

# Einführung in die Matrizenrechnung

Auszug aus dem Skriptum Ausgleichsrechnung I

von

Gerhard Navratil und Martin Staudinger

Institut für Geoinformation und Kartographie  
Technische Universität Wien  
Gußhausstraße 27-29  
1040 Vienna, Austria

## 1 Lineare Algebra

Die lineare Algebra beschäftigt sich mit der Lösung von linearen Gleichungen und linearen Gleichungssystemen. Lineare Gleichungssysteme sind für uns vor allem im Rahmen der „Methode der kleinsten Quadrate“ wichtig. Das Werkzeug für den Umgang mit linearen Gleichungssystemen ist die Matrizenrechnung. Viele der von uns verwendeten Berechnungsmethoden und Algorithmen werden in Matrixschreibweise angegeben und behandelt. Somit ist das Beherrschen der Matrizenrechnung in weiterer Folge unentbehrlich<sup>1</sup>.

Ein zentrales Problem der linearen Algebra ist die Lösung linearer Gleichungssysteme, wie z.B. des folgenden:

$$\begin{aligned}8x_1 + 1x_2 + 6x_3 &= 15 \\3x_1 + 5x_2 + 7x_3 &= 15 \\4x_1 + 9x_2 + 2x_3 &= 15\end{aligned}\tag{1}$$

oder (etwas allgemeiner ausgedrückt):

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3\end{aligned}\tag{2}$$

---

<sup>1</sup>Die Matrizenrechnung wurde übrigens in den 1950er-Jahren in die Geodäsie eingeführt. Zu den ersten Pionieren zählen dabei unter anderem Gotthardt (1951), (Hirvonen, 1958), Wolf (1959) und (Linkwitz, 1960)

Dabei treten die Koeffizienten  $a_{11}$  bis  $a_{33}$ , die Unbekannten  $x_1$  bis  $x_3$  und die Konstanten („rechte Seite“)  $b_1$  bis  $b_3$  auf. Lineare Gleichungssysteme sind dadurch definiert, dass die Unbekannten nur in der Potenz 0 und 1 vorkommen<sup>2</sup>. Ein Gleichungssystem heißt *inhomogen*, wenn mindestens ein  $b_i$  ungleich Null ist. Andernfalls heißt das Gleichungssystem *homogen*. Obiges Gleichungssystem besteht aus 3 Gleichungen in 3 Unbekannten und ist - unter bestimmten Voraussetzungen, die wir in weiterer Folge noch näher betrachten werden - eindeutig lösbar. Lineare Gleichungssysteme mit mehr Gleichungen als Unbekannten heißen *überbestimmt*, solche mit mehr Unbekannten als Gleichungen *unterbestimmt*.

Das Wort „Algebra“ (das im Übrigen im Deutschen auf der ersten Silbe betont wird, im Österreichischen hingegen oft auf der zweiten) kommt aus dem Arabischen. Die wörtliche Übersetzung heißt „Wiederherstellung“: Eines der ersten algebraischen Lehrbücher hieß *Hisab al-gabr w'al-muqabala* - „Wiederherstellen und Zusammenführen“ und wurde um 800 von *Al-Chwarismi*<sup>3</sup> geschrieben. Es behandelt das Auflösen von Gleichungen.

Al-Chwarismi's Buch über Algebra verdanken wir übrigens nicht nur das Wort „Algebra“ selbst. Als sein Buch ins Lateinische übersetzt wurde, wurde Al-Chwarismi zu „Algoritmi“ - unser Wort „Algorithmus“ kommt davon.

In weiterer Folge war „Algebra“ die Bezeichnung für die Lehre vom „Auflösen von Gleichungssystemen und Ungleichungssystemen“. Die klassische Algebra beschränkte sich dabei auf die elementaren Operationen Addition, Subtraktion, Multiplikation, Division, das Potenzieren und das Radizieren<sup>4</sup>. Nicht dazu gehören somit Exponentialgleichungen, Logarithmusgleichungen und trigonometrische (goniometrische) Gleichungen, also Gleichungen die z.B. Ausdrücke wie  $e^x$ ,  $\lg x$  oder  $\sin x$  enthalten. Sie werden auch als *transzendente* Gleichungen bezeichnet.

Die moderne Algebra beschäftigt sich heute nicht nur mit Gleichungssystemen und elementaren Operationen zu ihrer Auflösung, sondern generell und sehr formal mit den Beziehungen mathematischer Größen untereinander, ihren Strukturen, Regeln und Operationen. Die *lineare* Algebra befasst sie sich dabei speziell mit dem  $n$ -dimensionalen Vektorraum und mit linearen Transformationen in ihm.

Neben dieser Bedeutung des Wortes „Algebra“ als ein Teilgebiet der Mathematik gibt es auch noch eine weitere wichtige Bedeutung. Eine mathematische Struktur wird ebenfalls als Algebra bezeichnet, wenn sie bestimmte Eigenschaften erfüllt. Diese Eigenschaften betreffen unter anderem das Assoziativ-, das Kommutativ- und das Distributivgesetz, sowie das Vorhandensein eines neutralen und eines inversen Elements. Ein Beispiel dafür sind beispielsweise die rationalen Zahlen mit den Operationen Addition und Multiplikation. Unter diesen Gesichtspunkten können wir jedoch auch die Menge der Matrizen und ihrer Operationen als eine Algebra bezeichnen.

<sup>2</sup>Es kommen auch keine gemischten Terme vor, wie das bei quadratischen Gleichungen in mehreren Variablen der Fall ist.

<sup>3</sup>*Abu Ja'far Muhammad ibn Musa Al-Chwarismi*, persischer Gelehrter (Mathematiker, Astronom, Geograph und Historiker), ca. 780 - ca. 850

<sup>4</sup>Wurzelziehen

## 2 Matrizenalgebra

Der Begriff der Matrix wurde von Sylvester<sup>5</sup> geprägt (Sylvester, 1850). Mittlerweile gibt es eine Vielzahl einführender Bücher über Matrizenrechnung (z.B. Antosik, 1985; Ayres, 1978; Bellman, 1960; Eves, 1980; Horn, 1985; Schmidt, 1998; Zurmühl und Falk, 1984). Matrizenrechnung ist mittlerweile jedoch auch fester Bestandteil vieler allgemeiner Bücher über Mathematik (z.B. Reinhardt und Soeder, 1991) und vieler Bücher über Ausgleichsrechnung (z.B. Höpcke, 1980; Niemeier, 2002; Perović, 2005; Reißmann, 1976)

### 2.1 Definitionen

Eine  $(m, n)$ -Matrix ist eine (im Allgemeinen rechteckige) Anordnung von  $m \times n$  Elementen in  $m$  Zeilen und  $n$  Spalten:

$$(a_{ik}) := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} =_m \mathbf{A}_n. \quad (3)$$

Als Elemente einer Matrix erlaubt sind Variablen, Zahlen  $\in \mathbb{C}$  (oder Untermengen davon, also  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  oder  $\mathbb{R}$ ), Polynome, Differentiale, sonstige Operatoren (Funktionen) und Symbole aber auch selbst wieder Matrizen sein. Speziell Polynome und Differentiale als Elemente werden im Weiteren wichtig sein. Ist nicht anders angegeben, so werden die von uns betrachteten Matrizen immer reelle Zahlen als Elemente enthalten, oder Variablen, die an Stelle von reellen Zahlen stehen.

Der Typ der Matrix wird charakterisiert durch die Anzahl der Zeilen und Spalten (auch bezeichnet als *Dimension* oder *Größe*). Hat eine Matrix die gleiche Anzahl von Zeilen und Spalten so heißt sie *quadratische Matrix* (oder genauer: eine *n-reihige quadratische Matrix*). In allen anderen Fällen sprechen wir von einer *rechteckigen Matrix*. Die Dimension wird als Ausdruck (Zeilen x Spalten) angegeben. Eine  $(m \times 1)$ -Matrix hat somit nur eine Spalte und heißt *Spaltenvektor* und eine  $(1 \times n)$ -Matrix *Zeilenvektor*. *Skalare*, also „einzelne“ Zahlen, können - mit bestimmten Einschränkungen<sup>6</sup> - als  $(1 \times 1)$ -Matrizen aufgefasst werden.

Die einzelnen Elemente innerhalb einer Matrix werden über ihren *Index* angesprochen: Der *Zeilenindex* gibt die Zeile an, in der sich das Element befindet und der *Spaltenindex* die Spalte. Meist wird zuerst der Zeilen- und dann der Spaltenindex angegeben.  $a_{51}$  ist demnach das Element in der fünften Zeile und der ersten Spalte.

Wir können nun (ohne uns um die Art der Berechnung zu kümmern) das Gleichungssystem (1) mit Matrizen ausdrücken als

$$\mathbf{Ax} = \mathbf{b}. \quad (4)$$

Die auftretenden Matrizen sind dabei die quadratische *Koeffizientenmatrix*  $\mathbf{A}$ , den *Konstantenvektor*  $\mathbf{b}$  und den *Unbekanntenvektor*  $\mathbf{x}$ . Die Elemente der Matrizen entsprechen dabei jeweils den konkreten Elementen wie in (1) angegeben.

<sup>5</sup>James Joseph Sylvester, englischer Mathematiker, 1814 - 1897

<sup>6</sup>z.B. Matrizen können zwar mit einem Skalar, nicht aber mit einer  $(1 \times 1)$ -Matrix multipliziert werden, wie wir später noch sehen werden.

Im Unbekanntenvektor stehen Variablen für die Lösungen des Gleichungssystems. Unser Ziel ist es nun, einen oder mehrere Lösungsvektoren  $\mathbf{x}$  zu finden, der statt der Variablen solche reelle Zahlen enthält, dass Gleichung (4) erfüllt ist.

$$\mathbf{A} = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 15 \\ 15 \\ 15 \end{bmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (5)$$

Beim Anschreiben von Matrizen ist sowohl die Verwendung runder als auch eckiger Klammern erlaubt. Wir werden im Folgenden eckige Klammern für Matrizen mit Zahlen und runde Klammern für alle anderen Matrizen verwenden. Die Lesbarkeit erhöht sich auch, wenn blockweise auftretende Nullen in Matrizen nicht angeschrieben sind, also zum Beispiel

$$\mathbf{M} = \begin{bmatrix} 4 & & 1 & & \\ & & 5 & & 2 \\ & & & 6 & \\ & & & & 3 & 8 \end{bmatrix} \quad \text{an Stelle von } \mathbf{M} = \begin{bmatrix} 4 & 0 & 1 & 0 \\ 0 & 5 & 2 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 3 & 8 \end{bmatrix}.$$

Matrizen sind übrigens keine neue „Erfindung“ der Mathematik. Der Kupferstich „Die Melancholie“ (Melencolia I) von *Albrecht Dürer*<sup>7</sup> aus dem Jahr 1514 enthält bereits die Darstellung einer Matrix (siehe Abbildung 1). Bei der Matrix handelt es sich um eine spezielle Matrix, ein sogenanntes „magisches Quadrat“. Die Summen der Zahlen ist für jede Zeile oder Spalte und für jede Diagonale gleich groß (hier: 34). Die Matrix  $\mathbf{A}$  aus (5) war übrigens ebenfalls ein magisches Quadrat (mit der Summe 15). Zusätzlich enthält die Matrix in Abbildung 1 in der letzten Zeile das Entstehungsjahr des Werkes und Zahlen, welche nach Meinung von Astrologen - angeblich den Planeten Jupiter repräsentieren und dem „schädlichen“ Einfluss des Saturns (der durch andere Symbole auf dem Bild repräsentiert ist) entgegenwirken.

### 2.1.1 Submatrizen

Innerhalb einer  $(m, n)$ -Matrix kann man jeden  $(p, q)$ -Block von Elementen mit  $p \leq m$  und  $n \leq q$  selbst wieder als Matrix auffassen. Dieser (rechteckige oder quadratische) Block heißt *Submatrix* der ursprünglichen Matrix (Ausgangsmatrix). Wir können beispielsweise die Matrix  $\mathbf{A}$  aus (5) zerlegen in

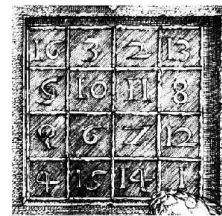
$$\mathbf{A} = \left[ \begin{array}{cc|c} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{array} \right] = \begin{pmatrix} \mathbf{P} & \mathbf{q} \\ \mathbf{r} & \mathbf{s} \end{pmatrix}.$$

Dabei entstehen die Submatrizen  $\mathbf{P}$ ,  $\mathbf{q}$  (ein Spaltenvektor),  $\mathbf{r}$  (ein Zeilenvektor), sowie die  $(1,1)$ -Matrix  $\mathbf{s}$  mit

$$\mathbf{P} = \begin{bmatrix} 8 & 1 \\ 3 & 5 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 6 \\ 7 \end{bmatrix}, \quad \mathbf{r} = [4 \quad 9], \quad \mathbf{s} = [2].$$

<sup>7</sup>*Albrecht Dürer*, deutscher Maler, Grafiker, Mathematiker und Kunsttheoretiker, 1471 - 1528

Abbildung 1: „Die Melancholie“ von *Albrecht Dürer* zeigt rechts oben die Darstellung einer Matrix (siehe vergrößerter Ausschnitt rechts). Die Matrix enthält ein „magisches Quadrat“.



### 2.1.2 Diagonal-, Dreiecks- und Treppenform einer Matrix

Als *Hauptdiagonale* einer  $(m,n)$ -Matrix bezeichnet man die Diagonale die durch die Elemente mit gleichem Zeilen- und Spaltenindex gebildet wird. Das sind somit die Elemente  $a_{11}, a_{22}, \dots, a_{mm}$  für eine Matrix mit  $m \leq n$  bzw. die Elemente  $a_{11}, a_{22}, \dots, a_{nn}$  für eine Matrix mit  $m \geq n$ . Bei einer quadratischen Matrix sind dies alle Elemente von der linken oberen bis zur rechten unteren Ecke.

Als *Diagonalmatrix* bezeichnen wir eine Matrix, bei der alle Elemente außerhalb der Hauptdiagonalen  $= 0$  sind.

$$a_{ij} = 0, \forall i \neq j \quad (6)$$

Ein Spezialfall der Diagonalmatrix ist die *Nullmatrix*, bei der auch die Elemente der Hauptdiagonale  $= 0$  sind.

Eine *Dreiecksmatrix* ist eine quadratische Matrix, deren Elemente unter- oder oberhalb der Hauptdiagonale alle Null sind. Die Elemente  $\neq 0$  bilden also ein Dreieck. In einer *oberen Dreiecksmatrix* sind nur die Hauptdiagonale und Elemente oberhalb von ihr belegt und alle Elemente unterhalb der Hauptdiagonalen  $= 0$ :

$$a_{ij} = 0, \forall i > j. \quad (7)$$

In einer *unteren Dreiecksmatrix* sind die Hauptdiagonale und Elemente unterhalb von ihr belegt:

$$a_{ij} = 0, \forall i < j. \quad (8)$$

Eine obere Dreiecksmatrix wird auch als *rechte Dreiecksmatrix* bezeichnet, da alle belegten Elemente im rechten, oberen Bereich angesiedelt sind. Für eine solche Matrix wird dann oft der Buchstabe **R** verwendet, bzw. entsprechend in der englischsprachigen Literatur die Bezeichnungen **U** („upper“) verwendet.

Eine untere Dreiecksmatrix wird auch als *linke Dreiecksmatrix* bezeichnet und dann - sowohl im Deutschen wie auch im Englischen - mit **L** („lower“) abgekürzt.

Wenn man die Definitionen 7 und 8 betrachtet, sieht man sofort, dass eine quadratische Diagonalmatrix gleichzeitig auch eine Dreiecksmatrix (und zwar sowohl eine obere, als auch eine untere Dreiecksmatrix) ist. Nicht-quadratische Matrizen, die die Forderung (8) erfüllen, haben keine strenge Dreiecksform. Solche Matrizen werden als Matrizen in *Treppenform* (engl. *echelon form*) bezeichnet.

### 2.1.3 Symmetrische und schief-symmetrische Matrizen

Als *symmetrisch* bezeichnet man eine Matrix, wenn gilt

$$a_{ij} = a_{ji} \quad \forall i, j \in \{1 \dots n\} \quad (9)$$

und als *schief-symmetrisch*, wenn gilt

$$a_{ij} = -a_{ji}, \quad \forall i, j \in \{1 \dots n\}. \quad (10)$$

Eine schief-symmetrische Matrix darf also der Hauptdiagonalen ausschließlich Nullen haben, da andernfalls die Forderung  $a_{ii} = -a_{ii}$  nicht erfüllt wäre.

### 2.1.4 Gleichheit von Matrizen

Zwei Matrizen A und B sind genau dann *gleich*, wenn sie gleichen Typ haben und die Elemente an den entsprechenden Positionen in beiden Matrizen gleich sind, d.h.

$$a_{ij} = b_{ij}, \quad \forall i \in \{1 \dots m\}, j \in \{1 \dots n\}. \quad (11)$$

## 2.2 Spur und Determinante

Spur und Determinante sind nur für quadratische Matrizen definiert. Sie sind Funktionen, jeder eine bestimmte Zahl zu.

### 2.2.1 Spur einer Matrix

Die *Spur* einer quadratischen Matrix wird abgekürzt mit  $\text{tr}(\mathbf{A})$  (entsprechend dem englischen Ausdruck *trace*). Sie ist die Summe der Hauptdiagonal-Elemente der Matrix. Für eine  $(n \times n)$ -Matrix gilt also

$$\text{tr}(\mathbf{A}) = a_{11} + a_{22} + \dots + a_{nn}. \quad (12)$$

### 2.2.2 Determinanten

Die *Determinante* (genauer: die *n-reihige Determinante*) einer quadratischen Matrix (abgekürzt mit  $\det(\mathbf{A})$ ) ist rekursiv definiert über den *Laplace'schen*<sup>8</sup> Entwicklungssatz:

---

<sup>8</sup>*Pierre-Simon Marquis de Laplace*, französischer Mathematiker, Astronom und Physiker, 1749 - 1827

$$\begin{aligned}\det(\mathbf{A}) = |\mathbf{A}| &= \sum_{i=1}^n a_{ik} \cdot (-1)^{i+k} |\mathbf{A}^{(ik)}| \\ &= \sum_{k=1}^n a_{ik} \cdot (-1)^{i+k} |\mathbf{A}^{(ik)}|\end{aligned}\tag{13}$$

für  $n > 1$  und

$$\det(\mathbf{A}) = a_{11}\tag{14}$$

für  $n = 1$ .

Die Matrix  $\mathbf{A}^{(ik)}$  erhalten wir aus  $\mathbf{A}$  durch Streichen der  $i$ -ten Zeile und  $k$ -ten Spalte. Diese Matrix heißt daher auch *Streichungsmatrix*. Die Determinante  $|\mathbf{A}^{(ik)}|$  der Streichungsmatrix wird *Minor* genannt und das Skalar  $(-1)^{i+k} |\mathbf{A}^{(ik)}|$  ist der *Kofaktor* (auch: *algebraisches Komplement*) von  $a_{ik}$  in  $\mathbf{A}$ .

Wie man aus (13) erkennt, kann man die Determinante auf zwei verschiedene Arten errechnen: *spaltenweise* (obere Formel, der Index  $i$  ist der Laufindex) oder *zeilenweise* (untere Formel,  $k$  ist der Laufindex). Der jeweils andere Index darf beliebig gewählt werden und bleibt während der Berechnung fest. Man spricht auch davon, dass man „die Determinante nach der  $k$ -ten Spalte entwickelt“ bzw. „nach der  $i$ -ten Zeile entwickelt“. Man wählt jene Spalte (Zeile), die die meisten Nullen enthält, um den Rechenaufwand zu minimieren. Die Determinante gibt Auskunft über einige wichtige Eigenschaften der Matrix; sie bestimmt (determiniert) sozusagen ihr Verhalten. Daher kommt auch ihr Name<sup>9</sup>.

### 2.2.3 Einige Regeln über Determinanten

- Beim Vertauschen zweier Zeilen (Spalten) in  $\mathbf{A}$  ändert sich das Vorzeichen der Determinante.
- Die Addition (Subtraktion) eines Vielfachen einer Zeile (Spalte) zu einer anderen Zeile (Spalte) lässt die Determinante unverändert.
- Die Determinante einer Dreiecks- oder Diagonalmatrix ist das Produkt der Elemente der Hauptdiagonalen.
- Wenn zwei Zeilen (Spalten) in  $\mathbf{A}$  gleich oder zueinander proportional sind, ist  $\det(\mathbf{A}) = 0$ . Hat eine Matrix eine Determinante gleich Null, so sagt man auch, sie hat eine „verschwindende Determinante“.

### 2.2.4 Reguläre und Singuläre Matrizen

Eine *reguläre* Matrix ist eine quadratische  $(n, n)$ -Matrix mit nichtverschwindender Determinante. Eine *singuläre* Matrix hat eine verschwindende Determinante  $\det(\mathbf{A}) = 0$ .

### 2.2.5 Zwei- und dreireihige Determinanten

Die Determinante einer  $(2,2)$ -Matrix kann direkt angegeben werden:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}\tag{15}$$

<sup>9</sup>vom lat. *determinare*: abgrenzen, bestimmen, festsetzen

Für (3,3)-Matrizen liefert die *Regel von Sarrus*<sup>10</sup> ebenfalls direkt das Ergebnis:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = +a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} \\ - a_{11} \cdot a_{23} \cdot a_{32} - a_{12} \cdot a_{21} \cdot a_{33} - a_{13} \cdot a_{22} \cdot a_{31} \quad (16)$$

Für alle anderen Matrizen gibt es keine solchen Formeln. Ab einer (4,4)-Matrix muss also der Laplace'sche Entwicklungssatz verwendet werden. Es wird sofort klar, dass der Aufwand für die Berechnung sehr schnell (nichtlinear) wächst.

## 2.3 Matrizenoperationen

### 2.3.1 Transposition

Die *Transposition* ist die einfachste Matrizenoperation. Sie hat eine einzige Matrix als Parameter und „stürzt“ der Matrix: Reihen und Spalten tauschen ihre Funktionen, Reihen werden also zu Spalten und umgekehrt. Die so entstandene *transponierte* Matrix erhält die Bezeichnung  $\mathbf{A}^T$  (wenn  $\mathbf{A}$  die Ausgangsmatrix war). Manchmal wird in der Literatur auch die Bezeichnung  $\mathbf{A}'$  verwendet.

$$(a_{ji}^T) := (a_{ij}), \quad \forall i \in \{1 \dots m\}, j \in \{1 \dots n\} \quad (17)$$

Bei elementweiser Betrachtung kann man auch sagen, dass die transponierte Matrix durch Vertauschen der Indizes der Elemente der Ausgangsmatrix entsteht.

Ein Problem der Matrixschreibweise ist die Unterscheidung zwischen einem Spalten- und einem Zeilenvektoren. Bei allen folgenden Betrachtungen werden wir die Transposition zur Unterscheidung verwenden. Ein Vektor  $\mathbf{x}$  bezeichnet dabei einen Spaltenvektor und  $\mathbf{x}^T$  einen Zeilenvektor.

Mit Hilfe der Transposition können wir auch die Definitionen (9) und (10) über symmetrische und schief-symmetrische Matrizen neu formulieren: Eine Matrix ist *symmetrisch*, wenn gilt

$$\mathbf{A} = \mathbf{A}^T \quad (18)$$

und *schief-symmetrisch*, wenn gilt

$$\mathbf{A} = -\mathbf{A}^T. \quad (19)$$

Jede beliebige quadratische Matrix  $\mathbf{A}$  kann als die Summe einer symmetrischen Matrix  $\mathbf{B}$  und einer schief-symmetrischen Matrix  $\mathbf{C}$  dargestellt werden:

$$\mathbf{A} = \mathbf{B}_{sym} + \mathbf{C}_{ssym} \quad (20)$$

mit der symmetrischen Matrix

$$\mathbf{B} = \frac{1}{2} (\mathbf{A} + \mathbf{A}^T) \quad (21)$$

und der schief-symmetrischen Matrix

$$\mathbf{C} = \frac{1}{2} (\mathbf{A} - \mathbf{A}^T). \quad (22)$$

Für die Berechnung benötigen wir allerdings noch Addition und Subtraktion von Matrizen.

<sup>10</sup> Pierre Frédérique Sarrus, französischer Mathematiker, 1798 - 1861



### 2.3.2 Addition und Subtraktion

Addition und Subtraktion von Matrizen sind definiert als Addition bzw. Subtraktion der jeweiligen Elemente (also der Elemente mit gleicher Position) der beiden Matrizen:

$$\mathbf{A} \pm \mathbf{B} = (a_{ik}) \pm (b_{ik}) = (a_{ik} \pm b_{ik}), \quad (23)$$

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

Formal genügen Matrixaddition und -subtraktion den bekannten Rechenregeln für Addition und Subtraktion reeller Zahlen. Eine Einschränkung ist jedoch, dass sie offensichtlich nur für Matrizen desselben Typs definiert sind.

Die Matrizenaddition ist assoziativ, d.h.

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \quad (24)$$

und kommutativ

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}. \quad (25)$$

Das Transponieren einer Summe kann auch summandenweise vorgenommen werden:

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T. \quad (26)$$

Die *Nullmatrix*  $\mathbf{0}$  ist das *neutrale Element* der Matrizenaddition. Sie hat als Elemente ausschließlich Nullen. Die Addition einer beliebigen Matrix zur Nullmatrix (oder umgekehrt) ergibt wieder die Ausgangsmatrix:

$$\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}. \quad (27)$$

### 2.3.3 Multiplikation einer Matrix mit einem Skalar

Die Multiplikation einer Matrix mit einem Skalar  $\alpha$  ist definiert als

$$\alpha \cdot \mathbf{A} = \alpha \cdot (a_{ik}) := (\alpha \cdot a_{ik}), \quad (28)$$

d.h. jedes Element aus  $\mathbf{A}$  wird mit  $\alpha$  multipliziert. Man kann somit natürlich auch aus einer Matrix einen Faktor<sup>11</sup> herausheben, der in allen Elementen enthalten ist.

Die Multiplikation einer Matrix mit einem Skalar ist kommutativ und assoziativ. Für die Multiplikation einer Matrix mit einem Skalar und die Matrizenaddition gilt auch das *Distributivgesetz*.

$$\alpha \mathbf{A} = \mathbf{A} \alpha \quad (29)$$

$$\alpha(\beta \mathbf{A}) = (\alpha\beta) \mathbf{A} \quad (30)$$

$$(\alpha + \beta) \mathbf{A} = \alpha \mathbf{A} + \beta \mathbf{A} \quad (31)$$

$$\alpha(\mathbf{A} + \mathbf{B}) = \alpha \mathbf{A} + \alpha \mathbf{B} \quad (32)$$

<sup>11</sup>Der Faktor kann ein Skalar oder auch eine beliebige Funktion sein. Es darf sich jedoch nicht um eine Matrix handeln.

### 2.3.4 Matrizenmultiplikation und -potenzen

Die Multiplikation zweier Matrizen ist definiert als

$$\mathbf{A} \cdot \mathbf{B} = (a_k^i) \cdot (b_j^k) := \left( \sum_{k=1}^n a_k^i \cdot b_j^k \right) = \mathbf{a}^i \cdot \mathbf{b}_j. \quad (33)$$

Das Produkt  $\mathbf{AB}$  einer  $(m, n)$ -Matrix  $\mathbf{A}$  mit einer  $(n, p)$ -Matrix  $\mathbf{B}$  ist also die  $(m, p)$ -Matrix  $\mathbf{C} = \mathbf{AB}$ , deren Elemente  $e_{ij}$  als skalares Produkt der  $i$ -ten Zeile von  $\mathbf{A}$  (des Zeilenvektors  $a^i$ ) mit der  $j$ -ten Spalte von  $\mathbf{B}$  (dem Spaltenvektor  $b_j$ ) gebildet werden.

Es ist offensichtlich, dass Matrizen nur dann miteinander multipliziert werden können, wenn die Spaltenzahl der ersten Matrix gleich der Zeilenzahl der zweiten Matrix ist. Daher sind auch *Potenzen* einer Matrix (z.B.  $\mathbf{A}^2$ ,  $\mathbf{A}^3$ , etc.) nur für *quadratische* Ausgangsmatrizen möglich. Zusätzlich wird auch klar, warum die Multiplikation mit einem Skalar unterschiedliche ist von der Multiplikation mit einer  $(1,1)$ -Matrix: Letztere ist nur in Ausnahmefällen (als Multiplikation mit einem Zeilen- bzw. Spaltenvektor) möglich.

Ergibt das Produkt  $\mathbf{AB} = \mathbf{0}$  (die *Nullmatrix*), heißt das *nicht*, dass  $\mathbf{A}$  oder  $\mathbf{B}$  (oder gar beide)  $= \mathbf{0}$  sind, sondern nur, dass mindestens eine der beiden Matrizen *singulär* ist. Ein Beispiel für einen solche Fall sind die beiden unten angegebenen Matrizen (wobei die Matrix  $\mathbf{B}$  die vierte Potenz der Matrix  $\mathbf{A}$  ist):

$$\mathbf{A} = \begin{bmatrix} -9 & 11 & -21 & 63 & -252 \\ 70 & -69 & 141 & -421 & 1684 \\ -575 & 575 & -1149 & 3451 & -13801 \\ 3891 & -3891 & 7782 & -23345 & 93365 \\ 1024 & -1024 & 2048 & -6144 & 24572 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -84 & 168 & -420 & 1344 & -5355 \\ 568 & -1136 & 2840 & -9088 & 36210 \\ -3892 & 7784 & -19460 & 62272 & -248115 \\ -1024 & 2048 & -5120 & 16384 & -65280 \end{bmatrix}$$

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Aus  $\mathbf{AB} = \mathbf{AC}$  kann außerdem nicht geschlossen werden, dass  $\mathbf{B} = \mathbf{C}$ ; der umgekehrte Schluss hingegen stimmt immer:

$$\mathbf{B} = \mathbf{C} \implies \mathbf{AB} = \mathbf{AC} \quad (34)$$

Das Matrizenprodukt ist nicht kommutativ, d.h. im Allgemeinen sind  $\mathbf{AB}$  und  $\mathbf{BA}$  verschiedene Matrizen (sofern sie überhaupt auf beide Arten multiplizierbar sind). Es gilt zu beachten, dass es zwei Arten der Multiplikation gibt: „von rechts“ und „von links“. Das wird insbesondere beim Arbeiten mit Matrixgleichungen wichtig - man multipliziert stets beide Seiten in gleicher Weise mit einer Matrix: entweder beide Seiten „von rechts“ oder beide Seiten „von links“.

Die Matrizenmultiplikation ist aber assoziativ, d.h.

$$(\mathbf{A}\mathbf{B})\mathbf{C} = \mathbf{A}(\mathbf{B}\mathbf{C}) \quad (35)$$

Das neutrale Element der Matrizenmultiplikation heißt *Einheitsmatrix*  $\mathbf{I}$  (von der englischen Bezeichnung *identity matrix*; manchmal - im Deutschen - auch als  $\mathbf{E}$  bezeichnet).  $\mathbf{I}$  ist eine quadratische Diagonalmatrix mit den Elementen

$$\mathbf{I}_{ik} = \delta_{ik}. \quad (36)$$

Dabei steht  $\delta_{ik}$  für das *Kroneckersymbol*<sup>12</sup>.

$$\delta_{ik} = \begin{cases} 1 & \text{für } i = k \\ 0 & \text{für } i \neq k \end{cases} \quad (37)$$

Die Multiplikation mit der Einheitsmatrix ist kommutativ und es gilt

$$\mathbf{I}\mathbf{A} = \mathbf{A}\mathbf{I} = \mathbf{A}. \quad (38)$$

Für Matrizenaddition und -multiplikation gilt das Distributivgesetz, d.h.

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{A}\mathbf{B} + \mathbf{A}\mathbf{C} \quad (39)$$

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{C}. \quad (40)$$

Für das *Potenzieren* von Matrizen gelten folgende Regeln:

$$\mathbf{A}^p \cdot \mathbf{A}^q = \mathbf{A}^{p+q} \quad (41)$$

$$(\mathbf{A}^p)^q = \mathbf{A}^{p \cdot q}. \quad (42)$$

Unter Verwendung der Einheitsmatrix können wir die Multiplikation mit einem Skalar doch noch als Matrizenmultiplikation definieren:

$$\alpha \cdot \mathbf{A} = (\alpha \cdot \mathbf{I}) \cdot \mathbf{A}. \quad (43)$$

Die Multiplikation mit einem Skalar kann also auch durch eine Matrizenmultiplikation mit einer Diagonalmatrix, deren Elemente auf der Hauptdiagonalen diesem Skalar entsprechen, erfolgen.

### 2.3.5 Transponieren von Matrizenprodukten

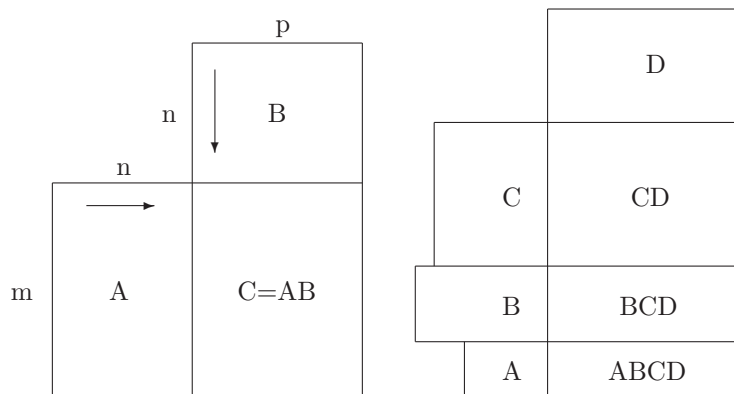
Wird ein Matrizenprodukt transportiert, so kann man stattdessen auch zunächst die Matrix transportieren und dann die Multiplikation in gestürzter Reihenfolge durchführen:

$$(\mathbf{A} \cdot \mathbf{B} \cdot \mathbf{C} \cdot \dots \cdot \mathbf{Z})^T = \mathbf{Z}^T \cdot \dots \cdot \mathbf{C}^T \cdot \mathbf{B}^T \cdot \mathbf{A}^T. \quad (44)$$

### 2.3.6 Das Falk'sche Schema

Manchmal kann es vorkommen, dass man zwei oder mehr Matrizen „von Hand“ multiplizieren muss (z.B. wenn die einzelnen Elemente nicht numerische Zahlen sondern beispielsweise Submatrizen sind). Dann ist eine von *Falk*<sup>13</sup> (1951) vorgeschlagene Anordnung nützlich, bei der jedes Produktelement  $c_{ik}$  genau im

Abbildung 2: Falk'sches Schema zur Matrizenmultiplikation



Kreuzungspunkt der  $i$ -ten Zeile von  $\mathbf{A}$  mit der  $k$ -ten Spalte von  $\mathbf{B}$  erscheint (Abbildung 2 links).

Die Falk'sche Anordnung empfiehlt sich insbesondere bei der Berechnung von Produkten aus mehr als zwei Matrizen (z.B.  $\mathbf{ABCD}$ ). Baut man das Schema dabei von oben nach unten auf (vgl. Abbildung 2 rechts), so beginnt man die Rechnung mit dem letzten Faktor und arbeitet sich somit „von hinten nach vorne“.

### 2.3.7 Determinante und Spur nach Matrizenoperationen

Die Determinante einer  $(n, n)$ -Matrix hat im Zusammenhang mit Matrizenoperationen folgende Eigenschaften:

$$|\alpha \mathbf{A}| = \alpha^n \cdot |\mathbf{A}|, \quad (45)$$

$$|-\mathbf{A}| = (-1)^n \cdot |\mathbf{A}|, \quad (46)$$

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|, \quad (47)$$

$$|\mathbf{A}| = |\mathbf{A}^T|. \quad (48)$$

Aus (47) wird ersichtlich, dass das Produkt zweier Matrizen genau dann singulär wird, wenn zumindest eine der miteinander multiplizierten Matrizen singulär ist.

Für die Spur einer Matrix gilt:

$$\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T), \quad (49)$$

$$\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B}), \quad (50)$$

$$\text{tr}(\alpha \mathbf{A}) = \alpha \text{tr}(\mathbf{A}). \quad (51)$$

### 2.3.8 Rechnen mit Submatrizen

Elementare Matrizenoperationen wie Addition, Subtraktion und Multiplikation können auch durchgeführt werden, wenn die Elemente der einzelnen Matrizen selbst wieder Matrizen (*Submatrizen*) sind. Dabei ist natürlich besonders darauf

<sup>12</sup>Leopold Kronecker, preußischer Mathematiker, 1823 - 1891

<sup>13</sup>Sigurd Falk, deutscher Mathematiker, \* 1920

zu achten, dass die Dimensionen der Submatrizen miteinander korrespondieren („Dimension“ einer Matrix ist hier sowohl im Sinne von „Anzahl der Zeilen mal Anzahl der Spalten“ zu verstehen als auch im Sinne der physikalischen Einheiten der einzelnen Elemente).

### 2.3.9 Die Gauß'sche Transformation

Unter der *Gauß'schen Transformation*<sup>14</sup> einer rechteckigen  $(m, n)$ -Matrix  $\mathbf{A}$  versteht man das Produkt

$$\mathbf{N} = \mathbf{A}^T \mathbf{A}. \quad (52)$$

Das Ergebnis ist eine quadratische, symmetrische  $(n, n)$ -Matrix  $\mathbf{N}$ . Die Elemente der Produktmatrix  $\mathbf{N}$  entstehen aus dem skalaren Produkt zweier Spaltenvektor von  $\mathbf{A}$  (oder eines Spaltenvektors mit sich selbst). Die Diagonalelemente sind stets positiv. Zusätzlich ist die Matrix *positiv definit* (bzw. *semidefinit*, wenn auch Diagonalelemente = 0 vorkommen). Positiv definit bedeutet, dass alle Subdeterminanten, die man durch Streichen der jeweils letzten  $k$  Spalten und Zeilen erhält (mit  $k=0$  bis  $n-1$ ; das sind also alle *Minoren*),  $\geq 0$  sind.

Es gibt einige Hinweise darauf, dass eine Matrix positiv definit ist:

- Die Diagonalelemente jeder positiv definiten Matrix sind positive reelle Zahlen.
- Jede Untermatrix einer positiv definiten Matrix ist positiv definit.
- Spur, Determinante und alle Minoren (Determinanten der Untermatrizen) einer positiv definiten Matrix sind positiv.
- Die Summe  $\mathbf{A} + \mathbf{B}$  zweier beliebiger positiv definiten Matrizen ist ebenfalls positiv definit.
- Eine symmetrische Matrix ist genau dann positiv definit, wenn ihre Eigenwerte positiv sind.

Man kann auch ein Matrizenprodukt, z.B.  $\mathbf{PA}$ , einer Gauß'schen Transformation unterziehen:

$$\mathbf{N} = \mathbf{A}^T \mathbf{PA}. \quad (53)$$

$\mathbf{N}$  ist dann ebenfalls wieder quadratisch, symmetrisch und positiv definit. Aus (47) und (48) folgt jedoch, dass  $\mathbf{N}$  genau dann singulär ist, wenn  $\mathbf{A}$  (oder  $\mathbf{P}$ ) singulär ist.

## 2.4 Orthogonale Matrizen

Orthogonale Matrizen sind quadratische Matrizen, bei denen die Spaltenvektoren (Zeilenvektoren) ein System orthonormaler Einheitsvektoren bilden. Das Skalarprodukt zweier beliebiger Spaltenvektoren (Zeilenvektoren) ergibt immer 0 oder 1:

$$q_i^T q_k = \delta_{ik}$$

mit  $\delta_{ik}$  dem Kroneckersymbol (siehe (34)).

<sup>14</sup>Johann Carl Friedrich Gauß, deutscher Mathematiker, Astronom und Geodät, 1777 - 1855

Anders ausgedrückt: Ist  $\mathbf{Q}$  eine orthogonale Matrix, so gilt

$$\mathbf{Q}^T \mathbf{Q} = \mathbf{I} \text{ bzw. } \mathbf{Q}^T = \mathbf{Q}^{-1}. \quad (54)$$

Ist  $\mathbf{Q}$  eine orthogonale Matrix, dann ist die Determinante  $\det(\mathbf{Q}) = \pm 1$ . Dieser Schluss gilt jedoch nicht umgekehrt, also nicht jede Matrix mit  $\det = \pm 1$  ist orthogonal. Orthogonale Matrizen mit  $\det = 1$  werden auch als *eigentlich orthogonal*, jene mit  $\det = -1$  als *uneigentlich orthogonal* bezeichnet.

Die Definition über die Orthonormalität gilt sowohl für die Spaltenvektoren als auch für die Zeilenvektoren, d.h. wenn  $\mathbf{Q}$  eine orthogonale Matrix ist, dann auch  $\mathbf{Q}^T$ . Die Multiplikation orthogonaler Matrizen ist außerdem kommutativ. Daher gilt:

$$\mathbf{Q}^T \mathbf{Q} = \mathbf{Q} \mathbf{Q}^T = \mathbf{I}. \quad (55)$$

Das Produkt zweier orthogonaler Matrizen ist wieder eine orthogonale Matrix.

In der Vektorrechnung heißen Vektoren, die aufeinander normal stehen *und* die Länge 1 haben, *orthonormal*. In der Matrizenrechnung ist hingegen für Matrizen nach obiger Definition nur der Begriff *orthogonal* üblich (was eigentlich nur auf die Orthogonalität hindeutet, nicht aber auf die Einheitslänge von 1). Sind jedoch Vektoren nur orthogonal und nicht auch gleichzeitig orthonormal, so bilden sie keine orthogonale Matrix, denn  $\det(\mathbf{Q}) \neq \pm 1$  und  $\mathbf{Q}^T \mathbf{Q} \neq \mathbf{I}$ .

Auch bei rechteckigen Matrizen können die Spalten oder Zeilen orthonormale Vektoren sein. Die Bezeichnung Orthogonalmatrix ist jedoch nur für quadratische Matrizen vorgesehen (für rechteckige Matrizen sind weder Determinante noch Inverse definiert).

## 2.5 Inversion

Als *inverse Matrix* oder *Kehrmatrix* einer quadratischen Matrix  $\mathbf{A}$  ist die Matrix  $\mathbf{A}^{-1}$  mit folgender Eigenschaft:

$$\mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}. \quad (56)$$

Es können nur quadratische Matrizen mit Determinante  $\neq 0$  invertiert werden. Zu jeder quadratischen, nicht singulären Matrix gibt es genau eine Inverse.

Die Berechnung der Inversen erfolgt mit Hilfe der auf Seite 7 eingeführten Kofaktoren. Diese werden in der *Kofaktorenmatrix* zusammengefasst. Die transponierte Kofaktorenmatrix heißt dann „die zu  $\mathbf{A}$  *adjungierte* Matrix  $\mathbf{A}^*$ “. Mit ihrer Hilfe können wir die Matrizeninversion definieren als *Regel von Cramer*:

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \mathbf{A}^*. \quad (57)$$

Für die Inversion einer regulären (2,2)-Matrix  $\mathbf{A}$  heißt das, dass die Elemente der Hauptdiagonalen vertauscht und die beiden anderen Elemente mit (-1) multipliziert werden. Die entstehende Matrix muss noch durch  $\det(\mathbf{A})$  dividiert werden:

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad (58)$$

Da für jeden Kofaktor eine Determinante berechnet werden muss, ist der Aufwand für die Inversion größerer Matrizen nach (57) relativ groß. Daher ist die Bedeutung der Regel von Cramer in der Praxis beschränkt.

### 2.5.1 Inversion eines Produktes

Die Inversion eines Matrizenproduktes kann ersetzt werden durch die Multiplikation der Inversen in umgekehrter Reihenfolge:

$$(\mathbf{A}\mathbf{B}\mathbf{C}\dots\mathbf{N})^{-1} = \mathbf{N}^{-1} \dots \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}. \quad (59)$$

Voraussetzung dafür ist natürlich, dass jede einzelne Matrix regulär und quadratisch ist. Das ist nicht notwendigerweise der Fall. Bei der Gauß'schen Transformation kann die Ausgangsmatrix rechteckig sein und das Ergebnis ist trotzdem eine quadratische Matrix.

Außerdem gilt:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I} &\Rightarrow (\mathbf{A}^{-1})^T \cdot \mathbf{A}^T = \mathbf{I}, \\ (\mathbf{A}^T)^{-1} &= (\mathbf{A}^{-1})^T. \end{aligned} \quad (60)$$

### 2.5.2 Weitere Regeln:

Ist  $\mathbf{A}$  eine *orthogonale* Matrix, so gilt wegen (54) und (56) offenbar:

$$\mathbf{A}^{-1} = \mathbf{A}^T. \quad (61)$$

Wenn  $\mathbf{A}$  eine *symmetrische* Matrix ist, dann ist auch  $\mathbf{Q} = \mathbf{A}^{-1}$  symmetrisch.

Ist  $\mathbf{A}$  eine *Diagonalmatrix*, so ist auch  $\mathbf{A}^{-1}$  eine Diagonalmatrix und ihre Elemente berechnen sich aus:

$$q_{ii} = \frac{1}{a_{ii}}. \quad (62)$$

Eine Diagonalmatrix kann demnach nur invertiert werden, wenn keines der Diagonalelemente  $=0$  ist. Das muss aber offensichtlich der Fall sein, da andernfalls die Determinante  $=0$  wäre. Somit kann die Inverse einer Diagonalmatrix direkt angegeben werden.

Auch von folgendem Sonderfall kann die Inverse sofort angeschrieben werden: Gegeben ist eine quadratische Matrix  $\mathbf{A}$ , bei der außer der Hauptdiagonalen nur noch genau eine Spalte oder Zeile besetzt ist. Dann hat auch die Inverse  $\mathbf{A}^{-1}$  diese Form und es gilt:

$$q_{ii} = \frac{1}{a_{ii}}; \quad q_{ik} = \frac{-a_{ik}}{a_{ii}a_{kk}}. \quad (63)$$

### 2.5.3 Unterteilung in Submatrizen

Manchmal ist es hilfreich, eine  $(n, n)$ -Matrix  $\mathbf{A}$  vor der Inversion in Submatrizen zu unterteilen. Das kann zu vereinfachten Berechnungen führen, wenn ganze Blöcke der Matrix nur aus Nullen bestehen. Ausgangspunkt ist also die Zerlegung

$$\mathbf{A} = \begin{pmatrix} p\mathbf{P}_p & p\mathbf{Q}_s \\ s\mathbf{R}_p & s\mathbf{S}_s \end{pmatrix} \quad (64)$$

mit  $p + s = n$ . Die Inverse hat dann eine ähnliche Gestalt:

$$\mathbf{A}^{-1} = \begin{pmatrix} p\tilde{\mathbf{P}}_p & p\tilde{\mathbf{Q}}_s \\ s\tilde{\mathbf{R}}_p & s\tilde{\mathbf{S}}_s \end{pmatrix} \quad (65)$$

mit den einzelnen Submatrizen

$$\tilde{\mathbf{P}} = (\mathbf{P} - \mathbf{Q}\mathbf{S}^{-1}\mathbf{R})^{-1} \quad (66)$$

$$\tilde{\mathbf{Q}} = -(\mathbf{P} - \mathbf{Q}\mathbf{S}^{-1}\mathbf{R})^{-1}(\mathbf{Q}\mathbf{S}^{-1}) \quad (67)$$

$$\tilde{\mathbf{R}} = -(\mathbf{S}^{-1}\mathbf{R})(\mathbf{P} - \mathbf{Q}\mathbf{S}^{-1}\mathbf{R})^{-1} \quad (68)$$

$$\tilde{\mathbf{S}} = \mathbf{S}^{-1} + (\mathbf{S}^{-1}\mathbf{R})(\mathbf{P} - \mathbf{Q}\mathbf{S}^{-1}\mathbf{R})^{-1}(\mathbf{Q}\mathbf{S}^{-1}). \quad (69)$$

Bei der praktischen Berechnung sollte man beachten, dass in den Formeln einige Klammerausdrücke öfters vorkommen. Diese werden günstigerweise nur einmal berechnet und dann mehrfach verwendet. Unter Verwendung der Abkürzungen  $\mathbf{U} = \mathbf{Q}\mathbf{S}^{-1}$  und  $\mathbf{V} = \mathbf{P} - \mathbf{U}\mathbf{R}$  kann man (66) bis (69) auch schreiben als

$$\left( \begin{array}{c|c} \tilde{\mathbf{P}} & \tilde{\mathbf{Q}} \\ \hline \tilde{\mathbf{R}} & \tilde{\mathbf{S}} \end{array} \right) = \left( \begin{array}{c|c} \mathbf{V}^{-1} & -\mathbf{V}^{-1}\mathbf{U} \\ \hline -\mathbf{S}^{-1}\mathbf{R}\mathbf{V}^{-1} & \mathbf{S}^{-1} + \mathbf{S}^{-1}\mathbf{R}\mathbf{V}^{-1}\mathbf{U} \end{array} \right). \quad (70)$$

Alternativ kann man auch folgende Formeln verwenden:

$$\tilde{\mathbf{P}} = \mathbf{P}^{-1} + (\mathbf{P}^{-1}\mathbf{Q})(\mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q})^{-1}(\mathbf{R}\mathbf{P}^{-1}) \quad (71)$$

$$\tilde{\mathbf{Q}} = -(\mathbf{P}^{-1}\mathbf{Q})(\mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q})^{-1} \quad (72)$$

$$\tilde{\mathbf{R}} = -(\mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q})^{-1}(\mathbf{R}\mathbf{P}^{-1}) \quad (73)$$

$$\tilde{\mathbf{S}} = (\mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q})^{-1}. \quad (74)$$

Unter Verwendung der Abkürzungen  $\mathbf{U} = \mathbf{P}^{-1}\mathbf{Q}$  und  $\mathbf{V} = \mathbf{S} - \mathbf{R}\mathbf{U}$  ergeben sich (71) bis (74) zu

$$\left( \begin{array}{c|c} \tilde{\mathbf{P}} & \tilde{\mathbf{Q}} \\ \hline \tilde{\mathbf{R}} & \tilde{\mathbf{S}} \end{array} \right) = \left( \begin{array}{c|c} \mathbf{P}^{-1} + \mathbf{U}\mathbf{V}^{-1}\mathbf{R}\mathbf{P}^{-1} & -\mathbf{U}\mathbf{V}^{-1} \\ \hline -\mathbf{V}^{-1}\mathbf{R}\mathbf{P}^{-1} & \mathbf{V}^{-1} \end{array} \right). \quad (75)$$

Auch gilt wieder, dass einige Klammerausdrücke öfters vorkommen und bei der praktischen Rechnung nur einmal berechnet werden sollten.

Weiters gilt:

$$\left( \begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{0} & \mathbf{C} \end{array} \right)^{-1} = \left( \begin{array}{c|c} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1} \\ \hline \mathbf{0} & \mathbf{C}^{-1} \end{array} \right) \quad (76)$$

$$\left( \begin{array}{c|c} \mathbf{A} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{B} \end{array} \right)^{-1} = \left( \begin{array}{c|c} \mathbf{A}^{-1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{B}^{-1} \end{array} \right) \quad (77)$$

$$\left( \begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{D} & \mathbf{I} \end{array} \right)^{-1} = \left( \begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline -\mathbf{D} & \mathbf{I} \end{array} \right). \quad (78)$$

#### 2.5.4 Weitere wissenswerte Regeln

Man kann die Matrizeninversion auch über eine Reihenentwicklung, die *Neumann'sche Reihe*, berechnen:

$$(\mathbf{I} + \mathbf{A})^{-1} = \mathbf{I} - \mathbf{A} + \mathbf{A}^2 - \mathbf{A}^3 + \mathbf{A}^4 - + \dots \quad (79)$$

Der Beweis der Richtigkeit dieser Gleichung ist einfach. Man multipliziert die Gleichung von links mit  $\mathbf{I} + \mathbf{A}$  und erhält auf der linken Seite die Einheitsmatrix. Auf der rechten Seite fallen alle Glieder außer der Einheitsmatrix weg.



Wenn  $\mathbf{A}^i$  mit wachsendem  $i$  zur Nullmatrix konvergiert, so konvergiert die Reihe offensichtlich ebenfalls. Die Angabe der dazu notwendigen Bedingung für die Matrix  $\mathbf{A}$  ist schwer. Eine mögliche (**nicht** notwendige) Bedingung ist:

Alle Elemente von  $\mathbf{A}$  sind kleiner als  $\frac{1}{u}$ , also z.B.  $\leq \frac{1}{u+1}$ . Ein Element von  $\mathbf{A}^2$  kann höchstens den Wert  $\frac{u}{(u+1)^2}$  erreichen, d.h. die obere Schranke der Absolutbeträge für die Elemente von  $\mathbf{A}^i$  ist  $\frac{u^{i-1}}{(u+1)^i}$  und strebt daher gegen Null.

Eine gute Konvergenz ist bereits dann gegeben, wenn die Elemente von  $\mathbf{A}$  kleiner als  $\frac{1}{2u}$  sind.

Das Aufsummieren der Elemente einer quadratischen Matrix  $\mathbf{A}$  kann ebenfalls mit Matrizenoperationen beschrieben werden:

|                                       |                                       |
|---------------------------------------|---------------------------------------|
| Zeilen summieren                      | $\mathbf{A}\mathbf{e}$ ,              |
| Spalten summieren                     | $\mathbf{e}^T\mathbf{A}$ ,            |
| alle Elemente summieren               | $\mathbf{e}^T\mathbf{A}\mathbf{e}$ ,  |
| die Quadrate aller Elemente summieren | $\text{tr}(\mathbf{A}^T\mathbf{A})$ . |

Der hier verwendete Vektor  $\mathbf{e}$  besteht dabei aus Einsen und er hat so viele Zeilen wie die Matrix  $\mathbf{A}$ . Die Formeln funktionieren prinzipiell auch mit nicht-quadratischen Matrizen, es müssen dann aber für  $\mathbf{e}$  und  $\mathbf{e}^T$  verschieden große Vektoren verwendet werden.

## 2.6 Auflösung von Gleichungssystemen durch Matrizenoperationen

Lineare Gleichungssysteme lassen sich mit Hilfe von Multiplikation, Inversion und Gauß'scher Transformation lösen. Gehen wir aus von einem Gleichungssystem der Form (4). Die Berechnung des unbekanntes Vektors  $\mathbf{x}$  kann dadurch erfolgen, dass beide Seiten der Gleichung von links mit  $\mathbf{A}^{-1}$  multipliziert werden

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \quad (80)$$

Da  $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$  gilt, folgt daraus:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \quad (81)$$

Die Auflösung setzt voraus, dass die einzelnen Matrizen in der angegebenen Form miteinander verknüpfbar sind, also:

- Anzahl der Zeilen in  $\mathbf{A}$  = Anzahl der Elemente in  $\mathbf{b}$  und
- Anzahl der Spalten in  $\mathbf{A}$  = Anzahl der Elemente in  $\mathbf{x}$ .

Außerdem muss die Matrix  $\mathbf{A}$  invertierbar sein, d.h. sie muss quadratisch und regulär sein. Man spricht dann von einem eindeutig lösbares Gleichungssystem. In der Ausgleichsrechnung haben wir es jedoch meist mit überbestimmten Gleichungssystemen zu tun. Die Matrix  $\mathbf{A}$  wird also nicht quadratisch („mehr Gleichungen als Unbekannte“) und somit nicht invertierbar sein. Das Gleichungssystem muss also anders gelöst werden.

Ohne auf die näheren Hintergründe einzugehen, wenden wir nun folgenden „Trick“ an: Wir unterwerfen das Gleichungssystem einer Gauß'schen Transformation, d.h. wir multiplizieren beiden Seiten von links mit  $\mathbf{A}^T$ :

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}. \quad (82)$$

(82) wird auch *Normalgleichungssystem* genannt und die Matrix  $\mathbf{N} = \mathbf{A}^T \mathbf{A}$  die *Normalgleichungsmatrix*. Die Matrix  $\mathbf{N}$  ist quadratisch und symmetrisch. Falls sie auch regulär ist, können wir sie invertieren und somit das Gleichungssystem lösen:

$$\mathbf{x} = \mathbf{N}^{-1} \mathbf{A}^T \mathbf{b} \quad (83)$$

mit  $\mathbf{N} = \mathbf{A}^T \mathbf{A}$ . Die Anwendung von Normalgleichungen ist eine zentrale Idee der „Methode der kleinsten Quadrate“, die von Gauß (1809, 1829).

## 2.7 Lineare Abhängigkeit und der Rang einer Matrix

In der Vektorrechnung gibt es folgende wichtige Definition: Ein  $k$ -Tupel von Vektoren  $(x_1, \dots, x_k)$  ist *linear abhängig*, wenn es reelle Zahlen  $(\lambda_1, \dots, \lambda_k)$  gibt für die gilt:

$$\sum_{i=1}^k \lambda_i x_i = 0. \quad (84)$$

Dabei dürfen nicht alle  $\lambda_i = 0$  sein. Dementsprechend gilt, dass Vektoren *linear unabhängig* sind, wenn aus (84) folgt

$$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0. \quad (85)$$

Betrachtet man nun die Zeilen (oder Spalten) einer Matrix als Vektoren, so können zwischen ihnen lineare Abhängigkeiten bestehen. Die maximale Anzahl der linear unabhängigen Zeilen (Spalten) wird *Rang* der Matrix  $\mathbf{A}$  bezeichnet. Manchmal wird der Rang auch als *Zeilenrang* bzw. *Spaltenrang* bezeichnet. Der Rang wird abgekürzt als  $\text{rank}(\mathbf{A})$  bzw.  $\text{rk}(\mathbf{A})$ . Die Anzahl der verschiedenen linearen Abhängigkeiten heißt *Rangdefekt* oder *Rangdefizit*  $d$ .

Bei einer quadratischen  $(n, n)$ -Matrix  $\mathbf{A}$  gilt

$$\text{rank}(\mathbf{A}) = n - d \quad (86)$$

mit dem Rangdefekt  $d$ .

Der Rang einer quadratischen Matrix  $\mathbf{A}$  kann maximal  $= n$  sein. In diesem Fall ist  $d = 0$  und  $\det(\mathbf{A}) \neq 0$  ( $\mathbf{A}$  ist also regulär und invertierbar). Wenn  $\text{rank}(\mathbf{A}) < n$  gilt, dann ist handelt es sich um eine singuläre (also nicht invertierbare) Matrix und es gilt somit  $\det(\mathbf{A}) = 0$ .

Für eine rechteckige  $(m, n)$ -Matrix kann der Rang maximal gleich der kleineren der beiden Zahlen  $m$  und  $n$  sein.

Zur eindeutigen Lösung des Gleichungssystems  $\mathbf{A}\mathbf{x} = \mathbf{b}$  darf die  $(n, n)$ -Matrix  $\mathbf{A}$  kein Rangdefizit aufweisen. Der Rang von  $\mathbf{A}$  muss außerdem gleich dem Rang des um den Vektor  $\mathbf{b}$  erweiterten Systems sein, also

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}, \mathbf{b}) = n. \quad (87)$$

Zeilen- und Spaltenrang einer Matrix sind immer gleich wie eine einfache Überlegung zeigt: Bei einer Matrix mit 3 Zeilen und 5 Spalten gibt es 3 Zeilenvektoren im  $\mathbb{R}^5$ . Es kann also maximal 3 linear unabhängige Zeilenvektoren geben. Andererseits gibt es 5 Spaltenvektoren im  $\mathbb{R}^3$ . Es kann aber im  $\mathbb{R}^3$  maximal 3 unabhängige Vektoren geben und somit sind Zeilen- und Spaltenrang gleich

groß. Da Zeilen- und Spaltenrang gleich sind, kann man Zeilen und Spalten zu vertauschen ohne den Rang zu ändern und es gilt:

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T). \quad (88)$$

Jetzt kann auch Regularität bzw. Singularität mit Hilfe des Ranges einer Matrix neu definiert werden: Eine  $(n, n)$ -Matrix  $\mathbf{A}$  ist regulär wenn sie *vollen Rang* hat, d.h. wenn kein Rangdefizit auftritt. Eine Matrix mit Rangdefekt ist singular.

### 2.7.1 Bestimmung des Ranges einer Matrix - Gauß'scher Algorithmus

Um den Rang einer Matrix zu bestimmen bringt man die Matrix durch *elementare Umformungen* auf Treppenform (bzw. quadratische Matrizen auf eine obere Dreiecksform). Man schaut also, dass in jeder Spalte alle Elemente unterhalb der Hauptdiagonalelemente = 0 werden. Die grundlegende Idee dabei ist, dass sich der Rang einer Matrix durch die angewendeten elementaren Umformungen nicht ändert. Zulässigen Umformungen dabei sind:

- Vertauschen zweier Zeilen (oder Spalten);
- Multiplizieren einer Zeile (Spalte) mit einem Faktor  $c \neq 0$ ;
- Addieren einer mit  $c$  multiplizierten Zeilen (Spalte) zu einer anderen Zeile (Spalte).

Nach Abschluss des Verfahrens (Treppen- bzw. Dreiecksform der Matrix) muss noch geprüft werden, ob eine der Zeilen (Spalten) das Vielfache einer anderen Zeile (Spalte) ist. Ist das der Fall, so beseitigt man eine der beiden betroffenen Zeilen (Spalten) durch die dritte der oben genannten elementaren Umformungen. Anschließend ist die Anzahl der nicht-verschwindenden Zeilen (Spalten) gleich dem Rang der Matrix („nicht-verschwindend“ heißt in diesem Fall, dass die Zeile/Spalte kein Nullvektor ist).

Diese Vorgangsweise lässt sich auch zur Lösung eines linearen Gleichungssystems anwenden und heißt *Gauß'scher Algorithmus* oder *Gauß'sche Elimination* (Zurmühl und Falk, 1984, S. 58-60). Die elementaren Umformungen können dabei auch als Matrizenmultiplikation dargestellt werden.

Bildet man eine Matrix  $\mathbf{J}_{ik}$ , indem man in einer Einheitsmatrix die  $i$ -te mit der  $k$ -ten Zeile vertauscht und multipliziert  $\mathbf{A}$  „von links“ mit dieser Matrix  $\mathbf{J}$ , so bewirkt dies ein Vertauschen der  $i$ -ten mit der  $k$ -ten Zeile in der Matrix  $\mathbf{A}$ . Bildet man eine Matrix  $\mathbf{K}_{ik}$ , indem man in einer Einheitsmatrix die 0 an der Stelle  $i, k$  ( $i \neq k$ ) durch eine Zahl  $c$  ersetzt, und multipliziert  $\mathbf{KA}$ , so bewirkt dies ein Addieren der  $c$ -fachen  $k$ -ten Zeile zur  $i$ -ten Zeile in  $\mathbf{A}$ . Ist  $c$  negativ, so entspricht dies einer Subtraktion.

Im Fall, dass man nicht zeilen- sondern spaltenweise vorgeht, multipliziert man einfach „von rechts“, d.h.  $\mathbf{AJ}$  vertauscht die  $i$ -te mit der  $k$ -ten Spalte, bzw.  $\mathbf{AK}$  addiert das  $c$ -fache der  $i$ -ten Spalte zur  $k$ -ten Spalte.

Der Gauß'sche Algorithmus ist hervorragend für die „händische“ Auflösung linearer Gleichungssysteme und zum Invertieren von Matrizen geeignet. Der Grund dafür ist die Übersichtlichkeit des Verfahrens. Der Nachteil ist die Anzahl der Reduktionsschritte, die bei einer  $(n, n)$ -Matrix  $\frac{n \cdot (n-1)}{2}$  betragen.

### 2.7.2 Zeilennormalform einer Matrix - Gauß-Jordan-Verfahren

Der Gauß'sche Algorithmus hat durch elementare Spalten- oder Zeilenumformungen zur Treppenform der Matrix geführt. *Jordan*<sup>15</sup> hat das Verfahren abgewandelt, indem er einerseits nur Umformungen der *Zeilen* zugelassen hat, andererseits aber als Ziel der Zeilenumformungen nicht eine Matrix in Treppenform, sondern eine in Zeilennormalform angegeben hat (Press et al., 1986, S. 24-29).

Eine  $(m, n)$ -Matrix  $\mathbf{A}$  besitzt *Zeilennormalform*, wenn sie folgende Eigenschaften erfüllt:

1. Unterhalb der Hauptdiagonalen stehen nur Nullen, d.h.

$$a_{ij} = 0, \forall i > j;$$

2. Das erste nicht-verschwindende Element jeder Zeile (von links gesehen) ist gleich 1;
3. Ist  $a_{ij}$  das erste nicht-verschwindende Element der  $i$ -ten Zeile, so ist

$$a_{kj} = 0, \forall k \neq i,$$

d.h. oberhalb und unterhalb des Elements  $a_{ij} = 1$  stehen lauter Nullen in der  $j$ -ten Spalte.

Bei einer quadratischen, regulären Matrix ergibt sich als Zeilennormalform die Einheitsmatrix. Es lässt sich jedoch jede beliebige Matrix durch elementare Zeilenumformungen in Zeilennormalform transformieren. Die Vorgangsweise dafür ist der *Gauß-Jordan-Algorithmus*:

1. Zunächst suchen wir - von links beginnend - in der gegebenen Matrix  $\mathbf{A}$  die erste vom Nullvektor verschiedene Spalte. Die ersten  $(l - 1)$  Spalten seien Nullvektoren, die  $l$ -te Spalte ist die erste vom Nullvektor verschiedene Spalte. Dann suchen wir innerhalb der  $l$ -ten Spalte ein von Null verschiedenes Element, und zwar - aus Gründen der numerischen Stabilität - das innerhalb dieser Spalte betragsmäßig größte Element. Dieses wird durch Zeilenvertauschung in die erste Zeile gebracht, also an die Stelle  $a_{1l}$ . Man nennt dieses Element, um das sich in den nächsten beiden Schritten alles „dreht“, das *Pivotelement*<sup>16</sup> bzw. den eben beschriebenen ersten Schritt *Pivotisieren*.
2. Wir dividieren die erste Zeile durch  $a_{1l}$ . Das erste Element dieser Zeile ist dann 1.
3. Nun eliminieren wir das Element der  $l$ -ten Spalte in jeder anderen Zeile indem wir von jeder Zeile die mit dem  $l$ -ten Element dieser Zeile multiplizierte erste Zeile subtrahieren. Es entsteht eine Matrix, deren  $l$ -te Spalte der Vektor

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

---

<sup>15</sup> *Wilhelm Jordan*, deutscher Geodät, 1842 - 1899

<sup>16</sup> vom franz. *le pivot* = Angelpunkt

ist.

4. Die Schritte 1) bis 3) werden jetzt für den Teil der Matrix wiederholt, den man erhält, wenn man die ersten  $l$  Spalten und die erste Zeile streicht; anschließend für den Teil der Matrix, der durch Streichen der ersten  $l + 1$  Spalten und der ersten beiden Zeilen entsteht usf. Man erhält zunächst eine Matrix, die die Eigenschaft (1) einer Matrix in Zeilennormalform erfüllt.
5.  $a_{ij}^*$  sei in der bisher gewonnenen Matrix das erste nicht-verschwindende Element der  $i$ -ten Zeile. Dann wird diese Zeile durch  $a_{ij}^*$  dividiert. Nach Abschluss dieses Schrittes für alle Zeilen ist die Eigenschaft (2) einer Matrix in Zeilennormalform erfüllt.
6.  $a_{ij}^*$  sei in der bisher gewonnenen Matrix das erste nicht-verschwindende Element der  $i$ -ten Zeile (es ist - nach dem letzten Schritt - gleich 1) und  $a_{kj}^*$  seien die anderen Elemente der  $j$ -ten Spalte. Man subtrahiert jetzt von jeder  $k$ -ten Zeile (für alle  $k \neq i$ ) die mit  $a_{kj}^*$  multiplizierte  $j$ -te Zeile.

Die letzten beiden Schritte werden, beginnend mit der zweiten Zeile, mehrfach durchgeführt, solange bis die entstandene Matrix Zeilennormalform hat.

Der Algorithmus lässt sich auch zur Matrizeninversion verwenden. Zu diesem Zweck wird die ursprüngliche Matrix auf der rechten Seite um eine Einheitsmatrix gleichen Typs erweitert und dann versucht, die ursprüngliche Matrix durch Zeilenumformungen in eine Einheitsmatrix zu verwandeln. Auf der rechten Seite ergibt sich dann die gesuchte inverse Matrix.

### 3 Das einfache Eigenwertproblem

Das einfache Eigenwertproblem geht von folgendem Ansatz aus: Für eine quadratische aber sonst beliebige  $(n, n)$ -Matrix  $\mathbf{A}$  sind Vektoren  $\mathbf{x}$  gesucht mit der Eigenschaft

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \quad (89)$$

Dabei ist  $\lambda$  ein Skalar. Statt  $\lambda\mathbf{x}$  kann man nach (38) auch schreiben:  $\lambda \cdot \mathbf{I}\mathbf{x}$  und dann die Gleichung (89) unter Anwendung des Distributivgesetzes umformen zu

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0. \quad (90)$$

Das Ergebnis ist eine homogene Matrixgleichung. Wenn die Matrix  $(\mathbf{A} - \lambda\mathbf{I})$  invertierbar ist, finden wir die Lösung von (90) indem wir zunächst auf beiden Seiten mit  $(\mathbf{A} - \lambda\mathbf{I})^{-1}$  multiplizieren:

$$(\mathbf{A} - \lambda\mathbf{I})^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = (\mathbf{A} - \lambda\mathbf{I})^{-1}0.$$

Das ist gleich bedeutend mit

$$\mathbf{x} = \mathbf{o}.$$

Diese Lösung ist aber trivial<sup>17</sup> (zumindest nach Meinung der Mathematiker). Eine nicht-triviale Lösung für  $\mathbf{x}$  wäre ein Vektor, der nicht gleich dem Nullvektor ist. Bei der Behandlung der Matrizenmultiplikation wurde angemerkt, dass ein Matrizenprodukt dann eine Nullmatrix ergibt, wenn einer der beiden Faktoren

---

<sup>17</sup>vom lat. *trivialis*: gewöhnlich

singulär ist (s. Seite 10). In unserem Fall muss also  $(\mathbf{A} - \lambda\mathbf{I})$  singulär sein, d.h. es muss gelten:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \quad (91)$$

Dieser Ausdruck heißt *charakteristische Determinante* von  $\mathbf{A}$ .

Wenn man dieses Problem als einfaches Eigenwertproblem bezeichnet, so muss es auch eine andere Form geben. Es handelt sich dabei um das allgemeine Eigenwertproblem (wobei das einfache Eigenwertproblem ein Spezialfall dieses Problems ist). Der Ansatz ist:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}. \quad (92)$$

Zusätzlich gibt es von beiden Arten des Eigenwertproblems auch noch jeweils eine zweite Art. Dabei wird berücksichtigt, dass in der Matrizenalgebra das Kommutativgesetz nicht gilt, also Multiplikation von links und von rechts zu unterschiedlichen Ergebnissen führen. Die entsprechenden Versionen für die Linkseigenvektoren sehen dann so aus:

$$\mathbf{x}\mathbf{A} = \lambda\mathbf{x} \text{ bzw. } \mathbf{x}\mathbf{A} = \lambda\mathbf{x}\mathbf{B}. \quad (93)$$

Wenn wir im folgenden vom Eigenwertproblem oder von den Eigenwerten sprechen, werden wir uns immer auf die durch (89) definierten rechten Eigenwerte 1. Art beziehen.

### 3.1 Eigenwerte

Man stellt eine Gleichung für die Berechnung der charakteristischen Determinante auf und setzt sie gleich Null (*charakteristische Gleichung*). Die Lösungen dieser Gleichung sind dann die *Eigenwerte* der Matrix  $\mathbf{A}$ . Es handelt sich bei der Gleichung um ein Polynom  $n$ -ten Grades in  $\lambda$  mit  $n$  Nullstellen (*charakteristisches Polynom*). Jede  $(n, n)$ -Matrix hat genau  $n$  Eigenwerte, die im Allgemeinen konjugiert komplex sind. Ist  $n$  ungerade, dann gibt es zumindest einen reellen Eigenwert. Je nach Anzahl der verschiedenen Nullstellen des Polynoms gibt es einfache, zweifache oder mehrfache Eigenwerte.

Zur Überprüfung der Eigenwerte können folgende einfache Zusammenhänge angewendet werden:

$$\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i, \quad (94)$$

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i. \quad (95)$$

Das Produkt aller Eigenwerte ergibt also die Determinante, ihre Summe die Spur der Matrix. Außerdem muss natürlich die Ausgangsgleichung (91) für jeden Eigenwert  $\lambda_i$  hinreichend genau erfüllt sein.

Bei symmetrischen Matrizen gilt zusätzlich: Das Produkt der Eigenwerte ist gleich dem Absolutwert des charakteristischen Polynoms.

#### 3.1.1 Weitere Eigenschaften

Ist  $\det(\mathbf{A}) = 0$ , so ist mindestens einer der Eigenwerte  $\lambda_i = 0$ . Genauer gilt: Hat die Matrix  $\mathbf{A}$  das Rangdefizit  $d$ , so sind genau  $d$  Eigenwerte gleich Null (Null ist somit eine  $d$ -facher Eigenwert).

Eine reell-symmetrische Matrix besitzt nur reelle Eigenwerte (aber nicht notwendigerweise verschiedene).

Die Eigenwerte einer (oberen oder unteren) Dreiecksmatrix sowie einer Diagonalmatrix sind genau die Elemente der Hauptdiagonalen.

Die Matrix  $\mathbf{A}$  und ihre Transponierte  $\mathbf{A}^T$  haben dieselben Eigenwerte.

Hat die Matrix  $\mathbf{A}$  die Eigenwerte  $\lambda_i$ , so hat die Matrix  $\mathbf{A}^m$  die Eigenwerte  $\lambda_i^m$ . Im Speziellen gilt: Hat die Matrix  $\mathbf{A}$  die Eigenwerte  $\lambda_i$  so hat ihre Inverse  $\mathbf{A}^{-1}$  die Eigenwerte  $\frac{1}{\lambda_i}$ .

Hat die Matrix  $\mathbf{A}$  die Eigenwerte  $\lambda_i$ , so hat die Matrix  $\mathbf{A} + c \cdot \mathbf{I}$  die Eigenwerte  $\lambda_i + c$  und die Matrix  $c \cdot \mathbf{A}$  die Eigenwerte  $c \cdot \lambda_i$

## 3.2 Eigenvektoren

Ist  $\lambda$  ein Eigenwert der Matrix  $\mathbf{A}$ , so hat die Gleichung (89) stets auch nicht-triviale (d.h. vom Nullvektor  $\mathbf{o}$  verschiedene) Lösungen für  $\mathbf{x}$ . Lösungen dieser Gleichung heißen „*Eigenvektoren* zum Eigenwert  $\lambda$ “. Zur Auflösung des homogenen Gleichungssystems kann man zum Beispiel das oben beschriebene Gauß-Jordan-Verfahren verwenden (siehe S. 20), und die Matrix  $(\mathbf{A} - \lambda\mathbf{I})$  für jedes  $\lambda_i$  auf Zeilennormalform bringen. Diese Zeilennormalform der Matrix wird aber eine Anzahl verschwindender Zeilen beinhalten; das Gleichungssystem ist nicht eindeutig lösbar. Zur Lösung muss zunächst über so viele Unbekannte frei verfügt werden, wie es verschwindende Zeilen gibt. Die restlichen Unbekannten können dann durch Auflösung des verbleibenden Gleichungssystems bestimmt werden. Um zu eindeutigen Ergebnissen zu kommen, werden die Lösungsvektoren normiert, d.h. auf Einheitslänge gebracht, indem man die einzelnen Elemente durch den Betrag des Vektors dividiert.

### 3.2.1 Spektralmatrix und Modalmatrix

Die Menge aller Eigenwerte von  $\mathbf{A}$  wird auch als das Spektrum<sup>18</sup> von  $\mathbf{A}$  bezeichnet; der (betragsmäßig) größte Eigenwert heißt *Spektralradius* von  $\mathbf{A}$ . Die Eigenwerte werden in der *Spektralmatrix*  $\Lambda$  zusammengefasst:  $\Lambda$  ist eine Diagonalmatrix; die Elemente der Hauptdiagonalen sind die Eigenwerte  $\lambda_1 \dots \lambda_n$ . (94) und (95) können dann auch so formuliert werden:

$$\det(\mathbf{A}) = \det(\Lambda), \quad (96)$$

$$\text{tr}(\mathbf{A}) = \text{tr}(\Lambda). \quad (97)$$

Die Eigenvektoren können - als Spalten nebeneinander geschrieben (Reihenfolge entsprechend der Reihenfolge der zugehörigen Eigenwerte in der Spektralmatrix) - in der *Modalmatrix*  $\mathbf{X}$  zusammengefasst werden. (Beachte: Die Eigenvektoren werden vorher *normiert*). Das Eigenwert-Problem kann dann auch so formuliert werden:

Gesucht sind zur Matrix  $\mathbf{A}$  eine *Spektralmatrix*  $\Lambda$  und eine *Modalmatrix*  $\mathbf{X}$ , sodass folgende gleich bedeutende Eigenschaften gelten:

$$\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}, \quad (98)$$

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \Lambda, \quad (99)$$

$$\mathbf{A}\mathbf{X} = \mathbf{X}\Lambda. \quad (100)$$

<sup>18</sup>vom lat. *spectrum*: Abbild (in der Seele), Vorstellung, genauer: ein Bild das man sich in der Vorstellung von etwas macht

(98) nennt man auch die *Eigenwertzerlegung* der Matrix  $\mathbf{A}$ .

(99) heißt *Hauptachsentransformation* (Erklärung siehe S. 24) bzw. „*Diagonalisieren* der Matrix  $\mathbf{A}$ “. Jede  $(n, n)$ -Matrix ist auf diese Art diagonalisierbar, wenn sie  $n$  linear unabhängige Eigenvektoren besitzt.

Die Matrizen  $\mathbf{A}$  und  $\Lambda$  sind im Übrigen *ähnliche* Matrizen. Zwei Matrizen  $\mathbf{R}$  und  $\mathbf{S}$  werden als *ähnliche Matrizen* bezeichnet, falls es eine Matrix  $\mathbf{U}$  gibt mit:

$$\mathbf{S} = \mathbf{U}^{-1} \mathbf{R} \mathbf{U}, \quad (101)$$

wobei diese Definition nicht voraussetzt, dass  $\mathbf{U}$  die Modalmatrix und  $\mathbf{S}$  die Spektralmatrix zu  $\mathbf{A}$  sind. Sie gilt auch für beliebig andere Matrizen  $\mathbf{S}$  und  $\mathbf{U}$ . Ähnliche Matrizen haben immer dieselbe Spektralmatrix (dieselben Eigenwerte).

Gleichung (101) ist die Gleichung für eine *Ähnlichkeitstransformation* (engl. *Similarity transformation*) der Matrix  $\mathbf{R}$ . Gleichung (99) ist ebenfalls die Gleichung einer Ähnlichkeitstransformation, d.h. das Diagonalisieren entspricht einer Ähnlichkeitstransformation.

Wenn  $\mathbf{A}$  symmetrisch ist, ist ihre Modalmatrix orthogonal. Für orthogonale Matrizen gilt bekanntlich  $\mathbf{X}^{-1} = \mathbf{X}^T$  und die Formeln zur Eigenwertzerlegung bzw. zum Diagonalisieren vereinfachen sich zu

$$\mathbf{A} = \mathbf{X} \Lambda \mathbf{X}^T, \quad (102)$$

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \Lambda \quad (103)$$

für alle  $\mathbf{A}$  für die gilt:  $\mathbf{A} = \mathbf{A}^T$ . Dem Diagonalisieren entspricht dann eine *orthogonale Ähnlichkeitstransformation*.

### 3.2.2 Weitere Eigenschaften

Ist die Matrix  $\mathbf{A}$  eine Diagonalmatrix, so ist  $\Lambda = \mathbf{A}$  und  $\mathbf{X} = \mathbf{I}$ .

Hat die Matrix  $\mathbf{A}$  die Eigenwerte  $\lambda_i$ , so hat die Matrix  $\mathbf{A}^m$  die Eigenwerte  $\mu_i = \lambda_i^m$  und jeder Eigenvektor von  $\mathbf{A}$  ist gleichzeitig Eigenvektor von  $\mathbf{A}^m$ .

Im Speziellen gilt für eine symmetrische Matrix  $\mathbf{A}_{sym}$ :  $\mathbf{A}^2 = \mathbf{A}^T \mathbf{A}$  und somit:  $\mu_i = \lambda_i^2$ ; außerdem haben in dem Fall  $\mathbf{A}$  und  $\mathbf{A}^T \mathbf{A}$  die selben Eigenvektoren.

### 3.2.3 Geometrische Interpretation

In einem  $(xy)$ -Koordinatensystem kann eine Ellipse in allgemeiner Lage (d.h. Mittelpunkt der Ellipse liegt im Koordinatenursprung, Hauptachsen in beliebiger Richtung) beschrieben werden durch

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1 \text{ bzw. } \mathbf{x}^T \mathbf{A} \mathbf{x} = 1. \quad (104)$$

Die Eigenvektoren der Matrix  $\mathbf{A}$  weisen dann genau in die Richtung der Hauptachsen; ihre Länge wird durch die Eigenwerte repräsentiert. Man kann nun das Koordinatensystem so drehen, dass die Ellipse im gedrehten  $(x'y')$ -Koordinatensystem in normaler Lage zu liegen kommt, d.h. die Hauptachsen der Ellipse fallen mit den Koordinatenachsen zusammen. Die entsprechende Rotation leistet die zu  $\mathbf{A}$  gehörige Modalmatrix  $\mathbf{X}$ ; die Ellipsengleichung hat die Form

$$\mathbf{X}^T \Lambda \mathbf{X}' = 1. \quad (105)$$



Daher auch der Name „Hauptachsentransformation“.

## 4 Singulärwertzerlegung

Ein der Eigenwertzerlegung ähnliches Verfahren ist die *Singulärwertzerlegung* einer Matrix (Press et al., 1986, S. 52-64). Sie ist im Gegensatz zur Eigenwertzerlegung nicht nur für quadratische sondern für beliebige Matrizen definiert. Dabei wird die  $(m, n)$ -Matrix  $\mathbf{A}$  in eine orthogonale  $(m, m)$ -Matrix  $\mathbf{U}^{19}$ , eine orthogonale  $(n, n)$ -Matrix  $\mathbf{V}^{20}$  und eine  $(m, n)$ -Diagonalmatrix  $\mathbf{S}$  zerlegt, sodass gilt:

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T. \quad (106)$$

Zur Bestimmung der drei Matrizen  $\mathbf{U}$ ,  $\mathbf{S}$  und  $\mathbf{V}$  führen wir zunächst eine Gauß'sche Transformation der Gleichung (106) durch:

$$\mathbf{A}^T\mathbf{A} = (\mathbf{U}\mathbf{S}\mathbf{V}^T)^T\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{V}\mathbf{S}^T\mathbf{U}^T\mathbf{U}\mathbf{S}\mathbf{V}^T. \quad (107)$$

$\mathbf{U}$  ist eine orthogonale Matrix, das Produkt  $\mathbf{U}^T\mathbf{U}$  ergibt daher eine Einheitsmatrix und kann weggelassen werden. Jetzt ersetzen wir, da  $\mathbf{S}$  eine Diagonalmatrix ist,

$$\mathbf{D} = \mathbf{S}^T\mathbf{S} = \mathbf{S}^2. \quad (108)$$

Damit vereinfacht sich (107) zu

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^T. \quad (109)$$

Auf der linken Seite steht dann die symmetrische Matrix  $(\mathbf{A}^T\mathbf{A})$ , und die rechte Seite entspricht formal einer Eigenwertzerlegung dieser symmetrischen Matrix in eine Modalmatrix  $\mathbf{V}$  und eine Spektralmatrix  $\mathbf{D}$  (vgl. (102)).

Damit können wir sofort  $\mathbf{V}$  angeben:  $\mathbf{V}$  wird aus den Eigenvektoren der Matrix  $(\mathbf{A}^T\mathbf{A})$  zusammengesetzt. Die Diagonalmatrix

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \quad (110)$$

enthält die Eigenwerte  $\lambda_i$  von  $(\mathbf{A}^T\mathbf{A})$ . Je nachdem, ob  $m < n$  oder  $m > n$  hat  $\mathbf{S}$  die Gestalt

$$\mathbf{S} = \left( \begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ \hline & & & 0 \end{array} \right) \text{ oder } \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ \hline & & & 0 \end{pmatrix} \quad (111)$$

und

$$\sigma_i = \sqrt{\lambda_i}. \quad (112)$$

Die  $\sigma_i$  werden die *Singulärwerte* der Matrix  $\mathbf{A}$  genannt. Ist der Rang der Matrix  $(\mathbf{A}^T\mathbf{A})$  gleich  $r$ , so gibt es  $r$  positive, von 0 verschiedene Singulärwerte der

<sup>19</sup>Somit gilt:  $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$ .

<sup>20</sup>Somit gilt:  $\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}$ .

Matrix  $\mathbf{A}$ . Zur Bestimmung von  $\mathbf{U}$  gehen wir ähnlich vor; diesmal multiplizieren wir (106) von rechts mit  $\mathbf{A}^T$ :

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{S}\mathbf{V}^T(\mathbf{U}\mathbf{S}\mathbf{V}^T)^T = \mathbf{U}\mathbf{S}\mathbf{S}^T\mathbf{U}^T. \quad (113)$$

Da  $\mathbf{S}$  eine Diagonalmatrix ist, können wir wieder setzen  $\mathbf{D} = \mathbf{S}\mathbf{S}^T = \mathbf{S}^2$  und damit (113) umschreiben in

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{D}\mathbf{U}^T. \quad (114)$$

Dieser Ausdruck entspricht dem aus Gleichung (109) und  $\mathbf{U}$  setzt sich demnach aus den Eigenvektoren von  $(\mathbf{A}\mathbf{A}^T)$  zusammen.

Die Singulärwertzerlegung wird übrigens auch im Deutschen oft mit SVD abgekürzt (entsprechend der englischen Bezeichnung *Singular Value Decomposition*).

## 5 Generalisierte Inverse

### 5.1 Moore-Penrose Pseudoinverse

Für manche Anwendungsfälle benötigt man auch für singuläre bzw. für rechteckige Matrizen so etwas Ähnliches wie eine Inverse. Dafür wurden die so genannten *generalisierten Inversen* definiert. Sie werden - im Gegensatz zur Bezeichnung  $\mathbf{A}^{-1}$  für die reguläre Inverse - mit  $\mathbf{A}^+$  bezeichnet. Eine generalisierte Inverse, die eine gewisse Bedeutung in unserem Anwendungsgebiet erlangt hat, ist die *Pseudoinverse*, nach ihren Autoren auch als *Moore-Penrose Inverse*<sup>21</sup> bezeichnet. Die Moore-Penrose Inverse ist jene  $(n, m)$ -Matrix  $\mathbf{A}^+$ , die bezüglich der  $(m, n)$ -Matrix  $\mathbf{A}$  folgende Eigenschaften erfüllt:

$$\mathbf{A}\mathbf{A}^+ = (\mathbf{A}\mathbf{A}^+)^T, \quad (115)$$

$$\mathbf{A}^+\mathbf{A} = (\mathbf{A}^+\mathbf{A})^T, \quad (116)$$

$$\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+, \quad (117)$$

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}. \quad (118)$$

Die Matrix  $\mathbf{A}^+$  ist für eine gegebene Matrix  $\mathbf{A}$  eindeutig. Wenn  $\mathbf{A}$  quadratisch und regulär ist, so erfüllt die „normale“ Inverse  $\mathbf{A}^{-1}$  ebenfalls die Gleichungen (115) bis (118), und zusätzlich auch noch (56).

### 5.2 Berechnung der Pseudoinversen mit Hilfe der Singulärwertzerlegung

Gehen wir zunächst von einer symmetrischen Matrix  $\mathbf{A} = \mathbf{A}^T$  aus. Wir können diese Matrix nach (102) in Eigenwerte zerlegen bzw. sie nach (103) diagonalisieren. Die Inverse zu  $\mathbf{A}$  kann gebildet werden durch die Inversion von (103):

$$\begin{aligned} (\mathbf{X}^T\mathbf{A}\mathbf{X})^{-1} &= \mathbf{X}^T\mathbf{A}^{-1}\mathbf{X} = \Lambda^{-1}, \\ \mathbf{A}^{-1} &= \mathbf{X}\Lambda^{-1}\mathbf{X}^T. \end{aligned} \quad (119)$$

<sup>21</sup> *Eliakim Hastings Moore*, amerikanischer Mathematiker, 1862 - 1932; *Roger Penrose*, britischer Mathematiker, \* 1931

Dies setzt voraus, dass  $\Lambda$  nicht singulär ist, d.h. keiner der Eigenwerte darf gleich Null sein. Wenn nun  $\Lambda = 0$  ein  $d$ -facher Eigenwert ist und  $r = (n - d)$  der Rang der Matrix  $\mathbf{A}$ , so führen wir zunächst eine Singulärwertzerlegung der symmetrischen Matrix  $\mathbf{A}$  durch:

$$\mathbf{A} = \mathbf{X}_r \Lambda_r \mathbf{X}_r^T \quad (120)$$

wobei  $\Lambda_r$  jene Submatrix aus der Spektralmatrix von  $\mathbf{A}$  ist, die nur die Eigenwerte  $\neq 0$  enthält, und  $\mathbf{X}_r$  die Matrix der zugehörigen Eigenvektoren.

Die *Pseudoinverse* ist dann definiert durch:

$$\mathbf{A}^+ = \mathbf{X}_r \Lambda_r^{-1} \mathbf{X}_r^T \quad (121)$$

mit  $\Lambda_r$ , und  $\mathbf{X}_r$ , in der Bedeutung wie oben.

Für nicht-symmetrische  $(n, n)$  und rechteckige  $(m, n)$ -Matrizen wenden wir die gleiche Vorgangsweise an und führen eine Singulärwertzerlegung durch. Die Pseudoinverse ist hier definiert als

$$\mathbf{A}^+ = \mathbf{V}_r \mathbf{S}_r^{-1} \mathbf{U}_r^T. \quad (122)$$

Es gibt jedoch auch noch andere Verfahren, mit denen Pseudoinverse berechnet werden können. Eines davon arbeitet mit einer additiv modifizierten Matrix (Keidel, 1975).

## 6 Iterative und numerische Verfahren

Für die praktische Bestimmung von Determinanten und Inversen, die direkte Auflösung von Gleichungssystemen, und die Bestimmung von Eigenwerten und Eigenvektoren ist es oft nicht ratsam die in den vorangegangenen Kapiteln beschriebenen Definitionen und Formeln direkt zu verwenden. Die oft große Anzahl von Einzelschritten führt zu einem für das praktische Rechnen unbrauchbar großem Aufwand. Außerdem ist die direkte Umsetzung der Definitionen oft numerisch nicht sehr günstig.

Es ist empfehlenswert, auf bekannte, getestete Programme zurückzugreifen. Im Folgenden werden wir einige der wichtigsten Verfahren und Algorithmen kurz erläutern. Eine nähere Auseinandersetzung mit den einzelnen Verfahren würde den Rahmen dieser Einführung sprengen; im Bedarfsfall sei auf entsprechende einschlägige Literatur verwiesen. Wir wollen uns aber doch vor Augen halten, welche Methoden Taschenrechner und Computer verwenden und welche Problematik dabei entstehen kann. Nur so können Ergebnisse richtig interpretiert und beurteilt werden.

### 6.1 Verfahren zur Lösung von Gleichungssystemen

#### 6.1.1 Eliminationsverfahren nach Gauß

Dieses Verfahren haben wir bereits beschrieben (siehe S. 19). Es transformiert eine Matrix durch elementare Zeilen- und Spaltenumformungen in eine obere Dreiecksmatrix bzw. Matrix in Treppenform. Das Verfahren kann verwendet werden, um den Rang einer Matrix zu bestimmen, aber auch zur Auflösung linearer Gleichungssysteme. In letzterem Fall ist aber bei Spaltenumformungen

besondere Vorsicht geboten, weil sich damit die Reihenfolge der Unbekannten ändert. Zur Lösung des Gleichungssystems wird die Koeffizientenmatrix noch um den Konstantenvektor (die „rechte Seite“ des Gleichungssystems) erweitert und in dieser erweiterten Matrix die Zeilenumformungen durchgeführt. Die letzte Zeile enthält dann nur eine Gleichung in einer Unbekannten und kann sofort aufgelöst werden. Durch sukzessives Einsetzen in die weiteren Gleichungen („Rückwärtseinsetzen“) können schließlich alle Unbekannten gefunden werden.

### 6.1.2 Eliminationsverfahren nach Gauß-Jordan

Dieses Verfahren haben wir ebenfalls bereits kennen gelernt (siehe S. 20). Es dient dazu, eine Matrix auf Zeilennormalform zu bringen bzw. in weiterer Folge als Verfahren zur Matrizeninversion.

### 6.1.3 LU-Verfahren (auch: LR-Verfahren)

Das LU-Verfahren („lower-upper-decomposition“) basiert auf dem Gauß-Verfahren und stellt eine  $(n, n)$ -Matrix  $\mathbf{A}$  als das Produkt einer unteren Dreiecksmatrix  $\mathbf{L}$  und einer oberen Dreiecksmatrix  $\mathbf{U}$  dar (Press et al., 1986, S. 31-38):

$$\mathbf{A} = \mathbf{LU} \quad (123)$$

Das Gleichungssystem (4) kann dann umgeschrieben werden in

$$\mathbf{Ax} = (\mathbf{LU})\mathbf{x} = \mathbf{L}(\mathbf{Ux}) = \mathbf{b} \quad (124)$$

Dann kann zunächst das Gleichungssystem

$$\mathbf{Ly} = \mathbf{b} \quad (125)$$

nach  $\mathbf{y}$  aufgelöst werden (beginnend mit der ersten Zeile und anschließendes „Vorwärtseinsetzen“) und danach

$$\mathbf{Ux} = \mathbf{y} \quad (126)$$

nach  $\mathbf{x}$  aufgelöst werden („Rückwärtseinsetzen“).

Gegenüber dem Gauß'schen Verfahren und dem Gauß-Jordan-Verfahren hat die LU-Zerlegung den Vorteil, dass eine einmal zerlegte Koeffizientenmatrix  $\mathbf{A}$  mit beliebigen weiteren Konstantenvektoren  $\mathbf{b}$  („rechten Seiten“) verwendet werden kann. Es lässt sich allerdings nicht jede Matrix in die Form (123) zerlegen.

### 6.1.4 Cholesky Verfahren

Das *Cholesky*<sup>22</sup>-Verfahren (Banachiewicz, 1939) ist eine Anwendung des LU-Verfahrens auf symmetrische, positiv-definite Matrizen. Es spaltet  $\mathbf{A}$  in eine obere Dreiecksmatrix  $\mathbf{U}$  und deren transponierte Matrix  $\mathbf{U}^T$  auf:

$$\mathbf{A} = \mathbf{U}^T\mathbf{U}. \quad (127)$$

---

<sup>22</sup> *Andre-Louis Cholesky*, französischer Geodät, 1875 - 1918

Die Elemente der Matrix  $\mathbf{U}$  berechnen sich aus

$$u_{ik} = \frac{a_{ik} - u_{1i}u_{1k} - u_{2i}u_{2k} - \dots - u_{i-1,i}u_{i-1,k}}{u_{ii}} \quad (128)$$

$$u_{ii}^2 = a_{ii} - u_{1i}^2 - u_{2i}^2 - \dots - u_{i-1,i}^2 \quad (129)$$

Angewendet auf ein Gleichungssystem

$$\mathbf{Ax} = \mathbf{b}$$

ergibt sich

$$\mathbf{Ux} = \mathbf{s} \quad (130)$$

mit

$$s_i = \frac{b_i - \sum_{k=1}^{i-1} u_{ik}s_k}{u_{ii}}. \quad (131)$$

Leider ist dieses Verfahren mathematisch nicht sehr stabil. Nur bei großen Residuen und kleinen Lösungsvektoren  $\mathbf{x}$  wird das System „gutmütig“. Eine nähere Beschreibung des Verfahrens findet man auch bei Zurmühl und Falk (1984, S. 77-79)

### 6.1.5 Partielle Reduktion mit dem Cholesky-Verfahren

Wenn man das ursprüngliche Gleichungssystem  $\mathbf{Ax} = \mathbf{b}$  aufspaltet in

$$\begin{aligned} \mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2 &= \mathbf{b}_1, \\ \mathbf{A}_{21}\mathbf{x}_1 + \mathbf{A}_{22}\mathbf{x}_2 &= \mathbf{b}_2, \end{aligned} \quad (132)$$

so können wir unsere Dreiecksmatrix ebenfalls aufspalten in

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{0} & \mathbf{U}_{22} \end{pmatrix}. \quad (133)$$

Aus Gleichung (127) können wir die folgenden Identitäten herleiten

$$\mathbf{U}_{11}^T \mathbf{U}_{11} = \mathbf{A}_{11}, \quad (134)$$

$$\mathbf{U}_{11}^T \mathbf{U}_{12} = \mathbf{A}_{12}, \quad (135)$$

$$\mathbf{U}_{12}^T \mathbf{U}_{12} + \mathbf{U}_{22}^T \mathbf{U}_{22} = \mathbf{A}_{22}. \quad (136)$$

Diese Identitäten wollen wir nun verwenden, um (132) zu lösen. Als ersten Schritt multiplizieren wir die erste Zeile mit  $(\mathbf{U}_{11}^T)^{-1}$  und erhalten wegen (134) und (135)

$$\mathbf{U}_{11}\mathbf{x}_1 + \mathbf{U}_{12}\mathbf{x}_2 = (\mathbf{U}_{11}^T)^{-1} \mathbf{b}_1 = \mathbf{s}_1 \quad (137)$$

Nun eliminieren wir  $\mathbf{x}_1$  aus der zweiten Zeile, indem wir die erste Zeile mit  $\mathbf{A}_{21}\mathbf{U}_{11}^{-1} = \mathbf{U}_{12}^T$  multiplizieren und von der zweiten Zeile abziehen. Wir erhalten

$$\mathbf{A}_{22}^{(p)} \mathbf{x}_2 = \mathbf{b}_2^{(p)} \quad (138)$$

mit

$$\mathbf{A}_{22}^{(p)} = \mathbf{A}_{22} - \mathbf{U}_{12}^T \mathbf{U}_{12} = \mathbf{U}_{22}^T \mathbf{U}_{22} = \mathbf{A}_{22} - \mathbf{A}_{12} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}, \quad (139)$$

$$\mathbf{b}_2^{(p)} = \mathbf{b}_2 - \mathbf{U}_{12}^T \mathbf{s}_1 = \mathbf{b}_2 - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{b}_1. \quad (140)$$

Zusammengefasst sieht unser Gleichungssystem also folgendermaßen aus:

$$\begin{aligned} \mathbf{U}_{11}\mathbf{x}_1 + \mathbf{U}_{12}\mathbf{x}_2 &= \mathbf{s}_1 \\ \mathbf{A}_{22}^{(p)}\mathbf{x}_2 &= \mathbf{b}_2^{(p)} \end{aligned} \quad (141)$$

Dieses System können wir nun lösen, indem wir zuerst  $\mathbf{x}_2$  und dann  $\mathbf{x}_1$  berechnen.

### 6.1.6 Gauß-Seidel-Verfahren

Sehr oft sind bei linearen Gleichungssystemen die Koeffizientenmatrizen sehr groß, viele Koeffizienten jedoch Null. Solche „schwach besetzten“ Systeme werden besser mit *iterativen* Verfahren gelöst, zum Beispiel mit dem *Gauß-Seidel-Verfahren*<sup>23</sup>. Dabei werden Näherungslösungen des Gleichungssystems so lange verbessert, bis eine vorgegebene gewünschte Genauigkeit erreicht ist. Der Nachteil des Verfahrens (wie jedes iterativen Verfahrens) ist, dass es nicht immer konvergiert (Press et al., 1986, S. 653-655).

## 6.2 Verfahren zur Bestimmung der Eigenwerte und Eigenvektoren

Zur numerischen Lösung des Eigenwertproblems gibt es spezielle Methoden, die nicht direkt auf den Nullstellen des charakteristischen Polynoms beruhen. Dazu gehören zum Beispiel der *QR-Algorithmus* und das *Jacobi-Verfahren*. Das *QR-Verfahren* setzt eine Matrix  $\mathbf{A}$  als Produkt einer orthogonalen Matrix  $\mathbf{Q}$  und einer oberen Dreiecksmatrix  $\mathbf{R}$  zusammen:

$$\mathbf{A} = \mathbf{QR}. \quad (142)$$

Es kann verwendet werden, um die Eigenwerte einer Matrix zu berechnen. Dazu wird zunächst die Matrix  $\mathbf{A}$  zerlegt in  $\mathbf{A}_k = \mathbf{Q}_k\mathbf{R}_k$  und anschließend die Matrix  $\mathbf{A}_{k+1} = \mathbf{R}_k\mathbf{Q}_k$  gebildet. Diese Matrix wird wieder einer QR-Zerlegung unterzogen usw. Die Diagonalelemente von  $\mathbf{A}_k$  konvergieren dann gegen die Eigenwerte von  $\mathbf{A}$ .

An Stelle einer oberen kann man für das Verfahren auch eine untere Dreiecksmatrix  $\mathbf{L}$  verwenden (*QL-Verfahren*).

Das *Jacobi-Verfahren*<sup>24</sup> ist ein indirektes Verfahren zur numerischen Bestimmung der Eigenwerte und Eigenvektoren einer *symmetrischen* Matrix. Dem Verfahren liegt der Gedanke zu Grunde, die symmetrische Matrix  $\mathbf{A}$  durch orthogonale Ähnlichkeitstransformationen näherungsweise in eine Diagonalmatrix überzuführen. Die gesuchten Eigenwerte sind dann gleich den Diagonalelementen der gefundenen Diagonalmatrix. Das Produkt aller auftretenden Transformationsmatrizen ist die gesuchte Modalmatrix, die spaltenweise die Eigenvektoren enthält.

Allen in den beiden letzten Abschnitten genannten Verfahren ist eines gemeinsam: Eine strenge Lösung ist in endlich vielen Schritten nicht möglich; sie liefern nur näherungsweise Lösungen. Das dafür aber oft schneller als strenge Lösungen. Der maximale „Fehler“, den man durch die Verwendung eines Näherungsverfahrens machen darf, kann oft vorgegeben werden oder es ist zumindest

<sup>23</sup> Philipp Ludwig von Seidel, deutscher Mathematiker, 1821 - 1896

<sup>24</sup> Carl Gustav Jacob Jacobi, deutscher Mathematiker, 1804 - 1851

sein Maximalwert bekannt. Daher sind diese Verfahren für technische Lösungen uneingeschränkt einsetzbar. Dabei ist es für uns von Interesse, weichen (relativen oder absoluten) Fehler - abhängig von der Fehlerschranke - der erhaltene Wert hat.

### 6.3 Matrixnormen

Unter der Matrixnorm  $\|\mathbf{A}\|$  versteht man eine (reelle) Funktion ihrer Elemente, welche bestimmte Eigenschaften aufweist:

$$\|\mathbf{A}\| \geq 0 \quad (143)$$

mit  $\|\mathbf{A}\| = 0$  nur für den Fall  $\mathbf{A} = \mathbf{0}$ .

$$\|c \cdot \mathbf{A}\| = |c| \cdot \|\mathbf{A}\|, \quad (144)$$

$$\|\mathbf{A} + \mathbf{B}\| = \|\mathbf{A}\| + \|\mathbf{B}\|, \quad (145)$$

$$\|\mathbf{A} \cdot \mathbf{B}\| = \|\mathbf{A}\| \cdot \|\mathbf{B}\|. \quad (146)$$

Es gibt mehrere Normen, die diese Eigenschaften erfüllen. Die wichtigsten, die für unsere Anwendungen in Frage kommen, sind:

Die *Frobenius-Norm*<sup>25</sup> (auch: *Schur Norm*<sup>26</sup> oder im Fall von Vektoren: *Euklidische Norm*<sup>27</sup>) ist definiert als:

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{k=1}^n |a_{ik}|^2}, \quad (147)$$

*Spektralnorm* (auch: *Hilbert-Norm*<sup>28</sup>) als

$$\|\mathbf{A}\|_S = \sqrt{\mu_{max}} \quad (148)$$

mit  $\mu_{max}$  = maximaler Eigenwert der Matrix  $\mathbf{A}^T \mathbf{A}$ .

### 6.4 Die Kondition eines Gleichungssystems

Die Norm einer Matrix spielt bei der Abschätzung der numerischen Genauigkeit von Lösungen eines linearen Gleichungssystems eine Rolle. Wir wollen uns die Problematik anhand eines so genannten „gestörten“ Gleichungssystems anschauen. „Gestört“ ist das Gleichungssystem deshalb, weil einer oder mehrere der Parameter um einen kleinen Wert abweicht. An Stelle des Gleichungssystems

$$\mathbf{Ax} = \mathbf{b} \quad (149)$$

mit der regulären Matrix  $\mathbf{A}$  betrachten wir das gestörte System

$$(\mathbf{A} + \Delta\mathbf{A})(\mathbf{x} + \Delta\mathbf{x}) = (\mathbf{b} + \Delta\mathbf{b}) \quad (150)$$

<sup>25</sup> Ferdinand Georg Frobenius, deutscher Mathematiker, 1849 - 1917

<sup>26</sup> Issai Schur, russisch-deutscher Mathematiker, 1875 - 1947

<sup>27</sup> Eukleides von Alexandrien, griechischer Mathematiker und Begründer der Elementargeometrie (der sog. Euklidischen Geometrie), ca. 300 v.Chr.

<sup>28</sup> David Hilbert, deutscher Mathematiker, 1862 - 1943

unter der Annahme, dass auch dieses System noch eine reguläre Koeffizientenmatrix  $(\mathbf{A} + \Delta\mathbf{A})$  besitzt. Betrachten wir als Beispiel das lineare Gleichungssystem

$$\begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad a \in \mathbb{R}, \quad |a| \neq 1.$$

Es besitzt die Lösung

$$x_0 = \frac{1}{1-a^2}, \quad y_0 = -\frac{a}{1-a^2}.$$

Nehmen wir nun ein „gestörtes System“ an:

$$\begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ \delta \end{pmatrix}.$$

Es hat die Lösung:

$$x_\delta = \frac{1-\delta a}{1-a^2}, \quad y_\delta = \frac{\delta-a}{1-a^2}$$

und somit als Differenzen gegenüber dem ursprünglichen System:

$$x_\delta - x_0 = \delta \frac{-a}{1-a^2}, \quad y_\delta - y_0 = \delta \frac{1}{1-a^2}.$$

Änderungen in der zweiten Komponente des Konstantenvektors auf der rechten Seite werden also mit dem Faktor  $a/(1-a^2)$  bzw.  $1/(1-a^2)$  verstärkt. Daraus folgt: Liegt  $a$  betragsmäßig nahe bei 1, so wächst der Verstärkungsfaktor stark an (Der Nenner geht gegen Null und damit der Bruch gegen  $\infty$ ). Umso größer  $a$  hingegen betragsmäßig ist, desto größer wird der Nenner, und desto kleiner der Bruch.

Sehen wir uns das Problem etwas allgemeiner an und fragen nach der Differenz  $\Delta x$  des Lösungsvektors  $\mathbf{x}$  in (150) und nach dem relativen Fehler bezüglich der „richtigen“ Lösung. Lösen wir (150) nach  $\Delta x$  auf, so erhalten wir für den absoluten Fehler

$$\Delta \mathbf{x} = (\mathbf{A} + \Delta\mathbf{A})^{-1}(\Delta \mathbf{b} - \Delta\mathbf{A}\mathbf{x}). \quad (151)$$

Zur Abschätzung des relativen Fehlers bedienen wir uns der in Abschnitt 2.7.3 eingeführten Matrixnorm und bestimmen  $\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|}$ . Ohne auf die etwas komplizierte Ableitung weiter einzugehen, geben wir das Ergebnis gleich an: Für (kleine) Störungen der Matrixelemente wird der relative Fehler der „rechten Seite“ und der Elemente der Koeffizientenmatrix noch um den Faktor  $\|\mathbf{A}^{-1}\| \cdot \|\mathbf{A}\|$  verstärkt. Wenn kleine Eingangsfehler auch nur kleine Ergebnisfehler bewirken, so sagt man, das Problem sei *gut konditioniert*; wenn hingegen kleine Eingangsfehler unverhältnismäßig große Ergebnisfehler bewirken, ist es *schlecht konditioniert*. Den Verstärkungsfaktor selber nennen wir die *Kondition* der Matrix  $\mathbf{A}$ . Genauer definieren wir, je nach verwendeten Norm,

$$\kappa_p = \|\mathbf{A}^{-1}\|_p \cdot \|\mathbf{A}\|_p \quad (152)$$

als die Kondition der Matrix  $\mathbf{A}$  bezüglich der Norm  $p$ . Üblicherweise wird dafür die Spektralnorm verwendet.

Unabhängig davon, welche Norm wir für die Kondition verwenden, gilt immer

$$\kappa(A) \geq 1. \quad (153)$$



Eine große Konditionszahl der Matrix  $\mathbf{A}$  weist auf unangenehme numerische Eigenschaften hin, die sich beim Auflösen von Gleichungssystemen (mit der Koeffizientenmatrix  $\mathbf{A}$ ) nachteilig auswirken. Einer singulären Matrix ( $\lambda_{\min} = 0$ ) wird die Konditionszahl  $\kappa = \infty$  zugeordnet.

Für symmetrische Matrizen gilt bekanntlich:  $\mu_i = \lambda_i^2$  und daher auch:

$$\|A\|_S = |\lambda_{\max}|, \quad (154)$$

d.h. die Spektralnorm ist gleich dem Absolutbetrag des größten Eigenwertes von  $\mathbf{A}$ . Außerdem hat die Inverse  $\mathbf{A}^{-1}$  einer symmetrischen Matrix die Eigenwerte  $\frac{1}{\lambda_i}$  (wenn  $\mathbf{A}$  die Eigenwerte  $\lambda_i$  hat). Daraus folgt für die Kondition symmetrischer Matrizen (Borre, 1973):

$$\kappa_S = \frac{|\lambda_{\max}|}{|\lambda_{\min}|}. \quad (155)$$

## Literatur

- Antosik, Piotr (1985): *Matrix Methods in Analysis*, Bd. 1113 von *Lecture Notes in Mathematics*. Springer, Berlin.
- Ayres, Frank (1978): *Matrizen: Theorie und Anwendung*. McGraw-Hill, New York.
- Banachiewicz, Tadeusz (1939): On the Computation of the Inverse Arrays. *Acta Astronomica*, Bd. 4.
- Bellman, Richard Ernest (1960): *Introduction to Matrix Analysis*. McGraw-Hill Series in Matrix Theory. McGraw-Hill, New York.
- Borre, Kai (1973): Über geodätische Netzformen und damit verknüpfte numerische Probleme. *Zeitschrift für Vermessungswesen (ZfV)*, 98. Jahrgang(2):67–75.
- Eves, Howard (1980): *Elementary Matrix Theory*. Dover Books on Advanced Mathematics. Dover Publications, New York, Unabridged and corr. republ. of the work orig. publ. in 1966 Aufl.
- Falk, Sigurd (1951): Ein übersichtliches Schema für die Matrizenmultiplikation. *Zeitschrift für Angewandte Mathematik und Mechanik*, 31.
- Gauß, Carl Friedrich (1809): *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium (Theorie der Bewegung der Himmelskörper, welche in Kegelschnittbahnen die Sonne umlaufen)*. F. Perthes und I. H. Besser, Hamburg.
- Gauß, Carl Friedrich (1829): Über ein neues allgemeines Grundgesetz der Mechanik. *Journal für die reine und angewandte Mathematik*, 4:232–235.
- Gotthardt, Ernst (1951): Über Matrizen und ihre Anwendung in der Ausgleichsrechnung. *Zeitschrift für Vermessungswesen (ZfV)*, 76. Jahrgang(4):97–100.

- Hirvonen, R.A. (1958): Ein Vorschlag für Matrizenbezeichnungen in der Ausgleichsrechnung. *Zeitschrift für Vermessungswesen (ZfV)*, 83. Jahrgang(5):154–166.
- Horn, Roger A. (1985): *Matrix Analysis*. Cambridge University Press, Cambridge.
- Höpcke, Walter (1980): *Fehlerlehre und Ausgleichsrechnung*. Walter de Gruyter.
- Keidel, Werner (1975): Die Berechnung der Pseudoinversen mit Hilfe einer additiv modifizierten Matrix. *Zeitschrift für Vermessungswesen (ZfV)*, 100. Jahrgang(2):62–67.
- Linkwitz, Klaus (1960): Über die Systematik verschiedener Formen der Ausgleichsrechnung. *Zeitschrift für Vermessungswesen (ZfV)*, 85. Jahrgang(5, 6, 7):156–166, 191–204, 243–254.
- Niemeier, Wolfgang (2002): *Ausgleichsrechnung*. Walter de Gruyter, Berlin.
- Perović, Gligorije (2005): *Least Squares*. University of Belgrade, Faculty of Civil Engineering.
- Press, William H.; Brian P. Flannery; Saul A. Teukolsky und William T. Vetterling (1986): *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge.
- Reißmann, Günter (1976): *Die Ausgleichsrechnung*. VEB Verlag für Bauwesen, Berlin, 5. Aufl.
- Reinhardt, F. und H. Soeder (1991): *dtv-Atlas zur Mathematik: Grundlagen, Algebra und Geometrie*. dtv, München.
- Schmidt, Karsten (1998): *Moderne Matrix-Algebra: mit Anwendungen in der Statistik*. Springer-Lehrbuch. Springer, Berlin.
- Sylvester, James Joseph (1850): Additions to the Articles „On a new Class of Theorems“, and „On Pascal’s Theorem“. *Philosophical Magazine*, S. 363–370.
- Wolf, Helmut (1959): Die Ausgleichung nach bedingten Beobachtungen mit Unbekannten im Matrizenkalkül. *Zeitschrift für Vermessungswesen (ZfV)*, 84. Jahrgang(4):105–108.
- Zurmühl, Rudolf und Sigurd Falk (1984): *Matrizen und ihre Anwendung*, Bd. 1 + 2. Springer, Heidelberg.