3.3 Das Rechnen mit Vektoren und Matrizen

Matrizen haben wir bisher nur benutzt, um die Zeilenoperationen beim Lösen linearer Gleichungssysteme übersichtlich darstellen zu können. Man kann aber auch sinnvolle Rechenoperationen für Matrizen und Vektoren definieren und dafür einen leistungsfähigen Kalkül entwickeln, der sich zur Lösung linearer Gleichungssysteme einsetzen lässt. Darum geht es in diesem Abschnitt. Dabei werden alle Rechenoperationen und Rechenregeln letztlich auf das Rechnen mit reellen Zahlen, mit den Einträgen der Matrizen, zurückgeführt. Der hauptsächliche Vorteil des Kalküls ist, dass man damit viele gleichartige Rechnungen / Gleichungen mit reellen Zahlen übersichtlich zu einer einzigen Rechnung / Gleichung für Matrizen oder Vektoren zusammenfassen kann. Daher wird die Matrix- und Vektorrechnung auch für die Modellierung komplexer (aber immer noch "linearer") ökonomischer Abläufe eingesetzt.

DEFINITION: (i) Die **Summe zweier Matrizen** ist positionsweise definiert, wenn beide dasselbe Format haben, d.h. man addiert die Einträge beider Matrizen in gleicher Position und trägt die Summen in die entsprechenden Positionen der Ergebnismatrix ein. Sind also $A = (a_{ij})$ und $B = (b_{ij})$ Matrizen desselben Formats (m, n), so ist ihre Summe die Matrix $C = A + B$ desselben Formats (m, n), die Einträge $c_{ij} = a_{ij} + b_{ij}$ hat. Analog ist die **Differenz zweier Matrizen** $A = (a_{ij})$, $B = (b_{ij})$ von gleichem Format (m, n) die Matrix $D = A - B$ desselben Formats (m, n), die Einträge $d_{ij} = a_{ij} - b_{ij}$ hat.

(ii) Die **Vervielfachung einer Matrix mit einem Faktor** $r \in \mathbb{R}$, auch **Multiplikation der Matrix mit einem Skalar** $r \in \mathbb{R}$ genannt (weil die Zahlen aus dem zu Grunde liegenden Zahlbereich in der Vektor- und Matrixrechnung auch als *Skalare* bezeichnet werden), ist definiert durch Multiplikation eines jeden Eintrags der Matrix A mit der Zahl r. Das Produkt rA ist also die Matrix desselben Formats wie $A = (a_{ij})$, welche die Einträge ra_{ij} hat. Die Matrix $(-1)A = (-a_{ij})$ heißt **das Negative der Matrix** A und wird $-A$ notiert. ■

DISKUSSION: 1) Man kann also nicht beliebige Matrizen addieren und subtrahieren:

- *Summe und Differenz sind nur für Matrizen gleichen Formats erklärt, und das Ergebnis ist wieder eine Matrix von diesem Format.*

Ebenso haben Vielfache rA einer Matrix A mit Faktoren $r \in \mathbb{R}$ dasselbe Format wie A. Wir schreiben skalare Faktoren meistens links vor die Matrizen bzw. die Zeilen- oder Spaltenvektoren. Das muss man nicht so halten; Ar ist dann dasselbe wir rA.

2) Ausführlicher geschrieben gehen die Rechenoperationen so:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & b_{ij} & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & a_{ij}+b_{ij} & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

$$r \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ra_{11} & \cdots & ra_{1n} \\ \vdots & ra_{ij} & \vdots \\ ra_{m1} & \cdots & ra_{mn} \end{pmatrix}$$

und entsprechend natürlich für die Differenzbildung (überall "$-$" statt "$+$") und für die Bildung des Negativen (überall "$-$" statt "r").

3) Für **Zeilen-** bzw. **Spaltenvektoren** sind diese Rechenoperationen natürlich auch erklärt, weil das ja spezielle Matrizen sind (mit nur einer Zeile bzw. nur einer Spalte). Dafür gilt also:

$$(a_1\ a_2 \ldots a_n) \pm (b_1\ b_2 \ldots b_n) = (a_1{\pm}b_1\ a_2{\pm}b_2 \ldots a_n{\pm}b_n)\,,$$

$$r(a_1\ a_2 \ldots a_n) = (ra_1\ ra_2 \ldots ra_n)\,,$$

bzw.

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \pm \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} a_1 \pm b_1 \\ a_2 \pm b_2 \\ \vdots \\ a_m \pm b_m \end{pmatrix}, \qquad r \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} ra_1 \\ ra_2 \\ \vdots \\ ra_m \end{pmatrix}.$$

Den Unterschied zwischen Zeilen- und Spaltenvektoren muss man nicht zu genau nehmen; schließlich enthält ein Zeilenvektor genau dieselbe Information wie der Spaltenvektor, der dieselben Einträge (in gleicher Reihenfolge) hat. Die Summe und Differenz ist aber nur für Vektoren gleichen Formats erklärt (also gleiche Komponentenzahl und beides Zeilenvektoren oder beides Spaltenvektoren), und das Ergebnis der Rechenoperationen ist wieder ein Vektor dieses Formats. Wir erinnern daran, dass wir statt Spaltenvektoren auch m-tupel schreiben, deren Einträge durch Kommas abgetrennt sind. Dafür sehen die Rechenoperationen dann aus wie bei Zeilenvektoren: $(a_1, a_2, \ldots, a_m) \pm (b_1, b_2, \ldots, b_m) = (a_1{\pm}b_1, a_2{\pm}b_2, \ldots, a_m{\pm}b_m)$ und $r(a_1, a_2, \ldots, a_m) = (ra_1, ra_2, \ldots, ra_m)$.

4) Natürlich kann man analog auch mehr als zwei Matrizen $A_1, A_2, \ldots, A_l$ addieren, wenn alle Summanden dasselbe Format haben. Und vor der Addition kann man sie noch mit reellen Faktoren $r_1, r_2, \ldots, r_l$ multiplizieren. So erhält man die sog. **Linearkombination der Matrizen A_k mit Koeffizienten r_k**, notiert

$$r_1A_1 + r_2A_2 + \ldots + r_lA_l = \sum_{k=1}^{l} r_kA_k\,.$$

Dies ist eine Matrix desselben Formats wie alle A_k, und sie hat in Zeile i und Spalte j den Eintrag $r_1a_{ij}^{(1)} + r_2a_{ij}^{(2)} + \ldots + r_la_{ij}^{(l)}$, wenn A_k die Einträge $a_{ij}^{(k)}$ besitzt für $k = 1 \ldots l$. Man bildet also auch Linearkombinationen positionsweise. Besonders oft kommen **Linearkombinationen von Vektoren** gleichen Formats vor, z.B. ist

$$r\mathbf{a} + s\mathbf{b} + t\mathbf{c} = r\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} + s\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} + t\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} ra_1 + sb_1 + tc_1 \\ ra_2 + sb_2 + tc_2 \\ \vdots \\ ra_m + sb_m + tc_m \end{pmatrix}$$

eine Linearkombination von drei Vektoren **a**, **b**, **c** aus $\mathbb{R}^m$. (Das Ergebnis ist wieder ein Spaltenvektor mit m Komponenten, auch wenn er recht "breit" geraten ist.) Lässt man die Parameter r, s, t hier beliebige reelle Zahlen durchlaufen, so bildet die entsprechende Menge aller Linearkominationen der drei Vektoren das, was wir in 3.2 eine 3-parametrige lineare Schar in $\mathbb{R}^m$ genannt hatten. Entsprechend bildet die Gesamtheit der Linearkombinationen $r_1\mathbf{a}_1 + \ldots + r_l\mathbf{a}_l$ von l Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_l \in \mathbb{R}^m$ eine l-parametrige lineare Schar in $\mathbb{R}^m$ (mit der speziellen Eigenschaft, dass sie den m-gliedrigen Nullvektor enthält; denn der entsteht, wenn man alle Koeffizienten $r_k = 0$ setzt). ■

BEISPIELE (*zur Matrix- / Vektoraddition und Multiplikation mit Skalaren*):

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

2) Es ist
$$(1 \;\; 2) + (-1 \;\; 1) + (0 \;\; -3) \; = \; 0$$
und
$$0\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0\,,$$
aber man darf *nicht* schreiben
$$(1 \;\; 2) + (-1 \;\; 1) + (0 \;\; -3) \; = \; 0 \; = \; 0\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad (???)\,;$$
denn die erste "0" steht für die 1×2–Nullmatrix (Zeilenvektor), die zweite aber für die 2×1–Nullmatrix (Spaltenvektor). *Nullmatrizen von verschiedenem Format sind verschieden* — wie generell Matrizen verschiedenen Formats! Wenn nötig, so kann man das Format einer Nullmatrix als Subskript angeben, etwa hier $0_{1\times 2}$, $0_{2\times 1}$ und allgemein $0_{m\times n}$ bei der $m\times n$–Nullmatrix.

3) Einige Linearkombinationen:
$$2(1 \;\; -1) - \frac{1}{2}(0 \;\; 4) + 3(2 \;\; 0) \; = \; (8 \;\; -4)\,,$$
$$\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} - \frac{1}{6}\begin{pmatrix} 0 & -1 \\ 1 & 6 \end{pmatrix} \; = \; 0 \; = \; \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$
$$x_1\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_m\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \; = \; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R}^m\,.$$
Die letzte Rechnung zeigt:

- *Jeder Vektor in $x \in \mathbb{R}^m$ ist eine Linearkombination der kanonischen Basisvektoren $e_1, \ldots, e_m$ von $\mathbb{R}^m$; die Koeffizienten sind dabei die Komponenten x_i von x und eindeutig bestimmt,*
$$x \; = \; \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \; = \; \sum_{i=1}^{m} x_i e_i\,.$$

4) Die Linearkombination von l Vektoren / Matrizen mit gleichen Koeffizienten $r_k = \frac{1}{l}$ für $k = 1 \ldots l$ nennt man ihr **arithmetisches Mittel.** Zum Beispiel sind
$$\frac{1}{2}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \; = \; \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \; = \; \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
und
$$\frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \; = \; \begin{pmatrix} 2/3 \\ 1/3 \\ -1/3 \end{pmatrix} \; = \; \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$
arithmetische Mittel von zwei bzw. drei Vektoren aus $\mathbb{R}^3$.

Die geometrische Bedeutung des arithmetischen Mittels von zwei Vektoren $a, b \in \mathbb{R}^m$ ist der *Mittelpunkt* der Verbindungsstrecke von a nach b in $\mathbb{R}^m$. Das arithmetische Mittel von drei Vektoren in $\mathbb{R}^m$ ist der *Schwerpunkt* des Dreiecks, das diese Vektoren als Eckpunkte hat. Das arithmetische Mittel der kanonischen Basisvektoren in $\mathbb{R}^m$ ist

$$\frac{1}{m}\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} + \frac{1}{m}\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix} + \ldots + \frac{1}{m}\begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix} = \begin{pmatrix}1/m\\1/m\\1/m\\\vdots\\1/m\end{pmatrix} = \frac{1}{m}\begin{pmatrix}1\\1\\1\\\vdots\\1\end{pmatrix} \in \mathbb{R}^m .$$

5) Wir betrachten eine *ökonomische Anwendung* (formaler Art, d.h. ökonomische Probleme werden nicht gelöst): In einem Unternehmen mit drei Teilbetrieben wird für jedes Quartal eines Jahres die **Verflechtungsmatrix** aufgestellt, welche die Leistungen der liefernden Stellen (Spalten) an die empfangenden Stellen (Zeilen) in Leistungseinheiten angibt. Die vier Verflechtungsmatrizen für die vier Quartale seien z.B.

$$A_1 = \begin{pmatrix}0&20&0\\30&0&10\\10&10&0\end{pmatrix}, \quad A_2 = \begin{pmatrix}0&15&0\\25&0&5\\10&5&0\end{pmatrix}, \quad A_3 = \begin{pmatrix}0&20&0\\25&0&15\\15&10&0\end{pmatrix}, \quad A_4 = \begin{pmatrix}0&25&0\\40&0&10\\5&15&0\end{pmatrix}.$$

Dann gibt die Summe der vier Matrizen

$$A_1 + A_2 + A_3 + A_4 = \begin{pmatrix}0&80&0\\120&0&40\\40&40&0\end{pmatrix}$$

die Verflechtungsmatrix für das ganze Jahr an und z.B. die Differenz

$$A_4 - A_3 = \begin{pmatrix}0&5&0\\15&0&-5\\-10&5&0\end{pmatrix}$$

die Mehr- bzw. Minderleistungen im vierten Quartal gegenüber dem dritten Quartal. (Beachte, dass hier kein Teilbetrieb Eigenleistungen verbraucht, also die A_k Nulleinträge auf der Diagonalen haben; das gilt dann natürlich auch für die Summe und für Differenzen.) Hätte man in jedem Quartal dieselben Leistungen verzeichnet wie im ersten, so hätte sich als Verflechtungsmatrix für des gesamte Jahr $4A_1$ ergeben. Hier ist "zufällig" $4A_1$ gleich der Summe $A_1 + A_2 + A_3 + A_4$, d.h. die im Vergleich mit dem ersten Quartal festzustellenden Mehr- und Minderleistungen im zweiten, dritten und vierten Quartal heben sich gerade auf. A_1 ist daher auch das arithmetische Mittel der vier Verflechtungsmatrizen.

6) Noch eine (formale) *ökonomische Anwendung*: Ein Betrieb stellt an Betriebsstätten (1)… (m) dasselbe Produkt her. Die monatlich produzierten Mengen erfasst ein Spaltenvektor $(x_1, \ldots, x_m)$, der sog. **Produktionsvektor**. Der Eintrag in Position (i) gibt dabei die an der Betriebsstätte (i) produzierte Produktmenge (in Mengeneinheiten) an. Für die 12 Monate eines Jahres hat man dann 12 solche Produktionsvektoren $\mathbf{x}_1, \ldots, \mathbf{x}_{12}$ (jeder mit m Einträgen). Werden die produzierten Mengen im k-ten Monat am Markt zum Preis p_k abgesetzt, so ist die Linearkombination

$$p_1\mathbf{x}_1 + p_2\mathbf{x}_2 + \ldots + p_{12}\mathbf{x}_{12}$$

ein Spaltenvektor mit m Komponenten, dessen i-ter Eintrag den von der i-ten Produktionsstätte erzielten Jahreserlös angibt; man könnte ihn also den *Jahreserlösvektor* des in die Stätten (1)… (m) aufgegliederten Betriebs nennen. Seine Komponentensumme (die Summe all seiner Einträge) ist dann der Jahreserlös des gesamten Betriebs. ■

Natürlich braucht man nicht wirklich Vektor- und Matrizenrechnung, um die ökonomischen Sachverhalte in den beiden letzten Beispielen zu erfassen. Man kann ja den Jahressaldo auch separat für jeden Teilbetrieb bzw. für jede Produktionsstätte ermitteln, und das läuft auf Dasselbe hinaus wie die in den Beispielen ausgeführte Addition der Verflechtungsmatrizen bzw. Linearkombination der Produktionsvektoren. Der Vorteil des Gebrauchs von Matrizen und Vektoren liegt hier nur in der konzentrierten Schreibweise und der Zusammenfassung vieler gleichartiger Vorgänge oder Berechnungen:

- *Mehrere gleichartige Rechnungen mit reellen Zahlen kann man oft zu einer einzigen Vektor- oder Matrixrechnung zusammenfassen.*

Aus den Rechengesetzen im Zahlbereich $\mathbb{R}$ und aus der Definition der Additon und Vervielfachung von Matrizen / Vektoren ergeben sich unmittelbar entsprechende Rechengesetze, die man zur Grundlage der Definition eines allgemeinen Vektorraums gemacht hat. Diese abstrakte Begriffsbildung ist in der Mathematik fundamental für die Lineare Algebra. Für Anwendungen der Linearen Algebra in der Ökonomie ist sie aber nicht so wichtig, weil man es dort immer mit konkreten Vektorräumen zu tun hat, z.B. mit dem Raum $\mathbb{R}^m$ der m-gliedrigen Spaltenvektoren (bzw. m-tupel) oder dem Raum $\mathbb{R}^{m\times n}$ der $m\times n$–Matrizen mit reellen Einträgen. Wir stellen deshalb hier auch nicht den Begriff des Vektorraums in den Vordergrund, sondern erklären ihn nur kurz (unter anderem deswegen, weil sich daraus die Antwort auf die oft gestellte Frage ergibt: "Was ist eigentlich ein Vektor?").

DISKUSSION (*Rechenregeln für die Addition und Vervielfachung; Vektorräume*)

1) Die **Rechenregeln für die Addition** von Matrizen sind:

$$\begin{aligned} A+B &= B+A && (\textit{Kommutativgesetz})\\ (A+B)+C &= A+(B+C) && (\textit{Assoziativgesetz})\\ A+0 &= A = 0+A && (\textit{neutrales Element})\\ A+B &= 0 \iff B = -A && (\textit{additives Inverses}) \end{aligned}$$

Hier sind — selbstverständlich — A, B, C und die Nullmatrix 0 Matrizen desselben Formats. Die beiden ersten Gesetze haben zur Folge, dass die Summe von beliebig vielen gleichformatigen Matrizen definiert ist und von der Klammerung und der Reihenfolge der Summanden nicht abhängt. Die Summe hat natürlich dasselbe Format wie alle Summanden. Die Rechengesetze hier gelten auch für Vektoren, also für einzeilige oder einspaltige Matrizen. Nur schreibt man dann meistens kleine Buchstaben a, b, $\ldots$, x, y, $\ldots$ (manchmal auch fett $\mathbf{a}$, $\mathbf{b}$, $\ldots$, $\mathbf{x}$, $\mathbf{y}$, $\ldots$) statt der für Matrizen üblichen großen Buchstaben.

2) Die **Rechenregeln für die Multiplikation mit einem Skalar** sind:

$$\begin{aligned} 1A &= A, \quad 0A = 0 &&\\ r(sA) &= (rs)A && (\textit{Assoziativgesetz})\\ (r+s)A &= rA+sA && (\textit{linkes Distributivgesetz})\\ r(A+B) &= rA+rB && (\textit{rechtes Distributivgesetz}) \end{aligned}$$

Hierbei sind $0, 1, r, s \in \mathbb{R}$ und A, B Matrizen desselben Formats wie die Nullmatrix 0. Für das Rechnen mit Zeilen- oder Spaltenvektoren gelten diese Gesetze natürlich auch. Da sie ganz analog zu Rechengesetzen für reelle Zahlen sind, ist es nicht schwer, damit richtig umzugehen.

3) Hat man eine Menge V mit einem ausgezeichneten Element 0 und ist für je zwei Elemente $v, w \in V$ sowie für Zahlen $r \in \mathbb{R}$ eine Summe $v + w \in V$ und das Vielfache $rv \in V$ definiert, derart dass die in 1) und 2) angegebenen Rechenregeln analog gelten, so nennt man V mit dieser Struktur einen **reellen Vektorraum**, die Elemente $v, w, \ldots$ von V **Vektoren** und 0 den **Nullvektor**. Der m-dimensionale Zahlenraum $\mathbb{R}^m$ der Spaltenvektoren mit m reellen Komponenten und der Raum $\mathbb{R}^{m\times n}$ der $m\times n$–Matrizen mit reellen Einträgen sind die hier wichtigsten Beispiele eines Vektorraums, aber es gibt viele weitere. Zum Beispiel bilden die Polynomfunktionen auf $\mathbb{R}$ oder die differenzierbaren Funktionen aus einem offenen Intervall I in $\mathbb{R}$ einen Vektorraum, wenn dafür die Addition und Vervielfachung wie üblich (wertweise) definiert und die Nullfunktion als Nullvektor genommen wird. Auch der Bereich der reellen Zahlen $\mathbb{R}$ selbst bildet einen reellen Vektorraum (weil die geforderten Rechenregeln in $\mathbb{R}$ alle gelten). Daher ist es nicht ganz korrekt zu sagen, ein Vektor sei ein "Objekt mit mehreren Komponenten" oder Einträgen wie eben ein m-tupel (mit $m > 1$) oder eine Matrix (mit mehr als einem Eintrag). Man kann eigentlich nicht definieren, was ein einzelner "Vektor" ist, sondern der richtige Begriff ist der des *Vektorraums* mit seinen Rechenoperationen und Rechengesetzen. Wenn man einen Vektorraum hat, dann nennt man seine Elemente eben auch "Vektoren", worum immer es sich dabei auch handeln mag (m-tupel, Matrizen, Polynome, Funktionen, ...).

4) Der Vorteil der Begriffsbildung des abstrakten Vektorraums ist, dass man damit gleichartige Konstruktionen und Rechnungen in vielen ganz verschiedenen Situationen einheitlich erfassen und ausführen kann. So lassen sich z.B. Linearkombinationen nicht nur von Spaltenvektoren und Matrizen, sondern auch von Polynomen, Funktionen, ... , eben von Vektoren in einem beliebigen Vektorraum V erklären, nämlich

$$r_1v_1 + r_2v_2 + \ldots + r_lv_l \in V$$

als **Linearkombination der Vektoren** $v_1, v_2, \ldots, v_l \in V$ mit **Koeffizienten** $r_1, r_2, \ldots, r_l \in \mathbb{R}$, und man kann damit genau so rechnen wie mit Linearkombinationen von Spaltenvektoren in $\mathbb{R}^m$ oder von $m\times n$–Matrizen.

Auch viele geometrische Vorstellungen, die man mit Vektoroperationen in der durch $\mathbb{R}^2$ modellierten "Zeichenebene" oder in dem durch $\mathbb{R}^3$ modellierten "Anschauungsraum" verbindet, lassen sich auf beliebige Vektorräume übertragen und erleichtern dann den Umgang mit und das Verständnis von Vektorrechnungen. Zum Beispiel bildet die Menge aller Linearkombinationen der Form $(1-t)v+tw$ zu zwei verschiedenen Vektoren $v, w \in V$ mit $t \in \mathbb{R}$, also die Menge aller Linearkombinationen dieser beiden Vektoren mit Koeffizientensumme 1, die sog. **Verbindungsgerade** der Punkte v, w in V. Und wenn man noch verlangt, dass die Koeffizienten nichtnegativ sind, so erhält man die **Verbindungsstrecke** $\{(1-t)v+tw \,:\, 0 \le t \le 1\}$ der beiden Punkte mit dem **Mittelpunkt** $\frac{1}{2}v + \frac{1}{2}w$. Allgemeiner nennt man eine Linearkombination von $v_1, v_2, \ldots, v_l \in V$ eine **affine Kombination** dieser Vektoren, wenn die Koeffizientensumme 1 ist, und eine **konvexe Kombination**, wenn außerdem noch alle Koeffizienten nichtnegativ sind. Eine konvexe Kombination kann man auch als **gewichtetes arithmetisches Mittel der Vektoren** ansehen,

$$s_1v_1 + s_2v_2 + \ldots + s_lv_l \in V \qquad (s_k \ge 0 \text{ für } k = 1 \ldots l,\ s_1+s_2+\ldots+s_l = 1)\,.$$

Für drei verschiedene Punkte $u, v, w \in V$, von denen keiner auf der Verbindungsgeraden der beiden anderen liegt, bilden die affinen Kombinationen anschaulich gesprochen die Ebene in V, die durch diese drei Punkte bestimmt ist, und die konvexen Kombinationen füllen das Dreieck darin mit den Ecken u, v, w aus; das arithmetische Mittel $\frac{1}{3}u + \frac{1}{3}v + \frac{1}{3}w$ ist z.B. der Schwerpunkt dieses Dreiecks.

All diese geometrischen Konstruktionen und Sprechweisen machen nicht nur in der Zeichenebene $\mathbb{R}^2$ und im Anschauungsraum $\mathbb{R}^3$ Sinn, sondern — weil sie nur die Rechenoperationen und Rechengesetze eines Vektorraums erfordern — in *jedem* Vektorraum, also auch im m-dimensionalen Zahlenraum $\mathbb{R}^m$ mit $m > 3$, im Raum der $m \times n$–Matrizen $\mathbb{R}^{m\times n}$, in Vektorräumen von Funktionen usw. Das ist eine der nützlichen Erkenntnisse, welche das Konzept des allgemeinen Vektorraums mit sich bringt. ■

Wichtiger als die Addition, aber nicht so einfach zu definieren, ist die Multiplikation von Matrizen (deren Formate in gewisser Weise zusammen "passen"). Zur Vorbereitung definieren wir das für die Anwendungen ebenfalls sehr wichtige Skalarprodukt zweier Zeilen- oder Spaltenvektoren mit gleicher Komponentenzahl. Der Name dieses Produkts kommt daher, dass das Ergebnis ein Skalar ist, also eine reelle Zahl. Das Skalarprodukt ist auch wichtig für die Geometrie, weil mit seiner Hilfe Grundgrößen wie Abstand und Winkel erklärt werden. Hier aber interessieren uns nur die formalen Eigenschaften dieses Produkts.

DEFINITION: Sind a und b zwei Zeilen- oder Spaltenvektoren mit gleich vielen reellen Einträgen $a_1, \ldots, a_n$ und $b_1, \ldots, b_n$ so ist das **Skalarprodukt der beiden Vektoren** (auch **inneres Produkt** genannt) definiert als die reelle Zahl

$$a \bullet b := a_1b_1 + a_2b_2 + \ldots + a_nb_n \in \mathbb{R}\ .$$

Um das Skalarprodukt $a \bullet b$ zu erhalten, multipliziert man also die Einträge von a und b mit gleicher Positionsnummer und addiert alle so entstehenden Produkte. ■

DISKUSSION: **1)** Bei der Bildung des Skalarprodukts ist es gleichgültig, ob die Faktoren Zeilen- oder Spaltenvektoren bzw. n-tupel sind. Die Faktoren müssen auch nicht von gleichem Format sein, d.h. es ist durchaus erlaubt, dass ein Faktor ein Zeilenvektor, der andere Faktor ein Spaltenvektor ist. Einzige Bedingung ist, dass die Faktoren Vektoren mit gleich vielen Komponenten sind.

- *Das Skalarprodukt ist nur für zwei (Zeilen- oder Spalten-) Vektoren mit gleicher Anzahl von Einträgen definiert; das Ergebnis ist dann ein Skalar, also eine reelle Zahl.*

Dementsprechend kann man auch nicht das Skalarprodukt $a \bullet b \bullet c$ von drei Vektoren mit gleicher Komponentenzahl bilden, weil $a \bullet b$ eine reelle Zahl und das Skalarprodukt von $a \bullet b$ mit c nicht erklärt ist. (Ausnahme: a, b, c haben jeweils nur eine Komponente, sind also reelle Zahlen; dann ist ihr Skalarprodukt das gewöhnliche Produkt reeller Zahlen.) Allenfalls kann man $(a \bullet b)c$ als Produkt sc des Skalars $s := a \bullet b$ mit dem Vektor c erklären.

2) Analog zur Addition könnte man ein Produkt "$*$" für Spaltenvektoren mit gleicher Komponentenzahl definieren durch positionsweise Multiplikation der Einträge, also durch $(a_1, a_2, \ldots, a_n) * (b_1, b_2, \ldots, b_n) := (a_1b_1, a_2b_2, \ldots, a_nb_n)$. Dieses Produkt würde auch den üblichen Rechenregeln für Produkte genügen. Nur ist es für die Anwendungen völlig bedeutungslos im Gegensatz zum Skalarprodukt von Vektoren und der damit später definierten Matrixmultiplikation. Deshalb wird dieses Produkt auch nicht eingeführt.

3) Wir haben die *Notation* $a \bullet b$ mit einem fetten Multiplikationspunkt für das Skalarprodukt gewählt, damit man es vom Produkt $a \cdot b = ab$ reeller Zahlen a, b unterscheiden und an der Notation sofort erkennen kann, dass es sich um das Skalarprodukt von Vektoren mit gleicher Komponentenzahl handelt. Andere gebräuchliche Bezeichnungen für das Skalarprodukt sind $\langle a, b\rangle$ oder (a, b). (Letzteres ist aber ungünstig, weil die Gefahr der Verwechselung mit dem Paar der Vektoren a und b besteht, das ebenfalls (a, b) bezeichnet wird.) Bei zwei Spaltenvektoren mit gleicher Komponentenzahl findet man auch die

Notation $a^{\top}b$ für ihr Skalarprodukt, wobei $a^{\top}$ der transponierte Vektor zu a ist, d.h. der Zeilenvektor mit denselben Einträgen (in gleicher Reihenfolge) wie der Spaltenvektor a. (Eigentlich ist $a^{\top}b$ dann das Produkt einer einzeiligen Matrix $a^{\top}$ mit einer einspaltigen Matrix b, und dieses Produkt hat als Ergebnis eine 1×1-Matrix mit dem Skalarprodukt $a\bullet b$ als einzigem Eintrag. Aber zwischen reellen Zahlen s und 1×1–Matrizen (s) macht man gewöhnlich keinen Unterschied.)

4) Aus der Definition des Skalarprodukts und den Rechenregeln im Zahlbereich $\mathbb{R}$ ergeben sich unmittelbar folgende **Rechengesetze für das Skalarprodukt:**

$$\begin{aligned} a\bullet b &= b\bullet a && \textit{(Symmetrie)} \\ (ra)\bullet b &= r(a\bullet b) = a\bullet(rb) && \textit{(Homogenität)} \\ (a\pm\widetilde{a})\bullet b &= a\bullet b\pm\widetilde{a}\bullet b && \textit{(linkes Distributivgesetz)} \\ a\bullet(b\pm\widetilde{b}) &= a\bullet b\pm a\bullet\widetilde{b} && \textit{(rechtes Distributivgesetz)} \end{aligned}$$

Hier ist r eine reelle Zahl und a, $\widetilde{a}$, b, $\widetilde{b}$ sind Zeilen- oder Spaltenvektoren mit gleicher Zahl von Einträgen, wobei a und $\widetilde{a}$ auch gleiches Format haben (d.h. beide sind Zeilenvektoren oder beide Spaltenvektoren; sonst wäre ja $a\pm\widetilde{a}$ nicht definiert) und ebenso auch b und $\widetilde{b}$. Die Tatsache, dass man Summen und reelle Zahlen aus jedem Faktor des Skalarprodukts "herausziehen" kann (die Homogenität und die Distributivität), wird kurz als *Bilinearität* des Skalarprodukts bezeichnet, und man sagt, das Skalarprodukt sei eine *symmetrische Bilinearform*, um diese Bilinearität und das Symmetriegesetz hervorzuheben. Wie beim Rechnen mit Zahlen folgen weitere Rechenregeln, wie z.B. die *binomischen Formeln für Skalarprodukte* von Zeilen- oder Spaltenvektoren a, b desselben Formats

$$(a\pm b)\bullet(a\pm b)=a\bullet a\pm 2a\bullet b+b\bullet b\,, \qquad (a+b)\bullet(a-b)=a\bullet a-b\bullet b\,,$$

wobei es hier nicht üblich ist, etwa a^2 für $a\bullet a$ zu schreiben, und wie das *Herausziehen beliebiger Linearkombinationen*

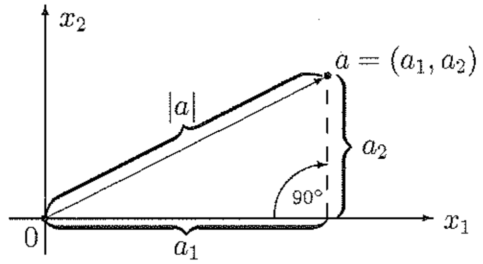
$$\begin{aligned} (r_1\mathbf{a}_1+\ldots+r_k\mathbf{a}_k)\bullet\mathbf{b} &= r_1(\mathbf{a}_1\bullet\mathbf{b})+\ldots+r_k(\mathbf{a}_k\bullet\mathbf{b})\,, \\ (r_1\mathbf{a}_1+\ldots+r_k\mathbf{a}_k)\bullet(s_1\mathbf{b}_1+\ldots+s_l\mathbf{b}_l) &= \sum_{i=1}^{k}\sum_{j=1}^{l} r_i s_j(\mathbf{a}_i\bullet\mathbf{b}_j)\,, \end{aligned}$$

wobei die r_i und s_j reelle Zahlen sind, alle Vektoren $\mathbf{a}_i$ dasselbe Format haben, alle $\mathbf{b}_j$ ebenfalls dasselbe Format besitzen und alle $\mathbf{a}_i$ und $\mathbf{b}_j$ dieselbe Anzahl von Komponenten haben. (Die Klammern rechts hätte man auch weglassen können, weil die Ausdrücke wegen der Homogenität des Skalarproduktes nicht von der Art der Klammerung abhängen.)

4) Das Skalarprodukt $a\bullet a = a_1^2+a_2^2+\ldots+a_n^2$ eines Vektors a mit Komponenten $a_1, a_2, \ldots, a_n$ ist nichtnegativ und nur dann gleich Null, wenn alle Komponenten Null sind. Die Bildung der (nichtnegativen) Wurzel aus $a\bullet a$ hat viele Eigenschaften mit dem Absolutbetrag reeller Zahlen gemeinsam und wird deshalb auch genau so notiert:

$$|a| := \sqrt{a\bullet a} = \sqrt{a_1^2+a_2^2+\ldots+a_n^2}\,,$$

genannt die **(Euklidische) Norm** oder die **Länge** des Vektors a. (Falls $n=1$ ist, also a eine reelle Zahl, so ist das der gewöhnliche Absolutbetrag von a. Wenn man die Norm in der Notation vom Absolutbetrag unterscheiden möchte, so kann man $\|a\|$ dafür schreiben. Nötig ist das aber nicht, weil sich aus dem Kontext ergibt, ob a eine reelle Zahl ist oder ein Vektor mit $n \geq 2$ Komponenten, und im letzteren Fall kann $|a|$ ja nur die Euklidische Norm bedeuten.) Mit dem elementargeometrischen Satz des Pythagoras kann man sehen, dass $|a|$ für $a = (a_1, a_2) \in \mathbb{R}^2$ tatsächlich die Länge der Strecke vom Ursprung $0 = (0,0)$ zum Punkt a ist. Oft stellt man Vektoren als Pfeile dar, hier z.B. den Vektor a als Pfeil mit Ende in 0 und Spitze in a; dann ist $|a|$ die elementargeometrische Länge dieses Pfeils. (Von dieser Darstellungsweise kommt auch der Name "Vektor" her, der eigentlich "gerichtete Größe" heißt.)



geometrische Interpretation von $|a| = \sqrt{a_1^2 + a_2^2}$ *mit dem Satz des Pythagoras*

Es gibt andere Möglichkeiten, die "Größe" von Vektoren a mit Komponenten $a_1, a_2, \ldots, a_n$ zu messen, z.B. sind für diesen Zweck auch die Summe der Komponentenbeträge $|a|_1 := |a_1| + |a_2| + \ldots + |a_n|$ oder ihr Maximum $|a|_\infty := \max(|a_1|, |a_2|, \ldots, |a_n|)$ gebräuchlich, oder allgemeiner die sog. p-Norm $|a|_p := (|a_1|^p + |a_2|^p + \ldots + |a_n|^p)^{1/p}$ zu einem Exponenten $p \in [1, \infty[$, die für $p = 2$ gerade die Euklidische Norm ist. Die Euklidische Norm $|a|$ ist ausgezeichnet durch ihre geometrische Bedeutung und ihren Zusammenhang mit dem Skalarprodukt. So kann man nicht nur $|a| = \sqrt{a \bullet a}$ durch das Skalarprodukt von a mit sich selbst ausdrücken, sondern auch umgekehrt Skalarprodukte von Vektoren $a, b \in \mathbb{R}^n$ durch Euklidische Normen

$$a \bullet a = |a|^2, \qquad a \bullet b = \tfrac{1}{2}|a+b|^2 - \tfrac{1}{2}|a|^2 - \tfrac{1}{2}|b|^2.$$

6) Sind a, b zwei vom Nullvektor verschiedene Vektoren in $\mathbb{R}^n$, so gibt die binomische Formel für Skalarprodukte folgenden Ausdruck für das Quadrat des Abstands von a zu einem Punkt tb auf der Geraden durch den Ursprung 0 und durch b:

$$|a - tb|^2 = |a|^2 - 2ta \bullet b + t^2|b|^2.$$

Diese quadratische Funktion von $t \in \mathbb{R}$ hat ihr Minimum bei $t = 0$, genau wenn $a \bullet b = 0$ ist. Genau dann ist also der Nullpunkt der zu a nächste Punkt auf der Geraden durch 0 und b. Aus der elementaren Geometrie in der Zeichenebene $\mathbb{R}^2$ bzw. im Anschauungsraum $\mathbb{R}^3$ weiß man aber, dass die Verbindungsstrecke von einem Punkt a zum nächsten Punkt auf einer Geraden das Lot von a auf diese Gerade ist, d.h. auf der Geraden senkrecht steht. Also gilt dort $a \bullet b = 0$, genau wenn die Pfeile, welche die Vektoren a und b darstellen, senkrecht zueinander sind. Daher ist es sinnvoll, allgemein zwei **Vektoren orthogonal zueinander** oder **senkrecht aufeinander** zu nennen, wenn ihr Skalarprodukt definiert ist und den Wert Null hat. Geometrisch stellt man sich einen rechten Winkel zwischen den die Vektoren darstellenden Pfeilen vor.

7) Eine zum Skalarprodukt gehörende *Division von Vektoren ist nicht definiert.* Der Quotient von b und a müsste ja, wenn er sinnvoll zu definieren wäre, die eindeutige Lösung x der Gleichung $a \bullet x = b$ sein. Dann müsste also x ein Vektor mit derselben Zahl n von Einträgen wie a sein und b ein Skalar. Aber es gibt (außer in dem hier nicht interessierenden Fall $n = 1$, wo a, b, x reelle Zahlen sind und $a \neq 0$) immer unendlich viele Lösungen x dieser Gleichung, wenn a mindestens eine von Null verschiedene Komponente a_j hat; denn es handelt sich ja um eine lineare Gleichung $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ für die n Komponenten von x. Und deshalb ist keine sinnvolle Definition der Division des Skalars b durch den Vektor a möglich, auch nicht wenn a kein Nullvektor ist. ■

BEISPIELE (*zum Skalarprodukt von Vektoren*):

1)
$$\begin{pmatrix}1\\2\\3\end{pmatrix} \bullet \begin{pmatrix}3\\-2\\1\end{pmatrix} = 1\cdot 3 + 2\cdot(-2) + 3\cdot 1 = 2\,,$$

$$(2 \quad 0 \quad -1 \quad 1)\bullet \begin{pmatrix}0\\1\\2\\3\end{pmatrix} = 2\cdot 0 + 0\cdot 1 + (-1)\cdot 2 + 1\cdot 3 = 1\,,$$

$$(1,1,1)\bullet(1,-2,1) = 1\cdot 1 + 1\cdot(-2) + 1\cdot 1 = 0\,.$$

Natürlich ist ein Skalarprodukt Null, wenn einer der Faktoren ein Nullvektor ist, also nur Einträge Null hat. Das letzte Beispiel zeigt, dass die Umkehrung nicht gilt. Geometrisch bedeutet $a\bullet b = 0$ für $a, b \in \mathbb{R}^n$ ja nur, dass die Vektoren a und b, als Pfeile vom Ursprung aus gedacht, orthogonal sind, also senkrecht zueinander; und das bedeutet (in $\mathbb{R}^n$ mit $n \geq 2$) natürlich nicht, dass a oder b der Nullvektor sein muss.

2) Das Skalarprodukt von einem Vektor a mit m Komponenten und einem Vektor, der m gleiche Einträge 1 hat, ist gerade die Summe der Einträge von a:
$$\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix} \bullet \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix} = a_1 + a_2 + \ldots + a_m\,.$$
Daher heißen Vektoren wie $(1,1,\ldots,1)$ auch *summierende Vektoren*. Entsprechend ist das Skalarprodukt von a mit $\frac{1}{m}(1,1,\ldots,1) \in \mathbb{R}^m$ das arithmetische Mittel der Einträge von a. Sind allgemeiner Gewichte $s_1, s_2, \ldots, s_m \geq 0$ mit $s_1 + s_2 + \ldots + s_m = 1$ gegeben, so ist das Skalarprodukt von a mit dem *Gewichtsvektor* $(s_1, s_2, \ldots, s_m)$ das gewichtete arithmetische Mittel der m Einträge von a mit den gegebenen Gewichten:
$$\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix} \bullet \begin{pmatrix}s_1\\s_2\\\vdots\\s_m\end{pmatrix} = s_1a_1 + s_2a_2 + \ldots + s_ma_m\,.$$

3) Das *Skalarprodukt eines Vektors x mit einem kanonischen Basisvektor* ist eine Komponente von x. Genauer ist $x \bullet e_i = x_i$, wenn x die Komponenten $x_1, \ldots, x_m$ hat und e_i wie üblich den kanonischen Basisvektor mit Eintrag 1 in i-ter Position und mit weiteren $m-1$ Einträgen 0 bezeichnet, also
$$x\bullet e_i = \begin{pmatrix}x_1\\x_2\\\vdots\\x_m\end{pmatrix} \bullet \overbrace{(0,\ \ldots\ ,0\ ,\underbrace{1}_{i\text{-te Position}},\ 0,\ \ldots\ ,0)}^{m} = x_i\,.$$
Insbesondere haben wir für *das Skalarprodukt von zwei kanonischen Basisvektoren* e_i und e_j in $\mathbb{R}^m$:
$$e_i\bullet e_j = \delta_{ij} := \begin{cases}1, & \text{wenn } i = j\,,\\ 0, & \text{wenn } i \neq j\,.\end{cases}$$
Dabei ist das sog. *Kronecker–Delta–Symbol* δ_{ij} eine in der Mathematik übliche Abkürzung für eine Größe, die den Wert 1 hat, wenn beide Indizes gleich sind, und den Wert 0 sonst.

4) Eine *ökonomische Anwendung* (formaler Art): Für m gefertigte Produkte können wir die in einem gewissen Zeitraum abgesetzten Mengen $x_1, x_2, \ldots, x_m$ (in Mengeneinheiten) zu einem **Absatzvektor x** zusammenfassen und die beim Verkauf dieser Produkte erzielten Preise $p_1, p_2, \ldots, p_n$ zu einem **Preisvektor p**. Dann gibt das Skalarprodukt den

$$\textbf{Umsatz / Erlös} \qquad \mathbf{x} \bullet \mathbf{p} = x_1 p_1 + x_2 p_2 + \ldots + x_m p_m$$

an, der mit den m Produkten im betrachteten Zeitraum insgesamt erzielt wurde. Entsprechend erhält man aus einem **Produktionsvektor x**, dessen Einträge x_i die Outputs (in Mengeneinheiten) von m Produkten angeben, und dem zugehörigen **Stückkostenvektor k**, dessen Einträge k_i die bei der Produktion pro Einheit des $i$ten Produkts entstandenen Kosten sind, durch Skalarproduktbildung die

$$\textbf{Gesamtkosten} \qquad \mathbf{x} \bullet \mathbf{k} = x_1 k_1 + x_2 k_2 + \ldots + x_m k_m \,.$$

Immer wenn, wie hier, die gleich positionierten Einträge zweier gleich langen Tabellenspalten miteinander multipliziert und die entstehenden Produkte aufsummiert werden, hat man in der Ökonomie eine Skalarproduktbildung.

5) *Anwendung auf eine lineare Gleichung*: Eine lineare Gleichung

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b$$

kann unter Verwendung des Skalarprodukts kurz

$$a \bullet x = b$$

geschrieben werden, wo $a = (a_1, a_2, \ldots, a_n)$ der *Koeffizientenvektor* der Gleichung ist und $x = (x_1, x_2, \ldots, x_n)$ der *Vektor der Unbekannten.* Die rechte Seite ist hier natürlich ein Skalar $b \in \mathbb{R}$. Die Lösungsmenge besteht also aus allen Vektoren $x \in \mathbb{R}^n$, die mit dem Koeffizientenvektor a ein Skalarprodukt mit dem gegebenen Wert b bilden. Im homogenen Fall $b = 0$ läßt sich die Lösungsmenge geometrisch als die Menge aller zu a senkrechten Vektoren in $\mathbb{R}^n$ beschreiben, genannt der zu *a orthogonale Unterraum von $\mathbb{R}^n$* und notiert $\{a\}^\perp$. Aus den Rechenregeln für das Skalarprodukt folgt für jede "spezielle" Lösung x^*:

$$a \bullet x = b \quad \Longleftrightarrow \quad a \bullet x = a \bullet x^* \quad \Longleftrightarrow \quad a \bullet (x - x^*) = 0\,,$$

d.h. die Lösungsmenge L zu $a \bullet x = b$ erhält man durch die Parallelverschiebung von $\{a\}^\perp$, die den Nullvektor in x^* überführt, $L = x^* + \{a\}^\perp$. In $\mathbb{R}^2$ ist das die zur Richtung von $a \neq 0$ senkrechte Gerade durch x^*, in $\mathbb{R}^3$ die zur Richtung von $a \neq 0$ senkrechte Ebene durch x^*.

Aus den Rechenregeln für das Skalarprodukt sieht man auch unmittelbar:

- *Linearkombinationen von Lösungsvektoren $\mathbf{x}_k \in \mathbb{R}^n$ der homogenen Gleichung sind wieder Lösungen der homogenen Gleichung;*
- *affine Kombinationen von Lösungsvektoren $\mathbf{x}_k \in \mathbb{R}^n$ der inhomogenen Gleichung sind wieder Lösungen der inhomogenen Gleichung.*

Beides folgt aus

$$a \bullet \sum_{k=1}^{l} t_k \mathbf{x}_k = \sum_{k=1}^{l} t_k (a \bullet \mathbf{x}_k)\,,$$

weil im ersten Fall $a \bullet \mathbf{x}_k = 0$ ist für alle k, also auch die rechte Seite der letzten Gleichung $= 0$, und im zweiten Fall $a \bullet \mathbf{x}_k = b$ für alle k sowie $\sum_{k=1}^{l} t_k = 1$, also die rechte Seite der letzten Gleichung wieder $= b$. Die Aussagen hier gelten übrigens nicht nur für lineare Gleichungen, sondern auch analog für lineare Gleichungssysteme und können (mühsam) durch direkte Rechnungen mit den Gleichungssystemen oder (einfacher) mit den folgenden Rechenregeln für die Multiplikation von Matrizen an Spaltenvektoren gezeigt werden. ∎

Mathematik für Wirtschaftswissenschaftler

(K. Steffen, Heinrich-Heine-Universität Düsseldorf, WS 2006/07)

Um zu motivieren, wie man das *Produkt einer Matrix mit einem Spaltenvektor* sinnvoll definiert, betrachten wir folgende ökonomische Situation: In einem Betrieb werden die Produkte $\boxed{1}, \ldots, \boxed{n}$ gefertigt, wobei die Rohstoffe $1, \ldots, m$ Verwendung finden. Die **Materialverflechtungsmatrix** $R = (r_{ij})$

$$
\begin{array}{c|cccc}
 & \boxed{1} & \boxed{2} & \ldots & \boxed{n} \\
\hline
1 & r_{11} & r_{12} & \ldots & r_{1n} \\
2 & r_{21} & r_{22} & \ldots & r_{2n} \\
\vdots & \vdots & \vdots & & \vdots \\
m & r_{m1} & r_{m2} & \ldots & r_{mn}
\end{array}
\quad \leftarrow \text{Produkte}
$$

Rohstoffe

hat in der i-ten Zeile und j-ten Spalte den **Rohstoffverbrauchskoeffizienten** r_{ij} als Eintrag, der angibt, wieviele Mengeneinheiten des Rohstoffs i für die Herstellung einer Einheit des Produkts $\boxed{j}$ benötigt werden. Die Frage ist nun, wie man zu einem gegebenen **Produktionsoutputvektor** $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ den zugehörigen **Rohstoffbedarfsvektor** $\mathbf{z} = (z_1, z_2, \ldots, z_m)$ berechnet, dessen Einträge z_i angeben, wieviele Einheiten vom Rohstoff i gebraucht werden, wenn der gegebene Produktionsplan von x_1 Outputeininheiten des Produkts $\boxed{1}$, x_2 Outputeinheiten des Produkts $\boxed{2}$, $\ldots$, x_n Outputeinheiten des Produkts $\boxed{n}$ realisiert werden soll. Da man für die Herstellung von x_j Outputeinheiten des Produkts $\boxed{j}$ gerade r_{ij} Einheiten des Rohstoffs i braucht, ist die Antwort

$$z_i = r_{i1}x_1 + r_{i2}x_2 + \ldots + r_{in}x_n \quad \text{für} \quad i = 1 \ldots m\,,$$

die i-te Komponente von $\mathbf{z}$ ist also das Skalarprodukt der i-ten Zeile der Materialverflechtungsmatrix R mit dem Produktionsvektor $\mathbf{x}$. Dies kann man durch eine einzige Vektorgleichung

$$\mathbf{z} = R\mathbf{x}$$

ausdrücken, wenn man das Produkt von der Matrix R (mit n Spalten) mit dem Spaltenvektor $\mathbf{x}$ (mit n Einträgen) gerade so definiert, wie wir es jetzt tun:

DEFINITION: Das **Produkt einer Matrix mit einem Spaltenvektor** ist definiert, wenn die Spaltenzahl n der Matrix A gleich der Anzahl der Komponenten des Spaltenvektors x ist; ist m die Zeilenzahl von A, so ist das Ergebnis Ax der Multiplikation der m-gliedrige Spaltenvektor, der als i-ten Eintrag das Skalarprodukt der i-ten Zeile von A mit der Spalte x hat für $1 \leq i \leq m$. ■

DISKUSSION: 1) Ausgeschrieben lautet die Definition:

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n \end{pmatrix} \in \mathbb{R}^m .$$

Dabei steht rechts ein Spaltenvektor mit ebenso vielen Einträgen, wie die Matrix links Zeilen besitzt, und die Spaltenzahl der Matrix ist gleich der Komponentenzahl des Spaltenvektors, an den sie multipliziert wird. Führt man die Abkürzung $\underline{a}_i := (a_{i1} \ a_{i2} \ \ldots \ a_{in})$ für die i-te Zeile der Matrix A ein, so kann man entsprechend der Definition schreiben

$$Ax = \begin{pmatrix} \underline{a}_1 \\ \underline{a}_2 \\ \vdots \\ \underline{a}_m \end{pmatrix} x = \begin{pmatrix} \underline{a}_1 \bullet x \\ \underline{a}_2 \bullet x \\ \vdots \\ \underline{a}_m \bullet x \end{pmatrix} .$$

2) Eine andere Möglichkeit der Beschreibung des Produkts Ax ist folgende:

$$Ax = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = x_1\overline{a}_1 + x_2\overline{a}_2 + \ldots + x_n\overline{a}_n ,$$

wo $\overline{a}_j$ die j-te Spalte der Matrix A bezeichnet. Es gilt also:

- *Das Produkt Ax ist die Linearkombination der Spaltenvektoren* $\overline{a}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ *von A mit den Komponenten x_j der Spalte x als Koeffizienten.*

Insbesondere hat der Ergebnisvektor Ax dasselbe Format wie die Spalten von A, und x muss ebenso viele Komponenten haben wie die Matrix A Spalten, damit das Produkt Ax definiert ist.

3) Wir betonen nochmals ausdrücklich, dass man nicht beliebige Matrizen an beliebige Vektoren multiplizieren kann, sondern dass die Formate zusammen "passen" müssen:

- *Das Produkt Ax ist nur definiert, wenn die Spaltenzahl der Matrix A gleich der Komponentenzahl des Spaltenvektors x ist;*
- *das Ergebnis der Multiplikation Ax ist dann ein Spaltenvektor mit so vielen Einträgen, wie A Zeilen hat;*
- *deshalb sind x und Ax im Allgemeinen Spalten mit unterschiedlich vielen Einträgen; nur wenn A eine quadratische Matrix ist, haben x und Ax dasselbe Format.*

4) Die grundlegenden **Rechenregeln für die Multiplikation von Matrizen an Spaltenvektoren** sind folgende:

$$\begin{aligned} 0x &= 0\,, \quad A0 = 0 \\ A(rx) &= r(Ax) = (rA)x && \textit{(Homogenität)} \\ (A+B)x &= Ax + Bx && \textit{(linkes Distributivgesetz)} \\ A(x+\widetilde{x}) &= Ax + A\widetilde{x} && \textit{(rechtes Distributivgesetz)} \end{aligned}$$

Dabei sind A, B und die Nullmatrix 0 Matrizen vom gleichen Format (m,n), die Vektoren x, $\widetilde{x}$ und der Nullvektor im Produkt $A0$ sind n-gliedrige Spaltenvektoren, $r \in \mathbb{R}$ ist ein Skalar und die Gleichungen sind als Gleichheit von Spaltenvektoren mit m Komponenten zu verstehen. Die beiden Distributivgesetze und die Homogenität fasst man zusammen, indem man sagt, dass das Produkt von $m\times n$–Matrizen mit n-gliedrigen Spaltenvektoren eine *bilineare Rechenoperation* ist. Die Bilinearität, also das erlaubte "Herausziehen" von Summenbildung und Multiplikation mit Skalaren aus den Faktoren, ist überhaupt die grundlegende Eigenschaft jeder Rechenoperation, die als "Produktbildung" bezeichnet wird. Natürlich lassen sich dann auch Summen mit mehr als zwei Summanden aus den Faktoren des Produkts herausziehen, und allgemeiner gilt:

- *Man kann bei der Produktbildung aus jedem Faktor beliebige Linearkombinationen "herausziehen",*

was hier für Matrizen $A, A_1, \ldots, A_k \in \mathbb{R}^{m\times n}$, Spaltenvektoren $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l \in \mathbb{R}^n$ und Skalare $r_1, r_2, \ldots, r_k, s_1, s_2, \ldots, s_l \in \mathbb{R}$ genau heißt:

$$\begin{aligned}(r_1A_1 + r_2A_2 + \ldots + r_kA_k)\mathbf{x} &= r_1(A_1\mathbf{x}) + r_2(A_2\mathbf{x}) + \ldots + r_k(A_k\mathbf{x}) ,\\ A(s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \ldots + s_l\mathbf{x}_l) &= s_1(A\mathbf{x}_1) + s_2(A\mathbf{x}_2) + \ldots + s_l(A\mathbf{x}_l) .\end{aligned}$$

Wenn beide Faktoren Linearkombinationen sind, so muss man diese nacheinander erst aus dem einen dann aus dem anderen Faktor "herausziehen", also

$$\begin{aligned}(A+B)(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) + B(\mathbf{x}+\mathbf{y}) &= A\mathbf{x} + A\mathbf{y} + B\mathbf{x} + B\mathbf{y} ,\\ (r_1A_1 + r_2A_2 + \ldots + r_kA_k)(s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \ldots + s_l\mathbf{x}_l) &= \sum_{i=1}^{k}\sum_{j=1}^{l} r_is_j(A_i\mathbf{x}_j) ,\end{aligned}$$

wenn auch noch $B \in \mathbb{R}^{m\times n}$ und $\mathbf{y} \in \mathbb{R}^n$ sind. ∎

BEISPIELE (*zum Produkt von Matrizen mit Spaltenvektoren*):

1)
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} ,$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \\ 5 \\ 4 \end{pmatrix} ,$$

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

Man beachte die Formate der Matrizen und Vektoren. Es darf einfach nicht passieren, dass man Matrizen an nicht passende Spaltenvektoren multipliziert oder Ergebnisvektoren mit falscher Komponentenzahl erhält! (Es kann auch nicht passieren, wenn man sich an die angegebene und einzig sinnvolle Definition des Produkts von Matrizen mit Spaltenvektoren hält und sich nicht seine eigene private Definition selbst "strickt".)

2) Im Fall einer einzeiligen Matrix $A = a = (a_1 \ a_2 \ \dots \ a_n)$ reduziert sich die Definition des Produkts Ax mit $x \in \mathbb{R}^n$ auf die Bildung des Skalarprodukts $a \bullet x$ von a mit x:

$$ax = (a_1 \ a_2 \ \dots \ a_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (a_1x_1 + a_2x_2 + \dots + a_nx_n) = (a \bullet x) .$$

Genau genommen steht rechts nicht das Skalarprodukt $a \bullet x$, sondern der 1-gliedrige Spaltenvektor bzw. die 1×1–Matrix mit diesem Skalarprodukt als einzigem Eintrag. Diesen Unterschied, der rein formal ist, kann man ignorieren und daher die Matrixklammern rechts weglassen. (Beachte aber, dass für Skalarprodukte $a \bullet x = x \bullet a$ gilt, während das Matrixprodukt xa der Spalte x mit der Zeile a, das später eingeführt wird, nicht gleich der 1×1–Matrix ax ist, sondern eine $n \times n$–Matrix.)

3) Für das Produkt der $\boldsymbol{n \times n}$**–Einheitsmatrix**

$$\mathbb{I} = \mathbb{I}_n := (\delta_{ij})_{1 \le i \le n,\, 1 \le j \le n} = \left.\underbrace{\begin{pmatrix} 1 & & \\ & 1 & 0 \\ 0 & & \ddots & \\ & & & 1 \end{pmatrix}}_{n}\right\} n$$

(wobei das schon erwähnte Symbol δ_{ij} den Wert 1 hat, wenn $i = j$ ist und den Wert 0 sonst, und das Subskript "n" bei $\mathbb{I}$ das Format der Einheitsmatrix anzeigt, wenn nötig) mit einem n-gliedrigen Spaltenvektor $x \in \mathbb{R}^n$ ist

$$\mathbb{I}x = x ;$$

denn die i-te Zeile von $\mathbb{I}$ ist der i-te kanonische n-gliedrige Basisvektor und dessen Skalarprodukt mit $x \in \mathbb{R}^n$ ist gerade die i-te Komponente von x.

- *Multiplikation der Einheitsmatrix an einen (passenden!) Spaltenvektor reproduziert diesen Vektor.*

4) Für eine **Diagonalmatrix**, d.h. eine quadratische Matrix der Form

$$D = (d_i\delta_{ij})_{1 \le i \le n,\, 1 \le j \le n} = \left.\underbrace{\begin{pmatrix} d_1 & & \\ & d_2 & 0 \\ 0 & & \ddots & \\ & & & d_n \end{pmatrix}}_{n}\right\} n$$

mit Einträgen d_i auf der Diagonalen und Nulleinträgen sonst, gilt

$$Dx = \underbrace{\begin{pmatrix} d_1 & & \\ & d_2 & 0 \\ 0 & & \ddots & \\ & & & d_n \end{pmatrix}}_{n} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} d_1x_1 \\ d_2x_2 \\ \vdots \\ d_nx_n \end{pmatrix} ,$$

d.h. die Multiplikation von D an $x \in \mathbb{R}^n$ läuft auf Dasselbe hinaus, wie jede Komponente von x mit dem entsprechenden Diagonaleintrag zu multiplizieren.

Sind alle Diagonaleinträge $d_i = d$ gleich, so ist D das d-fache der Einheitsmatrix, und $Dx = (d\mathbb{I})x = d(\mathbb{I}x) = dx$ das d-fache des Spaltenvektors x. Die Multiplikation der Matrix $d\mathbb{I}$ an einen (passenden) Spaltenvektor x läuft also auf Dasselbe hinaus wie die Multiplikation von x mit dem Skalar d; deshalb nennt man eine Matrix der Form $d\mathbb{I}$, also eine Diagonalmatrix mit lauter gleichen Diagonaleinträgen d, auch eine **skalare Matrix**.

5) Setzen wir für x den j-ten kanonischen Basisvektor $e_j \in \mathbb{R}^n$ ein, so ergibt sich

$$Ae_j = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \\ \\ j\text{-te Position (von } n) \\ \\ \end{array} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

- *Das Produkt einer $m \times n$–Matrix A mit dem j-ten kanonischen Basisvektor von $\mathbb{R}^n$ ist gerade die j-te Spalte von A.*

6) Eine ökonomische Anwendung findet die Multiplikation von Matrizen an Spaltenvektoren in der **Input–Output–Analyse.** Wir betrachten dazu wie vor der letzten Definition die Herstellung von Produkten $\boxed{1}, \ldots, \boxed{n}$ aus Rohstoffen / Vorprodukten $1, \ldots, m$, wobei die in der *$m \times n$–Materialverflechtungsmatrix* $R = (r_{ij})$ zusammengefassten sog. *Rohstoffverbrauchskoeffizienten* r_{ij} die Anzahl der Einheiten von i angeben, die für die Herstellung einer Einheit von $\boxed{j}$ benötigt werden. Ist $x = (x_1, x_2, \ldots, x_n)$ der *Produktionsvektor*, dessen Einträge die zu produzierenden Mengen der n Produkte angeben (in Mengeneinheiten), so ist der zugehörige *Rohstoffbedarfsvektor* $z = (z_1, z_2, \ldots, z_m)$ gegeben durch das Produkt $z = Rx$, weil die Anzahl der für den Produktionsplan benötigten Einheiten des Rohstoffs / Vorprodukts i ja gleich $z_i = r_{i1}x_1 + r_{i2}x_2 + \ldots + r_{in}x_n$ ist.

Wir nehmen nun an, dass die hergestellten Produkte zum Teil am Markt verkauft werden ("Output"), was durch einen *Nachfragevektor* oder *Absatzvektor* $y = (y_1, y_2, \ldots, y_n)$ beschrieben wird, und zum anderen Teil in die Produktion der anderen Produkte oder auch desselben Produkts als Vorprodukte eingehen ("Input"). Dieser Sachverhalt wird durch eine *Produktionsverflechtungsmatrix* $A = (a_{jk})$ vom Format (n, n) beschrieben, wobei der Eintrag a_{jk} angibt, wieviele Einheiten des Produkts $\boxed{j}$ für die Herstellung einer Einheit des Produkts $\boxed{k}$ gebraucht werden. Die a_{jk} heißen *Produktionskoeffizienten*, a_{jj} sind die *Eigenverbrauchskoeffizienten.* Die Gesamtproduktion des Produkts $\boxed{j}$ setzt sich dann zusammen aus der am Markt abgesetzten Menge y_j (*Output*) und der für die Herstellung der anderen Produkte bzw. desselben Produkts verbrauchten Menge $a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n$ (*endogener Input*), also hat man die Bilanz

$$x_j = y_j + a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n \qquad \text{für } j = 1 \ldots n,$$

die man in Form einer Vektorgleichung $x = y + Ax$ schreiben kann. Mit der $n \times n$–Einheitsmatrix $\mathbb{I}$ und den Rechenregeln kann man die Vektorgleichung $y = x - Ax = \mathbb{I}x - Ax = (\mathbb{I} - A)x$ umformen, also gilt für Produktions- und Absatzvektor die Gleichung:

$$(\mathbb{I} - A)x = y.$$

Die hier auftretende Matrix $\mathbb{I} - A$ heißt auch *Technologie–Matrix.* Die Vektorgleichung $(\mathbb{I} - A)x = y$ selbst wird als **statisches Input-Output-Modell** oder *Leontief-Modell* bezeichnet ("statisch", weil hier die Koeffizienten r_{ij} und a_{jk} als zeitlich konstant angenommen sind). Das Modell lässt sich auch auf ein in Teilbetriebe / Abteilungen gegliedertes Unternehmen anwenden und auf eine in Sektoren gegliederte Volkswirtschaft.

Für einen gegebenen Produktionsvektor x kann man nun den Rohstoffverbrauch und die für den Markt verfügbaren Mengen durch einfache Matrixmultiplikationen an x berechnen gemäß

$$z = Rx \in \mathbb{R}^m, \qquad y = (\mathbb{I} - A)x \in \mathbb{R}^n .$$

Dazu wäre natürlich Matrixrechnung nicht unbedingt erforderlich gewesen. Ihr Vorteil ist die übersichtliche "indexfreie" Schreibweise, in der auch die lineare Natur der Abhängigkeit von z und y vom Produktionsvektor x zum Ausdruck kommt (Multiplikation von x mit skalaren Faktoren und Ersetzen von x durch eine Summe von Produktionsvektoren wirken sich entsprechend auf y und z aus).

Interessanter ist die Frage, ob man einen vorgegebenen Nachfragevektor y realisieren kann. Dies erfordert die Auflösung der Gleichung $(\mathbb{I}-A)x = y$ nach $x \in \mathbb{R}^n$, was auf die Lösung eines Systems von n linearen Gleichungen für n Unbekannte mit Koeffizientenmatrix $\mathbb{I}-A$ und Inhomogenität y hinausläuft. Die *Struktur der Technologie-Matrix* $\mathbb{I}-A$ ist oft folgende: Die Produktionskoeffizienten a_{jk} sind nichtnegativ und klein, so dass in den Diagonalpositionen von $\mathbb{I}-A$ Einträge mit einem Wert nahe bei 1 und in den anderen Positionen nichtpositive Einträge von kleinem Betrag stehen. In dieser Situation kann man mathematisch beweisen (wozu aber mehr Theorie gehört; s. 3.4), dass die Auflösung von $(\mathbb{I}-A)x = y$ für jedes $y \in \mathbb{R}^n$ eindeutig möglich ist und dass zudem der Lösungsvektor $x \in \mathbb{R}^n$ nichtnegative Komponenten $x_k \geq 0$ hat, wenn alle Komponenten $y_j \geq 0$ sind, was ja allein ökonomisch sinnvoll ist. (Man kann schließlich eine Produktion nicht ohne Weiteres umkehren und durch Produktverbrauch wieder Rohstoffe zurückgewinnen.) Sind allerdings Produktionskoeffizienten zu groß, so lässt sich unter Umständen nicht mehr jeder Nachfragevektor y (mit $y_j \geq 0$ für alle j) durch einen Produktionsvektor x mit $x_k \geq 0$ für alle k realisieren, sondern die Lösung x hat — wenn überhaupt eine existiert — einen oder mehrere negative Einträge. Dies bedeutet, dass von dem entsprechenden Produkt für den endogenen Input mehr verbraucht wird, als überhaupt produziert werden kann. Das ist natürlich ökonomisch unsinnig, und die Konsequenz ist, dass in solchen Fällen eben nicht jede beliebige Nachfrage nach den Produkten realisiert werden kann.

In der Praxis kommen meist noch *Restriktionen für die Verfügbarkeit von Rohstoffen* hinzu, was durch *lineare Ungleichungen*

$$z_i \leq c_i \quad \text{für gewisse } i \in \{1, 2, \ldots, m\}$$

mit gegebenen Schranken $c_i \in \mathbb{R}_{>0}$ auszudrücken ist. Im Sinne der ökonomischen Fragestellung **zulässige Produktionsvektoren** sind dann nur solche $x \in \mathbb{R}^n$, für die erstens alle Komponenten nichtnegativ sind und zweitens die Komponenten z_i von $z = Rx$ den Ungleichungen $z_i \leq c_i$ genügen, also

$$x_k \geq 0 \quad \text{für } k = 1 \ldots n, \quad r_{i1}x_1 + r_{i2}x_2 + \ldots + r_{in}x_n \leq c_i \quad \text{für gewisse } i \in \{1, 2, \ldots, m\} .$$

Klar ist dann, dass Produktionsvektoren x mit hinreichend kleinen positiven Komponenten zulässig sind. Ob es aber zu einem gegebenen Nachfragevektor y (mit nichtnegativen Komponenten) einen *zulässigen* Produktionsvektor x gibt, der $(\mathbb{I}-A)x = y$ löst, und welche Nachfragevektoren in diesem Sinne realisierbar sind bzw. welche $\widetilde{y}$ möglichst nahe bei y realisiert werden können, wenn das für y selbst nicht möglich ist, — das sind durchaus schwierige und komplexe Fragen, auch in der mathematischen Theorie, mit denen sich die sog. **Input-Output-Analyse** befasst.

7) Noch eine ökonomische Anwendung der Multiplikation von Matrizen an Spaltenvektoren: Ein Unternehmen stelle die Produkte $\boxed{1}, \ldots, \boxed{n}$ her an verschiedenen Produktionsstätten ①, ... ⓜ. Die Verteilung der Produktionsoutputs wird beschrieben durch die $m \times n$–**Produktionsmatrix** $X = (x_{ij})$, deren Einträge x_{ij} angeben, wieviele Einheiten des Produkts $\boxed{j}$ an der Stätte ⓘ erzeugt werden (in einem betrachteten Zeitintervall, z.B. in einem Geschäftsjahr). Ist p_j der für den Verkauf einer Einheit von $\boxed{j}$ am Markt erzielte Preis, so bilden wir den *Preisvektor* $p = (p_1, p_2, \ldots, p_n)$ und können damit den *Umsatzvektor* (auch *Erlösvektor* genannt) $u = (u_1, u_2, \ldots, u_m)$ als Produkt der Produktionsmatrix mit dem Preisvektor darstellen gemäß

$$u = Xp\,, \qquad \text{also} \qquad \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}.$$

Der i-te Eintrag von u ist der bei der Produktionsstätte ⓘ erzielte Umsatz $u_i = \sum_{j=1}^{n} x_{ij}p_j = x_{i1}p_1 + x_{i2}p_2 + \ldots + x_{in}p_n$.

Um die Kostenstruktur zu beschreiben, führt man die $m \times n$–**Kostenproduktionsmatrix** $K = (k_{ij})$ ein, deren Einträge k_{ij} die bei Herstellung des Produkts $\boxed{j}$ an der Produktionssttte ⓘ entstehenden Stückkosten (Kosten pro hergestellte Mengeneinheit) angeben. Das Produkt $K\mathbf{x}$ der Produktionskostenmatrix K mit einem *Produktionsoutputvektor* $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ ist dann der Spaltenvektor, dessen i-ter Eintrag die Kosten $k_{i1}x_1 + k_{i2}x_2 + \ldots + k_{in}x_n$ für die Realisierung des durch $\mathbf{x}$ gegebenen Produktionsplans bei Produktion allein an der Stätte ⓘ angibt. Die eigentliche ökonomische Aufgabe ist aber, die Produktion mit gewünschten Outputs $x_1, x_2, \ldots, x_n$ so auf die m Produktionsstätten zu verteilen, dass die Gesamtkosten minimiert werden. Gesucht ist also eine Produktionsmatrix $X = (x_{ij})$, deren Spaltensummen $x_{1j} + x_{2j} + \ldots + x_{nk} = x_j$ vorgegeben sind, mit der Eigenschaft, dass die Summe $\sum_{i=1}^{m}\sum_{j=1}^{n} k_{ij}x_{ij}$, also die Gesamtkosten für die Produktion an allen Produktionsstätten, minimal wird. Dabei unterliegen die Einträge der zulässigen Produktionsmatrizen noch ökonomischen Restriktionen wie $x_{ij} \geq 0$ und $x_{ij} \leq c_{ij}$ mit gegebenen Kapazitätsgrenzen $c_{ij} \geq 0$ (maximaler möglicher Produktionsoutput für das Produkt $\boxed{j}$ an der Produktionsstätte ⓘ; kann auch $= 0$ sein, nämlich wenn $\boxed{j}$ an der Stätte ⓘ gar nicht gefertigt wird). Die Bestimmung einer Produktionsmatrix X zu gegebenem Produktionsoutputvektor $\mathbf{x}$, welche diese Restriktionen erfüllt und dabei zu kleinsten Gesamtkosten führt, ist dann eine anspruchsvolle mathematische Aufgabe, die in die sog. **Lineare Optimierung** gehört. Mehr dazu würde zu weit führen; das Rechnen mit Matrizen und Vektoren ist jedenfalls Voraussetzung für diese Theorie.

9) Ein letztes ökonomisches Beispiel für das Auftreten des Produkts einer Matrix mit einem Spaltenvektor: Wir betrachten eine $n \times n$–**Übergangsmatrix** $P = (p_{kj})$, die den Käuferwechsel bei n konkurrierende Produkten in einem Beobachtungszeitraum beschreibt, d.h $p_{kj} \cdot 100\,\%$ gibt den prozentualen Anteil der Kunden des Produkts $\boxed{j}$ an, die zum Produkt $\boxed{k}$ gewechselt (bzw. im Fall $k = j$ beim Produkt $\boxed{j}$ geblieben) sind. Der *Marktverteilungsvektor* $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ zu Anfang des Beobachtungszeitraums gibt die Anzahl x_j der Käufer des Produkts $\boxed{j}$ zu diesem Zeitpunkt an für $j = 1 \ldots n$. Die Käuferzahl des Produkts $\boxed{k}$ am Ende des Beobachtungszeitraums ist dann $y_k := p_{k1}x_1 + p_{k2}x_2 + \ldots + p_{kn}x_n$, d.h. der Marktverteilungsvektor $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ am Ende ist gleich dem Matrixprodukt $\mathbf{y} = P\mathbf{x}$. (Vorausgesetzt ist dabei natürlich, dass sich die Gesamtzahl der Käufer im Beobachtungszeitraum nicht geändert hat.) ■

Die Liste der Beispiele für eine prägnante Formulierung ökonomischer Sachverhalte mittels des Produkts von Matrizen und Spaltenvektoren ließe sich noch fortsetzen. Wir verzichten darauf, weil es sich dabei nur um die *Formulierung*, nicht um die *Lösung* ökonomischer Probleme handelt. Für die Lösung muss man letztlich doch meistens lineare Gleichungssysteme auflösen, was wir in 3.2 besprochen haben. Es gibt aber — neben seiner Nützlichkeit für übersichtliche Formulierungen komplexer Zusammenhänge — noch einen tieferen Grund dafür, dass das Produkt von Matrizen und Vektoren in der mathematischen Beschreibung ökonomischer Sachverhalte zwangsläufig auftaucht: Immer wenn eine vektorielle ökonomische Größe $\mathbf{y} = (y_1, \ldots, y_m)$ *lineare Abhängigkeit* von einer anderen vektoriellen Größe $\mathbf{x} = (x_1, \ldots, x_n)$ zeigt, d.h. bei Vervielfachung von $\mathbf{x}$ mit einem Faktor $r \in \mathbb{R}$ vervielfacht sich die abhängige Größe $\mathbf{y}$ mit demselben Faktor und bei Ersetzung von $\mathbf{x}$ durch eine Summe $\mathbf{x} + \widetilde{\mathbf{x}}$ wird $\mathbf{y}$ ersetzt durch die entsprechende Summe $\mathbf{y} + \widetilde{\mathbf{y}}$ der $\mathbf{x}$ und $\widetilde{\mathbf{x}}$ zugeordneten abhängigen Größen, immer wenn eine derartige einfache Gesetzmäßigkeit besteht, wird der Zusammenhang zwischen der abhängigen und der unabhängigen vektoriellen Größe durch eine Matrix–an–Vektor–Multiplikation $\mathbf{y} = A\mathbf{x}$ beschrieben! Wir erklären das in folgender

DEFINITION und DISKUSSION (*Lineare Operatoren und Matrizen*):

1) Eine Zuordnung $T: V \to W$, $V \ni x \mapsto T(x) \in W$, zwischen zwei reellen Vektorräumen V und W heißt **linearer Operator** oder **lineare Abbildung** von V nach W, wenn sie mit den Rechenoperationen in folgender Weise verträglich ist:

$$T(rx) = rT(x) \qquad \text{und} \qquad T(x + \widetilde{x}) = T(x) + T(\widetilde{x}) \qquad \text{in } W$$

für alle $x, \widetilde{x} \in V$ und alle $r \in \mathbb{R}$. Man sagt, dass man die Multiplikation mit einem Skalar bzw. die Summenbildung aus der linearen Abbildung "herausziehen" bzw. mit dem linearen Operator vertauschen kann. (Es ist gleich, ob man erst diese Rechenoperationen ausführt und dann die Abbildung T anwendet, oder ob man erst T ausführt und danach die Rechenoperationen — das Ergebnis ist dasselbe.) Da $0x = 0$ der Nullvektor in V ist, gilt dann insbesondere $T(0) = T(0x) = 0T(x) = 0$, also

$$T(0) = 0,$$

wobei natürlich links der Nullvektor von V und rechts der von W gemeint ist. Die linearen Operatoren sind die strukturerhaltenden Abbildungen zwischen Vektorräumen und für die mathematische Lineare Algebra ein genau so fundamentales Konzept wie der Begriff des Vektorraums. In der mathematischen Modellierung ökonomischer Sachverhalte tauchen lineare Operatoren auf, weil sie in gewisser Weise die einfachste Form der Abhängigkeit einer (vektoriellen) Größe $y = T(x)$ von einer unabhängigen (vektoriellen) Variablen x darstellen.

2) Ist $T: V \to W$ linear, so gilt allgemeiner

$$T(rx + \widetilde{r}\widetilde{x}) = T(rx) + T(\widetilde{r}\widetilde{x}) = rT(x) + \widetilde{r}T(\widetilde{x})$$

für $x, \widetilde{x} \in V$ und $r, \widetilde{r} \in \mathbb{R}$ und entsprechend für Linearkombinationen von mehr als zwei Vektoren $\mathbf{x}_1, \ldots, \mathbf{x}_k \in V$ mit Koeffizienten $r_1, \ldots, r_k \in \mathbb{R}$:

$$T(r_1\mathbf{x}_1 + r_2\mathbf{x}_2 + \ldots + r_k\mathbf{x}_k) = r_1T(\mathbf{x}_1) + r_2T(\mathbf{x}_2) + \ldots + r_kT(\mathbf{x}_k) .$$

- *Lineare Operatoren sind solche Abbildungen, aus denen man beliebige Linearkombinationsbildungen "herausziehen" kann.*

3) Aus den Rechengesetzen für die Multiplikation von $m \times n$–Matrizen A an n-gliedrige Spaltenvektoren folgt, dass der Multiplikationsoperator $\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^m$ ein linearer Operator ist; denn es gilt ja $A(rx) = r(Ax)$ und $A(x+\widetilde{x}) = Ax + A\widetilde{x}$ für $x, \widetilde{x} \in \mathbb{R}^n$ und $r \in \mathbb{R}$. Das Bemerkenswerte ist, dass wir auf diese Weise *jeden* linearen Operator T von $\mathbb{R}^n$ nach $\mathbb{R}^m$ erhalten. Dazu erinnern wir uns, dass sich jeder Vektor $x = (x_1, \ldots, x_n)$ als Linearkombination $x = x_1 e_1 + \ldots + x_n e_n$ der kanonischen Basisvektoren e_j von $\mathbb{R}^n$ schreiben lässt mit den Komponenten x_j von x als Koeffizienten. Daher gilt gemäß 2), wenn T linearer Operator von $\mathbb{R}^n$ nach $\mathbb{R}^m$ ist,

$$T(x) = T(x_1 e_1 + x_2 e_2 + \ldots + x_n e_n) = x_1 T(e_1) + x_2 T(e_2) + \ldots + x_n T(e_n) .$$

Andererseits haben wir für $m \times n$–Matrizen A gesehen, dass Ax gerade die Linearkombination der Spaltenvektoren von A ist mit den Komponenten von x als Koeffizienten. Wenn wir also die Bilder der kanonischen Basisvektoren $T(e_j)$ als Spalten der Matrix A nehmen, so haben wir $T(x) = Ax$ für alle $x \in \mathbb{R}^n$. Damit ist ein fundamentales Ergebnis bewiesen:

- *Die linearen Operatoren $T : \mathbb{R}^n \to \mathbb{R}^m$ sind genau die Abbildungen von der Form $T(x) = Ax$ mit einer $m \times n$-Matrix $A = (a_{ij})$.*
- *Dabei ist die j-te Spalte von A das Bild $T(e_j)$ des j-ten kanonischen Basisvektors in $\mathbb{R}^n$; der Matrixeintrag a_{ij} ist also die i-te Komponente von $T(e_j) \in \mathbb{R}^m$*

(für $i = 1 \ldots m$ und $j = 1 \ldots n$). Wegen dieses engen Zusammenhangs zwischen linearen Operatoren $T : \mathbb{R}^n \to \mathbb{R}^m$ und Matrizen $A \in \mathbb{R}^{m \times n}$ wird oft überhaupt kein Unterschied zwischen beiden Objekten gemacht, d.h. man bezeichnet eine Matrix auch als linearen Operator etc. Das ist natürlich nicht ganz korrekt — schließlich ist eine lineare Abbildung T von $\mathbb{R}^n$ nach $\mathbb{R}^m$ etwas anderes als eine $m \times n$–Matrix, ein rechteckiges Zahlenschema also —, aber es ist gerechtfertigt, weil sich T und A in so einfacher Weise gegenseitig bestimmen: T ist der Multiplikationsoperator $x \mapsto Ax$, der A an n-gliedrige Spaltenvektoren multipliziert, und A ist die Matrix mit den Bildern $T(e_j) \in \mathbb{R}^m$ der kanonischen Basisvektoren von $\mathbb{R}^n$ als Spalten. Man sagt, dass *der lineare Operator T durch die Matrix A dargestellt wird*, wenn dieser Zusammenhang zwischen T und A besteht.

Die eigentliche Bedeutung der Matrix–an–Vektor–Multiplikation und die mathematische Begründung der dazu getroffenen Definition liegt in der damit erreichten Darstellung linearer Operatoren. Die Multiplikation ist eben gerade so definiert, dass $x \mapsto Ax$ die linearen Operatoren von $\mathbb{R}^n$ nach $\mathbb{R}^m$ beschreibt. Gleichzeitig ist mit dem obigen Ergebnis auch die vor der Definition gemachte Aussage gerechtfertigt, dass ein linearer Zusammenhang $y = T(x)$ zwischen einer abhängigen vektoriellen Größe y (in $\mathbb{R}^m$) und einer unabhängigen vektoriellen Variablen x (in $\mathbb{R}^n$) *immer* durch eine Matix–an–Vektor-Multiplikation $y = Ax$ beschrieben werden kann.

4) Eine Folge des Ergebnisses aus 3), die keineswegs von vorneherein klar war, ist:

- *Jede lineare Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ ist durch die Angabe von $m \cdot n$ reellen Zahlen eindeutig bestimmt, nämlich durch die Einträge der $m \times n$–Matrix A, die T darstellt.*

Daher lassen sich lineare Operatoren $T(x) = Ax$ auch einfach programmieren: Man braucht nur die Matrixeinträge von A einzugeben und die Multiplikation von Matrizen an Spaltenvektoren zu programmieren, also die Rechenvorschrift $y_i := a_{i1} x_1 + a_{i2} x_2 + \ldots + a_{in} x_n$ für $i = 1 \ldots m$, dann liefert der Computer für jede Eingabe $x \in \mathbb{R}^n$ den zugehörigen Wert $y = T(x) = Ax$ der linearen Abbildung als Ausgabe. Beliebige Abbildungen von $\mathbb{R}^n$

nach $\mathbb{R}^m$ sind dagegen keineswegs durch endlich viele reelle Zahlen vollkommen bestimmt und insbesondere nicht durch ihre Werte in den kanonischen Basisvektoren von $\mathbb{R}^n$. Aus der Kenntnis der Abbildungswerte an irgendwelchen (endlich oder unendlich vielen) Stellen kann man nichts über die Werte an anderen Stellen in $\mathbb{R}^n$ schließen, wenn von der Abbildung keine besondere Struktur bekannt ist wie z.B. Linearität.

5) Eine Abbildung $T : V \longrightarrow W$ zwischen Vektorräumen heißt **affin lineare Abbildung**, wenn man affine Kombinationen herausziehen kann, d.h. es gilt

$$T((1-t)x + t\widetilde{x}) = (1-t)T(x) + t\,T(\widetilde{x}) \qquad \text{für } x, \widetilde{x} \in V \text{ und } t \in \mathbb{R}$$

und dann auch allgemeiner

$$T(t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \ldots + t_k\mathbf{x}_k) = t_1T(\mathbf{x}_1) + t_2T(\mathbf{x}_2) + \ldots + t_kT(\mathbf{x}_k)$$

für $k \in \mathbb{N}_{\geq 2}$, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in V$ und $t_1, t_2, \ldots, t_k \in \mathbb{R}$ mit $t_1 + t_2 + \ldots + t_k = 1$. Jede lineare Abbildung ist offenbar affin linear (man kann sogar beliebige Linearkombinationen herausziehen, nicht nur affine, wenn die Abbildung linear ist). Tatsächlich sind affin lineare Abbildungen nur wenig allgemeiner als lineare. Ist nämlich $T : V \longrightarrow W$ linear und $b \in W$ ein gegebenes Element von W, so ist die Abbildung $V \ni x \mapsto T(x) + b \in W$ affin linear, und umgekehrt ist jede affin lineare Abbildung aus V nach W von dieser Form, wobei b das Bild des Nullpunkts in Y ist. Das kann man leicht nachrechnen. Es gilt also:

- *Die affin linearen Abbildungen von V nach W sind genau die Abbildungen, die sich von einer linearen Abbildung nur um einen konstanten Summanden unterscheiden, also von der Form $V \ni x \mapsto T(x) + b \in W$ sind mit einer linearen Abbildung $T : V \longrightarrow W$ und einem Element $b \in W$.*

Die geometrische Bedeutung der affin linearen Abbildungen ist, dass sie Geraden in V in Geraden in W (oder ausnahmsweise auch in Punkte in W) transformieren und dabei Abstandsverhältnisse erhalten, so dass z.B. der Mittelpunkt $\frac{1}{2}x + \frac{1}{2}\widetilde{x}$ zu zwei Punkten $x, \widetilde{x}$ in V in den Mittelpunkt der Bildpunkte zu x und $\widetilde{x}$ in W abgebildet wird. Die linearen Abbildungen von V nach W sind durch dieselben Eigenschaften und die Zusatzbedingung, dass der Nullvektor in V auf den Nullvektor in W abgebildet wird, charakterisiert; sie bilden also Geraden durch 0 in V auf Geraden durch 0 in W ab (oder ausnahmsweise auf den Punkt $\{0\} \subset W$). Will man betonen, dass $T : V \longrightarrow W$ linear ist und nicht nur affin linear, dass also $T(0) = 0$ gilt und $T(rx) = rT(x)$ für beliebige $r \in \mathbb{R}$, so nennt man T eine **homogen lineare Abbildung**. ■

Die Interpretation der Matrix–an–Vektor–Multiplikation als linearen Operator bringt auch eine neue Sicht auf lineare Gleichungssysteme, die eine einfache und systematische Behandlung der Lösungsmenge und der Menge der zulässigen rechten Seiten ermöglicht. Wir diskutieren das, um den Nutzen des Konzepts der linearen Abbildung zu demonstrieren.

DISKUSSION (*Anwendungen der Linearität von Operatoren auf Gleichungssysteme*):

1) In Abschnitt 3.1 haben wir für ein lineares Gleichungssystem

$$(\text{LGS}) \qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \qquad \text{oder} \qquad \sum_{j=1}^{n} a_{ij}x_j = b_i \quad \text{für } i = 1 \ldots m$$

schon die abkürzende Schreibweise $Ax = b$ gebraucht, wobei $A = (a_{ij})$ die $m \times n$*–Koeffizientenmatrix* des Systems ist, $b = (b_1, b_2, \ldots, b_m) \in \mathbb{R}^m$ die *Spalte der rechten Seiten* und $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ der *Spaltenvektor der Unbekannten.* Jetzt können wir aber $Ax \in \mathbb{R}^m$ als Produkt der Matrix A mit x auffassen, dann ist die i-te Komponente von Ax gerade die linke Seite der i-ten linearen Gleichung in (LGS). Somit ist

$$Ax = b$$

nicht nur eine symbolische Abkürzung für das Gleichungssystem, sondern, wenn wir Ax als Produkt auffassen und die Gleichung als Vektorgleichung in $\mathbb{R}^m$, eine sinnvolle äquivalente Formulierung von (LGS).

2) Aus der Linearität des Operators $\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^m$ ergeben sich nun unmittelbar folgende Aussagen über das lineare Gleichungssystem (LGS):

(i) *Gilt $Ax = b$ und $A\widetilde{x} = \widetilde{b}$, so löst die Linearkombination $rx + \widetilde{r}\widetilde{x}$ mit Koeffizienten $r, \widetilde{r} \in \mathbb{R}$ das lineare Gleichungssystem mit derselben Koeffizientenmatrix A und mit rechter Seite $rb + \widetilde{r}\widetilde{b}$. (Entsprechend für Linearkombinationen $r_1\mathbf{x}_1 + \ldots + r_l\mathbf{x}_l$ von Lösungen $\mathbf{x}_1, \ldots, \mathbf{x}_l$ zu $A\mathbf{x}_k = \mathbf{b}_k$ für $k = 1 \ldots l$.)*

(ii) *Jede Linearkombination von Lösungen des homogenen Gleichungssystems $Ax = 0$ ist wieder eine Lösung des homogenen Systems. Jede affine Kombination von Lösungen des inhomogenen Gleichungssystems $Ax = b$ löst ebenfalls dieses System.*

(iii) *Jede Linearkombination von zulässigen rechten Spalten zur Koeffizientenmatrix A ist wieder eine zulässige rechte Spalte.*

Die erste Aussage folgt sofort aus $A(rx + \widetilde{r}\widetilde{x}) = r(Ax) + \widetilde{r}(A\widetilde{x}) = rb + \widetilde{r}\widetilde{b}$ und entsprechend für allgemeine Linearkombinationen. Die zweite erhält man daraus, indem man alle rechten Spalten Null setzt oder alle gleich b und dabei $r_1 + \ldots + r_l = 1$ für affine Kombinationen beachtet. Die dritte Aussage schließlich folgt sofort aus der ersten, wenn man sich daran erinnert, dass eine zulässige rechte Spalte ein Vektor $b \in \mathbb{R}^n$ ist, für den $Ax = b$ (mindestens) eine Lösung $x \in \mathbb{R}^n$ besitzt. Alle drei Aussagen sind also unmittelbare Konsequenzen der Linearität des Operators $x \mapsto Ax$. Man kann sie natürlich auch ohne Matrix- und Vektorrechnung — aber viel unübersichtlicher und weniger elegant — durch direkte Rechnungen mit dem ursprünglichen Gleichungssystem (LGS) herleiten. Als weitere Folgerung aus (i) notieren wir noch, was schon mehrfach bemerkt wurde:

(iv) *Die allgemeine Lösung des inhomogenen Systems $Ax = b$ ist, wenn es überhaupt eine Lösung gibt, die Summe irgendeiner ("speziellen") Lösung $x^* \in \mathbb{R}^n$ und der allgemeinen Lösung des homogenen Systems $Ax = 0$;*

denn wegen $Ax^* = b$ gilt ja $A(x+x^*) = b \iff Ax + Ax^* = b \iff Ax = 0$.

3) Man kann die Erkenntnisse aus 2) bei der Lösung linearer Gleichungssysteme praktisch nutzen. Hat man z.B. schon Lösungen für zwei rechte Spalten b und c und soll für eine weitere rechte Spalte d lösen, die man als Linearkombination $d = rb + sc$ von b und c erkennt, so braucht man nicht mehr zu rechnen, sondern kann gemäß (i) einfach die entsprechende Linearkombination der bereits bekannten Lösungen nehmen. Und wenn man zu $Ax = b$ schon die allgemeine Lösung bestimmt hat, so kennt man gemäß (iv) auch die allgemeine Lösung des homogenen Systems und damit wieder nach (iv) auch die allgemeine Lösung zu der neuen rechten Spalte d. ∎

Die Eigenschaft der Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$ und der Menge der zulässigen rechten Seiten b für $Ax = b$, dass beliebige Linearkombinationsbildung mit Elementen der Menge nicht aus der Menge herausführt, erweist sich als grundlegend für die Definition von "linearen Teilmengen" allgemeiner Vektorräume, der sogenannte *Unterräume.* Mit dieser Terminologie kann man dann (ii) dadurch ausdrücken, dass man sagt, die Lösungsmenge eines homogenen linearen Gleichungssystems $Ax = 0$ von m Gleichungen für n Unbekannte sei ein "Unterraum" von $\mathbb{R}^n$ und die Lösunmgsmenge des inhomogenen Systems $Ax = b$ ein "affiner Unterraum" (oder leer). Und (iii) wird ausgedrückt, indem man sagt, dass die Menge der zulässigen rechten Seiten ein Unterraum von $\mathbb{R}^m$ sei. Wir verschieben diese Diskussion aber auf einen späteren Abschnitt, weil wir nun zum *Produkt von zwei beliebigen, aber zusammen "passenden" Matrizen* kommen wollen.

Um diese Produktbildung zu motivieren, die alle bisher eingeführten Produkte im Matrix-Kalkül als Spezialfälle enthält, betrachten wir das Produkt $y = Ax$ einer $m\times n$–Matrix A mit einem n-gliedrigen Spaltenvektor x und das Produkt $z = By$ einer weiteren $l\times m$–Matrix B mit dem m-gliedrigen Spaltenvektor y (damit dieses Produkt By erklärt ist, muss B natürlich m Spalten haben!). Dann ist $z = B(Ax) \in \mathbb{R}^l$, und es ist klar, dass z linear von x abhängt. Also gibt es eine $l\times n$–Matrix C, so dass wir direkt $z = Cx$ als Produkt von C mit x ausgedrückt haben. Mit anderen Worten: Wenn die linearen Operatoren $T:\mathbb{R}^n \to \mathbb{R}^m$ und $U:\mathbb{R}^m \to \mathbb{R}^l$ durch $T(x) = Ax$ und $U(y) = By$ gegeben sind, so suchen wir die Matrix C, welche die *Hintereinanderausführung* (*Verkettung, Komposition*) $U \circ T(x) := U(T(x))$ darstellt. Diese Verkettung ist offenbar wieder ein linearer Operator von $\mathbb{R}^n$ nach $\mathbb{R}^l$. (Man kann skalare Faktoren und Summenbildung erst aus T und dann aus U "herausziehen" und damit aus $U \circ T$.) Also gibt es gemäß dem Zusammenhang zwischen linearen Operatoren und Matrizen auch genau eine Matrix C mit $U \circ T(x) = Cx$ für alle $x \in \mathbb{R}^n$, und diese Matrix C muss natürlich nun das Format (l, n) haben.

Um C zu bestimmen, genügt es, für x kanonische Basisvektoren e_j aus $\mathbb{R}^n$ einzusetzen; denn $U \circ T(e_j) = B(Ae_j)$ ist ja die j-te Spalte von C. Dann ist Ae_j die j-te Spalte von A, und $B(Ae_j)$ hat gemäß der Definition des Produkts von B mit dieser Spalte als h-ten Eintrag das Skalarprodukt der h-ten Zeile von B mit der j-ten Spalte von A, also

$$b_{h1}a_{1j} + b_{h2}a_{2j} + \ldots + b_{hm}a_{mj} = \sum_{i=1}^{m} b_{hi}a_{ij} ,$$

wenn $A = (a_{ij})$ und $B = (b_{hi})$ ist. Dies muss nun der Eintrag c_{hj} in der h-ten Zeile und j-ten Spalte der Matrix C sein, und damit ist sie bestimmt: c_{hj} ist das Skalarprodukt der h-ten Zeile von B mit der j-ten Spalte von A. Da die so definierte Matrix C bilinear von B und A abhängt, ist es sinnvoll, sie als Produkt $C = BA$ zu schreiben. Mit dieser Definition des Produkts BA (und nur mit dieser) gilt also $(BA)x = B(Ax)$, d.h. das Matrixprodukt BA stellt den linearen Operator dar, den man durch Hintereinanderausführung der linearen Operatoren zu A und B erhält.

Diese Überlegungen motivieren die folgende mathematische Definition des Matrixprodukts und haben durchaus auch ökonomische Bedeutung. Man stelle sich vor, dass gewisse Endprodukte Nr. $1 \ldots n$ aus Zwischenprodukten Nr. $1 \ldots m$ gefertigt werden gemäß einer Produktionsmatrix $S = (s_{ij})$ und dass die Zwischenprodukte ihrerseits aus $1 \ldots l$ Rohstoffen hergestellt werden gemäß einer Produktionsmatrix $R = (r_{hi})$. Für einen gegebenen Outputvektor $\mathbf{x} = (x_1, \ldots, x_n)$ zu den Endprodukten beschreibt dann $\mathbf{y} = S\mathbf{x}$ die dafür erforderliche Produktion der Zwischenprodukte und $\mathbf{z} = R\mathbf{y} = R(S\mathbf{x})$ wiederum die für die Realisierung dieser Zwischenproduktion erforderlichen Mengen an Rohstoffen. Das Matrixprodukt RS gibt also zu jedem Produktionsplan $\mathbf{x}$ direkt den zugehörigen Rohstoffbedarfsvektor $(RS)\mathbf{x}$ an und beschreibt damit auch in dieser realen ökonomischen Situation die Komposition von zwei linearen Abhängigkeiten $\mathbf{y} = S\mathbf{x}$ und $\mathbf{z} = R\mathbf{y}$. In analoger Weise kommen Matrixprodukte überall ins Spiel, wo (vektorielle) ökonomische Größen linear von anderen Größen zweiter Art abhängen, die ihrerseit wiederum von Größen einer dritten Art linear abhängig sind.

DEFINITION: Das **Produkt zweier Matrizen** ist definiert, wenn die Spaltenzahl des ersten Faktors gleich der Zeilenzahl des zweiten Faktors ist; das Ergebnis der Multiplikation ist dann die Matrix mit der Zeilenzahl des ersten und der Spaltenzahl des zweiten Faktors, deren Einträge die Skalarprodukte der Zeilen des ersten mit den Spalten des zweiten Faktors sind. Ist also $B = (b_{hi})$ eine $l{\times}m$–Matrix und $A = (a_{ij})$ eine $m{\times}n$–Matrix, so ist das Produkt BA definiert und gleich der $l{\times}n$–Matrix $C = (c_{hj})$ mit den Einträgen

$$\begin{aligned} c_{hj} &= \text{Skalarprodukt der } h\text{-ten Zeile von } B \text{ mit der } j\text{-ten Spalte von } A \\ &= b_{h1}a_{1j} + b_{h2}a_{2j} + \ldots + b_{hm}a_{mj} = \sum_{i=1}^{m} b_{hi}a_{ij}\,. \end{aligned}$$ ■

DISKUSSION: **1)** Bevor man eine Matrixmultiplikation ausführt, sollte man sich zunächst

- *Klarheit über die Formate der Matrixfaktoren und der Produktmatrix beschaffen!*
- *Das Produkt ist nur definiert, wenn die Matrixfaktoren "passen", d.h. wenn die Spaltenzahl des ersten Faktors gleich der Zeilenzahl des zweiten ist;*
- *das Matrixprodukt hat dann so viele Zeilen wie der erste Matrixfaktor und so viele Spalten wie der zweite.*

Die *schlimmsten Fehler*, die man bei der Matrixmultiplikation machen kann, sind das Multiplizieren von Matrizen, deren Formate nicht zusammen passen, oder die Angabe einer Produktmatrix, die nicht das richtige Format hat. (Das kann natürlich nur passieren, wenn man sich nicht an die Definition hält, sondern irgendwelche Vorschriften zur Matrixmultiplikation frei erfindet.) Hat A das Format (m,n) und B das Format $(l,\widetilde{m})$, so muss also zunächst $\widetilde{m} = m$ überprüft werden, damit das Produkt BA überhaupt gebildet werden kann, und dann ist das Ergebnis BA eine Matrix des Formats (l,n). Für die Bedingung $\widetilde{m} = m$ sagt man auch, dass die Matrizen B und A *komponierbar* sind oder *verkettet werden können*, weil diese Bedingung ja gerade bedeutet, dass die Hintereinanderausführung der linearen Operatoren $x \mapsto Ax$ und $y \mapsto By$ möglich ist. Dabei *kommt es auf die Reihenfolge der Matrixfaktoren an*; es ist durchaus möglich, dass B mit A verkettet werden kann ($\widetilde{m} = m$), aber nicht A mit B ($n \neq l$)!

2) Ist *einer der Matrixfaktoren quadratisch*, so hat das Matrixprodukt — wenn es definiert ist — dasselbe Format wie der andere Faktor. Sind *beide Faktoren quadratisch*, so ist das Produkt genau dann definiert, wenn sie dasselbe Format haben, und das Ergebnis der Multiplikation ist dann wieder eine Matrix von diesem quadratischen Format.

3) Bei etwas Übung multipliziert man Matrizen (mit ganzen Zahlen oder einfachen Brüchen als Einträgen) im Kopf, indem man "schielend" mit dem linken Auge eine Zeile der ersten Matrix und mit dem rechten Auge eine Spalte der zweiten Matrix abtastet, dabei das

Skalarprodukt der Zeile mit der Spalte ermittelt und diesen Wert in die richtige Position der Ergebnismatrix einträgt (dieselbe Zeilennummer wie die abgetastete Zeile und dieselbe Spaltennummer wie die abgetastete Spalte). Für weniger Geübte mag folgendes **Rechenschema für die Matrixmultiplikation** hilfreich sein (in der Mathematik für Ökonomen wird das *Falksches Schema* genannt — es ist aber keine so großartige Erfindung, dass man den Namen eines Erfinders hierzu nennen müsste):

$$
\begin{array}{r|l}
\overbrace{\hphantom{b_{11} \cdots b_{hi} \cdots b_{1m}}}^{m} \; m\left\{\vphantom{\begin{matrix} a \\ \vdots \\ a \\ \vdots \\ a \end{matrix}}\right. & \overbrace{\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \cdot & \cdots & a_{ij} & \cdots & \cdot \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}}^{n} = A \\
\hline
l\left\{ B = \begin{pmatrix} b_{11} & \cdots & \cdot & \cdots & b_{1m} \\ \vdots & & \vdots & & \vdots \\ b_{h1} & \cdots & b_{hi} & \cdots & b_{hm} \\ \vdots & & \vdots & & \vdots \\ b_{l1} & \cdots & \cdot & \cdots & b_{lm} \end{pmatrix}\right. & \begin{pmatrix} c_{11} & \cdots & & \cdots & c_{1n} \\ \vdots & & \downarrow & & \vdots \\ & \longrightarrow & c_{hj} & \cdots & \cdot \\ \vdots & & \vdots & & \vdots \\ c_{l1} & \cdots & \cdot & \cdots & c_{ln} \end{pmatrix} = BA
\end{array}
$$

Das Schema zeigt an, dass die Eintragung c_{hj} in Zeile h und Spalte j der Produktmatrix BA alleine aus der h-ten Zeile des ersten Faktors B und der der j-ten Spalte des zweiten Faktors A berechnet wird; allerdings muss man noch wissen *wie*, nämlich als Skalarprodukt von dieser Zeile und dieser Spalte.

4) Man kann die Matrixmultiplikation auch so beschreiben: *Der j-te Spaltenvektor des Produkts $C = BA$ ist das Produkt von B mit dem j-ten Spaltenvektor von A.*

$$
\begin{pmatrix} c_{1j} \\ \vdots \\ c_{lj} \end{pmatrix} = \begin{pmatrix} b_{11} & \ldots & b_{1m} \\ \vdots & & \vdots \\ b_{l1} & \ldots & b_{lm} \end{pmatrix} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.
$$

Für den Fall, dass der zweite Faktor eine $m\times 1$–Matrix ist, also ein Spaltenvektor, stimmt das nun definierte Matrixprodukt mit dem zuvor definierten Produkt einer Matrix mit einem (passenden) Spaltenvektor natürlich überein. Für die Zeilen des Produkts gilt:

Der h-te Zeilenvektor des Produkts $C = BA$ ist das Produkt der h-ten Zeile von B mit der Matrix A.

$$
(c_{h1} \ \ldots \ c_{hm}) = (b_{h1} \ \ldots \ b_{hm}) \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}.
$$

5) Rechenregeln für die Marixmultiplikation:

$$
\begin{aligned}
B(rA) &= r(BA) = (rB)A && (\textit{Homogenität}) \\
(B+\widetilde{B})A &= BA + \widetilde{B}A && (\textit{linkes Distributivgesetz}) \\
B(A+\widetilde{A}) &= BA + B\widetilde{A} && (\textit{rechtes Distributivgesetz})
\end{aligned}
$$

Diese Regeln zusammenfassend sagt man, dass *die Matrixmultiplikation eine bilineare Rechenoperation ist.* Man kann dann auch Linearkombinationen aus jedem Faktor herausziehen, also z.B. $B(rA + \tilde{r}\widetilde{A}) = rBA + \tilde{r}(B\widetilde{A})$. Natürlich müssen die Matrizen dabei zusammen passen, also A und $\widetilde{A}$ von einem Format (m, n) und B, $\widetilde{B}$ von einem Format

(l, m) sein. Der Beweis der Bilinearität ergibt sich unmittelbar aus der schon festgestellten Bilinearität der Skalarproduktbildung oder (weniger elegant und mühsamer) durch direktes Nachrechnen. Für die Multiplikation mit Nullmatrizen und Einheitsmatrizen gilt:

$$0A = 0, \quad B0 = 0; \qquad \mathbb{I}A = A, \quad B\mathbb{I} = B,$$

vorausgesetzt die Matrixprodukte sind definiert. (Dabei steht "0" für Nullmatrizen evtl. verschiedenen Formats und "$\mathbb{I}$" für Einheitsmatrizen evtl. verschiedenen Formats.) Die *Multiplikation mit einer passenden Einheitsmatrix von links oder rechts an eine Matrix reproduziert diese.* Das folgt daraus, dass die Multiplikation einer Einheitsmatrix an einen passenden Spaltenvektor diesen Vektor reproduziert. Und dass das Produkt mit einer passenden Nullmatrix wieder eine Nullmatrix als Ergebnis hat, ist klar.

6) Wenn sie zusammen passen, so kann man auch mehr als zwei Matrizen miteinander multiplizieren. Dabei hängt das Ergebnis nicht von der Klammerung der einzelnen Faktoren ab, d.h. es gilt

$$C(BA) = (CB)A \qquad (\textit{Assoziativgesetz der Matrixmultiplikation}),$$

wenn die Marixprodukte auf einer Seite der Gleichung definiert sind (dann sind die Produkte auf der anderen Seite auch definiert; wenn man also BA und $C(BA)$ bilden kann, so auch CB und $(CB)A$). Das kann direkt mit Matrixeinträgen nachgerechnet werden und sieht dann so aus:

$$\sum_{h=1}^{l} c_{gh} \left(\sum_{i=1}^{m} b_{hi} a_{ij} \right) = \sum_{h=1}^{l} \sum_{i=1}^{m} c_{gh} b_{hi} a_{ij} = \sum_{i=1}^{m} \left(\sum_{h=1}^{l} c_{gh} b_{hi} \right) a_{ij}.$$

Aber eigentlich ist diese Rechnung überflüssig; denn wir brauchen uns nur daran zu erinnern, dass die Matrixmultiplikation der Hintereinanderausführung der zugeordneten linearen Opearatoren entspricht, und die Hintereinanderausführung von Operatoren / Abbildungen ist assoziativ. (Es ist gleich, ob man sich von drei nacheinander auszuführenden Operationen die beiden ersten oder die beiden letzten zu einer einzigen Operation zusammengefasst denkt.) Daher gilt für alle passenden Spaltenvektoren x auch $(C(BA))x = C((BA)x) = C(B(Ax)) = (CB)(Ax) = ((CB)A)x$, und das bedeutet $C(BA) = (CB)A$, weil die linearen Operatoren ja die zugehörigen Matrizen eindeutig bestimmen. (Setze kanonische Einheitsvektoren für x ein.)

7) Wir betonen nochmals, was wir gerade benutzt haben und was die Grundlage des tieferen Verständnisses (und nicht nur der rechnerischen Beherrschung) der Matrixmultiplikation ist:

- *Das Produkt BA zweier Matrizen ist, wenn definiert, die Matrixdarstellung desjenigen linearen Operators, der durch Hintereinanderausführung des zu A gehörenden (ersten) und des zu B gehörenden (nachfolgenden) linearen Operators entsteht.*

Da $x \mapsto Ax$ die zu A gehörende lineare Abbildung ist und $y \mapsto By$ die zu B gehörende, heißt das gerade $(BA)x = B(Ax)$ für alle zu A passenden Spaltenvektoren x, und genau durch diese Bedingung hatten wir ja die Matrixmultiplikation definiert. Man beachte, dass im Produkt BA *der rechts stehende Faktor der zuerst auszuführenden linearen Abbildung entspricht.* Das liegt daran, dass wir Abbildungen — wie allgemein üblich — links vor die Argumente schreiben, also wird bei einer Verkettung $U \circ T(x) = U(T(x))$ auch die rechts stehende Abbildung T zuerst ausgeführt.

(Manche Autoren schreiben deshalb $(x)T$ oder x^T statt $T(x)$; dann erscheinen bei Hintereinanderausführung $((x)T)U = (x^T)^U$ die Abbildungen T und U von links nach rechts gelesen in der Reihenfolge, in der sie ausgeführt werden. Im Matrixkalkül werden nun die Vektoren als Zeilenvektoren geschrieben und die Matrizen von rechts daran multipliziert, um lineare Operatoren zu beschreiben. Dann entspricht im Matrixprodukt AB der erste Faktor auch der zuerst ausgeführten linearen Abbildung. Unter dem Gesichtspunkt, dass wir von links nach rechts lesen, mag diese Vorgehensweise gerechtfertigt sein. Sie ist aber unüblich, und wir folgen der Konvention, zur Beschreibung linearer Operatoren die Vektoren $x \in \mathbb{R}^n$ immer als Spalten zu schreiben und die Matrizen links davor.)

8) Ganz dringende **Warnung:**

- *Für die Matrixmultiplikation gilt kein Kommutativgesetz; die Existenz und der Wert eines Matrixprodukts hängen im Allgemeinen von der Reihenfolge der Faktoren ab!*

Für Matrizen A und B mit beliebigen Formaten (m, n) und (k, l) kann $AB = BA$ im Allgemeinen schon deswegen nicht richtig sein, weil mit der Existenz des einen Produkts AB (d.h. $n = k$) nicht auch das andere Produkt BA definiert (d.h. $l = m$) sein muss. Und auch wenn beide Produkte AB und BA definiert sind (also $n = k$ und $l = m$), so haben sie im Allgemeinen verschiedenes Format (nämlich (n, n) bei BA, aber (m, m) bei AB) und können schon deswegen nicht gleich sein. Schließlich zeigen einfachste Beispiele mit 2×2–Matrizen, dass selbst in dem Fall, wo A und B gleichformatige quadratische Matrizen sind und somit auch AB und BA dasselbe quadratische Format besitzen, trotzdem AB und BA nicht dieselben Einträge in denselben Positionen haben müssen, so dass $AB \neq BA$ ist, obwohl beide Produkte definiert und Matrizen des gleichen Formats sind.

Die Ungültigkeit des Kommutativgesetzes für die Matrizenmultiplikation ist eine häufige Fehlerquelle, weil geläufige Rechenregeln, welche auf der Vertauschbarkeit der Faktoren in einem Produkt beruhen, auf die Matrixmultiplikation angewendet werden, wo sie nicht gelten. Zum Beispiel ist für gleichformatige quadratische Matrizen $(A{+}B)^2 = (A{+}B)(A{+}B) = A^2 + BA + AB + B^2$, und das ist $\neq A^2 + 2AB + B^2$, wenn $AB \neq BA$ ist. (Bei quadratischen Matrizen A ist dabei A^2 natürlich eine Abkürzung für das Produkt AA von A mit sich selbst.) Aus analogem Grund ist für gleichformatige quadratische Matrizen im Allgemeinen $(A{+}B)(A{-}B) \neq A^2 - B^2$. Die geläufigen binomischen Formeln gelten also für die Matrixmultiplikation nicht! Und in einer Summe $AB + BC$ von Matrixprodukten kann man den Faktor B nicht einfach ausklammern (nur $AB + CB = (A{+}C)B$ und $BA + BC = B(A{+}C)$ sind korrekt). Eine *Ausnahme* ist die Multiplikation mit einer Einheitsmatrix $\mathbb{I}$ oder einer skalaren Matrix $r\mathbb{I}$; dafür gilt immer $(r\mathbb{I})A = rA = A(r\mathbb{I})$, wenn die Matrixprodukte erklärt sind, wenn also A und $\mathbb{I}$ dasselbe Format haben.

9) Noch eine Warnung: *Eine Division von Matrizen ist nicht definiert.* Der Quotient "B/A" zweier Matrizen A und B müsste nämlich eine Matrix-Lösung X der Gleichung $AX = B$ sein, aber diese Gleichung ist im Allgemeinen unlösbar, selbst wenn A keine Nullmatrix ist. Außerdem kann es unendlich viele Lösungsmatrizen X geben, und die gleichberechtigte Gleichung $XA = B$ kann ganz andere Lösungen haben. Welche dieser Lösungen X soll man dann zur Definition des Quotienten von B und A nehmen? Darauf gibt es keine sinnvolle Antwort, und deshalb bleibt die Division von Matrizen eben undefiniert.

In diesem Zusammenhang bemerken wir noch, dass *aus dem Verschwinden eines Matrixprodukts $AB = 0$ keineswegs das Verschwinden eines Faktors folgt.* Ein Beispiel erhalten wir, wenn wir $A \neq 0$ als Koeffizientenmatrix eines homogenen linearen Gleichungssystems mit einer nichttrivialen Lösung $x \neq 0$ nehmen; ist B die aus der einzigen Spalte x bestehende Matrix, so gilt auch $B \neq 0$, aber $AB = Ax = 0$.

10) Die Menge der quadratischen Matrizen eines gegebenen Formats (n, n) ist versehen mit einer Addition und einer Multiplikation mit Skalaren, aber auch noch mit einer zusätzlichen Multiplikation *in sich*, weil das Produkt von zwei $n \times n$–Matrizen wieder das Format (n, n) hat. Eine mit solchen Rechenoperationen versehene Menge, für die alle bei der Matrixrechnung oben angegebenen Rechengesetze gelten, heißt in der Mathematik eine *Algebra* (über dem Zahlbereich, dem die Skalare entnommen sind, hier also über $\mathbb{R}$). Der Raum $\mathbb{R}^{n \times n}$ der $n \times n$–Matrizen bildet somit eine Algebra, und zwar für $n \geq 2$ eine sog. *nichtkommutative Algebra*, weil eben für die Multiplikation das Kommutativgesetz $AB = BA$ dann nicht allgemein gilt. Beispiele für kommutative Algebren, in denen man dieselben Rechenoperationen und -gesetze und zusätzlich noch das Kommutativgesetz hat, sind der Raum der Polynomfunktionen auf $\mathbb{R}$, und der Raum der differenzierbaren reellwertigen Funktionen auf einem offenen Intervall in $\mathbb{R}$; hier gilt $fg = gf$ für das Produkt von Funktionen, also Kommutativität. ■

BEISPIELE (*zur Matrixmultiplikation*):

1)
$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 8 & 4 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 4 & 4 \\ 6 & 4 & 2 \end{pmatrix};$$

hier sind die Produkte AB und BA definiert, aber von verschiedenem Format.

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 2 \\ -2 & 2 & -2 \end{pmatrix}, \quad \text{aber} \quad \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \ \text{ist nicht definiert.}$$

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

hier ist $A \neq 0$, $B \neq 0$, $AB = 0$, $BA = 0$, aber $AB \neq BA$ wegen unterschiedlichen Formats.

2) *Produkte mit kanonischen Einheitsvektoren:*

$$\overbrace{(0 \ \ldots \ 0 \ 1 \ 0 \ \ldots \ 0)}^{m} \underset{\uparrow\, i\text{-te Pos.}}{} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (a_{i1} \ \ldots \ a_{in}) = i\text{-te Zeile von } A;$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \\ j\text{-te Position (von } n) \end{array} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = j\text{-te Spalte von } A.$$

3) *Produkt eines Spalten- mit einem Zeilenvektor* (sog. *dyadisches Produkt*):

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} (b_1 \ b_2 \ \dots \ b_n) = \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_mb_1 & a_mb_2 & \cdots & a_mb_n \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

Dies ist für beliebige m und n definiert (passt immer, da der erste Matrixfaktor eine Spalte und der zweite eine Zeile hat). In der anderen Reihenfolge ist das Matrixprodukt dagegen nur definiert, wenn beide Vektoren gleich viele Komponenten haben, und das Ergebnis ist dann das Skalarprodukt der beiden Vektoren a und b (bzw. genauer: die 1×1–Matrix mit dem Skalarprodukt als einzigem Eintrag):

$$(b_1 \ b_2 \ \dots \ b_n) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = (a_1b_1 + a_2b_2 + \ldots + a_nb_n) = (a \bullet b)\,.$$

4) Die *Multiplikation mit einer Diagonalmatrix* $\begin{pmatrix} d_1 & & 0 \\ & d_2 & \\ 0 & \ddots & d_n \end{pmatrix}$ an eine beliebige (passende!) Matrix A hat folgenden Effekt: Bei Multiplikation von D rechts an A werden die Spalten von A der Reihe nach mit den Diagonaleinträgen von D vervielfacht, bei Multiplikation von D links an A dagegen die Zeilen :

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} d_1 & & & & \\ & \ddots & & 0 & \\ & & d_j & & \\ & 0 & & \ddots & \\ & & & & d_n \end{pmatrix} = \begin{pmatrix} d_1a_{11} & \cdots & d_ja_{1j} & \cdots & d_na_{1n} \\ \vdots & & \vdots & & \vdots \\ d_1a_{m1} & \cdots & d_ja_{mj} & \cdots & d_na_{mn} \end{pmatrix},$$

$$\begin{pmatrix} d_1 & & & & \\ & \ddots & & 0 & \\ & & d_i & & \\ & 0 & & \ddots & \\ & & & & d_m \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} d_1a_{11} & \cdots & d_1a_{1n} \\ \vdots & & \vdots \\ d_ia_{i1} & \cdots & d_ia_{in} \\ \vdots & & \vdots \\ d_ma_{m1} & \cdots & d_ma_{mn} \end{pmatrix}.$$

Es ist also bei einer $n\times n$-Matrix A *nicht* gleichgültig, ob man A mit der Diagonalmatrix von rechts oder von links multipliziert. Ausnahme: D ist eine skalare Matrix, also eine Diagonalmatrix mit lauter gleichen Diagonaleinträgen $d_1 = \ldots = d_n = r$; dann ist $AD = DA$ für alle $A \in \mathbb{R}^{n\times n}$ (und *nur dann*, wie man beweisen kann). Weitere Ausnahme: A ist selbst eine Diagonalmatrix vom gleichen Format wie D; dann gilt $AD = DA$; denn:

- *das Produkt von zwei $n\times n$–Diagonalmatrizen ist wieder eine Diagonalmatrix, und ihre Einträge sind die Produkte der gleich positionierten Einträge beider Faktoren.*

5) *Zeilen- bzw. Spaltenoperationen mit einer Matrix A* kann man durch Multiplikation passender quadratischer Matrizen von links bzw. von rechts an A beschreiben. Diese Matrizen sehen folgendermaßen aus (alle quadratisch vom gleichen Format (n, n), die nicht angegebenen Einträge sind 0 außerhalb der Diagonalen und 1 auf der Diagonalen) und heißen

Elementarmatrizen:

$$M_i(r) := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & r & \cdots\cdots & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile, } r \in \mathbb{R}_{\neq 0} \\ \\ \\ \\ \end{matrix}$$

$$A_{ij}(s) := \begin{array}{c} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & s \cdots\cdots & \\ & & & \ddots \vdots & \\ & & & \vdots\; 1 & \\ & & & \vdots & 1 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow i\text{-te Zeile, } s \in \mathbb{R} \\ \\ \\ \\ \end{matrix} \\ \uparrow j\text{-te Spalte, } j \neq i \end{array}$$

$$V_{ij} := \begin{array}{c} \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & \cdots\cdots \\ & & \vdots & \ddots & \vdots & \\ & & 1 & \cdots & 0 & \cdots\cdots \\ & & \vdots & & \vdots & \ddots \\ & & \vdots & & \vdots & & 1 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \leftarrow j\text{-te Zeile } (j \neq i) \\ \\ \\ \end{matrix} \\ i\text{-te Spalte} \uparrow \qquad \uparrow j\text{-te Spalte} \end{array}$$

$M_i(r)$ ist die Diagonalmatrix mit Eintrag $r \neq 0$ in i-ter Diagonalposition und anderen Diagonaleinträgen 1, $A_{ij}(s)$ entsteht aus der Einheitsmatrix, indem man den Nulleintrag in Position (i, j) durch den Skalar s ersetzt, und V_{ij} erhält man aus der Einheitsmatrix durch Vertauschung der i-ten mit der j-ten Zeile (oder der i-ten mit der j-ten Spalte, das läuft hier auf Dasselbe hinaus). Allgemein bezeichnet man eine Matrix, die aus der Einheitsmatrix durch Vertauschungen von Zeilen und / oder Spalten hervorgeht, als *Permutationsmatrix*, weil ihre Multiplikation an einen passenden Spaltenvektor bewirkt, dass dessen Komponenten entsprechend permutiert (vertauscht) werden. Die Multiplikation einer Elementarmatrix von links an eine passende Matrix A bewirkt nun eine Zeilenoperation bei A wie folgt:

$M_i(r)A$ entsteht aus A durch Multiplikation der i-ten Zeile mit r,

$A_{ij}(s)A$ entsteht aus A durch Addition des s-fachen der j-ten Zeile zur i-ten;

$V_{ij}A$ entsteht aus A durch Vertauschung der i-ten mit der j-ten Zeile.

Entsprechend bewirkt die Multiplikation einer Elementarmetrix von rechts an eine passende Matrix B eine Spaltenoperation bei B:

$BM_j(r)$ entsteht aus B durch Multiplikation der j-ten Spalte mit r,

$BA_{ij}(s)$ entsteht aus B durch Addition des s-fachen der i-ten Spalte zur j-ten;

BV_{ij} entsteht aus B durch Vertauschung der i-ten mit der j-ten Spalte.

Nun haben wir in 3.1 gesehen, dass jede $m \times n$–Matrix A durch Zeilenoperationen auf Zeilen-Stufen-Form gebracht werden kann. Folglich gibt es $m \times m$–Elementarmatrizen $E_1, \ldots, E_p$, derart dass das Matrixprodukt $Z = E_p \ldots E_2 E_1 A$ eine $m \times n$–Matrix von Zeilen-Stufen-Form ist. Durch Spaltenoperationen, also durch Multiplikationen mit weiteren $n\times n$–Elementarmatrizen $F_1, \ldots, F_q$ von rechts, kann man die Zeilen-Stufen-Matrix Z in die Form einer $m\times n$–Blockmatrix bringen mit einer $l\times l$–Einheitsmatrix $\mathbb{I}_l$ als linker oberer Block und Nullmatrizen als weiteren Blöcken. Dabei ist l die Anzahl der Stufen (also der Zeilen mit gewissen Einträgen $\neq 0$) in der Zeilen-Stufen-Matrix Z, also der Rang der Matrix A. Resultat ist eine Darstellung

$$E_p \cdots E_2 E_1\, A\, F_1 F_2 \cdots F_q \;=\; \left(\begin{array}{c|c} \mathbb{I}_l & 0_{l,n-l} \\ \hline 0_{m-l,l} & 0_{m-l,n-l} \end{array}\right),$$

wo die Subskripte die Formate der Nullmatrizen anzeigen. Nun kann man die Zeilen- und Spaltenoperationen durch Multiplikation mit Elementarmatrizen von links und rechts wieder rückgängig machen und erhält so Gleichheit von A mit der entsprechend umgeformten Matrix der rechten Seite. Auf diese Weise bekommt man die Darstellung

$$A \;=\; E_1 E_2 \ldots E_p \left(\begin{array}{c|c} \mathbb{I}_l & 0_{l,n-l} \\ \hline 0_{m-l,l} & 0_{m-l,n-l} \end{array}\right) F_q \ldots F_2 F_1$$

einer beliebigen $m\times n$–Matrix A mit Rang l als Produkt von Elementarmatrizen $E_1, \ldots, E_p$, $F_1, \ldots, F_q$ (andere als oben) und der Blockmatrix aus einem Bolck $\mathbb{I}_l$ und Nullblöcken sonst. Im Fall $m = n = l$ ist diese Blockmatrix die Einheitsmatrix $\mathbb{I}_n$, und daher gilt:

- *Jede $n\times n$–Matrix von maximalem Rang $l = n$ lässt sich als Produkt von $n\times n$–Elementarmatrizen schreiben.*

6) Potenzen einer Matrix A kann man bilden, wenn sie quadratisch ist, also $A^1 := A$, $A^2 := AA$, $A^3 := AAA$ usw. Es gelten dann die Potenzgesetze $A^k A^l = A^{k+l}$ und $(A^k)^l = A^{kl}$ für natürliche Exponenten (aber nicht $(BA)^k = B^k A^k$, wenn $BA \neq AB$ ist, weil man dann z.B. in $(BA)^2 = BABA$ die beiden mittleren Faktoren nicht vertauschen darf; man definiert auch noch $A^0 := \mathbb{I}$ als die Einheitsmatrix desselben Formats wie A). Zum Beispiel gilt:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \Longrightarrow A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ A^3 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ \ldots,\ A^k = A \quad \text{für alle } k \in \mathbb{N};$$

Matrizen wie diese, deren Potenzen alle gleich der Matrix selbst sind, heißen *idempotente Matrizen.* Ein anderes Beispiel, in dem bei der Potenzbildung auch etwas Unerwartetes passiert, ist:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Longrightarrow A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{I}_2,\ A^3 = A,\ A^4 = \mathbb{I}_2,\ A^5 = A,\ A^6 = \mathbb{I}_2, \ldots.$$

Matrizen wie diese, deren Quadrat die Einheitsmatrix ist, heißen *involutorische Matrizen.* Ein drittes Beispiel:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ \Longrightarrow A^k = 0 \quad \text{für } k \geq 3.$$

Hier haben wir eine *nilpotente Matrix,* d.h. eine, die nicht selbst die Nullmatrix ist, aber eine Nullmatrix als Potenz hat. Die $n\times n$–Matrix A mit Einträgen $a_{i,i+1} = 1$ in der ersten Schrägzeile oberhalb der Diagonalen und Nulleinträgen sonst, hat allgemein den *Nilpotenzgrad* $n-1$, d.h. es gilt $A^k = 0$ für $k \geq n$, aber $A^k \neq 0$ für $1 \leq k < n$.

7) Bei sog. **gestaffelten Produktionsabläufen** werden gewisse Endprodukte Nr. $1 \dots n$ aus Zwischenprodukten Nr. $1 \dots m$ gefertigt gemäß einer Endproduktionsverflechtungsmatrix $S = (s_{ij})$, und die Zwischenprodukte werden ihrerseits aus $1 \dots l$ Rohstoffen hergestellt gemäß einer Zwischenproduktionsverflechtungsmatrix $R = (r_{hi})$. Die Einträge r_{hi} von R heißen *Rohstoffverbrauchskoeffizienten* und geben an, wieviele Mengeneinheiten des h-ten Rohstoffs für die Herstellung einer Einheit des Zwischenprodukts Nr. i benötigt werden. Entsprechend beschreiben die *Zwischenproduktverbrauchskoeffizienten* s_{ij}, wieviele Einheiten des i-ten Zwischenprodukts man zur Produktion einer Einheit des Endprodukts Nr. j braucht. Man kann dabei auch zulassen, dass ein Rohstoff direkt in die Endproduktion eingeht, indem man ihn einfach als Zwischenprodukt mit Verbrauchskoeffizient 1 für den betreffenden Rohstoff und mit Verbrauchskoeffizient 0 für die anderen Rohstoffe führt. Jedoch wollen wir hier den Fall ausschließen, dass ein Endprodukt in der Produktion von Zwischenprodukten oder von Endprodukten benötigt wird. (Eine derartige Situation haben wir in 3.2 für eine Produktion ohne Staffelung behandelt und ein lineares Gleichungssystem für den Produktionsvektor aufgestellt; ähnlich könnte man auch hier bei der gestaffelten Produktion verfahren, doch wird dann alles noch komplizierter.)

Ist $\mathbf{x} \in \mathbb{R}^n$ ein gegebener *Outputvektor der Endproduktion (Produktionsplan)*, dessen Komponenten x_j die zu produzierenden Mengeneinheiten des j-ten Endprodukts angeben, so ist $y_i = s_{i1}x_1 + s_{i2}x_2 + \ldots + s_{in}x_n$ die für die Endproduktion bereitzustellende Menge des i-ten Zwischenprodukts. Die y_i sind die Komponenten des *Inputvektors für die Endproduktion* $\mathbf{y} \in \mathbb{R}^m$, der gleichzeitig der *Outputvektor der Zwischenproduktion* ist (wenn ohne Überschuss zwischenproduziert wird). In Matrixschreibweise kann der lineare Zusammenhang zwischen $\mathbf{x}$ und $\mathbf{y}$ durch $\mathbf{y} = S\mathbf{x}$ ausgedrückt werden. Entsprechend ist der Bedarf z_h an Einheiten des h-ten Rohstoffs für die Zwischenproduktion gegeben durch $z_h = r_{h1}y_1 + r_{h2}y_2 + \ldots + r_{hm}y_m$, und der *Rohstoffverbrauchsvektor (Rohstoffbedarfsvektor)* $\mathbf{z} \in \mathbb{R}^l$, also der Inputvektor für die Zwischenproduktion, berechnet sich gemäß $\mathbf{z} = R\mathbf{y}$ in linearer Weise aus $\mathbf{y}$. Da $\mathbf{y}$ seinerseits linear von $\mathbf{x}$ abhängt, ist $\mathbf{z}$ eine mittelbar von $\mathbf{x}$ linear abhängige Variable, d.h. man erhält $\mathbf{z}$ aus $\mathbf{x}$ durch Hintereinanderschaltung von zwei linearen Operatoren: $\mathbf{z} = R\mathbf{y} = R(S\mathbf{x})$. Nun ist das Matrixprodukt RS gerade so motiviert und definiert worden, dass $R(S\mathbf{x}) = (RS)\mathbf{x}$ gilt, d.h. die Hintereinanderausführung der beiden Multiplikationen mit S und R kann ersetzt werden durch eine einzige lineare Operation, nämlich die Multiplikation mit RS.

- *Der unmittelbare lineare Zusammenhang zwischen dem Endproduktoutputvektor* $\mathbf{x}$ *und dem Rohstoffverbrauchsvektor* $\mathbf{z}$ *ist gegeben durch das Produkt der Zwischenproduktionsmatrix* R *mit der Endproduktionsmatrix* S, *d.h.* $\mathbf{z} = (RS)\mathbf{x}$.

Die Überlegungen lassen sich ausdehnen auf einen *mehrstufigen Produktionsprozess*, bei dem man $k \geq 2$ Produktionsstufen hat, derart dass der Outputvektor $\mathbf{x}_{h-1}$ der $(h-1)$-ten Stufe jeweils der Inputvektor (Bedarfsvektor) für die Produktion der nächsten Stufe h ist. Der Inputvektor für die erste Produktionsstufe ist dann der Rohstoffverbrauchsvektor $\mathbf{x}_0$, und der Outputvektor der letzten Stufe ist der Endproduktoutputvektor $\mathbf{x}_k$. Man hat dann k Produktionsverflechtungsmatrizen $A_1, A_2, \ldots, A_k$, wobei die Einträge von $A_h = (a_{ij}^{(h)})$ angeben, wieviele Einheiten des i-ten Produkts der Stufe $h-1$ für die Herstellung einer Einheit des j-ten Produkts der Stufe h erforderlich sind. Der Zusammenhang zwischen $\mathbf{x}_h$ und $\mathbf{x}_{h-1}$ ist also wie oben $\mathbf{x}_{h-1} = A_h\mathbf{x}_h$ für $h = 1 \dots k$. Somit gilt

$$\mathbf{x}_0 = A_1\mathbf{x}_1 = A_1A_2\mathbf{x}_2 = \ \ldots \ = A_1A_2 \cdots A_k\mathbf{x}_k\,,$$

d.h. der unmittelbare lineare Zusammenhang zwischen dem Endproduktoutputvektor $\mathbf{x}_k$ und dem Rohstoffbedarfsvektor $\mathbf{x}_0$ ist durch das Produkt aller k Produktionsmatrizen (in der Reihenfolge von links nach rechts aufsteigender Produktionsstufen) gegeben.

8) In analoger Weise kommen, wie schon gesagt, Matrixprodukte überall ins Spiel, wo (vektorielle) ökonomische Größen mittelbar über eine Kette von linearen Abhängigkeitsgesetzen durch andere ökonomische (vektorielle) Variablen bestimmt sind. Noch ein Beispiel dieser Art: Wird die *Kundenwanderung zwischen n konkurrierenden Produkten* in einer ersten Beobachtungsperiode durch die Übergangsmatrix $P_1 \in \mathbb{R}^{n\times n}$ beschrieben und in den folgenden Perioden durch $P_2, \ldots, P_k \in \mathbb{R}^{n\times n}$, so ist das Matrixprodukt $P = P_k \cdots P_2 P_1$ die Übergangsmatrix, welche den Kundenwechsel im gesamten Zeitraum vom Beginn der ersten bis zum Ende der k-ten Beobachtungsperiode darstellt. (Daraus folgt übrigens, dass ein Produkt von stochastischen Matrizen wie die P_h hier, also Matrizen mit der Zeilensumme 1, wieder eine stochastische Matrix ist. Das sieht man auch durch Multiplikation der Matrizen von links an einen summierenden Spaltenvektor, also einen mit lauter Einträgen 1; denn Stochastizität der Matrix bedeutet gerade, dass ein solches Matrix–an–Vektor–Produkt als Ergebnis wieder einen summierenden Spaltenvektor hat.)

9) Matrixprodukte treten in ökonomischem Zusammenhang auch dann auf, wenn man mehrere Vorgänge zusammenfasst, die durch Multiplikation einer festen Matrix an verschiedene Spaltenvektoren beschrieben werden. Sind z.B. in einer Produktion für k verschiedene Perioden Produktionsoutputvektoren $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ vorgesehen und die zugehörigen Rohstoffverbrauchsvektoren $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k \in \mathbb{R}^m$ berechnet, so erhebt sich die Frage nach dem *günstigsten Rohstoffeinkauf*. Kommen dafür Lieferanten $L_1, \ldots, L_l$ in Frage, wobei der h-te Lieferant für den i-ten Rohstoff den Preis p_{hi} verlangt, so gibt das Produkt der *Preismatrix* $P = (p_{hi})$ mit einem Rohstoffverbrauchsvektor $\mathbf{z}$ den zugehörigen *Rohstoffkostenvektor* $P\mathbf{z}$, dessen h-te Komponente $p_{h1}z_1 + \ldots + p_{hm}z_m$ die Kosten beim Bezug aller in $\mathbf{z}$ angegebenen Rohstoffmengen vom Lieferanten L_h angibt für $h = 1 \ldots l$.

Hier kann man nun die aus dem Produktionsplan resultierenden Rohstoffverbrauchsvektoren $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k$ als Spalten zu einer $m\times k$–*Verbrauchsmatrix* Z zusammenstellen. Dann gibt die j-te Eintragung in der h-ten Zeile des Matrixprodukts PZ die Kosten der für die Realisierung des j-ten Produktionsoutputvektors benötigten Rohstoffe an, wenn diese alle beim Lieferanten L_h bezogen werden. Unterstellt, der Bezug aller Rohstoffe bei einem einzigen Lieferanten ist (wegen Rabatten bei größeren Bestellungen) in jeder Periode günstiger als die Beschaffung bei verschiedenen Lieferanten, so wird man also in jeder Spalte von PZ die kleinste Zahl aufsuchen und dann, wenn diese bei der j-ten Spalte etwa in h-ter Position steht, die Rohstoffe für die j-te Produktionsperiode beim Lieferanten L_h ordern.

Bei Beispielen dieser Art ist die Verwendung eines Matrixprodukts allerdings weder notwendig, noch besonders hilfreich — man kann genau so gut die einzelnen Vorgänge für jede Periode separat durchrechnen. Die mathematische Grundlage der *Zusammenfassung von mehreren Matrix–an–Vektor–Multiplikationen zu einer Matrix–Matrix–Multiplikation* ist folgende schon früher gemachte Beobachtung:

- *Die j-te Spalte des Matrixprodukts BA ist das Produkt der Matrix B mit der j-ten Spalte von A;*
- *die i-te Zeile des Matrixprodukts BA ist das Produkt der i-ten Zeile von B mit der Matrix A.*

10) Ein letztes Beispiel zur Matrixmultiplikation aus der Ökonomie: In einem Unternehmen werden an Produktionsstätten ①, …, ⓜ Produkte $\boxed{1}, \ldots, \boxed{n}$ gefertigt gemäß einer $m\times n$–*Mengenproduktionsmatrix* $X = (x_{ij})$, wobei x_{ij} die Zahl der von Produkt Nr. j an der Produktionsstätte Nr. i hergestellten Einheiten ist. Dazu gehört die gleichformatige *Kostenproduktionsmatrix* $K = (k_{ij})$, wobei der *Kostenfaktor* k_{ij} die Kosten der Produktion einer Einheit des Produkts Nr. j an der Produktionsstätte (i) angibt, also die Stückkosten für das Produkt an dieser Produktionsstätte. Die Summe

$$k_{i1}x_{i1} + k_{i2}x_{i2} + \ldots + k_{in}x_{in} = \sum_{j=1}^{n} k_{ij}x_{ij}$$

gibt dann die gesamten Kosten der Produktionsstätte (i) an, wenn gemäß dem durch X gegebenen Produktionsplan produziert wird.

Um diese Summe als Eintrag in einem Matrixprodukt zu interpretieren bilden wir die **transponierte Matrix** ("gekippte Matrix")

$$K^{\mathsf{T}} = \begin{pmatrix} k_{11} & k_{21} & \ldots\ldots & k_{m1} \\ k_{12} & k_{22} & \ldots\ldots & k_{m2} \\ \vdots & \vdots & & \vdots \\ k_{1n} & k_{2n} & \ldots\ldots & k_{mn} \end{pmatrix} \quad \text{zu} \quad K = \begin{pmatrix} k_{11} & k_{12} & \ldots & k_{1n} \\ k_{21} & k_{22} & \ldots & k_{2n} \\ \vdots & \vdots & & \vdots \\ k_{m1} & k_{m2} & \ldots & k_{mn} \end{pmatrix}.$$

Die Zeilen von K^{T} sind also die Spalten von K, und die Spalten von K^{T} sind die Zeilen von K. (Hat K das Format (m,n) so ist K^{T} eine Matrix vom Format (n,m). Nur bei quadratischen Matrizen K hat K^{T} dasselbe Format. Ist sogar $K = K^{\mathsf{T}}$, also $k_{ij} = k_{ji}$ für alle i,j, so heißt K eine – zur Diagonalen – **symmetrische Matrix**. Andere gebräuchliche Bezeichnungen für die transponierte Matrix zu K sind K^t oder K'; manchmal wird das Superskript auch vorangestellt, also ${}^{\mathsf{T}}K$ bzw. ${}^{t}K$ geschrieben.)

Nun ist das Matrixprodukt

$$XK^{\mathsf{T}} \in \mathbb{R}^{m\times m}$$

definiert, und sein i-ter Eintrag auf der Diagonalen ist gerade die obige Summe $\sum_{j=1}^{n} x_{ij}k_{ij}$, welche die Kosten der Produktionsstätte (i) angibt. In diesem Sinne können wir sagen;

- *Die Diagonaleinträge der $m\times m$–Matrix XK^{T} schlüsseln die Kosten auf die m Produktionsstätten auf.*

(Die in nichtdiagonalen Positionen (h,i) befindlichen Einträge der Matrix sind weniger relevant; sie geben die Kosten an der Produktionsstätte (i) an, wenn dort nach dem eigentlich für (h) vorgesehenen Produktionsplan produziert würde.) Auch in der anderen Reihenfolge ist das Produkt von X und K^{T} definiert, nun allerdings eine $n\times n$–Matrix:

$$K^{\mathsf{T}}X \in \mathbb{R}^{n\times n}.$$

Der j-te Diagonaleintrag dieser Matrix ist

$$k_{1j}x_{1j} + k_{2j}x_{2j} + \ldots + k_{mj}x_{mj} = \sum_{i=1}^{m} k_{ij}x_{ij}$$

und gibt die Kosten für die Erzeugung des Produkts Nr. j saldiert über alle m Produktionsstätten an. In diesem Sinne gilt also:

- *Die Diagonaleinträge der $n\times n$–Matrix $K^{\mathsf{T}}X$ schlüsseln die Kosten auf die n Produkte auf.*

(Die Nichtdiagonaleinträge beschreiben hier, was die Produktion des Produkts kosten würde, wenn man dafür die Stückkosten eines anderen Produkts hätte; diese Einträge haben also keine ökonomische Bedeutung.)

Die Summe aller Diagonaleinträge ist bei $XK^\top$ und bei $K^\top X$ dieselbe,

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n} x_{ij}k_{ij}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} k_{ij}x_{ij}\right),$$

und das ist auch ökonomisch klar; denn beide Summen geben ja die *Gesamtkosten* der Produktion an, nur einmal aufgegliedert in die Kosten der einzelnen Produktionsstätten und einmal in die Gesamtkosten für die einzelnen Produkte. Man nennt die Summe der Diagonaleinträge einer quadratischen Matrix ihre *Spur* und kann daher formulieren

- *Die Spur von $XK^\top$ und die von $K^\top X$ geben die Gesamtkosten $\sum_{i=1}^{m}\sum_{j=1}^{n} k_{ij}x_{ij}$ der Produktion an.* ∎

DEFINITION: Die **Spur einer quadratischen Matrix** $A = (a_{ij})_{1\le i,j\le n}$ ist die Summe ihrer Diagonaleinträge, $\mathrm{Spur}(A) := a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii}$. ∎

DISKUSSION: **1)** In den Überlegungen im letzten Beispiel war X eine eigentlich ganz beliebige $m\times n$–Matrix und $K^\top$ eine beliebige $n\times m$–Matrix. Daher zeigen diese Überlegungen ganz allgemein den **Spursatz**:

$$\mathrm{Spur}(AB) = \mathrm{Spur}(BA) \qquad \text{für alle } A \in \mathbb{R}^{m\times n} \text{ und } B \in \mathbb{R}^{n\times m}.$$

Obwohl also im Allgemeinen $AB \neq BA$ ist, im Fall $m \neq n$ haben diese beiden quadratischen Matrizen ja sogar verschiedenes Format, besitzen AB und BA doch stets dieselbe Summe von Diagonaleinträgen, wenn beide Matrixprodukte definiert sind. In der Tat, ist $A = (a_{ij})$ und $B = (b_{jk})$, so hat AB die Diagonaleinträge $\sum_{j=1}^{n} a_{ij}b_{ji}$ für $i = 1\ldots m$ und BA die Diagonaleinträge $\sum_{i=1}^{m} b_{ji}a_{ij}$ für $j = 1\ldots n$, und die Summe aller Diagonaleinträge ist in beiden Fällen

$$\mathrm{Spur}(AB) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji}.$$

2) Bildet man zu einer Matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ die Spur des Produkts $AA^\top$, so erhält man die Quadratsumme aller Einträge der Matrix:

$$\mathrm{Spur}(AA^\top) = \sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij})^2.$$

Daher ist die Wurzel aus der Spur ein Analogon der Euklidischen Norm von Vektoren aus $\mathbb{R}^n$, die ja als Quadratwurzel aus der Quadratsumme der Komponenten definiert ist. Als ein sinnvolles Maß für die "absolute Gesamtgröße" einer Matrix (eines von vielen verschiedenen in der Mathematik gebräuchlichen) definieren wir daher die **Euklidische Norm einer Matrix** als Quadratwurzel aus der Summe der Quadrate aller Einträge:

$$|A| := \sqrt{\mathrm{Spur}(AA^\top)} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij})^2}.$$

Man kann durch $A \bullet B := \mathrm{Spur}(AB^\top) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji}$ auch ein Skalarprodukt auf dem Raum $\mathbb{R}^{m\times n}$ der $m\times n$–Matrizen einführen, das zum üblichen Skalarprodukt auf $\mathbb{R}^n$ analog ist. Dann gilt auch hier, dass die Norm die Wurzel aus dem Skalarprodukt einer Matrix mit sich selbst ist, $|A| = \sqrt{A \bullet A}$. ∎