ ist gerade die konstante $u(x) \equiv y$ auf Ω.

(2) Für die Nullabbildung $0 \in L(X, Y)$ ist $N(0) = X$ und $R(0) = \{0\}$, für $X \neq \{0\}$ ist 0 nicht injektiv. Für $Y \neq \{0\}$ ist 0 nicht surjektiv.

(3) Mit einer Matrix $A \in \mathbb{K}^{m \times n}$ ist $T_A \in L(\mathbb{K}^n, \mathbb{K}^m)$ aus Beispiel 3.1.5 genau dann

- injektiv, wenn die linear homogene Gleichung (L_0) nur die triviale Lösung hat.
- surjektiv, wenn es für jede Inhomogenität $b \in \mathbb{K}^m$ mindestens eine Lösung $x \in \mathbb{K}^n$ von (L_b) gibt.

(4) Bei der Differenziation $D : C^1([a, b], \mathbb{R}) \to C([a, b], \mathbb{R})$ aus Beispiel 3.1.6 (2) besteht der Kern $N(D)$ aus allen konstanten Funktionen. Für das Bild $R(D)$ erhalten wir dagegen $C([a, b], \mathbb{R})$, denn für ein beliebiges $v \in C([a, b], \mathbb{R})$ gilt nach dem Hauptsatz der Differential- und Integralrechnung $Du = v$ mit $u(t) := v(a) + \int_a^t v(s)ds$. Damit ist D nicht injektiv, aber surjektiv.

3.1.12 Satz (Dimensionssatz)

Für jede $T \in L(X, Y)$ mit $\dim X < \infty$ gilt $\dim N(T) + \dim R(T) = \dim X$.
Beweis: Es sei $\{x_1, \cdots, x_m\}$ eine Basis von $N(T)$ und $\{y_1, \cdots, y_n\}$ eine Basis von $R(T)$. Wir wählen $\hat{x}_1, \cdots, \hat{x}_n \in X$ derart, dass $T\hat{x}_i = y_i, y \leq i \leq n$ gilt und weisen nach, dass $\mathcal{X} := \{x_1, \cdots, x_m, \hat{x}_1, \cdots, \hat{x}_n\}$ eine Basis von X ist.

(i) $\mathcal{X}$ ist linear unabhängig: $\dim N(T) + \dim R(T) = \dim X$

(ii) $\mathcal{X}$ ist ein EZS: $\dim m + \dim n = \dim X$

3.1.13 Korollar

Sei $T \in L(X, Y)$ ist $\{x_1, \cdots, x_n\}$ eine Basis von $N(T)$ und $\{x_1, \cdots, x_n, x_{n+1}, \cdots, x_d\}$ eine Basis von X mit $n < d$. Das Bild $R(T)$ hat folgende Basis:

$$\{Tx_{n+1}, \cdots, Tx_d\}$$

Beweis: Sei $d = \dim X$,$n = \dim N(T)$. Nach Dimensionssatz 3.1.12 gilt: $\dim R(T) = d - n$. Wir suchen $d - n$ linear unabhängige Vektoren in $R(T)$. Die Vektoren $\{Tx_{n+1}, \cdots, Tx_d\}$ sind $d - n$ Vektoren in $R(T)$. Wir zeigen , dass diese linear unabhängig sind. Hierzu gehen wir indirekt vor. Wir nehmen an, dass $\{Tx_{n+1}, \cdots, Tx_d\}$ linear abhängig sind.
$\Rightarrow$ Es existiert ein Index $j^*, n < j^* \leq d$, so dass $Tx_{j^*} = \sum_{\substack{j=n+1 \\ j \neq j^*}}^d \eta_j Tx_j$. Das heißt $\sum_{j=n+1}^d \eta_j Tx_j = 0$

mit $\eta_{j^*} = -1 \circledast\circledast$. Aus der Linearität von T folgt: $T(\sum_{j=n+1}^{d} \eta_j x_j) = 0$, d.h. $\sum_{j=n+1}^{d} \eta_j x_j \in N(T)$. Weil $\{x_1, \cdots, x_n\}$ Basis von $N(T)$ ist gilt: $\sum_{j=n+1}^{d} \eta_j x_j = \sum_{j=1}^{n} \eta_j x_j$ für geeignete $n_1, \cdots, n_n \in \mathbb{K}$, also: $\sum_{j=1}^{n} \eta_j x_j - \sum_{j=n+1}^{d} \eta_j x_j = 0 \circledast$
Weil nach Voraussetzung $x_1, \cdots x_d$ Basis von X ist, folgt aus $\circledast$, dass $\eta_j = 0 \; \forall 1 \leq j \leq d$ (Definition der linearen Unabhängigkeit). Das ist ein Widerspruch zu $\circledast\circledast$. Das heißt $\{Tx_{n+1}, \cdots, Tx_d\}$ sind linear unabhängig.

3.1.14 Satz (Prinzip der linearen Fortsetzung)

Sei $\{x_1, \cdots, x_n\}$ ein Basis von X und $\{\hat{y}_1, \cdots, \hat{y}_n\} \in Y$.

(a) Sind $T, S \in L(X, Y)$ zwei linearen Abbildungen mit $Tx_i = Sx_i \; \forall 1 \leq i \leq n$. Dann gilt $T = S$

(b) Es existiert genau eine lineare Abbildung $T \in L(X, Y)$ mit $Tx_i = \hat{y}_i, \; \forall 1 \leq i \leq n$

3.1.15 Bemerkung

Für gegebenes $x = \sum_{k=1}^{n} \xi_k x_k, \; \xi_k \in \mathbb{K} \; \forall 1 \leq k \leq n$, gilt $Tx = T(\sum_{k=1}^{n} \xi_k x_k) \overset{(3.1a)}{=} \sum_{k=1}^{n} \xi_k Tx_k$. Kenntnis der Koeffizienten ξ_k und der Werte $Tx_i, 1 \leq i \leq n$ erlaubt uns den Wert von Tx zu bestimmen.
Beweis (Satz 3.1.14):

(a) Sei $Tx_i = Sx_i; \; \forall 1 \leq i \leq n$. Sei $x \in X$ mit $x = \sum_{i=1}^{n} \xi_i x_i, \; 1 \leq i \leq n \; \xi_i \in \mathbb{K}$.

$$Tx = \sum_{i=1}^{n} \xi_i Tx_i = \sum_{i=1}^{n} \xi_i Sx_i = S(\sum_{i=1}^{n} \xi_i x_i) = Sx$$

.

(b) Wir definieren T wie folgt. Der Vektor x habe die Darstellung $x = \sum_{i=1}^{n} \xi_i x_i \; \xi_i \in \mathbb{K}$. Wir definieren $Tx := \sum_{i=1}^{n} \xi_i \hat{y}_i$. Dann gilt $Tx_j = \sum_{j=1}^{n} \xi_{i,j} \hat{y}_i = \hat{y}_j$. Zeige noch: T ist linear. Sei $z \in X$ dargestellt als $z = \sum_{i=1}^{n} \beta_i x_i$ und $\lambda \in \mathbb{K}$.
zeige: $T(x+z) = T(x) + T(z)$ und $T(\lambda x) = \lambda T(x)$

$$T(x+z) = T(\sum_{i=1}^{n} \xi_i x_i + \sum_{i=1}^{n} \beta_i x_i) = T(\sum_{i=1}^{n} (\xi_i + \beta_i) x_i) \overset{def.}{=} \sum_{i=1}^{n} (\xi_i + \beta_i) \hat{y}_i = \sum_{i=1}^{n} \xi_I \hat{y}_i + \sum_{i=1}^{n} \beta_i \hat{y}_i = Tx + Tz$$

$$T(\lambda x) = \lambda Tx \text{ analog}$$

3.2 Isomorphismen

3.2.1 Definition

Eine bijektive Abbildung $T \in L(X, Y)$ heißt Isomorphismus, und wir definieren $GL(X, Y) = \{T \in L(X, Y) : T \text{ bijektiv}\}$. Lineare Räume X, Y werden als isomorph bezeichnet, wenn es einen Isomorphismus $T \in L(X, Y)$ gibt. Schreibweise: $X \cong Y$.

3.2.2 Bemerkung

(a) Wenn Z ein weiterer $\mathbb{K}$-Vektorraum ist, und $T \in GL(X,Y)$ und $S \in GL(Y,Z)$, dann ist $S \circ T \in GL(X,Z)$. bildlich: $X \stackrel{T}{\to} Y \stackrel{S}{\to} Z$. Wir schreiben $GL(X)$ für $GL(X,X)$. Mit neutralem Element id_x wird $GL(X)$ zu einer Gruppe, der sogenannten General Linear Group. Achtung: $GL(X)$ ist kein Untervektorraum von L(X)!

(b) Durch $A = \{(X,Y) : X \text{ und } Y \text{ sind Isomorph}\}$ wird eine Äquivalenzrelation auf der Menge aller linearen Räume erklärt.

3.2.3 Beispiel (Transponierte)

Die Abbildung $\circ^T : K^{n\times m} \to K^{m\times n}$

$$A = (a_{i,j})_{\substack{1\leq i\leq n \\ 1\leq j\leq m}} \mapsto A^T = (a_{j,i})_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$$

(Zeilen und Spalten vertauschen!)

ist ein Isomorphismus. Es gilt das Inverse des Transponieren ist das Transponieren selbst, d.h. $((A)^T)^T = A$ (für $n = m$). Damit ist der Raum der n-Spalten isomorph zum Raum der n-Zeilen.

3.2.4 Beispiel (Polynome)

(1) Sei $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $P_n(\mathbb{K})$ die Polynome von maximalem Grad $n \in \mathbb{N}_0$. $P_n(\mathbb{K})$ ist isomorph zu $\mathbb{K}^{n+1}$ via den Isomorphismus $T : P_n(\mathbb{K}) \to \mathbb{K}^{n+1}$, $p \to (\alpha_0, \cdots, \alpha_n)$, wobei $p(t) = \sum_{k=0}^{n} \alpha_k t^k$. Zum Beispiel: $p(t) = t^2 + t$ $p \mapsto (0,1,1,0,\cdots,0)$.

(2) $l_{0,0} = \{(\alpha_k)_{k\in\mathbb{N}_0} : \exists n_0 \ \forall n \geq n_0 \ \alpha_n = 0\}$ bezeichnet die Menge aller Folgen, die schließlich 0 sind.
$l_{0,0}$ ist Isomorph zum Raum der Polynome $P(\mathbb{K})$ via den Isomorphismus $T \cdot CP(\mathbb{K}) \to l_{0,0}$; $p \mapsto (\alpha_0, \cdots, \alpha_n, 0, \cdots, 0, \cdots) : p = \sum_{k=0}^{n} \alpha_k t^k$

3.2.5 Lemma

Für $T \in L(X,Y)$ ist mit $\mathcal{S} \subseteq X$ auch $T\mathcal{S}$ linear abhängig (in Y).

3.2.6 Satz

Es sei $T \in GL(X,Y)$. Eine Menge $\mathcal{S} \subseteq X$ ist genau dann linear abhängig, wenn $T\mathcal{S} \subseteq Y$ linear abhängig ist.

3.2.7 Bemerkung

Als logische Kontraposition erhalten wir, dass Isomorphismen linear unabhängige Mengen (oder Basen) auf ebensolche abbilden.
Beweis:

- ($\Rightarrow$) Folgt aus Lemma 3.2.5
- ($\Leftarrow$) Nun sei $T\mathcal{S} \subseteq Y$ linear abhängig. Wegen $T \in GL(X,Y)$ ist auch $T^{-1}(T\mathcal{S})$ nach Lemma 3.2.5 linear abhängig.

3.2.8 Proposition

Jeder n-dimensionale lineare Raum ist isomorph zu $\mathbb{K}^n$, $n \in \mathbb{N}_0$.
Beweis: Es sei X ein linearer Raum mit $m := \dim X$ und der Basis $\mathcal{X} = \{x_1, \cdots, x_n\}$. Wir definieren $T : \mathbb{K}^n \to X$, $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \mapsto \sum_{k=1}^n \xi_k x_k$. Aufgrund der linearen Unabhängigkeit von $\mathcal{X}$ ist $N(T) = \{0\}$, nach Satz 3.1.10(a) ist T dann injektiv. Da $\mathcal{X}$ ein EZS ist, muss T auch surjektiv sein.

3.2.9 Satz

Endlich dimensionale lineare Räume X, Y sind genau dann isomorph, wenn $\dim X = \dim Y$.
Beweis

- ($\Leftarrow$) Es sei $n := \dim X = \dim Y$. Nach Proposition 3.2.8 existieren Isomorphismen $\Phi : X \to \mathbb{K}^n, \Psi : Y \to \mathbb{K}$, womit $\Psi^{-1} \circ \Phi \in GL(X,Y)$ ein Isomorphismus ist.
- ($\Rightarrow$) Sei $T \in GL(X,Y)$, $\mathcal{X} = \{x_1, \cdots, x_n\}$. Dann ist $y_i := Tx_i,\ 1 \leq i \leq n$ nach Satz 3.2.6 Basis von Y.

3.2.10 Satz

Für jedes $T \in L(X,Y)$ zwischen linearen Räumen X, Y mit $\dim X = \dim Y$ sind äquivalent:

(a) T ist ein Isomorphismus (d.h. $T \in GL(X,Y)$)

(b) T ist injektiv

(c) T ist surjektiv

Beweis: Es ist nachzuweisen, dass Surjektivität und Injektivität äquivalent sind. Es sei $T \in L(X,Y)$ injektiv. Wegen Satz 3.1.10(a) ist dies äquivalent zu $N(T) = \{0\}$. Mit dem Dimensionssatz3.1.12 ist dann $\dim R(T) = \dim X - \dim N(T) = \dim Y$ und T ist genau dann injektiv, wenn $\dim R(T) = \dim Y$, d.h. $R(T) = Y$ gilt. Aufgrund von Satz 3.1.10(b) folgt die Behauptung.

3.3 Lineare Abbildungen und Matrizen

X, Y lineare Räume, $\mathcal{X} = \{x_1, \cdots, x_n\}$, $\mathcal{Y} := \{y_1, \cdots, y_m\}$. Wir folgern aus Satz 3.1.14, dass eine lineare Abbildung $T \in L(X, Y)$ durch die Bilder Tx_i der Basisvektoren von X bestimmt ist. Sind etwa:

$$(3.3a)\ Tx_i = \sum_{j=1}^{m} a_{i,j} y_j \text{ für } 1 \leq i \leq n$$

und $x = \sum_{k=1}^{n} \xi_k x_k$ mit $Tx = \sum_{j=1}^{m} \eta_j y_j$, so resultiert $Tx = T(\sum_{k=1}^{n} \xi_k x_k) = \sum_{k=1}^{n} \xi_k T x_k \overset{3.3a}{=} \sum_{k=1}^{n} \xi_k \sum_{j=1}^{m} a_{k,j} y_j = \sum_{i=1}^{m} (\sum_{k=1}^{n} a_{i,k} \xi_k) y_i$. Für die Koordinaten $\eta_i, \cdots, \eta_m \in \mathbb{K}$ von Tx bzgl. $\mathcal{Y}$ erhalten wir:

$$(3.3b)\ \eta_i = \sum_{k=1}^{n} a_{i,k} \xi_k \text{ für } 1 \leq i \leq m$$

3.3.1 Satz (darstellende Matrix)

Jedes $T \in L(X, Y)$ wird eindeutig durch eine Matrix $T_{\mathcal{X}}^{\mathcal{Y}} \in \mathbb{K}^{m \times n}$ beschreiben, in deren $k-$ter Spalte gerade die Koordinaten von Tx_k bzgl der Basis $\mathcal{Y}$ stehen. Man nennt $T_{\mathcal{X}}^{\mathcal{Y}}$ die T darstellende Matrix in den Basen $\mathcal{X}$ und $\mathcal{Y}$; im Fall $X = Y$ und $\mathcal{X} = \mathcal{Y}$ schreiben wir $T_{\mathcal{X}} := T_{\mathcal{X}}^{\mathcal{X}}$.

3.3.2 Bemerkung

Wir versehen $X = \mathbb{K}^n$ und $Y = \mathbb{K}^m$ mit Standardbasen $\mathcal{E}_n$ bzw. $\mathcal{E}_m$ aus Beispiel 2.4.3. Für jede Abbildung $T \in L(\mathbb{K}^n, \mathbb{K}^m)$ mit darstellender Matrix

$$T_{\mathcal{E}_n}^{\mathcal{E}_m} \text{ gilt dann } T = T_{T_{\mathcal{E}_n}^{\mathcal{E}_m}} \text{ (Erinnerung } T_A x = Ax)$$

3.3.3 Beispiel (Polynome)

Es sei $n \in \mathbb{N}$ und $X = P_n(\mathbb{R})$ ausgestattet mit der monomialen Basis $\mathcal{M}_n := \{m_0, \cdots, m_n\}$ aus Beispiel 2.4.4. Als lineare Abbildung betrachten wir die Ableitung $D : P_n(\mathbb{R}) \to P_n(\mathbb{R})$ mit den Bildern

$$D_{m_0} = 0,\ D_{m_k} = k m_{k-1} \text{ für alle } k \in \mathbb{N}$$

und erhalten aus Satz 3.3.1 die darstellende Matrix

$$D_{\mathcal{M}_n} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots \\ 0 & 0 & 0 & 0 & \cdots & n-1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

3.3.4 Proposition

Zu jedem $A \in K^{m\times n}$ gibt es ein $T \in L(X,Y)$ mit $A = T_{\mathcal{X}}^{\mathcal{Y}}$.
Beweis: Es sei $A \in \mathbb{K}^{m\times n}$. Zu beliebigen $x \in X$ finden wir $\xi_i, \cdots, \xi_n \in \mathbb{K}$ mit $x = \sum_{k=1}^{n} \xi_k x_k$. Die gesuchte Abbildung $T \in L(X,Y)$ ist dann

$$Tx := \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j}\xi_j \right) y_i$$

3.3.5 Korollar

Für endlich dimensionale Räume X, Y sind $L(X,Y)$ und $\mathbb{K}^{\dim Y \times \dim X}$ isomorph. Insbesondere ist $\dim L(X,Y) = \dim X \cdot \dim Y$.
Beweis: Es sei $m = \dim Y$,$n = \dim X$. Wir verwenden die lineare Abbildung $\Phi : L(X,Y) \to \mathbb{K}^{\dim Y \times \dim X}$,$\Phi(T) = T_{\mathcal{X}}^{\mathcal{Y}}$ die in Proposition 3.3.4 konstruiert wurde.

3.3.6 Satz

Es sei Z ein weiterer linearer Raum mit Basis $\mathcal{Z}$. Für $T \in L(X,Y),\ S \in L(Y,Z)$ ist

$$(S \circ T)_{\mathcal{X}}^{\mathcal{Z}} = S_{\mathcal{Y}}^{\mathcal{Z}} T_{\mathcal{X}}^{\mathcal{Y}}$$

3.3.7 Bemerkung

Für $A \in \mathbb{K}^{l\times m},\ B \in \mathbb{K}^{m\times n}$ gilt

$$(3.3c)\ T_A \circ T_b = T_{AB}$$

3.3.2 Die Abbildung T_A

In diesem Abschnitt sei $A \in \mathbb{K}^{m\times n}$ und $T_A \in L(\mathbb{K}^n, \mathbb{K}^m)$ mit $T_A x := Ax$.

3.3.8 Satz

Die Ränge von $A \in \mathbb{K}^{m\times n}$ und $T_A \in L(\mathbb{K}^n, \mathbb{K}^m)$ stimmen überein: rkT_A =rk$A$.

3.3.9 Bemerkung

Mit den Spalten $a_1, \cdots a_n \in \mathbb{K}^m$ von A gilt $R(T_A) = \operatorname{span}\{a_1, \cdots, a_n\}$, weshalb insbesondere $R(T_A)$ und der Spaltenraum von A gleiche Dimension haben. Andererseits war rk$A$ nach Definition 2.6.1 die Dimension des Zeilenraumes von A. Daher wird Satz 3.3.8 auch formuliert als: „Spaltenrang=Zeilenrang“.

Beweis: Zunächst merken wir an, dass $N(T_A)$ mit dem Lösungsraum $L_0 \subseteq \mathbb{K}^n$ einer homogenen Gleichung (L_0) übereinstimmt. Nach Proposition 2.6.3 gilt also $\dim N(T) = n - \mathrm{rk}A$ Andererseits liefert der Dimensionssatz 3.1.12, dass $n = \dim N(T_A) + \dim N(T_A) = \dim N(T_A) + \mathrm{rk}T_A$ und folglich $\mathrm{rk}T_A = \mathrm{rk}A$.

Nun sei $A \in \mathbb{K}^{n\times n}$ quadratisch mit $\mathrm{rk}A = n$. Dies ist mit Satz 3.2.10 äquivalent dazu, dass $T_A \in L(\mathbb{K}^n)$ ein Isomorphismus ist. Wir interessieren uns nun für die simultane Lösbarkeit von $Ax = e$; für alle $1 \leq i \leq n$ welche wir mittels der augmentierten Matrix $(A, e_1, \cdots e_n) = (A, I_n) \in \mathbb{K}^{n\times(2n)}$ notieren. Vermöge des Gauß-Verfahrens lässt sich (A, I_n) auf die Form (I_n, B) mit einem $B \in \mathbb{K}^{n\times n}$ bringen; nach Konstruktion ist $A \cdot B = I_n$.

3.3.10 Definition (inverse Matrix)

Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt invertierbar, falls ein $B \in \mathbb{K}^{n\times n}$ mit der Eigenschaft $A \cdot B = I_n$ existiert. Man nennt B die Inverse von A und schreibt $A^{-1} := B$.

3.3.11 Beispiel

Wir wollen die Matrix $A := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}$ invertieren. Nach obigen Schema

$$\left.\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \end{array}\right. \Leftrightarrow \left.\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 \\ 0 & 2 & 3 & 3 & 0 & -1 \end{array}\right. \Leftrightarrow \left.\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right. \Leftrightarrow \left.\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & 6 & -3 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right.$$

$$\Leftrightarrow \left.\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right.$$

und erhalten $A^{-1} = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{pmatrix}$.
Eine testweise Multiplikation ergibt $AA^{-1} = A^{-1}A = I_3$.

3.3.12 Korollar

Die inverse Matrix $A^{-1} \in \mathbb{K}^{n\times n}$ ist eindeutig bestimmt mit $A^{-1}A = I_n$. Ferner ist A genau dann invertierbar, wenn $T_A \in GL(\mathbb{K}^n)$ ist.

3.3.13 Definition (regulär, singulär)

Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt regulär, falls $\mathrm{rk}A = n$ gilt, andernfalls nennen wir sie singulär.

3.3.14 Satz (Charakterisierung regulärer Matrizen)

Folgende Aussagen sind äquivalent für jedes $A \in \mathbb{K}^{n\times n}$

(a) $T_A \in GL(\mathbb{K}^n)$

(b) A ist regulär

(c) Die Zeilen von A sind linear unabhängig

(d) Die Spalten von A sind linear unabhängig

(e) Die homogene Gleichung (L_0) hat nur die triviale Lösung

(f) Für jedes $b \in \mathbb{K}^n$ ist (L_0) eindeutig lösbar.

Beweis: Die Äquivalenz von (b) und (c) ist Definition 3.3.13. Aufgrund von Satz 3.3.8 und Bemerkung 3.3.9 sind auch (c) und (d) gleichwertig.
(d) $\Rightarrow$ (e) resultiert aus Definition 2.3.6 der linearen Unabhängigkeit.
(e) $\Rightarrow$ (f) ergibt sich aus Satz 1.4.9(a).
(f) $\Rightarrow$ (a) Nach unserer Voraussetzung (f) ist $T_A^{-1}(\{b\})$ für jedes $b \in \mathbb{K}^n$ einpunktig. Also ist $T_A \in L(\mathbb{K}^n)$ bijektiv.
(a) $\Rightarrow$ (b) Übung!
Zum Abschluss des Unterabschnittes beschäftigen wir uns mit linear inhomogenen Gleichungen: Ihre augmentierte Koeffizientenmatrix $(A, b) \in \mathbb{K}^{m\times(n+1)}$ definieren wir durch

$$(A, b) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} & b_m \end{pmatrix}$$

3.3.15 Satz

Es sei $b \in \mathbb{K}^m$. Eine inhomogene Gleichung (L_b) hat genau dann eine Lösung, wenn gilt

$$\mathrm{rk} A = \mathrm{rk}(A, b)$$

Beweis: Übungsaufgabe.

3.3.3 Basiswechsel

Auf einem endlich dimensionalen Raum X seien die Basen $\mathcal{X} := \{x_1, \cdots, x_n\}$ und $\mathcal{X}' := \{x'_1, \cdots, x'_n\}$ gegeben. Wie verhält sich die Darstellungsmatrix von $T \in L(X)$ wenn man von $\mathcal{X}$ auf die Basis $\mathcal{X}'$ übergeht?

- Zunächst lassen sich die Elemente von $\mathcal{X}'$ darstellen durch die Elemente von $\mathcal{X}$

  $$(3.3e)\ X_j' = \sum_{i=1}^{n} s_{ij} x_i \text{ für alle } 1 \leq j \leq n$$

  Hieraus wird die sogenannte Basiswechselmatrix $S = (s_{ij})_{1 \leq i,j \leq n}$ gebildet, welche den Übergang von $\mathcal{X}$ nach $\mathcal{X}'$ liefert:

  Spalten von S = Koordinatenvektoren der „neuen“ Basisvektoren

- Umgekehrt lassen sich die x_j durch die x_i' ausdrücken. Aus $x_j = \sum_{i=1}^n s_{ij} x_i'$ erhalten wir $x_j = \sum_{i=1}^n s_{ij} (\sum_{k=1}^n s_{ki} x_k) = \sum_{k=1}^n (\sum_{i=1}^n s_{ki} s_{ij}') x_k$. Der Ausdruck in der letzten Klammer ist genau das (k, j)-te Element von SS'. Wegen der linearen Unabhängigkeit von $\mathcal{X}$ folgern wir, dass $SS' = I_n$ und somit $S' = S^{-1}$ sein muss. Schließlich definiert man auch jede invertierbare Matrix S einen Basiswechsel gemäß (3.3e) bzw. $x_j' = Sx_j$.

3.3.16 Satz

Ist $S \in \mathbb{K}^{n \times n}$ die Basiswechselmatrix zwischen $\mathcal{X}$ und $\mathcal{X}'$, so gilt $T_{\mathcal{X}'} = S^{-1} T_{\mathcal{X}} S$, für alle $T \in L(X)$.

3.3.17 Definition (Ähnlichkeit)

Zwei Matrizen $A, B \in \mathbb{K}^{n \times n}$ heißen ähnlich, falls es eine reguläre Matrix $S \in \mathbb{K}^{n \times n}$ derart gibt, dass $B = S^{-1} AS$.

3.3.18 Bemerkung

(1) Laut Satz 3.3.16 sind Darstellungsmatrizen bezüglich verschiedener Basen ähnlich vermöge der Basiswechselmatrix

(2) Die Ähnlichkeit von Matrizen definiert eine Äquivalenzrelation auf $\mathbb{K}^{n \times n}$.

Ziel: Finde „einfache“ Repräsentanten.

4 Eigenwerte

Im gesamten Kapitel sei X ein linearer Raum über dem Körper $\mathbb{K}$.

4.1 Determinanten

Wir beginnen mit einem Exkurs über Permutationen. Dazu sei $(S_n, \circ)$ die in Beispiel 2.1.7 eingeführte symmetrische Gruppe aller bijektiven Selbstabbildungen von $\{1, \cdots, n\}$, $n \in \mathbb{N}$. Ihre Elemente σ werden als Permutationen bezeichnet (von $\{1, \cdots, n\}$) und man notiert sie als Schema:

$$\begin{bmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{bmatrix}$$

oder als $n-$Tupel $(\sigma(1), \cdots, \sigma(n))$. Weil σ bijektiv ist, kommt jede Zahl $j \in \{1, \cdots, n\}$ genau einmal in $(\sigma(1), \cdots, \sigma(n))$ vor; ferner gibt es genau $n!$ solche Permutationen.

4.1.1 Definition (Signum)

Es sei $\sigma \in S_n$ und $s(\sigma)$ bezeichnet die Anzahle der Paare $(i, j) \in \mathbb{N}^2$ mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$. Dann ist das Signum einer Permutation σ definiert durch

$$\operatorname{sgn} \sigma := (-1)^{s(\sigma)}$$

4.1.2 Beispiel

Für die identische Permutation $\mathrm{id} = (1, 2, \cdots, n)$ gilt $s(\sigma) = 0$ und folglich $\operatorname{sgn} \mathrm{id} = 1$. Weiter erhält man

$$\begin{aligned} \sigma &= (2, 1, 3, 4, \cdots, n),\ s(\sigma) = 1,\ \operatorname{sgn} \sigma = -1 \\ \sigma &= (n, n-1, \cdots, 2, 1),\ s(\sigma) = \frac{(n-1)n}{2}, \operatorname{sgn} \sigma = (-1)^{s(\sigma)} \end{aligned}$$

4.1.3 Proposition

Für alle $\sigma, \tau \in S_n$ gilt $\operatorname{sgn} \sigma \circ \tau = \operatorname{sgn} \sigma \cdot \operatorname{sgn} \tau$.

4.1.4 Bemerkung

Mittels Beispiel 4.1.2 ist $1 = \operatorname{sgn} \mathrm{id} = \operatorname{sgn} \sigma \circ \sigma^{-1}$ und damit erhalten wir

$$(4.1.a)\ \operatorname{sgn} \sigma^{-1} = \operatorname{sgn} \sigma \text{ für alle } \sigma \in S_n$$

Beweis: Es seien $\sigma, \tau \in S_n$ und $x_1, \cdots, x_n \in \mathbb{Q}$ paarweise verschieden. Dann sind auch $y_i := x_{\sigma(i)}$ mit $1 \leq i \leq n$ paarweise verschieden.

(I) Zunächst gilt die Identität

$$(4.1b)\ \operatorname{sgn}\sigma = \prod_{1\leq i<j\leq n} \frac{x_{\sigma(i)} - x_{\sigma(j)}}{x_i - x_j}$$

denn Zähler und Nenner des Produkts stimmen bis auf ihr Vorzeichen überein. Im Zähler tritt ein Faktor $x_k - x_l$ mit $k > l$ genau $s(\sigma)$-mal auf, während dies im Nenner nicht vorkommt.

(II) Aufgrund von $4.1b$ erhalten wir

$$\operatorname{sgn}\tau = \prod_{1\leq i<j\leq n} \frac{y_{\tau(i)} - y_{\tau(j)}}{y_i - y_j} = \prod_{1\leq i<j\leq n} \frac{x_{\sigma\circ\tau(i)} - x_{\sigma\circ\tau(j)}}{x_{\sigma(i)} - x_{\sigma(j)}}$$

und folglich resultiert die Behauptung aus

$$\begin{aligned}
\operatorname{sgn}\sigma\circ\tau &\overset{(4.1b)}{=} \prod_{1\leq i<j\leq n} \frac{x_{\sigma\circ\tau(i)} - x_{\sigma\circ\tau(j)}}{x_i - x_j} \\
&= \left(\prod_{1\leq i<j\leq n} \frac{x_{\sigma\circ\tau(i)} - x_{\sigma\circ\tau(j)}}{x_{\sigma(i)} - x_{\sigma(j)}}\right)\left(\prod_{1\leq i<j\leq n} \frac{x_{\sigma(i)} - x_{\sigma(j)}}{x_i - x_j}\right) \\
&= \operatorname{sgn}\tau \cdot \operatorname{sgn}\sigma
\end{aligned}$$

4.1.5 Definition (Determinante)

Die durch $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$

$$(4.1c)\ \det A := \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \prod_{i=1}^{n} a_{i\sigma(i)}$$

4.1.6 Bemerkung

(1) ($4.1c$) heißt auch Leibniz-Formel

(2) Es gilt die Beziehung $\det(\alpha A) = \alpha^n \cdot \det A$ für alle $\alpha \in \mathbb{K}, A \in \mathbb{K}^{n\times n}$, folglich ist die Determinante für $n \geq 2$ nicht linear.

4.1.7 Beispiel

Wir erhalten $\det 0 = 0$ und $\det I_n = 1$

4.1.8 Beispiel

In Dimensionen $n \leq 3$ kann die Determinante einer Matrix $A \in \mathbb{K}^{n \times n}$ verhältnismäßig einfach berechnet werden.

(1) Für $n = 1$ gilt $\det A = a_{1,1}$

(2) Für $n = 2$ ist $S_2 = \{(1,2),(2,1)\}$ und wir erhalten $\det A = \det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}$

(3) Für $n = 3$ gilt $S_3 = \{(1,2,3),(2,3,1),(3,1,2),(2,1,3),(3,2,1),(1,3,2)\}$ wobei die ersten drei Permutationen das Signum 1 besitzen und die weiteren das Signum -1 besitzen. Dies liefert die Regel von Sarrus

$$\begin{aligned}\det A = \det \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} &\left| \begin{matrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \leftarrow \text{Bildlich dargestellt wie die Regel funktioniert} \\ a_{3,1} & a_{3,2} \end{matrix} \right| \\ &= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1}\end{aligned}$$

4.1.9 Lemma

Für alle $A, B \in \mathbb{K}^{n \times n}$ gilt

(a) $\det A = \det A^T$

(b) Entsteht B durch eine Permutation $\sigma \in S_n$ der Spalten von A (d.h. ist formal $B = (a_{\sigma(1)}\, a_{\sigma(2)}, \cdots, a_{\sigma(n)})$), oder der Zeilen von A $\left(\text{d.h. formal } B = \begin{pmatrix} a^{\sigma(1)} \\ \vdots \\ a^{\sigma(n)} \end{pmatrix} \right)$, so gilt $\det B = \operatorname{sgn} \sigma \det A$

(c) Falls zwei Spalten oder Zeilen von A übereinstimmen, so ist $\det A = 0$.

Beweis: Es seien $A, B \in \mathbb{K}^{n \times n}$

(a) Durch direktes Nachrechnen und 4.1a folgt

$$\begin{aligned}\det A^T &\overset{(4.1c)}{=} \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{i=1}^{n} a_{\sigma(i)i} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn} \sigma^{-1} \prod_{j=1}^{n} a_{j\sigma^{-1}(j)} \\ &= \det A\end{aligned}$$

(b) folgt ähnlich und (c) ist etwas involvierter.

Eine zentrale Eigenschaft von Determinanten ist ihre Multiplikativität. Allerdings ist sie in Dimensionen $n > 1$ nicht additiv: Beispiel $\det(I_n + I_n) = 2^n$, aber $\det I_n = 1$.

4.1.10 Satz (Multiplikativität der Determinante)

Es gilt

$$(4.1d)\, \det(AB) = \det A \cdot \det B \text{ für alle } A, B \in \mathbb{K}^{n\times n}.$$

Zusätzlich zu Satz 3.3.14: Charakterisierung regulärer Matrizen:

4.1.11 Satz (Regularität und die Determinante)

Eine Matrix $A \in \mathbb{K}^{n\times n}$ ist genau dann regulär, wenn $\det A \neq 0$. Dann gilt

$$(4.1e)\ \det A^{-1} = \frac{1}{\det A}$$

4.1.12 Korollar

Für ähnliche Matrizen $A, B \in \mathbb{K}^{n\times n}$ ist $\det A = \det B$.

4.1.13 Bemerkung

Es sei X ein linearer Raum, mit $\dim X < \infty$. Auf Basis von Korollar 4.1.12 lässt sich auch die Determinante $\det\colon L(X) \to \mathbb{K}$ einer linearen Abbildung $T \in L(X)$. Mit ihrer darstellenden Matrix $T_{\mathcal{X}}$ ist

$$\det T := \det T_{\mathcal{X}}$$

Hierbei ist $\det T$ unabhängig von der Basis $\mathcal{X}$, denn nach Satz 3.3.16 sind alle darstellenden Matrizen ähnlich und haben gleiche Determinante.
Beweis: Mit $A, B \in \mathbb{K}^{n\times n}$ und einer regulären Matrix $S \in \mathbb{K}^{n\times n}$ mit $B = S^{-1}AS$ gilt aufgrund von Satz 4.1.10 und Satz 4.1.11

$$\det B = \det S^{-1}AS \overset{(4.1d)}{=} \det S^{-1} \det A \det S = \det A$$

Problem mit der Leibniz (4.1c): Sehr aufwändig!
Lösung: Zu gegebenem $A \in \mathbb{K}^{n\times n}$ und $1 \leq k, l \leq n$ sei $A_{kl} := (a_{ij})_{\substack{1\leq i\leq n, i\neq k\\ 1\leq j\leq n, i\neq l}} \in \mathbb{K}^{(n-1)\times n(-1)}$ diejenige Matrix, welche aus A durch Streichen der $k-$ten Zeile und der $l-$ten Spalte entsteht.

4.1.14 Proposition (Entwicklung von det)

Es sei $A \in \mathbb{K}^{n\times n}$. Für alle Indizes $k, l \in \{1, \cdots, n\}$ gilt dann:

(a) die <u>Entwicklung nach der $k-$ten Zeile</u>

$$\det A = \sum_{i=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj}$$

(b) die <u>Entwicklung nach der $l-$ten Spalte</u>

$$\det A = \sum_{i=1}^{n} (-1)^{i+l} a_{il} \det A_{il}$$

4.1.15 Beispiel

Durch Entwicklung nach der ersten Zeile erhalten wir:

$$\begin{aligned}\det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} &= 0 \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} + (-1) \det \begin{pmatrix} 3 & 5 \\ 6 & 8 \end{pmatrix} + 2 \det \begin{pmatrix} 3 & 4 \\ 6 & 7 \end{pmatrix} \\ &= 0\end{aligned}$$

Entwicklung nach der ersten Spalte liefert entsprechend:

$$\begin{aligned}\det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} &= 0 \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} + (-1) 3 \det \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} + 6 \det \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \\ &= 0\end{aligned}$$

4.1.16 Beispiel (Dreiecksmatrizen)

Ist $A \in \mathbb{K}^{n\times n}$ eine Dreiecksmatrix, so gilt $\det A = \prod_{i=1}^{n} a_{ii}$.

4.1.17 Proposition (Inverse und det)

Ist $A \in \mathbb{K}^{n\times n}$ regulär.

$$A^{-1} = \frac{1}{\det A} \left((-1)^{i+j} \det A_{ji} \right)_{1\leq j, i \leq n}$$

4.1.18 Beispiel

In $\mathbb{K}^{2\times 2}$ gilt

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Beweis: Mittels der Matrix $B := ((-1)^{i+j} \det A_{ji})_{1\leq i,j\leq n}$ erhalten wir durch Nachrechnen $AB = BA = \det A \cdot I_n$. Wegen Korollar 3.3.12 ist $(\det A)^{-1}B$ die Inverse von A.

4.2 Eigenwerte und Eigenvektoren

Es sei X ein linearer Raum über $\mathbb{K}$.
Ziel: Finde zu $T \in L(X)$ eine Basis $\mathcal{X}$ von X derart, dass $T_{\mathcal{X}} \in \mathbb{K}^{n\times n}$ möglichst „einfach" ist. Hilfreich sind hierbei diejenigen Vektoren, welche von T auf ein Vielfaches abgebildet werden.

4.2.1 Definition (Eigenwert, Eigenvektor und Eigenraum)

Existiert zu $T \in L(X)$ ein Skalar $\lambda \in \mathbb{K}$ und ein Vektor $x \in X \setminus \{0\}$ mit

$$(4.2a)\ Tx = \lambda x$$

So nennt man x den zum Eigenwert λ gehörigen Eigenvektor von T. Der Kern $E_\lambda := N(T - \lambda \operatorname{id} x)$ wird Eigenraum von T und dessen Dimension die geometrische Vielfachheit von λ genannt. Das Spektrum $\sigma(T) \subseteq \mathbb{K}$ ist die Menge aller Eigenwerte.

4.2.2 Bemerkung

(1) Eine Abbildung $T \in L(X)$ besitzt genau dann einen nichttrivialen Kern, falls $0 \in \sigma(T)$ gilt. Im Fall $\dim X < \infty$ ist T genau dann invertierbar, wenn $0 \notin \sigma(T)$.

(2) Für jedes $\lambda \in \sigma(T)$ ist der zugehörige Eigenraum E_λ invariant bezüglich T, d.h.

$$x \in E_\lambda \Rightarrow Tx \in E_\lambda$$

Ist insbesondere $m := \dim E_\lambda < \infty$, so besitzt $S := T\,|_{E_\lambda} \in L(E_\lambda)$ bezüglich jeder Basis $\mathcal{X}$ von E_λ die Darstellung $S_{\mathcal{X}} = \operatorname{diag}(\lambda, \cdots, \lambda) \in \mathbb{K}^{m\times m}$.

(3) Mit $A \in \mathbb{K}^{n\times n}$ besteht das Spektrum $\sigma(T_A)$ aus allen $\lambda \in \mathbb{K}$ derart, dass der Lösungsraum der homogenen Gleichung $[A - \lambda I_n]x = 0$ nichttrivial ist; letzterer stimmt mit E_λ überein.

4.2.3 Beispiel

(1) Mit der Nullabbildung $0 \in L(X)$ gilt $\overset{\in L(X)}{0} \quad x = \overset{\in X}{0}$ für alle $x \in X$. Folglich ist $\sigma(0) = \{0\}$, jedes $x \neq 0$ ist Eigenvektor und $E_0 = X$. Für die Identität $\mathrm{id}\, x$ gilt $\sigma(\mathrm{id}\, x) = \{1\}$ und $E_1 = X$.

(2) Wir betrachten die von $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ induzierte Abbildung $T_A \in L(\mathbb{R}^2)$:

- Wegen $T_A \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ist -1 ein Eigenwert, $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ zugehörigen Eigenvektor und $E_{-1} = \mathbb{R} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ Eigenraum zu -1; also hat -1 die geometrische Vielfachheit 1.
- Aufgrund von $T_A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist ferner 3 ein Eigenwert $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ zugehöriger Eigenvektor und $E_3 = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Geometrische Vielfachheit ist 1.

4.2.4 Beispiel (Shift-Operator)

Auf dem Folgenraum $X := F(\mathbb{Z}, \mathbb{K})$ betrachten wir den Vorwärts-Shift $(Tx)_k := x_{k+1},\ T \in L(X)$. Im Fall $\lambda \neq 0$ gilt die Eigenwert - Eigenvektor - Beziehung $(4.2a)$ genau dann, wenn

$$[x_{k+1} = (Tx)_k = \lambda x_k], \text{ für alle } k \in \mathbb{Z}$$

dies ist wiederum für jede Folge $x^\lambda \in X,\ x_k^\lambda = \lambda^k$ erfüllt. Im Fall $\lambda = 0$ gibt es dagegen keine Folge $X \neq 0$, welche $(4.2a)$ erfüllt. Daher besitzt der Shift-Operator T das Spektrum $\sigma(T) = \mathbb{K} \setminus \{0\}$ und x^λ ist ein zu $\lambda \in \sigma(T)$ gehöriger Eigenvektor. Aufgrund von $E_\lambda = \mathrm{span}\{x^\lambda\}$ besitzt jedes λ die geometrische Vielfachheit 1.

4.2.5 Proposition

Sind $\lambda_i,\ 1 \leq i \leq m$, paarweise verschiedene Eigenwerte von $T \in L(X)$ mit zugehörigen Eigenvektoren $x_i \in X$, so ist $\{x_1, \cdots, x_m\}$ linear unabhängig.
<u>Beweis</u>: Induktion über m: Im Fall $m = 1$ ist wegen $x_1 \neq 0$ nichts zu zeigen. Es gelte nun die Aussage für m und wir machen den Ansatz

$$(4.2b) \sum_{i=1}^{m+1} \xi_i x_i = 0 \text{ mit } \xi_1, \cdots, \xi_{m+1} \in \mathbb{K}$$

Es resultiert hieraus die Beziehungen

$$0 \overset{(4.2b)}{=} T(\sum_{i=1}^{m} \xi_i x_i) \overset{(3.1a)}{=} \sum_{i=1}^{m+1} \xi_i T x_i \overset{(4.2a)}{=} \sum_{i=1}^{m+1} \xi_i \lambda_i x_i$$

$$0 = \lambda_{m+1} \sum_{i=1}^{m+1} \xi_i x_i = \sum_{i=1}^{m+1} \xi_i \lambda_{m+1} x_i$$

und durch Subtraktion folgt $0 = \sum_{i=1}^{m} \xi_i(\lambda_i - \lambda_{m+1})x_i$. Laut Induktionsannahme ist $\{x_1, \cdots, x_m\}$ linear unabhängig und damit $\xi_i(\lambda_i - \lambda_{m+1}) = 0$. Da die Eigenwerte paarweise verschieden sind, folgt zunächst $\xi_i = 0,\ 1 \leq i \leq m$. Mit (4.2$b$) folgt $\xi_{m+1}x_{m+1} = 0$. Also Eigenvektor ist $x_{m+1} \neq 0$ und daher $\xi_{m+1} = 0$.

4.2.6 Korollar

Ist $n = \dim X < \infty$, so gilt:

(a) T hat höchstens n verschiedene Eigenwerte.

(b) Besitzt T genau n verschiedene Eigenwerte λ_i mit zugehörigen Eigenvektoren x_i, so ist $\mathcal{X} := \{x_1, \cdots, x_n\}$ eine Basis von X und $T_{\mathcal{X}} = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$.

Beweis:

(a) Hätte T mehr als n verschiedene Eigenwerte, so gäbe es mehr als n linear unabhängige Vektoren in X.

(b) Ergibt sich aus Korollar 2.4.15 und Bemerkung 4.2.2(2)

4.3 Das charakteristische Polynom

Im Folgendem gilt: χ =griechischer Buchstabe „chi“, und $\mathcal{X}$ ein Skript-X
X sei linearer Raum über $\mathbb{K}$, $\dim X = n < \infty$, $\mathcal{X}$ sei Basis. Dann hat jede $T \in L(X)$ eine darstellende Matrix $T_{\mathcal{X}} \in \mathbb{K}^{n \times n}$. Bezeichnet $\mathcal{X}'$ eine weitere Basis von X, so folgern wir aus Satz 3.3.16 die Existenz einer regulären Basiswechselmatrix S mit $T_{\mathcal{X}'} = S^{-1}T_{\mathcal{X}}S$ und nach den Sätzen 4.1.10, 4.1.11:

$$\det(T_{\mathcal{X}'} - tI_n) = \det(S^{-1}T_{\mathcal{X}}S - tS^{-1}S) \overset{(4.1d)}{=} \det S^{-1} \det(T_{\mathcal{X}} - tI_n) \det S \overset{(4.1e)}{=} \det(T_{\mathcal{X}} - tI_n)$$

für alle $t \in \mathbb{K}$. Also hängen die Werte von $t \mapsto \det(T_{\mathcal{X}} - tI_n)$ nur von $T \in L(X)$ und nicht von $\mathcal{X}$ ab.

4.3.1 Definition (charakteristisches Polynom)

Besitzt $T \in L(X)$ eine darstellende Matrix $T_{\mathcal{X}} \in \mathbb{K}^{n \times n}$, so ist das charakteristisches Polynom von T gegeben durch $\chi_T(t) = \det(T_{\mathcal{X}} - tI_n)$.

4.3.2 Bemerkung

Es seien $A \in \mathbb{K}^{n \times n}$ und die induzierte Abbildung $T_A \in L(\mathbb{K}^n)$ gegeben.

(1) Im Fall $X = \mathbb{K}^n$ bezeichnet man χ_{T_A} auch als charakteristisches Polynom von A.

(2) Die Spur von A ist definiert durch (vgl. Aufgabe)

$$trA = \sum_{i=1}^{n} a_{ii}$$

und mit $\chi_{T_A}(t) = c_n t^n + \cdots + c_1 t + c_0$ gilt $c_n = (-1)^n$, $c_{n-1} = (-1)^{n-1} trA$, $c_0 = \det A$.

Man sagt ein Polynom $\chi \in P_n(\mathbb{K})$ zerfällt in Linearfaktoren falls es Koeffizienten $\mu \in \mathbb{K},\ \lambda_1, \cdots, \lambda_n \in \mathbb{K}$ gibt, mit

$$(4.3a)\chi(t) = \mu \prod_{i=1}^{n} (t - \lambda_i) \text{ auf } \mathbb{K}$$

Die Werte λ_i sind dann die Nullstellen von χ, ihre algebraische Vielfachheit gibt an, wie oft der Linearfaktor $t - \lambda_i$ in (4.3a) vorkommt.

4.3.3 Satz

Für jedes $T \in L(X)$ gilt $\sigma(T) = \chi_T^{-1}(\{0\})$, und die Vielfachheit von $\lambda \in \sigma(T)$ heißt algebraische Vielfachheit des Eigenwerts λ.
Beweis: Mittels Definition 4.2.1 gelten die Äquivalenzen

$$\lambda \notin \sigma(T) \Leftrightarrow E_\lambda = N(T - \lambda \,\mathrm{id}\ \ x) = \{0\} \Leftrightarrow [T_{\mathcal{X}} - \lambda I_n]x = 0 \text{ hat nur die triviale Lösung.}$$

$$\Leftrightarrow T_{\mathcal{X}} - \lambda I_n \text{ ist regulär (Satz 3.3.14)}$$

$$\Leftrightarrow \chi_T(\lambda) = \det(T_{\mathcal{X}} - \lambda I_n) \neq 0 \text{ nach Satz 4.1.11}$$

4.3.4 Beispiel

Es seien $\alpha \in [0, 2\pi),\ A = \begin{pmatrix} cos\alpha & -sin\alpha \\ sin\alpha & cos\alpha \end{pmatrix}$

(1) Über dem Körper $\mathbb{K} = \mathbb{R}$ betrachten wir $T = T_A \in L(\mathbb{R}^2)$. Sie hat das charakteristische Polynom $\chi_T(t) = t^2 - 2cos\alpha \cdot t + 1$. Für $\alpha = 0$ ist $\sigma(T) = \{1\}$ und für $\alpha = \pi$ gilt $\sigma(T) = \{-1\}$. Die Eigenwerte sind doppelte Nullstellen von χ_T und damit von algebraischer Vielfachheit 2. Ansonsten besitzt χ_T für $\alpha \notin \{0, \pi\}$ keine reellen Nullstellen, d.h. $\sigma(T) = \emptyset$.

(2) Mit $\mathbb{K} = \mathbb{C}$ besitzt $T = T_A \in L(\mathbb{C}^2)$ für alle α das Spektrum

$$\sigma(T) = \{cos\alpha - i\sqrt{1 - cos^2\alpha}, cos\alpha + i\sqrt{1 - cos^2\alpha}\}.$$

Seine Elemente haben für $\alpha \notin \{0, \pi\}$ die algebraische Vielfachheit 1.

Wir nennen einen Körper $\mathbb{K}$ algebraisch abgeschlossen, falls jedes Polynom $p \in P(\mathbb{K})$ mit $\deg p > 0$ mindestens eine Nullstelle hat.

4.3.5 Proposition

Ist $\mathbb{K}$ algebraisch abgeschlossen, so hat jedes $T \in L(X)$ einen Eigenwert.

4.3.6 Beispiel

(1) $\mathbb{R}$ ist nicht algebraisch abgeschlossen (z.B. $t^2 + 1$). Der Fundamentalsatz der Algebra besagt gerade, dass $\mathbb{C}$ algebraisch abgeschlossen ist; folglich zerfällt auch jedes Polynom über $\mathbb{C}$ in Linearfaktoren.

(2) $\mathbb{Q}$ ist nicht algebraisch abgeschlossen (z.B. $t^2 - 2$).

4.3.7 Satz

Die geometrische Vielfachheit jedes Eigenwertes ist kleiner oder gleich seiner algebraischen Vielfachheit.
Beweis: Es sei $T \in L(X)$ und $\lambda \in \sigma(T)$ mit $m := \dim E_\lambda$. Dann existieren m linear unabhängige Eigenvektoren von T in E_λ. Wählt man diese als ersten Teil einer Basis $\mathcal{X}$ von X, so gilt (vgl. Bemerkung 4.2.2(2))

$$T_{\mathcal{X}} = \begin{pmatrix} \Lambda & B \\ 0 & C \end{pmatrix}, \Lambda = \operatorname{diag}(\lambda, \cdots, \lambda) \in \mathbb{K}^{n \times m},\ B \in \mathbb{K}^{m \times (m-n)},\ C \in \mathbb{K}^{(n-m) \times (n-m)}$$

und wir erhalten nach Satz 4.1.14, dass

$$\chi_T(t) = (\lambda - t)^m \det(C - tI_{n-m})$$

Daher ist die algebraische Vielfachheit von λ mindestens m.

4.4 Diagonalisierung und Trigonalisierung

Wir betrachten $n-$dimensionale lineare Räume X mit $n < \infty$.

4.4.1 Definition (diagonalisierbar, trigonalisierbar)

Eine Abbildung $T \in L(X)$ heißt

(a) diagonalisierbar, falls eine Basis $\mathcal{X}$ von X existiert, in der $T_{\mathcal{X}}$ eine Diagonalmatrix ist.

(b) trigonalisierbar, falls eine Basis $\mathcal{X}$ von X existiert, in der $T_{\mathcal{X}}$ obere Dreiecksform besitzt.

4.4.2 Bemerkung

(1) Entsprechend wird eine Matrix $A \in \mathbb{K}^{n\times n}$ diagonalisierbar (trigonalisierbar) genannt, falls $T_A \in L(\mathbb{K}^n)$ diese entsprechende Eigenschaft besitzt, bzw. ähnlich zu einer Diagonalmatrix (obere Dreiecksmatrix) ist.

(2) Mit oberen Dreiecksmatrizen zu arbeiten ist reine Konvention und keine mathematische Notwendigkeit.

4.4.3 Satz (Charakterisierung von Diagonalisierbarkeit)

Eine Abbildung $T \in L(X)$ ist genau dann Diagonalisierbar, wenn gilt:

(i) Ihr charakteristisches Polynom zerfällt in Linearfaktoren.

(ii) Die geometrische Vielfachheit jedes Eigenwertes stimmt mit seiner algebraische Vielfachheit überein.

Beweis: Wir zeigen zwei Richtungen:

($\Rightarrow$) Es sei $T \in L(X)$ diagonalisierbar und $\mathcal{X}$ eine Basis von X derart, dass

$$T_{\mathcal{X}} = \operatorname{diag}(a_1, \cdots, a_n) \text{ mit } a_1, \cdots, a_n \in \mathbb{K}.$$

Das charakteristisches Polynom ergibt sich daher zu

$$(4.4a)\ \chi_T = \prod_{k=1}^{n}(a_k - t) \left[= \det(T_{\mathcal{X}} - tI_n) = \det \begin{pmatrix} a_1 - t & & \\ & \ddots & \\ & & a_n - t \end{pmatrix} \right]$$

und zerfällt offenbar in Linearfaktoren. Nach Satz 4.3.7 bleibt also nachzuweisen, dass die geometrische Vielfachheit eines $\lambda \in \sigma(T)$ mindestens gleich seiner algebraischen Vielfachheit ist. Bezeichnet r die Algebraische Vielfachheit von λ, so kommt der Linearfaktor $\lambda - t$ genau $r-$mal in (4.4a) vor, d.h. r Diagonalelemente von $T_{\mathcal{X}}$ sind gleich λ. Damit werden r Elemente der Basis $\mathcal{X}$ auf ihr $\lambda-$faches abgebildet und folglich ist $\dim E_\lambda \geq r$

($\Leftarrow$) Umgekehrt zerfalle χ_T in Linearfaktoren, es gelte die Bedingung (ii) und es sei $\sigma(T) = \{\lambda_1, \cdots, \lambda_m\}$ mit paarweise verschiedenen Eigenwerten $\lambda_i,\ 1 \leq i \leq m \leq n$. Zu jedem Eigenraum E_{λ_i} wählen wir eine Basis

$$\mathcal{X}_i = \{x_1^i, \cdots, x_{r_i}^i\} \text{ mit } 1 \leq i \leq m, r_1 + \cdots r_m = n.$$

und erhalten somit n Vektoren im Raum X. Wir wollen zeigen, dass diese linear unabhängig sind. Dazu

$$(4.4b)\ 0 = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \xi_j^i x_j^i = \sum_{i=1}^{m} y_i \text{ mit } y_i = \sum_{j=1}^{r_i} \xi_j^i x_j^i \text{ und Koeffizienten } \xi_j^i$$

Die Invarianz der E_{λ_i} zeigt, dass jedes solcher $y_i \in E_{\lambda_i}$ auf sein λ_i-faches abgebildet wird, also Eigenvektor von T ist. Mit Proposition 4.2.5 liefert dies die lineare Unabhängigkeit von $\{y_1, \cdots, y_m\}$. Dann muss aber $0 = y_i = \sum_{j=1}^{r_1} \xi_j^i x_j^i,\ 1 \leq i \leq m$ gelten, denn andernfalls wäre (4.4b) eine nichttriviale Darstellung der 0. Da $\mathcal{X}_i$ Basis ist, folgt $\xi_j^i = 0$ und somit sind die n Vektoren aus $\mathcal{X}_1, \cdots, \mathcal{X}_n$ linear unabhängig, bilden also eine Basis von X, in welcher $T_{\mathcal{X}}$ diagonal ist.

4.4.4 Beispiel

(1) Wir betrachten die lineare Abbildung

$$T \in L(\mathbb{R}^3),\ Tx = \begin{pmatrix} 0 & 1 & -1 \\ 3 & 2 & -3 \\ 2 & 2 & -3 \end{pmatrix} x$$

und dem charakteristischen Polynom $\chi_T(t) = t^3 + t^2 - t - 1 = (t+1)^2(t-1)$. Es zerfällt in Linearfaktoren und der Eigenwert -1 besitzt die algebraische Vielfachheit 2. Der Lösungsraum der homogenen linearer Gleichung $[T + \mathrm{id}]x = 0$ lautet

$$E_{-1} = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

und ist der zugehörige Eigenraum zum Eigenwert-1. Somit hat -1 die geometrische Vielfachheit 2. Der weitere Eigenwert 1 ist algebraisch einfach. Seine geometrische Vielfachheit 1 ergibt sich aus der Gleichung $[T - \mathrm{id}]x = 0$, deren Lösungsraum

$$E_1 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$$

gerade der zum Eigenwert 1 gehörige Eigenraum ist. Deshalb garantiert Satz 4.4.3 die Diagonalisierbarkeit von T. In der Tat, mit der aus den Eigenvektoren von T gebildeten Basiswechselmatrix $S = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix}$ (gebildet von E_{-1} und E_1 als Spalten) gibt $S^{-1}TS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(2) Wir betrachten $T \in L(\mathbb{R}^3)$

$$Tx := \begin{pmatrix} -2 & 1 & 3 \\ 2 & 1 & -1 \\ -7 & 2 & 7 \end{pmatrix} x$$

mit dem charakteristischen Polynom $\chi_T(t) = -(t-2)^3$, das in Linearfaktoren zerfällt. Der Eigenwert 2 von T hat die algebraische Vielfachheit 3. Man verifiziert leicht $E_2 = N(T - 2\,\mathrm{id}) = \mathbb{R}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, weshalb die geometrische Vielfachheit 1 ist. Also ist T nicht diagonalisierbar.

4.4.5 Satz (Charakterisierung von Trigonalisierbarkeit)

Eine Abbildung $T \in L(X)$ ist genau dann trigonalisierbar, wenn ihr charakteristisches Polynom in Linearfaktoren zerfällt.

4.4.6 Bemerkung (Jordan-Normalform)

Der Satz 4.4.5 lässt sich wesentlich präziser fassen - wenn auch mit deutlich aufwändigeren Beweis: zu jedem $T \in L(X)$, deren charakteristisches Polynom in Linearfaktoren zerfällt, existiert eine Basis $\mathcal{X}$ von X, so dass die darstellende Matrix $T_{\mathcal{X}}$ in Jordan-Normalform ist.

$$T_{\mathcal{X}} = \mathrm{diag}(J_1, \cdots, J_r) \text{ mit } 1 \leq i \leq r$$

Hierbei besitzt jeder Jordan-Block die Form

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \in \mathbb{K}^{n_i \times m_i}, \ 1 \leq i \leq r$$

mit $\sigma(T) = \{\lambda_1, \cdots, \lambda_r\}$ und $n_1 + \cdots + n_r = n$.

4.4.7 Beispiel

(1) Mit $\alpha \in [0, 2\pi)$ und der aus Beispiel 4.3.4 bekannten Matrix:

$$A := \begin{pmatrix} cos\alpha & -sin\alpha \\ sin\alpha & cos\alpha \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

betrachten wir die induzierte lineare Abbildung $T = T_A \in L(\mathbb{R}^2)$. Ihr charakteristisches Polynom $\chi_T(t) := t^2 - 2cos(\alpha)t + 1$ hat die Diskriminante $4(cos^2\alpha - 1)$. Daher ist T genau für $\alpha \in \{0, \pi\}$ trigonalisierbar über $\mathbb{R}$.

(2) Weiter betrachten wir die in Beispiel 4.4.4(2) diskutierte Abbildung $T \in L(\mathbb{R}^3)$

$$Tx := \begin{pmatrix} -2 & 1 & 3 \\ 2 & 1 & -1 \\ -7 & 2 & 7 \end{pmatrix} x$$

Da ihr charakteristisches Polynom $\chi_T(t) = (t-2)^3$ in Linearfaktoren zerfällt, ist sie nach Satz 4.4.5 trigonalisierbar. In der Tat liefert die Matrix $S := \begin{pmatrix} -3 & -4 & 1 \\ -3 & 2 & 0 \\ -3 & -7 & 0 \end{pmatrix}$ ihre Jordan-Normalform $J = S^{-1}TS = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

4.4.8 Korollar

Jede lineare Abbildung $T \in L(X)$ auf einem endlich-dimensionalen Raum X über $\mathbb{C}$ ist trigonalisierbar.
Beweis: Mit dem Fundamentalsatz der Algebra zerfällt jedes Polynom $p \in P(\mathbb{C})$ in Linearfaktoren, also insbesondere χ_T.

4.4.9 Korollar

Zerfällt χ_T in Linearfaktoren, so gilt

$$\det T = \prod_{j=1}^{n} \lambda_j, \ tr\ T = \sum_{j=1}^{n} \lambda_j \text{ mit } \sigma(T) = \{\lambda_1, \cdots, \lambda_n\}.$$

Beweis: Nach Satz 4.4.5 ist $T \in L(X)$ trigonalisierbar und folglich ist jede Darstellung $T_\mathcal{X}$ ähnlich zu einer Dreiecksmatrix $D \in \mathbb{K}^{n\times n}$ vermöge S. Dann folgt mit Korollar 4.1.12, dass $\det T = \det D = \lambda_1, \cdots, \lambda_n$. Nach einer Übungsaufgabe ist $trT = trS^{-1}(DS) = tr(DS)S^{-1} = trD =$

$\lambda_1 + \cdots + \lambda_n$.
Wir definieren die Potenzen $T^k \in L(X),\ k \in \mathbb{N}_0$, einer Abbildung $T \in L(X)$ rekursiv.

$$T^0 := \operatorname{id} x,\ T^{k+1} := T \circ T^k \text{ für } k \in \mathbb{N}_0$$

und entsprechend für Matrizen $A \in \mathbb{K}^{n\times n}$. Wegen $\dim L(X) = n^2$(Korollar 3.3.5) ist die (n^2+1)-elementige Menge $\{T^0, T^1, \cdots, T^{n^2}\}$ linear abhängig. Tatsächlich hat sogar $\{T^0, \cdots, T^n\}$ diese Eigenschaft:

4.4.10 Satz (von Cayley-Hamilton)

Jede lineare Abbildung $T \in L(X)$ erfüllt ihr charakteristisches Polynom, d.h. formal $\chi_T(T) = 0$.

4.4.11 Bemerkung

(1) Bezeichnet $\mathcal{X}$ eine Basis von X, so gilt für die darstellende Matrix $\chi_T(T_\mathcal{X}) = 0$

(2) Für jedes $A \in \mathbb{K}^{2\times 2}$ gilt $A^2 = (trA)A - (\det A)I_2$

5 Innere Produkte

In diesem Kapitel sei stets $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Wir beschränken uns weiter auf reelle oder komplexe lineare Räume X. Insbesondere im Fall $\mathbb{K} = \mathbb{C}$ erinnern wir an das komplex-konjugierte

$$\overline{z} = (x, -y) = x - iy$$

einer komplexen Zahl $z = (x, y) = x + iy$, sowie ihren Betrag $|z| := \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}$. Wir betrachten $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ und erhalten: $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$.

5.1 Skalarprodukte und Orthogonalität

5.1.1 Definition (inneres Produkt)

Ein <u>inneres Produkt</u> auf X ist eine Abbildung $\langle\cdot,\cdot\rangle\colon X \times X \to \mathbb{K}$ mit

(i) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$ (<u>Linearität im 1. Argument</u>)

(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ (<u>konjugierte Symmetrie</u>)

(iii) $\langle x, x\rangle \geq 0$ und Gleichheit genau für $x = 0$ (<u>positive Definitheit</u>)

Für alle $x, y, z \in X,\ \alpha, \beta \in \mathbb{K}$. Ein linearer Raum mit innerem Produkt heißt auch <u>Prä-Hilbert-Raum</u>. Statt innerem Produkt sagt man auch <u>Skalarprodukt</u>.

5.1.2 Bemerkung

(1) Aufgrund der konjugierten Symmetrie ist stets $\langle x, x\rangle \in \mathbb{R}$, während die positive Definitheit $\langle x, x\rangle > 0$ für $x \neq 0$ garantiert.

(2) Ein inneres Produkt ist <u>Semilinear</u> im 2. Argument:

$$(5.1a)\ \langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle \text{ für alle } x, y, z \in X,\ \alpha, \beta \in \mathbb{K}$$

(3) Unter einer <u>Norm</u> auf X versteht man die Funktion

$$\|\cdot\|\colon X \to \mathbb{R},\ \|x\| := \sqrt{\langle x, x\rangle}$$

Insbesondere gilt $\|x\| = 0 \Leftrightarrow x = 0$. Zu jedem $x \neq 0$ nennen wir $y := \frac{1}{\|x\|}x$ den <u>normierten Vektor</u> zu x, denn $\|y\| = 1$.

5.1.3 Bemerkung (Orthogonalität)

(1) Die Elemente $x, y \in X$ heißen orthogonal, falls $\langle x, y\rangle = 0$. Die resultierende Relation

$$x \perp y :\Leftrightarrow \langle x, y\rangle = 0$$

ist symmetrisch aber nicht transitiv. Wegen $\langle x, 0\rangle = \langle 0, x\rangle = 0$ ist $0 \in X$ orthogonal zu jedem $x \in X$.

(2) Die Teilmengen $Y_1, Y_2 \subseteq X$ heißen orthogonal, insofern

$$\langle y_1, y_2\rangle = 0 \text{ für alle } y_1 \in Y_1,\ y_2 \in Y_2$$

5.1.4 Beispiel (Euklidischer Raum)

$\mathbb{R}^n$ mit $\langle x, y\rangle := \sum_{j=1}^n x_j y_j$, $\|x\| = \sqrt{\sum_{j=1}^n x_j^2}$

5.1.5 Beispiel (Unitärer Raum)

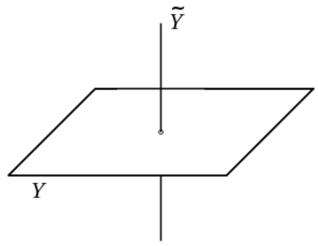
$\mathbb{C}^n$ mit $\langle x, y\rangle = \sum_{j=1}^n x_j \overline{y_j}$, $n = 1$, $\langle i, i\rangle = i \cdot (-i) = 1$.

5.1.6 Beispiel

(1) Mit sogenannten Gewichten $\omega_1, \ldots, \omega_n > 0$ ist auch $\langle x, y\rangle_\omega = \sum_{j=1}^n \omega_j x_j \overline{y}_j$ ein inneres Produkt auf $\mathbb{K}^n$ mit induzierter Norm $\|x\|_\omega = \sqrt{\sum_{j=1}^n \omega_j |x_j|^2}$

(2) Mit einer stetigen Gewichtsfunktion $\omega\colon [a, b] \to (0, \infty)$, $a < b$ sind auch die stetigen Funktionen $X = C([a, b], \mathbb{K})$ ein linearer Raum mit innerem Produkt

$$(5.1b)\ \langle x, y\rangle_\omega := \int_a^b \omega(t)x(t)\overline{y(t)}dt$$



5.1.7 Definition (Orthogonales Komplement)

Es sei $S \subseteq X$. Dann heißt $S^\perp := \{x \in X : \langle x, y\rangle = 0 \text{ für alle } y \in S\}$ das orthogonale Komplement von S (in X)

5.1.8 Beispiel

(1) Es ist $\{0\}^\perp = X$ und $X^\perp = \{0\}$

(2) Im Raum $X = \mathbb{R}^3$ haben wir in Beispiel 2.5.2 nachgewiesen, dass die Gerade $\tilde{Y} := \mathbb{R}\begin{pmatrix}1\\1\\1\end{pmatrix}$ ein Komplement zur Ebene $Y := \{x \in X, x_1 - x_2 + x_3 = 0\} = \mathbb{R}\begin{pmatrix}1\\1\\0\end{pmatrix} + \mathbb{R}\begin{pmatrix}1\\0\\-1\end{pmatrix}$ ist. Betrachtet man $\mathbb{R}^3$ als Euklidischen Raum, so ist $\tilde{Y}$ wegen $\left\langle \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix} \right\rangle = 2$ jedoch kein orthogonales Komplement von Y, vielmehr gilt $Y^\perp = \mathbb{R}\begin{pmatrix}1\\-1\\1\end{pmatrix}$

5.1.9 Proposition

Für jedes $S \subseteq X$ ist $S^\perp$ ein Unterraum von X mit $(\operatorname{span} S) \cap S^\perp = \{0\}$.
Beweis: Es seien $x_1, x_2 \in S^\perp$ und wir erhalten $\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle = 0$ für alle $\alpha_1, \alpha_2 \in \mathbb{K}$ $y \in S$ dies garantiert $\alpha_1 x_1 + \alpha_2 x_2 \in S^\perp$ und $S^\perp$ ist ein linearer Raum. Weiter sei $x \in S^\perp$ und $x \in \operatorname{span} S$, d.h. $x = \sum_{i=1}^n \xi_i x_i$ mit Koeffizienten $\xi_i \in \mathbb{K}$ und $x_i \in S$. Dies impliziert $\langle x, x\rangle = \langle \sum_{i=1}^n \xi_i x_i, x\rangle = \sum_{i=1}^n \xi_i \underbrace{\langle x_i, x\rangle}_{=0 \text{ weil } x \in S^\perp} = 0$ und folglich $x = 0$.

5.1.10 Satz

Ist $\dim X < \infty$ und Y ein Unterraum von X, so gilt

(a) $X = Y \oplus Y^\perp$

(b) $(Y^\perp)^\perp = Y$

(c) $\dim X = \dim Y + \dim Y^\perp$

Beweis (a): $Y \cap Y^\perp = \{0\}$ gilt nach Proposition 5.1.9. Es bleibt $X = Y \oplus Y^\perp$ zu zeigen. Zu jedem $x \in X$ muss es also $y \in Y$ und $y^\perp \in Y^\perp$ mit $x = y + y^\perp$ geben. Ist $\{y_1, \ldots, y_n\}$ eine Orthonormalbasis von Y, so definieren wir

$$y := \sum_{i=1}^m \langle x, y_i\rangle \cdot y_i, \qquad y^\perp := x - y \text{ usw.}$$

5.1.11 Definition (Orthogonal- und Orthonormalbasis)

Eine Familie $\mathcal{S} \subseteq X$ heißt

(a) orthogonal, falls $\langle x, y \rangle = 0$, für alle $x, y \in \mathcal{S}$ $x \neq y$

(b) orthonormal, falls $\langle x, y \rangle = \begin{cases} 1, x = y \\ 0, x \neq y \end{cases}$, für $x, y \in \mathcal{S}$

Eine Orthogonal- bzw. Orthonormalbasis von X ist eine orthogonale bzw. orthonormale Basis von X.
Jede orthonormale Familie ist auch orthogonal.

5.1.12 Beispiel

(1) Die Standardbasis $\epsilon_n = \{e_1, \ldots, e_n\}$ aus Beispiel 2.4.3 ist eine Orthonormalbasis des Euklidischen Raums $\mathbb{R}^n$ wie auch des unitären Raums $\mathbb{C}^n$.

(2) Die Legendre Polynome $p_n \in P(\mathbb{R})$, $p_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n}(t^2 - 1)^n$ für $n \in \mathbb{N}_0$ definieren eine orthogonale Familie $\{p_n\}_{n \in \mathbb{N}_0}$ bzgl. dem inneren Produkt (5.1b) aus Beispiel 5.1.6(2) mit $\omega(t) = 1$ auf $[-1, 1]$

(3) Wir definieren die trigonometrischen Funktionen $c_n, s_n \colon [-\pi, \pi] \to \mathbb{R}$, $c_n(t) := cos(nt)$, $s(t) := sin(nt)$ mit $n \in \mathbb{N}_0$. Dann ist $\mathcal{S} := \{c_n\}_{n \in \mathbb{N}_0} \cup \{s_n\}_{n \in \mathbb{N}}$ eine orthogonale Familie in $C([-\pi, \pi], \mathbb{R})$ bezüglich (5.1b) mit $\omega(t) = 1$ auf $[-\pi, \pi]$.

5.1.13 Proposition

Jede orthogonle Familie $\mathcal{S} \subseteq X$ mit $0 \notin \mathcal{S}$ ist linear unabhängig.
Beweis: Es sei $\{x_1, \ldots, x_n\}$ eine endliche Teilfamilie von $\mathcal{S}$ und $\sum_{j=1}^n \xi_j x_j = 0$ mit Koeffizienten $\xi_j \in \mathbb{K}$. Dann resultiert $0 = \langle 0, x_k \rangle = \langle \sum_{j=1}^n \xi_j x_j, x_k \rangle = \sum_{j=1}^n \xi_j \langle x_j, x_k \rangle = \xi_k \langle x_j, x_k \rangle$ für alle $1 \leq k \leq n$. Wegen $x_k \neq 0$ folgt $\xi_k = 0$.
Vorteil von Orthonormalbasen: Koordinaten einfach berechenbar

5.1.14 Proposition

Ist $\{x_1, \ldots, x_n\}$ eine Orthonormalbasis von X so gilt $x = \sum_{j=1}^n \langle x, x_j \rangle x_j$ für alle $x \in X$.
Beweis: Da $\{x_1, \ldots, x_n\}$ orthonormal ist, gilt $\langle x_j, x_k \rangle = \delta_{jk}$ für $1 \leq j, k \leq n$. Die Koeffizienten $\xi_j \in \mathbb{K}$ eines beliebigen $x = \sum_{j=1}^n \xi_j x_j$ $\quad \xi_k = \sum_{j=1}^n \xi_k \underbrace{\langle x_j, x_k \rangle}_{\delta_{jk}} = \langle \sum_{j=1}^n \xi_j x_j, x_k \rangle = \langle x, x_k \rangle$

Berechnung von Orthogonal- bzw. Orthonormalbasis aus einer gegebenen Basis: Orthogonalisierungsverfahren von Gram-Schmidt:

(0) Es sei eine linear unabhängige Familie $\mathcal{S} := \{x_1, x_2, x_3, \dots\}$ gegeben auf X mit innerem Produkt $\langle \cdot, \cdot \rangle$

(1) Setze $y_1 := x_1$

(2) Für $k \geq 1$ definiere

$$(5.1c)\ y_{k+1} := x_{k+1} - \sum_{j=1}^{k} \frac{\overline{\langle y_j, x_{k+1} \rangle}}{\langle y_j, y_j \rangle} y_j$$

für alle $k \in \mathbb{N}$ solange $x_{k+1} \in \mathcal{S}$. Wir erhalten aus $\{x_1, x_2, \cdots\}$ eine Familie orthogonaler Vektoren $\{y_1, y_2, \cdots\} = \mathcal{Y}$ und normiert man die Elemente von $\mathcal{Y}$ gemäß

$$(5.1d)\ z_k := \frac{1}{\|y_k\|} y_k,$$

so erhalten wir sogar eine orthonormale Familie.

5.1.15 Satz

Jeder endliche-dimensionale lineare Raum mit innerem Produkt hat eine Orthonormalbasis.
Beweis: Nach Satz 2.4.8 besitzt X eine Basis $\{x_1, \cdots, x_n\}$. Auf diese wenden wir obiges Gram-Schmidt-Verfahren an: Es sei $y_1 := x_1$, y_{k+1} für $1 \leq k \leq n$ durch $(5.1c)$ gegeben und $\{y_1, \cdots, y_k\}$ bereits orthogonal. Wir erhalten dann

$$\langle y_i, y_{k+1} \rangle \overset{(5.1c)}{=} \left\langle y_i, x_{k+1} - \sum_{j=1}^{k} \frac{\overline{\langle y_j, x_{k+1}}}{\langle y_j, y_j \rangle} y_j \right\rangle = \langle y_i, x_{k+1} \rangle - \left\langle y_i, \sum_{j=1}^{k} \frac{\overline{\langle y_j, x_{k+1}}}{\langle y_j, y_j \rangle} y_j \right\rangle$$

$$\langle y_i, x_{k+1} \rangle - \sum_{j=1}^{k} \frac{\langle y_j, x_{k+1} \rangle}{\langle y_j, y_j \rangle} \overbrace{\langle y_i, y_j \rangle}^{\delta_{ij}}$$

$$\langle y_i, x_{k+1} \rangle - \langle y_i, x_{k+1} \rangle = 0$$

dass y_i für alle $1 \leq i \leq k$ orthogonal auf y_{k+1} ist. Wir haben also induktiv eine orthogonale Familie $\{y_1, \cdots, y_k, y_{k+1}\}$ erzeugt. Demnach ist $\{y_1, \cdots y_n\}$ eine Orthogonalbasis von X. Die gesuchte Orthonormalbasis $\{z_1, \cdots, z_n\}$ ergibt sich aus $(5.1d)$.

5.1.16 Beispiel (Legendre-Polynome)

Aufgrund von Beispiel 2.4.4 sind die Monome $\mathcal{M}_3 := \{m_0, m_1, m_2, m_3\}$ eine Basis von $P_3(\mathbb{R})$. Verwenden wir auf $P_3(\mathbb{R})$ nun das innere Produkt (vgl. $(5.1b)$)

$$\langle p, q \rangle := \int_{-1}^{1} p(t)q(t)dt,$$

so ist $\mathcal{M}_3$ wegen $\langle m_k, m_l\rangle = \dfrac{1-(-1)^{k+l+1}}{1+k+l}$ jedoch keine Orthogonalbasis. Um eine solche zu bestimmen, wenden wir das Gram-Schmidt-Verfahren an. Zunächst sei dazu $p_0 = m_0$ und wegen

$$\langle p_0, m_1\rangle = 0,\ \langle p_0, m_2\rangle = \frac{2}{3},\ \langle p_0, m_3\rangle = 0$$

erhalten wir $p_1 = m_1 - \dfrac{\langle p_0, m_1\rangle}{\langle p_0, p_0\rangle} p_0 = m_1$, also $p_1(t) = t$. Mittels der Beziehungen

$$\langle p_1, m_2\rangle = 0,\ \langle p_1, m_3\rangle = \frac{2}{5}$$

ist weiter $p_2 = m_2 - \dfrac{\langle p_0, m_2\rangle}{\langle p_0, p_0\rangle} p_0 - \dfrac{\langle p_1, m_2\rangle}{\langle p_1, p_1\rangle} p_1 = m_2 - \dfrac{\langle p_0, m_2\rangle}{\langle p_0, p_0\rangle} p_0$, also $p_2(t) = t^2 - \frac{1}{3}$. Ebenso erhält man $p_3(t) = t^3 - \frac{3}{5}t$ und hieraus ergibt sich die Orthonormalbasis $\{q_0, q_1, q_2, q_3\}$ des $P_3(\mathbb{R})$ mit

$$q_0(t) \equiv \sqrt{\frac{1}{2}},\ q_1(t) = \sqrt{\frac{3}{2}}t,\ q_2(t) = \frac{3}{2}\sqrt{\frac{5}{2}}(t^2 - \frac{1}{3}),\ q_3(t) = \frac{5}{2}\sqrt{\frac{7}{2}}(t^3 - \frac{3}{5}t)$$

Fordert man nicht die Normierung $\langle q_k, q_k\rangle = 1$, sondern $q_k(1) = 1$, so ergibt sich durch Multiplikation der p_k mit einem geeigneten Faktor, dass

$$q_0(t) \equiv 1,\ q_1(t) = t,\ q_2(t) = \frac{1}{2}(3t^2 - 1),\ q_3(t) = \frac{1}{2}(5t^3 - 3t);$$

dies sind die Legendre-Polynome aus Beispiel 5.1.12(2)

5.2 Adjungierte Abbildungen

Sei X ein linearer Raum über $\mathbb{K}$ mit innerem Produkt $\langle\cdot,\cdot\rangle$. Eine lineare Abbildung $S\colon X \to \mathbb{K}$ heißt Funktional. Mit gegebenem $y \in X$ definieren wir das Funktional $y'\colon X \to \mathbb{K}$ durch

$$(5.2a)\ y'(x)\colon = \langle x, y\rangle \text{ für alle } x \in X$$

Aufgrund von Definition 5.1.1(i) gilt nämlich $y'(\alpha_1 x_1 + \alpha_2 x_2) = \langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle = \alpha_1 y'(x_1) + \alpha_2 y'(x_2)$ für alle $\alpha_1, \alpha_2 \in \mathbb{K},\ x_1, x_2 \in X$, also $y' \in L(X, \mathbb{K})$

5.2.1 Satz (Riesz'scher Darstellungssatz)

Ist $\dim X < \infty$, so gibt es zu jedem Funktional $S \in L(X, \mathbb{K})$ ein eindeutiges $y \in X$ mit $Sx = \langle x, y\rangle$ für alle $x \in X$, d.h. $S = y'$.
Beweis: Nach Satz 5.1.15 gibt es eine Orthonormalbasis $\{x_1, \cdots, x_n\}$ von X.

(I) <u>Existenz</u>: Zu beliebig gegebenem $S \in L(X, \mathbb{K})$ sei

$$(5.2b)\ y\colon = \sum_{i=1}^{n} \overline{Sx_i} x_i \in X$$

und es gilt

$$\langle x_j, y\rangle \overset{(5.2b)}{=} \left\langle x_j, \sum_{i=1}^{n} \overline{Sx_i} x_i \right\rangle = \sum_{i=1}^{n} \overline{\overline{Sx_i}}\, \overbrace{\langle x_j, x_i\rangle}^{\delta_{ij}} = Sx_j \text{ für } 1 \leq j \leq n$$

Damit stimmen S und y' auf einer Basis von X überein, weshalb Satz 3.1.14(a) sofort $S = y'$ garantiert.

(II) <u>Eindeutigkeit</u>: Um nachzuweisen, dass obiges y aus $(5.2b)$ eindeutig bestimmt ist, erfülle auch $z \in X$ die Beziehung $Sx = \langle x, y\rangle = \langle x, z\rangle$ für alle $x \in X$. Dies impliziert $\langle x, y - z\rangle = 0$ für alle $x \in X$ und damit $y - z = 0$

Nun seien $T \in L(X)$, $y \in X$. Wir definieren $S_y\colon X \to \mathbb{K}$

$$S_y x\colon = \langle Tx, y\rangle$$

und es folgt sofort $S_y \in L(X, \mathbb{K})$. Bei festem T gibt es nach Satz 5.2.1 zu jedem $y \in X$ ein eindeutiges $y_0 \in X$ mit $S_y x = \langle x, y_0\rangle$ für alle $x \in X$. Wir bezeichnen diese Zuordnung $y \mapsto y_0$ mit T' und die entsprechende Abbildung $T'\colon X \to X$ ist festgelegt durch

$$\langle Tx, y\rangle = \langle x, T'y\rangle \text{ für alle } x, y \in X$$

Gegeben $T = L(X)$, $y \in X$. Wir definieren:

$$S_y x := \langle Tx, y\rangle$$

und mittels $(3.1a)$, sowie Definition 5.1.1(i) folgt sofort $S_y \in L(X, \mathbb{K})$. Bei festen $T \in L(X)$ gibt es nach Satz 5.2.1 dann ein eindeutig bestimmtes $y_0 \in X$mit $S_y x = \langle x, y_0\rangle$. Wir bezeichnen diese Zuordnung $y \mapsto y_0$ mit T' und die Abbildung $T'\colon X \to X$ ist festgelegt durch $\langle Tx, y\rangle = \langle x, T'y\rangle$ für alle $x, y \in X$. Aufgrund von Definition 5.1.1 erhalten wir mit beliebigen $x_1, x_2 \in X$,$\alpha_1, \alpha_2 \in \mathbb{K}$, dass

$$\langle x, T'(\alpha_1 x_1 + \alpha_2 x_2)\rangle = \langle Tx, \alpha_1 x_1 + \alpha_2 x_2\rangle = \overline{\alpha_1}\langle Tx, x_1\rangle + \overline{\alpha_2}\langle Tx, x_2\rangle = \overline{\alpha_1}\langle x, T'x_1\rangle + \overline{\alpha_2}\langle x, T'x_2\rangle$$

$$\langle x, \alpha_1 T'x_1 + \alpha_2 T'x_2\rangle$$

Damit ist $T' \in L(X)$.

5.2.2 Definition (adjungierte Abbildung)

Die zu $T \in L(X)$ adjungierte Abbildung $T' \in L(X)$ ist definiert durch

$$(5.2c)\ \langle Tx, y\rangle = \langle x, T'y\rangle \text{ für alle } x, y \in X$$

Man beachte, dass T' von konkret auf X verwendete inneren Produkt abhängt.

5.2.3 Bemerkung

Direkt aus (5.2c) folgt

(1) Die Abbildung $\cdot'\colon L(X) \to L(X)$ ist semilinear, d.h.

$$(\alpha_1 T_1 + \alpha_2 T_2)' = \overline{\alpha_1} T_1' + \overline{\alpha_2} T_2' \text{ für } \alpha_1, \alpha_2 \in \mathbb{K}, T_1, T_2 \in L(X)$$

und erfüllt $T'' = T$, sowie $(ST)' = T'S'$ für $S, T \in L(X)$

(2) Für $T \in GL(X)$ ist auch $T' \in GL(X)$ mit

$$(T')^{-1} = (T^{-1})'$$

5.2.4 Beispiel (Transponierte)

Es sei $X = \mathbb{R}^n$ (Euklidisch) ausgestattet mit der Standardbasis $\mathcal{E}_n = \{e_1, \cdots, e_n\}$ und $A \in \mathbb{R}^{n\times n}$ die darstellende Matrix von $T \in L(X)$, d.h. $Tx = Ax$ für alle $X \in \mathbb{R}^n$. Bezeichnet dann $A' \in \mathbb{R}^{n\times n}$ die darstellende Matrix von T', so ist

$$a'_{ij} = \left\langle e_i, \begin{pmatrix} a'_{1j} \\ \vdots \\ a'_{nj} \end{pmatrix} \right\rangle = \langle e_i, T'e_j\rangle \overset{(5.2c)}{=} \langle Te_i, e_j\rangle = \left\langle \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}, e_j \right\rangle = a_{ij} \text{ für alle } 1 \le i, j \le n.$$

Folglich ist die adjungierte Abbildung T' in Euklidischen Räumen gerade durch die Transponierte $A' = A^T$ gegeben (vergleiche Beispiel 3.2.3).

5.2.5 Beispiel (Hermite'sche)

Im unitärer Raum $\mathbb{C}^n$ mit seiner Standardbasis $\mathcal{E}_n$ ergibt sich entsprechend:

$$\overline{a'_{ij}} = \left\langle e_i, \begin{pmatrix} a'_{1j} \\ \vdots \\ a'_{nj} \end{pmatrix} e_i, T'e_j \right\rangle \overset{(5.2c)}{=} \langle Te_i, e_j\rangle = \left\langle \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}, e_j \right\rangle = a_{ji} \text{ für alle } 1 \le i, j \le n$$

und folglich $a'_{ij} = \overline{a_{ji}}$. Für die adjungierte Abbildung in unitären Räumen ist also $A' = A^*$ mit der sogenannten Hermite'schen

$$A^* = \overline{(A^T)} = (\overline{a_{ji}})_{1\le i,j\le n}$$

5.2.6 Satz

Ist $\dim X < \infty$ und $T \in L(X)$, so gilt

$$N(T') = R(T)^\perp \quad \mathrm{rk}T = \mathrm{rk}T'$$

<u>Beweis</u>: Wir erhalten

$$N(T') = \{x \in X : T'x = 0\} = \{x \in X : \langle y, T'x\rangle = 0 \text{ für alle } y \in X\}$$

$$\overset{(5.2c)}{=} \{x \in X : \langle Ty, x\rangle = 0 \text{ für alle } y \in X\} = R(T)^\perp$$

Mit Satz 5.1.10(c), der eben gezeigten Aussage und Satz 3.1.12

$$\mathrm{rk}T = \dim R(T) =^1 \dim X - \dim R(T)^\perp = \dim X - \dim N(T') = \dim R(T') = \mathrm{rk}T'$$

5.3 Sebstadjungierte Abbildungen

Wieder sei X ein linearer Raum mit inneren Produkt $\langle \cdot, \cdots \rangle$

5.3.1 Definition (Selbstadjungiert)

Eine Abbildung $T \in L(X)$ heißt <u>selbstadjungiert</u>, falls gilt $T = T'$, d.h.

$$(5.3a)\langle Tx, y\rangle = \langle x, Ty\rangle \text{ für alle } x, y \in X$$

5.3.2 Beispiel (Euklidischer Raum)

Auf dem Euklidischen Raum $\mathbb{R}^n$ ist eine Abbildung $T \in L(X)$ genau dann selbstadjungiert, falls ihre darstellende Matrix $T_{\mathcal{E}_n} = A$ <u>symmetrisch</u> ist, d.h. $A = A^T$

$$\text{z.B. } \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ nicht.}$$

5.3.3 Beispiel (unitärer Raum)

In einem unitären Raum $\mathbb{C}^n$ ist $T \in L(\mathbb{C}^n)$ dann und nur dann selbstadjungiert, wenn $A = T_{\mathcal{E}_n}$ Hermite'sch ist (vergleiche Beispiel 5.2.5), d.h. $A = A^*$.

[1]gilt wegen: $X = Y \oplus Y^\perp$
$\dim X = \dim Y + \dim Y^\perp$

5.3.5 Satz

Es sei $\dim X < \infty$. Ist $T \in L(X)$ selbstadjungiert, so gilt:

(a) $\sigma(T) \subseteq \mathbb{R}$

(b) Es gibt eine Orthonormalbasis von X aus Eigenvektoren von T

5.3.6 Bemerkung

Auf Matrizen bezogen besagt dies gerade, dass symmetrische und Hermite'sche Matrizen reelles Spektrum haben und der Raum $\mathbb{K}^n$ eine Orthonormalbasis aus Eigenvektoren besitzt. Beweis:

(a) Es sei $\lambda \in \mathbb{K}$ ein Eigenwert von T mit zugehörigen Eigenvektor $x \in X$. Da T selbstadjungiert ist, gilt

$$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle \overset{(4.2a)}{=} \langle Tx, x\rangle = \langle x, Tx\rangle \overset{(4.2a)}{=} \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle$$

und wegen $x \neq 0$ folgt $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$.

(b) Induktion über $n = \dim X$. Für $n = 1$ ist nichts nachzuweisen. Für $n > 1$ gilt den Fundamentalsatz der Algebra

$$\chi_T(t) = \prod_{i=1}^{n}(\lambda_i - t) \text{ mit } \lambda_i \in \mathbb{C}$$

und wegen Satz 4.3.3 ist jedes λ_i Eigenwert von T; Aussage (a) garantiert sogar $\lambda_i \in \mathbb{R}$. Es sei nun x_1 ein normierter Eigenvektor zu λ_1 und $X_1 = (\operatorname{span}\{x_1\})^\perp$. Mittels Satz 5.1.10(c) ist dann $\dim X_1 = n - 1$. Für jedes $x \in X_1$ ist dann

$$\langle x_1, Tx\rangle \overset{(5.3a)}{=} \langle Tx_1, x\rangle \overset{(4.2a)}{=} \lambda_1 \overbrace{\langle x_1, x\rangle}^{0} = 0$$

d.h. $Tx \in X_1$. Wir haben damit gezeigt, dass X_1 invariant unter T ist. Damit induziert $T \in L(X)$ eine selbstadjungierte Abbildung $T_1 : = T|_{X_1} \in L(X_1)$. Nach Induktionsvoraussetzung gilt es eine Orthonormalbasis $\mathcal{X}_1 : = \{x_2, \cdots, x_n\}$ von X_1 aus nomierten Eigenvektoren x_i von T. Da jeder Eigenvektor von T_1 auch Eigenvektor von T ist, muss $\{x_1, x_2, \cdots, x_n\}$ eine Orthonormalbasis von X sein.

5.3.7 Korollar

Jede selbstadjungierte Abbildung ist diagonalisierbar. Beweis: Wähle Basis aus Eigenvektoren.