

Kapitel 5 Lineare Algebra

§ 1 Algebraische Strukturen

Inhalt:

Gruppen, Restklassen, Permutationen, Körper, Vektorräume, lineare Abbildungen, Matrizen und Matrizenprodukt.

Eine *binäre Operation* auf einer Menge M ist eine Abbildung $m : M \times M \rightarrow M$. Von Fall zu Fall werden ganz unterschiedliche Bezeichnungen für $m(x, y)$ gewählt.

Beispiele.

1. M eine Menge von Zahlen, $m(x, y) = x + y$ oder $= x \cdot y$.
2. M die Menge aller endlichen Gruppen von Zeichen (Buchstaben, Ziffern, Satzzeichen, Leerräume), zwei Gruppen werden einfach hintereinander gestellt.
3. M eine Menge von Abbildungen $f : X \rightarrow X$, für eine feste Menge X . Als binäre Operation kann man die Verknüpfung $f \circ g$ nehmen.
4. Das Skalarprodukt auf dem \mathbb{R}^n ist für $n \geq 2$ **keine** binäre Operation, denn das Produkt zweier Vektoren ist ein Skalar, kein Vektor.

Definition:

Eine Menge G mit einer binären Operation $*$ heißt eine *Gruppe*, falls folgende Eigenschaften erfüllt sind:

1. $(a * b) * c = a * (b * c)$ für alle $a, b, c \in G$.
2. $\exists e \in G$, so daß $a * e = e * a = a$ für alle $a \in G$.
3. Zu jedem $a \in G$ gibt es ein Element $a^{-1} \in G$ mit

$$a * a^{-1} = a^{-1} * a = e.$$

Man nennt dann $*$ die *Gruppenoperation*, e das *neutrale Element* und a^{-1} das *Inverse* zu a .

Eine Gruppenoperation muß also immer assoziativ sein, sie braucht aber nicht kommutativ zu sein. Ist sie kommutativ, so spricht man von einer *kommutativen* oder *abelschen Gruppe*.

Beispiele.

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ sind abelsche Gruppen mit dem neutralen Element 0, das Inverse ist in diesem Falle das Negative. Auch $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$ und $(\mathbb{C} \setminus \{0\}, \cdot)$ sind abelsche Gruppen, mit dem neutralen Element 1. Das Inverse ist das Reziproke. Aber $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist keine Gruppe mehr, weil das Inverse fehlt.
2. Sei X eine beliebig nicht-leere Menge und G die Menge aller **bijektiven** Abbildungen $f : X \rightarrow X$. Mit der Verknüpfung von Abbildungen wird G zu einer Gruppe. Das neutrale Element ist die identische Abbildung id_X , das Inverse zu f ist die Umkehrabbildung f^{-1} .
3. Die Vektoren des \mathbb{R}^3 bilden mit der Vektoraddition eine abelsche Gruppe. Neutrales Element ist der Nullvektor, das Inverse zu \mathbf{x} ist $-\mathbf{x}$.

Dagegen bildet der \mathbb{R}^3 mit dem Vektorprodukt $\mathbf{v} \times \mathbf{w}$ keine Gruppe, denn schon das Assoziativgesetz ist nicht erfüllt.

Im Folgenden sollen noch zwei etwas umfangreichere Beispiele betrachtet werden.

Weil ganze Zahlen i.a. kein ganzzahliges Inverses besitzen, gibt es die Teilbarkeits-theorie. Bekanntlich ist x ein *Teiler* von y , wenn es eine ganze Zahl q gibt, so daß $y = x \cdot q$ ist. Man schreibt dann auch: $x | y$. Jede positive ganze Zahl a hat die *trivialen Teiler* ± 1 und $\pm a$. Hat sie nur diese trivialen Teiler und ist sie > 1 , so nennt man sie eine *Primzahl*.

Definition:

Es sei $n > 1$ eine feste natürliche Zahl. Zwei Zahlen $a, b \in \mathbb{Z}$ heißen *kongruent modulo n* (in Zeichen: $a \equiv b \pmod{n}$), falls gilt:

$$n | (a - b).$$

So ist z.B. $23 \equiv 1 \pmod{11}$ und $23 \equiv -2 \pmod{25}$.

Die Kongruenz modulo n besitzt folgende Eigenschaften:

1. $a \equiv a \pmod{n}$;
2. wenn $a \equiv b \pmod{n}$, dann auch $b \equiv a \pmod{n}$;
3. wenn $a \equiv b \pmod{n}$ und $b \equiv c \pmod{n}$, dann auch $a \equiv c \pmod{n}$.

Ist $a \geq 0$, so gibt es ein q und genau ein r mit $0 \leq r \leq n - 1$, so daß gilt:

$$a = q \cdot n + r.$$

Dann ist r der *Rest*, der bei der (ganzzahligen) Division von a durch n bleibt. Offensichtlich ist $a \equiv r \pmod{n}$.

Insgesamt gibt es unendlich viele Zahlen x , die modulo n zu a kongruent sind, nämlich alle Zahlen $x = a + q \cdot n$, $q \in \mathbb{Z}$. Man nennt die Menge dieser Zahlen die *Kongruenzklasse* von a modulo n . Das kleinste Element ≥ 0 in dieser Klasse ist der Rest r . Weil die Kongruenzklasse von a schon durch den Rest r bestimmt ist, nennt man sie auch die *Restklasse* von a modulo n und bezeichnet sie mit $[r]$ oder $[r]_n$. Man darf sie aber auch mit $[a]$ oder $[a]_n$ bezeichnen, wenn man folgende Verabredung trifft:

$$[a]_n = [b]_n : \iff a \equiv b \pmod{n}.$$

Halten wir den „Modul“ n fest, so werden durch die n möglichen Reste $0, 1, \dots, n - 1$ alle Restklassen modulo n bestimmt, mehr kann es nicht geben. Und da es zu jeder ganzen Zahl einen eindeutig bestimmten Rest bei der Division durch n gibt, gehört jede ganze Zahl zu genau einer Restklasse.

Beispiel.

Sei $n = 3$. Es gibt drei mögliche Reste, nämlich $r = 0$, $r = 1$ und $r = 2$. Also gibt es genau drei Restklassen:

$$\begin{aligned} [0] &= \{0, \pm 3, \pm 6, \pm 9, \dots\}, \\ [1] &= \{1, -2, 4, -5, 7, -8, \dots\}, \\ [2] &= \{2, -1, 5, -4, 8, -7, \dots\}. \end{aligned}$$

Zusammen ergibt das die Menge aller ganzen Zahlen.

Es ist dann z.B. $[1] = [-5] = [7]$.

Man kann sehr leicht zeigen: Ist $a \equiv b \pmod{n}$ und $c \equiv d \pmod{n}$, so ist auch

$$a + c \equiv b + d \pmod{n} \quad \text{und} \quad a \cdot c \equiv b \cdot d \pmod{n}.$$

Das bedeutet, daß man Kongruenzklassen addieren und multiplizieren kann:

$$[a] + [b] := [a + b] \quad \text{und} \quad [a] \cdot [b] := [a \cdot b].$$

Dabei passieren aber merkwürdige Dinge.

Ist etwa $n = 7$, so ist $[2] + [3] = [5]$, wie man es erwarten würde. Aber es ist auch $[5] + [3] = [1]$, denn $5 + 3 = 8$ ist kongruent 1 modulo 7. Natürlich wäre auch die Gleichung $[5] + [3] = [8]$ richtig gewesen, aber wir wollen möglichst nur die Bezeichnungen $[0], [1], \dots, [6]$ für die Restklassen modulo 7 verwenden.

Ein Vorteil der Kongruenzrechnung liegt darin, daß man Rechnungen mit großen Zahlen auf Rechnungen mit kleinen Zahlen reduzieren kann. Man verliert zwar etwas Information, behält aber oft noch genügend viel übrig. Ein Beispiel sind die bekannten Teilbarkeitsregeln:

Eine natürliche Zahl hat im Dezimalsystem eine Darstellung

$$n = a_0 + 10a_1 + 100a_2 + \cdots + 10^k a_k, \text{ mit } 0 \leq a_i \leq 9.$$

Weil $10 \equiv 1 \pmod{9}$ ist, folgt:

$$n \equiv a_0 + a_1 + \cdots + a_k \pmod{9}.$$

Daraus folgt: Ist die rechte Seite der Kongruenz, also die Quersumme von n , durch 9 teilbar, so ist auch die linke Seite, also n selbst, durch 9 teilbar. In der „Zahlentheorie“ lernt man noch viel mehr solcher Tricks.

Was hat das alles mit Gruppen zu tun?

Sei $n > 1$ und $G = \{[0], [1], \dots, [n-1]\}$ die Menge aller Restklassen modulo n . Wir haben oben schon gelernt, wie man Restklassen addiert. Diese Addition ist assoziativ und kommutativ, das rechnet man leicht nach. Die „Nullklasse“ $[0]$ ist das neutrale Element, denn es ist ja $[a] + [0] = [a + 0] = [a]$. Und wie steht es mit dem Inversen, das hier ja ein Negatives sein sollte? Ist $1 \leq r \leq n-1$, so ist auch $1 \leq n-r \leq n-1$. Weil $r + (n-r) = n \equiv 0 \pmod{n}$ ist, ist $-[r] = [n-r]$. Also ist G eine abelsche Gruppe, die wir mit \mathbb{Z}_n bezeichnen wollen.

Zum Beispiel ist $\mathbb{Z}_2 = \{[0], [1]\}$, mit

$$\begin{aligned} [0] + [0] &= [0], \\ [0] + [1] &= [1], \\ [1] + [0] &= [1] \\ \text{und } [1] + [1] &= [0]. \end{aligned}$$

Wir kommen jetzt zum nächsten Beispiel.

Unter einer *Permutation* versteht man eine bijektive Abbildung von $\{1, \dots, n\}$ auf sich. Wir schreiben eine Permutation σ in der Form

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

Die Werte $\sigma(1), \dots, \sigma(n)$ bestimmen – in dieser Anordnung – die Abbildung σ . Es gibt bekanntlich genau $n!$ Permutationen der Zahlen $1, \dots, n$. Wir bezeichnen die Menge aller dieser Permutationen mit S_n . So ist z.B.

$$S_3 = \left\{ \begin{pmatrix} 123 \\ 123 \end{pmatrix}, \begin{pmatrix} 123 \\ 132 \end{pmatrix}, \begin{pmatrix} 123 \\ 213 \end{pmatrix}, \begin{pmatrix} 123 \\ 231 \end{pmatrix}, \begin{pmatrix} 123 \\ 312 \end{pmatrix}, \begin{pmatrix} 123 \\ 321 \end{pmatrix} \right\}.$$

Wird n groß, so wird der Umgang mit den Permutationen aus S_n kompliziert. Eine gewisse Hilfe ist dann die „Zykel-Schreibweise“. Wir demonstrieren das an einem Beispiel:

$$\text{Sei } \sigma := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 4 & 5 & 1 & 2 & 6 \end{pmatrix} \in S_7.$$

Beginnen wir mit der 1. Die 1 wird auf 3 abgebildet, die 3 auf die 4, die 4 auf die 5 und die 5 wieder auf die 1. Das ergibt einen abgeschlossenen *Zykel*, den man mit $(1, 3, 4, 5)$ bezeichnet. Damit ist aber die Abbildung σ noch nicht abgehandelt. Die nächste Zahl nach der 1, die wir noch nicht berücksichtigt haben, ist die 2. Mit ihr beginnen wir das Spiel erneut: Die 2 wird auf die 7 abgebildet, die 7 auf die 6 und die 6 wieder auf die 2. Das ergibt den Zykel $(2, 7, 6)$. Da nun alle Zuordnungen von σ berücksichtigt wurden, schreiben wir:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 4 & 5 & 1 & 2 & 6 \end{pmatrix} = (1, 3, 4, 5)(2, 7, 6).$$

Wird eine Zahl i auf sich abgebildet, so ergibt das einen „Einer-Zykel“ (i). Solche Zykel schreibt man normalerweise gar nicht hin. Wenn allerdings σ die Identität ist, so besteht σ nur aus Einer-Zykeln, und man muß wenigstens einen davon hinschreiben:

$$\text{id}_{\{1, \dots, n\}} = (1).$$

Sind $\sigma := \begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix}$ und $\tau := \begin{pmatrix} 1 & \dots & n \\ \tau(1) & \dots & \tau(n) \end{pmatrix}$ zwei Permutationen, so kann man sie zu einer neuen Permutation $\sigma \circ \tau$ miteinander verknüpfen, mit

$$\sigma \circ \tau = \begin{pmatrix} 1 & \dots & n \\ \sigma(\tau(1)) & \dots & \sigma(\tau(n)) \end{pmatrix}.$$

So ist z.B.

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Man beachte, daß die Permutationen **von rechts nach links** abzarbeiten sind! Die rechte (innere) Permutation bildet z.B. die 1 auf die 3 ab, und die linke (äußere) Permutation bildet dann die 3 auf die 2 ab. Im Ergebnis auf der anderen Seite der Gleichung muß deshalb unter der 1 die 2 stehen.

Die Menge S_n aller Permutationen von $\{1, \dots, n\}$ bildet offensichtlich eine Gruppe. Normalerweise ist die Verknüpfung von Permutationen nicht kommutativ. Wenn wir allerdings die Menge $\{1, \dots, n\}$ in zwei disjunkte Teilmengen M und N aufteilen können, so daß die Permutation σ nur die Zahlen aus M verändert und τ nur die Zahlen aus N , dann spielt die Reihenfolge keine Rolle, es ist $\sigma \circ \tau = \tau \circ \sigma$. Das hat Konsequenzen für die Zykel-Schreibweise. Die Aufteilung einer Permutation in mehrere (disjunkte) Zykel bedeutet, daß sich die gegebene Permutation aus mehreren einfacheren Permutationen verknüpfen läßt, und die Reihenfolge spielt dabei keine Rolle. Es ist also z.B.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 4 & 5 & 1 & 2 & 6 \end{pmatrix} = (1, 3, 4, 5) \circ (2, 7, 6) = (2, 7, 6) \circ (1, 3, 4, 5).$$

Eine Permutation heißt *Transposition* oder *Vertauschung*, wenn nur zwei Zahlen miteinander vertauscht werden und alle anderen fest bleiben, wenn sie also nur aus einem „Zweier-Zykel“ besteht, z.B.

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

Satz

Jede Permutation läßt sich als endliche Verknüpfung von Transpositionen schreiben.

BEWEIS: Da wir jede Permutation als Folge von Zyklen schreiben können, reicht es zu zeigen, daß sich jeder beliebige Zykel aus endlich vielen Zweier-Zykeln zusammensetzen läßt.

Betrachten wir also einen beliebigen Zykel (i_1, i_2, \dots, i_k) in einer Permutation σ . Dann ist

$$\sigma(i_\nu) = i_{\nu+1} \text{ für } \nu = 1, 2, \dots, k-1, \text{ und } \sigma(i_k) = i_1.$$

Der Zweier-Zykel (i_1, i_2) vertauscht nur i_1 mit i_2 und läßt alles andere fest. Führt man anschließend den Zweier-Zykel (i_1, i_3) aus, so wird insgesamt i_1 auf i_2 , i_2 auf i_3 und i_3 auf i_1 abgebildet. So kann man fortfahren bis zu der Kombination

$$(i_1, i_k) \circ \dots \circ (i_1, i_3) \circ (i_1, i_2),$$

die i_1 auf i_2 , i_2 auf i_3 usw. und schließlich i_{k-1} auf i_k und i_k auf i_1 abbildet. Also ist

$$(i_1, i_2, \dots, i_k) = (i_1, i_k) \circ \dots \circ (i_1, i_3) \circ (i_1, i_2).$$

Man beachte aber, daß es hier auf die Reihenfolge ankommt, weil die einzelnen Zweier-Zykel nicht paarweise disjunkt sind! ■

Das oben angegebene Verfahren zur Auflösung einer Permutation in Transpositionen liefert ein eindeutiges Ergebnis. Leider sind auch andere Verfahren denkbar, und die Zerlegung einer Permutation in Transpositionen ist i.a. nicht eindeutig bestimmt.

Definition:

Ist $\sigma \in S_n$, so nennt man die Zahl

$$\text{sign}(\sigma) := \prod_{i < j} \frac{\sigma(i) - \sigma(j)}{i - j}$$

das *Signum* von σ .

Das ist eine seltsame Definition. Um sie besser zu verstehen, muß man sich überlegen, daß oberhalb und unterhalb des Bruchstriches die gleichen numerischen Werte stehen, nur in anderer Reihenfolge (was keine Rolle spielt) und mit anderen Vorzeichen: Unten steht vor jeder Zahl ein Minuszeichen, oben steht (für $i < j$) das Vorzeichen

$$s_{i,j}(\sigma) := \begin{cases} +1 & \text{falls } \sigma(i) < \sigma(j) \\ -1 & \text{falls } \sigma(i) > \sigma(j). \end{cases}$$

Also ist $\text{sign}(\sigma) = \prod_{i < j} s_{i,j}(\sigma) = \pm 1$.

Ist z.B. τ eine Transposition mit $\tau(i) = i + 1$, $\tau(i + 1) = i$ und $\tau(k) = k$ für alle anderen k , so ist $s_{i,i+1}(\tau) = -1$ und $s_{k,l}(\tau) = +1$ in allen anderen Fällen, also insgesamt $\text{sign}(\tau) = -1$.

Satz (Multiplikativität des Signums)

Für $\sigma, \tau \in S_n$ ist $\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \cdot \text{sign}(\tau)$.

BEWEIS: Ist $k \neq l$, so ist $\frac{\sigma(k) - \sigma(l)}{k - l} = \frac{\sigma(l) - \sigma(k)}{l - k}$. Da bei der Berechnung des Signums ein Produkt über alle Paare (k, l) mit $k \neq l$ gebildet wird, folgt:

$$\begin{aligned} \text{sign}(\sigma \circ \tau) &= \prod_{i < j} \frac{\sigma(\tau(i)) - \sigma(\tau(j))}{i - j} \\ &= \prod_{i < j} \frac{\sigma(\tau(i)) - \sigma(\tau(j))}{\tau(i) - \tau(j)} \cdot \prod_{i < j} \frac{\tau(i) - \tau(j)}{i - j} \\ &= \text{sign}(\sigma) \cdot \text{sign}(\tau). \end{aligned}$$

■

Da jede Permutation Produkt von Transpositionen ist, können wir das Signum nun sehr einfach berechnen, und insbesondere folgt:

Satz

Auch wenn eine Permutation auf verschiedene Weisen durch eine Folge von Transpositionen erzeugt wird, so ist die Anzahl der dabei benötigten Vertauschungen modulo 2 eindeutig bestimmt.

Eine Permutation $\sigma \in S_n$ heißt *gerade*, wenn sie sich aus einer geraden Anzahl von Vertauschungen zusammensetzt. Andernfalls heißt sie *ungerade*. Ist σ gerade, so ist $\text{sign}(\sigma) = +1$, andernfalls $= -1$. So ist z.B.

$$\begin{aligned} \text{sign} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{pmatrix} &= \text{sign}[(3, 4)] = -1, \\ \text{sign} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 2 & 5 \end{pmatrix} &= \text{sign}[(1, 3) \circ (2, 4)] = +1 \\ \text{und } \text{sgn} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix} &= \text{sign}[(1, 3, 4) \circ (2, 5)] \\ &= \text{sign}[(1, 4) \circ (1, 3) \circ (2, 5)] = -1. \end{aligned}$$

Das Signum einer Permutation wird uns später bei den Determinanten wieder begegnen.

Definition:

Ein *Körper* ist eine Menge K mit zwei binären Operationen $+$ und \cdot , so daß gilt:

1. $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.
2. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1.
3. Es gilt das Distributivgesetz:

$$a \cdot (x + y) = a \cdot x + a \cdot y.$$

Wir kennen schon die Körper \mathbb{Q} , \mathbb{R} und \mathbb{C} . Weitere Beispiele sind die Gruppen \mathbb{Z}_2 und \mathbb{Z}_3 .

Allerdings sind nicht alle Restklassengruppen Körper. Zum Beispiel gilt in \mathbb{Z}_6 :

$$[2] \cdot [3] = [0].$$

Würde es zu $[2]$ ein Inverses geben, also eine Restklasse $[r]$ mit $0 \leq r \leq 5$ und $[2] \cdot [r] = [1]$, so wäre

$$[3] = [1] \cdot [3] = ([2] \cdot [r]) \cdot [3] = ([r] \cdot [2]) \cdot [3] = [r] \cdot ([2] \cdot [3]) = [r] \cdot [0] = [0].$$

Das kann nicht sein.

Ist p eine Primzahl und $1 \leq r \leq p - 1$, so ist $\text{ggT}(p, r) = 1$. Mit Hilfe des Euklidischen Algorithmus kann man dann zeigen, daß es ganze (eventuell negative) Zahlen a, b gibt, so daß $a \cdot r + b \cdot p = \text{ggT}(r, p) = 1$ ist. Das bedeutet, daß $a \cdot r \equiv 1 \pmod{p}$ ist, also $[a] \cdot [r] = [1]$. Jedes Element $[r] \neq [0]$ in \mathbb{Z}_p besitzt ein Inverses. Damit ist \mathbb{Z}_p ein Körper.

Definition:

Ein *Ring* ist eine Menge R mit zwei binären Operationen $+$ und \cdot , so daß gilt:

1. $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.
2. Die Multiplikation ist assoziativ.
3. Es gelten die Distributivgesetze:

$$a \cdot (x + y) = a \cdot x + a \cdot y$$

und $(a + b) \cdot x = a \cdot x + b \cdot x.$

Der Ring heißt *kommutativ*, falls $a \cdot b = b \cdot a$ für alle a, b gilt. Gibt es ein neutrales Element 1 für die Multiplikation, so spricht man von einem *Ring mit Eins*.

Ein Körper ist also ein kommutativer Ring mit Eins, in dem jedes Element $x \neq 0$ ein multiplikatives Inverses besitzt.

\mathbb{Z} und \mathbb{Z}_6 sind Beispiele von Ringen, die keine Körper sind, aber auch die Menge aller stetigen Funktionen auf einem Intervall.

Ringstrukturen werden für uns künftig keine große Rolle spielen, aber gelegentlich kommen sie vor.

Definition:

Sei K ein beliebiger fester Körper. Ein K -*Vektorraum* besteht aus einer abelschen Gruppe $(V, +)$ und einer Multiplikation, die jedem Paar $(\lambda, x) \in K \times V$ eindeutig ein Element $\lambda x \in V$ zuordnet, so daß für $\alpha, \beta \in K$ und $x, y \in V$ gilt:

$$\begin{aligned} (\alpha\beta)x &= \alpha(\beta x), \\ \alpha(x + y) &= \alpha x + \alpha y, \\ (\alpha + \beta)x &= \alpha x + \beta x, \\ 1x &= x. \end{aligned}$$

Die Rechengesetze in einem beliebigen K -Vektorraum sind die gleichen wie in einem \mathbb{R} -Vektorraum. Dementsprechend übertragen sich auch die meisten Begriffe und Aussagen. Unterräume, Linearkombinationen, Erzeugendensysteme, Basen werden in der bekannten Weise definiert, und es gelten die Sätze über Existenz und Konstruktion von Basen (z.B. der Austauschsatz) wie im Falle $K = \mathbb{R}$. Man kann zeigen, daß jeder Vektorraum eine Basis besitzt, aber wir interessieren uns hauptsächlich für endlich-dimensionale Räume.

Wichtigstes Beispiel für einen K -Vektorraum ist natürlich der Raum

$$K^n = \{\mathbf{x} = (x_1, \dots, x_n) : x_i \in K \text{ für } i = 1, \dots, n\}.$$

Wir schreiben die Elemente des K^n also als „Zeilenvektoren“.

Mit $M_{n,m}(K)$ bezeichnen wir die Menge der Matrizen

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Auch $M_{n,m}(K)$ ist ein K -Vektorraum. Speziell ist $M_{1,m}(K) = K^m$ und $M_{n,1}(K)$ der Raum der n -reihigen Spaltenvektoren.

Ist $A \in M_{n,m}(K)$, so ist die *transponierte Matrix* $A^t \in M_{m,n}(K)$ definiert durch

$$A^t = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{nm} \end{pmatrix}.$$

Beim Transponieren werden also die Zeilen der alten Matrix zu den Spalten der neuen Matrix, und umgekehrt. Ist $\mathbf{x} = (x_1, \dots, x_n) \in K^n = M_{1,n}(K)$ ein Zeilenvektor, so ist

$$\vec{x} := \mathbf{x}^t = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(K)$$

der entsprechende Spaltenvektor.

Definition:

Es seien V und W zwei K -Vektorräume. Eine Abbildung $f : V \rightarrow W$ heißt K -linear, falls für $\alpha \in K$ und $x, y \in V$ gilt:

$$\begin{aligned} f(x + y) &= f(x) + f(y), \\ f(\alpha x) &= \alpha f(x). \end{aligned}$$

Man beachte, daß es lineare Abbildungen immer nur über dem gleichen Körper geben kann.

Satz

Es seien V und W zwei K -Vektorräume, $\{a_1, \dots, a_m\}$ eine Basis von V . Dann gibt es zu jedem System von m Vektoren $w_1, \dots, w_m \in W$ genau eine lineare Abbildung $f : V \rightarrow W$ mit $f(a_i) = w_i$ für $i = 1, \dots, m$.

BEWEIS: Jeder Vektor $v \in V$ besitzt eine eindeutige Darstellung

$$v = \alpha_1 a_1 + \dots + \alpha_m a_m, \text{ mit } \alpha_i \in K.$$

Gibt es eine lineare Abbildung f von der gewünschten Art, so muß wegen der Linearität von f gelten:

$$\begin{aligned} f(v) &= f(\alpha_1 a_1 + \dots + \alpha_m a_m) \\ &= \alpha_1 f(a_1) + \dots + \alpha_m f(a_m) \\ &= \alpha_1 w_1 + \dots + \alpha_m w_m. \end{aligned}$$

Das zeigt die Eindeutigkeit!

Umgekehrt kann f auf diese Weise definiert werden:

$$f(\alpha_1 a_1 + \dots + \alpha_m a_m) := \alpha_1 w_1 + \dots + \alpha_m w_m.$$

Die Linearität läßt sich dann leicht nachrechnen. Ist

$$v = \alpha_1 a_1 + \dots + \alpha_m a_m \quad \text{und} \quad w = \beta_1 a_1 + \dots + \beta_m a_m,$$

so ist

$$\begin{aligned} f(v+w) &= f((\alpha_1 + \beta_1)a_1 + \dots + (\alpha_m + \beta_m)a_m) \\ &= (\alpha_1 + \beta_1)w_1 + \dots + (\alpha_m + \beta_m)w_m \\ &= (\alpha_1 w_1 + \dots + \alpha_m w_m) + (\beta_1 w_1 + \dots + \beta_m w_m) \\ &= f(v) + f(w) \end{aligned}$$

und

$$\begin{aligned} f(\lambda v) &= f((\lambda \alpha_1)a_1 + \dots + (\lambda \alpha_m)a_m) \\ &= (\lambda \alpha_1)w_1 + \dots + (\lambda \alpha_m)w_m \\ &= \lambda(\alpha_1 w_1 + \dots + \alpha_m w_m) \\ &= \lambda f(v). \end{aligned}$$

■

Beispiel.

Wir schreiben die Elemente von K^m als Spaltenvektoren. Etwas vornehmer könnten wir auch sagen, wir *identifizieren* den Vektorraum $V = K^m$ mit dem Matrizenraum $M_{m,1}(K)$, indem wir jeden Zeilenvektor \mathbf{v} durch den Spaltenvektor $\vec{v} = \mathbf{v}^t$ ersetzen.

In K^m haben wir die Standardbasis $\{\vec{e}_1, \dots, \vec{e}_m\}$. Eine lineare Abbildung $f : K^m \rightarrow K^n$ ist dann durch die Bilder der Einheitsvektoren $\vec{w}_j := f(\vec{e}_j) \in K^n$ festgelegt.

Schreibt man die m Spalten $\vec{w}_1, \dots, \vec{w}_m$ nebeneinander, so erhält man eine Matrix $M(f) = (\vec{w}_1, \dots, \vec{w}_m) \in M_{n,m}(K)$. Wir setzen

$$M(f) \cdot \vec{v} := f(\vec{v}).$$

Auch jeder Vektor aus K^n kann als Linearkombination von Einheitsvektoren dargestellt werden, die wir (zur Unterscheidung von den Einheitsvektoren im K^m) hier mit \vec{k}_i bezeichnen wollen. Dann gibt es Elemente $a_{ij} \in K$, so daß gilt:

$$\vec{w}_j = \sum_{i=1}^n a_{ij} \vec{k}_i = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}, \text{ also } M(f) = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Ist $\vec{x} = x_1 \vec{e}_1 + \cdots + x_m \vec{e}_m$, so ist

$$\begin{aligned} f(\vec{x}) &= x_1 \vec{w}_1 + \cdots + x_m \vec{w}_m \\ &= \sum_{j=1}^m x_j \left(\sum_{i=1}^n a_{ij} \vec{k}_i \right) \\ &= \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij} x_j \right) \vec{k}_i, \end{aligned}$$

also

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{pmatrix}.$$

So ist das Produkt einer Matrix mit einem Vektor definiert.

Wie früher definieren wir für eine lineare Abbildung $f : V \rightarrow W$:

$$\begin{aligned} \text{Ker}(f) &:= \{v \in V : f(v) = 0\} \\ \text{und } \text{Im}(f) &:= f(V) = \{y \in W : \exists x \in V \text{ mit } y = f(x)\}. \end{aligned}$$

Die lineare Abbildung ist genau dann injektiv, wenn $\text{Ker}(f) = \{0\}$ ist, und sie ist genau dann surjektiv, wenn $\text{Im}(f) = W$ ist.

Satz (Dimensionsformel)

Sei $f : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen K -Vektorräumen. Dann ist

$$\dim_K \text{Ker}(f) + \dim_K \text{Im}(f) = \dim_K V.$$

Der BEWEIS von früher kann praktisch unverändert übernommen werden.

Definition:

Eine bijektive lineare Abbildung $f : V \rightarrow W$ heißt ein *Isomorphismus*. Man schreibt dann auch: $V \cong W$.

Satz

Ist $f : V \rightarrow W$ ein Isomorphismus, so ist auch $f^{-1} : W \rightarrow V$ linear.

BEWEIS: Seien $w_1, w_2 \in W$, $\alpha \in K$. Dann gilt:

$$\exists v_1, v_2 \in V \text{ mit } f(v_1) = w_1 \text{ und } f(v_2) = w_2.$$

Also ist

$$\begin{aligned} f^{-1}(w_1 + w_2) &= f^{-1}(f(v_1) + f(v_2)) \\ &= f^{-1}(f(v_1 + v_2)) \\ &= v_1 + v_2 \\ &= f^{-1}(w_1) + f^{-1}(w_2), \end{aligned}$$

und

$$\begin{aligned} f^{-1}(\alpha w_1) &= f^{-1}(\alpha f(v_1)) \\ &= f^{-1}(f(\alpha v_1)) \\ &= \alpha v_1 \\ &= \alpha f^{-1}(w_1). \end{aligned}$$

■

Beispiel.

Sei

$$V := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$$

und $W := \{(y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : 2y_1 + 3y_2 = 0 \text{ und } y_3 + y_4 = 0\}$.

Beides sind reelle Vektorräume. Wie man das zeigt, haben wir schon im ersten Semester gesehen. Sei nun $f : V \rightarrow \mathbb{R}^4$ definiert durch

$$f(x_1, x_2, x_3) := (x_1, -\frac{2}{3}x_1, x_2, -x_2).$$

f ist offensichtlich Einschränkung einer linearen Abbildung $F : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ mit

$$F(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{2}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Wie man leicht sieht, ist $F(V) \subset W$. Also induziert F eine lineare Abbildung $f : V \rightarrow W$.

Offensichtlich ist

$$\text{Ker}(f) = \{(x_1, x_2, x_3) \in V : x_1 = x_2 = 0\} = \{(0, 0, 0)\},$$

also f injektiv.

Ist $\mathbf{y} = (y_1, y_2, y_3, y_4) \in W$, so ist $y_2 = -\frac{2}{3}y_1$ und $y_4 = -y_3$. Setzen wir $\mathbf{x} := (y_1, y_3, -y_1 - y_3)$, so ist $\mathbf{x} \in V$ und

$$f(\mathbf{x}) = (y_1, -\frac{2}{3}y_1, y_3, -y_3) = (y_1, y_2, y_3, y_4).$$

Also ist f surjektiv und damit bijektiv, also ein Isomorphismus.

Die Umkehrabbildung erhält man, indem man die Gleichung $f(\mathbf{x}) = \mathbf{y}$ nach \mathbf{x} auflöst. Also ist

$$f^{-1}(y_1, y_2, y_3, y_4) = (y_1, y_3, -y_1 - y_3).$$

Dabei haben wir die letzte Komponente aus den ersten beiden errechnet, es muß ja $x_1 + x_2 + x_3 = 0$ sein.

Für ein weiteres Beispiel müssen wir etwas ausholen.

Definition:

Sind V, W zwei K -Vektorräume, so setzen wir

$$L(V, W) := \{f : V \rightarrow W \mid f \text{ } K\text{-linear}\}.$$

Dann ist $L(V, W)$ selbst wieder ein K -Vektorraum, durch

$$\begin{aligned} (f + g)(v) &:= f(v) + g(v) \\ \text{und } (\lambda f)(v) &:= \lambda f(v). \end{aligned}$$

Das neutrale Element ist die „Null-Abbildung“ $o : V \rightarrow W$ mit $o(v) = 0$ für alle $v \in V$. Genau genommen müßte allerlei nachgerechnet werden, nämlich

1. Die Abbildungen $f + g$ und λf sind wieder linear.
2. In $L(V, W)$ sind alle Vektorraum-Axiome erfüllt.

Wir ersparen uns hier die Rechnerei.

Im Falle $V = K^m$ und $W = K^n$ gibt es eine Abbildung

$$M : L(K^m, K^n) \rightarrow M_{n,m}(K)$$

mit $M(f) \cdot \vec{x} = f(\vec{x})$. Da die Bilder der Einheitsvektoren gerade die Spalten der Matrix sind, ist $M(f) = (f(\vec{e}_1), \dots, f(\vec{e}_m))$.

Behauptung: M ist ein Isomorphismus.

BEWEIS: Um die Gleichungen $M(f + g) = M(f) + M(g)$ und $M(\lambda f) = \lambda M(f)$ zu beweisen, rechnen wir jeweils die Spalten aus:

$$\begin{aligned} M(f + g) \cdot \vec{e}_\nu &= (f + g)(\vec{e}_\nu) = f(\vec{e}_\nu) + g(\vec{e}_\nu) \\ &= M(f) \cdot \vec{e}_\nu + M(g) \cdot \vec{e}_\nu \\ &= (M(f) + M(g)) \cdot \vec{e}_\nu \\ \text{und } M(\lambda f) \cdot \vec{e}_\nu &= (\lambda f)(\vec{e}_\nu) = \lambda f(\vec{e}_\nu) \\ &= \lambda(M(f) \cdot \vec{e}_\nu) \\ &= (\lambda M(f)) \cdot \vec{e}_\nu. \end{aligned}$$

Weiter ist

$$\begin{aligned} \text{Ker}(M) &= \{f \in L(K^m, K^n) : M(f) = 0\} \\ &= \{f \in L(K^m, K^n) : M(f) \cdot \vec{e}_\nu = \vec{0} \text{ für alle } \nu\} \\ &= \{f \in L(K^m, K^n) : f(\vec{e}_\nu) = \vec{0} \text{ für alle } \nu\} \\ &= \{f \in L(K^m, K^n) : f(\vec{x}) = \vec{0} \text{ für alle } \vec{x}\} = \{o\}, \end{aligned}$$

also M injektiv.

Ist eine Matrix $A \in M_{m,n}(K)$ gegeben, so setzen wir $f = f_A$, mit $f_A(\vec{x}) := A \cdot \vec{x}$. Dann ist $M(f) \cdot \vec{e}_\nu = f(\vec{e}_\nu) = A \cdot \vec{e}_\nu$ für alle ν , also $M(f) = A$. Damit ist M auch surjektiv. ■

Satz

Sei $f : V \rightarrow W$ ein Isomorphismus zwischen zwei K -Vektorräumen. Ist $\dim(V) = n < \infty$, so ist auch $\dim(W) = n$.

BEWEIS: Ist $\{a_1, \dots, a_n\}$ eine Basis von V , so wird $\text{Im}(f)$ von den Vektoren $f(a_1), \dots, f(a_n)$ erzeugt, ist also wieder endlich-dimensional. Nach der Dimensionsformel ist $\dim \text{Im}(f) = n - \dim \text{Ker}(f) = n$. Aber weil f surjektiv ist, ist $\text{Im}(f) = W$. ■

Aus diesem Satz folgt:

$$\dim_K L(K^m, K^n) = \dim M_{n,m}(K) = n \cdot m.$$

Die Beziehungen zwischen dem Raum der linearen Abbildungen und dem Raum der Matrizen gehen noch weiter:

Es seien zwei lineare Abbildungen $f : K^m \rightarrow K^n$ und $g : K^n \rightarrow K^r$ gegeben. Dann ist auch $g \circ f : K^m \rightarrow K^r$ linear, und wir definieren:

$$M(g) \cdot M(f) := M(g \circ f).$$

Das ergibt ein Produkt von Matrizen aus $M_{r,n}(K)$ und $M_{n,m}(K)$. Die Anzahl der Spalten der ersten Matrix muß mit der Anzahl der Zeilen der zweiten Matrix übereinstimmen. Ist $M(g) = (a_{ij})$ und $M(f) = (b_{jk})$, so gilt für die Einheitsvektoren $\vec{e}_k \in K^m$, $\vec{k}_j \in K^n$ und $\vec{l}_i \in K^r$:

$$\begin{aligned} M(g) \cdot M(f) \cdot \vec{e}_k &= M(g \circ f) \cdot \vec{e}_k \\ &= g(f(\vec{e}_k)) \\ &= g\left(\sum_{j=1}^n b_{jk} \vec{k}_j\right) \\ &= \sum_{j=1}^n \sum_{i=1}^r a_{ij} b_{jk} \vec{l}_i \\ &= \sum_{i=1}^r \left(\sum_{j=1}^n a_{ij} b_{jk}\right) \vec{l}_i. \end{aligned}$$

Also ist

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix} = \begin{pmatrix} \cdots & \sum_{j=1}^n a_{ij}b_{jk} & \cdots \\ \vdots & & \vdots \end{pmatrix}$$

Um das Element in der i -ten Zeile und der k -ten Spalte des Matrizenproduktes zu erhalten, legt man die i -te Zeile der linken Matrix auf die k -te Spalte der rechten Matrix, multipliziert die übereinander liegenden Elemente miteinander und summiert diese Produkte auf. Das geht genau dann, wenn die Anzahl der Spalten der linken Matrix mit der Anzahl der Zeilen der rechten Matrix übereinstimmt.

Es folgen nun einige Rechenregeln für das Matrizenprodukt:

Satz

Voraussetzung ist immer, daß die Matrizen multiplizierbar sind. Dann gilt:

1. $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.
2. $A \cdot (B + B') = A \cdot B + A \cdot B'$.
3. $(A + A') \cdot B = A \cdot B + A' \cdot B$.
4. $\alpha(A \cdot B) = (\alpha A) \cdot B = A \cdot (\alpha B)$, für $\alpha \in K$.

Bemerkung. Auch wenn man das Matrizenprodukt $A \cdot B$ bilden kann, braucht das für $B \cdot A$ nicht zu gelten.

Sind A und B beides quadratische Matrizen aus $M_{n,n}(K)$, so kann man $A \cdot B$ und $B \cdot A$ bilden, aber i.a. ist $A \cdot B \neq B \cdot A$. So ist z.B.

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 4 & 4 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 6 \end{pmatrix}.$$

Auch kann man i.a. in Matrixgleichungen nicht kürzen! Es ist

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix},$$

aber natürlich

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

§ 2 Lineare Koordinaten und Basiswechsel

Inhalt:

Lineare Koordinatensysteme, Basiswechsel, invertierbare Matrizen, die allgemeine lineare Gruppe, Invarianz des Ranges, Invertierung von Matrizen, Beschreibung linearer Abbildungen durch Matrizen.

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

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

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

Durch $\Phi_A(x) := (x_1, \dots, x_n)$ wird also eine Abbildung $\Phi_A : V \rightarrow K^n$ definiert. Wir nennen Φ_A das durch A bestimmte *lineare Koordinatensystem* für V , und die Koeffizienten x_1, \dots, x_n heißen die *Koordinaten* von x bezüglich der Basis A . Zum Rechnen (mit Matrizen) sind Spaltenvektoren oftmals besser geeignet. Deshalb setzen wir noch

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

Beispiele.

1. Ist V selbst der K^n und $E = \{e_1, \dots, e_n\}$ die Standardbasis, so ist $\Phi_E = \text{id}_V$, und $[\mathbf{x}]_E = \mathbf{x}^t = \vec{x}$.
2. Nun sei $V = K^n$, aber $A = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ eine beliebige Basis, mit $\mathbf{a}_i = (a_{1i}, \dots, a_{ni})$ für $i = 1, \dots, n$.

Wir können die Vektoren \mathbf{a}_i , als Spalten $\vec{a}_i = \mathbf{a}_i^t$ geschrieben, zu einer Matrix $A \in M_{n,n}(K)$ zusammenfassen:

$$A = (\vec{a}_1, \dots, \vec{a}_n) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Wir haben hier einen Fall von „Notations-Mißbrauch“. Das Symbol A bezeichnet sowohl die Basis, als auch die Matrix, die aus den Basisvektoren

gebildet werden kann. Es muß jeweils aus dem Kontext entnommen werden, welche Bedeutung gerade benutzt wird.

Sei jetzt $\mathbf{y} = (y_1, \dots, y_n)$ ein beliebiger Vektor in V . Dann gilt :

$$\mathbf{y} = \sum_{i=1}^n x_i \mathbf{a}_i \iff y_k = \sum_{i=1}^n x_i a_{ki} \text{ für } k = 1, \dots, n,$$

also

$$\Phi_A(\mathbf{y}) = \mathbf{x} \iff \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = y_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = y_n. \end{cases} \iff A \cdot \vec{x} = \vec{y}.$$

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

Der Vektor $\vec{x} = [\mathbf{y}]_A$ ist die eindeutig bestimmte Lösung des Linearen Gleichungssystems $A \cdot \vec{x} = \vec{y}$.

Speziell ist $[\mathbf{a}_\nu]_A = \vec{e}_\nu$ für $\nu = 1, \dots, n$.

3. Sei jetzt $V \subset K^n$ ein r -dimensionaler Untervektorraum. Das ist schon fast der allgemeinste Fall, mit dem wir uns zu befassen haben. Die Elemente $\mathbf{a}_i = (a_{1i}, \dots, a_{ni})$ einer Basis von V , $i = 1, \dots, r$, sind immer noch Vektoren im K^n , aber wir haben nur r davon. Also ist

$$A = (\vec{a}_1, \dots, \vec{a}_r) \in M_{n,r}(K).$$

Die Berechnung der Koordinaten funktioniert aber genauso wie oben.

Ist $\mathbf{y} \in V$, so ist der Vektor $\vec{x} = [\mathbf{y}]_A \in K^r$ die eindeutig bestimmte Lösung des Linearen Gleichungssystems $A \cdot \vec{x} = \vec{y}$.

Es ist $\text{rg}(A) = r$, und da jeder Vektor $\vec{y} \in V$ Linearkombination der Spalten von A ist, hat auch die erweiterte Matrix (A, \vec{y}) den Rang r . Deshalb ist das LGS tatsächlich lösbar.

4. Sei $V = M_{2,2}(K)$. Dann bilden die Matrizen

$$E_{11} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

eine Basis A von V , und $\Phi_A : M_{2,2}(K) \rightarrow K^4$ ist gegeben durch

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

5. Wir behandeln hier eigentlich nur endlich-dimensionale Vektorräume. Trotzdem soll angedeutet werden, wie es im unendlichdimensionalen Fall aussehen kann. Es sei $V = \mathbb{R}[X]$ der Raum aller Polynome mit reellen Koeffizienten. Die Polynome $1, X, X^2, X^3$, usw. bilden eine (abzählbare) Basis A von $\mathbb{R}[X]$. Dann setzen wir

$$\Phi_A \left(\sum_{i=0}^n a_i X^i \right) := (a_0, a_1, \dots, a_n, 0, 0, \dots).$$

Wir erhalten zwar Vektoren mit unendlich vielen Komponenten, aber nur endlich viele davon sind jeweils $\neq 0$.

Sehr wichtig ist die Feststellung, daß das lineare Koordinatensystem Φ_A von der Basis $A = \{a_1, \dots, a_n\}$ abhängt. Was passiert, wenn man eine zweite Basis $B = \{b_1, \dots, b_n\}$ von V betrachtet. Welcher Zusammenhang besteht für einen Vektor $x \in V$ zwischen $\Phi_A(x)$ und $\Phi_B(x)$?

Man sieht sehr leicht, daß Φ_A eine lineare Abbildung ist. Und außerdem ist Φ_A bijektiv, die Umkehrabbildung ist durch

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

gegeben.

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

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_B \circ \Phi_A^{-1}} & K^n \\ \Phi_A \swarrow & & \nearrow \Phi_B \\ & V & \end{array}$$

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

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

In der Sprache der Matrizen heißt das:

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

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

Für die Spalten der Basiswechsel-Matrix gilt:

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

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

Beispiel.

Sei $V \subset K^n$ ein r -dimensionaler Untervektorraum, und es seien zwei Basen

$$A = \{\mathbf{a}_1, \dots, \mathbf{a}_r\} \quad \text{und} \quad B = \{\mathbf{b}_1, \dots, \mathbf{b}_r\}$$

gegeben. Die zugehörigen Matrizen in $M_{n,r}(K)$ bezeichnen wir ebenfalls mit A und B . Die Basiswechsel-Matrix $W_{B,A}$ liegt in diesem Falle in $M_{r,r}(K)$, und es gilt:

$$(B \cdot W_{B,A}) \cdot \vec{e}_\nu = B \cdot (W_{B,A} \cdot \vec{e}_\nu) = B \cdot [\mathbf{a}_\nu]_B = \vec{a}_\nu = A \cdot \vec{e}_\nu,$$

also

$$B \cdot W_{B,A} = A.$$

Die ν -te Spalte von $W_{B,A}$ ist die Lösung $\vec{x} = \vec{x}_\nu$ des LGS $B \cdot \vec{x} = \vec{a}_\nu$.

Kommen wir zurück zu einem allgemeinen Basiswechsel in einem n -dimensionalen K -Vektorraum V . Es gilt:

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

und

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

Bezeichnen wir die n -reihige Einheitsmatrix $(\vec{e}_1, \dots, \vec{e}_n)$ mit E_n , so erhalten wir (weil $[x]_A$ und $[x]_B$ jeweils völlig beliebig gewählt werden können und z.B. $[a_\nu]_A = \vec{e}_\nu$ ist):

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

Definition:

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

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

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

Eigenschaften invertierbarer Matrizen

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

1. A ist invertierbar.
2. $f_A : K^n \rightarrow K^n$ ist ein Isomorphismus.
3. $\text{Ker}(f_A) = \{0\}$.
4. Das LGS $A \cdot \vec{x} = \vec{0}$ ist eindeutig lösbar.
5. Das LGS $A \cdot \vec{x} = \vec{b}$ ist für jedes $\vec{b} \in K^n$ eindeutig lösbar.
6. $\text{rg}(A) = n$.

BEWEIS:

(1) \implies (2):

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

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

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

(2) \implies (3):

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

(3) \implies (4):

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

(4) \implies (5):

Daß $A \cdot \vec{x} = \vec{0}$ eindeutig lösbar ist, bedeutet, daß $f_A : K^n \rightarrow K^n$ injektiv ist. Wegen der Dimensionsformel ist dann f_A auch surjektiv, und jedes LGS $A \cdot \vec{x} = \vec{b}$ lösbar. Ist $A \cdot \vec{x}_1 = A \cdot \vec{x}_2 = \vec{b}$, so ist $A \cdot (\vec{x}_1 - \vec{x}_2) = \vec{0}$, also $\vec{x}_1 = \vec{x}_2$.

(5) \implies (6):

Nach der Dimensionsformel ist $\text{rg}(A) + \dim_K(\text{Ker}(f_A)) = n$. Aus der eindeutigen Lösbarkeit des LGS $A \cdot \vec{x} = \vec{0}$ folgt, daß $\text{Ker}(f_A) = \{0\}$ ist. Also ist $\text{rg}(A) = n$.

(6) \implies (1):

Ist $\text{rg}(A) = n$, so bilden die Spalten $\vec{a}_j := \vec{s}_j(A)$ eine Basis von K^n . Weil $f(\vec{e}_j) = \vec{a}_j$ ist, für $j = 1, \dots, n$, wird durch $\vec{a}_j \mapsto \vec{e}_j$ eine Umkehrabbildung zu f_A definiert. Mit f_A ist auch A invertierbar. ■

Definition:

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

Die Menge aller n -reihigen regulären Matrizen wird mit $GL_n(K)$ bezeichnet („General Linear Group“ oder „allgemeine lineare Gruppe“).

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

Gruppeneigenschaft der GL_n

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

BEWEIS:

- 1) E_n liegt in $GL_n(K)$ und spielt die Rolle des neutralen Elements.
- 2) Jedes $A \in GL_n(K)$ besitzt ein Inverses, das wieder in $GL_n(K)$ liegt.
- 3) Seien $A, B \in GL_n(K)$. Dann ist $A \cdot B \in M_{n,n}(K)$, und es gilt:

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

und genauso

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

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

- 4) Offensichtlich ist $A \cdot (B \cdot C) = (A \cdot B) \cdot C$. ■

Beispiele.

1. $n = 1$:

$$GL_1(K) = \{a \in K \mid ax = 0 \text{ eindeutig lösbar}\} = K^* = K \setminus \{0\}.$$

2. $n = 2$:

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist genau dann regulär, wenn $\text{rg}(A) = 2$ ist, wenn also $ad - bc = \det(A) \neq 0$ ist (vgl. Mathematik 1, Kap. 4, §8).

$$GL_2(K) = \{A \in M_{2,2}(K) \mid \det(A) \neq 0\}.$$

3. Eine Diagonalmatrix

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

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

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

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

Satz von der Verkleinerung des Ranges

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

Dann ist $\text{rg}(A \cdot B) \leq \text{rg}(A)$ und $\text{rg}(A \cdot B) \leq \text{rg}(B)$.

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

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

$$\text{rg}(A \cdot B) = \dim \text{Im}(f \circ g) \leq \dim \text{Im}(f) = \text{rg}(A).$$

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

$$\text{rg}(A \cdot B) = \dim(f(\text{Im}(g))) \leq k = \dim \text{Im}(g) = \text{rg}(B).$$

■

Invarianz des Ranges

Die Multiplikation mit einer regulären Matrix ändert den Rang nicht:

Ist $A \in M_{n,m}(K)$, $P \in \text{GL}_n(K)$ und $Q \in \text{GL}_m(K)$, so ist $\text{rg}(P \cdot A \cdot Q) = \text{rg}(A)$.

BEWEIS: Nach dem Satz von der Verkleinerung des Ranges ist $\text{rg}(P \cdot A \cdot Q) \leq \text{rg}(A)$. Aber da $A = P^{-1} \cdot (P \cdot A \cdot Q) \cdot Q^{-1}$ ist, gilt auch die umgekehrte Ungleichung. Zusammen ergibt das die Gleichheit. ■

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

Eine Zeilenoperation vom Typ (I) ist eine lineare Abbildung der Form

$$\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} 1 & \cdots & 0 \\ & \ddots & \\ \vdots & & \lambda & \vdots \\ & & & \ddots \\ 0 & \cdots & & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ \lambda x_i \\ \vdots \\ x_n \end{pmatrix},$$

und eine Zeilenoperation vom Typ (II) ist eine lineare Abbildung der Form

$$\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} 1 & \cdots & & & 0 \\ & \ddots & & & \\ \vdots & & 1 & & 1 & \vdots \\ & & & \ddots & & \\ \vdots & & & & 1 & \vdots \\ & & & & & \ddots \\ 0 & \cdots & & & & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_i + x_k \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix},$$

Eine Folge von Zeilenoperationen ist also eine bijektive lineare Abbildung der Gestalt

$$\vec{x} \mapsto P \cdot \vec{x},$$

simultan angewandt auf alle Spalten von A , und eine Folge von Spaltenvertauschungen ist eine bijektive lineare Abbildung der Gestalt

$$(x_1, \dots, x_n) \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (x_1, \dots, x_n) \cdot Q,$$

simultan angewandt auf alle Zeilen von A , mit einer Permutation $\sigma \in S_n$.

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

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

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

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

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

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

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

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

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

Zusammengefaßt ergibt sich so:

Verfahren zur Invertierung von Matrizen

Sei $A \in \text{GL}_n(K)$.

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

BEWEIS: Daß A allein durch eine Folge von Zeilenoperationen in die Einheitsmatrix verwandelt werden kann, haben wir oben gesehen. Wenn diese Folge von Zeilenoperationen durch die invertierbare Matrix P verwirklicht wird, hat man die Gleichung

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

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

Beispiel.

Sei $B := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Offensichtlich ist $\text{rg}(B) = 3$. Wir versuchen jetzt, auf

$$(B | E_3) = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right)$$

so lange Zeilenoperationen anzuwenden, bis in der linken Hälfte die Einheitsmatrix steht.

1	1	0	1	0	0	
1	0	1	0	1	0	
0	1	1	0	0	1	
1	1	0	1	0	0	
0	-1	1	-1	1	0	(- 1. Zeile)
0	1	1	0	0	1	
1	0	1	0	1	0	(+ 2. Zeile)
0	-1	1	-1	1	0	
0	0	2	-1	1	1	(+ 2. Zeile)
1	0	1	0	1	0	
0	1	-1	1	-1	0	($\times(-1)$)
0	0	1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	($\times\frac{1}{2}$)
1	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	(- 3. Zeile)
0	1	0	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	(+ 3. Zeile)
0	0	1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	

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

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

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

$$\vec{x} = A^{-1} \cdot \vec{b}.$$

2. Sei $A = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ eine Basis von K^n . Wie üblich bezeichnen wir die aus den Spaltenvektoren $\vec{a}_1, \dots, \vec{a}_n$ gebildete Matrix auch mit A . Dann ist A invertierbar.

Ist nun $\mathbf{y} \in K^n$, so ist $[\mathbf{y}]_A$ die eindeutig bestimmte Lösung des LGS $A \cdot \vec{x} = \vec{y}$. Daraus folgt:

$$[\mathbf{y}]_A = A^{-1} \cdot \vec{y}.$$

Dies funktioniert nicht mehr, wenn wir einen Vektor \mathbf{y} aus einem r -dimensionalen Untervektorraum $V \subset K^n$ bezüglich einer Basis $A = \{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ darstellen wollen, denn in diesem Falle ist $A \in M_{n,r}(K)$ nicht einmal eine quadratische Matrix, kann also nicht invertiert werden. Gültig bleibt aber die Gleichung

$$A \cdot [\mathbf{y}]_A = \vec{y}.$$

3. Sind $A = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ und $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ Basen von K^n , so ist

$$B \cdot W_{B,A} = A,$$

also

$$W_{B,A} = B^{-1} \cdot A.$$

Auch diese explizite Formel gilt nicht mehr, wenn wir es mit Basen eines Untervektorraums von K^n zu tun haben.

Beispiele.

1. Die Vektoren $\mathbf{b}_1 := (1, 1, 0)$, $\mathbf{b}_2 := (1, 0, 1)$ und $\mathbf{b}_3 := (0, 1, 1)$ bilden eine Basis B des \mathbb{R}^3 . Die Matrix $B := (b_1, b_2, b_3)$ haben wir oben schon einmal betrachtet.

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

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

Probe: Tatsächlich ist

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

2. Die Vektoren $\mathbf{a}_1 = (1, 3, 2)$, $\mathbf{a}_2 = (0, -1, 5)$ und $\mathbf{a}_3 = (4, 2, -3)$ bilden auch eine Basis des \mathbb{R}^3 . Dann ist

$$\begin{aligned} W_{B,A} &= B^{-1} \cdot A = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 4 \\ 3 & -1 & 2 \\ 2 & 5 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -3 & 9/2 \\ 0 & 3 & -1/2 \\ 2 & 2 & -5/2 \end{pmatrix}. \end{aligned}$$

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

$$M(f) \cdot \vec{x} = f(\vec{x}).$$

Die Beschreibung von f durch die Matrix $M(f)$ benutzt die Darstellung von Vektoren mit Hilfe der Standardbasen. Haben wir es mit beliebigen (endlich-dimensionalen) Vektorräumen und beliebigen Basen zu tun, so benutzen wir Koordinatensysteme, um auch in diesem Fall lineare Abbildungen durch Matrizen zu beschreiben.

Gegeben seien also

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

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

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

Mit $M_{B,A}(f) \in M_{n,m}(K)$ bezeichnen wir die Matrix der α_{ij} . Man beachte dabei die Stellung der Indizes. Links steht die Basis des Zielraums und rechts die Basis des Ausgangsraums.

Mit Hilfe von Koordinaten kann man die Situation folgendermaßen beschreiben:

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

Die Abbildung $f_{B,A} := \Phi_B \circ f \circ \Phi_A^{-1} : K^m \rightarrow K^n$ wird erst durch das vorgelegte Diagramm definiert. Sie ist ebenfalls linear und wird in der üblichen Weise durch eine Matrix $M(f_{B,A})$ beschrieben. Wir zeigen, daß $M(f_{B,A}) = M_{B,A}(f)$ ist. Es gilt nämlich:

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

Definition:

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

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

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

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

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

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

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

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

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

Ein Sonderfall liegt vor, wenn f eine lineare Abbildung von K^m nach K^n ist. Dann kann f bezüglich der Standardbasen durch eine Matrix $M := M(f) \in M_{n,m}(K)$ beschrieben werden:

$$f(\vec{x}) = M \cdot \vec{x}.$$

Wir wollen f aber bezüglich anderer Basen $A = \{\vec{a}_1, \dots, \vec{a}_m\}$ von K^m und $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ von K^n beschreiben. Dann ist

$$\begin{aligned} M_{B,A}(f) &= ([f(\vec{a}_1)]_B, \dots, [f(\vec{a}_m)]_B) \\ &= ([M \cdot \vec{a}_1]_B, \dots, [M \cdot \vec{a}_m]_B) \\ &= (B^{-1} \cdot M \cdot \vec{a}_1, \dots, B^{-1} \cdot M \cdot \vec{a}_m) \\ &= B^{-1} \cdot M \cdot A. \end{aligned}$$

$$f = f_M : K^m \rightarrow K^n \text{ wird beschrieben durch } M_{B,A}(f) = B^{-1} \cdot M \cdot A.$$

Sehr häufig wird auch der Fall $V = W$ betrachtet. Eine lineare Abbildung $f : V \rightarrow V$ bezeichnet man als einen *Endomorphismus* von V . Man kann dann versuchen, mit einiger einzigen Basis A auszukommen und f durch $M_A(f) := M_{A,A}(f)$ zu beschreiben. Ist $V = K^n$, so liegt es nahe, die Standardbasis zu benutzen. Aber oft zeigt es sich, daß man durch geschickte Wahl der Basis zu einer viel einfacheren Matrix kommen kann. Wie das geht, werden wir am Ende von Kapitel 7 in der Eigenwert-Theorie sehen.

Beispiele.

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

$$\Phi_A(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_0, a_1, a_2, a_3).$$

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

$$D(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2.$$

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

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

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

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

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

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

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

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

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

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

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

§ 3 Dualität und Orthogonalität

Inhalt:

Bilinearformen, Dualitäten, Dualraum und duale Basis, Orthogonalität, ON-Basen, Orthogonalisierungsverfahren, orthogonales Komplement, orthogonale Projektion, adjungierte Abbildung, hermitesche Formen und Skalarprodukte.

Definition:

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

1. Für alle $v \in V$ ist die Abbildung $w \mapsto b(v, w)$ linear.
2. Für alle $w \in W$ ist die Abbildung $v \mapsto b(v, w)$ linear.

Beispiele.

1. Sei $V = W = K^n$ und $b(\mathbf{v}, \mathbf{w}) := \mathbf{v} \cdot \mathbf{w}^t = v_1 w_1 + \dots + v_n w_n$. Wir bezeichnen dies als die *kanonische Bilinearform* auf K^n .

Im Falle $K = \mathbb{R}$ erhalten wir so das euklidische Skalarprodukt $\mathbf{v} \bullet \mathbf{w}$. Ist dann $\mathbf{v} \neq \mathbf{0}$, so ist $b(\mathbf{v}, \mathbf{v}) = \|\mathbf{v}\|^2 > 0$. Ist K dagegen ein anderer Körper, so braucht das nicht zu gelten. Im Falle $K = \mathbb{C}$ und $\mathbf{v} = (j, 0, \dots, 0)$ erhält man z.B. $b(\mathbf{v}, \mathbf{v}) = -1$.

2. Es sei wieder $V = W = K^n$. Weiter sei $A \in M_{n,n}(K)$. Dann ist auch

$$b_A(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot A \cdot \mathbf{w}^t = (v_1, \dots, v_n) \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

eine Bilinearform. Die Eigenschaften von b_A können – abhängig von A – sehr unterschiedlich sein.

Ist z.B. $K = \mathbb{R}$, $n = 2$ und $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, so ist

$$b_A((x_1, x_2), (y_1, y_2)) = x_1 y_2 - x_2 y_1 = \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}.$$

Dann ist $b_A(\mathbf{x}, \mathbf{x}) = 0$ für alle $\mathbf{x} \in \mathbb{R}^2$.

Die kanonische Bilinearform erhält man auch auf diese Weise, durch die Matrix $A = E_n$ (Einheitsmatrix).

Tatsächlich ist **jede** Bilinearform auf dem K^n von der Form b_A , mit einer geeigneten Matrix A . Ist eine Bilinearform b gegeben, so erhält man die zugehörige Matrix $A = (a_{ij} : i, j = 1, \dots, n)$ durch die Formel

$$a_{ij} = b_A(\mathbf{e}_i, \mathbf{e}_j), \text{ für } i, j = 1, \dots, n.$$

Man nennt eine Bilinearform $b : V \times V \rightarrow K$ *symmetrisch*, falls $b(v, w) = b(w, v)$ für alle $v, w \in V$ gilt. Die kanonische Bilinearform auf K^n (und damit auch das euklidische Skalarprodukt) ist symmetrisch. Die Bilinearform $b((x_1, x_2), (y_1, y_2)) = x_1 y_2 - x_2 y_1$ ist dagegen nicht symmetrisch. Und bei Bilinearformen $b : V \times W \rightarrow K$ mit $V \neq W$ macht der Begriff der „Symmetrie“ keinen Sinn.

Wir wollen uns noch überlegen, wann eine Bilinearform $b = b_A$ auf dem K^n symmetrisch ist. Dazu brauchen wir noch eine Aussage über das Transponieren von Matrizen:

Behauptung: Ist $A \in M_{n,m}(K)$ und $B \in M_{m,k}(K)$, so ist

$$(A \cdot B)^t = B^t \cdot A^t.$$

BEWEIS: Ist $A = (a_{ik})$ und $B = (b_{kj})$, so ist

$$((A \cdot B)^t)_{ij} = (A \cdot B)_{ji} = \sum_{k=1}^m a_{jk} b_{ki}$$

und

$$(B^t \cdot A^t)_{ij} = \sum_{k=1}^m (B^t)_{ik} (A^t)_{kj} = \sum_{k=1}^m a_{jk} b_{ki}.$$

Ein Skalar $\alpha \in K$ kann auch als (1×1) -Matrix aufgefaßt werden. Dann ist $\alpha^t = \alpha$. Ist nun $b = b_A$ die gegebene Bilinearform, so ist

$$b(\mathbf{w}, \mathbf{v}) = b(\mathbf{w}, \mathbf{v})^t = (\mathbf{w} \cdot A \cdot \mathbf{v}^t)^t = \mathbf{v} \cdot A^t \cdot \mathbf{w}^t.$$

Ist b zudem symmetrisch, so ist

$$b(\mathbf{w}, \mathbf{v}) = b(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot A \cdot \mathbf{w}^t.$$

Beides zusammen kann nur gelten, wenn $A = A^t$ ist.

Eine Matrix $A \in M_{n,n}(K)$ wird *symmetrisch* genannt, wenn $A^t = A$ ist. Eine Bilinearform $b = b_A$ auf dem K^n ist genau dann symmetrisch, wenn die zugehörige Matrix A symmetrisch ist.

Definition:

Eine *Dualität* zwischen zwei Vektorräumen V und W ist eine Bilinearform $b : V \times W \rightarrow K$ mit folgenden Eigenschaften:

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

Ist $V = W$, so sprechen wir von einer *Dualität auf V* .

Was bedeutet diese Definition? Eine Bilinearform $b : V \times W \rightarrow K$ liefert zu jedem $w \in W$ eine Linearform $L_w \in L(V, K)$ durch $L_w(v) := b(v, w)$ und zu jedem $v \in V$ eine Linearform $R_v \in L(W, K)$ durch $R_v(w) := b(v, w)$. Dadurch erhalten wir zwei lineare Abbildungen $L : W \rightarrow L(V, K)$ und $R : V \rightarrow L(W, K)$, mit

$$L : w \mapsto L_w \quad \text{und} \quad R : v \mapsto R_v.$$

Eine Dualität liegt genau dann vor, wenn $\text{Ker}(L) = \{0\}$ und $\text{Ker}(R) = \{0\}$ ist, wenn also beide Abbildungen injektiv sind. Manchmal nennt man eine Dualität auch eine *nicht ausgeartete Bilinearform*.

Um dieses recht abstrakte Konzept zu verstehen, betrachten wir mehrere Beispiele. Im wesentlichen muß man mit diesen Beispielen zurechtkommen.

Beispiele.

1. Es seien $\alpha_1, \dots, \alpha_n$ Elemente $\neq 0$ von K , und $\alpha = (\alpha_1, \dots, \alpha_n)$. Wir definieren $b_\alpha : K^n \times K^n \rightarrow K$ durch

$$b_\alpha(\mathbf{x}, \mathbf{y}) := \alpha_1 x_1 y_1 + \dots + \alpha_n x_n y_n.$$

Ist $b_\alpha(\mathbf{x}, \mathbf{y}) = 0$ für alle \mathbf{y} , so ist insbesondere

$$0 = b_\alpha(\mathbf{x}, \mathbf{e}_i) = \alpha_i x_i, \text{ also } x_i = 0, \text{ für alle } i.$$

Das gleiche gilt, wenn man \mathbf{x} und \mathbf{y} vertauscht. Also ist b_α eine Dualität auf dem K^n .

Spezialfälle sind das euklidische Skalarprodukt, aber auch das „Minkowski-Skalarprodukt“ auf dem \mathbb{R}^4 :

$$m(\mathbf{x}, \mathbf{y}) := x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4.$$

Letzteres spielt eine bedeutende Rolle in der Relativitätstheorie.

2. Die Bilinearform $b : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $b((x_1, x_2), (y_1, y_2)) := x_1 y_1$ ist **keine** Dualität, denn es ist $b((x_1, x_2), (0, 1)) = 0$ für alle $\mathbf{x} = (x_1, x_2)$.

3. Sei V ein endlich-dimensionaler K -Vektorraum. Die Menge aller Linearformen auf V bildet den Vektorraum

$$V^* := L(V, K).$$

Ist etwa $V = K^n$, so ist $L(K^n, K) \cong M_{1,n}(K) = K^n$. Der Isomorphismus funktioniert folgendermaßen: Jeder $1 \times n$ -Matrix (also jedem Vektor) $\mathbf{a} \in K^n$ wird eine lineare Abbildung $\lambda_{\mathbf{a}} : K^n \rightarrow K$ zugeordnet, durch

$$\lambda_{\mathbf{a}}(\mathbf{x}) := \mathbf{a} \cdot \mathbf{x}^t = a_1x_1 + \cdots + a_nx_n.$$

Im Falle $K = \mathbb{R}$ haben wir diese Zuordnung von Vektoren auf Linearformen schon im ersten Semester kennen gelernt.

Ist V beliebig, so wählen wir eine Basis A von V und erhalten ein Koordinatensystem $\Phi_A : V \rightarrow K^n$. Dies ist ein Isomorphismus, und wir erhalten einen weiteren Isomorphismus $(K^n)^* \rightarrow V^*$, durch

$$\lambda \mapsto \lambda \circ \Phi_A.$$

Daß die letztere Abbildung linear ist, läßt sich leicht nachrechnen, und die Umkehrabbildung dazu ist durch $\varrho \mapsto \varrho \circ \Phi_A^{-1}$ gegeben. Damit ist

$$V \cong K^n \cong (K^n)^* \cong V^*.$$

Ein endlich-dimensionaler Vektorraum ist also immer zu seinem Dualraum isomorph.

Behauptung: Die Abbildung $b : V \times V^* \rightarrow K$ mit $b(v, f) := f(v)$ ist eine Dualität.

BEWEIS:

1) Die Linearität im 1. Argument ist trivial, denn für festes f ist natürlich $v \mapsto f(v)$ linear.

2) Linearität im 2. Argument: Ist $v \in V$ fest, so gilt:

$$b(v, f_1 + f_2) = (f_1 + f_2)(v) = f_1(v) + f_2(v) = b(v, f_1) + b(v, f_2)$$

und

$$b(v, \alpha f) = (\alpha f)(v) = \alpha \cdot f(v) = \alpha \cdot b(v, f).$$

3) Ist $b(v, f) = 0$ für alle v , so ist $f(v) = 0$ für alle v , also $f = 0$ die Nullabbildung.

Sei jetzt $b(v, f) = 0$ für alle f . Wir nehmen an, es sei $v \neq 0$. Dann gibt es eine Basis $\{v_1, \dots, v_n\}$ von V mit $v = v_1$ (Basis-Ergänzungssatz). Nun definieren wir eine Linearform f_0 auf V durch

$$f_0(v_i) := \begin{cases} 1 & \text{falls } i = 1, \\ 0 & \text{sonst.} \end{cases}$$

Damit ist $b(v, f_0) = f_0(v_1) = 1$, und das ist ein Widerspruch. Also muß $v = 0$ sein. ■

Man nennt $V^* = L(V, K)$ den *Dualraum* von V .

Satz (Existenz der „dualen Basis“)

V und W seien endlich-dimensionale K -Vektorräume. Gibt es eine Dualität

$$b : V \times W \rightarrow K,$$

so sind V und W isomorph, und zu jeder Basis $\{a_1, \dots, a_n\}$ von V gibt es genau eine Basis $\{b_1, \dots, b_n\}$ von W , so daß gilt:

$$b(a_i, b_j) = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst.} \end{cases} \quad (\text{Kronecker-Symbol})$$

BEWEIS: Wir benutzen folgende Beobachtung: Ist $f : A \rightarrow B$ eine injektive lineare Abbildung (zwischen endlich-dimensionalen Vektorräumen), so ist

$$\dim(\text{Im}(f)) = \dim(A) - \dim \text{Ker}(f) = \dim \text{Im}(f) \leq \dim(B).$$

Ist $b : V \times W \rightarrow K$ eine Dualität, so sind die zugehörigen linearen Abbildungen $L : W \rightarrow V^*$ (mit $w \mapsto b(\cdot, w)$) und $R : V \rightarrow W^*$ (mit $v \mapsto b(v, \cdot)$) injektiv. Also ist

$$\dim(V) \leq \dim(W^*) = \dim(W) \leq \dim(V^*) = \dim(V).$$

Das ist nur möglich, wenn überall Gleichheitszeichen stehen. Also sind L und R Isomorphismen, und es ist $V \cong W$.

Ist $\{a_1, \dots, a_n\}$ eine Basis von V , so gibt es eindeutig bestimmte Linearformen $\lambda_j \in V^*$ mit

$$\lambda_j(a_i) = \delta_{ij}, \text{ für alle } i, j,$$

und zu jeder Linearform $\lambda_j \in V^*$ gibt es einen eindeutig bestimmten Vektor $b_j \in W$ mit $L(b_j) = \lambda_j$. Dann ist

$$b(a_i, b_j) = L(b_j)(a_i) = \lambda_j(a_i) = \delta_{ij}.$$

Zur Eindeutigkeit: Ist $\{b_1^*, \dots, b_n^*\}$ eine weitere Basis mit der gewünschten Eigenschaft, so ist

$$b(a_i, b_j - b_j^*) = \delta_{ij} - \delta_{ij} = 0 \text{ für alle } i, j,$$

also $b_j = b_j^*$ für alle j . ■

Beispiele.

1. Sei $V = W = \mathbb{R}^n$. Zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ heißen *orthogonal*, falls $\mathbf{x} \bullet \mathbf{y} = 0$ ist. Unter einer *Orthonormalbasis* von \mathbb{R}^n versteht man eine Basis $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ mit $\mathbf{a}_i \bullet \mathbf{a}_j = \delta_{ij}$ für alle i, j . Sie ist zu sich selbst dual.
2. Sei $W = V^*$. Ist $\{a_1, \dots, a_n\}$ eine Basis von V , so ist die duale Basis ein System $\{\lambda_1, \dots, \lambda_n\}$ von Linearformen mit $\lambda_i(a_j) = \delta_{ij}$.

Ist speziell $V = \mathbb{R}^n$, so besteht die duale Basis zur Standardbasis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ aus den Linearformen ε^i mit

$$\varepsilon^i(x_1, \dots, x_n) = x_i, \text{ für } i = 1, \dots, n.$$

Die Berechnung der dualen Basis zu einer (beliebigen) Basis des \mathbb{R}^n ist übrigens ganz einfach. Ist $A = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ eine Basis, so bilden wir wie üblich die Matrix $A = (\vec{a}_1, \dots, \vec{a}_n)$. Die Elemente von $L(\mathbb{R}^n, \mathbb{R}) \cong M_{1,n}(\mathbb{R}) = \mathbb{R}^n$ schreiben wir grundsätzlich als Zeilenvektoren. Sind $\alpha_1, \dots, \alpha_n$ die Elemente der dualen Basis, so wirken sie auf einen Vektor \mathbf{x} durch

$$\alpha_i(\mathbf{x}) = \alpha_i \bullet \mathbf{x}.$$

Fassen wir die α_i zu einer Matrix $A^* := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ zusammen, so ist

$$A^* \cdot A = \left(\alpha_i(\mathbf{a}_j) \mid \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, n \end{array} \right) = E_n, \text{ also } A^* = A^{-1}.$$

Definition:

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

Orthogonalsysteme sind linear unabhängig

Sei $S \subset \mathbb{R}^n$ ein Orthogonalsystem. Dann ist S linear unabhängig und enthält insbesondere höchstens n Elemente.

BEWEIS:

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

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

Also ist $\lambda_l = 0$ für $l = 1, \dots, k$. Damit sind die Vektoren linear unabhängig, und da sie beliebig ausgewählt wurden, ist ganz S linear unabhängig. ■

Eine *Orthonormalbasis* (kurz: *ON-Basis*) im \mathbb{R}^n ist demnach ein maximales Orthogonalsystem $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$, in dem die Vektoren \mathbf{a}_i zusätzlich normiert sind (d.h., $\|\mathbf{a}_i\| = 1$ für alle i). Analog kann man auch Orthonormalbasen von Untervektorräumen des \mathbb{R}^n definieren.

Beispiele.

1. Die Standardbasis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ ist eine ON-Basis des \mathbb{R}^n .
2. Die Vektoren $\mathbf{a}_1 := (1, 1)$ und $\mathbf{a}_2 := (1, -1)$ bilden ein Orthogonalsystem im \mathbb{R}^2 , sie sind aber nicht normiert.

Die Vektoren $\mathbf{b}_1 := \frac{1}{\sqrt{2}}(1, 1)$ und $\mathbf{b}_2 := \frac{1}{\sqrt{2}}(1, -1)$ bilden eine ON-Basis.

Die Überprüfung, ob ein System von Vektoren eine Basis bildet, ist bei ON-Systemen besonders einfach. Und auch die Ermittlung der Koordinaten eines Vektors bezüglich einer ON-Basis ist einfach:

Die Koordinaten bezüglich einer ON-Basis

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

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

durch die Formel

$$x_i = \mathbf{x} \bullet \mathbf{a}_i, \text{ für } i = 1, \dots, k.$$

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

Der BEWEIS ist trivial, die Aussage sehr nützlich.

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

Schmidt'sches Orthogonalisierungsverfahren

Sei $U \subset \mathbb{R}^n$ ein k -dimensionaler Untervektorraum und $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ eine Basis von U .

Dann besitzt U eine ON-Basis $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$, so daß gilt:

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

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

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

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

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

Es gilt:

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

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

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

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

Beispiel.

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

$$\mathbf{a}_1 = \frac{1}{\sqrt{2}}(1, 1, 0),$$

$$\mathbf{b}_2 = \mathbf{x}_2 - (\mathbf{x}_2 \bullet \mathbf{a}_1) \mathbf{a}_1 = \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$

Dann ist $\|\mathbf{b}_2\| = \frac{1}{\sqrt{2}}$, und wir setzen

$$\mathbf{a}_2 = \sqrt{2} \cdot \mathbf{b}_2 = \frac{1}{\sqrt{2}}(-1, 1, 0).$$

Definition:

Sei $b : V \times W \rightarrow K$ eine Dualität und $U \subset V$ ein K -Unterraum. Dann nennen wir

$$U^\perp := \{y \in W : b(x, y) = 0 \text{ für alle } x \in U\}$$

das *orthogonale Komplement* von U .

Ist $T \subset W$, so definiert man analog $T^\perp := \{x \in V : b(x, y) = 0 \text{ für alle } y \in T\}$.

Beispiele.

1. Ist $U \subset \mathbb{R}^n$ ein Untervektorraum, so besteht das orthogonale Komplement von U (bezüglich des Skalarproduktes) aus allen Vektoren, die auf U senkrecht stehen:

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

2. Ist V ein endlich-dimensionaler K -Vektorraum, $\{\lambda_1, \dots, \lambda_r\}$ ein System linear unabhängiger Linearformen auf V und T der von den λ_i erzeugte Unterraum von V^* , so ist

$$T^\perp = \{x \in V : \lambda_1(x) = \dots = \lambda_r(x) = 0\}$$

die gemeinsame Nullstellenmenge von $\lambda_1, \dots, \lambda_r$.

Ist $U \subset V$ ein Untervektorraum und $b : V \times W \rightarrow K$ eine Dualität, so ist auch $U^\perp \subset W$ ein Untervektorraum.

Satz (Eigenschaften des orthogonalen Komplements)

Sei $b : V \times W \rightarrow K$ eine Dualität.

1. Ist $U_1 \subset U_2 \subset V$, so ist $U_2^\perp \subset U_1^\perp \subset W$.
2. Ist $n = \dim(V) = \dim(W)$ und $U \subset V$, so ist $\dim(U^\perp) = n - \dim(U)$.
3. Es ist $U^{\perp\perp} = U$.

BEWEIS: 1) Sei $y \in U_2^\perp$, also $b(x, y) = 0$ für alle $x \in U_2$. Dann gilt dies natürlich erst recht für alle $x \in U_1$, d.h., y liegt in U_1^\perp .

2) Ist $U = \{0\}$, so ist $U^\perp = W$ und die Aussage trivial. Sei also $U \neq \{0\}$ und $\{a_1, \dots, a_r\}$ eine Basis von U . Man kann sie zu einer Basis $\{a_1, \dots, a_r, a_{r+1}, \dots, a_n\}$ von V ergänzen. Dann gibt es dazu eine eindeutig bestimmte duale Basis $\{b_1, \dots, b_n\}$ von W .

Sei $S := \langle b_{r+1}, \dots, b_n \rangle$. Weil $b(a_i, b_j) = 0$ für $i = 1, \dots, r$ und $j = r+1, \dots, n$ ist, ist $S \subset U^\perp$. Sei nun umgekehrt $y \in U^\perp$. Es gibt eine Darstellung $y = \sum_{j=1}^n y_j b_j$. Daraus folgt:

$$0 = b(a_i, y) = \sum_{j=1}^n y_j b(a_i, b_j) = \sum_{j=1}^n y_j \delta_{ij} = y_i \text{ für } i = 1, \dots, r,$$

also $y = \sum_{j=r+1}^n y_j b_j \in S$. Das bedeutet, daß $S = U^\perp$ ist, sowie

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

3) Wegen (2) ist $\dim(U^{\perp\perp}) = \dim(U)$. Außerdem ist $U \subset U^{\perp\perp}$. Zusammen ergibt das die Gleichheit. ■

Sei $U \subset \mathbb{R}^n$ ein Untervektorraum. Wir wollen jetzt den Begriff der *orthogonalen Projektion* auf U einführen.

Jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ kann wie folgt eindeutig zerlegt werden:

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

Wir können nämlich eine ON-Basis $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ von \mathbb{R}^n finden, so daß $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ eine ON-Basis von U und $\{\mathbf{a}_{r+1}, \dots, \mathbf{a}_n\}$ eine ON-Basis von U^\perp ist. Dann ist $\mathbf{x} = \mathbf{u} + \mathbf{v}$ mit

$$\mathbf{u} := \sum_{i=1}^r (\mathbf{x} \bullet \mathbf{a}_i) \mathbf{a}_i \quad \text{und} \quad \mathbf{v} := \sum_{j=r+1}^n (\mathbf{x} \bullet \mathbf{a}_j) \mathbf{a}_j$$

die gewünschte Zerlegung.

Nun definiert man $\text{pr}_U : \mathbb{R}^n \rightarrow U$ durch

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

Es ist also

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

Diese Abbildung ist vernünftig definiert und – wie man leicht nachrechnen kann – auch linear. Sie hat folgende Eigenschaften:

1. $\text{pr}_U : \mathbb{R}^n \rightarrow U$ ist surjektiv.
2. $\text{Ker}(\text{pr}_U) = U^\perp$.
3. $\|\mathbf{x} - \text{pr}_U(\mathbf{x})\| \leq \|\mathbf{x} - \mathbf{u}\|$ für alle $\mathbf{u} \in U$ und $\mathbf{x} \in \mathbb{R}^n$.

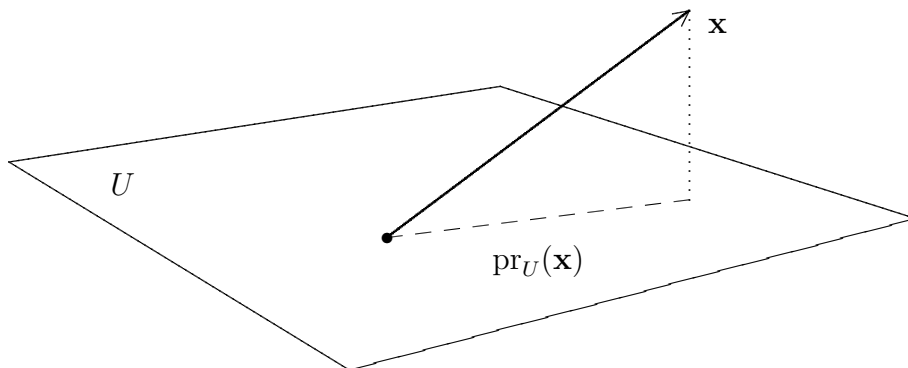
Zum BEWEIS: (1) ist klar.

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

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

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

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



Eine **Anwendung** der orthogonalen Projektion ist die näherungsweise Lösung von linearen Gleichungssystemen. Ist $f = f_A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung, so ist die Gleichung

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

nur dann lösbar, wenn $\mathbf{b} \in \text{Im}(f_A)$ ist.

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

Bei einer exakten Lösung \mathbf{x}_0 der Original-Gleichung wäre $f_A(\mathbf{x}_0) = \mathbf{b}$ und daher $\|f_A(\mathbf{x}_0) - \mathbf{b}\| = 0$. Ist das nicht erreichbar, so sucht man unter allen möglichen $\mathbf{x} \in \mathbb{R}^m$ denjenigen Vektor $\mathbf{x} = \mathbf{x}_0$, für den $\|f_A(\mathbf{x}) - \mathbf{b}\|$ minimal wird. Dann wird der Abstand von \mathbf{b} vom Unterraum $U := \text{Im}(f_A)$ genau in $\mathbf{y}_0 = f_A(\mathbf{x}_0)$ angenommen, und das ist genau dann der Fall, wenn $\mathbf{y}_0 = \text{pr}_U(\mathbf{b})$ ist. Wie man \mathbf{x}_0 nun bestimmen kann, werden wir weiter unten sehen.

Zuvor müssen wir noch ein anderes Thema behandeln. Ist $A \in M_{n,m}(K)$, so ist durch A eine lineare Abbildung $f_A : K^m \rightarrow K^n$ bestimmt.

- Verwenden wir die **Spaltenschreibweise**, so ist

$$f_A(\vec{x}) := A \cdot \vec{x} \quad (\text{oder – anders geschrieben – } f_A(\mathbf{x}^t) = A \cdot \mathbf{x}^t).$$

- Verwenden wir dagegen für alle Vektoren die **Zeilenschreibweise**, so müssen wir das Ergebnis noch transponieren, um wieder eine Zeile zu erhalten:

$$f_A(\mathbf{x}) = (A \cdot \mathbf{x}^t)^t = \mathbf{x} \cdot A^t.$$

Definition:

Ist $A \in M_{n,m}(K)$ und $f = f_A : K^m \rightarrow K^n$ die zugehörige lineare Abbildung, so heißt die durch $f^*(\mathbf{y}) := \mathbf{y} \cdot A$ (bzw. $\vec{y} \mapsto A^t \cdot \vec{y}$) definierte lineare Abbildung $f^* : K^n \rightarrow K^m$ die zu f *konjugierte Abbildung*.

Die konjugierte Abbildung f^* ist entgegengesetzt zu f gerichtet, und es ist

$$(f_A)^* = f_{A^t}.$$

Wir bezeichnen jetzt die kanonischen Dualitäten auf K^n bzw. K^m beide mit b . Für $\mathbf{x}, \mathbf{y} \in K^n$ ist dann $b(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}^t$. Orthogonale Komplemente sollen sich im Folgenden auf diese Dualität beziehen. Ist $U \subset K^n$ ein Unterraum, so ist

$$U^\perp = \{\mathbf{x} \in K^n : b(\mathbf{x}, \mathbf{y}) = 0 \text{ für alle } \mathbf{y} \in U\}$$

ebenfalls ein Unterraum von K^n .

Insbesondere ist

$$b(\mathbf{x}, f(\mathbf{y})) = \mathbf{x} \cdot (\mathbf{y} \cdot A^t)^t = \mathbf{x} \cdot (A \cdot \mathbf{y}^t) = (\mathbf{x} \cdot A) \cdot \mathbf{y}^t = b(f^*(\mathbf{x}), \mathbf{y}), \text{ für } \mathbf{x} \in K^n, \mathbf{y} \in K^m.$$

Weil b symmetrisch ist, ist genauso

$$b(f(\mathbf{y}), \mathbf{x}) = b(\mathbf{y}, f^*(\mathbf{x})).$$

Satz (Eigenschaften der konjugierten Abbildung)

Sei $A \in M_{n,m}(K)$, $f = f_A : K^m \rightarrow K^n$ und $f^* : K^n \rightarrow K^m$ die zu f konjugierte Abbildung. Dann gilt:

1. $\text{Ker}(f^*) = \text{Im}(f)^\perp$.
2. $\text{rg}(A) = \text{rg}(A^t)$ (Spaltenrang = Zeilenrang).

BEWEIS: 1) Sei $\mathbf{y} \in \text{Ker}(f^*)$. Dann ist $f^*(\mathbf{y}) = \mathbf{0}$.

Ist $\mathbf{z} \in \text{Im}(f)$ beliebig, so ist $\mathbf{z} = f(\mathbf{x})$, für ein $\mathbf{x} \in K^m$, und

$$b(\mathbf{z}, \mathbf{y}) = b(f(\mathbf{x}), \mathbf{y}) = b(\mathbf{x}, f^*(\mathbf{y})) = b(\mathbf{x}, \mathbf{0}) = 0.$$

Also liegt \mathbf{y} in $\text{Im}(f)^\perp$, und es ist $\text{Ker}(f^*) \subset \text{Im}(f)^\perp$. Da alle Schlüsse umkehrbar sind, gilt sogar die Gleichheit.

2) Es ist

$$\text{rg}(A^t) = \dim \text{Im}(f^*) = n - \dim(\text{Ker } f^*) = n - \dim \text{Im}(f)^\perp = \dim \text{Im}(f) = \text{rg}(A).$$

Die Zahl $\text{rg}(A)$ gibt die Anzahl der linear unabhängigen Spalten von A an. Da die Zeilen von A die Spalten von A^t sind, hat A auch $\text{rg}(A)$ linear unabhängige Zeilen.

■

Wir kommen jetzt noch einmal auf die näherungsweise Lösung von Linearen Gleichungssystemen zu sprechen.

Es sei \mathbf{x}_0 die beste Näherungslösung des LGS $f_A(\mathbf{x}) = \mathbf{b}$. Wir haben gesehen, daß dann $f_A(\mathbf{x}_0) = \text{pr}_U(\mathbf{b})$ ist, mit $U = \text{Im}(f_A)$.

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

$$\text{pr}_U(f_A(\mathbf{x}_0) - \mathbf{b}) = \text{pr}_U(\mathbf{b}) - \text{pr}_U(\mathbf{b}) = \mathbf{0},$$

also

$$f_A(\mathbf{x}_0) - \mathbf{b} \in \text{Ker}(\text{pr}_U) = (\text{Im}(f_A))^\perp = \text{Ker}(f_{A^t}).$$

Daraus folgt: \mathbf{x}_0 ist genau dann die beste Näherungslösung der Gleichung $A \cdot \mathbf{x}^t = \mathbf{b}^t$, wenn $A^t \cdot (A \cdot \mathbf{x}_0^t - \mathbf{b}^t) = \mathbf{0}^t$ ist, also \mathbf{x}_0 eine exakte Lösung der sogenannten „Normal-Gleichung“

$$(A^t \cdot A) \cdot \mathbf{x}^t = A^t \cdot \mathbf{b}.$$

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

Wir wollen jetzt noch das euklidische Skalarprodukt verallgemeinern. Auf dem \mathbb{C}^n haben wir zwar die kanonische Dualität

$$b(\mathbf{z}, \mathbf{w}) = z_1 w_1 + \cdots + z_n w_n,$$

aber das ist kein Skalarprodukt. Das Ergebnis kann komplex sein, und selbst wenn $b(\mathbf{z}, \mathbf{z})$ reell ist, so braucht diese Zahl nicht positiv zu sein. Mit einer Bilinearform auf dem \mathbb{C}^n ist das nicht zu erreichen. Deshalb müssen wir über \mathbb{C} mit einem anderen Begriff arbeiten.

Definition:

Es sei V ein \mathbb{C} -Vektorraum. Eine *hermitesche Form* auf V ist eine Abbildung $h : V \times V \rightarrow \mathbb{C}$ mit folgenden Eigenschaften:

1. h ist \mathbb{R} -bilinear.
2. $h(\alpha v, w) = \alpha \cdot h(v, w)$ für $v, w \in V$ und $\alpha \in \mathbb{C}$.
3. Es ist $h(v, w) = \overline{h(w, v)}$ für alle $v, w \in V$.

Behauptung: Ist $h : V \times V \rightarrow \mathbb{C}$ eine hermitesche Form, so gilt:

1. $h(v, \alpha w) = \overline{\alpha} \cdot h(v, w)$ für $v, w \in V$ und $\alpha \in \mathbb{C}$.
2. Für $v \in V$ ist $h(v, v)$ reell.

BEWEIS: Es ist $h(v, \alpha w) = \overline{h(\alpha w, v)} = \overline{\alpha \cdot h(w, v)} = \overline{\alpha} \cdot \overline{h(w, v)} = \overline{\alpha} \cdot h(v, w)$, und es ist $\overline{h(v, v)} = h(v, v)$, also $h(v, v)$ reell. ■

Definition:

Eine hermitesche Form $h : V \times V \rightarrow \mathbb{C}$ heißt *positiv definit* (oder ein *hermitesches Skalarprodukt*), falls $h(v, v) > 0$ für jedes $v \neq 0$ ist.

Beispiel.

Das kanonische hermitesche Skalarprodukt auf dem \mathbb{C}^n ist definiert durch

$$\langle \mathbf{z}, \mathbf{w} \rangle := z_1 \overline{w_1} + \cdots + z_n \overline{w_n}.$$

Ist $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{C}^n$ und $z_\nu = x_\nu + j y_\nu$, so ist

$$\langle \mathbf{z}, \mathbf{z} \rangle = z_1 \overline{z_1} + \cdots + z_n \overline{z_n} = (x_1)^2 + (y_1)^2 + \cdots + (x_n)^2 + (y_n)^2.$$

Also ist $\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle}$ die gewöhnliche euklidische Norm.

Die Begriffe „Orthogonalität“, „Orthonormalisierung“ und „orthogonales Komplement“ übertragen sich sinngemäß.

Jede hermitesche Form h auf dem \mathbb{C}^n hat die Gestalt

$$h(\mathbf{v}, \mathbf{w}) = h_B(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot B \cdot \overline{\mathbf{w}}^t,$$

mit einer Matrix $B \in M_{n,n}(\mathbb{C})$. Die Einträge b_{ij} in der Matrix $B = (b_{ij})$ gewinnt man durch $b_{ij} = h(\mathbf{e}_i, \mathbf{e}_j)$. Daß h hermitesch ist, ist gleichbedeutend damit, daß $\overline{B}^t = B$ ist.

h ist ein Skalarprodukt, falls $\mathbf{z} \cdot B \cdot \overline{\mathbf{z}}^t > 0$ ist, für alle $\mathbf{z} \neq \mathbf{0}$.

Wie man diese Eigenschaft mit einem handlichen Kriterium nachprüft, können wir im Augenblick noch nicht sagen. Die Theorie der Determinanten wird uns dabei später weiterhelfen.