Jacobi 方法

张晓平

2018 年 10 月 28 日

Jacobi 方法是求实对称矩阵全部特征值和特征向量的一种古老方法。众所周知，任一实对称矩阵都可通过正交相似变换约化成对角阵。Jacobi 方法正是利用这一特点，用一系列的 Givens 变换将一个给定的实对称矩阵逐步约化为对角阵。从收敛速度来看，Jacobi 方法与对称 QR 算法相比相差甚远，但其优点是编程简单，并行效率很高。

1 Jacobi 方法介绍

设 $\boldsymbol{A}=(a_{ij})\in\mathbb{R}^{n\times n}$ 是实对称矩阵，Jacobi 方法的目标是将 $\boldsymbol{A}$ 的非对角"范数"

$$E(\boldsymbol{A})=\left(\|\boldsymbol{A}\|_F^2-\sum_{i=1}^n a_{ii}^2\right)^{1/2}=\left(\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n a_{ij}^2\right)^{1/2}$$

逐步约化为零。

Jacobi 方法做一次约化的基本步骤是

1. 选择旋转平面 $(p,q), 1\le p<q\le n$；

2. 确定旋转角 θ 使得

$$\begin{bmatrix} b_{pp} & b_{pq}\\ b_{qp} & b_{qq}\end{bmatrix}=\begin{bmatrix} c & s\\ -s & c\end{bmatrix}^T\begin{bmatrix} a_{pp} & a_{pq}\\ a_{qp} & a_{qq}\end{bmatrix}\begin{bmatrix} c & s\\ -s & c\end{bmatrix} \tag{1}$$

   是对角阵（即 $b_{pq}=b_{qp}=0$），其中 $c=\cos\theta, s=\sin\theta$；

3. 对 $\boldsymbol{A}$ 做相似变换：

$$\boldsymbol{B}=(b_{ij})=\boldsymbol{G}^T\boldsymbol{A}\boldsymbol{G},\quad \boldsymbol{G}=\boldsymbol{G}(p,q,\theta).$$

注意 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 只在第 p,q 行与第 p,q 列不同，它们之间的关系为

$$\begin{cases} b_{ip}=b_{pi}=ca_{ip}-sa_{iq}, & i\neq p,q,\\ b_{iq}=b_{qi}=sa_{ip}+ca_{iq}, & i\neq p,q,\\ b_{pp}=c^2a_{pp}-2csa_{pq}+s^2a_{qq}, & \\ b_{qq}=s^2a_{pp}+2csa_{pq}+c^2a_{qq}, & \\ b_{pq}=b_{qp}=(c^2-s^2)a_{pq}+cs(a_{pp}-a_{qq}). & \end{cases} \tag{2}$$

1.1 假设已经选定好平面 (p,q)，如何确定 θ 使得 $b_{pq}=b_{qp}=0$?

由(2)的最后一个式子可知，确定 θ 相当于计算 c 和 s 使得

$$(c^2-s^2)a_{pq}+cs(a_{pp}-a_{qq})=0. \tag{3}$$

若 $a_{pq}=0$，则只需取 $c=1,s=0$ 即可。若 $a_{pq}\neq 0$，则令

$$\tau=\frac{a_{qq}-a_{pp}}{2a_{pq}},\quad t=\tan\theta=\frac{s}{c},$$

代入(3)可知 t 是方程

$$t^2+2\tau t-1=0$$

的解，即

$$t=-\tau\pm\sqrt{\tau^2+1}.$$

这里选绝对值较小的根，即

$$t=\frac{sgn(\tau)}{|\tau|+\sqrt{\tau^2+1}}, \tag{4}$$

这可保证 $|t|\leq 1$，从而 $|\theta|\leq\frac{\pi}{4}$。确定了 t 之后，c,s 可由如下公式确定

$$c=\frac{1}{\sqrt{1+t^2}},\quad s=tc. \tag{5}$$

旋转变换

```
function [c, s] = givens_rotation(A, p, q)
%% A should be symmetric
if A(p, q) == 0
    c = 1; s = 0;
else
    tau = (A(q, q) - A(p, p)) / (2 * A(p, q));
    t = sign(tau) / (abs(tau) + sqrt(tau^2+1));
    c = 1 / sqrt(1+t^2); s = t*c;
end

end
```

Jacobi 约化

```
function [A, Q] = Jacobi(A, p, q)
n = size(A, 1); Q = eye(n);
[c, s] = givens_rotation(A, p, q);
G = [c s; -s c];
A([p q], 1:end) = G' * A([p q], 1:end);
A(1:end, [p q]) = A(1:end, [p q]) * G;
Q(1:end, [p q]) = Q(1:end, [p q]) * G;
end
```

1.2 如何选取旋转平面?

1.2.1 经典 Jacobi 方法

由 F 范数的定义

$$\|\boldsymbol{A}\|_F = \sum_{i,j=1}^{n} a_{ij}^2 = tr(\boldsymbol{A}^T\boldsymbol{A}),$$

$\forall$ 正交矩阵 $\boldsymbol{P}$ 有

$$\|\boldsymbol{PA}\|_F = tr((\boldsymbol{PA})^T(\boldsymbol{PA})) = tr(\boldsymbol{A}^T\boldsymbol{P}^T\boldsymbol{PA}) = \mathrm{tr}(\boldsymbol{A}^T\boldsymbol{A}) = \|\boldsymbol{A}\|_F,$$

这说明 F 范数在正交变换下保持不变。

由于 F 范数在正交变换下保持不变，故 $\|\boldsymbol{B}\|_F = \|\boldsymbol{A}\|_F$。另一方面，由(1)可知，

$$a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2 = b_{pp}^2 + b_{qq}^2,$$

于是

$$\begin{aligned}
E(\boldsymbol{B})^2 &= \|\boldsymbol{B}\|_F^2 - \sum_{i=1}^{n} b_{ii}^2 \\
&= \|\boldsymbol{A}\|_F^2 - \left(\sum_{\substack{i=1\\ i\neq p,q}}^{n} b_{ii}^2 + b_{pp}^2 + b_{qq}^2\right) \\
&\xlongequal[i\neq p,q]{b_{ii}^2=a_{ii}^2} \|\boldsymbol{A}\|_F^2 - \left(\sum_{\substack{i=1\\ i\neq p,q}}^{n} a_{ii}^2 + a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2\right) \\
&= \|\boldsymbol{A}\|_F^2 - \sum_{i=1}^{n} a_{ii}^2 - 2a_{pq}^2 \\
&= E(\boldsymbol{A})^2 - 2a_{pq}^2.
\end{aligned}$$

我们希望 $E(B)$ 尽可能地小，因此 (p,q) 的最佳选择应使得

$$|a_{pq}| = \max_{1\le i<j\le n} |a_{ij}|, \tag{6}$$

即选取非对角元中绝对值最大者所在的行列为旋转平面。

按照(6)确定旋转平面 (p,q)，再由(4)和(5)来确定 c 和 s 的方法就是经典 Jacobi 方法。Jacobi 方法的基本迭代格式为

$$\boldsymbol{A}_k = \boldsymbol{G}_k^T\boldsymbol{A}_{k-1}\boldsymbol{G}_k, \quad k = 1,2,\cdots, \tag{7}$$

其中 $\boldsymbol{G}_k = \boldsymbol{G}(p_k,q_k,\theta_k)$。

经典 Jacobi 方法

```
clear;
n = 5;
B = rand(5);
A = (B + B')/2

%% choose p and q
for k = 1:5
    for i = 1:n*(n-1)/2
        B = abs(A - diag(diag(A)));
        [p, q]=find(B==max(max(B)));
```

```
        p = p(1); q = q(1);
        A = Jacobi(A, p, q);
    end
    A
end
```

定理. *对于经典 Jacobi 方法，有*

$$E(\boldsymbol{A}_k) \le \left(1 - \frac{2}{n(n-1)}\right)^{k/2} E(\boldsymbol{A}) \to 0, \quad as \quad k \to \infty. \tag{8}$$

证明. 由之前的讨论可知

$$E(\boldsymbol{A}_k)^2 = E(\boldsymbol{A}_{k-1})^2 - 2\left(a_{pq}^{(k-1)}\right)^2,$$

其中

$$|a_{pq}^{(k-1)}| = \max_{1\le i<j\le n} |a_{ij}^{(k-1)}|.$$

又因

$$E(\boldsymbol{A}_{k-1})^2 \le n(n-1)\left(a_{pq}^{(k-1)}\right)^2$$

于是有

$$E(\boldsymbol{A}_k)^2 \le \left(1 - \frac{2}{n(n-1)}\right) E(\boldsymbol{A}_{k-1})^2.$$

由此可知(8)成立。

□

1.2.2 循环 Jacobi 方法

对于经典 Jacobi 方法，

- 每做一次相似变换，运算量仅为 $O(n)$；
- 确定旋转平面 (p,q) 却需要进行 $n(n-1)/2$ 个元素之间的比较。

因此，经典 Jacobi 方法的大部分时间都在寻找最佳的旋转平面，这是得不偿失的。

为避免此问题，一种简单的办法就是不寻找最佳旋转平面，而是在一次扫面中，按照某种预先指定的顺序对每个非对角元约化一次，这就是所谓的循环 Jacobi 方法。最自然的循环 Jacobi 方法就是按如下顺序来扫描：

$$(p,q) = (1,2),\cdots,(1,n); \quad (2,3),\cdots,(2,n); \quad \cdots; \quad (n-1,n).$$

2 Jacobi 方法的优点

Jacobi 方法计算特征向量特别方便。如果经过 k 次变换后迭代停止，则

$$\boldsymbol{A}_k = \boldsymbol{G}_k^T \cdots \boldsymbol{G}_2^T \boldsymbol{G}_1^T A \boldsymbol{G}_1 \boldsymbol{G}_2 \cdots \boldsymbol{G}_k.$$

记

$$\boldsymbol{Q}_k = \boldsymbol{G}_1 \boldsymbol{G}_2 \cdots \boldsymbol{G}_k,$$

则有

$$\boldsymbol{A}\boldsymbol{Q}_k = \boldsymbol{Q}_k \boldsymbol{A}_k.$$

由于 $\boldsymbol{A}_k$ 的非对角元非常小，其对角元素就是 $\boldsymbol{A}$ 的近似特征值，而上式表明 $\boldsymbol{Q}_k$ 的列向量就是 $\boldsymbol{A}$ 的近似特征向量，并且所有近似特征向量都是正交规范的。这样要计算 $\boldsymbol{A}$ 的特征向量，只需将 $\boldsymbol{G}_i$ 累积起来即可，累积可以在迭代过程中同时进行。