

矩阵特征值和特征向量计算

本章主要学习如何通过数值方法求解矩阵的特征值和特征向量，主要介绍了幂法和反幂法。对于矩阵特征值 λ 有：

$$\mathbf{Ax} = \lambda \mathbf{x}$$

幂法

对于矩阵 \mathbf{A} ，假设其存在 n 个特征值和对应的特征向量 \mathbf{x}_i ：

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$$

给定任意初始向量 \mathbf{u}_0 ，能写成各个特征向量的线性组合：

$$\mathbf{u}_0 = a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \cdots + a_n \mathbf{x}_n$$

那么给定下述迭代方程，展开可以发现：

$$\begin{aligned}\mathbf{u}_k &= \mathbf{Au}_{k-1} = \mathbf{A}^2 \mathbf{u}_{k-2} = \cdots = \mathbf{A}^k \mathbf{u}_0 \\ &= \lambda_1^k [a_1 \mathbf{x}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{x}_2 + a_3 \left(\frac{\lambda_3}{\lambda_1} \right)^k \mathbf{x}_3 + \cdots]\end{aligned}$$

在 k 足够大的情况下， λ_2/λ_1 等项趋于零：

$$\mathbf{u}_k \approx \lambda_1^k \alpha_1 \mathbf{x}_1$$

这样我们可以有基本的幂法迭代方法：

$$\begin{aligned}\mathbf{y}_{k-1} &= \frac{\mathbf{u}_{k-1}}{\|\mathbf{u}_{k-1}\|} \\ \mathbf{u}_k &= \mathbf{Ay}_{k-1} \\ \beta_k &= \mathbf{y}_{k-1}^T \mathbf{u}_k\end{aligned}$$

其中 β_k 为当前求解迭代步解出的 λ_1 特征值， \mathbf{y}_{k-1} 为解出的特征向量。

幂法的迭代收敛速度取决于 $|\frac{\lambda_m}{\lambda_1}|$ 的比值，这个比值越小，收敛速度越快。

反幂法

矩阵非奇异的情况下有：

$$\mathbf{Ax}_i = \lambda_i \mathbf{x}_i \rightarrow \mathbf{A}^{-1} \mathbf{x} = \frac{1}{\lambda_i} \mathbf{x}_i$$

利用上式右侧的格式带入到幂法的迭代方程中，可以得到：

$$\begin{aligned}\mathbf{y}_{k-1} &= \frac{\mathbf{u}_{k-1}}{\|\mathbf{u}_{k-1}\|} \\ \mathbf{u}_k &= \mathbf{A}^{-1} \mathbf{y}_{k-1} \\ \beta_k &= \mathbf{y}_{k-1}^T \mathbf{u}_k\end{aligned}$$

这即为反幂法，可以求解得到矩阵最小的特征值，注意若使用上述迭代方法，求解的特征值为 $\lambda_i = 1/\beta_k$ 。

上述的两个方法在实际的计算求解中发现了 β_k 在求解过程中可能最终和实际的特征值符号相反，和豆包沟通后发现可能需要用更可靠的 Rayleigh 商，即上述的迭代过程使用这个迭代方法，而且反幂求解的特征值无需再求倒数。

$$\beta_k = \frac{\mathbf{y}_k^T \mathbf{A} \mathbf{y}_k}{\mathbf{y}_k^T \mathbf{y}_k}$$

在实际过程中，可以采用原点平移的方法求解靠近某个 p 值的特征值，这是由于：

$$\begin{aligned}\mathbf{Ax} &= \lambda \mathbf{x} \\ (\mathbf{A} - p\mathbf{I})\mathbf{x} &= (\lambda - p)\mathbf{x}\end{aligned}$$

这样，只需要将 $\mathbf{A} - p\mathbf{I}$ 作为新的矩阵 \mathbf{A} 带入反幂法求解即可。

上述三种求解矩阵特征值和特征向量的方法写在了Power_Method.py中