15. Lineare Gleichungssysteme und Matrizen

Wie wir nun schon in einigen Fällen gesehen haben, laufen viele Fragestellungen der linearen Algebra auf rechnerischer Seite am Ende auf *lineare Gleichungssysteme* hinaus – z. B. die Überprüfung von Erzeugendensystemen und der linearen Unabhängigkeit in Beispiel 14.2 (b) bzw. Beispiel 14.5 (b), oder die Berechnung von Durchschnitten und Summen von Unterräumen in Algorithmus 14.27. Mit anderen Worten müssen wir also zu gegebenen $m,n \in \mathbb{N}$ sowie $a_{i,j}, b_i \in K$ für $i = 1,\dots,m$ und $j = 1,\dots,n$ alle $x_1,\dots,x_n \in K$ bestimmen, die simultan die Gleichungen

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m \end{aligned} \tag{$*$}$$

erfüllen. Wir wollen daher in diesem Kapitel studieren, wann solche linearen Gleichungssysteme lösbar sind, und wie im Fall der Lösbarkeit die Lösungsmengen aussehen und konkret berechnet werden können.

15.A Matrizen

Bevor wir mit der eigentlichen Untersuchung linearer Gleichungssysteme beginnen, sollten wir als Erstes eine effizientere Notation dafür einführen, da die Schreibweise $(*)$ in der Einleitung oben natürlich sehr unübersichtlich und fehleranfällig ist. Dies ist mit Hilfe von sogenannten *Matrizen* möglich. Die Idee besteht dabei einfach darin, die $m \cdot n$ Zahlen $a_{i,j}$ in $(*)$ in einem einzigen mathematischen Objekt zusammenzufassen.

Definition 15.1 (Matrizen)**.** Es seien $m,n \in \mathbb{N}$.

(a) Eine $m \times n$-**Matrix** über K ist ein rechteckiges Schema

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \quad \text{mit } a_{i,j} \in K \text{ für alle } 1 \leq i \leq m \text{ und } 1 \leq j \leq n$$

mit m Zeilen und n Spalten. Analog zur Schreibweise für Zahlenfolgen in Definition **?? ??** bezeichnen wir eine solche Matrix auch mit

$$\big(a_{i,j}\big)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \quad \text{oder einfach} \quad (a_{i,j})_{i,j}.$$

Es ist eine Konvention, dass wir dabei hinter den Klammern immer zuerst den Zeilenindex und dann den Spaltenindex angeben (Merkregel: „Zeile zuerst, Spalte später“).

Die Menge aller $m \times n$-Matrizen mit Einträgen in K wird mit $K^{m \times n}$ bezeichnet – auch hier steht in der Bezeichnung also zuerst die Anzahl der Zeilen und dann die Anzahl der Spalten.

(b) Eine Matrix $A = (a_{i,j})_{i,j} \in K^{m \times n}$ mit genauso vielen Zeilen wie Spalten (also $m = n$) heißt **quadratisch**. In diesem Fall bezeichnet man die Einträge $a_{i,i}$ mit $i = 1,\dots,n$ (also die Einträge von links oben nach rechts unten) als die **Diagonaleinträge** der Matrix.

Definition 15.2 (Matrixoperationen)**.** Für zwei Matrizen $A = (a_{i,j})_{i,j}$ und $B = (b_{i,j})_{i,j}$ in $K^{m \times n}$ sowie $\lambda \in K$ definieren wir

(a) die Addition $A + B := (a_{i,j} + b_{i,j})_{i,j} \in K^{m \times n}$;

(b) die Skalarmultiplikation $\lambda A := \lambda \cdot A := (\lambda \cdot a_{i,j})_{i,j} \in K^{m \times n}$;

(c) die **transponierte Matrix** $A^\mathsf{T} := (a_{j,i})_{i,j} \in K^{n\times m}$.

Beispiel 15.3. Die reelle Matrix (z. B. der Größe 2×3), bei der in jeder Zeile i alle Einträge gleich i sind, können wir schreiben als

$$(i)_{i,j} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} \quad \in \mathbb{R}^{2\times 3}.$$

Addition und Skalarmultiplikation für Matrizen (und damit auch für Vektoren) sind einfach komponentenweise definiert, es ist also z. B.

$$\begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{pmatrix} \quad \text{und} \quad 2 \cdot \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 4 \\ 6 & 8 & 10 \end{pmatrix}$$

in $\mathbb{R}^{2\times 3}$. Die Transposition hingegen vertauscht die Rolle von Zeilen und Spalten in der Matrix, wie z. B. in

$$\begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix}^\mathsf{T} = \begin{pmatrix} 0 & 3 \\ 1 & 4 \\ 2 & 5 \end{pmatrix} \quad \in \mathbb{R}^{3\times 2}.$$

Bemerkung 15.4.

(a) Offensichtlich ist $K^{m\times n}$ mit der Addition und Skalarmultiplikation aus Definition 15.2 ein K-Vektorraum. In der Tat unterscheidet sich dieser Raum von K^{mn} ja nur dadurch, dass wir die $m \cdot n$ Einträge der Matrix nicht untereinander, sondern in einem rechteckigen Schema anordnen. Dementsprechend erhält man also auch eine Basis von $K^{m\times n}$, indem man alle Matrizen mit einem Eintrag 1 und allen anderen Einträgen 0 nimmt, im Fall $K^{2\times 2}$ also z. B.

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Insbesondere gilt damit $\dim K^{m\times n} = mn$.

(b) Der Nullvektor im Vektorraum $K^{m\times n}$ ist offensichtlich die Matrix, in der alle Einträge gleich 0 sind. Diese Matrix wird dementsprechend auch die **Nullmatrix** genannt und einfach als 0 geschrieben.

(c) Matrizen in $K^{m\times 1}$ mit nur einer Spalte haben in ihrer Schreibweise die gleiche Form wie Vektoren in K^m. In der Tat werden wir $m\times 1$-Matrizen im Folgenden in der Regel mit Vektoren in K^m identifizieren.

Bisher gibt es bis auf die Art der Anordnung der Zahlen keinen nennenswerten Unterschied zwischen den Matrizen in $K^{m\times n}$ und den Vektoren in K^{mn}. Es gibt jedoch eine sehr wichtige weitere Operation, die auf Matrizen, jedoch nicht auf Vektoren in K^{mn} definiert ist, nämlich die sogenannte *Matrixmultiplikation*:

Definition 15.5 (Matrixmultiplikation)**.** Für natürliche Zahlen m, n, p seien $A = (a_{i,j})_{i,j} \in K^{m\times n}$ und $B = (b_{j,k})_{j,k} \in K^{n\times p}$, d. h. die Matrix B habe so viele Zeilen wie A Spalten. Dann definieren wir das Matrixprodukt AB als

$$AB := A \cdot B := \left(\sum_{j=1}^{n} a_{i,j} b_{j,k} \right)_{i,k} \quad \in K^{m\times p}$$

(merke: Es wird über die „mittleren Indizes“ summiert, also über den Spaltenindex der ersten und den Zeilenindex der zweiten Matrix). Das Produkt AB hat also so viele Zeilen wie die erste Matrix und so viele Spalten wie die zweite.

Beispiel 15.6.

(a) Es ist z. B.

$$\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 0\cdot 4+1\cdot 7 & 0\cdot 5+1\cdot 8 & 0\cdot 6+1\cdot 9 \\ 2\cdot 4+3\cdot 7 & 2\cdot 5+3\cdot 8 & 2\cdot 6+3\cdot 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 29 & 34 & 39 \end{pmatrix}$$

(hier ist $m = n = 2$ und $p = 3$). Im Gegensatz dazu ist das Matrixprodukt

$$\begin{pmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$$

nicht definiert, weil die erste Matrix drei Spalten, die zweite aber nur zwei Zeilen hat. Der Einfachheit halber werden wir in Zukunft bei einem Matrixprodukt AB stets voraussetzen, dass die zweite Matrix so viele Zeilen hat wie die erste Spalten, und dies nicht jedes Mal wieder erwähnen.

(b) Sind $A = (a_{i,j})_{i,j} \in K^{m \times n}$ und $x \in K^n$ mit Koordinaten $x_1, \dots, x_n$, so können wir x gemäß Bemerkung 15.4 (c) als Matrix mit nur einer Spalte auffassen und erhalten das Matrixprodukt

$$Ax = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{pmatrix} \in K^{m \times 1} = K^m.$$

Beachte, dass dieser Ausdruck genau die linke Seite eines linearen Gleichungssystems wie in der Einleitung zu diesem Kapitel ist. Wir können lineare Gleichungssysteme in Zukunft also einfach als $Ax = b$ schreiben, wobei $A \in K^{m \times n}$ eine gegebene Matrix, $b \in K^m$ ein gegebener Vektor und $x \in K^n$ der gesuchte Vektor sind.

(c) Ein einfacher, aber oft vorkommender und daher wichtiger Spezialfall von (b) ist, wenn $x = e_j$ für $j = 1, \dots, n$ der j-te Einheitsvektor ist: In diesem Fall ist Ae_j gerade die j-te Spalte von A.

Wie üblich nach dem Einführen einer neuen Struktur wollen wir auch hier zunächst einmal die grundlegenden Eigenschaften der Matrixmultiplikation angeben bzw. beweisen.

Lemma 15.7 (Eigenschaften der Matrixmultiplikation)**.** *Für alle Matrizen A, B, C passender Größe (d. h. so dass die betrachteten Summen und Produkte definiert sind) sowie $\lambda \in K$ gilt:*

(a) *(Distributivität)* $(A+B)C = AC + BC$ *und* $A(B+C) = AB + AC$.

(b) *(Verträglichkeit mit der Skalarmultiplikation)* $(\lambda A)B = A(\lambda B) = \lambda(AB)$.

(c) *(Assoziativität)* $(AB)C = A(BC)$.

Bei der Multiplikation mehrerer Matrizen werden wir die Klammern daher oft weglassen; das n-fache Matrixprodukt $A \cdots A$ für $n \in \mathbb{N}_{>0}$ schreiben wir als A^n.

(d) *(Verträglichkeit mit der Transposition)* $(AB)^\mathsf{T} = B^\mathsf{T} A^\mathsf{T}$.

Das Matrixprodukt ist jedoch im Allgemeinen nicht kommutativ (aufgrund der Größenbedingung ist das Produkt AB ja auch nicht einmal genau dann definiert, wenn BA es ist).

Beweis. Der Beweis ergibt sich durch einfaches Nachrechnen. Wir zeigen exemplarisch den zweiten Teil von (a): Für $A = (a_{i,j})_{i,j} \in K^{m \times n}$ und $B = (b_{j,k})_{j,k}, C = (c_{j,k})_{j,k} \in K^{n \times p}$ gilt

$$\begin{aligned} A(B+C) &= \Big(\sum_{j=1}^n a_{i,j}(b_{j,k} + c_{j,k}) \Big)_{i,k} \\ &= \Big(\sum_{j=1}^n a_{i,j}b_{j,k} + \sum_{j=1}^n a_{i,j}c_{j,k} \Big)_{i,k} \\ &= \Big(\sum_{j=1}^n a_{i,j}b_{j,k} \Big)_{i,k} + \Big(\sum_{j=1}^n a_{i,j}c_{j,k} \Big)_{i,k} \\ &= AB + AC. \end{aligned}$$

□

Bemerkung 15.8 (**Blockmatrixmultiplikation**)**.** Es seien $A \in K^{m \times n}$ und $B \in K^{n \times p}$ zwei Matrizen, die in „Blockform“

$$A = \left(\begin{array}{c|c} A^{(1,1)} & A^{(1,2)} \\ \hline A^{(2,1)} & A^{(2,2)} \end{array} \right) \quad \text{bzw.} \quad B = \left(\begin{array}{c|c} B^{(1,1)} & B^{(1,2)} \\ \hline B^{(2,1)} & B^{(2,2)} \end{array} \right)$$

mit $A^{(1,1)} \in K^{m_1 \times n_1}$ und $B^{(1,1)} \in K^{n_1 \times p_1}$ gegeben sind. Nach Definition der Matrixmultiplikation können wir das Produkt AB dann ebenfalls in Blockform

$$AB = \left(\begin{array}{c|c} A^{(1,1)}B^{(1,1)} + A^{(1,2)}B^{(2,1)} & A^{(1,1)}B^{(1,2)} + A^{(1,2)}B^{(2,2)} \\ \hline A^{(2,1)}B^{(1,1)} + A^{(2,2)}B^{(2,1)} & A^{(2,1)}B^{(1,2)} + A^{(2,2)}B^{(2,2)} \end{array}\right)$$

schreiben, wobei die Blöcke formal genauso multipliziert und addiert werden, als würde man das Produkt zweier Matrizen der Größe 2×2 berechnen: Ist $A^{(r,s)} = \left(a_{i,j}^{(r,s)}\right)_{i,j}$ und $B^{(r,s)} = \left(b_{j,k}^{(r,s)}\right)_{j,k}$ für $r,s \in \{1,2\}$, so ist z. B. der Eintrag von AB in Zeile $i \leq m_1$ und Spalte $k \leq p_1$ (also im Block links oben) gegeben durch

$$\sum_{j=1}^{n_1} a_{i,j}^{(1,1)} b_{j,k}^{(1,1)} + \sum_{j=1}^{n-n_1} a_{i,j}^{(1,2)} b_{j,k}^{(2,1)},$$

was genau der Eintrag von $A^{(1,1)}B^{(1,1)} + A^{(1,2)}B^{(2,1)}$ an dieser Stelle ist.

Analog funktioniert diese Regel auch für eine Aufteilung in mehr oder weniger als zwei Blöcke entlang der Zeilen bzw. Spalten. Häufig kommt z. B. der Fall vor, in dem wir die Matrix B spaltenweise als $B = \left(b^{(1)} \,|\, \cdots \,|\, b^{(p)}\right)$ mit $b^{(1)}, \ldots, b^{(p)} \in K^n$ schreiben; in diesem Fall ist dann

$$AB = A\left(b^{(1)} \,|\, \cdots \,|\, b^{(p)}\right) = \left(Ab^{(1)} \,|\, \cdots \,|\, Ab^{(p)}\right).$$

Wir werden diese Rechenregeln zur Blockmatrixmultiplikation im Folgenden oft verwenden, ohne jedesmal wieder darauf hinzuweisen.

Aufgabe 15.9.

(a) Finde eine Matrix $A \in \mathbb{R}^{2\times 2} \backslash \{0\}$ mit $A^2 := A \cdot A = 0$.

(b) Finde Matrizen $A, B \in \mathbb{R}^{2\times 2}$ mit $AB = 0$ und $BA \neq 0$.

Mit Hilfe des Matrixprodukts können wir nun zu jeder Matrix $A \in K^{m\times n}$ zwei Unterräume definieren; einen von K^m und einen von K^n:

Definition 15.10 (Bild und Kern einer Matrix)**.** Für eine Matrix $A \in K^{m\times n}$ mit $m,n \in \mathbb{N}$ nennen wir

(a) $\mathrm{Im}A := \{Ax : x \in K^n\} \subset K^m$ das **Bild** von A;

(b) $\mathrm{Ker}A := \{x \in K^n : Ax = 0\} \subset K^n$ den **Kern** von A.

Die Bezeichnungen kommen dabei von den englischen Worten „image“ und „kernel“.

Lemma und Definition 15.11 (Rang einer Matrix)**.** *Es sei $A \in K^{m\times n}$ mit $m,n \in \mathbb{N}$.*

(a) $\mathrm{Im}A$ *ist der von den Spalten von A erzeugte Unterraum von K^m.*

Man nennt seine Dimension $\mathrm{rk}A := \dim \mathrm{Im}A$ *den* ***Rang*** *von A; die Bezeichnung kommt dabei vom englischen Wort „rank“.*

(b) $\mathrm{Ker}A$ *ist ein Unterraum von K^n.*

Beweis.

(a) Schreiben wir $A = (a_1 \,|\, \cdots \,|\, a_n)$ mit $a_1, \ldots, a_n \in K^m$, so ist nach Definition des Matrixprodukts

$$\mathrm{Im}A = \{Ax : x \in K^n\} = \{x_1a_1 + \cdots + x_na_n : x_1, \ldots, x_n \in K\} = \mathrm{Lin}(a_1, \ldots, a_n).$$

Also ist $\mathrm{Im}A$ der von den Spalten von A erzeugte Unterraum.

(b) Wir überprüfen die Unterraumkriterien aus Definition 13.8: Zunächst ist natürlich $A \cdot 0 = 0$, also $0 \in \mathrm{Ker}A$. Für alle $x, y \in \mathrm{Ker}A$ und $\lambda \in K$ gilt außerdem $Ax = Ay = 0$, also auch

$$A(x+y) = Ax + Ay = 0 \quad \text{und} \quad A \cdot \lambda x = \lambda \cdot Ax = 0,$$

und damit $x + y \in \mathrm{Ker}A$ und $\lambda x \in \mathrm{Ker}A$. □

33