1 Euklidische Vektorräume

1.1 Vorbemerkung.

Dieses Kapitel nimmt endlich die Physiker ernst, die von Anfang an einen Vektor als eine gerichtete Größe betrachtet haben, als einen Pfeil, dessen Länge den Betrag der physikalischen Größe bezeichnet und dessen Richtung eben ihre Wirkrichtung. In dem Sinne, in dem eine Kraft verschiedenen Betrag haben kann, gemessen in Newton, zudem aber auch in verschiedenen Richtungen ziehen oder drücken kann – im Gegensatz etwa zur Temperatur eines Gasvolumens, über die alles gesagt ist, wenn man ihren Betrag angibt, gemessen in Kelvin. Diese Vorstellungen und Konzepte, man muss das kulturhistorisch an dieser Stelle einfach einmal würdigend erwähnen, sind deutlich älter als die[1] axiomatischen Definitionen der Linearen Algebra, so wie wir sie kennengelernt haben.

Aber Wirkrichtungen, Winkel zwischen ihnen, und Beträge, also "Längen", haben wir bislang nicht betrachtet. Letztlich nicht betrachten *können*, denn in den Konzepten des Linearen Raumes und der Linearen Abbildungen zwischen ihnen sind sie nicht angelegt, sind darin sozusagen nicht enthalten oder nicht verborgen. Anders formuliert, bedarf es einer zusätzlichen Struktur, um Winkel und Beträge einzuführen, einer Struktur, die auf die bisher eingeführten noch draufgepackt wird. Das Vorgehen ist uns mittlerweile vertraut, denn wo wir hinwollen, ist klar: als den Betrag eines Vektors $(3, -4, 1)$ würde der Physiker sofort, und gut begründbar, die Größe $\sqrt{3^2 + (-4)^2 + 1^2} = \sqrt{26}$ ansehen, einfach weil er sich den Vektor als Pfeil im dreidimensionalen Anschauungsraum denkt und diese Größe dann die gemäß des Satzes von PYTHAGORAS bestimmte Länge jenes Pfeils darstellt. Winkel zwischen Pfeilen liefert der ebenfalls aus der Schule bekannte, wenn vielleicht auch nicht mehr unbedingt vertraute[2] Cosinussatz.

Die neu einzuführende Struktur ist unter dem Namen *Skalarprodukt* bekannt; sie heißt so, weil es sich um eine Abbildung handelt, deren Zielwert stets ein Skalar ist, also ein Element aus dem Körper $\mathbb{K}$, auf dem wir unsere Vektorräume aufbauen. Die Bezeichnung kann möglicherweise verwechselt werden mit dem *Skalaren Produkt*, auch *Äußeres Produkt* genannt, mittels dessen Vektoren skaliert werden. Eine Verwechslungsgefahr besteht nur dann nicht, wenn der Leser sich mit beiden Konzepten sorgfältig vertraut gemacht hat. Ein $\mathbb{K}$-Vektorraum $(\mathrm{V}, +, \cdot)$, der mit der

[1] tatsächlich überraschend neuzeitlichen!

[2] Der Leser registriere die in dieser Bemerkung implizit enthaltene Aufforderung.

Zusatzstruktur des Skalarprodukts ausgestattet ist, heißt allerdings nicht etwa “Skalarer Vektorraum” oder so ähnlich, das wäre dann doch etwas verwirrend, sondern vielmehr *Euklidischer Vektorraum*. Mit diesen also werden wir uns im vorliegenden Kapitel befassen.

1.2 Skalarprodukte.

Zur axiomatischen Definition des Begriffs des Skalarprodukts lassen wir uns von den intuitiv gewonnen Konzepten der Physiker leiten. Welche Eigenschaften sollen die Begriffe von “Länge eines Vektors” und “Winkel zwischen zwei Vektoren” haben? Wir zielen auf eine Produktbildung ab, bei der je zwei Vektoren eine *Zahl* zugewiesen wird, also eine Abbildung $\mathrm{V} \times \mathrm{V} \to \mathbb{K}$. In der Physik soll das Skalarprodukt dazu dienen, festzustellen, welchen *Anteil* ein Vektor $\mathbf{v}$ *in Richtung* eines anderen Vektors $\mathbf{w}$ hat, denn dies ist bei der Analyse der Wirkungen etwa von Kräften wichtig, die in verschiedene Richtungen ziehen. Betrachten wir zum Beispiel die Euklidische Ebene, die von den beiden Basisvektoren $\mathbf{e}_1 = (1, 0)$ und $\mathbf{e}_2 = (0, 1)$ aufgespannt wird[3]. Der Vektor $\mathbf{v} = (3, \sqrt{3})$ schließt mit der x-Achse, das heißt mit dem Vektor $\mathbf{e}_1$, den Winkel 30° beziehungsweise $\frac{\pi}{6}$ ein. Seine Länge wird der Physiker nach Pythagoras mit $\|\mathbf{v}\| = \sqrt{3^2 + \sqrt{3}^2} = \sqrt{12} = 2\sqrt{3}$ angeben. Welchen Anteil in x-Richtung hat dieser Vektor? Elementargeometrisch gesehen[4] ist der Cosinus des Winkels zwischen $\mathbf{v}$ und der x-Achse das Verhältnis aus dem gesuchten Anteil a (der Ankathete des Dreiecks) und der Länge des Vektors (der Hypothenuse des Dreiecks), also $\cos \sphericalangle(\mathbf{v}, \mathbf{e}_1) = \frac{a}{\|\mathbf{v}\|} = \frac{a}{2\sqrt{3}}$. Andererseits wissen wir[5], dass $\cos \frac{\pi}{6} = \frac{1}{2}\sqrt{3}$ ist. Aus beidem schließen wir $a = 3$. Der Vektor $\mathbf{v}$ hat in $\mathbf{e}_1$-Richtung also genau den Anteil 3, was nicht verwundert, wenn wir uns $\mathbf{v} = (3, \sqrt{3})$ vor Augen halten. Das gesuchte Skalarprodukt wird daher so aussehen, dass wir für irgeneinen Vektor $\mathbf{v} = (v_1, v_2)$ als Skalarprodukt $\langle \mathbf{v}, \mathbf{e}_1 \rangle = v_1$ erhalten möchten (und entsprechend natürlich auch $\langle \mathbf{v}, \mathbf{e}_2 \rangle = v_2$, so dass wir aus elementargeometrischen Betrachtungen heraus $\langle \mathbf{v}, \mathbf{e}_1 \rangle = \|\mathbf{v}\| \cdot \|\mathbf{e}_1\| \cdot \cos \sphericalangle(\mathbf{v}, \mathbf{e}_1)$ (und entsprechend wieder $\langle \mathbf{v}, \mathbf{e}_2 \rangle = \|\mathbf{v}\| \cdot \|\mathbf{e}_2\| \cdot \cos \sphericalangle(\mathbf{v}, \mathbf{e}_2)$ haben. Allgemein resultiert, wenn wir nicht $\mathbf{e_1}$ oder $\mathbf{e_2}$ hernehmen, sondern irgendeinen daraus linearkombinierten Vektor $\mathbf{w}$, als Ziel unserer Überlegungen das Produkt

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos \sphericalangle(\mathbf{v}, \mathbf{w}) \ ,$$

[3] Das hochgestellte Transpositionszeichen lassen wir weg, wie das die Physiker ja auch tun.

[4] an dieser Stelle mache sich der Leser eine entsprechende Zeichnung!

[5] oder schlagen es nach

und die linke Seite wird gebildet, indem die Komponenten der beiden Vektoren paarweise multipliziert und das Ergebnis aufsummiert wird. Übrigens können wir dann natürlich auch $\langle \mathbf{v}, \mathbf{v} \rangle$ nach dieser Regel bilden und erhalten nichts anderes als das Quadrat $\|\mathbf{v}\|^2$ der "pythagoräischen" Länge von $\mathbf{v}$, sowohl auf der linken Seite dieser Gleichung (wo wir die Quadrate der Komponenten des Vektors aufsummieren) als auch auf der rechten (wo der Cosinus von 0 eben 1 ergibt). Diese Konstruktion wird weiter unten den Namen des "Standardskalarprodukts" auf dem Vektorraum $\mathbb{R}^n$ verliehen bekommen, wenn sie in offensichtlicher Weise von 2 auf allgemein n Dimensionen erweitert wurde.

Zunächst abstrahieren wir von der konkreten Produktbildung in $\mathbb{R}^n$, um ein entsprechendes Skalarprodukt in allen Vektorräumen zur Verfügung zu haben. Was ist an der gerade eingeführten Skalarproduktbildung wesentlich? Zunächst einmal ist sie symmetrisch in den beiden Argumenten. Wäre das nicht der Fall, würde die "Winkel-zwischen-Vektoren"-Interpretation nicht funktionieren. Dann ist sie in jedem der beiden Argumente linear, denn zum Beispiel verdoppelt sich der Anteil von $\mathbf{v}$ in x-Richtung, wenn die Länge von $\mathbf{v}$ verdoppelt wird. Aus dem Umstand, dass wir $\langle \mathbf{v}, \mathbf{v} \rangle$ als Längenquadrat von $\mathbf{v}$ ansprechen möchten, folgt schließlich, dass für einen nichtverschwindenden Vektor $\mathbf{v}$ der Ausdruck $\langle \mathbf{v}, \mathbf{v} \rangle$ nicht Null sein sollte. Diese Eigenschaften des oben betrachteten Produkts machen wir zu den definierenden Axiomen des Skalarprodukts in der folgenden

Definition. Sei $(\mathrm{V}, +, \cdot)$ ein $\mathbb{R}$-Vektorraum. Eine Abbildung $\langle \cdot, \cdot \rangle : \mathrm{V} \times \mathrm{V} \to \mathbb{R}$ heißt **Skalarprodukt** auf V, wenn die folgenden Bedingungen für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathrm{V}$ und $\lambda \in \mathbb{R}$ erfüllt sind:

$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ ("Symmetrie")

$\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$ und $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = 0$ ("positive Definitheit")

$\langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle$ und $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ ("Bilinearität").

In diesem Fall heißt $(\mathrm{V}, +, \cdot, \langle \cdot, \cdot \rangle)$ *Euklidischer Vektorraum* über $\mathbb{R}$. Für ein $\mathbf{v} \in \mathrm{V}$ heißt die Zahl $\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ *Norm* von $\mathbf{v}$.

Zwei Bemerkungen dazu. Aus den ersten Blick wäre die Bezeichnung "Bilinearität" eigenartig, wenn wir doch scheinbar nur die Linearität im *ersten* der beiden Argumente von $\langle \cdot, \cdot \rangle$ gefordert haben. Allerdings zieht die Symmetrie des Produkts sofort nach sich, dass Linearität dann automatisch auch im *zweiten* Argument vorliegt. Zweitens: die Definitheitsforderung könnten wir kürzer auch als $\mathbf{v} \neq \mathbf{0} \Rightarrow \langle \mathbf{v}, \mathbf{v} \rangle > 0$ schreiben. Denn die Implikation $\mathbf{v} = \mathbf{0} \Rightarrow \langle \mathbf{v}, \mathbf{v} \rangle = 0$ ist sowieso klar, wenn wir uns an das akribisch hergeleitete $(-1) \cdot \mathbf{w} = -\mathbf{w}$ aus dem Kapitel über Lineare Räume, Seite 3, erinnern und die Linearität im ersten Argument verwenden: $\langle \mathbf{0}, \mathbf{0} \rangle = \langle \mathbf{w} - \mathbf{w}, \mathbf{0} \rangle = \langle \mathbf{w}, \mathbf{0} \rangle - \langle \mathbf{w}, \mathbf{0} \rangle = 0$. Die übrig bleibende Bedingung $\langle \mathbf{v}, \mathbf{v} \rangle = \mathbf{0} \Rightarrow \mathbf{v} = \mathbf{0}$ ist dann gleichbedeutend mit $\mathbf{v} \neq \mathbf{0} \Rightarrow \langle \mathbf{v}, \mathbf{v} \rangle \neq 0$.

Auch bei abstrakten Skalarprodukten, die wir nicht mehr so einfach elementargeometrisch interpretieren können, wird zuweilen die Größe $\sphericalangle(\mathbf{v}, \mathbf{w}) = \arccos\left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \, \|\mathbf{w}\|}\right)$ als *Winkel zwischen den Vektoren* $\mathbf{v}$ *und* $\mathbf{w}$ angesprochen. In diesem Sinne kann man zum Beispiel bei der Entwicklung periodischer Funktionen in Fourierreihen davon sprechen, dass die in diesem Zusammenhang als Basisvektoren auftretenden trigonometrischen Funktionen "senkrecht aufeinander stehen", auch wenn das die Anschauung etwas strapazieren mag. Aber wir entwickeln sowieso im Lauf der Beschäftigung mit mathematischen Objekten eine Art von "höherer Anschauung" – da wird uns eine solche Aussage ganz einleuchtend vorkommen. Nun nimmt bekanntermaßen die Cosinusfunktion nur Werte zwischen -1 und $+1$ an – der Ausdruck $\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \, \|\mathbf{w}\|}$ sollte das bitteschön daher auch tun, sonst funktioniert diese Setzung ja nicht. Das ist aber nicht unbedingt unmittelbar einsichtig, sondern vielmehr eine beweiswürdige Aussage. Man nennt sie *Cauchy-Schwarz*sche Ungleichung:

Bemerkung. Für jedes Skalarprodukt $\langle \cdot, \cdot \rangle$ auf V gilt für alle $\mathbf{v}, \mathbf{w} \in$ V die **Cauchy-Schwarzsche Ungleichung**

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leqslant \|\mathbf{v}\| \cdot \|\mathbf{w}\| \; .$$

Interessanterweise spielt diese Ungleichung insbesondere bei den in der Analysis von Funktionenräumen verwendeten Skalarprodukten eine enorme Rolle. Aber auch für bescheidene endlichdimensionale $\mathbb{R}^n$ ist sie wichtig, wie wir gleich bei der sogenannten *Dreiecksungleichung* sehen werden. Ein kurzer Blick auf den Beweis – weil diese Ungleichung so wichtig und weil der Beweis, zwei schöne kleine Kunstgriffe nutzend, so kurz ist. Also: seien $\mathbf{v}, \mathbf{w}$ Vektoren aus V. Wegen der Definitheit gilt dann, egal welchen Wert die Zahl $\lambda \in \mathbb{R}$ hat, die Ungleichung $\langle \mathbf{v} - \lambda \cdot \mathbf{w}, \mathbf{v} - \lambda \cdot \mathbf{w} \rangle \geqslant 0$ (*das* hinzuschreiben ist der erste Kunstgriff). Wir können das ausmultiplizieren unter Nutzung der Bilinearität und erhalten $\langle \mathbf{v}, \mathbf{v} \rangle - 2\lambda \langle \mathbf{v}, \mathbf{w} \rangle + \lambda^2 \langle \mathbf{w}, \mathbf{w} \rangle \geqslant 0$ beziehungsweise $\|\mathbf{v}\|^2 - 2\lambda \langle \mathbf{v}, \mathbf{w} \rangle + \lambda^2 \|\mathbf{w}\|^2 \geqslant 0$. Jetzt kommt der zweite Kunstgriff: diese Aussage gilt ja für jedes λ, niemand kann uns also davon abhalten, ein spezielles λ zu wählen, nämlich $\lambda := \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}$. Dieses spezielle λ eingesetzt macht die Ungleichung zu $\|\mathbf{v}\|^2 - 2\frac{\langle \mathbf{v}, \mathbf{w} \rangle^2}{\|\mathbf{w}\|^2} + \frac{\langle \mathbf{v}, \mathbf{w} \rangle^2}{\|\mathbf{w}\|^4} \cdot \|\mathbf{w}\|^2 \geqslant 0$

beziehungsweise, nach Durchmultiplizieren mit $\|\mathbf{w}\|^2$ dann $\|\mathbf{v}\|^2 \cdot \|\mathbf{w}\|^2 - \langle \mathbf{v}, \mathbf{w} \rangle^2 \geqslant 0$. Das ist aber gerade die CAUCHY-SCHWARZsche Ungleichung, wenn man noch die Wurzel zieht.

Mit Hilfe der CAUCHY-SCHWARZschen Ungleichung bekommen wir die bereits erwähnte **Dreiecksungleichung** für die durch $\langle \cdot, \cdot \rangle$ vermittelte Norm geliefert. Sie lautet einfach $\|\mathbf{v} + \mathbf{w}\| \leqslant \|\mathbf{v}\| + \|\mathbf{w}\|$ und ist eminent anschaulich: wenn wir zwei Pfeile aneinanderhängen, ist der so entstehende, vom Anfangspunkt des *ersten* zum Endpunkt des *zweiten* reichende, Pfeil nie länger als die Summe der Längen der einzelnen Pfeile. Wie uns die CAUCHY-SCHWARZschen Ungleichung das zeigt? Ganz einfach: $(\|\mathbf{v}\| + \|\mathbf{w}\|)^2 = \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2$ (Binomische Gleichung) $\geqslant \|\mathbf{v}\|^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2$ (CAUCHY-SCHWARZsche Ungleichung) $= \|\mathbf{v} + \mathbf{w}\|^2$ und daraus $\|\mathbf{v}\| + \|\mathbf{w}\| \geqslant \|\mathbf{v} + \mathbf{w}\|$ durch Wurzelziehen. Aus der Anschauung heraus dürfte auch klar sein, wann diese Ungleichung zur Gleichung wird: nämlich dann, wenn die beiden aneinander anzuhängenden Pfeile in die selbe Richtung weisen, das namensgebende Dreieck also eigentlich gar nicht richtig entsteht beziehungsweise degeneriert ist. Wer sich mit dieser Anschauung nicht zufriedengeben will, sondern auch dafür noch einen Beweis braucht, dem sei die entsprechende Übungsaufgabe (mit Anleitung) ans Herz gelegt und er muss sich Mathematiker nennen lassen...

Die versammelten Eigenschaften dieser Normbildung $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle$ sind zusammengefasst $\|\mathbf{v}\| \geqslant 0$ mit $\|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$, $\|\lambda \cdot \mathbf{v}\| = |\lambda| \cdot \|\mathbf{v}\|$ und eben die Dreiecksungleichung. Niemanden wird es jetzt wundern, wenn diese Eigenschaften wiederum axiomatisiert werden können: unabhängig von irgendeinem Skalarprodukt wird ein Linearer Raum zu einem *Normierten Raum*, wenn eine Abbildung $\| \cdot \| : \mathrm{V} \to \mathbb{R}$ auf ihm definiert ist, die den genannten Axiomen gehorcht. Eine Norm – also eine Art Maß für Größe oder Länge und damit auch für Abstand und Nähe in einem Vektorraum – wird eingesetzt zur Definition des Begriffs der *Konvergenz*. Die ganze Analysis baut darauf auf. Hier in der Linearen Algebra werden wir den Begriff des Normierten Raums jedoch nicht weiter vertiefen und lassen es bewenden bei einer Übungsaufgabe (mit Anleitung); wenn im Folgenden von der Norm eines Vektors die Rede ist, ist immer die gemeint, die das betreffende Skalarprodukt in der genannten Weise erzeugt.

Im Vektorraum $\mathbb{R}^n$ der n-Tupel reeller Zahlen ist das übliche, das sogenannte *Standardskalarprodukt*, dasjenige, das wir weiter oben für die Motivation und Einführung des Begriffs als solchem verwendet haben:

Definition. Gegeben der Vektorraum $\mathbb{R}^n$ der n-Tupel reeller Zahlen. Die Abbildung

$\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}; (\mathbf{v},\mathbf{w}) \mapsto \langle\mathbf{v},\mathbf{w}\rangle := \sum_{i=1}^n v_i w_i$

ist ein Skalarprodukt auf $\mathbb{R}^n$. Es heißt **Standardskalarprodukt** auf $\mathbb{R}^n$.

Das ist wieder sehr formal notiert – man nehme sich die paar Minuten, genau nachzuvollziehen, was da steht. Weil wir ja vereinbart haben, Vektoren aus $\mathbb{R}^n$ als *Spalten*vektoren aufzufassen, also als Matrizen aus $M(n \times 1, \mathbb{R})$, können wir das Standardskalarprodukt auch als Matrixmultiplikation schreiben. Wie? Ganz einfach: seien $\mathbf{v}$ und $\mathbf{w}$ zwei Spaltenvektoren aus $\mathbb{R}^n$ respektive zwei Matrizen aus $M(n \times 1, \mathbb{R})$, dann ist $\mathbf{v}^\mathsf{T}$ ein *Zeilen*vektor oder eben eine Matrix aus $M(1 \times n, \mathbb{R})$. Und nach den Regeln der Matrixmultiplikation ist sodann $\mathbf{v}^\mathsf{T}\mathbf{w}$ gerade die 1×1-Matrix (also *Zahl*) $\mathbf{v}^\mathsf{T}\mathbf{w} = \sum_i v_i w_i$ – und das ist eben das Standardskalarprodukt. Manchmal ist es einfach praktisch, sich $\langle\mathbf{v},\mathbf{w}\rangle$ als $\mathbf{v}^\mathsf{T}\mathbf{w}$ zu denken...

Zum Schluss dieses Abschnitts aber noch eine wichtige Bemerkung: das Standardskalarprodukt auf ist beileibe nicht das einzige Skalarprodukt auf $\mathbb{R}^n$. Man konsultiere die Übungen zu diesem Kapitel, um zu lernen, was hier alles möglich ist.

1.3 Orthogonalität.

Der Begriff "orthogonal" kommt (wieder einmal) aus dem Altgriechischen: "orthos" heißt "richtig" oder "recht", und "gonia" bedeutet "Winkel". Zwei zueinander orthogonale Vektoren stehen also senkrecht aufeinander – mehr heißt das nicht. Den Winkel zwischen zwei Vektoren eines Euklidischen Raumes haben wir im vorangegangenen Abschnitt zu $\sphericalangle(\mathbf{v},\mathbf{w}) = \arccos\left(\frac{\langle\mathbf{v},\mathbf{w}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|}\right)$ definiert, woraus sich sofort die Bedingung für Orthogonalität ergibt: der Winkel soll $\frac{\pi}{2}$ sein, der Cosinus mithin den Wert 0 haben, der Zähler in $\frac{\langle\mathbf{v},\mathbf{w}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$ also 0 sein – Orthogonalität von $\mathbf{v}$ und $\mathbf{v}$ bedeutet demnach nichts anderes als $\langle\mathbf{v},\mathbf{w}\rangle = 0$. Man bemerke, dass Orthogonalität ein Begriff ist, der relativ zum Skalarprodukt eines Euklidischen Raumes definiert ist. Ohne Skalarprodukt kein Winkel und ohne Winkel keine Orthogonalität!

Im Hauptteil dieses Abschnitts wollen wir auf das Ziel hinaus, Vektorraumbasen zu konstruieren, deren Vektoren wechselseitig senkrecht aufeinander stehen, also orthogonal sind. Wenn wir eine solche Basis haben, ist es natürlich sinnvoll, die Basisvektoren unter Zuhilfenahme des Skalarprodukts dann auch noch zu *normieren*, ihnen also die Länge 1 zu geben. Das ist dann nicht schwierig, denn aus irgeneinem Vektor $\mathbf{v}$ wird mittels Division durch seine Länge $\|\mathbf{v}\|$ der

Vektor $\frac{\mathbf{v}}{\|\mathbf{v}\|}$, der dann die Länge $\|\frac{\mathbf{v}}{\|\mathbf{v}\|}\| = \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|} = 1$ hat, ansonsten aber die Eigenschaften von $\mathbf{v}$ behält. Insbesondere können wir ein k-Tupel linear unabhängiger Vektoren $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ auf diese Weise zu einem k-Tupel linear unabhängiger *normierter* Vektoren machen und ein Erzeugendensystem von k Vektoren zu einem Erzeugendensystem von k *normierten* Vektoren. Man mache sich klar, dass die Skalierung eines jeden jeweils enthaltenen Vektors mit dem Inversen seiner Norm (oder mit irgeneinem anderen Faktor) keinen Einfluss darauf hat, dass das gegebene Tupel ein Erzeugendensystem ist oder linear unabhängig!

Nun wünschen wir uns – nach dem Vorbild der kanonischen Basis der Einheitsvektoren $\mathbf{e}_i$ im $\mathbb{R}^n$ – für unseren Vektorraum V aber nicht nur eine Basis, deren Vektoren normiert sind, sondern eine, deren Vektoren auch noch wechselseitig orthogonal sind. Man spricht dann von einer Basis, die ein Ortho*normal*system ist, und meint folgendes:

> Ein **Orthonormalsystem** in einem Euklidischen Vektorraum V ist ein k-Tupel $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, für dessen Vektoren gilt $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ für $i \neq j$ sowie $\|\mathbf{v}_i\| = 1$ für alle i. Wir können dies auch kurz als $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$ schreiben, wenn δ_{ij} das KRONECKER-Symbol meint mit $\delta_{ij} = 0$ für $i \neq j$ und $\delta_{ii} = 1$ für alle i.

Der Begriff des Orthonormalsystems beinhaltet noch nicht weitere Eigenschaften wie zum Beispiel die, ein Erzeugendensystem zu sein, aber linear unabhängig sind die Vektoren eines Orthonormalsystems schon. Anschaulich ist das klar, denn wie will man einen der Vektoren aus den anderen zusammenbauen, wenn alle wechselseitig senkrecht aufeinander stehen? Der formale Nachweis ist zudem extrem einfach: wir müssen nur die Linearkombination $\lambda_1\mathbf{v}_1 \cdot \cdots \cdot \lambda_k\mathbf{v}_k$ mit einem der $\mathbf{v}_i$ im Skalarprodukt zusammenbringen und $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$ nutzen: aus $\lambda_1\mathbf{v}_1 \cdots \cdot \lambda_k\mathbf{v}_k = 0$ folgt dann $\langle \lambda_1\mathbf{v}_1 \cdots \cdot \lambda_k\mathbf{v}_k, \mathbf{v}_i \rangle = \lambda_i = 0$ und so für alle i; sämtliche λ_i sind also 0. Haben wir jetzt aber eine Basis gegeben, also ein Erzeugendensystem linear unabhängiger Vektoren, wie machen wir daraus eine **Orthonormalbasis**, will sagen eine Basis, die gleichzeitig Orthonormalsystem ist?

Das Prozedere dazu ist sehr anschaulich. Nehmen wir den einfachsten Fall: in der Euklidischen Ebene $\mathbb{R}^2$ haben wir zwei Vektoren $\mathbf{v}$ und $\mathbf{w}$, die linear unabhängig sein mögen (mithin eine Basis des $\mathbb{R}^2$ bilden). Wir nehmen uns den ersten, $\mathbf{v}$, und normieren ihn. Der zweite, $\mathbf{w}$, mag irgendeinen Winkel zu $\mathbf{v}$ einschließen. Wie konstruieren wir daraus einen zu $\mathbf{v}$ senkrechten Vektor? Man mache sich eine Zeichnung (oder schaue nochmal auf die von oben): zunächst nehmen wir den Anteil von $\mathbf{w}$ in $\mathbf{v}$-Richtung, das ist $\langle\mathbf{w},\mathbf{v}\rangle\cdot\mathbf{v}$. Und dann verketten wir diesen mit $\mathbf{w}$ in der Form $\mathbf{w}-\langle\mathbf{w},\mathbf{v}\rangle\cdot\mathbf{v}$. Das Resultat ist zwar nicht normiert, steht aber jedenfalls, wie die Anschauung zeigt, senkrecht auf $\mathbf{v}$. Letzteres lässt sich natürlich auch formal einsehen: es ist $\langle\mathbf{w}-\langle\mathbf{w},\mathbf{v}\rangle\cdot\mathbf{v},\mathbf{v}\rangle=\langle\mathbf{w},\mathbf{v}\rangle-\langle\mathbf{w},\mathbf{v}\rangle\langle\mathbf{v},\mathbf{v}\rangle$. Weil aber $\langle\mathbf{v},\mathbf{v}\rangle=1$ ist – schließlich hatten wir $\mathbf{v}$ im ersten Schritt normiert –, ist dieses Skalarprodukt 0. Im letzten Schritt normieren wir dann auch den neuen $\mathbf{w}$. In drei Dimensionen geht das genauso: dort nehmen wir an, die beiden Vektoren $\mathbf{u}$ und $\mathbf{v}$ wären bereits normiert und orthogonal, und $\mathbf{w}$ stünde irgendwie schräg dazu, voraussetzungsgemäß natürlich linear unabhängig. Dann bilden wir aus $\mathbf{w}$ den Vektor $\mathbf{w}-\langle\mathbf{w},\mathbf{u}\rangle\mathbf{u}-\langle\mathbf{w},\mathbf{v}\rangle\mathbf{v}$. Wie in zwei Dimensionen sehen wir anschaulich oder rechnerisch[6], dass dieser neue $\mathbf{w}$ senkrecht steht auf den andern beiden, und wir brauchen ihn schließlich nur noch zu normieren. Auf diese Weise bauen wir aus jeder vorgegebenen Basis eine Orthonormalbasis. Das ist die folgende Bemerkung wert:

> Jeder endlichdimensionale Euklidische Vektorraum besitzt eine Orthonormalbasis.

Das geschilderte Verfahren ist offenbar *konstruktiv*, das heißt es kann rechnerisch beziehungsweise algorithmisch[7] tatsächlich durchgeführt werden und liefert nach endlich vielen Schritten in der Tat eine konkret gegebene Orthonormalbasis. Man kann sich vorstellen, wie praktisch es ist, konkrete Rechnungen in einer Orthonormalbasis durchführen zu können im Vergleich zu irgendwelchen Basisvektoren, von denen man nur weiß, dass es genug davon gibt und dass sie linear unabhängig sind. Gelegentlich spricht man vom GRAM-SCHMIDTschen Orthonormalisierungsverfahren.

1.4 Orthogonale Abbildungen.

Wir haben das Grundschema bereits kennengelernt: sobald ein sogenannter “Raum” geschaffen wurde, indem einer Menge gewisse Zusatzstrukturen aufgenötigt wurden, betrachtet man in der Folge Abbildungen zwischen diesen Räumen, die die Zusatzstrukturen “respektieren”. So haben wir die Linearen Abbildungen und die ihnen zugehörigen Matrizen eingeführt als diejenigen Abbildungen, die die lineare, das heißt: die Vektorraumstruktur beachten. Und auch für die Euklidischen Vektorräume gibt es wieder einen spezifischen strukturerhaltenden Abbildungstyp, genannt *Orthogonale Abbildung*. Orthogonale Abbildungen sind natürlich *auch* Lineare Abbildungen, weil

[6] natürlich ausnutzend, dass bereits $\langle\mathbf{u},\mathbf{v}\rangle=0$ *ist*.

[7] Es sei allerdings bemerkt, dass das Verfahren numerisch gesehen nicht gut konditioniert ist.

ja Euklidische Vektorräume eben *auch* Vektorräume sind. Zusätzlich respektieren sie aber auch noch das Skalarprodukt. Und zwar in der spontan einsichtigen Weise, wie die folgende Definition zeigt. Wie andernorts bereits praktiziert, schleppen wir zumindest temporär Subskripte mit, die andeuten, in welchem der beiden beteiligten Vektorräume die betreffenden Objekte "leben"; wenn ein gewisser Gewöhnungseffekt eingetreten ist, lassen wir sie wieder weg.

> **Definition.** Seien $(\mathrm{V}, +_{\mathrm{V}}, \cdot_{\mathrm{V}}, \langle\cdot,\cdot\rangle_{\mathrm{V}})$ und $(\mathrm{W}, +_{\mathrm{W}}, \cdot_{\mathrm{W}}, \langle\cdot,\cdot\rangle_{\mathrm{W}})$ zwei Euklidische Vektorräume und $L : \mathrm{V} \to \mathrm{W}$ eine Lineare Abbildung. L heißt **orthogonal**, wenn für alle $\mathbf{v}, \mathbf{w} \in \mathrm{V}$ gilt: $\langle L\mathbf{v}, L\mathbf{w}\rangle_{\mathrm{W}} = \langle \mathbf{v}, \mathbf{w}\rangle_{\mathrm{V}}$.

Man gehe dieser Bezeichnungsweise nicht auf den Leim. Wenn eine Lineare Abbildung respektive die zugehörige Matrix "orthogonal" sind, bedeutet das nicht, sie stünden in irgend einer Weise "senkrecht" auf was auch immer. Es heißt vielmehr, dass sie ein Paar von Vektoren, die senkrecht aufeinander stehen, in ein Paar von Vektoren überführen, die wiederum senkrecht aufeinander stehen.

Als Konsequenz davon haben Orthogonale Abbildungen die vorteilhafte Eigenschaft, Orthonormalsysteme in Orthonormalsysteme zu überführen. Diese Aussage wollen wir noch ein bisschen ausbauen, sie gilt nämlich auch umgekehrt. Sei dazu $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ eine Orthonormalbasis von V (siehe letzter Abschnitt: das heißt eine Basis, deren Vektoren normiert sind und die wechselseitig senkrecht aufeinander stehen, will heißen verschwindendes Skalarprodukt haben). Die Standardbasis des $\mathbb{R}^n$ zum Beispiel, die aus den Vektoren $\mathbf{e}_1 = (1, 0, 0, \ldots, 0)^\mathsf{T}, \mathbf{e}_2 = (0, 1, 0, \ldots, 0)^\mathsf{T}, \ldots, \mathbf{e}_n = (0, 0, 0, \ldots, 1)^\mathsf{T}$ besteht, hat diese Eigenschaft bezüglich des Standardskalarprodukts. Unter Benutzung des oben genannten KRONECKER-*Symbols* δ_{ij}, das wie gesagt durch $\delta_{ii} = 1$ und $\delta_{ij} = 0$ für $i \neq j$ definiert ist, kann man die Eigenschaft einer Orthonormalbasis kurz und knapp so schreiben: $\langle \mathbf{e}_i, \mathbf{e}_j\rangle = \delta_{ij}$. Oder in Worten: das Skalarprodukt eines Basisvektors mit sich selbst ist 1, er hat also auch die Länge 1. Das Skalarprodukt eines Basisvektors mit irgend einem *anderen* Basisvektor ist hingegen 0. Und jetzt gilt folgendes: seien V und W zwei Euklidische Vektorräume, und sei $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ irgendeine Orthonormalbasis von V. Dann ist eine lineare Abbildung $L : \mathrm{V} \to \mathrm{W}$ *genau dann* orthogonal, wenn $\{L\mathbf{e}_1, \ldots, L\mathbf{e}_n\}$ ein Orthonormalsystem von W ist. Nicht notwendig eine Basis, denn $\dim W$ darf in dieser Aussage ja durchaus von $\dim V$ abweichen.

Die eine Beweisrichtung ist simpel: *wenn* L orthogonal ist, *dann* gilt $\langle L\mathbf{e_i}, L\mathbf{e_j}\rangle = \langle \mathbf{e_i}, \mathbf{e_j}\rangle = \delta_{ij}$. Nun sei umgekehrt $\{L\mathbf{e}_1, \ldots L\mathbf{e}_n\}$ als Orthonormalsystem in W vorausgesetzt, also $\langle L\mathbf{e_i}, L\mathbf{e_j}\rangle = \delta_{ij}$.

Wir müssen zeigen, dass dann L orthogonal ist. Um das einzusehen, wenden wir es auf irgendwelche Vektoren $\mathbf{v}, \mathbf{w} \in \mathrm{V}$ an, entwickeln diese beiden dann aber jeweils in die Orthonormalbasis der $\mathbf{e}_i$ von V. Dann rechnen wir unter Benutzung der Bilinearität des Skalarprodukts $\langle L\mathbf{v}, L\mathbf{w}\rangle = \langle L(\sum_i \lambda_i \mathbf{e_i}), L(\sum_j \mu_j \mathbf{e_j})\rangle = \sum_i \sum_j \lambda_i \mu_j \langle L\mathbf{e_i}, L\mathbf{e_j}\rangle = \sum_i \sum_j \lambda_i \mu_j \delta_{ij} = \sum_i \sum_j \lambda_i \mu_j \langle \mathbf{e_i}, \mathbf{e_j}\rangle =$
$\langle \sum_i \lambda_i \mathbf{e_i}, \sum_j \mu_j \mathbf{e_j}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$. Damit ist L als orthogonal erkannt.— Das ist nicht schwierig, übt aber ein wenig den Umgang mit Summationssymbolen; zudem sollte der Leser zumindest in Gedanken die jeweiligen Subskripte "V" und "W" dazuschreiben, um sich darüber klarzuwerden, in welchem der beiden Räume er sich gerade befindet.

Die für die konkrete Matrizenrechnung daraus abgeleitete Konsequenz ist die folgende: die Spalten einer orthogonalen quadratischen Matrix – als "Bilder der Einheitsvektoren"! – sind bezüglich des Standardskalarprodukts in $\mathbb{R}^n$ eine Orthonormalbasis. Weil die Multiplikation $A \cdot B$ zweier Matrizen $A, B \in \mathbb{R}^{n\times n}$ darauf hinausläuft, immer mit den *Zeilen* von A und den *Spalten* von B Standardskalarprodukte zu bilden, können wir für ein orthogonales $A \in \mathbb{R}^{n\times n}$ schlussfolgern: $A^\mathsf{T} \cdot A = E$. Zudem ist leicht ersichtlich, dass eine orthogonale Abbildung L *immer* injektiv ist. Sei nämlich $\mathbf{v} \in \operatorname{Kern} L$, also $L\mathbf{v} = \mathbf{0}$, dann gilt für den Betrag von $\mathbf{v}$ ja $\|\mathbf{v}\|_\mathrm{V} = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle_\mathrm{V}} = \sqrt{\langle L\mathbf{v}, L\mathbf{v}\rangle_\mathrm{W}} = \sqrt{\langle \mathbf{0}, \mathbf{0}\rangle_\mathrm{W}} = 0$, also muss $\mathbf{v} = \mathbf{0}$ sein und $\dim \operatorname{Kern} L = 0$. Mit anderen Worten – man erinnere sich an die Dimensionsformel $\dim \operatorname{Kern} L + \dim \operatorname{Bild} L = \dim \mathrm{W}$! –: *orthogonale quadratische Matrizen sind immer invertierbar.*

Und dieses $A^\mathsf{T} \cdot A = E$ ist ein wichtiges und bemerkenswertes Ergebnis, denn das heißt ja nichts anderes als $A^\mathsf{T} = A^{-1}$: *das Inverse einer orthogonalen quadratischen Matrix ist gleich ihrem Transponierten.* Wir fassen zusammen. Für eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ sind die folgenden Aussagen gleichwertig:

(i) A ist orthogonal.

(ii) Die Spalten von A bilden eine orthonormale Basis von $\mathbb{R}^n$.

(iii) Die Zeilen von A bilden eine orthonormale Basis von $\mathbb{R}^n$.

(iv) Es gilt $A \cdot A^\mathsf{T} = E$.

(v) Es gilt $A^\mathsf{T} \cdot A = E$.

(vi) A ist invertierbar mit $A^{-1} = A^\mathsf{T}$

Auch für die Determinante einer quadratischen orthogonalen Matrix gibt es eine zusätzliche interessante Aussage. Nicht nur, dass sie nicht 0 ist, können wir formulieren (was aus ihrer Invertierarkeit folgt), sondern sogar noch mehr: weil wir ja wissen, dass $\det(A \cdot B) = \det A \cdot \det B$ und stets $\det A^{\mathsf{T}} = \det A$ ist, können wir für *orthogonale* quadratische Matrizen A schlussfolgern $1 = \det(E) = \det(A \cdot A^{-1}) = \det(A \cdot A^{\mathsf{T}}) = \det(A) \cdot \det(A^{\mathsf{T}}) = (\det A)^2$, also $|\det A| = 1$. Eine orthogonale quadratische Matrix hat Determinante 1 oder -1. Diejenigen orthogonalen Matrizen A, die speziell $\det A = 1$ aufweisen, heißen *spezielle orthogonale Matrizen*, aber das muss man sich nicht unbedingt merken. Auch wenn es für sich genommen interessant ist: die Menge dieser Matrizen, mit $\mathrm{SO}(n)$ für n Dimensionen abgekürzt, bildet mit der Matrixmultiplikation als Verknüpfung selbst wieder eine Gruppe, die "Spezielle Orthogonale Gruppe". Man überzeuge sich beispielsweise, dass das Matrixprodukt zweier solcher Matrizen wiederum in $\mathrm{SO}(n)$ liegt. — Im Übrigen sei warnend vermerkt, dass aus der Orthogonalität einer Matrix zwar folgt, dass der Betrag ihrer Determinante den Wert 1 hat, umgekehrt aus $\det M = 1$ oder $\det M = -1$ aber keineswegs bereits geschlossen werden kann, dass es sich bei M um eine orthogonale Matrix handelt! Ein abschreckendes Beispiel dazu in den Übungen.

Die für viele Anwendungen wichtigsten orthogonalen Transformationen sind die im dreidimensionalen Anschauungsraum stattfindenden Drehungen und Spiegelungen. Bei beiden Operationen ist anschaulich klar, dass sie winkelerhaltend sind. Ein konkretes Beispiel für den Fall der Drehung in einer Ebene um einen gegebenen Winkel findet sich in den Übungen zu diesem Kapitel.

1.5 Standardskalarprodukt und Matrixtransposition.

Zum Schluss betrachten wir noch einmal die Transposition als solche im Zusammenhang mit dem Skalarprodukt. Sei $M = (M_{ij})$ eine beliebige – nicht mehr notwendig orthogonale – Matrix, dann bilden wir ihre Transponierte bekanntlich dadurch, dass wir ihre Zeilen zu Spalten machen und ihre Spalten zu Zeilen: $M^{\mathsf{T}}_{ij} = M_{ji}$. Nun bezeichne $\langle \cdot, \cdot \rangle$ das *Standardskalarprodukt* auf $\mathbb{R}^n$, das heißt – und hier üben wir uns ein weiteres Mal in der Verwendung des Summenzeichens – wir betrachten $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} v_i w_i$, wenn v_i die Komponenten des Spaltenvektors $\mathbf{v}$ sind und w_i die von $\mathbf{w}$. In dieser Schreibweise hat der Vektor $M\mathbf{v}$ die Komponenten $(M\mathbf{v})_i$ mit $(M\mathbf{v})_i = \sum_j M_{ij} v_j$; die Summationen laufen ja stets von $i, j, k, \ldots = 1$ bis n und wir lassen die immer gleichen Summationsgrenzen daher einfach weg. Das Skalarpodukt von $M\mathbf{v}$ mit

$\mathbf{w}$ schreibt sich so: $\langle M\mathbf{v}, \mathbf{w}\rangle = \sum_i (M\mathbf{v})_i w_i = \sum_i \sum_j M_{ij} v_j w_i$. In einer solchen Formel können wir die Summationsindizes natürlich nennen, wie wir wollen, i oder j oder α oder wie auch immer – nur müssen *verschiedene* Indizes mit *verschiedenen* Buchstaben bezeichnet sein. Niemand kann uns daher hindern, in dieser Formel i in j umzubenennen und j in i, so dass wir $\langle M\mathbf{v}, \mathbf{w}\rangle = \sum_j \sum_i M_{ji} v_i w_j$ bekommen. Aber definitionsgemäß ist ja $M_{ji} = M^{\mathsf{T}}_{ij}$, wir haben also $\langle M\mathbf{v}, \mathbf{w}\rangle = \sum_j \sum_i M^{\mathsf{T}}_{ij} v_i w_j = \sum_i \sum_j M^{\mathsf{T}}_{ij} w_j v_i = \sum_i (M^{\mathsf{T}}\mathbf{w})_i v_i = \langle \mathbf{v}, M^{\mathsf{T}}\mathbf{w}\rangle$. Die quasi koordinatenunabhängige Charakterisierung der Matrizentransposition, die wir hier gefunden haben, ist verallgemeinerbar (zum Beispiel für das Standardskalarprodukt auf $\mathbb{C}^n$ und die entsprechenden Matrizenoperationen “komplex-konjugiert-Transponieren”) und schon von daher eine Bemerkung wert:

> **Bemerkung.** Für eine Matrix $M \in M(n \times n, \mathbb{R})$, ihre Transponierte M^{T} und zwei beliebige Vektoren $\mathbf{v}$ und $\mathbf{w}$ aus $\mathbb{R}^n$ gilt $\langle M\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, M^{\mathsf{T}}\mathbf{w}\rangle$, wenn $\langle\cdot,\cdot\rangle$ das Standardskalarprodukt auf $\mathbb{R}^n$ ist.

Mittels der weiter oben eingeführten Schreibweise des Standardskalarprodukts als Matrixmultiplikation $\langle \mathbf{v}, \mathbf{w}\rangle \equiv \mathbf{v}^{\mathsf{T}}\mathbf{w}$ ist diese Beziehung noch eingängiger: $\langle M\mathbf{v}, \mathbf{w}\rangle \equiv (M\mathbf{v})^{\mathsf{T}}\mathbf{w} = (\mathbf{v}^{\mathsf{T}} M^{\mathsf{T}})\mathbf{w} = \mathbf{v}^{\mathsf{T}}(M^{\mathsf{T}}\mathbf{w}) \equiv \langle \mathbf{v}, M^{\mathsf{T}}\mathbf{w}\rangle$, wenn wir uns an die Matrixformel $(KM)^{\mathsf{T}} = M^{\mathsf{T}} K^{\mathsf{T}}$ erinnern[8]. Man beachte aber bei diesen schönen und einfachen Rechnungen immer, dass es sich dabei um das *Standard*skalarprodukt handelt und nicht um irgendeine abstrakte Konstruktion.

Für den speziellen Fall, dass die vorliegende Matrix A auch noch *orthogonal* ist, geben uns diese Schreibweisen eine Möglichkeit, $A^{-1} = A^{\mathsf{T}}$ einzusehen, die genauso einfach ist wie das auf Seite 10 vorgebrachte Argument mit der “Zeilen auf Spalten”-Skalarproduktbildung bei der Matrixmultiplikation. Denn die Orthogonalität von A heißt definitionsgemäß ja $\langle \mathbf{v}, \mathbf{w}\rangle = \langle A\mathbf{v}, A\mathbf{w}\rangle$ für alle möglichen Wahlen von $\mathbf{v}$ und $\mathbf{w}$, also ist $\mathbf{v}^{\mathsf{T}}\mathbf{w} = (A\mathbf{v})^{\mathsf{T}}(A\mathbf{w}) = \mathbf{v}^{\mathsf{T}}(A^{\mathsf{T}}A)\mathbf{w}$. Und weil dies für *alle* Paarungen von Vektoren $\mathbf{v}$, $\mathbf{w}$ gilt – man wähle zum Beispiel hintereinander verschiedene Vektoren der Standardbasis –, muss offenbar $A^{\mathsf{T}}A = E$ sein, mit anderen Worten $A^{\mathsf{T}} = A^{-1}$.

Stichwort “Standardbasis”: zusammen mit dem Standardskalarprodukt liefert sie eine schöne Möglichkeit, die Komponenten einer vorgegebenen allgemeinen Matrix M zu identifizieren. Man muss sich dazu lediglich vor Augen halten, dass die k-te Komponente des i-ten Standardein-

[8] und, das sollte man vielleicht explizit auch mal wieder erwähnen, an die Assoziativität der Matrixmultiplikation.

heitsvektors $\mathbf{e}_i = (0, \dots, 1, \dots, 0)^\mathsf{T}$ (die "1" natürlich an der i-ten Stelle) definitionsgemäß 1 ist, falls $k = i$, und 0 sonst – mittels des KRONECKER-Symbols auf den Punkt gebracht $(\mathbf{e}_i)_k = \delta_{ik}$. Der Rest ist wieder nur eine Geduldsübung in Summationssachen. Denn der Vektor $M\mathbf{e}_j$ hat dann die Komponenten $(M\mathbf{e}_j)_i = \sum_k M_{ik}(\mathbf{e}_j)_k = \sum_k M_{ik}\delta_{jk} = M_{ij}$; für das letzte Gleichheitszeichen muss man sich vergegenwärtigen, welche Terme der Summation auf Grund der Definition des KRONECKER-Symbols wegfallen – fast alle nämlich[9] Nutzen wir $(M\mathbf{e}_j)_k = M_{kj}$ innerhalb eines Skalarprodukts mit $\mathbf{e}_i$ aus: $\langle \mathbf{e}_i, M\mathbf{e}_j \rangle = \sum_k (\mathbf{e}_i)_k (M\mathbf{e}_j)_k = \sum_k \delta_{ik} M_{kj} = M_{ij}$. Das simple Ergebnis unserer Rechnung können wir also wie folgt formulieren:

> Für das Standardskalarprodukt $\langle \cdot, \cdot \rangle$ und die Standardeinheitsbasis $\{\mathbf{e}_i\}$, $i = 1 \dots n$, auf $\mathbb{R}^n$ gilt für eine beliebige quadratische Matrix M und ihre Komponenten M_{ij}, dass $\langle \mathbf{e}_i, M\mathbf{e}_j \rangle = M_{ij}$ ist.

Indem wir M im Skalarprodukt mit Einheitsvektoren in die Zange nehmen, können wir jede ihrer Komponenten quasi herauspräparieren. Für das glatte Hin- und Herwechseln zwischen komponentenweisen Ausdrücken und koordinatenfreien Schreibweisen sind solche Beziehungen das Schmiermittel[10]. Eine gewisse Übung bei solchen Manövern vorausgesetzt, gelangt man sehr schnell zu den Aussagen, auf die man abzielt. Zum Beispiel ist für eine *orthogonale* Matrix A die Gleichung $\delta_{ij} = \langle \mathbf{e}_i, \mathbf{e}_i \rangle = \langle A\mathbf{e}_i, A\mathbf{e}_j \rangle = \langle \mathbf{e}_i, (A^\mathsf{T} A)\mathbf{e}_j \rangle = (AA^\mathsf{T})_{ij}$ eine weitere rasche Möglichkeit, einzusehen, dass $A^\mathsf{T} = A^{-1}$ ist, denn die Zahlen δ_{ij} sind ja nichts anderes als E_{ij}, die Komponenten der Einheitsmatrix E.

---

[9]... wobei $(M\mathbf{e}_j)_i = M_{ij}$ ja nun wieder einmal nichts anderes ist als eine Formelversion des wohlbekannten Merkspruchs "Die Spalten / sind Bilder / der Einheits- / Vektoren".

[10]Man sollte so etwas gut beherrschen. Die Physiker sowieso, aber auch die Informatiker, denn die Index- respektive Korrdinatenschreibweisen sind ja die, die in den Schleifen entsprechender Algorithmen verwendet werden. Sicher bieten moderne Programmiersprachen generell die Möglichkeit, koordinatenfrei zu coden, also Matrizen und Vektoren als Variablen einzusetzen und die für solche Typen vorgefertigten Rechenoperationen schleifenfrei zu nutzen. Manchmal, wenn man etwas Besonderes vorhat, muss man diese Variablen aber doch in Einzelkomponenten aufbrechen...