2.3 Orthogonale Matrizen – QR-Zerlegung

Im Folgenden soll für eine gegebene Matrix $A \in \mathbb{R}^{n \times m}$, $1 \le m \le n$, eine Faktorisierung der Form

$$A = QS \qquad (2.15)$$

bestimmt werden mit einer orthogonalen Matrix Q, d.h.

$$Q \in \mathbb{R}^{n \times n}, \quad Q^{-1} = Q^T,$$

und einer verallgemeinerten Oberen Dreiecksmatrix

$$S = \left[\begin{array}{c} R \\ \hline 0 \end{array}\right] \in \mathbb{R}^{n \times m}, \quad R = \left[\begin{array}{c} \diagdown \\ 0 \end{array}\right] \in \mathbb{R}^{m \times m};$$

$$(2.16) \qquad 0 = (0) \in \mathbb{R}^{(n-m) \times m} \text{ (Nullmatrix)}$$

Solche Zerlegungen ermöglichen z.B. die stabile Lösung von möglicherweise schlecht konditionierten lösbaren lin. Gl.-Systemen $A\vec{x} = \vec{b}$ $(m = n)$ oder die stabile Lösung von Ausgleichsproblemen $\|A\vec{x} - \vec{b}\|_2 \to \min$, $\vec{x} \in \mathbb{R}^m$.

Wir erinnern uns an die Eigenschaften orthogonaler Matrizen:

i) $\|Q\vec{x}\|_2 = \|\vec{x}\|_2 = \|Q^T\vec{x}\|_2$, $\vec{x} \in \mathbb{R}^n$ (2.17),

ii) $\mathrm{cond}_2(QA) = \mathrm{cond}_2(A)$ (2.18)

für $A \in \mathbb{R}^{n\times n}$, regulär, und $Q \in \mathbb{R}^{n\times n}$, orthogonal,

iii) für $Q_1, Q_2 \in \mathbb{R}^{n\times n}$, orthogonal, gilt

$Q_1 Q_2$ ist orthogonal

Faktorisierung $A = QR$ mittels Gram-Schmidt-Orthogonalisierung

Für quadratische reguläre Matrizen A $(m=n)$ hat (2.15), (2.16) die Form

$$A = QR \tag{2.19}$$

mit Q orthogonal und R obere Dreiecksmatrix vom Typ $(n\times n)$. Schreiben A, Q, R in der Form

$$A = \left[\vec{a}_1 \middle| \vec{a}_2 \middle| \dots \middle| \vec{a}_n\right], \quad Q = \left[\vec{q}_1 \middle| \vec{q}_2 \middle| \dots \middle| \vec{q}_n\right], \quad R = \begin{bmatrix} r_{11} & \dots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{bmatrix}$$

mit den Spaltenvektoren $\vec{a}_k, \vec{q}_k \in \mathbb{R}^n$, $k = 1, \dots, n$

(2.19) bedeutet dann

$$\vec{a}_j = \sum_{i=1}^{j} r_{ij} \vec{q}_i \quad , \quad j = 1, \dots, n, \quad \vec{q}_1, \dots, \vec{q}_n \in \mathbb{R}^n \tag{2.20}$$

paarweise orthogonal

Algorithmus von Gram-Schmidt zur Orthogonalisierung:

a) Ausgangspunkt

Man hat $j-1$ orthonormale Vektoren $\vec{q}_1, \dots, \vec{q}_{j-1} \in \mathbb{R}^n$ mit

$$\mathrm{span}(\vec{a}_1, \dots, \vec{a}_{j-1}) = \mathrm{span}(\vec{q}_1, \dots, \vec{q}_{j-1}) =: M_{j-1}$$

b) man bestimmt im Schritt $j \geq 1$
das Lot von $\vec{a}_j$ auf den lin. Unterraum $M_{j-1} \subset \mathbb{R}^n$

$$\hat{\vec{q}}_j := \vec{a}_j - \sum_{i=1}^{j-1} (\vec{a}_j^T \vec{q}_i) \vec{q}_i \qquad (2.21)$$

und nach der Normierung

$$\vec{q}_j = \frac{\hat{\vec{q}}_j}{\|\hat{\vec{q}}_j\|}$$

sind die Vektoren $\vec{q}_1, \dots, \vec{q}_j \in \mathbb{R}^n$ paarweise orthonormal und es gilt

$$\operatorname{span}(\vec{a}_1, \dots, \vec{a}_j) = \operatorname{span}(\vec{q}_1, \dots, \vec{q}_j).$$

Aus der Gleichung (2.21) folgt

$$\vec{a}_j = \underbrace{\|\hat{\vec{q}}_j\|_2}_{=:r_{jj}} \vec{q}_j + \sum_{i=1}^{j-1} \underbrace{(\vec{a}_j^T \vec{q}_i)}_{=:r_{ij}} \vec{q}_i, \quad j = 1, \dots, n$$

$$= \sum_{i=1}^{j} r_{ij} \vec{q}_i \qquad (2.22)$$

Nach Abschluss der Gram-Schmidt-Orthogonalisierung hat man damit mit (2.22)

$$\left[\vec{a}_1 | \vec{a}_2 | \dots | \vec{a}_n\right] = \left[\vec{q}_1 | \vec{q}_2 | \dots | \vec{q}_n\right] \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ & r_{22} & \dots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}$$

als QR-Zerlegung.

Bemerkung 2.10

Der beschriebene Algorithmus kann im ungünstigen Fall Probleme bereiten (nicht gutartig sein), wenn z.B. $\|\hat{\vec{q}}_j\|$ sehr klein wird (Lösung folgt)

Householder-Matrizen / Transformationen

Definition 2.11

Eine Abb. $\mathcal{H}: \mathbb{R}^n \to \mathbb{R}^n$, $\vec{x} \longmapsto \mathcal{H}\vec{x}$ mit einer Matrix $\mathcal{H} = E - 2\vec{w}\vec{w}^T$, $\vec{w} \in \mathbb{R}^n$, $\|\vec{w}\|_2 = \vec{w}^T\vec{w} = 1$, (2.23)

bezeichnet man als Householder-Transformation und $\mathcal{H}$ als Householder-Matrix

Eigenschaften von $\mathcal{H}$:

$\mathcal{H}^T = \mathcal{H}$	Symmetrie
$\mathcal{H}^2 = E$	($\mathcal{H}$ ist involutorisch)
$\mathcal{H}^T\mathcal{H} = E$	Orthogonalität

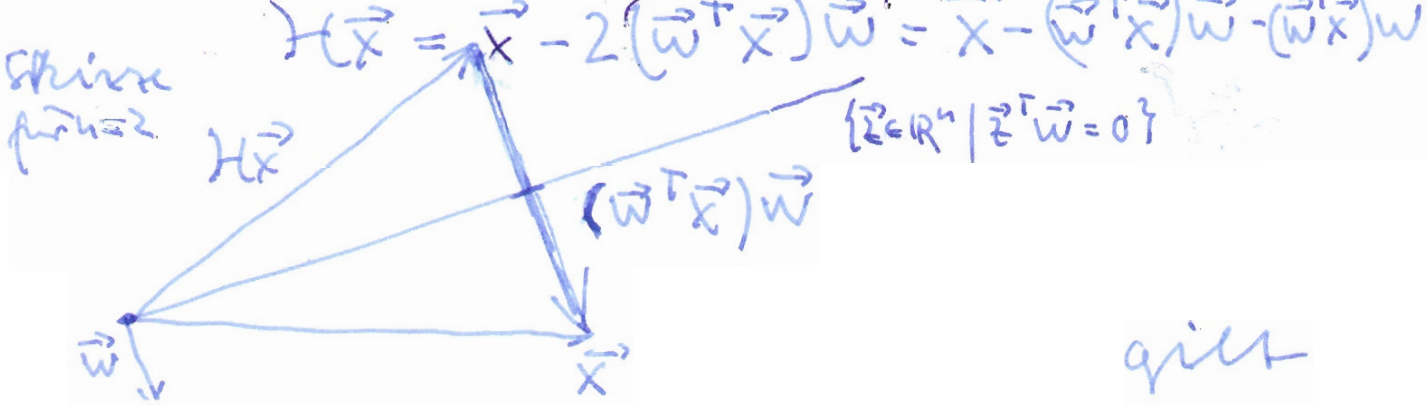
(Nachweis als Übung)

Wirkung der Householder-Transformation

Spiegelung von $\vec{x} \in \mathbb{R}^n$ an der Hyperebene $\{\vec{z} \in \mathbb{R}^n \mid \vec{z}^T\vec{w} = 0\}$, da die Identität

$$\mathcal{H}\vec{x} = \vec{x} - 2\overbrace{(\vec{w}^T\vec{x})\vec{w}}^{=\vec{w}\vec{w}^T\vec{x}} = \vec{x} - (\vec{w}^T\vec{x})\vec{w} - (\vec{w}^T\vec{x})\vec{w}$$

Skizze für $n=2$



gilt

Hilfssatz 2.12

Gegeben sei $0 \neq \vec{x} \in \mathbb{R}^n$ mit $\vec{x} \notin \operatorname{span}\{\vec{e}_1\}$.

Für $\vec{w} = \dfrac{\vec{x} + \sigma \vec{e}_1}{\|\vec{x} + \sigma \vec{e}_1\|_2}$ mit $\sigma = \pm \|\vec{x}\|_2$ (2.24)

gilt $\|\vec{w}\|_2 = 1$,

$$H\vec{x} = (E - 2\vec{w}\vec{w}^T)\vec{x} = -\sigma \vec{e}_1 \quad \Big\} \ (2.25)$$

Beweis

$\|\vec{w}\|_2 = 1$, weil $\vec{x} + \sigma \vec{e}_1 \neq 0$ und damit (2.24) wohl definiert ist und damit gilt

Für den Nachweis von (2.25) erhält man

$$\|\vec{x} + \sigma \vec{e}_1\|_2^2 = \|\vec{x}\|_2^2 + 2\sigma \vec{e}_1^T \vec{x} + \sigma^2$$
$$= 2(\vec{x} + \sigma \vec{e}_1)^T \vec{x}$$

und mit (2.24), d.h. $\dfrac{(\vec{x} + \sigma \vec{e}_1)^T}{\|\vec{x} + \sigma \vec{e}_1\|_2} = \vec{w}^T$

folgt

$$2\vec{w}^T \vec{x} = \frac{2(\vec{x} + \sigma \vec{e}_1)^T \vec{x}}{\|\vec{x} + \sigma \vec{e}_1\|_2} = \|\vec{x} + \sigma \vec{e}_1\|_2,$$

die normalige Nutzung von (2.24) ergibt

$$2\vec{w}\vec{w}^T \vec{x} = \vec{x} + \sigma \vec{e}_1$$
$$\Longleftrightarrow \vec{x} - 2\vec{w}\vec{w}^T \vec{x} = -\sigma \vec{e}_1$$

was zu zeigen war. $\square$

Bemerkung

Um Stellenauslöschungen zu vermeiden, wird in (2.24) $\sigma = \operatorname{sgn}(x_1)\|\vec{x}\|_2$ gewählt, d.h. z.B. für $\vec{x} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$ ist $\sigma = -\sqrt{3^2}$.

Algorithmus zur Konstruktion der Faktorisierung mittels Householder-Transformationen

Ausgehend von $A = A^{(1)} \in \mathbb{R}^{n \times m}$ sollen sukzessive Matrizen der Form

$$A^{(j)} = \begin{pmatrix} a_{11}^{(j)} & a_{12}^{(j)} & & \cdots & & a_{1m}^{(j)} \\ & \ddots & & & & \\ & & a_{j-1\,j-1}^{(j)} & \cdots & & a_{j-1\,m}^{(j)} \\ & & & a_{jj}^{(j)} & \cdots & a_{jm}^{(j)} \\ & & & \vdots & & \vdots \\ & & & a_{nj}^{(j)} & \cdots & a_{nm}^{(j)} \end{pmatrix} \qquad j = 2, \dots, m+1 \qquad (2.26)$$

berechnet werden, so dass am Ende mit $A^{(m+1)} = S$ die verallgemeinerte obere Dreiecksmatrix vorliegt.

Die Matrizen der Form (2.26) erhält man für $j = 1, \dots, m$ durch Transformationen der Form

$$A^{(j+1)} = \hat{H}_j A^{(j)}, \quad \hat{H}_j = \left[\begin{array}{c|c} E_{j-1} & \Theta \\ \hline \Theta & H_j \end{array}\right]$$

$$H_j = E_{n-(j-1)} - 2\vec{w}_j \vec{w}_j^T, \quad \|\vec{w}_j\|_2 = 1,$$

E_l ist Einheitsmatrix aus $\mathbb{R}^{l \times l}$, und $\vec{w}_j \in \mathbb{R}^{n-(j-1)}$ ist so zu wählen, dass gilt

$$H_j \underbrace{\begin{pmatrix} a_{jj}^{(j)} \\ \vdots \\ a_{nj}^{(j)} \end{pmatrix}}_{=:\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_{n-(j-1)} \end{pmatrix}} = \sigma \underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=:\vec{e}_1}, \quad \vec{w}_j = \frac{\vec{a} + \operatorname{sgn}(a_1)\|\vec{a}\|_2 \vec{e}_1}{\|\vec{a} + \operatorname{sgn}(a_1)\|\vec{a}\|_2 \vec{e}_1\|_2}$$

$$\vec{a}, \vec{e}_1 \in \mathbb{R}^{n-(j-1)},$$

Die Matrizen $\hat{H}_1, \dots, \hat{H}_m$ sind aufgrund der Eigenschaften der Matrizen $H_1, \dots, H_m$ orthogonal und symmetrisch, so dass man mit

$$S = \hat{H}_m \hat{H}_{m-1} \cdot \ldots \cdot \hat{H}_1 A \quad , \quad Q = \hat{H}_1 \hat{H}_2 \cdot \ldots \cdot \hat{H}_m$$

die Faktorisierung $A = QS$ konstruiert hat, da Q als Produkt von orthogonalen Matrizen auch eine orthogonale Matrix ist.