14. Der Gauß-Algorithmus für lineare Gleichungssysteme

Im letzten Kapitel haben wir lineare Gleichungssysteme

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

über einem gegebenen Körper K betrachtet und sie zunächst einmal in Matrixform als

$$Ax = b \quad \text{mit} \quad A \in K^{m \times n}, x \in K^n \text{ und } b \in K^m \tag{2}$$

umgeschrieben, um leichter mit ihnen umgehen zu können. Wir wollen nun untersuchen, wie die Lösungsmengen solcher Gleichungssysteme aussehen, und dabei auch Algorithmen (d. h. Rechenverfahren) entwickeln, um diese Lösungsmengen auch konkret berechnen zu können. Außerdem werden wir mit diesen Verfahren auch bestimmen können, ob eine gegebene quadratische Matrix im Sinne von Definition 13.15 invertierbar ist, und in diesem Fall ihre inverse Matrix berechnen können.

14.A Elementarmatrizen und Zeilenstufenformen

Die Idee für die Lösung linearer Gleichungssysteme ist sehr naheliegend: Wir wollen die gegebenen Gleichungen in (1) oben so umformen und miteinander kombinieren, dass die neuen Gleichungen im Idealfall nur noch jeweils eine Variable beinhalten und ihre Lösung damit leicht abgelesen werden kann. Dies können wir z. B. tun, indem wir auf geschickte Art mehrfach hintereinander ein geeignetes Vielfaches einer Gleichung zu einer anderen addieren, so dass dabei möglichst viele Variablen herausfallen.

Beachte, dass in der Matrixschreibweise (2) des Gleichungssystems jede Gleichung aus (1) einer Zeile der Matrix entspricht. Umformungen, die in (1) Gleichungen miteinander kombinieren, werden in (2) also Zeilen der Matrix miteinander kombinieren. Wie wir jetzt sehen werden, lassen sich solche Zeilenumformungen sehr effizient als Matrixprodukte schreiben.

Konstruktion 14.1 (Elementarmatrizen)**.** Es seien $m, n \in \mathbb{N}$ und $A = (a_{i,j})_{i,j} \in K^{m \times n}$.

(a) Für $k \in \{1, \dots, m\}$ und $\lambda \in K \backslash \{0\}$ setzen wir

$$F_k(\lambda) := \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in K^{m \times m},$$

wobei der Eintrag λ in Zeile und Spalte k steht. Die Matrix $F_k(\lambda)$ ist also nichts weiter als die Einheitsmatrix, bei der der Eintrag 1 in der k-ten Zeile und Spalte durch ein $\lambda \neq 0$ ersetzt wurde. Mit dieser Matrix ist

$$F_k(\lambda) \cdot A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ \lambda a_{k,1} & \cdots & \lambda a_{k,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix},$$

d. h. die Multiplikation von A mit $F_k(\lambda)$ von links entspricht genau der Multiplikation der k-ten Zeile von A mit λ.

(b) Für $k,l \in \{1,\dots,m\}$ mit $k \neq l$ und $\lambda \in K$ setzen wir

$$F_{k,l}(\lambda) := \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \lambda & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in K^{m\times m},$$

wobei der Eintrag λ in Zeile k und Spalte l steht, d. h. diesmal haben wir in der Einheitsmatrix den Eintrag 0 in Zeile k und Spalte l durch λ ersetzt. (Beachte, dass der Eintrag λ für $k < l$ oberhalb und für $k > l$ unterhalb der Diagonale steht; wir haben in der Matrix oben der Einfachheit halber nur den Fall $k < l$ dargestellt.) In diesem Fall ist

$$F_{k,l}(\lambda)\cdot A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \lambda & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,n} \\ \vdots & & \vdots \\ a_{l,1} & \cdots & a_{l,n} \\ \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & & \vdots \\ a_{k,1}+\lambda a_{l,1} & \cdots & a_{k,n}+\lambda a_{l,n} \\ \vdots & & \vdots \\ a_{l,1} & \cdots & a_{l,n} \\ \vdots & & \vdots \end{pmatrix},$$

d. h. die Multiplikation von A mit $F_{k,l}(\lambda)$ von links entspricht der Addition des λ-fachen von Zeile l zu Zeile k.

Die Matrizen $F_k(\lambda)$ für $\lambda \in K\backslash\{0\}$ sowie $F_{k,l}(\lambda)$ für $k \neq l$ und $\lambda \in K$ heißen **Elementarmatrizen**. Es gibt sie in jeder (quadratischen) Größe $m \times m$; zur Vereinfachung der Schreibweise deuten wir diese Größe in der Notation $F_k(\lambda)$ bzw. $F_{k,l}(\lambda)$ aber nicht an.

Man sagt, dass $F_k(\lambda)\cdot A$ und $F_{k,l}(\lambda)\cdot A$ aus A durch eine **elementare Zeilenumformung** entstehen.

Bemerkung 14.2.

(a) Die Elementarmatrizen sind invertierbar mit

$$(F_k(\lambda))^{-1} = F_k\left(\frac{1}{\lambda}\right) \quad \text{und} \quad (F_{k,l}(\lambda))^{-1} = F_{k,l}(-\lambda).$$

Dies folgt direkt aus Konstruktion 14.1: Wenn wir z. B. das Matrixprodukt

$$F_k\left(\frac{1}{\lambda}\right)\cdot F_k(\lambda) = F_k\left(\frac{1}{\lambda}\right)\cdot F_k(\lambda)\cdot E$$

bilden, multiplizieren wir die k-te Zeile in der Einheitsmatrix zuerst mit λ und dann mit $\frac{1}{\lambda}$, d. h. es kommt insgesamt wieder die Einheitsmatrix heraus – also ist $F_k(\frac{1}{\lambda})\cdot F_k(\lambda) = E$. Genauso ergibt sich auch $F_k(\lambda)\cdot F_k(\frac{1}{\lambda})$, und damit ist $F_k(\frac{1}{\lambda})$ das Inverse von $F_k(\lambda)$.

Analog zeigt man die Aussage über das Inverse von $F_{k,l}(\lambda)$.

(b) Das Vertauschen von zwei Zeilen $k,l \in \{1,\dots,m\}$ mit $k \neq l$ in einer Matrix $A \in K^{m\times n}$ lässt sich als Folge von elementaren Zeilenumformungen realisieren: Sind $a_1,\dots,a_m$ die Zeilen von A, so erhalten wir

$$A = \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l \\ \vdots \end{pmatrix} \xrightarrow{F_{k,l}(1)\cdot} \begin{pmatrix} \vdots \\ a_k+a_l \\ \vdots \\ a_l \\ \vdots \end{pmatrix} \xrightarrow{F_{l,k}(-1)\cdot} \begin{pmatrix} \vdots \\ a_k+a_l \\ \vdots \\ -a_k \\ \vdots \end{pmatrix} \xrightarrow{F_{k,l}(1)\cdot} \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ -a_k \\ \vdots \end{pmatrix} \xrightarrow{F_l(-1)\cdot} \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ a_k \\ \vdots \end{pmatrix}.$$

(c) Führen wir mit einer Matrix $A \in K^{m\times n}$ mehrere elementare Zeilenumformungen aus, so ist das Ergebnis nach Konstruktion 14.1 eine Matrix $F_1\cdot\cdots\cdot F_r\cdot A$ mit Elementarmatrizen $F_1,\dots,F_r$. Da diese Elementarmatrizen nach (a) invertierbar sind, ist das Produkt

$F := F_1 \cdot \dots \cdot F_r$ nach Lemma 13.16 ebenfalls wieder invertierbar. Das Ergebnis der Zeilenumformungen ist also die Matrix FA, d. h. wir haben A von links mit einer invertierbaren Matrix multipliziert. In der Tat werden wir in Lemma 14.11 noch sehen, dass *jede* Multiplikation mit einer invertierbaren Matrix von links als Abfolge von elementaren Zeilenumformungen geschrieben werden kann.

Mit derartigen Zeilenumformungen wollen wir jetzt eine Matrix bzw. ein lineares Gleichungssystem auf eine möglichst einfache Form bringen. Die folgende Form wird dabei herauskommen.

Definition 14.3 (Zeilenstufenform)**.** Es seien $m,n,r \in \mathbb{N}$ und $A = (a_{i,j})_{i,j} \in K^{m\times n}$.

(a) Wir sagen, dass A in **Zeilenstufenform** mit r **Stufen** ist, wenn

$$\begin{array}{cc} & \begin{array}{ccccc} k_1 & k_2 & & k_r & \\ \downarrow & \downarrow & & \downarrow & \end{array} \\ A = & \left(\begin{array}{ccccc} 1 & & & & \\ & 1 & & & * \\ & & \ddots & & \\ 0 & & & 1 & \\ & & & & \end{array}\right) \end{array}$$

ist, wobei das Symbol „$*$" für beliebige Einträge steht. Mit anderen Worten gibt es also Spalten $1 \le k_1 < \dots < k_r \le n$ (die sogenannten **Stufenspalten**), so dass gilt:

- $a_{i,k_i} = 1$ für alle $i = 1,\dots,r$;
- rechts hiervon stehen beliebige Einträge (d. h. $a_{i,j}$ ist beliebig falls $i \le r$ und $j > k_i$);
- alle anderen Einträge sind 0.

(b) Sind zusätzlich noch außer den geforderten Einsen alle anderen Einträge in den Stufenspalten gleich Null, ist A also sogar von der Form

$$\left(\begin{array}{cccccc} 1 & *\cdots* & 0 & *\cdots* & 0 & *\cdots* \\ & & 1 & *\cdots* & 0 & *\cdots* \\ & & & \ddots & \vdots & \vdots \\ & 0 & & & 1 & *\cdots* \\ & & & & & \end{array}\right)$$

(d. h. ist $a_{l,k_i} = 0$ für alle $l < i \le r$), so sagt man, dass A in **reduzierter Zeilenstufenform** ist.

(c) Ist $S \in \mathrm{GL}(m,K)$ eine invertierbare Matrix, so dass SA in (reduzierter) Zeilenstufenform ist, so heißt SA eine (reduzierte) Zeilenstufenform von A.

Der wesentliche Satz in diesem Abschnitt besagt nun, dass jede Matrix eine (reduzierte) Zeilenstufenform besitzt. Dabei sind beide Formen in der Praxis wichtig: Die reduzierte Zeilenstufenform wird sich in Satz 14.6 als eindeutig herausstellen und ist damit aus theoretischer Sicht natürlich deutlich schöner – andererseits werden wir aber auch sehen, dass die normale, nicht-reduzierte Zeilenstufenform für viele Zwecke ausreichend und mit weniger Rechenaufwand zu erzielen ist. In der Tat ist der Beweis des folgenden Satzes konstruktiv in dem Sinne, dass er auch ein Verfahren angibt, wie eine (reduzierte) Zeilenstufenform mit elementaren Zeilenumformungen erreicht werden kann. Dieser Algorithmus im Beweis ist mindestens genauso wichtig wie die eigentliche Aussage:

Satz 14.4 (**Gauß-Algorithmus**)**.** *Jede Matrix $A \in K^{m\times n}$ lässt sich mit elementaren Zeilenumformungen auf (reduzierte) Zeilenstufenform bringen.*

Insbesondere besitzt A im Sinne von Definition 14.3 (c) also eine (reduzierte) Zeilenstufenform.

Beweis. Wir beweisen den Satz mit Induktion über die Anzahl n der Spalten von A. Da für $n = 0$ nichts zu zeigen ist, müssen wir nur den Induktionsschritt durchführen. Es sei also $A \in K^{m\times(n+1)}$ beliebig vorgegeben. Wir wollen A mit elementaren Zeilenumformungen auf (reduzierte) Zeilenstufenform bringen und nehmen nach Induktionsvoraussetzung an, dass wir dies für alle Matrizen mit n Spalten bereits können. Wir unterscheiden dabei zwei Fälle:

Fall 1: Alle Einträge in der ersten Spalte von A sind 0, d. h. A hat die Form

$$A = \left(\begin{array}{c|c} 0 & \\ \vdots & A' \\ 0 & \end{array}\right)$$

für eine Matrix $A' \in K^{m\times n}$. Nach Induktionsvoraussetzung können wir A' dann durch elementare Zeilenumformungen auf (reduzierte) Zeilenstufenform bringen. Da diese Zeilenumformungen an der ersten Nullspalte aber nichts ändern, haben wir damit auch A auf (reduzierte) Zeilenstufenform gebracht.

Fall 2: Es sind nicht alle Einträge in der ersten Spalte von A gleich 0.

(a) Falls der Eintrag $a_{1,1}$ links oben in A gleich Null ist, vertauschen wir zwei Zeilen von A so, dass dieser Eintrag nicht mehr gleich 0 ist (nach Bemerkung 14.2 (b) ist dies durch elementare Zeilenumformungen machbar).

(b) Wir dividieren die erste Zeile durch $a_{1,1}$ und erhalten eine Matrix der Form

$$\left(\begin{array}{c|c} 1 & \\ a_{2,1} & \\ \vdots & * \\ a_{m,1} & \end{array}\right).$$

(c) Wir subtrahieren von jeder Zeile $k \in \{2,\dots,m\}$ das $a_{k,1}$-fache der ersten Zeile und bekommen dadurch

$$\left(\begin{array}{c|c} 1 & * \cdots * \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array}\right)$$

mit einer Matrix $A' \in K^{(m-1)\times n}$.

(d) Nach Induktionsvoraussetzung können wir jetzt A' durch elementare Zeilenumformungen auf Zeilenstufenform bringen. Damit hat aber auch die gesamte Matrix bereits Zeilenstufenform – wollen wir also nur diese normale, nicht-reduzierte Zeilenstufenform erreichen, so sind wir damit fertig. Andernfalls bringen wir A' gemäß der Induktionsvoraussetzung sogar auf reduzierte Zeilenstufenform und bekommen eine Matrix der Form

$$\left(\begin{array}{c|cccccc} 1 & *\cdots* & * & *\cdots* & & * & *\cdots* \\ \hline 0 & & 1 & *\cdots* & & 0 & *\cdots* \\ \vdots & & & & \ddots & & \\ \vdots & & 0 & & & & \\ \vdots & & & & & 1 & *\cdots* \\ 0 & & & & & & \end{array}\right).$$

Subtrahieren wir nun geeignete Vielfache der Zeilen $2,\dots,m$ von der ersten Zeile, so können wir damit dann noch die Einträge in den Stufenspalten der ersten Zeile zu Null machen und erhalten so auch die reduzierte Zeilenstufenform. □

Beispiel 14.5. Wir wollen die Matrix

$$A = \begin{pmatrix} -1 & 2 & 4 \\ -1 & 3 & 7 \\ 1 & 1 & 5 \end{pmatrix} \quad \in \mathbb{R}^{3\times 3}$$

mit elementaren Zeilenumformungen auf Zeilenstufenform bringen. Dazu können wir nach dem Algorithmus im Beweis von Satz 14.4 wie folgt vorgehen (die Notation $Z3-3Z2 \to Z3$ bedeutet z. B., dass wir das Dreifache der zweiten Zeile von der dritten subtrahieren):

$$\begin{pmatrix} -1 & 2 & 4 \\ -1 & 3 & 7 \\ 1 & 1 & 5 \end{pmatrix} \xrightarrow{-Z1\to Z1} \begin{pmatrix} 1 & -2 & -4 \\ -1 & 3 & 7 \\ 1 & 1 & 5 \end{pmatrix} \xrightarrow{\substack{Z2+Z1\to Z2 \\ Z3-Z1\to Z3}} \begin{pmatrix} 1 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 3 & 9 \end{pmatrix}$$

$$\xrightarrow{Z3-3Z2\to Z3} \begin{pmatrix} 1 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Möchten wir eine reduzierte Zeilenstufenform, so addieren wir schließlich noch das Doppelte der zweiten Zeile zur ersten und erhalten

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Satz 14.6 (Eindeutigkeit der reduzierten Zeilenstufenform)**.** *Die reduzierte Zeilenstufenform einer Matrix $A \in K^{m\times n}$ ist eindeutig.*

Beweis. Wie in Satz 14.4 zeigen wir diese Aussage wieder mit Induktion über die Anzahl n der Spalten von A, wobei auch hier der Induktionsanfang für $n=0$ trivial ist.

Für den Induktionsschritt sei $n \in \mathbb{N}$ so, dass die reduzierte Zeilenstufenform jeder Matrix in $K^{m\times n}$ eindeutig ist. Wir betrachten nun eine Matrix in $K^{m\times(n+1)}$, die wir als $(A\,|\,b)$ mit $A \in K^{m\times n}$ und $b \in K^m$ schreiben können. Es seien

$$S(A\,|\,b) = (SA\,|\,Sb) \quad \text{und} \quad \tilde{S}(A\,|\,b) = (\tilde{S}A\,|\,\tilde{S}b)$$

mit $S, \tilde{S} \in \mathrm{GL}(m,K)$ zwei reduzierte Zeilenstufenformen dieser Matrix. Wir müssen zeigen, dass sie übereinstimmen.

Da eine reduzierte Zeilenstufenform durch Streichen der letzten Spalte wieder in eine reduzierte Zeilenstufenform übergeht, sind SA und $\tilde{S}A$ damit reduzierte Zeilenstufenformen von $A \in K^{m\times n}$. Nach Induktionsvoraussetzung gilt also bereits $SA = \tilde{S}A$, d. h. wir müssen nur noch $Sb = \tilde{S}b$ zeigen.

Es sei dazu r die Anzahl der Stufen in der Zeilenstufenform $SA = \tilde{S}A$. Insbesondere enthält diese Matrix dann die Einheitsvektoren $e_1, \dots, e_r$ in den Stufenspalten $k_1, \dots, k_r$ wie in Definition 14.3; nach Beispiel 13.10 (d) gilt also

$$SAe_{k_i} = \tilde{S}Ae_{k_i} = e_i \quad \Rightarrow \quad S^{-1}e_i = \tilde{S}^{-1}e_i \qquad \text{für alle } i = 1, \dots, r. \tag{$*$}$$

Wir unterscheiden nun zwei Fälle:

(a) Hat die reduzierte Zeilenstufenform $(SA\,|\,Sb)$ ebenfalls nur r Stufen, so hat Sb die Form $\lambda_1 e_1 + \cdots + \lambda_r e_r$ für gewisse $\lambda_1, \dots, \lambda_r \in K$. Wir erhalten dann wie behauptet

$$\tilde{S}b = \tilde{S}S^{-1}Sb = \tilde{S}S^{-1}(\lambda_1 e_1 + \cdots \lambda_r e_r) \overset{(*)}{=} \tilde{S}\tilde{S}^{-1}(\lambda_1 e_1 + \cdots \lambda_r e_r) = \lambda_1 e_1 + \cdots \lambda_r e_r = Sb.$$

Analog folgt dies natürlich, wenn die reduzierte Zeilenstufenform $(\tilde{S}A\,|\,\tilde{S}b)$ aus r Stufen besteht.

(b) Haben beide reduzierten Zeilenstufenformen $(SA\,|\,Sb)$ und $(\tilde{S}A\,|\,\tilde{S}b)$ mehr als r Stufen, so muss die letzte Spalte dieser Matrizen eine Stufenspalte sein. Nach Definition 14.3 (b) ist dies nur möglich für $Sb = \tilde{S}b = e_{r+1}$. □

Bemerkung 14.7 (Die reduzierte Zeilenstufenform als Normalform)**.** Oft betrachtet man in der linearen Algebra Matrizen mit einer bestimmten Art erlaubter Umformungen (in unserem Fall Multiplikationen mit invertierbaren Matrizen von links) und versucht damit, eine eindeutig bestimmte und in gewissem Sinne möglichst einfache Form der Matrix zu erreichen. Eine solche Form bezeichnet man dann als *Normalform*.

In dieser Sprechweise ist die reduzierte Zeilenstufenform nach Satz 14.4 und Satz 14.6 also eine Normalform bezüglich der Multiplikation mit invertierbaren Matrizen von links. Wir werden in dieser Vorlesung noch einige weitere Normalformen bezüglich anderer Umformungen kennenlernen (siehe **??**).

14.B Der Rang von Matrizen

Da jede Matrix A nach Satz 14.4 und Satz 14.6 eine eindeutig bestimmte reduzierte Zeilenstufenform besitzt, ist insbesondere natürlich auch die Anzahl der Stufen in dieser Form eindeutig durch A festgelegt. Diese Zahl, die für die Untersuchung von A ganz besonders wichtig ist, wollen wir in diesem Abschnitt nun genauer betrachten.

Definition 14.8 (Rang einer Matrix)**.** Für eine Matrix $A \in K^{m\times n}$ heißt die Anzahl der Stufen in der reduzierten Zeilenstufenform von A der **Rang** von A. Wir schreiben ihn als $\operatorname{rk} A$.

Bemerkung 14.9. Es sei $A \in K^{m\times n}$.

(a) Offensichtlich ist $\operatorname{rk} A \le m$ und auch $\operatorname{rk} A \le n$, da nicht mehr Stufen in einer Matrix der Größe $m \times n$ Platz haben.

(b) Wenden wir den Gauß-Algorithmus an, um eine Matrix in Zeilenstufenform auf eine Matrix in reduzierter Zeilenstufenform zu bringen, so wird die Matrix im Beweis von Satz 14.4 nur noch in Schritt (d) umgeformt, der die Anzahl der Stufen aber nicht mehr verändert.

Wollen wir den Rang einer Matrix A berechnen, genügt uns also schon eine *beliebige* Zeilenstufenform von A: Ihre Anzahl Stufen ist ebenfalls gleich $\operatorname{rk} A$. Wir werden dies im Folgenden oft verwenden, ohne jedes Mal explizit darauf hinzuweisen.

Beispiel 14.10.

(a) Für alle $m, n \in \mathbb{N}$ sind sowohl die Nullmatrix $0 \in K^{m\times n}$ als auch die Einheitsmatrix $E_n \in K^{n\times n}$ bereits in (reduzierter) Zeilenstufenform mit 0 bzw. n Stufen. Also ist $\operatorname{rk} 0 = 0$ und $\operatorname{rk} E_n = n$.

(b) Für die Matrix
$$A = \begin{pmatrix} -1 & 2 & 4 \\ -1 & 3 & 7 \\ 1 & 1 & 5 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$
aus Beispiel 14.5 gilt $\operatorname{rk} A = 2$, da die dort berechnete Zeilenstufenform zwei Stufen hat.

(c) Hat $A \in K^{m\times n}$ nur r Zeilen, in denen ein Eintrag ungleich 0 vorkommt, so ist $\operatorname{rk} A \le r$: Wir können dann diese r Zeilen nach oben tauschen und den Gauß-Algorithmus nur mit diesen obersten r Zeilen durchführen, wodurch sich natürlich auch eine Zeilenstufenform mit höchstens r Stufen ergibt.

Mit Hilfe des Rangs einer Matrix können wir nun bereits das Problem aus Kapitel 13 lösen, wie man von einer quadratischen Matrix bestimmen kann, ob sie invertierbar ist, und in diesem Fall die inverse Matrix berechnen kann.

Lemma 14.11 (Äquivalente Kriterien für Invertierbarkeit)**.** *Für eine quadratische Matrix $A \in K^{n\times n}$ sind äquivalent:*

(a) $A \in \mathrm{GL}(n, K)$, *d. h. A ist invertierbar.*

(b) $\operatorname{rk} A = n$.

(c) *Die reduzierte Zeilenstufenform von A ist E_n.*

(d) *A ist ein Produkt von Elementarmatrizen.*

Beweis.

(a) $\Rightarrow$ (b): Mit dem Gauß-Algorithmus finden wir ein Produkt F von Elementarmatrizen, so dass FA in Zeilenstufenform ist. Aber FA ist auch invertierbar (da F und A es sind) und kann damit nach Beispiel 13.18 (c) keine Zeile haben, deren Einträge alle 0 sind. Also muss es n Stufen in der Zeilenstufenform FA geben, d. h. es ist $\operatorname{rk} A = n$.

(b) $\Rightarrow$ (c): Dies ist klar, da E_n die einzige $n \times n$-Matrix in reduzierter Zeilenstufenform mit n Stufen ist.

(c) $\Rightarrow$ (d): Nach Voraussetzung können wir A mit dem Gauß-Algorithmus auf die Einheitsmatrix bringen, d. h. es gibt Elementarmatrizen $F_1, \dots, F_k$ mit $F_1 \cdots F_k A = E_n$. Nach Bemerkung 14.2 (a) sind Elementarmatrizen aber invertierbar und ihre Inversen wieder Elementarmatrizen, und damit ist $A = F_k^{-1} \cdots F_1^{-1}$ ein Produkt von Elementarmatrizen.

(d) $\Rightarrow$ (a): Da Elementarmatrizen invertierbar sind, sind nach Lemma 13.16 auch Produkte von Elementarmatrizen wieder invertierbar. □

Algorithmus 14.12 (Inverse Matrix)**.** Gegeben sei eine quadratische Matrix $A \in K^{n \times n}$. Wir wollen überprüfen, ob A invertierbar ist, und in diesem Fall die inverse Matrix A^{-1} berechnen. Dazu starten wir mit der $n \times (2n)$-Matrix $(A \,|\, E_n)$, in der wir A und die Einheitsmatrix der gleichen Größe nebeneinander schreiben. Wir bringen dann die Matrix A mit dem Gauß-Algorithmus auf reduzierte Zeilenstufenform, machen die dafür benötigten Zeilenumformungen aber in allen $2n$ Spalten der Matrix. Ist $F \in \mathrm{GL}(n,K)$ das Produkt der Elementarmatrizen, das den durchgeführten Umformungen entspricht, so erhalten wir also die Matrix $(FA \,|\, FE_n) = (FA \,|\, F)$, wobei in der linken Hälfte die reduzierte Zeilenstufenform FA von A steht.

Nach Lemma 14.11 ist A genau dann invertierbar, wenn für diese reduzierte Zeilenstufenform die Einheitsmatrix herausgekommen ist, wenn also $FA = E_n$ ist. In diesem Fall ist aber offensichtlich $F = A^{-1}$ die gesuchte inverse Matrix – und diese steht nach den Umformungen genau in der rechten Hälfte unseres Diagramms.

Als konkretes Beispiel wollen wir die Matrix

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

wählen. Wir wenden also wie gewohnt den Gauß-Algorithmus auf die Matrix A an, führen aber alle Umformungen mit der 2×4-Matrix $(A \,|\, E_2)$ durch:

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right) \overset{Z2-3Z1 \to Z2}{\longrightarrow} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & 1 \end{array}\right) \overset{Z1-Z2 \to Z1}{\longrightarrow} \left(\begin{array}{cc|cc} 1 & 0 & 4 & -1 \\ 0 & 1 & -3 & 1 \end{array}\right).$$

Da die herausgekommene reduzierte Zeilenstufenform von A (die linke Hälfte dieser Matrix) die Einheitsmatrix ist, ist A invertierbar. Die inverse Matrix steht nun in der rechten Hälfte des Diagramms: Es ist

$$A^{-1} = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}.$$

Bisher haben wir mit einer gegebenen Matrix $A \in K^{m \times n}$ nur *Zeilen*umformungen durchgeführt, weil wir damit – wie wir am Anfang von Abschnitt 14.A schon gesehen hatten und gleich noch genauer untersuchen werden – gerade die möglichen Umformungen eines linearen Gleichungssystems der Form $Ax = b$ mit $x \in K^n$ und $b \in K^m$ beschreiben können. Diese Umformungen konnten wir durch Multiplikation mit einer invertierbaren Matrix $S \in \mathrm{GL}(m,K)$ von links erreichen, nämlich mit einem Produkt von Elementarmatrizen.

Natürlich können wir mit A aber analog auch *Spalten*umformungen durchführen, was dann wie in Konstruktion 14.1 einer Multiplikation mit einem Produkt von Elementarmatrizen bzw. einer invertierbaren Matrix $T \in \mathrm{GL}(n,K)$ von rechts entspricht. Erlauben wir sowohl Zeilen- als auch Spaltenumformungen, würden wir vermutlich erwarten, dass wir A auf eine noch einfachere Form als eine reduzierte Zeilenstufenform bringen können, da wir ja mehr Möglichkeiten für die Umformungen haben. In der Tat wollen wir jetzt sehen, dass wir dann eine noch viel einfachere Normalform im Sinne von Bemerkung 14.7 erhalten können: Der Rang kann durch diese Umformungen zwar nicht geändert werden, aber für einen gegebenen Rang erhalten wir immer dieselbe Normalform.

Diese Umformung von Matrizen durch Multiplikation mit invertierbaren Matrizen von links und rechts hat übrigens eine besondere Bedeutung, die wir in **??** noch sehen werden. Sie hat daher einen speziellen Namen.

Definition 14.13 (Äquivalente Matrizen)**.** Es seien $m,n \in \mathbb{N}$. Zwei Matrizen $A,B \in K^{m\times n}$ heißen **äquivalent** zueinander, wenn es invertierbare Matrizen $S \in \mathrm{GL}(m,K)$ und $T \in \mathrm{GL}(n,K)$ gibt mit $B = SAT$.

Man prüft leicht nach, dass dies in der Tat eine Äquivalenzrelation auf $K^{m\times n}$ gemäß Definition 2.27 ist.

Lemma 14.14. *Für jede Matrix $A \in K^{m\times n}$ gilt:*

(a) *Für alle $S \in \mathrm{GL}(m,K)$ gilt $\mathrm{rk}(SA) = \mathrm{rk}\,A$.*

(b) *Für alle $T \in K^{n\times p}$ gilt $\mathrm{rk}(AT) \le \mathrm{rk}\,A$, mit Gleichheit falls $T \in \mathrm{GL}(n,K)$.*

Insbesondere haben äquivalente Matrizen also denselben Rang.

Beweis. Es sei $F \in \mathrm{GL}(m,K)$ ein Produkt von Elementarmatrizen, so dass FA in Zeilenstufenform mit $r := \mathrm{rk}\,A$ Stufen ist.

(a) Wegen $FS^{-1} \cdot SA = FA$ ist die Matrix FA auch eine Zeilenstufenform von SA, es ist also auch $\mathrm{rk}(SA) = r = \mathrm{rk}\,A$.

(b) Da die unteren $m-r$ Zeilen von FA nur Nullen enthalten, gilt dies nach Definition der Matrixmultiplikation auch für das Matrixprodukt $FA \cdot T$. Nach Beispiel 14.10 (c) ist damit $\mathrm{rk}(AT) \overset{(a)}{=} \mathrm{rk}(FAT) \le r = \mathrm{rk}\,A$.

Ist T sogar invertierbar, folgt daraus auch $\mathrm{rk}\,A = \mathrm{rk}(AT \cdot T^{-1}) \le \mathrm{rk}(AT)$, und damit die Gleichheit $\mathrm{rk}(AT) = \mathrm{rk}\,A$. □

Satz 14.15 (Normalform von Matrizen bezüglich Äquivalenz)**.** *Zu jeder Matrix $A \in K^{m\times n}$ gibt es ein $r \in \mathbb{N}$ und invertierbare Matrizen $S \in \mathrm{GL}(m,K)$ und $T \in \mathrm{GL}(n,K)$ mit*

$$SAT = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ \hline & 0 & & 0 \end{array}\right) \in K^{m\times n}.$$

Dabei ist r eindeutig bestimmt als $r = \mathrm{rk}\,A$.

*Analog zu Bemerkung 14.7 sagt man, dass eine solche Matrix in **Normalform** bezüglich der Äquivalenz von Matrizen ist.*

Beweis. Nach Lemma 14.14 ist $r = \mathrm{rk}(SAT) = \mathrm{rk}\,A$; eine Normalform wie im Satz ist also höchstens für $r = \mathrm{rk}\,A$ möglich.

Andererseits können wir eine solche Normalform aber auch einfach durch Zeilen- und Spaltenumformungen erreichen: Dazu bringen wir dei Matrix A zunächst mit elementaren Zeilenumformungen auf ihre reduzierte Zeilenstufenform. Da bei dieser Zeilenstufenform genau die Einheitsvektoren $e_1,\dots,e_r$ in den Stufenspalten stehen, können wir nun durch Subtraktion geeigneter Vielfacher dieser Stufenspalten von den anderen Spalten alle übrigen Einträge der Matrix zu 0 machen. Durch Spaltenvertauschung können wir schließlich noch die Stufenspalten ganz nach links schieben und so die gewünschte Normalform erreichen. □

Beispiel 14.16. Da die Matrix

$$A = \begin{pmatrix} -1 & 2 & 4 \\ -1 & 3 & 7 \\ 1 & 1 & 5 \end{pmatrix} \quad \in \mathbb{R}^{3\times 3}$$

aus Beispiel 14.5 den Rang 2 hat, gibt es nach Satz 14.15 invertierbare Matrizen $S,T \in \mathrm{GL}(3,\mathbb{R})$ mit

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

Eine wichtige Folgerung aus Satz 14.15 ist, dass sich der Rang einer Matrix durch Transponieren nicht ändert: Für Matrizen in einer solchen Normalform ist dies nämlich offensichtlich, und jede andere ist ja äquivalent dazu. Exakt aufgeschrieben sieht dieser Beweis dann wie folgt aus.

Folgerung 14.17. *Für jede Matrix $A \in K^{m\times n}$ gilt $\mathrm{rk}(A^\mathsf{T}) = \mathrm{rk}\,A$.*

Beweis. Nach Satz 14.15 gibt es $S \in \mathrm{GL}(m,K)$ und $T \in \mathrm{GL}(n,K)$, so dass

$$SAT = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

mit $r = \mathrm{rk}\,A$. Mit Lemma 13.11 (d) gilt dann auch

$$T^\mathsf{T} A^\mathsf{T} S^\mathsf{T} = (SAT)^\mathsf{T} = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right)^\mathsf{T} = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

(wobei sich beim letzten Gleichheitszeichen die Größe der Matrix von $m\times n$ auf $n\times m$ ändert). Nach Folgerung 13.17 (d) sind mit S und T ferner auch S^T und T^T invertierbar, d. h. $T^\mathsf{T}A^\mathsf{T}S^\mathsf{T}$ ist äquivalent zu A^T. Damit ergibt sich aus Lemma 14.14

$$\mathrm{rk}\,A^\mathsf{T} = \mathrm{rk}(T^\mathsf{T}A^\mathsf{T}S^\mathsf{T}) = \mathrm{rk}\left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) = r = \mathrm{rk}\,A. \qquad \square$$

Eine weitere Folgerung hieraus ist die in Beispiel 13.18 (a) schon angekündigte Aussage, dass es für die Bestimmung der Invertierbarkeit einer quadratischen Matrix ausreicht, eine links- oder rechtsinverse Matrix zu finden.

Folgerung 14.18 (Rang eines Matrixprodukts)**.** *Es seien $A \in K^{m\times n}$ und $B \in K^{n\times p}$.*

(a) *Es gilt $\mathrm{rk}(AB) \le \mathrm{rk}\,A$ und $\mathrm{rk}(AB) \le \mathrm{rk}\,B$.*

(b) *Für quadratische Matrizen (also $m = n = p$) genügt bereits eine der beiden Gleichungen $AB = E_n$ und $BA = E_n$ dafür, dass A invertierbar ist mit $A^{-1} = B$.*

Beweis.

(a) Die erste Ungleichung ist genau Lemma 14.14 (b). Die zweite erhalten wir daraus nun einfach mit Folgerung 14.17 durch Transponieren, denn

$$\mathrm{rk}(AB) = \mathrm{rk}\left((AB)^\mathsf{T}\right) = \mathrm{rk}(B^\mathsf{T}A^\mathsf{T}) \overset{14.14\,(b)}{\le} \mathrm{rk}(B^\mathsf{T}) = \mathrm{rk}\,B.$$

(b) Es sei zunächst $AB = E_n$. Nach (a) ist dann $n = \mathrm{rk}\,E_n = \mathrm{rk}(AB) \le \mathrm{rk}\,A$, also $\mathrm{rk}\,A = n$. Nach Lemma 14.11 ist A damit invertierbar, und Multiplikation der Gleichung $AB = E_n$ mit A^{-1} von links liefert sofort $B = A^{-1}$.

Den Fall $BA = E_n$ zeigt man natürlich analog mit der zweiten Ungleichung aus (a). $\square$

14.C Lösungsmengen linearer Gleichungssysteme

Nach unseren Vorarbeiten in diesem Kapitel können wir jetzt untersuchen, wie die Lösungsmengen linearer Gleichungssysteme aussehen und konkret berechnet werden können. Dazu führen wir zunächst ein paar nützliche Notationen ein.

Definition 14.19 ($L(A,b)$, Bild und Kern einer Matrix)**.** Es seien $A \in K^{m\times n}$ eine Matrix und $b \in K^m$.

(a) Wir bezeichnen die Lösungsmenge des linearen Gleichungssystems $Ax = b$ mit

$$L(A,b) := \{x \in K^n : Ax = b\} \quad \subset K^n.$$

(b) Das **Bild** von A ist definiert als

$$\operatorname{Im} A := \{Ax : x \in K^n\} \quad \subset K^m$$

(die Schreibweise kommt vom englischen Wort „image“). Nach Definition ist das Gleichungssystem $Ax = b$ also genau dann lösbar, d. h. $L(A,b) \neq \emptyset$, wenn $b \in \operatorname{Im} A$.

(c) Der **Kern** von A ist definiert als

$$\operatorname{Ker} A := \{x \in K^n : Ax = 0\} \quad \subset K^n$$

(die Schreibweise kommt vom englischen Wort „kernel“). Offensichtlich ist $\operatorname{Ker} A = L(A,0)$.

Bemerkung 14.20. Schreiben wir in Definition 14.19 die Matrix A spaltenweise als $A = (a_1 \,|\, \cdots \,|\, a_n)$ mit $a_1,\dots,a_n \in K^m$, sowie den Vektor $x \in K^n$ als $x = (\lambda_j)_j$, so ist das dort auftretende Matrixprodukt Ax nach Definition gleich

$$Ax = (a_1 \,|\, \cdots \,|\, a_n) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 a_1 + \cdots + \lambda_n a_n.$$

Formulieren wir die Konzepte vom Bild und Kern einer Matrix mit Hilfe dieser Spaltenvektoren $a_1,\dots,a_n$, so erhalten wir die folgenden sehr wichtigen Begriffe.

Definition 14.21 (Linearkombinationen und lineare Abhängigkeit)**.** Es seien $a_1,\dots,a_n \in K^m$ gegebene Vektoren und $A := (a_1 \,|\, \cdots \,|\, a_n) \in K^{m\times n}$.

(a) Für $\lambda_1,\dots,\lambda_n \in K$ heißt der Vektor

$$\lambda_1 a_1 + \cdots + \lambda_n a_n \quad \in K^m$$

eine **Linearkombination** von $a_1,\dots,a_n$. Die Menge aller dieser Linearkombinationen bezeichnen wir mit $\operatorname{Lin}(a_1,\dots,a_n)$; nach Definition 14.19 (b) und Bemerkung 14.20 ist offensichtlich $\operatorname{Lin}(a_1,\dots,a_n) = \operatorname{Im} A$.

(b) Die Vektoren $a_1,\dots,a_n$ heißen **linear abhängig**, wenn man aus ihnen eine Linearkombination des Nullvektors

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0 \qquad (*)$$

bilden kann, in der mindestens ein $\lambda_1,\dots,\lambda_n$ ungleich 0 ist (man nennt dies auch eine *nichttriviale Linearkombination* des Nullvektors).

Ist das Gegenteil der Fall, folgt aus der Linearkombination $(*)$ des Nullvektors also bereits, dass $\lambda_1 = \cdots = \lambda_n = 0$ gelten muss, so heißen die Vektoren $a_1,\dots,a_n$ **linear unabhängig**. In der Matrixschreibweise von Definition 14.19 (c) und Bemerkung 14.20 bedeutet dies offensichtlich genau $\operatorname{Ker} A = \{0\}$.

Beispiel 14.22.

(a) Die Einheitsvektoren $e_1, \dots, e_n \in K^n$ sind linear unabhängig: Sind nämlich $\lambda_1, \dots, \lambda_n \in K$ mit

$$\lambda_1 e_1 + \cdots + \lambda_n e_n = 0, \quad \text{also} \quad \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

so folgt daraus natürlich sofort $\lambda_1 = \cdots = \lambda_n = 0$.

(b) Enthalten die Vektoren $a_1, \dots, a_n \in K^m$ den Nullvektor, also ist $a_i = 0$ für ein $i \in \{1, \dots, n\}$, so sind diese Vektoren immer linear abhängig, denn dann ist ja $1 \cdot a_i = 0$ eine nicht-triviale Linearkombination des Nullvektors.

Ebenso sind die Vektoren $a_1, \dots, a_n$ immer linear abhängig, wenn ein Vektor unter ihnen mehrfach vorkommt, also wenn $a_i = a_j$ für gewisse $i, j \in \{1, \dots, n\}$ mit $i \neq j$ gilt: Dann ist nämlich $1 \cdot a_i - 1 \cdot a_j = 0$ eine nicht-triviale Linearkombination des Nullvektors.

(c) Es seien $a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $a_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $a_3 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \in \mathbb{R}^2$.

Dann sind a_1 und a_2 linear unabhängig, denn aus

$$\lambda_1 a_1 + \lambda_2 a_2 = 0 \quad \Rightarrow \quad \lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \lambda_1 + 2\lambda_2 = \lambda_1 + \lambda_2 = 0$$

folgt sofort $\lambda_1 = \lambda_2 = 0$. Genauso sieht man, dass a_1 und a_3 linear unabhängig sind, und ebenso a_2 und a_3.

Aber alle drei Vektoren a_1, a_2, a_3 zusammen sind linear abhängig, denn es gilt

$$1 \cdot a_1 - 2 \cdot a_2 + 1 \cdot a_3 = 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Die lineare Unabhängigkeit von Vektoren $a_1, \dots, a_n$ ist also eine Eigenschaft *aller dieser Vektoren zusammen* – wir können sie nicht überprüfen, indem wir uns die Vektoren nacheinander einzeln oder in Paaren anschauen.

Bemerkung 14.23 (Untervektorräume und Dimension)**.** Die Mengen $\mathrm{Lin}(a_1, \dots, a_n)$ aus Definition 14.21 (a) bestehen nach Definition aus allen skalaren Vielfachen und Summen der gegebenen Vektoren $a_1, \dots, a_n$. Gemäß Beispiel 13.4 (b) können wir sie uns damit vorstellen als Geraden, Ebenen oder entsprechende „höherdimensionale Mengen" durch den Nullpunkt. Solche Mengen werden wir in **??** als Untervektorräume kennenlernen und noch genau untersuchen.

Sind die Vektoren $a_1, \dots, a_n$ dabei linear unabhängig, so bedeutet dies anschaulich, dass sie „in unabhängige Richtungen zeigen" und damit „einen n-dimensionalen Raum aufspannen" – auch diesen Dimensionsbegriff werden wir in **??** noch exakt einführen. Für den Moment wollen wir nur festhalten, dass jeder Vektor in $\mathrm{Lin}(a_1, \dots, a_n)$ dann *auf eindeutige Art* eine Linearkombination der gegebenen Vektoren $a_1, \dots, a_n$ ist, so dass wir uns $\mathrm{Lin}(a_1, \dots, a_n)$ also als eine „Menge mit n unabhängigen Parametern" vorstellen können:

Lemma 14.24 (Eindeutigkeit von Linearkombinationen)**.** *Es seien $a_1, \dots, a_n \in K^m$ linear unabhängige Vektoren. Dann gibt es zu jedem Vektor $v \in \mathrm{Lin}(a_1, \dots, a_n)$ eindeutig bestimmte $\lambda_1, \dots, \lambda_n \in K$ mit $v = \lambda_1 a_1 + \cdots + \lambda_n a_n$.*

Beweis. Es seien $\lambda_1, \dots, \lambda_n \in K$ sowie $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n \in K$ mit

$$v = \lambda_1 a_1 + \cdots + \lambda_n a_n = \tilde{\lambda}_1 a_1 + \cdots + \tilde{\lambda}_n a_n.$$

Daraus folgt natürlich sofort

$$(\lambda_1 - \tilde{\lambda}_1) a_1 + \cdots + (\lambda_n - \tilde{\lambda}_n) a_n = 0,$$

wegen der linearen Unabhängigkeit der Vektoren also $\lambda_i - \tilde{\lambda}_i = 0$ und damit auch $\lambda_i = \tilde{\lambda}_i$ für alle $i \in \{1, \dots, n\}$. □

Wir kommen nun wie angekündigt zur Hauptaussage über die Lösungsmengen linearer Gleichungssysteme. Genau wie beim Gauß-Algorithmus ist auch der Beweis dieses Satzes konstruktiv – er gibt also auch an, wie die dort mit $v_1,\dots,v_{n-r}$ und v bezeichneten Vektoren konkret berechnet werden können.

Satz 14.25 (Lösungsmengen linearer Gleichungssysteme)**.** *Es seien $A \in K^{m\times n}$ eine Matrix vom Rang $r := \operatorname{rk} A$ und $b \in K^m$.*

(a) *Es gibt $n-r$ linear unabhängige Vektoren $v_1,\dots,v_{n-r} \in K^n$ mit*

$$L(A,0) = \operatorname{Ker} A = \operatorname{Lin}(v_1,\dots,v_{n-r}).$$

(b) *Das Gleichungssystem $Ax = b$ hat genau dann eine Lösung $v \in L(A,b)$, wenn die sogenannte* erweiterte Koeffizientenmatrix *$(A\,|\,b) \in K^{m\times(n+1)}$ ebenfalls den Rang r hat.*

(c) *Ist $v \in L(A,b)$ eine Lösung des Gleichungssystems $Ax = b$, so ist die gesamte Lösungsmenge gegeben durch $L(A,b) = v + \operatorname{Ker} A$.*

Insbesondere ist das lineare Gleichungssystem $Ax = b$ also genau dann eindeutig lösbar, wenn $\operatorname{rk}(A\,|\,b) = \operatorname{rk} A = n$ gilt.

Beweis. Wir bringen die erweiterte Koeffizientenmatrix $(A\,|\,b)$ mit dem Gauß-Algorithmus auf reduzierte Zeilenstufenform $S(A\,|\,b) = (SA\,|\,Sb)$ mit $S \in \operatorname{GL}(m,K)$. Dies ändert weder den Rang der betrachteten Matrizen (siehe Lemma 14.14 (a)) noch die gesuchten Lösungsmengen (wegen $Ax = b \Leftrightarrow SAx = Sb$). Das neue Gleichungssystem hat nun die erweiterte Koeffizientenmatrix

$$\begin{array}{c}
\begin{array}{ccccccc}
k_1 & & k_2 & & & k_r & \\
\downarrow & & \downarrow & & & \downarrow &
\end{array} \\
\left(\begin{array}{ccccccc|c}
1 & * \cdots * & 0 & * \cdots * & & 0 & * \cdots * & b_1 \\
 & & 1 & * \cdots * & & 0 & * \cdots * & b_2 \\
 & & & & \ddots & \vdots & \vdots & \vdots \\
 & \mathbf{0} & & & & 1 & * \cdots * & b_r \\
 & & & & & & & \vdots \\
 & & & & & & & b_m
\end{array}\right)
\end{array}, \tag{1}$$

wobei $(b_i)_i := Sb$ ist und „$*$“ die Einträge von $(a_{i,j})_{i,j} := SA$ bezeichnet, die ungleich 0 sein können. Wir bezeichnen die Stufenspalten von SA mit $k_1 \le \dots \le k_r$, und die Nichtstufenspalten von SA mit $j_1 \le \dots \le j_{n-r}$.

(a) Ist $b = 0$, so ist die i-te Zeile des umgeformten Gleichungssystems für $i \in \{1,\dots,r\}$

$$x_{k_i} + \sum_{l=1}^{n-r} a_{i,j_l} x_{j_l} = 0 \quad \Leftrightarrow \quad x_{k_i} = -\sum_{l=1}^{n-r} a_{i,j_l} x_{j_l}. \tag{2}$$

Die Gleichungen bestimmen also die Stufenvariablen $x_{k_1},\dots,x_{k_r}$ eindeutig aus den Nichtstufenvariablen $x_{j_1},\dots,x_{j_{n-r}}$. Wir können daher $\lambda_l := -x_{j_l} \in K$ für $l = 1,\dots,n-r$ als freie Parameter wählen, und erhalten damit aus (2) als Vektor geschrieben die allgemeine Lösung

$$x = \lambda_1 v_1 + \dots + \lambda_{n-r} v_{n-r}, \tag{3}$$

wobei v_l der Vektor ist mit j_l-ter Koordinate -1 und k_i-ter Koordinate a_{i,j_l} für $i = 1,\dots,r$ (alle anderen Koordinaten sind 0).

Damit ist also $L(A,0) = \operatorname{Lin}(v_1,\dots,v_{n-r})$. Außerdem ergibt sich in (3) für x nur dann der Nullvektor, wenn alle $\lambda_l = -x_{j_l}$ gleich 0 sind. Damit sind $v_1,\dots,v_{n-r}$ auch linear unabhängig.

(b) Die Matrix $(SA\,|\,Sb)$ in reduzierter Zeilenstufenform in (1) kann nur Rang r oder $r+1$ haben.

Hat sie Rang $r+1$, so ist $b_{r+1} = 1$, und die $(r+1)$-te Zeile des umgeformten Gleichungssystems liefert den Widerspruch $0 = 1$. In diesem Fall ist also $L(A,b) = \emptyset$.

Hat sie Rang r, so erhalten wir offensichtlich eine Lösung, indem wir die Nichtstufenvariablen $x_{j_1}, \dots, x_{j_{n-r}}$ gleich 0 setzen, und die Stufenvariablen gleich $x_{k_i} = b_i$ für $i \in \{1, \dots, r\}$.

(c) Ist $Av = b$, so ist für ein $x \in K^n$ genau dann $Ax = b$, wenn $A(x-v) = 0$ gilt, also $x - v \in \operatorname{Ker} A$ und damit $x \in v + \operatorname{Ker} A$ ist. □

Bemerkung 14.26.

(a) Die konkrete Form der Vektoren $v_1, \dots, v_{n-r}$ und v in Satz 14.25 lässt sich gut zu einer Merkregel zusammenfassen: Nach der Berechnung der reduzierten Zeilenstufenform $(SA\,|\,Sb) \in K^{m\times(n+1)}$ bilden wir eine neue Matrix $C \in K^{n\times(n+1)}$, bei der wir für alle $i \in \{1, \dots, r\}$ die Zeile i von $(SA\,|\,Sb)$ in die Zeile k_i von C schreiben. Die übrigen Einträge von C setzen wir gleich -1 auf der Diagonalen und 0 sonst.

Nach dem Beweis von (a) stehen dann in den Spalten $j_1, \dots, j_{n-r}$ von C, die den Nichtstufenspalten von SA entsprechen, die Vektoren $v_1, \dots, v_{n-r}$: In Spalte j_l steht in Zeile k_i der Eintrag a_{i,j_l}, in Zeile j_l der Eintrag -1, und sonst überall 0.

In der letzten Spalte steht nach dem Beweis von (b) außerdem der Vektor v.

(b) Falls wir nur ein Gleichungssystem $Ax = 0$ mit rechter Seite 0 lösen (also $\operatorname{Ker} A$ bestimmen) wollen, müssen wir in (a) die rechte Spalte für b gar nicht erst hinschreiben, da die Einträge dort ohnehin alle gleich 0 sind und während des Verfahrens auch bleiben. In diesem Fall ist das Gleichungssystem natürlich auch in jedem Fall lösbar mit $x = 0$.

(c) Alternativ zum Verfahren im Beweis des Satzes kann man die erweiterte Koeffizientenmatrix $(A\,|\,b)$ auch nur auf nicht notwendig reduzierte Zeilenstufenform bringen (was natürlich weniger Rechenarbeit ist). Im Fall der Lösbarkeit bestimmt dann die i-te Gleichung des umgeformten Gleichungssystems für $i \in \{1, \dots, r\}$ die Stufenvariable x_{k_i} aus den Nichtstufenvariablen *und den weiter rechts stehenden Stufenvariablen* $x_{k_{i+1}}, \dots, x_{k_r}$. Wir müssen diese Gleichungen dann also von unten nach oben durchgehen und so der Reihe nach $x_{k_r}, \dots, x_{k_1}$ aus den freien Variablen *und den bereits bestimmten Stufenvariablen* berechnen (was dann wieder etwas mehr Arbeit ist).

Beispiel 14.27. Wir wollen $\operatorname{Ker} A = L(A, 0)$ und $L(A, b)$ für

$$A = \begin{pmatrix} 1 & 4 & 0 & 5 \\ 2 & 8 & 1 & 7 \\ 0 & 0 & 2 & -6 \end{pmatrix} \in \mathbb{R}^{3\times 4} \quad \text{und} \quad b = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \in \mathbb{R}^3$$

bestimmen. Dazu bringen wir $(A\,|\,b)$ auf reduzierte Zeilenstufenform:

$$\left(\begin{array}{cccc|c} 1 & 4 & 0 & 5 & 1 \\ 2 & 8 & 1 & 7 & 3 \\ 0 & 0 & 2 & -6 & 2 \end{array}\right) \xrightarrow{Z2-2Z1\to Z2} \left(\begin{array}{cccc|c} 1 & 4 & 0 & 5 & 1 \\ 0 & 0 & 1 & -3 & 1 \\ 0 & 0 & 2 & -6 & 2 \end{array}\right) \xrightarrow{Z3-2Z2\to Z3} \left(\begin{array}{cccc|c} 1 & 4 & 0 & 5 & 1 \\ 0 & 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Das gegebene Gleichungssystem ist also äquivalent zu

$$\begin{aligned} x_1 + 4x_2 \quad + 5x_4 &= 1, \\ x_3 - 3x_4 &= 1, \\ 0 &= 0. \end{aligned}$$

Offensichtlich lässt sich dieses umgeformte Gleichungssystem nun leicht lösen: Die Variablen x_2 und x_4 können wir frei wählen; die erste Gleichung bestimmt daraus dann eindeutig x_1 und die zweite x_3.

Um die so entstehende Lösung mit dem Verfahren aus Bemerkung 14.26 (a) direkt hinschreiben zu können, schreiben wir die beiden Zeilen der umgeformten Matrix wie unten grau hinterlegt in die Zeilen 1 und 3 einer 4×5-Matrix C (da die Stufen in den Spalten 1 und 3 liegen), und füllen die restlichen Einträge mit -1 auf der Diagonalen und 0 sonst. Da die Nichtstufenspalten der obigen Zeilenstufenform die zweite und vierte sind, lesen wir dann die Vektoren v_1, v_2 und v in Satz 14.25 direkt in der Matrix C in den eingekreisten Spalten 2 und 4 bzw. der letzten Spalte ab: Mit

$$C = \left(\begin{array}{ccccc|c} 1 & 4 & 0 & 5 & 1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right) \quad \Rightarrow \quad v_1 = \begin{pmatrix} 4 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \; v_2 = \begin{pmatrix} 5 \\ 0 \\ -3 \\ -1 \end{pmatrix}, \; v = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

ist $\operatorname{Ker} A = \operatorname{Lin}(v_1, v_2)$ (mit linear unabhängigen Vektoren v_1 und v_2) und $L(A,b) = v + \operatorname{Lin}(v_1, v_2)$.

Zum Abschluss dieses Kapitels wollen wir noch einige einfache, aber sehr wichtige Folgerungen aus unseren Ergebnissen festhalten. Die erste betrifft die sogenannte universelle Lösbarkeit linearer Gleichungssysteme – d. h. die Frage nach der Lösbarkeit eines solchen Gleichungssystems $Ax = b$ für eine beliebige rechte Seite b.

Folgerung 14.28 (Universelle Lösbarkeit linearer Gleichungssysteme). *Für jede Matrix $A \in K^{m \times n}$ gilt:*

(a) $\operatorname{rk} A = n \;\Leftrightarrow\;$ *Für alle $b \in K^m$ hat das Gleichungssystem $Ax = b$ höchstens eine Lösung.*

(b) $\operatorname{rk} A = m \;\Leftrightarrow\;$ *Für alle $b \in K^m$ hat das Gleichungssystem $Ax = b$ mindestens eine Lösung.*

Beweis.

(a) Nach Satz 14.25 (a) ist $\operatorname{rk} A = n$ genau dann, wenn $\operatorname{Ker} A = \{0\}$ gilt, also wenn das Gleichungssystem $Ax = 0$ genau die Lösung $x = 0$ hat. Nach Satz 14.25 (c) hat dann auch jedes Gleichungssystem $Ax = b$ höchstens eine Lösung.

(b) „$\Rightarrow$“: Ist $\operatorname{rk} A = m$, so ist wegen $(A\,|\,b) \in K^{m \times (n+1)}$ auch $\operatorname{rk}(A\,|\,b) = m$ nach Bemerkung 14.9 (a). Damit ist $L(A,b) \neq \emptyset$ nach Satz 14.25 (b).

„$\Leftarrow$“: Ist $\operatorname{rk} A < m$, so hat die reduzierte Zeilenstufenform SA von A unten (mindestens) eine Nullzeile. Damit hat das Gleichungssystem $SAx = e_m \Leftrightarrow Ax = S^{-1} e_m$ dann keine Lösung. □

Bemerkung 14.29. Nach Folgerung 14.28 hat das Gleichungssystem $Ax = b$ mit $A \in K^{m \times n}$ und $b \in K^m$ genau dann für alle b eine eindeutige Lösung, wenn $\operatorname{rk} A = m = n$, also wenn A quadratisch und invertierbar ist (siehe Lemma 14.11). Diesen einfachen Fall hatten wir schon in Folgerung 13.17 (d) betrachtet: Die eindeutige Lösung ist dann $x = A^{-1} b$.

Unsere Ergebnisse zeigen auch, dass die lineare Unabhängigkeit von Vektoren $a_1, \dots, a_n \in K^m$ eng mit dem Rang der daraus gebildeten Matrix $(a_1\,|\cdots|\,a_n) \in K^{m \times n}$ zusammenhängt. In der Tat liefert Folgerung 14.31 zusammen mit der geometrischen Interpretation der linearen Unabhängigkeit aus Bemerkung 14.23 vermutlich eine der anschaulichsten Arten, sich den Rang einer Matrix (ohne Bezug auf algorithmische Konzepte wie den Gauß-Algorithmus oder eine Zeilenstufenform) anschaulich vorzustellen.

Folgerung 14.30 (Rangkriterium für lineare Unabhängigkeit). *Es seien $a_1, \dots, a_n \in K^m$.*

(a) *Die Vektoren $a_1, \dots, a_n$ sind genau dann linear unabhängig, wenn $\operatorname{rk}(a_1\,|\cdots|\,a_n) = n$ gilt.*

(b) *Für $n > m$ sind die Vektoren $a_1, \dots, a_n$ in jedem Fall linear abhängig.*

Beweis. Es sei $A = (a_1\,|\cdots|\,a_n) \in K^{m \times n}$ mit Rang $r := \operatorname{rk} A$.

(a) Die Vektoren $a_1, \dots, a_n$ sind nach Definition 14.21 (b) genau dann linear unabhängig, wenn $\operatorname{Ker} A = \{0\}$. Dies ist nach Satz 14.25 (a) äquivalent zu $r = n$.

(b) Aus $n > m$ folgt mit Bemerkung 14.9 (a) sofort $\operatorname{rk} A \le m < n$. Nach (a) sind $a_1, \dots, a_n$ dann also linear abhängig. □

Folgerung 14.31 (Rang = Maximalzahl linear unabhängiger Spalten). *Für jede Matrix $A \in K^{m \times n}$ ist $\operatorname{rk} A$ die Maximalzahl linear unabhängiger Spalten in A.*

Beweis. Für beliebige Spalten $1 \le k_1 \le \cdots \le k_s \le n$ sei $A' = (a_{k_1}\,|\cdots|\,a_{k_s}) \in K^{m \times s}$ die Teilmatrix von A, die durch Auswahl dieser s Spalten entsteht.

Wir bringen A nun auf Zeilenstufenform $SA = (Sa_1 \,|\, \cdots \,|\, Sa_n)$ mit $r := \operatorname{rk} A$ Stufen, und betrachten die Matrix

$$SA' = (Sa_{k_1} \,|\, \cdots \,|\, Sa_{k_s}) \quad \in K^{m \times s},$$

die ebenfalls durch Auswahl der entsprechenden s Spalten in SA entsteht:

- Ist $s = r$ und wählen wir für $k_1, \dots, k_r$ genau die Stufenspalten, so ist SA' ebenfalls in Zeilenstufenform mit r Stufen. Also ist dann $\operatorname{rk} A' = r = s$, und damit sind die s ausgewählten Spalten von A nach Folgerung 14.30 (a) linear unabhängig.
- Ist $s > r$, so hat SA' genau wie SA höchstens r Zeilen mit Einträgen ungleich 0, und damit folgt $\operatorname{rk} A' \leq r < s$ aus Beispiel 14.10 (c). Also sind die s ausgewählten Spalten nach Folgerung 14.30 (a) diesmal linear abhängig.

Insgesamt ist r damit die Maximalzahl linear unabhängiger Spalten in A. □

Bemerkung 14.32. Der Beweis von Folgerung 14.31 ist in dem Sinne konstruktiv, dass er auch angibt, wie man r linear unabhängige Spalten aus einer Matrix A vom Rang r auswählen kann: Man bringt A auf Zeilenstufenform und wählt dann die Spalten der Ausgangsmatrix A, die zu den Stufenspalten gehören. So haben wir z. B. für die Matrix

$$A = \begin{pmatrix} -1 & 2 & 4 \\ -1 & 3 & 7 \\ 1 & 1 & 5 \end{pmatrix} = (a_1 \,|\, a_2 \,|\, a_3) \in \mathbb{R}^{3 \times 3} \quad \text{mit} \quad a_1 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \, a_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \, a_3 = \begin{pmatrix} 4 \\ 7 \\ 5 \end{pmatrix}$$

in Beispiel 14.5 eine Zeilenstufenform mit 2 Stufen und Stufenspalten 1 und 2 (und damit $\operatorname{rk} A = 2$) bestimmt. Dementsprechend sind die Vektoren a_1 und a_2 also nach Folgerung 14.31 linear unabhängig, während alle drei Vektoren a_1, a_2, a_3 linear abhängig sind.

Wir haben in Folgerung 14.30 (b) auch gesehen, dass in K^m bis zu maximal m Vektoren linear unabhängig sein können. Wenn wir also $n < m$ linear unabhängige Vektoren $a_1, \dots, a_n$ in K^m haben (so wie hier im Beispiel a_1 und a_2 in K^3), können wir uns daher fragen, ob wir diese Vektoren immer zu m linear unabhängigen Vektoren ergänzen können. Wie wir jetzt sehen werden, ist dies in der Tat immer möglich – sogar mit Einheitsvektoren, und wiederum mit einem konstruktiven Verfahren.

Lemma 14.33 (Ergänzung linear unabhängiger Vektoren)**.** *Es seien $a_1, \dots, a_n \in K^m$ linear unabhängige Vektoren. Dann gibt es Einheitsvektoren $e_{j_1}, \dots, e_{j_{m-n}} \in K^m$, so dass die m Vektoren $a_1, \dots, a_n, e_{j_1}, \dots, e_{j_{m-n}}$ linear unabhängig sind.*

Beweis. Wir setzen wieder $A = (a_1 \,|\, \cdots \,|\, a_n) \in K^{m \times n}$; nach Folgerung 14.30 (a) ist also $\operatorname{rk} A = n$. Diesmal bringen wir A mit elementaren *Spalten*umformungen auf eine Spaltenstufenform mit n Stufen. Sind nun $j_1, \dots, j_{m-n}$ die Nichtstufenzeilen dieser Spaltenstufenform, so können wir aus A durch Hinzufügen von Spalten mit den Einheitsvektoren $e_{j_1}, \dots, e_{j_{m-n}}$ eine $m \times m$-Matrix in Spaltenstufenform mit m Stufen erzeugen. Da diese Matrix dann Rang m hat, sind die m Vektoren $a_1, \dots, a_n, e_{j_1}, \dots, e_{j_{m-n}}$ damit nach Folgerung 14.30 (a) linear unabhängig. □

Beispiel 14.34. Wir hatten in Bemerkung 14.32 gesehen, dass die dort angegebenen Vektoren $a_1, a_2 \in \mathbb{R}^3$ linear unabhängig sind. Wollen wir a_1 und a_2 durch Hinzunahme eines weiteren Vektors zu drei linear unabhängigen Vektoren ergänzen, bringen wir die Matrix $A = (a_1 \,|\, a_2)$ wie in Lemma 14.33 auf Spaltenstufenform mit 2 Stufen:

$$A = \begin{pmatrix} -1 & 2 \\ -1 & 3 \\ 1 & 1 \end{pmatrix} \overset{-S1 \to S1}{\longrightarrow} \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ -1 & 1 \end{pmatrix} \overset{S2-2S1 \to S2}{\longrightarrow} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 3 \end{pmatrix}.$$

Da die dritte Zeile hierbei keine Stufenzeile ist, sind die Vektoren a_1, a_2, e_3 linear unabhängig.

Beachte, dass dies nicht die einzige Möglichkeit wäre: In diesem Beispiel könnte man leicht nachrechnen, dass auch a_1, a_2, e_1 und a_1, a_2, e_2 linear unabhängig sind. Der Algorithmus aus Lemma 14.33 liefert nur eine mögliche Art, die gegebenen linear unabhängigen Vektoren mit Einheitsvektoren zu maximal vielen linear unabhängigen Vektoren zu ergänzen.