Beweis: Jeder Vektor $\vec{v} \in V$ läßt sich in eindeutiger Weise als Linearkombination $\vec{v} = \lambda_1 \vec{b}_1 + \cdots + \lambda_n \vec{b}_n$ mit $\vec{b}_1, \ldots, \vec{b}_n \in \mathcal{B}$ darstellen, und für eine lineare Abbildung φ muß dann gelten

$$\varphi(\vec{v}) = \lambda_1 \varphi(\vec{b}_1) + \cdots + \lambda_n \varphi(\vec{b}_n) .$$

■

Besonders nützlich ist dies im Fall endlichdimensionaler Vektorräume.

Sei etwa V ein m-dimensionaler k-Vektorraum und W ein n-dimensionaler; wir wählen Basen

$$\mathcal{B} = \left(\vec{b}_1, \ldots, \vec{b}_m\right) \quad \text{und} \quad \mathcal{C} = \left(\vec{c}_1, \ldots, \vec{c}_n\right) .$$

Eine lineare Abbildung $\varphi\colon V \to W$ ist, wie wir gerade gesehen haben, eindeutig bestimmt durch die Bilder der Basisvektoren $\vec{b}_j$; diese wiederum lassen sich als Linearkombinationen der Basisvektoren $\vec{c}_i$ schreiben:

$$\varphi\left(\vec{b}_j\right) = a_{1j}\vec{c}_1 + a_{2j}\vec{c}_2 + \cdots + a_{nj}\vec{c}_n \quad \text{mit} \quad a_{ij} \in k .$$

Somit ist φ bei gegebenen Basen eindeutig bestimmt durch die $n \cdot m$ Skalare a_{ij}. Wir fassen diese zusammen zu einer *Matrix:*

Definition: *a)* Eine $n \times m$-Matrix A über dem Körper k ist eine zweidimensionale Anordnung von Körperelementen $a_{ij} \in k$ in n Zeilen und m Spalten, d.h.

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix} .$$

b) Die Menge aller $n \times m$-Matrizen A über k bezeichnen wir mit $k^{n \times m}$.

Die Matrix A zur linearen Abbildung φ bezeichen wir als *Abbildungsmatrix* von φ bezüglich der (geordneten) Basen $\mathcal{B}$ und $\mathcal{C}$; für eine lineare Abbildung $\varphi\colon V \to V$, deren Zielraum gleich derUrbildmenge ist, wählen wir im allgemeinen nur eine Basis $\mathcal{B}$ von V, setzen also $\mathcal{C} = \mathcal{B}$, und reden dann von der Abbildungsmatrix bezüglich der Basis $\mathcal{B}$.

Wenn $\mathcal{B}$ und $\mathcal{C}$ vorgegeben sind, gibt es offensichtlich für jede Matrix $A \in k^{n \times m}$ eine lineare Abbildung $\varphi\colon V \to W$ mit A als Abbildungsmatrix, nämlich diejenige lineare Abbildung, für die

$$\varphi\left(\vec{b}_j\right) = a_{1j}\vec{c}_1 + a_{2j}\vec{c}_2 + \cdots + a_{nj}\vec{c}_n$$

ist. Bei gegebenen Basen entsprechen sich lineare Abbildungen und Matrizen also eineindeutig: Zu jeder linearen Abbildung gibt es *genau* eine Matrix und zu jeder Matrix *genau* eine lineare Abbildung.

Ein wesentlicher Punkt ist hier, daß es sich bei einer linearen Abbildung $\varphi\colon V \to W$ im allgemeinen um eine Abbildung zwischen *unendlichen* Mengen handelt und solche Abbildungen nur selten mit endlichem Aufwand beschrieben werden können. (Wie sieht etwa eine „allgemeine" Abbildung $\varphi\colon \mathbb{R} \to \mathbb{R}$ aus?) Eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen ist, wie wir gerade gesehen haben, nach Wahl von Basen durch die endlich vielen Einträge der Abbildungsmatrix eindeutig bestimmt und damit einer algorithmischen Behandlung zugänglich.

Matrizen als zweidimensionale Zahlenschemata sind natürlich erheblich älter als Vektorräume und lineare Abbildungen; erste Spuren aus dem zweiten vorchristlichen Jahrhundert finden sich bereits in den *Neun Büchern der Rechenkunst* 九章算術 aus der chinesischen Han-Dynastie. Rechenregeln für den Umgang mit Matrizen tauchen ab dem 16. Jahrhundert bei den verschiedensten Autoren auf.

Als Beispiel betrachten wir den Vektorraum V aller reeller Polynome vom Grad höchstens vier und den Vektorraum W aller reeller Polynome vom Grad höchstens drei zusammen mit der linearen Abbildung

$$\varphi\colon V \to W; \quad f \mapsto f' .$$

Bevor wir eine Abbildungsmatrix berechnen können, brauchen wir zunächst Basen der beiden Vektorräume, zum Beispiel die „üblichen" Basen $\mathcal{B} = (1, X, X^2, X^3, X^4)$ und $\mathcal{C} = (1, X, X^2, X^3)$. Dann ist

$$\begin{aligned}
\varphi(1) &= 0 &&= 0 \cdot 1 + 0 \cdot X + 0 \cdot X^2 + 0 \cdot X^3 \\
\varphi(X) &= 1 &&= 1 \cdot 1 + 0 \cdot X + 0 \cdot X^2 + 0 \cdot X^3 \\
\varphi(X^2) &= 2X &&= 0 \cdot 1 + 2 \cdot X + 0 \cdot X^2 + 0 \cdot X^3 \\
\varphi(X^3) &= 3X^2 &&= 0 \cdot 1 + 0 \cdot X + 3 \cdot X^2 + 0 \cdot X^3 \quad \text{und} \\
\varphi(X^4) &= 4X^3 &&= 0 \cdot 1 + 0 \cdot X + 0 \cdot X^2 + 4 \cdot X^3 ,
\end{aligned}$$

die Abbildungsmatrix ist also

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

Hier wie auch im allgemeinen Fall stehen in den *Spalten* der Abbildungsmatrix die Koeffizienten der Bilder der Basisvektoren von V, ausgedrückt bezüglich der Basis von W.

b) Rechenregeln für Matrizen

Wir haben Matrizen eingeführt, um mit linearen Abbildungen konkret rechnen zu können; als erstes sollten wir uns dazu überlegen, wie man mit *Matrizen* rechnen kann.

Zu zwei $n \times m$-Matrizen

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \text{ und } B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix}$$

können wir deren Summe

$$A + B \underset{\text{def}}{=} \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2m}+b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{nm}+b_{nm} \end{pmatrix}$$

definieren, und für einen Skalar $\lambda \in k$ auch das Produkt

$$\lambda A \underset{\text{def}}{=} \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix} .$$

Das legt die Vermutung nahe, daß $k^{n\times m}$ mit diesen beiden Verknüpfungen ein Vektorraum sein könnte, und in der Tat gilt:

Lemma: $k^{n\times m}$ ist ein k-Vektorraum der Dimension nm.

Beweis: Eigentlich gibt es nichts zu beweisen, denn wir haben einfach die Elemente aus dem Vektorraum $k^{(nm)}$ anders hingeschrieben, ohne daß dabei an den Rechenoperationen etwas zu ändern. Für diejenigen, die den Umgang mit Vektorraumaxiomen und das Rechnen mit Matrizen noch etwas üben wollen, sei aber trotzdem ein ausführlicher Beweis gegeben:

Beide Rechenoperationen sind so definiert, daß, wenn wir den ij-Eintrag für sich alleine betrachten, dort die entsprechende Rechenoperation im Grundkörper k ausgeführt wird. Da für die Rechenoperationen im Grundkörper alle bei der Definition eines Vektorraums geforderten Rechenregeln gelten, gelten sie auch in $k^{n\times m}$. Beispielsweise ist also

$$\begin{aligned} A + B &= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2m}+b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{nm}+b_{nm} \end{pmatrix} \\ &= \begin{pmatrix} b_{11}+a_{11} & b_{12}+a_{12} & \dots & b_{1m}+a_{1m} \\ b_{21}+a_{21} & b_{22}+a_{22} & \dots & b_{2m}+a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1}+a_{n1} & b_{n2}+a_{n2} & \dots & b_{nm}+a_{nm} \end{pmatrix} = B + A \end{aligned}$$

und

$$\begin{aligned} \lambda(\mu A) &= \lambda \begin{pmatrix} \mu a_{11} & \mu a_{12} & \dots & \mu a_{1m} \\ \mu a_{21} & \mu a_{22} & \dots & \mu a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \mu a_{n1} & \mu a_{n2} & \dots & \mu a_{nm} \end{pmatrix} \\ &= \begin{pmatrix} \lambda(\mu a_{11}) & \lambda(\mu a_{12}) & \dots & \lambda(\mu a_{1m}) \\ \lambda(\mu a_{21}) & \lambda(\mu a_{22}) & \dots & \lambda(\mu a_{2m}) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda(\mu a_{n1}) & \lambda(\mu a_{n2}) & \dots & \lambda(\mu a_{nm}) \end{pmatrix} \\ &= \begin{pmatrix} (\lambda\mu)a_{11} & (\lambda\mu)a_{12} & \dots & (\lambda\mu)a_{1m} \\ (\lambda\mu)a_{21} & (\lambda\mu)a_{22} & \dots & (\lambda\mu)a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ (\lambda\mu)a_{n1} & (\lambda\mu)a_{n2} & \dots & (\lambda\mu)a_{nm} \end{pmatrix} = (\lambda\mu)A . \end{aligned}$$

Nullelement der Addition ist natürlich die Nullmatrix

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

deren sämtliche Einträge Null sind, und

$$-\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = \begin{pmatrix} -a_{11} & -a_{12} & \dots & -a_{1m} \\ -a_{21} & -a_{22} & \dots & -a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & -a_{nm} \end{pmatrix}.$$

Schließlich müssen wir uns noch überlegen, daß $k^{n\times m}$ die Dimension nm hat; wir wir am Ende von §1*h)* gesehen haben, ist das gleichbedeutend damit, daß es eine Basis aus nm Matrizen gibt.

Als eine solche Basis wählen wir die Menge aller Matrizen E_{ij}, die so definiert sind, daß E_{ij} an der Stelle ij eine Eins stehen hat und sonst lauter Nullen.

In $k^{4\times 5}$ wäre also etwa $E_{23} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

In einem beliebigen $k^{n\times m}$ läßt sich jede Matrix eindeutig als Linearkombination der E_{ij} schreiben, denn

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}E_{ij},$$

und ist

$$\sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_{ij}E_{ij} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \dots & \lambda_{nm} \end{pmatrix}$$

die Nullmatrix, so müssen offensichtlich alle λ_{ij} verschwinden, die E_{ij} sind also auch linear unabhängig. Damit bilden sie eine Basis von $k^{n\times m}$, und $\dim_k k^{n\times m} = nm$. ∎

Die Basis mit den Matrizen E_{ij} ist zwar sicherlich die einfachste Basis für den Vektorraum aller $n \times m$-Matrizen, aber nicht immer die beste: Matrizen bieten sich beispielsweise auch an, um digitalisierte Bilder darzustellen, und zumindest in digitalen Kameras oder Scannern entsteht das Bild wirklich als eine Matrix von Grauwerten *bzw.* als drei Matrizen von Farb- oder sonstigen Werten, dargestellt in der Basis aus den E_{ij}. Für die Übertragung oder Speicherung ist das aber selten optimal, da hier für jede Komponente der Basisdarstellung dieselbe Genauigkeit erforderlich ist. Daher werden die Bilder etwa für die Speicherung im JPEG-Format zunächst in eine andere Basis umgerechnet, bezüglich derer viele Koeffizienten nahe bei Null liegen. Bei der Diskretisierung und Quantisierung enstehen dann viele Koeffizienten, für die nur wenige oder gar keine Bit benötigt werden, was im Zusammenspiel mit anderen Verfahren wie *run length encoding* und HUFFMAN-Codierung zu Komprimierungsfaktoren um die zwanzig oder dreißig ohne nennenswerten Qualitätsverlust führt.

Auch für zwei lineare Abbildungen $\varphi, \psi\colon V \to W$ können wir eine Summe definieren, und für $\lambda \in k$ auch ein Produkt $\lambda\varphi$ durch

$$\varphi + \psi\colon \begin{cases} V \to W \\ \vec{v} \mapsto \varphi(\vec{v}) + \psi(\vec{v}) \end{cases} \quad \text{und} \quad \lambda\varphi\colon \begin{cases} V \to W \\ \vec{v} \mapsto \lambda\varphi(\vec{v}) \end{cases};$$

sind V und W endlichdimensional und sind A, B die Abbildungsmatrizen von φ, ψ, so hat $\varphi + \psi$ offenbar die Abbildungsmatrix $A + B$ und λA ist die von $\lambda\varphi$.

Auch hier ist klar, daß die sämtlichen linearen Abbildungen $V \to W$ einen Vektorraum bilden, da einfach für jeden Vektor $\vec{v} \in V$ die Vektorraumoperationen von W auf die Bildvektoren angewendet werden; dieser Vektorraum wir mit $\mathrm{Hom}_k(V, W)$ bezeichnet nach dem Wort *Homomorphismus,* das man gelegentlich anstelle von *lineare Abbildung* gebraucht.

Im endlichdimensionalen Fall hat $\mathrm{Hom}_k(V,W)$ wegen der eineindeutigen Entsprechung von linearen Abbildungen und Matrizen als Dimension das Produkt der Dimensionen von V und von W. Die Basismatrix $E_{ij} \in k^{n\times m}$ entspricht dabei bezüglich der Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W jener linearen Abbildung, die alle Vektoren aus $\mathcal{B}$ mit Ausnahme des j-ten auf den Nullvektor abbildet; der j-te Basisvektor geht auf den i-ten Basisvektor aus $\mathcal{C}$. Bei den linearen Abbildungen $k^5 \to k^4$ etwa entspräche obige Matrix E_{23} der linearen Abbildung

$$k^5 \to k^4; \quad \begin{pmatrix} u \\ v \\ w \\ x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ w \\ 0 \\ 0 \end{pmatrix} .$$

Lineare Abbildungen lassen sich nicht nur addieren und mit Skalaren multiplizieren; sie lassen sich auch, wie alle Abbildungen, hintereinanderausführen: Sind $\psi\colon U \to V$ und $\varphi\colon V \to W$ lineare Abbildungen, so ist auch

$$\varphi \circ \psi\colon U \to W; \quad \vec{v} \mapsto \varphi\big(\psi(\vec{v})\big)$$

eine lineare Abbildung.

Falls alle beteiligten Vektorräume endlichdimensional sind, können wir endliche Basen wählen; das seien etwa

$$\mathcal{B} = \big(\vec{b}_1, \dots, \vec{b}_m\big) \text{ für } U, \quad \mathcal{C} = \big(\vec{c}_1, \dots, \vec{c}_n\big) \text{ für } V$$
$$\text{und} \quad \mathcal{D} = \big(\vec{d}_1, \dots, \vec{d}_p\big) \text{ für } W ,$$
$$\text{d.h.} \quad \dim_k U = m, \quad \dim_k V = n \quad \text{und} \quad \dim_k W = p .$$

Dann haben wir Abbildungsmatrizen $A \in k^{p\times n}$ von φ und $B \in k^{n\times m}$ von ψ; wir wollen die Abbildungsmatrix $C \in k^{p\times m}$ von $\varphi\circ\psi\colon U \to W$ berechnen.

Nach Definition der Abbildungsmatrizen $A = (a_{ij})$ von φ und $B = (b_{j\ell})$

von ψ ist

$$\varphi \circ \psi(\vec{b}_i) = \varphi\big(\psi(\vec{b}_i)\big) = \varphi\left(\sum_{j=1}^{n} b_{ji}\vec{c}_j\right) = \sum_{j=1}^{n} b_{ji}\varphi(\vec{c}_j)$$
$$= \sum_{j=1}^{n} b_{ji} \sum_{\ell=1}^{p} a_{\ell j}\vec{d}_\ell = \sum_{\ell=1}^{p}\left(\sum_{j=1}^{n} a_{\ell j} b_{ji}\right)\vec{d}_\ell .$$

Für die Abbildungsmatrix $C = (c_{i\ell})$ von $\varphi\circ\psi$ ist nach Definition

$$\varphi \circ \psi(\vec{b}_i) = \sum_{\ell=1}^{p} c_{\ell i}\vec{d}_\ell , \text{ also ist } \quad c_{\ell i} = \sum_{j=1}^{n} a_{\ell j} b_{ji} .$$

Definition: Für zwei Matrizen $A = (a_{i\ell}) \in k^{p\times n}$ und $B = (b_{\ell j}) \in k^{n\times m}$ bezeichnen wir die Matrix $C = (c_{ij}) \in k^{p\times m}$ mit

$$c_{ij} = \sum_{\ell=1}^{n} a_{i\ell} b_{\ell j}$$

als *Produktmatrix* $C = AB$ von A und B.

Für praktische Rechnungen empfiehlt es sich als Eselsbrücke, den zweiten Faktor des Produkts höher zu stellen nach dem Schema

$$\begin{array}{cc} & \begin{pmatrix} \cdots & b_{1j} & \cdots \\ & \vdots & \\ \cdots & b_{nj} & \cdots \end{pmatrix} \\ \begin{pmatrix} \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{pmatrix} & \begin{pmatrix} \cdots & \sum_{\ell=1}^{n} a_{i\ell} b_{\ell j} & \cdots \end{pmatrix} \end{array} ;$$

dadurch behält man den Überblick, welcher Rechenschritt jeweils als nächster auszuführen ist.

Im Gegensatz zu den meisten bislang aufgetretenen Produkten ist dieses Matrixprodukt im allgemeinen *nicht* kommutativ: Falls nicht zufälligerweise $n = p$ sein sollte, ist das Matrixprodukt BA nicht einmal

definiert, geschweige denn gleich AB. Allgemein ist Kommutativität bei der Hintereinanderausführung von Abbildungen eine sehr seltene Ausnahmeerscheinung; schließlich ist auch $\sin(x^2) \neq \sin^2 x$ für fast jedes x, und so ist auch bei Matrizen, selbst wenn beide Produkte definiert sind, im allgemeinen $AB \neq BA$. Beispielsweise ist

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

aber

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

Ansonsten gelten aber doch die meisten bekannten Rechenregeln, beispielsweise das *Assoziativitätsgesetz*

$$A(BC) = (AB)C \quad \text{für alle } A \in k^{n\times m},\ B \in k^{m\times p},\ C \in k^{p\times q} .$$

Es ist durchaus möglich (und verglichen mit manch anderen Dingen sogar nicht einmal so extrem aufwendig), dieses Gesetz nach obiger Formel explizit nachzurechnen. Bevor wir uns das antun, sollten wir uns aber daran erinnern, wo das Matrixprodukt eigentlich herkommt: Matrizen entsprechen umkehrbar eindeutig linearen Abbildungen, und das Matrixprodukt entspricht deren Hintereinanderausführung. Für die Hintereinanderausführung von Abbildungen (egal ob linear oder nicht) ist das Assoziativgesetz aber trivial: Sind etwa

$$\varphi\colon k^m \to k^n, \quad \psi\colon k^n \to k^p \quad \text{und} \quad \omega\colon k^q \to k^p$$

drei lineare Abbildungen mit Abbildungsmatrizen A, B, C, so ist für jeden Vektor $\vec{v} \in k^q$ sowohl

$$\big(\varphi \circ (\psi \circ \omega)\big)(\vec{v}) = \varphi\big(\psi \circ \omega)(\vec{v})\big) = \varphi\Big(\psi\big(\omega(\vec{v})\big)\Big)$$

als auch

$$\big((\varphi \circ \psi) \circ \omega\big)\big(\vec{v}\big) = (\varphi \circ \psi)\big(\omega(\vec{v})\big) = \varphi\Big(\psi\big(\omega(\vec{v})\big)\Big) ,$$

d.h. für die Hintereinanderausführung von Abbildungen (egal ob linear oder nicht) ist das Assoziativgesetz

$$\varphi \circ (\psi \circ \omega) = (\varphi \circ \psi) \circ \omega$$

automatisch erfüllt.

Da nun $A(BC)$ die Abbildungsmatrix von $\varphi \circ (\psi \circ \omega)$ ist und $(AB)C$ die von $(\varphi \circ \psi) \circ \omega$, und da diese beiden Abbildungen übereinstimmen, müssen auch die Abbildungsmatrizen gleich sein, wir haben also gezeigt, daß

$$A(BC) = (AB)C \quad \text{für alle } A \in k^{n\times m},\ B \in k^{m\times p},\ C \in k^{p\times q} ,$$

ohne daß wir ein einziges Matrixprodukt explizit ausrechnen mußten.

Genauso folgen auch die Rechenregeln

$$A(B_1 + B_2) = AB_1 + AB_2 \quad \text{und} \quad (A_1 + A_2)B = A_1B + A_2B$$

aus den entsprechenden Rechenregeln

$$\varphi \circ (\psi_1 + \psi_2) = \varphi \circ \psi_1 + \varphi \circ \psi_2 \quad \text{und} \quad (\varphi_1 + \varphi_2) \circ \psi = \varphi_1 \circ \psi + \varphi_2 \circ \psi ,$$

die sich wiederum leicht und ohne Rechnung überprüfen lassen:

$$\begin{aligned} \big(\varphi \circ (\psi_1 + \psi_2)\big)(\vec{v}) &= \varphi\big(\psi_1(\vec{v}) + \psi_2(\vec{v})\big) = \varphi\big(\psi_1(\vec{v})\big) + \varphi\big(\psi_2(\vec{v})\big) \\ &= (\varphi \circ \psi_1)(\vec{v}) + (\varphi \circ \psi_2)(\vec{v}) = (\varphi \circ \psi_1 + \varphi \circ \psi_2)(\vec{v}) \end{aligned}$$

und

$$\begin{aligned} \big((\varphi_1 + \varphi_2) \circ \psi\big)(\vec{v}) &= (\varphi_1 + \varphi_2)\big(\psi(\vec{v})\big) = \varphi_1\big(\psi(\vec{v})\big) + \varphi_2\big(\psi(\vec{v})\big) \\ &= (\varphi_1 \circ \psi)(\vec{v}) + (\varphi_2 \circ \psi)(\vec{v}) = (\varphi_1 \circ \psi + \varphi_2 \circ \psi)(\vec{v}) . \end{aligned}$$

Da in der obigen Formel für die Matrixmultiplikation alle $b_{\ell j}$ linear in den Ausdrücken für c_{ij} vorkommen *usw.,* hätten sich die beiden letzen Rechenregeln auch einfach direkt nachrechnen lassen. Ebenfalls durch direktes Nachrechnen überzeugt man sich von der Formel

$$(\lambda A)B = \lambda(AB) = A(\lambda B) \quad \text{für alle } \lambda \in k .$$

folgt wohl am einfachsten durch direktes Nachrechnen und für die *Einheitsmatrix*

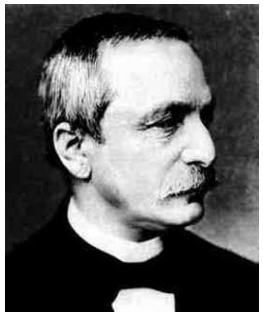
$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in k^{n\times n}$$

folgt ebenfalls sofort durch Nachrechnen wie auch durch Interpretation von E als Abbildungsmatrix der identischen Abbildung $k^n \to k^n$, die jeden Vektor auf sich selbst abbildet, daß

$$A \cdot E = A \quad \text{und} \quad E \cdot A = A \qquad \text{für alle } A \in k^{m \times n}$$

ist. Den Eintrag an der Stelle ij der Einheitsmatrix bezeichnet man als das KRONECKER-δ:

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} .$$



LEOPOLD KRONECKER (1823–1891) ist heute zwar Vielen nur im Zusammenhang mit dem KRONECKER-δ bekannt, er war aber einer der bedeutendsten deutschen Mathematiker seiner Zeit. Seine Arbeiten befaßten sich mit Algebra, Zahlentheorie und Analysis, wobei er insbesondere die Verbindungen zwischen der Analysis und den beiden anderen Gebieten erforschte. Bekannt ist auch seine Ablehnung jeglicher mathematischer Methoden, die, wie die Mengenlehre oder Teile der Analysis, unendliche Konstruktionen verwenden. Er war deshalb mit vielen anderen bedeutenden Mathematikern seiner Zeit verfeindet, z.B. mit CANTOR und mit WEIERSTRASS

Bei den reellen Zahlen und auch sonst in jedem Körper gibt es zu jedem Element $a \neq 0$ ein inverses Element a^{-1}, so daß $aa^{-1} = a^{-1}a = 1$ ist. Bei Matrizen muß es das selbst für quadratische Matrizen nicht geben: Für eine beliebige Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$

und

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} ,$$

was beides offensichtlich nie die Einheitsmatrix sein kann.

Definition: Eine $n \times n$-Matrix $A \in k^{n \times n}$ heißt *invertierbar,* wenn es eine Matrix $B \in k^{n \times n}$ gibt, so daß $AB = BA = E$ ist. B heißt inverse Matrix von A; in Zeichen $B = A^{-1}$.

(Es wäre theoretisch möglich, Invertierbarkeit auch für nicht-quadratische Matrizen zu definieren, aber das hat keinen sonderlichen Nutzen.)

Um zu sehen, wann eine Matrix $A \in k^{n \times n}$ invertierbar ist, betrachten wir wieder die Situation bei den linearen Abbildungen: Zu einer linearen Abbildung $\varphi: k^n \to k^n$ gibt es genau dann eine Umkehrabbildung $\psi: k^n \to k^n$, so daß $\varphi \circ \psi$ und $\psi \circ \varphi$ beide die identische Abbildung sind, wenn φ bijektiv ist. Nach dem Korollar am Ende von §1*i)* ist dies genau dann der Fall, wenn φ surjektiv ist, wenn also das Bild von φ Dimension n hat. Dieses Bild wird aber erzeugt von den Bildern der Einheitsvektoren, und das sind gerade die Spalten der Abbildungsmatrix. Diese n Vektoren erzeugen genau dann ganz k^n, wenn sie linear unabhängig sind.

Definition: Der (Spalten-)Rang einer Matrix $A \in k^{n \times m}$ ist die maximale Anzahl linear unabhängiger Spaltenvektoren von A.

Nach obiger Diskussion gilt daher

Lemma: Eine Matrix $A \in k^{n \times n}$ ist genau dann invertierbar, wenn sie Rang n hat. Die inverse Matrix $B = A^{-1}$ ist sowohl durch die Bedingung $AB = E$ als auch durch die Bedingung $BA = E$ eindeutig bestimmt.

Die Eindeutigkeit der inversen Matrix folgt dabei natürlich aus der Eindeutigkeit der Umkehrabbildung. Ebenfalls klar ist das folgende

Lemma: Sind $A, B \in k^{n \times n}$ invertierbar, so auch ihr Produkt, und $(AB)^{-1} = B^{-1}A^{-1}$.

Beweis: $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$; also ist AB invertierbar und $B^{-1}A^{-1}$ ist die inverse Matrix. ∎

Man beachte, daß im allgemeinen $A^{-1}B^{-1}$ *nicht* invers zu AB ist: Für

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \quad \text{und} \quad AB = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

sind

$$A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad B^{-1} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$$

die inversen Matrizen, und Multiplikation zeigt, daß

$$(AB)^{-1} = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} \neq \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} = A^{-1}B^{-1}$$

ist. Insbesondere unterscheidet sich auch

$$ABA^{-1}B^{-1} = \begin{pmatrix} 21 & -8 \\ 8 & -3 \end{pmatrix}$$

deutlich von der Einheitsmatrix.

c) Matrixdarstellung der komplexen Zahlen

Die komplexen Zahlen bilden mit ihrer üblichen Addition und der Einschränkung der üblichen Multiplikation zu einer Abbildung

$$\cdot : \begin{cases} \mathbb{R} \times \mathbb{C} \to \mathbb{C} \\ (r, z) \mapsto rz \end{cases}$$

einen $\mathbb{R}$-Vektorraum, und die Abbildung

$$\mathbb{C} \to \mathbb{C}; \quad z \mapsto cz$$

ist für jede komplexe Zahl $c = a + ib \in \mathbb{C}$ insbesondere eine lineare Abbildung von $\mathbb{R}$-Vektorräumen. Wählen wir $(1, i)$ als $\mathbb{R}$-Basis von $\mathbb{C}$, so bildet sie die beiden Basisvektoren 1 und i ab auf

$$c \cdot 1 = a + bi \quad \text{und} \quad c \cdot i = ai + bi^2 = -b + ai\,;$$

sie hat also die Abbildungsmatrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Die Hintereinanderausführung zweier solcher Abbildungen entspricht der Multiplikation der entsprechenden komplexen Zahlen; insbesondere gehört also das Produkt $(a + ib)(a' + ib')$ zweier komplexer Zahlen zur Produktmatrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = \begin{pmatrix} aa' - bb' & -(a'b + ab') \\ a'b + ab' & aa' - bb' \end{pmatrix}.$$

Der Körper der komplexen Zahlen kann damit auch identifiziert werden mit der Menge aller reeller 2×2-Matrizen der obigen Form mit der

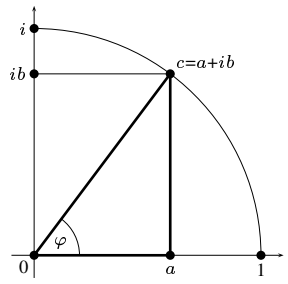
Matrixaddition und dem Matrixprodukt. Man beachte, daß die Multiplikation zweier Matrizen dieser speziellen Form kommutativ ist, denn die Multiplikation komplexer Zahlen ist kommutativ.

Umgekehrt wissen wir, daß die Multiplikation vom Matrizen assoziativ ist; damit folgt ganz ohne Rechnung das in §1*b)* dem Leser zum Nachrechnen überlassene Assoziativgesetz für die Multiplikation komplexer Zahlen. Entsprechend folgen auch die Distributivgesetze aus denen für die Matrizenaddition und -multiplikation.

Betrachten wir speziell den Fall, daß $c = a + ib$ den Betrag eins hat. Dann ist $|cz| = |c| \cdot |z| = |z|$, und allgemeiner ist für zwei beliebige komplexe Zahlen z, w auch

$$|cz - cw| = |c(z - w)| = |c| \cdot |z - w| = |z - w|\,,$$

der EUKLIDische Abstand zwischen den Bildpunkten cz und zw ist also gleich dem Abstand zwischen z und w. Die Abbildung $z \mapsto cz$ ist somit eine Kongruenzabbildung. Da sie für $c \neq 1$ den Nullpunkt als einzign Fixpunkt hat, ist sie entweder eine Drehung oder eine Drehspiegelung. Letzteres kommt nicht in Frage, da man kontinuierlich auf dem Einheitskreis von eins zu c gehen kann, also haben wir eine Drehung um einen Winkel φ.



Zur Bestimmung von φ können wir ausnutzen, daß diese Drehung den Punkt Eins in den Punkt c überführt: Betrachten wir dazu das rechtwinklige Dreieck mit Ecken $0, a$ und $c = a + ib$! Seine Hypothenuse hat die Länge $|c| = 1$, seine Katheten sind a und b. Nach Definition der Winkelfunktionen als Ankathete *bzw.* Gegenkathete durch Hypothenuse ist offensichtlich $a = \cos\varphi$ und $b = \sin\varphi$. Somit ist $c = \cos\varphi + i\sin\varphi$.

Schreiben wir für einen beliebigen Winkel φ

$$c_\varphi = a_\varphi + ib_\varphi \quad \text{mit} \quad a_\varphi = \cos\varphi \quad \text{und} \quad b_\varphi = \sin\varphi\,,$$

so ist offensichtlich $c_\varphi c_\psi = c_{\varphi+\psi}$, denn die Hintereinanderausführung zweier Drehungen um den Nullpunkt ist eine Drehung mit der Summe

der beiden Drehwinkel. Ausmultipliziert ergibt dies

$$c_{\varphi+\psi} = (a_\varphi + ib_\varphi)(a_\psi + ib_\psi) = (a_\varphi a_\psi - b_\varphi b_\psi) + i(a_\varphi b_\psi + a_\psi b_\varphi)\,,$$

d.h.

$$\cos(\varphi+\psi) = \cos\varphi\cos\psi - \sin\varphi\sin\psi \quad \text{und}$$

$$\sin(\varphi+\psi) = \sin\varphi\cos\psi + \cos\varphi\sin\psi\,;$$

wir haben also die Additionstheoreme für Sinus und Kosinus gefunden.

Außerdem zeigt die Beziehung $c_\varphi c_\psi = c_{\varphi+\psi}$, daß sich c_φ als Funktion von φ wie eine Potenz verhält; wir schreiben deshalb, im Augenblick noch formal,

$$c_\varphi = e^{i\varphi} \underset{\text{def}}{=} \cos\varphi + i\sin\varphi\,.$$

Offensichtlich läßt sich jede komplexe Zahl in der Form $z = re^{i\varphi}$ schreiben, wobei $r = |z|$ der Betrag von z ist. Für $z = 0$ können wir für φ jeden beliebigen Wert nehmen; für $z \neq 0$ ist φ modulo 2π eindeutig bestimmt. In diesem Fall bezeichnen wir $\varphi = \arg z$ als das *Argument* von z und (r,φ) als die *Polarkoordinaten* von z.

In Polarkoordinaten wird die Multiplikation komplexer Zahlen deutlich einfacher als in den bislang benutzten kartesischen Koordinaten:

$$(re^{i\varphi})\cdot(se^{i\psi}) = rs\cdot e^{i(\varphi+\psi)}\,.$$

Dies läßt sich auch benutzen, um Wurzeln zu ziehen: Offensichtlich ist $\sqrt[n]{r}\cdot e^{i\varphi/n}$ eine n-te Wurzel aus $re^{i\varphi}$; weitere Wurzeln sind die Zahlen $\sqrt[n]{r}\cdot e^{i(\varphi+2k\pi)/n}$ für $k = 1,\ldots,n-1$. (Für $k = n$ bekommen wir wieder dieselbe Zahl wie für $k = 0$.)

Beispielsweise ist in Polarkoordinaten $i = e^{\pi i/2}$, da i aus der Eins durch Drehung um 90° oder $\frac{\pi}{2}$ entsteht. Also ist $e^{\pi i/4}$ eine Quadratwurzel aus i. Der Winkel $\frac{\pi}{4}$ oder 45° tritt in gleichschenkligen rechtwinkligen Dreiecken auf, die man auch auffassen kann als entlang der Diagonale halbierte Quadrate; da die Diagonale eines Quadrats das $\sqrt{2}$-fache der Seite ist, sind also Sinus und Cosinus gleich $\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$, und somit ist

$$\cos\frac{\pi}{4} + i\sin\frac{\pi}{4} = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$$

eine Quadratwurzel aus i. Die andere ist natürlich einfach das Negative davon.

Nicht nur komplexe Zahlen lassen sich mit Matrizen identifizieren; Ganz entsprechend kann man auch die Elemente der Körper $\mathbb{F}_{2^n}$ mit $n\times n$-Matrizen über $\mathbb{F}_2$ identifizieren. Nimmt man etwa $1,\alpha$ als Basisvektoren des $\mathbb{F}_2$-Vektorraums $\mathbb{F}_4$, so entsprechen die vier Elemente $0,1,\alpha$ und $\alpha+1$ des Körpers $\mathbb{F}_4$ den Matrizen

$$\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad \begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \begin{pmatrix}0&1\\1&1\end{pmatrix}\quad\text{und}\quad \begin{pmatrix}1&1\\1&0\end{pmatrix},$$

und für $\mathbb{F}_{32}$ mit $\mathbb{F}_2$-Basis $1,\alpha,\alpha^2,\alpha^3,\alpha^4$ und Relation $\alpha^5 = \alpha^2+1$ entspricht α der Matrix

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

d) Das Gaußsche Eliminationsverfahren

Die meisten kennen wohl aus der Schule zumindest Verfahren zur Lösung linearer Gleichungssysteme in bis zu drei Unbekannten, teilweise vielleicht auch für Systeme aus beliebig vielen Gleichungen in beliebig vielen Unbekannten.

Der GAUSS-Algorithmus, mit dem wir uns hier beschäftigen wollen, bestimmt die Lösungsmenge eines beliebigen linearen Gleichungssystems, und falls das Gleichungssystem nicht gerade eine sehr spezielle Gestalt hat, liefert er sie im allgemeinen auf die effizienteste Art und Weise.

Seine Grundidee ist sehr einfach: Im Falle einer einzigen Gleichung mit einer einzigen Unbekannten x können wir das „Gleichungssystem“

$$ax = b$$

sofort lösen: Für $a \neq 0$ ist $x = b/a$, d.h. $\mathcal{L} = \{b/a\}$; ansonsten gibt es für $b \neq 0$ keine Lösung, d.h. $\mathcal{L} = \emptyset$, und für $a = b = 0$ ist jedes x aus k eine Lösung, d.h. $\mathcal{L} = k$.

Das GAUSSsche Eliminationsverfahren führt ein allgemeines lineares Gleichungssystem sukzessive zurück auf solche lineare Gleichungen in einer Unbekannten, ausgehend von zwei trivialen Beobachtungen:

1. Die Lösungsmenge eines linearen Gleichungssystems ändert sich nicht, falls wir zwei Gleichungen miteinander vertauschen.
2. Die Lösungsmenge ändert sich auch dann nicht, wenn wir ein Vielfaches einer Gleichung zu einer anderen addieren, d.h. wenn wir die Gleichung

$$\ell_j(x_1, \ldots, x_m) \underset{\text{def}}{=} a_{j1}x_1 + \ldots + a_{jm}x_m = b_j$$

ersetzen durch

$$\ell_j(x_1, \ldots, x_m) + \lambda\ell_i(x_1, \ldots, x_m) = b_j + \lambda b_i\,, \qquad (*)$$

denn unter der Nebenbedingung

$$\ell_i(x_1, \ldots, x_m) = b_i$$

ist $(*)$ äquivalent zu $\ell_j(x_1, \ldots, x_m) = 0$.

Mit Hilfe dieser beiden Beobachtungen läßt sich nun die Variablenanzahl wie folgt sukzessive reduzieren: Beginnen wir mit der Elimination von x_1. Falls x_1 im Gleichungssystem überhaupt nicht vorkommt, falls also alle $a_{i1} = 0$ sind, gibt es nichts zu tun: Wir haben ein Gleichungssystem in $x_2, \ldots, x_m$, und für jede Lösung $(x_2, \ldots, x_m)$ dieses Systems sowie jedes beliebige $x_1 \in k$ ist $(x_1, x_2, \ldots, x_m)$ eine Lösung des ursprünglichen Systems.

Ansonsten können wir, indem wir nötigenfalls zwei Gleichungen miteinander vertauschen, annehmen, daß $a_{11} \neq 0$ ist. Dann lassen wir die erste Gleichung so stehen, wie sie ist, und ersetzen jede weitere Gleichung $\ell_j(x_1, \ldots x_m) = b_j$ durch

$$\ell_j(x_1, \ldots x_m) - \frac{a_{j1}}{a_{11}}\ell_1(x_1, \ldots x_m) = b_j - \frac{a_{j1}}{a_{11}}b_1\,;$$

in diesen Gleichungen kommt x_1 offenbar nicht mehr vor. Wir haben somit ein Gleichungssystem in einer Variablen weniger, plus der Gleichung

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1\,.$$

Sobald wir das Gleichungssytem für $x_2, \ldots, x_m$ gelöst haben, wird diese Gleichung nach Einsetzen einer Lösung $(x_2, \ldots, x_m)$ zu einer linearen Gleichung für x_1, die wir lösen können:

$$x_1 = \frac{b_1 - (a_{12}x_2 + \cdots + a_{1m}x_m)}{a_{11}}\,.$$

Das Gleichungssystem für $x_2, \ldots, x_m$ wird nun, falls $m > 2$ ist, nach genau derselben Methode weiterreduziert: Falls x_2 in keiner der Gleichungen vorkommt, haben wir tatsächlich ein Gleichungssystem in $x_3, \ldots, x_m$, andernfalls können wir durch Vertauschen zweier Gleichungen annehmen, daß x_2 in der ersten Gleichung mit einem von Null verschiedenen Vorfaktor auftritt, und wir können wie oben x_2 aus allen weiteren Gleichungen eliminieren *usw.*

Da in jedem Eliminationsschritt ein Nenner auftritt, können die Nenner vor allem bei großem m gelegentlich schnell unübersichtlich groß werden. Obwohl es *im Prinzip* nicht notwendig ist, kann man um dies zu vermeiden noch als dritte Operation die Multiplikation einer Gleichung mit einem Körperelement (z.B. einem Hauptnenner der Koeffizienten) zulassen. Dies ist gleichbedeutend damit, daß man anstelle der Operation $(*)$ allgemeiner die Ersetzung von $\ell_j(x_1, \ldots x_m)$ durch

$$\mu\ell_j(x_1, \ldots x_m) + \lambda\ell_i(x_1, \ldots x_m)$$

mit beliebigem $\mu \neq 0$ aus k zuläßt. ($\mu = 0$ muß hier natürlich unbedingt ausgeschlossen werden, denn sonst läßt sich die Gleichung $\ell_j(x_1, \ldots, x_m) = 0$ aus dem neuen Gleichungssystem nicht mehr herleiten, d.h. wir erhalten auch „Lösungen“, die diese Gleichung nicht erfüllen.) Ein spezieller Fall der Multiplikation einer Gleichung mit einem Körperelement ist das Kürzen durch einen gemeinsamen Teiler der Koeffizienten, wodurch ein Gleichungssystem (egal ob das ursprüngliche oder ein Zwischenresultat) und vor allem der weitere Rechengang oft erheblich übersichtlicher wird.

e) Erste Beispiele

Betrachten wir dazu einige Beispiele; der Grundkörper k sei dabei jeweils der Körper der reellen Zahlen oder einer seiner Teilkörper, etwa $k = \mathbb{Q}$.

Sei zunächst

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ 5x_1 & - & 3x_2 & + & 7x_2 & = & 19 \end{array}$$

das zu lösende Gleichungssystem. Da x_1 in der ersten Gleichung tatsächlich vorkommt, müssen wir nichts vertauschen; allerdings müssen wir ein Drittel der ersten Gleichung von der zweiten und fünf Drittel der ersten Gleichung von der dritten subtrahieren, um x_2 aus diesen beiden Gleichungen zu eliminieren, was auf das etwas unangenehme Gleichungssystem

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ & & \frac{4}{3}x_2 & + & \frac{11}{3}x_3 & = & \frac{4}{3} \\ & & -\frac{19}{3}x_2 & + & \frac{16}{3}x_3 & = & \frac{32}{3} \end{array}$$

führt. Solche Gleichungssysteme sind zwar nicht immer vermeidbar, aber hier hätten wir es auch einfacher haben können: Wenn wir im ursprünglichen Gleichungssystem die ersten beiden Gleichungen vertauschen, wird es zu

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ 5x_1 & - & 3x_2 & + & 7x_2 & = & 19\,, \end{array}$$

und hier müssen wir stattdessen das Dreifache *bzw.* Fünffache der ersten Gleichung von der zweiten *bzw.* dritten subtrahieren, was auf das deutlich angenehmere Gleichungssystem

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ & & -4x_2 & - & 11x_3 & = & -4 \\ & & -13x_2 & - & 13x_3 & = & 4 \end{array}$$

führt. Etwas ähnliches hätten wir auch bekommen, wenn wir die letzten beiden Gleichungen des anderen Systems einfach mit drei multipliziert hätten, aber grundsätzlich ist es meistens günstiger, die Gleichung mit dem einfachsten führenden Koeffizienten an erster Stelle zu haben. Auch hier wird das Gleichungssystem zumindest optisch etwas angenehmer, wenn wir die zweite und die dritte Gleichung mit (-1) multiplizieren:

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ & & 4x_2 & & 11x_3 & = & 4 \\ & & 13x_2 & + & 13x_3 & = & -4 \end{array}$$

Uns interessieren zunächst nur die letzten beiden Zeilen. Diese bilden ein lineares Gleichungssystem in x_2 und x_3, aus dem wir x_2 in einer der beiden Gleichungen eliminieren möchten.

Da x_2 in der zweiten Gleichung wirklich vorkommt, subtrahieren wir $13/4$ mal diese Gleichung von der dritten und erhalten als neues System

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ & & 4x_2 & + & 11x_3 & = & 4 \\ & & & & -\frac{91}{4}x_3 & = & -17\,. \end{array}$$

Hier ist die letzte Gleichung eine gewöhnliche lineare Gleichung für x_3, aus der wir sofort ablesen, daß

$$x_3 = \frac{68}{91}$$

ist. Dies setzen wir in die vorletzte Gleichung ein:

$$4x_2 + \frac{748}{91} = 4$$

ist eine lineare Gleichung für x_2 mit Lösung

$$x_2 = -\frac{96}{91}\,.$$

Dies sowie x_3 setzen wir schließlich in die erste Gleichung ein:

$$x_1 + \frac{80}{91} = 3$$

hat die Lösung

$$x_1 = \frac{193}{91}\,,$$

so daß ingesamt

$$\left(\frac{193}{91}, -\frac{96}{91}, \frac{68}{91}\right)$$

die (einzige) Lösung des Gleichungssystems ist.

Als nächstes wollen wir ein Beispiel betrachten, bei dem nicht mit einer eindeutigen Lösung zu rechnen ist: Wir nehmen ein System aus drei Gleichungen in sieben Unbekannten. Zumindest intuitiv schein klar, daß

man mit nur drei Bedingungen sieben Variable nicht eindeutig festlegen kann; wir wollen sehen, was der GAUSS-Algorithmus in so einer Situation liefert.

Das Gleichungssystem ist

$$\begin{array}{rcrcrcrcrcrcr} & & 5x_3 & & & - & 2x_6 & & & = & 3 \\ & 4x_2 & & + & 2x_4 & - & 3x_6 & - & 7x_7 & = & -2 \\ x_1 & & & - & 3x_4 & + & x_5 & & & = & 0 \end{array}$$

Hier kommt x_1 in der ersten Gleichung nicht vor, wohl aber in der dritten. Wir vertauschen daher die erste Gleichung mit der dritten und erhalten

$$\begin{array}{rcrcrcrcrcrcr} x_1 & & & - & 3x_4 & + & x_5 & & & = & 0 \\ & 4x_2 & & + & 2x_4 & - & 3x_6 & - & 7x_7 & = & -2 \\ & & 5x_3 & & & - & 2x_6 & & & = & 3\,. \end{array}$$

In diesem Gleichungssystem kommt x_1 nur in der ersten Gleichung vor, x_2 nicht mehr hinter der zweiten und x_3 (natürlich) nicht mehr hinter der dritten, wir haben also bereits die Treppengestalt erreicht.

Die dritte Gleichung $5x_3 - 2x_6 = 3$ hat offensichtlich keine eindeutig bestimmte Lösung: Wir können entweder nach x_3 oder nach x_6 auflösen, so daß eine der beiden Unbekannten beliebig bewählt werden kann. Wählen wir etwa für x_6 irgendein Element $\alpha \in k$, so wird

$$x_6 = \alpha, \quad x_3 = \frac{2}{5}\alpha + \frac{3}{5}\,.$$

Die zweite Gleichung wird nach Einsetzen von $x_6 = \alpha$ zu

$$4x_2 + 2x_4 - 3\alpha - 7x_7 = -2\,,$$

also zu einer Beziehung zwischen x_2, x_4 und x_7, denn α ist ein bereits fest gewähltes und somit bekanntes Element des Grundkörpers k. Wieder können wir eine der beiden Variablen willkürlich festlegen, etwa x_7 auf ein $\beta \in k$ setzen und x_4 auf ein $\gamma \in k$; wir erhalten

$$x_7 = \beta, \quad x_4 = \gamma, \quad x_2 = \frac{3}{4}\alpha + \frac{7}{4}\beta - \frac{\gamma}{2} - \frac{1}{2}\,.$$

In der ersten Gleichung schließlich müssen wir alle bereits berechneten oder festgelegten Unbekannten einsetzen und erhalten dann die Gleichung

$$x_1 - 3\gamma + x_5 = 0\,.$$

Hier können wir noch $x_5 = \delta$ beliebig festlegen und erhalten dann für die letzte noch verbleibende Variable

$$x_1 = 3\gamma - \delta\,.$$

Insgesamt hängt die Lösung dieses Gleichungssystems also ab von vier willkürlich wählbaren Körperelementen $\alpha, \beta, \gamma, \delta \in k$ oder, wie wir auch sagen werden, von vier *Parametern*. Die *Lösungsmenge* $\mathcal{L}$ ist daher die Menge

$$\left\{ \left(3\gamma - \delta, \frac{3}{4}\alpha + \frac{7}{4}\beta - \frac{\gamma}{2} - \frac{1}{2}, \frac{2}{5}\alpha + \frac{3}{5}, \gamma, \delta, \alpha, \beta\right) \;\middle|\; \alpha, \beta, \gamma, \delta \in k \right\}\,.$$

Im ersten Beispiel brauchten wir von den beiden Operationen des GAUSS-Algorithmus nur den Eliminationsschritt, im zweiten nur den Vertauschungsschritt. Als nächstes betrachten wir ein Beispiel, in dem beide notwendig sind, das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcrcrcr} & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\ 2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\ 6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15 \\ 4x_1 & + & x_2 & & & - & 3x_4 & & & + & x_6 & = & 0 \\ 6x_1 & - & 6x_2 & + & x_3 & + & 10x_4 & + & 12x_5 & - & 6x_6 & = & -8 \end{array}$$

Hier kommt x_1 ausgerechnet in der ersten Gleichung nicht vor, dafür aber in allen folgenden; wir müssen also die erste Gleichung mit einer der anderen vertauschen, etwa der zweiten:

$$\begin{array}{rcrcrcrcrcrcr} 2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\ & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\ 6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15 \\ 4x_1 & - & 7x_2 & & & - & 3x_4 & & & + & x_6 & = & 0 \\ 2x_1 & - & 7x_2 & + & x_3 & + & 13x_4 & + & 12x_5 & - & 6x_6 & = & -8 \end{array}$$

Nun kommt x_1 in der zweiten Gleichung nicht mehr vor; aus der dritten, vierten *bzw.* fünften Gleichung kann x_1 eliminiert werden, indem man dreimal, zweimal *bzw.* einmal die erste Gleichung subtrahiert:

$$\begin{array}{rcrcrcrcrcrcr} 2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\ & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\ & & -9x_2 & + & x_3 & + & 17x_4 & + & 18x_5 & - & 9x_6 & = & -21 \\ & & -3x_2 & & & + & 5x_4 & + & 12x_5 & - & 3x_6 & = & -24 \\ & & -12x_2 & + & x_3 & + & 22x_4 & + & 30x_5 & - & 12x_6 & = & -44 \end{array}$$

Als nächstes muß x_2 aus der dritten bis fünften Gleichung eliminiert werden; da x_2 in der zweiten Gleichung mit Koeffizient eins vorkommt, müssen wir einfach jenes Vielfache der zweiten Gleichung subtrahieren, das dem jeweiligen x_2-Koeffizienten entspricht. Da hier alle diese Koeffizienten negativ sind, bedeutet dies, daß wir dasjenige Vielfache der zweiten Gleichung *addieren,* das dem Betrag des Koeffizienten entspricht:

$$\begin{array}{rcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & x_3 & + & 82x_4 & + & 6x_5 & & & = & 76\,.
\end{array}$$

Hier muß x_3 aus der letzten Gleichung eliminiert werden; dazu muß offenbar einfach die dritte Gleichung subtrahiert werden:

$$\begin{array}{rcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 7\,.
\end{array}$$

Eigentlich sollte man hier schon sehen, was los ist, aber wir rechnen zur Veranschaulichung des GAUSS-Algorithmus trotzdem stur weiter nach Schema F: Danach muß x_4 aus der letzten Gleichung eliminiert werden durch Subtraktion der vorletzten:

$$\begin{array}{rcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & & & & & & & 0 & = & 1\,.
\end{array}$$

Damit wird endgültig klar, daß jede Lösung $(x_1, \ldots, x_6)$ des gegebenen Gleichungssystems insbesondere auch die Gleichung $0 = 1$ erfüllen muß, d.h. die Lösungsmenge ist leer.

Nachdem wir soviel Arbeit in dieses Beispiel investiert haben, sollten wir zumindest einen Teil der Rechnungen recyclen zu einem Beispiel, in dem es Lösungen gibt. Dazu muß nur die letzte der fünf ursprünglichen Gleichungen auf der rechten Seite leicht abgeändert werden: Wir

betrachten nun das System

$$\begin{array}{rcrcrcrcrcrcr}
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15 \\
4x_1 & + & x_2 & & & - & 3x_4 & & & + & x_6 & = & 0 \\
6x_1 & - & 6x_2 & + & x_3 & + & 10x_4 & + & 12x_5 & - & x_6 & = & \mathbf{-9}\,.
\end{array}$$

Hierauf lassen sich genau dieselben Umformungen anwenden wie oben, anstelle des Systems mit fünfter Gleichung $0 = 1$ führen diese nun aber auf

$$\begin{array}{rcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & & & & & & & 0 & = & 0\,.
\end{array}$$

Diese letzte Gleichung ist natürlich für jedes Tupel $(x_1, \ldots, x_6)$ erfüllt. Die vorletzte gibt eine Beziehung zwischen x_4 und x_5, wir können also

$$x_5 = \alpha \in k$$

beliebig wählen und erhalten dann

$$x_4 = -\frac{3}{10}\alpha + \frac{3}{10}\,.$$

Wenn wir dies in die dritte Gleichung einsetzen, bleibt dort nur noch x_3 als Variable stehen, und aus

$$x_3 - \frac{93}{5}\alpha + \frac{93}{5} = 69$$

lesen wir sofort ab, daß

$$x_3 = \frac{93}{5}\alpha + \frac{252}{5}$$

ist. Damit gehen wir in die zweite Gleichung:

$$x_2 - \frac{7}{5}\alpha + x_6 + \frac{3}{2} = 10\,.$$

Hier können wir wieder eine der beiden noch verbliebenen Variablen auf einen beliebigen Wert setzen, etwa

$$x_6 = \beta \in k\,.$$

Dann wird

$$x_2 = \frac{7}{2}\alpha - \beta + \frac{17}{2}\,,$$

was wir schließlich zusammen mit all den anderen bereits berechneten x_i in die erste Gleichung einsetzen können:

$$2x_1 + \frac{11}{5}\alpha + \frac{79}{5} = 12$$

hat die Lösung

$$x_1 = -\frac{11}{10}\alpha - \frac{19}{10}\,.$$

Damit hängt also die allgemeine Lösung dieses linearen Gleichungssystems von zwei frei wählbaren Parametern α und β ab. Sie wird geringfügig übersichtlicher, wenn wir α als $\alpha = 10\gamma$ schreiben; dann ist $\mathcal{L}$ gleich der Menge

$$\left\{\left(-11\gamma - \tfrac{19}{10}, 35\gamma - \beta + \tfrac{17}{2}, 186\gamma + \tfrac{252}{5}, -3\gamma + \tfrac{3}{10}, 10\gamma, \beta\right) \,\middle|\, \beta, \gamma \in k\right\}.$$

Als nächstes Beispiel wollen wir ein System betrachten, daß von zwar festen, aber nicht numerisch gegebenen Parametern abhängt: Wir betrachten das Gleichstromnetzwerk aus Abbildung zwölf mit bekannten Widerständen $R_1, \ldots, R_5$ und bekanntem Eingangs- und Ausgangsstrom I; gesucht sind die Ströme $I_1, \ldots, I_5$.

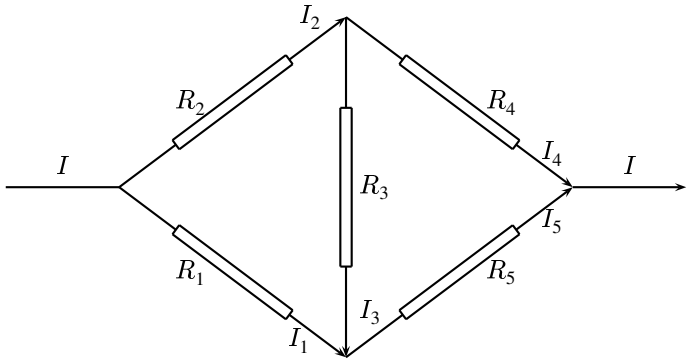


Abb. 12: Ein Gleichstromnetzwerk

Nach den KIRCHHOFFschen Gesetzen müssen sich zunächst an allen Knoten die eingehenden und die ausgehenden Ströme zu Null ergänzen, d.h. wir erhalten die Gleichungen

$$I_1 + I_2 = I = I_4 + I_5$$

für Anfang und Ende sowie

$$I_2 = I_3 + I_4 \quad \text{und} \quad I_1 + I_3 = I_5$$

für oben und unten. Außerdem müssen sich die Spannungen in den beiden Dreiecken zu Null summieren; nach dem OHMschen Gesetz führt dies auf die beiden Gleichungen

$$R_2 I_2 + R_3 I_3 - R_1 I_1 = 0 \quad \text{und} \quad R_3 I_3 + R_5 I_5 - R_4 I_4 = 0\,.$$

Ordnen wir dies nach den Strömen I_ν, erhalten wir also das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcrcl}
I_1 & + & I_2 & & & & & & & = & I \\
 & & & & & & I_4 & + & I_5 & = & I \\
 & & I_2 & - & I_3 & - & I_4 & & & = & 0 \\
I_1 & & & + & I_3 & & & - & I_5 & = & 0 \\
-R_1 I_1 & + & R_2 I_2 & + & R_3 I_3 & & & & & = & 0 \\
 & & & & -R_3 I_3 & + & R_4 I_4 & - & R_5 I_5 & = & 0\,.
\end{array}$$

Die Elimination von I_1 aus allen Gleichungen ab der zweiten ist einfach: Wir müssen die erste Gleichung von der vierten subtrahieren und ihr R_1-faches zur fünften addieren; das neue System ist

$$\begin{array}{rcrcrcrcrcr}
I_1 & + & I_2 & & & & & & & = & I \\
 & & & & & & I_4 & + & I_5 & = & I \\
 & & I_2 & - & I_3 & - & I_4 & & & = & 0 \\
 & & -I_2 & + & I_3 & & & - & I_5 & = & -I \\
 & & (R_1 + R_2) I_2 & + & R_3 I_3 & & & & & = & R_1 I \\
 & & & & -R_3 I_3 & + & R_4 I_4 & - & R_5 I_5 & = & 0\,,
\end{array}$$

Da I_2 in der zweiten Gleichung nicht vorkommt, vertauschen wir diese mit der dritten; danach können wir letztere zur vierten addieren und ihr $(R_1 + R_2)$-faches von der fünften subtrahieren, um I_2 aus allen weiteren Gleichungen zu eliminieren.

Damit weiterhin jede Gleichung in eine Zeile paßt, setzen wir zur Abkürzung

$$R_{12} \underset{\text{def}}{=} R_1 + R_2 \quad \text{und} \quad R_{123} \underset{\text{def}}{=} R_1 + R_2 + R_3$$

und erhalten

$$\begin{array}{rcrcrcrcr}
I_1 & + & I_2 & & & & & = & I \\
 & & I_2 & - & I_3 & - & I_4 & = & 0 \\
 & & & & & & I_4 \;+\; I_5 & = & I \\
 & & & & & & -I_4 \;+\; -I_5 & = & -I \\
 & & & & R_{123}I_3 & + & R_{12}I_4 & = & R_1 I \\
 & & & & -R_3 I_3 & + & R_4 I_4 \;-\; R_5 I_5 & = & 0\,.
\end{array}$$

Zur Elimination von I_3 bietet sich an, die dritte Gleichung, in der I_3 nicht vorkommt, mit der sechsten zu vertauschen (diese hat einen einfacheren I_3-Koeffizienten als die fünfte), und dann das R_{123}/R_3-fache dieser Gleichung zur fünften zu addieren. Dazu müssen wir uns allerdings zunächst überlegen, ob das überhaupt zulässig ist: Falls $R_3 = 0$ ist, dürfen wir natürlich nicht durch R_3 dividieren; wenn dieser Fall nicht ausgeschlossen werden kann, muß er also ab hier getrennt behandelt werden.

Im vorliegenden Beispiel wollen wir aber davon ausgehen, daß alle fünf Widerstände tatsächlich vorhanden sind, so daß alle R_i positive reelle Zahlen sind. Dann können wir durch R_3 dividieren und das System wird zu

$$\begin{array}{rcrcrcrcr}
I_1 & + & I_2 & & & & & = & I \\
 & & I_2 & - & I_3 & - & I_4 & = & 0 \\
 & & & & -R_3 I_3 & + & R_4 I_4 \;-\; R_5 I_5 & = & 0 \\
 & & & & & & -I_4 \;+\; -I_5 & = & -I \\
 & & & & & & \alpha I_4 \;+\; \beta I_5 & = & R_1 I \\
 & & & & & & I_4 \;+\; I_5 & = & I
\end{array}$$

mit

$$\alpha = R_{12} + \frac{R_4 R_{123}}{R_3} \quad \text{und} \quad \beta = \frac{-R_5 R_{123}}{R_3}\,.$$

Schließlich muß noch I_4 aus den letzten beiden Gleichungen eliminiert werden; wir addieren also die vierte Gleichung unverändert zur letzten

und ihr α-faches zur vorletzten:

$$\begin{array}{rcrcrcrcr}
I_1 & + & I_2 & & & & & = & I \\
 & & I_2 & - & I_3 & - & I_4 & = & 0 \\
 & & & & -R_3 I_3 & + & R_4 I_4 \;-\; R_5 I_5 & = & 0 \\
 & & & & & & -I_4 \;+\; -I_5 & = & -I \\
 & & & & & & (\beta-\alpha) I_5 & = & (R_1-\alpha) I \\
 & & & & & & 0 & = & 0
\end{array}$$

Damit können wir nacheinander

$$I_5 = \frac{R_1 - \alpha}{\beta - \alpha} \cdot I\,, \qquad I_4 = I - I_5\,, \qquad I_3 = \frac{R_4 I_4 - R_5 I_5}{R_3}\,,$$
$$I_2 = I_3 + I_4 \qquad \text{und} \qquad I_1 = I - I_2$$

bestimmen, wobei sich der Leser noch überlegen sollte, warum die Division durch $\beta - \alpha$ unproblematisch ist.

Zum Berechnen der Ströme in konkreten Beispielen reicht diese Lösungsformel aus; ist man allerdings an einem symbolischen Ausdruck interessiert, muß man die Definitionen von $\alpha, \beta, R_{12}, R_{123}$ einsetzen und nacheinander alle rechte Seiten auf Ausdrücke nur in I und den R_i reduzieren. Dies ist eine *im Prinzip* einfache Übungsaufgabe in Bruchrechnung, die man allerdings im vorliegenden Fall besser einem Computeralgebrasystem überläßt. Als Ergebnis erhält man die expliziten Formeln

$$I_1 = \frac{\big(R_2R_5 + R_2R_3 + R_2R_4 + R_3R_4\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$

$$I_2 = \frac{\big(R_1R_4 + R_1R_5 + R_3R_5 + R_1R_3\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$

$$I_3 = \frac{\big(R_1R_4 - R_2R_5\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$

$$I_4 = \frac{\big(R_1R_3 + R_1R_5 + R_2R_5 + R_3R_5\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$

$$I_5 = \frac{\big(R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}\,,$$

die für die meisten *konkreten* Anwendungen erheblich weniger nützlich sind als die obige Form des Ergebnisses.

Als letztes Beispiel schließlich betrachten wir eines, das auch von einem Parameter abhängt, bei dem man aber *nicht* wie im obigen Beispiel einfach durch Ausdrücke im Parameter dividieren darf. (Eigentlich hätten wir es da auch nicht immer dürfen, aber wir haben einfach angenommen, daß alle Widerstände wirklich vorhanden und damit positiv sind; da Kurzschlüsse immer wieder vorkommen, ist diese Annahme nicht hundertprozentig realistisch.) Das Gleichungssystem hänge ab von einem Paramter $a \in k$ und habe die Form

$$\begin{aligned} x_1 + ax_2 + x_3 &= 1 \\ x_1 + x_2 &= 1 \\ -2x_1 - 2ax_2 - ax_3 &= 1\,. \end{aligned}$$

Ein Computeralgebrasystem findet unschwer die Lösung

$$x_1 = \frac{a^2 - 3a - 1}{a^2 - 3a + 2}, \qquad x_2 = \frac{3}{a^2 - 3a + 2}, \qquad x_3 = \frac{-3}{a - 2}\,.$$

Diese „Lösung“ hat aber für $a = 2$ und auch für $a = 1$ Nullen im Nenner, ist dort also nicht erklärt. Für ein Computeralgebrasystem ist das kein Problem: Wie der Name sagt, rechnet es *algebraisch,* und da ist a keine reelle Zahl, sondern ein Symbol, das nichts mit irgendwelchen Zahlen zu tun hat. Damit ist $a - 2$ ein formaler Ausdruck, der nie Null sein kann, denn das *Symbol* a ist schließlich verschieden von der *Zahl* zwei.

Dieses Rechnen in sogenannten Funktionenkörpern ist mathematisch problemlos, ist aber nicht das, was in den meisten Anwendungen gefragt ist: Dort steht a im allgemeinen für eine variable Größe, in Abhängigkeit von der das Gleichungssystem gelöst werden soll. Man kann sich beispielsweise vorstellen, daß das Gleichungssystem ein lineares Regelungsproblem beschreibt in Abhängigkeit von steuerbaren Größen x_1, x_2 und x_3, wobei die zu steuernden Größen zu

$$x_1 + ax_2 + x_3, \quad x_1 + x_2 \quad \text{und} \quad -2x_1 - 2ax_2 - ax_3$$

werden. Der Paramter a wäre dann zu interpretieren als eine von außen vorgegebene Umgebungsbedingung (z.B. die Temperatur), und das lineare Gleichungssystem besagt, daß wir das System so regeln wollen,

daß die drei steuerbaren Größen allesamt eins werden – unabhängig von der Außentemperatur.

Bei einer solchen Interpretation können wir natürlich *nicht* einfach durch $a - 2$ dividieren; ein Ergebnis, wie $x_3 = -3/(a - 2)$ besagt in so einem Fall, daß das Ziel im Falle $a = 2$ nicht erreichbar ist, und das ist ein sehr wichtiges Ergebnis. Bei einem System, das einen Stromkreis beschreibt, könnte das zum Beisiel bedeuten, daß beim Parameterwert $a = 2$ ein Kurzschluß entsteht, so daß dieser Parameterwert unbedingt verhindert werden muß. Deshalb muß man bei einem Gleichungssystem, das ein reales Problem beschreibt, vor jeden Division durch einen parameterabhängigen Ausdruck garantieren, daß dieser Ausdruck von Null verschieden ist, und man muß die Fälle, in denen er Null wird, gesondert diskutieren.

Im vorliegenden Beispiel (einer Vordiplomsaufgabe vom April 1999) führt dies auf folgende Lösung:

Subtraktion der ersten Gleichung von der zweiten sowie Addition der zweifachen ersten Gleichung zur dritten ergibt

$$\begin{aligned} (a-1)x_2 + x_3 &= 0 \\ (2-a)x_3 &= 3\,. \end{aligned}$$

Für $a = 2$ ist die letzte dieser beiden Gleichungen unlösbar, ansonsten ist

$$x_3 = \frac{3}{2-a} \quad \text{falls} \quad a \neq 2\,.$$

Für $a = 1$ wird die vorletzte Gleichung zu $x_3 = 0$, was der schon gefundenen Lösung

$$x_3 = \frac{3}{2-a} = 1$$

widerspricht; auch dann ist also das Gleichungssystem unlösbar. In allen anderen Fällen erhalten wir

$$x_2 = \frac{3}{(a-1)(a-2)} \quad \text{falls} \quad a \neq 1, 2\,.$$

Schließlich läßt sich noch, beispielsweise aus der zweiten Gleichung des ursprünglichen Systems, x_1 berechnen und wir erhalten als Ergebnis:

Die Lösungsmenge ist

$$\mathcal{L} = \left\{ \left(1 - \frac{3}{(a-1)(a-2)}, \frac{3}{(a-1)(a-2)}, \frac{3}{a-2} \right) \right\} ,$$

falls $a \neq 1, 2$, und

$$\mathcal{L} = \emptyset , \quad \text{falls } a = 1 \text{ oder } a = 2$$

ist.

In einem solchen Fall wird man durch die Nullstellen des Nenners gewarnt, daß hier etwas schiefgehen muß; es gibt aber auch Beispiele, in denen ein Computeralgebrasystem Lösungen schlichtweg „übersieht“: Bertrachten wir etwa das lineare Gleichungssystem

$$\begin{aligned} x_1 + \phantom{3x_2 +{}} 2x_3 &= 9 \\ 2x_1 + 3x_2 + x_3 &= 9 \\ - x_2 + ax_3 &= a+2 . \end{aligned}$$

Ein Computer findet leicht die Lösung

$$x = 7, \quad y = -2 \quad \text{und} \quad z = 1 .$$

Der GAUSSalgorithmus führt uns aber über das Gleichungssystem

$$\begin{aligned} x_1 + \phantom{3x_2 -{}} 2x_3 &= 9 \\ 3x_2 - 3x_3 &= -9 \\ - x_2 + ax_3 &= a+2 \end{aligned}$$

oder

$$\begin{aligned} x_1 + \phantom{x_2 -{}} 2x_3 &= 9 \\ x_2 - x_3 &= -3 \\ - x_2 + ax_3 &= a+2 \end{aligned}$$

auf die Endgestalt

$$\begin{aligned} x_1 + \phantom{x_2 -{}} 2x_3 &= 9 \\ x_2 - x_3 &= -3 \\ (a-1)x_3 &= a-1 , \end{aligned}$$

in der man nur für $a \neq 1$ aus der letzten Gleichung schließen darf, daß $z = 1$ ist; für $a = 1$ haben wir die immer erfüllte Gleichung $0z = 0$. Die Lösungsmenge ist hier also

$$\mathcal{L} = \big\{ (7, -2, 1) \big\} \quad \text{für } a \neq 1$$

und

$$\mathcal{L} = \big\{ (-2\lambda + 9, \lambda - 3, \lambda) \;\big|\; \lambda \in \mathbb{R} \big\} \quad \text{für } a = 1 .$$

Jemand, der Maple hinreichend gut kennt, hätte natürlich auch diese vollständige Lösung damit ermitteln können, aber der einfachstmögliche Befehl reicht definitiv nicht aus – zumindest ein Grund, warum man auch heute noch lernen muß, lineare Gleichungssysteme von Hand zu lösen.

Ein anderer Grund, warum speziell Technische Informatiker das lernen müssen, liegt in der Natur vieler Anwendungen: Lineare Gleichungssysteme müssen beispielsweise gelöst werden bei Steuerungs- und Regelungsproblemen. In vielen Fällen wird diese Steuerung nicht von einem leistungsfähigen Universalrechner durchgeführt, sondern von einer eigens dafür entwickelten Schaltung, die mit möglichst wenig Aufwand arbeiten soll – sei es aus Kostengründen oder wegen des Raumbedarfs oder der Wärmeentwicklung. In solchen Fällen geht es dann darum, die Lösung möglichst effizient zu ermitteln, und bei der Definition des Wortes „effizient“ können hier durchaus auch nichtmathematische Gesichtspunkte eine Rolle spielen. Daher ist es wichtig, das volle Instrumentarium der Lösungstheorie linearer Gleichungssysteme zu beherrschen, um die jeweils beste Methode implementieren zu können. Deshalb werden wir auch noch Alternativen zum GAUSS-Algorithmus betrachten, und in der *Numerik I* werden weitere Verfahren folgen.

f) Die Struktur der Lösungsmenge

Nach diesen Beispielen ist es an der Zeit, wieder zu den theoretischen Grundlagen zurückzukehren. Sei also wieder

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ \vdots \qquad\qquad &= \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n , \end{aligned}$$

ein allgemeines lineares Gleichungssystem. Wenn wir die a_{ij} zu einer

Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

zusammenfassen und die Unbekannten und rechten Seiten zu Vektoren

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

läßt sich das Gleichungssystem auch kurz als $A\vec{x} = \vec{b}$ schreiben, wobei wir das Produkt der Matrix A mit dem Vektor $\vec{x}$ so definieren, daß es jener Vektor aus k^n sein soll, dessen i-te Komponente die Summe

$$\sum_{j=1}^{m} a_{ij} x_j$$

sein soll. Identifiziert man Vektoren aus k^m mit $m \times 1$-Matrizen, ist das gerade das Produkt der $n \times m$-Matrix A mit der $m \times 1$-Matrix der Variablen.

Wir bezeichen A als die *Matrix des Gleichungssystems* und

$$\left(A \mid \vec{b}\right) \underset{\text{def}}{=} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_m \end{pmatrix}$$

als die *erweiterte Matrix.*

Damit können wir das Gleichungssystem in die Sprache der Vektorräume und linearen Abbildungen übersetzen: Wir betrachten die lineare Abbildung

$$\varphi\colon k^m \to k^n; \quad \vec{v} \mapsto A\vec{v},$$

und die Lösungsmenge des linearen Gleichungssystems ist

$$\mathcal{L} = \left\{\vec{v} \in k^m \;\middle|\; \varphi(\vec{v}) = \vec{b}\right\}.$$

Für zwei Lösungsvektoren $\vec{v}, \vec{w} \in k^m$ ist

$$\varphi(\vec{v} - \vec{w}) = \varphi(\vec{v}) - \varphi(\vec{w}) = \vec{b} - \vec{b} = \vec{0},$$

die Differenz zweier Lösungen ist also eine Lösung des entsprechenden linearen Gleichungssystems mit lauter Nullen auf der rechter Seite.

Definition: Ein lineares Gleichungssystem $A\vec{x} = \vec{b}$ heißt *homogen,* wenn $\vec{b} = \vec{0}$ ist; ansonsten heißt es *inhomogen.* $A\vec{x} = \vec{0}$ heißt das zu $A\vec{x} = \vec{b}$ gehörige homogene Gleichungssystem.

Damit wissen wir also

Lemma: Zwei Lösungen des linearen Gleichungssystems $A\vec{x} = \vec{b}$ unterscheiden sich durch eine Lösung des homogenen Gleichungssystems. ■

Die Lösung eines homogenen Gleichungssystems besteht aus allen Vektoren aus k^m, die von der oben definierten linearen Abbildung φ auf den Nullvektor abgebildet werden, ist also gerade der Kern von φ und damit ein Untervektorraum von k^n. Insbesondere ist die Lösungsmenge eines homogenen linearen Gleichungssystems also nie leer: Wie man sofort auch direkt sieht, gibt es immer die sogenannte *triviale* Lösung, bei der alle Variablen gleich Null sind.

Nach der Dimensionsformel aus §1*i)* ist

$$\dim \operatorname{Kern} \varphi = m - \dim \operatorname{Bild} \varphi,$$

und das Bild wird erzeugt von den m Vektoren $\varphi(\vec{e}_i) = A\vec{e}_i$. Das sind aber gerade die Spaltenvektoren der Matrix A; die Dimension des Bildes ist also gleich der Anzahl linear unabhängiger Spaltenvektoren von A, und diese Zahl hatten wir oben als den (Spalten-)Rang der Matrix definiert. Damit wissen wir also

Lemma: Die Lösungsmenge des homogenen linearen Gleichungssystems $A\vec{x} = \vec{0}$ in m Variablen ist ein Untervektorraum der Dimension $m - \operatorname{Rang} A$. ■

Im Falle eines inhomogenen Gleichungssystems ist die Sache nicht ganz so einfach: Schließlich hatten wir bei den Beispielen schon gesehen,

daß von zwei Gleichungssystemen, die sich nur in ihrer rechten Seite unterscheiden, das eine unlösbar sein kann, während das andere eine oder mehrere Lösungen hat.

Der Grund dafür wird klar, wenn wir das Gleichungssystem wieder über die lineare Abbildung φ interpretieren: Das Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn die rechte Seite $\vec{b}$ im Bild von φ liegt. Damit muß sie linear abhängig sein von den Spaltenvektoren von A, die ja den Bildraum erzeugen, d.h. der Rang der *erweiterten* Matrix, die außer den Spaltenvektoren von A noch den Vektor $\vec{b}$ enthält, darf nicht größer sein, als der von A. Kleiner kann er natürlich nicht sein, also haben wir gezeigt

Lemma: Ein inhomogenes Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn der Rang r seiner erweiterten Matrix gleich dem von A ist. Es ist genau dann eindeutig lösbar, wenn dieser gemeinsame Rang gleich der Anzahl m der Variablen ist. ■

Ist der gemeinsame Rang von Matrix und erweiterter Matrix kleiner als m, so ist das Gleichungssystem nicht eindeutig lösbar. Da sich zwei Lösungen um eine Lösung des zugehörigen homogenen Gleichungssystems unterscheiden und diese wiederum einen Vektorraum der Dimension $m - \text{Rang}\, A$ bilden, hängt die Lösung dann von $m - \text{Rang}\, A$ Parametern ab, ist aber für ein echt inhomogenes Gleichungssystem kein Vektorraum: Ein Vektorraum enthält stets den Nullvektor, und wenn es eine Lösung gibt, bei der *alle* Variablen verschwinden, stehen auf der rechten Seite des Gleichungssystems lauter Nullen, d.h. das Gleichungssystem ist homogen.

Das obige Lemma wird nur selten von praktischem Nutzen sein, denn wenn Matrix und erweiterte Matrix keine sehr spezielle Gestalt haben, wird der GAUSS-Algorithmus im allgemeinen die einfachste Möglichkeit sein, um die Ränge von Matrix und erweiterter Matrix zu bestimmen. Man muß ihn allerdings nicht ganz zu Ende durchführen, denn die Ränge sind schon klar, wenn auf der linken Seite Trepengestalt erreicht ist. Der Vollständigkeit halber sei hier kurz angegeben, wie diese im allerallgemeinsten Fall aussieht:

Die verschiedenen Schritte des GAUSS-Algorithmus wirken sich auf die Matrix wie auch auf die erweiterte Matrix so aus, daß entweder zwei Zeilen miteinander vertauscht werden, oder aber ein Vielfaches einer Zeile zu einer anderen Zeile addiert wird – Addition bedeutet dabei die gewöhnliche komponentenweise Addition, d.h. die Addition von Zeilenvektoren. Am Ende entsteht eine erweiterte Matrix der Form

$$\left(\begin{array}{cccccccccccc}
[0] & \bullet & [\star] & \star & [\star] & \star & \dots & [\star] & \star & [\star] & \star \\
[0] & 0 & [0] & \bullet & [\star] & \star & \dots & [\star] & \star & [\star] & \star \\
[0] & 0 & [0] & 0 & [0] & \bullet & \dots & [\star] & \star & [\star] & \star \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & \bullet & [\star] & \star \\
\hline
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & \bullet \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & \bullet \\
\hline
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & 0
\end{array}\right),$$

wobei $\bullet$ für ein von Null verschiedenes und $\star$ für ein beliebiges Element des Körpers k steht; [0] und $[\star]$ stehen für keine, eine oder mehrere Nullen *bzw.* beliebige Elemente, deren Anzahl in untereinanderstehenden Termen jeweils die gleiche sein soll.

Die Zeilen zwischen den beiden eingezeichneten Linien, in denen alle Einträge der Matrix des Gleichungssystems verschwinden, müssen natürlich nicht wirklich auftreten, genauso wenig die Zeilen unterhalb der zweiten Linie, wo sogar alle Einträge der erweiterten Matrix verschwinden, so daß auch die rechten Seiten der entsprechenden Gleichungen in der Endgestalt nach Anwendung des GAUSS-Algorithmus Null sind.

Falls zwischen den beiden eingezeichneten Linien wirklich eine oder mehrere Zeilen stehen, ist das Gleichungssystem *unlösbar,* denn dann

gibt es ja in der Endform des Gleichungssystems Gleichungen, in denen links alle Koeffizienten Null sind, während rechts eine von Null verschiedene Zahl steht. Indem man gegebenenfalls ein Vielfaches der ersten solchen Gleichung von den etwa vorhandenen weiteren Gleichungen subtrahiert, kann man erreichen, daß alle weiteren Gleichungen zu $0 = 0$ werden, d.h. wir können erreichen, daß zwischen den beiden Linien *höchstens eine* Zeile steht, und das lineare Gleichungssystem ist genau dann lösbar, wenn keine dort steht.

Damit ist klar, daß die Anzahl der Zeilen oberhalb der ersten Linie der *Rang der Matrix A* ist und die Anzahl der Zeilen oberhalb der zweiten Linie (nachdem man dafür Sorge getragen hat, daß höchstens eine Zeile zwischen den beiden Linien steht) der *Rang der erweiterten Matrix.*

Noch etwas läßt sich diesem Diagramm ansehen: Der Rang der Matrix, die maximale Anzahl linear unabhängiger Spalten also, ist offenbar gerade gleich der Anzahl der Zeichen • oberhalb des ersten Strichs. Genau das ist aber auch die maximale Anzahl linear unabhängiger Zeilen der Matrix, d.h. wir können am Rande noch notieren, daß der *Zeilrang* einer Matrix gleich dem *Spaltenrang* ist, was die Kurzbezeichnung *Rang* rechtfertigt.

Betrachten wir zum Abschluß noch ein Beispiel für das Rangkriterium zur Lösbarkeit eines linearen Gleichungssystems: Im vorigen Abschnitt hatten wir das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcrcrcr}
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15 \\
4x_1 & - & 7x_2 & & & - & 3x_4 & & & + & x_6 & = & 0 \\
2x_1 & - & 7x_2 & + & x_3 & + & 13x_4 & + & 12x_5 & - & 7x_6 & = & -8
\end{array}$$

als unlösbar erkannt. Seine erweiterte Matrix ist

$$\begin{pmatrix} 0 & 1 & 0 & 5 & -2 & 1 & 10 \\ 2 & 2 & 0 & -4 & -6 & 2 & 12 \\ 6 & -3 & 1 & 5 & 0 & -3 & 15 \\ 4 & -7 & 0 & -3 & 0 & 1 & 0 \\ 2 & -7 & 1 & 13 & 12 & -7 & -8 \end{pmatrix},$$

und nach Anwendung des GAUSS-Algorithmus kommen wir auf die Endgestalt

$$\left(\begin{array}{rrrrrrr} 2 & 2 & 0 & -4 & -6 & 2 & 12 \\ 0 & 1 & 0 & 5 & -2 & 1 & 10 \\ 0 & 0 & 1 & 62 & 0 & 0 & 69 \\ 0 & 0 & 0 & 20 & 6 & 0 & 6 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

Somit hat die Matrix des linearen Gleichungssystems den Rang vier, wohingegen die erweiterte Matrix den Rang fünf hat.

Ändern wir im Gleichungssystem die rechte Seite der letzten Gleichung von -8 zu -9, so wird es lösbar. In der Endgestalt ändert sich, wie wir oben gesehen haben, auch nur die rechte Seite der letzten Gleichung; die erweiterte Matrix wird also schließlich zu

$$\left(\begin{array}{rrrrrrr} 2 & 2 & 0 & -4 & -6 & 2 & 12 \\ 0 & 1 & 0 & 5 & -2 & 1 & 10 \\ 0 & 0 & 1 & 62 & 0 & 0 & 69 \\ 0 & 0 & 0 & 20 & 6 & 0 & 6 \\ \hline\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Hier haben Matrix und erweiterte Matrix denselben Rang vier.

g) Affine Räume

Für inhomogene Gleichungssysteme kennen wir bislang noch nicht die Struktur des Lösungsraums selbst; wir wissen nur, daß die *Differenzen* von Lösungen (so es welche gibt) einen Vektorraum bilden.

Dies sollte etwas an die Situation in §1*a)* erinnern, wo wir zwei Punkten deren *Verbindungsvektor* zugeordnet hatten und zwei Vektoren unabhängig von ihren Anfangs- und Endpunkten als gleich bezeichnet hatten, wenn sie nur die gleiche Länge und Richtung hatten. Eine ähnliche Situation haben wir auch hier: Lösungen eines inhomogenen Gleichungssystems (so es welche gibt) verhalten sich wie Punkte, Lösungen eines homogenen Gleichungssystems wie Vektoren.

Beispielsweise gibt es keine sinnvolle Addition zweier Punkte, genauso wie auch die Summe zweier Lösungen einen inhomogenen Gleichungssystems keine Lösung mehr ist. Andererseits kann man einen Vektor an

einem Punkt abtragen, um einen weiteren Punkt zu bekommen, und genauso kann man auch eine Lösung des homogenen Gleichungssystems zu einer Lösung des inhomogenen addieren, um eine weitere Lösung des inhomogenen zu erhalten.

Die Mathematik sucht hinter analogen Sachverhalten gemeinsame Strukturen; in diesem Fall bezeichnet sie diese als *affine Räume:*

Definition: Ein affiner Raum über dem Körper k ist gegeben durch ein Paar (A, V) bestehend aus einer nichtleeren Menge A, deren Elemente wir als Punkte bezeichnen, und einem Vektorraum V, so daß gilt:

1.) Es gibt eine Abbildung $\varphi\colon \begin{cases} A \times A \to V \\ (\boldsymbol{x}, \boldsymbol{y}) \mapsto \overrightarrow{\boldsymbol{x}\boldsymbol{y}} \end{cases}$, mit der Eigenschaft, daß

$$\overrightarrow{\boldsymbol{x}\boldsymbol{y}} + \overrightarrow{\boldsymbol{y}\boldsymbol{z}} = \overrightarrow{\boldsymbol{x}\boldsymbol{z}} \quad \text{für alle } \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in A\,.$$

2.) Für jeden Punkt $O \in A$ ist die Abbildung

$$\varphi_O\colon A \to V; \quad \boldsymbol{x} \mapsto \varphi(O, \boldsymbol{x}) = \overrightarrow{O\boldsymbol{x}}$$

bijektiv; ihre Umkehrabbildung sei mit ψ_O bezeichnet.

Die Dimension des Vektorraums V bezeichnen wir auch als Dimension des affinen Raums.

Anschaulich gesehen ordnet die Abbildung φ zwei Punkten $\boldsymbol{x}, \boldsymbol{y} \in A$ den Verbindungsvektor $\overrightarrow{\boldsymbol{x}\boldsymbol{y}}$ zu, und ψ_O beschreibt das *Abtragen* eines Vektors: In §1 hatten wir Vektoren eingeführt als Äquivalenzklassen von Pfeilen gleicher Länge und gleicher Richtung; nimmt man zum Vektor $\vec{v}$ den Pfeil, der im Punkt O beginnt, so hat dieser den Endpunkt $\psi_O(\vec{v})$.

Standardbeispiel eines affinen Raums ist der $\mathbb{R}^n$; hier ist A gleich der *Menge* $\mathbb{R}^n$, deren Elemente wir als Punkte $\boldsymbol{x} = (x_1, \ldots, x_n)$ schreiben, und V ist gleich dem *Vektorraum* $\mathbb{R}^n$, dessen Elemente wir als Spaltenvektoren schreiben. φ ist die Abbildung, die den Punkten

$$\boldsymbol{x} = (x_1, \ldots, x_n) \quad \text{und} \quad \boldsymbol{y} = (y_1, \ldots, y_n)$$

deren Verbindungsvektor

$$\varphi(\boldsymbol{x}, \boldsymbol{y}) = \overrightarrow{\boldsymbol{x}\boldsymbol{y}} = \begin{pmatrix} y_1 - x_1 \\ \vdots \\ y_n - x_n \end{pmatrix}$$

zuordnet, und für $O = (a_1, \ldots, a_n)$ ist

$$\psi_O\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = (a_1 + x_1, \ldots, a_n + x_n)\,.$$

Ganz entsprechend läßt sich für jeden beliebigen Körper k die Menge k^n zu einem affinen Raum machen; der zugehörige Vektorraum ist natürlich der *Vektorraum* k^n.

Weitere interessante Beispiele sind vor allem die Geraden, Ebenen *usw.* in diesem Raum; diese fassen wir zusammen unter der

Definition: Ein affiner Unterraum des affinen Raums (A, V) ist ein affiner Raum (B, U) mit $B \subseteq A$ und $U \leq V$, wobei auch die Abbildungen φ und ψ_O Einschränkungen der entsprechenden Abbildungen von (A, V) sind.

Betrachten wir etwa die Gerade

$$\left\{(1 + \lambda, \lambda) \;\middle|\; \lambda \in \mathbb{R}\right\} \subseteq \mathbb{R}^2$$

durch den Punkt $(1, 0)$ mit Steigung 1. Hier ist B natürlich gleich der Menge aller Punkte auf der Geraden, also die gerade angegebene Menge selbst, und als Vektorraum nehmen wir den vom Richtungsvektor $\binom{1}{1}$ der Geraden aufgespannten Untervektorraum des $\mathbb{R}^2$.

Dieselbe Gerade können wir auch durch die Gleichung

$$y = x + 1 \quad \text{oder} \quad -x + y = 1$$

beschreiben. Allgemein können wir, wie wir uns als nächstes überlegen wollen, jeden affinen Unterraum eines k^n als Lösungsmenge eines linearen Gleichungssystems in n Unbekannten interpretieren, und umgekehrt ist auch jede solche Lösungsmenge, falls sie nicht leer ist, affiner

Unterraum von k^n. (Mit k^n ist hier natürlich das Paar (k^n, k^n) gemeint, das wir kurz, wenn auch schlampig, als affinen Raum k^n bezeichnen.)

Beginnen wir mit der Lösungsmenge L eines linearen Gleichungssystems

$$A\vec{x} = \vec{b} \quad \text{mit} \quad A \in k^{m\times n}, \quad \vec{b} \in k^m$$

für einen Vektor $\vec{x} \in k^n$; mit anderen Worten mit der Lösungsmenge von m linearen Gleichungen in n Unbekannten. (Hier schreiben wir, aus alter Gewohnheit, die Punkte eines affinen Raums ausnahmsweise als Vektoren.) *Falls* es Lösungen gibt, ist die Differenz $\vec{y} = \vec{x}^{(1)} - \vec{x}^{(2)}$ zweier (nicht notwendigerweise verschiedener) Lösungen $\vec{x}^{(1)}$ und $\vec{x}^{(2)}$ eine Lösung des homogenen Gleichungssystems $A\vec{y} = \vec{0}$.

Die Lösungsmenge eines *homogenen* linearen Gleichungssystems in n Variablen ist aber Kern einer linearen Abbildung und daher (und auch aus vielen anderen Gründen) ein Untervektorraum U des k^n. Zusammen mit diesem Untervektorraum wird L zu einem affinen Raum, wobei die Abbildung $\varphi\colon L \times L \to U$ zwei Lösungen $\vec{x}^{(1)}$ und $\vec{x}^{(2)}$ deren Differenz $\vec{y} = \vec{x}^{(1)} - \vec{x}^{(2)}$ als „Verbindungsvektor“ zuordnet; entsprechend macht für jede Lösung $\vec{x}$ des inhomogenen Gleichungssystems die Abbildung $\psi_{\vec{x}}\colon U \to L$ eine Lösung $\vec{y}$ des homogenen Gleichungssystems durch Addition von $\vec{x}$ zu einer Lösung des inhomogenen Gleichungssystems.

Ist umgekehrt (B, U) ein affiner Unterraum von (k^n, k^n), so können wir eine Basis $\{\vec{b}_1, \ldots, \vec{b}_r\}$ von U wählen und diese zu einer Basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$ des k^n ergänzen. Bezüglich dieser neuen Basis des k^n ist dann U die Lösungsmenge des homogenen linearen Gleichungssystems

$$x_{r+1} = 0, \quad x_{r+2} = 0, \quad \ldots, \quad x_n = 0\,.$$

Sind nun $(c_1, \ldots, c_n)$ und $(d_1, \ldots, d_n)$ zwei Punkte aus B, so liegt ihre Differenz (d.h. ihr „Verbindungsvektor“) in U, und somit muß für alle $i > r$ notwendigerweise $c_i = d_i$ sein. Damit ist B bezüglich der Basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$ die Lösungsmenge des linearen Gleichungssystems

$$x_{r+1} = d_{r+1}, \quad x_{r+2} = d_{r+2}, \quad \ldots, \quad x_n = d_n\,.$$

Durch Rücktransformation auf die Standardbasis des k^n kann man dieses lineare Gleichungssystem in ein lineares Gleichungssystem bezüglich der ursprünglichen Basis umformen.

Um den Kontakt an die Schulgeometrie zu verdeutlichen, sei für Interessenten noch kurz dargestellt, was man mit affinen Unterräumen *geometrisch* anstellen kann.

Spätestens die Interpretation als Lösungsmenge eines linearen Gleichungssystems sollte jedem klargemacht haben, daß der Durchschnitt zweier affiner Unterräume entweder leer ist oder wieder ein affiner Unterraum; in beiden Fällen kann er durch Lösen eines linearen Gleichungssystems ermittelt werden. Seine Dimension hängt offensichtlich von einer ganzen Reihe von Faktoren ab: Schon in der Ebene kann der Durchschnitt zweier Geraden entweder leer sein (parallele Geraden) oder nulldimensional (ein Schnittpunkt) oder eindimensional (zusammenfallende Geraden).

Definition: Für zwei affine Unterräume (A, U) und (B, V) eines affinen Raums (C, W) ist der von A und B aufgespannte affine Unterraum von (C, W) der kleinste affine Unterraum (S, T) von (C, W), der beide enthält.

Bei dieser Definition stellt sich als erstes die Frage, ob sie überhaupt sinnvoll ist, ob es einen solchen kleinsten Raum also auch wirklich gibt. Dazu bilden wir den Durschnitt aller affiner Unterräume, die (A, U) und (B, V) als Unterräume enthalten. Der Durchschnitt einer beliebigen Menge affiner Unteräume ist entweder leer oder wieder ein affiner Unterraum. Leer kann er hier nicht sein, da wir nur Unterräume bertrachten, die die beiden gegebenen Räume enthalten, also ist er ein affiner Unterraum und offensichtlich der kleinste unter denen, die (A, U) und (B, V) als Unterräume enthalten.

Seine Dimension kann allerdings größer werden als die Summe der Dimensionen der beiden Ausgangsräume; ein wahrscheinlich aus der Schule bekanntes Beispiel dafür sind *windschiefe Geraden* im $\mathbb{R}^3$, d.h zwei Geraden, die zwar nicht parallel zueinander sind, die aber in zwei voneinander verschiedenen parallelen Ebenen liegen.

Ein Beispiel hierfür sind etwa die x-Achse eines kartesischen Koordiantensystems im $\mathbb{R}^3$ und die um eins in z-Richtung verschobene y-Achse. Diese beiden Geraden schneiden sich nicht, denn alle Punkte der ersten

haben z-Koordinate Null, während auf der zweiten $z = 1$ ist, und sie sind auch offensichtlich nicht parallel.

Wir wollen uns überlegen, daß es keine Ebene gibt, die beide Geraden enthält. Dazu gehen wir aus von der allgemeinsten Gleichung

$$ax + by + cz = d$$

für eine Ebene im $\mathbb{R}^3$. Wenn diese Ebene die x-Achse enthalten soll, also die Menge aller Punkte der Form $(x, 0, 0)$, muß $a = d = 0$ sein. Die um eins in z-Richtung verschobene y-Achse besteht aus allen Punkten der Form $(0, y, 1)$; sie liegt in der Ebenen, wenn $b = 0$ und $c = d$ ist. Für eine Ebene, die beide Geraden enthält, müßte also $a = b = c = d = 0$ sein, und dann haben wir keine Ebene mehr. Da es zwischen einer Ebenen und dem gesamten $\mathbb{R}^3$ keine weiteren affinen Unterräume mehr gibt, ist der kleinste affine Unterraum, der die beiden Geraden enthält, der gesamte $\mathbb{R}^3$.

Damit drängt sich die Frage auf, ob es möglicherweise auch im $\mathbb{R}^4$ oder in Räumen noch höherer Dimension zwei Geraden geben kann, die den gesamten Raum aufspannen. Da die Anschauung hier leider nicht weiterhilft, müssen wir abstrakt mathematisch argumentieren. Das wird am einfachsten, wenn wir rechnen können, und dazu führen wir Koordinatensysteme ein:

Definition: Ein Koordinatensystem des affinen Raums (A, V) besteht aus einem Punkt $O \in A$, genannt der Ursprung oder auch Nullpunkt des Koordinatensystems, und einer Basis von V. Die Geraden durch O mit einem Basisvektor als Richtungsvektor heißen *Koordinatenachsen.*

Bezüglich eines solchen Koordinatensystems hat dann in der Tat jeder Punkt $\boldsymbol{x}$ eines n-dimensionalen affinen Raums A ein n-tupel von Koordinaten, nämlich die Komponenten des Vektors $\overrightarrow{O\boldsymbol{x}}$. Bevor wir aber zeigen können, daß der affine Raum dadurch isomorph wird zu k^n, müssen wir zunächst definieren, was strukturerhaltende Abbildungen zwischen affinen Räumen überhaupt sein sollen:

Definition: Eine affine Abbildung zwischen den affinen Räumen (A, V) und (B, W) besteht aus einer Abbildung $\varphi\colon A \to B$ und einer linearen

Abbildung $\psi\colon V \to W$ derart, daß für alle Punkte $\boldsymbol{x}, \boldsymbol{y} \in A$ gilt:

$$\overrightarrow{\varphi(\boldsymbol{x})\,\varphi(\boldsymbol{y})} = \psi\left(\overrightarrow{\boldsymbol{x}\boldsymbol{y}}\right) .$$

Die Abbildung heißt *Isomorphismus* oder *Affinität,* falls φ bijektiv ist.

Offensichtlich ist φ genau dann bijektiv, wenn ψ ein Isomorphismus von Vektorräumen ist, so daß wir dies nicht zusätzlich fordern müssen. Es ist auch klar, daß nach Wahl eines Koordinatensystems eines n-dimensionalen affinen Raums die Koordinaten einen Isomorphismus mit dem affinen Raum k^n definieren.

Da wir wissen, wie lineare Abbildungen zwischen Vektorräumen bezüglich zweier Basen aussehen, können wir auch sofort hinschreiben, wie die Abbildung $\varphi\colon A \to B$ bezüglich zweier Koordinatensysteme aussieht:

$$(x_1, \ldots, x_n) \mapsto \Big(\sum_{j=1}^{n} a_{1j}x_j + b_1, \ldots, \sum_{j=1}^{n} a_{mj}x_j + b_m\Big) ,$$

im Gegensatz zu den homogenen linearen Abbildungen zwischen Vektorräumen sind hier also auch inhomogene lineare Abbildungen möglich. In Matrixschreibweise betrachtet man am besten die Spaltenvektoren: Ist

$$\varphi\big((x_1, \ldots, x_n)\big) = (y_1, \ldots, y_m) ,$$

so ist

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \vec{b} \quad \text{mit} \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} .$$

Geometrisch betrachtet beschreibt der konstante Vektor $\vec{b}$ also eine Verschiebung.

Vor allem in der Computergraphik faßt man die Matrix A und den Vektor $\vec{b}$ oft auch zusammen zu einer Matrix A^* mit $m + 1$ Zeilen und $n + 1$ Spalten, indem man zunächst den Vektor $\vec{b}$ als $(n + 1)$-te Spalte an die Matrix anhängt und dann als $(m + 1)$-te Zeile lauter Nullen einsetzt mit Ausnahme einer Eins an der letzten Stelle:

$$A^* = \begin{pmatrix} A & \vec{b} \\ \mathbf{0} & 1 \end{pmatrix} ,$$

wobei die fette Null für n gewöhnliche Nullen steht. Mit dieser Bezeichnung ist dann

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \\ 1 \end{pmatrix} = A^* \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} ,$$

eine Formel, die zwar rechnerisch aufwendiger ist als die obige, dafür aber etwas kompakter. Sie hat auch den Vorteil, daß sich die Hintereinanderausführung affiner Abbildungen so durch ein Matrixprodukt beschreiben läßt.

Als Anwendung von Koordinatensystemem wollen wir, wie bereits erwähnt, die Dimension des Erzeugnisses zweier affiner Unterräume berechnen.

Betrachten wir zunächst die Dimension des Erzeugnisses zweier Untervektorräume:

Definition: Die Summe $U + V$ zweier Untervektorräume eines festen Vektorraums ist der kleinste Untervektorraum, der beide enthält.

Mit der Notation aus §1*f)* können wir dies auch als $U + V = [U \cup V]$ schreiben, denn für jede Teilmenge M eines Vektorraums hatten wir $[M]$ als kleinsten Untervektorraum definiert, der M enthält.

Lemma: Sind U und V endlichdimensional, so ist

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V) .$$

Beweis: Wir starten mit einer Basis $\left(\vec{b}_1, \ldots, \vec{b}_r\right)$ des Durchschnitts $U \cap V$ und ergänzen diese zu einer Basis $\left(\vec{b}_1, \ldots, \vec{b}_n\right)$ von U und zu einer Basis $\left(\vec{b}_1, \ldots, \vec{b}_r, \vec{c}_{r+1}, \ldots, \vec{c}_m\right)$ von V. Dann ist $U + V$ der von den Vektoren $\vec{b}_i$ und $\vec{c}_j$ erzeugte Untervektorraum, denn einen kleineren Untervektorraum, der U und V enthält, kann es nicht geben. Außerdem sind diese Vektoren linear unabhängig: Ist nämlich

$$\lambda_1 \vec{b}_1 + \cdots + \lambda_n \vec{b}_n + \mu_{r+1} \vec{c}_{r+1} + \cdots + \mu_m \vec{c}_m = \vec{0} ,$$

so können wir dies auch schreiben als

$$\vec{v} \underset{\text{def}}{=} \lambda_1 \vec{b}_1 + \cdots + \lambda_n \vec{b}_n = -\left(\mu_{r+1} \vec{c}_{r+1} + \cdots + \mu_m \vec{c}_m\right) .$$

Der Vektor $\vec{v}$ läßt sich damit sowohl als Linearkombination von Vektoren aus U darstellen wie auch als Linearkombination von Vektoren aus V; er liegt also in $U \cap V$. Da dieser Durchschnitt die Vektoren $\vec{b}_1, \ldots, \vec{b}_r$ als Basis hat, müssen daher λ_{r+1} bis λ_n verschwinden. Damit ist also

$$\lambda_1 \vec{b}_1 + \cdots + \lambda_{\mathbf{r}} \vec{b}_{\mathbf{r}} + \mu_{r+1} \vec{c}_{r+1} + \cdots + \mu_m \vec{c}_m = \vec{0} ,$$

und hier haben wir eine Darstellung des Nullvektors als Linearkombination von Basisvektoren des Vektorraums V. Also müssen auch alle restlichen λ_i sowie die sämtlichen μ_j verschwinden.

Somit ist $\left(\vec{b}_1, \ldots, \vec{b}_n, \vec{c}_{r+1}, \ldots, \vec{c}_m\right)$ eine Basis von $U + V$, d.h.

$$\dim(U + V) = n + (m - r) = \dim U + \dim V - \dim(U \cap V) .$$

■

Bei Untervektorräumen ist damit alles klar; kehren wir zurück zur den affinen Unterräumen; wir betrachten also zwei affinen Unterräume (A, U) und (B, V) eines affinen Raums (C, W) sowie den davon erzeugten affinen Unterraum (S, T).

Falls A und B nichtleeren Durchschnitt haben, können wir Koordinatensysteme mit einem gemeinsamen Nullpunkt finden und W ist der Vektorraum $U + V$, d.h. in diesem Fall ist

$$\begin{aligned} \dim S = \dim T = \dim(U + V) &= \dim U + \dim V - \dim(U \cap V) \\ &= \dim A + \dim B - \dim(A \cap B) , \end{aligned}$$

denn $A \cap B$ hat ein Koordinatensystem mit demselben Nullpunkt und daher ist der zugehörige Untervektorraum $U \cap V$.

Falls A und B aber disjunkt sind, müssen wir verschiedene Punkte O_A und O_B als Nullpunkte der entsprechenden Koordinatensysteme wählen; jetzt ist der Vektorraum T des von A und B aufgespannten affinen Raums (S, T) ganz sicher nicht mehr $U + V$, denn er muß nun auch den Verbindungsvektor von O_A und O_B enthalten. Wäre dieser als Summe $\vec{u} + \vec{v}$ von Vektoren $\vec{u} \in U$ und $\vec{v} \in V$ darstellbar, könnten wir den Vektor $\vec{u}$ an O_A abtragen und $-\vec{v}$ an O_B; beide Vektoren

hätten denselben Endpunkt, der somit im (leeren) Durchschnitt von A und B liegen müßte. Also ist T mindestens gleich $U+V$ plus dem vom Verbindungsvektor aufgespannten eindimensionalen Untervektorraum. Das reicht aber auch schon, denn nimmt man etwa O_A als Ursprung des Koordinatensystems, kommt man nun durch Abtragen der Vektoren aus diesem Untervektorraum zu jedem Punkt von A oder B. Also ist in diesem Fall

$$\begin{aligned}\dim S &= \dim T = 1 + \dim(U+V)\\ &= \dim U + \dim V - \dim(U \cap V) + 1\,.\end{aligned}$$

Insgesamt haben wir damit gezeigt

Lemma: (A,U) und (B,V) seien affine Unterräume eines affinen Raum (C,W), und (S,T) sei der von beiden erzeugte affine Unterraum. Dann ist

$$\dim S = \begin{cases} \dim A + \dim B - \dim(A \cap B) & \text{falls } A \cap B \neq \emptyset \\ \dim A + \dim B - \dim(U \cap V) + 1 & \text{falls } A \cap B = \emptyset\,. \end{cases}$$

■

Betrachten wir als Beispiel die beiden Unterräume

$$A = \left\{ (1,0,1,0,1) + \lambda \begin{pmatrix}1\\2\\3\\4\\5\end{pmatrix} + \mu \begin{pmatrix}5\\4\\3\\2\\1\end{pmatrix} \;\middle|\; \lambda,\mu \in \mathbb{R} \right\}$$

und

$$B = \left\{ (0,1,0,1,0) + \lambda \begin{pmatrix}0\\1\\2\\3\\4\end{pmatrix} + \mu \begin{pmatrix}0\\1\\4\\9\\16\end{pmatrix} \;\middle|\; \lambda,\mu \in \mathbb{R} \right\}$$

von $\mathbb{R}^5$.

Der Untervektorraum zu A ist

$$\begin{aligned} U &= \left[\begin{pmatrix}1\\2\\3\\4\\5\end{pmatrix}, \begin{pmatrix}5\\4\\3\\2\\1\end{pmatrix} \right] = \left[\begin{pmatrix}1\\2\\3\\4\\5\end{pmatrix}, \begin{pmatrix}6\\6\\6\\6\\6\end{pmatrix} \right] = \left[\begin{pmatrix}1\\2\\3\\4\\5\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} \right] \\ &= \left[\begin{pmatrix}0\\1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} \right], \end{aligned}$$

der zu B ist

$$V = \left[\begin{pmatrix}0\\1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}0\\1\\4\\9\\16\end{pmatrix} \right] = \left[\begin{pmatrix}0\\1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}0\\0\\2\\6\\12\end{pmatrix} \right].$$

Der Durchschnitt der beiden Untervektorräume ist also eindimensional und wird vom gemeinsamen Basisvektor aufgespannt.

Zur Berechnung des Durchschnitts von A und B müssen wir untersuchen, ob es reelle Zahlen $\lambda_1, \mu_1, \lambda_2, \mu_2$ gibt, so daß

$$(1,0,1,0,1) + \lambda_1 \begin{pmatrix}1\\2\\3\\4\\5\end{pmatrix} + \mu_1 \begin{pmatrix}5\\4\\3\\2\\1\end{pmatrix} = (0,1,0,1,0) + \lambda_2 \begin{pmatrix}0\\1\\2\\3\\4\end{pmatrix} + \mu_2 \begin{pmatrix}0\\1\\4\\9\\16\end{pmatrix}$$

ist. Mit den gerade berechneten neuen Basen von U und V können wir stattdessen auch das etwas einfachere Gleichungssystem

$$(1,0,1,0,1) + \lambda_3 \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} + \mu_3 \begin{pmatrix}0\\1\\2\\3\\4\end{pmatrix} = (0,1,0,1,0) + \lambda_4 \begin{pmatrix}0\\1\\2\\3\\4\end{pmatrix} + \mu_4 \begin{pmatrix}0\\0\\2\\6\\12\end{pmatrix}$$

betrachten mit neuen Variablen $\lambda_3, \mu_3, \lambda_4$ und μ_4, denn der Durchschnitt zweier affiner Unterräume hängt natürlich nicht vom Koordinatensystem

ab. Hier können wir alle Basisvektoren auf die linke Seite bringen und die beiden Basispunkte in Gestalt ihres Verbindungsvektors auf die rechte; mit den neuen Variablen $\lambda = \lambda_3$, $\mu = \mu_3 - \lambda_4$ und $\nu = -\mu_4$ wird das Gleichungssystem dann zu

$$\lambda \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 0 \\ 2 \\ 6 \\ 12 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} .$$

Bei diesem System kann das für die erste Komponente nur dann gutgehen, wenn $\lambda = -1$ ist; setzen wir diesen Wert ein und betrachten dann die zweite Komponente, folgt, daß $\mu = 2$ sein muß. Für ν bleibt noch das Gleichungssystem

$$\nu \begin{pmatrix} 0 \\ 0 \\ 2 \\ 6 \\ 12 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -4 \\ -4 \\ -8 \end{pmatrix} ,$$

das offensichtlich unlösbar ist. Damit war auch das ursprüngliche Gleichungssystem unlösbar, wir sind hier also im Fall $A \cap B = \emptyset$.

Nach obiger Formel hat das Erzeugnis von A und B daher die Dimension

$$\dim A + \dim B - \dim(U \cap V) + 1 = 2 + 2 - 1 + 1 = 4 ;$$

wie wir oben gesehen haben, enthält der zugehörige Untervektorraum außer $U + V$ auch noch den Verbindungsvektor

$$\begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

der Nullpunkte $(0, 1, 0, 1, 0)$ von A und $(1, 0, 1, 0, 1)$ von B.

Da wir $U + V$ und auch den Verbindungsvektor kennen, läßt sich das Erzeugnis von A und B nun leicht explizit angeben: Es ist der affine

Unterraum

$$\left\{ (1,0,1,0,1) + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 1 \\ 4 \\ 9 \\ 16 \end{pmatrix} + \rho \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \;\middle|\; \begin{matrix} \lambda, \mu, \nu, \\ \rho \in \mathbb{R} \end{matrix} \right\} .$$

h) Ausblick: Numerische Lösung linearer Gleichungssysteme

Mancher Leser mag sich in die Mittelstufe zurückversetzt gefühlt haben, als bei den Beispielen dieses Paragraphen plötzlich Brüche auftauchten anstelle der Dezimalzahlen, mit denen der „echte Praktiker“ seinen Taschenrechner füttert.

Tatsächlich sollte der GAUSS-Algorithmus, so wie er hier vorgestellt wurde, nur für Körper benutzt werden, in denen man *exakt* rechnen kann; beim Rechnen mit reellen Zahlen, vor allem wenn es per Computer oder Taschenrechner geschieht, muß man sich meist mit Näherungslösungen begnügen, und leider können in ungünstigen Fällen selbst kleine Rundungsfehler große Auswirkungen auf das Ergebnis haben.

Der Umgang mit Gleitkommazahlen ist Gegenstand der Numerik (und teilweise auch der Praktischen Informatik); hier sei nur anhand weniger Beispiele gezeigt, daß Probleme auftreten *können:*

Beginnen wir mit dem Gleichungssystem

$$\begin{array}{rcrcr} 1{,}35x & - & 2{,}768y & = & -10 \\ 4{,}241x & - & 8{,}69y & = & -31{,}4 + \varepsilon . \end{array}$$

Seine Lösung liegt für $\varepsilon = 0$ bei

$$x \approx 41{,}527 \quad \text{und} \quad y \approx 23{,}866 .$$

Stört man aber die rechte Seite nur im 0,01 in der zweiten Gleichung, ersetzt man dort also die 31,4 durch 31,41, wird die Lösung zu

$$x \approx 74{,}558 \quad \text{und} \quad y \approx 39{,}976 ,$$

und ersetzt man sie gar durch 31,5, erhält man

$$x \approx 371{,}838 \quad \text{und} \quad y \approx 184{,}964 .$$

Schon kleine Störungen der Konstanten, wie sie durch allfällige Rundungsfehler immer wieder auftreten, können also das Ergebnis stark verfälschen. Im vorliegenden Fall läßt sich dieser Effekt leicht quantifizieren: Für allgemeines ε hat das obige Gleichungssystem die Lösung

$$x \approx 41{,}527 + 3303{,}10\varepsilon \quad \text{und} \quad y \approx 23{,}866 + 1610{,}98\varepsilon\,,$$

jede Störung der rechten Seite der zweite Gleichung führt also zu einer mehr als 3000-fachen Störung des Werts von x.

In diesem einfachen Beispiel kann man relativ schnell sehen, was passiert; außerdem deuten die im Vergleich zur rechten Seite etwas kleinen Koeffizienten auf der linken Seite darauf hin, daß es Probleme geben könnte. Leider ist die Situation nicht immer so klar: Beim linearen Gleichungssystem

$$\begin{aligned}
\frac{x_1}{2} + \frac{x_2}{3} + \frac{x_3}{4} + \frac{x_4}{5} + \frac{x_5}{6} + \frac{x_6}{7} &= 1\\
\frac{x_1}{3} + \frac{x_2}{4} + \frac{x_3}{5} + \frac{x_4}{6} + \frac{x_5}{7} + \frac{x_6}{8} &= 1\\
\frac{x_1}{4} + \frac{x_2}{5} + \frac{x_3}{6} + \frac{x_4}{7} + \frac{x_5}{8} + \frac{x_6}{9} &= 1\\
\frac{x_1}{5} + \frac{x_2}{6} + \frac{x_3}{7} + \frac{x_4}{8} + \frac{x_5}{9} + \frac{x_6}{10} &= 1\\
\frac{x_1}{6} + \frac{x_2}{7} + \frac{x_3}{8} + \frac{x_4}{9} + \frac{x_5}{10} + \frac{x_6}{11} &= 1\\
\frac{x_1}{7} + \frac{x_2}{8} + \frac{x_3}{9} + \frac{x_4}{10} + \frac{x_5}{11} + \frac{x_6}{12} &= 1
\end{aligned}$$

sind die Koeffizienten im Vergleich zur rechten Seite durchaus in vernünftigen Größenordnungen, und die Lösung

$$\begin{aligned}
&x_1 = -42, && x_2 = 840, && x_3 = -5040,\\
&x_4 = 12600, && x_5 = -13860, && x_6 = 5544
\end{aligned}$$

ist auch nicht sonderlich exotisch.

Berechnen wir die Lösung allerdings numerisch mit nur drei geltenden Ziffern, so erhalten wir die numerische Lösung

$$\begin{aligned}
&x_1 = -19{,}0, && x_2 = 171 && x_3 = -359,\\
&x_4 = 216 && x_5 = -83{,}8, && x_6 = 98{,}2,
\end{aligned}$$

die mit der korrekten Lösung offensichtlich nicht viel zu tun hat.

Bei diesem wie bei allen folgenden Beispielen von Gleitkommarechnungen wird ein Leser, der die Ergebnisse überprüfen möchte, mit hoher Wahrscheinlichkeit andere als die angegebenen Zahlenwerte erhalten. Der Grund liegt darin, daß das Ergebnis einer Rechnung mit Gleitkommazahlen nicht nur von der Anzahl der mitgeführten Dezimalstellen abhängt, sondern auch von der Reihenfolge der Rechenschritte. Das Assoziativgesetz gilt bei Gleitkommazahlen weder für die Addition noch für die Multiplikation: Beispielsweise ist bei drei geltenden Ziffern

$$(4{,}21 + 6{,}82) - 2{,}13 = 11{,}0 - 2{,}13 = 8{,}87\,,$$

aber

$$4{,}21 + (6{,}82 - 2{,}13) = 4{,}21 + 4{,}69 = 8{,}90\,,$$

wobei hier so gerechnet wurde, daß jedes Ergebnis *einer* Rechenoperation zunächst exakt bestimmt wurde und dann auf drei geltende Ziffern gerundet wurde. Die meisten heutigen Computer runden bei Gleitkommarechnungen auf diese Weise, auch wenn sie natürlich im allgemeinen nicht wirklich das exakte Zwischenergebnis berechnen. Es gibt aber Algorithmen, mit denen sich zu einer vorgegebenen Rechengenauigkeit eine Zahl berechnen läßt, die dasselbe Rundungsergebnis hat wie das exakte Zwischenergebnis. Taschenrechnern treiben nur selten einen so hohen Aufwand; hier können noch zusätzliche Fehler entstehen.

Die Nichtassoziativität von Gleitkommaoperationen hat noch eine weitere unerwartete Konsequenz: Zumindest bei Lösungsmethoden, die bei der Wahl des nächsten Eliminationsschritts auch Zufallsentscheidungen einfließen lassen (wie dies beispielsweise bei Maple der Fall ist), kann auch die mehrfache Lösung desselben Gleichungssystems mit demselben Algorithmus auf demselben Computer zu verschiedenen Ergebnissen führen, das Ergebnis ist also nicht nur nicht richtig, sondern sogar nicht einmal konsistent falsch.

Hier sind beispielsweise für das gerade betrachtete Beispiel die Ergebnisse aus zehn Lösungsversuchen mit Maple, gerechnet mit jeweils vier geltenden Ziffern:

x_1	x_2	x_3	x_4	x_5	x_6
−33,900	136,700	−120,800	510,500	−1455,000	1001,000
−20,400	235,900	−956,200	2039,000	−2304,000	1035,000
−33,900	136,700	−120,800	510,500	−1455,000	1001,000
−34,000	136,700	−120,000	509,700	−1456,000	1002,000
−37,200	172,500	−226,300	613,100	−1461,000	977,300
−34,000	136,700	−120,000	509,700	−1456,000	1002,000
17,600	−208,200	689,700	−636,100	−306,300	472,300
19,500	−249,300	874,000	−936,800	−120,000	442,500
8,000	−143,900	587,400	−692,700	−80,040	349,400
−20,400	235,900	−956,200	2039,000	−2304,000	1035,000

Egal wie stark die Zahlenwerte, die ein Leser beim Rechnen mit drei oder vier geltenden Ziffern bekommt, von den obigen abweichen, werden sie also mit ziemlicher Sicherheit eines gemeinsam haben: Sie haben nicht einmal entfernt mit den korrekten Zahlen zu tun und sind somit völlig unbrauchbar.

Eine Erhöhung der Ziffernzahl führt nur langsam zu besseren Ergebnissen: Bei einer Genauigkeit von vier bis fünfzehn Ziffern erhalten wir nacheinander die folgenden „Lösungen“:

Ziffern	x_1	x_2	x_3
4	-21,100	236,600	-940,200
5	8,550	-29,800	-458,970
6	19,806	-159,680	2,252
7	1,683	137,057	-1502,936
8	-33,359	699,124	-4325,108
9	-41,766	835,783	-5017,165
10	-42,004	840,010	-5039,853
11	-41,998	839,974	-5039,864
12	-41,999	839,988	-5039,936
13	-42,000	839,998	-5039,992
14	-42,000	840,000	-5039,999
15	-42,000	840,000	-5040,000
korrekt	-42	840	-5040

und

Ziffern	x_4	x_5	x_6
4	613,100	-1461,000	977,300
5	-312,910	-702,410	663,730
6	1427,120	-2813,060	1530,740
7	4960,798	-6457,628	2897,409
8	11062,780	-12367,780	5009,726
9	12418,887	-13683,876	5480,857
10	12599,672	-13859,628	5543,852
11	12600,048	-13860,032	5544,007
12	12599,855	-13859,858	5543,949
13	12599,981	-13859,982	5543,994
14	12599,998	-13859,998	5543,999
15	12600,000	-13860,000	5544,000
korrekt	12600	-13860	5544

Wie man sieht, werden die Ergebnisse erst ab zehnstelliger Genauigkeit halbwegs korrekt, h.h. trotz der kleinen Koeffizienten müßte man im vorliegenden Fall mit doppeltgenauen reellen Zahlen rechnen. (Bei einer Gleitkommaarithmetik nach IEEE Standard 754, wie sie in den meisten heutigen Prozessoren realisiert ist, sind Gleitkommazahlen einfacher Genauigkeit mit einem relativen Fehlerrn von bis zu etwa $5{,}96 \cdot 10^{-8}$ behaftet, was zwischen sieben und acht geltenden Ziffern entspricht; bei doppelter Genauigkeit liegt die Schranke bei etwa $1{,}11 \cdot 10^{-16}$; das entspricht nicht ganz sechzehn geltenden Ziffern.)

Daß selbst doppelte Genauigkeit nicht immer ausreicht, zeigt die Vergrößerung des obigen Gleichungssystems auf 15 Variable:

$$
\begin{aligned}
\tfrac{x_1}{2}+\tfrac{x_2}{3}+\tfrac{x_3}{4}+\tfrac{x_4}{5}+\tfrac{x_5}{6}+\tfrac{x_6}{7}+\tfrac{x_7}{8}+\tfrac{x_8}{9}+\tfrac{x_9}{10}+\tfrac{x_{10}}{11}+\tfrac{x_{11}}{12}+\tfrac{x_{12}}{13}+\tfrac{x_{13}}{14}+\tfrac{x_{14}}{15}+\tfrac{x_{15}}{16} &= 1\\
\tfrac{x_1}{3}+\tfrac{x_2}{4}+\tfrac{x_3}{5}+\tfrac{x_4}{6}+\tfrac{x_5}{7}+\tfrac{x_6}{8}+\tfrac{x_7}{9}+\tfrac{x_8}{10}+\tfrac{x_9}{11}+\tfrac{x_{10}}{12}+\tfrac{x_{11}}{13}+\tfrac{x_{12}}{14}+\tfrac{x_{13}}{15}+\tfrac{x_{14}}{16}+\tfrac{x_{15}}{17} &= 1\\
\tfrac{x_1}{4}+\tfrac{x_2}{5}+\tfrac{x_3}{6}+\tfrac{x_4}{7}+\tfrac{x_5}{8}+\tfrac{x_6}{9}+\tfrac{x_7}{10}+\tfrac{x_8}{11}+\tfrac{x_9}{12}+\tfrac{x_{10}}{13}+\tfrac{x_{11}}{14}+\tfrac{x_{12}}{15}+\tfrac{x_{13}}{16}+\tfrac{x_{14}}{17}+\tfrac{x_{15}}{18} &= 1\\
\tfrac{x_1}{5}+\tfrac{x_2}{6}+\tfrac{x_3}{7}+\tfrac{x_4}{8}+\tfrac{x_5}{9}+\tfrac{x_6}{10}+\tfrac{x_7}{11}+\tfrac{x_8}{12}+\tfrac{x_9}{13}+\tfrac{x_{10}}{14}+\tfrac{x_{11}}{15}+\tfrac{x_{12}}{16}+\tfrac{x_{13}}{17}+\tfrac{x_{14}}{18}+\tfrac{x_{15}}{19} &= 1\\
\tfrac{x_1}{6}+\tfrac{x_2}{7}+\tfrac{x_3}{8}+\tfrac{x_4}{9}+\tfrac{x_5}{10}+\tfrac{x_6}{11}+\tfrac{x_7}{12}+\tfrac{x_8}{13}+\tfrac{x_9}{14}+\tfrac{x_{10}}{15}+\tfrac{x_{11}}{16}+\tfrac{x_{12}}{17}+\tfrac{x_{13}}{18}+\tfrac{x_{14}}{19}+\tfrac{x_{15}}{20} &= 1\\
\tfrac{x_1}{7}+\tfrac{x_2}{8}+\tfrac{x_3}{9}+\tfrac{x_4}{10}+\tfrac{x_5}{11}+\tfrac{x_6}{12}+\tfrac{x_7}{13}+\tfrac{x_8}{14}+\tfrac{x_9}{15}+\tfrac{x_{10}}{16}+\tfrac{x_{11}}{17}+\tfrac{x_{12}}{18}+\tfrac{x_{13}}{19}+\tfrac{x_{14}}{20}+\tfrac{x_{15}}{21} &= 1\\
\tfrac{x_1}{8}+\tfrac{x_2}{9}+\tfrac{x_3}{10}+\tfrac{x_4}{11}+\tfrac{x_5}{12}+\tfrac{x_6}{13}+\tfrac{x_7}{14}+\tfrac{x_8}{15}+\tfrac{x_9}{16}+\tfrac{x_{10}}{17}+\tfrac{x_{11}}{18}+\tfrac{x_{12}}{19}+\tfrac{x_{13}}{20}+\tfrac{x_{14}}{21}+\tfrac{x_{15}}{22} &= 1
\end{aligned}
$$

$$
\begin{aligned}
\tfrac{x_1}{9}+\tfrac{x_2}{10}+\tfrac{x_3}{11}+\tfrac{x_4}{12}+\tfrac{x_5}{13}+\tfrac{x_6}{14}+\tfrac{x_7}{15}+\tfrac{x_8}{16}+\tfrac{x_9}{17}+\tfrac{x_{10}}{18}+\tfrac{x_{11}}{19}+\tfrac{x_{12}}{20}+\tfrac{x_{13}}{21}+\tfrac{x_{14}}{22}+\tfrac{x_{15}}{23} &= 1\\
\tfrac{x_1}{10}+\tfrac{x_2}{11}+\tfrac{x_3}{12}+\tfrac{x_4}{13}+\tfrac{x_5}{14}+\tfrac{x_6}{15}+\tfrac{x_7}{16}+\tfrac{x_8}{17}+\tfrac{x_9}{18}+\tfrac{x_{10}}{19}+\tfrac{x_{11}}{20}+\tfrac{x_{12}}{21}+\tfrac{x_{13}}{22}+\tfrac{x_{14}}{23}+\tfrac{x_{15}}{24} &= 1\\
\tfrac{x_1}{11}+\tfrac{x_2}{12}+\tfrac{x_3}{13}+\tfrac{x_4}{14}+\tfrac{x_5}{15}+\tfrac{x_6}{16}+\tfrac{x_7}{17}+\tfrac{x_8}{18}+\tfrac{x_9}{19}+\tfrac{x_{10}}{20}+\tfrac{x_{11}}{21}+\tfrac{x_{12}}{22}+\tfrac{x_{13}}{23}+\tfrac{x_{14}}{24}+\tfrac{x_{15}}{25} &= 1\\
\tfrac{x_1}{12}+\tfrac{x_2}{13}+\tfrac{x_3}{14}+\tfrac{x_4}{15}+\tfrac{x_5}{16}+\tfrac{x_6}{17}+\tfrac{x_7}{18}+\tfrac{x_8}{19}+\tfrac{x_9}{20}+\tfrac{x_{10}}{21}+\tfrac{x_{11}}{22}+\tfrac{x_{12}}{23}+\tfrac{x_{13}}{24}+\tfrac{x_{14}}{25}+\tfrac{x_{15}}{26} &= 1\\
\tfrac{x_1}{13}+\tfrac{x_2}{14}+\tfrac{x_3}{15}+\tfrac{x_4}{16}+\tfrac{x_5}{17}+\tfrac{x_6}{18}+\tfrac{x_7}{19}+\tfrac{x_8}{20}+\tfrac{x_9}{21}+\tfrac{x_{10}}{22}+\tfrac{x_{11}}{23}+\tfrac{x_{12}}{24}+\tfrac{x_{13}}{25}+\tfrac{x_{14}}{26}+\tfrac{x_{15}}{27} &= 1\\
\tfrac{x_1}{14}+\tfrac{x_2}{15}+\tfrac{x_3}{16}+\tfrac{x_4}{17}+\tfrac{x_5}{18}+\tfrac{x_6}{19}+\tfrac{x_7}{20}+\tfrac{x_8}{21}+\tfrac{x_9}{22}+\tfrac{x_{10}}{23}+\tfrac{x_{11}}{24}+\tfrac{x_{12}}{25}+\tfrac{x_{13}}{26}+\tfrac{x_{14}}{27}+\tfrac{x_{15}}{28} &= 1\\
\tfrac{x_1}{15}+\tfrac{x_2}{16}+\tfrac{x_3}{17}+\tfrac{x_4}{18}+\tfrac{x_5}{19}+\tfrac{x_6}{20}+\tfrac{x_7}{21}+\tfrac{x_8}{22}+\tfrac{x_9}{23}+\tfrac{x_{10}}{24}+\tfrac{x_{11}}{25}+\tfrac{x_{12}}{26}+\tfrac{x_{13}}{27}+\tfrac{x_{14}}{28}+\tfrac{x_{15}}{29} &= 1\\
\tfrac{x_1}{16}+\tfrac{x_2}{17}+\tfrac{x_3}{18}+\tfrac{x_4}{19}+\tfrac{x_5}{20}+\tfrac{x_6}{21}+\tfrac{x_7}{22}+\tfrac{x_8}{23}+\tfrac{x_9}{24}+\tfrac{x_{10}}{25}+\tfrac{x_{11}}{26}+\tfrac{x_{12}}{27}+\tfrac{x_{13}}{28}+\tfrac{x_{14}}{29}+\tfrac{x_{15}}{30} &= 1
\end{aligned}
$$

Die Koeffizienten der linken und der rechten Seiten unterscheiden sich hier höchstens um den (moderaten) Faktor dreißig, und auch die Anzahl der Gleichungen ist im Vergleich zu den Systemen mit mehreren Tausend Variablen, die in vielen Anwendungen auftreten, eher gering. Daß es trotzdem numerische Probleme gibt, sieht man an der folgenden Tabelle; sie zeigt die korrekten und die mit fünfzehnstelliger Genauigkeit berechneten Werte der Variablen x_i:

x_1	240	-3173,68185135567
x_2	-28560	228151,786322960
x_3	1113840	-5303344,76558434
x_4	-21162960	58391350,3363271
x_5	232792560	-353947335,741599
x_6	-1629547920	1247733630,21701
x_7	7682154480	-2469397672,05011
x_8	-25241364720	1965727944,18240
x_9	58896517680	2105462770,85717
x_{10}	-98160862800	-6442117291,59057
x_{11}	116008292400	4902607321,68726
x_{12}	-94915875600	1334651674,35581
x_{13}	51108548400	-4588357739,33491
x_{14}	-16287339600	2863533974,79321
x_{15}	2326762800	-619210289,163625

Offensichtlich hat das berechnete Ergebnis nichts mit dem tatsächlichen zu tun, und auch hier werden verschiedene Computer oder sogar (im Falle etwa von Maple) derselbe Computer bei verschiedenen Versuchen völlig verschiedene Ergebnisse liefern.

Beim Lösen von linearen Gleichungssystemen muß man also unbedingt

darauf achten, daß Rundungsfehler auf das absolut notwendige Minimum beschränkt werden, und selbst dann muß man noch sehr vorsichtig sein. In der Numerik werden daher oft Iterationsverfahren bevorzugt, da diese ihre Rundungsfehler (zumindest tendenziell) selbst korrigieren.

Hier seien nur zwei Faustregeln für den GAUSS-Algorithmus erwähnt:

1.) Die verschiedenen Gleichungen des Systems sollten Koeffizienten in ähnlichen Größenordnungen enthalten. Da sich nichts an der Lösungsmenge ändert, wenn man eine der Gleichungen mit einer von Null verschiedenen Konstanten multipliziert, kann man dies etwa dadurch erreichen, daß man alle Zeilenvektoren der Matrix des Gleichungssystems auf dieselbe Länge (z.B. Eins) normiert.

2.) Falls man ein Vielfaches einer Gleichung zu einer anderen addiert, sollte der Koeffizient, mit dem die erste Gleichung multipliziert wird, möglichst klein sein, damit die zweite Gleichung nicht zu stark gestört wird. Dies läßt sich etwa dadurch erreichen, daß man bei der Elimination einer Variablen genau die Gleichung unverändert läßt, in der diese Variable mit dem betragsgrößten Koeffizienten vorkommt.

Für alles weitere sei auf die *Numerik* verwiesen; wer nicht so lange warten möchte, findet sehr viel mehr zu diesem Thema, als in jede Numerikvorlesung paßt, in der einschlägigen Spezialliteratur, z.B. im klassischen Buch

J.H. WILKINSON: Rounding errors in algebraic processes, *Prentice Hall,* 1963 oder *Dover,* 1994

oder in neueren Büchern wie

MICHAEL L. OVERTON: Numerical Computing with IEEE Floating Point Arithmetic – *Including One Theorem, One Rule of Thumb and One Hundred and One Exercises,* SIAM, 2001,

wo die Probleme des numerischen Rechnens in sehr elementarer und praxisnaher Darstellung behandelt werdet; mehr theoretischen Hintergrund findet man bei

FRANÇOISE CHAITIN-CHATELIN, VALÉRIE FRAYSSÉ: Lectures on finite precision computations, SIAM, 1996

und in dem sehr ausführlichen Buch

NICHOLAS J. HIGHAM: Accuracy and stability of numerical algorithms, SIAM, 1996.

i) Matrixgleichungen und die Berechnung der Inversen

Wir haben inzwischen recht gut verstanden, wann eine Matrix invertierbar ist und können dies auch leicht rechnerisch überprüfen; wir kennen aber noch keine Methode, mit dem wir die inverse Matrix wirklich berechnen könnten. Wie wir gleich sehen werden, stellt und der GAUSS-Algorithmus eine entsprechendes Verfahren zur Verfügung, das uns nicht nur die Bestimmung der Inversen erlaubt, sondern allgemeiner die Berechnung der Lösungsmenge einer beliebigen (linearen) Matrixgleichung.

Wir gehen aus von einer $n \times m$-Matrix $A = (a_{ij}) \in k^{n\times m}$ sowie einer $n \times p$-Matrix $B = (b_{i\ell}) \in k^{n\times p}$; gesucht ist eine Lösung der Matrixgleichung $AX = B$, wobei $X = (x_{j\ell})$ dann offensichtlich eine $m \times p$-Matrix sein muß. Besonders interessant ist der Fall $n = m = p$ und $B = E$, denn in diesem Fall ist die $X = A^{-1}$ die inverse Matrix.

Sind $$\vec{x}^{\,(\ell)} = \begin{pmatrix} x_{1\ell} \\ \vdots \\ x_{m\ell} \end{pmatrix} \quad \text{und} \quad \vec{b}^{\,(\ell)} = \begin{pmatrix} b_{1\ell} \\ \vdots \\ b_{n\ell} \end{pmatrix}$$

die Spaltenvektoren von X und B, so ist die Matrixgleichung $AX = B$ äquivalent zu den p linearen Gleichungssystemen

$$A\vec{x}^{\,(\ell)} = \vec{b}^{\,(\ell)} \quad \text{für} \quad \ell = 1, \ldots, p\,,$$

die allesamt dieselbe Matrix A haben. Falls diese Gleichungen alle lösbar sind, gibt es also eine Matrix X mit $AX = B$.

Bei der Lösung dieser Gleichungssysteme wird es im allgemeinen vorteilhaft sein, auch bei der Rücksubstitution alle rechten Seiten simultan zu behandeln, d.h. anstelle des Einsetzens werden links durch Zeilenoperationen so lange Koeffizienten eliminiert, bis in jeder Zeile möglichst nur noch ein einziger von Null verschiedener Eintrag steht. (Dies ist natürlich nur dann erreichbar, wenn das Gleichungssystem höchstens

eine Lösung hat.) Durch Division läßt sich der verbliebene Eintrag noch auf Eins normieren, so daß sich rechts leicht die Lösung ablesen läßt.

Speziell im Falle quadratisher Matrizen kann man bei eindeutig lösbaren Gleichungen durch Zeilenvertauschungen sogar erreichen, daß links die Einheitsmatrix steht; da alle angewandten Operationen die Matrixgleichung „links $\times$ X = rechts“ erhalten, steht dann rechts die Lösungsmatrix X.

Betrachten wir als Beispiel die Inversion der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ 8 & 9 & 12 \end{pmatrix}.$$

Wir müssen das lineare Gleichungssystem $A\vec{x} = \vec{b}$ lösen für die drei rechten Seiten

$$\vec{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dazu wenden wir den GAUSS-Algorithmus simultan auf alle drei Gleichungssysteme an, indem wir die drei rechten Seiten nebeneinander schreiben, d.h. wir gehen aus von einer um alle drei rechten Seiten erweiterten Matrix

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 7 & 0 & 1 & 0 \\ 8 & 9 & 12 & 0 & 0 & 1 \end{pmatrix}$$

und wenden darauf die üblichen Eliminationsschritte an:

Als erstes müssen wir in der ersten Spalte alle Koeffizienten bis auf einen zu Null machen; dazu subtrahieren wir das vier- *bzw.* achtfache der ersten Zeile von der zweiten *bzw.* dritten und erhalten

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -5 & -4 & 1 & 0 \\ 0 & -7 & -12 & -8 & 0 & 1 \end{pmatrix}.$$

Als nächstes sollte der zweite Eintrag der letzten Zeile eliminiert werden; dazu können wir $7/3$ der mittleren Zeile von der letzten subtrahieren

oder aber, um Nenner zu vermeiden, sieben mal die zweite Gleichung von dreimal der dritten subtrahieren; dies führt auf

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -5 & -4 & 1 & 0 \\ 0 & 0 & -1 & 4 & -7 & 3 \end{pmatrix}.$$

In dieser Gestalt lassen sich die drei Gleichungssysteme nun leicht durch Rücksubstitution lösen; wir wollen aber stattdessen lieber weiter mit Zeilenoperationen arbeiten. (Kurzes Nachdenken zeigt, daß dies tatsächlich nur eine andere Betrachtungsweise für genau dieselben Rechenoperationen ist.)

Subtrahieren wir fünfmal die dritte Zeile von der zweiten und addieren sie dreimal zur ersten, erhalten wir

$$\begin{pmatrix} 1 & 2 & 0 & 13 & -21 & 9 \\ 0 & -3 & 0 & -24 & 36 & -15 \\ 0 & 0 & -1 & 4 & -7 & 3 \end{pmatrix}.$$

Diese Matrix wird deutlich übersichtlicher, wenn wir die zweite Zeile durch -3 dividieren und die dritte durch -1 und schließlich noch zweimal die neue zweite Zeile von der ersten subtrahieren; sie bekommt dann die Form

$$\begin{pmatrix} 1 & 0 & 0 & -3 & 3 & -1 \\ 0 & 1 & 0 & 8 & -12 & 5 \\ 0 & 0 & 1 & -4 & 7 & -3 \end{pmatrix}.$$

Damit haben wir linke eine Einheitsmatrix und können die Lösungsvektoren sofort ablesen: Es sind gerade die rechten Seiten, d.h. die rechte Matrix

$$X = \begin{pmatrix} -3 & 3 & -1 \\ 8 & -12 & 5 \\ -4 & 7 & -3 \end{pmatrix}$$

ist die Lösung von $AX = E$, wovon man sich zur Vorsicht vielleicht noch einmal durch Nachrechnen überzeugen sollte.

Daß so etwas auch beim Rechnen mit Computern kein Schaden sein kann, zeigt das folgende Beispiel einer Matrixinversion in MATLAB (*Mat*rix *Lab*oratory), einen häufig verwendeten System für numerische Berechnungen mit Matrizen:


```
seiler@hilbert:~ $ matlab

                        < M A T L A B >
           Copyright 1984-2000 The MathWorks, Inc.
                  Version 6.0.0.88 Release 12
                            Sep 21 2000


To get started, type one of these: helpwin, helpdesk, or
demo.
For Product information, visit www.mathworks.com.

>> A = [1,2,3;4,5,6;7,8,9]

A =
        1     2     3
        4     5     6
        7     8     9

>> B = inv(A)

Warning: Matrix is close to singular or badly scaled.
         Results may be inaccurate. RCOND = 1.541976e-18.
B =
     1.0e+16 *
      -0.4504    0.9007   -0.4504
       0.9007   -1.8014    0.9007
      -0.4504    0.9007   -0.4504
>> A*B
ans =
        2     0     0
        8     0    -4
       16     0     8
>> B*A
ans =
        0    -4     0
        0     8     0
        4     0     0
```

Wie man auch hier wieder einmal sieht, sind Computer zwar sehr nützliche Werkzeuge, aber eben nur für Anwender, die mitdenken und wissen, was sie tun. Computer können uns das Rechnen abnehmen, nicht aber das Denken. Wenn es um reelle Zahlen geht, kommt hinzu, daß Computer mit dieser überabzählbaren Menge genauso wenig umgehen können wie wir. Der beim numerischen Rechnen übliche Ausweg besteht darin, statt im Körper der reellen Zahlen in der Menge der Gleitkommazahlen zu rechnen. Dies ist aber grob fahrlässig, sofern man nicht sich nicht *vorher* genau überlegt, welche Genauigkeit man vom Ergebnis erwarten kann – was im allgemeinen durchaus nichttriviale Mathematik erfordert. Theoretisch gibt es einen Ausweg: Mit der sogenannten Intervallarithmetik lassen sich Berechnungen so durchführen, daß das Ergebnis ein Intervall ist, in dem das theoretisch richtige Resultat *mit Sicherheit* liegt. Leider werden dessen Schranken aber sehr schnell sehr pessimistisch, so daß das Intervall nach umfangreichen Rechnungen oft kaum noch praktisch verwertbare Information liefert.

j) Spezielle Matrizen

Matrizen mit vielen Einträgen werden schnell unhandlich, insbesondere wenn viele der Einträge von Null verschieden sind. In diesem Abschnitt wollen wir einige Matrizen spezieller Form betrachten und uns überlegen, ob und gegebenenfalls wie wir deren Gestalt beim Rechnen ausnutzen können.

1) Diagonalmatrizen: Nach der Nullmatrix und der Einheitsmatrix am einfachsten sind die Diagonalmatrizen:

Definition: Eine quadratische Matrix $D = (d_{ij}) \in k^{n\times n}$ heißt Diagonalmatrix, wenn sämtliche Einträge außerhalb der Diagonale verschwinden, d.h. $d_{ij} = 0$ für $i \neq j$.

Eine $n \times n$-Diagonalmatrix ist somit gegeben durch ihre n Diagonaleinträge d_{ii}, anstelle von n^2 Werten müssen also nur n gespeichert werden. Addition und Multiplikation von Diagonalmatrizen sind denkbar einfach: Wir müssen nur die einander entsprechenden Einträge addieren *bzw.* multiplizieren.

Auch mit Inversen gibt es keine Probleme: Offensichtlich ist der Rang einer Diagonalmatrix gleich der Anzahl nichtverschwindender Diagonaleinträge, die Matrix ist also genau dann invertierbar, wenn keiner der Diagonaleinträge verschwindet. Die inverse Matrix ist dann einfach die Diagonalmatrix mit den Inversen der Diagonaleinträge der gegebenen Matrix.

Diagonalmatrizen sind nicht nur einfach, wenn man sie untereinander verknüpft, auch die Multiplikation einer Diagonalmatrix mit einer beliebigen anderen Matrix ist problemlos:

Sei etwa $A \in k^{n\times m}$ eine beliebige Matrix und $D \in k^{m\times m}$ eine Diagonalmatrix mit Diagonaleinträgen $d_1, \ldots, d_m$. Wir wollen das Produkt AD berechnen. Seine i-te Spalte ist das Produkt von AD mit dem i-ten Einheitsvektor $\vec{e}_i$ von k^m. Das Produkt $D\vec{e}_i$ ist entsprechend der i-te Spaltenvektor von D, also $d_i\vec{e}_i$, und $A\vec{e}_i$ ist der i-te Spaltenvektor $\vec{a}_i$ von A. Also ist

$$(AD)\vec{e}_i = A(D\vec{e}_i) = A(d_i\vec{e}_i) = d_i(A\vec{e}_i) = d_i\vec{a}_i\,,$$

die Matrix AD entsteht somit aus A einfach daraus, daß der i-te Spaltenvektor von A mit dem i-ten Diagonaleintrag d_i von D multipliziert wird für alle i.

Ist D stattdessen eine $n \times n$-Matrix, ist das Produkt DA definiert und wir können es auf dieselbe Weise berechnen: Zunächst ist $(DA)\vec{e}_i = D(A\vec{e}_i) = D\vec{a}_i$. Multiplikation einer Diagonalmatrix mit einem Vektor multipliziert die j-te Komponente dieses Vektors mit dem j-ten Diagonaleintrag von D, also wird der j-te Eintrag von $\vec{a}_i$ mit d_j multipliziert. Dies passiert für jeden Spaltenvektor $\vec{a}_i$, also entsteht die Matrix DA aus A, indem man für jedes j ihre j-te Zeile mit d_j multipliziert.

Als *Regel* können wir damit festhalten: Multipliziert man eine $n \times m$-Matrix A von *links* mit einer $n \times n$-Diagonalmatrix D mit Einträgen $d_1, \ldots, d_n$, so wird die i-te *Zeile* von A mit d_i multipliziert. Multipliziert man A von *rechts* mit einer $m \times m$-Diagonalmatrix D mit Einträgen $d_1, \ldots, d_m$, so wird die i-te *Spalte* von A mit d_i multipliziert.

2) Dreiecksmatrizen: Nachdem wir das Rechnen mit Diagonalmatrizen gut im Griff haben, können wir zu etwas komplizierteren Matrizen

übergehen, z.B. den Dreiecksmatrizen:

Definition: Eine Matrix $A = (a_{ij}) \in k^{n\times n}$ heißt *untere Dreiecksmatrix,* falls $a_{ij} = 0$ für $i < j$; sie heißt *obere Dreiecksmatrix,* falls $a_{ij} = 0$ wann immer $i > j$.

Eine 2×2-Matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ ist also genau dann eine obere Dreiecksmatrix, wenn $a_{21} = 0$ ist, und genau dann eine untere Dreiecksmatrix, wenn $a_{12} = 0$ ist, d.h.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} \quad \textit{bzw.} \quad A = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix} .$$

Zählen wir zunächst, wie viele Einträge einer $n \times n$-Dreiecksmatrix nach Definition verschwinden müssen: Für eine untere Dreiecksmatrix ist das in der ersten Spalte keiner, in der zweiten einer, *usw.;* in der n-ten Spalte schließlich sind es alle bis auf einen, also $n - 1$. Insgesamt sind also

$$0 + 1 + \cdots + (n-1) = \frac{(n-1)n}{2}$$

Einträge notwendigerweise gleich Null, so daß nur

$$n^2 - \frac{(n-1)n}{2} = \frac{n(n+1)}{2}$$

Einträge gespeichert werden müssen. Für eine obere Dreiecksmatrix kommen wir auf dieselben Zahlen, wir müssen nur die Spalten in umgekehrter Reihenfolge betrachten.

Es ist klar, daß die Summen von oberen *bzw.* unteren Dreiecksmatrizen wieder obere *bzw.* untere Dreiecksmatrizen sind; für Produkte und – so sie existieren – inverse Matrizen müssen wir uns allerdings etwas mehr Gedanken machen.

Wie bei Diagonalmatrizen sieht man sofort, daß eine Dreiecksmatrix genau dann invertierbar ist, wenn keiner ihrer Diagonaleinträge verschwindet: Falls etwa der i-te Diagonaleintrag einer oberen Dreiecksmatrix verschwindet, liegen die ersten i Spaltenvektoren im von den ersten $i-1$ Koordinateneinheitsvektoren erzeugten $(i-1)$-dimensionalen

Untervektorraum von k^n und können somit unmöglich zusammen mit den restlichen $n - i$ Spaltenvektoren einen n-dimensionalen Vektorraum aufspannen. Bei unteren Dreiecksmatrizen argumentiert man im wesentlichen genauso mit dem vom i-ten bis zum n-ten Spaltenvektor aufgespannten Untervektorraum.

Am einfachsten geht das, wenn wir wieder die lineare Abbildung

$$\varphi\colon k^n \to k^n; \quad \vec{v} \mapsto A\vec{v}$$

betrachten. Da in den Spalten der Abbildungsmatrix die Bilder der Koordinateneinheitsvektoren $\vec{e}_i$ stehen, muß im Fall einer oberen Dreiecksmatrix A der erste dieser Vektoren auf ein Vielfaches von sich selbst abgebildet werden, der zweite auf eine Linearkombination von $\vec{e}_1$ und $\vec{e}_2$ *usw.;* allgemein ist $A\vec{e}_i$ eine Linearkombination der Vektoren $\vec{e}_1, \ldots, \vec{e}_i$. Die lineare Abbildung φ bildet also jeden der Untervektorräume $U_i = [\vec{e}_1, \ldots, \vec{e}_i]$ auf sich selbst ab, und diese Bedingung reicht auch aus um sicherzustellen, daß die Abbildungsmatrix von φ bezüglich der Basis $(\vec{e}_1, \ldots, \vec{e}_n)$ eine obere Dreiecksmatrix ist.

Genau dann, wenn die Matrix invertierbar ist, haben wir eine bijektive Abbildung φ; diese ist auch bijektiv, wenn man sie auf die Untervektorräume $U_i = [\vec{e}_1, \ldots, \vec{e}_i]$ einschränkt, denn natürlich bleibt sie injektiv, und wie wir aus §2*i)* wissen, ist jede injektive Abbildung eines endlichdimensionalen Vektorraums auf sich selbst auch surjektiv.

Im Falle der unteren Dreiecksmatrizen haben wir im wesentlichen dasselbe Ergebnis, nur daß wir jetzt die Basisvektoren von hinten her betrachten müssen: A ist also genau dann eine untere Dreiecksmatrix, wenn φ jeden der Untervektorräume $V_i = [\vec{e}_i, \ldots, \vec{e}_n]$ auf sich selbst abbildet.

Als Charakterisierung einer Dreiecksmatrix ist das Kriterium, das wir gerade hergeleitet haben, sicherlich nicht sehr nuützlich: Die direkte Definition ist sehr viel einfacher nachzuprüfen. Dafür bekommen wir aber mit diesem Kriterium fast gratis das folgende

Lemma: *a)* Das Produkt zweier $\left\{\begin{matrix}\text{unterer}\\ \text{oberer}\end{matrix}\right\}$ Dreiecksmatrizen ist wieder eine $\left\{\begin{matrix}\text{untere}\\ \text{obere}\end{matrix}\right\}$ Dreiecksmatrix; deren Diagonaleinträge sind die Produkte der Diagonaleinträge der beiden Faktoren.

b) Eine $\left\{\begin{matrix}\text{untere}\\ \text{obere}\end{matrix}\right\}$ Dreiecksmatrix ist genau dann invertierbar, wenn keines ihrer Diagonalelemente verschwindet. Alsdann ist auch ihre inverse Matrix wieder eine $\left\{\begin{matrix}\text{untere}\\ \text{obere}\end{matrix}\right\}$ Dreiecksmatrix.

Beweis: a) A und B seien die beiden Matrizen, und

$$\varphi\colon k^n \to k^n;\quad \vec{v} \mapsto A\vec{v} \qquad \text{und} \qquad \psi\colon k^n \to k^n;\quad \vec{v} \mapsto A\vec{v}$$

seien die zugehörigen linearen Abbildungen.

Wie wir gerade gesehen haben, lassen sich die Eigenschaften „obere Dreiecksmatrix“ *bzw.* „untere Dreiecksmatrix“ dadurch charakterisieren, daß gewisse Untervektorräume U_i *bzw.* V_i von k^n auf sich sebst abgebildet werden. Ist aber für zwei Abbildungen φ und ψ sowohl $\varphi(U) \subseteq U$ als auch $\psi(U) \subseteq U$, so ist auch $(\varphi \circ \psi)(U) \subseteq U$, d.h. auch die Abbildungsmatrix AB von $\varphi \circ \psi$ bildet U auf sich selbst ab. Somit ist auch AB eine $\left\{\begin{matrix}\text{untere}\\ \text{obere}\end{matrix}\right\}$ Dreiecksmatrix.

Den i-ten Diagonaleintrag erhalten wir als $\vec{e}_i$-Komponente des Bilds von $\vec{e}_i$. Im Falle zweier oberer Dreiecksmatrix A, B mit i-tem Diagonaleintrag a_i *bzw.* b_i ist $B\vec{e}_i = b_i\vec{e}_i + \vec{u}_{i-1}$ mit $\vec{u}_{i-1} \in U_{i-1}$; da auch A sowohl U_i als auch U_{i-1} auf sich selbst abbildet, ist also auch $(AB)\vec{e}_i = a_ib_i + \vec{w}_{i-1}$ mit $\vec{w}_{i-1} \in U_{i-1}$, d.h. die Diagonaleinträge werden miteinander multipliziert. Ganz entsprechend argumentiert man für untere Dreiecksmatrizen: Hier ist $A\vec{e}_i = a_i\vec{e}_i + \vec{v}_{i+1}$ mit $\vec{v}_{i+1} \in V_{i+1}$.

b) Eine Dreiecksmatrix A ist, wie wir oben gesehen haben, genau dann invertierbar, wenn keiner ihrer Diagonaleinträge verschwindet; alsdann ist auch jede der Abbildungen $U \to U;\ \vec{v} \mapsto A\vec{v}$, die eine $\left\{\begin{matrix}\text{untere}\\ \text{obere}\end{matrix}\right\}$ Dreiecksmatrix charakterisieren, bijektiv, also invertierbar. Somit bildet auch die lineare Abbildung $\psi\colon U \to k^n;\ \vec{v} \mapsto A^{-1}\vec{v}$ jeden der Untervektorräume U auf sich selbst ab, d.h, auch A^{-1} ist eine $\left\{\begin{matrix}\text{untere}\\ \text{obere}\end{matrix}\right\}$ Dreiecksmatrix. ∎

Um nur Teil *a)* dieses Lemmas zu beweisen, wäre der Umweg über lineare Abbildungen nicht notwendig gewesen; das hätten wir direkt auch

billiger haben können. Beispielsweise ist für zwei untere Dreiecksmatrizen $A = (a_{i\ell})$ und $B = (b_{\ell j})$ mit Produkt $C = (c_{ij})$ nach Definition der Matrixmultiplikation

$$c_{ij} = \sum_{\ell=1}^{n} a_{i\ell}b_{\ell j}\,.$$

Für $i > j$ ist für jedes $\ell < i$ der Eintrag $a_{i\ell} = 0$, da A eine untere Dreiecksmatrix ist, und für $\ell \geq i > j$ ist $b_{\ell j} = 0$, da B eine untere Dreiecksmatrix ist. Also ist auch c_{ij} als Summe von lauter Nullen gleich Null. Für die Diagonalelemente erhalten wir

$$c_{ii} = \sum_{\ell=1}^{n} a_{i\ell}b_{\ell i}\,,$$

wobei für $\ell < i$ der Eintrag $a_{i\ell}$ verschwindet und für $\ell > i$ der Eintrag $b_{\ell i}$. Somit bleibt nur der Summand $a_{ii}b_{ii}$ übrig.

Ähnlich könnte man auch die entsprechende Aussage für obere Dreiecksmatrizen beweisen.

Für inverse Matrizen kennen wir jedoch keine brauchbaren Formeln; hier ist der Umweg über die linearen Abbildungen der einzige Beweis, der mit unseren Kenntnissen möglich ist, und selbst wenn man die (ziemlich unangenehmen) Formeln kennt, nach denen man die Einträge von A^{-1} durch die von A ausdrücken kann, ist der obige Beweis kürzer und verständlicher.

3) Matrizen mit nur einem Eintrag: Bei der numerischen Behandlung partieller Differentialgleichungen oder auch in der Kontrolltheorie hat man es oft mit riesigen Matrizen zu tun, in denen dann aber nur wenige, unregelmäßig verteilte Einträge von Null verschieden sind. Man spricht hier von *spärlich besetzten Matrizen,* im Englischen *sparse matrices.* Natürlich speichert man eine solche Matrix nicht als ein Feld von $n \times m$ Einträgen, sondern als eine Liste von Tripeln (i, j, a_{ij}) zu solchen Indizes, für die $a_{ij} \neq 0$ ist. Ein eigenes Teilgebiet der numerischen Mathematik beschäftigt sich mit der Suche nach effizienten Algorithmen für solche Matrizen.

Hier wollen wir uns beschränken auf den Umgang mit den nach der Nullmatrix am spärlichsten besetzten Matrizen, also denen mit nur einem nichtverschwindenden Eintrag. Der Einfachheit halber setzen wir diesen Eintrag auf Eins; um einen anderen Wert a zu erhalten, müssen wir einfach die gesamte Matrix mit a multiplizieren.

Definition: Für eine vorgegebene Größe $n \times m$ bezeichnen wir mit E_{ij} jene $n \times m$-Matrix, die an der Stelle (i, j) den Eintrag Eins hat und sonst lauter Nullen:

$$E_{ij} = \begin{array}{c} \begin{array}{ccccccc} & & & j & & & \\ & & & \downarrow & & & \end{array} \\ \begin{pmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{array}$$

Die Abbildung $\varphi\colon k^m \to k^n;\ \ \vec{v} \mapsto E_{ij}\vec{v}$ bildet offensichtlich alle Koordinateneinheitsvektoren des k^m auf den Nullvektor ab mit Ausnahme des j-ten; dieser wird auf den i-ten Koordinateneinheitsvektor des k^n abgebildet. Für einen beliebigen Vektor $\vec{v} \in k^m$ ist somit $\varphi(\vec{v}) = v_j\vec{e}_i$ jener Vektor aus k^n, der an der i-ten Stelle den j-ten Eintrag von $\vec{v}$ stehen hat und sonst überall Nullen.

Für eine $n \times p$-Matrix A sei $\psi\colon k^p \to k^n$ die lineare Abbildung zu A; dann bildet ψ den ℓ-ten Koordinateneinheitsvektor von k^p ab auf den ℓ-ten Spaltenvektor von A, und dieser wird durch φ abgebildet auf $a_{j\ell}\vec{e}_i$. Die Abbildungsmatrix $E_{ij}A$ von $\varphi \circ \psi$ hat also als i-te Zeile die j-te Zeile von A; alle anderen Zeilen sind Null.

Für eine $q \times m$-Matrix B sei entsprechend $\omega\colon k^m \to k^q$ die lineare Abbildung; dann bildet $\omega \circ \varphi$ alle Koordinateneinheitsvektoren von k^n mit Ausnahme des j-ten auf den Nullvektor ab, da bereits φ diese Eigenschaft hat. Der j-te Koordinateneinheitsvektor wird von φ abgebildet auf den i-ten Koordinateneinheitsvektor von k^m, der wiederum von ω abgebildet wird auf den i-ten Spaltenvektor von B. Die Abbildungsmatrix BE_{ij} von $\omega \circ \varphi$ hat also als j-te Spalte die i-te Spalte von A; alle anderen Spalten sind Null.

Kurz können wir diese beiden Resultate zusammenfassen zu folgender Regel:

- Multiplikation von links mit E_{ij} führt zu einer Matrix, deren i-te Zeile die j-te Zeile der Ausgangsmatrix ist; alle anderen Zeilen sind Null.
- Multiplikation von rechts mit E_{ij} führt zu einer Matrix, deren i-te Spalte die j-te Spalte der Ausgangsmatrix ist; alle anderen Zeilen sind Null.

4) Permutationen und Permutationsmatrizen: Zu den einfachsten linearen Abbildungen gehören jene, die einfach die Reihenfolge der Basisvektoren ändern; um ihre Abbildungsmatrizen soll es hier gehen.

Definition: *a)* Eine Permutation ist eine bijektive Abbildung

$$\pi\colon \{1, \dots, n\} \to \{1, \dots, n\}\,.$$

Sie wird auch kurz als $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$ geschrieben.

b) Eine Transposition ist eine Permutation τ, die zwei Elemente von $\{1, \dots, n\}$ vertauscht und den Rest festläßt. Sind i, j die beiden Elemente, die von τ vertauscht werden, so schreiben wir kurz $\tau = (i\ j)$.

c) Die Menge aller Permutationen $\pi\colon \{1, \dots, n\} \to \{1, \dots, n\}$ wird als *symmetrische Gruppe* $\mathfrak{S}_n$ bezeichnet.

Die Matrixdarstellung einer Permutation hat natürlich nichts mit Abbildungsmatrizen von linearen Abbildungen zu tun; hier verwenden wir die Matrix nur als Darstellung der Wertetabelle von π.

Das Wort *Gruppe* tauchte bereits in §1 auf im Zusammenhang mit der Definition eines Vektorraum; es wird langsam Zeit, es wirklich zu definieren. Intuitiv versteht man unter einer Gruppe eine Struktur, deren Elemente sich so miteinander verknüpfen lassen, daß die „üblichen“ Rechenregeln gelten – mit Ausnahme, eventuell, des Kommutativgesetzes, denn dieses gilt ja beispielsweise schon bei der Matrixmultiplikation nicht mehr. Die exaktes Definition ist die folgende:

Definition: Eine *Gruppe* $(G, \circ)$ ist eine Menge G zusammen mit einer inneren Verknüpfung $\circ\colon G \times G \to G$, für die gilt:

1. $g \circ (h \circ k) = (g \circ h) \circ k$ für alle $g, h, k \in G$
2. Es gibt ein *Neutralelement* $e \in G$, so daß für alle $g \in G$ gilt $e \circ g = g \circ e = g$.
3. Zu jedem Element $g \in G$ gibt es ein *inverses Element* $g' \in g$, so daß $g \circ g' = g' \circ g = e$ ist.

Die Gruppe heißt *kommutativ* oder *abelsch,* wenn zusätzlich gilt

4. $g \circ h = h \circ g$ für alle $g, h \in G$.

Der norwegische Mathematiker NILS HENRIK ABEL (1802–1829) ist trotz seinen frühen Todes (an Tuberkulose) Initiator vieler Entwicklungen der Mathematik des neunzehnten Jahrhunderts; Begriffe wie abelsche Gruppen, abelsche Integrale, abelsche Funktionen, abelsche Varietäten, die auch in der heutigen Mathematik noch allgegenwärtig sind, verdeutlichen seinen Einfluß. Zu seinem 200. Geburtstag stiftete die norwegische Regierung zu seinen Ehren einen ABEL-Preis für Mathematik, der in ähnlicher Ausstattung und ähnlicher Weise wie die Nobelpreise verliehen wird. Erster Preisträger war 2003 JEAN-PIERRE SERRE (* 1926).

Im Sinne dieser Definition bilden beispielsweise die ganzen Zahlen mit der Addition als Verknüpfung eine Gruppe, sogar eine abelsche Gruppe, nicht aber die natürlichen Zahlen, denn es gibt beispielsweise keine natürliche Zahl n, so daß $n + 2 = 0$ ist, und auch die Null liegt zumindest nach der in dieser Vorlesung zugrundegelegten Konvention nicht in $\mathbb{N}$. Entsprechend bilden die ganzen Zahlen bezüglich der *Multiplikation* keine Gruppe, da die Gleichung $5x = 1$ in $\mathbb{Z}$ nicht lösbar ist, aber die positiven rationalen Zahlen bilden eine multiplikative Gruppe. Die rationalen Zahlen bilden keine, da $0x = 1$ unlösbar ist, aber die rationalen Zahlen ohne Null sind eine multiplikative Gruppe, genauso auch die *invertierbaren* $n \times n$-Matrizen.

Da wir die Menge aller Permutationen als *symmetrische Gruppe* bezeichnen, sollten auch diese eine Gruppe bilden.

Die Gruppenoperation ist natürlich die Hintereinanderausführung von Abbildungen, das Produkt zweier Permutationen π und ω ist also die Abbildung $\pi \circ \omega$, die eine Zahl i abbildet auf $(\pi \circ \omega)(i) = \pi(\omega(i))$.

Die Assoziativität der Verknüpfung ist, wie stets bei der Hintereinanderausführung von Abbildungen, klar; Einselement ist die identische Permutation, und Inverse gibt es, da eine Permutation nach Definition bijektiv ist und somit eine Umkehrabbildung hat. Speziell für Transpositionen ist die Bestimmung dieser Umkehrabbildung besonders einfach: Da eine Transposition nicht anderes tut, als zwei Elemente miteinander zu vertauschen, macht sie sich selbst rückgängig und ist somit ihr eigenes Inverses.

Wie meist bei der Hintereinanderausführung von Abbildungen ist auch bei Permutationen das Kommutativgesetz nicht erfüllt. Als Beispiel betrachten wir die beiden Transpositionen (1 2) und (1; 3).

Beim Produkt (1 2) ∘ (1 3) wird zuerst die Transposition (1 3) ausgeführt, die die Zahlen 1 und 3 vertauscht; sodann vertauscht (1 2) die Zahlen 1 und 2:

$$\begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 2 & 1 \\ \downarrow & \downarrow & \downarrow \\ 3 & 1 & 2 \end{array}$$

Somit ist

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

Beim Produkt (1 3) ∘ (1 2) dagegen wird zuerst die Transposition (1 2) ausgeführt und dann erst (1 3); hier ist der Gang der Vertauschungen also

$$\begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 1 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1 \end{array},$$

d.h.

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

Aus der Darstellung

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$$

einer Permutation läßt sich leicht die Anzahl möglicher Permutationen von n Elementen ablesen: Füllen wir die untere Zeile von links aus systematisch auf, so gibt es für $\pi(1)$ noch die volle Auswahl unter allen n Elementen, für $\pi(2)$ kommen alle Elemente in Frage *außer* $\pi(1)$, also nur noch $n - 1$ Stück, für $\pi(3)$ gibt es entsprechend nur noch $n - 2$, bis es schließlich bei $\pi(n)$ überhaupt nichts mehr zu wählen gibt, denn

$\pi(n)$ ist die einzige Zahl, die bis dahin noch nicht in der zweiten Zeile der Matrix steht. Insgesamt gibt es also

$$n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1 = n!$$

Permutationen von $\{1, \ldots, n\}$, die symmetrische Gruppe $\mathfrak{S}_n$ enthält somit $n!$ Elemente.

Uns interessieren derzeit keine Abbildungen der Menge $\{1, \ldots, n\}$ auf sich selbst, sondern lineare Abbildungen zwischen Vektorräumen; die bekommen wir, indem wir Permutationen auf die Basisvektoren anwenden. Wir betrachten also zu einer Permutation $\pi \in \mathfrak{S}_n$ die lineare Abbildung $\varphi_\pi \colon k^n \to k^n$, die dem i-ten Standardbasisvektor $\vec{e}_i$ von k^n den Vektor $\vec{e}_{\pi(i)}$ zuordnet; ihre Abbildungsmatrix (bezüglich der Standardbasis von k^n) bezeichnen wir mit P_π. Da φ_π den i-ten Koordinateneinheitsvektor von k^n auf den $\pi(i)$-ten abbildet, hat P_π in jeder Zeile und jeder Spalte genau einen nichtverschwindenden Eintrag: In der i-ten Spalte steht an der Position $\pi(i)$ eine Eins, überall sonst stehen Nullen. Der Eintrag $p_{k\ell}$ ist also genau dann gleich Eins, wenn $k = \pi(l)$ ist, ansonsten ist er Null.

Die Hintereinanderausführung der Abbildungen φ_π und damit die Multiplikation der Matrizen P_π ist vollständig bestimmt durch die Hintereinanderausführung der zugehörigen Permutationen: Wegen

$$(\varphi_\pi \circ \varphi_\omega)(\vec{e}_i) = \varphi_\pi\big(\varphi_\omega(\vec{e}_i)\big) = \varphi_\pi\big(\vec{e}_{\omega(i)}\big) = \vec{e}_{\pi(\omega(i))} = \vec{e}_{(\pi\circ\omega)(i)}$$

ist $\varphi_\pi \circ \varphi_\omega = \varphi_{\pi\circ\omega}$ und $P_\pi \cdot P_\omega = P_{\pi\circ\omega}$. Damit ist auch klar, daß die Matrizen P_π stets invertierbar sind: Da Permutationen als bijektive Abbildungen invertierbar sind, ist einfach $P_\pi^{-1} = P_{\pi^{-1}}$.

Auch die Multiplikation der Permutationsmatrix P_π mit einer beliebigen Matrix (passender Größe) läßt sich leicht ausführen: Für $\pi \in \mathfrak{S}_n$ und $A \in k^{n\times m}$ betrachten wir zu A die lineare Abbildung $\psi \colon k^m \to k^n$, die einen Vektor $\vec{v} \in k^m$ auf $A\vec{v}$ abbildet; dann ist $P_\pi A$ die Abbildungsmatrix von $\varphi_\pi \circ \psi$. Ihre Spalten sind die Bilder der Koordinateneinheitsvektoren von k^m. Der i-te dieser Vektoren wird von ψ auf den i-ten Spaltenvektor von A abgebildet, und auf dessen Komponenten wird die Permutation π angewandt. Dies passiert für jede Spalte von A,

insgesamt entsteht $P_\pi A$ somit aus A, indem man die Permutation π auf dessen *Zeilen* anwendet. Für die Transposition $\pi = (12) \in \mathfrak{S}_3$ etwa ist

$$P_\pi = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und

$$P_\pi \cdot \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

Entsprechend können wir für eine $p \times n$-Matrix B das Produkt BP_π ausrechnen: φ_π bildet den i-ten Koordinateneinheitsvektor von k^n ab auf den $\pi(i)$-ten, und dieser wird von der linearen Abbildung zu B abgebildet auf den $\pi(i)$-ten Spaltenvektor von B. Die Matrix BP_π entsteht also aus B daraus, daß die Permutation π auf dessen *Spalten* angewandt wird. Beispielsweise ist für $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix}$

$$P_\pi = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

und

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

k) Die LR-Zerlegung einer Matrix

Ein lineares Gleichungssystem $A\vec{x} = \vec{b}$ mit einer invertierbaren Matrix A läßt sich zumindest formal sehr leicht auflösen: Durch Multiplikation mit A^{-1} folgt, daß $\vec{x} = A^{-1} \cdot \vec{b}$ ist. Entsprechend einfach läßt sich eine Matrixgleichung der Form $AX = B$ auflösen: Hier ist $X = A^{-1}B$.

Für ein einziges lineares Gleichungssystem ist dieser Lösungsweg nicht sonderlich interessant, denn schon der Aufwand zur Berechnung von A^{-1} ist größer als der zur direkten Lösung des Gleichungssystems:

Wir müssen dazu schließlich mehrere lineare Gleichungssysteme simultan lösen. Hat man aber viele lineare Gleichungssysteme, die sich nur durch ihre rechte Seiten unterscheiden, wie dies zum Beispiel bei linearen Steuerungsproblemen der Fall ist. so kann man sehr viel Zeit sparen, wenn man zunächst ein für alle Mal die inverse Matrix berechnet und dann die einzelnen Gleichungssysteme für den Preis einer Matrix-Vektor-Multiplikation lösen kann. Hinzu kommt, daß man beispielsweise bei eingebauten Steuerungen die Matrixinversion vorab auf einem leistungsfähigen Compter durchführen kann und dann für die eigentliche Steuerung nur die Matrix-Vektor-Multiplikation implementieren muß.

Falls A nicht invertierbar ist, falls man also beispielsweise eine Steuerung mit Redundanz hat, kann man nicht so vorgehen, aber eine fast genauso effiziente Modifikation führt auch hier ans Ziel. Das entsprechende Verfahren wird je nach Lehrbuch als LR-Zerlegung oder LU-Zerlegung bezeichnet, wobei LR für *links/rechts* und LU für *lower/upper* steht.

Die Grundidee der LR-Zerlegung ist dieselbe wie die zur Berechnung der inversen Matrix: Wie wenden den GAUSS-Algorithmus simultan an auf mehrere rechte Seiten.

Um die Struktur der entstehenden Zerlegung besser zu verstehen, betrachten wir den wesentlichen Schritt des GAUSS-Algorithmus, die Addition von Vielfachen einer Gleichung zu einer anderen, als Multiplikation mit einer geeigneten Matrix.

Konkret sei $AX = B$ mit $A \in k^{n\times m}$ und $B \in k^{n\times p}$ eine Gleichung für die unbekannte Matrix $X \in k^{m\times p}$ über dem Körper k. Zur Lösung dieser Gleichung arbeiten wir mit Zeilenumformungen der erweiterten Matrix $M = (A \mid B) \in k^{n\times(m+p)}$, d.h. wir addieren ein Vielfaches einer Zeile zu einer anderen.

Angenommen, wir möchten das c-fache der j-ten Zeile zur i-ten addieren. Dazu betrachten wir die aus aus dem letzten Abschnitt bekannte Matrix $E_{ij} \in k^{n\times n}$, d.h. jene $n \times n$-Matrix, die an der Stelle (i, j) den Eintrag Eins hat und sonst lauter Nullen. Das Produkt $E_{ij}M$ hat, wie wir uns dort überlegt haben, als i-te Zeile die j-te Zeile von M und alle

anderen Zeilen sind Null. Das Produkt $cE_{ij}M$ hat somit als i-te Zeile das c-fache der j-ten Zeile von M (und sonst lauter Nullzeilen). Setzen wir

$$Z_{ij}(c) = E + cE_{ij} \in k^{n\times n}\,,$$

so transformiert die Addition des c-fachen der j-ten Zeile zur i-ten also die Gleichung $AX = B$ in die neue Gleichung

$$\big(Z_{ij}(c)A\big)\,X = Z_{ij}(c)B\,.$$

Da wir beim GAUSS-Algorithmus jeweils Vielfache von weiter oben stehenden Zeilen zu weiter unten stehenden addieren, ist hierbei stets $j < i$, d.h. $Z_{ij}(c)$ ist eine untere Dreiecksmatrix mit lauter Einsen in der Hauptdiagonale.

Führen wir nacheinander mehrere solche Eliminationsschritte aus, so wird M insgesamt mit einem Produkt von mehreren Matrizen der Form $Z_{ij}(c)$ multipliziert; da das Produkt von unteren Dreiecksmatrizen mit Einsen in der Hauptdiagonale wieder eine untere Dreiecksmatrix mit Einsen in der Hauptdiagonale ist, wird M also insgesamt mit einer solchen Matrix Z multipliziert: $(ZA)X = ZB$. Falls wir allein durch Zeilenumformungen die Endgestalt des GAUSS-Algorithmus erreichen können, gibt es also eine untere Dreiecksmatrix Z, die den Gesamteffekt dieser Zeilenumformungen beschreibt, und $ZA = R$ ist eine obere Dreiecksmatrix.

Diese Matrix Z ist bei einem vollbesetzten System aus n Gleichungen in m Unbekannten mit $m \geq n$ im allgemeinen ein Produkt von $\frac{1}{2}(n-1)(n-2)$ Dreiecksmatrizen, denn so viele Koeffizienten müssen wir eliminieren. Es ist klar, daß wir selbst für moderat große Werte von n eine alternative Berechnungsweise finden sollten.

Dazu betrachten wir den Spezialfall $B = E \in k^{n\times n}$, d.h. wir nehmen als rechte Seite eine $n\times n$-Einheitsmatrix. Dann führen wir, ohne uns weiter um die Matrizen $Z_{ij}(c)$ zu kümmern, Zeilenumformungen durch, wie wir es von linearen Gleichungssystemen und von der Matrixinversion her gewohnt sind, bis links eine obere Dreiecksmatrix erscheint.

Der Gesamteffekt dieser Umformungen kann auch so beschrieben werden, daß wir die Ausgangsgleichung $AX = E$ mit einer unteren Dreiecksmatrix Z mit Einsen in der Hauptdiagonale multipliziert haben,

wobei als Koeffizientenmatrix eine obere Dreiecksmatrix $R = ZA$ entstand. Die umgeformte Gleichung ist also

$$(ZA)X = ZE \quad \text{oder} \quad RX = Z\,,$$

d.h. wir können die Matrizen R und Z direkt ablesen: R steht dort, wo am Anfang A stand, und Z steht an der Stelle der Einheitsmatrix.

Die Gleichung $R = ZA$ läßt sich nach A auflösen, denn als untere Dreiecksmatrix mit Einsen in der Hauptdiagonalen ist Z insbesondere invertierbar, wobei auch die inverse Matrix $L = Z^{-1}$ eine untere Dreiecksmatrix mit Einsen in der Hauptdiagonalen ist. Damit haben wir

$$A = Z^{-1}R = LR$$

als Produkt einer unteren und einer oberen Dreiecksmatrix dargestellt; dies bezeichnet man als die LR-Zerlegung von A.

Im allgemeinen erreichen wir die Endgestalt beim GAUSS-Algorithmus allerdings nur, wenn wir zusätzlich zu den Eliminationsschritten auch noch Zeilenvertauschungen vornehmen. Diese entsprechen, wie wir im letzten Abschnitt gesehen haben, der Multiplikation mit Permutationsmatrizen, jetzt wird also der Gesamteffekt der Umformungen ein gemischtes Produkt aus Dreiecksmatrizen und Permutationsmatrizen beschrieben, und über solche Produkte können wir im allgemeinen nichts aussagen.

Wir können aber versuchen, die Zeilen von A gleich am Anfang so zu permutieren, daß anschließend keine Zeilenvertauschungen mehr nötig sind. In diesem Fall haben wir also A mit einer Permutationsmatrix P multipliziert und finden dann Dreiecksmatrizen Z, R und $L = Z^{-1}$, so daß

$$ZPA = R \quad \text{oder} \quad A = P^{-1}LR$$

ist. Strategien, um eine geeignete Permutation P zu finden, werden in der Numerik unter dem Stichwort *Pivotsuche* behandelt; bei kleinen Matrizen wird man sie im allgemeinen ohne Schwierigkeiten auch so finden.

Betrachten wir dazu ein konkretes Beispiel:

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}.$$

Hier müssen wir vor der Anwendung des GAUSS-Algorithmus offensichtlich Zeilen vertauschen, z.B. die erste und die zweite. Dies entspricht einer Multiplikation mit

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und führt auf

$$A' = PA = \begin{pmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 6 & 7 & 8 \end{pmatrix}.$$

Wir schreiben die Einheitsmatrix daneben:

$$\left(\begin{array}{ccc|ccc} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 6 & 7 & 8 & 0 & 0 & 1 \end{array}\right).$$

Die Sechs links unten wird eliminiert durch Subtraktion der zweifachen ersten Zeile von der letzten:

$$\left(\begin{array}{ccc|ccc} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -1 & -2 & -2 & 0 & 1 \end{array}\right)$$

Die Endgestalt entsteht daraus, wenn man nun noch die zweite Zeile zur dritten addiert:

$$\left(\begin{array}{ccc|ccc} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 & 1 \end{array}\right)$$

Hier steht links die Matrix R, rechts steht Z, d.h.

$$R = \begin{pmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix}.$$

In der Tat rechnet man leicht nach, daß

$$ZPA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}$$
$$= \begin{pmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = R$$

ist. Die Inversion von Z geht hier sehr schnell: Im Rechenschema

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -2 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

erhalten wir links die Einheitsmatrix, indem wir von der dritten Zeile zweimal die erste subtrahieren und und einmal die zweite addieren:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right)$$

Somit ist

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 0 \end{pmatrix}$$

und $A = PLR$, denn als Permutationsmatrix zu einer Transposition ist P zu sich selbst invers.

Falls man die LR-Zerlegung einer quadratischen Matrix A kennt, folgt beispielsweise, daß A genau dann invertierbar ist, wenn R invertierbar ist, denn P und L sind immer invertierbar. Alsdann ist

$$A = P^{-1}LR \Longrightarrow A^{-1} = R^{-1}L^{-1}P = R^{-1}ZP\,.$$

Auch für unser Ausgangsproblem, die Lösung von Matrixgleichungen $AX = B$ ist die Kenntnis der LR-Zerlegung nützlich: Durch Multiplikation mit Z erhalten wir die neue Gleichung $RX = ZB$, die man wegen der Treppengestalt der linken Seite leicht durch sukzessives Einsetzen lösen kann.

§4: Basiswechsel, Eigenvektoren und Determinanten

a) Eigenwerte und Eigenvektoren

Die Matrix einer linearen Abbildung $\varphi\colon V \to V$ bezüglich einer Basis $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$ ist genau dann eine Diagonalmatrix, wenn jeder der Basisvektoren $\vec{b}_i$ von φ auf ein Vielfaches $\lambda_i \vec{b}_i$ von sich selbst abgebildet wird; alsdann ist die Abbildungsmatrix gleich der Diagonalmatrix mit Einträgen $\lambda_1, \ldots, \lambda_n$. Somit hängt es sehr von der Basis ab, ob die Abbildungsmatrix Diagonalgestalt hat oder nicht.

Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 1 & -2 \\ 2 & 1 & -2 & 1 \\ 1 & -2 & 1 & 2 \\ -2 & 1 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

etwa ist ganz sicher keine Diagonalmatrix. Betrachten wir die lineare Abbildung

$$\varphi\colon \mathbb{R}^4 \to \mathbb{R}^4; \quad \vec{v} \mapsto A\vec{v}$$

aber bezüglich der Basis $\mathcal{B} = (\vec{b}_1, \vec{b}_2, \vec{b}_3, \vec{b}_4)$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad \vec{b}_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{b}_4 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix},$$

so rechnet man leicht nach, daß

$$\varphi(\vec{b}_1) = A\vec{b}_1 = 2\vec{b}_1, \quad \varphi(\vec{b}_2) = 2\vec{b}_2, \quad \varphi(\vec{b}_3) = -4\vec{b}_3 \quad \text{und} \quad \varphi(\vec{b}_4) = 4\vec{b}_4$$

ist, bezüglich $\mathcal{B}$ hat φ also die Diagonalmatrix

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

als Abbildungsmatrix. Wenn keine schwerwiegenden anderen Gründe dagegensprechen, wird es bei umfangreichen Rechnungen mit der Matrix A meist eine gute Idee sein, statt mit der Standardbasis von $\mathbb{R}^4$