第四章 矩阵分解

第 11 讲 QR 分解

黄定江

DaSE @ ECNU

djhuang@dase.ecnu.edu.cn

11.1.1 QR 分解

- 矩阵的 QR 分解也称正交三角分解，是一种特殊的三角分解。
- QR 分解在解决最小二乘问题、矩阵特征值的计算等问题中起到重要作用，也是目前计算一般矩阵的全部特征值和特征向量的最有效方法之一。
- 矩阵 $\boldsymbol{A}$ 的 QR 分解可以通过 Gram-Schmidt 正交化、Householder 变换和 Givens 变换等方法实现。

QR 分解

定义 1

设矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}(m \geq n)$，如果存在 m 阶正交矩阵 Q 和 n 阶上三角矩阵 R，使得

$$\boldsymbol{A} = \boldsymbol{Q}\begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{pmatrix},$$

则称之为 $\boldsymbol{A}$ 的 QR 分解或正交三角分解。

在上述定义中，当 $\boldsymbol{A} \in \mathbb{C}^{m\times n}(m \geq n)$ 且 Q 为 m 阶酉矩阵，则称之为 $\boldsymbol{A}$ 的酉三角分解。

QR 分解

定理 1

对任意一个列满秩的实矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}(m \geq n)$，都存在正交三角分解

$$\boldsymbol{A} = \boldsymbol{Q}\begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{pmatrix},$$

其中 $\boldsymbol{Q}$ 为 m 阶正交矩阵，$\boldsymbol{R}$ 具有正的对角元的上三角矩阵；而且当 $m = n$ 且 $\boldsymbol{A}$ 非奇异时，上述分解还是唯一的。

上述定理对于复矩阵也成立，此时 $\boldsymbol{Q}$ 为酉矩阵。

证明.

设 $\boldsymbol{A}$ 是一个列满秩的实矩阵, $\boldsymbol{A}$ 的 n 个列向量为 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, 由于 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ 线性无关，将它们用 Schmidt 正交化方法得标准正交向量 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$ 即

$$\begin{cases} \boldsymbol{a}_1 & = r_{11}\boldsymbol{q}_1 \\ \boldsymbol{a}_2 & = r_{12}\boldsymbol{q}_1 + r_{22}\boldsymbol{q}_2 \\ & \ldots \\ \boldsymbol{a}_n & = r_{1n}\boldsymbol{q}_1 + r_{2n}\boldsymbol{q}_2 + \cdots + r_{nn}\boldsymbol{q}_n \end{cases}$$

其中 $r_{ii} > 0, i = 1, 2, \ldots, n$ 从而有

$$\begin{pmatrix} \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots \boldsymbol{a}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{q}_1, \boldsymbol{q}_2, \ldots \boldsymbol{q}_n \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & \ldots & r_{1n} \\ 0 & r_{22} & r_{23} & \ldots & r_{2n} \\ 0 & 0 & r_{33} & \ldots & r_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & r_{nn} \end{pmatrix}$$

如果给 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots \boldsymbol{q}_n$ 补上 $m-n$ 个标准正交的向量 $\boldsymbol{q}_{n+1}, \boldsymbol{q}_{n+2}, \ldots \boldsymbol{q}_m$ 就有

$$\left(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots \boldsymbol{a}_n\right) = \left(\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots \boldsymbol{q}_m\right) \begin{pmatrix} r_{11} & r_{12} & r_{13} & \ldots & r_{1n} \\ 0 & r_{22} & r_{23} & \ldots & r_{2n} \\ 0 & 0 & r_{33} & \ldots & r_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & r_{nn} \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

令 $\boldsymbol{Q} = \left(\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots \boldsymbol{q}_m\right), \boldsymbol{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \ldots & r_{1n} \\ 0 & r_{22} & r_{23} & \ldots & r_{2n} \\ 0 & 0 & r_{33} & \ldots & r_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & r_{nn} \end{pmatrix}$，则 $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$

再证唯一性。

如果

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R} = \boldsymbol{Q}_1\boldsymbol{R}_1,$$

由此得 $\boldsymbol{Q} = \boldsymbol{Q}_1\boldsymbol{R}_1\boldsymbol{R}^{-1}$，令 $\boldsymbol{D} = \boldsymbol{R}_1\boldsymbol{R}^{-1}$，那么 $\boldsymbol{D}$ 仍为具有正对角元的上三角矩阵。

由于

$$\boldsymbol{I} = \boldsymbol{Q}^T\boldsymbol{Q} = (\boldsymbol{Q}_1\boldsymbol{D})^T(\boldsymbol{Q}_1\boldsymbol{D}) = \boldsymbol{D}^T\boldsymbol{D}$$

即 $\boldsymbol{D}$ 为正交矩阵，因此 $\boldsymbol{D}$ 为单位矩阵（正交上三角为对角阵）

故

$$\boldsymbol{Q} = \boldsymbol{Q}_1\boldsymbol{D} = \boldsymbol{Q}_1, \boldsymbol{R}_1 = \boldsymbol{D}\boldsymbol{R} = \boldsymbol{R}$$

11.1.2 Schmidt 正交化方法

例 1

求下列矩阵的正交三角分解（QR）表达式：

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Schmidt 正交化方法

记 $\boldsymbol{a}_1=(0,1,1)^T \boldsymbol{a}_2=(1,1,0)^T \quad \boldsymbol{a}_3=(1,0,1)^T$

由 Gram-Schmidt 正交化方法。先正交化得

$$\begin{cases} \boldsymbol{b}_1=\boldsymbol{a}_1=(0,1,1)^T \\ \boldsymbol{b}_2=\boldsymbol{a}_2-\frac{\langle \boldsymbol{a}_2,\boldsymbol{b}_1\rangle}{\langle \boldsymbol{b}_1,\boldsymbol{b}_1\rangle}\boldsymbol{b}_1=(1,\frac{1}{2},-\frac{1}{2})^T \\ \boldsymbol{b}_3=\boldsymbol{a}_3-\frac{\langle \boldsymbol{a}_3,\boldsymbol{b}_1\rangle}{\langle \boldsymbol{b}_1,\boldsymbol{b}_1\rangle}\boldsymbol{b}_1-\frac{\langle \boldsymbol{a}_3,\boldsymbol{b}_2\rangle}{\langle \boldsymbol{b}_2,\boldsymbol{b}_2\rangle}\boldsymbol{b}_2=(\frac{2}{3},-\frac{2}{3},\frac{2}{3})^T \end{cases}$$

然后单位化

$$\begin{cases} \boldsymbol{q}_1=\frac{1}{\sqrt{2}}(0,1,1)^T \\ \boldsymbol{q}_2=\frac{1}{\sqrt{6}}(2,1,-1)^T \\ \boldsymbol{q}_3=\frac{1}{\sqrt{3}}(1,-1,1)^T \end{cases}$$

Schmidt 正交化方法

整理得

$$\begin{cases} \boldsymbol{a}_1 = |\boldsymbol{b}_1|\boldsymbol{q}_1 \\ \boldsymbol{a}_2 = \langle \boldsymbol{a}_2, \boldsymbol{q}_1\rangle \boldsymbol{q}_1 + |\boldsymbol{b}_2|\boldsymbol{q}_2 \\ \boldsymbol{a}_3 = \langle \boldsymbol{a}_3, \boldsymbol{q}_1\rangle \boldsymbol{q}_1 + \langle \boldsymbol{a}_3, \boldsymbol{q}_2\rangle \boldsymbol{q}_2 + |\boldsymbol{b}_3|\boldsymbol{q}_3 \end{cases}$$

于是

$$\boldsymbol{Q} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) = \begin{bmatrix} 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$\boldsymbol{R} = \begin{bmatrix} |\boldsymbol{b}_1| & \langle \boldsymbol{a}_2, \boldsymbol{q}_1\rangle & \langle \boldsymbol{a}_3, \boldsymbol{q}_1\rangle \\ 0 & |\boldsymbol{b}_2| & \langle \boldsymbol{a}_3, \boldsymbol{q}_2\rangle \\ 0 & 0 & |\boldsymbol{b}_3| \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{6}}{2} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}$$

那么 $\boldsymbol{A} = \boldsymbol{QR}$ 即为所求表达式。

注记

- 在实际数值计算中，Gram-Schmidt 正交化是数值不稳定的，计算中累积的舍入误差会使最终结果的正交性变得很差；
- 因此常用一种修正的 Gram-Schmidt 正交化方法，它是对经典 Gram-Schmidt 正交化法的修正，使上三角矩阵 $\boldsymbol{R}$ 的元素不是按列，而是按行计算，这时舍入误差将变小。

11.2.1 Householder 变换：定义和性质回顾

定义 2

设 $\boldsymbol{w} \in \mathbb{R}^n$ 满足 $\|\boldsymbol{w}\|_2 = 1$，定义 $\boldsymbol{H} \in \mathbb{R}^{n\times n}$ 为

$$\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}} \tag{1}$$

则称 $\boldsymbol{H}$ 为 *Householder* 变换。

Householder 变换性质回顾

设 $\boldsymbol{H}$ 是由(1)定义的一个 Householder 变换，那么 $\boldsymbol{H}$ 满足 (1) 对称性：$\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{H}$；(2) 正交性：$\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} = \boldsymbol{I}$；(3) 对合性：$\boldsymbol{H}^2 = \boldsymbol{I}$；(4) 反射性：对任意的 $\boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{H}\boldsymbol{x}$ 是 $\boldsymbol{x}$ 关于 $\boldsymbol{w}$ 的垂直超平面的镜像反射；(5) $\mathrm{diag}(\boldsymbol{I}, \boldsymbol{H})$ 也是一个 Householder 矩阵；(6) $\det \boldsymbol{H} = -1$。

Householder 变换：性质

定理 2

设 $\mathbf{0} \neq \boldsymbol{x} \in \mathbb{R}^n$，则可构造单位向量 $\boldsymbol{w} \in \mathbb{R}^n$, 使由(1)定义的 *Householder* 变换 $\boldsymbol{H}$ 满足

$$\boldsymbol{H}\boldsymbol{x} = \alpha \boldsymbol{e}_1, \tag{2}$$

其中 $\alpha = \pm\|\boldsymbol{x}\|_2$。

证明.

由于 $\boldsymbol{Hx} = \left(\boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\right)\boldsymbol{x} = \boldsymbol{x} - 2\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}\right)\boldsymbol{w}$,

故欲使 $\boldsymbol{Hx} = \alpha\boldsymbol{e}_1$，则 $\boldsymbol{w}$ 应为

$$\boldsymbol{w} = \frac{\boldsymbol{x} - \alpha\boldsymbol{e}_1}{\|\boldsymbol{x} - \alpha\boldsymbol{e}_1\|_2}$$

$$\begin{aligned}
\boldsymbol{Hx} &= \boldsymbol{x} - 2\frac{(\boldsymbol{x} - \alpha\boldsymbol{e}_1)(\boldsymbol{x} - \alpha\boldsymbol{e}_1)^{\mathrm{T}}}{\|\boldsymbol{x} - \alpha\boldsymbol{e}_1\|_2^2}\boldsymbol{x} \\
&= \boldsymbol{x} - \frac{2(\boldsymbol{x} - \alpha\boldsymbol{e}_1)^{\mathrm{T}}\boldsymbol{x}}{\|\boldsymbol{x} - \alpha\boldsymbol{e}_1\|_2^2}(\boldsymbol{x} - \alpha\boldsymbol{e}_1) \\
&= \boldsymbol{x} - \frac{2\|\boldsymbol{x}\|_2^2 - 2\alpha\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2 - 2\alpha\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{x} + \alpha^2}(\boldsymbol{x} - \alpha\boldsymbol{e}_1) \\
&= \boldsymbol{x} - \frac{2\alpha^2 - 2\alpha\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{x}}{\alpha^2 - 2\alpha\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{x} + \alpha^2}(\boldsymbol{x} - \alpha\boldsymbol{e}_1) \\
&= \boldsymbol{x} - (\boldsymbol{x} - \alpha\boldsymbol{e}_1) = \alpha\boldsymbol{e}_1
\end{aligned}$$

又因 $\boldsymbol{H}$ 是正交矩阵，必须有

$$\begin{aligned}
\|\boldsymbol{x}\|_2 &= \|\boldsymbol{Hx}\|_2 \\
&= \|\alpha\boldsymbol{e}_1\|_2 \\
&= |\alpha| \cdot \|\boldsymbol{e}_1\|_2 \\
&= |\alpha|
\end{aligned}$$

即 $\alpha = \pm\|\boldsymbol{x}\|_2$。

□

- 上述定理说明，对任意的 $\boldsymbol{x} \in \mathbb{R}^n (\boldsymbol{x} \neq \boldsymbol{0})$，通过适当选取单位向量 $\boldsymbol{w}$，可构造出 Householder 变换矩阵 $\boldsymbol{H}$，使 $\boldsymbol{Hx}$ 的后 $n-1$ 分量变为零。这类似于 Gauss 变换，可以把一个给定向量的若干个指定的分量变为零。
- 另外也说明，可以利用 Householder 变换将任意向量 $\boldsymbol{x}$ 化为与第一自然基向量 $\boldsymbol{e}_1$ 平行的向量（共线）。
- 而且其证明亦告诉我们，可按如下的步骤来构造确定 $\boldsymbol{H}$ 的单位向量 $\boldsymbol{w}$:

  (1) 计算 $\boldsymbol{v} = \boldsymbol{x} \pm \|\boldsymbol{x}\|_2 \boldsymbol{e}_1$;

  (2) 计算 $\boldsymbol{w} = \boldsymbol{v}/\|\boldsymbol{v}\|_2$。
- 此外，在实际计算中，α 取正还是取负根据具体情况来决定。

例 2

用 *Householder* 变换将向量 $\boldsymbol{x} = (0,3,4)^T$ 化为与 $\boldsymbol{e} = (1,0,0)^T$ 平行的向量。

解

由于 $\|\boldsymbol{x}\|_2 = 5$，不妨取 $\alpha = \|\boldsymbol{x}\|_2 = 5$。令

$$\boldsymbol{\omega} = \frac{\boldsymbol{x} - \alpha\boldsymbol{e}}{\|\boldsymbol{x} - \alpha\boldsymbol{e}\|_2} = \frac{1}{5\sqrt{2}}\begin{pmatrix} -5 \\ 3 \\ 4 \end{pmatrix},$$

则

$$\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{\omega}\boldsymbol{\omega}^T = \frac{1}{25}\begin{pmatrix} 0 & 15 & 20 \\ 15 & 16 & -12 \\ 20 & -12 & 91 \end{pmatrix}$$

因此 $\boldsymbol{H}\boldsymbol{x} = 5\boldsymbol{e}$。

11.2.2 基于 Householder 变换的矩阵 QR 分解

利用 Householder 变换求矩阵的 QR 分解的步骤：

[1] 将矩阵 $\boldsymbol{A}$ 按列分块 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_n)$ ，取

$$\boldsymbol{\omega}_1=\frac{\boldsymbol{\alpha}_1-a_1\boldsymbol{e}}{\|\boldsymbol{\alpha}_1-a_1\boldsymbol{e}\|_2},\ a_1=\|\boldsymbol{\alpha}_1\|_2$$

则

$$\boldsymbol{H}_1=\boldsymbol{I}-2\boldsymbol{\omega}_1\boldsymbol{\omega}_1^T$$

那么

$$\boldsymbol{H}_1\boldsymbol{A}=(\boldsymbol{H}_1\boldsymbol{\alpha}_1,\boldsymbol{H}_1\boldsymbol{\alpha}_2,\dots,\boldsymbol{H}_1\boldsymbol{\alpha}_n)=\begin{pmatrix} a_1 & * & \dots & * \\ 0 & & & \\ \vdots & & \boldsymbol{B}_1 & \\ 0 & & & \end{pmatrix}$$

[2] 将矩阵 $\boldsymbol{B}_1$ 按列分块，$\boldsymbol{B}_1=(\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\ldots,\boldsymbol{\beta}_n)$

取

$$\boldsymbol{\omega}_2=\frac{\boldsymbol{\beta}_2-b_2\boldsymbol{e}}{\|\boldsymbol{\beta}_2-b_2\boldsymbol{e}\|_2},\ b_2=\|\boldsymbol{\beta}_2\|_2$$

则

$$\tilde{\boldsymbol{H}}_2=\boldsymbol{I}-2\boldsymbol{\omega}_2\boldsymbol{\omega}_2^T$$

并且令

$$\boldsymbol{H}_2=\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \tilde{\boldsymbol{H}}_2\end{pmatrix}$$

故有

$$\boldsymbol{H}_2(\boldsymbol{H}_1\boldsymbol{A})=\begin{pmatrix}a_1 & * & * & \ldots & *\\ 0 & a_2 & * & \ldots & *\\ 0 & 0 & & & \\ \vdots & \vdots & & \boldsymbol{C}_1 & \\ 0 & 0 & & & \end{pmatrix}$$

依次进行下去，得到第 $n-1$ 个 n 阶的 Householder 矩阵 $\boldsymbol{H}_{n-1}$，使得

$$\boldsymbol{H}_{n-1}\dots\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A}=\begin{pmatrix} a_1 & * & \dots & * \\ & a_2 & \dots & * \\ & & \ddots & \vdots \\ & & & a_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}=\begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{pmatrix}$$

因 $\boldsymbol{H}_i$ 是自逆矩阵，令 $\boldsymbol{Q}=\boldsymbol{H}_1\boldsymbol{H}_2\dots\boldsymbol{H}_{n-1}$，则 $\boldsymbol{A}=\boldsymbol{Q}\begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{pmatrix}$

例 3

已知矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 1\end{pmatrix}$，利用 *Householder* 变换求 $\boldsymbol{A}$ 的 $\boldsymbol{QR}$ 分解。

已知矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 1\end{pmatrix}$，利用 Householder 变换求 $\boldsymbol{A}$ 的 $\boldsymbol{QR}$ 分解。

解

因为 $\boldsymbol{\alpha}_1=(0,0,2)^{\mathrm{T}}$，记 $a_1=\|\boldsymbol{\alpha}_1\|_2=2$，令 $\boldsymbol{w}_1=\frac{\boldsymbol{\alpha}_1-a_1\boldsymbol{e}_1}{\|\boldsymbol{\alpha}_1-a_1\boldsymbol{e}_1\|_2}=\frac{1}{\sqrt{2}}(-1,0,1)^{\mathrm{T}}$，则

$$\boldsymbol{H}_1=\boldsymbol{I}-2\boldsymbol{w}_1\boldsymbol{w}_1^{\mathrm{T}}=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}$$

从而

$$\boldsymbol{H}_1\boldsymbol{A}=\begin{pmatrix}2 & 1 & 1\\ 0 & 4 & -2\\ 0 & 3 & 1\end{pmatrix}$$

已知矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 1\end{pmatrix}$，利用 Householder 变换求 $\boldsymbol{A}$ 的 $\boldsymbol{QR}$ 分解。

解

记 $\boldsymbol{\beta}_2=(4,3)^{\mathrm{T}}$，则 $b_2=\|\boldsymbol{\beta}_2\|_2=5$，令 $\boldsymbol{w}_2=\frac{\boldsymbol{\beta}_2-b_2\boldsymbol{e}_2}{\|\boldsymbol{\beta}_2-b_2\boldsymbol{e}_2\|_2}=\frac{1}{\sqrt{10}}(-1,3)^{\mathrm{T}}$

$$\hat{\boldsymbol{H}}_2=\boldsymbol{I}-2\boldsymbol{w}_2\boldsymbol{w}_2^{\mathrm{T}}=\begin{pmatrix}\frac{4}{5} & \frac{3}{5}\\ \frac{3}{5} & -\frac{4}{5}\end{pmatrix}$$

记

$$\boldsymbol{H}_2=\begin{pmatrix}1 & \boldsymbol{0}^{\mathrm{T}}\\ 0 & \hat{\boldsymbol{H}}_2\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 0 & \frac{3}{5} & -\frac{4}{5}\end{pmatrix}$$

已知矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 1\end{pmatrix}$，利用 Householder 变换求 $\boldsymbol{A}$ 的 $\boldsymbol{QR}$ 分解。

解

则

$$\boldsymbol{H}_2(\boldsymbol{H}_1\boldsymbol{A})=\begin{pmatrix}2 & 1 & 1\\ 0 & 5 & -1\\ 0 & 0 & -2\end{pmatrix}=\boldsymbol{R}$$

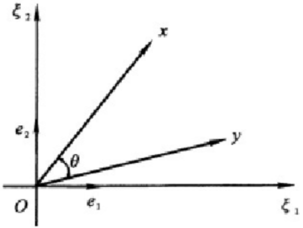
取

$$\boldsymbol{Q}=\boldsymbol{H}_1\boldsymbol{H}_2=\begin{pmatrix}0 & \frac{3}{5} & -\frac{4}{5}\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 1 & 0 & 0\end{pmatrix}$$

则 $\boldsymbol{A}=\boldsymbol{QR}$。

11.3.1 Givens 变换：Givens 矩阵回顾

在平面解析几何中，已知使向量 $\boldsymbol{x}$ 逆时针旋转角度 θ 后变为向量 $\boldsymbol{y}$ 的旋转变换（见下图）如下：



$$\boldsymbol{y}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\boldsymbol{x}=\boldsymbol{T}\boldsymbol{x}$$

其中 $\boldsymbol{T}$ 是正交矩阵，称为平面旋转矩阵。将其推广到一般的 n 维空间中，可以得到初等旋转变换，也称为 Givens 变换。Givens 矩阵是正交矩阵，而且其行列式为 1。

Givens 变换：Givens 矩阵回顾

定义 3

设 $c, s \in \mathbb{R}$，$c^2+s^2=1$，$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 n 维向量空间 $\mathbb{R}^n$ 的一个标准正交基，记 n 阶矩阵：

$$\boldsymbol{T}_{kl} = \boldsymbol{I} + s(\boldsymbol{e}_k\boldsymbol{e}_l^{\mathrm{T}} - \boldsymbol{e}_l\boldsymbol{e}_k^{\mathrm{T}}) + (c-1)(\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}} + \boldsymbol{e}_l\boldsymbol{e}_l^{\mathrm{T}}) = \begin{array}{cc} \begin{bmatrix} 1 & & \vdots & & \vdots & & \\ & \ddots & \vdots & & \vdots & & \\ \cdots & \cdots & c & \cdots & s & \cdots & \cdots \\ & & \vdots & & \vdots & & \\ \cdots & \cdots & -s & \cdots & c & \cdots & \cdots \\ & & \vdots & & \vdots & \ddots & \\ & & \vdots & & \vdots & & 1 \end{bmatrix} & \begin{array}{c} \\ \\ k \\ \\ l \\ \\ \\ \end{array} \\ \begin{array}{ccccccc} \quad & \quad & k & \quad\quad & l & \quad & \quad \end{array} & \end{array}$$

称 $\boldsymbol{T}_{kl}$ 为 *Givens* 矩阵。由 $\boldsymbol{T}_{kl}$ 所确定的线性变换称为 *Givens* 变换或初等旋转变换。

定理 3

对于任意向量 $x \in \mathbb{R}^n$，存在 Givens 变换 $\boldsymbol{T}_{kl}$ 使得 $\boldsymbol{T}_{kl}\boldsymbol{x}$ 的第 l 个分量为 0，第 k 个分量为非负实数，其余分量不变。

证明.

记 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$, $\boldsymbol{T}_{kl}x = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$

由 Givens 矩阵的定义可得

$$\begin{cases} y_k & = cx_k + sx_l \\ y_l & = -sx_k + cx_l \\ y_j & = x_j, (j \neq k, l) \end{cases}$$

(i) 当 $|x_k|^2 + |x_l|^2 = 0$ 时，取 $c = 1, s = 0$，则 $\boldsymbol{T}_{kl} = \boldsymbol{I}$, 此时

$$y_k = y_l = 0, y_j = x_j (j \neq k, l)$$

结论成立

(ii) 当 $|x_k|^2+|x_l|^2\neq 0$ 时，取

$$c=\frac{x_k}{\sqrt{|x_k|^2+|x_l|^2}},\ s=\frac{x_l}{\sqrt{|x_k|^2+|x_l|^2}},$$

则

$$\begin{cases} y_k &= \frac{x_k^2}{\sqrt{|x_k|^2+|x_l|^2}}+\frac{x_l^2}{\sqrt{|x_k|^2+|x_l|^2}}=\sqrt{|x_k|^2+|x_l|^2}>0 \\ y_l &= -\frac{x_kx_l}{\sqrt{|x_k|^2+|x_l|^2}}+\frac{x_lx_k}{\sqrt{|x_k|^2+|x_l|^2}}=0 \\ y_j &= x_j, (j\neq k, l) \end{cases}$$

结论成立

推论 1

给定一个向量 $x \in \mathbb{R}^n$，则存在一组 *Givens* 矩阵 $\boldsymbol{T}_{12}, \boldsymbol{T}_{13}, \ldots, \boldsymbol{T}_{1n}$，使得

$$\boldsymbol{T}_{1n} \ldots \boldsymbol{T}_{13} \boldsymbol{T}_{12} x = \|\boldsymbol{x}\|_2 \boldsymbol{e}_1,$$

称为用 *Givens* 变换化向量 $\boldsymbol{x} \in \mathbb{R}^n$ 与第一自然基向量 $\boldsymbol{e}_1$ 共线。

证明.

设 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$，由定理 3 知存在 Givens 矩阵 $\boldsymbol{T}_{12}$，使得

$$\boldsymbol{T}_{12}\boldsymbol{x} = (\sqrt{|x_1|^2 + |x_2|^2}, 0, x_3, \ldots, x_n)^{\mathrm{T}}$$

对于 $\boldsymbol{T}_{12}\boldsymbol{x}$ 又存在 Givens 矩阵 $\boldsymbol{T}_{13}$ 使得

$$\boldsymbol{T}_{13}(\boldsymbol{T}_{12}\boldsymbol{x}) = (\sqrt{|x_1|^2 + |x_2|^2 + |x_3|^2}, 0, 0, x_4, \ldots, x_n)^{\mathrm{T}}$$

依此继续下去，可以得出

$$\boldsymbol{T}_{1n} \ldots \boldsymbol{T}_{13} \boldsymbol{T}_{12}\boldsymbol{x} = (\sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}, 0, 0, \ldots, 0)^{\mathrm{T}} = \|\boldsymbol{x}\|_2 \boldsymbol{e}_1$$

$\square$

例 4

用 *Givens* 变换化向量 $x=(1,2,2)$ 与第一自然基向量共线

解

由于 $x_1=1, x_2=2, \sqrt{|x_1|^2+|x_2|^2}=\sqrt{5}$ 取

$$c_1=\frac{1}{\sqrt{5}}, s_1=\frac{2}{\sqrt{5}}$$

构造 *Givens* 矩阵

$$\boldsymbol{T}_{12}=\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

故

$$\boldsymbol{T}_{12}\boldsymbol{x}=\begin{pmatrix} \sqrt{5} \\ 0 \\ 2 \end{pmatrix}$$

对于 $\boldsymbol{T}_{12}\boldsymbol{x}$ 取

$$c_2 = \frac{\sqrt{5}}{3}, s_2 = \frac{2}{3}$$

则

$$\boldsymbol{T}_{13} = \begin{pmatrix} \frac{\sqrt{5}}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \\ -\frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix}, \boldsymbol{T}_{13}\boldsymbol{T}_{12}\boldsymbol{x} = 3\boldsymbol{e}_1$$

11.3.2 基于 Givens 变换的矩阵 QR 分解

利用 Givens 变换求矩阵 QR 分解的步骤

先将矩阵 $\boldsymbol{A}$ 按列分块，

$$\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)$$

[1] 对于 $\boldsymbol{\alpha}_1$ 存在一组 Givens 矩阵 $\boldsymbol{T}_{12}, \boldsymbol{T}_{13}, \ldots, \boldsymbol{T}_{1n}$ 使得

$$\boldsymbol{T}_{1n} \ldots \boldsymbol{T}_{13} \boldsymbol{T}_{12} \boldsymbol{\alpha}_1 = \|\boldsymbol{\alpha}_1\|_2 \boldsymbol{e}_1$$

于是

$$\boldsymbol{T}_{1n} \ldots \boldsymbol{T}_{13} \boldsymbol{T}_{12} \boldsymbol{A} = \begin{pmatrix} a_1 & * \\ 0 & \boldsymbol{B}_1 \end{pmatrix}, a_1 = \|\boldsymbol{\alpha}_1\|_2$$

利用 Givens 变换求矩阵 QR 分解的步骤

[2] 将矩阵 $\begin{pmatrix} * \\ \boldsymbol{B}_1 \end{pmatrix}$ 按列分块

$$\begin{pmatrix} * \\ \boldsymbol{B}_1 \end{pmatrix} = \begin{pmatrix} * & * & \cdots & * \\ \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3 & \cdots & \boldsymbol{\beta}_n \end{pmatrix}$$

又存在一组 Givens 矩阵 $\boldsymbol{T}_{23}, \boldsymbol{T}_{24}, \ldots, \boldsymbol{T}_{2n}$ 使得

$$\boldsymbol{T}_{2n} \ldots \boldsymbol{T}_{24} \boldsymbol{T}_{23} \begin{pmatrix} * \\ \boldsymbol{\beta}_2 \end{pmatrix} = (*, b_2, 0, \ldots, 0)^{\mathrm{T}}, b_2 = \|\beta_2\|_2$$

因此

$$\boldsymbol{T}_{2n}\dots\boldsymbol{T}_{24}\boldsymbol{T}_{23}\boldsymbol{T}_{1n}\dots\boldsymbol{T}_{13}\boldsymbol{T}_{12}\boldsymbol{A}=\begin{pmatrix} a_1 & * & * & \dots & * \\ 0 & b_1 & * & \dots & * \\ 0 & 0 & & \boldsymbol{C}_2 & \end{pmatrix}$$

依次进行下去得到

$$\boldsymbol{T}_{n-1,n}\dots\boldsymbol{T}_{2n}\dots\boldsymbol{T}_{23}\boldsymbol{T}_{1n}\dots\boldsymbol{T}_{12}\boldsymbol{A}=\begin{pmatrix} a_1 & * & \dots & * \\ & a_2 & \dots & * \\ & & \ddots & \vdots \\ & & & a_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}=\begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{pmatrix}$$

[3] 令 $\boldsymbol{Q}=\boldsymbol{T}_{12}^T\dots\boldsymbol{T}_{1n}^T\boldsymbol{T}_{23}^T\dots\boldsymbol{T}_{2n}^T\dots\boldsymbol{T}_{n-1,n}^T$，则 $\boldsymbol{A}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{R}\\\boldsymbol{O}\end{pmatrix}$

说明

利用 Givens 变换进行 QR 分解，需要作 $\frac{n(n-1)}{2}$ 个初等旋转矩阵的连乘积，当 n 较大时，计算量较大，因此常用镜像变换来进行 QR 分解。

例 5

已知矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 1\end{pmatrix}$，利用 *Givens* 变换求 $\boldsymbol{A}$ 的 $\boldsymbol{QR}$ 分解。

解

因为 $a_{11}=0, a_{31}=2$，取 $c=0, s=1$，构造

$$\boldsymbol{T}_{13}=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ -1 & 0 & 0\end{pmatrix}$$

则有

$$\boldsymbol{A}^{(1)}=\boldsymbol{T}_{13}\boldsymbol{A}=\begin{pmatrix}2 & 1 & 1\\ 0 & 4 & -2\\ 0 & -3 & -1\end{pmatrix}$$

解

因为 $a_{22}^{(1)} = 4, a_{32}^{(1)} = -3$，取 $c = \frac{4}{5}, s == -\frac{3}{5}$，构造

$$\boldsymbol{T}_{23} = \begin{pmatrix} 1 & & \\ & \frac{4}{5} & -\frac{3}{5} \\ & \frac{3}{5} & \frac{4}{5} \end{pmatrix}$$

则

$$\boldsymbol{A}^{(2)} = \boldsymbol{T}_{23}\boldsymbol{A}^{(1)} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 5 & -1 \\ 0 & 0 & -2 \end{pmatrix} = \boldsymbol{R}$$

$$\boldsymbol{Q} = \boldsymbol{T}_{13}^T \boldsymbol{T}_{23}^T = \begin{pmatrix} 0 & \frac{3}{5} & -\frac{4}{5} \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 1 & 0 & 0 \end{pmatrix}$$

则 $\boldsymbol{A} = \boldsymbol{QR}$。

本讲小结

QR 分解

- Gram-Schmidt 正交化
- Householder 变换
- Givens 变换

对于复矩阵，也有相应的上述三种方法。

将用于最小二乘问题、特征值的计算求解等！