Kapitel 3

Vektoren und Matrizen

In diesem Kapitel sollen die Grundlagen aus der linearen Algebra bereitgestellt werden, die später bei der Konstruktion effizienter Algorithmen für die Lösung linearer Gleichungssysteme benötigt werden. Neben den grundlegenden Begriffen wie zum Beispiel Normen von Vektoren und Matrizen wird die Singulärwertzerlegung beliebiger Matrizen hergeleitet. Das Orthogonalisierungsverfahren nach Gram–Schmidt bildet den Ausgangspunkt für die Herleitung von modernen Iterationsverfahren für lineare Gleichungssysteme.

3.1 Normen von Vektoren und Matrizen

Für $n \in \mathbb{N}$ ist $\mathbb{R}^n$ der Raum der n–dimensionalen Vektoren $\underline{u} \in \mathbb{R}^n$ mit Komponenten $u_i \in \mathbb{R}$ für $i = 1, \ldots, n$. Mit

$$\langle \cdot, \cdot \rangle \; : \; \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

wird ein beliebiges Skalarprodukt im Vektorraum $\mathbb{R}^n$ bezeichnet, das heißt es gilt die Distributivität

$$\langle \underline{u} + \underline{v}, \underline{w} \rangle \; = \; \langle \underline{u}, \underline{w} \rangle + \langle \underline{v}, \underline{w} \rangle,$$

die Kommutativität

$$\langle \underline{u}, \underline{v} \rangle \; = \; \langle \underline{v}, \underline{u} \rangle,$$

die Homogenität

$$\langle \alpha \underline{u}, \underline{v} \rangle \; = \; \alpha \, \langle \underline{u}, \underline{v} \rangle$$

für alle Vektoren $\underline{u}, \underline{v}, \underline{w} \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$ sowie die positive Definitheit

$$\langle \underline{u}, \underline{u} \rangle \; > \; 0$$

für alle $\underline{u} \in \mathbb{R}^n$ mit $\underline{u} \neq \underline{0}$. Insbesondere definiert

$$\langle \underline{u}, \underline{v} \rangle_2 \; := \; (\underline{u}, \underline{v}) \; = \; \sum_{i=1}^{n} u_i v_i$$

das Euklidische Skalarprodukt. Für einen Vektor $\underline{u} \in \mathbb{R}^n$ bezeichnet

$$\| \cdot \|_V \,:\, \mathbb{R}^n \to \mathbb{R}$$

eine beliebige Vektornorm, für welche die Normaxiome gelten, das heißt die positive Definitheit

$$\|\underline{u}\|_V \,\geq\, 0 \quad \text{für alle } \underline{u} \in \mathbb{R}^n, \quad \|\underline{u}\|_V \,=0 \quad \text{genau dann, wenn } \underline{u} = \underline{0},$$

die Homogenität

$$\|\alpha\, \underline{u}\|_V \,=\, |\alpha|\, \|\underline{u}\|_V \quad \text{für alle } \underline{u} \in \mathbb{R}^n \text{ und } \alpha \in \mathbb{R},$$

sowie die Dreiecksungleichung

$$\|\underline{u} + \underline{v}\|_V \,\leq\, \|\underline{u}\|_V + \|\underline{v}\|_V \quad \text{für alle } \underline{u}, \underline{v} \in \mathbb{R}^n.$$

Beispiele für Vektornormen sind die Euklidische Norm

$$\|\underline{u}\|_2 \,:=\, \left(\sum_{i=1}^n u_i^2 \right)^{1/2},$$

die Maximumnorm

$$\|\underline{u}\|_\infty \,:=\, \max_{i=1,\ldots,n} |u_i|\,,$$

sowie die Summennorm

$$\|\underline{u}\|_1 \,:=\, \sum_{i=1}^n |u_i|\,.$$

Nach Definition ist

$$\|\underline{u}\|_2^2 \,=\, \sum_{i=1}^n u_i^2 \,=\, (\underline{u}, \underline{u}) \quad \text{für alle } \underline{u} \in \mathbb{R}^n$$

und es gilt die Cauchy–Schwarz–Ungleichung

$$(\underline{u}, \underline{v}) \,=\, \sum_{i=1}^n u_i v_i \,\leq\, \left(\sum_{i=1}^n u_i^2 \right)^{1/2} \left(\sum_{i=1}^n v_i^2 \right)^{1/2} \,=\, \|\underline{u}\|_2 \|\underline{v}\|_2 \tag{3.1}$$

für alle $\underline{u}, \underline{v} \in \mathbb{R}^n$.

Zwei Vektornormen $\| \cdot \|_{V_1}$ und $\| \cdot \|_{V_2}$ heißen zueinander äquivalent, wenn unabhängig von $\underline{u} \in \mathbb{R}^n$ zwei positive Konstanten c_1 und c_2 existieren, so daß die Äquivalenzungleichungen

$$c_1\, \|\underline{u}\|_{V_1} \,\leq\, \|\underline{u}\|_{V_2} \,\leq\, c_2\, \|\underline{u}\|_{V_1} \quad \text{für alle } \underline{u} \in \mathbb{R}^n \tag{3.2}$$

erfüllt sind. Die Äquivalenzungleichungen sind scharf, wenn im allgemeinen unterschiedliche Vektoren $\underline{u} \in \mathbb{R}^n$ existieren, für die in (3.2) jeweils die Gleichheit gilt.

Lemma 3.1. *Für beliebiges $\underline{u} \in \mathbb{R}^n$ gelten die Äquivalenzungleichungen*

$$\begin{aligned} \|\underline{u}\|_\infty &\leq \|\underline{u}\|_1 \leq n\,\|\underline{u}\|_\infty, \\ \|\underline{u}\|_\infty &\leq \|\underline{u}\|_2 \leq \sqrt{n}\,\|\underline{u}\|_\infty, \\ \|\underline{u}\|_2 &\leq \|\underline{u}\|_1 \leq \sqrt{n}\,\|\underline{u}\|_2. \end{aligned}$$

Alle Abschätzungen sind scharf.

Beweis: Zunächst ist

$$\|\underline{u}\|_\infty = \max_{i=1,\ldots,n} |u_i| \leq \sum_{i=1}^n |u_i| = \|\underline{u}\|_1,$$

wobei die Gleichheit zum Beispiel für $\underline{u} = (1, 0, \ldots, 0)^\top$ angenommen wird. Für die Abschätzung in umgekehrter Richtung folgt

$$\|\underline{u}\|_1 = \sum_{i=1}^n |u_i| \leq n \max_{i=1,\ldots,n} |u_i| = n\,\|\underline{u}\|_\infty.$$

Diese ist scharf zum Beispiel für $\underline{u} = (1, \ldots, 1)^\top$.
Die Äquivalenz zwischen Maximumnorm $\|\cdot\|_\infty$ und Euklidischer Norm $\|\cdot\|_2$ folgt analog, das heißt

$$\|\underline{u}\|_\infty^2 = \left(\max_{i=1,\ldots,n} |u_i|\right)^2 = \max_{i=1,\ldots,n} |u_i|^2 \leq \sum_{i=1}^n |u_i|^2 = \|\underline{u}\|_2^2$$

sowie

$$\|\underline{u}\|_2^2 = \sum_{i=1}^n |u_i|^2 \leq n \max_{i=1,\ldots,n} |u_i|^2 = n\,\|\underline{u}\|_\infty^2 .$$

Gleichheit gilt beispielsweise für $\underline{u} = (1, 0, \ldots, 0)^\top$ sowie für $\underline{u} = (1, \ldots, 1)^\top$.
Die Kombination der bereits gezeigten Ungleichungen ergibt für die Äquivalenz der Euklidischen Norm $\|\cdot\|_2$ zur Summennorm $\|\cdot\|_1$ die Ungleichungen

$$\frac{1}{n}\,\|\underline{u}\|_1 \leq \|\underline{u}\|_2 \leq \sqrt{n}\,\|\underline{u}\|_1. \tag{3.3}$$

Da aber in den einzelnen Äquivalenzungleichungen die Gleichheit jeweils für unterschiedliche Vektoren $\underline{u} \in \mathbb{R}^n$ angenommen wird, sind die resultierenden Äquivalenzungleichungen (3.3) nicht scharf und daher nicht optimal.
Mit der Cauchy–Schwarz–Ungleichung (3.1) folgt

$$\|\underline{u}\|_1 = \sum_{i=1}^n |u_i| = \sum_{i=1}^n (1 \cdot |u_i|) \leq \left(\sum_{i=1}^n 1^2\right)^{1/2} \left(\sum_{i=1}^n |u_i|^2\right)^{1/2} = \sqrt{n}\,\|\underline{u}\|_2 .$$

Diese ist scharf für $\underline{u} = (1, \ldots, 1)^\top$. Andererseits ist

$$\|\underline{u}\|_2^2 = \sum_{i=1}^n |u_i|^2 \le \left(\sum_{i=1}^n |u_i| \right)^2 = \|\underline{u}\|_1^2$$

mit Gleichheit für $\underline{u} = (1, 0, \ldots, 0)^\top$. ■

Sei $B \in \mathbb{R}^{m \times n}$ eine beliebig gegebene Matrix mit Einträgen $B[k, \ell] = b_{k\ell} \in \mathbb{R}$ für $k = 1, \ldots, m$ und $\ell = 1, \ldots, n$. Mit

$$\| \cdot \|_M : \mathbb{R}^{m \times n} \to \mathbb{R}$$

wird eine beliebige Matrixnorm bezeichnet. Beispiele für Matrixnormen sind die Zeilensummennorm

$$\|B\|_\infty := \max_{k=1,\ldots,m} \sum_{\ell=1}^n |b_{k\ell}|,$$

die Spaltensummennorm

$$\|B\|_1 := \max_{\ell=1,\ldots,n} \sum_{k=1}^m |b_{k\ell}|$$

sowie die Frobenius–Norm (Hilbert–Schmidt–Norm)

$$\|B\|_F := \left(\sum_{k=1}^m \sum_{\ell=1}^n b_{k\ell}^2 \right)^{1/2}.$$

Für eine sowohl in $\mathbb{R}^n$ als auch in $\mathbb{R}^m$ gegebene Vektornorm $\| \cdot \|_V$ kann durch

$$\|B\|_M := \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n, B\underline{x} \in \mathbb{R}^m} \frac{\|B\underline{x}\|_V}{\|\underline{x}\|_V}$$

stets eine induzierte Matrixnorm definiert werden. Insbesondere induziert die Euklidische Vektornorm die Euklidische Matrixnorm

$$\|B\|_2 := \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n} \frac{||B\underline{x}||_2}{||\underline{x}||_2}.$$

Lemma 3.2. *Die Zeilensummennorm $\|B\|_\infty$ wird durch die Maximumnorm $\|\underline{x}\|_\infty$ induziert.*

Beweis: Für die Maximumnorm von $B\underline{x} \in \mathbb{R}^m$ für einen beliebigen Vektor $\underline{x} \in \mathbb{R}^n$ ergibt sich

$$\|B\underline{x}\|_\infty = \max_{k=1,\ldots,m} \left| \sum_{\ell=1}^n b_{k\ell} x_\ell \right| \le \|\underline{x}\|_\infty \max_{k=1,\ldots,m} \sum_{\ell=1}^n |b_{k\ell}|.$$

Für alle $\underline{x} \in \mathbb{R}^n$ mit $\|\underline{x}\|_\infty \neq 0$ ist somit

$$\frac{\|B\underline{x}\|_\infty}{\|\underline{x}\|_\infty} \leq \max_{k=1,\dots,m} \sum_{\ell=1}^{n} |b_{k\ell}| = \|B\|_\infty ,$$

woraus

$$\sup_{0 \neq \underline{x} \in \mathbb{R}^n} \frac{\|B\underline{x}\|_\infty}{\|\underline{x}\|_\infty} \leq \|B\|_\infty$$

folgt. Für den Nachweis der umgekehrten Ungleichung bezeichne k_0 den Index, für welchen die Zeilensummennorm angenommen wird, das heißt

$$\|B\|_\infty = \max_{k=1,\dots,m} \sum_{\ell=1}^{n} |b_{k\ell}| = \sum_{\ell=1}^{n} |b_{k_0\ell}| .$$

Sei $\underline{\widetilde{x}} \in \mathbb{R}^n$ definiert durch

$$\widetilde{x}_\ell = \begin{cases} \dfrac{b_{k_0\ell}}{|b_{k_0\ell}|} & \text{für } b_{k_0\ell} \neq 0, \\ 1 & \text{für } b_{k_0\ell} = 0 \end{cases}$$

und $\ell = 1, \dots, n$. Nach Konstruktion ist $\|\underline{\widetilde{x}}\|_\infty = 1$. Dann ergibt sich

$$\|B\underline{\widetilde{x}}\|_\infty = \max_{k=1,\dots,m} \Big|\sum_{\ell=1}^{n} b_{k\ell}\widetilde{x}_\ell\Big| \geq \Big|\sum_{\ell=1}^{n} b_{k_0\ell}\widetilde{x}_\ell\Big| = \sum_{\ell=1}^{n} |b_{k_0\ell}| = \|B\|_\infty ,$$

und wegen $\|\underline{\widetilde{x}}\|_\infty = 1$ folgt

$$\|B\|_\infty \leq \frac{\|B\underline{\widetilde{x}}\|_\infty}{\|\underline{\widetilde{x}}\|_\infty} \leq \sup_{0 \neq \underline{x} \in \mathbb{R}^n} \frac{\|B\underline{x}\|_\infty}{\|\underline{x}\|_\infty} \leq \|B\|_\infty$$

und somit die Gleichheit. ∎

Lemma 3.3. *Die Spaltensummennorm $\|B\|_1$ wird durch die Summennorm $\|\underline{x}\|_1$ induziert.*

Beweis: Für die Summennorm von $B\underline{x} \in \mathbb{R}^m$ ergibt sich

$$\begin{aligned} \|B\underline{x}\|_1 &= \sum_{k=1}^{m} \left| \sum_{\ell=1}^{n} b_{k\ell} x_\ell \right| \leq \sum_{k=1}^{m} \sum_{\ell=1}^{n} |b_{k\ell}| \, |x_\ell| \\ &\leq \left(\max_{\ell=1,\dots,n} \sum_{k=1}^{m} |b_{k\ell}| \right) \sum_{\ell=1}^{n} |x_\ell| = \|B\|_1 \, \|\underline{x}\|_1 \end{aligned}$$

für alle $\underline{x} \in \mathbb{R}^n$, und für $\|\underline{x}\|_1 \neq 0$ folgt

$$\sup_{0 \neq \underline{x} \in \mathbb{R}^n} \frac{\|B\underline{x}\|_1}{\|\underline{x}\|_1} \leq \|B\|_1 .$$

Sei nun ℓ_0 der Index, für den die Spaltensummennorm angenommen wird,

$$\|B\|_1 \;=\; \max_{\ell=1,\dots,n} \sum_{k=1}^{m} |b_{k\ell}| \;=\; \sum_{k=1}^{m} |b_{k\ell_0}|\,,$$

und sei $\underline{\widetilde{x}} = (\delta_{1\ell_0}, \dots, \delta_{n\ell_0})^\top$ mit $\|\underline{\widetilde{x}}\|_1 = 1$. Hierbei bezeichnet

$$\delta_{k\ell} \;=\; \begin{cases} 1 & \text{für } k = \ell, \\ 0 & \text{für } k \neq \ell \end{cases}$$

das Kroneckersymbol. Dann folgt

$$\|B\|_1 \;=\; \sum_{k=1}^{m} |b_{k\ell_0}| \;=\; \sum_{k=1}^{m} \left| \sum_{\ell=1}^{n} b_{k\ell} \widetilde{x}_\ell \right| \;=\; \|B\underline{\widetilde{x}}\|_1 \;=\; \frac{\|B\underline{\widetilde{x}}\|_1}{\|\underline{\widetilde{x}}\|_1} \;\leq\; \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n} \frac{\|B\underline{x}\|_1}{\|\underline{x}\|_1}$$

und somit insgesamt die Behauptung

$$\|B\|_1 \;=\; \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n} \frac{\|B\underline{x}\|_1}{\|\underline{x}\|_1}.$$

■

Eine Matrixnorm $\|\cdot\|_M$ heißt kompatibel beziehungsweise verträglich zur Vektornorm $\|\cdot\|_V$, wenn für beliebige Matrizen $B \in \mathbb{R}^{m\times n}$ und beliebige Vektoren $\underline{x} \in \mathbb{R}^n$ die Ungleichung

$$\|B\underline{x}\|_V \;\leq\; \|B\|_M \|\underline{x}\|_V$$

gilt. Für eine durch eine Vektornorm $\|\cdot\|_V$ induzierte Matrixnorm $\|\cdot\|_M$ folgt

$$\|B\|_M \;=\; \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n} \frac{\|B\underline{x}\|_V}{\|\underline{x}\|_V} \;\geq\; \frac{\|B\underline{x}\|_V}{\|\underline{x}\|_V} \quad \text{für alle } \underline{x} \in \mathbb{R}^n, \|x\|_V \neq 0,$$

das heißt eine induzierte Matrixnorm $\|\cdot\|_M$ ist stets verträglich zu der sie erzeugenden Vektornorm $\|\cdot\|_V$. Ist eine Matrixnorm $\|\cdot\|_M$ durch eine Vektornorm $\|\cdot\|_V$ induziert, so ergibt sich für die Norm der Einheitsmatrix $I \in \mathbb{R}^{n\times n}$

$$\|I\|_M \;=\; \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n} \frac{\|I\underline{x}\|_V}{\|\underline{x}\|_V} \;=\; \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n} \frac{\|\underline{x}\|_V}{\|\underline{x}\|_V} \;=\; 1.$$

Abschließend soll ein Beispiel einer zu einer Vektornorm $\|\cdot\|_V$ verträglichen Matrixnorm $\|\cdot\|_M$ betrachtet werden, die durch keine Vektornorm induziert wird.

Beispiel 3.1. *Sei zunächst $m = n$. Für die Einheitsmatrix $I \in \mathbb{R}^{n\times n}$ gilt dann in der Frobenius–Norm $\|I\|_F = \sqrt{n}$, dies steht aber für $n > 1$ im Widerspruch zu $\|I\|_M = 1$ für eine induzierte Matrix–Norm $\|\cdot\|_M$. Deshalb kann die Frobenius–Norm $\|A\|_F$ durch keine Vektornorm $\|\underline{x}\|_V$ induziert sein.*

Für $B \in \mathbb{R}^{m\times n}$ folgt andererseits mit der Cauchy–Schwarz–Ungleichung (3.1)

$$\|B\underline{x}\|_2^2 \,=\, \sum_{k=1}^m \left(\sum_{\ell=1}^n b_{k\ell}x_\ell\right)^2 \,\leq\, \sum_{k=1}^m \left(\sum_{\ell=1}^n b_{k\ell}^2\right)\left(\sum_{\ell=1}^n x_\ell^2\right) \,=\, \|B\|_F^2\|\underline{x}\|_2^2$$

und somit die Verträglichkeit der Frobenius–Norm $\|B\|_F$ zur Euklidischen Vektornorm $\|\underline{x}\|_2$.

Eine invertierbare Matrix $V \in \mathbb{R}^{n\times n}$ (beziehungsweise $U \in \mathbb{R}^{m\times m}$) heißt orthogonal, wenn ihre inverse Matrix V^{-1} durch die transponierte Matrix $V^\top$ gegeben ist, das heißt

$$V^\top V \,=\, VV^\top \,=\, I_n \in \mathbb{R}^{n\times n}, \quad U^\top U \,=\, UU^\top \,=\, I_m \in \mathbb{R}^{m\times m}.$$

Wegen

$$\|\underline{x}\|_2^2 \,=\, (\underline{x},\underline{x})_2 \,=\, (\underbrace{V^\top V}_{=I}\underline{x},\underline{x})_2 \,=\, (V\underline{x},V\underline{x})_2 \,=\, \|V\underline{x}\|_2^2$$

für beliebige Vektoren $\underline{x} \in \mathbb{R}^n$ folgt mit der Substitution $\underline{x} = V\underline{z}$

$$\|B\|_2 \,=\, \sup_{\underline{0}\neq\underline{x}\in\mathbb{R}^n} \frac{\|B\underline{x}\|_2}{\|\underline{x}\|_2} \,=\, \sup_{\underline{0}\neq\underline{x}=V\underline{z}\in\mathbb{R}^n} \frac{\|BV\underline{z}\|_2}{\|V\underline{z}\|_2} \,=\, \sup_{\underline{0}\neq\underline{z}\in\mathbb{R}^n} \frac{\|BV\underline{z}\|_2}{\|\underline{z}\|_2} \,=\, \|BV\|_2.$$

Analog ergibt sich

$$\|B\|_2 \,=\, \sup_{\underline{0}\neq\underline{x}\in\mathbb{R}^n} \frac{\|B\underline{x}\|_2}{\|\underline{x}\|_2} \,=\, \sup_{\underline{0}\neq\underline{x}\in\mathbb{R}^n} \frac{\|UB\underline{x}\|_2}{\|\underline{x}\|_2} \,=\, \|UB\|_2\,.$$

Insgesamt gilt also für eine beliebige Matrix $B \in \mathbb{R}^{m\times n}$ und orthogonale Matrizen $V \in \mathbb{R}^{n\times n}$ beziehungsweise $U \in \mathbb{R}^{m\times m}$ die Gleichheit

$$\|B\|_2 \,=\, \|UB\|_2 \,=\, \|BV\|_2 \,=\, \|UBV\|_2, \tag{3.4}$$

das heißt die Euklidische Matrixnorm $\|B\|_2$ ist invariant bezüglich orthogonaler Transformationen.
Für $\ell = 1,\ldots,n$ bezeichne $\underline{b}^\ell = (b_{k\ell})_{k=1}^m$ die Spaltenvektoren der Matrix $B \in \mathbb{R}^{m\times n}$ mit der Euklidischen Vektornorm

$$\|\underline{b}^\ell\|_2^2 \,=\, \sum_{k=1}^m b_{k\ell}^2\,.$$

Damit ergibt sich für die Frobenius–Norm der Matrix B die Darstellung

$$\|B\|_F^2 \,=\, \sum_{k=1}^m\sum_{\ell=1}^n b_{k\ell}^2 \,=\, \sum_{\ell=1}^n \|\underline{b}^\ell\|_2^2\,.$$

Andererseits gilt für das Matrixprodukt UB mit einer orthogonalen Matrix $U \in \mathbb{R}^{m\times m}$

$$UB \,=\, \left(U\underline{b}^1,\ldots,U\underline{b}^n\right).$$

Aus der Invarianz der Euklidischen Vektornorm ergibt sich in der Frobenius–Norm

$$\|UB\|_F^2 \;=\; \sum_{\ell=1}^{n} \|U\underline{b}^\ell\|_2^2 \;=\; \sum_{\ell=1}^{n} \|\underline{b}^\ell\|_2^2 \;=\; \|B\|_F^2$$

und somit

$$\|UB\|_F \;=\; \|B\|_F\,.$$

Damit folgt auch, jeweils durch Übergang zur transponierten Matrix, für eine orthogonale Matrix $V \in \mathbb{R}^{n\times n}$

$$\|B\|_F \;=\; \|B^\top\|_F \;=\; \|V^\top B^\top\|_F \;=\; \|(V^\top B^\top)^\top\|_F \;=\; \|BV\|_F.$$

Insgesamt gilt für eine beliebige Matrix $B \in \mathbb{R}^{m\times n}$ und orthogonale Matrizen $V \in \mathbb{R}^{n\times n}$ und $U \in \mathbb{R}^{m\times m}$ die Gleichheit

$$\|B\|_F \;=\; \|UB\|_F \;=\; \|BV\|_F \;=\; \|UBV\|_F, \tag{3.5}$$

das heißt die Invarianz der Frobenius–Norm bezüglich orthogonaler Transformationen.

Ist eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ invertierbar, so definiert

$$\kappa_M(A) \;:=\; \|A\|_M \|A^{-1}\|_M \tag{3.6}$$

die Konditionszahl bezüglich der Matrixnorm $\|\cdot\|_M$. Insbesondere bezeichnet

$$\kappa_2(A) \;=\; \|A\|_2 \|A^{-1}\|_2 \tag{3.7}$$

die spektrale Konditionszahl. Eine Matrix $A \in \mathbb{R}^{n\times n}$ (beziehungsweise die Familie von Matrizen $A \in \mathbb{R}^{n\times n}$ für verschiedene $n \in \mathbb{N}$) heißt schlecht konditioniert, wenn ihre spektrale Konditionszahl $\kappa_2(A)$ proportional zur Dimension n anwächst.

3.2 Eigenwerte und Singulärwerte

Eine komplexe Zahl $\lambda(A) \in \mathbf{C}$ heißt Eigenwert der quadratischen Matrix $A \in \mathbb{R}^{n\times n}$, wenn das lineare Gleichungssystem

$$A\underline{x} \;=\; \lambda(A)\,\underline{x} \tag{3.8}$$

eine nicht triviale Lösung $\underline{x} \in \mathbb{R}^n$ mit $\|\underline{x}\|_V > 0$ besitzt. Diese heißt Eigenvektor zum Eigenwert $\lambda(A)$. Als notwendige Bedingung für die Existenz nichttrivialer Lösungen von (3.8) ergeben sich die μ voneinander verschiedenen Eigenwerte $\lambda_k(A)$ für $k = 1, \ldots, \mu \le n$ als Nullstellen des charakteristischen Polynoms

$$p(\lambda) \;:=\; \det(A - \lambda I) \;=\; (\lambda_1(A) - \lambda)^{\alpha_1} \ldots (\lambda_\mu(A) - \lambda)^{\alpha_\mu} \;=\; \prod_{k=1}^{\mu} (\lambda_k(A) - \lambda)^{\alpha_k}.$$

Die Potenzen $\alpha_k \in \mathbb{N}$ beschreiben die algebraische Vielfachheit des Eigenwertes $\lambda_k(A)$, und es gilt

$$\sum_{k=1}^{\mu} \alpha_k \,=\, n\,.$$

Durch Koeffizientenvergleich des charakteristischen Polynoms folgen

$$\mathrm{spur}(A) \,=\, \sum_{i=1}^{n} a_{ii} \,=\, \sum_{k=1}^{\mu} \alpha_k \lambda_k(A), \quad \det(A) \,=\, \prod_{k=1}^{\mu} [\lambda_k(A)]^{\alpha_k}.$$

Da ein Eigenwert $\lambda_k(A)$ Nullstelle des charakteristischen Polynoms $\det(A - \lambda I)$ ist, so ist auch sein konjugiert komplexer Wert $\overline{\lambda_k(A)}$ Nullstelle und somit Eigenwert von A. Wegen $\det(A - \lambda I) = \det(A^\top - \lambda I)$ sind diese auch Eigenwerte der transponierten Matrix $A^\top$.
Die zum Eigenwert $\lambda_k(A)$ gehörenden Eigenvektoren bilden einen linearen Teilraum,

$$\mathcal{L}(\lambda_k(A)) \,:=\, \{\underline{x} \in \mathbb{R}^n \,:\, A\underline{x} = \lambda_k(A)\underline{x}\}\,, \quad \beta_k \,:=\, \dim \mathcal{L}(\lambda_k(A)),$$

dessen Dimension β_k die Anzahl der linear unabhängigen Eigenvektoren zum Eigenwert $\lambda_k(A)$ angibt. Diese heißt geometrische Vielfachheit des Eigenwerts $\lambda_k(A)$.
Durch

$$\varrho(A) \,:=\, \max_{k=1,\ldots,\mu\leq n} |\lambda_k(A)|$$

wird schließlich der Spektralradius der Matrix A definiert.

Für symmetrische Matrizen $A = A^\top \in \mathbb{R}^{n\times n}$ sind die Eigenwerte $\lambda_k(A)$ für $k = 1,\ldots,n$ reell und die zugehörigen Eigenvektoren $\{\underline{v}^k\}_{k=1}^n$ bilden eine Orthonormalbasis mit

$$(\underline{v}^k, \underline{v}^\ell) \,=\, \delta_{k\ell} \quad \text{für alle } k, \ell = 1,\ldots,n.$$

Ein beliebiges Element $\underline{x} \in \mathbb{R}^n$ kann deshalb durch

$$\underline{x} \,=\, \sum_{k=1}^{n} \xi_k \underline{v}^k \quad \text{mit } \xi_k \,=\, (\underline{x}, \underline{v}^k) \tag{3.9}$$

dargestellt werden, und es gilt

$$\|\underline{x}\|_2^2 \,=\, (\underline{x}, \underline{x}) \,=\, (\sum_{k=1}^{n} \xi_k \underline{v}^k, \sum_{\ell=1}^{n} \xi_\ell \underline{v}^\ell) \,=\, \sum_{k=1}^{n}\sum_{\ell=1}^{n} \xi_k \xi_\ell (\underline{v}^k, \underline{v}^\ell) \,=\, \sum_{k=1}^{n} \xi_k^2$$

sowie

$$(A\underline{x}, \underline{x}) \,=\, \sum_{k=1}^{n}\sum_{\ell=1}^{n} \xi_k \xi_\ell (A\underline{v}^k, \underline{v}^\ell) \,=\, \sum_{k=1}^{n}\sum_{\ell=1}^{n} \xi_k \xi_\ell \lambda_k(A) (\underline{v}^k, \underline{v}^\ell) \,=\, \sum_{k=1}^{n} \lambda_k(A) \xi_k^2.$$

Eine symmetrische Matrix $A = A^\top \in \mathbb{R}^{n\times n}$ heißt positiv definit, falls alle Eigenwerte $\lambda_k(A)$ positiv sind. In diesem Fall folgt

$$(A\underline{x},\underline{x}) \;=\; \sum_{k=1}^n \lambda_k(A)\xi_k^2 \;\geq\; \min_{k=1,\ldots,n} \lambda_k(A) \sum_{k=1}^n \xi_k^2 \;=\; \min_{k=1,\ldots,n} \lambda_k(A)\, \|\underline{x}\|_2^2$$

für alle $\underline{x} \in \mathbb{R}^n$. Weiterhin kann der Rayleigh–Quotient durch die extremalen Eigenwerte von A abgeschätzt werden, das heißt für alle $\underline{x} \in \mathbb{R}^n$ mit $\|\underline{x}\|_V > 0$ gilt

$$\min_{k=1,\ldots,n} \lambda_k(A) \;\leq\; \frac{(A\underline{x},\underline{x})}{(\underline{x},\underline{x})} \;\leq\; \max_{k=1,\ldots,n} \lambda_k(A).$$

Damit folgt

$$\lambda_{\min}(A) \;=\; \min_{0\neq\underline{x}\in\mathbb{R}^n} \frac{(A\underline{x},\underline{x})}{(\underline{x},\underline{x})}, \quad \lambda_{\max}(A) \;=\; \max_{0\neq\underline{x}\in\mathbb{R}^n} \frac{(A\underline{x},\underline{x})}{(\underline{x},\underline{x})}.$$

Gelten die Spektraläquivalenzungleichungen

$$c_1^A\,(\underline{x},\underline{x}) \;\leq\; (A\underline{x},\underline{x}) \;\leq\; c_2^A\,(\underline{x},\underline{x}) \tag{3.10}$$

für alle $\underline{x} \in \mathbb{R}^n$ mit positiven Konstanten c_1^A und c_2^A, so folgt

$$c_1^A \;\leq\; \lambda_{\min}(A) \;\leq\; \lambda_{\max}(A) \;\leq\; c_2^A\,,$$

das heißt, die Konstanten c_1^A und c_2^A sind untere beziehungsweise obere Schranken der extremalen Eigenwerte der positiv definiten Matrix A.
Für eine symmetrische und positiv definite Matrix $A \in \mathbb{R}^{n\times n}$ kann durch

$$\langle \underline{u},\underline{v}\rangle_A \;:=\; (A\underline{u},\underline{v}) \;=\; (\underline{u},A\underline{v}) \;:\; \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{3.11}$$

das A–energetische Skalarprodukt erklärt werden. Die durch dieses Skalarprodukt induzierte Vektornorm

$$||\underline{x}||_A \;:=\; [\langle \underline{x},\underline{x}\rangle_A]^{1/2} \;=\; (A\underline{x},\underline{x})^{1/2} \tag{3.12}$$

wird als A–energetische Vektornorm bezeichnet.
Die durch die Eigenvektoren von $A = A^\top \in \mathbb{R}^{n\times n}$ gebildete Matrix

$$V \;=\; (\underline{v}^1,\ldots,\underline{v}^n) \in \mathbb{R}^{n\times n}$$

ist orthogonal, und es gilt

$$AV \;=\; \big(A\underline{v}^1,\ldots,A\underline{v}^n\big) \;=\; \big(\lambda_1(A)\underline{v}^1,\ldots,\lambda_n(A)\underline{v}^n\big) \;=\; VD$$

mit der durch die Eigenwerte von A definierten Diagonalmatrix

$$D \;=\; \operatorname{diag}\,(\lambda_k(A))_{k=1}^n\,.$$

Multiplikation mit $V^\top$ von links ergibt wegen der Orthogonalität der Eigenvektoren

$$V^\top A V \,=\, D \tag{3.13}$$

beziehungsweise durch die Multiplikation mit $V^\top$ von rechts folgt die bekannte Faktorisierung der Matrix A,

$$A \,=\, V\,D\,V^\top \,=\, \sum_{k=1}^{n} \lambda_k(A)\,\underline{v}^k \underline{v}^{k,\top}. \tag{3.14}$$

Die Darstellung (3.14) ist einerseits Grundlage für die Definition einer Niedrig–Rang Approximation von A, andererseits ermöglicht sie die symmetrische Vorkonditionierung eines linearen Gleichungssystems $A\underline{x} = \underline{f}$ zur Verbesserung der spektralen Konditionszahl der vorkonditionierten Systemmatrix. Hierzu wird die Wurzel einer symmetrischen und positiv definiten Matrix A benötigt: Für positive Eigenwerte $\lambda_k(A) > 0$, $k = 1, \ldots, n$, kann die Diagonalmatrix

$$D^{1/2} \,=\, \operatorname{diag}\left(\sqrt{\lambda_k(A)}\right)_{k=1}^{n}$$

und somit die symmetrische und positiv definite Matrix

$$A^{1/2} \,=\, V D^{1/2} V^\top \tag{3.15}$$

erklärt werden. Nach Konstruktion gilt

$$A^{1/2} A^{1/2} \,=\, V D^{1/2} \underbrace{V^\top V}_{=I} D^{1/2} V^\top \,=\, V D V^\top \,=\, A\,.$$

Entsprechend kann

$$A^{-1/2} \,=\, (A^{1/2})^{-1} \,=\, V D^{-1/2} V^\top, \quad D^{-1/2} \,=\, \operatorname{diag}\left(\frac{1}{\sqrt{\lambda_k(A)}}\right)_{k=1}^{n}$$

definiert werden. Mit der Transformation $\underline{x} = A^{-1/2}\underline{z}$ folgt aus den Spektraläquivalenzungleichungen (3.10) auch die Gültigkeit der Spektraläquivalenzungleichungen

$$\frac{1}{c_2^A}\,(\underline{z},\underline{z}) \,\le\, (A^{-1}\underline{z},\underline{z}) \,\le\, \frac{1}{c_1^A}\,(\underline{z},\underline{z}) \tag{3.16}$$

für alle $\underline{z} \in \mathbb{R}^n$.

Der Rang einer Matrix A beschreibt die Anzahl der linear unabhängigen Zeilen beziehungsweise Spalten von A. Die Darstellung (3.14) zeigt, daß der Rang einer symmetrischen Matrix $A \in \mathbb{R}^{n\times n}$ mit der Anzahl der nicht verschwindenden Eigenwerte zusammenfällt, das heißt es gilt

$$A \,=\, \sum_{k=1}^{\operatorname{rang} A} \lambda_k(A) \underline{v}^k \underline{v}^{k,\top},$$

falls eine entsprechende Nummerierung der Eigenwerte mit $\lambda_k(A) = 0$ für $k > \text{rang}A$ vorausgesetzt wird.
Aus der Norminvarianz (3.4) folgt schließlich

$$\|A\|_2 \,=\, \|VDV^\top\|_2 \,=\, \|D\|_2 \,=\, \max_{k=1,\ldots,n} |\lambda_k(A)| \,=\, \varrho(A),$$

beziehungsweise gilt mit der Invarianz (3.5) der Frobenius–Norm

$$\|A\|_F \,=\, \|VDV^\top\|_F \,=\, \|D\|_F \,=\, \sqrt{\sum_{k=1}^n [\lambda_k(A)]^2}.$$

Ist die Matrix A invertierbar, so sind die Eigenwerte der Inversen A^{-1} durch

$$\lambda_k(A^{-1}) = [\lambda_k(A)]^{-1}$$

gegeben. Ist A zusätzlich symmetrisch und positiv definit, so folgt für die spektrale Konditionszahl

$$\kappa_2(A) \,=\, \|A\|_2\|A^{-1}\|_2 \,=\, \varrho(A)\varrho(A^{-1}) \,=\, \frac{\max\limits_{k=1,\ldots,n} |\lambda_k(A)|}{\min\limits_{k=1,\ldots,n} |\lambda_k(A)|} \,=\, \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

Bei den obigen Überlegungen wurden Matrizen $A \in \mathbb{R}^{n\times n}$ betrachtet. Sei nun $B \in \mathbb{R}^{m\times n}$ eine beliebig gegebene Matrix mit $\text{rang}\, B \le \min\{m,n\}$. Dann definiert $A := B^\top B \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix mit $\text{rang}\, A \le \min\{m,n\}$ und n reellen Eigenwerten

$$\lambda_k(A) = \lambda_k(B^\top B)$$

sowie einem zugehörigen orthonormalen System $\{\underline{v}^k\}_{k=1}^n$ von Eigenvektoren. Dieses bildet eine Basis des $\mathbb{R}^n$, so daß jedes Element $\underline{x} \in \mathbb{R}^n$ wie in (3.9) dargestellt werden kann,

$$\underline{x} \,=\, \sum_{k=1}^n \xi_k \underline{v}^k \quad \text{mit } \xi_k \,=\, (\underline{x}, \underline{v}^k).$$

Wegen

$$\begin{aligned} 0 \;&\le\; \|B\underline{x}\|_2^2 \,=\, (B\underline{x}, B\underline{x}) \,=\, (B^\top B\underline{x}, \underline{x}) \,=\, (A\underline{x}, \underline{x}) \\ &=\; \sum_{k=1}^n \sum_{\ell=1}^n \xi_k \xi_\ell (A\underline{v}^k, \underline{v}^\ell) \,=\, \sum_{k=1}^n \sum_{\ell=1}^n \xi_k \xi_\ell \lambda_k(A)(\underline{v}^k, \underline{v}^\ell) \,=\, \sum_{k=1}^n \lambda_k(A)\, \xi_k^2 \end{aligned}$$

folgt $\lambda_k(A) \ge 0$ für alle $k = 1, \ldots, n$. Ohne Einschränkung der Allgemeinheit gelte $\lambda_k(A) > 0$ für alle $k = 1, \ldots, \mu = \text{rang}\, A \le \min\{m,n\}$ und $\lambda_k(A) = 0$ für $k = \mu+1, \ldots, n$. Nach (3.13) gilt die Faktorisierung

$$V^\top AV \,=\, V^\top B^\top BV \,=\, D \,=\, \text{diag}\,(\lambda_k(A))_{k=1}^n\,. \tag{3.17}$$

Wegen $\lambda_k(A) \geq 0$ für $k = 1, \ldots, \min\{m, n\}$ existieren die Singulärwerte

$$\sigma_k(B) \;=\; \sqrt{\lambda_k(A)} \;=\; \sqrt{\lambda_k(B^\top B)} \geq 0 \quad \text{für } k = 1, \ldots, \min\{m, n\}.$$

Insbesondere gilt $\sigma_k(B) > 0$ für $k = 1, \ldots, \mu \leq \min\{m, n\}$. Die Singulärwerte definieren eine Diagonalmatrix

$$\Sigma \;=\; \operatorname{diag}\left(\sigma_k(B)\right)_{k=1}^{\min\{m,n\}} \in \mathbb{R}^{m \times n} \tag{3.18}$$

und es gilt

$$D \;=\; \Sigma^\top \Sigma \in \mathbb{R}^{n \times n}.$$

Wird durch

$$\Sigma^+ \;=\; \begin{pmatrix} \frac{1}{\sigma_1(B)} & & & & & \\ & \ddots & & & & \\ & & \frac{1}{\sigma_\mu(B)} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \in \mathbb{R}^{n \times m} \tag{3.19}$$

die Pseudoinverse zu Σ definiert, das heißt

$$\Sigma^+ \Sigma \;=\; \begin{pmatrix} I_\mu & \\ & 0 \end{pmatrix} \in \mathbb{R}^{n \times n},$$

dann folgt aus der Faktorisierung (3.17) durch Multiplikation mit der Pseudoinversen $\Sigma^{+,\top}$ von links

$$\Sigma^{+,\top} V^\top B^\top B V \;=\; \Sigma \in \mathbb{R}^{m \times n}$$

beziehungsweise

$$U^\top B V \;=\; \Sigma \tag{3.20}$$

mit

$$U \;=\; B V \Sigma^+ \in \mathbb{R}^{m \times m}.$$

Wegen

$$U^\top U \;=\; \Sigma^{+,\top} V^\top B^\top B V^\top \Sigma^+ \;=\; \Sigma^{+,\top} D \Sigma^+ \;=\; \begin{pmatrix} I_\mu & \\ & 0 \end{pmatrix} \in \mathbb{R}^{m \times m}$$

ist $U^\top$ die Pseudoinverse zu U. Damit folgt aus (3.20) die Singulärwertzerlegung von $B \in \mathbb{R}^{m \times n}$,

$$B \;=\; U \Sigma V^\top \;=\; \sum_{k=1}^{\mu} \sigma_k(B) \underline{u}^k \underline{v}^{k,\top}, \tag{3.21}$$

das heißt $\mu = \operatorname{rang} B$ beschreibt die Anzahl der nicht verschwindenden Singulärwerte von B. Aus der Invarianz (3.4) der Euklidischen Matrixnorm folgt schließlich

$$\|B\|_2 = \|U \Sigma V^\top\|_2 = \|\Sigma\|_2 = \max_{k=1,\ldots,\mu} \sigma_k(B) = \max_{k=1,\ldots,\mu} \sqrt{\lambda_k(B^\top B)} = \sqrt{\varrho(B^\top B)}$$

beziehungsweise ist mit der Invarianz (3.5) der Frobenius–Norm

$$\|B\|_F = \|U\Sigma V^\top\|_F = \|\Sigma\|_F = \sqrt{\sum_{k=1}^{\mu}[\sigma_k(B)]^2}.$$

Multiplikation der Gleichung (3.20) von rechts mit $V^\top$ und Übergang zur Transponierten ergibt

$$B^\top U = V\Sigma$$

und somit folgt durch Vergleich der Spaltenvektoren

$$B^\top \underline{u}_k = \sigma_k(B)\, \underline{v}_k \quad \text{für } k = 1, \ldots, \min\{m, n\}.$$

Multiplikation der Gleichung (3.20) von links mit U liefert

$$BV = U\Sigma$$

und somit

$$B\underline{v}_k = \sigma_k(B)\, \underline{u}_k \quad \text{für } k = 1, \ldots, \min\{m, n\}.$$

3.3 Orthogonalisierung von Vektorsystemen

Für $m \in \mathbb{N}$ mit $m \leq n$ heißt ein System[1] $\{\underline{w}^k\}_{k=0}^{m-1}$ von m nicht verschwindenden Vektoren $\underline{w}^k \in \mathbb{R}^n$, das heißt es gilt $\|\underline{w}^k\|_V > 0$, linear unabhängig, wenn die Gleichheit

$$\sum_{k=0}^{m-1} \alpha_k \underline{w}^k = \underline{0}$$

nur für die triviale Lösung

$$\alpha_0 = \cdots = \alpha_k = \cdots = \alpha_{m-1} = 0$$

erfüllt ist. Die Vektoren $\{\underline{w}^k\}_{k=0}^{m-1}$ heißen zueinander orthogonal bezüglich dem Skalarprodukt $\langle \cdot, \cdot \rangle$, falls

$$\langle \underline{w}^k, \underline{w}^\ell \rangle = 0 \quad \text{für alle } k, \ell = 0, \ldots, m-1 \text{ und } k \neq \ell$$

gilt, und orthonormal, wenn

$$\langle \underline{w}^k, \underline{w}^\ell \rangle = \delta_{k\ell} \quad \text{für alle } k, \ell = 0, \ldots, m-1$$

erfüllt ist. Für $m = n$ heißt das System $\{\underline{w}^k\}_{k=0}^{n-1}$ von n linear unabhängigen Vektoren Basis des $\mathbb{R}^n$, das heißt ein beliebiges Element $\underline{u} \in \mathbb{R}^n$ kann als Linearkombination der Basisvektoren $\{\underline{w}^k\}_{k=0}^{n-1}$ dargestellt werden.

[1] Im Hinblick auf die später beschriebenen Iterationsverfahren zur Lösung linearer Gleichungssysteme werden Vektorsysteme $\{\underline{w}^k\}_{k=0}^{n-1}$ stets von $k = 0, \ldots, n-1$ indiziert.

Beispiel 3.2. *Die Einheitsvektoren*

$$\underline{e}^k \;=\; \left(\delta_{(k+1)j}\right)_{j=1}^n \quad \textit{für } k = 0, \ldots, n-1$$

bilden eine Basis des $\mathbb{R}^n$*. Diese wird als kanonische Basis bezeichnet. Die Einheitsvektoren* $\underline{e}^k$ *sind orthonormal bezüglich dem Euklidischen Skalarprodukt, und für einen beliebigen Vektor* $\underline{u} = (u_1, \ldots, u_n)^\top \in \mathbb{R}^n$ *gilt die Darstellung*

$$\underline{u} \;=\; \sum_{k=0}^{n-1} u_{k+1}\, \underline{e}^k \in \mathbb{R}^n.$$

Gegeben sei jetzt eine beliebige Basis $\{\underline{w}^k\}_{k=0}^{n-1}$ des $\mathbb{R}^n$, gesucht ist eine bezüglich dem Skalarprodukt $\langle \cdot, \cdot \rangle$ orthogonale Basis $\{\underline{p}^k\}_{k=0}^{n-1}$ mit

$$\langle \underline{p}^k, \underline{p}^\ell \rangle \;=\; 0 \quad \text{für } k, \ell = 0, \ldots, n-1 \text{ und } k \neq \ell.$$

Diese kann durch das Gram–Schmidtsche Orthogonalisierungsverfahren wie folgt konstruiert werden:

Setze

$$\underline{p}^0 := \underline{w}^0.$$

Für $k = 0, \ldots, n-2$ berechne

$$\underline{p}^{k+1} := \underline{w}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell}\underline{p}^\ell, \quad \beta_{k\ell} = \frac{\langle \underline{w}^{k+1}, \underline{p}^\ell \rangle}{\langle \underline{p}^\ell, \underline{p}^\ell \rangle}.$$

Algorithmus 1.1: Orthogonalisierungsverfahren nach Gram–Schmidt.

Lemma 3.4. *Sei* $\{\underline{w}^k\}_{k=0}^{n-1}$ *ein System linear unabhängiger Vektoren. Dann ist das durch das Gram–Schmidtsche Orthogonalisierungsverfahren* (Algorithmus 1.1) *erzeugte Vektorsystem* $\{\underline{p}^k\}_{k=0}^{n-1}$ *orthogonal, das heißt es gilt*

$$\langle \underline{p}^k, \underline{p}^\ell \rangle \;=\; 0 \quad \textit{für } k, \ell = 0, \ldots, n-1, \; k \neq \ell$$

und

$$\langle \underline{p}^k, \underline{p}^k \rangle \;>\; 0 \quad \textit{für } k = 0, \ldots, n-1.$$

Beweis: Der Nachweis erfolgt durch vollständige Induktion nach k. Für $k = 0$ ist $\underline{p}^0 = \underline{w}^0$ und es gilt $\langle \underline{p}^0, \underline{p}^0 \rangle > 0$. Dann ist $\underline{p}^1$ durch

$$\underline{p}^1 \;=\; \underline{w}^1 - \beta_{10}\underline{p}^0, \quad \beta_{10} \;=\; \frac{\langle \underline{w}^1, \underline{p}^0 \rangle}{\langle \underline{p}^0, \underline{p}^0 \rangle}$$

wohldefiniert, und die Orthogonalität folgt aus

$$\langle \underline{p}^1, \underline{p}^0 \rangle \;=\; \langle \underline{w}^1 - \beta_{10}\underline{p}^0, \underline{p}^0 \rangle \;=\; \langle \underline{w}^1, \underline{p}^0 \rangle - \frac{\langle \underline{w}^1, \underline{p}^0 \rangle}{\langle \underline{p}^0, \underline{p}^0 \rangle} \langle \underline{p}^0, \underline{p}^0 \rangle \;=\; 0.$$

Zu zeigen bleibt $\langle \underline{p}^1, \underline{p}^1 \rangle > 0$. Dieser Nachweis erfolgt **indirekt**, das heißt aus der Annahme $\langle \underline{p}^1, \underline{p}^1 \rangle = 0$ folgt

$$\underline{0} \,=\, \underline{p}^1 \,=\, \underline{w}^1 - \beta_{10}\underline{p}^0 \,=\, \underline{w}^1 - \beta_{10}\underline{w}^0$$

im Widerspruch zur linearen Unabhängigkeit der Vektoren $\underline{w}^0$ und $\underline{w}^1$.
Für $k = 1$ gelten somit die Induktionsvoraussetzungen

$$\langle \underline{p}^\ell, \underline{p}^j \rangle \,=\, 0 \quad \text{für alle } \ell, j = 0, \ldots, k \text{ mit } \ell \neq j$$

und

$$\langle \underline{p}^\ell, \underline{p}^\ell \rangle > 0 \quad \text{für alle } \ell = 0, \ldots, k.$$

Aus der Induktionsvoraussetzung für $k \in \mathbb{N}$ folgt durch Einsetzen der Koeffizienten β_{kj} für den Induktionsschritt $k+1$ die Orthogonalität

$$\langle \underline{p}^{k+1}, \underline{p}^j \rangle \,=\, \langle \underline{w}^{k+1}, \underline{p}^j \rangle - \sum_{\ell=0}^{k} \beta_{k\ell} \langle \underline{p}^\ell, \underline{p}^j \rangle \,=\, \langle \underline{w}^{k+1}, \underline{p}^j \rangle - \beta_{kj} \langle \underline{p}^j, \underline{p}^j \rangle \,=\, 0$$

für $j = 0, \ldots, k$. Zu zeigen bleibt $\langle \underline{p}^{k+1}, \underline{p}^{k+1} \rangle > 0$. Nach Konstruktion gilt

$$\underline{p}^\ell \in \operatorname{span}\left\{\underline{w}^0, \ldots, \underline{w}^\ell\right\} \quad \text{für alle } \ell = 0, \ldots, k+1.$$

Die Annahme $\underline{p}^{k+1} = \underline{0}$ führt dann wegen

$$\underline{0} \,=\, \underline{p}^{k+1} \,=\, \underline{w}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell} \underline{p}^\ell \,=\, \underline{w}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell} \sum_{j=0}^{\ell} \alpha_{\ell j} \underline{w}^j$$

zum Widerspruch zur Voraussetzung der linearen Unabhängigkeit des Vektorsystems $\{\underline{w}^\ell\}_{\ell=0}^{k+1}$.
Damit ist Algorithmus 1.1 wohldefiniert. ■

Sei $A \in \mathbb{R}^{n \times n}$ eine invertierbare Matrix mit $\operatorname{rang} A = n$. Dann bilden die Spaltenvektoren von A,

$$A \,=\, \left(\underline{a}^1, \ldots, \underline{a}^n\right) \in \mathbb{R}^{n \times n},$$

ein linear unabhängiges Vektorsystem $\{\underline{a}^k\}_{k=1}^n$. Die Anwendung des Orthogonalisierungsverfahrens nach Gram–Schmidt bezüglich dem Euklidischen Skalarprodukt und bei gleichzeitiger Normierung,

$$\underline{\hat{v}}^k = \underline{a}^k - \sum_{\ell=1}^{k-1} (\underline{a}^k, \underline{v}^\ell)\, \underline{v}^\ell, \quad \underline{v}^k = \frac{1}{\|\underline{\hat{v}}^k\|_2}\, \underline{\hat{v}}^k \quad \text{für } k = 1, \ldots, n,$$

liefert für die Spaltenvektoren von A die Darstellung

$$\underline{a}^k \,=\, \|\underline{\hat{v}}^k\|_2\, \underline{v}^k + \sum_{\ell=1}^{k-1} (\underline{a}^k, \underline{v}^\ell)\, \underline{v}^\ell \quad \text{für } k = 1, \ldots, n.$$

In Matrixschreibweise lautet diese

$$A \,=\, Q\,R \tag{3.22}$$

mit

$$Q \,=\, \big(\underline{v}^1, \ldots, \underline{v}^n\big) \in \mathbb{R}^{n \times n}, \quad Q^\top Q = I,$$

und der durch

$$R[\ell, k] \,=\, \left\{ \begin{array}{cl} (\underline{a}^k, \underline{v}^\ell) & \text{für } \ell = 1, \ldots, k-1, \\ \|\hat{\underline{v}}^k\|_2 & \text{für } \ell = k, \\ 0 & \text{für } \ell = k+1, \ldots, n. \end{array} \right.$$

für $k = 1, \ldots, n$ definierten oberen Dreiecksmatrix R. Durch das Orthogonalisierungsverfahren von Gram–Schmidt kann also die QR–Zerlegung (3.22) einer regulären Matrix $A \in \mathbb{R}^{n \times n}$ berechnet werden.

Wird ein gegebenes linear unabhängiges Vektorsystem $\{\underline{w}^\ell\}_{\ell=0}^{n-1}$ bezüglich dem A–energetischen Skalarprodukt (3.11) orthogonalisiert, so nennt man das resultierende Vektorsystem $\{\underline{p}^\ell\}_{\ell=0}^{n-1}$ A–orthogonal beziehungsweise konjugiert, das heißt es gilt

$$\langle \underline{p}^k, \underline{p}^\ell \rangle_A \,=\, (A\underline{p}^k, \underline{p}^\ell) \,=\, (\underline{p}^k, A\underline{p}^\ell) \,=\, 0 \quad \text{für } k \neq \ell.$$