7. Großübung

1 QR-Zerlegung

Als QR-Zerlegung wird die Zerlegung

$$A = QR$$

der Matrix $A \in \mathbb{R}^{m \times n}$ in die rechte obere Dreiecksmatrix $R \in \mathbb{R}^{m \times n}$ und die orthogonale Matrix $Q \in \mathbb{R}^{m \times m}$ bezeichnet.

Die Lösung des Gleichungssystems $Ax = b$ kann in der Form

$$\begin{aligned} Ax =& QRx = b \\ \Rightarrow Rx =& Q^{-1}b = Q^T b \end{aligned}$$

durch Rückwärtseinsetzen gewonnen werden. Die rechte Seite $Q^T b$ wird im Laufe der QR-Zerlegung gewonnen. Eine explizite Bestimmung der Matrix Q ist daher i.d.R. nicht erforderlich.

Wir behandeln im Folgenden zwei Verfahren zur Erzeugung der QR-Zerlegung:

- Givens-Rotationen
- Householder-Transformationen

2 Givens-Rotationen

Grundaufgabe

Das Prinzip der Givens-Rotationen beruht darauf, die Spalten von A schrittweise durch Drehungen in die (wahlweise positive oder negative) Richtung der Einheitsvektoren des Koordinatensystems abzubilden, wodurch in den Spaltenvektoren entsprechende Nulleinträge entstehen. Ggfs. wird die rechte Seite b ebenfalls mitbehandelt.

Die Grundaufgabe der Givens-Rotationen besteht folglich darin, eine orthogonale Matrix G zu finden, welche den gegebenen Spaltenvektor $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ in Richtung des Einheitsvektors e^1 abbildet.

$$G \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$$

r entspricht, ggfs. mit abweichendem Vorzeichen, der euklidische Länge des Vektors $\begin{pmatrix} a \\ b \end{pmatrix}$, $r = \pm\sqrt{a^2+b^2}$, $r \in \mathbb{R}$.

Wie im Folgenden gezeigt wird, kann die Matrix G durch die Drehmatrix

$$G = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$

beschrieben werden.

Bemerkung: In der Praxis ist es nicht nötig den Drehwinkel ϕ explizit zu bestimmen. Es müssen nur Werte für c und s bestimmt werden.

Bestimmung von c und s (rechnerische Variante):

$$G = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \Rightarrow G^{-1} = \frac{1}{c^2+s^2} \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \stackrel{!}{=} G^T = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

$$\Rightarrow c^2 + s^2 = det(G) = 1 \ \ (\star)$$

$$\begin{pmatrix} r \\ 0 \end{pmatrix} \stackrel{!}{=} G \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ca + sb \\ -sa + cb \end{pmatrix}$$

$$\Rightarrow s = c\frac{b}{a}$$

$$(\star) \Rightarrow 1 = c^2\left(1 + \frac{b^2}{a^2}\right)$$

$$\Rightarrow c = \frac{a}{\pm\sqrt{a^2+b^2}}$$

$$\Rightarrow s = \frac{b}{\pm\sqrt{a^2+b^2}}$$

$$r = ca + sb = \pm\sqrt{a^2+b^2}$$

Bei der Herleitung wurde $a \neq 0$ angenommen, andernfalls muss statt nach s zunächst nach c umgeformt werden. Die Vorzeichenwahl ($\pm$) ist irrelevant, hier wird ein positives Vorzeichen gewählt.

Anwendung

Die Givens-Rotationsmatrizen $G_{i,k}$ entstehen durch Einbettung ebener Drehungen in $m \times m$-Matrizen (s. Beispiel). Die Anwendung auf den (zur Bestimmung der Rotationsmatrix verwendeten) Spaltenvektor von A erzeugt eine Null in der k-ten Zeile.

Die Givens-Rotationsmatrizen müssen auf alle Spaltenvektoren von A und ggfs. b angewendet werden. Es gilt:

$$G_{i_N,k_N} \cdot \ldots \cdot G_{i_1,k_1} \cdot A = R$$

$$\Rightarrow A = G^T_{i_1,k_1} \cdot \ldots \cdot G^T_{i_N,k_N} \cdot R = QR$$

Der Aufwand der Givens-Rotationen beträgt $\frac{4}{3}n^3$ Operationen für ein vollbesetztes $A \in \mathbb{R}^{n \times n}$.

Beispiel

$$[A|b] = \left[\begin{array}{ccc|c} 3 & -1 & 5 & -2 \\ 4 & 2 & -3 & 1 \\ -2 & 6 & 1 & 3 \end{array}\right]$$

$$a = 3,\ b = 4,\ r = \sqrt{a^2 + b^2} = 5,\ c = \frac{a}{r} = 0.6,\ s = \frac{b}{r} = 0.8$$

$$G_{12} = \begin{pmatrix} 0.6 & 0.8 & 0 \\ -0.8 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\to G_{12}\,[A|b] = \left[\begin{array}{ccc|c} 5 & 1 & 0.6 & -0.4 \\ 0 & 2 & -5.8 & 2.2 \\ -2 & 6 & 1 & 3 \end{array}\right]$$

$$a = 5,\ b = -2,\ r = \sqrt{29},\ c = \frac{a}{r} = 0.9285,\ s = \frac{b}{r} = -0.3714$$

$$G_{13} = \begin{pmatrix} 0.9285 & 0 & -0.3714 \\ 0 & 1 & 0 \\ 0.3714 & 0 & 0.9285 \end{pmatrix}$$

$$\to G_{13}G_{12}\,[A|b] = \left[\begin{array}{ccc|c} 5.385 & -1.3 & 0.1857 & -1.486 \\ 0 & 2 & -5.8 & 2.2 \\ 0 & 5.942 & 1.151 & 2.637 \end{array}\right]$$

$$a = 2,\ b = 5.942,\ r = 6.270,\ c = \frac{a}{r} = 0.3190,\ s = \frac{b}{r} = 0.9478$$

$$G_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.3190 & 0.9478 \\ 0 & -0.9478 & 0.3190 \end{pmatrix}$$

$$\to G_{23}G_{13}G_{12}\,[A|b] = \left[\begin{array}{ccc|c} 5.385 & -1.3 & 0.1857 & -1.486 \\ 0 & 6.270 & -0.7590 & 3.201 \\ 0 & 0 & 5.864 & -1.244 \end{array}\right]$$

Lösen durch Rückwärtseinsetzen:

$$\to x_3 = -0.2121,\ x_2 = 0.4848,\ x_1 = -0.1515$$

Bemerkung: In der Praxis ist das explizite Aufstellen der Rotationsmatrizen G_{ik} nicht erforderlich. Es genügt die Werte für c und s zu kennen, um diese, ohne formale Matrix-Matrix-Multiplikation, anzuwenden.

3 Householder-Transformationen

Grundaufgabe

Das Prinzip der Householder-Transformation ist ähnlich dem der Givens-Rotation. Statt Drehungen werden beim Householder-Verfahren jedoch Spiegelungen verwendet, und die transformierten Spaltenvektoren y sind Elemente des $\mathbb{R}^m$. Die Grundaufgabe besteht darin, eine orthogonale Matrix $Q_V \in \mathbb{R}^{m\times m}$ mit

$$Q_V y = \begin{pmatrix} \pm \parallel y \parallel_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

zu gegebenem $y \in \mathbb{R}^m$ zu finden.

Die Grundaufgabe wird durch die folgdende Darstellung der Transformationsmatrix Q_V erfüllt (Herleitung: Dahmen, Reusken, *Numerik für Ingenieure und Naturwissenschaftler*, S. 103):

$$Q_V = I - \frac{2}{v^T v} v v^T$$

$$v = y + \alpha e^1$$

$$\alpha = sign(y_1) \parallel y \parallel_2$$

$$Q_V y = -\alpha e^1$$

Anwendung

Die QR-Zerlegung wird durch sukzessive Anwendung der Transformationsmatrizen Q_V gewonnen:

$$R = Q_{V_N} \cdot \ldots \cdot Q_{V_1} A = Q^T A$$

Die QR-Zerlegung der Matrix $A \in \mathbb{R}^{n\times n}$ durch Householder-Transformationen benötigt $\frac{2}{3}n^3$ Operationen.

Beispiel

$$[A|b] = \left[\begin{array}{ccc|c} 3 & -1 & 5 & -2 \\ 4 & 2 & -3 & 1 \\ -2 & 6 & 1 & 3 \end{array}\right]$$

$$y = \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix},\ \alpha = sign(y_1) \parallel y \parallel_2 = 1 \cdot \sqrt{3^2 + 4^2 + (-2)^2} = 5.385$$

$$v = y + \alpha e^1 = \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix} + \begin{pmatrix} 5.385 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 8.385 \\ 4 \\ -2 \end{pmatrix}$$

$$\beta := \frac{2}{v^T v} = \frac{2}{(8.385,\ 4,\ -2) \begin{pmatrix} 8.385 \\ 4 \\ -2 \end{pmatrix}} = 2.215 \cdot 10^{-2}$$

$$h := v^T [A|b] = (8.385,\ 4,\ -2) \left[\begin{array}{ccc|c} 3 & -1 & 5 & -2 \\ 4 & 2 & -3 & 1 \\ -2 & 6 & 1 & 3 \end{array} \right] = [45.155,\ -12.385,\ 27.925| -18.77]$$

$$r := \beta v h = 2.215 \cdot 10^{-2} \begin{pmatrix} 8.385 \\ 4 \\ -2 \end{pmatrix} [45.16,\ -12.385,\ 27.925| -18.77]$$

$$\rightarrow r = \left[\begin{array}{ccc|c} 8.385 & -2.30 & 5.186 & -3.486 \\ 4 & -1.097 & 2.474 & -1.663 \\ -2 & 0.549 & -1.237 & 0.8314 \end{array} \right]$$

$$[A|b] := [A|b] - r = \left[\begin{array}{ccc|c} -5.385 & 1.30 & -0.1857 & 0.1486 \\ 0 & 3.097 & -5.474 & 2.663 \\ 0 & 5.451 & 2.237 & 2.169 \end{array} \right]$$

$$y = \begin{pmatrix} 3.097 \\ 5.451 \end{pmatrix},\ \alpha = sign(y_1) \parallel y \parallel_2 = 6.270$$

$$v = y + \alpha e^1 = \begin{pmatrix} 3.097 \\ 5.451 \end{pmatrix} + \begin{pmatrix} 6.270 \\ 0 \end{pmatrix} = \begin{pmatrix} 9.367 \\ 5.451 \end{pmatrix}$$

$$\beta := \frac{2}{v^T v} = 1.70 \cdot 10^{-2}$$

$$h := v^T [A|b] = (9.367,\ 5.451) \left[\begin{array}{cc|c} 3.097 & -5.474 & 2.663 \\ 5.451 & 2.237 & 2.169 \end{array} \right] = [58.7286,\ -39.08|\ 36.76]$$

$$r := \beta v h = 1.70 \cdot 10^{-2} \begin{pmatrix} 9.367 \\ 5.451 \end{pmatrix} [58.7286,\ -39.08|\ 36.76] = \left[\begin{array}{cc|c} 9.367 & -6.233 & 5.864 \\ 5.451 & -3.627 & 3.4126 \end{array} \right]$$

$$[A|b] := [A|b] - r = \left[\begin{array}{ccc|c} -5.385 & 1.30 & -0.1857 & 1.486 \\ 0 & -6.270 & 0.7590 & -3.201 \\ 0 & 0 & 5.864 & -1.244 \end{array} \right]$$

Lösen durch Rückwärtseinsetzen:

$$\rightarrow x_3 = -0.2121,\ x_2 = 0.4848,\ x_1 = -0.1515$$

Bemerkung: In der Praxis ist es nicht erforderlich, den jeweils ersten Spaltenvektor mitzuberechnen. Dieser ist stets $(-\alpha,\ 0, \ldots, 0)^T$. Ferner muss die Matrix r nicht explizit aufgestellt werden. Stattdessen wird direkt die Matrix $[A|b] - r$ berechnet.

Householder-Tableau

Das Householder-Tableau ermöglicht es, die oben durchgeführte Rechnung kompakter aufzuschreiben.

Allgemeine Darstellung

$$
\begin{array}{c|cc}
 & [A|b] & \\
 & & \alpha = \dots \\
\hline
v^T & h = v^T\,(A|b) & \beta = \dots \\
\hline
r = \beta v & H = r\,h & \\
\hline
 & [A|b] - H & \\
\hline
\end{array}
$$

oder kürzer

$$
\begin{array}{c|c}
 & [A|b] \\
\hline
v^T & h = v^T\,(A|b) \\
\hline
r = \beta v & [A|b] - r\,h \\
\hline
\end{array}
$$

Beispiel

$$
\begin{array}{c||c|c}
[A|b]^{(0)} = & \left[\begin{array}{ccc|c} 3 & -1 & 5 & -2 \\ 4 & 2 & -3 & 1 \\ -2 & 6 & 1 & 3 \end{array}\right] & \alpha_1 = 5.385 \\
\hline
(8.385 \quad 4 \quad -2) & \left[\begin{array}{ccc|c} 45.155 & -12.385 & 27.925 & -18.77 \end{array}\right] & \beta_1 = 2.215 \cdot 10^{-2} \\
\hline
\begin{pmatrix} 0.1857 \\ 0.0886 \\ -0.0443 \end{pmatrix} & \left[\begin{array}{ccc|c} 8.385 & -2.30 & 5.186 & -3.486 \\ 4 & -1.097 & 2.474 & -1.663 \\ -2 & 0.549 & -1.237 & 0.8314 \end{array}\right] & \\
\hline
[A|b]^{(1)} = & \left[\begin{array}{ccc|c} -5.385 & 1.30 & -0.1857 & 1.486 \\ 0 & 3.097 & -5.474 & 2.663 \\ 0 & 5.451 & 2.237 & 2.169 \end{array}\right] & \alpha_2 = 6.270 \\
\hline
(\quad 9.367 \quad 5.451) & \left[\begin{array}{ccc|c} & 58.7286 & -39.08 & 36.76 \end{array}\right] & \beta_2 = 1.70 \cdot 10^{-2} \\
\hline
\begin{pmatrix} \\ 0.1592 \\ 0.09267 \end{pmatrix} & \left[\begin{array}{ccc|c} & & & \\ & 9.367 & -6.233 & 5.864 \\ & 5.451 & -3.627 & 3.4126 \end{array}\right] & \\
\hline
[A|b]^{(2)} = & \left[\begin{array}{ccc|c} -5.385 & 1.30 & -0.1857 & 1.486 \\ & -6.270 & 0.7590 & -3.201 \\ & 0 & 5.864 & -1.244 \end{array}\right] & \\
\hline
\end{array}
$$