Lineare Algebra II für Informatiker

Joachim Ohser

5. Juni 2005

Dieses Skript soll die Vorlesung ergänzen; die Verwendung des Skrips bei der Vorbereitung der Prüfung ersetzt nicht den Vorlesungsbesuch. Der Inhalt des Skripts stimmt weitgehend mit dem Inhalt des in der Vorlesung behandelten Stoffs überein. Die Dozenten sind in der Gestaltung ihrer Vorlesung jedoch nicht an den Inhalt das Skripts gebunden. Prüfungsgegenstand ist stets der Inhalt der Vorlesung.

Inhaltsverzeichnis

Kapitel 1

Das Eigenwertproblem für quadratischer Matrizen

1.1 Einführung und Definitionen

Wir betrachten das lineare Gleichungssystem

$$Ax = \lambda x,$$

wobei die quadratische Matrix A als lineare Transformation (Rotation, Spiegelung, ...) interpretiert wird und λ eine reelle oder komplexe Zahl ist. Eine Lösung x des Gleichungssystems ist ein Vektor, für den die lineare Transformation Ax lediglich eine eine Streckung (Stauchung) λx des Vektors x bedeutet.

Natürlich hat das lineare Gleichungssystem stets die Triviallösung $x = 0$. Interessant sind daher nicht triviale Lösungen $x \neq 0$.

Definition. Eine komplexe Zahl λ heißt *Eigenwert* der quadratischen Matrix A, wenn es Vektoren $x \neq 0$ gibt, die Lösung des linearen Gleichungssystems

$$Ax = \lambda x \tag{1.1}$$

sind. Die Lösungsvektoren x heißen *Eigenvektoren* der Matrix A zum Eigenwert λ.

Bemerkung. Zu jedem Eigenwert gibt es unendlich viele Eigenvektoren, denn wenn x Lösung von (1.1) ist, dann ist wegen

$$\begin{aligned} Ax &= \lambda x \\ tAx &= t\lambda x \\ A \cdot (tx) &= \lambda \cdot (tx) \end{aligned}$$

auch jedes Vielfache tx mit $t \in \mathbb{R} \setminus \{0\}$ Eigenvektor von A zum Eigenwert λ.

Beispiel. Die Zahl $\lambda = 1$ ist Eigenwert der Matrix

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

Der Vektor

$$x = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

ist Eigenvektor von A zum Eigenwert $\lambda = 1$, denn

$$\begin{aligned} \begin{pmatrix} -2 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} &= 1 \cdot \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}. \end{aligned}$$

Aber auch jeder andere Vektor

$$tx = \begin{pmatrix} 0 \\ -t \\ t \end{pmatrix}$$

mit $t \in \mathbb{R} \setminus \{0\}$ ist ebenfalls Eigenvektor von A zum Eigenwert λ.

Satz. Sind $\lambda_1, \ldots, \lambda_m$, $m \leq n$, voneinander verschiedene Eigenwerte einer (n,n)-Matrix, dann sind die dazugehörigen Eigenvektoren $x_1, \ldots, x_m$ linear unabhängig.

Abschließend treffen wir noch eine Aussage über einen Spezialfall.

Satz. Die Eigenwerte einer symmetrischen Matrix sind reell. Die zu verschiedenen Eigenwerten einer symmetrischen Matrix gehörenden Eigenvektoren sind othogonal zueinander.

1.2 Die Berechnung von Eigenwerten und Eigenvektoren

Ausgehend von Gleichung (1.1) kann das Eigenwertproblem einer beliebigen (n,n)-Matrix A wie folgt behandelt werden:

$$\begin{aligned} Ax &= \lambda x \\ Ax - \lambda x &= 0 \\ Ax - \lambda Ix &= 0, \qquad \text{da} \quad Ix = x, \quad \text{wobei } I \text{ die Einheitsmatrix bezeichnet,} \\ (A - \lambda I)x &= 0 \end{aligned}$$

Die letzte Gleichung ist ein homogenes lineares Gleichungssystem für den Vektor x mit der Koeffizientenmatrix $B = A - \lambda I$. Dieses homogene lineare Gleichungssystem hat genau dann nicht triviale Lösungen $x \neq 0$, wenn der Rang von B kleiner als n ist, d. h. $\mathrm{rg}\,(A - \lambda I) < n$. Das ist gleichbedeutend damit, dass $\det(A - \lambda I) = 0$ ist. Die Determinante $\det(A - \lambda I)$ ist ein Polynom vom Grad n in λ; $\det(A - \lambda I)$ heißt *charakteristisches Polynom* der Matrix A. Die Eigenwerte von A sind folglich die Nullstellen des charakteristischen Polynoms.

Definition. Sei A eine quadratische Matrix. Die Gleichung

$$\det(A - \lambda I) = 0 \tag{1.2}$$

heißt *charakteristische Gleichung* oder *Eigenwertgleichung* von A.

Satz. Die charakteristische Gleichung $\det(A - \lambda I) = 0$ einer (n, n)-Matrix A hat genau n Lösungen $\lambda_1, \ldots, \lambda_n$.

In anderen Worten: Eine (n, n)-Matrix hat genau n Eigenwerte, die als Nullstellen des chrakteristischen Polynoms (d. h. als Lösung der chrakteristischen Gleichung) bestimmt werden können. Ein zu λ_i gehörender Eigenvektor x_i ist (nicht triviale) Lösung des homogenen linearen Gleichungssystems

$$(A - \lambda_i I)x_i = 0, \tag{1.3}$$

$i = 1, \ldots, n$.

Bemerkungen

(i) Die Lösungen x_i von (1.3) sind Parameterlösungen, d. h. auch alle tx_i mit $t \in \mathbb{R}$ sind Lösungen von (1.3). Da nur nicht triviale Lösungen zulässig sind, schließen wir $t = 0$ aus.

(ii) Oft ist es zweckmäßig, nur diejenigen Eigenvektoren anzugeben, deren Länge gleich Eins ist (normierte Eigenvektoren).

Definition. Die Lösungsmenge des homogenen linearen Gleichungssystems (1.3) heißt der zum Eigenvektor λ_i gehörennde *Eigenraum* E_i.

Ist λ_i eine einfache Nullstelle des charakteristischen Polynoms (d. h. λ_i ist ein einfacher Eigenvektor), dann dann ist der zu λ_i gehörige Eigenraum

$$E_i = \{tx_i \,:\, t \in \mathbb{R}\}$$

eindimensional. Der Eigenraum E_i entspricht dann einer Geraden im $\mathbb{R}^n$. Ist λ_i eine doppelte Nullstelle des charakteristischen Polynoms, dann ist der dazugehörige Eigenraum E_i meist zweidimensional (also z. B. eine Ebene im $\mathbb{R}^3$). Allgemein gilt für m-fache Nullstellen des charakteristischen Polynoms die folgende Aussage:

Satz. Seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte einer symmetischen (n, n)-Matrix mit $\lambda_k = \ldots = \lambda_{k+m-1}$. Dann ist E_k der Eigenraum zu $\lambda_k, \ldots, \lambda_{k+m-1}$, d. h. es gibt m linear unabhängige Vektoren $x_k, \ldots, x_{k+m-1} \in \mathbb{R}^n$ mit der Eigenschaft

$$E_k = \left\{ x \,:\, x = \sum_{i=1}^{m} t_i x_{i+k-1}, \quad t_1, \ldots, t_m \in \mathbb{R} \right\}.$$

Das bedeutet, der Eigenraum E_k wird durch die Vektoen $x_k, \ldots, k_{k+m-1}$ „aufgespannt“, und die Dimension des Eigenraums stimmt mit der Vielfachheit m des Eigenwerts λ_k überein. Für symmetrische Matrizen gilt also stets

$$\mathrm{rg}\,(A - \lambda_k I) = m.$$

Die Vektoren $x_k, \ldots, k_{k+m-1}$ lassen sich so wählen, dass sie zueinander orthogonal sind.

Beispiele

(i) Wir betrachten noch einmal das Beispiel von Seite 6. Die Eigenwerte von A sind Lösungen der charakteristischen Gleichung (1.2), d. h.

$$\begin{aligned}
\det \left[\begin{pmatrix} -2 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & -2 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] &= 0 \\
\det \begin{pmatrix} -2-\lambda & 1 & 1 \\ 0 & 1-\lambda & 0 \\ 0 & -2 & -1-\lambda \end{pmatrix} &= 0 \\
(-2-\lambda) \cdot (1-\lambda) \cdot (-1-\lambda) &= 0.
\end{aligned}$$

Lösungen des chrakteristischen Polynoms sind

$$\lambda_1 = -2, \qquad \lambda_2 = 1, \qquad \lambda_3 = -1.$$

Zum Eigenwert λ_1 erhält man den Eigenvektor x_1 als Lösung des homogenen linearen Gleichungssystems (1.3). Es gilt

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & -2 & 1 \end{pmatrix} \cdot x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{d. h.} \qquad x_1 = \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix}, \quad t \in \mathbb{R} \setminus \{0\}.$$

Analog erhält man für den Eigenwert $\lambda_2 = 1$

$$\begin{pmatrix} -3 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & -2 & -2 \end{pmatrix} \cdot x_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{d. h.} \qquad x_2 = \begin{pmatrix} 0 \\ t \\ -t \end{pmatrix}, \quad t \in \mathbb{R} \setminus \{0\},$$

und aus $\lambda_3 = -1$ folgt

$$\begin{pmatrix} -1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & -2 & 0 \end{pmatrix} \cdot x_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{d. h.} \qquad x_3 = \begin{pmatrix} t \\ 0 \\ t \end{pmatrix}, \quad t \in \mathbb{R} \setminus \{0\}.$$

(ii) Die Matrix

$$A = \begin{pmatrix} 5 & -\sqrt{12} \\ -\sqrt{12} & 1 \end{pmatrix}$$

ist symmetrisch. Erwartungsgemäß sind ihre Eigenwerte λ_1 und λ_2 reelle Zahlen. Aus der charakteristischen Gleichung

$$\begin{aligned} \det\left[\begin{pmatrix} 5 & -\sqrt{12} \\ -\sqrt{12} & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] &= 0 \\ \det\begin{pmatrix} 5-\lambda & -\sqrt{12} \\ -\sqrt{12} & 1-\lambda \end{pmatrix} &= 0 \\ (5-\lambda)\cdot(1-\lambda) - (-\sqrt{12})\cdot(-\sqrt{12}) &= 0 \\ \lambda^2 - 6\lambda - 7 &= 0 \end{aligned}$$

erhalten wir

$$\begin{aligned} \lambda_{1/2} &= 3 \pm \sqrt{9+7} \\ \lambda_1 &= 7, \qquad \lambda_2 = -1. \end{aligned}$$

Die Eigenvektoren x_1 und x_2 sind Lösungen von

$$\begin{pmatrix} 5-7 & -\sqrt{12} \\ -\sqrt{12} & 1-7 \end{pmatrix} x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{bzw.} \qquad \begin{pmatrix} 5-(-1) & -\sqrt{12} \\ -\sqrt{12} & 1-(-1) \end{pmatrix} x_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Eine Lösung von

$$\begin{pmatrix} -2 & -\sqrt{12} \\ -\sqrt{12} & -6 \end{pmatrix} x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ist

$$x_1 = \begin{pmatrix} \sqrt{3} \\ -1 \end{pmatrix},$$

und aus

$$\begin{pmatrix} 6 & -\sqrt{12} \\ -\sqrt{12} & 2 \end{pmatrix} x_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

erhalten wir

$$x_2 = \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}.$$

Die beiden Vektoren x_1 und x_2 sind orthogonal; ihr Skalarprodukt ist Null, $x_1 \cdot x_2 = 0$.

(iii) Die Matrix

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

ist nicht symmetrisch. Dennoch sind in diesem Fall die Eigenwerte reellwertig. Es gilt

$$\begin{aligned} \det \begin{pmatrix} 1-\lambda & -2 \\ 0 & 2-\lambda \end{pmatrix} &= 0 \\ (1-\lambda)(2-\lambda) &= 0. \end{aligned}$$

Damit sind $\lambda_1 = 1$ und $\lambda_2 = 2$ die Eigenwerte von A. Die Eigenvektorn x_1, x_2 sind Lösungen von

$$\begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{bzw.} \qquad \begin{pmatrix} -1 & -2 \\ 0 & 0 \end{pmatrix} x_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{bzw.}$$

d. h.

$$x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Es gilt $x_1 \cdot x_2 = 2$, d. h. die Eigenvektoren sind nicht orthogonal.

(iv) Das letzte Beispiel lässt sich leicht so verändern, dass das charakteristische Polynom eine doppelte Nullstelle hat. Für

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

ist $\lambda_{1/2} = 1$ doppelte Nullstelle das charakteristischen Polynoms. Außerdem ist jeder Vektor

$$x_{1/2} = t \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Lösung des inhomogenen linearen Gleichungssystems

$$\begin{pmatrix} 0 & -2 \\ 0 & 0 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Zwei voneinander verschiedene Vektoren x_1 und x_2 wären also parallel. Der zu der doppelten Nullstelle $\lambda_{1/2}$ des charakteristischen Polynoms gehörende Eigenraum $E_{1/2}$ ist eine Gerade im $\mathbb{R}^2$. In diesem Fall stimmt die Dimension des Eigenraums $E_{1/2}$ nicht mit der Vielfachheit des Eigenvektors $\lambda_{1/2}$ überein. Es gilt $\operatorname{rg}(A - \lambda_{1/2} I) = 1$.

(v) Für

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

sind die Eigenwerte komplexe Zahlen. Wir erhalten aus der charakteristischen Gleichung

$$\begin{aligned}
\det\begin{pmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{pmatrix} &= 0 \\
(1-\lambda)^2+4 &= 0 \\
\lambda^2-2\lambda+5 &= 0 \\
\lambda_{1/2} &= 1 \pm \sqrt{1-5}.
\end{aligned}$$

Folglich sind

$$\lambda_1 = 1+2\mathrm{i}, \qquad \lambda_2 = 1-2\mathrm{i}.$$

die Eigenwerte von A. Die Koeffizienten der dazugehörigen Eigenvektoren sind komplewertig.

Aus den Beispielen (iii) und (iv) ist ersichtlich, dass für obere und untere Dreiecksmatrizen $A = (a_{ik})$ die Eigenwerte λ_i identisch mit den Diagonalelementen a_{ii} sind,

$$\lambda_i = a_{ii}, \qquad i = 1, \ldots, n.$$

Das gilt natürlich auch für Diagonalmatrizen, da jede Diagonalmatrix gleichzeitig obere und untere Dreiecksmatrix ist.

Am Ende dieses Abschnittes soll noch ein einfacher Satz über die Eigenwerte von Matrizen formuliert werden. Dafür wird zunächst die Spur einer quadratischen Matrix eingeführt:

Definition. Die *Spur* $\operatorname{tr} A$ einer (n,n)-Matrix $A = (a_{ik})$ ist die Summe ihrer Hauptdiagonalelemente,

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}.$$

Die Bezeichnung tr ist von dem englischen Begriff „trace“ für „Spur“abgeleitet.

Satz. Es sei A eine (n,n)-Matrix mit den Eigenwerten $\lambda_1, \ldots, \lambda_n$, dann gelten für die Spur $\operatorname{tr} A$ bzw. die Determinante $\det A$ die Beziehungen

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i, \qquad \det A = \prod_{i=1}^{n} \lambda_i.$$

Daraus folgt unter anderem, dass eine reguläre Matrix nur von Null verschiedene Eigenwerte hat. Außerdem können die beiden Gleichungen leicht zur Überprüfung der Ergebnissen von Eigenwertberechnungen verwendet werden.

Beispiele

(i) Wir betrachten noch einmal das Beispiel von Seite 6. Die Determinante der Matrix

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

mit den Eigenwerten $\lambda_1 = -2$, $\lambda_2 = 1$ und $\lambda_3 = -1$ errechnet sich aus

$$\det A = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 = (-2) \cdot 1 \cdot (-1) = 2.$$

(ii) Gegen seien die Spur $\operatorname{tr} A = 0$ und die Determinante $\det A = -1$ einer $(2,2)$-Matrix. Ihre Eigenwerte λ_1 und λ_2 sind folglich Lösungen der Gleichungen

$$\lambda_1 + \lambda_2 = 0, \qquad \lambda_1 \cdot \lambda_2 = -1.$$

Daraus folgt $\lambda_1 = 1$ und $\lambda_2 = -1$.

1.3 Lineare Abbildungen

Ähnlich wie lineare, reelle Funktionen des Typs $f(x) = ax + b$ als Abbildungen $F : \mathbb{R} \mapsto \mathbb{R}$ der Menge der reellen Zahlen $\mathbb{R}$ in die Menge der reellen Zahlen aufgefasst werden können, identifizieren wir in diesem Abschnitt eine quadratische (n,n)-Matrix $A = (a_{ij})$ mit der durch sie vermittelten linearen Abbildung $\varphi(x) = Ax$, d. h.

$$\varphi : \mathbb{R}^n \mapsto \mathbb{R}^n.$$

Einer linearen Abbildung φ wird meist die Basis aus den Einheitsvektoren $e_1, \ldots, e_n$ zugrunde gelegt.

Wir nehmen der Einfachheit halber an, dass A eine symmetrische (n,n)-Matrix mit paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$ ist. Dann entspricht der Abbildung φ eine Abbildung

$$\psi : \mathbb{R}^n \mapsto \mathbb{R}^n$$

mit $\psi(y) = \Lambda y$, wobei Λ eine Diagonalmatrix mit den Eigenwerten der Matrix A in der Hauptdiagonalen ist,

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & \lambda_i & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Das System der normierten Eigenvektoren $x_1, \ldots, x_n$ von A bildet eine Basis der Abbildung ψ. (Da wir vorausgesetzt haben, dass A symmetrisch ist, bildet das Eigenvektorsystem sorgar eine *Orthonormalbasis*, d. h. die Basisvektorn sind orthogonal und ihre Länge ist Eins.)

Die Tatsache, dass ψ eine lineare Abbildung bezüglich der neuen Basis $x_1, \ldots, x_n$ ist, folgt unmittelbar aus Gleichung (1.1). Sei y ein beliebiger Vektor, $y \in \mathbb{R}^n$. Dann gilt für alle $i = 1, \ldots, n$

$$\begin{aligned} Ax_i &= \lambda_i x_i \\ (Ax_i)' &= \lambda_i x_i' \\ x_i' A' &= \lambda_i x_i' \\ x_i' A &= \lambda_i x_i', && \text{da } A \text{ symmetrisch ist} \\ x_i' A y &= \lambda_i x_i' y. \end{aligned}$$

(Dabei entspricht das Skalarprodukt $x_i'y$ der i-ten Koordinate im neuen Koordnatensystem; $|x_i'y|$ ist die Länge der orthogonalen Projektion von y auf x_i.)

Sei nun C die Matrix, deren Spaltenvektorn die (normierten) Eigenvektoren von A sind, dann lässt sich die letzte Gleichung in der Form

$$\begin{aligned} C'Ay &= \Lambda C'y \\ C'\varphi(y) &= \psi(C'y) \end{aligned}$$

schreiben. Das bedeutet, es ist egal, ob wir die Abbildung $\varphi(y)$ eines Vektors y bezüglich der Basis von ψ betrachten ober ob der Vektor y bzüglich des neuen Koordinatensystems mit ψ abgebildet wird, siehe auch Abbildung 1.1.

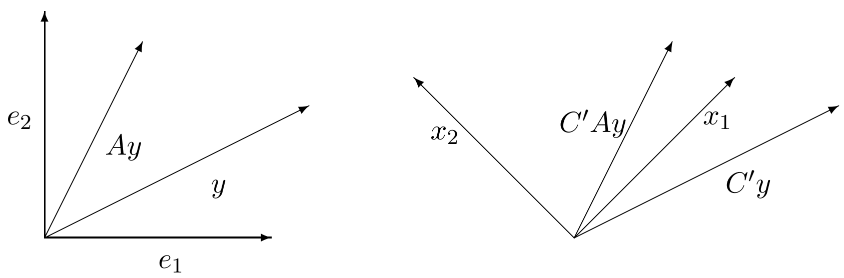

Abbildung 1.1: Darstellung von y und Ay, links: im kartesischen Koordinatensystem, d. h. bzüglich der Basis $e_1, \ldots, e_n$ und rechts: im Koordinatensystem bezüglich der Basis $x_1, \ldots, x_n$.

Bedeutung. Wird im Zusammenhang mit einer Anwendung stets die gleiche Abbildung φ verwendet, dann erscheint es zweckmäßig, das Koordinatensystem zu wechseln und statt φ die einfachere Abbildung ψ anzuwenden.

Beispiel. Wir betrachten noch einmal das Beispiel (ii) von Seite 9,

$$A = \begin{pmatrix} 5 & -\sqrt{12} \\ -\sqrt{12} & 1 \end{pmatrix}, \qquad \lambda_1 = 7, \quad \lambda_2 = -1, \qquad x_1 = \frac{1}{2}\begin{pmatrix} \sqrt{3} \\ -1 \end{pmatrix}, \quad x_2 = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix},$$

wobei hier die Eigenvektoren x_1 und x_2 als normierte Vektoren angegeben sind. Wir wählen willkürlich einen Vektor y, z. B.

$$y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Dann gilt

$$\varphi(y) = Ay = \begin{pmatrix} 5 - \sqrt{12} \\ 1 - \sqrt{12} \end{pmatrix}.$$

Jetzt drücken wir das Ergebnis dieser Abbildung in den Koordinaten des neuen Koordinatensystems aus,

$$C'\varphi(y) = C'Ay = \frac{1}{2}\begin{pmatrix} 7\sqrt{3}-7 \\ -1-\sqrt{12} \end{pmatrix}.$$

Nun gehen wir den umgekehrten Weg. Wir drücken zuerst y in den neuen Koordinaten aus,

$$C'y = \begin{pmatrix} \sqrt{3}-1 \\ 1+\sqrt{3} \end{pmatrix},$$

und wenden anschließend die Abbildung ψ an,

$$\psi(C'y) = \frac{1}{2}\begin{pmatrix} 7\sqrt{3}-7 \\ -1-\sqrt{12} \end{pmatrix}.$$

Die Ergebnisse für $C'\varphi(y)$ und $\psi(C'y)$ sind also identisch.

1.4 Orthogonale Matrizen

Wir betrachten im folgenden spezielle Matrizen, d. h. spezielle lineare Abbildungen.

Definition. Eine (n,n)-Matrix A heißt *orthogonal*, wenn für alle $x, y \in \mathbb{R}^n$ gilt

$$(Ax)'Ay = x'y. \tag{1.4}$$

Bemerkungen

(i) Aus Gleichung (1.4) folgt für $x = y$ damit unmittelbar

$$|Ax| = \sqrt{(Ax)'(Ax)} = \sqrt{x'A'Ax} = \sqrt{x'x} = |x|.$$

Das bedeutet, die zu einer orthogonalen Matrix A gehörige lineare Abbildung $x \mapsto Ax$ lässt die Länge $|x| = \sqrt{x'x}$ eines Vektors x unverändert.

(ii) Wegen

$$\cos\angle(Ax, Ay) = \frac{(Ax)'(Ay)}{|Ax|\cdot|Ay|} = \frac{x'A'Ay}{|x|\cdot|y|} = \frac{x'y}{|x|\cdot|y|} = \cos\angle(x,y)$$

ändern sich auch die Winkel zwischen zwei Vektoren bei einer linearen Abbildung nicht, wenn A orthogonal ist.

(iii) Verwenden wir statt x und y die Einheitsvektoren e_i und e_j, für die die i-te bzw. j-te Komponente gleich Eins sind; alle anderen Komponenten sind gleich Null. Dann gilt

$$Ae_i = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} e_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}.$$

Damit erhalten wir aus Gleichung (1.4)

$$
\begin{aligned}
(Ae_i)'(Ae_j) &= e_i'e_j \\
e_i'A'Ae_j &= e_i'e_j \\
e_i' \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} e_j &= e_i'e_j \\
(a_{1i}, \ldots, a_{ni}) \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} &= \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} .
\end{aligned}
$$

Das bedeutet, die Spaltenvektoren von A sind orthogonal zueinander und haben die Länge 1. Ein solches System von Spaltenvektoren heißt *Orthonormalsystem*.

(iv) Wegen $\det A = \det A'$ gilt für orthogonale Matrizen

$$(\det A)^2 = (\det A')(\det A) = \det(A'A) = \det I = 1.$$

Damit ist $\det A \neq 0$, und es existiert auch stets die zu A inverse Matrix A^{-1}. Daraus folgt

$$
\begin{aligned}
A'A &= I \\
A'A \cdot A^{-1} &= I \cdot A^{-1} \\
A' &= A^{-1} \\
A \cdot A' &= A \cdot A^{-1} \\
AA' &= I.
\end{aligned}
$$

Folglich bilden auch die Zeilenvektoren einer orthogonalen Matrix ein Orthonormalsystem.

Die letzten beiden Bemerkungen fassen wir wegen ihrer Bedeutung in einem Satz zusammen:

Satz. Die Systeme der Spalten- bzw. Zeilenvektoren einer quadratischen Matrix bilden genau dann Orthonormalsysteme, wenn A orthogonal ist. Für orthogonale Matrizen gelten die Gleichungen

$$A'A = I, \qquad AA' = I, \qquad A' = A^{-1}.$$

Beispiele

(i) Die Matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

ist orthogonal, denn das System der Spaltenvektoren bildet ein Orthonormalsystem,

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}' \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 1, \qquad \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}' \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = 0, \qquad \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}' \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = 1.$$

Da A außerdem symmetrisch ist, bildet auch das System der Zeilenvektoren ein Orthogonalsystem.

(ii) Die Drehmatrix

$$A_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, \qquad \varphi \in [0, 2\pi)$$

bewirkt für einen gegebenen Drehwinkel φ weder eine Längenänderung noch die Änderung des Winkels zwischen zwei Vektoren. Die Matrix A_φ ist folglich orthogonal (für jedes φ). Das kann natürlich leicht überprüft werden, indem man zeigt, dass die Spaltenvektoren von A_φ ein Orthonormalsystem bilden. Übrigens ist die Matrix A_φ für $\varphi = \frac{\pi}{4}$ identisch mit der Matrix A aus Beispiel (i).

(iii) Die Matrizen

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

sind orthogonale Matrizen. Sie bewirken Spiegelungen an der y-Achse, an der x-Achse bzw. am Koordinatenursprung.

1.5 Transformation einer Matrix auf Diagonalgestalt

In diesem Abschnitt wird folgendes Problem behandelt: Gibt es zu einer quadratischen Matrix A eine Matrix C mit der Eigenschaft, dass die Matrix $C^{-1}AC$ Diagonalgestalt hat. Einen ersten Hinweis zur Lösung des Problems liefert die folgende Aussage:

Satz. Sei A eine reguläre (n, n)-Matrix mit den Eigenwerten $\lambda_1, \dots, \lambda_n$. Wenn es zu den Eigenwerten von A ein System von n unabhängigen Eigenvektorn gibt, dann gilt

$$C^{-1}AC = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & \lambda_i & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}, \tag{1.5}$$

wobei C die Matrix der normierten Eigenvektoren von A ist.

Die Transformation $C^{-1}AC$ heißt *Hauptachsentransformation* von A.

Für welche Matrizen ist die Voraussetzung des Satzes erfüllt? Aus Abschnitt 1.1 wissen wir, dass das eingangs beschriebene Problem lösbar ist, wenn A symmetrisch und alle Eigenwerte voneinander verschieden sind. Aber auch dann, wenn die Eigenwerte einer symmetrischen Matrix nicht alle voneinander verschieden sind, lässt sich eine Matrix C finden, so dass (1.5) gilt. Die Matrix C lässt sich in diesem Fall sogar so wählen, dass sie orthogonal ist, d. h. $C^{-1} = C'$.

Satz. Wenn die Matrix A symmetrisch ist, dann gibt es zu A eine orthogonale Matrix C, so dass die Gleichung (1.5) erfüllt ist, d. h.

$$C'AC = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & \lambda_i & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}. \tag{1.6}$$

Beispiele

(i) Die symmetrische Matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

hat die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 3$. Dazu gehörige Eigenvektoren sind

$$x_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \text{und} \qquad x_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Nach einer Normierung der Eigenvektoren erhält man daraus die Matrix

$$C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Es gilt

$$\begin{aligned} C'AC &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \cdot \left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \right] \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \end{aligned}$$

(ii) Die Matrix

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

ist singulär. Mindestens einer der Eigenwerte ist gleich Null. Die Eigenwertgleichung

$$(1 - \lambda)^2 - 1 = 0$$

hat die Lösungen $\lambda_1 = 0$ und $\lambda_2 = 2$. Die dazu gehörenden normierten Eigenvektoren sind

$$x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad x_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Mit

$$C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

erhalten wir erwartungsgemäß

$$C'AC = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

(iii) Noch ein etwas komplizierteres Beispiel, in dem die verschiedenen Facetten des Problems besser zum Ausdruck kommen. Gegeben sei die Matrix

$$A = \begin{pmatrix} 5 & -2 & 2 \\ -2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix}.$$

Ihre Eigenwertgleichung

$$\begin{aligned} \left[(5-\lambda)(2-\lambda)^2 + 4 + 4\right] - \left[4(2-\lambda) + 4(2-\lambda) + 5 - \lambda)\right] &= 0 \\ \lambda^3 - 9\lambda^2 + 15\lambda - 7 &= 0 \end{aligned}$$

hat die Lösungen $\lambda_{1/2} = 1$ und $\lambda_3 = 7$. Zum doppelten Eigenwert $\lambda_{1/2} = 1$ erhält man aus dem homogenen linearen Gleichungssystem

$$\begin{pmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

z. B. die beiden Eigenvektoren

$$x_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}.$$

Mit $\lambda_3 = 7$ lässt sich aus

$$\begin{pmatrix} -2 & -2 & 2 \\ -2 & -5 & -1 \\ 2 & -1 & -5 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

der Vektor

$$x_3 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$$

bestimmen. Da die Matrix A symmetrisch ist, müssen die beiden Vektoren x_1 und x_2 orthogonal zu x_3 sein, da der Eigenraum zu $\lambda_{1/2}$ orthogonal zum Eigenraum von λ_3 ist. Jedoch sind die beiden Vektoren x_1 und x_2 nicht zueinander orthogonal, $x_1'x_2 = -2$. Wir ersetzen daher x_2 durch einen Vektor, der sowohl zu x_1 als auch zu x_3 orthogonal ist,

$$x_2 = x_1 \times x_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix}.$$

Nun sind die drei Vektoren x_1, x_2, x_3 paarweise orthogonal (d. h. sie bilden ein Orthogonalsystem). Normierung liefert schließlich die orthogonale Transformationsmatrix

$$C = \frac{1}{\sqrt{6}} \begin{pmatrix} 0 & \sqrt{2} & 2 \\ \sqrt{3} & \sqrt{2} & -1 \\ \sqrt{3} & -\sqrt{2} & 1 \end{pmatrix},$$

von der wir wissen, dass durch $C'AC$ eine Hauptachsentransformation von A erfolgt.

1.6 Anwendungen

Für die Eigenwerttheorie gibt es eine Vielzahl von Anwendungen, von denen hier nur drei betrachtet werden, für deren Behandlung nur wenige neue Begriffe zusätzlichg eingeführt werden müssen.

1.6.1 Quadratische Formen

Wir betrachten zunächst die Gleichung einer Ellipse im $\mathbb{R}^2$ mit den Längen a und b der Halbachsen,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1. \tag{1.7}$$

Eine *Ellipse* E ist also die Menge der Punkte des $\mathbb{R}^2$, die diese Ungleichung erfüllen,

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,:\, \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\}.$$

Diese Ellipse ist im Koordinatenursprung zentriert; die Halbachse mit der Länge a ist in x-Richtung ausgerichtet und die Halbachse mit der Länge b in y-Richung, siehe Abbildung

Die linke Seite der Ungleichung (1.7) kann unter Zuhilfenahme der Matrixschreibweise in der Form

$$\begin{pmatrix} x \\ y \end{pmatrix}' \begin{pmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

geschrieben werden. Eine Darstellung dieser Art wird quadratische Form genannt. Im folgenden soll der allgemeine Fall einer quadratischen Form betrachtet werden.

Definition. Sei A eine symmetrische (n, n)-Matrix, dann heißt der Ausdruck

$$x'Ax, \qquad x \in \mathbb{R}^n$$

quadratische Form.

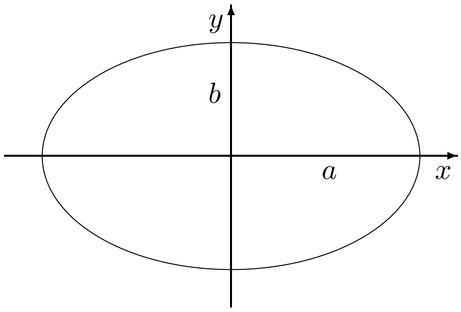

Abbildung 1.2: Eine Ellipse mit den Halbachsenlängen a und b.

In Anwendungen sind solche quadratischen Formen besonders interessant, für die die Matrix A nur positive Eigenwerte hat.

Definition. Eine symmetrische Matrix A heißt *positiv definit*, wenn alle ihre Eigenwerte größer als Null sind, und A heißt *positiv semidefinit*, wenn die Eigenwerte nicht negativ sind.

Jede positiv definite Matrix ist also zugleich auch positiv semidefinit.

Damit lassen sich n-dimensionale Ellipsen einführen – so genannte Ellipsoide.

Definition. Sei A eine positiv semidefinite (n, n)-Matrix mit den Eigenwerten $\lambda_1, \ldots, \lambda_n$ und den Eigenvektoren $x_1, \ldots, x_n$. Die Menge

$$E = \left\{ x \in \mathbb{R}^n \, : \, x'Ax \leq 1 \right\}$$

heißt *Ellipsoid* mit den Halbachsenlängen

$$a_i = \begin{cases} \frac{1}{\sqrt{\lambda_i}} & \text{für} \quad \lambda_i > 0 \\ 0 & \text{für} \quad \lambda_i = 0 \end{cases},$$

wobei die Richtungen der Halbachsen den Richtungen der Eigenvektoren x_i entsprechen.

Bemerkungen

(i) Klar, eine positiv definite Matrix ist auch positiv semidefinit.

(ii) Eine positiv definite (n, n)-Matrix A ist regulär, d. h. $\operatorname{rg} A = n$, $\det A > 0$.

(iii) Für jede positiv semidefinite (n, n)-Matrix A gibt es eine (m, n)-Matrix B, so dass $A = B'B$ gilt. Wenn der Rang von B größer oder gleich n ist, $\operatorname{rg} B \geq n$, dann ist $\operatorname{rg} A = n$.

(iv) Das n-dimensionale Volumen $V(E)$ eines Ellipsoids mit den Eigenwerten $\lambda_1, \ldots, \lambda_n$ errechnet sich aus

$$V(E) = \frac{\pi^{n/2}}{\Gamma\left(1 + \frac{n}{2}\right)} \prod_{i=1}^{n} \frac{1}{\sqrt{\lambda_i}} \quad = \frac{\pi}{\Gamma\left(1 + \frac{n}{2}\right)} \frac{1}{\sqrt{\det A}},$$

wobei Γ die Eulersche Gamma-Funktion bezeichnet. Im 2D-Fall erhält man für die Ellipsenfläche $A(E) = \pi / \sqrt{\lambda_1 \lambda_2}$, und für 3D-Ellipsoide gilt

$$V(E) = \frac{4\pi}{3} \frac{1}{\sqrt{\lambda_1 \lambda_2 \lambda_3}}.$$

Beispiel. Die im Beispiel (i) auf Seite 17 betrachtete Matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

hat die positiven Eigenwerte $\lambda_1 = 3$ und $\lambda_2 = 1$; A ist folglich positiv definit, und die Ungleichung

$$x' \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} x < 1$$

beschreibt eine Ellipse. Die Längen der Hauptachsen betragen $1/\sqrt{3}$ und 1. Somit hat die Ellipse eine Fläche von $A(E) = \pi/\sqrt{3}$.

Die Richtung der Hauptachsen sind durch die zu λ_1 und λ_2 gehörenden Eigenvektoren

$$x_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

gegeben. Die Winkel φ_1 und φ_2 zwischen der x-Achse den Eigenvektoren berechnen sich aus

$$\alpha_1 = \arccos \angle(x_1, e_1), \qquad \alpha_2 = \arccos \angle(x_2, e_1) \qquad \text{mit} \quad e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Wir erhalten $\alpha_1 = 3\pi/4$ für die kleine und $\alpha_2 = \pi/4$ für die große Halbachse.

1.6.2 Die Hauptkomponente in Daten

Wir gehen in diesem Abschnitt der Einfachheit halber davon aus, dass nur zwei Merkmale X_1 und X_2 (z. B. zwei physikalische, biologische, soziologische, ... Kennwerte) untersucht werden. $x_{11}, \ldots, x_{1n}$ bzw. $x_{21}, \ldots, x_{2n}$ seien die Messwerte von X_1 bzw. X_2. Ihre Zentrierungen $x_{1i} - \bar{x}_1$ bzw. $x_{2i} - \bar{x}_2$ werden zu einer Matrix A zusammengefasst,

$$A = \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 \\ \vdots & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 \end{pmatrix},$$

wobei

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_{i1}, \qquad \bar{x}_2 = \frac{1}{n} \sum_{i=1}^{n} x_{i2}$$

die Mittelwerte der x_{i1} bzw. x_{i2} bezeichnen.

Es gilt

$$
\begin{aligned}
A'A &= \begin{pmatrix} \sum\limits_{i=1}^{n}(x_{i1}-\bar{x}_1)^2 & \sum\limits_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2) \\ \sum\limits_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2) & \sum\limits_{i=1}^{n}(x_{i2}-\bar{x}_2)^2 \end{pmatrix} \\
&= \begin{pmatrix} \sum\limits_{i=1}^{n} x_{i1}^2 & \sum\limits_{i=1}^{n} x_{i1}x_{i2} \\ \sum\limits_{i=1}^{n} x_{i1}x_{i2} & \sum\limits_{i=1}^{n} x_{i2}^2 \end{pmatrix} - n \cdot \begin{pmatrix} \bar{x}_1^2 & \bar{x}_1\bar{x}_2 \\ \bar{x}_1\bar{x}_2 & \bar{x}_2^2 \end{pmatrix}.
\end{aligned}
$$

Der Eigenvektor zum größten Eigenwert von $A'A$ heißt *Hauptkomponente* der (zentrierten) Daten A.

Bemerkungen

(i) Diese Form der Hauptkomponentenanalyse lässt sich leicht auf Daten A von mehr als zwei Merkmalen übertragen. An Stelle einer Hauptkomponenten werden dann, mehrere Hauptkomponenten bestimmt, die den Eigenvektoren von $A'A$ zu den größten Eigenwerten bestimmt werden.

(ii) Neben der Zentrierung der Daten ist häufig noch eine Normierung mit der Standardabweichungen zweckmäßig. Sei

$$
s_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2} \qquad \text{mit} \qquad \bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}
$$

die Standardabweichung im Merkmal X_j, dann werden die zum Merkmal X_j gehörenden Daten mit Hilfe von

$$
\xi_{ij} = \frac{x_{ij}-\bar{x}_j}{s_j}, \qquad i = 1, \ldots, n
$$

normiert. Das ist vor allem dann sinnvoll, wenn sich die zu untersuchenden Merkmale in ihren Standardabweichungen stark unterscheiden.

Beispiel. Die Daten $1, 2, 3, 4$ eines Merkmals X_1 und die Daten $2, 3, 3, 4$ eines Merkmals X_2 haben die Mittelwerte $\bar{x}_1 = \frac{3}{2}$ bzw. $\bar{x}_2 = 3$ Es gilt

$$
\begin{aligned}
A'A &= \begin{pmatrix} 14 & 21 \\ 21 & 38 \end{pmatrix} - 4 \cdot \begin{pmatrix} \frac{9}{4} & \frac{9}{2} \\ \frac{9}{2} & 9 \end{pmatrix} \\
&= \begin{pmatrix} 14 & 21 \\ 21 & 38 \end{pmatrix} - \begin{pmatrix} 9 & 18 \\ 18 & 36 \end{pmatrix} \\
&= \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}
\end{aligned}
$$

Die Eigenwerte von $A'A$ errechnen sich aus

$$\begin{aligned} \det(A'A - \lambda I) &= \det \begin{pmatrix} 5-\lambda & 3 \\ 3 & 2-\lambda \end{pmatrix} \\ &= (5-\lambda)(2-\lambda) - 9 \\ &= \lambda^2 - 7\lambda + 1 = 0 \\ \lambda_{1/2} &= \frac{7}{2} \pm \sqrt{\frac{49}{4} - 1}. \end{aligned}$$

Von Interesse ist lediglich der größte Eigenwert $\lambda_1 = 6,654$. Der dazu gehörige Eigenvektor $x_1 = \begin{pmatrix} u \\ v \end{pmatrix}$ ist Lösung der Gleichung

$$\begin{pmatrix} 5-\lambda_1 & 3 \\ 3 & 2-\lambda_1 \end{pmatrix} x_1 = 0.$$

Wir erhalten den Eigenvektor

$$x_1 = \frac{1}{1,902} \begin{pmatrix} 1,616 \\ 1 \end{pmatrix} = \begin{pmatrix} 0,850 \\ 0,526 \end{pmatrix}$$

mit der Richtung $\varphi_1 = \arctan \frac{v}{u} = 31,8°$, siehe Abbildung 1.3. Der Vektor x_1 ist die Hauptkomponente der Daten.

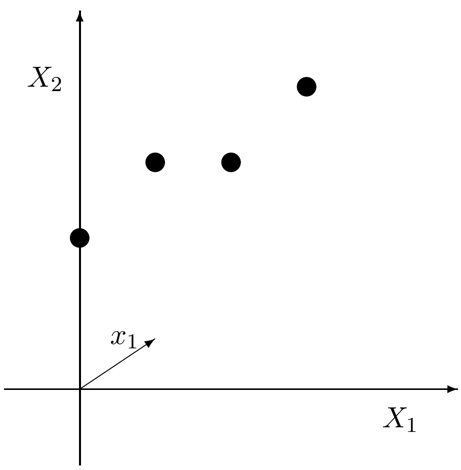

Abbildung 1.3: Die Richtung des Eigenvektors x_1 zum größten Eigenwert λ_1 der Matrix $A'A$ ist gleich der Richtung einer (gedachten) Ausgleichsgeraden durch die Datenpunkte.

1.6.3 Die Hauptrichtung von Objekten

Wir betrachten das Problem zur Vereinfachung der Darstellung nur für diskretisierte Teilmengen (Objekte) des $\mathbb{R}^2$. Sei also X eine beschränkte Teilmenge des $\mathbb{R}^2$. In der Bildanalyse wird von einem

Bild der Menge X auf einem 2D-Gitter $\mathbb{L}^2 = c\mathbb{Z}^2$ ausgegangen, wobei wie bisher $\mathbb{Z}$ die Menge der ganzen Zahlen bezeichnet und c der Gitterabstand ist, $c > 0$. Mit dem Durchschnitt

$$X \cap \mathbb{L}^2$$

wird häufig der Bildvordergrund bezeichnet, siehe Abbildung 1.4.

Fasst man die (zentrierten) Koordinaten der zu einem Objekt gehörenden Pixel zu einer Matrix A zusammen, dann kann die Hauptrichtung des Objekts als Richtung des Eigenvektors x_1 zum größten Eigenwert λ_1 der Matrix $A'A$ bestimmt werden.

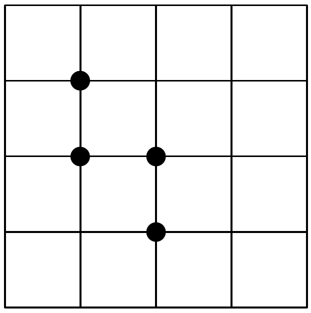

Abbildung 1.4: Ein diskretes Objekt X auf einem Gitter.

Beispiel. Die Matrix A der zentrierten Pixelkoordinaten des Objekte in Abbildung 1.4 ist

$$A = \begin{pmatrix} -\frac{1}{2} & 1 \\ -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 \\ \frac{1}{2} & -1 \end{pmatrix}.$$

Die Rechnung läuft wie folgt:

$$\begin{aligned} A'A &= \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \\ \det(A'A - \lambda I) &= (1-\lambda)(2-\lambda) - 1 \\ &= \lambda^2 - 3\lambda + 1 = 0 \\ \lambda_{1/2} &= \frac{3}{2} \pm \sqrt{\frac{9}{4} - 1} \end{aligned}$$

Der größte Eigenwert ist also $\lambda_1 = \frac{3+\sqrt{5}}{2}$. Der dazu gehörige Eigenvektor x_1 ist Lösung des homogenen linearen Gleichungssystems

$$\begin{pmatrix} -\frac{1+\sqrt{5}}{2} & -1 \\ -1 & \frac{1-\sqrt{5}}{2} \end{pmatrix} x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Daraus folgt unmittelbar

$$x_1 = \begin{pmatrix} 1 \\ -\frac{1+\sqrt{5}}{2} \end{pmatrix}.$$

Die Richtung von x_1 und damit die Hauptrichtung des Objekts ist

$$\varphi_1 = \arctan\left(-\frac{1+\sqrt{5}}{2}\right) + 2\pi = 301,7^\circ.$$

Kapitel 2

Lineare Gleichungssysteme II

In diesem Abschnitt werden Faktorisierungen der Koeffizientenmatrix von linearen Gleichungssystemen behandelt, d. h. Darstellungen der Koeffizientenmatrix als Produkt von Matrizen. Solche Faktorisierungen sind für die Lösung von linearen Gleichungssystemen nützlich und daher die Grundlage für ihre numerische Behandlung.

Wir wollen hier zunächst an das Gauß-Verfahren zur Lösung linearer Gleichungssysteme anknüpfen und die LU-Zerlegung der Koeffizientenmatrix einführen. Es zeigt sich sehr schnell, dass die LU-Zerlegung viele Unzulänglichkeiten hat, die z. B. durch Pivotisierung behoben werden können. Dennoch ist die LU-Zerlegung mit Pivotisierung an quadratische Koeffizientenmatrizen gebunden.

Um auch über- und unterbestimmte lineare Gleichungssysteme behandeln zu können, wird die Singulärwertzerlegung (SVD) für allgemeine Rechteckmatrizen eingeführt. Die Lösung eines linearen Gleichungssystems auf der Grundlage der SVD ist sicher das Verfahren, das in den meisten Fällen zum Erfolg führen wird. Es liefert darüber hinaus präzise Aussagen zur Lösbarkeit des Gleichungssystems und zur Kondition der Koeffizientenmatrix sowie eine Beschreibung der Lösungsmenge.

Es gibt weitere Verfahren zur Lösung linearer Gleichungssysteme, die auf Faktorisierungen der Koeffizientenmatrix basieren, z. B. das Cholesky-Verfahren, die im Rahmen der Vorlesung nicht behandelt werden können. Diese Verfahren bieten in wichtigen Spezialfällen Vorteile gegenüber der SVD. Diesbezüglich wird auf die Numerikvorlesung bzw. auf einschlägige Literatur über numerische Mathematik verwiesen.

2.1 Die elementare Gauß-Elimination

Wir wiederholen zur Einführung zunächst das einfache Gauß-Verfahren zur Lösung eines linearen Gleichungssystems $Ax = b$. Bei der Anwendung des Gauß-Verfahrens soll aus Gründen der Systematik zunächst auf die Vertauschung von Zeilen verzichtet werden (*elementare* oder *naive Gauß-Eliminiation*).

Beispiele

(i) Das lineare Gleichungssystem

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{pmatrix} x = \begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}$$

kann mit dem elementaren Gauß-Verfahren gelöst werden. Durch die Vorwärtssubstitution erhält die Koeffizientenmatrix eine Dreicksform,

$$\begin{array}{ll}
 & \begin{array}{rrr|l} 10 & -7 & 0 & 7 \\ -3 & 2 & 6 & 4 \\ 5 & -1 & 5 & 6 \end{array} \quad \begin{array}{c} \\ (\mathrm{II} + 0,3\,\mathrm{I}) \\ (\mathrm{III} + 0,5\,\mathrm{I}) \end{array} \\
\Longrightarrow & \begin{array}{rrr|r} 10 & -7 & 0 & 7 \\ 0 & -0,1 & 6 & 6,1 \\ 0 & 2,5 & 5 & 2,5 \end{array} \quad \begin{array}{c} \\ \\ (\mathrm{III} - 25\,\mathrm{II}) \end{array} \\
\Longrightarrow & \begin{array}{rrr|r} 10 & -7 & 0 & 7 \\ 0 & -0,1 & 6 & 6,1 \\ 0 & 0 & 155 & 155 \end{array} \quad .
\end{array}$$

Dabei werden die Multiplikatoren $m_{21} = 0,3$, $m_{31} = 0,5$ und $m_{32} = 25$ verwendet. Rückwärtssystitution liefert nacheinander die Komponenten $x_3 = 1$, $x_2 = -1$ und $x_1 = 0$ des Lösungsvektors x, d. h.

$$x = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

(ii) Das in Beispiel (i) betrachtete lineare Gleichungssystem wird nun geringfügig modifiziert, um die Effekte von Rundungsfehlern zu demonstrieren,

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2,099 & 6 \\ 5 & -1 & 5 \end{pmatrix} x = \begin{pmatrix} 7 \\ 3,901 \\ 6 \end{pmatrix}.$$

Vorwärtssubstitution mit den Multiplikatoren $m_{21} = 0,3$, $m_{31} = 0,5$ und $m_{32} = -2\,500$ liefert

$$\begin{array}{rrr|c} 10 & -7 & 0 & 7 \\ 0 & -0.001 & 6 & 6,001 \\ 0 & 0 & 15\,005 & 15002,5 + 2,5 \end{array} \quad .$$

Die Rückwärtssubstitution liefert die gleiche (exakte) Lösung x wie in (i). Verwendet man bei der Rechnung jedoch nur 5 Dezimalstellen, dann liefert die Rückwärtssubstitution

$$
\begin{aligned}
15\,005 x_3 &= 15\,003 + 2,5 = 15\,006 \\
x_3 &= \frac{15\,006}{15\,006} = 1,000\,1 \\
-0,001 x_2 + \underbrace{6 \cdot 1,000\,1}_{6,000\,6} &= 6,001 \\
-0,001 x_2 &= 0,000\,4 \\
x_2 &= \frac{0,000\,4}{-0,001} = -0,4 \\
10 x_1 - 7 \cdot (-0,4) &= 7 \\
10 x_1 + 2,8 &= 7 \\
x_1 &= 0,42.
\end{aligned}
$$

Die Lösung

$$
\tilde{x} = \begin{pmatrix} 0,42 \\ -0,4 \\ 1,000\,1 \end{pmatrix}
$$

weicht also als Folge von Rundungsfehlern erheblich von der exakten Lösung x ab.

Mit Hilfe der eingeführten Multiplikatoren m_{ik} kann das Gauß-Verfahren algorithmisch in der folgenden Form notiert werden:

Vorwärtssubstitution.

$$
\text{für } k = 1, \ldots, n-1 \\
\left.\begin{array}{rcll}
m_{ik} &=& -a_{ik} & \\
a_{ij} &:=& a_{ij} + m_{ik} a_{ik}, & j = k+1, \ldots, n \\
b_i &:=& b_i + a_{ik} b_k &
\end{array}\right\}, \qquad i = k+1, \ldots, n.
$$

Rückwärtssubstitution.

$$
\begin{aligned}
x_n &= b_b / a_{nn} \\
x_i &= \frac{1}{a_{ii}} \left(b_i - \sum_{j=i+1}^{n} a_{ij} x_j \right), \qquad i = n-1, n-2, \ldots, 1.
\end{aligned}
$$

Bemerkungen

(i) Die elementare Gauß-Elimination setzt voraus, dass in jedem Schritt der Vorwärtssubstitution (also für jedes k) das Diagonalelement verschieden von Null ist, $a_{ii} \neq 0$. Andernfalls muss der Algorithmus im i-ten Schritt abgebrochen werden.

(ii) Für singuläre Koeffizientenmatrizen A gibt es stets einen Index k, so dass $a_{ii} = 0$ ist.

(iii) Leider kann $a_{ii} = 0$ auch für reguläre Matrizen erhalten werden. Der Algorithmus wird abgebrochen, obwohl das lineare Gleichungssystem eine eindeutig bestimmte Lösung hat.

(iv) Auch dann, wenn $a_{ii} \neq 0$ für alle i, kann die mit der elementaren Gauß-Elimination erhaltene Lösung des linearen Gleichungssystems wegen unkontrollierter Fehlerfortpflanzung stark von der exakten Lösung abweichen.

2.2 Die LU-Zerlegung einer quadratischen Matrix

Wir führen in diesem Abschnitt eine Faktorisierung der Koeffizientenmatrix A des linearen Gleichungssystems $Ax = b$ ein, die der elementaren Gauß-Elimination entspricht. Diese Faktorisierung, d. h. die Zerlegung von A in ein Produkt aus einer unteren (*lower*) Dreiecksmatrix L und eine obere (*upper*) Dreiecksmatrix U wird LU-Zerlegung (*LU decpomposition*) genannt.

Wir betachten zunächst zur Einführung der Matrixschreibweise der elementaren Gauß-Elimination wieder den 3D-Fall mit der Koeffizientenmatrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Da sich die Matrix A im Verlauf der Gauß-Elimination ändert, führen wir einen oberen Index ein, der den Eliminationschritten der Vorwärtssubstitution entspricht. Wir schreiben daher für die Ausgangsmatrix $A = A^0$ und übertragen diese Indexierung dann auch auf die durch den Eliminationsschritte geänderten Koeffizienten

Der erste Eliminationsschritt entspricht einer Multiplikation der Matrix A mit der Matrix

$$M_{21} = \begin{pmatrix} 1 & 0 & 0 \\ m_{12} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \text{wobei} \qquad m_{21} = -a_{21}/a_{11}.$$

Damit erhalten wir

$$M_{21}A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} + m_{21}a_{12} & a_{23} + m_{21}a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Das Resultat der Multiplikation setzen wir gleich A^1, d. h. $M_{12}A = A^1$, und wir schreiben zur Abkürzung

$$A^1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a^1_{22} & a^1_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Analog erhalten wir für

$$M_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ m_{31} & 0 & 1 \end{pmatrix} \qquad \text{mit} \qquad m_{31} = -a_{31}/a_{11}$$

im nächsten Eliminationsschritt

$$M_{31}A^1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^1 & a_{23}^1 \\ 0 & a_{32} + m_{31}a_{12} & a_{33} + m_{31}a_{13} \end{pmatrix} := \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^1 & a_{23}^1 \\ 0 & a_{32}^2 & a_{33}^2 \end{pmatrix} = A^2.$$

Der dritte Eliminationsschritt mit

$$M_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ m_{31} & 0 & 1 \end{pmatrix} \qquad \text{und} \qquad m_{32} = -a_{31}^2/a_{22}^1$$

vervollständigt die Transformation von A auf Dreiecksgestalt, d. h., wir erhalten die obere Dreiecksmatrix U durch

$$M_{32}A^2 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^1 & a_{23}^1 \\ 0 & 0 & a_{33} + m_{32}a_{23} \end{pmatrix} := \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^1 & a_{23}^1 \\ 0 & 0 & a_{33}^3 \end{pmatrix} = U.$$

Die drei Transformationen können zusammenfassend in der Form

$$M_{23}M_{31}M_{21}A = U$$

geschrieben werden. Daraus erhält man die Darstellung

$$\begin{aligned} A &= (M_{23}M_{31}M_{21})^{-1}U \\ &= M_{21}^{-1}M_{31}^{-1}M_{32}^{-1}U. \end{aligned}$$

Man kann zeigen, dass

$$M_{21}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad M_{31}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -m_{31} & 0 & 1 \end{pmatrix}, \qquad M_{32}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{31} & 1 \end{pmatrix}$$

und

$$M_{21}^{-1}M_{31}^{-1}M_{32}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ -m_{31} & -m_{31} & 1 \end{pmatrix},$$

d. h., die letzte Matrix ist eine untere Dreiecksmatrix, die in der Literatur mit L bezeichnet wird, $L = M_{21}^{-1}M_{31}^{-1}M_{32}^{-1}$. Ihre Koeffizienten entsprechen bis auf das Vorzeichen den Multiplikatoren, die bei der Gauß-Elimination erhalten werden.

Die Gauß-Elimination führt im 3D-Fall also zu einer Zerlegung der Koeffizientenmatrix A des linearen Gleichungssystems $Ax = b$ in eine untere und eine obere Dreiecksmatrix, $A = LU$. Diese

Aussage gilt unter bestimmten Bedingungen für beliebige (n, n)-Matrizen:

Satz. Sei A eine reguläre (n, n)-Matrix und sind alle Diaginalelemente a_{ii}^j, die im Verlaufe der elementaren Gauß-Elimination erhalten werden, verschieden von Null, dann gibt es eine untere Dreiecksmatrix L, deren Diagonalelemente gleich Eins sind, und eine obere Dreiecksmatrix U, so dass

$$A = LU \qquad (LU\text{-Zerlegung von } A). \tag{2.1}$$

Falls eine LU-Zerlegung von A existiert, ist sie eindeutig.

Bemerkungen

(i) Der Algorithmus für die LU-Zerlegung einer quadratischen Matrix entspricht im wesentlichen der elementaren Gauß-Elimination.

(ii) Die LU-Zerlegung einer quadratischen Matrix A wird durch entsprechende Software realisiert, wobei in der Regel aus Speicherplatzgründen die Koeffizienten der Matrizen L und U auf den Koffizienten von A zurück gegeben werden. Genauer: die Koeffizienten von A unterhalb der Hauptdiagonalen werden durch die negativen Multiplikatoren $-m_{ij}$ überschrieben $(i > j)$, und die von Null verschiedenen Koeffizienten u_{ij} von U werden auf den a_{ij} zurück gegeben, $a_{ij} = u_{ij}$ für $i \leq j$.

(iii) Falls während der Ausführung der LU-Zerlegung ein Diagonalelement $a_{ii} = 0$ auftritt, muss der Algorithmus abgebrochen werden. Die Fehlermeldung, d. h. der Grund für den Abbruch, kann nicht spezifiziert werden. (Der Grund für den Abbruch könnte die Singularität von A sein, aber nicht in jedem Fall.)

(iv) Die LU-Zerlegung der Matrix A ist von besonderem Vorteil, wenn Lösungen x für mehrere rechte Seiten b des linearen Gleichungssystems $Ax = b$ bestimmt werden müssen.

Lösung eines linearen Gleichungssystems bei gegebener LU-Zerlegung

Gegeben sei die LU-Zerlegung der Koeffizientenmatrix des linearen Gleichungssystems $Ax = b$. Dann gilt

$$\begin{aligned} Ax &= b \\ LUx &= b, \qquad \text{Substitution}: \quad Ux = y \\ Ly &= b. \end{aligned}$$

Daraus folgt unmittelbar, dass bei gegebener LU-Zerlegung ein lineares Gleichungssystem durch folgende Schritte gelöst werden kann:

1. *Vorwärtssubstitution.* Der Vektor y wird als Lösung der Gleichung

$$Ly = b$$

berechnet.

2. *Rückwärtssubstitution.* Die Lösung x des linearen Gleichungssystems wird aus

$$Ux = y$$

erhalten.

Beispiele

(i) Für die Koeffizientenmatrix A des linearen Gleichungssystems

$$\begin{pmatrix} 1 & -1 & 3 \\ 2 & -4 & 7 \\ -1 & -5 & 3 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

sei die LU-Zerlegung gegeben,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & 3 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Zu berechnen ist x.

1. *Vorwärtssubstitution.* Für die Komponenten des Vektors y folgt unmittelbar

$$\begin{aligned} & & y_1 &= 1, \\ 2 + y_2 &= 1, & y_2 &= -1, \\ -1 - 3 + y_3 &= 1, & y_3 &= 5. \end{aligned}$$

2. *Rückwärtssubstitution.* Aus den berechneten Werten für die Komponenten von y erhält man

$$\begin{aligned} & & 3x_3 &= 5, & x_3 &= \frac{5}{3}, \\ -2x_2 + \frac{5}{3} &= -1, & -2x_2 &= -\frac{8}{3}, & x_2 &= \frac{4}{3}, \\ & & x_1 - \frac{4}{3} + 5 &= 1, & x_1 &= -\frac{8}{3}. \end{aligned}$$

Der Vektor

$$x = \frac{1}{3} \begin{pmatrix} -8 \\ 4 \\ 5 \end{pmatrix}$$

ist also Lösung des linearen Gleichungssystems $Ax = b$.

(ii) Zu berechnen ist die Inverse A^{-1} der Matrix

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

bei gegebener LU-Zerlegung

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -2 & 1 \\ 0 & 1 \end{pmatrix}.$$

Für A^{-1} gilt $AA^{-1} = I$, d. h., die erste Spalte von A^{-1} ist Lösung des linearen Gleichungssystems $Ax = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Mit Hilfe der LU-Zerlegung von A erhält man zunächst durch Vorwärtssubstitution den Vektor y aus

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{d. h.} \qquad y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die Rückwärtssubstitution liefert x aus

$$\begin{pmatrix} -2 & 1 \\ 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \text{d. h.} \qquad x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Analog erhält man den zweiten Spaltenvektor von A^{-1} als Lösung des linearen Gleichungssystems $Ax = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$,

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} y &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \text{d. h.} \qquad y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ \begin{pmatrix} -2 & 1 \\ 0 & 1 \end{pmatrix} x &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \text{d. h.} \qquad x = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}. \end{aligned}$$

Damit ist

$$A^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & 1 \end{pmatrix}.$$

Schließlich soll noch darauf hingewiesen werden, dass sich auch die Determinante von A leicht aus einer gegebenen LU-Zerlegung bestimmen lässt. Es gilt

$$\det A = \det(LU) = \det L \cdot \det U = 1 \cdot \det U.$$

Dieses Ergebnis soll noch als Satz formuliert werden.

Satz. Sei A eine reguläre (n, n)-Matrix und existiere eine LU-Zerlegung von A, dann gilt für die Determinante von A

$$\det A = \det U = \prod_{i=1}^{n} u_{ii}.$$

Beispiele

(i) Aus der LU-Zerlegung

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

kann die Determinante von A direkt abgelesen werden, $\det A = -2$.

(ii) Aus

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

folgt unmittelbar $\det A = -6$.

2.3 Das Gauß-Jordan-Verfahren

In der elementaren Gauß-Elimination wurde die Möglichkeit der Vertauschung von Zeilen in einem linearen Gleichungssystem nicht genutzt. Der Vertauschung von Zeilen kann aber von Vorteil oder sogar unverzichtbar sein, wie die folgenden Beispiele zeigen.

Beispiele

(i) Wir machen noch einmal einen Rückgriff auf Beispiel (ii) von Seite 28, vertauschen jetzt aber während der Vorwärtssubstitution der Gauß-Elimination die 2. und 3. Zeile des linearen Gleichungssystems.

$$\begin{array}{ccc|c} \boxed{10} & -7 & 0 & 7 \\ -3 & 2,099 & 6 & 3,901 \\ 5 & -1 & 5 & 6 \end{array} \qquad \begin{array}{l} m_{21} = -\frac{-3}{10} = 0,3, \quad m_{31} = -\frac{5}{10} = -0,5 \\ \mathrm{II} + m_{21} \cdot \mathrm{I} \\ \mathrm{III} + m_{31} \cdot \mathrm{I} \end{array}$$

$$\begin{array}{ccc|c} 10 & -7 & 0 & \\ -3 & 2,099 & 6 & 7 \\ 0 & -0,001 & 6 & 6,001 \\ 0 & \boxed{2,5} & 5 & 2.5 \end{array} \qquad \text{Vertauschen der 2. und 3. Zeile}$$

$$\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ 0 & \boxed{2,5} & 5 & 2,5 \\ 0,001 & 0 & 6 & 6,001 \end{array} \qquad \begin{array}{l} m_{32} = 0,000\,4 \\ \\ \mathrm{III} + m_{32} \cdot \mathrm{II} \end{array}$$

$$\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ 0 & \boxed{2,5} & 5 & 2,5 \\ 0 & 0 & 6,002 & 6,002 \end{array}$$

In diesem Fall wird durch die Rückwärtssubstitution unabhängig von der Rundung (nach der 5. Dezimalstelle) die exakte Lösung

$$x = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

erhalten.

In diesem Beispiel ist der Divisor der im Eliminationsverfahren verwendeten Multiplikatoren durch Einrahmung markiert. Offensichtlich ist es also vorteilhaft, die Zeilen so zu vertauschen, dass der Betrag des Divisors möglichst groß ist.

(ii) Zu lösen ist das (sehr einfache) lineare Gleichungssystem

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

Die elementare Gauß-Elimination bricht sofort ab, obwohl eine eindeutig bestimmte Lösung existiert. Vertauscht man jedoch die beiden Zeilen, erhält man aus

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

mit der Gauß-Elimination problemlos die Lösung $x = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Beispiele dieser Art sind die Begründung für die Anwendung des *Gauß-Jordan-Verfahrens*, das einer elementaren Gauß-Elimination mit *Spaltenpivotisierung* entspricht. Die Spaltenpivotisierung (auch *Spaltenpivotsuche* genannt) wird in der elementaren Gauß-Elimination unmittelbar vor der Ausführung des k-ten Schrittes in der Vorwärtssubstitution auf Seite 29 eingefügt.

Spaltenpivitisierung.

(i) Suche nach dem kleinsten Index ℓ mit

$$|a_{\ell k}| \geq |a_{ik}|, \qquad \text{für} \quad i = k, \ldots, n.$$

Die Zahl $a_{\ell k}$ heißt k-tes *Pivotelement*.

(ii) Vertauschen der k-tem mit der ℓ-ten Zeile von A (wobei A bereits in vorangegangenen Schritten modifiziert sein könnte).

(iii) Vertauschen von b_k mit b_ℓ.

Dazu sollen noch zwei weiter Beispiele gerechnet werden.

Beispiele

(iii) Die Lösung des linearen Gleichungssystems

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

mit Spaltenpivotisierung liefert

$$\begin{array}{cc|c} 1 & 3 & 1 \\ \boxed{2} & 4 & 1 \end{array} \quad \Longrightarrow \quad \begin{array}{cc|c} \boxed{2} & 4 & 1 \\ 1 & 3 & 1 \end{array} \qquad \begin{array}{l} m_{21} = -\frac{1}{2} \\ \mathrm{II} + m_{21} \cdot \mathrm{I} \end{array}$$

$$\begin{array}{cc|c} 2 & 4 & 1 \\ 0 & 1 & \frac{1}{2} \end{array}$$

$$x = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

(iv) Zu lösen ist das lineare Gleichungssystem

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

Die Vorwärtssubstitution mit Spaltenpivotisierung liefert

$$\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ \boxed{3} & 1 & 2 & 0 \\ 2 & 1 & 1 & 1 \end{array} \quad \Longrightarrow \quad \begin{array}{ccc|c} \boxed{3} & 1 & 2 & 0 \\ 1 & 2 & 3 & 1 \\ 2 & 1 & 1 & 1 \end{array} \qquad \begin{array}{l} m_{21} = -\frac{1}{3} \\ m_{31} = -\frac{2}{3} \end{array}$$

$$\begin{array}{ccc|c} 3 & 1 & 2 & 0 \\ 0 & \boxed{\frac{5}{3}} & \frac{7}{3} & 1 \\ 0 & \frac{1}{3} & -\frac{1}{3} & 1 \end{array} \qquad m_{32} = -\frac{1/3}{5/3} = -\frac{1}{5}$$

$$\begin{array}{ccc|c} 3 & 1 & 2 & 0 \\ 0 & \frac{5}{3} & \frac{7}{3} & 1 \\ 0 & 0 & -\frac{4}{5} & \frac{4}{5} \end{array}.$$

Mit Rückwärtssubstitution erhält man $x_3 = -1$, $x_2 = 2$ und $x_1 = 0$, d. h. $x = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}$.

Bemerkungen

(i) Das Gauß-Jordan-Verfahren liefert stets eine Lösung, wenn A regulär ist. Für singuläre Matrizen ist ein definierter Abbruch möglich (Fehlerausschrift „Matrix singulär“).

(ii) Das Gauß-Jordan-Verfahren ist numerisch stabiler als die elementare Gauß-Elimination (bzw. die LU-Zerlegung).

2.4 Die LU-Zerlegung mit Pivotisierung

Die Bedeutung der Pivitisierung soll noch einmal an einem (besonders drastischen) Beispiel verdeutlicht werden.

Beispiele

(i) Dazu modifizieren wir die Koeffizientenmatrix von Beispiel (ii) auf Seite 36 geringfügig,

$$\begin{pmatrix} 10^{-16} & 1 \\ 1 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Die LU-Zerlegung der Koeffizientenmatrix ist

$$\begin{pmatrix} 10^{-16} & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 10^{16} & 1 \end{pmatrix} \cdot \begin{pmatrix} 10^{-16} & 1 \\ 0 & 1 - 10^{16} \end{pmatrix},$$

Vorwärtssubstitution liefert $y = \begin{pmatrix} 1 \\ -10^{16} \end{pmatrix}$, und durch Rückwärtssubstitution erhält aus

$$\begin{pmatrix} 10^{-16} & 1 \\ 0 & 1 - 10^{16} \end{pmatrix} x = \begin{pmatrix} 1 \\ -10^{16} \end{pmatrix}$$

die Lösung

$$x = \begin{pmatrix} -\frac{10^{16}}{10^{16}-1} \\ \frac{10^{16}}{10^{16}-1} \end{pmatrix} \approx \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

(ii) Rundet man im letzten Beispiel jedoch bereits in der LU-Zerlegung,

$$\begin{pmatrix} 10^{-16} & 1 \\ 1 & 1 \end{pmatrix} \approx \begin{pmatrix} 1 & 0 \\ 10^{-16} & 1 \end{pmatrix} \cdot \begin{pmatrix} 10^{-16} & 1 \\ 0 & -10^{16} \end{pmatrix},$$

dann führt die Rückwärtssubstitution zu

$$\tilde{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(iii) Vertauscht man dagegen die beiden Zeilen bereits vor der LU-Zerlegung,

$$\begin{pmatrix} 1 & 1 \\ 10^{-16} & 1 \end{pmatrix} x = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

dann bekommt man

$$\begin{pmatrix} 1 & 1 \\ 10^{-16} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 10^{-16} & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 - 10^{-16} \end{pmatrix}.$$

Aus

$$\begin{pmatrix} 1 & 0 \\ 10^{-16} & 1 \end{pmatrix} y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

folgt $y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, und aus

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 - 10^{-16} \end{pmatrix} x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

erhält man stets die exakte Lösung $x \approx \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, unabhängig davon, ob bereits die Koeffizienten von U gerundet wurden oder nicht.

2.4.1 Permutationsmatrizen

Die Matrixnotation des Gauß-Jordan-Verfahrens erfordert die Einführung von Permutationsmatrizen. Dazu betrachten wir zunächst einen Permutationsvektor

$$\pi = \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_n \end{pmatrix},$$

der Zeilenindizes $1, \ldots, n$, der aus der Spaltenpivotisierung von A während des Gauß-Jordan-Verfahrens erhalten wird. Es gilt $\pi \in \mathbb{N}^n$, und die Koeffizienten von π erfüllen die Bedingung

$$\{\pi_1 \ldots, \pi_n\} = \{1, \ldots, n\}.$$

Der Permutationsvektor π wird im folgenden abkürzend auch *Permutation* genannt.

Definition. Sei π eine (beliebige) Permutation der Zahlen $1, \ldots, n$, dann heißt die (n, n)-Matrix $P = (p_{ik})$ mit

$$p_{ik} = \begin{cases} 1, & \text{falls} \quad k = \pi_i \\ 0, & \text{sonst} \end{cases}$$

die *Permutationsmatrix* zur Permutation π.

Die Permutationsmatrix P ist orthogonal; es gilt $P'P = I$, $PP' = I$ und $P = P^{-1}$. Insbesondere ist

$$P \cdot \begin{pmatrix} 1 \\ \vdots \\ n \end{pmatrix} = \pi \qquad \text{und} \qquad P'\pi = \begin{pmatrix} 1 \\ \vdots \\ n \end{pmatrix}.$$

Beispiele

(i) Sei $\pi = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, d. h. die beiden Zeilen eines linearen Gleichungssystems mit zwei Gleichungen (und zwei Unbekannten) werden vertauscht. Dann ist

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

und z. B. für $A = \begin{pmatrix} 10^{-16} & 1 \\ 1 & 1 \end{pmatrix}$ ist

$$PA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 10^{-16} & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 10^{-16} & 1 \end{pmatrix}.$$

(ii) Für $\pi = (3, 1, 4, 2)'$ ist

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

Sei

$$A = \begin{pmatrix} 4 & 0 & -3 & 1 \\ -2 & -5 & 4 & 1 \\ 2 & -1 & 3 & 3 \\ 6 & -1 & -1 & 6 \end{pmatrix}, \qquad \text{dann ist} \qquad PA = \begin{pmatrix} 2 & -1 & 3 & 3 \\ 4 & 0 & -3 & 1 \\ 6 & -1 & -1 & 6 \\ -2 & -5 & 4 & 1 \end{pmatrix},$$

d. h., in PA sind die Zeilen entsprechend der Permutation π vertauscht. (Die Matrix AP' würde einer Vertauschung der Spalten von A bezüglich π entsprechen.)

2.4.2 Matrixschreibweise des Gauß-Jordan-Verfahrens

Mit Hilfe der Permutationsmatrix P kann die LU-Zerlegung so modifiziert werden, dass sie für beliebige reguläre (n, n)-Matritzen anwendbar ist.

Satz. Sei A eine reguläre (n, n)-Matrix und π die durch das Gauß-Jordan-Verfahren initierte Pernutation. Dann exisitieren eine Permutationsmatrix P bezüglich π, eine untere Dreiecksmatrix L, deren Diagonalelemente gleich Eins sind, $\ell_{ii} = 1$, und eine obere Dreiecksmatrix U, so dass

$$PA = LU. \tag{2.2}$$

Diese Darstellung ist equivalent zu der Faktorisierung $A = P'LU$, d. h. die Koeffizientenmatrix A kann als Produkt der drei Matrizen P', L und U geschrieben werden (*LU-Zerlegung mit Pivotisierung* oder *$P'LU$-Zerlegung*). Damit gilt

$$\begin{aligned} Ax &= b \\ PAx &= Pb \\ LUx &= Pb, \qquad \text{Substitution : } \quad Ux = y \\ Ly &= Pb, \end{aligned}$$

und davon leitet sich unmittelbar die Matrixschreibweise des Gauß-Jordan-Verfahrens ab.

1. *Vorwärtssubstitution.* Der Vektor y wird als Lösung der Gleichung

   $$Ly = Pb$$

   berechnet (wobei Pb der Permutation von b bezüglich π entspricht).

2. *Rückwärtssubstitution.* Die Lösung x des linearen Gleichungssystems wird aus

   $$Ux = y$$

   erhalten.

Bemerkung. Für die Determinante der Matrix P gilt $|\det P| = 1$. Ist k die Anzahl der Permutationen, dann ist $\det P = (-1)^k$. Folglich ist

$$\det A = \det(P'LU) = \det P' \cdot \det L \cdot \det U = \det P \cdot \det U = (-1)^k \cdot \prod_{i=1}^{n} u_{ii}.$$

Beispiel. Wir beizehen uns noch einmal auf Beispiel (ii) von Seite 28, vgl. auch Beispiel (i) von Seite 35. Für die Koeffizientenmatrix A des linearen Gleichungssystems gilt $A = P'LU$, d. h.

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2,099 & 6 \\ 5 & -1 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ -0,3 & -0,000\,4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 10 & -7 & 0 \\ 0 & 2,5 & 5 \\ 0 & 0 & 6,002 \end{pmatrix}.$$

Angenommen, die $P'LU$-Zerlegung von A ist gegeben (d. h. durch ein entsprechendes Programm berechnet), dann erhält man die Lösung des linearen Gleichungssystems durch Vorwärtssubstitution,

$$\begin{aligned} \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ -0,3 & -0,000\,4 & 1 \end{pmatrix} y &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ 3,901 \\ 6 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \\ 3,901 \end{pmatrix} \\ \Longrightarrow \qquad y &= \begin{pmatrix} 7 \\ 2,5 \\ 6,002 \end{pmatrix}, \end{aligned}$$

und durch Rückwärtssubstitution,

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & 2,5 & 5 \\ 0 & 0 & 6,002 \end{pmatrix} x = \begin{pmatrix} 7 \\ 2,5 \\ 6,002 \end{pmatrix} \qquad \Longrightarrow \qquad x = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

2.5 Die Singulärwertzerlegung

2.5.1 Die Methode der kleinsten Quadrate

Ein überbestimmtes lineares Gleichungssystem $Ax = b$ mit mehr Gleichungen als Unbekannten, d. h. mit $m > n$ ist in der Regel wegen $\operatorname{rg} A < \operatorname{rg}(A|b)$ nicht lösbar. Es gilt jedoch

$$\begin{aligned} Ax &= b \\ A'Ax &= A'b \\ x &= (A'A)^{-1}A'b. \end{aligned} \tag{2.3}$$

Die Gültigkeit der letzten Zeile setzt vorraus, dass die Matrix $A'A$ regulär ist. Der Vektor x kann jedoch keine Lösung von $Ax = b$ sein. Zur Bedeutung von x kann aber die folgende Aussage getroffen werden:

Satz. Sei A eine (m, n)-Matrix mit $m \geq n$ und $\operatorname{rg} A = n$, dann ist die Matrix $A'A$ regulär und die quadratische Abweichung

$$|Ax - b|$$

ist für $x = (A'A)^{-1}A'b$ mit $b \in \mathbb{R}^m$ minimal.

Bemerkungen

(i) „Minimal“ bedeutet hier, dass es keinen Vektor $y \in \mathbb{R}^n$ gibt, für den

$$|Ay - b| < |Ax - b|$$

ist. Der Vektor x ist Lösung des linearen Gleichungssystems

$$A'Ax = A'b \qquad \text{(Normalgleichung)}.$$

(ii) Gegeben seien Daten (x_i, y_i), $i = 1, \ldots, m$, wobei die Werte y_i als (fehlerbehaftete) Messwerte an den Stellen x_i interpretiert werden. An die Daten soll eine Gerade $y = ax + b$ angepasst werden. Ist $m > 2$, dann ist mit Abweichungen $\varepsilon_i = ax_i + b - y_i$ der y_i von der Ausgleichsgeraden zu rechnen. Falls eine linearer Zusammenhang zwischen den x- und y-Werten besteht, können die ε_i als Messfehler der y_i interpretiert werden. Der Zuammenhang zwischen den Daten kann in der Form

$$\begin{aligned} ax_1 + b &= y_1 + \varepsilon_1 \\ &\vdots \\ ax_m + b &= y_m + \varepsilon_n \end{aligned}, \qquad \text{d. h.} \qquad \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{pmatrix}$$

geschrieben werden. Die Bestimmung des Koeffizientenvektors $\begin{pmatrix} a \\ b \end{pmatrix}$ als Lösung der Normalgleichung

$$\begin{pmatrix} a \\ b \end{pmatrix} = \left[\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix}' \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \right]^{-1} \cdot \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix}' \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

minimiert die Fehlerquadratsumme $\sum_{i=1}^m \varepsilon_i^2$ und heißt *Methode der kleinsten Quadrate* (MKQ).[1]

Beispiele

(i) Für das lineare Gleichungssystem

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

ist mit der Methode der kleinsten Quadrate der Vektor x zu bestimmen. Für die Koeffizientenmatriy A und die rechte Seite b gilt

$$A'A = \begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix}, \quad (A'A)^{-1} = \frac{1}{20} \begin{pmatrix} 4 & -6 \\ -6 & 14 \end{pmatrix} = \begin{pmatrix} 0,2 & -0,3 \\ -0,3 & 0,7 \end{pmatrix}, \quad A'b = \begin{pmatrix} 15 \\ 8 \end{pmatrix},$$

und damit ist

$$x = (A'A)^{-1} A'b = \begin{pmatrix} 0,2 & -0,3 \\ -0,3 & 0,7 \end{pmatrix} \cdot \begin{pmatrix} 15 \\ 8 \end{pmatrix} = \begin{pmatrix} 0,6 \\ 1,1 \end{pmatrix}.$$

Es ist

$$Ax = \begin{pmatrix} 1,1 \\ 1,7 \\ 2,3 \\ 2,9 \end{pmatrix},$$

und daraus folgt

$$|Ax - b| = \sqrt{0,01 + 0,09 + 0,09 + 0.01} = \sqrt{0,2}.$$

Es gibt keinen Vektor $y \in \mathbb{R}^2$ mit $|Ay - b| < \sqrt{0,2}$.

[1] Hier ist die MKQ nur für den einfachsten Fall behandelt. In der Literatur finden sich allgemeinere Fassungen der MKQ.

(ii) Analog erhält man aus

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

mit Hilfe der Methode der kleinsten Quadrate

$$A'A = \begin{pmatrix} 96 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{pmatrix}, \quad (A'A)^{-1} = \frac{1}{20} \begin{pmatrix} 5 & -15 & 5 \\ -15 & 49 & -21 \\ 5 & -21 & 19 \end{pmatrix}, \quad A'b = \begin{pmatrix} 19 \\ 9 \\ 6 \end{pmatrix},$$

die Lösung der Normalgleichung

$$x = \begin{pmatrix} -\frac{1}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix}$$

In diesem Beispiel ist $|Ax - b| = 0$, d. h., x ist nicht nur Lösung der Normalgleichung sondern wegen $\operatorname{rg} A = \operatorname{rg}(A|b)$ gleichzeitig Lösung von $Ax = b$.

2.5.2 Allgemeine Lösung der Normalgleichung

Um die Normalgleichung zu lösen, wird im vorangegangenen Abschnitt gefordert, dass $\operatorname{rg} A = n$. Das ist aber nicht immer der Fall. Ausserdem ist die Lösung der Normalgleichung durch (2.3) oft numerisch instabil. Der numerischen Instabilität könnte mit der LU-Zerlegung mit Pivotisierung begegnet werden. Wir führen in diesem Abschnitt jedoch ein Verfahren ein, das sich auf alle Formen von linearen Gleichungssystemen übertragen lässt. Dieses Verfahren basiert wie die LU-Zerlegung auf einer Faktorisierung der Koeffitientenmatrix A.

Satz. Sei A eine (m, n)-Matrix mit $m \geq n$, dann existieren eine orthogonale (m, n)-Matrix U, eine orthogonale (n, n)-Matrix V sowie eine (n, n)-Diagonalmatrix W, so dass

$$A = UWV' \qquad \text{(Singulärwertzerlegung).} \tag{2.4}$$

Abkürzend wird für „Singulärwertzerlegung“ auch „SVD“ geschrieben (*singular value decomposition*).

Berechnung von U, V, W

(i) Berechnung der Eigenwerte $\lambda_1, \ldots, \lambda_n$ der Matrix $A'A$, die nach ihrer Größe geordnet sind

$$\lambda_1, \leq \ldots \leq \lambda_n.$$

Da die Matrix $A'A$ positiv semidefinit ist, sind alle ihre Eigenwerte größer gleich Null. Die Wurzeln der Eigenwerte von $A'A$ bilden die Hauptdiagonale von W,

$$W = \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}) = \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}.$$

Es wird angemerkt, dass die ersten n Eigenwerte der Matrix AA' identisch mit den Eigenwerten von $A'A$ sind. Die Eigenwerte $\lambda_{n+1}, \ldots, \lambda_m$ sind gleich Null.

(ii) Berechnung der normierten Eigenvektoren $u_1, \ldots, u_n$ von AA' zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Da AA' symmetrisch ist, sind die Eigenvektoren orthogonal. Die Eigenvektoen von AA' bilden die Salten der Matrix U,

$$U = (u_1, \ldots, u_n).$$

(iii) Berechnung der normierten Eigenvektoren $v_1, \ldots, v_n$ von $A'A$ zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Die Matrix $A'A$ ist ebenfalls symmetrisch. Folglich sind auch ihre Eigenvektoren orthogonal. Die Matrix V wird aus den Vektoren $v_1, \ldots, v_n$ gebildet,

$$V = (v_1, \ldots, v_n).$$

Bemerkung. Die Singärwertzerlegung kann für beliebige Rechteckmatrizen eingeführt werden, also auch für den Fall $m < n$. Ist $m < n$, dann wird das oben beschriebene Verfahren auf die Matrix A' angewendet, d. h. wir erhalten die Faktorisierung $A' = UWV'$, und daraus folgt unmittelbar $A = VWU'$. Die SVD einer Matrix A mit $m < n$ hat also die gleiche Form wie für eine Matrix mit $m \geq n$. Damit kann auch die Lösung linearer Gleichungssysteme mit weniger Gleichungen als Unbekannten auf der Grundlage der SVD von A erhalten werden.

Definition. Sei A eine (m, n)-Matrix mit $m \geq n$. Die Wurzeln $w_1, \ldots, w_n$ der Eigenwerte von $A'A$ heißen *Singulärwerte.*

Der folgende Satz ist die Grundlage der Lösung eines linearen Gleichungssystems mit Hilfe der SVD.

Satz. Sei $A = UWV'$ die SVD einer (m, n)-Matrix A mit $r = \operatorname{rg} A \leq \min\{m, n\}$. Dann ist

$$x = \sum_{i=1}^{r} \frac{u_i'b}{w_i} v_i \tag{2.5}$$

die eindeutig bestimmte Lösung der Normalgleichung $A'Ax = A'b$.

Bemerkungen

(i) Wenn die Lösung des linearen Gleichungssystems $Ax = b$ eindeutig bestimmt ist, d. h. $\operatorname{rg} A = \operatorname{rg}(A|b) = n$, dann ist (2.5) identisch mit dieser Lösung.

(ii) Hat das lineare Gleichungssystem $Ax = b$ keine Lösung, d. h. $\operatorname{rg} A < \operatorname{rg}(A|b)$, dann ist für den mit (2.5) berechneten Vektor x der Abstand $|Ax - b|$ minimal.

(iii) Wenn die Lösung von $Ax = b$ nicht eindeutig bestimmt ist, d. h. $\operatorname{rg} A = \operatorname{rg}(A|b) < n$, dann ist x aus (2.5) diejenige Lösung mit der kleinsten Norm. Das beteutet $|x| \leq |y|$ für alle $y \in \mathbb{R}^n$ mit $Ay = b$.

Beispiele

(i) Wir betrachten zur Demonstration der Vorgehensweise zunächst ein möglichst einfaches Beispiel und machen eine SVD der Matrix

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

Es gilt

$$\begin{aligned} A'A &= \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \\ \det(A'A - \lambda I) &= \det \begin{pmatrix} 5-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\ &= (5-\lambda)(1-\lambda) - 4 \\ &= \lambda^2 - 6\lambda + 1 = 0 \\ \lambda_{1/2} &= 3 \pm \sqrt{9-1}, \qquad \lambda_1 = 3 + 2\sqrt{2}, \quad \lambda_2 = 3 - 2\sqrt{2}. \end{aligned}$$

Damit ist

$$W = \begin{pmatrix} 1+\sqrt{2} & 0 \\ 0 & -1+\sqrt{2} \end{pmatrix}.$$

Die (normierten) Eigenvektoren von $A'A$ sind Lösungen von

$$\begin{pmatrix} 2-2\sqrt{2} & 2 \\ 2 & -2-2\sqrt{2} \end{pmatrix} x_1 = 0, \qquad \text{d. h.} \quad v_1 = \frac{1}{\sqrt{4+2\sqrt{2}}} \begin{pmatrix} 1+\sqrt{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0,924 \\ 0,383 \end{pmatrix}$$

bzw.

$$\begin{pmatrix} 2+2\sqrt{2} & 2 \\ 2 & -2+2\sqrt{2} \end{pmatrix} x_2 = 0, \qquad \text{d. h.} \quad v_2 = \frac{1}{\sqrt{4-2\sqrt{2}}} \begin{pmatrix} \sqrt{2}-1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0,383 \\ -0,924 \end{pmatrix},$$

und hieraus folgt

$$V = \begin{pmatrix} 0,924 & 0,383 \\ 0,383 & -0,924 \end{pmatrix}.$$

Analog erhält man aus

$$AA' = \begin{pmatrix} 5 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

die Eigenwerte $\lambda_1 = 3+2\sqrt{2}$, $\lambda_2 = 3-2\sqrt{2}$ und $\lambda_3 = 0$ sowie die dazu gehörigen Eigenvektoren

$$u_1 = \begin{pmatrix} 0,924 \\ 0,383 \\ 0 \end{pmatrix}, \qquad u_2 = \begin{pmatrix} -0,383 \\ 0,924 \\ 0 \end{pmatrix}, \qquad u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Die ersten beiden Eigenwerte von AA' sind also identisch mit denen von $A'A$; der dritte Eigenvektor wird in der SVD nicht verwendet. Somit gilt

$$U = \begin{pmatrix} 0,924 & -0,383 \\ 0,383 & 0,924 \\ 0 & 0 \end{pmatrix}$$

Die SVD von A hat damit die Form

$$\begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0,924 & -0,383 \\ 0,383 & 0,924 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2,414 & 0 \\ 0 & 0,414 \end{pmatrix} \cdot \begin{pmatrix} 0,924 & 0,383 \\ 0,383 & -0,924. \end{pmatrix}.$$

(ii) Zu lösen ist das lineare Gleichungssystem $Ax = b$ mit

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \qquad \text{und} \qquad b = \begin{pmatrix} 2 \\ 3 \\ 3 \\ 4 \end{pmatrix}.$$

Es gilt

$$\begin{aligned} A'A &= \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \\ \det(A'A - \lambda I) &= \det \begin{pmatrix} 4-\lambda & 6 \\ 6 & 14-\lambda \end{pmatrix} \\ &= (4-\lambda)(14-\lambda) - 36 \\ &= \lambda^2 - 18\lambda + 20 = 0 \\ \lambda_{1/2} &= 9 \pm \sqrt{81-20}, \qquad \lambda_1 = 16,810, \quad \lambda_2 = 1,190. \end{aligned}$$

Damit haben wir $w_1 = 4,100$, $w_2 = 1,091$ und

$$W = \begin{pmatrix} 4,100 & 0 \\ 0 & 1,091 \end{pmatrix}.$$

Die zu λ_1 bzw. λ_2 gehörigen Eigenvektoren v_1 bzw. v_2 erhält man wie folgt:

$$v_1 = \begin{pmatrix} s \\ t \end{pmatrix} \quad \text{ist Lösung von} \quad \begin{aligned} (A'A - \lambda_1 I)v_1 &= 0 \\ (4 - 16,810)s + 6t &= 0 \\ v_1 = \frac{1}{2,358}\begin{pmatrix} 1 \\ 2,135 \end{pmatrix} &= \begin{pmatrix} 0,424 \\ 0,906 \end{pmatrix}, \end{aligned}$$

und

$$v_2 = \begin{pmatrix} s \\ t \end{pmatrix} \quad \text{ist Lösung von} \quad \begin{aligned} (A'A - \lambda_1 I)v_1 &= 0 \\ (4 - 1{,}190)s + 6t &= 0 \\ v_2 = \frac{1}{1{,}104}\begin{pmatrix} -1 \\ 0{,}468 \end{pmatrix} &= \begin{pmatrix} -0{,}906 \\ 0{,}424 \end{pmatrix}. \end{aligned}$$

Das ergibt die Matrix $V = \begin{pmatrix} 0{,}424 & -0{,}906 \\ 0{,}906 & 0{,}424 \end{pmatrix}$.

Die Eigenwerte λ_1 und λ_2 der Matrix $A'A$ sind auch Eigenwerte der Matrix

$$AA' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & 7 \\ 1 & 4 & 7 & 10 \end{pmatrix}.$$

(Weitere Eigenwerte von AA' sind $\lambda_{3/4} = 0$, die jedoch zur SVD von A nicht benötigt werden.) Die beiden Eigenvektoren u_1 und u_2 kann man zum Beispiel mit Hilfe des Gauß-Verfahrens als Lösungen der beiden linearen Gleichungssysteme

$$(AA' - \lambda_1)u_1 = 0, \qquad (AA' - \lambda_2)u_2 = 0$$

erhalten,

$$u_1 = \begin{pmatrix} 0{,}103 \\ 0{,}324 \\ 0{,}545 \\ 0{,}766 \end{pmatrix}, \qquad u_2 = \begin{pmatrix} -0{,}830 \\ -0{,}441 \\ -0{,}051 \\ 0{,}336 \end{pmatrix}. \qquad \text{Daraus folgt} \quad U = \begin{pmatrix} 0{,}103 & -0{,}830 \\ 0{,}324 & -0{,}441 \\ 0{,}545 & -0{,}052 \\ 0{,}766 & 0{,}336 \end{pmatrix}.$$

Die SVD der Matix A hat damit die Form

$$A = \begin{pmatrix} 0{,}103 & -0{,}830 \\ 0{,}324 & -0{,}441 \\ 0{,}545 & -0{,}052 \\ 0{,}766 & 0{,}336 \end{pmatrix} \cdot \begin{pmatrix} 4{,}100 & 0 \\ 0 & 1{,}091 \end{pmatrix} \cdot \begin{pmatrix} 0{,}424 & 0{,}906 \\ -0{,}906 & 0{,}424 \end{pmatrix}.$$

Damit erhalten wir den Vektor x mit Hilfe von Gleichung (2.5) aus

$$\begin{aligned} x &= \frac{u_1'b}{w_1}v_1 + \frac{u_2'b}{w_2}v_2 \qquad \text{mit} \qquad u_1'b = 5{,}880, \quad U_2'b = -1{,}797 \\ &= 5{,}880 \cdot \begin{pmatrix} 0{,}424 \\ 0{,}906 \end{pmatrix} - 1{,}797 \cdot \begin{pmatrix} 0{,}424 \\ 0{,}906 \end{pmatrix} \\ &= \begin{pmatrix} 2{,}1 \\ 0{,}6 \end{pmatrix}, \end{aligned}$$

vgl. auch mit Beispiel (i) auf Seite 43.

2.5.3 Die Matrixnorm und die Konditionszahl

Wir wollen nun noch einige Begriffe einführen, die sich direkt aus der SVD einer Matrix A ableiten.

Zunächst ist klar, dass der Rang einer Matrix A gleich der Anzahl ihrer von Null verschiedenen Singulärwerte ist. Die Determinante $\det A$ ist durch die Singlärwerte bis auf ihr Vorzeichen bestimmt:

Satz. Sei A eine (n,n)-Matrix A mit den Singulärwerten $w_1,\dots,w_n$, dann gilt

$$|\det A| = \prod_{i=1}^{n} w_i.$$

Nun führen wir die zur Euklidischen Vektornorm passende Matrixnorm ein.

Definition. Die zur Euklidischen Vektornorm gehörende *Norm* $\|A\|$ einer quadratischen Matrix ist definiert als ihr maximaler Singulärwert,

$$\|A\| = w_1$$

(*Spektralnorm*).

Die Norm einer Matrix ist die Grundlage für die Charakterisierung ihrer Kondition und damit für die Fehlerabschätzung bei der Lösung linearer Gleichungssysteme.

Definition. Die *Kondition* $\operatorname{cond} A$ einer (n,n)-Matrix ist definiert durch

$$\operatorname{cond} A = \begin{cases} \|A\|\cdot\|A^{-1}\|, & \text{falls } A \text{ regulär} \\ \infty, & \text{falls } A \text{ singulär} \end{cases}.$$

Bemerkungen

(i) Ist A eine reguläre Matrix mit den Singulärwerten $w_1,\dots,w_n$, dann hat ihre Inverse A^{-1} die Singulärwerte $1/w_n,\dots,1/w_1$. Daraus folgr für reguläre Matrizen unmittelbar

$$\operatorname{cond} A = w_1/w_n.$$

(ii) Eine Matrix heißt *schlecht konditioniert*, wenn ihre Kondition zu groß ist, d. h. $\operatorname{cond} A > 10^6$ bei der Verendung des Datentyps `float` für die Koeffizienten von A und $\operatorname{cond} A > 10^{12}$ für `double`.

Satz. Sei Δb der Fehler der rechten Seite b des linearen Gleichungssystems $Ax = b$. Dann gilt für den sich daraus ergebenden Fehler Δx der Lösung x die Abschätzung

$$|\Delta x| \le \operatorname{cond} A \cdot |\Delta b|.$$

Beispiele

(i) Zu berechnen sind die Spektralnorm und die Kondition der Matrix

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

Es gilt

$$\begin{aligned}
A'A &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\
\det(A'A - \lambda I) &= (1-\lambda)(2-\lambda) - 1 \\
&= \lambda^2 - 3\lambda + 1 = 0 \\
\lambda_{1/2} &= \frac{3}{2} \pm \sqrt{\frac{9}{4} - 1} = \frac{3}{2} \pm \frac{\sqrt{5}}{2}, \qquad \lambda_1 = 2{,}618, \quad \lambda_2 = 0{,}382.
\end{aligned}$$

Daraus folgt $w_1 = 1{,}618$, $w_2 = 0{,}618$ und $\operatorname{cond} A = \|A\| \cdot \|A^{-1}\| = w_1/w_2 = 2{,}618$.

(ii) Wir untersuchen nun ein lineares Gleichungssystem $Ax = b$ mit der Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 1{,}001 \\ 1 & 1 \end{pmatrix}.$$

Für die rechte Seite $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ hat das Gleichungssystem die Lösung $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Bereits eine geringe Änderung in der rechten Seite, z. B. $\tilde{b} = \begin{pmatrix} 0{,}9 \\ 1{,}001 \end{pmatrix}$ liefert eine völlig andere Lösung $\tilde{x} = \begin{pmatrix} 111{,}01 \\ -110 \end{pmatrix}$. Um diesen Effekt zu erklären berechnen wir die Kondition der Matrix A. Es gilt

$$\begin{aligned}
A'A &= \begin{pmatrix} 2 & 2{,}001 \\ 2{,}001 & 2{,}002 \end{pmatrix} \\
\det(A'A - \lambda I) &= \lambda^2 - 4{,}002\lambda + 0{,}000\,001 = 0 \\
\lambda_{1/2} &= 2{,}001 \pm \sqrt{4{,}004\,001 - 0{,}000\,001} = 2{,}001 \pm \sqrt{4{,}004} \\
&= 2{,}001 \pm 2{,}000\,999\,75 \\
\lambda_1 &= 4{,}002, \qquad \lambda_2 = 2{,}5 \cdot 10^{-7} \\
w_1 &= 2{,}000\,5 \qquad w_2 = 0{,}000\,5 \\
\operatorname{cond} A &= w1/w2 = 4\,001.
\end{aligned}$$

Bei einer so großen Kondition der Koeffizientenmatrix sind große Abweichungen in der Lösung zu erwarten,

$$\begin{aligned}
|\Delta x| &\leq \operatorname{cond} A \cdot |\Delta b| \\
&= 4\,001 \cdot \sqrt{0{,}1^2 + 0{,}01^2} \\
&= 402{,}1.
\end{aligned}$$

Index