Kapitel 12

Lineare Algebra

12.1 Vektorräume

↓20.1.04

Definition 12.1: (Vektorräume)

Ein **Vektorraum** V *über dem* **Skalarenkörper** K *ist eine nichtleere Menge von Objekten (***Vektoren** *genannt). Es muss eine Addition*

$$v_1, v_2 \in V \;\Rightarrow\; v_1 + v_2 \in V$$

mit den Eigenschaften

$$v_1 + v_2 = v_2 + v_1 \quad (\textit{Kommutativität}),$$

$$(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3) \quad (\textit{Assoziativität})$$

für alle $v_i \in V$ *gegeben sein, es muss ein neutrales Element* 0 *(der Nullvektor) mit* $v + 0 = v \ \forall v \in V$ *existieren, zu jedem* $v \in V$ *muss es einen negativen Vektor* $-v$ *mit* $v + (-v) = 0$ *geben.*

Weiterhin muss jedes $v \in V$ *mit jedem „Skalar"* $\lambda \in K$ *multiplizierbar sein, so dass* $\lambda \cdot v$ *wieder in* V *liegt und*

$$\lambda_1 \cdot (\lambda_2 \cdot v) = (\lambda_1 \cdot \lambda_2) \cdot v \quad (\textit{Assoziativität}),$$

$$1 \cdot v = v \quad (\textit{neutrales Element der Multiplikation}),$$

$$(\lambda_1 + \lambda_2) \cdot v = \lambda_1 \cdot v + \lambda_2 \cdot v \quad (\textit{Distributivität})$$

für alle λ_1, $\lambda_2 \in K$ *und* $v \in V$ *gilt. Hierbei sei* 1 *das neutrale Element der Multiplikation im Körper* K.

Bemerkung 12.2: *Als Skalarenkörper* K *interessiert uns hier nur* $\mathbb{R}$ *oder* $\mathbb{C}$*, wir sprechen dann von einem* **„reellen"** *oder einem* **„komplexen"** *Vektorraum.*

Beispiel 12.3: 1) Der Raum[1]

$$\mathbb{R}^n := \Big\{(x_1, x_2, \ldots, x_n)^T \;;\; x_1, \ldots x_n \in \mathbb{R}\Big\}$$

bzw.

$$\mathbb{C}^n := \Big\{(x_1, x_2, \ldots, x_n)^T \;;\; x_1, \ldots x_n \in \mathbb{C}\Big\}$$

mit den Operationen

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

und

$$\lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$$

ist ein reeller bzw. komplexer Vektorraum.

2) Die Menge

$$\left\{\lambda_1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} ;\; \lambda_1, \lambda_2 \in \mathbb{R}\right\} \subset \mathbb{R}^3$$

mit den Operationen des $\mathbb{R}^3$ (Addition und skalare Multiplikation) ist ein in den $\mathbb{R}^3$ eingebetteter Vektorraum. Man nennt diese auch einen **„Unterraum“** des $\mathbb{R}^3$. Es handelt sich um eine den Ursprung enthaltende Ebene im $\mathbb{R}^3$. Man beachte, dass die Summe zweier Ebenenvektoren wieder in der Ebene liegt.

3) Der Raum der stetigen Funktionen

$$C([0,1]) = \{f : [0,1] \to \mathbb{R};\; f \text{ ist eine stetige Funktion}\}$$

mit den Operationen der punktweisen Addition

$$f + g :\; x \to f(x) + g(x)$$

und der Multiplikation

$$\lambda \cdot f :\; x \to \lambda \cdot f(x)$$

[1] Vektoren des $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ sind bei uns immer **Spalten**. Aus Platzgründen verwendet man im Text lieber eine Zeilendarstellung $(x_1, \ldots, x_n)$. Damit die Bezeichnungen konsistent sind, wird das Symbol T (Transposition)

$$(x_1, \ldots, x_n)^T := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

benutzt, das anzeigt, dass dieser Vektor als Spalte zu interpretieren ist.

ist ein reeller Vektorraum (man beachte, dass die Summe stetiger Funktionen wieder stetig ist und dass mit f auch $\lambda \cdot f$ stetig ist).

Notation 12.4:

Für allgemeine Vektorräume V benutzen wir die Notation $v \in V$ ohne Vektorpfeil über dem Vektor v. Für $V = \mathbb{R}^n$ bzw. $V = \mathbb{C}^n$ schreiben wir Vektoren mit einem Vektorpfeil (Physikernotation), also $\vec{v} \in \mathbb{R}^n$ bzw. $\vec{v} \in \mathbb{C}^n$.

Definition 12.5: (lineare Unabhängigkeit)

Eine Menge von Vektoren $v_1, \ldots, v_n$ eines Vektorraums heißt **„linear unabhängig“**, *wenn eine Linearkombination $\sum_{i=1}^n \lambda_i \cdot v_i$ nur dann der Nullvektor sein kann, wenn alle „Linearkoeffizienten“ λ_i verschwinden:*

$$\sum_{i=1}^{n} \lambda_i \cdot v_i = 0 \quad \Rightarrow \quad \lambda_1 = \ldots = \lambda_n = 0.$$

Sind Vektoren nicht linear unabhängig, nennt man sie **„linear abhängig“**.

Bemerkung 12.6: *In einer Menge linear unabhängiger Vektoren $\{v_1, \ldots, v_n\}$ kann keiner der Vektoren 0 sein. Angenommen, $v_1 = 0$. Dann wäre*

$$\lambda_1 \cdot v_1 + \ldots + \lambda_n \cdot v_n = 0$$

für beliebiges λ_1 und $\lambda_2 = \ldots = \lambda_n = 0$ erfüllt.

Beispiel 12.7: Die Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sind linear unabhängige Vektoren im $\mathbb{R}^3$, denn aus

$$\sum_{i=1}^{3} \lambda_1 \cdot \vec{v}_i = \lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \overset{(!)}{=} \vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

folgt sofort $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Die Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

hingegen sind nicht linear unabhängig, denn z.B. die Linearkombination

$$\vec{v}_1 + \vec{v}_2 - \vec{v}_3 = \vec{0}$$

verschwindet, obwohl die Linearfaktoren $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = -1$ nicht verschwinden.

Definition 12.8: (von Vektoren aufgespannte Räume)

Seien $v_1, \ldots, v_n$ Vektoren eines Vektorraums über dem Skalarenkörper K. Die Menge der Vektoren

$$\mathrm{span}(v_1, \ldots, v_n) = \left\{ \sum_{i=1}^{n} \lambda_i \cdot v_i \; ; \; \lambda_1, \ldots, \lambda_n \in K \right\}$$

nennt man **„den von den Vektoren** $v_1, \ldots, v_n$ **aufgespannten (Unter-)Raum“**. *Bezeichnung: die Vektoren $v_1, \ldots, v_n$ bilden ein* **„Erzeugendensystem“** *des (Unter-)Raums.*

Definition 12.9: (Basis eines Vektorraums)

Eine Menge $\{v_1, \ldots, v_n\}$ von Vektoren eines Vektorraums V heißt **„Basis“**, *wenn die Vektoren linear unabhängig sind und $V = \mathrm{span}(v_1, \ldots, v_n)$ gilt.*

Satz und Definition 12.10:

Besitzt ein Vektorraum eine Basis mit endlich vielen Elementen, so hat jede alternative Basis die selbe Anzahl von Elementen. Diese Anzahl von Basiselementen wird die **„Dimension“** *des Vektorraums genannt.*

Beweis: ist technisch (z.B. über den „Steinitzschen Austauschsatz“). Siehe Bücher über Lineare Algebra.

Satz 12.11:

a) *Seien $\{v_1, \ldots, v_m\}$ linear unabhängig. Für jeden Vektor v, der sich nicht als Linearkombination von $v_1, \ldots, v_m$ schreiben läßt, sind die Vektoren $\{v_1, \ldots, v_m, v\}$ linear unabhängig.*

b) *In einem Raum der Dimension n bildet jede Menge mit n linear unabhängigen Vektoren eine Basis.*

c) *In einem Raum der Dimension n gibt es maximal n linear unabhängige Vektoren.*

Beweis: (für technisch Interessierte)
a) Sei $\{v_1, \ldots, v_m\}$ eine beliebige Menge linear unabhängiger Vektoren. Sei v ein Vektor, der nicht Linearkombination von $v_1, \ldots, v_m$ ist, d.h., für jede Wahl von skalaren Linearfaktoren $\mu_1, \ldots, \mu_m$ gilt

$$v - \sum_{i=1}^{m} \mu_i \cdot v_i \neq 0. \qquad (\#)$$

Betrachte nun die Vektorgleichung

$$\lambda_0 \cdot v + \sum_{i=1}^{m} \lambda_i \cdot v_i = 0. \qquad (\#\#)$$

Es folgt, dass $\lambda_0 = 0$ gelten muss, denn sonst würde

$$v - \sum_{i=1}^{m} \Big(- \frac{\lambda_i}{\lambda_0} \Big) \cdot v_i = 0$$

folgen, was im Widerspruch zu (#) steht. Damit wird aus (##) die Gleichung $\sum_{i=1}^{m} \lambda_i \cdot v_i = 0$, woraus wegen der linearen Unabhängigkeit von $v_1, \ldots, v_m$ folgt, dass $\lambda_1 = \ldots = \lambda_m = 0$ gilt. Also impliziert (##) insgesamt $\lambda_0 = \lambda_1 = \ldots = \lambda_m = 0$, d.h., die $m+1$ Vektoren $v_1, \ldots, v_m, v$ sind linear unabhängig.

b) Seien $v_1, \ldots, v_n$ linear unabhängige Vektoren. Sei $\{e_1, \ldots, e_n\}$ irgendeine Basis. Seien $e_{j_1}, \ldots, e_{j_m}$ die Basisvektoren, die sich nicht als Linearkombination von $v_1, \ldots, v_n$ schreiben lässen. Nach a) sind $\{v_1, \ldots, v_n, e_{j_1}, \ldots, e_{j_m}\}$ linear unabhängig. Jeder Vektor v des Vektorraums kann als Linearkombination der Basiselemente e_i geschrieben werden. Jedes e_i läßt sich aber nach Konstruktion als Linearkombination von $v_1, \ldots, v_n, e_{j_1}, \ldots, e_{j_m}$ ausdrücken, so dass sich insgesamt jedes v als Linearkombination der Vektoren $v_1, \ldots, v_n, e_{j_1}, \ldots, e_{j_m}$ darstellen läßt. Damit ist $\mathrm{span}(v_1, \ldots, v_n, e_{j_1}, \ldots, e_{j_m})$ der gesamte Vektorraum, also ist $\{v_1, \ldots, v_n, e_{j_1}, \ldots, e_{j_m}\}$ eine Basis mit $n+m$ Elementen. Dies darf nach Satz 12.10 aber nur für $m = 0$ sein, d.h., die n linear unabhängigen Vektoren $v_1, \ldots, v_n$ bildeten schon eine Basis.

c) Von m linear unabhängigen Vektoren (mit $m > n$) bilden die ersten n Vektoren nach b) schon eine Basis, der $(n+1)$-te Vektor kann damit als Linearkombination der ersten n geschrieben werden und ist damit nicht linear unabhängig von den ersten n.

Q.E.D.

Beispiel 12.12: 1) Der reelle Raum

$$\mathbb{R}^n = \Big\{ (x_1, x_2, \ldots, x_n)^T \ ; \ x_1, \ldots, x_n \in \mathbb{R} \Big\}$$

und der komplexe Raum

$$\mathbb{C}^n = \Big\{ (x_1, x_2, \ldots, x_n)^T \ ; \ x_1, \ldots, x_n \in \mathbb{C} \Big\}$$

haben jeweils die Dimension n. Die **„Standardbasis“** besteht jeweils aus den Vektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots \ , \ \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Beachte dazu

$$\sum_{i=1}^{n} x_i \cdot \vec{e}_i = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Jede Menge von n linear unabhängigen Vektoren bildet jeweils eine alternative Basis.

2) Seien $\vec{v}_1$, $\vec{v}_2$ zwei linear unabhängige Vektoren im $\mathbb{R}^n$ (mit $n \geq 2$). Sie bilden eine Basis der 2-dimensionalen Ebene

$$E = \mathrm{span}(\vec{v}_1, \vec{v}_2) = \Big\{ \lambda_1 \cdot \vec{v}_1 + \lambda_2 \cdot \vec{v}_2 \ ; \ \lambda_1, \lambda_2 \in \mathbb{R} \Big\}.$$

Eine alternative Basis des Vektorraums E (Unterraum des $\mathbb{R}^n$) ist z.B.

$$\vec{v}_1' = \vec{v}_1 + \vec{v}_2, \ \vec{v}_2' = \vec{v}_1 - \vec{v}_2,$$

denn $\vec{v}_1'$, $\vec{v}_2'$ sind wieder linear unabhängig und jeder von $\vec{v}_1$, $\vec{v}_2$ aufgespannte Vektor läßt sich auch von $\vec{v}_1'$, $\vec{v}_2'$ aufspannen:

$$\vec{v} = \lambda_1 \cdot \vec{v}_1 + \lambda_2 \cdot \vec{v}_2 = \Big(\frac{\lambda_1}{2} + \frac{\lambda_2}{2} \Big) \cdot \vec{v}_1' + \Big(\frac{\lambda_1}{2} - \frac{\lambda_2}{2} \Big) \cdot \vec{v}_2'.$$

3) Sei V die Menge aller Polynome vom Grad $\leq n$. In Blatt 13 ist zu zeigen, dass die Monome $\{1, x, x^2, \ldots, x^n\}$ eine Basis dieses $n+1$ dimensionalen Vektorraums sind.

4) Sei V die Menge aller Polynome (mit beliebigem Grad). Die Dimension dieses Raums ist ∞, denn es gibt beliebig viele linear unabhängige Elemente (z.B. die Monome x^n mit beliebig großem n).

Für analytische Betrachtungen auf dem $\mathbb{R}^n$ bzw. dem $\mathbb{C}^n$ braucht man einen Längenbegriff, mit dem man den Abstand $\|\vec{v} - \vec{w}\|$ zwischen Vektoren messen kann:

Definition 12.13: (Norm)

Sei V ein Vektorraum über dem Skalarenkörper K. Eine **„Norm"** *auf V ist eine Abbildung $\| . \| : V \to [0, \infty)$ mit den Eigenschaften:*

a) $\|v + w\| \leq \|v\| + \|w\|$ *für alle $v, w \in V$ (***„Dreiecksungleichung"***).*

b) $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$ *für alle $v \in V$, $\lambda \in K$ (***„Homogenität"***).*

c) $\|v\| = 0 \;\Leftrightarrow\; v = 0$.

Beispiel 12.14: Auf dem $\mathbb{R}^n$ bzw. dem $\mathbb{C}^n$ existieren die sogenannten p-Normen:

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\|_p = \Big(\sum_{i=1}^{n} |x_i|^p \Big)^{1/p}, \quad p \in [1, \infty).$$

Die Norm-Eigenschaften b) und c) sind offensichtlich, der Beweis der Dreiecksungleichung

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

ist jedoch nicht-trivial. Zusätzlich kann man die Norm

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\|_\infty = \max_{i=1,\ldots,n} |x_i|$$

betrachten. Die Namensgebung entstammt der Tatsache, dass

$$\|\vec{x}\|_\infty = \lim_{p \to \infty} \|\vec{x}\|_p$$

für jedes $\vec{x} \in \mathbb{R}^n$ bzw. $\vec{x} \in \mathbb{C}^n$ gilt.

Eine wichtige Struktur auf dem $\mathbb{R}^n$ bzw. dem $\mathbb{C}^n$ ist das Skalarprodukt:

Definition 12.15: (Skalarprodukt)

Ein **„Skalarprodukt"** *auf einem Vektorraum V über dem Skalarenkörper K ist eine Abbildung $\langle ., . \rangle : V \times V \to \mathbb{C}$ mit den Eigenschaften*

a) Linearität:

$$\langle v, \lambda_1 \cdot w_1 + \lambda_2 \cdot w_2 \rangle = \lambda_1 \cdot \langle v, w_1 \rangle + \lambda_2 \cdot \langle v, w_2 \rangle,$$

$$\langle \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2, w \rangle = \overline{\lambda_1} \cdot \langle v_1, w \rangle + \overline{\lambda_2} \cdot \langle v_2, w \rangle,$$

für alle $v, v_1, v_2, w, w_1, w_2 \in V$, $\lambda_1, \lambda_2 \in K$.

b) $\langle v, w \rangle = \overline{\langle w, v \rangle}$ *(komplexe Konjugation) für alle $v, w \in V$.*

c) $\langle v, v \rangle \geq 0$ *für alle $v \in V$. (Beachte: b) impliziert $\langle v, v \rangle \in \mathbb{R}$.)*

d) $\langle v, v \rangle = 0 \;\Leftrightarrow\; v = 0$.

Beispiel 12.16: 1) Auf dem $\mathbb{R}^n$ ist das interessanteste Skalarprodukt

$$\vec{x} \cdot \vec{y} = \langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{i=1}^{n} x_i \cdot y_i.$$

Offensichtlich sind alle geforderten Eigenschaften a) – d) des Skalarprodukts erfüllt. Die komplexe Konjugation ist dabei irrelevant.

2) Auf dem $\mathbb{C}^n$ ist das interessanteste Skalarprodukt

$$\vec{x} \cdot \vec{y} = \langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{i=1}^{n} \overline{x_i} \cdot y_i.$$

Offensichtlich sind alle geforderten Eigenschaften a) – d) des Skalarprodukts erfüllt, wobei die komplexe Konjugation wesentlich ist.

Definition 12.17: (Orthogonalität)
Zwei Vektoren $\vec{x}, \vec{y}$ im $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ heissen **„orthogonal“**, *wenn $\langle \vec{x}, \vec{y}\,\rangle = 0$ gilt.*

Bemerkung 12.18: *Offensichtlich erzeugt das Skalarprodukt die p-Norm aus Bemerkung 12.14 mit $p = 2$:*

$$\|\vec{x}\|_2 = \sqrt{\langle \vec{x}, \vec{x}\,\rangle}.$$

Es gilt die Cauchy-Schwarzsche Ungleichung

$$|\langle \vec{x}, \vec{y}\,\rangle| \leq \|\vec{x}\|_2 \cdot \|\vec{y}\|_2,$$

die wir schon in Satz 11.5 kennengelernt haben (für den Beweis siehe Aufgabe 99 auf Blatt 12). Das Gleichheitszeichen gilt in der Cauchy-Schwarzschen Ungleichung genau dann, wenn die Vektoren $\vec{x}$ und $\vec{y}$ parallel (= linear abhängig) sind.

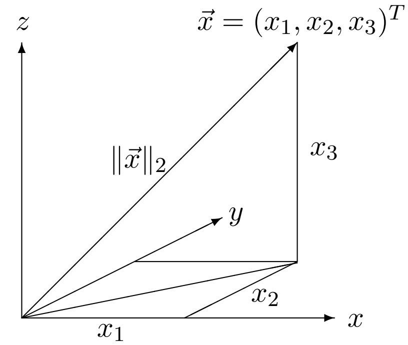
12.2 Etwas Geometrie im $\mathbb{R}^3$

12.2.1 Skalar-, Vektor- und Spatprodukt

Auf dem $\mathbb{R}^3$ (und dem darin eingebetteten $\mathbb{R}^2$) gelten einige wichtige geometrische Interpretationen der eingeführten Struktuen.

Satz 12.19: (Geometrische Interpretation der 2-Norm)
Die 2-Norm $\|\vec{x}\|_2$ eines Vektors $\vec{x} \in \mathbb{R}^2$ bzw. $\vec{x} \in \mathbb{R}^3$ ist seine geometrische Länge.

Beweis:



Pythagoras liefert als geometrische Länge $\sqrt{x_1^2 + x_2^2 + x_3^2} = \|\vec{x}\|_2$.

Q.E.D.

Satz 12.20: (Geometrische Interpretation des Skalarprodukts)
Für beliebige Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^2$ bzw. $\vec{x}, \vec{y} \in \mathbb{R}^3$ gilt

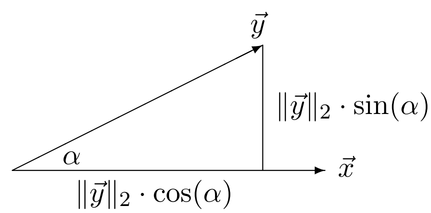
$$\langle \vec{x}, \vec{y}\,\rangle = \|x\|_2 \cdot \|y\|_2 \cdot \cos(\alpha),$$

wo α der Winkel zwischen den Vektoren $\vec{x}$ und $\vec{y}$ ist.

Beweis: Das Skalarprodukt ist invariant unter Drehungen, d.h., es gilt

$$\langle \vec{x}, \vec{y}\,\rangle = \langle \vec{x}\,', \vec{y}\,'\,\rangle,$$

wenn $\vec{x}\,' = O\,\vec{x}$ $\vec{y}\,' = O\,\vec{y}$ durch die selbe Drehung (beschrieben durch eine orthogonal Matrix O) aus $\vec{x}$ und $\vec{y}$ hervorgehen (wir erläutern dies, sobald wir orthogonale Matrizen eingeführt haben). Damit können wir o.B.d.A. die Vektoren so drehen, dass $\vec{x} = (\|\vec{x}\|_2, 0, 0)^T$ und $\vec{y} = (y_1, y_2, 0)^T = (\|\vec{y}\|{\cdot}\cos(\alpha), \|\vec{y}\|{\cdot}\sin(\alpha), 0)^T$ gilt:



Die Definition des Skalarprodukts liefert dann sofort

$$\langle \vec{x}, \vec{y}\,\rangle = \langle \begin{pmatrix} \|x\|_2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \|y\|_2 \cdot \cos(\alpha) \\ \|y\|_2 \cdot \sin(\alpha) \\ 0 \end{pmatrix} \rangle = \|x\|_2 \cdot \|y\|_2 \cdot \cos(\alpha).$$

Q.E.D.

Speziell auf dem $\mathbb{R}^3$ ist ein interessantes Produkt zweier Vektoren definiert, das einen Vektor liefert, der auf den zwei gegebenen Vektoren senkrecht steht:

Definition 12.21: (Das Kreuzprodukt auf dem $\mathbb{R}^3$)

Das Kreuzprodukt $\vec{x} \times \vec{y}$ zweier Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^3$ ist der Vektor

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2 \cdot y_3 - x_3 \cdot y_2 \\ x_3 \cdot y_1 - x_1 \cdot y_3 \\ x_1 \cdot y_2 - x_2 \cdot y_1 \end{pmatrix}.$$

Bemerkung 12.22: *a) Wie leicht nachzurechnen ist, gilt*

$$\langle \vec{x}, \vec{x} \times \vec{y} \rangle = \langle \vec{y}, \vec{x} \times \vec{y} \rangle = 0,$$

d.h., das Kreuzprodukt zweier Vektoren steht senkrecht auf jedem der beiden Vektoren.
b) Das Kreuzprodukt verschwindet genau dann, wenn die beiden Vektoren parallel sind (oder einer der Vektoren verschwindet).
c) Das Kreuzprodukt ist weder kommutativ noch assoziativ. Es gilt

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}$$

und die „Jacobi-Identität“

$$\vec{x} \times (\vec{y} \times \vec{z}) + \vec{y} \times (\vec{z} \times \vec{x}) + \vec{z} \times (\vec{x} \times \vec{y}) = 0$$

für alle $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$.

Satz 12.23: (Geometrische Interpretation des Kreuzprodukts)

Für beliebige Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^3$ gilt

$$\|\vec{x} \times \vec{y}\|_2 = \|x\|_2 \cdot \|y\|_2 \cdot |\sin(\alpha)|,$$

wo α der Winkel zwischen den Vektoren $\vec{x}$ und $\vec{y}$ ist. Die Länge des Kreuzprodukts ist damit die Fläche des von $\vec{x}, \vec{y}$ aufgespannten Parallelogramms.

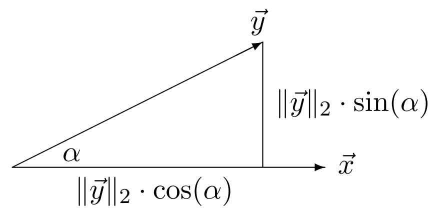
Beweis: Wir argumentieren ähnlich wie bei Satz 12.20. Das Kreuzprodukt ist invariant unter Drehungen: es gilt[2]

$$(O\vec{x}) \times (O\vec{y}) = \pm O(\vec{x} \times \vec{y}),$$

wobei O eine orthogonale Drehmatrix ist (das Vorzeichen hängt davon ab, ob O eine Spiegelung enthält). Da Drehungen die Längen der Vektoren und den Winkel zwischen den Vektoren nicht ändern, können wir o.B.d.A. die Vektoren so drehen, dass

$$\vec{x} = (\|\vec{x}\|_2, 0, 0)^T \quad \text{und} \quad \vec{y} = (y_1, y_2, 0)^T = (\|\vec{y}\| \cdot \cos(\alpha), \|\vec{y}\| \cdot \sin(\alpha), 0)^T$$

gilt:



Die Definition des Kreuzprodukts liefert dann sofort

$$\vec{x} \times \vec{y} = \begin{pmatrix} \|x\|_2 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} \|y\|_2 \cdot \cos(\alpha) \\ \|y\|_2 \cdot \sin(\alpha) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \|x\|_2 \cdot \|y\|_2 \cdot \sin(\alpha) \end{pmatrix}.$$

Die Fläche des von $\vec{x}, \vec{y}$ aufgespannten Parallelogramms ist die Seitenlänge $\|\vec{x}\|_2$ mal der Höhe $\|\vec{y}\|_2 \cdot \sin(\alpha)$.

Q.E.D.

Bemerkung 12.24: *Die Beobachtung, dass $\vec{x} \times \vec{y}$ senkrecht auf $\vec{x}$ und $\vec{y}$ steht, legt die Richtung von $\vec{x} \times \vec{y}$ bis auf ein Vorzeichen geometrisch fest. Der letzte Satz legt die Länge von $\vec{x} \times \vec{y}$ geometrisch fest. Das Vorzeichen von $\vec{x} \times \vec{y}$ ist geometrisch durch die* **„Rechte-Hand-Regel"** *bestimmt: das aus $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$ bestehende Dreibein entspricht Daumen, Zeigefinger und (gekrümmtem) Mittelfinger der rechten Hand. Dies legt in der Tat die Richtung von $\vec{x} \times \vec{y}$ fest, wenn man sich nicht den Mittelfinger brechen will.*

Und noch ein interessantes Produkt auf dem $\mathbb{R}^3$:

Definition 12.25: (Das Spatprodukt auf dem $\mathbb{R}^3$)

Das Spatprodukt $|[\vec{x}, \vec{y}, \vec{z}]|$ dreier Vektoren $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ ist der skalare Wert

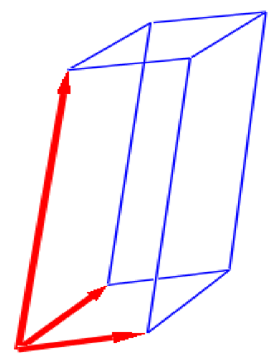
$$|[\vec{x}, \vec{y}, \vec{z}]| = \langle \vec{x}, \vec{y} \times \vec{z} \rangle = \langle \vec{y}, \vec{z} \times \vec{x} \rangle = \langle \vec{z}, \vec{x} \times \vec{y} \rangle.$$

Ein „Dreibein" $\vec{x}, \vec{y}, \vec{z}$ heißt **„rechtshändig"**, *wenn $|[\vec{x}, \vec{y}, \vec{z}]| > 0$ gilt. Für $|[\vec{x}, \vec{y}, \vec{z}]| < 0$ heißt das Dreibein* **„linkshändig"**.

[2] Dies ist etwas mühselig nachzurechnen.

Satz 12.26: (Geometrische Interpretation des Spatprodukts)

Für beliebige Vektoren $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ gilt $|[\vec{x}, \vec{y}, \vec{z}]| = \pm V$, wo V das Volumen des von den Vektoren $\vec{x}, \vec{y}, \vec{z}$ aufgespannten Spats ist:



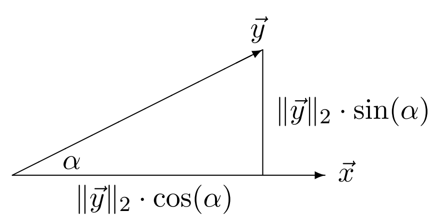
Beweis: Wir argumentieren ähnlich wie bei Satz 12.20 und Satz 12.23. Das Spatprodukt ist invariant unter Drehungen: es gilt

$$|[\vec{x}, \vec{y}, \vec{z}]| = \pm |[O\vec{x}, O\vec{y}, O\vec{z}]|,$$

wobei O eine orthogonale Drehmatrix ist (das Vorzeichen hängt davon ab, ob O eine Spiegelung enthält). O.B.d.A. können wir die Vektoren so drehen, dass

$$\vec{x} = (\|\vec{x}\|_2, 0, 0)^T \quad \text{und} \quad \vec{y} = (y_1, y_2, 0)^T = (\|\vec{y}\| \cdot \cos(\alpha), \|\vec{y}\| \cdot \sin(\alpha), 0)^T$$

gilt:



Die Definition des Kreuzprodukts liefert dann sofort

$$|[\vec{x}, \vec{y}, \vec{z}]| = \langle \vec{z}, \vec{x} \times \vec{y} \rangle = \langle \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \begin{pmatrix} \|x\|_2 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} \|y\|_2 \cdot \cos(\alpha) \\ \|y\|_2 \cdot \sin(\alpha) \\ 0 \end{pmatrix} \rangle$$

$$= \langle \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \|x\|_2 \cdot \|y\|_2 \cdot \sin(\alpha) \end{pmatrix} \rangle = z_3 \cdot \|x\|_2 \cdot \|y\|_2 \cdot \sin(\alpha).$$

Hierbei ist $\|x\|_2 \cdot \|y\|_2 \cdot \sin(\alpha)$ die Fläche des von $\vec{x}, \vec{y}$ aufgespannten Parallelogramms (die Grundfläche des Spats) und z_3 ist die Höhe des Spats (die

Grundfläche liegt hier in der x-y-Ebene). Das Spatvolumen ist Grundfläche mal Höhe.

Q.E.D.

12.2.2 Geraden und Ebenen

Definition 12.27: (Geraden)

Zu gegebenen $\vec{x}_0 \in \mathbb{R}^2$ bzw. $\mathbb{R}^3$ und $0 \neq \vec{d} \in \mathbb{R}^2$ bzw. $\mathbb{R}^3$ definieren wir die **Gerade durch $\vec{x}_0$ mit dem Richtungsvektor** $\vec{d}$ *als die Menge aller Punkte der Form*

$$\Big\{\vec{x}_0 + \lambda \cdot \vec{d};\ \lambda \in \mathbb{R}\Big\}.$$

Definition 12.28: (Ebenen)

Zu gegebenem $\vec{x}_0 \in \mathbb{R}^3$ und linear unabhängigen Vektoren $\vec{d}_1, \vec{d}_2 \in \mathbb{R}^3$ definieren wir die **Ebene durch $\vec{x}_0$ mit den Tangentialvektoren** $\vec{d}_1, \vec{d}_2$ *als die Menge aller Punkte der Form*

$$\Big\{\vec{x}_0 + \lambda_1 \cdot \vec{d}_1 + \lambda_2 \cdot \vec{d}_2\,;\ \lambda_1, \lambda_2 \in \mathbb{R}\Big\}.$$

In dieser Definition werden 9 Parameter (jeweils 3 Komponenten der Vektoren $\vec{x}_0, \vec{d}_1, \vec{d}_2$) zur Definition der Ebene vorgegeben. Alternativ kann man die Ebene auch durch insgesamt nur 3 Parameter festlegen, denn eine Ebene ist als die Menge aller Punkte $(x_1, x_2, x_3)^T$ bestimmt, die eine lineare Gleichung

$$\langle \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rangle = n_1 \cdot x_1 + n_2 \cdot x_2 + n_3 \cdot x_3 = d$$

bestimmen, wobei o.B.d.A. durch Multiplikation der Gleichung mit einer geeigneten Zahl der Wert von d auf 0 oder 1 festgelegt werden kann. Sei

$$\vec{n} = \vec{d}_1 \times \vec{d}_2$$

der „Normalenvektor“ der Ebene, der senkrecht auf den beiden Tangentenvektoren $\vec{d}_1, \vec{d}_2$ steht. In der Tat folgt aus der Parameterform

$$\vec{x} = \vec{x}_0 + \lambda_1 \cdot \vec{d}_1 + \lambda_2 \cdot \vec{d}_2$$

der Ebenenpunkte durch skalare Multiplikation mit $\vec{n}$:

$$\langle \vec{n}, \vec{x}\,\rangle = \langle \vec{n}, \vec{x}_0\,\rangle + \lambda_1 \cdot \underbrace{\langle \vec{n}, \vec{d}_1\,\rangle}_{=0} + \lambda_2 \cdot \underbrace{\langle \vec{n}, \vec{d}_2\,\rangle}_{=0} = \langle \vec{n}, \vec{x}_0\,\rangle,$$

d.h., $\langle \vec{n}, \vec{x}\,\rangle$ hat für alle Ebenenpunkte den selben Wert $\langle \vec{n}, \vec{x}_0 \rangle$ (=: d). Die Gleichung

$$\langle \vec{n}, \vec{x}\,\rangle = d$$

heißt **„Hessesche Normalform“** der Ebene. Der Vektor $\vec{n}$ ist dabei der **Normalenvektor** der Ebene. Der Parameter d hat folgende geometrische Interpretation: Wir starten am Nullpunkt und laufen orthogonal auf die Ebene zu, d.h., wir betrachten die Punkte auf der Geraden $\vec{x} = \lambda \cdot \vec{n}$. Für den Durchstoßpunkt durch die Ebene muß

$$\langle \vec{n}, \vec{x}\,\rangle = \langle \vec{n}, \lambda \cdot \vec{n}\,\rangle = \lambda \cdot \langle \vec{n}, \vec{n}\,\rangle = d \qquad \Rightarrow \qquad \lambda = \frac{d}{\langle \vec{n}, \vec{n}\,\rangle}$$

gelten, der Durchstoßpunkt ist also

$$\vec{x} = \frac{d}{\langle \vec{n}, \vec{n}\,\rangle} \cdot \vec{n}.$$

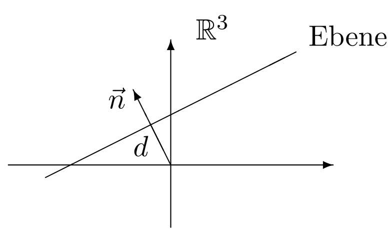
Der Betrag dieses Vektors

$$\|\vec{x}\|_2 = \frac{d}{\langle \vec{n}, \vec{n}\,\rangle} \cdot \|\vec{n}\|_2 = \frac{d}{\|\vec{n}\|_2}$$

ist der Abstand der Ebene zum Nullpunkt:

Für $\|\vec{n}\|_2 = 1$ ist der Parameter d der durch $\langle \vec{n}, \vec{x}\,\rangle = d$ gegebenen Ebene der Abstand der Ebene vom Nullpunkt.

Ist $d > 0$, zeigt der Normalenvektor vom Ursprung in Richtung der Ebene:



12.3 Matrizen

Betrachten wir den Fall eines Vektors $\vec{x} = (x_1, \ldots, x_n)^T$, der mit der Standardbasis

$$\vec{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

als $\vec{x} = \sum_j x_j \cdot \vec{e}_j$ gegeben ist. Angenommen, wir haben eine alternative Basis $\vec{e}_i'$ gegeben, die durch die Linearkombinationen

$$\vec{e}_j = \sum_{i=1}^{n} \vec{e}_i' \cdot a_{ij}, \quad j = 1, \ldots, n$$

mit der Standardbasis verknüpft ist. In der neuen Basis gilt $\vec{x} = \sum_i x_i' \cdot \vec{e}_i'$, wobei sich die neuen Koeffizienten x_i' durch

$$\vec{x} = \sum_{i=1}^{n} x_i' \cdot \vec{e}_i' = \sum_{j=1}^{n} x_j \cdot \vec{e}_j = \sum_{j=1}^{n} x_j \cdot \sum_{i=1}^{n} \vec{e}_i' \cdot a_{ij} = \sum_{i=1}^{n} \underbrace{\Big(\sum_{j=1}^{n} a_{ij} \cdot x_j\Big)}_{x_i'} \cdot \vec{e}_i'$$

ergeben, also

$$\boxed{x_i' = \sum_{j=1}^{n} a_{ij} \cdot x_j, \quad i = 1, \ldots, n.}$$

Wir schreiben dies formal als

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

wobei wir die Koeffizienten (a_{ij}) in ein rechteckiges Schema („Matrix“) einordnen. Wir führen solche Rechteck-Schemata nun formal ein und vereinbaren, wie wir damit rechnen.

Definition 12.29: (Matrizen)

Sei K ein Skalarenkörper (also für uns $K = \mathbb{R}$ bzw. $K = \mathbb{C}$). Die Menge

$$K^{n \times m} = \left\{ \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} ;\ a_{11}, \ldots, a_{nm} \in K \right\}$$

wird als **Raum der** $(n \times m)$**-Matrizen** *bezeichnet. Wir benutzen die Notation $A = (a_{ij}) \in K^{n \times m}$. Der Raum $K^{n \times m}$ wird mit den Operationen*

$$(A + B)_{ij} = a_{ij} + b_{ij}, \quad (\lambda \cdot A)_{ij} = \lambda \cdot a_{ij}, \quad i = 1, \ldots, n,\ j = 1, \ldots, m$$

für $A = (a_{ij}) \in K^{n \times m}$, $B = (b_{ij}) \in K^{n \times m}$, $\lambda \in K$ zu einem Vektorraum über K.

Neben der Addition von Matrizen und der Multiplikation mit Skalaren definieren wir eine Multiplikation von Matrizen:

$$\boxed{(n \times p)\text{-Matrix} \;\times\; (p \times m)\text{-Matrix} = (n \times m)\text{-Matrix.}}$$

Definition 12.30: (Matrizenmultiplikation)
Für $A = (a_{ij}) \in K^{n\times p}$, $B = (b_{ij}) \in K^{p\times m}$ definieren wir das Produkt $C = AB$ als die Matrix $C = (c_{ij}) \in K^{n\times m}$ mit

$$c_{ij} = \sum_{k=1}^{p} a_{ik} \cdot b_{kj}, \quad i = 1, \ldots, n, \; j = 1, \ldots, m.$$

Spaltenvektoren entsprechen Matrizen mit nur einer Spalte

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^{n\times 1},$$

Zeilenvektoren entsprechen Matrizen mit nur einer Zeile:

$$(x_1, \ldots, x_n) \in K^{1\times n}.$$

Das Matrixprodukt einer Zeile mit einer Spalte entspricht dem Skalarprodukt:

$$(a_1, \ldots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \sum_{k=1}^{n} a_k \cdot b_k.$$

Bemerkung 12.31: *Das Matrixprodukt läßt sich damit folgendermassen verinnerlichen:*

$$\begin{pmatrix} a_{11} & \ldots & a_{1p} \\ \vdots & \ddots & \vdots \\ \boxed{a_{i1} \;\; \ldots \;\; a_{ip}} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{np} \end{pmatrix} \begin{pmatrix} b_{11} & \ldots & \boxed{\begin{matrix} b_{1j} \\ \vdots \\ b_{pj} \end{matrix}} & \ldots & b_{1m} \\ \vdots & \ddots & & \ddots & \vdots \\ b_{p1} & \ldots & & \ldots & b_{pm} \end{pmatrix} = \begin{pmatrix} c_{11} & \ldots & c_{1m} \\ \vdots & \boxed{c_{ij}} & \vdots \\ c_{n1} & \ldots & c_{nm} \end{pmatrix},$$

wobei

$$\begin{aligned} c_{ij} &= \textit{Skalarprodukt}\Big((i\textit{-te Zeile von } A) \times (j\textit{-te Spalte von } B)\Big) \\ &= (a_{i1}, \ldots, a_{ip}) \begin{pmatrix} b_{1j} \\ \vdots \\ b_{pj} \end{pmatrix} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}. \end{aligned}$$

Bemerkung 12.32: *Das Matrixprodukt (mit Matrizen A, B, C zulässiger Dimension) ist assoziativ:*

$$(AB)C = A(BC),$$

denn

$$\begin{array}{rclcl}
\Big((AB)C\Big)_{ij} & = & \displaystyle\sum_{k_2}\left(\sum_{k_1} a_{ik_1}\cdot b_{k_1k_2}\right)\cdot c_{k_2j} & = & \displaystyle\sum_{k_1}\sum_{k_2} a_{ik_1}\cdot b_{k_1k_2}\cdot c_{k_2j} \\
 & & & & \Vert \\
\Big(A(BC)\Big)_{ij} & = & \displaystyle\sum_{k_1} a_{ik_1}\left(\sum_{k_2} b_{k_1k_2}\cdot c_{k_2j}\right) & = & \displaystyle\sum_{k_1}\sum_{k_2} a_{ik_1}\cdot b_{k_1k_2}\cdot c_{k_2j}.
\end{array}$$

Warnung: *Das Matrixprodukt ist nicht kommutativ, i.A. gilt $AB \neq BA$. Beispiel:*

$$\begin{array}{rcccccc}
 & & A & & B & & \\
AB & = & \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} & & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} & = & \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \\
 & & & & & & \not\parallel \\
BA & = & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} & & \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} & = & \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.
\end{array}$$

Wir führen weiterhin ein:

Definition 12.33: (Transposition)

a) *Die* **Transponierte** $A^T = (b_{ij})$ *einer $(n \times m)$-Matrix $A = (a_{ij})$ ist die $(m \times n)$-Matrix*

$$b_{ij} = a_{ji}, \quad i = 1, \ldots, m, \; j = 1, \ldots, n,$$

b) *Eine Matrix heißt* **symmetrisch** *(oder auch* **selbstadjungiert***), wenn $A = A^T$ gilt, also $a_{ij} = a_{ji}$.*

   Sie heißt **antisymmetrisch** *(oder auch* **schiefsymmetrisch***), wenn $A = -A^T$ gilt, also $a_{ij} = -a_{ji}$.*

c) *Die* **hermitesch Transponierte** $A^H = (c_{ij})$ *einer (komplexen) $(n \times m)$-Matrix $A = (a_{ij})$ ist die $(m \times n)$-Matrix*

$$c_{ij} = \overline{a_{ji}}, \quad i = 1, \ldots, m, \; j = 1, \ldots, n,$$

   (Transposition plus komplexe Konjugation).

d) *Eine Matrix heißt* **hermitesch***, wenn $A = A^H$ gilt, also $a_{ij} = \overline{a_{ji}}$. Sie heißt* **antihermitesch** *(oder auch* **schiefhermitesch***), wenn $A = -A^H$ gilt, also $a_{ij} = -\overline{a_{ji}}$.*

Bemerkung 12.34: *Für das reelle Skalarprodukt* $\langle \vec{x}, \vec{y}\,\rangle = \sum_k x_k \cdot y_k$ *auf dem* $\mathbb{R}^n$ *gilt offensichtlich*

$$\langle \vec{x}, A\,\vec{y}\,\rangle = \langle A^T \vec{x}, \vec{y}\,\rangle = \textstyle\sum_{ij} x_i \cdot a_{ij} \cdot y_j$$

für jedes $\vec{x}, \vec{y} \in \mathbb{R}^n$ *und jedes* $A \in \mathbb{R}^{n\times n}$.
Für das komplexe Skalarprodukt $\langle \vec{x}, \vec{y}\,\rangle = \sum_k \overline{x_k} \cdot y_k$ *auf dem* $\mathbb{C}^n$ *gilt offensichtlich*

$$\langle \vec{x}, A\,\vec{y}\,\rangle = \langle A^H \vec{x}, \vec{y}\,\rangle = \textstyle\sum_{ij} \overline{x_i} \cdot a_{ij} \cdot y_j$$

für jedes $\vec{x}, \vec{y} \in \mathbb{C}^n$ *und jedes* $A \in \mathbb{C}^{n\times n}$.

Durch (hermitesche) Transposition kann eine Matrix von einer Seite des Skalarprodukts auf die andere Seite bewegt werden.

Definition 12.35: (Einheitsmatrix)

Die Matrix

$$\mathbb{1} = (\delta_{ij}) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

auf dem $K^{n\times n}$ *heißt* $(n \times n)$-**Einheitsmatrix**. *Es gilt*

$$\mathbb{1}\vec{x} = \vec{x}, \quad \mathbb{1}A = A, \quad B\,\mathbb{1} = B$$

für alle $\vec{x} \in K^n$, *alle* $A \in K^{n\times m}$ *und alle* $B \in K^{m\times n}$.

Definition 12.36: (Inverse einer Matrix)

Existiert zu einer quadratischen Matrix $A \in K^{n\times n}$ *eine Matrix* $A^{-1} \in K^{n\times n}$ *mit*

$$A^{-1}A = \mathbb{1},$$

so heißt A^{-1} **Inverse von** A. *Es gilt dann automatisch auch* $AA^{-1} = \mathbb{1}$.

Wegen der Nichtkommutativität des Matrixprodukts ist dabei kurz zu überlegen, dass die „Linksinverse“ mit der Eigenschaft $A^{-1}A = \mathbb{1}$ auch automatisch eine „Rechtsinverse“ mit $AA^{-1} = \mathbb{1}$ ist: Angenommen, eine Rechtsinverse B mit $AB = \mathbb{1}$ existiere. Multiplizieren wir von links mit A^{-1}, erhalten wir $A^{-1}AB = B = A^{-1}\,\mathbb{1} = A^{-1}$, also: $B = A^{-1}$.

Es ist eine fundamentale Aufgabe, Kriterien zu suchen, wann eine Matrix invertierbar ist. Wir werden sehen, dass z.B. eine Matrix genau dann invertierbar ist, wenn die Determinante von A nicht verschwindet. Weiterhin ist es wichtig, einen Algorithmus zur Berechnung der Inversen anzugeben. Dies wird in den Abschnitten 12.3 und 12.5 geschehen.

12.4 Determinanten

Definition 12.37: (Permutation)

Eine invertierbare Abbildung $p : \{1, \ldots, n\} \to \{1, \ldots, n\}$ *heißt* **„Permutation"** *der Zahlen* $1, \ldots, n$.

Die **„Inversionszahl"** *einer Permutation ist*

$$I_p = \Big| \Big\{ (j,k);\ j,k \in \{1, \ldots, n\},\ j < k,\ p(j) > p(k) \Big\} \Big|$$

(also: I_p *ist die Anzahl aller Paare* (j,k) *mit* $j < k$*, bei denen die Permutation die Anordung* $j < k$ *zu* $p(j) > p(k)$ *vertauscht).*

Das **„Signum"** *einer Permutation ist* $\mathrm{sign}(p) = (-1)^{I_p}$.

Satz 12.38:

Es gibt $n!$ *verschiedene Permutationen der Zahlen* $1, \ldots, n$.

Beweis: Für $p(1)$ gibt es n verschiedene Möglichkeiten, für $p(2)$ gibt es $n-1$ verschiedene Möglichkeiten (da $p(2) \neq p(1)$ gelten muss), für $p(3)$ gibt es $n-2$ verschiedene Möglichkeiten (da $p(3) \neq p(1)$ und $p(3) \neq p(2)$ gelten muss), usw. Damit gibt es insgesamt $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1 = n!$ Möglichkeiten, eine Permutation von $1, \ldots, n$ zu konstruieren.

Q.E.D.

$\boxed{\downarrow 27.1.04}$

Konstruiert man alle Permutationen, indem man Schritt für Schritt einzelne Zahlenpaare in einer Permutation austauscht, kann man leicht das Signum nachhalten: mit jeder Vertauschung ändert das Signum sein Vorzeichen.

Satz 12.39:

Unterscheiden sich zwei Permutationen p *und* $\tilde{p}$ *nur durch den Austausch eines einzelnen Paares* (j_0, k_0)*, also*

$$\tilde{p}(j_0) = p(k_0),\ \tilde{p}(k_0) = p(j_0),$$

$$\tilde{p}(i) = p(i)\ \forall i \in \{1, \ldots, n\} \setminus \{j_0, k_0\},$$

so gilt

$$\mathrm{sign}(p) = -\mathrm{sign}(\tilde{p}).$$

Beweis: (für technisch Interessierte)
Es gelte o.B.d.A. $j_0 < k_0$ und $p(k_0) > p(j_0)$ (für $p(k_0) < p(j_0)$ sind in den folgenden Betrachtungen die Mengen A und $\tilde{A}$ auszutauschen).
Die Paare (j,k) mit $j,k \in \{1,\dots,n\} \setminus \{j_0,k_0\}$ sind jeweils in beiden Mengen

$$A = \Big\{(j,k);\ j,k \in \{1,\dots,n\},\ j<k,\ p(j)>p(k)\Big\}$$

und

$$\tilde{A} = \Big\{(j,k);\ j,k \in \{1,\dots,n\},\ j<k,\ \tilde{p}(j)>\tilde{p}(k)\Big\}$$

gemeinsam vorhanden oder nicht vorhanden, da für diese Zahlen $\tilde{p}(j) = p(j)$ und $\tilde{p}(k) = p(k)$ gilt. Für $j \in \{j_0,k_0\}$ bzw. $k \in \{j_0,k_0\}$ enthält A die Teilmenge

$$\begin{aligned}
&\Big\{(k_0,k);\ k \in \{k_0+1,\dots,n\},\ p(k_0) > p(k)\Big\}\\
\cup\ &\Big\{(j_0,k);\ k \in \{k_0+1,\dots,n\} \setminus \{k_0\},\ p(j_0) > p(k)\Big\}\\
\cup\ &\Big\{(j,k_0);\ j \in \{1,\dots,k_0-1\} \setminus \{j_0\},\ p(j) > p(k_0)\Big\}\\
\cup\ &\Big\{(j,j_0);\ j \in \{1,\dots,j_0-1\},\ p(j) > p(j_0)\Big\},
\end{aligned}$$

während $\tilde{A}$ die Teilmenge

$$\begin{aligned}
\Big\{(j_0,k_0)\Big\}\ \cup\ &\Big\{(k_0,k);\ k \in \{k_0+1,\dots,n\},\ \underbrace{\tilde{p}(k_0)}_{=p(j_0)} > \underbrace{\tilde{p}(k)}_{=p(k)}\Big\}\\
\cup\ &\Big\{(j_0,k);\ k \in \{k_0+1,\dots,n\} \setminus \{k_0\},\ \underbrace{\tilde{p}(j_0)}_{=p(k_0)} > \underbrace{\tilde{p}(k)}_{=p(k)}\Big\}\\
\cup\ &\Big\{(j,k_0);\ j \in \{1,\dots,j_0-1\} \setminus \{j_0\},\ \underbrace{\tilde{p}(j)}_{=p(j)} > \underbrace{\tilde{p}(k_0)}_{=p(j_0)}\Big\}\\
\cup\ &\Big\{(j,k_0);\ j \in \{1,\dots,j_0-1\},\ \underbrace{\tilde{p}(j)}_{=p(j)} > \underbrace{\tilde{p}(j_0)}_{=p(k_0)}\Big\}
\end{aligned}$$

enthält. Unter der Annahme $p(k_0) > p(j_0)$ folgt damit:

$$\begin{aligned}
\tilde{A} \setminus A = \Big\{(j_0,k_0)\Big\}\ \cup\ &\Big\{(j_0,k);\ k \in \{j_0+1,\dots,n\},\ p(j_0) < p(k) < p(k_0)\Big\}\\
\cup\ &\Big\{(j,k_0);\ j \in \{1,\dots,k_0-1\},\ p(j_0) < p(j) < p(k_0)\Big\}
\end{aligned}$$

und

$$\begin{aligned}
A \setminus \tilde{A} = \ &\Big\{(k_0,k);\ k \in \{k_0+1,\dots,n\},\ p(j_0) < p(k) < p(k_0)\Big\}\\
\cup &\Big\{(j,j_0);\ j \in \{1,\dots,j_0-1\},\ p(j_0) < p(j) < p(k_0)\Big\}.
\end{aligned}$$

Mit $A = (A \cap \tilde{A}) \cup (A \setminus \tilde{A})$ und $\tilde{A} = (A \cap \tilde{A}) \cup (\tilde{A} \setminus A)$ folgt

$$|\tilde{A}| - |A| = 1 + 2 \cdot \Big|\Big\{j;\ j \in \{j_0 + 1, \ldots, k_0 - 1\},\ p(j_0) < p(j) < p(k_0)\Big\}\Big|.$$

Damit unterscheiden sich $|\tilde{A}|$ und $|A|$ um eine ungerade ganze Zahl, womit $\mathrm{sign}(\tilde{p}) = (-1)^{|\tilde{A}|} = -(-1)^{|A|} = -\mathrm{sign}(p)$ folgt.

Q.E.D.

Beispiel 12.40: Wir schreiben eine Permutation in der Form

$$\frac{1 \quad 2 \quad \ldots \quad n}{p(1) \quad p(2) \quad \ldots \quad p(n)}.$$

Für die $3! = 6$ möglichen Permutationen der Zahlen $1, 2, 3$ gilt:

$$p_1 : \quad \frac{1\ 2\ 3}{1\ 2\ 3}, \quad \mathrm{sign}(p_1) = 1,$$

$$p_2 : \quad \frac{1\ 2\ 3}{1\ 3\ 2}, \quad \mathrm{sign}(p_2) = -1,$$

$$p_3 : \quad \frac{1\ 2\ 3}{3\ 1\ 2}, \quad \mathrm{sign}(p_3) = 1,$$

$$p_4 : \quad \frac{1\ 2\ 3}{3\ 2\ 1}, \quad \mathrm{sign}(p_4) = -1,$$

$$p_5 : \quad \frac{1\ 2\ 3}{2\ 3\ 1}, \quad \mathrm{sign}(p_5) = 1,$$

$$p_6 : \quad \frac{1\ 2\ 3}{2\ 1\ 3}, \quad \mathrm{sign}(p_6) = -1.$$

Definition 12.41: (Determinante einer Matrix)

Die Determinante einer quadratischen Matrix $A = (a_{i,j}) \in K^{n \times n}$ ist

$$\det(A) = \sum_p \mathrm{sign}(p) \cdot a_{1p(1)} \cdot a_{2p(2)} \cdot \ldots \cdot a_{np(n)},$$

wobei sich diese Summe über alle $n!$ Permutationen p der Zahlen $1, \ldots, n$ erstreckt.

Beispiel 12.42: Für die Fälle $n = 1, 2, 3$ ergibt sich

$$\det((a_{11})) = a_{11},$$

$$\det\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21},$$

$$\det\left(\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\right) = \begin{array}{c} a_{11}\cdot a_{22}\cdot a_{33} + a_{12}\cdot a_{23}\cdot a_{31} + a_{13}\cdot a_{21}\cdot a_{32} \\ -a_{11}\cdot a_{23}\cdot a_{32} - a_{12}\cdot a_{21}\cdot a_{33} - a_{13}\cdot a_{22}\cdot a_{31}. \end{array}$$

Bezeichnung 12.43:

Seien

$$\vec{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \ \ldots\ , \ \vec{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$$

Vektoren im K^n. Die aus diesen Spalten aufgebaute Matrix wird mit $[\vec{a}_1, \ldots, \vec{a}_n]$ bezeichnet:

$$[\vec{a}_1, \ldots, \vec{a}_n] = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}.$$

Bemerkung 12.44: *Die Definition der Matrixmultiplikation ist so, dass*

$$A\,[\vec{b}_1, \ldots, \vec{b}_n] = [A\,\vec{b}_1, \ldots, A\,\vec{b}_n]$$

gilt, d.h., in einem Produkt AB ergeben sich die Spalten des Produkts als A wirkend auf die einzelnen Spalten von B.

Satz 12.45: (Linearität der Determinante bzgl. der Spalten)

Die Determinante ist linear in jeder Matrixspalte:

$$\det([\vec{a}_1, \ldots, \lambda\cdot\vec{a}_i + \lambda'\cdot\vec{a}'_i, \ldots, \vec{a}_n]) =$$

$$\lambda\cdot\det([\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n]) + \lambda'\cdot\det([\vec{a}_1, \ldots, \vec{a}'_i, \ldots, \vec{a}_n]).$$

Beweis: Dies folgt unmittelbar aus der Definition

$$\det(A) = \sum_p \mathrm{sign}(p)\cdot a_{1p(1)}\cdot a_{2p(2)}\cdot\ldots\cdot a_{np(n)},$$

denn in jedem Summanden taucht genau eine Komponente jeder Spalte genau einmal auf.

Q.E.D.

Satz 12.46: (Antisymmetrie der Determinante bzgl. der Spalten)

Vertauscht man zwei Spalten, so wechselt die Determinante das Vorzeichen:

$$\det([\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n]) = -\det([\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n]).$$

Beweis: Dies folgt unmittelbar aus der Definition

$$\det(A) = \sum_p \mathrm{sign}(p) \cdot a_{1p(1)} \cdot a_{2p(2)} \cdot \ldots \cdot a_{np(n)}.$$

Sei A die Ausgangsmatrix, sei $\tilde{A}$ die Matrix, die durch Austausch zweier Spalten entsteht. Sei τ die Permutation, die die beiden Spaltenindizes austauscht und alle anderen Indizes beibehält. Es gilt

$$\det(\tilde{A}) = \sum_p \mathrm{sign}(p) \cdot \tilde{a}_{1p(1)} \cdot \tilde{a}_{2p(2)} \cdot \ldots \cdot \tilde{a}_{np(n)},$$

wobei $\tilde{a}_{ij} = a_{i\tau(j)}$. Damit folgt

$$\det(\tilde{A}) = \sum_p \mathrm{sign}(p) \cdot a_{1\tau(p(1))} \cdot a_{2\tau(p(2))} \cdot \ldots \cdot a_{n\tau(p(n))}.$$

Läuft p über alle Permutationen, läuft auch $\tilde{p} = \tau \circ p$ über alle Permutationen, wobei nach Satz 12.39 $\mathrm{sign}(\tilde{p}) = -\mathrm{sign}(p)$ gilt:

$$\det(\tilde{A}) = \sum_{\tilde{p}} (-\mathrm{sign}(\tilde{p})) \cdot a_{1\tilde{p}(1)} \cdot a_{2\tilde{p}(2)} \cdot \ldots \cdot a_{n\tilde{p}(n)} = -\det(A).$$

Q.E.D.

Folgerung 12.47:

a) *Stimmen in einer Matrix zwei Spalten überein, so verschwindet ihre Determinante.*

b) *Addiert man zu einer Spalte ein beliebiges Vielfaches einer anderen Spalte hinzu, so verändert sich der Wert der Determinante nicht.*

Beweis: a) Vertauscht man die beiden übereinstimmenden Spalten, muss die Determinante das Vorzeichen wechseln, andererseits aber gleich bleiben. Das kann nur sein, wenn die Determinante 0 ist.

b) Mit der Linearität nach Satz 12.45 gilt

$$\det([\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j + \lambda \cdot \vec{a}_i, \ldots, \vec{a}_n])$$
$$= \det([\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j \ldots, \vec{a}_n]) + \lambda \cdot \det([\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_i \ldots, \vec{a}_n]),$$

wobei die 2-te Determinante nach a) verschwindet.

Q.E.D.

Satz 12.48: (Invarianz unter Transposition)

Für jedes $A \in K^{n\times n}$ *gilt* $\det(A^T) = \det(A)$.

Beweis: Zu jeder Permutation p gibt es die inverse Permutation p^{-1} mit $p \circ p^{-1} = p^{-1} \circ p =$ identische Abbildung : $i \to i$, wobei $\mathrm{sign}(p^{-1}) = \mathrm{sign}(p)$ gilt. Läuft p durch alle Permutationen, läuft auch p^{-1} durch alle Permutationen. Damit gilt

$$\det(A) = \sum_p \mathrm{sign}(p) \cdot a_{1p(1)} \cdot a_{2p(2)} \cdot \ldots \cdot a_{np(n)},$$

$$= \sum_p \mathrm{sign}(p^{-1}) \cdot a_{1p^{-1}(1)} \cdot a_{2p^{-1}(2)} \cdot \ldots \cdot a_{np^{-1}(n)}.$$

Für $j = p^{-1}(i)$ ergibt sich mit $a_{ip^{-1}(i)} = a_{p(j)j}$ nach geeigneter Umsortierung der Reihenfolge der Faktoren im Produkt:

$$\det(A) = \sum_p \underbrace{\mathrm{sign}(p^{-1})}_{=\mathrm{sign}(p)} \cdot a_{p(1)1} \cdot a_{p(2)2} \cdot \ldots \cdot a_{p(n)n}. \qquad (\#)$$

Dies ist mit $a_{ji} = (A^T)_{ij}$ die Formel für $\det(A^T)$.

Q.E.D.

Folgerung 12.49:

$\boxed{\text{\textit{Die Rolle von Zeilen und Spalten sind bei der Determinantenberechnung austauschbar.}}}$

Beispielsweise verschwindet analog zu Folgerung 12.47.a) die Determinante, wenn zwei Zeilen übereinstimmen. Man kann ein Vielfaches einer Zeile zu einer anderen addieren, ohne die Determinante zu verändern usw.

Definition 12.50: (Minoren)

Zu $A \in K^{n\times n}$ *und* $i, j \in \{1, \ldots, n\}$ *definiere den* **„Minor"**

$$m_{ij} = \det\left(\begin{pmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nj} & \ldots & a_{nn} \end{pmatrix}\right),$$

d.h., die (Unter-)Determinante, die durch Streichen der Zeile i *und der Spalte* j *entsteht.*

Der folgende Satz liefert eine rekursive Berechnungsmöglichkeit für Determinanten. Durch „Entwicklung nach einer Spalte bzw. Zeile“ ergibt sich eine $(n \times n)$-Determinante als Summe über Minoren (also $((n-1)\times(n-1))$-Determinanten), so dass man rekursiv die Matrixdimension reduzieren kann, bis sich eine große Summe von (1×1)-Determinanten ergibt, die durch den entsprechenden Matrixeintrag gegeben sind.

Satz 12.51: (Laplacescher Entwicklungssatz)

Seien m_{ij} die Minoren der Matrix $A \in K^{n\times n}$. Für jede Spalte j gilt

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot m_{ij}.$$

Man nennt diese Formel **„Entwicklung der Determinante nach der j-ten Spalte“**.

Für jede Zeile i gilt

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot m_{ij}.$$

Man nennt diese Formel **„Entwicklung der Determinante nach der i-ten Zeile“**.

Beweisskizze: (Technisch) Wir entwickeln zunächst nach der letzten Spalte, wobei wir die Darstellung (#) aus dem Beweis von Satz 12.48 verwenden. In

$$\det(A) = \sum_{p} \operatorname{sign}(p) \cdot a_{p(1)1} \cdot a_{p(2)2} \cdot \ldots \cdot a_{p(n)n}$$

fixieren wir beim Durchlauf durch alle Permutationen zunächst den Bildwert für n, d.h., wir geben $p(n) = i$ mit $i \in \{1, \ldots, n\}$ vor:

$$\det(A) = \sum_{i=1}^{n} a_{in} \cdot \sum_{p_i} \operatorname{sign}(p_i) \cdot a_{p_i(1)1} \cdot \ldots \cdot a_{p_i(n-1),n-1},$$

wobei im i-ten Summanden die (Teil-)Permutation p_i alle invertierbaren Abbildungen $\{1, 2, \ldots, n-1\} \to \{1, \ldots, n\} \setminus \{i\}$ zu durchlaufen hat. (In $\operatorname{sign}(p_i)$ ist p_i als Permutation aller Zahlen $\{1, \ldots, n\}$ mit vorgebenem $p_i(n) = i$ aufzufassen.) Offensichtlich enthält

$$\sum_{p_i} \operatorname{sign}(p_i) \cdot a_{p_i(1)1} \cdot \ldots \cdot a_{p_i(n-1),n-1}$$

alle Terme, die in der Definition des Minors m_{in} auftauchen (der Zeilenindex nimmt alle Werte in $\{1,\ldots,n\}\setminus\{i\}$ an, der Spaltenindex n taucht im Produkt $a_{p_i(1)1}\cdot\ldots\cdot a_{p_i(n-1),n-1}$ nicht auf). Es verbleibt, die Vorzeichen $\mathrm{sign}(p_i)$ nachzuhalten. Sei

$$\tilde{p}_i(k)=\begin{cases} p_i(k) & \text{falls } p_i(k)\in\{1,\ldots,i-1\},\\ i & \text{falls } p_i(k)=i, \text{ also } k=n,\\ p_i(k)-1 & \text{falls } p_i(k)\in\{i+1,\ldots,n\}.\end{cases}$$

Diese Permutation unterscheidet sich von p_i durch die $n-i$ Vertauschungen $i\leftrightarrow i+1$, $i+1\leftrightarrow i+2$, …, $n-1\leftrightarrow n$, also gilt $\mathrm{sign}(p_i)=(-1)^{n-i}\cdot\mathrm{sign}(\tilde{p}_i)$:

$$\sum_{p_i}\mathrm{sign}(p_i)\cdot a_{p_i(1)1}\cdot\ldots\cdot a_{p_i(n-1),n-1}=(-1)^{n-i}\sum_{\tilde{p}_i}\mathrm{sign}(\tilde{p}_i)\cdot\tilde{a}_{\tilde{p}_i(1)1}\cdot\ldots\cdot\tilde{a}_{\tilde{p}_i(n-1),n-1},$$

wobei $(\tilde{a}_{ij})$ die Matrix ist, die durch Streichen der i-ten Zeile aus $A=(a_{ij})$ entsteht (die ursprüngliche i-te Zeile kann man sich als n-te Zeile von $(\tilde{a}_{ij})$ vorstellen).
Die Permutation $\tilde{p}_i:\{1,\ldots,n-1\}\to\{1,\ldots,n-1\}$ ist als Permutation der Zahlen $\{1,\ldots,n-1\}$ interpretierbar. Damit ist die obige Summe über die $\tilde{p}_i$ nichts anderes als die Determinate der Ausgangsmatrix A, aus der die i-te Zeile und die n-te Spalte gestrichen wurde (also der Minor m_{in} der Matrix (a_{ij})):

$$\sum_{p_i}\mathrm{sign}(p_i)\cdot a_{p_i(1)1}\cdot\ldots\cdot a_{p_i(n-1),n-1}=(-1)^{n-i}\cdot m_{in}=(-1)^{n+i}\cdot m_{in}.$$

Insgesamt erhält man also die „Entwicklung nach der n-ten Spalte“

$$\det(A)=\sum_{i=1}^{n}(-1)^{i+n}\cdot a_{in}\cdot m_{in}.$$

Will man nicht nach der n-ten Spalte sondern nach der j-ten Spalte entwickeln, lösche man in der Ausgangsmatrix die j-te Spalte, lasse die Spalten $j+1$ bis n nach links rücken und stecke die ursprünglich j-te Spalte in die n-te Spalte (dies entspricht den $n-j$ Spaltenvertauschungen $j\leftrightarrow j+1$, dann $j+1\leftrightarrow j+2$ usw. bis $n-1\leftrightarrow n$. Entwickle nach der neuen n-ten Spalte (dies ist die alte j-te Spalte). Nach den obigen Überlegungen gilt

$$\det(A')=\sum_{i=1}^{n}(-1)^{i+n}\cdot a'_{in}\cdot m'_{in},$$

wo $A'=(a'_{ij})$ durch die obigen Spaltenvertauschungen aus A entsteht. Speziell gilt also $a'_{in}=a_{ij}$. Für die entsprechenden Minoren gilt $m'_{in}=m_{ij}$. Nach $n-j$ Spaltenvertauschungen gilt

$$\det(A')=(-1)^{n-j}\cdot\det(A)=(-1)^{j-n}\cdot\det(A),$$

also

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot m_{ij}.$$

Damit haben wir die „Entwicklung nach der j-ten Spalte“.

Da nach Satz 12.48 die Rolle von Zeilen und Spalten gegeneinander ausgetauscht werden können, kann man statt nach Spalten auch nach Zeilen entwickeln.

Q.E.D.

Beispiel 12.52: Wir entwickeln eine (2×2)-Determinante nach der ersten Spalte:

$$\det\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right) = a_{11} \cdot \det((a_{22})) - a_{21} \cdot \det((a_{12})) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.$$

Entwicklung nach der zweiten Zeile:

$$\det\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right) = -a_{21} \cdot \det((a_{12})) + a_{22} \cdot \det((a_{11})) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.$$

Beispiel 12.53: Wir entwickeln eine (3×3)-Determinante nach der zweiten Spalte:

$$\det\left(\begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \\ 0 & 4 & 5 \end{pmatrix}\right) =$$

$$-2 \cdot \det\left(\begin{pmatrix} \cancel{1} & \cancel{2} & \cancel{1} \\ 0 & \cancel{3} & 2 \\ 0 & \cancel{4} & 5 \end{pmatrix}\right) + 3 \cdot \det\left(\begin{pmatrix} 1 & \cancel{2} & 1 \\ \cancel{0} & \cancel{3} & \cancel{2} \\ 0 & \cancel{4} & 5 \end{pmatrix}\right) - 4 \cdot \det\left(\begin{pmatrix} 1 & \cancel{2} & 1 \\ 0 & \cancel{3} & 2 \\ \cancel{0} & \cancel{4} & \cancel{5} \end{pmatrix}\right)$$

$$= -2 \cdot \det\left(\begin{pmatrix} 0 & 2 \\ 0 & 5 \end{pmatrix}\right) + 3 \cdot \det\left(\begin{pmatrix} 1 & 1 \\ 0 & 5 \end{pmatrix}\right) - 4 \cdot \det\left(\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\right)$$

$$= -2 \cdot 0 + 3 \cdot 5 - 4 \cdot 2 = 7.$$

Entwicklung nach der ersten Spalte:

$$\det\left(\begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \\ 0 & 4 & 5 \end{pmatrix}\right) = 1 \cdot \det\left(\begin{pmatrix} \cancel{1} & \cancel{2} & \cancel{1} \\ \cancel{0} & 3 & 2 \\ \cancel{0} & 4 & 5 \end{pmatrix}\right) - 0 \cdot (...) + 0 \cdot (...) = 1 \cdot (15 - 8) = 7.$$

Beobachtung: man spart sich Rechenarbeit, wenn man nach Zeilen oder Spalten mit möglichst vielen Nullen entwickelt, da die entsprechenden Minoren gar nicht ausgerechnet werden brauchen (sie werden sowieso mit 0 multipliziert und tragen zur Determinante nichts bei).

Es ist rechentechnisch günstig, nach Zeilen oder Spalten mit möglichst vielen Nullen zu entwickeln.

↓28.1.04

Mit Folgerung 12.47.b) hat man eine Möglichkeit, Nullen in der Matrix zu erzeugen, indem man geeignete Vielfache einer Zeile/Spalte von einer anderen Zeile/Spalte abzieht. Hier ein Beispiel. In

$$\det\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}\right)$$

subtrahieren wir zunächst die zweite Zeile von der dritten:

$$\det\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}\right) = \det\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 3 & 4 \end{pmatrix}\right).$$

Wir subtrahieren die erste Zeile von der zweiten:

$$\det\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 3 & 4 \end{pmatrix}\right) = \det\left(\begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 4 \end{pmatrix}\right).$$

Wir subtrahieren die erste Spalte von der zweiten:

$$\det\left(\begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 4 \end{pmatrix}\right) = \det\left(\begin{pmatrix} 1 & 1 & 3 \\ 3 & 0 & 3 \\ 3 & 0 & 4 \end{pmatrix}\right).$$

Entwicklung nach der zweiten Spalte liefert:

$$\det\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}\right) = \det\left(\begin{pmatrix} 1 & 1 & 3 \\ 3 & 0 & 3 \\ 3 & 0 & 4 \end{pmatrix}\right) = -1 \cdot \det\left(\begin{pmatrix} 3 & 3 \\ 3 & 4 \end{pmatrix}\right) = -3.$$

Definition 12.54: (Dreiecksmatrizen)

Eine Matrix der Form

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$$

heißt **„obere Dreiecksmatrix“**. *Eine Matrix der Form*

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix}$$

heißt **„untere Dreiecksmatrix“**.

Satz 12.55:

Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente.

Beweis: Entwickle bei einer oberen Dreiecksmatrix die Determinanten rekursiv jeweils nach der ersten Spalte:

$$\begin{aligned}
&\det\left(\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}\right) \\
&= a_{11} \cdot \det\left(\begin{pmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}\right) \\
&= a_{11} \cdot a_{22} \cdot \det\left(\begin{pmatrix} a_{33} & a_{34} & \dots & a_{3n} \\ 0 & a_{44} & \dots & a_{4n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}\right)
\end{aligned}$$

usw. Bei unteren Dreiecksmatrizen entwickelt man analog rekursiv nach der ersten Zeile.

Q.E.D.

Nun ein fundamentaler Satz, der das eindeutige Indiz dafür ist, dass der Determinantenbegriff vernünftig mit Matrizen und ihrer Multiplikation zusammenspielt:

Satz 12.56: (Determinante eines Matrixprodukts)

Für $A, B \in K^{n\times n}$ gilt

$$\det(A\,B) = \det(A) \cdot \det(B).$$

Beweis: Wir benutzen die Spaltenzerlegung $A = [\vec{a}_1, \dots, \vec{a}_n]$. Mit

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

folgt die Spaltenzerlegung

$$AB = \left[\sum_{k_1=1}^{n} \vec{a}_{k_1} \cdot b_{k_1 1}, \sum_{k_2=1}^{n} \vec{a}_{k_2} \cdot b_{k_2 2}, \dots, \sum_{k_n=1}^{n} \vec{a}_{k_n} \cdot b_{k_n n}\right].$$

Da die Determinante linear in jeder Spalte ist, folgt

$$\det(AB) = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \det([\vec{a}_{k_1}, \vec{a}_{k_2}, \ldots, \vec{a}_{k_n}]) \cdot b_{k_1 1} \cdot b_{k_2 2} \cdot \ldots \cdot b_{k_n n}.$$

In dieser n-fachen Summe (mit n^n Termen) sind nur diejenigen Terme vorhanden, in denen die Summationsindizes $k_1, \ldots, k_n$ alle verschieden sind (bei den anderen Termen taucht in der Determinante der Spalten von A mindestens eine Spalte mehrfach auf und liefert 0). Die Mehrfachsumme reduziert sich damit auf die $n!$ Summanden, in denen $k_1, \ldots, k_n$ eine Permutation von $\{1, \ldots, n\}$ darstellt:

$$\det(AB) = \sum_{p} \det([\vec{a}_{p(1)}, \vec{a}_{p(2)}, \ldots, \vec{a}_{p(n)}]) \cdot b_{p(1)1} \cdot b_{p(2)2} \cdot \ldots \cdot b_{p(n)n}.$$

Durch Vertauschen der Spalten in der Determinante der Spalten von A erhält man

$$\det([\vec{a}_{p(1)}, \vec{a}_{p(2)}, \ldots, \vec{a}_{p(n)}]) = \mathrm{sign}(p) \cdot \det([\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]) = \mathrm{sign}(p) \cdot \det(A),$$

so dass sich ergibt:

$$\begin{aligned} \det(AB) &= \det(A) \cdot \sum_{p} \mathrm{sign}(p) \cdot b_{p(1)1} \cdot b_{p(2)2} \cdot \ldots \cdot b_{p(n)n} \\ &= \det(A) \cdot \det(B^T) = \det(A) \cdot \det(B). \end{aligned}$$

Q.E.D.

Definition 12.57: (Adjunktenmatrix)

Einer Matrix $A \in K^{n \times n}$ mit den Minoren m_{ij} gemäß Definition 12.50 wird die **“Adjunktenmatrix“** *$A_{adj} \in K^{n \times n}$ zugeordnet:*

$$(A_{adj})_{ij} = (-1)^{i+j} \cdot m_{ji}.$$

Satz 12.58:

Für jedes $A \in K^{n \times n}$ mit der Adjunktenmatrix A_{adj} gilt:

$$A_{adj} A = A\, A_{adj} = \det(A) \cdot \mathbb{I}.$$

Beweis: Sei $A = (a_{ij}) = [\vec{a}_1, \ldots, \vec{a}_n]$, $A_{adj} = (b_{ij}) = ((-1)^{i+j} \cdot m_{ji})$. Es gilt

$$(A_{adj}A)_{ij} = \sum_{k=1}^{n} b_{ik} \cdot a_{kj} = \sum_{k=1}^{n} (-1)^{i+k} \cdot m_{ki} \cdot a_{kj}.$$

Diese Summe ist die Entwicklung der Determinante

$$\det([\vec{a}_1, \ldots, \overset{\overset{i}{\downarrow}}{\vec{b}}, \ldots, \vec{a}_n])$$

mit $\vec{b} = \vec{a}_j$ nach der i-ten Zeile. Für $i \neq j$ taucht in dieser Determinante zweimal die Spalte $\vec{a}_j$ auf und liefert 0. Für $i = j$ ergibt sich $\det([\vec{a}_1, \ldots, \vec{a}_n]) = \det(A)$:

$$(A_{adj}A)_{ij} = \begin{cases} \det(A) & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Um zu zeigen, dass auch $A\,A_{adj} = \det(A) \cdot \mathbb{1}$ gilt, betrachte man $A_{adj}^T A^T$. Nach Zerlegung von A^T in Spalten ergibt sich analog zu oben $A_{adj}^T A^T = \det(A^T) \cdot \mathbb{1}$. Durch Transposition folgt dann

$$(A_{adj}^T A^T)^T = AA_{adj} = (\det(A^T) \cdot \mathbb{1})^T = \det(A) \cdot \mathbb{1}.$$

Q.E.D.

Folgerung 12.59:

Eine quadratische Matrix A ist genau dann invertierbar, wenn $\det(A) \neq 0$ gilt. Mit der Adjunktenmatrix A_{adj} gilt:

$$A^{-1} = \frac{1}{\det(A)} \cdot A_{adj}.$$

Satz 12.60: (Die Cramersche Regel)

Wir betrachten das System

$$\begin{array}{ccccccc} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1, \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} \cdot x_1 & + & \cdots & + & a_{nn} \cdot x_n & = & b_n \end{array}$$

von n linearen Gleichungen für die n Unbekannten $x_1, \ldots, x_n$. Mit $A = (a_{ij}) \in K^{n \times n}$, $\vec{x} = (x_1, \ldots, x_n)^T$, $\vec{b} = (b_1, \ldots, b_n)^T$ ergibt sich die kompakte Schreibweise

$$A\vec{x} = \vec{b}.$$

Es gibt genau dann eine eindeutige Lösung $\vec{x} = A^{-1}\vec{b}$, wenn $\det(A) \neq 0$ gilt. Die Komponenten der Lösung sind gegeben durch

$$x_i \;=\; \frac{\begin{array}{c} \qquad\qquad i \\ \qquad\qquad \downarrow \\ \det\left(\begin{pmatrix} a_{11} & \ldots & b_1 & \ldots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \ldots & b_n & \ldots & a_{nn} \end{pmatrix}\right) \end{array}}{\det\left(\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}\right)}$$

mit $i = 1, \ldots, n$.

Beweis: Wir benutzen $(A^{-1})_{ij} = \frac{1}{\det(A)} \cdot (-1)^{i+j} \cdot m_{ji}$ mit den Minoren m_{ij} von A. Es folgt

$$x_i = \left(A^{-1}\vec{b}\right)_i = \sum_{j=1}^{n} (A^{-1})_{ij} \cdot b_j = \frac{1}{\det(A)} \cdot \sum_{j=1}^{n} (-1)^{i+j} \cdot m_{ji} \cdot b_j .$$

Hierbei ist $\sum_{j=1}^{n} (-1)^{i+j} \cdot m_{ji} \cdot b_j$ die Entwicklung von

$$\det([\vec{a}_1, \ldots, \vec{b}, \ldots, \vec{a}_n])$$

nach der i-ten Spalte, wo $\vec{a}_1, \ldots, \vec{a}_n$ die Spalten von A sind und die Spalte $\vec{a}_i$ durch die Spalte $\vec{b}$ ersetzt wurde.

Q.E.D.

12.5 Lineare Gleichungssysteme

Im letzten Abschnitt haben wir ein System $A\vec{x} = \vec{b}$ von n linearen Gleichungen für n Unbekannte $\vec{x} = (x_1, \ldots, x_n)^T$ gelöst, vorausgesetzt dass die Koeffizientenmatrix A invertierbar ist:

$$\vec{x} = A^{-1}\vec{b}.$$

Hierbei wurden explizite Lösungsformeln gefunden: die Inverse A^{-1} läßt sich über die Adjunktenmatrix von A mittels Determinanten bestimmen, die Lösung $\vec{x}$ ist direkt über die Cramersche Regel 12.60 mittels Determinanten auszudrücken. Dies ist jedoch noch nicht die endgültige Antwort auf die Aufgabenstellung „Lösung linearer Gleichungssysteme“. Es verbleiben die Probleme:

- Was passiert bei nicht-invertierbaren Koeffizientenmatrizen? Was ist, wenn n Gleichungen für m Unbekannte (mit $n \neq m$) zu lösen sind?

- Die Determinantenformeln sind uneffektiv: die bisher vorgestellte Berechnung von Determinanten (z.B. rekursiv über Laplace-Entwicklung nach Zeilen oder Spalten) braucht viel zu viele Rechenschritte. (Eine $(n \times n)$-Determinante besteht aus einer Summe von $n!$ Termen. Die Stirling-Näherung 9.58 sagt z.B. für $n = 100$ etwa $n! \approx 9.3 \times 10^{157}$ Summanden für die Determinante voraus).

Das Determinantenkalkül des letzten Abschnitts ist in der Tat mehr als allgemeine Strukturtheorie anzusehen, deren wesentliches Ergebnis die Aussage

$A\vec{x} = \vec{b}$ mit invertierbarem A hat genau dann eine eindeutige Lösung $\vec{x} = A^{-1}\vec{b}$, wenn $\det(A) \neq 0$ gilt.

In der Praxis geht man anders vor. Gleichungssysteme mit z.B. $n = 100$ Unbekannten lassen sich in der Tat in Sekundenbruchteilen lösen, wobei Determinanten keinerlei Rolle spielen.

12.5.1 Quadratische Dreieckssysteme

Ist die Koeffizientenmatrix A in einem linearen Gleichungssystem $A\vec{x} = \vec{b}$ von Dreiecksform, so ist die Berechnung der Lösung sehr einfach (und mit wenig Rechenaufwand verbunden):

Satz 12.61: (Rücksubstitution)
Ein lineares System $A\vec{x} = \vec{b}$

$$\begin{array}{lllllllll} a_{11} \cdot x_1 & & & & & & & = & b_1, \\ a_{21} \cdot x_1 & + & a_{22} \cdot x_2 & & & & & = & b_2, \\ \vdots & & \vdots & & \ddots & & & & \vdots \\ a_{n1} \cdot x_1 & + & a_{n2} \cdot x_2 & + & \cdots & + & a_{nn} \cdot x_n & = & b_n \end{array}$$

mit einer invertierbaren unteren $(n \times n)$-Dreiecksmatrix $A = (a_{ij})$ wird durch **„Rücksubstitution“**

$$\boxed{\begin{array}{l} x_1 \ := \ b_1 \,/\, a_{11} \ ; \\ \textbf{for } i := 2 \textbf{ to } n \textbf{ do } \ x_i \ := \ \left(b_i - \sum_{j=1}^{i-1} a_{ij} \cdot x_j \right) / \, a_{ii} \ ; \end{array}}$$

mit

$$1 + \sum_{i=2}^{n} i \ = \ \frac{n(n+1)}{2} \ = \ \frac{n^2}{2} \left(1 + O\left(\frac{1}{n}\right) \right)$$

Multiplikationen/Divisionen (und ähnlich vielen Additionen) gelöst.

Bei einem System

$$\begin{array}{ccccccccc} a_{11}\cdot x_1 & + & \cdots & + & a_{1,n-1}\cdot x_{n-1} & + & a_{1n}\cdot x_n & = & b_1 \\ & & \ddots & & \vdots & & \vdots & & \vdots \\ & & & & a_{n-1,n-1}\cdot x_{n-1} & + & a_{n-1,n}\cdot x_n & = & b_{n-1} \\ & & & & & & a_{nn}\cdot x_n & = & b_n \end{array}$$

mit invertierbarer oberer $(n \times n)$-Dreiecksmatrix $A = (a_{ij})$ lautet die Rücksubstitution:

$$\boxed{\begin{array}{l} x_n := b_n \,/\, a_{nn}\ ; \\ \textbf{for } i := n-1 \textbf{ downto } 1 \textbf{ do } \ x_i := \left(b_i - \displaystyle\sum_{j=i+1}^{n} a_{ij}\cdot x_j \right) / \, a_{ii}\ ; \end{array}}$$

Beweis: Im Fall eines unteren Dreieckssystems liefert die erste Gleichung sofort $x_1 = b_1/a_{11}$. Eingesetzt in die zweite Gleichung erhält man $a_{22}\cdot x_2 = b_2 - a_{21}\cdot x_1$, wodurch sich sofort x_2 ergibt. Allgemein: hat man $x_1, \ldots, x_{i-1}$ bereits bestimmt, liefert die i-te Gleichung

$$a_{ii}\cdot x_i = b_i - \sum_{j=1}^{i-1} a_{ij}\cdot x_j$$

den Wert für x_i. Beachte: die Koeffizientenmatrix A ist genau dann invertierbar, wenn keines der Diagonalelmente verschwindet, man kann also durch die a_{ii} teilen!

Bei einem oberen Dreieckssystemen startet man mit der letzten Gleichung, um x_n zu bestimmen, setzt diesen Wert dann in die vorletzte Gleichung ein, aus der dann x_{n-1} bestimmbar ist usw.

Q.E.D.

Bemerkung 12.62: *Der Rechenaufwand von etwa $n^2/2$ Multiplikationen entspricht exakt dem Aufwand, eine Multiplikation einer $(n \times n)$-Dreicksmatrix mit einem Vektor mit n Komponenten durchzuführen. Die Lösung von $A\vec{x} = \vec{b}$ bekommt man bei Dreiecksmatrizen also genauso schnell wie die Multiplikation $A\vec{x}$.*

29.1.04↓

12.5.2 Der Gauß-Algorithmus

Nun wird der Standardalgorithmus zur Lösung linearer Gleichungssysteme vorgestellt, der

- die allgemeine Lösung von n Gleichungen mit m Unbekannten liefert (auch für $n \neq m$),
- mit einer akzeptablen Anzahl von Rechenoperationen auskommt. Faustregel: für n Gleichungen mit n Unbekannten braucht man $O(n^3)$ Rechenoperationen.

Es handelt sich dabei um den **„Gauss-Algorithmus“**, der je nach Variante als „**Gauss-**“, „**Gauss-Jordan-**“, „**Entzerrungsalgorithmus**“ oder auch als „LR**-Zerlegung**“ bezeichnet wird.

Gegeben seien n lineare Gleichungen für m Unbekannte

$$\begin{array}{ccccccc} a_{11}\cdot x_1 & + & \cdots & + & a_{nm}\cdot x_m & = & b_1, \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}\cdot x_1 & + & \cdots & + & a_{nm}\cdot x_m & = & b_n, \end{array}$$

die wir bequemer als $(n \times (m+1))$-Matrix

$$[A \mid \vec{b}\,] = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1m} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \dots & a_{nm} & b_n \end{array}\right)$$

kodieren (die **„erweiterte Koeffizientenmatrix“**). Jede Zeile

$$\left(\begin{array}{ccc|c} a_{i1} & \dots & a_{im} & b_i \end{array}\right)$$

ist dabei jeweils als Gleichung zu lesen:

$$a_{i1}\cdot x_1 + \ldots + a_{im}\cdot x_m = b_i.$$

Elementaroperationen 12.63:

Wir wollen das Gleichungssystem so umschreiben (vereinfachen), dass die Lösung möglichst direkt abgelesen werden kann. Dabei sollen lediglich folgende **Elementaroperationen** *eingesetzt werden:*

i) Vertauschen von Gleichungen (= Zeilen von $[A \mid \vec{b}\,]$)

ii) Multiplikation einer Gleichung (= Zeile von $[A \mid \vec{b}\,]$) mit einer Zahl $\neq 0$, *iii)* *Addieren eines Vielfachen einer Gleichung (= Zeile von $[A \mid \vec{b}\,]$) zu einer anderen Gleichung (= Zeile von $[A \mid \vec{b}\,]$).*

Diese Operationen lassen sich als Matrixmultiplikation mit **„Elementarmatrizen“** *interpretieren (nicht angezeigte Einträge sind 0):*

i) *Die Multiplikation $A \to EA$ mit*

$$E = \begin{array}{cc} \begin{array}{ccccccccccc} &&& i_1 &&&& i_2 &&& \\ &&& \downarrow &&&& \downarrow &&& \end{array} & \\ \begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& 0 &&&& 1 &&& \\ &&&& 1 &&&&&& \\ &&&&& \ddots &&&&& \\ &&&&&& 1 &&&& \\ &&& 1 &&&& 0 &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix} & \begin{array}{c} \\ \\ \\ \leftarrow i_1 \\ \\ \\ \\ \leftarrow i_2 \\ \\ \\ \\ \end{array} \end{array}$$

vertauscht die Zeilen i_1 und i_2 in A. Es gilt $\det(E) = -1$ für $i_1 \neq i_2$ bzw. $\det(E) = 1$ für $i_1 = i_2$.

ii) *Die Multiplikation $A \to EA$ mit*

$$E = \begin{array}{cc} \begin{array}{ccccccc} &&& i &&& \\ &&& \downarrow &&& \end{array} & \\ \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} & \begin{array}{c} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{array} \end{array}$$

multipliziert die i-te Zeile von A mit dem Faktor λ. Es gilt $\det(E) = \lambda$.

iii) *Die Multiplikation $A \to EA$ mit*

$$E = \begin{array}{cc} \begin{array}{cccccc} && i_1 &&& \\ && \downarrow &&& \end{array} & \\ \begin{pmatrix} 1 &&&&& \\ & 1 &&&& \\ && \ddots &&& \\ &&& \ddots && \\ && \lambda && 1 & \\ &&&&& 1 \end{pmatrix} & \begin{array}{c} \\ \\ \\ \\ \leftarrow i_2 \\ \\ \end{array} \end{array}$$

addiert das λ-fache der Zeile i_1 zur Zeile i_2. Es gilt $\det(E) = 1$.

Mit der Forderung $\lambda \neq 0$ in ii) haben alle Elementarmatrizen nichtverschwindende Determinanten, sind also invertierbar. Die Anwendung dieser Matrizen entspricht also einer Äquivalenztransformation, die die Lösung des Gleichungssystems nicht verändert.

Mit diesen Operationen soll das Ausgangsschema

$$[A \mid \vec{b}\,] = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1m} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \dots & a_{nm} & b_n \end{array}\right)$$

in eine neue (einfachere) Form $[A' \mid \vec{b}'\,]$ gebracht werden. Welche Struktur sollte die vereinfachte Koeffizientenmatrix haben? Wenn es uns gelingt, das System $A\vec{x} = \vec{b}$ in ein äquivalentes System $A'\vec{x} = \vec{b}'$ mit $A' = \mathbb{1}$ umzuschreiben, dann sind wir fertig. Nach Abschnitt 12.5.1 reicht es jedoch, als Endform A' = Dreiecksmatrix anzusteuern, da das verbleibende Dreieckssystem $A'\vec{x} = \vec{b}'$ schnell per Rücksubstitution 12.61 gelöst werden kann.

Hier nun der **Gauß-Algorithmus** zur Lösung von n Gleichungen

$$\begin{array}{ccccccccc} a_{11} \cdot x_1 & + & a_{12} \cdot x_2 & + & \cdots & + & a_{1m} \cdot x_m & = & b_1 \\ a_{21} \cdot x_1 & + & a_{22} \cdot x_2 & + & \cdots & + & a_{2m} \cdot x_m & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{n1} \cdot x_1 & + & a_{n2} \cdot x_2 & + & \cdots & + & a_{nm} \cdot x_m & = & b_n \end{array}$$

für m Unbekannte:

1.ter Eliminationsschritt: ziehe geeignete Vielfache der ersten Zeile von den anderen ab, um Nullen unter a_{11} zu erzeugen:

$$\begin{array}{cc|cccccccc} a_{11} \cdot x_1 & + & a_{12} \cdot x_2 & + & \cdots & + & a_{1m} \cdot x_m & = & b_1 \\ \hline a_{21}/a_{11} & & a'_{22} \cdot x_2 & + & \cdots & + & a'_{2m} \cdot x_m & = & b'_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{n1}/a_{11} & & a'_{n2} \cdot x_2 & + & \cdots & + & a'_{nm} \cdot x_m & = & b'_n \end{array}$$

mit

$$a'_{ik} = a_{ik} - \frac{a_{i1}}{a_{11}} \cdot a_{1k}\,, \quad b'_i = b_i - \frac{a_{i1}}{a_{11}} \cdot b_1\,, \quad i = 2, \dots, n\,, k = 2, \dots, m\,.$$

Bezeichnung: a_{i1}/a_{11} = **"Eliminationsfaktoren"**. Im Prinzip stehen unter a_{11} Nullen, d.h., diese Speicherplätze sind frei geworden. Diese werden mit den Eliminationsfaktoren belegt, was nichts kostet, sich aber im Abschnitt 12.6.2

als äußerst wertvoll herausstellen wird.

2.ter Eliminationsschritt: erzeuge Nullen unter a'_{22}:

$$
\begin{array}{ccccccccccc}
a_{11}\cdot x_1 & + & a_{12}\cdot x_2 & + & a_{13}\cdot x_3 & + & \cdots & + & a_{1n}\cdot x_m & = & b_1 \\
\cline{1-1}
a_{21}/a_{11} & \multicolumn{1}{|c}{} & a'_{22}\cdot x_2 & + & a'_{23}\cdot x_3 & + & \cdots & + & a'_{2n}\cdot x_m & = & b'_2 \\
\cline{3-3}
\vdots & & a'_{32}/a'_{22} & \multicolumn{1}{|c}{} & a''_{33}\cdot x_3 & + & \cdots & + & a''_{3m}\cdot x_m & = & b''_3 \\
\vdots & & \vdots & \multicolumn{1}{|c}{} & \vdots & & \ddots & & \vdots & & \vdots \\
a_{n1}/a_{11} & & a'_{n2}/a'_{22} & \multicolumn{1}{|c}{} & a''_{n3}\cdot x_3 & + & \cdots & + & a''_{nm}\cdot x_m & = & b''_n
\end{array}
$$

mit

$$
a''_{ik} = a'_{ik} - \frac{a'_{i2}}{a'_{22}} \cdot a'_{2k}\ , \quad b''_i = b'_i - \frac{a'_{i2}}{a'_{22}} \cdot b'_2\ , \quad i = 3, \ldots, n\,,\ k = 3, \ldots, m\,.
$$

j**.ter Eliminationsschritt:** ziehe Vielfaches der j.ten Zeile von den restlichen Zeilen ab, um unter dem j.ten Diagonalelement Nullen zu erzeugen.

Einzig mögliches Problem: das j.te Diagonalelement verschwindet.

1. Fall: Darunter stehen nur Nullen. Es ist nichts mehr zu tun, gehe zum $(j+1)$.ten Schritt. Dieser Fall tritt genau dann irgendwann ein, wenn A nicht invertierbar ist.

2. Fall: Es gibt mindestens einen nichtverschwindenden Eintrag unter dem j.ten Diagonalelement. Wähle irgendeinen dieser Einträge (das **"Pivotelement"**), bringe es in die j.te Zeile, indem die gesamte Pivotzeile (inklusive der abzuspeichernden Eliminationsfaktoren!) mit der j.ten Zeile ausgetauscht wird. Dann erzeuge die gewünschten Nullen.

Algorithmus 12.64: (Gauß-Elimination = LR-Zerlegung, siehe Satz 12.75)

Erweitere $A = (a_{ij}) \in K^{n\times m}$ zum $n \times (m+1)$-Schema:

$$
\left(\begin{array}{ccc|c}
a_{11} & \ldots & a_{1m} & a_{1,m+1} \\
\vdots & \ddots & \vdots & \vdots \\
a_{n1} & \ldots & a_{nm} & a_{n,m+1}
\end{array}\right) \quad \textit{mit } a_{i,n+1} = b_i, \quad i = 1, \ldots, n\ .
$$

for $j := 1$ **to** $\min(m, n)$ **do begin**

  falls $a_{jj} = a_{j+1,j} = \cdots = a_{nj} = 0$**, so gehe zum** $(j+1)$**.ten Schritt** ;

  $p :=$ **Pivotindex(**j**)**$\in \{j, j+1, \ldots, n\}$ **;**

  Vertauschung der Zeilen j **und** p **im** $n \times (m+1)$**-Schema** ;

  for $i := j+1$ **to** n **do begin**

    $a_{ij} := a_{ij}/a_{jj}$;   (∗ **Abspeichern der Eliminationsfaktoren** ∗)

    for $k := j+1$ **to** $m+1$ **do** $a_{ik} := a_{ik} - a_{ij}\, a_{jk}$;

  end ;

end ;

Hierbei ist `Pivotindex(j)` der Aufruf einer Prozedur, welche unter den Werten $\{a_{jj}, a_{j+1,j}, \ldots, a_{nj}\}$ ein Element sucht, das nicht verschwindet, und den entsprechenden Zeilenindex p zurückgibt. Im Prinzip kann irgendein nichtverschwindendes Pivotelement gewählt werden. Aus Gründen der numerischen Stabilität (minimiere den Einfluss von Rundungsfehlern bei Gleitpunktarithmetik) wählt man meist $p =$`Pivotindex`(j) so, dass

$$|a_{pj}| = \max(|a_{jj}|, |a_{j+1,j}|, \ldots, |a_{nj}|)$$

gilt (d.h., das Pivotelement a_{pj} ist das betragsgrößte Element im unteren Teil der j-ten Spalte).

Resultat: zuletzt ist das zu $A\vec{x} = \vec{b}$ äquivalente obere Dreieckssystem $R\vec{x} = \vec{c}$

$$\left(\begin{array}{ccc|c} r_{11} & \cdots & r_{1n} & c_1 \\ & \ddots & \vdots & \vdots \\ * & & r_{nn} & c_n \end{array}\right) \qquad \text{(für } n = m\text{)}$$

bzw.

$$\left(\begin{array}{ccccc|c} r_{11} & \cdots & r_{1n} & \ldots & r_{1m} & c_1 \\ & \ddots & \vdots & \ddots & \vdots & \vdots \\ * & & r_{nn} & \ldots & r_{nm} & c_n \end{array}\right) \qquad \text{(für } n < m\text{)}$$

bzw.

$$\left(\begin{array}{ccc|c} r_{11} & \cdots & r_{1m} & c_1 \\ * & \ddots & \vdots & \vdots \\ \vdots & \ddots & r_{mm} & c_m \\ \vdots & \ddots & * & c_{m+1} \\ \vdots & \ddots & \vdots & \vdots \\ * & \ddots & * & c_n \end{array}\right) \qquad \text{(für } n > m\text{)}$$

im $(n \times (m+1))$-Schema gespeichert. Mit $*$ sind dabei die abgespeicherten Eliminationsfaktoren angedeutet, die erst im Abschnitt 12.6.2 von Bedeutung sein werden.

Rechenaufwand zur Erzeugung des oberen Dreieckssystems:

$$\sum_{j=1}^{\min(m,n)} \sum_{i=j+1}^{n} \Big(1 + \sum_{k=j+1}^{m+1} 1\Big) =$$

$$\begin{cases} \dfrac{n \cdot (n-1) \cdot (3 \cdot m - n + 5)}{6} & \text{für } n \leq m \\ \dfrac{m \cdot (3 \cdot m \cdot n - m \cdot n - 6 \cdot m + 9 \cdot n - 5)}{6} & \text{für } n \geq m \end{cases}$$

Multiplikationen (und ähnlich viele Additionen). Für den in der Praxis wichtigsten Fall $m = n$ sind dies

$$\frac{n\,(n-1)\,(2 \cdot n + 5)}{6} = \frac{n^3}{3} \cdot \Big(1 + O\Big(\frac{1}{n}\Big)\Big)$$

Multiplikationen.

Bemerkung 12.65: *Die Umformung $A\vec{x} = \vec{b} \Longleftrightarrow R\vec{x} = \vec{c}$ funktioniert immer, selbst wenn A nicht invertierbar ist (wobei dann auch R nicht invertierbar ist).*

Bemerkung 12.66: *Wir diskutieren den Fall, dass die Anzahl n der Gleichungen nicht mit der Anzahl m der Unbekannten übereinstimmt:*

Fall $n < m$ (weniger Gleichungen als Unbekannte):

Das vom Gauß-Algorithmus gelieferte obere Dreieckssystem hat die Form

$$\left(\begin{array}{cccccc|c} r_{11} & \cdots & r_{1n} & \ldots & \ldots & r_{1m} & c_1 \\ & \ddots & \vdots & \ddots & \ddots & \vdots & \vdots \\ & & r_{nn} & \ldots & \ldots & r_{nm} & c_n \end{array}\right).$$

Durch Hinzufügen von $m - n$ triviale Gleichungen $0 = 0$

$$\left(\begin{array}{cccccc|c} r_{11} & \cdots & r_{1n} & \ldots & \ldots & r_{1m} & c_1 \\ & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\ & & r_{nn} & \ldots & \ldots & r_{nm} & c_n \\ & & & 0 & \ldots & 0 & 0 \\ & & & & \ddots & \vdots & \vdots \\ & & & & & 0 & 0 \end{array}\right)$$

ergibt sich ein quadratisches oberes $(m \times m)$-System, dessen allgemeine Lösung im nächsten Abschnitt 12.5.3 beschrieben wird.

Fall $n > m$ (mehr Gleichungen als Unbekannte)*:*
Das vom Gauß-Algorithmus gelieferte obere Dreieckssystem hat die Form

$$\left(\begin{array}{ccc|c} r_{11} & \cdots & r_{1m} & c_1 \\ & \ddots & \vdots & \vdots \\ & & r_{mm} & c_m \\ & & & c_{m+1} \\ & & & \vdots \\ & & & c_n \end{array}\right).$$

Ist irgendeiner der Werte $c_{m+1}, \ldots, c_n$ ungleich 0, ist das Gleichungssystem nicht lösbar (dort stehen die Gleichungen $0 = c_p$ mit $p = m+1, \ldots n$). Sei nun $c_{m+1} = \ldots = c_n = 0$. Man wirft diese trivialen Gleichungen fort und erhält ein $(m \times m)$-Gleichungsystem

$$\left(\begin{array}{ccc|c} r_{11} & \cdots & r_{1m} & c_1 \\ & \ddots & \vdots & \vdots \\ & & r_{mm} & c_m \end{array}\right),$$

dessen allgemeine Lösung im nächsten Abschnitt 12.5.3 beschrieben wird. Es sollte im Fall $n > m$ erwähnt werden, dass man in der Praxis nicht wirklich den Gauß-Algorithmus verwendet, da man wegen der Rundungsfehler bei Gleitpunktarithmetik die Entscheidung, ob $c_{m+1} = \ldots = c_n = 0$ gilt, nicht fundiert treffen kann (es stehen dort kleine Zahlenwerte, die eventuell reine Rundungsartefakte sein können). Stattdessen benutzt man sogenannte „Singulärwertzerlegungen" zum Lösen überbestimmter Gleichungssysteme.

12.5.3 Die allgemeine Lösung des Dreieckssystems

$\boxed{\downarrow 3.2.04}$

Der Gauß-Algorithmus hat aus den ursprünglichen n Gleichungen $A\vec{x} = \vec{b}$ für m Unbekannte ein äquivalentes oberes Dreieckssystem $R\vec{x} = \vec{c}$ gemacht. Hierbei haben wir gemäß Bemerkung 12.66 für $n < m$ triviale Gleichungen hinzugenommen bzw. für $n > m$ triviale Gleichungen entfernt, um zu einem quadratischen System zu kommen (genausoviele Gleichungen wie Unbekannte). Wir haben nun also noch die Situtation eines quadratischen Systems ($m = n$) in oberer Dreiecksform

$$\begin{array}{ccccccc} r_{11} \cdot x_1 & + & \ldots & + & r_{1n} \cdot x_n & = & c_1 \\ & & \ddots & & \vdots & & \vdots \\ & & & & r_{nn} \cdot x_n & = & c_n \end{array}$$

zu diskutieren, wobei allerdings evtl. einige Diagonalelemente verschwinden können.

Fall 1, kein Diagonalelement verschwindet:
In diesem Fall gilt $\det(R) \neq 0$. Da R durch Anwendungen von invertierbaren Elementartransformationen 12.63 aus A entsteht, muss auch $\det(A) \neq 0$ gelten, d.h., wir lesen an den Diagonalelementen von R ab, dass das Ausgangssystem $A\vec{x} = \vec{b}$ wegen Invertierbarkeit von A eindeutig lösbar ist. Das vorliegende äquivalente Dreieckssystem $R\vec{x} = \vec{c}$ kann schnell durch Rücksubstitution 12.61 gelöst werden.

Fall 2, einige Diagonalelemente verschwinden („Entzerrungsalgorithmus“):
Wie oben folgt $\det(R) = 0$ genau dann, wenn $\det(A) = 0$ gilt. Wir lesen also an den verschwindenden Diagonalelementen von R ab, dass das Ausgangssystem $A\vec{x} = \vec{b}$ wegen Nichtinvertierbarkeit von A nicht eindeutig lösbar ist.

<u>**Reduktion auf Treppenform:**</u> Wir haben ein quadratisches oberes Dreieckssystem vor uns, in dem längs der Diagonalen einige Elemente verschwinden. Mit $*$ deuten wir nicht-verschwindende Elemente an, der rechte obere Teil des Schemas ist mit irgendwelchen Elementen besetzt:

$$
\left(\begin{array}{cccccccccc|c}
* & & & & & * & * & & & & * \\
 & \ddots & & & & & \ddots & \ddots & & & \vdots \\
 & & * & & & & & \ddots & \ddots & & \vdots \\
 & & & 0 & & & * & & \ddots & * & * \\
 & & & & * & & & & & * & * \\
 & & & & & * & & & & & * \\
 & & & & & & 0 & * & & & * \\
 & & & & & & & * & & & * \\
 & & & & & & & & \ddots & & * \\
 & & & & & & & & & \ddots & \vdots
\end{array}\right)
$$

Unser Ziel ist es zunächst, alle Zeilen mit 0 auf der Diagonalen zu Nullzeilen zu reduzieren. Dazu mache für jede Zeile mit einer 0 auf der Diagonale die folgenden Schritte. Sei i der Zeilenindex dieser Zeile.

- Ist die Zeile eine Nullzeile: tue gar nichts und gehe zur nächsten Zeile.
- Ist die i-te Zeile keine Nullzeile: Suche das erste nicht-verschwindende Element in dieser Zeile. Es habe den Spaltenindex j. Ist das j-te Diagonalelement des Schemas 0, so vertausche die Zeilen i und j. Ist das j-te

Diagonalelement des Schemas $\neq 0$, so ziehe ein geeignetes Vielfaches der j-ten Zeile von der i-ten Zeile ab, damit das j-te Element der i-ten Zeile verschwindet.

- Wiederhole die obigen Schritte für die i-te Zeile, bis diese eine Nullzeile ist.

Das Gleichungssystem hat nun **„obere Treppenform"**:

$$\left(\begin{array}{ccccccccc|c}
* & & & & * & * & & & & * \\
 & \ddots & & & & \ddots & \ddots & & & \vdots \\
 & & * & & & & \ddots & \ddots & & * \\
 & & & 0 & 0 & 0 & \dots & 0 & 0 & * \\
 & & & & * & * & \dots & \dots & * & * \\
 & & & & & 0 & 0 & \dots & 0 & * \\
 & & & & & & * & & & * \\
 & & & & & & & \ddots & & \vdots \\
 & & & & & & & & \ddots & \vdots
\end{array}\right) \tag{\#}$$

Rechts neben den Nullen auf der Diagonale können jetzt nur noch Nullen stehen. Steht in der letzten Spalte des Schemas (eine der Komponenten d_i der rechten Seite $\vec{d}$ von $R\vec{x} = \vec{d}$) rechts neben einer Nullzeile ein nicht-verschwindender Eintrag, so ist das Gleichungssystem insgesamt nicht lösbar: diese Gleichung besagt $0 = d_i \neq 0$.

Elimination „nach oben": Laufe von rechts unten nach links oben durch die Diagonale:

- Ist das Diagonalelement ungleich Null, mache es zu 1, indem die gesamte Zeile durch das Diagonalelement geteilt wird. Dann erzeuge über dem Diagonalelement Nullen, in dem geeignete Vielfache diese Zeile von den oberen Zeilen abgezogen werden.

Es verbleiben längs der Diagonalen nur noch Nullen und Einsen. Über den Einsen stehen nur noch Nullen, über den Nullen auf der Diagonale verbleiben irgendwelche Einträge:

$$
\left(\begin{array}{ccccccccccc|c}
1 & 0 & & * & & 0 & & * & & 0 & & d_1 \\
 & \ddots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
 & & 1 & * & & \vdots & & * & & \vdots & & d_{i_1-1} \\
 & & & 0 & & \vdots & & 0 & & \vdots & & 0 \\
 & & & & 1 & 0 & & * & & \vdots & & d_{i_1+1} \\
 & & & & & \ddots & & \vdots & & \vdots & & \vdots \\
 & & & & & & 1 & * & & \vdots & & d_{i_2-1} \\
 & & & & & & & 0 & & \vdots & & 0 \\
 & & & & & & & & 1 & 0 & & d_{i_2+1} \\
 & & & & & & & & & \ddots & & \vdots \\
 & & & & & & & & & & \ddots & \vdots
\end{array}\right). \qquad (\#\#)
$$

<u>Entzerrung:</u>

- Ersetze die Nullen auf der Diagonale durch -1.

Das Schema hat nun die Endform

$$
\left(\begin{array}{ccccccccccc|c}
1 & 0 & & * & & 0 & & * & & 0 & & d_1 \\
 & \ddots & & * & & \vdots & & * & & \vdots & & \vdots \\
 & & 1 & * & & \vdots & & * & & \vdots & & d_{i_1-1} \\
 & & & -1 & & \vdots & & 0 & & \vdots & & 0 \\
 & & & & 1 & 0 & & * & & \vdots & & d_{i_1+1} \\
 & & & & & \ddots & & * & & \vdots & & \vdots \\
 & & & & & & 1 & * & & \vdots & & d_{i_2-1} \\
 & & & & & & & -1 & & \vdots & & 0 \\
 & & & & & & & & 1 & 0 & & d_{i_2+1} \\
 & & & & & & & & & \ddots & & \vdots \\
 & & & & & & & & & & \ddots & \vdots
\end{array}\right), \quad (\#\#\#)
$$

an dem sich die allgemeine Lösung unmittelbar ablesen läßt:

$$\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_{i_1-1} \\ 0 \\ d_{i_1+1} \\ \vdots \\ d_{i_2-1} \\ 0 \\ d_{i_2+1} \\ \vdots \end{pmatrix} + \lambda_1 \cdot \begin{pmatrix} * \\ \vdots \\ * \\ \boxed{-1} \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ * \\ \vdots \\ * \\ \boxed{-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots$$

mit beliebigen freien Konstanten λ_1, λ_1 etc.

Definition 12.67: (Kernvektoren einer Matrix)

Ein Vektor $\vec{k} \in K^m$ einer Matrix $A \in K^{n \times m}$ mit $A\,\vec{k} = 0$ heißt **„Kernvektor“** *von A. Die Menge aller Kernvektoren bilden entweder die leere Menge oder einen linearen Unterraum im K^m (den* **„Kern“** *von A). Die Dimension des Kerns heißt* **„Defekt“** *der Matrix. Der* **„Rang“** *einer $(n \times m)$-Matrix ist*

$$rg(A) = \min(n, m) - \mathit{Defekt}(A).$$

Satz 12.68: (Die allgemeine Lösung linearer Systeme)

Die allgemeine Lösung eines lösbaren linearen Gleichungssystems $A\vec{x} = \vec{b}$ ist

$$\vec{x} = \vec{x}_0 + \sum_j \lambda_j \cdot \vec{k}_j,$$

wobei $\vec{x}_0$ irgendeine spezielle Lösung von $A\vec{x} = \vec{b}$ ist, $\vec{k}_1$, $\vec{k}_2$ etc. eine Basis des Kerns von A bilden und λ_1, λ_2 etc. beliebige freie Parameter sind.

Beweis: Sei $\vec{x}$ eine beliebige Lösung von $A\vec{x} = \vec{b}$, sei $\vec{x}_0$ eine spezielle Lösung. Die Differenz $\vec{k} = \vec{x} - \vec{x}_0$ erfüllt dann $A\vec{k} = A(\vec{x} - \vec{x}_0) = A\vec{x} - A\vec{x}_0 = \vec{b} - \vec{b} = 0$, ist also ein Kernvektor.

Q.E.D.

Jedes markierte Diagonalelement -1 in (###) entspricht einem Kernvektor der Ausgangsmatrix, der sich als die entsprechende Spalte des Endschemas ablesen läßt. Existieren mehrere solcher eingefügten Diagonalelemente -1, sind die entsprechenden Kernvektoren automatisch linear unabhängig, denn sie haben als untersten Eintrag jeweils -1 an unterschiedlichen Stellen:

Die Dimension des Kerns der Matrix ist durch die Anzahl der Diagonal--Nullen in (#) bzw. (##) gegeben.

Fazit:

Der Gauß-Algorithmus wie oben beschrieben liefert nach Entzerrung des oberen Gleichungssystems sowohl eine spezielle Lösung $\vec{x}_0 = (d_1, d_2, \dots)^T$ (mit 0 in den Komponenten, die den Kernindizes entsprechen) sowie einen Satz von linear unbhängigen Kernvektoren (mit -1 in den Komponenten, die den Kernindizes entsprechen). Diese Kernvektoren bilden eine Basis des Kerns von A. Wir haben damit die allgemeine Lösung des Ausgangssystems $A\vec{x} = \vec{b}$ gefunden.

Beispiel 12.69: Wir berechnen die allgemeine Lösung des linearen Gleichungssystems (4 Gleichungen für 5 Unbekannte):

$$\begin{pmatrix} 0 & 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 3 & 1 \\ 1 & 2 & 3 & 5 & 2 \\ 1 & 2 & 3 & 6 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 6 \\ 10 \\ 16 \\ 16 \end{pmatrix}. \qquad (\#)$$

Wir bilden die erweiterte Koeffizientenmatrix:

$$\left(\begin{array}{ccccc|c} 0 & 1 & 1 & 2 & 1 & 6 \\ 1 & 1 & 2 & 3 & 1 & 10 \\ 1 & 2 & 3 & 5 & 2 & 16 \\ 1 & 2 & 3 & 6 & 1 & 16 \end{array}\right).$$

Gauß-Elimination der ersten Spalte. Zunächst vertauschen der Zeilen 1 und 2:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 1 & 10 \\ 0 & 1 & 1 & 2 & 1 & 6 \\ 1 & 2 & 3 & 5 & 2 & 16 \\ 1 & 2 & 3 & 6 & 1 & 16 \end{array}\right).$$

Erzeugen der Nullen in der ersten Spalte:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 1 & 10 \\ 0 & 1 & 1 & 2 & 1 & 6 \\ 0 & 1 & 1 & 2 & 1 & 6 \\ 0 & 1 & 1 & 3 & 0 & 6 \end{array}\right).$$

Erzeuge Nullen in der zweiten Spalte:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 1 & 10 \\ 0 & 1 & 1 & 2 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right).$$

Obere Dreiecksform ist erreicht. Hinzufügen einer trivialen Gleichung, damit ein quadratisches (5×5)-System entsteht:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 1 & 10 \\ 0 & 1 & 1 & 2 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Reduktion auf Treppenstufenform (die Zeilen mit 0 auf der Diagonalen sind zu Nullzeilen zu machen). Es ist hier nichts zu tun. Wir markieren lediglich die Nullen längs der Diagonale:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 1 & 10 \\ 0 & 1 & 1 & 2 & 1 & 6 \\ 0 & 0 & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{0} & 0 \end{array}\right).$$

Erzeuge Nullen über den nicht-verschwindenden Diagonalelementen. Zunächst die Nullen in der 4-ten Spalte. Ziehe dazu das 2-fache der 4-ten Zeile von der ersten und das 3-fache der 4-ten Zeile von der zweiten ab:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 0 & 4 & 10 \\ 0 & 1 & 1 & 0 & 3 & 6 \\ 0 & 0 & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{0} & 0 \end{array}\right).$$

Erzeuge Nullen in der 2-ten Spalte. Subtrahiere dazu die zweite Zeile von der ersten:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & 0 & 3 & 6 \\ 0 & 0 & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{0} & 0 \end{array}\right).$$

Die 3-te und die 5-te Spalte haben 0 auf der Diagonale und liefern damit 2 Kernvektoren, indem man -1 auf der Diagonale einträgt:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & 0 & 3 & 6 \\ 0 & 0 & \boxed{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{-1} & 0 \end{array}\right).$$

Die allgemeine Lösung des Ausgangssystems (#) hat damit zwei beliebige freie Parameter λ_1 und λ_2:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \cdot \begin{pmatrix} 1 \\ 1 \\ \boxed{-1} \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 1 \\ 3 \\ 0 \\ -1 \\ \boxed{-1} \end{pmatrix}.$$

12.6 Inverse von Matrizen

Zu einer invertierbaren $n \times n$-Matrix A soll die Inverse A^{-1} berechnet werden. Zwar läßt sich dies mit der im nächsten Abschnitt vorgestellten Gauß-Jordan-Variante des Gauß-Algorithmus leicht bewerkstelligen, in der Praxis wird eine explizite Inverse von A jedoch selten benötigt (man hat in der Regel $A\vec{x} = \vec{b}$ mit einer konkret vorgegebenen rechten Seite $\vec{b}$ zu lösen). In einigen Fällen hat man immer wieder $A\vec{x} = \vec{b}$ zu lösen, wobei die Matrix A unverändert bleibt und sich im Laufe einer Iteration die rechte Seite $\vec{b}$ immer wieder neu ergibt. In diesem Fall mag man versucht sein, A^{-1} explizit zu berechnen und zu speichern, um $A\vec{x} = \vec{b}$ durch eine Matrix-Vektor-Multiplikation $A^{-1}\vec{b}$ mit $O(n^2)$ Operationen berechnen zu können, anstatt immer wieder erneut den Gauß-Algorithmus mit einem Aufwand von $O(n^3)$ Operationen auf $A\vec{x} = \vec{b}$ loszulassen. In Abschnitt 12.6.2 zeigt es sich jedoch, dass mit einem Durchlauf des Gauß-Algorithmus eine sogenannte LR-Zerlegung vorliegt, die genauso gut ist wie eine explizite Inverse A^{-1}.

12.6.1 Die Gauß-Jordan-Variante

Statt einer einzelnen Gleichung $A\vec{x} = \vec{b}$ mit $A \in K^{n\times n}$ und einer einzelnen rechten Seite $\vec{b} \in K^n$ lösen wir simultan $A\vec{x} = \vec{e}_i$ mit $i = 1, \ldots, n$, wobei $\vec{e}_i$ die Standardbasis des K^n darstellt:

Bezeichnung 12.70:

Die Standardbasis des K^n wird mit $\vec{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ ($i = 1, \ldots, n$) bezeichnet.

Dies läuft darauf hinaus, dass wir das erweiterte Koeffizientenschema $[A, \vec{e}_1, \ldots, \vec{c}_n] = [A, \mathbb{1}]$ betrachten. Wir lassen den Gauß-Algorithmus 12.64 auf das Schema los, wobei die Elementaroperationen auf das Gesamtschema angewendet werden (d.h., die Schleife über k in 12.64 ist bis zur Obergrenze $2 \cdot n$ (statt $m + 1$) auszudehnen). Als Zwischenergebnis erhält man das Schema

$$[R, B]$$

mit einer oberen Dreiecksmatrix R. Verschwindet mindestens ein Diagonalelement von R, so ist R nicht invertierbar (und damit ist auch die Ausgangsmatrix A nicht invertierbar), und wir können die Berechnung von A^{-1} abbrechen. Seien also nun alle Diagonalelemente von R ungleich 0. Wir dividieren nun jede Zeile von $[R, B]$ durch das entsprechende Diagonalelement von R, so dass längs der Diagonale nur Einsen stehen. Dann eliminieren wir mit der letzten Spalte von R

beginnend „nach oben“, indem wir durch Subtraktion geeigneter Vielfache einer Zeile von den darüber liegenden Zeilen Nullen über den Diagonalelementen erzeugen. Diese Elementaroperationen sind dabei jeweils auf das gesamte Schema anzuwenden. Zuletzt erhalten wir im linken Teil des Schemas die Einheitsmatrix:

$$[\mathbb{1}, C].$$

Satz 12.71:

Die Matrix C im rechten Teils des Schemas ist die Inverse A^{-1} der Ausgangsmatrix A.

Beweis: Alle Schritte des Gauß-Algorithmus zur Erzeugung der oberen Dreiecksform R als auch die Normierung der Diagonalelemente auf 1 und die folgende „Elimination nach oben“ sind durch Multiplikation mit Elementarmatrizen 12.63 E_1, E_2 etc. zu beschreiben, die auch auf den rechten Teil des Schemas angewendet werden:

$$[\mathbb{1}, C] = [E_k \cdots E_2 E_1 \, A, E_k \cdots E_2 E_1 \mathbb{1}].$$

Vergleich der linken Teilschemata liefert

$$\mathbb{1} = E_k \cdots E_2 E_1 \, A,$$

d.h., es gilt $E_k \cdots E_2 E_1 = A^{-1}$. Vergleich der rechten Teilschemata liefert $C = E_k \cdots E_2 E_1 \mathbb{1} = A^{-1}$.

Q.E.D.

12.6.2 LR-Zerlegung

Aus Zeitgründen wird dieser Abschnitt über die LR-Zerlegung in der Vorlesung nicht vorgestellt und hiermit als „nicht prüfungsrelevant“ deklariert. `;-)`

Wir betrachten noch einmal die Zeilenvertauschungen, die im Gauß-Algorithmus durchzuführen sind. Diese sind als Matrixmultiplikationen mit mit

$$P = \begin{array}{cc} \begin{pmatrix} \mathbb{1} & & & & \\ & 0 & & 1 & \\ & & \mathbb{1} & & \\ & 1 & & 0 & \\ & & & & \mathbb{1} \end{pmatrix} & \begin{array}{l} \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \end{array} \\ \begin{array}{ccccc} \quad & \uparrow & \quad & \uparrow & \quad \\ & i & & j & \end{array} & \end{array} = \mathbb{1} - \vec{e}_i \, \vec{e}_i^T - \vec{e}_j \, \vec{e}_j^T + \vec{e}_i \, \vec{e}_j^T + \vec{e}_j \, \vec{e}_i^T$$

zu interpretieren: $\begin{array}{lcl} A \to PA & : & \text{Vertauschung der Zeilen } i \text{ und } j, \\ A \to AP & : & \text{Vertauschung der Spalten } i \text{ und } j, \end{array}$

Definition 12.72:

Eine **Permutationsmatrix** *ist eine quadratische Matrix, die in jeder Zeile und jeder Spalte mit genau einer* 1 *und ansonsten mit* 0 *besetzt ist.*

Lemma 12.73:

a) Produkte von Permutationsmatrizen sind wieder Permutationsmatrizen.

b) Permutationsmatrizen sind orthogonal: $PP^T = P^TP = \mathbb{1}$.

c) Linksmultiplikation mit einer Permutationsmatrix entspricht einer Zeilenvertauschung, Rechtsmultiplikation einer Spaltenvertauschung.

Beweis: offensichtliche Übung.

Beispiel 12.74: Sei $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $A = \begin{pmatrix} \vec{z}_1^T \\ \vec{z}_2^T \\ \vec{z}_3^T \end{pmatrix} = (\vec{s}_1, \vec{s}_2, \vec{s}_3)$ (Zerlegung von A nach Zeilen bzw. Spalten). Es gilt $PA = \begin{pmatrix} \vec{z}_2^T \\ \vec{z}_3^T \\ \vec{z}_1^T \end{pmatrix}$, $AP = (\vec{s}_3, \vec{s}_1, \vec{s}_2)$.

Satz 12.75: (LR-Zerlegung)

Das System $A\vec{x} = \vec{b}$ *wird im Gauß-Algorithmus 12.64 auf die Endform*

$$\left(\begin{array}{cccc|c} r_{11} & \ldots & \ldots & r_{1n} & c_1 \\ l_{21} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ l_{n1} & \ldots & l_{n,n-1} & r_{nn} & c_n \end{array}\right)$$

gebracht (der untere Dreiecksanteil (l_{ij}) *entsteht aus den abgespeicherten Eliminationsfaktoren). Es seien* $P_1, \ldots, P_{n-1}$ *(eventuell mit* $P_j = \mathbb{1}$*) die bei der Pivotierung durchgeführten Zeilenvertauschungen. Mit*

$$L = \begin{pmatrix} 1 & & & 0 \\ l_{11} & \ddots & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \ldots & l_{n,n-1} & 1 \end{pmatrix} , \quad R = \begin{pmatrix} r_{11} & \ldots & \ldots & r_{1n} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \vdots \\ 0 & & & r_{nn} \end{pmatrix}$$

und $P = P_{n-1} \cdots P_1$ *gilt:* $PA = LR$ *(“*LR**-Faktorisierung**”*).*

Beweis: (für technisch Interessierte)
Zu Beginn des j.ten Schrittes liegt das Zwischenschema

$$\left(\begin{array}{cccc|ccc} r_{11} & \cdots & \cdots & r_{1,j-1} & r_{1j} & \cdots & r_{1n} \\ l_{21} & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ l_{j-1,j} & \cdots & l_{j-1,j-2} & r_{j-1,j-1} & r_{j-1,j} & \cdots & r_{j-1,n} \\ \hline l_{j1}^{(j)} & \cdots & \cdots & l_{j,j-1}^{(j)} & a_{jj}^{(j)} & \cdots & a_{jn}^{(j)} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ l_{n1}^{(j)} & \cdots & \cdots & l_{n,j-1}^{(j)} & a_{nj}^{(j)} & \cdots & a_{nn}^{(j)} \end{array}\right) = L^{(j)} + R^{(j)}$$

mit

$$L^{(j)} = \left(\begin{array}{ccc|c} 0 & & 0 & \\ l_{21} & \ddots & & 0 \\ \vdots & \ddots & 0 & \\ \hline l_{j1}^{(j)} & \ddots & l_{j,j-1}^{(j)} & \\ \vdots & \ddots & \vdots & 0 \\ l_{n1}^{(j)} & \cdots & l_{n,j-1}^{(j)} & \end{array}\right)$$

und

$$R^{(j)} = \left(\begin{array}{ccc|ccc} r_{11} & \cdots & \cdots & \cdots & \cdots & r_{1n} \\ & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & & r_{j-1,j-1} & \cdots & \cdots & r_{j-1,n} \\ \hline & & & a_{jj}^{(j)} & \cdots & a_{jn}^{(j)} \\ & 0 & & \vdots & \ddots & \vdots \\ & & & a_{nj}^{(j)} & \cdots & a_{nn}^{(j)} \end{array}\right)$$

vor. Die nächste Zeilenvertauschung P_j liefert $\tilde{L}^{(j)} = P_j L^{(j)}$ und

$$\tilde{R}^{(j)} = P_j R^{(j)} = \left(\begin{array}{ccc|ccc} r_{11} & \cdots & \cdots & \cdots & \cdots & r_{1n} \\ & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & & r_{j-1,j-1} & \cdots & \cdots & r_{j-1,n} \\ \hline & & & r_{jj} & \cdots & r_{jn} \\ \cline{4-6} & & & \tilde{a}_{j+1,j}^{(j)} & \cdots & \tilde{a}_{j+1,n}^{(j)} \\ & 0 & & \vdots & \ddots & \vdots \\ & & & \tilde{a}_{nj}^{(j)} & \cdots & \tilde{a}_{nn}^{(j)} \end{array}\right).$$

Die neuen Eliminationsfaktoren $\vec{l}_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \tilde{a}^{(j)}_{j+1,j}/r_{jj} \\ \vdots \\ \tilde{a}^{(j)}_{n,j}/r_{jj} \end{pmatrix}$ werden als j.te Spalte an $\tilde{L}^{(j)}$ angefügt:

$$L^{(j+1)} = \tilde{L}^{(j)} + \vec{l}_j\, \vec{e}_j^{\,T} \ .$$

Vielfache der j.ten Zeile $\vec{e}_j^{\,T} \tilde{R}^{(j)} = (0, \ldots, 0, r_{jj}, \ldots, r_{jn})$ von $\tilde{R}^{(j)}$ werden von den unteren Zeilen von $\tilde{R}^{(j)}$ abgezogen:

$$R^{(j+1)} = \tilde{R}^{(j)} - \vec{l}_j\, \vec{e}_j^{\,T} \tilde{R}^{(j)} \ .$$

Beachte hierzu

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ l_{j+1} \\ \vdots \\ l_n \end{pmatrix} (0, \ldots, 0, r_{jj}, \ldots, r_{jn}) = \left(\begin{array}{c|ccc} 0 & & 0 & \\ \hline & l_{j+1}r_{jj} & \ldots & l_{j+1}r_{jn} \\ 0 & \vdots & \ddots & \vdots \\ & l_n r_{jj} & \ldots & l_n r_{jn} \end{array}\right) .$$

Ergebnis: der Gauß-Algorithmus wird beschrieben durch

$$L^{(j+1)} = P_j L^{(j)} + \vec{l}_j e_j^T \ , \quad R^{(j+1)} = P_j R^{(j)} - \vec{l}_j e_j^T\, P_j R^{(j)}$$

mit $L^{(1)} = 0$, $R^{(1)} = A$.

Behauptung: In jedem Schritt gilt $(\mathbb{1} + L^{(j)}) R^{(j)} = P_{j-1} \cdots P_1\, A$.

(Mit $L = \mathbb{1} + L^{(n)}$, $R = R^{(n)}$, $P = P_{n-1} \cdots P_1$ ist der Satz damit bewiesen).

Beweis: Induktion nach j:

$j = 1$: $(\mathbb{1} + L^{(1)}) R^{(1)} = A$ gilt offensichtlich.

Schritt $j \to j+1$:

$$\begin{aligned}
& (\mathbb{1} + L^{(j+1)})\, R^{(j+1)} \\
= \ & (\mathbb{1} + P_j L^{(j)} + \vec{l}_j e_j^T)\, (\mathbb{1} - \vec{l}_j e_j^T)\, P_j R^{(j)} \\
= \ & (\mathbb{1} - \cancel{\vec{l}_j \vec{e}_j^{\,T}} + P_j L^{(j)} - P_j L^{(j)} \vec{l}_j \vec{e}_j^{\,T} + \cancel{\vec{l}_j \vec{e}_j^{\,T}} - \vec{l}_j \vec{e}_j^{\,T} \vec{l}_j \vec{e}_j^{\,T})\, P_j R^{(j)} \\
= \ & (\mathbb{1} + P_j L^{(j)}) P_j R^{(j)} \qquad\qquad (\text{beachte } L^{(j)} \vec{l}_j = 0 \ , \ \vec{e}_j^{\,T} \vec{l}_j = 0) \\
= \ & P_j (\mathbb{1} + L^{(j)} P_j)\, R^{(j)} \overset{(*)}{=} P_j (\mathbb{1} + L^{(j)})\, R^{(j)} \ = \ P_j\, (P_{j-1} \cdots P_1\, A) \ .
\end{aligned}$$

Für $(*)$ beachte $L^{(j)}P_j = L^{(j)}$, da Rechtsmultiplikation mit P_j die *Spalten* $j, \ldots, n$ von $L^{(j)}$ vertauscht, die alle gleich (nämlich 0) sind.

Q.E.D.

Bemerkung 12.76: $L = \begin{pmatrix} 1 & & \\ * & \ddots & \\ * & * & 1 \end{pmatrix}$ *ist immer invertierbar. Damit ist* $R = L^{-1}PA$ *genau dann invertierbar, wenn* A *invertierbar ist.*

Bemerkung 12.77: *Zu jeder quadratischen Matrix existiert damit eine Zeilenvertauschung* P*, so daß* PA *in der angegebenen Form* LR *faktorisiert werden kann. Bei invertierbarem* A *und fixiertem* P *sind die* LR*-Faktoren eindeutig:*

$$LR = \tilde{L}\tilde{R} \quad \Longrightarrow \quad \underbrace{\tilde{L}^{-1}L}_{\substack{\text{untere}\\ \text{Dreiecksmatrix}}} \quad = \quad \underbrace{\tilde{R}R^{-1}}_{\substack{\text{obere}\\ \text{Dreiecksmatrix}}} \quad = \ \text{Diagonalmatrix } D\ .$$

Mit $L = \tilde{L}D$ *folgt* $D = \mathbb{1}$*, da die Diagonalen von* L *und* $\tilde{L}$ *mit* 1 *besetzt sind.*

Bemerkung 12.78: *Es gibt Matrizen, für die keine Faktorisierung* $A = LR$ *existiert, man braucht i.a.* P*. Beispiel:*

$$\begin{pmatrix} 0 & * \\ 1 & * \end{pmatrix} \stackrel{(?)}{=} \begin{pmatrix} 1 & 0 \\ l_{21} & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} \\ l_{21} \cdot r_{11} & l_{21} \cdot r_{12} + r_{22} \end{pmatrix}.$$

Aus der ersten Spalte ergeben sich die Gleichungen

$$0 \stackrel{(!)}{=} r_{11}, \quad 1 \stackrel{(!)}{=} l_{21} \cdot r_{11},$$

die keine Lösung für l_{21} *und* r_{11} *zulassen. In der Tat braucht der Gauß-Algorithmus für die Matrix* $\begin{pmatrix} 0 & * \\ 1 & * \end{pmatrix}$ *gleich im ersten Schritt wegen der* 0 *links oben eine Zeilenvertauschung, ohne die der Algorithmus zum Erliegen kommen würde.*

Bemerkung 12.79: P *ist die Gesamtheit aller im Gauß-Algorithmus durchgeführten Zeilenvertauschungen. Man speichert* P *nicht als Matrix, sondern als "***Buchhaltungsvektor***"* $\vec{p}$*. Dazu initialisiert man die Spalte* $\vec{p} = (1, 2, \ldots, n)^T$ *und wendet die Zeilenvertauschungen auch auf* $\vec{p}$ *an. Zuletzt hat man* $\vec{p} = (p_1, \ldots, , p_n)^T$*. Interpretation: die ursprüngliche Zeile* p_i *ist in die* i*.te Zeile gewandert. Hiermit ergibt sich für beliebiges* $\vec{b}$ *dann* $\vec{c} = P\vec{b}$ *durch*

$$\boxed{\textbf{for } i := 1 \textbf{ to } n \textbf{ do } \ c_i := b_{p_i}\ ;}$$

Bemerkung 12.80: *Mit* $\det(L) = 1$ *und leicht berechenbarem*

$$\det(R) = \text{ Produkt der Diagonalelemente}$$

kann $\det(A)$ *aus* $\det(P)\det(A) = \det(R)$ *berechnet werden. Mit* $\det(P) = \det(P_{n-1}) \cdots \det(P_1)$ *und*

$$\det(P_j) = \begin{cases} \ \ 1 & , \text{falls } P_j = \mathbb{1} \\ -1 & , \text{falls } P_j = \text{ Austausch zweier Zeilen} \end{cases}$$

braucht nur die Anzahl der nichttrivialen Zeilenvertauschungen im Gauß-Algorithmus mitgezählt zu werden.

Bemerkung 12.81: Die Faktorisierungsdaten P, L, R sind genauso gut wie eine explizite Inverse A^{-1}! *Die Lösung von*

$$A\vec{x} = \vec{b} \quad \Longrightarrow \quad PA\vec{x} = P\vec{b}\ , \quad \text{d.h.,} \quad L\,\underbrace{R\,\vec{x}}_{\vec{y}} = P\vec{b}$$

ist durch 2 Rücksubstitutionen

$$\begin{array}{rl} & \text{bestimme } \vec{y} \text{ aus } L\vec{y} = P\vec{b}, \\ \text{dann} & \text{bestimme } \vec{x} \text{ aus } R\vec{x} = \vec{y} \end{array}$$

schnell zu berechnen. Der Aufwand (etwa n^2 Multiplikationen) entspricht dem der Matrix-Vektor-Multiplikation $A^{-1}\vec{b}$ bei explizit vorliegendem A^{-1}.

Merke: Der Gauß-Algorithmus liefert mit $\approx n^3/3$ Multiplikationen die Zerlegung $PA = LR$ einer $n \times n$-Matrix. Liegen P, L, R vor, dann ist $A\vec{x} = \vec{b}$ durch doppelte Rücksubstitution mit etwa n^2 Multiplikationen schnell zu lösen.

12.7 Eigenwerte und -vektoren

4.2.04↓

Die „Spektralanalyse“ = „Diagonalisierung“ linearer Abbildungen (das Bestimmen von Eigenwerten und -vektoren) ist speziell für physikalische Anwendungen von zentraler Bedeutung.

12.7.1 Definitionen

Definition 12.82: (Eigenwerte und -vektoren)

Zu einer quadratischen Matrix $A \in K^{n\times n}$ gebe es $\lambda \in \mathbb{C}$ und $\vec{x} \in \mathbb{C}^n$ mit $\vec{x} \neq 0$, so dass

$$A\,\vec{x} = \lambda \cdot \vec{x}.$$

Der Wert λ heißt **„Eigenwert"** *von A, der Vektor $\vec{x}$ wird als* **„Eigenvektor"** *zum Eigenwert λ bezeichnet. Die Vektormenge*

$$E_\lambda = \Big\{ \vec{x} \in \mathbb{C}^n;\; A\vec{x} = \lambda \cdot \vec{x} \Big\}$$

heißt **„Eigenraum"** *zum Eigenvektor λ.*

Bemerkung 12.83: *Man überzeugt sich leicht davon, dass zu einem gegebenen Eigenwert λ von A der Eigenraum E_λ ein Unterraum des $\mathbb{C}^n$ ist. Speziell heißt dies, dass Eigenvektoren nicht eindeutig sind. Jedes Vielfache eines Eigenvektors ist wieder ein Eigenvektor (falls das Vielfache nicht gerade 0 ist). Ist die Dimension des Eigenraums E_λ größer als 1, können auch linear unabhängige Eigenvektoren zu λ zu einem neuen Eigenvektor kombiniert werden.*

Definition 12.84: (Das charakteristische Polynome einer Matrix)

Zu $A = (a_{ij}) \in K^{n\times n}$ definieren wir das **„charakteristische Polynom"**

$$p_A(\lambda) = \det(A - \lambda \cdot \mathbb{1}) = \det\left(\begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \ldots & \ldots & a_{nn}-\lambda \end{pmatrix}\right)$$

Man hat sich hierbei schnell davon überzeugt, dass das charakteristische Polynom

$$p_A(\lambda) = (-1)^n \cdot \lambda^n + c_{n-1} \cdot \lambda^{n-1} + \ldots + c_1 \cdot \lambda + c_0$$

in der Tat ein Polynom vom Grad n ist (um den führenden Koeffizienten $(-1)^n$ einzusehen, betrachte man einfach $A = 0$). Einige weitere der Koeffizienten c_i des Polynoms sind ebenfalls leicht der Matrix zuzuordnen. Z.B. gilt:

$$c_{n-1} = (-1)^n \cdot \underbrace{\sum_{i=1}^{n} a_{ii}}_{\mathrm{Spur}(A)}, \qquad c_0 = \det(A).$$

Zu c_0: Werte einfach das charakteristische Polynom an der Stelle $\lambda = 0$ aus, wodurch sich nach Definition $c_0 = p_A(0) = \det(A - 0 \cdot \mathbb{1}) = \det(A)$ ergibt.

Satz 12.85: (Charakterisierung der Spektraldaten einer Matrix)

Die Eigenwerte einer Matrix $A \in K^{n\times n}$ sind die Nullstellen des charakteristischen Polynoms. Die zum Eigenwert λ gehörigen Eigenvektoren sind die nicht-verschwindenden Kernvektoren der Matrix $A - \lambda \cdot \mathbb{1}$.

Beweis: Die Gleichung $A\vec{x} = \lambda \cdot \vec{x}$ kann als $(A - \lambda \cdot \mathbb{1})\,\vec{x} = 0$ geschrieben werden. Ist $A - \lambda \cdot \mathbb{1}$ invertierbar, gibt es nur die eindeutige Lösung $\vec{x} = 0$, die wir nicht als Eigenvektor zulassen. Da $A - \lambda \cdot \mathbb{1}$ genau dann nicht invertierbar ist, wenn $p_A(\lambda) = \det(A - \lambda \cdot \mathbb{1})$ verschwindet, sind die Eigenwerte die Nullstellen von $p_A(\lambda)$. Für nicht-invertierbares $A_\lambda := A - \lambda \cdot \mathbb{1}$ existiert ein Kern der Matrix A_λ, der aus Eigenvektoren von A besteht.

Q.E.D.

Wir haben damit unmittelbar eine algorithmische Möglichkeit, sowohl die Eigenwerte als auch die dazugehörigen Eigenvektoren zu bestimmen:

a) Bestimme die Eigenwerte als die Nullstellen λ_1, λ_2 etc. des charakteristischen Polynoms.

b) Berechne zum Eigenwert λ_i eine Basis des Kerns von $A - \lambda_i \cdot \mathbb{1}$. Dies kann beispielsweise durch den Gauß-Algorithmus angewendet auf die Gleichung $(A - \lambda \cdot \mathbb{1})\vec{x} = 0$ geschehen, wobei die Kernvektoren sich gemäß Abschnitt 12.5.3 algorithmisch ergeben. Dies liefert eine Basis des Eigenraums E_{λ_i}.

Beispiel 12.86: Zur Matrix

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

ergibt sich das charakteristische Polynom

$$p_A(\lambda) = \det\left(\begin{pmatrix} -\lambda & 1 \\ -2 & 3-\lambda \end{pmatrix}\right) = -\lambda \cdot (3 - \lambda) + 2 = \lambda^2 - 3 \cdot \lambda + 2$$

mit den Nullstellen (Eigenwerten)

$$\lambda_1 = 1, \quad \lambda_2 = 2.$$

Die allgemeine Lösung von $(A - \lambda_1 \cdot \mathbb{1})\,\vec{x} = 0$ ist der Kern der Matrix

$$A - \lambda_1 \cdot \mathbb{1} = \begin{pmatrix} -1 & 1 \\ -2 & 2 \end{pmatrix}.$$

Über den Entzerrungsalgorithmus ergibt sich die allgemeine Lösung $\vec{x} = \alpha \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ mit beliebigen $\alpha \in \mathbb{C}$. Wir wählen als Basis des 1-dimensionalen Kerns den Eigenvektor

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die allgemeine Lösung von $(A - \lambda_2 \cdot \mathbb{1})\,\vec{x} = 0$ ist der Kern der Matrix

$$A - \lambda_2 \cdot \mathbb{1} = \begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix}.$$

Über den Entzerrungsalgorithmus ergibt sich die allgemeine Lösung $\vec{x} = \alpha \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ mit beliebigen $\alpha \in \mathbb{C}$. Wir wählen als Basis des 1-dimensionalen Kerns den Eigenvektor

$$\vec{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Wir gehen nun der Frage nach, wieviele Eigenvektoren man zu einer $n \times n$-Matrix A finden kann. Speziell: kann ich die Eigenvektoren als Basis des $\mathbb{C}^n$ benutzen?

Satz 12.87:

Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.

Beweis: Seien $\vec{x}_1, \ldots, \vec{x}_k$ Eigenvektoren von $A \in K^{n \times n}$ zu den Eigenwerten $\lambda_1, \ldots, \lambda_k$, die alle paarweise verschieden seien. Wir betrachten die Linearkombination

$$\sum_{i=1}^{k} c_i \cdot \vec{x}_i = 0 \qquad (\#)$$

und müssen hieraus schliessen, dass $c_1 = \ldots = c_k = 0$ folgt.
Wir machen dazu Induktion nach k. Für $k = 1$ folgt sicherlich $c_1 = 0$, da $\vec{x}_1$ als Eigenvektor nicht verschwindet. Wir machen nun den Induktionsschritt $k-1 \to k$. Seien also $\vec{x}_1, \ldots, \vec{x}_{k-1}$ linear unabhängige Eigenvektoren zu $\lambda_1, \ldots, \lambda_{k-1}$, sei $\vec{x}_k$ ein zusätzlicher Eigenvektor zum Eigenwert λ_k. Aus (#) folgt durch Anwendung von A

$$\sum_{i=1}^{k} c_i \cdot \lambda_i \cdot \vec{x}_i = 0$$

und durch Multiplikation von (#) mit λ_k

$$\sum_{i=1}^{k} c_i \cdot \lambda_k \cdot \vec{x}_i = 0.$$

Durch Differenzbildung folgt

$$\sum_{i=1}^{k-1} c_i \cdot (\lambda_i - \lambda_k) \cdot \vec{x}_i = 0,$$

wobei der Term mit $\vec{v}_k$ herausfällt. Nach Induktionsvoraussetzung folgt

$$c_i \cdot (\lambda_i - \lambda_k) = 0$$

wegen der linearen Unabhängigkeit von $\vec{v}_1, \ldots, \vec{v}_{k-1}$ und damit $c_1 = \ldots = c_{k-1} = 0$, da $\lambda_i - \lambda_k \neq 0$. Damit reduziert sich (#) zu $\lambda_k \cdot \vec{v}_k = 0$, woraus auch $c_k = 0$ folgt.

Q.E.D.

Die Frage, ob man den gesamten $\mathbb{C}^n$ durch Eigenvektoren einer Matrix aufspannen kann, läßt sich damit für den Spezialfall leicht beantworten, dass alle Eigenwerte unterschiedlich sind:

Folgerung 12.88:

Hat eine $n \times n$-Matrix n unterschiedliche Eigenwerte, so bilden die entsprechenden Eigenvektoren eine Basis des $\mathbb{C}^n$.

Beweis: Nach Satz 12.87 gehören zu den n unterschiedlichen Eigenwerten n linear unabhängige Eigenvektoren. Mit Satz 12.11.b) bilden diese Vektoren eine Basis des $\mathbb{C}^n$ und spannen den gesamten Raum auf.

Q.E.D.

Die Sache wird deutlich komplizierter, wenn es „entartete Eigenwerte" gibt.

Definition 12.89: (mehrfache Eigenwerte)

Sei $\lambda \in \mathbb{C}$ ein Eigenwert von $A \in K^{n\times n}$. Die Vielfachheit der Nullstelle λ des charakteristischen Polynoms von A im Sinne von Definition 1.24 heißt **„algebraische Vielfachheit"** *des Eigenwerts. Im Gegensatz dazu wird die Dimension des Kerns von $A - \lambda \cdot \mathbb{1}$ als* **„geometrische Vielfachheit"** *bezeichnet.*

Satz 12.90:

Für jeden Eigenwert einer Matrix gilt:

$$1 \leq \textit{geometrische Vielfachheit} \;\leq\; \textit{algebraische Vielfachheit}.$$

Beweis: Die Vielfachheit einer Polynomnullstelle (algebraische Vielfachheit des Eigenwerts) ist nach Definition ≥ 1. Für jeden Eigenwert λ_0 einer Matrix A ist $A - \lambda_0 \cdot \mathbb{1}$ nicht invertierbar und hat damit einen nicht-trivialen Kern. Damit gilt, dass auch für die geometrische Vielfachheit eines Eigenwerts stets ≥ 1 gilt. Nun zur Abschätzung

$$\text{geometrische Vielfachheit} \;\leq\; \text{algebraische Vielfachheit}.$$

Zunächst beobachten wir, dass eine Matrix $A \in K^{n\times n}$ und $T^{-1}AT$ (mit einer beliebigen invertierbaren Matrix $T \in K^{n\times n}$) das selbe charakteristische Polynom haben:

$$\begin{aligned}\det(T^{-1}AT - \lambda\cdot \mathbb{1}) &= \det(T^{-1}(A-\lambda\cdot\mathbb{1})T)\\ &= \det(T^{-1})\cdot\det(T)\cdot\det(A-\lambda\cdot\mathbb{1})\\ &= \det(T^{-1}T)\cdot\det(A-\lambda\cdot\mathbb{1})T) = \det(A-\lambda\cdot\mathbb{1}).\end{aligned}$$

Seien $\vec{k}_1,\ldots,\vec{k}_g$ Basisvektoren des Kerns von $A-\lambda_0\cdot\mathbb{1}$ für einen Eigenwert λ_0 mit geometrischer Vielfachheit g. Wir wählen irgendwelche linear unabhängigen Vektoren $\vec{v}_{g+1},\ldots,\vec{v}_n$, die zusammen mit den Kernvektoren eine Basis des $\mathbb{C}^n$ bilden. Betrachte nun

$$\begin{aligned}\tilde{A} &= \underbrace{[\vec{k}_1,\ldots,\vec{k}_g,\vec{v}_{g+1},\ldots,\vec{v}_n]^{-1}}_{T^{-1}}\ A\ \underbrace{[\vec{k}_1,\ldots,\vec{k}_g,\vec{v}_{g+1},\ldots,\vec{v}_n]}_{T}\\ &= T^{-1}[A\cdot\vec{k}_1,\ldots,A\cdot\vec{k}_g,A\vec{v}_{g+1},\ldots,A\vec{v}_n]\\ &= T^{-1}[\lambda_0\cdot\vec{k}_1,\ldots,\lambda_0\cdot\vec{k}_g,A\vec{v}_{g+1},\ldots,A\vec{v}_n]\\ &= [\lambda_0\cdot T^{-1}\vec{k}_1,\ldots,\lambda_0\cdot T^{-1}\vec{k}_g,T^{-1}A\vec{v}_{g+1},\ldots,T^{-1}A\vec{v}_n].\end{aligned}$$

Nach Konstruktion gilt

$$T^{-1}\underbrace{[\vec{k}_1,\ldots,\vec{k}_g,\vec{v}_{g+1},\ldots,\vec{v}_n]}_{T} = [T^{-1}\vec{k}_1,\ldots,T^{-1}\vec{k}_g,T^{-1}\vec{v}_{g+1},\ldots,T^{-1}\vec{v}_n]$$

$$= \left(\begin{array}{ccc|ccc} 1 & 0 & \ldots & 0 & \ldots & 0\\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & 1 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \ddots & 0\\ 0 & \ldots & \ldots & \ldots & 0 & 1\end{array}\right).$$

Betrachtet man die ersten g Spalten, folgt $T^{-1}\vec{k}_i = \vec{e}_i$ (Standardbasis), also

$$\tilde{A} = [\lambda_0\cdot T^{-1}\vec{k}_1,\ldots,\lambda_0\cdot T^{-1}\vec{k}_g,\vec{*},\ldots,\vec{*}] = \left(\begin{array}{ccc|cc} \lambda_0 & 0 & \ldots & * & \ldots\\ 0 & \ddots & \ddots & \vdots & \ddots\\ \vdots & \ddots & \lambda_0 & \vdots & \ddots\\ \vdots & \ddots & 0 & \vdots & \ddots\\ \vdots & \ddots & \vdots & \vdots & \ddots\\ 0 & \ldots & 0 & * & \ldots\end{array}\right).$$

Das charakteristische Polynom von $\tilde{A}$ ergibt sich durch rekursive Entwicklung von $\det(\tilde{A} - \lambda \cdot \mathbb{1})$ nach den ersten Spalten zu

$$\det(\tilde{A} - \lambda \cdot \mathbb{1}) = (\lambda_0 - \lambda)^g \cdot \tilde{p}(\lambda)$$

mit irgendeinem Faktorpolynom, das durch den rechten unteren Block von $\tilde{A}$ bestimmt wird. Das obige charakteristische Polynom von $\tilde{A}$, das mit dem von A übereinstimmt, hat den Eigenwert λ_0 mit der Vielfachheit von mindestens g.

Q.E.D.

Seien $\lambda_1, \ldots, \lambda_k$ die unterschiedlichen Nullstellen des charakteristischen Polynoms vom Grad n, ihre algebraischen Vielfachheiten seien $n_1, \ldots, n_k$. Nach dem Fundamentalsatz der Algebra 1.29 addieren sich die algebraischen Vielfachheiten zu Dimension n des Raums auf: $n_1 + \ldots + n_k = n$. Wenn dies auch für die geometrischen Vielfachheiten $g_1, \ldots, g_k$ gilt, so läßt sich aus den Eigenvektoren eine Basis bilden. Man wähle dazu aus jedem Kern von $A - \lambda_i \cdot \mathbb{1}$ jeweils g_i Basisvektoren, die insgesamt ein System von $g_1 + \ldots + g_k$ linear unabhängigen Vektoren im $\mathbb{C}^n$ bilden. Gilt $g_1 + \ldots + g_k = n$, so haben wir eine Basis des $\mathbb{C}^n$ vor uns. Die entscheidende Frage ist also:

Gilt immer geometrische Vielfachheit =algebraische Vielfachheit?

Die Antwort ist leider:

Nein!

Beispiel 12.91: Die Matrix

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

hat das charakteristische Polynom

$$p_A(\lambda) = \det\left(\begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix}\right) = \lambda^2.$$

Der Wert $\lambda = 0$ ist eine doppelte Nullstelle dieses Polynoms, damit ist $\lambda = 0$ ein Eigenwert mit der algebraischen Vielfachheit 2. Die Kernvektoren $\vec{k} = (k_1, k_2)^T$ von $A - 0 \cdot \mathbb{1} = A$ sind durch

$$A\vec{k} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

charakterisiert, also durch die einzelne Gleichung $k_2 = 0$. Der Kern von A besteht also aus dem 1-dimensionalen Raum

$$\left\{ \begin{pmatrix} k_1 \\ 0 \end{pmatrix} ; k_1 \in \mathbb{C} \right\}.$$

Man findet also die algebraische Vielfachheit 2 von $\lambda = 0$, aber die geometrische Vielfachheit 1. Mit dem einzigen Eigenvektor $(1, 0)^T$, der den Kern aufspannt, kann man nicht den $\mathbb{C}^2$ aufspannen.

12.7.2 Symmetrische/hermitesche Matrizen

$\downarrow$5.2.04

Zunächst einige einfache Beobachtungen zur den Eigenwerten und -vektoren zur (hermitesch) transponierten Matrix A^H (Definition 12.33) einer Matrix A:

Satz 12.92:

Ist λ Eigenwert von A, so ist $\overline{\lambda}$ Eigenwert von A^H.

Beweis: Wegen

$$\overline{\det(A - \lambda \cdot \mathbb{1})} = \overline{\det(A^T - \lambda \cdot \mathbb{1})} = \det(A^H - \overline{\lambda} \cdot \mathbb{1})$$

stimmten die charakteristischen Polynome von A und A^H bis auf komplexe Konjugation überein: $\overline{p_A(\lambda)} = p_{A^H}(\overline{\lambda})$.

Q.E.D.

Satz 12.93:

Sei $\vec{x}$ ein Eigenvektor von A zum Eigenwert λ, sei $\vec{y}$ ein Eigenvektor von A^H zum Eigenwert $\mu \neq \overline{\lambda}$. Dann gilt $\langle \vec{y}, \vec{x} \rangle = 0$.

Beweis: Wir betrachten das komplexe Skalarprodukt $\langle \vec{a}, \vec{b} \rangle = \sum_i \overline{a_i} \cdot b_i$ auf dem $\mathbb{C}^n$. Es gilt

$$\begin{aligned} \langle \vec{y}, A\vec{x} \rangle &= \langle \vec{y}, \lambda \cdot \vec{x} \rangle = \lambda \cdot \langle \vec{y}, \vec{x} \rangle \\ = \langle A^H \vec{y}, \vec{x} \rangle &= \langle \mu \cdot \vec{y}, \vec{x} \rangle = \overline{\mu} \cdot \langle \vec{y}, \vec{x} \rangle. \end{aligned}$$

Es folgt

$$(\lambda - \overline{\mu}) \cdot \langle \vec{y}, \vec{x} \rangle = 0,$$

womit das Skalarprodukt zwischen $\vec{x}$ und $\vec{y}$ für $\lambda \neq \overline{\mu}$ verschwinden muss.

Q.E.D.

Mit der selben Argumentation wie im obigen Beweis bekommt man im Fall hermitescher Matrizen die wichtige Aussage, dass alle Eigenwerte reell sind:

Satz 12.94:

Alle Eigenwerte hermitescher Matrizen $A = A^H$ sind reell.

Beweis: Wir betrachten das komplexe Skalarprodukt $\langle \vec{a}, \vec{b} \rangle = \sum_i \overline{a_i} \cdot b_i$ auf dem $\mathbb{C}^n$. Mit $A\vec{x} = \lambda \cdot \vec{x}$ folgt

$$\begin{aligned} \langle \vec{x}, A\vec{x} \rangle &= \langle \vec{x}, \lambda \cdot \vec{x} \rangle = \lambda \cdot \langle \vec{x}, \vec{x} \rangle \\ = \langle A^H \vec{x}, \vec{x} \rangle &= \langle \lambda \cdot \vec{x}, \vec{x} \rangle = \overline{\lambda} \cdot \langle \vec{x}, \vec{x} \rangle, \end{aligned}$$

also

$$(\lambda - \overline{\lambda}) \cdot \langle \vec{x}, \vec{x} \rangle = 0,$$

wobei $\langle \vec{x}, \vec{x} \rangle \neq 0$ gilt, da ein Eigenvektor nicht 0 sein kann. Aus $\lambda = \overline{\lambda}$ folgt, dass λ reell ist.

Q.E.D.

Satz 12.95:

Eigenvektoren hermitescher Matrizen $A = A^H$ zu verschiedenen Eigenwerten sind orthogonal.

Beweis: Dies ist der Spezialfall von Satz 12.93 für hermitesche Matrizen. Sei $A\vec{x} = \lambda \cdot \vec{x}$, $A\vec{y} = \mu \cdot \vec{y}$ mit reellen Eigenwerten $\lambda \neq \mu$. Wir betrachten das komplexe Skalarprodukt $\langle \vec{a}, \vec{b}\,\rangle = \sum_i \overline{a_i} \cdot b_i$ auf dem $\mathbb{C}^n$. Es gilt

$$\langle \vec{y}, A\vec{x}\,\rangle = \langle \vec{y}, \lambda \cdot \vec{x}\,\rangle = \lambda \cdot \langle \vec{y}, \vec{x}\,\rangle$$

$$= \langle A^H\vec{y}, \vec{x}\,\rangle = \langle \mu \cdot \vec{y}, \vec{x}\,\rangle = \overline{\mu} \cdot \langle \vec{y}, \vec{x}\,\rangle = \mu \cdot \langle \vec{y}, \vec{x}\,\rangle.$$

Es folgt

$$(\lambda - \mu) \cdot \langle \vec{y}, \vec{x}\,\rangle = 0,$$

womit das Skalarprodukt zwischen $\vec{x}$ und $\vec{y}$ für $\lambda \neq \mu$ verschwinden muss.

Q.E.D.

Hier die zentrale Besonderheit des Eigenwertproblems für hermitesche Matrizen:

Satz 12.96:

a) Für jeden Eigenwert einer hermiteschen Matrix gilt

$$1 \leq \textit{geometrische Vielfachheit} \ = \ \textit{algebraische Vielfachheit}.$$

b) Hermitesche Matrizen besitzen eine Basis von orthogonalen Eigenvektoren.

Beweis: Wir beweisen a) und b) simultan für hermitesches $A \in K^{n\times n}$. Seien dazu $\lambda_1, \ldots, \lambda_k$ die paarweise unterschiedlichen Eigenwerte mit den geometrischen Vielfachheiten $g_1, \ldots, g_k$. Seien $\vec{v}_j^{(i)}$ mit $j = 1, \ldots, g_i$ eine orthogonale Basis des Eigenraums E_{λ_i} (nehme z.B. irgendeine Basis des Kerns von $A - \lambda_i \cdot 1\!\!1$ und orthogonalisiere sie per Gram-Schmidt (Aufgabe 110, Blatt 13)). Da nach Satz 12.95 die Eigenvektoren zu unterschiedlichen Eigenwerten orthogonal sind, haben wir insgesamt eine Sammlung von $g = \sum_{i=1}^k g_i$ orthogonalen Eigenvektoren

$$\underbrace{\vec{v}_1^{(1)}, \ldots, \vec{v}_{g_1}^{(1)},}_{\text{Eigenwert } \lambda_1} \ \underbrace{\vec{v}_1^{(2)}, \ldots, \vec{v}_{g_2}^{(2)},}_{\text{Eigenwert } \lambda_2} \ \ldots \ , \ \underbrace{\vec{v}_1^{(k)}, \ldots, \vec{v}_{g_k}^{(k)},}_{\text{Eigenwert } \lambda_k}$$

wobei

$$g = \sum_{i=1}^{k} g_i \leq \sum_{i=1}^{k} n_i = n$$

gilt. Wir wählen weitere $n - g$ Vektoren $\vec{w}_1, \ldots, \vec{w}_{n-g}$ hinzu, die zusammen mit den Eigenvektoren den K^n aufspannen. Nach Orthogonalisierung per Gram-Schmidt (Aufgabe 110, Blatt 13), können wir sie als paarweise orthogonal und orthogonal zu den Eigenvektoren $\vec{v}_j^{(i)}$ annehmen.

Aus diesen insgesamt n Vektoren wird spaltenweise die Matrix

$$T = [\vec{v}_1^{(1)}, \dots, \vec{v}_{g_1}^{(1)},\ \vec{v}_1^{(2)}, \dots, \vec{v}_{g_2}^{(2)},\ \dots\ ,\ \vec{v}_1^{(k)}, \dots, \vec{v}_{g_k}^{(k)}, \vec{w}_1, \dots, \vec{w}_{n-g}]$$

aufgebaut, die wegen der Orthogonalität der Spalten $T^H T = \mathbb{1}$ erfüllt (also eine orthogonale Matrix mit $T^{-1} = T^H$ ist). Analog zum Beweis von Satz 12.90 betrachten wir die Matrix

$$\tilde{A} = T^{-1} A T,$$

die mit den selben Argumenten wie im Beweis von Satz 12.90 von der Form

$$\tilde{A} \quad = \quad \begin{array}{c}
\begin{array}{cccccccccc}
 & & g_1 & & & & g & g+1 & & n \\
 & & \downarrow & & & & \downarrow & \downarrow & & \downarrow
\end{array} \\
\left(\begin{array}{ccccccc|ccc}
\lambda_1 & 0 & \dots & \dots & \dots & \dots & 0 & * & \dots & * \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \lambda_1 & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \lambda_k & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots & \ddots & \vdots \\
0 & \dots & \dots & \dots & \dots & 0 & \lambda_k & * & \dots & * \\
0 & \dots & \dots & \dots & \dots & \dots & 0 & * & \dots & * \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & 0 & * & \dots & *
\end{array}\right)
\end{array} .$$

sein muss. Wegen $\tilde{A}^H = T^H A^H (T^{-1})^H = T^{-1} A T = \tilde{A}$ ist $\tilde{A}$ wieder hermitesch, damit muss $\tilde{A}$ speziell von folgender Form sein:

$$\tilde{A} \quad = \quad \begin{array}{c}
\begin{array}{cccccccccc}
 & & g_1 & & & & g & g+1 & & n \\
 & & \downarrow & & & & \downarrow & \downarrow & & \downarrow
\end{array} \\
\left(\begin{array}{ccccccc|ccc}
\lambda_1 & 0 & \dots & \dots & \dots & \dots & 0 & 0 & \dots & 0 \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \lambda_1 & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \lambda_k & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots & \ddots & \vdots \\
0 & \dots & \dots & \dots & \dots & 0 & \lambda_k & 0 & \dots & 0 \\
\hline
0 & \dots & \dots & \dots & \dots & \dots & 0 & * & \dots & * \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & 0 & * & \dots & *
\end{array}\right)
\end{array} .$$

Das charakteristische Polynom von $\tilde{A}$ (das mit dem von A übereinstimmt, siehe den Beweis von Satz 12.90) ist

$$p_{\tilde{A}}(\lambda) = (\lambda_1 - \lambda)^{g_1} \cdot \ldots \cdot (\lambda_k - \lambda)^{g_k} \cdot \det(S - \lambda \cdot \mathbb{1}),$$

wo S der (symmetrische) Matrixblock der Dimension $(n-g) \times (n-g)$ in der rechten unteren Ecke von $\tilde{A}$ ist. Vergleich mit

$$p_A(\lambda) = (\lambda_1 - \lambda)^{n_1} \cdot \ldots \cdot (\lambda_k - \lambda)^{n_k}$$

liefert

$$\det(S - \lambda \cdot \mathbb{1}) = (\lambda_1 - \lambda)^{n_1 - g_1} \cdot \ldots \cdot (\lambda_k - \lambda)^{n_k - g_k},$$

d.h., dieser Block hat einen Eigenwert λ_i, falls für mindestens einen Index $g_i < n_i$ gilt. Dazu gehört ein Eigenvektor $(x_1, \ldots x_{n-g})^T \in K^{n-g}$ des Blocks S, der zu einem Eigenvektor

$$\vec{x} = (0, \ldots, 0, x_1, \ldots, x_{n-g})^T \in K^n$$

von $\tilde{A}$ zum Eigenwert λ_i erweitert werden kann. Hieraus wird ein Eigenvektor $T\vec{x}$ von A, denn

$$A(T\vec{x}) = T\tilde{A}T^{-1}(T\vec{x}) = T\tilde{A}\vec{x} = T\lambda_i \cdot \vec{x} = \lambda_i \cdot (T\vec{x}).$$

Man überzeugt sich leicht davon, dass der Eigenvektor $\vec{x}$ linear unabhängig von den Kernvektoren $\vec{v}_1^{(i)}, \ldots, \vec{v}_{g_i}^{(i)}$ ist, denn nach Konstruktion gilt $T^{-1}\vec{v}_{g_i}^{(i)} = (0, \ldots, 0, 1, 0, \ldots 0)^T$, wo die 1 unter den ersten g Komponenten des Vektors zu finden sind. (Im Gegensatz dazu hat $\vec{x}$ mindestens einen von 0 verschiedenen Eintrag in den letzten $n-g$ Komponenten.) Damit haben wir einen Widerspruch zu $g_i < n_i$, da wir nun einen weiteren linear unabhängigen Eigenvektor zu λ_i gefunden haben.

Q.E.D.

12.7.3 Diagonalisierung und Funktionalkalkül

Der wesentliche Grund, warum man sich für die Eigenwerte und -vektoren von Matrizen interessiert, ist die Tatsache, dass man mit Hilfe dieser „spektralen Daten“ eine Matrix durch eine „Ähnlichkeitstransformation“ $T^{-1}AT$ auf Diagonalgestalt bringen kann, was das Hantieren mit der Matrix einfacher macht. Dies läuft darauf hinaus, statt der Standardbasis des K^n eine Basis aus Eigenvektoren der Matrix A zu benutzen, um die Vektoren des K^n darzustellen.

Satz 12.97: (Diagonalisierung von Matrizen)

Sei A eine reelle oder komplexe $n \times n$-Matrix mit den (eventuell komplexen) Eigenwerten $\lambda_1, \ldots, \lambda_n$ und entsprechenden Eigenvektoren $\vec{x}_1, \ldots, \vec{x}_n$, also $A\,\vec{x}_k = \lambda_k\,\vec{x}_k$. Sei $T = [\vec{x}_1, \ldots, \vec{x}_n]$ die Matrix, deren Spalten aus diesen Eigenvektoren besteht. Es gilt

$$A\,T = T\,D \quad mit \quad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Sind die Eigenvektoren linear unabhängig, so ist T invertierbar, und es folgt

$$A = T\,D\,T^{-1}.$$

Beweis: Nach Definition der Matrixmultiplikation gilt

$$A\,T = A\,[\vec{x}_1, \ldots, \vec{x}_n] = [A\,\vec{x}_1, \ldots, A\,\vec{x}_n].$$

Mit

$$T\,D = [\vec{x}_1, \ldots, \vec{x}_n] \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = [\lambda_1 \cdot \vec{x}_1, \ldots, \lambda_n \cdot \vec{x}_n]$$

und $A\,\vec{x}_k = \lambda_k \cdot \vec{x}_k$ folgt $A\,T = T\,D$.

Q.E.D.

Folgerung 12.98:

Mit einer Diagonalisierung $A = T\,D\,T^{-1}$ gilt

$$A^n = \underbrace{T\,D\,T^{-1}}_{A}\;\underbrace{T\,D\,T^{-1}}_{A}\;\cdots\;\underbrace{T\,D\,T^{-1}}_{A} = T\,D\,D \cdots D\,T^{-1} = T\,D^n\,T^{-1}.$$

Die Potenzen von A sind damit auf Potenzen der Diagonalform D von A zurückgeführt, wobei D^n sich ohne große Rechnung durch Potenzieren der Diagonalelemente ergibt:

$$\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{pmatrix}^n = \begin{pmatrix} \lambda_1^n & & \\ & \lambda_2^n & \\ & & \ddots \end{pmatrix}.$$

Bemerkung 12.99: *Hat man einen Eigenwert λ_k der zu diagonalisierenden Matrix gefunden, so kann man nach 12.97* **irgendeinen** *dazugehörigen Eigenvektor $\vec{x}_k$ benutzen, die entsprechende Spalte von T zu besetzen. Nun sind Eigenvektoren aber nicht eindeutig: ist $\vec{x}_k$ ein Eigenvektor, so ist jedes Vielfaches $\vec{y}_k = c_k\,\vec{x}_k$ wieder ein Eigenvektor zum selben Eigenwert λ_k. Es gibt damit viele unterschiedliche Transformationsmatrizen T, während die Diagonalmatrix bis auf Umnummerierung der Eigenwerte eindeutig ist. Wie kann das sein? Antwort: die Transformationsmatrizen unterscheiden sich nur um eine Diagonalmatrix:*

$$\tilde{T} = [\vec{y}_1, \ldots, \vec{y}_n] = [c_1\,\vec{x}_1, \ldots, c_n\,\vec{x}_n] = \underbrace{[\vec{x}_1, \ldots, \vec{x}_n]}_{T} \underbrace{\begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}}_{C} = T\,C.$$

In der Diagonalisierung fällt die diagonale „Skalierungsmatrix“ C heraus, da die Multiplikation von Diagonalmatrizen kommutativ ist:

$$A = \tilde{T}\,D\,\tilde{T}^{-1} = T\,C\,D\,(T\,C)^{-1} = T\,C\,D\,C^{-1}\,T^{-1}$$
$$= T\,D\,C\,C^{-1}\,T^{-1} = T\,D\,T^{-1}.$$

Bemerkung 12.100: *Da bei der Diagonalisierung die Invertierbarkeit von T wichtig ist, nützt es überhaupt nichts, triviale Eigenvektoren $\vec{x}_k = 0$ zu betrachten (die ja auch strenggenommen per Definition von Eigenvektoren gar nicht als Eigenvektoren zulässig sind).*

Die Aufgabenstellung der Diagonalisierung läuft darauf hinaus, die Eigenvektoren zu *allen* Eigenwerten zu finden. Sobald man eine Basis von linear unabhängigen Eigenvektoren gefunden hat, hat man die invertierbare Transformationsmatrix T gefunden, welche die betrachtete Matrix auf Diagonalform bringt. Sind alle Eigenwerte verschieden, so existiert immer eine Basis von linear unabhängigen Eigenvektoren. Bei entarteten Eigenwerten (mehrfachen Nullstellen des charakteristischen Polynoms) ist dies jedoch nicht garantiert, und in der Tat gibt es nicht-diagonalisierbare Matrizen (siehe Beispiel 12.91). Symmetrische reelle bzw. hermitesche komplexe Matrizen sind immer diagonalisierbar, selbst wenn die (bei Symmetrie/Hermitzität der Matrix automatisch reellen) Eigenwerte entartet sind.

Beispiel 12.101: Mittels Diagonalisierung können wir für

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

eine explizite Darstellung von A^n für jedes $n \in \mathbb{Z}$ ermitteln. Nach Satz 12.97 sind alle Eigenwerte und –vektoren zu bestimmen. Die Eigenwerte sind die Nullstellen des charakteristischen Polynoms

$$\det(\lambda\,\mathbb{1} - A) = \det\begin{pmatrix} \lambda - 1 & \frac{1}{2} \\ -2 & \lambda - 1 \end{pmatrix} = \lambda^2 - 2\,\lambda + 2$$

(hier ist $\mathbb{1}$ die 2×2-Einheitsmatrix). Nach der Standardformel für die Nullstellen quadratischer Polynome ergeben sich die komplex konjugierten Eigenwerte

$$\lambda_{1,2} = 1 \pm \sqrt{-1} = 1 \pm i.$$

Die Eigenvektoren zu λ_1 bzw. λ_2 findet man, indem man die Kernvektoren von $A - \lambda_k\,\mathbb{1}$ ermittelt, d.h., die Gleichungen $(A - \lambda_k\,\mathbb{1})\,\vec{x}_k = A\,\vec{x}_k - \lambda_k\,\vec{x}_k = \vec{0}$ löst. Mit dem Gauß-Algorithmus 12.64 gefolgt vom Entzerrungsalgorithmus aus Abschnitt 12.5.3 findet man die Lösungen

$$\vec{x}_1 = \begin{pmatrix} i \\ 2 \end{pmatrix} \quad \text{bzw.} \quad \vec{x}_2 = \begin{pmatrix} 1 \\ 2 \cdot i \end{pmatrix}.$$

Die Transformationsmatrix T wird spaltenweise aus diesen Eigenvektoren aufgebaut, die Inverse wird berechnet:

$$T = \begin{pmatrix} i & 1 \\ 2 & 2 \cdot i \end{pmatrix}\ , \quad T^{-1} = \begin{pmatrix} -\frac{i}{2} & \frac{1}{4} \\ \frac{1}{2} & -\frac{i}{4} \end{pmatrix}.$$

Damit folgt die Diagonalisierung

$$A = \begin{pmatrix} 1 & -\frac{1}{2} \\ 2 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} i & 1 \\ 2 & 2 \cdot i \end{pmatrix}}_{T} \underbrace{\begin{pmatrix} 1 + i & 0 \\ 0 & 1 - i \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} -\frac{i}{2} & \frac{1}{4} \\ \frac{1}{2} & -\frac{i}{4} \end{pmatrix}}_{T^{-1}}.$$

Dies liefert

$$A^n = \underbrace{\begin{pmatrix} i & 1 \\ 2 & 2 \cdot i \end{pmatrix}}_{T} \underbrace{\underbrace{\begin{pmatrix} (1+i)^n & 0 \\ 0 & (1-i)^n \end{pmatrix}}_{D^n} \underbrace{\begin{pmatrix} -\frac{i}{2} & \frac{1}{4} \\ \frac{1}{2} & -\frac{i}{4} \end{pmatrix}}_{T^{-1}}}$$

$$= \begin{pmatrix} i & 1 \\ 2 & 2 \cdot i \end{pmatrix} \begin{pmatrix} -\frac{i}{2} \cdot (1+i)^n & \frac{1}{4} \cdot (1+i)^n \\ \frac{1}{2} \cdot (1-i)^n & -\frac{i}{4} \cdot (1-i)^n \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} \cdot (1+i)^n + \frac{1}{2} \cdot (1-i)^n & \frac{i}{4} \cdot (1+i)^n - \frac{i}{4} \cdot (1-i)^n \\ -i \cdot (1+i)^n + i \cdot (1-i)^n & \frac{1}{2} \cdot (1+i)^n + \frac{1}{2} \cdot (1-i)^n \end{pmatrix}.$$

Das Ergebnis muß als Potenz der reellen Matrix A reell sein. Dies ist in dieser Darstellung leider alles andere als offensichtlich. Über die Polardarstellungen

$$1 \pm i = \sqrt{2} \cdot e^{\pm i \cdot \frac{\pi}{4}}$$

ergibt sich mit $|1 \pm i| = \sqrt{2}$ und den Polarwinkeln $\pm\frac{\pi}{4}$:

$$(1 \pm i)^n = 2^{n/2} \cdot e^{\pm n \cdot i \cdot \frac{\pi}{4}} = 2^{n/2} \cdot \Big(\cos\Big(n \cdot \frac{\pi}{4}\Big) \pm i \cdot \sin\Big(n \cdot \frac{\pi}{4}\Big)\Big).$$

Damit folgt

$$\begin{array}{rcrcr} (A^n)_{11} & = & \frac{1}{2} \cdot (1+i)^n + \frac{1}{2} \cdot (1-i)^n & = & 2^{n/2} \cdot \cos(n \cdot \frac{\pi}{4}), \\ (A^n)_{12} & = & \frac{i}{4} \cdot (1+i)^n - \frac{i}{4} \cdot (1-i)^n & = & -2^{n/2-1} \cdot \sin(n \cdot \frac{\pi}{4}), \\ (A^n)_{21} & = & -i \cdot (1+i)^n + i \cdot (1-i)^n & = & 2^{n/2+1} \cdot \sin(n \cdot \frac{\pi}{4}), \\ (A^n)_{22} & = & \frac{1}{2} \cdot (1+i)^n + \frac{1}{2} \cdot (1-i)^n & = & 2^{n/2} \cdot \cos(n \cdot \frac{\pi}{4}). \end{array}$$

Insgesamt erhalten wir also in der Tat eine explizite (und nun recht einfache) reelle Darstellung beliebiger Potenzen von A:

$$\begin{pmatrix} 1 & -\frac{1}{2} \\ 2 & 1 \end{pmatrix}^n = \begin{pmatrix} 2^{\frac{n}{2}} \cdot \cos(n \cdot \frac{\pi}{4}) & -2^{\frac{n}{2}-1} \cdot \sin(n \cdot \frac{\pi}{4}) \\ 2^{\frac{n}{2}+1} \cdot \sin(n \cdot \frac{\pi}{4}) & 2^{\frac{n}{2}} \cdot \cos(n \cdot \frac{\pi}{4}) \end{pmatrix}.$$

Satz 12.102: (Hamilton-Cayley-Theorem)

Das charakteristische Polynom von $A \in K^{n \times n}$ sei

$$p_A(\lambda) = (-1)^n \cdot \lambda^n + c_{n-1} \cdot \lambda^{n-1} + \ldots + c_1 \cdot \lambda + c_0.$$

Dann gilt die Matrix-Gleichung

$$p_A(A) = (-1)^n \cdot A^n + c_{n-1} \cdot A^{n-1} + \ldots + c_1 \cdot A + c_0 \cdot \mathbb{1} = 0 \in K^{n \times n}.$$

Beweis: Für diagonalisierbares $A = TDT^{-1}$ mit $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ ist der Beweis einfach. Mit $A^k = TD^kT^{-1}$ und $D^k = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)$ folgt

$$p_A(A) = p_A(TDT^{-1}) = Tp_A(D)T^{-1} = T \begin{pmatrix} p_A(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & p_A(\lambda_n) \end{pmatrix} T^{-1}$$

$$= T \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix} T^{-1} = \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

Die Aussage gilt auch für nicht-diagonalisierbare Matrizen A, wobei der Beweis aber deutlich aufwendiger wird (das lassen wir hier).

Q.E.D.

Für Diagonalmatrizen $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ gilt $D^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. Man nimmt dies zum Anlass, für beliebige Funktion $f : \mathbb{R} \to \mathbb{R}$ bzw. $f : \mathbb{C} \to \mathbb{C}$ zu definieren: $f(D) = \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))$.

Definition 12.103: (Funktionalkalkül)

Für eine Funktionen $f : \mathbb{R} \to \mathbb{R}$ bzw. $f : \mathbb{R} \to \mathbb{R}$ definieren wir die Matrix-Auswertung dieser Funktion für eine diagonalisierbare Matrix $A = TDT^{-1} \in K^{n\times n}$ mit $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ mittels

$$f(A) = Tf(D)T^{-1} = T \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} T^{-1}.$$

Beispiel 12.104: Wir betrachten $f(x) = e^x$. Stellt man sich die exp-Funktion als die Potenzreihe

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

vor, so liegt es nahe, die Matrixauswertung für $A \in K^{n\times n}$ durch

$$e^A = \mathbb{1} + A + \frac{1}{2!} \cdot A^2 + \frac{1}{3!} \cdot A^3 + \ldots \qquad (\#)$$

zu definieren. Für diagonalisierbares $A = TDT^{-1}$ mit $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ stimmt dies mit der obigen Definition 12.103 überein:

$$\begin{aligned} e^A &= \mathbb{1} + TDT^{-1} + \frac{1}{2!} \cdot (TDT^{-1})^2 + \frac{1}{3!} \cdot (TDT^{-1})^3 + \ldots \\ &= T\Big(\mathbb{1} + D + \frac{1}{2!} \cdot D^2 + \frac{1}{3!} \cdot D^3 + \ldots\Big)T^{-1} \\ &= T \begin{pmatrix} 1 + \lambda_1 + \frac{\lambda_1^2}{2!} + \ldots & & 0 \\ & \ddots & \\ 0 & & 1 + \lambda_n + \frac{\lambda_n^2}{2!} + \ldots \end{pmatrix} T^{-1} \\ &= T \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix} T^{-1}. \end{aligned}$$

Die Definition (#) über die Potenzreihe ist dabei auch für nicht-diagonalisierbares A sinnvoll.

Hier eine Anwendung:

Satz 12.105: (dynamische Systeme)

Sei $A \in K^{n\times n}$ eine konstante Matrix. Die Lösung des Differentialgleichungssystems

$$\frac{d}{dt}\,\vec{x}(t) = A\,\vec{x}(t)$$

ist

$$\vec{x}(t) = e^{t\cdot A}\,\vec{x}(0).$$

Beweis: Es gilt

$$\frac{d}{dt}\,e^{tA} = \frac{d}{dt}\Big(\mathbb{1} + t\cdot A + \frac{t^2}{2!}\cdot A^2 + \frac{t^3}{3!}\cdot A^3 + \ldots\Big) = A + t\cdot A^2 + \frac{t^2}{2!}\cdot A^3 + \ldots$$

$$= A\Big(\mathbb{1} + t\cdot A + \frac{t^2}{2!}\cdot A^2 + \frac{t^3}{3!}\cdot A^3 + \ldots\Big) = A\,e^{t\cdot A}.$$

Damit ist $\vec{x}(t) = e^{t\cdot A}\,\vec{x}(0)$ in der Tat eine Lösung der DGL:

$$\frac{d}{dt}\,\vec{x}(t) = \frac{d}{dt}\,e^{t\cdot A}\vec{x}(0) = A\,e^{tA}\,\vec{x}(0) = A\,\vec{x}(t).$$

Q.E.D.

Beispiel 12.106: Betrachte das DGL-System

$$\frac{dx_1}{dt} = x_1 - \frac{x_2}{2}, \quad \frac{dx_2}{dt} = 2\cdot x_1 + x_2,$$

das als

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

geschrieben wird. In Beispiel 12.101 war die Diagonalisierung

$$\begin{pmatrix} 1 & -\frac{1}{2} \\ 2 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} i & 1 \\ 2 & 2\cdot i \end{pmatrix}}_{T}\underbrace{\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}}_{D}\underbrace{\begin{pmatrix} -\frac{i}{2} & \frac{1}{4} \\ \frac{1}{2} & -\frac{i}{4} \end{pmatrix}}_{T^{-1}}$$

gefunden worden. Die Lösung der DGL ist damit

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \underbrace{\underbrace{\begin{pmatrix} i & 1 \\ 2 & 2\cdot i \end{pmatrix}}_{T}\underbrace{\begin{pmatrix} e^{t\cdot(1+i)} & 0 \\ 0 & e^{t\cdot(1-i)} \end{pmatrix}}_{e^{t\cdot D}}\underbrace{\begin{pmatrix} -\frac{i}{2} & \frac{1}{4} \\ \frac{1}{2} & -\frac{i}{4} \end{pmatrix}}_{T^{-1}}}_{e^{t\cdot A}}\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}$$

$$= \begin{pmatrix} \cos(t)\cdot e^t & -\frac{1}{2}\cdot\sin(t)\cdot e^t \\ 2\cdot\sin(t)\cdot e^t & \cos(t)\cdot e^t \end{pmatrix}\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix},$$

also

$$x_1(t) = \cos(t) \cdot e^t \cdot x_1(0) - \frac{x_2(0)}{2} \cdot \sin(t) \cdot e^t,$$
$$x_2(t) = 2 \cdot \sin(t) \cdot e^t \cdot x_1(0) + x_2(0) \cdot \cos(t) \cdot e^t.$$