

Kapitel 0

Lineare Gleichungssysteme

Schon die Nummer dieses Kapitels deutet an, daß es zur Motivation und Vorbereitung der in Kapitel 1 beginnenden systematischen Darstellung dient. Wir haben dafür das wichtigste Problem der elementaren linearen Algebra gewählt, nämlich lineare Gleichungssysteme. Dabei kann man sehr schön die wesentlichen Aspekte vorführen: den geometrischen Hintergrund und die algorithmische Methode. Was auf die späteren Kapitel verschoben wird, sind die präzisen Beweise mit Hilfe der üblichen theoretischen Hilfsmittel.

Wer mit den verwendeten Notationen von Mengen und Abbildungen nicht vertraut ist, kann bei Bedarf in 1.1 nachsehen.

0.1 Der reelle n -dimensionale Raum

Ein großer Fortschritt in der Geometrie gelang durch Einführung von Koordinaten, die man zur Erinnerung an R. DESCARTES auch *kartesische* Koordinaten nennt. Dadurch kann man Punkte in den Räumen der Geometrie beschreiben durch Systeme von Zahlen, und mit den Zahlen kann man rechnen. Diese Methode zur Behandlung geometrischer Fragen nennt man *analytische Geometrie*. Die elementarsten Begriffe hierzu seien kurz erklärt.

0.1.1. Wir gehen aus von den *reellen Zahlen*, deren Gesamtheit wir mit $\mathbb{R}$ bezeichnen. Ihre Einführung ist Gegenstand der Analysis, in der Geometrie dienen sie als *Zahlengerade*, und diese Zahlen kann man nach den üblichen Regeln addieren und multiplizieren.

Punkte der Ebene sind festgelegt durch Paare, Punkte des gewöhnlichen Raumes durch Tripel von reellen Zahlen. Für die Theorie macht es keine Probleme, gleich n -Tupel zu betrachten, wobei n eine beliebige natürliche Zahl ist. Damit erhält man den *reellen Standardraum der Dimension n*

$$\mathbb{R}^n = \{x = (x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\},$$

d.h. die Menge der *geordneten n -Tupel* (oder *Vektoren*) von reellen Zahlen. Geordnet heißt, daß die Reihenfolge wichtig ist, d.h. zwei n -Tupel $(x_1, \dots, x_n)$ und $(y_1, \dots, y_n)$ sind genau dann gleich, wenn $x_1 = y_1, \dots, x_n = y_n$. Die Zahlen $x_1, \dots, x_n$ heißen *Komponenten* von x .

Der Fall $n = 0$ ist sinnlos, $\mathbb{R}^1$ ist die Zahlengerade, $\mathbb{R}^2$ entspricht der Ebene und $\mathbb{R}^3$ dem „Raum“. Für größere n hat man zunächst keine unmittelbare geometrische Vorstellung mehr, dafür aber eine ganz andere und sehr realistische Interpretation. Hat etwa eine Bank n Kunden, so kann man deren Kontostände zu einem bestimmten Zeitpunkt mit $x_1, \dots, x_n$ bezeichnen, alle zusammen (und

geordnet!) sind ein „Punkt“

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Die Entwicklung der Kontostände im Laufe der Zeit wird dann durch eine „Kurve“ im $\mathbb{R}^n$ beschrieben, ihre Beschreibung geht schon über den Rahmen der linearen Algebra hinaus. Eine lineare Operation ist etwa die Berechnung der augenblicklichen Bilanz. Haben die Einlagen neben dem Nennwert x_i einen Börsenkurs a_i , so ist das bewertete Kapital gegeben durch

$$a_1 x_1 + \dots + a_n x_n = b,$$

ein typischer Fall für eine *lineare Gleichung*.

In der Praxis hat man mehrere Bedingungen, und diese sind meist nicht durch Gleichungen, sondern durch Ungleichungen gegeben, von der Art, daß für die obige Summe Begrenzungen vorgegeben sind. Und das Problem besteht darin, einen „Gewinn“ zu optimieren. Zur Lösung solcher Aufgaben in der *linearen Optimierung* (vgl. etwa [Fi]) benötigt man genügende Kenntnisse über lineare Gleichungen, was vielleicht auch einen gewinnorientierten Leser eine Zeit lang bei der Stange halten kann.

0.1.2. In der linearen Algebra muß man mit n -Tupeln rechnen. Die grundlegenden Operationen sind eine *Addition*

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

und eine *Multiplikation mit einer Zahl* $\lambda \in \mathbb{R}$

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda \cdot x_1, \dots, \lambda \cdot x_n).$$

Man kann diese Operationen geometrisch deuten, wenn man die n -Tupel als Vektoren ansieht, d.h. naiv als Pfeile vom *Ursprung* $0 = (0, \dots, 0)$ mit Spitze in $x = (x_1, \dots, x_n)$. Für $n = 2$ kann man das einfach zeichnen:

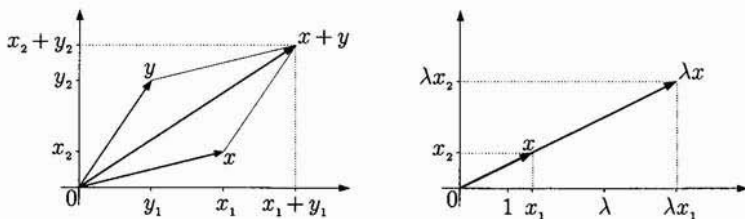


Bild 0.1

Der Ursprung selbst heißt auch *Nullvektor*, wenn man ihn addiert, hat das keine Wirkung. Multipliziert man mit $\lambda = 0$, so wird jedes x zum Nullvektor. Das Negative von x ist gegeben durch

$$-x := (-x_1, \dots, -x_n),$$

es gilt $x + (-x) = 0$. Statt $x + (-y)$ schreibt man kürzer $x - y$.

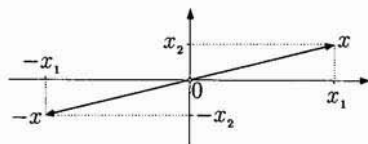


Bild 0.2

Nach diesen wenigen Formalitäten können wir nun die einfachsten Beispiele von linearen Gleichungen behandeln. Um die geometrische Anschauung dabei zu benutzen, betrachten wir zunächst ausführlich die Fälle $n = 2$ und $n = 3$.

0.2 Geraden in der Ebene

0.2.1. Durch zwei verschiedene Punkte geht genau eine *Gerade*, das gehört zu den wenigen Tatsachen der Geometrie, die auch Nicht-Mathematikern einleuchten. Mit Hilfe von Vektoren kann man das so beschreiben: Sind $v, v' \in \mathbb{R}^2$ die beiden Punkte, so ist $w := v' - v \neq 0$. Die Punkte auf der Geraden L durch v und v' kann man unter Benutzung eines reellen *Parameters* λ darstellen als

$$L = \{x \in \mathbb{R}^2 : \text{es gibt ein } \lambda \in \mathbb{R} \text{ mit } x = v + \lambda w\} =: v + \mathbb{R}w.$$

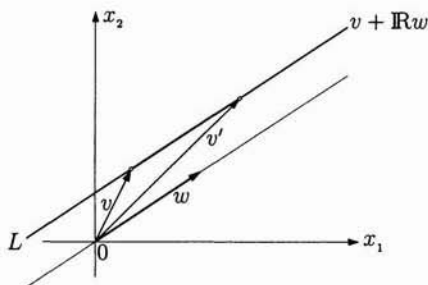


Bild 0.3

Man kann L auch ansehen als Bild der Zahlengerade $\mathbb{R}$ unter der Abbildung

$$\Phi: \mathbb{R} \rightarrow L \subset \mathbb{R}^2, \quad \lambda \mapsto v + \lambda w.$$

Das nennt man eine *Parametrisierung* der Geraden.

0.2.2. Die zweite Möglichkeit der Beschreibung benutzt eine *lineare Gleichung* der Form

$$a_1 x_1 + a_2 x_2 = b.$$

Dabei gelten x_1, x_2 als *Unbestimmte* und $a_1, a_2 \in \mathbb{R}$ als *Koeffizienten*. Die Unbestimmten sind variabel, die Koeffizienten fest. Man betrachtet die Menge der *Lösungen*

$$L := \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\}.$$

Ist $a_1 = a_2 = 0$, so ist $L = \emptyset$ für $b \neq 0$ und $L = \mathbb{R}^2$ für $b = 0$. Dieser Fall gilt als „entartet“. Andernfalls müßte man im Prinzip alle Paare (x_1, x_2) in die Gleichung einsetzen und feststellen, ob sie erfüllt ist. Wenn man das ohne System tut, wird man dabei sehr selten Glück haben.

Ein gutes System ist eine Parametrisierung, mit deren Hilfe sich alle Lösungen produzieren lassen. Das ist in diesem Fall leicht zu erhalten.

1) Ist $a_2 = 0$ und $a_1 \neq 0$, so wird die gegebene Gleichung zu

$$x_1 = \frac{b}{a_1},$$

das ist eine zur x_2 -Achse parallele Gerade, und eine Parametrisierung ist gegeben durch

$$\mathbb{R} \rightarrow L, \quad \lambda \mapsto \left(\frac{b}{a_1}, \lambda \right).$$

Hier ist also die erste Koordinate fest, die zweite frei wählbar.

Ist $a_1 = 0$, aber $a_2 \neq 0$, so hat man nur die Rollen von x_1 und x_2 zu vertauschen.

2) Ist $a_1 \neq 0$ und $a_2 \neq 0$, so kann man die Gerade leicht zeichnen, indem man die Schnittpunkte mit den Achsen $x_1 = 0$ und $x_2 = 0$ berechnet:

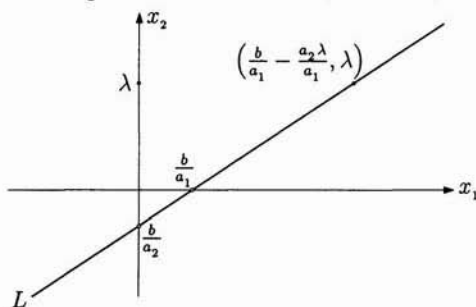


Bild 0.4

Wählt man wieder die x_2 -Koordinate eines Punktes der Geraden als Parameter λ , so kann man daraus

$$x_1 = \frac{b}{a_1} - \frac{a_2 \lambda}{a_1}$$

berechnen, und eine Parametrisierung der zunächst durch die Gleichung gegebenen Geraden ist gefunden:

$$\mathbb{R} \rightarrow L, \quad \lambda \mapsto \left(\frac{b}{a_1} - \frac{a_2 \lambda}{a_1}, \lambda \right).$$

0.2.3. Zwei Geraden in der Ebene schneiden sich in genau einem Punkt, es sei denn, sie sind gleich oder parallel. Sind sie durch Gleichungen gegeben, so stellt sich die Frage, wie man entscheiden kann, welcher Fall vorliegt, und wie man eventuell den eindeutigen Schnittpunkt findet. Dazu einige

Beispiele. a) Die Geraden seien gegeben durch

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

Der Schnittpunkt p ist ganz einfach zu finden, indem man $x_2 = 2$ aus der zweiten Gleichung in die erste einsetzt: $x_1 = 1 + 2 = 3$, also $p = (3, 2)$.

Eine Variante sind die Gleichungen

$$\begin{aligned}x_1 - x_2 &= 1, \\x_1 + 3x_2 &= 9.\end{aligned}$$

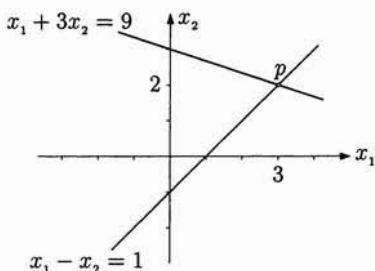
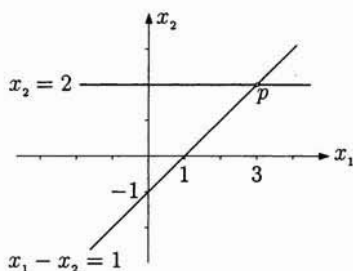


Bild 0.5

Zieht man die erste Gleichung von der zweiten ab und dividiert die Differenz durch vier, so erhält man wieder die obigen Gleichungen, und man sieht an den Zeichnungen, daß der Schnittpunkt der gleiche ist.

b) Die Geraden seien gegeben durch

$$\begin{aligned}x_1 - x_2 &= 1, \\2x_1 - 2x_2 &= b,\end{aligned}$$

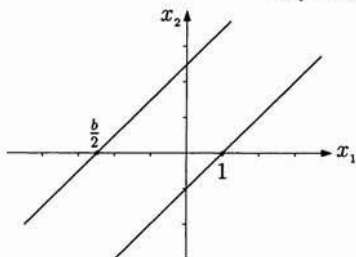


Bild 0.6

mit beliebigem b . Man sieht sofort, daß sie für $b = 2$ gleich und für $b \neq 2$ parallel, also ohne Schnittpunkt sind. Darauf kommt man auch durch formales

Rechnen, wenn man wieder die 2-fache erste Gleichung von der zweiten abzieht. Das ergibt

$$\begin{aligned}x_1 - x_2 &= 1, \\ 0x_1 - 0x_2 &= b - 2.\end{aligned}$$

Die zweite Gleichung lautet in jedem Fall $b = 2$. Ist b so gewählt, kann man sie weglassen und es bleibt die erste. Ist $b \neq 2$ gewählt, so ist die zweite Gleichung nie erfüllt und man kann die erste weglassen.

c) Nun nehmen wir zu den zwei Geraden aus Beispiel a) eine dritte hinzu:

$$\begin{aligned}x_1 - x_2 &= 1, & \text{I} \\ x_1 + 3x_2 &= 9, & \text{II} \\ x_1 + x_2 &= 2. & \text{III}\end{aligned}$$

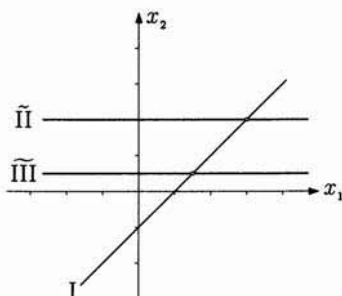
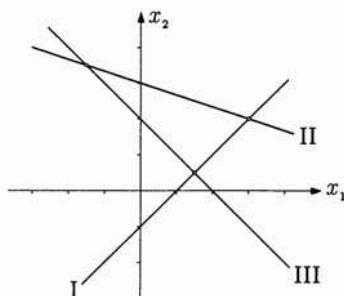


Bild 0.7

Daß sie keinen gemeinsamen Schnittpunkt haben, sieht man an Bild 0.7, wir wollen es auch durch formales Umformen zeigen. Wie umgeformt wurde, ist rechts vermerkt.

$$\begin{aligned}x_1 - x_2 &= 1, & \text{I} \\ 4x_2 &= 8, & \widetilde{\text{II}} = \text{II} - \text{I} \\ 2x_2 &= 1. & \widetilde{\text{III}} = \text{III} - \text{I}\end{aligned}$$

Die Gleichungen $\widetilde{\text{II}}$ und $\widetilde{\text{III}}$ verlangen $x_2 = 2$ und $x_2 = \frac{1}{2}$, das ist ein Widerspruch.

Wie man sieht, sind derartige Umformungen von Gleichungssystemen sehr wirksam, wir werden in 0.4.6 eine Rechtfertigung dafür geben. In obigem Beispiel sollte man bemerken, daß die gemeinsamen Schnittpunkte von I mit II und I mit III erhalten bleiben.

0.2.4. Den Begriff Gerade hatten wir bisher noch nicht präzise erklärt, das soll schleunigst nachgeholt werden:

Definition. Eine Teilmenge $L \subset \mathbb{R}^2$ heißt *Gerade*, wenn es $a_1, a_2, b \in \mathbb{R}$ mit $(a_1, a_2) \neq (0, 0)$ gibt, so daß

$$L = \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\}.$$

Daß man eine Gerade genauso gut mit Hilfe einer Parametrisierung beschreiben kann, ist die Aussage von folgendem

Satz. Eine Teilmenge $L \subset \mathbb{R}^2$ ist genau dann eine Gerade, wenn es $v, w \in \mathbb{R}^2$ mit $w \neq 0$ gibt, so daß

$$L = v + \mathbb{R}w.$$

Dieser Satz folgt später sehr leicht aus der allgemeinen Theorie. Ohne weitere Hilfsmittel ist der Beweis etwas mühsam, wir wollen es dennoch vorführen:

1) Ist L eine Gerade im Sinne der obigen Definition, so gibt es v und w mit den angegebenen Eigenschaften.

Sei also L gegeben durch a_1, a_2, b mit $(a_1, a_2) \neq (0, 0)$. Wir führen den Fall $a_1 \neq 0$ aus, der Fall $a_2 \neq 0$ geht analog. Indem man im Ergebnis von 0.2.2 $\lambda = 0$ und $\lambda = a_1$ setzt (siehe Bild 0.8), kommt man zu der Definition

$$v := \left(\frac{b}{a_1}, 0 \right), \quad w := (-a_2, a_1), \quad L' := v + \mathbb{R}w,$$

und es ist zu zeigen, daß $L = L'$. Dazu ist $L \subset L'$ und $L' \subset L$ zu beweisen.

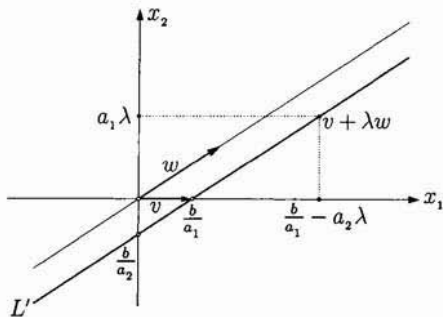


Bild 0.8

a) $L \subset L'$: Ist $(x_1, x_2) \in L$, so ist $a_1x_1 + a_2x_2 = b$. Also gilt für $\lambda := x_2/a_1$

$$(x_1, x_2) = \left(\frac{b}{a_1} - a_2\lambda, a_1\lambda \right) = v + \lambda w.$$

Somit ist $(x_1, x_2) \in L'$.

b) $L' \subset L$: Ist $x \in L'$, so gibt es ein λ mit

$$x = v + \lambda w = \left(\frac{b}{a_1} - a_2\lambda, a_1\lambda \right) = (x_1, x_2).$$

Setzt man $x_1 = \frac{b}{a_1} - a_2\lambda$ und $x_2 = a_1\lambda$ in die Gleichung von L ein, so erhält man

$$a_1x_1 + a_2x_2 = b - a_1a_2\lambda + a_1a_2\lambda = b.$$

Also ist $x \in L$.

2) Ist $L = v + \mathbb{R}w$, so ist eine Gleichung zu finden. Ist $v = (v_1, v_2)$ und $w = (w_1, w_2)$ mit $w_1 \neq 0$, so zeigt eine einfache Überlegung (siehe Bild 0.9), daß

$$\frac{x_2 - v_2}{x_1 - v_1} = \frac{w_2}{w_1}, \quad \text{also} \quad w_2x_1 - w_1x_2 = w_2v_1 - w_1v_2$$

sein muß. Wir definieren daher $a_1 := w_2$, $a_2 := -w_1$, $b := w_2v_1 - w_1v_2$ und

$$L' := \{(x_1, x_2) \in \mathbb{R}^2 : a_1x_1 + a_2x_2 = b\}.$$

a) $L \subset L'$: Dazu muß man $x = v + \lambda w = (v_1 + \lambda w_1, v_2 + \lambda w_2) \in L$ in die Gleichung einsetzen. Das ergibt

$$w_2(v_1 + \lambda w_1) - w_1(v_2 + \lambda w_2) = w_2v_1 - w_1v_2.$$

Also ist $x \in L'$.

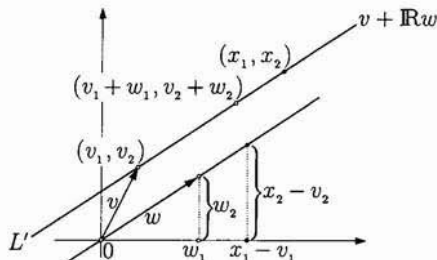


Bild 0.9

b) $L' \subset L$: Sei $x = (x_1, x_2) \in L'$. Dazu muß man ein $\lambda \in \mathbb{R}$ finden, so daß

$$x = (x_1, x_2) = v + \lambda w = (v_1 + \lambda w_1, v_2 + \lambda w_2).$$

Wegen $w_1 \neq 0$ kann man aus der ersten Komponente ausrechnen, daß

$$\lambda = \frac{x_1 - v_1}{w_1}$$

sein muß. Definiert man λ durch diese Gleichung, so folgt aus $x \in L'$, daß auch

$$x_2 = \frac{b - a_1 x_1}{a_2} = \frac{w_2 v_1 - w_1 v_2 - w_2 x_1}{-w_1} = v_2 + \lambda w_2.$$

Also ist $x \in L$.

□

0.3 Ebenen und Geraden im Standardraum $\mathbb{R}^3$

0.3.1. Als „Raum“ betrachten wir den dreidimensionalen reellen Raum. Wie in der Ebene geht durch zwei verschiedene Punkte $v, v' \in \mathbb{R}^3$ genau eine Gerade L . Ist $w := v' - v$, so ist wie in 0.2.1

$$L = \{v + \lambda w : \lambda \in \mathbb{R}\} = v + \mathbb{R}w$$

eine *Parameterdarstellung* von L . Es gibt jedoch zwei wesentliche Unterschiede zur Ebene:

- 1) Zwei Geraden sind im allgemeinen *windschief*, d.h. ohne Schnittpunkt und nicht parallel.
- 2) Eine lineare Gleichung im $\mathbb{R}^3$ beschreibt eine Ebene, zur Beschreibung einer Geraden benötigt man zwei lineare Gleichungen (d.h. man stellt sie als Schnitt von zwei Ebenen dar).

Das wird im folgenden genauer ausgeführt.

0.3.2. Wir betrachten eine lineare Gleichung der Form

$$a_1x_1 + a_2x_2 + a_3x_3 = b$$

und die Menge der *Lösungen*

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : a_1x_1 + a_2x_2 + a_3x_3 = b\}.$$

Ist $a_1 = a_2 = a_3 = 0$, so ist $E = \emptyset$ für $b \neq 0$ und $E = \mathbb{R}^3$ für $b = 0$. Für $a_1 = a_2 = 0$, aber $a_3 \neq 0$ lautet die Gleichung

$$x_3 = \frac{b}{a_3},$$

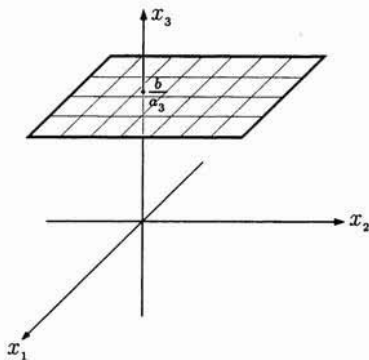


Bild 0.10

das ist eine Ebene parallel zur Ebene $x_3 = 0$.

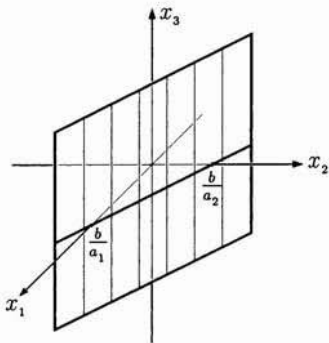


Bild 0.11

Ist $a_3 = 0$ und $(a_1, a_2) \neq (0, 0)$, so haben wir eine Gerade

$$L = \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\},$$

und E entsteht daraus, indem man alle im Raum darüber- und darunterliegenden Punkte dazunimmt, d.h. (siehe Bild 0.11)

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : (x_1, x_2) \in L\}.$$

Sind alle drei Koeffizienten $a_i \neq 0$, so schneidet E die Achsen an den Stellen b/a_i .

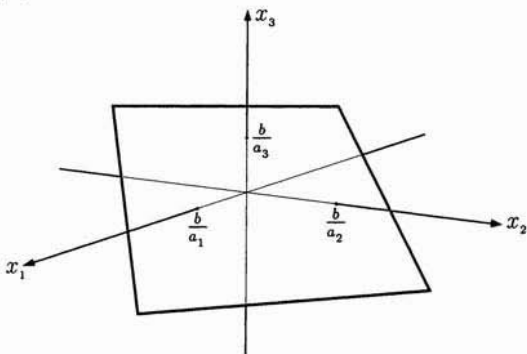


Bild 0.12

Nach diesen Beispielen eine präzise

Definition. Eine Teilmenge $E \subset \mathbb{R}^3$ heißt *Ebene*, wenn es $a_1, a_2, a_3, b \in \mathbb{R}$ mit $(a_1, a_2, a_3) \neq (0, 0, 0)$ gibt, so daß

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : a_1 x_1 + a_2 x_2 + a_3 x_3 = b\}.$$

0.3.3. Zur Parametrisierung einer Geraden reicht ein Parameter, bei einer Ebene geht es mit zweien, und zwar wie folgt. In der Ebenengleichung nehmen wir $a_3 \neq 0$ an (andernfalls vertausche man die Indizes). Wir betrachten die *Standardebene* $\mathbb{R}^2$ und bezeichnen die Punkte mit (λ_1, λ_2) . Setzt man $x_1 = \lambda_1, x_2 = \lambda_2$ in die Ebenengleichung ein, so erhält man

$$x_3 = \frac{1}{a_3} (b - a_1 \lambda_1 - a_2 \lambda_2).$$

Abstrakter ausgedrückt erhält man eine Abbildung

$$\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad (\lambda_1, \lambda_2) \mapsto \left(\lambda_1, \lambda_2, \frac{b}{a_3} - \frac{a_1 \lambda_1}{a_3} - \frac{a_2 \lambda_2}{a_3} \right).$$

Eine triviale Rechnung ergibt

$$a_1 \lambda_1 + a_2 \lambda_2 + a_3 \left(\frac{b}{a_3} - \frac{a_1 \lambda_1}{a_3} - \frac{a_2 \lambda_2}{a_3} \right) = b,$$

also ist $\Phi(\mathbb{R}^2) \subset E$, und mit etwas mehr Rechnung als in 0.2.4 kann man zeigen, daß

$$\Phi: \mathbb{R}^2 \rightarrow E$$

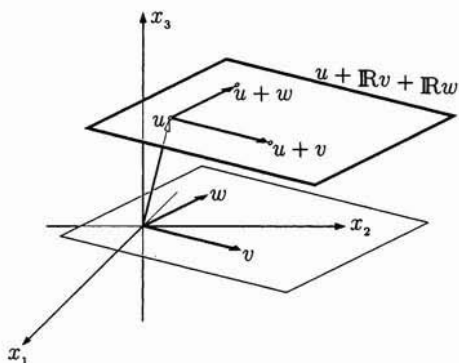
sogar bijektiv ist. Wir verzichten darauf, das hier vorzurechnen, weil es in 2.3 aus der allgemeinen Theorie folgt. Die so erhaltene Abbildung Φ von der Standardebene $\mathbb{R}^2$ auf die in $\mathbb{R}^3$ liegende Ebene E heißt *Parametrisierung*. In der oben angegebenen Form hat sie eine einfache geometrische Interpretation: Φ ist die Umkehrung der Projektion

$$\pi: E \rightarrow \mathbb{R}^2, \quad (x_1, x_2, x_3) \mapsto (x_1, x_2),$$

auf die Koordinatenebene. Natürlich gibt es viele andere Parametrisierungen von E .

Die oben mit Hilfe der Abbildung Φ angegebene Parametrisierung der Ebene E kann man noch etwas anders schreiben. Sind $u, v, w \in \mathbb{R}^3$ gegeben, so sei

$$u + \mathbb{R}v + \mathbb{R}w := \{x \in \mathbb{R}^3 : \text{es gibt } \lambda_1, \lambda_2 \in \mathbb{R} \text{ mit } x = u + \lambda_1 v + \lambda_2 w\}.$$

**Bild 0.13**

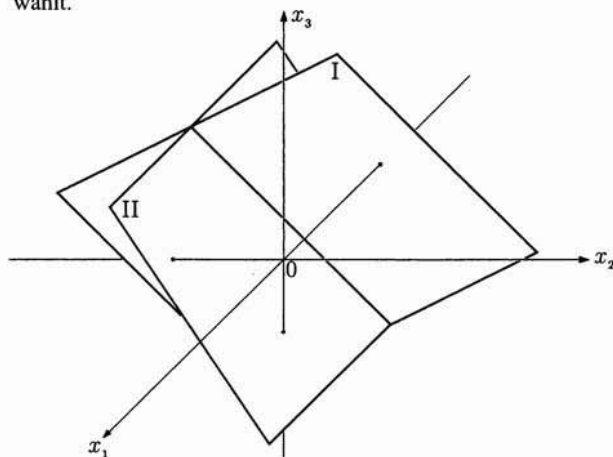
Das ist das Bild von Φ , wenn man

$$u = \Phi(0, 0) = \left(0, 0, \frac{b}{a_3}\right),$$

$$v = \Phi(1, 0) - u = \left(1, 0, -\frac{a_1}{a_3}\right),$$

$$w = \Phi(0, 1) - u = \left(0, 1, -\frac{a_2}{a_3}\right)$$

wählt.

**Bild 0.14**

0.3.4. Nun wollen wir zwei Ebenen schneiden. Dazu zunächst ein

Beispiel. Wir betrachten die Gleichungen

$$x_1 + x_2 + x_3 = -6, \quad \text{I}$$

$$x_1 + 2x_2 + 3x_3 = -10, \quad \text{II}$$

und formen sie um zu

$$x_1 + x_2 + x_3 = -6, \quad \text{I}$$

$$x_2 + 2x_3 = -4. \quad \tilde{\text{II}} = \text{II} - \text{I}$$

Der Schnitt der beiden Ebenen ist eine Gerade L , ein Punkt darauf ist festgelegt durch seine x_3 -Koordinate. Übersetzt in eine Rechnung bedeutet das: man wählt einen reellen Parameter λ , setzt $x_3 = \lambda$ und berechnet erst mit $\tilde{\text{II}}$ und dann mit I

$$x_2 = -2\lambda - 4,$$

$$x_1 = \lambda - 2.$$

Das ergibt eine Parametrisierung von L

$$\Phi: \mathbb{R} \rightarrow L \subset \mathbb{R}^3, \quad \lambda \mapsto (\lambda - 2, -2\lambda - 4, \lambda).$$

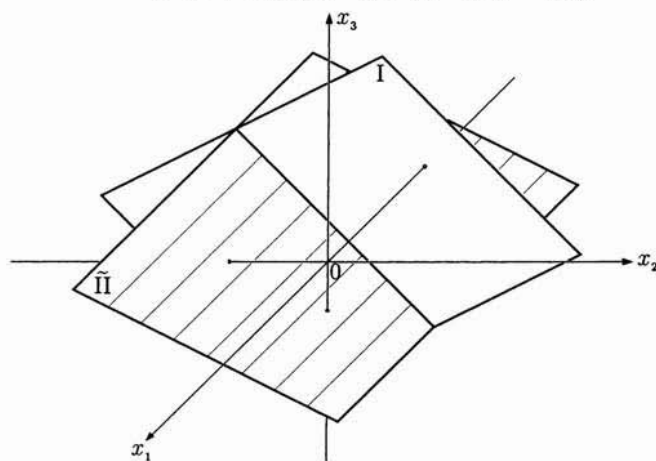


Bild 0.15

Hat man allgemein zwei Ebenengleichungen

$$a_1x_1 + a_2x_2 + a_3x_3 = b, \quad \text{I}$$

$$a'_1x_1 + a'_2x_2 + a'_3x_3 = b', \quad \text{II}$$

und ist $a_1 \neq 0$, so führt eine Umformung wie oben zu einem Unglück, wenn es ein $\varrho \in \mathbb{R}$ gibt mit

$$(a'_1, a'_2, a'_3) = (\varrho a_1, \varrho a_2, \varrho a_3). \quad (*)$$

Dann wird nämlich die linke Seite von $\tilde{\Pi} := \Pi - \varrho \cdot \text{I}$ gleich Null. Das hat einen geometrischen Hintergrund, denn $(*)$ bedeutet, daß die durch I und II beschriebenen Ebenen parallel oder gleich sind. Ein präziser Beweis davon kann später leicht nachgeholt werden (vgl. Aufgabe 8 zu 2.3).

0.3.5. Nun schneiden wir schließlich drei Ebenen im $\mathbb{R}^3$.

Beispiel. Wir nehmen zu dem Beispiel aus 0.3.4 noch eine Gleichung dazu:

$$\begin{array}{rcl} x_1 + x_2 + x_3 & = & -6, \quad \text{I} \\ x_1 + 2x_2 + 3x_3 & = & -10, \quad \text{II} \\ 2x_1 + 3x_2 + 6x_3 & = & -18. \quad \text{III} \end{array}$$

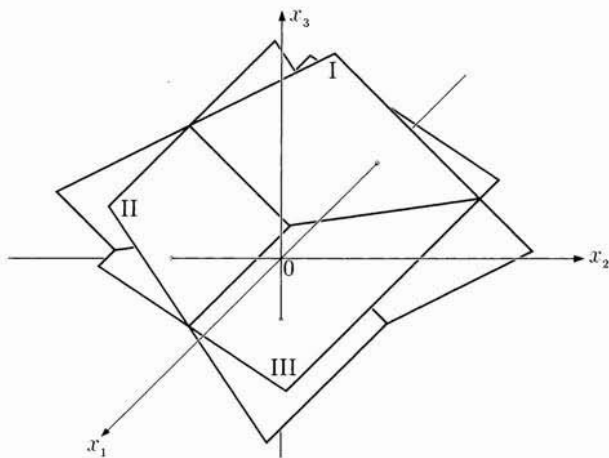


Bild 0.16

Die drei Ebenen schneiden sich in einem Punkt. Um ihn zu berechnen, formen wir wieder um. Die erste Runde ergibt

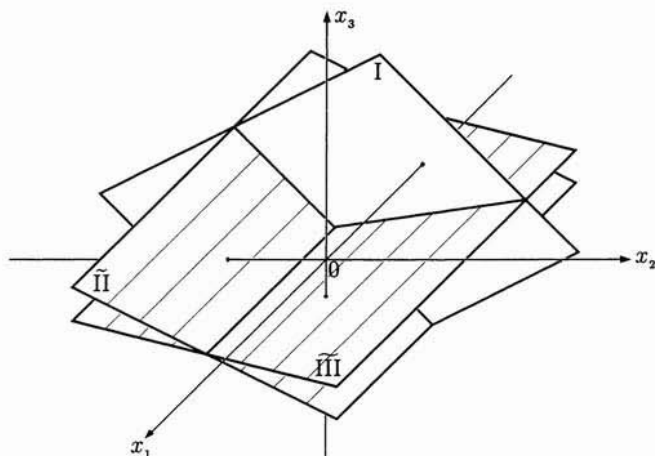


Bild 0.17

$$\begin{array}{rcl}
 x_1 + x_2 + x_3 & = & -6, \quad \text{I} \\
 x_2 + 2x_3 & = & -4, \quad \tilde{\text{II}} = \text{II} - \text{I} \\
 x_2 + 4x_3 & = & -6, \quad \tilde{\text{III}} = \text{III} - 2\text{I}
 \end{array}$$

Um den Schnittpunkt ganz schematisch ausrechnen zu können, formen wir $\tilde{\text{III}}$ noch einmal um zu

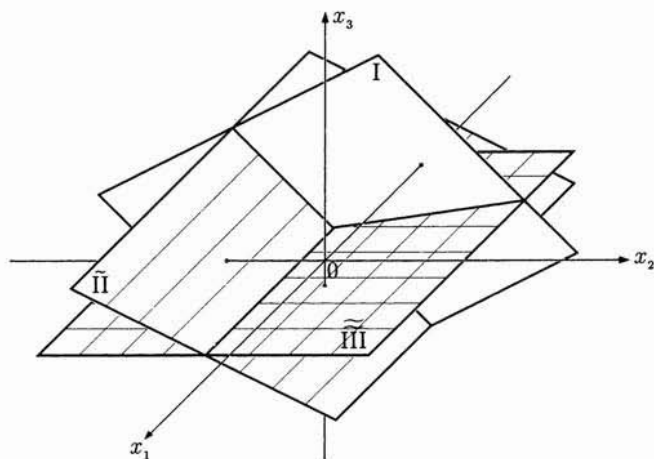
$$2x_3 = -2. \quad \tilde{\tilde{\text{III}}} = \tilde{\text{III}} - \tilde{\text{II}}$$

Damit ergibt sich durch Einsetzen von unten nach oben

$$\begin{array}{rcl}
 x_3 & = & -1 \quad \text{nach } \tilde{\tilde{\text{III}}}, \\
 x_2 & = & -2 \quad \text{nach } \tilde{\text{II}}, \\
 x_1 & = & -3 \quad \text{nach I.}
 \end{array}$$

Der einzige Schnittpunkt der drei Ebenen ist also $(-3, -2, -1)$.

Ob dieses Verfahren im allgemeinen einen einzigen Schnittpunkt liefert, hängt ganz von den Koeffizienten der drei Gleichungen ab (vgl. 2.3.6).

**Bild 0.18**

Aufgaben zu 0.3

1. Zeigen Sie, dass für zwei Punkte $v, w \in \mathbb{R}^n$ die folgenden Bedingungen äquivalent sind:

- i) $v \neq 0$, und es gibt kein $\varrho \in \mathbb{R}$ mit $w = \varrho \cdot v$.
- ii) $w \neq 0$, und es gibt kein $\varrho \in \mathbb{R}$ mit $v = \varrho \cdot w$.
- iii) Sind $\lambda, \mu \in \mathbb{R}$ mit $\lambda v + \mu w = 0$, so folgt notwendigerweise $\lambda = \mu = 0$.

Man nennt v und w *linear unabhängig*, falls eine der obigen Bedingungen erfüllt ist. v und w heißen *linear abhängig*, falls sie nicht linear unabhängig sind. Im untenstehenden Bild sind v und w linear unabhängig, v und w' linear abhängig.

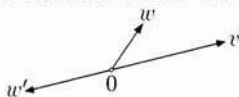


Bild 0.19

2. a) Beweisen Sie, dass eine Teilmenge E des $\mathbb{R}^3$ genau dann eine Ebene ist, wenn es Vektoren $u, v, w \in \mathbb{R}^3$ gibt, so dass v und w linear unabhängig sind und

$$E = u + \mathbb{R}v + \mathbb{R}w.$$

b) Finden Sie für die Ebene $E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 3x_1 - 2x_2 + x_3 = -1\}$ eine Parametrisierung.

c) Geben Sie für die in Parameterdarstellung gegebene Ebene

$$E = (1, 2, 3) + \mathbb{R} \cdot (4, 5, 6) + \mathbb{R} \cdot (7, 8, 9)$$

eine beschreibende lineare Gleichung an.

3. Zeige Sie: Sind $x, y, z \in \mathbb{R}^3$ drei Punkte, die nicht auf einer Geraden liegen, so gibt es genau eine Ebene $E \subset \mathbb{R}^3$, die x, y und z enthält, nämlich

$$E = x + \mathbb{R} \cdot (x - y) + \mathbb{R} \cdot (x - z).$$

0.4 Das Eliminationsverfahren von GAUSS

0.4.1. Nach den vielen Spezialfällen und Beispielen ist der Leser hoffentlich gerüstet für den allgemeinen Fall. Die Zahl $n \in \mathbb{N}$ ist die Zahl der Unbestimmten, davon unabhängig kann man eine Zahl $m \in \mathbb{N}$ für die Anzahl der linearen Bedingungen wählen. Da m die Anzahl der verfügbaren Buchstaben überschreiten kann, verwenden wir für die Koeffizienten a_{ij} doppelte Indizes. Das Gleichungssystem lautet also:

$$\begin{array}{ccccccc} a_{11}x_1 + & \dots & + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 + & \dots & + a_{mn}x_n & = & b_m, \end{array} \quad (*)$$

und gesucht ist die Menge der $(x_1, \dots, x_n) \in \mathbb{R}^n$, die alle Gleichungen erfüllen.

Das System $(*)$ ist mühsam aufzuschreiben. Ein Meister in übersichtlichen Rechenverfahren war A. CAYLEY, der auch erstmals systematisch *Matrizen* verwendete. Das hilft hier sofort. Die Koeffizienten a_{ij} schreibt man, wie sie in $(*)$ vorkommen, als rechteckiges Schema (*Matrix* genannt)

$$A := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Nun ist der Kniff, nicht die *liegenden* Vektoren (oder *Zeilen*) $(x_1, \dots, x_n)$ und $(b_1, \dots, b_m)$ zu betrachten, sondern entsprechend der Anordnung der b_i in $(*)$ die *stehenden* Vektoren (oder *Spalten*)

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Zwischen der Matrix A und der Spalte x der Höhe n erklärt man ein Produkt, das eine Spalte der Höhe m ergibt:

$$A \cdot x := \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}.$$

Dabei ist entscheidend, daß x so viele Zeilen wie A Spalten hat. Das *lineare Gleichungssystem* $(*)$ kann man dann in der Form

$$A \cdot x = b \quad (*')$$

schreiben, wobei das eine Gleichheit von zwei Spalten mit jeweils m Zeilen bedeutet. Diese geschickte Schreibweise ist gewöhnungsbedürftig und auch etwas gefährlich, weil man leicht vergessen kann, was sie explizit bedeutet.

Man nennt A die *Koeffizientenmatrix* des linearen Gleichungssystems. Hängt man die Spalte b noch an, so erhält man die Matrix

$$(A, b) := \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix},$$

sie heißt *erweiterte Koeffizientenmatrix*. Darin ist alle Information über das Gleichungssystem enthalten.

Hat eine Matrix A insgesamt m Zeilen und n Spalten, so spricht man zur Abkürzung von einer $(m \times n)$ -Matrix. Man schreibt dafür $A = (a_{ij})$, die reellen Zahlen a_{ij} heißen *Einträge* von A .

Eine andere Methode, das Gleichungssystem $(*)$ kürzer aufzuschreiben, benutzt das *Summenzeichen* $\sum$. Allgemein ist

$$\sum_{j=1}^n c_j := c_1 + \dots + c_n.$$

Dabei heißt j der *Summationsindex*, man kann ihn durch jeden anderen Buchstaben ersetzen. In dieser Schreibweise lautet die i -te Gleichung

$$\sum_{j=1}^n a_{ij} x_j = b_i,$$

das ganze System also

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad \text{für } i = 1, \dots, m. \quad (*'')$$

Welche Schreibweise man bevorzugt, ist nebensächlich. Wir werden vorwiegend $A \cdot x = b$ benutzen, weil dabei die geringste Anzahl von Buchstaben erforderlich ist.

0.4.2. Nachdem eine energiesparende Schreibweise für Gleichungssysteme vereinbart ist, können wir das Problem der Lösung in Angriff nehmen. Die *Lösungsmenge* ist nach Definition gleich

$$\text{Lös}(A, b) := \{x \in \mathbb{R}^n : A \cdot x = b\},$$

wobei x als Spaltenvektor geschrieben ist. (Diese suggestive Bezeichnungsweise habe ich bei [J] entdeckt; ich konnte nicht der Versuchung widerstehen, sie zu

übernehmen.) Das System zu *lösen* heißt, eine effiziente Methode zur Beschreibung der Menge $\text{Lös}(A, b)$ anzugeben. Was wir schließlich erhalten werden, ist eine Zahl $k \in \mathbb{N}$ und eine explizit angebbare bijektive Abbildung

$$\Phi: \mathbb{R}^k \rightarrow \text{Lös}(A, b) \subset \mathbb{R}^n,$$

sie heißt *Parametrisierung*. Die Berechnung von Φ mit Hilfe des nach C.F. GAUSS benannten *Eliminationsverfahrens* ist recht einfach, das ist Ziel dieses Kapitels. Der Nachweis der guten Eigenschaften von Φ erfordert etwas Theorie und wird in Kapitel 2 nachgeholt. Der abstrakte Hintergrund von linearen Gleichungssystemen wird schließlich in Kapitel 6 erläutert.

0.4.3. In den Beispielen aus 0.2 und 0.3 hatten wir Gleichungssysteme so lange umgeformt, bis eine Parametrisierung schrittweise „von unten nach oben“ berechnet werden konnte. Beispiele für so umgeformte Koeffizientenmatrizen A waren

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Die Nullen zu Beginn der Zeilen haben dabei eine typische Staffelung, die Trennlinie von den anderen Einträgen hat Stufenform.

Definition. Eine $m \times n$ -Matrix $A = (a_{ij})$ heißt in *Zeilenstufenform*, wenn sie von der folgenden Form ist:

$$A = \left(\begin{array}{ccccccc} & & \circledast & & & & \\ & & & \circledast & & & \\ & & & & \circledast & & \\ & & & & & \ddots & \\ & & & & & & \circledast \\ & 0 & & & & & \\ & & & & & & \end{array} \right) \left. \vphantom{\begin{pmatrix} \\ \\ \\ \\ \\ \end{pmatrix}} \right\} r.$$

Dabei müssen die Einträge an den mit $\circledast$ markierten Stellen ungleich Null sein, und unterhalb der eingezeichneten „Stufenlinie“ dürfen nur Nullen stehen.

Damit auch die Grenzfälle klar geregelt sind, kann man diese Definition noch präziser aufschreiben. A ist in Zeilenstufenform, wenn folgendes gilt:

1. Es gibt eine Zahl r mit $0 \leq r \leq m$, so daß in den Zeilen mit Index 1 bis r jeweils nicht nur Nullen stehen und in den Zeilen mit Index $r + 1$ bis m nur Nullen stehen.

2. Für jedes i mit $1 \leq i \leq r$ betrachten wir den niedrigsten Index j_i der Spalte, in der ein Eintrag ungleich Null steht, in Zeichen

$$j_i := \min\{j: a_{ij} \neq 0\}.$$

Offensichtlich ist $1 \leq j_i \leq n$, und die zusätzliche *Stufenbedingung* lautet

$$j_1 < j_2 < \dots < j_r.$$

Man beachte, daß der Fall $r = 0$ zugelassen ist; dann sind alle Einträge von A gleich Null. Die besonders ausgezeichneten und oben durch $\circledast$ gekennzeichneten Einträge

$$a_{1j_1}, \dots, a_{rj_r}$$

heißen *Pivots* (auf deutsch *Angelpunkte*) von A . Sie sind nach Definition von Null verschieden.

Beispiel. Für

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

ist $m = 4$, $n = 7$, $r = 3$, $j_1 = 2$, $j_2 = 3$, $j_3 = 6$.

Besonders einfach aufzuschreiben ist der Spezialfall, in dem

$$j_1 = 1, \quad j_2 = 2, \quad \dots, \quad j_r = r.$$

Dann hat A die Form

$$A = \left(\begin{array}{ccccccc} \circledast & & & & & & \\ & \circledast & & & & & \\ & & \ddots & & & & \\ & & & 0 & & \circledast & \\ & & & & & & \end{array} \right) \Bigg\}^r.$$

Durch eine Umordnung der Spalten von A , d.h. eine andere Numerierung der Unbekannten des entsprechenden Gleichungssystems, kann man das stets erreichen. Für die Praxis ist das nebensächlich, aber für die Theorie kann man sich dadurch die lästigen Doppelindizes j_i ersparen.

0.4.4. Nun geben wir ein Lösungsverfahren für ein lineares Gleichungssystem an, bei dem die Koeffizientenmatrix A in Zeilenstufenform ist. Zur Vereinfachung nehmen wir an, daß die Pivots in den ersten r Spalten sitzen. Dann hat die

erweiterte Koeffizientenmatrix die Gestalt

$$(A, b) = \left(\begin{array}{cccc|c} a_{11} & & & & b_1 \\ & a_{22} & & & \vdots \\ & & \ddots & & \vdots \\ & & & a_{rr} & b_r \\ & 0 & & & b_{r+1} \\ & & & & \vdots \\ & & & & b_m \end{array} \right)$$

mit $a_{11} \neq 0, \dots, a_{rr} \neq 0$. Die Einträge $b_{r+1}, \dots, b_m$ sind entscheidend für die Frage, ob es überhaupt eine Lösung gibt.

Bemerkung. Gibt es ein $b_i \neq 0$ mit $r+1 \leq i \leq m$, so ist $\text{Lös}(A, b)$ leer.

Beweis. Die i -te Gleichung lautet

$$0 \cdot x_1 + \dots + 0 \cdot x_n = b_i \neq 0.$$

Diese Bedingung kann kein x erfüllen. □

Im gegenteiligen Fall

$$b_{r+1} = \dots = b_m = 0$$

geben wir nun eine Methode an, Lösungen zu konstruieren. Dazu unterscheiden wir zwischen zwei Arten von Variablen:

$x_{r+1}, \dots, x_n$ sind *freie Variablen*, sie können alle beliebigen Werte annehmen.

$x_1, \dots, x_r$ sind *gebundene Variablen*, sie sind eindeutig dadurch festgelegt, für welche Werte sich die freien Variablen entschieden haben.

Das kann man so beschreiben. Man setzt $k := n - r$, das ist die Zahl der freien Variablen, wählt $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ als *Parameter*, und setzt

$$x_{r+1} = \lambda_1, \quad x_{r+2} = \lambda_2, \dots, x_n = \lambda_k.$$

Zur Berechnung der $x_1, \dots, x_r$ daraus beginnt man mit der r -ten Gleichung

$$a_{rr}x_r + a_{r,r+1}\lambda_1 + \dots + a_{rn}\lambda_k = b_r.$$

Daraus erhält man

$$x_r = \frac{1}{a_{rr}}(b_r - a_{r,r+1}\lambda_1 - \dots - a_{rn}\lambda_k).$$

Setzt man das in die $(r-1)$ -te Gleichung ein, erhält man analog

$$x_{r-1} = d_{r-1,r-1}b_{r-1} + d_{r-1,r}b_r + c_{r-1,1}\lambda_1 + \dots + c_{r-1,k}\lambda_k,$$

wobei die auftretenden Zahlen c und d von den Einträgen der Zeilen $r-1$ und r aus der Matrix A abhängen. Führt man so fort, erhält man schließlich

$$x_1 = d_{11}b_1 + \dots + d_{1r}b_r + c_{11}\lambda_1 + \dots + c_{1k}\lambda_k.$$

Insgesamt ergibt sich eine Abbildung

$$\begin{aligned}\Phi: \mathbb{R}^k &\rightarrow \text{Lös}(A, b) \subset \mathbb{R}^n, \\ (\lambda_1, \dots, \lambda_k) &\mapsto (x_1, \dots, x_r, \lambda_1, \dots, \lambda_k),\end{aligned}$$

wobei für $x_1, \dots, x_r$ die oben berechneten von $\lambda_1, \dots, \lambda_k$, den Einträgen von A und $b_1, \dots, b_r$ abhängigen Ausdrücke einzusetzen sind. Da hierfür nach den obigen Rechnungen alle r Gleichungen erfüllt sind, liegen die Werte von Φ in der Lösungsmenge $\text{Lös}(A, b)$. Man hat also für beliebig gewählte Parameter $\lambda_1, \dots, \lambda_k$ eine Lösung des gegebenen Gleichungssystems erhalten. $\square$

Beispiel. Ist A die (4×7) -Matrix aus 0.4.3, und wählt man

$$b_1 = 3, \quad b_2 = 1, \quad b_3 = 2, \quad b_4 = 0,$$

so sind die Lösungen von $A \cdot x = b$ also in Spalten geschrieben gegeben durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \frac{3}{2} - 2\lambda_2 - 3\lambda_3 - \frac{5}{2}\lambda_4 \\ \frac{1}{3} - 3\lambda_2 - 2\lambda_3 + \frac{1}{3}\lambda_4 \\ \lambda_2 \\ \lambda_3 \\ \frac{2}{3} - \frac{1}{3}\lambda_4 \\ \lambda_4 \end{pmatrix} \\ = \begin{pmatrix} 0 \\ \frac{3}{2} \\ \frac{1}{3} \\ 0 \\ 0 \\ \frac{2}{3} \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -2 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ -3 \\ -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ -\frac{5}{2} \\ \frac{1}{3} \\ 0 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix}.$$

0.4.5. Einen wichtigen Spezialfall wollen wir noch erwähnen: Ist die Matrix A quadratisch, so hat man ebensoviele Gleichungen wie Unbekannte. Ist speziell

A auf Zeilenstufenform mit $r = n$, so ist

$$A = \begin{pmatrix} \boxed{\circledast} & & & \\ & \boxed{\circledast} & & \\ & & \ddots & \\ 0 & & & \boxed{\circledast} \end{pmatrix},$$

und es gibt wegen $k = n - r = 0$ keinen freien Parameter, also eine einzige Lösung

$$x = (x_1, \dots, x_n),$$

die man wieder von unten nach oben berechnet. Ist überdies

$$b_1 = \dots = b_n = 0, \quad \text{so ist} \quad x_n = \dots = x_1 = 0,$$

man erhält also nur die *triviale Lösung*. Beispiele für eindeutig lösbare Gleichungssysteme findet man in 0.2.3 a), 0.3.5 und Aufgabe 2.

0.4.6. Nachdem wir gesehen haben, wie sich ein Gleichungssystem in Zeilenstufenform lösen läßt, versuchen wir nun, ein beliebiges System auf diese Form zu bringen. Dazu benutzen wir zwei Arten von *elementaren Zeilenumformungen* der erweiterten Koeffizientenmatrix:

- 1) Vertauschung von zwei Zeilen.
- 2) Addition der λ -fachen i -ten Zeile zur k -ten Zeile, wobei $0 \neq \lambda \in \mathbb{R}$ und $i \neq k$ ist.

Diese Umformungen sind gerechtfertigt durch den

Satz. Sei (A, b) die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems, und $(\tilde{A}, \tilde{b})$ aus (A, b) durch endlich viele elementare Zeilenumformungen entstanden. Dann haben die Systeme $A \cdot x = b$ und $\tilde{A} \cdot x = \tilde{b}$ gleiche Lösungsräume, in Zeichen $\text{Lös}(A, b) = \text{Lös}(\tilde{A}, \tilde{b})$.

Vorsicht! Man beachte, daß Spaltenumformungen eine völlig andere Wirkung haben, weil dadurch die Unbekannten „gemischt“ werden. Das ist unerwünscht. Nur Vertauschungen in den ersten n Spalten sind ungefährlich, sie bewirken lediglich eine Umnummerierung der Unbekannten.

Beweis. Es genügt zu beweisen, daß der Lösungsraum bei einer einzigen elementaren Zeilenumformung unverändert bleibt, denn dann ändert auch Wiederholung nichts.

Typ 1) ist völlig unproblematisch, weil alle Gleichungen simultan erfüllt sein müssen, die Reihenfolge ist gleichgültig.

Bei Typ 2) muß man etwas rechnen. Da nur die Zeilen i und k betroffen sind, genügt es zu zeigen, daß die beiden aus jeweils zwei Gleichungen bestehenden Systeme

$$\begin{aligned} a_{i1}x_1 + \dots + a_{in}x_n &= b_i \\ a_{k1}x_1 + \dots + a_{kn}x_n &= b_k \end{aligned} \quad (*)$$

und

$$\begin{aligned} a_{i1}x_1 + \dots + a_{in}x_n &= b_i \\ (a_{k1} + \lambda a_{i1})x_1 + \dots + (a_{kn} + \lambda a_{in})x_n &= b_k + \lambda b_i \end{aligned} \quad (\tilde{*})$$

gleiche Lösungsräume haben. Erfüllt $x = (x_1, \dots, x_n)$ die Gleichungen $(*)$, so auch die erste von $(\tilde{*})$, und durch Addition der λ -fachen ersten Gleichung von $(*)$ zur zweiten die zweite Gleichung von $(\tilde{*})$. Umgekehrt folgt durch Subtraktion der λ -fachen ersten Gleichung aus $(\tilde{*})$ von der zweiten auch die zweite Gleichung aus $(*)$. $\square$

Was bei Umformungen vom Typ 2) geometrisch vorgeht, sieht man am einfachsten in der Ebene. Zwei Gleichungen beschreiben zwei Geraden, die Lösungsmenge besteht aus den Schnittpunkten (keiner, einer, oder eine ganze Gerade, vgl. 0.2). Was verschiedene Faktoren λ bewirken, wollen wir am besten an einem Beispiel zeigen: Gegeben seien Geraden

L_i durch $x_1 = 1$ und L_k durch $x_1 - x_2 = 2$.

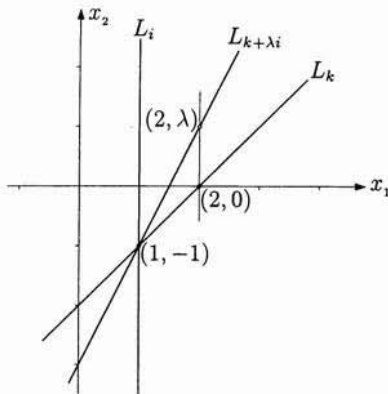


Bild 0.20

Dann ist

$$L_{k+\lambda i} \text{ gegeben durch } (1 + \lambda)x_1 - x_2 = 2 + \lambda.$$

Diese Schar von Geraden mit Parameter λ geht durch $(1, -1)$, sie enthält alle Geraden durch $(1, -1)$ mit Ausnahme von L_i , und die Zahl λ ist am Schnittpunkt mit der Geraden $x_1 = 2$ zu sehen.

0.4.7. Der letzte und am schwierigsten in allgemeiner Form aufzuschreibende Schritt ist enthalten in dem

Satz. Jede Matrix A kann man durch elementare Zeilenumformungen in eine Matrix $\tilde{A}$ in Zeilenstufenform überführen.

Beweis. Wir geben ein konkretes Verfahren an, das schrittweise durchgeführt wird und so aufgebaut ist, daß daraus ohne große Schwierigkeiten ein Computerprogramm gemacht werden kann. Wer durch die vielen Indizes verwirrt ist, möge zunächst das unten angegebene Beispiel studieren, bei dem drei Runden nötig sind.

Sei A eine $m \times n$ -Matrix. Ist $A = 0$, so hat A nach Definition schon Zeilenstufenform mit $r = 0$.

Ist $A \neq 0$, so gibt es mindestens einen Eintrag $\neq 0$. Also gibt es mindestens eine von Null verschiedene Spalte, wir wählen die mit dem kleinsten Index j_1 , in Zeichen

$$j_1 = \min\{j : \text{es gibt ein } i \text{ mit } a_{ij} \neq 0\}.$$

Ist $a_{1j_1} \neq 0$, so können wir es als Pivot wählen. Andernfalls suchen wir uns ein $a_{i_1j_1} \neq 0$ und vertauschen die Zeile 1 mit der Zeile i_1 . Das ist schon die erste Zeile von $\tilde{A}$, also gilt für den ersten Pivot

$$\tilde{a}_{1j_1} = a_{i_1j_1}.$$

Durch Umformungen vom Typ 2) kann man alle unterhalb von $\tilde{a}_{1j_1}$ stehenden Einträge zu Null machen. Ist a einer davon, so soll

$$a + \lambda \tilde{a}_{1j_1} = 0$$

werden, also hat man

$$\lambda = -\frac{a}{\tilde{a}_{1j_1}} \quad (*)$$

zu wählen. Das Ergebnis dieser Umformungen ist von der Gestalt

$$\tilde{A}_1 = \left(\begin{array}{cccc|cccc} 0 & \cdots & 0 & \tilde{a}_{1j_1} & * & \cdots & * & * \\ \vdots & & \vdots & 0 & \hline \vdots & & \vdots & \vdots & & & & \\ 0 & \cdots & 0 & 0 & & & & A_2 \end{array} \right),$$

wobei an den mit * markierten Stellen irgendwelche Einträge stehen. Die Matrix A_2 hat $m - 1$ Zeilen und $n - j_1$ Spalten.

Im zweiten Schritt macht man mit A_2 das Gleiche wie oben im ersten Schritt mit $A = A_1$: Ist $A_2 = 0$, so hat $\tilde{A}_1$ schon Zeilenstufenform; andernfalls suche man $j_2 > j_1$ und den Pivot $\tilde{a}_{2j_2}$. Die dabei nötigen Zeilenumformungen von A_2 kann man auf die Zeilen 2 bis m von $\tilde{A}_1$ ausdehnen, ohne daß sich in den Spalten 1 bis j_1 etwas ändert, denn dort stehen nur Nullen.

Ist A_2 umgeformt, so erhält man A_3 , u.s.w. Das Verfahren muß abbrechen, weil die Zeilen- und Spaltenzahlen der Matrizen A_k abnehmen, oder weil im Lauf des Verfahrens eine Matrix $A_k = 0$ entsteht. Das Endergebnis ist

$$\tilde{A} = \begin{pmatrix} & \tilde{a}_{1j_1} & * & \cdots & \cdots & \cdots & * \\ & & \tilde{a}_{2j_2} & & & & \vdots \\ & & & \ddots & & & \vdots \\ & 0 & & & \tilde{a}_{rj_r} & \cdots & * \end{pmatrix}.$$

□

Beispiel. Damit der Gang der Rechnung mit dem bloßen Auge zu erkennen ist, sind die Einträge so gewählt, daß sie ganzzahlig bleiben.

$$A = \begin{array}{ccccc|ccccc} 0 & 0 & 1 & 2 & 9 & 0 & 3 & 4 & 5 & 9 & \overline{0} & \underline{3} & \underline{4} & \underline{5} & \underline{9} \\ 0 & 3 & 4 & 5 & 9 & \rightsquigarrow & 0 & 0 & 1 & 2 & 9 & 0 & 0 & 1 & 2 & 9 \\ 0 & 6 & 7 & 8 & 9 & & 0 & 6 & 7 & 8 & 9 & 0 & 0 & -1 & -2 & -9 \\ 0 & 9 & 9 & 9 & 9 & & 0 & 9 & 9 & 9 & 9 & 0 & 0 & -3 & -6 & -18 \end{array}$$

$$\rightsquigarrow \begin{array}{ccccc|ccccc} \overline{0} & \underline{3} & \underline{4} & \underline{5} & \underline{9} & \overline{0} & \underline{3} & \underline{4} & \underline{5} & \underline{9} \\ 0 & 0 & \underline{1} & \underline{2} & \underline{9} & 0 & 0 & \underline{1} & \underline{2} & \underline{9} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 \end{array} = \tilde{A}$$

Bei dem oben allgemein beschriebenen Verfahren wird aus r verschiedenen Spalten jeweils ein Eintrag als Pivot ausgewählt, Kandidaten sind alle von Null verschiedenen Einträge. Für die Theorie wird sich später zeigen, daß das Ergebnis nicht von der Wahl abhängt. Für die Praxis ist es vorteilhaft, den vom Betrag her größten Eintrag zu wählen, weil entsprechend (*) durch den Pivot dividiert wird, und kleine Nenner zu großen Schwankungen führen können (vgl. Aufgabe 4).

0.4.8. Nun ist das *Eliminationsverfahren* von GAUSS für ein System von m linearen Gleichungen und n Unbestimmten mit reellen Koeffizienten komplett, wir fassen die einzelnen Schritte noch einmal zusammen:

- 1) Man schreibe die erweiterte Koeffizientenmatrix (A, b) auf.
- 2) Man bringe A auf Zeilenstufenform und forme dabei die Spalte b mit um. Ergebnis ist $(\tilde{A}, \tilde{b})$, insbesondere die Zahl r . Beachte, daß in der b -Spalte kein Pivot gesucht wird!
- 3) Man lese an $\tilde{b}$ ab, ob es Lösungen gibt und wenn ja, berechne man die Parametrisierung

$$\Phi: \mathbb{R}^{n-r} \rightarrow \text{Lös}(\tilde{A}, \tilde{b}) = \text{Lös}(A, b) \subset \mathbb{R}^n$$

der Lösungsmenge.

Nun zu der entscheidenden Frage nach den Eigenschaften der so erhaltenen Abbildung Φ , der *Parametrisierung* der Lösungen. Zu verschiedenen k -Tupeln von Parametern gehören auch verschiedene Lösungen, da die Parameter direkt in die letzten k Komponenten der Lösung eingetragen werden (vgl. 0.4.4). Also ist Φ injektiv. Bleibt die Frage, ob Φ surjektiv ist, d.h. ob alle Lösungen von der Parametrisierung erreicht werden.

Zur Beantwortung der letzten Frage muß man etwas arbeiten. Insbesondere ist zu zeigen, daß die mit Hilfe der Umformungen erhaltene Zahl r nur von der Matrix A und nicht vom Verfahren (etwa der Auswahl der Pivots) abhängt. Man kann das alles relativ direkt angehen, aber viel schöner geht es mit Hilfe von etwas Theorie, die ja ohnehin schmackhaft gemacht werden soll. Die Zahl r wird sich als *Rang* der Matrix A erweisen, und die Lösungsmenge $\text{Lös}(A, b)$ hat die Struktur eines *affinen Raumes* der *Dimension* $k = n - r$. In dieser Situation folgt die Surjektivität von Φ aus der Injektivität, die Begründungen werden in 2.3 nachgeholt. Das soll aber nicht daran hindern, schon vorher Gleichungssysteme zu lösen, wenn sie auftreten.

Aufgaben zu 0.4

1. Lösen Sie folgende lineare Gleichungssysteme:

a)

$$\begin{array}{rrrrrcl} & x_2 & +2x_3 & +3x_4 & = & 0 \\ x_1 & +2x_2 & +3x_3 & +4x_4 & = & 0 \\ 2x_1 & +3x_2 & +4x_3 & +5x_4 & = & 0 \\ 3x_1 & +4x_2 & +5x_3 & +6x_4 & = & 0 \end{array}$$

b)

$$\begin{array}{rrrrr} -6x_1 & +6x_2 & +2x_3 & -2x_4 & = & 2 \\ -9x_1 & +8x_2 & +3x_3 & -2x_4 & = & 3 \\ -3x_1 & +2x_2 & +x_3 & & = & 1 \\ -15x_1 & +14x_2 & +5x_3 & -4x_4 & = & 5 \end{array}$$

2. Geben Sie die Lösung des linearen Gleichungssystems an, das durch die folgende erweiterte Koeffizientenmatrix gegeben ist:

$$\left(\begin{array}{cccc|c} 1 & -1 & 2 & -3 & 7 \\ 4 & 0 & 3 & 1 & 9 \\ 2 & -5 & 1 & 0 & -2 \\ 3 & -1 & -1 & 2 & -2 \end{array} \right).$$

3. Bestimmen Sie, für welche $t \in \mathbb{R}$ das folgende lineare Gleichungssystem in Matrixdarstellung lösbar ist und geben Sie gegebenenfalls die Lösung an.

$$\left(\begin{array}{ccc|c} 2 & 4 & 2 & 12t \\ 2 & 12 & 7 & 12t + 7 \\ 1 & 10 & 6 & 7t + 8 \end{array} \right)$$

4. Lösen Sie das folgende lineare Gleichungssystem auf einem Taschenrechner mit einer Rechengenauigkeit von n Stellen hinter dem Komma (Abschneiden weiterer Stellen ohne Rundung!) für $\varepsilon = 10^{-k}$ für größer werdendes $k \leq n$, und zwar einmal mit dem Pivot ε und einmal mit dem „maximalen Zeilenpivot“ 1 der ersten Spalte.

$$\begin{array}{rcl} x + y & = & 2, \\ \varepsilon x + y & = & 1. \end{array}$$

Beschreiben Sie den geometrischen Hintergrund dieser Umformungen.