

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

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

- 1.) Die verschiedenen Gleichungen des Systems sollten Koeffizienten in ähnlichen Größenordnungen enthalten. Da sich nichts an der Lösungsmenge ändert, wenn man eine der Gleichungen mit einer von Null verschiedenen Konstanten multipliziert, kann man dies etwa dadurch erreichen, daß man alle Zeilenvektoren der Matrix des Gleichungssystems auf dieselbe Länge (z.B. Eins) normiert.
- 2.) Falls man ein Vielfaches einer Gleichung zu einer anderen addiert, sollte der Koeffizient, mit dem die erste Gleichung multipliziert wird, möglichst klein sein, damit die zweite Gleichung nicht zu stark gestört wird. Dies läßt sich etwa dadurch erreichen, daß man bei der Elimination einer Variablen genau die Gleichung unverändert läßt, in der diese Variable mit dem betragsgrößten Koeffizienten vorkommt.

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

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

oder in neueren Büchern wie

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

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

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

und in dem sehr ausführlichen Buch

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

### i) Matrixgleichungen und die Berechnung der Inversen

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

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

Sind

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

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

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

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

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

natürlich nur dann erreichbar, wenn das Gleichungssystem höchstens eine Lösung hat.) Durch Division läßt sich der verbliebene Eintrag noch auf Eins normieren, so daß sich dann rechts leicht die Lösung ablesen läßt.

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

Betrachten wir als Beispiel die Inversion der Matrix

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

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

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

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

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

und wenden darauf die üblichen Eliminationsschritte an:

Als erstes müssen wir in der ersten Spalte alle Koeffizienten bis auf einen zu Null machen; dazu subtrahieren wir das vier-  $b_{2W}$ -achtfache der ersten Zeile von der zweiten  $b_{2W}$ -dritten und erhalten

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

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

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

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

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

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

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

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

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

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

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

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

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

```
seiler@hilbert:~$ matlab
```

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

To get started, type one of these: helpwin, helpdesk, or demo.

For Product information, visit [www.mathworks.com](http://www.mathworks.com).

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

```
A =
     1     2     3
     4     5     6
     7     8     9
```

```
>> B = inv(A)
```

Warning: Matrix is close to singular or badly scaled.

Results may be inaccurate. RCOND = 1.541976e-18.

```
B =
     1.0e+16 *
    -0.4504     0.9007    -0.4504
     0.9007    -1.8014     0.9007
    -0.4504     0.9007    -0.4504
```

```
>> A*B
```

```
ans =
```

```
     2     0     0
     8     0    -4
    16     0     8
```

```
>> B*A
```

```
ans =
```

```
     0    -4     0
     0     8     0
     4     0     0
```

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

## j) Spezielle Matrizen

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

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

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

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

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

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

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

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

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

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

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

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

übergehen, z.B. den Dreiecksmatrizen:

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

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

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

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

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

Einträge notwendigerweise gleich Null, so daß nur

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

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

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

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

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

Am einfachsten geht das, wenn wir wieder die lineare Abbildung

$$\varphi: k^n \rightarrow k^n; \quad \vec{v} \mapsto A\vec{v}$$

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

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

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

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

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

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

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

$$\varphi: k^n \rightarrow k^n; \quad \vec{v} \mapsto A\vec{v} \quad \text{und} \quad \psi: k^n \rightarrow k^n; \quad \vec{v} \mapsto B\vec{v}$$

seien die zugehörigen linearen Abbildungen.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$E_{ij} = \begin{pmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix} \begin{matrix} j \\ \downarrow \\ \leftarrow i \end{matrix}$$

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

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

Für eine  $q \times m$ -Matrix  $B$  sei entsprechend  $\omega: k^m \rightarrow k^q$  die lineare Abbildung; dann bildet  $\omega \circ \varphi$  alle Koordinateneinheitsvektoren von  $k^n$  mit Ausnahme des  $j$ -ten auf den Nullvektor ab, da bereits  $\varphi$  diese Eigenschaft hat. Der  $j$ -te Koordinateneinheitsvektor wird von  $\varphi$  abgebildet auf den  $i$ -ten Koordinateneinheitsvektor von  $k^m$ , der wiederum von  $\omega$

abgebildet wird auf den  $i$ -ten Spaltenvektor von  $B$ . Die Abbildungsmatrix  $BE_{ij}$  von  $\omega \circ \varphi$  hat also als  $j$ -te Spalte die  $i$ -te Spalte von  $A$ ; alle anderen Spalten sind Null.

Kurz können wir diese beiden Resultate zusammenfassen zu folgender Regel:

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

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

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

$$\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}.$$

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

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

c) Die Menge aller Permutationen  $\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  wird als *symmetrische Gruppe*  $S_n$  bezeichnet.

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

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

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

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

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

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



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

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

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

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

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

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

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

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

Somit ist

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

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

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

d.h.

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

Aus der Darstellung

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

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

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

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

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

Uns interessieren derzeit keine Abbildungen der Menge  $\{1, \dots, n\}$  auf sich selbst, sondern lineare Abbildungen zwischen Vektorräumen; die bekommen wir, indem wir Permutationen auf die Basisvektoren anwenden. Wir betrachten also zu einer Permutation  $\pi \in \mathfrak{S}_n$  die lineare Abbildung

$$\varphi_\pi: k^n \rightarrow k^n; \quad \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto \begin{pmatrix} v_{\pi(1)} \\ \vdots \\ v_{\pi(n)} \end{pmatrix};$$

ihre Abbildungsmatrix (bezüglich der Standardbasis von  $k^n$ ) bezeichnen wir mit  $P_\pi$ . Da  $\varphi_\pi$  den  $i$ -ten Koordinateneinheitsvektor von  $k^n$  auf den  $\pi(i)$ -ten abbildet, hat  $P_\pi$  in jeder Zeile und jeder Spalte genau einen nichtverschwindenden Eintrag: In der  $i$ -ten Spalte steht an der Position  $\pi(i)$  eine Eins, überall sonst stehen Nullen. Der Eintrag  $p_{k\ell}$  ist also genau dann gleich Eins, wenn  $k = \pi(\ell)$  ist, ansonsten ist er Null.

Die Hintereinanderausführung der Abbildungen  $\varphi_\pi$  und damit die Multiplikation der Matrizen  $P_\pi$  ist vollständig bestimmt durch die Hintereinanderausführung der zugehörigen Permutationen: Wegen

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

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

Auch die Multiplikation der Permutationsmatrix  $P_\pi$  mit einer beliebigen Matrix (passender Größe) läßt sich leicht ausführen: Für  $\pi \in \mathfrak{S}_n$  und  $A \in k^{n \times m}$  betrachten wir zu  $A$  die lineare Abbildung  $\psi: k^m \rightarrow k^n$ , die einen Vektor  $\vec{v} \in k^m$  auf  $A\vec{v}$  abbildet; dann ist  $P_\pi A$  die Abbildungsmatrix von  $\varphi_\pi \circ \psi$ . Ihre Spalten sind die Bilder der Koordinateneinheitsvektoren von  $k^m$ . Der  $i$ -te dieser Vektoren wird von  $\psi$  auf



den  $i$ -ten Spaltenvektor von  $A$  abgebildet, und auf dessen Komponenten wird die Permutation  $\pi$  angewandt. Dies passiert für jede Spalte von  $A$ , insgesamt entsteht  $P_\pi A$  somit aus  $A$ , indem man die Permutation  $\pi$  auf dessen *Zeilen* anwendet. Für die Transposition  $\pi = (12) \in \mathfrak{S}_3$  etwa ist

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

und

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

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

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

und

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

### k) Die LR-Zerlegung einer Matrix

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

Für ein einziges lineares Gleichungssystem ist dieser Lösungsweg nicht sonderlich interessant, denn schon der Aufwand zur Berechnung

von  $A^{-1}$  ist größer als der zur direkten Lösung des Gleichungssystems: Wir müssen dazu schließlich mehrere lineare Gleichungssysteme simultan lösen. Hat man aber viele lineare Gleichungssysteme, die sich nur durch ihre rechte Seiten unterscheiden, wie dies zum Beispiel bei linearen Steuerungsproblemen der Fall ist, so kann man sehr viel Zeit sparen, wenn man zunächst ein für alle Mal die inverse Matrix berechnet und dann die einzelnen Gleichungssysteme für den Preis einer Matrix-Vektor-Multiplikation lösen kann. Hinzu kommt, daß man beispielsweise bei eingebauten Steuerungen die Matrixinversion vorab auf einem leistungsfähigen Computer durchführen kann und dann für die eigentliche Steuerung nur die Matrix-Vektor-Multiplikation implementieren muß.

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

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

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

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

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

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

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

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

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

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

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

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

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

Der Gesamteffekt dieser Umformungen kann auch so beschrieben werden, daß wir die Ausgangsgleichung  $AX = E$  mit einer unteren Dreiecksmatrix  $Z$  mit Einsen in der Hauptdiagonale multipliziert haben, wobei als Koeffizientenmatrix eine obere Dreiecksmatrix  $R = ZA$  entstand. Die umgeformte Gleichung ist also

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

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

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

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

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

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

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

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

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

Betrachten wir dazu ein konkretes Beispiel:

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

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

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

und führt auf

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

Wir schreiben die Einheitsmatrix daneben:

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

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

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

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

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

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

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

In der Tat rechnet man leicht nach, daß

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

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

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

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

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

Somit ist

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

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

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

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

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