

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

j) Spezielle Matrizen

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

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

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

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

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

Die Matrix einer linearen Abbildung $\varphi: V \rightarrow V$ bezüglich einer Basis $\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_n)$ ist genau dann eine Diagonalmatrix, wenn jeder der Basisvektoren \vec{b}_i von φ auf ein Vielfaches $\lambda_i \vec{b}_i$ von sich selbst abgebildet wird; alsdann ist die Abbildungsmatrix gleich der Diagonalmatrix mit Einträgen $\lambda_1, \dots, \lambda_n$. Somit hängt es sehr von der Basis ab, ob die Abbildungsmatrix Diagonalgestalt hat oder nicht.

Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 1 & -2 \\ 2 & 1 & -2 & 1 \\ 1 & -2 & 1 & 2 \\ -2 & 1 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

etwa ist ganz sicher keine Diagonalmatrix. Betrachten wir die lineare Abbildung

$$\varphi: \mathbb{R}^4 \rightarrow \mathbb{R}^4; \quad \vec{v} \mapsto A\vec{v}$$

aber bezüglich der Basis $\mathcal{B} = (\vec{b}_1, \vec{b}_2, \vec{b}_3, \vec{b}_4)$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad \vec{b}_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{b}_4 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix},$$

so rechnet man leicht nach, daß

$$\varphi(\vec{b}_1) = A\vec{b}_1 = 2\vec{b}_1, \quad \varphi(\vec{b}_2) = 2\vec{b}_2, \quad \varphi(\vec{b}_3) = -4\vec{b}_3 \quad \text{und} \quad \varphi(\vec{b}_4) = 4\vec{b}_4$$

ist, bezüglich \mathcal{B} hat φ also die Diagonalmatrix

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

als Abbildungsmatrix. Wenn keine schwerwiegenden anderen Gründe dagegensprechen, wird es bei umfangreichen Rechnungen mit der Matrix A meist eine gute Idee sein, statt mit der Standardbasis von \mathbb{R}^4 mit der Basis \mathcal{B} zu rechnen.

Wir werden im weiteren Verlauf dieses Paragraphen noch lernen, wie man das Rechnen mit der Matrix A auf das Rechnen mit der sehr viel angenehmeren Diagonalmatrix D zurückführen kann und wie man, wo dies möglich ist, eine Basis \mathcal{B} wie im obigen Beispiel finden kann. Im Augenblick wollen wir uns damit begnügen, die dabei auftretenden Begriffe zu definieren:

Definition: Ein Vektor $\vec{v} \in k^n \setminus \{0\}$ heißt *Eigenvektor* der Matrix $A \in k^{n \times n}$, wenn es eine Zahl $\lambda \in k$ gibt, so daß $A\vec{v} = \lambda\vec{v}$ ist. Dieses λ bezeichnen wir als einen *Eigenwert* von A .

Falls es also zu einer Matrix A eine Basis aus Eigenvektoren gibt, können wir sie bezüglich dieser Basis als Diagonalmatrix darstellen; wie wir später sehen werden, gibt es solche Basen häufig, aber nicht immer.

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

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

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

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

Ist D stattdessen eine $n \times n$ -Matrix, ist das Produkt DA definiert und wir können es auf dieselbe Weise berechnen: Zunächst ist $(DA)\vec{e}_i =$

$D(A\vec{e}_i) = D\vec{a}_i$. Multiplikation einer Diagonalmatrix mit einem Vektor multipliziert die j -te Komponente dieses Vektors mit dem j -ten Diagonaleintrag von D , also wird der j -te Eintrag von \vec{a}_i mit d_j multipliziert. Dies passiert für jeden Spaltenvektor \vec{a}_i , also entsteht die Matrix DA aus A , indem man für jedes j ihre j -te Zeile mit d_j multipliziert.

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

2) Dreiecksmatrizen: Nachdem wir das Rechnen mit Diagonalmatrizen gut im Griff haben, können wir zu etwas komplizierteren Matrizen übergehen, z.B. den Dreiecksmatrizen:

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

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

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

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

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

Einträge notwendigerweise gleich Null, so daß nur

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

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

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

Wie bei Diagonalmatrizen sieht man sofort, daß eine Dreiecksmatrix genau dann invertierbar ist, wenn keiner ihrer Diagonaleinträge verschwindet: Falls etwa der i -te Diagonaleintrag einer oberen Dreiecksmatrix verschwindet, liegen die ersten i Spaltenvektoren im von den ersten $i-1$ Koordinateneinheitsvektoren erzeugten $(i-1)$ -dimensionalen Untervektorraum von k^n und können somit unmöglich zusammen mit den restlichen $n-i$ Spaltenvektoren einen n -dimensionalen Vektorraum aufspannen. Bei unteren Dreiecksmatrizen argumentiert man im wesentlichen genauso mit dem vom i -ten bis zum n -ten Spaltenvektor aufgespannten Untervektorraum.

Am einfachsten geht das, wenn wir wieder die lineare Abbildung

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

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

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

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

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

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

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

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

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

seien die zugehörigen linearen Abbildungen.

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

Den i -ten Diagonaleintrag erhalten wir als \vec{e}_i -Komponente des Bilds von \vec{e}_i . Im Falle zweier oberer Dreiecksmatrix A, B mit i -tem Diagonaleintrag a_i bzw. b_i ist $B\vec{e}_i = b_i\vec{e}_i + \vec{u}_{i-1}$ mit $\vec{u}_{i-1} \in U_{i-1}$; da auch A sowohl U_i als auch U_{i-1} auf sich selbst abbildet, ist also auch

$(AB)\vec{e}_i = a_i b_i + \vec{w}_{i-1}$ mit $\vec{w}_{i-1} \in U_{i-1}$, d.h. die Diagonaleinträge werden miteinander multipliziert. Ganz entsprechend argumentiert man für untere Dreiecksmatrizen: Hier ist $A\vec{e}_i = a_i \vec{e}_i + \vec{v}_{i+1}$ mit $\vec{v}_{i+1} \in V_{i+1}$.

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

Um nur Teil a) dieses Lemmas zu beweisen, wäre der Umweg über lineare Abbildungen nicht notwendig gewesen; das hätten wir direkt auch billiger haben können. Beispielsweise ist für zwei untere Dreiecksmatrizen $A = (a_{i\ell})$ und $B = (b_{\ell j})$ mit Produkt $C = (c_{ij})$ nach Definition der Matrixmultiplikation

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

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

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

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

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

Für inverse Matrizen kennen wir jedoch keine brauchbaren Formeln; hier ist der Umweg über die linearen Abbildungen der einzige Beweis, der mit unseren Kenntnissen möglich ist, und selbst wenn man die

(ziemlich unangenehmen) Formeln kennt, nach denen man die Einträge von A^{-1} durch die von A ausdrücken kann, ist der obige Beweis kürzer und verständlicher.

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

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

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

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

Die Abbildung $\varphi: k^n \rightarrow k^n$; $\vec{v} \mapsto E_{ij}\vec{v}$ bildet offensichtlich alle Koordinateneinheitsvektoren des k^n auf den Nullvektor ab mit Ausnahme

des j -ten; dieser wird auf den i -ten Koordinateneinheitsvektor des k^n abgebildet. Für einen beliebigen Vektor $\vec{v} \in k^m$ ist somit $\varphi(\vec{v}) = v_j \vec{e}_i$ jener Vektor aus k^n , der an der i -ten Stelle den j -ten Eintrag von \vec{v} stehen hat und sonst überall Nullen.

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

Für eine $q \times m$ -Matrix B sei entsprechend $\omega: k^m \rightarrow k^q$ die lineare Abbildung; dann bildet $\omega \circ \varphi$ alle Koordinateneinheitsvektoren von k^n mit Ausnahme des j -ten auf den Nullvektor ab, da bereits φ diese Eigenschaft hat. Der j -te Koordinateneinheitsvektor wird von φ abgebildet auf den i -ten Koordinateneinheitsvektor von k^m , der wiederum von ω abgebildet wird auf den i -ten Spaltenvektor von B . Die Abbildungsmatrix BE_{ij} von $\omega \circ \varphi$ hat also als j -te Spalte die i -te Spalte von A ; alle anderen Spalten sind Null.

Kurz können wir diese beiden Resultate zusammenfassen zu folgender Regel:

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

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

Definition: a) Eine Permutation ist eine bijektive Abbildung

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

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

- b) Eine Transposition ist eine Permutation τ , die zwei Elemente von $\{1, \dots, n\}$ vertauscht und den Rest festläßt. Sind i, j die beiden Elemente, die von τ vertauscht werden, so schreiben wir kurz $\tau = (i j)$.
- c) Die Menge aller Permutationen $\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ wird als *symmetrische Gruppe* S_n bezeichnet.

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

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

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

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

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

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



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

Im Sinne dieser Definition bilden beispielsweise die ganzen Zahlen mit der Addition als Verknüpfung eine Gruppe, sogar eine abelsche Gruppe, nicht aber die natürlichen Zahlen, denn es gibt beispielsweise keine natürliche Zahl n , so daß $n + 2 = 0$ ist, und auch die Null

liegt zumindest nach der in dieser Vorlesung zugrundegelegten Konvention nicht in \mathbb{N} . Entsprechend bilden die ganzen Zahlen bezüglich der *Multiplikation* keine Gruppe, da die Gleichung $5x = 1$ in \mathbb{Z} nicht lösbar ist, aber die positiven rationalen Zahlen bilden eine multiplikative Gruppe. Die rationalen Zahlen bilden keine, da $0x = 1$ unlösbar ist, aber die rationalen Zahlen ohne Null sind eine multiplikative Gruppe, genauso auch die *invertierbaren* $n \times n$ -Matrizen.

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

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

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

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

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

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

Somit ist

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

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

also

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

d.h.

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

Aus der Darstellung

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

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

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

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

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

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

ihre Abbildungsmatrix (bezüglich der Standardbasis von k^n) bezeichnen wir mit P_π . Da φ_π den i -ten Koordinateneinheitsvektor von k^n auf den

$\pi(i)$ -ten abbildet, hat P_π in jeder Zeile und jeder Spalte genau einen nichtverschwindenden Eintrag: In der i -ten Spalte steht an der Position $\pi(i)$ eine Eins, überall sonst stehen Nullen. Der Eintrag $p_{k\ell}$ ist also genau dann gleich Eins, wenn $k = \pi(\ell)$ ist, ansonsten ist er Null.

Die Hintereinanderausführung der Abbildungen φ_π und damit die Multiplikation der Matrizen P_π ist vollständig bestimmt durch die Hintereinanderausführung der zugehörigen Permutationen: Wegen

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

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

Auch die Multiplikation der Permutationsmatrix P_π mit einer beliebigen Matrix (passender Größe) läßt sich leicht ausführen: Für $\pi \in \mathfrak{S}_n$ und $A \in k^{n \times m}$ betrachten wir zu A die lineare Abbildung $\psi: k^m \rightarrow k^n$, die einen Vektor $\vec{v} \in k^m$ auf $A\vec{v}$ abbildet; dann ist $P_\pi A$ die Abbildungsmatrix von $\varphi_\pi \circ \psi$. Ihre Spalten sind die Bilder der Koordinateneinheitsvektoren von k^m . Der i -te dieser Vektoren wird von ψ auf den i -ten Spaltenvektor von A abgebildet, und auf dessen Komponenten wird die Permutation π angewandt. Dies passiert für jede Spalte von A , insgesamt entsteht $P_\pi A$ somit aus A , indem man die Permutation π auf dessen *Zeilen* anwendet. Für die Transposition $\pi = (12) \in \mathfrak{S}_3$ etwa ist

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

und

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

Entsprechend können wir für eine $p \times n$ -Matrix B das Produkt BP_π ausrechnen: φ_π bildet den i -ten Koordinateneinheitsvektor von k^n ab auf den $\pi(i)$ -ten, und dieser wird von der linearen Abbildung zu B abgebildet auf den $\pi(i)$ -ten Spaltenvektor von B . Die Matrix BP_π entsteht also

aus B daraus, daß die Permutation π auf dessen *Spalten* angewandt wird. Beispielsweise ist für $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix}$

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

und

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

k) Die LR-Zerlegung einer Matrix

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

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

Falls A nicht invertierbar ist, falls man also beispielsweise eine Steuerung mit Redundanz hat, kann man nicht so vorgehen, aber eine fast

genauso effiziente Modifikation führt auch hier ans Ziel. Das entsprechende Verfahren wird je nach Lehrbuch als LR-Zerlegung oder LU-Zerlegung bezeichnet, wobei LR für *links/rechts* und LU für *lower/upper* steht.

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

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

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

Angenommen, wir möchten das c -fache der j -ten Zeile zur i -ten addieren. Dazu betrachten wir die aus dem letzten Abschnitt bekannte Matrix $E_{ij} \in k^{n \times n}$, d.h. jene $n \times n$ -Matrix, die an der Stelle (i, j) den Eintrag Eins hat und sonst lauter Nullen. Das Produkt $E_{ij}M$ hat, wie wir uns dort überlegt haben, als i -te Zeile die j -te Zeile von M und alle anderen Zeilen sind Null. Das Produkt $cE_{ij}M$ hat somit als i -te Zeile das c -fache der j -ten Zeile von M (und sonst lauter Nullzeilen). Setzen wir

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Betrachten wir dazu ein konkretes Beispiel:

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

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

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

und führt auf

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

Wir schreiben die Einheitsmatrix daneben:

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

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

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

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

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

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

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

In der Tat rechnet man leicht nach, daß

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

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

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

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

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

Somit ist

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

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

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

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

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

1) Beispiel eines Basiswechsels

In den meisten (endlichdimensionalen) Vektorräumen, die wir bislang betrachtet hatten, gab es offensichtliche Basen wie etwa die Standardbasis aus den Koordinateneinheitsvektoren in \mathbb{R}^n oder die reinen Potenzen bei Vektorräumen von Polynomen. Diese Basen sind allerdings für rechnerische Zwecke nicht immer optimal: Sowohl in vielen Anwendungen der Mathematik als auch innerhalb der Mathematik ist es oft günstiger, mit anderen Basen zu rechnen, in denen beispielsweise wichtige Matrizen zu Diagonal- oder Dreiecksmatrizen werden. Das letzte Thema diesen Paragraphen wird sich mit dem Problem befassen, Matrizen und Vektoren bezüglich einer Basis in eine andere Basis umzurechnen.

In diesem Abschnitt betrachten wir zur Einstimmung einen Vektorraum, in dem es keine irgendwie ausgezeichnete Basis gibt.

Wir gehen aus von dem Problem, Farben quantitativ zu charakterisieren: Physikalisch gesehen hängen Farben mit den verschiedenen Wellenlängen des sichtbaren Lichts zusammen, eine Farbe müßte also eigentlich durch eine Funktion auf dem Intervall von etwa 380 bis etwa

780 nm beschrieben werden, wobei keineswegs nur stetige Funktionen in Betracht kommen.

Nun werden Farben aber nur selten mit dem Spektrometer betrachtet; was wirklich interessiert ist der Eindruck auf das menschliche Auge. Dieses hat vier Arten von Photorezeptoren: Die sehr empfindlichen Stäbchen, die nur bei schwachem Licht eine Rolle spielen und die auch nur Helligkeitsinformationen liefern können, sowie drei Arten von weiniger lichtempfindlichen Zapfchen, k , ℓ und m , die bei gutem Licht ein Bild mit Farbinformation liefern, da jede Art ihre eigene spektrale Empfindlichkeitskurve hat. Klassische wie auch digitale Photographie sowie Farbdarstellungen auf Fernseh- und Computermonitoren benutzen allesamt darauf, daß dem Auge etwas vorgesetzt wird, was die k , ℓ und m -Zapfchen zur gleichen Reaktion veranlaßt wie das „echte“ Farbsignal.

Damit reicht es aus, Farben als Elemente eines dreidimensionalen Raums zu beschreiben. Da sich die Empfindlichkeitsbereiche insbesondere der ℓ - und der m -Zapfchen stark überlappen, wäre es allerdings weder sinnvoll noch sonderlich praktikabel, eine Basis über die Ausgabewerte der drei Arten von Zapfchen zu definieren. Stattdessen werden je nach Hauptanwendungsziel mehrere Farbmodelle betrachtet, die ausgehend von den unterschiedlichsten Ansätzen allesamt denselben dreidimensionalen \mathbb{R} -Vektorraum beschreiben. Wir wollen uns die beiden einfachsten etwas genauer anschauen.

Am bekanntesten ist wohl das RGB-Modell, das Farben aus den drei Grundfarben Rot, Grün und Blau kombiniert. Basis des Vektorraums sind hierbei also drei Vektoren \vec{r}, \vec{g} und \vec{b} , die einem Rot, Grün bzw. Blau einer vorgegebenen Frequenz und Intensität entsprechen, und alle anderen Farben werden als Linearkombinationen

$$R\vec{r} + G\vec{g} + B\vec{b}$$

dargestellt. Diese Darstellung ist insbesondere gut geeignet für die Farbdarstellung auf einem Monitor oder auch Fernsehschirm, wo Farben in genau dieser Weise erzeugt werden. Bei digitalen Bildern betrachtet man i.a. nur Koeffizienten R, G, B aus dem Intervall $[0, 1]$, so daß gewisse Grenzhelligkeiten nicht überschritten werden können. Bei der

Darstellung mit einer Genauigkeit von acht Bit betrachtet man oft auch das 255-fache der entsprechenden Werte, gerundet zur nächsten ganzen Zahl.

Bei anderen Farbmodellen sieht man nicht so sehr auf die *Erzeugung* der Farben am Bildschirm, sondern auf deren Eigenschaften, wie sie ein menschlicher Betrachter wahrnimmt: Die Ausgabeimpulse der Zäpfchen werden noch im Sehapparat sofort weiterverarbeitet, und was wir bewußt zur Kenntnis nehmen, sind definitiv nicht die R-, G- und B-Anteile der Farben, sondern eher Helligkeiten, Farbsättigungen und ähnliche Eigenschaften.

Eine für die digitale Speicherung und Übermittlung interessante Tatsache ist, daß wir Helligkeiten viel feiner unterscheiden und viel höher auflösen können als Farbnuancen. Es liegt daher nahe, eine Basis zu wählen, in der auch die Gesamthelligkeit als Komponente auftritt, wobei diese Komponente entweder mit höherer Genauigkeit digitalisiert wird als die beiden anderen oder (häufiger) nur diese Komponente für *jedes* Pixel gespeichert wird, die beiden anderen aber nur für jedes zweite Pixel oder gar nur einmal pro Quadrat aus 2×2 Pixel. Die Wahl der Helligkeit als Basiskomponente hat auch den Vorteil, daß man dann in einfacher Weise diese Komponente für Schwarz-Weiß-Versionen der Bilder verwenden kann, was zum Beispiel beim Fernsehen eine wichtige Rolle spielt.

Wenn wir die Helligkeit als eine Basiskomponente wählen, stehen für die eigentliche chromatische Information nur noch zwei Basisvektoren zur Verfügung; diese können entweder aus zwei Farbanteilen des RGB-Systems abgeleitet werden oder (was für die Gestaltung photorealistischer Bilder gern angewandt wird) aus einer weiteren achromatischen Komponente wie etwa der Farbsättigung und nur *einer* chromatischen Komponente.

Bei den drei derzeit gebräuchlichen Fernsehstandards PAL, SECAM und NTSC sowie auch beim HDTV und beim JPEG-Format für digitale Bilder geht man den ersten Weg: Der erste Basisvektor \vec{y} beschreibt die Helligkeit oder *Luminanz*, die beiden weiteren Vektoren sind gewichtete Differenzen zwischen dieser Helligkeit und dem Rot- oder Blauanteil

des Farbvektors. Die Gewichte unterscheiden sich dabei in den einzelnen Fällen; da für die Hörer dieser Vorlesung das JPEG-Format interessanter sein dürfte als Fernsehstandards, betrachten wir das dort verwendete YCbCr-Modell.

Die Helligkeit ist, wie in allen diesen Systemem, als gewichtetes Mittel der drei Farbanteile definiert; die Gewichte L_R , L_G und $L_B \in (0, 1)$ bezeichnet man als *Lumared*, *Lumagreen* und *Lumablue*. Mit diesen Bezeichnungen ist die Helligkeit

$$Y = L_R R + L_G G + L_B B \quad \text{mit} \quad L_R + L_G + L_B = 1;$$

Standardwerte sind

$$L_R = \frac{299}{1000}, \quad L_G = \frac{587}{1000} \quad \text{und} \quad L_B = \frac{114}{1000}.$$

Die beiden Chrominanz sind festgelegt durch

$$C_b = \frac{B - Y}{2 - 2L_B} \quad \text{und} \quad C_r = \frac{R - Y}{2 - 2L_R}.$$

Damit haben wir eine neue Basis $\{\vec{y}, \vec{c}_b, \vec{c}_r\}$, mittels derer wir dieselben Farben beschreiben wie bezüglich der Basis $\{\vec{r}, \vec{g}, \vec{b}\}$.

Ganz offensichtlich ist es wichtig, zwischen den beiden Basen hin- und herrechnen zu können, denn schließlich werden auch JPEG-Bilder auf RGB-Monitoren betrachtet, und CCD Chips in Digitalkameras messen zunächst einmal RGB-Komponenten, aus denen dann oft ein JPEG-Bild erzeugt wird.

Im vorliegenden Fall ist es einfach, konkrete Formeln zu finden:

$$\begin{aligned} Y &= L_R R + L_G G + L_B B, \\ C_b &= \frac{B - Y}{2 - 2L_B} = -\frac{L_R}{2 - 2L_B} \cdot R - \frac{L_G}{2 - 2L_B} \cdot G + \frac{1}{2} \cdot B, \\ C_r &= \frac{R - Y}{2 - 2L_R} = \frac{1}{2} \cdot R - \frac{L_G}{2 - 2L_R} \cdot G - \frac{L_B}{2 - 2L_R} \cdot B. \end{aligned}$$

Auch die Umkehrung läßt sich leicht ausrechnen:

$$\begin{aligned} R &= Y + (2 - 2L_R)C_r, \\ B &= Y + (2 - 2L_B)C_b \quad \text{und} \\ G &= \frac{Y - L_R R - L_B B}{L_G} \\ &= \frac{(1 - L_R - L_B)Y - L_B(2 - 2L_B)C_b - L_R(2 - 2L_R)C_r}{L_G} \\ &= Y - \frac{L_B(2 - 2L_B)}{L_G}C_b - \frac{L_R(2 - 2L_R)}{L_G}C_r, \end{aligned}$$

denn $1 - L_R - L_B = L_G$. Dies können wir auch mit Matrizen formulieren:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} \frac{L_R}{2 - 2L_B} & \frac{L_G}{2 - 2L_B} & \frac{L_B}{2 - 2L_R} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-L_G}{2 - 2L_B} & \frac{-L_B}{2 - 2L_R} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 - 2L_R \\ 1 & -\frac{L_B(2 - 2L_B)}{L_G} & -\frac{L_R(2 - 2L_R)}{L_G} \\ 1 & 2 - 2L_B & 0 \end{pmatrix} \begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix}.$$

und

Der Übergang zwischen den beiden Basen kann also jeweils als Multiplikation mit einer Matrix interpretiert werden, und natürlich sind die beiden zugehörigen Matrizen invers zueinander.

Wenn wir die erste der beiden Matrizen mit A bezeichnen, so sind die Spalten von A die Bilder der Einheitsvektoren des RGB-Systems, umgerechnet ins YCbCr-System; wir erhalten also die Beziehungen

$$\begin{aligned} \vec{r} &= L_R \vec{y} - \frac{L_R}{2 - 2L_B} \vec{c}_b + \frac{1}{2} \vec{c}_r \\ \vec{g} &= L_G \vec{y} - \frac{L_G}{2 - 2L_B} \vec{c}_b - \frac{L_G}{2 - 2L_R} \vec{c}_r \quad \text{und} \\ \vec{b} &= L_B \vec{y} + \frac{1}{2} \vec{c}_b - \frac{L_B}{2 - 2L_R} \vec{c}_r. \end{aligned}$$

Entsprechend sind die Spalten von A^{-1} die Bilder der Einheitsvektoren des YCbCr-Systems, umgerechnet ins RGB-System, d.h.

$$\begin{aligned} \vec{y} &= \vec{r} + \vec{g} + \vec{b}, \\ \vec{c}_b &= -\frac{2 - 2L_B}{L_G} \vec{y} + (2 - 2L_B) \vec{b} \quad \text{und} \\ \vec{c}_r &= (2 - 2L_R) \vec{r} - \frac{2 - 2L_B}{L_G} \vec{y}. \end{aligned}$$

Man beachte die Unterschiede zwischen der Darstellung der Basisvektoren in der jeweils anderen Basis und den Umrechnungsformeln für die Koeffizienten: Beim Umrechnen der RGB-Werte in YCbCr-Werte haben wir als Koeffizienten der einzelnen Gleichungen die *Zeilen* von A ; die Basisvektoren $\vec{r}, \vec{g}, \vec{b}$ selbst sind aber im YCbCr-System ausgedrückt durch die *Spalten* der inversen Matrix A^{-1} .

Zur Verdeutlichung seien die Gleichungen nochmals angegeben mit numerischen Koeffizienten, näherungsweise berechnet für die Standardwerte von L_R, L_G und L_B : Damit ist

$$A = \begin{pmatrix} 0,2990 & 0,5870 & 0,1140 \\ -0,1687 & -0,3313 & 0,5000 \\ 0,5000 & -0,4187 & -0,0813 \end{pmatrix}$$

und

$$A^{-1} = \begin{pmatrix} 1 & 0 & 1,4020 \\ 1 & -0,3441 & -0,7141 \\ 1 & 1,7720 & 0 \end{pmatrix},$$

also erhalten wir die Gleichungen

$$\begin{aligned} Y &= 0,2990 R + 0,5870 G + 0,1140 B \\ C_b &= -0,1687 R - 0,3313 G + 0,5000 B \\ C_r &= 0,5000 R - 0,4187 G - 0,0813 B \end{aligned}$$

und

$$\begin{aligned} \vec{r} &= 0,2990 \vec{y} - 0,1687 \vec{c}_b + 0,5000 \vec{c}_r \\ \vec{g} &= 0,5870 \vec{y} - 0,3313 \vec{c}_b - 0,4187 \vec{c}_r \\ \vec{b} &= 0,1140 \vec{y} + 0,5000 \vec{c}_b - 0,0813 \vec{c}_r. \end{aligned}$$

Entsprechend liefern Zeilen und Spalten von A^{-1} die Beziehungen

$$\begin{aligned} R &= Y + 1,402 Cr \\ G &= Y - 0,3441 Cb - 0,7141 Cr \\ B &= Y + 1,772 Cb \end{aligned}$$

und

$$\begin{aligned} \vec{y} &= \vec{r} + \vec{g} + \vec{b} \\ \vec{c}_b &= -0,3441 \vec{g} + 10,772 \vec{b} \\ \vec{c}_r &= 1,402 \vec{r} - 0,7141 \vec{g}. \end{aligned}$$

m) Basiswechsel im allgemeinen Fall

Gehen wir über zur allgemeinen Situation!

Wir gehen aus von einem n -dimensionalen k -Vektorraum V und betrachten darin beiden Basen

$$B = (\vec{b}_1, \dots, \vec{b}_n) \quad \text{und} \quad C = (\vec{c}_1, \dots, \vec{c}_n).$$

Da B eine Basis ist, lassen sich die Vektoren \vec{c}_i als Linearkombinationen

$$\vec{c}_i = a_{i1} \vec{b}_1 + \dots + a_{in} \vec{b}_n = \sum_{j=1}^n a_{ij} \vec{b}_j$$

der \vec{b}_j schreiben; entsprechen lassen sich natürlich auch die \vec{b}_j als Linearkombinationen

$$\vec{b}_j = m_{j1} \vec{c}_1 + \dots + m_{jn} \vec{c}_n = \sum_{\ell=1}^n m_{j\ell} \vec{c}_\ell$$

schreiben. Dies können wir in die Darstellung von \vec{c}_i einsetzen und erhalten

$$\vec{c}_i = \sum_{j=1}^n a_{ij} \vec{b}_j = \sum_{j=1}^n a_{ij} \sum_{\ell=1}^n m_{j\ell} \vec{c}_\ell = \sum_{\ell=1}^n \left(\sum_{j=1}^n a_{ij} m_{j\ell} \right) \vec{c}_\ell.$$

Da C eine Basis ist, bezüglich derer \vec{c}_i die eindeutige Basisdarstellung

$$\vec{c}_i = 1 \cdot \vec{c}_i$$

hat, folgt

$$\sum_{j=1}^n a_{ij} m_{j\ell} = \begin{cases} 1 & \text{falls } i = \ell \\ 0 & \text{falls } i \neq \ell \end{cases}.$$

Fassen wir die a_{ij} zu einer Matrix A zusammen und die m_{ij} zu einer Matrix M , ist also $AM = E$, d.h. die beiden Matrizen sind invers zueinander (was eigentlich niemanden erstaunen sollte). Formal können wir dies schreiben als

$$\begin{pmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{pmatrix} = A \begin{pmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_n \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_n \end{pmatrix} = A^{-1} \begin{pmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{pmatrix},$$

wobei die \vec{b}_j und \vec{c}_i bei der Auswertung dieser Formel nur als Symbole zu betrachten sind: Wir wollen nicht wirklich einen Vektor von Vektoren definieren, sondern diese Formel einfach als kompakte Merkmittel für den Zusammenhang zwischen den beiden Basen betrachten.

Beim konkreten Rechnen in Vektorräumen geht es nicht so sehr um die Basisvektoren selbst, sondern um die Koeffizienten der Basisdarstellung; sei also der Vektor

$$\vec{v} = v_1 \vec{b}_1 + \dots + v_n \vec{b}_n = w_1 \vec{c}_1 + \dots + w_n \vec{c}_n$$

bezüglich beider Basis dargestellt; wir suchen einen Zusammenhang zwischen den v_j und den w_i .

Mit der obigen Matrix $M = (m_{ij})$ ausgedrückt ist

$$\vec{v} = \sum_{j=1}^n v_j \vec{b}_j = \sum_{j=1}^n v_j \sum_{i=1}^n m_{ji} \vec{c}_i = \sum_{i=1}^n \left(\sum_{j=1}^n v_j m_{ji} \right) \vec{c}_i,$$

also folgt wegen der Eindeutigkeit der Basisdarstellung

$$w_i = \sum_{j=1}^n v_j m_{ji} = \sum_{i=1}^n m_{ji} v_j.$$

Letzteres läßt sich leider nicht als Produkt von M mit dem Spaltenvektor der v_j schreiben: Die Indizes von m_{ji} haben die falsche Reihenfolge. Da die Formel trotzdem richtig und nützlich ist, definieren wir eine neue

Matrix, die aus der alten durch Vertauschung der Indizes, d.h. also von Zeilen und Spalten, entsteht:

Definition: Die *transponierte* Matrix $N = {}^tM$ zur einer Matrix $M = (m_{ij})_{i=1, \dots, n}^{j=1, \dots, m} \in k^{n \times m}$ ist jene Matrix $N = (n_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in k^{m \times n}$ mit $n_{ij} = m_{ji}$ für alle i, j .

Somit braucht man zur Umrechnung der Koeffizienten ineinander also die transponierte Matrix zu $M = A^{-1}$. Diese Matrix bezeichnen wir als die *Matrix des Basiswechsels*; sie wird auch gelegentlich als die zu A *kontragrediente* Matrix bezeichnet. Für sie gilt:

Satz: Sind $\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_n)$ und $\mathcal{C} = (\vec{c}_1, \dots, \vec{c}_n)$ zwei Basen eines n -dimensionalen k -Vektorraums und ist

$$\vec{c}_i = \sum_{j=1}^n a_{ij} \vec{b}_j \quad \text{mit} \quad A = (a_{ij}) \in k^{n \times n},$$

so berechnen sich für

$$\vec{v} = v_1 \vec{b}_1 + \dots + v_n \vec{b}_n = w_1 \vec{c}_1 + \dots + w_n \vec{c}_n$$

die Koeffizienten w_i der Basisdarstellung bezüglich \mathcal{C} aus den v_j nach der Formel

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = {}^t(A^{-1}) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix};$$

die Matrix des Basiswechsels ist also ${}^t(A^{-1})$. ■

Betrachten wir dazu ein Beispiel: Im Vektorraum $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ aller stetiger Funktionen von \mathbb{R} nach \mathbb{R} sei der Untervektorraum U erzeugt von den beiden Funktionen e^x und e^{-x} , die wir zur Basis \mathcal{B} von U zusammenfassen. Eine weitere Basis \mathcal{C} von U bestehe aus den Funktionen $\sinh x$ und $\cosh x$. Dann ist

$$\sinh x = \frac{1}{2}e^x - \frac{1}{2}e^{-x} \quad \text{und} \quad \cosh x = \frac{1}{2}e^x + \frac{1}{2}e^{-x},$$

die Matrix A , mittels derer wir \mathcal{C} durch \mathcal{B} ausdrücken, ist also

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

Die Berechnung der inversen Matrix ist hier nicht schwierig; man überzeugt sich leicht, daß gilt

$$A^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{und somit} \quad {}^t(A^{-1}) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Damit kennen wir die Matrix des Basiswechsels; ihr Produkt mit einem Koeffizientenvektor $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ ist

$${}^t(A^{-1}) \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a-b \\ a+b \end{pmatrix}.$$

Also sollte gelten

$$ae^x + ae^{-x} = (a-b) \sinh x + (a+b) \cosh x,$$

was man in der Tat leicht nachrechnet.

Als nächstes wollen wir uns überlegen, daß oben in ${}^t(A^{-1})$ die Klammern überflüssig waren; genauer gilt

Lemma: Für $A \in k^{n \times m}$ und $B \in k^{m \times p}$ ist ${}^t(AB) = {}^tB {}^tA$ und ${}^t(A^{-1}) = ({}^tA)^{-1}$.

Beweis: Mit $A = (a_{ij})$ und $B = (b_{j\ell})$ hat AB an der Stelle $i\ell$ den Eintrag $\sum_{j=1}^m a_{ij} b_{j\ell}$, und ihre Transponierte hat denselben Eintrag an der Stelle ℓi .

tB hat an der Stelle ℓj den Eintrag $b_{j\ell}$ und tA hat an der Stelle ji den Eintrag a_{ij} , das Produkt ${}^tB {}^tA$ hat somit an der Stelle ℓi den Eintrag $\sum_{j=1}^m b_{j\ell} a_{ij}$, was wegen der Kommutativität der Multiplikation in k mit dem oben berechneten Ausdruck übereinstimmt.

Damit ist die erste Formel bewiesen. Wenden wir sie an auf $B = A^{-1}$, so folgt die Beziehung

$${}^t(A^{-1}) {}^tA = {}^t(AA^{-1}) = {}^tE = E,$$

die beiden Matrizen sind also invers zueinander. ■

Als letztes Thema im Zusammenhang mit Basiswechseln wollen wir uns überlegen, wie sich die Abbildungsmatrix einer linearen Abbildung unter Basiswechseln verhält.

Wir betrachten also eine lineare Abbildung $\varphi: V \rightarrow W$; dabei sei V ein n -dimensionaler k -Vektorraum, und W sei ein m -dimensionaler. In jedem der beiden Vektorräume seien zwei Basen gegeben; in V seien dies

$$\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_n) \quad \text{und} \quad \mathcal{C} = (\vec{c}_1, \dots, \vec{c}_m),$$

in W seien es

$$\mathcal{D} = (\vec{d}_1, \dots, \vec{d}_m) \quad \text{und} \quad \mathcal{E} = (\vec{e}_1, \dots, \vec{e}_m).$$

Um nicht ganz die Übersicht zu verlieren, ordnen wir jedem Vektor

$$\vec{v} = v_1 \vec{b}_1 + \dots + v_n \vec{b}_n = w_1 \vec{c}_1 + \dots + w_m \vec{c}_m$$

aus V die beiden Spaltenvektoren

$$\vec{v}_{\mathcal{B}} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{und} \quad \vec{v}_{\mathcal{C}} = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}$$

aus k^m zu; entsprechend haben wir für jeden Vektor $\vec{w} \in W$ zwei Vektoren $\vec{w}_{\mathcal{D}}$ und $\vec{w}_{\mathcal{E}}$ aus k^m .

Die Abbildungsmatrix $M_{\mathcal{B}, \mathcal{D}}$ von φ bezüglich der Basen \mathcal{B} von V und \mathcal{D} von W hat dann die Eigenschaft, daß

$$\varphi(\vec{v})_{\mathcal{D}} = M_{\mathcal{B}, \mathcal{D}} \vec{v}_{\mathcal{B}}$$

ist; für die Abbildungsmatrix $M_{\mathcal{C}, \mathcal{E}}$ von φ bezüglich der Basen \mathcal{C} von V und \mathcal{E} von W gilt entsprechend

$$\varphi(\vec{v})_{\mathcal{E}} = M_{\mathcal{C}, \mathcal{E}} \vec{v}_{\mathcal{C}}.$$

Für den Übergang von einer Basis zur anderen rechnen wir der Einfachheit halber nicht mit der zu Beginn dieses Abschnitts betrachteten Matrix, die die Basisvektoren durcheinander ausdrückt, sondern gleich mit

der Matrix des Basiswechsels; die Matrizen $A \in k^{n \times n}$ und $B \in k^{m \times m}$ seien also so gewählt, daß für Vektoren $\vec{v} \in V$ und $\vec{w} \in W$ gilt

$$\vec{v}_{\mathcal{C}} = A \vec{v}_{\mathcal{B}} \quad \text{und} \quad \vec{w}_{\mathcal{E}} = B \vec{w}_{\mathcal{D}}.$$

Dann ist

$$\varphi(\vec{v})_{\mathcal{E}} = B \varphi(\vec{v})_{\mathcal{D}} = B M_{\mathcal{B}, \mathcal{D}} \vec{v}_{\mathcal{B}} = B M_{\mathcal{B}, \mathcal{D}} A^{-1} \vec{v}_{\mathcal{C}},$$

d.h.

$$M_{\mathcal{C}, \mathcal{E}} = B M_{\mathcal{B}, \mathcal{D}} A^{-1}.$$

Im nächsten Semester werden wir vor allen den Fall $V = W$ benötigen; hier wählt man natürlich fast immer $\mathcal{D} = \mathcal{B}$ und $\mathcal{C} = \mathcal{E}$, so daß jede Abbildungsmatrix von nur *einer* Basis von V abhängt.

Dann ist auch $A = B$ und die obige Formel vereinfacht sich zu

$$M_{\mathcal{C}} = A M_{\mathcal{B}} A^{-1},$$

wobei wir hier in $M_{\mathcal{B}} = M_{\mathcal{B}, \mathcal{B}}$ den zweiten Index natürlich weglassen.

Betrachten wir als Beispiel die Differentiation im von e^x und e^{-x} erzeugten Vektorraum; ihre Abbildungsmatrix bezüglich dieser Basis \mathcal{B} ist

$$M_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Die Matrix des Basiswechsels Basis \mathcal{C} aus $\sinh x$ und $\cosh x$ ist, wie wir oben gesehen haben,

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

und

$$A M_{\mathcal{B}} A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

ganz in Übereinstimmung mit der Tatsache, daß die Differentiation $\sinh x$ zu $\cosh x$ macht und umgekehrt.