immer, wenn von Basen die Rede ist, zur Vermeidung logischer Schwierigkeiten auf den endlichdimensionalen Fall beschränken:

Satz: Jeder endlichdimensionale EUKLIDische oder HERMITEsche Vektorraum V hat eine Orthonormalbasis.

Beweis: Zunächst wissen wir, daß V überhaupt eine Basis $\{\vec{b}_1, \dots, \vec{b}_n\}$ hat; daraus konstruieren wir schrittweise eine *Orthogonalbasis* nach dem sogenannten GRAM-SCHMIDTschen Orthogonalisierungsverfahren. Dieses liefert in seinem r-ten Schritt eine Orthogonalbasis $\{\vec{c}_1, \dots, \vec{c}_r\}$ des von dern ersten r Basisvektoren aufgespannten Untervektorraums $[\vec{b}_1, \dots, \vec{b}_r]$.

Der erste Schritt ist der einfachste: Da es noch keine Orthogonalitätsbedingung für $\vec{c}_1$ gibt, können wir einfach $\vec{c}_1 = \vec{b}_1$ setzen.

Nachdem wir $r \geq 1$ Schritte durchgeführt haben, haben wir r linear unabhängige Vektoren $\vec{c}_1, \dots, \vec{c}_r$ mit $\vec{c}_i \cdot \vec{c}_j = 0$ für $i \neq j$ aus dem von $\vec{b}_1, \dots, \vec{b}_r$ aufgespannten Untervektorraum. Ist $r = n$, haben wir eine Orthogonalbasis; andernfalls muß ein auf den bisher konstruierten $\vec{c}_i$ senkrecht stehender Vektor $\vec{c}_{r+1}$ gefunden werden, der zusammen mit diesen den von $\vec{b}_1$ bis $\vec{b}_{r+1}$ erzeugten Untervektorraum erzeugt.

Da $\vec{c}_1, \dots, \vec{c}_r$ und $\vec{b}_1, \dots, \vec{b}_r$ denselben Untervektorraum erzeugen, gilt dasselbe für $\vec{c}_1, \dots, \vec{c}_r, \vec{b}_{r+1}$ und $\vec{b}_1, \dots, \vec{b}_{r+1}$; das Problem ist, daß $\vec{b}_{r+1}$ im allgemeinen nicht orthogonal zu den $\vec{c}_i$ sein wird. Wir dürfen $\vec{b}_{r+1}$ aber abändern um einen beliebigen Vektor aus dem von $\vec{c}_1, \dots, \vec{c}_r$ aufgespannten Untervektorraum; also setzen wir

$$\vec{c}_{r+1} = \vec{b}_{r+1} + \lambda_1 \vec{c}_1 + \cdots + \lambda_r \vec{c}_r$$

und versuchen, die λ_i so zu bestimmen, daß dieser Vektor orthogonal zu $\vec{c}_1, \dots, \vec{c}_r$ wird.

Wegen der Orthogonalität der $\vec{c}_i$ ist

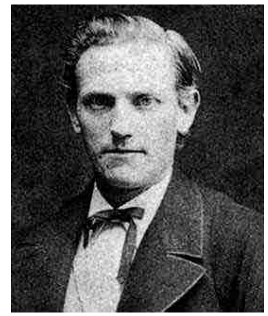
$$\vec{c}_{r+1} \cdot \vec{c}_i = \vec{b}_{r+1} \cdot \vec{c}_i + \sum_{j=1} \lambda_j (\vec{c}_j \cdot \vec{c}_i) = \vec{b}_{r+1} \cdot \vec{c}_i + \lambda_i (\vec{c}_i \cdot \vec{c}_i)\,;$$

setzen wir daher $\lambda_i = -\dfrac{\vec{b}_{r+1} \cdot \vec{c}_i}{\vec{c}_i \cdot \vec{c}_i}$, so ist $\vec{v} \cdot \vec{c}_i = 0$ für alle $i = 1, \dots, r$.

Nach dem n-ten Schritt haben wir eine Orthogonalbasis $\{\vec{c}_1, \dots, \vec{c}_n\}$ von V konstruiert. Daraus wird die gewünschte *Orthonormalbasis* $\{\vec{e}_1, \dots, \vec{e}_n\}$, wenn wir jeden Vektor durch seine Länge dividieren, d.h.

$$\vec{e}_i = \frac{\vec{c}_i}{\|\vec{c}_i\|}\,.$$

■



Der dänische Mathematiker JØRGAN PEDERSEN GRAM (1850–1916) lehrte an der Universität Kopenhagen, war aber gleichzeitig auch noch geschäftsführender Direktor einer Versicherungsgesellschaft und Präsident des Verbands der dänischen Versicherungsunternehmen. Er publizierte anscheinend nur eine einzige mathematische Arbeit *Sur quelque théorèmes fondamentaux de l'algèbre moderne,* die 1874 erschien. Das GRAM-SCHMIDTsche Orthogonalisierungsverfahren, durch das er heute hauptsächlich bekannt ist, stammt wohl von LAPLACE (1749–1827) und wurde auch schon 1836 von CAUCHY verwendet.

ERHARD SCHMIDT (1876–1959) wurde in Estland geboren; er studierte in Berlin bei SCHWARZ und promovierte in Göttingen bei HILBERT.

ERHARD SCHMIDT ist einer der Begründer der modernen Funktionalanalysis; insbesondere geht die Verallgemeinerung EUKLIDischer und HERMITEscher Vektorräume zu sogenannten HILBERT-Räumen, mit der wir uns im nächsten Semester im Zusammenhang mit der FOURIER-Analysis und der Theorie der Differentialgleichungen beschäftigen werden, auf ihn zurück. Er war Professor in Zürich, Erlangen, Breslau und Berlin.

Um auch den Umgang mit HERMITEschen Skalarprodukten und das Rechnen mit komplexen Zahlen zu üben, betrachten wir als Beispiel den von

$$\vec{b}_1 = \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 1+2i \\ -2+i \\ -i \\ -1 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} 1+2i \\ -i \\ -4-i \\ 1+2i \end{pmatrix}$$

aufgespannten Untervektorraum von $\mathbb{C}^4$.

Wie oben im Beweis können wir bei der Anwendung des GRAM-SCHMIDTschen Orthogonalisierungsverfahrens $\vec{c}_1 = \vec{b}_1$ setzen; im zweiten Schritt suchen wir einen Vektor $\vec{c}_2 = \vec{b}_2 + \lambda_1\vec{c}_1$ mit der Eigenschaft, daß

$$\vec{c}_2 \cdot \vec{c}_1 = (\vec{b}_2 + \lambda_1\vec{c}_1) \cdot \vec{c}_1 = \vec{b}_2 \cdot \vec{c}_1 + \lambda_1(\vec{c}_1 \cdot \vec{c}_1) = 0$$

ist.

(Wir können natürlich auch fordern, daß

$$\vec{c}_1 \cdot \vec{c}_2 = \vec{c}_1 \cdot (\vec{b}_2 + \lambda_1\vec{c}_1) = \vec{c}_1 \cdot \vec{b}_2 + \overline{\lambda_1}(\vec{c}_1 \cdot \vec{c}_1)$$

verschwindet, aber dann erhalten wir zunächst eine Formel für $\overline{\lambda_1}$ und müssen noch komplex konjugieren, Daher ist zumindest im komlexen Fall die Forderung $\vec{c}_2 \cdot \vec{c}_1 = 0$ rechnerisch etwas bequemer.)

Wir brauchen also die HERMITEschen Skalarprodukte

$$\vec{b}_2 \cdot \vec{c}_1 = \begin{pmatrix} 1+2i \\ -2+i \\ -i \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix} \quad \text{und } \vec{c}_1 \cdot \vec{c}_1 = \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix} .$$

Bei deren Berechnung darf man auf keinen Fall vergessen, die Einträge des jeweils zweiten Vektors komplex zu konjugieren, d.h.

$$\begin{aligned} \vec{b}_2 \cdot \vec{c}_1 &= (1+2i) \cdot \overline{1} + (-2+i) \cdot \overline{i} + (-i) \cdot \overline{(-i)} + (-1) \cdot \overline{(-1)} \\ &= (1+2i) \cdot 1 + (-2+i) \cdot (-i) + (-i) \cdot i + (-1) \cdot (-1) . \\ &= (1+2i) + (1+2i) + 1 + 1 = 4 + 4i \end{aligned}$$

Entsprechend ist

$$\begin{aligned} \vec{c}_1 \cdot \vec{c}_1 &= 1 \cdot \overline{1} + i \cdot \overline{i} + (-i) \cdot \overline{(-i)} + (-1) \cdot \overline{(-1)} \\ &= 1 \cdot 1 + i \cdot (-i) + (-i) \cdot i + (-1) \cdot (-1) = 1 + 1 + 1 + 1 = 4 , \end{aligned}$$

wobei das allerdings auch mit weniger Mühe einzusehen ist: Da für einen Vektor $\vec{v} \in \mathbb{C}^n$ mit Komponenten $v_1, \ldots, v_n$

$$\vec{v} \cdot \vec{v} = \sum_{i=1}^{n} v_i \cdot \overline{v_i} = \sum_{i=1}^{n} |v_i|^2$$

ist, hätte es genügt, einfach die Betragsquadrate der vier Einträge aufzusummieren, wobei im vorliegenden Fall trivial ist, daß diese allesamt eins sind.

Einsetzen der beiden Skalarprodukte in die Gleichung für λ_1 zeigt, daß

$$(4+4i) + 4\lambda_1 = 0 \quad \text{und damit} \quad \lambda_1 = \frac{-4-4i}{4} = -1 - i$$

sein muß. Somit ist

$$\vec{c}_2 = \vec{b}_2 + \lambda_1\vec{c}_1 = \begin{pmatrix} 1+2i \\ -2+i \\ -i \\ -1 \end{pmatrix} - \begin{pmatrix} 1+i \\ -1+i \\ 1-i \\ -1-i \end{pmatrix} = \begin{pmatrix} i \\ -1 \\ -1 \\ i \end{pmatrix} .$$

Im dritten Schritt muß ein Vektor

$$\vec{c}_3 = \vec{b}_3 + \mu_1\vec{c}_1 + \mu_2\vec{c}_2$$

berechnet werden mit

$$\vec{c}_3 \cdot \vec{c}_1 = \vec{c}_3 \cdot \vec{c}_2 = 0 .$$

Hier ist

$$\vec{c}_3 \cdot \vec{c}_1 = \vec{b}_3 \cdot \vec{c}_1 + \mu_1\vec{c}_1 \cdot \vec{c}_1 + \mu_2\vec{c}_2 \cdot \vec{c}_1 = \vec{b}_3 \cdot \vec{c}_1 + \mu_1\vec{c}_1 \cdot \vec{c}_1 ,$$

da der zweite Schritt sicherstellte, daß $\vec{c}_2 \cdot \vec{c}_1 = 0$ ist. Das HERMITEsche Skalarprodukt $\vec{c}_1 \cdot \vec{c}_1 = 4$ kennen wir bereits, und eine kurze Rechnung zeigt, daß

$$\vec{b}_3 \cdot \vec{c}_1 = (1+2i) + (-1) + (1-4i) + (-1-2i) = -4i$$

ist. Also folgt

$$\mu_1 = -\frac{\vec{b}_3 \cdot \vec{c}_1}{\vec{c}_1 \cdot \vec{c}_1} = -\frac{-4i}{4} = i$$

und entsprechend folgt auch

$$\mu_2 = -\frac{\vec{b}_3 \cdot \vec{c}_2}{\vec{c}_2 \cdot \vec{c}_2} = -\frac{8}{4} = -2 ,$$

wobei diesesmal die genaue Berechnung der beiden HERMITEschen Skalarprodukte dem Leser als Übungsaufgabe überlassen sei. Der dritte Basisvektor $\vec{c}_3 = \vec{b}_3 + i\vec{c}_1 - 2\vec{c}_2$ ist daher gleich

$$\begin{pmatrix} 1+2i \\ -i \\ -4-i \\ 1+2i \end{pmatrix} + i \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix} - 2 \begin{pmatrix} i \\ -1 \\ -1 \\ i \end{pmatrix} = \begin{pmatrix} 1+i \\ 1-i \\ -1-i \\ 1-i \end{pmatrix} .$$

Damit ist eine *Orthogonalbasis* gefunden; wenn wir eine *Orthonormalbasis* wollen, müssen wir die Vektoren noch durch ihre Längen dividieren: Wir wissen bereits, daß $\vec{c}_1 \cdot \vec{c}_1 = \vec{c}_2 \cdot \vec{c}_2 = 4$ ist, so daß die ersten beiden Vektoren die Länge zwei haben; $\vec{c}_3$ hat vier Komponenten vom Betrag zwei, also ist

$$\vec{c}_3 \cdot \vec{c}_3 = 4 \cdot 2 = 8 \quad \text{und} \quad \|\vec{c}_2\| = \sqrt{8} = 2\sqrt{2} .$$

Die gesuchte Orthonormalbasis besteht daher aus den Vektoren

$$\frac{1}{2} \begin{pmatrix} 1 \\ i \\ -i \\ 1 \end{pmatrix}, \quad \frac{1}{2} \begin{pmatrix} i \\ -1 \\ -1 \\ i \end{pmatrix} \quad \text{und} \quad \frac{\sqrt{2}}{4} \begin{pmatrix} 1+i \\ 1-i \\ -1-i \\ 1-i \end{pmatrix} .$$

Die Nützlichkeit von Orthonormal- und Orthogonalbasen ergibt sich vor allem aus den beiden folgenden Sätzen:

Satz: $\{\vec{e}_1, \dots, \vec{e}_n\}$ sei eine Orthonormalbasis eines EUKLIDischen oder HERMITEschen Vektorraums V.
a) Sind $\vec{v} = a_1\vec{e}_1 + \cdots + a_n\vec{e}_n$ und $\vec{w} = b_1\vec{e}_1 + \cdots + b_n\vec{e}_n$ zwei Vektoren aus V, so ist

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} a_i \overline{b_i} .$$

b) Für einen Vektor $\vec{v} \in V$ ist

$$\vec{v} = a_1\vec{e}_1 + \cdots + a_n\vec{e}_n \quad \text{mit} \quad a_i = \vec{v} \cdot \vec{e}_i .$$

Der *Beweis* ist in beiden Fällen ein Einzeiler:

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \overline{b_j}\, \vec{e}_i \cdot \vec{e}_j = \sum_{i=1}^{n} a_i \overline{b_i} , \quad \text{denn} \quad \vec{e}_i \cdot \vec{e}_j = \delta_{ij}$$

$$\vec{v} \cdot \vec{e}_i = (a_1\vec{e}_1 + \cdots + a_n\vec{e}_n) \cdot \vec{e}_i = \sum_{j=1}^{n} a_j \vec{e}_j \cdot \vec{e}_i = a_i .$$

■

Bezüglich einer Orthonormalbasis sieht also jedes (gewöhnliche oder HERMITEsche) Skalarprodukt aus wie das entsprechende Standardskalarprodukt – und das unabhängig von der gewählten Orthonormalbasis. Sobald man Vektoren bezüglich einer Orthonormalbasis dargestellt hat, lassen sich also alle Skalarprodukte sowie auch die Basisdarstellung eines beliebigen Vektors auf die einfachst denkbare Weise darstellen.

Da die Vektoren einer Orthonormalbasis oftmals Wurzelausdrücke als Komponenten enthalten, sind vor allem beim Rechnen ohne Hilfsmittel Orthogonalbasen oft übersichtlicher als Orthonormalbasen; wir wollen den Satz also der Vollständigkeit halber auch für Othogonalbasen formulieren:

Satz: $\{\vec{e}_1, \dots, \vec{e}_n\}$ sei eine Orthogonalbasis eines EUKLIDischen oder HERMITEschen Vektorraums V.
a) Sind $\vec{v} = a_1\vec{e}_1 + \cdots + a_n\vec{e}_n$ und $\vec{w} = b_1\vec{e}_1 + \cdots + b_n\vec{e}_n$ zwei Vektoren aus V, so ist

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} a_i \overline{b_i}\, |\vec{e}_i|^2 .$$

b) Für einen Vektor $\vec{v} \in V$ ist

$$\vec{v} = a_1\vec{e}_1 + \cdots + a_n\vec{e}_n \quad \text{mit} \quad a_i = \frac{(\vec{v} \cdot \vec{e}_i)}{|\vec{e}_i|^2} .$$

Die *Beweise* sind fast dieselben Einzeiler:

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \overline{b_j}\, \vec{e}_i \cdot \vec{e}_j = \sum_{i=1}^{n} a_i \overline{b_i}\, |ve_i|^2 , \quad \text{denn} \quad \vec{e}_i \cdot \vec{e}_j = \delta_{ij}\, |ve_i|^2$$

$$\vec{v} \cdot \vec{e}_i = (a_1\vec{e}_1 + \cdots + a_n\vec{e}_n) \cdot \vec{e}_i = \sum_{j=1}^{n} a_j \vec{e}_j \cdot \vec{e}_i = a_i\, |ve_i|^2 .$$

■

f) Die QR-Zerlegung einer Matrix

Wir betrachten den Vektorraum $V = \mathbb{R}^n$ oder $V = \mathbb{C}^n$ mit seiner Standardbasis und seiner üblichen Struktur als EUKLIDischer *bzw.* HERMITEscher Vektorraum zusammen mit einer Basis $\{\vec{b}_1, \dots, \vec{b}_n\}$. Nach dem GRAM-SCHMIDTschen Orthogonalisierungsverfahren können wir daraus eine Orthogonalbasis $\{\vec{c}_1, \dots, \vec{c}_n\}$ konstruieren, wobei gilt

$$\vec{c}_k = \vec{b}_k + \lambda_{k,k-1}\vec{b}_{k-1} + \cdots + \lambda_{k1}\vec{c}_1 \quad \text{mit} \quad \lambda_{kj} = -\frac{\vec{b}_k \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j}\,.$$

Lösen wir auf nach $\vec{b}_k$, erhalten wir

$$\vec{b}_k = \vec{c}_k - \lambda_{k,k-1}\vec{b}_{k-1} - \cdots - \lambda_{k1}\vec{c}_1 = \sum_{j=1}^{n} r_{jk}\vec{c}_j$$

$$\text{mit } r_{jk} = \begin{cases} -\lambda_{kj} = \dfrac{\vec{b}_k \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j} & \text{für } j < k \\ 1 & \text{für } j = k \\ 0 & \text{für } j > k \end{cases}\,.$$

Fassen wir die Vektoren $\vec{b}_1, \dots \vec{b}_n$ auf als Spaltenvektoren einer Matrix B und die $\vec{c}_j$ als Spalten von C, so gilt also für die i-ten Komponenten

$$b_{ik} = \sum_{j=1}^{n} r_{jk}c_{ij} = \sum_{j=1}^{n} c_{ij}r_{jk} \quad \text{oder} \quad B = CR \quad \text{mit} \quad R = (r_{ij})\,.$$

Da r_{ij} für $i > j$ verschwindet und alle $\lambda_{ii} = 1$ sind, ist R eine obere Dreiecksmatrix mit Einsen in der Hauptdiagonalen.

Ist B eine beliebige invertierbare $n \times n$-Matrix, so bilden die Spaltenvektoren $\vec{b}_i$ von B eine Basis des $\mathbb{R}^n$ *bzw.* $\mathbb{C}^n$; wie wir gerade gesehen haben, gibt es also eine obere Dreiecksmatrix R und eine Matrix C, deren Spalten eine Orthogonalbasis bilden, so daß $B = CR$ ist.

Normieren wir die Spalten von C auf Länge eins, erhalten wir eine $n \times n$-Matrix Q, deren Spalten eine Orthonormalbasis bilden; um die Beziehung $B = QR$ zu erhalten, müssen wir dann allerdings auch noch für jedes $i = 1, \dots, n$ die i-te Zeile von R mit der Länge des i-ten Spaltenvektors $\vec{c}_i$ von C multiplizieren. Dabei bleibt R zwar eine obere

Dreiecksmatrix, in der Hauptdiagonalen stehen nun aber im allgemeinen keine Einsen mehr.

Was passiert, wenn B nicht invertierbar ist? Selbst wenn wir für B eine beliebige $n \times m$-Matrix nehmen, die Anzahl m der Spaltenvektoren $\vec{b}_i$ also nicht mehr gleich der Anzahl der Zeilen ist, lassen sich die Rechenschritte des GRAM-SCHMIDTschen Orthogonalisierungsverfahrens weiterhin durchführen – mit einem wesentlichen Unterschied:

Betrachten wir eine Folge $\vec{b}_1, \dots, \vec{b}_m$ von Vektoren aus $\mathbb{R}^n$ *bzw.* $\mathbb{C}^n$, versehen mit dem EUKLIDischen oder HERMITEschen Standardskalarprodukt. Wir konstruieren wieder im k-ten Schritt zunächst eine Orthogonalbasis des von $\vec{b}_1$ bis $\vec{b}_k$ aufgespannten Untervektorraums.

Schon im ersten Schritt kann es eine Änderung geben: Der Vektor $\vec{b}_1$ könnte der Nullvektor sein und damit definitiv nicht als erster Basisvektor in Frage kommen.

Wir können also im ersten Schritt nicht einfach den Vektor $\vec{b}_1$ als ersten Vektor der zu konstruierenden Orthogonalbasis nehmen, sondern wir müssen den ersten Vektor $\vec{b}_\ell$ nehmen, der vom Nullvektor verschieden ist.

Die eventuell davor stehenden Nullvektoren können wir allerdings nicht ganz vergessen, denn sie sind schließlich Spalten der Matrix und müssen am Ende auch im Produkt CR wieder auftauchen. Wir müssen deshalb in der Matrix R die ersten $\ell - 1$ Zeilen auf Null setzen.

Danach wird $\vec{c}_1 = \vec{b}_\ell$ gesetzt; gegebenenfalls kann man den Vektor auch gleich noch auf Länge eins normieren: Zumindest wenn man von Hand rechnet, wird das davon abhängen, wie kompliziert die Länge ist. Die ℓ-te Zeile von R beginnt mit einer Eins (oder der Länge von $\vec{b}_\ell$, falls $\vec{c}_1$ auf Länge Eins normiert wurde) und enthält ansonsten lauter Nullen.

Im $(k + 1)$-ten Schritt, für $k \geq 1$, gehen wir davon aus, daß wir eine Orthogonalbasis $\vec{c}_1, \dots, \vec{c}_k$ des von $\vec{b}_1$ bis $\vec{b}_\ell$ aufgespannten Untervektorraums gefunden haben für ein $\ell \leq m$. Falls $\ell < m$ ist, machen wir den Ansatz

$$\vec{c}_{k+1} = \vec{b}_{\ell+1} - r_{\ell+1,1}\vec{c}_1 - \cdots - r_{\ell+1,k}\vec{c}_k$$

und bestimmen die Koeffizienten $\vec{r}_{ij}$ so, daß $\vec{c}_{k+1}\vec{c}_j = 0$ ist für alle $j \le k$, d.h.

$$r_{\ell+1,j} = \frac{\vec{b}_{\ell+1} \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j} .$$

Falls $\vec{b}_{\ell+1}$ linear unabhängig ist von $\vec{b}_1, \dots, \vec{b}_\ell$, ist dann

$$\vec{c}_{k+1} = \vec{b}_{\ell+1} - r_{\ell+1,1}\vec{c}_1 - \cdots - r_{\ell+1,k}\vec{c}_k$$

linear unabhängig von $\vec{c}_1, \dots, \vec{c}_k$, denn diese Vektoren erzeugen denselben Untervektorraum wie $\vec{b}_1, \dots, \vec{b}_\ell$.

Wenn allerdings $\vec{b}_{\ell+1}$ linear abhängig ist von $\vec{b}_1, \dots, \vec{b}_\ell$, ist $\vec{b}_{\ell+1}$ auch linear abhängig von $\vec{c}_1, \dots, \vec{c}_k$; da diese Vektoren eine Orthogonalbasis des von ihnen erzeugten Vektorraums sind, ist daher

$$\vec{b}_{\ell+1} = \lambda_1\vec{c}_1 + \cdots + \lambda_k\vec{c}_k \quad \text{mit} \quad \lambda_j = \frac{\vec{b}_{\ell+1} \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j} .$$

Vergleicht man dies mit der obigen Formel für $r_{\ell+1,j}$, so sieht man, daß in diesem Fall $\vec{c}_{k+1}$ der Nullvektor ist, wir bekommen also kein neues Element der Orthogonalbasis.

In diesem Fall müssen wir daher mit dem nächsten Vektor $\vec{b}_{\ell+2}$ – so es einen gibt – einen neuen Ansatz

$$\vec{c}_{k+1} = \vec{b}_{\ell+2} - r_{\ell+1,1}\vec{c}_1 - \cdots - r_{\ell+2,k}\vec{c}_k$$

machen, und wieder die Koeffizienten $\vec{r}_{ij}$ so bestimmen, daß $\vec{c}_{k+1}\vec{c}_j = 0$ ist für alle $j \le k$, d.h.

$$r_{\ell+2,j} = \frac{\vec{b}_{\ell+2} \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j} ,$$

wobei der entstehende Vektor $\vec{c}_{k+1}$ wieder der Nullvektor sein kann und so weiter, bis wir entweder einen vom Nullvektor verschiedenen Vektor $\vec{c}_{k+1}$ erhalten oder aber kein neuer Vektor $\vec{b}_j$ mehr existiert.

Im ersteren Fall machen wir weiter mit dem $(k+2)$-ten Schritt, andernfalls haben wir eine Basis $\vec{c}_1, \dots, \vec{c}_k$ gefunden für den von $\vec{b}_1, \dots, \vec{b}_m$ aufgespannten Untervektorraum von $\mathbb{R}^n$ *bzw.* $\mathbb{C}^n$.

Das muß allerdings noch keine Basis des gesamten $\mathbb{R}^n$ *bzw.* $\mathbb{C}^n$ sein, denn offensichtlich ist die Anzahl der Vektoren $\vec{c}_k$, die wir so erhalten, gleich dem Rang von A, und der könnte auch kleiner als n sein – für $m < n$ muß das sogar so sein.

Wenn wir eine quadratische Matrix C suchen, müssen wir daher gegebenenfalls noch die Vektoren $\vec{c}_1, \dots, \vec{c}_k$ zu einer vollen Orthogonalbasis ergänzen; dazu können wir beispielsweise so lange GRAM-SCHMIDT auf Vektoren aus der Standardbasis anwenden, bis wir n orthogonale Vektoren $\vec{c}_i$ gefunden haben.

Falls alle Vektoren $\vec{c}_i$ gleich auf Länge eins normiert wurden, bilden sie sogar eine Orthonormalbasis. Da eine solche Normierung allerdings oft Nenner mit Wurzeln produziert, ist es oft rechnerisch angenehmer, die $\vec{c}_i$ unnormiert zu lassen. In diesem Fall müssen sie zum Schluß auf die Länge eins gebracht werden; wir setzen dazu

$$\vec{q}_i = \frac{1}{|\vec{c}_i|}\vec{c}_i$$

und definieren Q als die $n \times n$-Matrix mit Spaltenvektoren $\vec{q}_1$ bis $\vec{q}_n$.

Da die Einträge der Matrix R bislang so berechnet waren, daß $CR = A$ ist, müssen wir auch diese noch modifizieren: Um von C auf Q zu kommen, haben wir die i-te Spalte durch $|\vec{c}_i|$ dividiert; zur Kompensation müssen wir daher in R die i-te *Zeile* mit $|\vec{c}_i|$ *multiplizieren;* dann haben wir mit der so modifizierten Matrix R die Produktzerlegung $A = QR$, die sogenannte QR-Zerlegung von A.

Da die Matrix R hier nicht mehr quadratisch ist, können wir sie nicht als Dreiecksmatrix bezeichnen; trotzdem teilt sie mit jeder oberen Dreiecksmatrix die Eigenschaft, daß r_{ij} verschwindet für $i > j$.

Definition: Eine $n \times m$-Matrix $R = (r_{ij})$ heißt (obere) Trapezmatrix oder einfach trapezförmig, wenn r_{ij} verschwindet wann immer $i > j$ ist.

Insgesamt haben wir also gezeigt:

Satz: Zu jeder reellen oder komplexen $n \times m$-Matrix A gibt es eine $n \times n$-Matrix Q, deren Spalten eine Orthonormalbasis von $\mathbb{R}^n$ *bzw.* $\mathbb{C}^n$

bilden, sowie eine trapezförmige $n \times m$-Matrix R mit $r_{ij} = 0$, so daß $A = QR$ ist.
Hat A einen Rang $k < n$, so enthalten die letzten $n - k$ Zeilen von R nur Nullen; daher ist auch $A = Q'R'$, wobei Q' die $n \times k$-Matrix aus den ersten k Spalten von Q und R' die $k \times m$-Matrix aus den ersten k Zeilen von R ist. ■

Als Beispiel betrachten wir die Matrix

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

und bezeichnen ihre Spaltenvektoren mit $\vec{a}_1, \dots, \vec{a}_5$.

$\vec{a}_1$ hat bereits die Länge eins, kann also als erster Vektor $\vec{q}_1$ der Orthonormalbasis genommen werden, so daß der erste Spaltenvektors $\vec{r}_1$ der Matrix R einfach der erste Koordinateneinheitsvektor ist.

Der zweite Spaltenvektor $\vec{a}_2$ von A hat mit $\vec{q}_1$ das Skalarprodukt eins; da $\vec{q}_1$ selbst bereits Länge eins hat, setzen wir also

$$\vec{q}_2 = \vec{a}_2 - \vec{q}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{r}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

denn $\vec{a}_2 = \vec{q}_1 + \vec{q}_2$.

Im dritten Schritt suchen wir zunächst einen Vektor der Form

$$\vec{c}_3 = \vec{a}_3 + \lambda\vec{q}_1 + \mu\vec{q}_3,$$

der auf $\vec{q}_1$ und $\vec{q}_2$ senkrecht steht. Da $\vec{q}_1$ und $\vec{q}_2$ die Länge eins haben, ist hier einfach $\lambda = -\vec{a}_3 \cdot \vec{q}_1 = -3$ und $\mu = -\vec{a}_3 \cdot \vec{q}_2 = -2$, der gesuchte Vektor ist also der dreifache vierte Einheitsvektor, und als zugehörigen Vektor $\vec{q}_3$ der Länge eins nehmen wir den vierten Einheitsvektor selbst. Dann ist

$$\vec{a}_3 = 3\vec{q}_1 + 2\vec{q}_2 + 3\vec{q}_3, \quad \text{also} \quad \vec{r}_3 = \begin{pmatrix} 3 \\ 2 \\ 3 \\ 0 \end{pmatrix}.$$

Beim nächsten Spaltenvektor $\vec{a}_4$ sieht man eigentlich bereits ohne Rechnung, daß er im Erzeugnis von $\vec{q}_1$ bis $\vec{q}_3$ liegt, denn offensichtlich ist $\vec{a}_4 = \vec{a}_3 + \vec{q}_2 + \vec{q}_3$. Wenn wir trotzdem stur nach Schema F losrechnen mit dem Ansatz

$$\vec{c}_4 = \vec{a}_4 + \lambda_1\vec{q}_1 + \lambda_2\vec{q}_2 + \lambda_3\vec{q}_3 \quad \text{mit} \quad \lambda_i = -\vec{a}_4 \cdot \vec{q}_i,$$

erhalten wir erwartungsgemäß $\vec{c}_4 = \vec{0}$, also keinen neuen Vektor der Orthogonalbasis. Wir bekommen aber, da wir auch $\vec{a}_4$ als Linearkombination der $\vec{q}_i$ darstellen müssen, eine neue Spalte $\vec{r}_4$ der Matrix R; wegen

$$\vec{a}_4 = 3\vec{q}_1 + 3\vec{q}_2 + 4\vec{q}_3 \quad \text{ist} \quad \vec{r}_4 = \begin{pmatrix} 3 \\ 3 \\ 4 \\ 0 \end{pmatrix}.$$

In der letzten Spalte von A stehen lauter Nullen; also brauchen wir auch hier keinen neuen Basisvektor und sehen auch ohne jede Rechnung, daß $\vec{r}_5$ gleich dem Nullvektor ist.

Damit kennen wir die Matrix R vollständig, allerdings fehlt für eine vollständige QR-Zerlegung noch eine Spalte von Q. Diese kann offensichtlich völlig unabhängig von A gewählt werden als irgendein Vektor, der auf $\vec{q}_1$ bis $\vec{q}_3$ senkrecht steht. Die kanonische Wahl ist natürlich der dritte Einheitsvektor; die einzige Alternative dazu wäre sein negatives.

Damit ist $A = QR$ mit

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

Natürlich ist auch

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 & 3 & 4 & 0 \\ 0 & 1 & 2 & 3 & 4 & 0 \\ 0 & 0 & 3 & 4 & 5 & 0 \end{pmatrix},$$

denn die vierte Spalte von Q wird ja bei der Multiplikation mit R stets mit Null gewichtet.

g) Orthogonale und unitäre Matrizen

Bei der Lektüre des letzten Abschnitts hat sich sicherlich mancher mehrfach die Frage gestellt, warum wir selbst bei kleinem Rang von A an einer Zerlegung interessiert sind, bei der Q eine quadratische Matrix ist, deren Spalten eine Orthonormalbasis des gesamten $\mathbb{R}^n$ *bzw.* $\mathbb{C}^n$ bilden. Der Grund liegt darin, daß solche Matrizen Q eine ganze Reihe interessanter Eigenschaften haben, die zur Lösung mehrerer wichtiger Probleme verwendet werden können. Dies wollen wir uns im folgenden etwas genauer ansehen.

Betrachten wir zunächst den reellen Fall.

Sind $\vec{q}_i$ die Spaltenvektoren von Q, so ist die Tatsache, daß die $\vec{q}_i$ eine Orthonormalbasis bilden, äquivalent dazu, daß $\vec{q}_i \cdot \vec{q}_j = \delta_{ij}$ für alle i, j. Da wir $\vec{q}_j$ nicht nur als j-te Spalte von Q, sondern auch als j-te Zeilte der transponierten Matrix ${}^tQ$ auffassen können, können wir diese n^2 Gleichungen zusammenfassen zu der einen Matrixgleichung ${}^tQQ = E$.

Auch im HERMITEschen Fall ist $\vec{q}_i \cdot \vec{q}_j = \delta_{ij}$ für alle i, j, aber nun ist das HERMITEsche Skalarprodukt nicht mehr durch dieselbe Formel definiert, die auch bei der Matrixmultiplikation Anwendung findet, sondern enthält noch eine komplexe Konjugation im zweiten Faktor. Daher sind die Gleichungen $\vec{q}_i \cdot \vec{q}_j = \delta_{ij}$ hier äquivalent dazu, daß ${}^tQ\overline{Q}$ die Einheitsmatrix ist. Transponieren macht dies zu ${}^t({}^tQ\overline{Q}) = {}^t\overline{Q}Q = {}^tE = E$.

Die Schreibweise ${}^t\overline{Q}$ ist etwas umständlich; deshalb führen wir eine Abkürzung ein:

Definition: Die *adjungierte Matrix* zu einer Matrix $A \in \mathbb{C}^{n\times m}$ ist die Matrix $A^* = {}^t\overline{A} = \overline{{}^tA}$.

Damit ist also beispielsweise

$$A^* = \begin{pmatrix} 1-i & 2-i & 3-i \\ 4-i & 5-i & 6-i \end{pmatrix} \quad \text{für} \quad A = \begin{pmatrix} 1+i & 4+i \\ 2+i & 5+i \\ 3+i & 6+i \end{pmatrix}.$$

Für eine reelle Matrix A ist natürlich $A^* = {}^tA$ einfach die transponierte Matrix.

Definition: *a)* Eine Matrix $Q \in \mathbb{R}^{n\times n}$ heißt *orthogonal*, wenn ${}^tQQ = E$ ist.
b) Eine Matrix $U \in \mathbb{C}^{n\times n}$ heißt *unitär*, wenn $U^*U = E$ ist.

Die Matrix Q aus der QR-Zerlegung einer reellen Matrix ist somit orthogonal; im Falle einer komplexen Matrix ist Q unitär.

Lemma: *a)* Eine Matrix $Q \in \mathbb{R}^{n\times n}$ ist genau dann orthogonal, wenn $Q^{-1} = {}^tQ$ ist. Insbesondere ist jede orthogonale Matrix invertierbar.
b) Eine Matrix $U \in \mathbb{C}^{n\times n}$ ist genau dann unitär, wenn $U^{-1} = U^*$ ist; insbesondere ist jede unitäre Matrix invertierbar.
c) Die inverse Matrix einer orthogonalen *bzw.* unitären Matrix ist wieder orthogonal *bzw.* unitär.
d) Das Produkt zweier orthogonaler *bzw.* unitärer Matrizen ist wieder orthogonal *bzw.* unitär.

Beweis: Da Unitarität und Orthogonalität für reelle Matrizen äquivalent sind, können wir uns auf unitäre Matrizen beschränken.

a) und *b)* sind klar, denn nach Definition einer unitären Matrix U ist U^* invers zu U. Für *c)* muß somit gezeigt werden, daß U^* wieder unitär ist. Nach Definition ist U invers zu dieser Matrix, wegen $(U^*)^* = U$ folgt die Unitarität.

Für *d)* schließlich seien U_1 und U_2 zwei unitäre Matrizen; dann ist

$$(U_1U_2)^* = U_2^*U_1^* = U_2^{-1}U_1^{-1} = (U_1U_2)^{-1},$$

und damit ist auch U_1U_2 unitär. ■

Insbesondere die Aussagen *a)* und *b)* zeigen einen großen praktischen Vorteil orthogonaler und unitärer Matrizen: Man kann sie mit minimalem Aufwand invertieren.

Ist etwa $A\vec{x} = \vec{b}$ ein lineares Gleichungssystem mit einer $n \times m$-Koeffizientenmatrix A, und ist $A = QR$ die QR-Zerlegung von A, so ist

$$A\vec{x} = \vec{b} \iff QR\vec{x} = \vec{b} \iff R\vec{x} = Q^{-1}\vec{b} = Q^*\vec{b}.$$