计算方法

Ch08 矩阵特征值和特征向量

程勇

buctcourse@163.com

http://www.buct.edu.cn

Dept. of Computer
Beijing University of Chemical Technology

Sept. 19, 2019

提纲

本章提纲

基本概念

A 为 n 阶矩阵，若数 λ 和 n 维非 0 列向量 x 满足 $Ax=\lambda x$，那么数 λ 称为 A 的特征值，x 称为 A 的对应于特征值 λ 的特征向量。式 $Ax=\lambda x$ 也可写成 $(A-\lambda E)x=0$，并且 $|\lambda E-A|$ 叫做 A 的特征多项式。当特征多项式等于 0 的时候，称为 A 的特征方程，特征方程是一个齐次线性方程组，求解特征值的过程其实就是求解特征方程的解。

$$Ax=\lambda x \Rightarrow Ax=\lambda Ex \Rightarrow (\lambda E-A)x=0$$

$$|\lambda E-A|=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{m1} & -a_{11} & \dots & \lambda-a_{mn} \end{vmatrix}=0$$

特征值和特征向量例子

计算 A 的特征值和特征向量。

$$A = \begin{pmatrix} 4 & 2 & -5 \\ 6 & 4 & -9 \\ 5 & 3 & -7 \end{pmatrix}$$

计算如下行列式得特征多项式：

$$|\lambda E - A| = \begin{vmatrix} \lambda - 4 & -2 & 5 \\ -6 & \lambda - 4 & 9 \\ -5 & -3 & \lambda + 7 \end{vmatrix} = 0$$

$$\begin{aligned}
&(\lambda - 4)(\lambda - 4)(\lambda + 7) + (-2) * 9 * (-5) + 5 * (-6) * (-3) \\
&- 5 * (\lambda - 4) * (-5) - (-2) * (-6) * (\lambda + 7) \\
&- (\lambda - 4) * 9 * (-3) = 0
\end{aligned}$$

特征值和特征向量例子 (续)

进一步化简可得：

$$\lambda^2 * (\lambda - 1) = 0$$

得到特征值：

$$\lambda_1 = 1, \lambda_2 = \lambda_3 = 0$$

当 $\lambda_1 = 1$ 时，可求得特征向量：

$$\xi_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

同理，当 $\lambda_2 = \lambda_3 = 0$ 时，可求得特征向量：

$$\xi_2 = \xi_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

乘幂法

设 $\mathbf{A}$ 是 n 阶矩阵，如果数 λ 和 n 维非零列向量 $\mathbf{x}$ 使关系式：

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x},$$

成立，则称数 λ 为方阵 $\mathbf{A}$ 的特征值，非零向量 $\mathbf{x}$ 称为 $\mathbf{A}$ 的对应于特征值 λ 的特征向量。

特征值问题是对方阵而言的，对应的特征向量 $\mathbf{x} \neq 0$；n 阶方阵 $\mathbf{A}$ 的特征值，就是使其次线性方程组 $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$ 有非零解的 λ 值，即满足方程 $|\mathbf{A} - \lambda\mathbf{I}| = 0$ 的 λ 都是矩阵 $\mathbf{A}$ 的特征值。

$$|\mathbf{A} - \mathbf{I}| = 0 \Leftrightarrow \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0$$

乘幂法 (续)

一般计算方法是第一步先求解特征值 λ，求 $\mathbf{A}$ 的特征值 λ 就是求 $|\mathbf{A}-\lambda\mathbf{I}|=0$ 的根；接下来再求解特征向量 $\mathbf{x}$，求 $\mathbf{A}$ 的相应于 λ 的特征向量就是求 $(\mathbf{A}-\lambda\mathbf{I})\mathbf{x}=0$ 的非零解向量。求矩阵 $\mathbf{A}$ 的特征值及特征向量问题就转化为求解多项式方程以及齐次线性方程组的通解问题。

不过在 n 非常大时，直接求解特征值及其对应的特征向量开销会很大，因此可以用乘幂法解其数值。假定 n 阶矩阵 $\mathbf{A}$ 对应的 n 个特征值按模从大到小的排序为：

$$|\lambda_1|\geq|\lambda_2|\geq|\lambda_3|\geq\cdots\geq|\lambda_n|,$$

称 λ_1 为 A 的按模最大特征值 (主特征值)。

乘幂法 (续)

任取非零向量 $\mathbf{x}_0 = \left(x_1^{(0)}, x_x^{(0)}, \cdots, x_n^{(0)}\right)$。特征向量构成 n 维空间的一组基。任取的初始向量 $\mathbf{x}_0$ 由它们的线性组合给出：

$$\mathbf{x}_0 = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n$$

假设 $a_1 \neq 0$，由此构造向量序列 $\mathbf{x}_k = A\mathbf{x}_{k-1}$，$k = 1, 2, 3, \cdots$。其中：

$$\begin{aligned}
\mathbf{x}_1 &= A\mathbf{x}_0 = a_1 A v_1 + a_2 A v_2 + \cdots + a_n A v_n \\
&= a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \cdots + a_n \lambda_n v_n \\
\mathbf{x}_2 &= A\mathbf{x}_1 = a_1 \lambda_1 A v_1 + a_2 \lambda_2 A v_2 + \ldots + a_n \lambda_n A v_n \\
&= a_1 \lambda_1^2 v_1 + a_2 \lambda_2^2 v_2 + \ldots + a_n \lambda_n^2 v_n
\end{aligned}$$

乘幂法 (续)

得到向量序列：

$$
\begin{aligned}
\mathbf{x}_1 &= a_1\lambda_1 v_1 + a_2\lambda_2 v_2 + \ldots + a_n\lambda_n v_n \\
\mathbf{x}_2 &= a_1\lambda_1^2 v_1 + a_2\lambda_2^2 v_2 + \ldots + a_n\lambda_n^2 v_n \\
&\ldots \\
\mathbf{x}_k &= a_1\lambda_1^k v_1 + a_2\lambda_2^k v_2 + \ldots + a_n\lambda_n^k v_n
\end{aligned}
$$

下面按模最大特征值 λ_1 是单根的情况讨论：

$$
\begin{aligned}
\mathbf{x}_k &= a_1\lambda_1^k v_1 + a_2\lambda_2^k v_2 + \cdots + a_n\lambda_n^k v_n \\
&= \lambda_1^k \left[a_1 v_1 + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + a_n \left(\frac{\lambda_n}{\lambda_1}\right)^2 v_n \right] \\
&= \lambda_1^k (a_1 v_1 + \varepsilon_k)
\end{aligned}
$$

$$
\varepsilon_k = a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + a_n \left(\frac{\lambda_n}{\lambda_1}\right)^2 v_n
$$

乘幂法 (续)

由于 $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|$，故 $\left|\frac{\lambda_j}{\lambda_1}\right| < 1(j = 2, 3, \cdots n)$，得到

$$\lim_{k \to +\infty} \varepsilon_k = 0$$

因此有：

$$\lim_{k \to +\infty} \frac{1}{\lambda_1^k} \mathbf{x}_k = \lim_{k \to +\infty} \frac{\lambda_1^k (a_1 v_1 + \varepsilon_k)}{\lambda_1^k} = a_1 v_1$$

此即为 λ_1 对应的特征向量。
因此得幂方法的迭代公式：

$$\mathbf{x}_k = A\mathbf{x}_{k-1} \quad k = 1, 2, \cdots$$

当 k 充分大时，有 $\mathbf{x}_k \approx \lambda_1^k a_1 v_1$

乘幂法改进

存在的问题是：当 $|\lambda_1| > 1$ 时，$\mathbf{x}_k$ 中不为 0 的分量会随着 $k \to \infty$ 而趋于 ∞，计算机计算时会造成溢出，当 $|\lambda_1| < 1$ 时，$\mathbf{x}_k$ 中的分量会随着 $k \to \infty$ 而趋于 0。实际计算时，为了避免计算过程中出现绝对值过大或过小的数参加运算，通常在每步迭代时，将向量“归一化”处理，修改计算公式：

$$\begin{cases} \mathbf{y}_k = A\mathbf{x}_{k-1} \\ m_k = \max(y_k) \quad\quad (k = 1, 2, \cdots) \\ \mathbf{x}_k = \mathbf{y}_k / m_k \end{cases}$$

上式中 m_k 是向量 $\mathbf{y}_k$ 中绝对值最大的一个分量，这时 $\mathbf{x}_k$ 分量的模最大为 1，当 k 充分大时，有：

$$\begin{cases} \lambda_1 \approx m_k \\ v_1 \approx x_k \end{cases}$$

乘幂法例子

求矩阵 A 的主特征值和其对应的特征向量。

$$A=\begin{bmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{bmatrix}$$

根据迭代公式：

$$\begin{cases} \mathbf{y}_k = A\mathbf{x}_{k-1} \\ m_k = \max(y_k) \qquad (k=1,2,\cdots) \\ \mathbf{x}_k = \mathbf{y}_k/m_k \end{cases}$$

取 $x_0=(0,0,1)^\top$，则 $y_1=(2,4,1)^\top$，$m_1=4$，归一化后得到 $x_1=\frac{1}{m_1}y_1=(0.5,1,0.25)^\top$。

乘幂法例子 (续)

依此类推，结果如下表：

k	y_k^T	m_k	x_k^T
1	$(2, 4, 1)$	4	$(0.5, 1, 0.25)$
2	$(4.5, 9, 7.75)$	9	$(0.5, 1, 0.8611)$
3	$(5.7222, 11.4444, 8.361)$	11.4444	$(0.5, 1, 0.7360)$
4	$(5.4621, 10.9223, 8.2306)$	10.9223	$(0.5, 1, 0.7536)$
5	$(5.5075, 11.0142, 8.2576)$	11.0142	$(0.5, 1, 0.7494)$
6	$(5.4987, 10.9974, 8.2494)$	10.9974	$(0.5, 1, 0.7501)$
7	$(5.5002, 11.0005, 8.2501)$	11.0005	$(0.5, 1, 0.7500)$
8	$(5.5000, 11.0000, 8.2500)$	11.0000	$(0.5, 1, 0.7500)$

反幂法

如果 A 非奇异，用其逆矩阵代替 A 进行乘幂法，称为反幂法。逆矩阵的特征值与 A 互为倒数。即为：

$$\left|\frac{1}{\lambda_n}\right| \geq \left|\frac{1}{\lambda_{n-1}}\right| \geq \cdots \geq \left|\frac{1}{\lambda_1}\right|$$

用 A 的逆进行乘幂法，得到 $\frac{1}{\lambda_n}$
迭代格式为：

$$\begin{cases} \mathbf{y}_k = A^{-1}\mathbf{z}_{k-1} \\ m_k = \max\left(\mathbf{y}_k\right) \quad (k = 1, 2, \cdots) \\ \mathbf{z}_k = \mathbf{y}_k / m_k \end{cases}$$

$\mathbf{z}_0$ 任意给定，m_k 为 $\mathbf{y}_k$ 中绝对值最大的分量。当 $k \to \infty$，有：

$$\lim_{k\to\infty} m_k = \left|\frac{1}{\lambda_n}\right|$$

反幂法 (续)

为避免求逆矩阵运算，通常采取如下的算法：

$$\begin{cases} A\mathbf{y}_k = \mathbf{z}_{k-1} \\ m_k = \max(\mathbf{y}_k) \\ \mathbf{z}_k = \mathbf{y}_k / m_k \end{cases} \quad (k = 1, 2, \cdots)$$

每次迭代需要解 $A\mathbf{y}_k = \mathbf{z}_{k-1}$。为此，可将 A 进行 LU 分解，则每次迭代只需要解两个三角方程组：

$$\begin{cases} \mathbf{L}\mathbf{x} = \mathbf{z}_{k-1} \\ \mathbf{U}\mathbf{y}_k = \mathbf{x} \end{cases}$$

最后求得：

$$\frac{1}{\lambda_n} = \lim_{k\to\infty} m_k$$

$$\lim_{k\to\infty} \mathbf{x}_k = \frac{x_n}{\max(x_n)}$$

本章小结

- 基本概念；
- 乘幂法；
- 反幂法；

练习题

1. 编程实现乘幂法;
2. 编程实现反幂法;

谢谢!

AUTHOR: Cheng Yong
ADDRESS: Dept. of Computer
Beijing University of Chemical Technology
Beijing, 100029, China
EMAIL: buctcourse@163.com