Merkblatt Matrixdarstellungen

Franz Embacher

In diesem kleinen Text werden einige praktisch relevante Sachverhalte im Zusammenhang mit Darstellungsmatrizen von Vektoren und linearen Abbildungen zusammengestellt.

1 Konventionen

Im Folgenden sei V ein n-dimensionaler Vektorraum und W ein m-dimensionaler Vektorraum, beide über demselben Körper $\mathbb{K}$. (Wir sagen dann kurz, dass V und W $\mathbb{K}$-Vektorräume sind.) In V seien geordnete Basen

$$\mathcal{B} = (f_1, \ldots, f_n) \qquad \text{und} \qquad \mathcal{B}' = (f'_1, \ldots, f'_n), \tag{1}$$

und in W seien geordnete Basen

$$\mathcal{C} = (g_1, \ldots, g_m) \qquad \text{und} \qquad \mathcal{C}' = (g'_1, \ldots, g'_m) \tag{2}$$

gewählt. Mit „Basis" ist immer eine geordnete Basis gemeint, also ein Tupel von Vektoren, keine Menge. Den j-ten Vektor der Standardbasis jedes $\mathbb{K}^p$ wollen wir mit e_j bezeichnen.

Das Wort „Vektor" bezeichnet ganz allgemein ein Element eines Vektorraums, während mit dem Ausdruck „p-Vektor" ein Element von $\mathbb{K}^p$ gemeint ist. Generell wollen wir p-Vektoren als Spaltenvektoren (d.h. als $p \times 1$-Matrizen) auffassen[1].

Weiters sei $\varphi : V \longrightarrow W$ eine lineare Abbildung. Für $v \in V$ nennen wir $\varphi(v)$ das Bild (oder Bildelement) von v.

Im Spezialfall $V = \mathbb{K}^n$ und $W = \mathbb{K}^m$ ist die Wirkung von φ unmittelbar durch die Anwendung einer Matrix $A \in \mathbb{K}^{m\times n}$ gegeben[2], d.h.

$$\begin{aligned} \varphi : \mathbb{K}^n &\longrightarrow \mathbb{K}^m \\ \varphi(x) &= Ax. \end{aligned} \tag{3}$$

Dieser Zusammenhang zwischen linearen Abbildungen $\mathbb{K}^n \longrightarrow \mathbb{K}^m$ und $m\times n$-Matrizen über $\mathbb{K}$ ist so eng, dass man zwischen φ und A oft keinen großen Unterschied macht und sie durchaus auch mit demselben Symbol bezeichnet. Dann heißt die lineare Abbildung etwa A, sie führt x in $A(x)$ über, was salopp als Ax angeschrieben und als Matrizenmultiplikation gedeutet wird – womit das Symbol A nun auch für die Matrix steht. Wir wollen in diesem Text ein klein

[1] p-Vektoren sind p-Tupeln, also Listen von Elementen von $\mathbb{K}$, unabhängig davon, wie man sie anschreibt. Sobald aber die Matrzenrechnung ist Spiel kommt, muss man zwischen Spalten- und Zeilenvektoren unterscheiden. In *diesem* Sinn fassen wir p-Vektoren als Spaltenvektoren auf. Ist x ein solcher Spaltenvektor, so ist x^T ein Zeilenvektor, wobei das Symbol ^{T} das Transponieren bezeichnet.

[2] Wie Sie sehen, bezeichnen wir die Menge aller $m \times n$-Matrizen über $\mathbb{K}$ mit $\mathbb{K}^{m\times n}$.

wenig pingeliger sein und diese Identifizierung der Klarheit halber nicht machen, lediglich in Abschnitt 7 klingt sie wieder an.

Die individuellen Komponenten (Koeffizienten, Eintragungen) einer Matrix schreiben wir bei Bedarf an, indem dem Symbol für die Matrix die entsprechenen Indizes beigefügt werden. So bezeichnet A_{jk} die Eintragung in der j-ten Zeile und k-ten Spalte der Matrix A. Für festgehaltenes k sind A_{jk} die Komponenten der k-ten Spalte von A. Die j-te Komponente eines p-Vektors x bezeichnen wir mit x_j.

Die identische Abbildung eines Vektorraums V (also $v \longmapsto v$) bezeichnen wir mit id_V, die $n \times n$-Einheitsmatrix mit I_n und deren Komponenten wie üblich mit δ_{jk} (gleich 1 für $j = k$ und 0 für $j \neq k$).

2 Matrixdarstellung eines Vektors

Jeder Vektor $v \in V$ kann in die Basis $\mathcal{B}$ entwickelt (d.h. als Linearkombination der Basisvektoren geschrieben) werden,

$$v = \sum_{j=1}^{n} c_j \, f_j, \tag{4}$$

wobei die Entwickungskoeffizienten $c_j \in \mathbb{K}$ durch v eindeutig bestimmt sind. Den n-Vektor

$$[v]_{\mathcal{B}} := \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{K}^n \tag{5}$$

der Entwicklungskoeffizienten nennen wir die Matrixdarstellung von v bezüglich der Basis $\mathcal{B}$. Die Zuordnung $v \longmapsto [v]_{\mathcal{B}}$ bezeichnen wir mit $\Psi_{\mathcal{B}}$:

$$\begin{aligned} \Psi_{\mathcal{B}} : V &\longrightarrow \mathbb{K}^n \\ \Psi_{\mathcal{B}}(v) &= [v]_{\mathcal{B}}. \end{aligned} \tag{6}$$

Sie ist linear und bijektiv, also ein (Vektorraum-)Isomorphismus. Das bedeutet, dass jeder n-dimensionale $\mathbb{K}$-Vektorraum zu $\mathbb{K}^n$ isomorph ist, d.h. hinsichlich der Vektorraumstruktur mit $\mathbb{K}^n$ identifiziert werden kann. (In manchen Lehrbüchern wird statt $\Psi_{\mathcal{B}}$ die dazu inverse Abbildung $\Phi_{\mathcal{B}} : \mathbb{K}^n \longrightarrow V$ betrachtet, die also dem n-Vektor der Entwicklungskoeffizienten das zugehörige Element von V zuordnet. Übersetzen Sie dann an jeder Stelle einfach $\Psi_{\mathcal{B}} = {\Phi_{\mathcal{B}}}^{-1}$ und ${\Psi_{\mathcal{B}}}^{-1} = \Phi_{\mathcal{B}}$.)

In W kann man dasselbe machen, die Matrixdarstellung von $w \in W$ bezüglich der Basis $\mathcal{C}$ wird mit $[w]_{\mathcal{C}}$ bezeichnet, die Abbildung $w \longmapsto [w]_{\mathcal{C}}$ mit $\Psi_{\mathcal{C}}$:

$$\begin{aligned} \Psi_{\mathcal{C}} : W &\longrightarrow \mathbb{K}^m \\ \Psi_{\mathcal{C}}(w) &= [w]_{\mathcal{C}}. \end{aligned} \tag{7}$$

3 Eine triviale, aber wichtige Eigenschaft jeder linearen Abbildung $\varphi : \mathbb{K}^n \longrightarrow \mathbb{K}^m$

Eine lineare Abbildung $\varphi : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ ist, wie bereits besprochen, durch die Anwendung einer Matrix $A \in \mathbb{K}^{m\times n}$ gegeben, siehe (3).

Merkregel 1: Die j-te Spalte von A ist das Bild des j-ten Vektors der Standardbasis von $\mathbb{K}^n$.

Beweis: Das Bild des ersten Vektors e_1 der Standardbasis von $\mathbb{K}^n$ berechnet sich zu

$$\varphi(e_1) = Ae_1 = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{nm} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{m1} \end{pmatrix}, \tag{8}$$

und analog funktioniert es mit dem j-ten Basisvektor e_j.

Wieso wird eine solch triviale Eigenschaft überhaupt angeschrieben? Sie ist nützlich, wenn man die Matrix A noch *nicht* kennt, sondern aus der Wirkung von φ erschließen möchte, und wir werden sie im Folgenden oft benutzen, um andere, weniger triviale Merkregeln zu beweisen.

4 Matrixdarstellung einer linearen Abbildung

Sei $\varphi : V \longrightarrow W$ eine lineare Abbildung, und seien die Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W fix gewählt. Da V durch die Abbildung $\Psi_\mathcal{B}$ mit $\mathbb{K}^n$ und W durch die Abbildung $\Psi_\mathcal{C}$ mit $\mathbb{K}^m$ identifiziert werden kann, kann φ – sozusagen aus dem Blickwinkel dieser beiden Basen – als Abbildung $\mathbb{K}^n \longrightarrow \mathbb{K}^m$ dargestellt werden und daher auch durch eine $m \times n$-Matrix über $\mathbb{K}$. Das macht man sich am besten mit Hilfe eines kommutativen Diagramms klar:

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\Psi_\mathcal{B}"'] & W \arrow[d, "\Psi_\mathcal{C}"] \\ \mathbb{K}^n \arrow[r, "\varphi_{\mathcal{C},\mathcal{B}}"'] & \mathbb{K}^m \end{tikzcd} \tag{9}$$

Die lineare Abbildung $\mathbb{K}^n \longrightarrow \mathbb{K}^m$, die sich durch den unteren Pfeil ergibt, wurde hier mit $\varphi_{\mathcal{C},\mathcal{B}}$ bezeichnet. Wir lesen ab:

$$\varphi_{\mathcal{C},\mathcal{B}} = \Psi_\mathcal{C} \circ \varphi \circ {\Psi_\mathcal{B}}^{-1} \qquad \text{oder, umgekehrt:} \qquad \varphi = {\Psi_\mathcal{C}}^{-1} \circ \varphi_{\mathcal{C},\mathcal{B}} \circ \Psi_\mathcal{B}\,. \tag{10}$$

Die Wirkung von $\varphi_{\mathcal{C},\mathcal{B}}$ ist durch die Anwendung einer $m \times n$-Matrix über $\mathbb{K}$ gegeben, die wir mit $[\varphi]_{\mathcal{C},\mathcal{B}}$ bezeichnen wollen. Sie heißt die Matrixdarstellung (oder Darstellungsmatrix) von φ

bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$. Im Spezialfall $W = V$ und $\mathcal{C} = \mathcal{B}$ wollen wir für sie einfach $[\varphi]_{\mathcal{B}}$ anstelle von $[\varphi]_{\mathcal{B},\mathcal{B}}$ schreiben.

Zeichnen wir das kommutative Diagramm (9) noch einmal auf, aber nun für einen konkreten Ausgangsvektor $v \in V$:

$$\begin{tikzcd} v \arrow[r, mapsto, "\varphi"] \arrow[d, mapsto, "\Psi_{\mathcal{B}}"'] & \varphi(v) \arrow[d, mapsto, "\Psi_{\mathcal{C}}"] \\ {[v]_{\mathcal{B}}} \arrow[r, mapsto, "\varphi_{\mathcal{C},\mathcal{B}}"'] & {[\varphi(v)]_{\mathcal{C}}} \end{tikzcd} \tag{11}$$

Daraus ergibt sich unter Verwendung der Matrixdarstellung $[\varphi]_{\mathcal{C},\mathcal{B}}$ von $\varphi_{\mathcal{C},\mathcal{B}}$

$$[\varphi(v)]_{\mathcal{C}} = [\varphi]_{\mathcal{C},\mathcal{B}}\,[v]_{\mathcal{B}}, \tag{12}$$

also, wie es sein muss, m-Vektor $=$ $m \times n$-Matrix mal n-Vektor. Damit ist die Wirkung der linearen Abbildung φ, durch die Brille der verwendeten Basen betrachtet, vollständig auf Matrizenrechnung (also auf die Manipulation von Zahlen) zurückgeführt. Alle sonstigen Eigenschaften der Vektorräume V, W und ihrer Elemente sowie der linearen Abbildung φ spielen in (12) keine Rolle mehr.

Wie die Matrixdarstellung einer linearen Abbildung zustande kommt, kann man auch durch die Entwicklungskoeffizienten bezüglich der jeweiligen Basis einsehen: Mit

$$v = \sum_{j=1}^{n} c_j\, f_j \qquad \text{(Entwicklung von } v \text{ in die Basis } \mathcal{B}\text{)} \tag{13}$$

und

$$\varphi(v) = \sum_{k=1}^{m} d_k\, g_k \qquad \text{(Entwicklung von } \varphi(v) \text{ in die Basis } \mathcal{C}\text{)} \tag{14}$$

ist $[\varphi]_{\mathcal{C},\mathcal{B}}$ gerade jene $m \times n$-Matrix, die

$$[v]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{K}^n \qquad \text{in} \qquad [\varphi(v)]_{\mathcal{C}} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} \in \mathbb{K}^m \qquad \text{überführt.} \tag{15}$$

Und wie finden wir in der Praxis die Darstellungsmatrix $[\varphi]_{\mathcal{C},\mathcal{B}}$, wenn φ, $\mathcal{B}$ und $\mathcal{C}$ gegeben sind? Mit Hilfe von

Merkregel 2: Die j-te Spalte von $[\varphi]_{\mathcal{C},\mathcal{B}}$ besteht aus den Entwicklungskoeffizienten von $\varphi(f_j)$ bezüglich der Basis $\mathcal{C}$.

Beweis: Wegen

$$f_j = 0 \cdot f_1 + \cdots + 1 \cdot f_j + \cdots + 0 \cdot f_n \tag{16}$$

ist $[f_j]_{\mathcal{B}} = e_j$. Wird $v = f_j$ in (12) eingesetzt, so ergibt sich

$$[\varphi(f_j)]_{\mathcal{C}} = [\varphi]_{\mathcal{C},\mathcal{B}}\, e_j \overset{\text{Merkregel 1}}{=} j\text{-te Spalte von } [\varphi]_{\mathcal{C},\mathcal{B}}. \tag{17}$$

Da der m-Vektor $[\varphi(f_j)]_{\mathcal{C}}$ aus den Entwicklungskoeffizienten von $\varphi(f_j) \in W$ bezüglich der Basis $\mathcal{C}$ besteht, folgt die behauptete Aussage.

5 Matrixdarstellungen in einem euklidischen oder unitären Vektorraum

Ist $\mathbb{K}$ geich $\mathbb{R}$ oder $\mathbb{C}$, und sind V und W eukldische oder unitäre Vektorräume (also jeweils mit einem Skalarprodukt $\langle\, .\, ,\, .\, \rangle$ ausgestattet, das im unitären Fall im zweiten Argument linear sei), so wählt man zweckmäßigerweise $\mathcal{B}$ und $\mathcal{C}$ als Orthonormalbasen. In diesem Fall gilt natürlich alles in Abschnitt 4 Gesagte, und es gelten zusätzlich einige Beziehungen, die die Ermittlung von Darstellungsmatrizen stark vereinfachen. Die Entwicklungskoeffizienten in (13) und (14) sind dann durch

$$c_j = \langle f_j, v \rangle \tag{18}$$

und

$$d_k = \langle g_k, \varphi(v) \rangle \tag{19}$$

gegeben. Daraus folgt sofort

Merkregel 3: Die kj-Komponente der Darstellungsmatrix $[\varphi]_{\mathcal{C},\mathcal{B}}$ ist gegeben durch

$$\big([\varphi]_{\mathcal{C},\mathcal{B}}\big)_{kj} = \langle g_k, \varphi(f_j) \rangle. \tag{20}$$

Beweis:

$$d_k = \langle g_k, \varphi(v) \rangle = \left\langle g_k, \varphi\left(\sum_{j=1}^{n} c_j\, f_j\right)\right\rangle = \sum_{j=1}^{n} \underbrace{\langle g_k, \varphi(f_j) \rangle}_{\left([\varphi]_{\mathcal{C},\mathcal{B}}\right)_{kj}} c_j\,. \tag{21}$$

Beachten Sie, dass man die Komponenten $\big([\varphi]_{\mathcal{C},\mathcal{B}}\big)_{kj}$ hier durch simples „Ablesen“ bekommt: Ist $x \mapsto y$ linear, also durch Anwendung einer Matrix M dargestellt, und ist $y_k = \sum_j M_{kj} x_j$ für alle x und für alle k, so sind M_{kj} die Komponenten von M.

6 Basiswechsel

Manche Fragestellungen an lineare Abbildungen werden durch die Verwendung geeigneter Basen stark vereinfacht. Da man solche geeignete Basen aber in der Regel zunächst nicht kennt, beginnt man mit irgendwelchen Basen $\mathcal{B}$ und $\mathcal{C}$, um später zu anderen (geeigneteren) Basen $\mathcal{B}'$ und $\mathcal{C}'$ überzugehen. Damit stellt sich die Frage, wie sich die Darstellungsmatrix einer linearen Abbildung unter einem Basiswechsel (einer Basistransformation) ändert. Wir betrachten also den Übergang von den „alten“ Basen $\mathcal{B}$ und $\mathcal{C}$ zu den „neuen“ Basen $\mathcal{B}'$ und $\mathcal{C}'$.

Grundlegend für das Folgende sind die Formeln und Diagramme von Abschnitt 4 und dieselben Formeln und Diagramme für die Basen $\mathcal{B}'$ und $\mathcal{C}'$ anstelle von $\mathcal{B}$ und $\mathcal{C}$. Der komplette Sachverhalt des Basiswechsels kann in einem alle relevanten Abbildungen einschließenden kommutativen Diagramm dargestellt werden:

$$\begin{tikzcd}
\mathbb{K}^n \arrow[r, "\varphi_{\mathcal{C}',\mathcal{B}'}"'] & \mathbb{K}^m \\
V \arrow[u, "\Psi_{\mathcal{B}'}"] \arrow[r, "\varphi"] \arrow[d, "\Psi_{\mathcal{B}}"'] & W \arrow[u, "\Psi_{\mathcal{C}'}"'] \arrow[d, "\Psi_{\mathcal{C}}"] \\
\mathbb{K}^n \arrow[r, "\varphi_{\mathcal{C},\mathcal{B}}"'] & \mathbb{K}^m
\end{tikzcd} \tag{22}$$

Das Ziel ist es, $\varphi_{\mathcal{C}',\mathcal{B}'}$ durch $\varphi_{\mathcal{C},\mathcal{B}}$ und damit $[\varphi]_{\mathcal{C}',\mathcal{B}'}$ durch $[\varphi]_{\mathcal{C},\mathcal{B}}$ auszudrücken. Dazu kann man direkt das obige Diagramm nutzen oder die Beziehungen (10) und die entsprechende Variante für die neuen Basen, also

$$\varphi_{\mathcal{C}',\mathcal{B}'} = \Psi_{\mathcal{C}'} \circ \varphi \circ {\Psi_{\mathcal{B}'}}^{-1} \qquad \text{oder, umgekehrt:} \qquad \varphi = {\Psi_{\mathcal{C}'}}^{-1} \circ \varphi_{\mathcal{C}',\mathcal{B}'} \circ \Psi_{\mathcal{B}'}\,, \tag{23}$$

verwenden. Wie auch immer, es ergibt sich

$$\varphi_{\mathcal{C}',\mathcal{B}'} = \Psi_{\mathcal{C}'} \circ {\Psi_{\mathcal{C}}}^{-1} \circ \varphi_{\mathcal{C},\mathcal{B}} \circ \Psi_{\mathcal{B}} \circ {\Psi_{\mathcal{B}'}}^{-1}. \tag{24}$$

Beachten Sie, dass die hier auftretenden linearen Abbildungen

$$\Psi_{\mathcal{C}'} \circ {\Psi_{\mathcal{C}}}^{-1} : \mathbb{K}^m \longrightarrow \mathbb{K}^m \tag{25}$$
$$\Psi_{\mathcal{B}} \circ {\Psi_{\mathcal{B}'}}^{-1} : \mathbb{K}^n \longrightarrow \mathbb{K}^n \tag{26}$$

nur von den verwendeten Basen (also nicht von φ) abhängen und beide bijektiv (also invertierbar) sind. Sie können daher durch die Wirkungen invertierbarer quadratischer Matrizen dargestellt werden. Die Matrix, deren Wirkung (25) darstellt, wollen wir T^{-1} nennen, die Matrix, deren Wirkung (26) darstellt, wollen wir S nennen. Damit können wir (24) in Matrixnotation anschreiben:

$$[\varphi]_{\mathcal{C}',\mathcal{B}'} = T^{-1}\,[\varphi]_{\mathcal{C},\mathcal{B}}\,S. \tag{27}$$

Das ist die gesuchte Basistransformationsformel für Darstellungsmatrizen linearer Abbildungen. Machen wir uns noch klar, wie die Matrizen S und T mit dem Basiswechsel zusammenhängen: Für vorgegebene $v \in V$ und $w \in W$ wirken die Abbildungen (26) und (25) so:

$$\Psi_{\mathcal{B}} \circ {\Psi_{\mathcal{B}'}}^{-1} : [v]_{\mathcal{B}'} \longmapsto [v]_{\mathcal{B}} = S\,[v]_{\mathcal{B}'} \tag{28}$$
$$\Psi_{\mathcal{C}'} \circ {\Psi_{\mathcal{C}}}^{-1} : [w]_{\mathcal{C}} \longmapsto [w]_{\mathcal{C}'} = T^{-1}\,[w]_{\mathcal{C}}\,, \tag{29}$$

und durch Invertieren dieser Beziehungen ergibt sich

$$\Psi_{\mathcal{B}'} \circ {\Psi_{\mathcal{B}}}^{-1} : [v]_{\mathcal{B}} \longmapsto [v]_{\mathcal{B}'} = S^{-1}\,[v]_{\mathcal{B}} \tag{30}$$
$$\Psi_{\mathcal{C}} \circ {\Psi_{\mathcal{C}'}}^{-1} : [w]_{\mathcal{C}'} \longmapsto [w]_{\mathcal{C}} = T\,[w]_{\mathcal{C}'}\,. \tag{31}$$

S und T führen Entwicklungskoeffizienten bezüglich der neuen Basen in Entwicklungskoeffizienten bezüglich der alten Basen über (und S^{-1} und T^{-1} genau umgekehrt). Explizit durch die Entwicklungskoeffizienten ausgedrückt sehen (28) und (31) so aus:

$$v = \sum_{j=1}^{n} c_j\,f_j = \sum_{j=1}^{n} c'_j\,f'_j \implies c_j = \sum_{i=1}^{n} S_{ji}\,c'_i \tag{32}$$
$$w = \sum_{j=1}^{n} d_k\,g_k = \sum_{k=1}^{n} d'_k\,g'_k \implies d_k = \sum_{l=1}^{n} T_{kl}\,d'_l\,. \tag{33}$$

Für den Übergang von den alten zu den neuen Basisvektoren bedeutet das

$$f'_j = \sum_{i=1}^{n} f_i\,S_{ij} \tag{34}$$
$$g'_k = \sum_{l=1}^{m} g_l\,T_{lk}\,. \tag{35}$$

Die Komponenten von S und T sind also jeweils genau die Koeffizienten der Entwicklung der neuen Basisvektoren in die alte Basis. (Falls Sie oft mit diesen Dingen zu tun haben, sehen Sie sich die Unterschiede zwischen (32) – (33) und (34) – (35) hinsichtlich der Rolle der Indizes von S und T und der Rolle der gestrichenen/ungestrichenen Größen genau an!) Daraus ergibt sich unmittelbar

Merkregel 4: Die j-te Spalte von S ist $[f'_j]_{\mathcal{B}}$ (sie besteht also aus den Entwicklungskoeffizienten von f'_j bezüglich der Basis $\mathcal{B}$). Die k-te Spalte von T ist $[g'_k]_{\mathcal{C}}$, sie besteht also aus den Entwicklungskoeffizienten von g'_k bezüglich der Basis $\mathcal{C}$.

> Beweis: Setzen Sie $j = 1$ in (34). Die Entwicklungskoeffizienten von f'_1 bezüglich der Basis $\mathcal{B}$ sind dann durch S_{i1} gegeben – das sind aber genau die Komponenten der ersten Spalte von S. Für alle anderen j geht das analog: S_{ij} sind die Komponenten der j-ten Spalte von S.

Sie können statt dessen, wenn Ihnen das mehr zusagt, auch $v = f'_j$ in (28) einsetzen. Da $[f'_j]_{\mathcal{B}'} = e_j$ ist, folgt $[f'_j]_{\mathcal{B}} = S\,[f'_j]_{\mathcal{B}'} = Se_j$, was wegen Merkregel 1 gleich der j-ten Spalte von S ist.

Für die k-te Spalte von T funktioniert alles genauso.

Wir wollen noch erwähnen, dass man S und T auch auf eine andere, etwas raffinierte Weise darstellen kann:

Merkregel 5: Es gilt

$$S = [\mathrm{id_V}]_{\mathcal{B},\mathcal{B}'} \tag{36}$$

$$T = [\mathrm{id_W}]_{\mathcal{C},\mathcal{C}'}\,, \tag{37}$$

d.h. S ist Darstellungsmatrix der identischen Abbildung $V \longrightarrow V$, aber bezogen auf zwei unterschiedliche Basen von V, und T analog in W.

Beweis: Aus $v = \mathrm{id}_V(v)$ folgt $[v]_{\mathcal{B}} = [\mathrm{id}_V]_{\mathcal{B},\mathcal{B}'}\,[v]_{\mathcal{B}'}$, und mit $v = f'_j$ und $[f'_j]_{\mathcal{B}'} = e_j$ ergibt sich (Merkregel 1), dass $[f'_j]_{\mathcal{B}}$ die j-te Spalte von $[\mathrm{id}_V]_{\mathcal{B},\mathcal{B}'}$ ist. Nach Merkregel 4 ist das aber auch die j-te Spalte von S. Da $[\mathrm{id}_V]_{\mathcal{B},\mathcal{B}'}$ und S in allen Spalten übereinstimmen, sind sie gleich. Daher gilt auch (37), weil in W der analoge Sachverhalt vorliegt.

Im Spezialfall $W = V$ (daher $m = n$), also

$$\varphi : V \longrightarrow V, \tag{38}$$

einer Situation, die in vielen Anwendungsfällen vorliegt, kann man $\mathcal{C} = \mathcal{B}$ und $\mathcal{C}' = \mathcal{B}'$ (d.h. $g_j = f_j$ und $g'_j = f'_j$ für $i = 1, \ldots, n$) wählen. Der Basiswechsel besteht dann nur darin, von $\mathcal{B}$ zu $\mathcal{B}'$ überzugehen. In diesem Fall gilt $T = S$, und (27) reduziert sich auf die eminent wichtige Beziehung

$$[\varphi]_{\mathcal{B}'} = S^{-1}\,[\varphi]_{\mathcal{B}}\,S. \tag{39}$$

7 „Matrixdarstellung einer Matrix“

Das im vorigen Abschnitt Gesagte kann natürlich auch auf eine lineare Abbildung $\varphi : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, deren Wirkung unmittelbar durch die Anwendung einer Matrix $A \in \mathbb{K}^{m\times n}$ gegeben ist (siehe Abschnitt 3), angewandt werden. Gemäß der bisher entwickelten Sprache ist dann A die Darstellungsmatrix von φ bezüglich der Standardbasen von $\mathbb{K}^n$ und $\mathbb{K}^m$. Wie in Abschnitt 1 besprochen, wird nicht immer zwischen φ und A unterschieden. So gesehen ist die Matrix A ihre eigene Darstellungsmatrix, aber auf die Standardbasen von $\mathbb{K}^n$ und $\mathbb{K}^m$ bezogen. Wechselt man zu anderen Basen $\mathcal{B}'$ von $\mathbb{K}^n$ und $\mathcal{C}'$ von $\mathbb{K}^m$, so wird die Abbildung, die ursprünglich durch A dargestellt war, durch eine Matrix B dargestellt, womit (27) die gebräuchlichere Form

$$B = T^{-1}AS. \tag{40}$$

annimmt. Wenn man also salopperweise A statt φ sagt, muss man B als Darstellungsmatrix von A bezüglich der Basen $\mathcal{B}'$ und $\mathcal{C}'$ ansehen.

Im Spezialfall $m = n$ und $\mathcal{C}' = \mathcal{B}'$ reduziert sich (39) auf

$$B = S^{-1}AS, \tag{41}$$

was sofort Anlass zur Definition gibt, dass zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ zueinander ähnlich heißen, wenn es eine invertierbare Matrix $S \in \mathbb{K}^{n\times n}$ gibt, für die (41) gilt. A und B stellen dann dieselbe lineare Abbildung dar, nur aus dem Blickwinkel verschiedener Basen betrachtet. Sie teilen viele Eigenschaften, insbesondere haben sie die gleiche Determinante, die gleiche Spur, den gleichen Rang und die gleichen Eigenwerte (mit gleicher algebraischer und geometrischer Vielfachheit).

8 Diagonalisierung

Eine lineare Abbildung $\varphi : V \longrightarrow V$ heißt diagonalisierbar, wenn V eine Basis aus Eigenvektoren von φ besitzt. Die Darstellungsmatrix von φ bezüglich einer solchen Basis ist dann eine $n \times n$-Diagonalmatrix mit den Eigenwerten von φ auf der Hauptdiagonale.

Analog heißt eine Matrix $A \in \mathbb{K}^{n\times n}$ diagonalisierbar, wenn $\mathbb{K}^n$ eine Basis aus Eigenvektoren von A besitzt. Auf diese Basis bezogen, wird die Wirkung der durch A definierten linearen Abbildung dann durch eine $n \times n$-Diagonalmatrix mit den Eigenwerten von A auf der Hauptdiagonale dargestellt.

Aufgrund der strukturellen Gleichheit dieser beiden Situationen beschränken wir uns auf den zweiten Fall: Wie diagonalisiert man eine Matrix? Ist A diagonalisierbar, so ist ein Basiswechsel von der Standardbasis des $\mathbb{K}^n$ zu einer Basis aus Eigenvektoren durchzuführen. Im Lichte der Transformationsformel (41) bedeutet das, eine invertierbare Matrix S zu finden, für die $S^{-1}AS$ diagonal ist. Wie findet man eine solche Matrix, und wie hängt sie mit der Basis aus Eigenvektoren zusammen?

Merkregel 6: Sind $A, S, D \in \mathbb{K}^{n\times n}$ und D diagonal, so gilt $AS = SD$ genau dann, wenn jede Spalte von S entweder gleich 0 (also der Nullvektor) oder ein Eigenvektor von A ist.

> Beweis: Wir bezeichnen $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ und schreiben die Beziehung $AS = SD$ in Komponenten an:
>
> $$\sum_{k=1}^{n} A_{ik} S_{kj} = \lambda_j S_{ij}\,, \tag{42}$$
>
> was geichbedeutend damit ist, dass die j-te Spalte von S entweder gleich 0 oder ein Eigenvektor von A zum Eigenwert λ_j ist.

Merkregel 7: Sind $A, S \in \mathbb{K}^{n\times n}$ und S invertierbar, so ist $S^{-1}AS$ genau dann diagonal, wenn die Spalten von S eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A bilden.

> Beweis: Wir schreiben $S^{-1}AS = D$ in der Form $AS = SD$ und wenden Merkregel 6 an, nun unter der Voraussetzung, dass S invertierbar ist. Der Fall, dass eine Spalte von S gleich 0 ist, kann nun wegen der Invertierbarkeit von S nicht mehr auftreten. Die j-te Spalte von S ist ein Eigenvektor von A zum Eigenwert λ_j. Da S invertierbar ist, bilden die Spalten von S eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A.

Daraus folgt unmittelbar

Merkregel 8: Ist $S^{-1}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, so ist jedes λ_j ein Eigenwert von A, und die j-te Spalte von S ist ein Eigenvektor von A zum Eigenwert λ_j.

Beweis: Merkregeln 6 und 7.

9 Diagonalisierung in $\mathbb{R}^{n\times n}$ bezüglich einer Orthonormalbasis

Ist $A \in \mathbb{R}^{n\times n}$ eine diagonalisierbare Matrix, so kann es vorkommen, dass $\mathbb{R}^n$, ausgestattet mit dem Standard-Skalarprodukt

$$\langle x, y\rangle = \sum_{j=1}^{n} x_j\, y_j\,, \tag{43}$$

eine Orthonormalbasis aus Eigenvektoren von A besitzt. Das ist insbesondere dann der Fall, wenn A symmetrisch ist, d.h. wenn $A^T = A$ gilt.

Wird in einem solchen Fall die Matrix S, für die, wie in Abschnitt 8 besprochen, $S^{-1}AS$ diagonal ist, so gewählt, dass ihre Spalten eine Orthonormalbasis von $\mathbb{R}^n$ aus Eigenvektoren von A bilden, so ist $S^TS = I_n$, denn $(S^TS)_{jk}$ ist nach den Regeln der Matrizenmultiplikation nichts anderes als das Skalarprodukt der j-ten Spalte mit der k-ten Spalte von S und somit, da die Spalten von S eine Orthonormalbasis bilden, gleich δ_{jk}. Eine Matrix mit dieser Eigenschaft heißt orthogonale Matrix, und es folgt

$$S^{-1} = S^T. \tag{44}$$

In diesem Fall existiert also eine orthogonale Matrix S, für die S^TAS diagonal ist. Man nennt die Diagonalisierung dann auch Hauptachsentransformation.

10 Diagonalisierung in $\mathbb{C}^{n\times n}$ bezüglich einer Orthonormalbasis

Ist $A \in \mathbb{C}^{n\times n}$ eine diagonalisierbare Matrix (auch wenn sie nur reelle Komponenten besitzt, kann sie als komplexe Matrix aufgefasst werden), so kann es vorkommen, dass $\mathbb{C}^n$, ausgestattet mit dem Standard-Skalarprodukt

$$\langle x, y\rangle = \sum_{j=1}^{n} x_j^*\, y_j \qquad \text{(der Stern bezeichnet komplexe Konjugation),} \tag{45}$$

eine Orthonormalbasis aus Eigenvektoren von A besitzt. Ein wichtiger Satz besagt (den wir hier nicht beweisen), dass das genau dann der Fall ist, wenn A eine normale Matrix ist, d.h. wenn

$$A^\dagger A = AA^\dagger \tag{46}$$

gilt, wobei $A^\dagger = (A^T)^* = (A^*)^T$ (ausgesprochen „A *dagger*“) die zu A adjungierte Matrix ist. Darunter fallen mehrere wichtige Klassen von Matrizen:

- Reelle symmetrische Matrizen ($A^T = A = A^*$).
  Alle Eigenwerte einer reellen symmaterischen Matrix sind reell.
- Hermitesche (selbstadjungierte) Matrizen ($A^\dagger = A$).
  Alle Eigenwerte einer hermiteschen Matrix sind reell.
- Reelle Orthogonalprojektionen ($A^2 = A = A^* = A^T$).
  Jeder Eigenwert einer reellen Orthogonalprojektion ist gleich 0 oder 1.
- Komplexe Orthogonalprojektionen ($A^2 = A = A^\dagger$).
  Jeder Eigenwert einer komplexen Orthogonalprojektion ist gleich 0 oder 1.
- Antihermitesche (antiselbstadjungierte) Matrizen ($A^\dagger = -A$).
  Alle Eigenwerte einer antihermiteschen Matrix sind imaginär.
- Orthogonale Matrizen ($A^* = A$ invertierbar und $A^{-1} = A^T$).
  Jeder Eigenwert einer orthogonalen Matrix ist gleich -1 oder 1.
- Unitäre Matrizen (A invertierbar und $A^{-1} = A^\dagger$).
  Jeder Eigenwert einer unitären Matrix ist komplex und hat Betrag 1, ist also von der Form $e^{i\,\alpha}$ für ein $\alpha \in \mathbb{R}$.

Wird für eine normale Matrix A eine Matrix S, für die, wie in Abschnitt 8 besprochen, $S^{-1}AS$ diagonal ist, so gewählt, dass ihre Spalten eine Orthonormalbasis von $\mathbb{C}^n$ aus Eigenvektoren von A bilden, so ist $S^\dagger S = I_n$, denn $(S^\dagger S)_{jk}$ ist nach den Regeln der Matrizenmultiplikation nichts anderes als das Skalarprodukt der j-ten Spalte mit der k-ten Spalte von S (Achtung: nicht vergessen, das erste Argument gemäß (45) komplex zu konjugieren!) und somit, da die Spalten von S eine Orthonormalbasis bilden, gleich δ_{jk}. Daher ist S eine unitäre Matrix, d.h. es gilt

$$S^{-1} = S^\dagger. \tag{47}$$

In diesem Fall existiert also eine unitäre Matrix S, für die $S^\dagger AS$ diagonal ist. Man nennt die Diagonalisierung dann ebenso wie in dem in Abschnitt 9 behandelten reellen Fall Hauptachsentransformation.