第七章-特征值和特征向量

(2)：QR 分解和 QR 算法

目录

幂法及其变式可以求解特定的特征值和特征向量。

目标：同时求解所有的特征值和特征向量。

目录

QR 分解

定理 (QR 分解)

- 对于任何的非奇异 n 阶矩阵 $\boldsymbol{A}$，则 $\boldsymbol{A}$ 可分解为 $\boldsymbol{A}=\boldsymbol{QR}$；其中 $\boldsymbol{Q}$ 为正交矩阵（即 $\boldsymbol{Q}^T=\boldsymbol{Q}^{-1}\leftrightarrow\boldsymbol{QQ}^T=\boldsymbol{I}$, $\boldsymbol{Q}^T\boldsymbol{Q}=\boldsymbol{I}$），$\boldsymbol{R}$ 为上三角矩阵。
- 当 $\boldsymbol{R}$ 的对角元素都为正数时，该 QR 分解唯一。

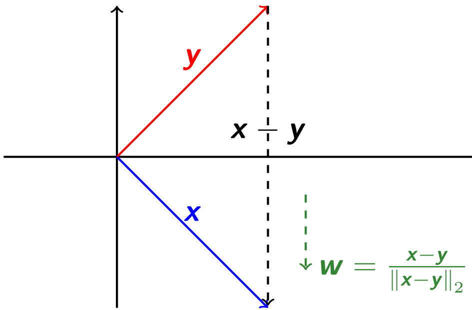
接下来，我们会学习如何找到这样的 QR 分解。

探究关系 $\boldsymbol{Q}^T\boldsymbol{A} = \boldsymbol{R}$ 的第一列，我们至少需要找到某个 $\boldsymbol{Q}$ 使得

$$\boldsymbol{Q}^T\boldsymbol{a}_1 = \begin{bmatrix} r_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$，其中 $\boldsymbol{a}_1$ 是矩阵 $\boldsymbol{A}$ 的第一列。

问题： $|r_1|$ 和 $\|\boldsymbol{a}_1\|_2$ 需要有什么关系吗?

我们探究一般的情况，如何通过一个正交矩阵实现变换 $\boldsymbol{Qx} = \boldsymbol{y}$，其中 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是给定的向量，且 $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2$



我们观察到：

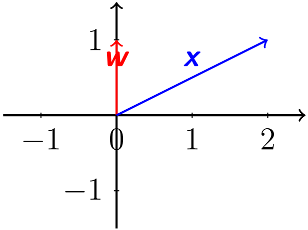
$$\boldsymbol{y} = \boldsymbol{x} - (\boldsymbol{x} - \boldsymbol{y}) = \boldsymbol{x} - \|\boldsymbol{x} - \boldsymbol{y}\|\boldsymbol{w} = \boldsymbol{x} - 2\langle \boldsymbol{w}, \boldsymbol{x}\rangle \boldsymbol{w} = \left(\boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^T\right)\boldsymbol{x}.$$

练习： 请验证 $\boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^T$ 是对称的正交矩阵。

定义 (初等反射阵)

假设向量 $\|\boldsymbol{w}\|_2 = 1$，矩阵 $\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^T$ 称为初等反射阵，又记为 $\boldsymbol{H}(\boldsymbol{w})$。

练习： 请在图中大致画出 $\boldsymbol{H}(\boldsymbol{w})\boldsymbol{x}$



定理

初等反射阵 $\boldsymbol{H}$ 为对称阵 $(\boldsymbol{H}^\top = \boldsymbol{H})$，正交阵 $(\boldsymbol{H}^\top\boldsymbol{H} = \boldsymbol{I})$ 和对合阵 $(\boldsymbol{H}^2 = \boldsymbol{I})$。

总结

定理

设 $\boldsymbol{x} \neq \boldsymbol{y} \in \mathbb{R}^n$，且 $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2$，则存在一个初等反射阵 $\boldsymbol{H}$，使得 $\boldsymbol{Hx} = \boldsymbol{y}$。

证明：取 $\boldsymbol{w} = \frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2}$

定理 (推论)

设 $\boldsymbol{x} \in \mathbb{R}^n$，令 $\sigma = \pm\|\boldsymbol{x}\|_2$，且 $\boldsymbol{x} \neq -\sigma\boldsymbol{e}_1$，则存在一个初等反射阵 $\boldsymbol{H} = \boldsymbol{I} - 2\frac{\boldsymbol{u}\boldsymbol{u}^T}{\|\boldsymbol{u}\|_2^2}$，使得 $\boldsymbol{Hx} = -\sigma\boldsymbol{e}_1$，其中 $\boldsymbol{u} = \boldsymbol{x} + \sigma\boldsymbol{e}_1$。

证明：在最上面的定理中，取 $\boldsymbol{y} = -\sigma\boldsymbol{e}_1$。

提示： 为避免有效数字的缺失，我们会取 $\sigma = \mathrm{sign}(x_1)\|\boldsymbol{x}\|_2$；但实际情况中，除非 $\boldsymbol{x}$ 平行于 $\boldsymbol{e}_1$，否则 $\sigma = \pm\|\boldsymbol{x}\|_2$ 都是可以的。

考虑矩阵 $\boldsymbol{A}$ 的列 $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \ldots & \boldsymbol{A}_n \end{pmatrix}$ 则

$$\boldsymbol{HA} = \begin{pmatrix} \boldsymbol{HA}_1 & \boldsymbol{HA}_2 & \cdots & \boldsymbol{HA}_n \end{pmatrix}$$

$$= \begin{pmatrix} -\sigma_1 & & & \\ 0 & & & \\ \vdots & \boldsymbol{HA}_2 & \cdots & \boldsymbol{HA}_n \\ 0 & & & \end{pmatrix} = \left(\begin{array}{c|ccc} -\sigma_1 & \star & \cdots & \star \\ \hline 0 & & & \\ \vdots & & \boldsymbol{A}^{(2)} & \\ 0 & & & \end{array}\right)$$

若我们希望系数为 σ_1，则可以通过

$$(-\boldsymbol{H})\boldsymbol{A} = \begin{pmatrix} \sigma_1 & & & \\ 0 & & & \\ \vdots & -\boldsymbol{HA}_2 & \cdots & -\boldsymbol{HA}_n \\ 0 & & & \end{pmatrix} = \left(\begin{array}{c|ccc} \sigma_1 & \star & \cdots & \star \\ \hline 0 & & & \\ \vdots & & \boldsymbol{A}^{(2)} & \\ 0 & & & \end{array}\right)$$

对于 $\boldsymbol{A}^{(2)}$，我们可以重复过程找到某个对称且正交矩阵 $\boldsymbol{H}^{(2)}$，使得

$$\boldsymbol{H}^{(2)}\boldsymbol{A}^{(2)} = \left(\begin{array}{c|ccc} -\sigma_2 & \star & \cdots & \star \\ \hline 0 & & & \\ \vdots & & \boldsymbol{A}^{(3)} & \\ 0 & & & \end{array}\right)$$

因此，

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & & & & \\ 0 & & \boldsymbol{H}^{(2)} & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix}}_{\widetilde{\boldsymbol{H}}^{(2)}} \left(\begin{array}{c|cccc} -\sigma_1 & \star & \cdots & \cdots & \star \\ \hline 0 & & & & \\ \vdots & & \boldsymbol{A}^{(2)} & & \\ \vdots & & & & \\ 0 & & & & \end{array}\right) = \begin{pmatrix} -\sigma_1 & \star & \cdots & \cdots & \star \\ 0 & -\sigma_2 & \star & \cdots & \star \\ 0 & 0 & & & \\ 0 & \vdots & & \boldsymbol{A}^{(3)} & \\ 0 & 0 & & & \end{pmatrix}$$

通过不断继续，可得 $\widetilde{\boldsymbol{H}}^{(n)}\widetilde{\boldsymbol{H}}^{(n-1)}\cdots\widetilde{\boldsymbol{H}}^{(2)}\widetilde{\boldsymbol{H}}^{(1)}\boldsymbol{A}=\boldsymbol{R}$ 一个上三角矩阵，其中 $\widetilde{\boldsymbol{H}}^{(k)}$ 为第 k 层基于 Householder 构造的对称正交阵。在这个过程中，我们可以保证 $\boldsymbol{R}_{ii}\geq 0$。

最后，令 $\boldsymbol{Q}^{-1}=\widetilde{\boldsymbol{H}}^{(n)}\widetilde{\boldsymbol{H}}^{(n-1)}\cdots\widetilde{\boldsymbol{H}}^{(2)}\widetilde{\boldsymbol{H}}^{(1)}$，则 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$。

练习： 请验证这样定义出来的矩阵 $\boldsymbol{Q}$ 为正交矩阵，且

$$\boldsymbol{Q}=\widetilde{\boldsymbol{H}}^{(1)}\widetilde{\boldsymbol{H}}^{(2)}\cdots\widetilde{\boldsymbol{H}}^{(n-1)}\widetilde{\boldsymbol{H}}^{(n)}$$

练习： 请计算矩阵 $\boldsymbol{A}$ 的 QR 分解，其中 $\boldsymbol{A}=\begin{bmatrix} 1 & -2 & 1 \\ 2 & 2 & 1 \\ -2 & 1 & 0 \end{bmatrix}$

答案：$\boldsymbol{Q}=\begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \end{bmatrix}$, $\boldsymbol{R}=\begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

计算复杂度

矩阵 $\boldsymbol{Q}$ 可由如下算法实现:

$$\boldsymbol{Q} = \boldsymbol{I}_n$$
$$\boldsymbol{Q} = \boldsymbol{Q}\widetilde{\boldsymbol{H}}^{(k)}, \qquad k = 1, 2, \cdots, n$$

这个过程中，计算量为 $\mathcal{O}(n^2 + n(n-1) + \cdots + n) = \mathcal{O}(n^3)$

计算上三角矩阵 $\boldsymbol{R}$ 的复杂度为 $\mathcal{O}(n^2 + (n-1)^2 + \cdots + 1) = \mathcal{O}(n^3)$

总体而言，基于 Householder 的 QR 算法的复杂度为 $\mathcal{O}(n^3)$。

目录

QR 算法

给定 $\boldsymbol{A} \overset{分解}{=} \boldsymbol{Q}\boldsymbol{R}$，我们取 $\boldsymbol{B} \overset{定义}{=} \boldsymbol{R}\boldsymbol{Q} = \boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q}$

我们可以观察到 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 具有相同的特征值。

$$
\begin{array}{ll}
\boldsymbol{A} \equiv \boldsymbol{A}_1 \overset{分解}{=} \boldsymbol{Q}_1\boldsymbol{R}_1 & \\
\boldsymbol{A}_2 \overset{定义}{=} \boldsymbol{R}_1\boldsymbol{Q}_1 = \boldsymbol{Q}_1^T\boldsymbol{A}_1\boldsymbol{Q}_1 & \boldsymbol{A}_2 \overset{分解}{=} \boldsymbol{Q}_2\boldsymbol{R}_2 \\
\boldsymbol{A}_3 \overset{定义}{=} \boldsymbol{R}_2\boldsymbol{Q}_2 & \boldsymbol{A}_3 \overset{分解}{=} \boldsymbol{Q}_3\boldsymbol{R}_3 \\
\vdots & \\
 & \boldsymbol{A}_k \overset{分解}{=} \boldsymbol{Q}_k\boldsymbol{R}_k \\
\boldsymbol{A}_{k+1} \overset{定义}{=} \boldsymbol{R}_k\boldsymbol{Q}_k & \cdots
\end{array}
$$

定理 (基本性质)

令 $\widetilde{\boldsymbol{Q}}_k = \boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k$，$\widetilde{\boldsymbol{R}}_k = \boldsymbol{R}_k\boldsymbol{R}_{k-1}\cdots\boldsymbol{R}_1$，则

- $\boldsymbol{A}_{k+1}$ 相似于 $\boldsymbol{A}_k$，即 $\boldsymbol{A}_{k+1} = \boldsymbol{Q}_k^T\boldsymbol{A}_k\boldsymbol{Q}_k$
- $\boldsymbol{A}_{k+1} = \widetilde{\boldsymbol{Q}}_k^T\boldsymbol{A}_1\widetilde{\boldsymbol{Q}}_k$
- $\boldsymbol{A}^k$ 的 QR 分解为 $\boldsymbol{A}^k = \widetilde{\boldsymbol{Q}}_k\widetilde{\boldsymbol{R}}_k$

证明该结论将作为练习。

QR 算法的收敛性

定理 (定理 9.18 QR 算法的收敛性)

设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$,

1° 如果 $\boldsymbol{A}$ 的特征值满足: $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$;

2° $\boldsymbol{A}$ 满足一些其他条件:【$\boldsymbol{A}$ 有标准形 $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}$ 其中 $\boldsymbol{D} = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$, 且设 $\boldsymbol{X}^{-1}$ 有 LU 分解】

则由 QR 算法产生的 $\{\boldsymbol{A}_k\}_{k=1}^{\infty}$ 收敛于上三角

$$\boldsymbol{A}_k \xrightarrow{\text{当}\,k\to\infty} \begin{pmatrix} \lambda_1 & \times & \cdots & \times \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & \times \\ & & & \lambda_n \end{pmatrix}.$$

对于对称矩阵的 QR 算法

定理 (定理 9.19)

若对称阵 $\boldsymbol{A}$ 满足定理 9.18 的条件, 则由 QR 算法产生的 $\{\boldsymbol{A}_k\}$ 收敛于对角阵。

目录

总结

- 知道 QR 分解是什么
- 能通过 Householder 方法计算 QR 分解，并了解其运算复杂度为 $\mathcal{O}(n^3)$
- 知道 QR 算法是什么，以及定理 9.18 和 9.19 的结论