《数值计算方法》课程

矩阵的特征值和奇异值

（幂法与反幂法）

胡建芳

（研究方向：计算机视觉）

http://sdcs.sysu.edu.cn/content/5143

计算机学院

课程回顾

- **解方程组：**

线性方程组：

……

非线性方程组：

……

线性方程组中，特殊矩阵，含上三角矩阵，下三角矩阵对角矩阵，对称正定矩阵

- **课程内容：**

1，绪论+基础知识；
2，求解方程　（重点，必考）
3，方程组（重点，必考）
4，插值
5，最小二乘法
6，数值微分与积分
7，特征值与奇异值　（重点，必考）
8，最优化

第三大重点

9，课程实践

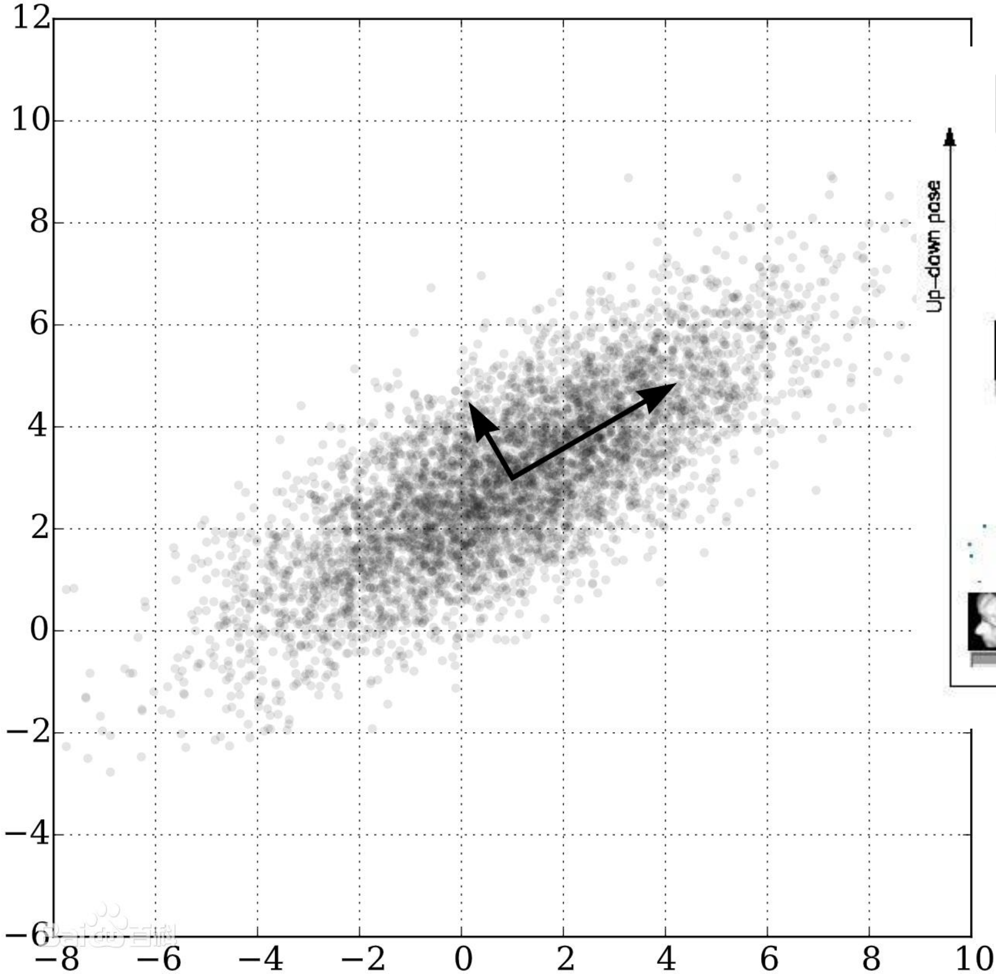
课程内容背景

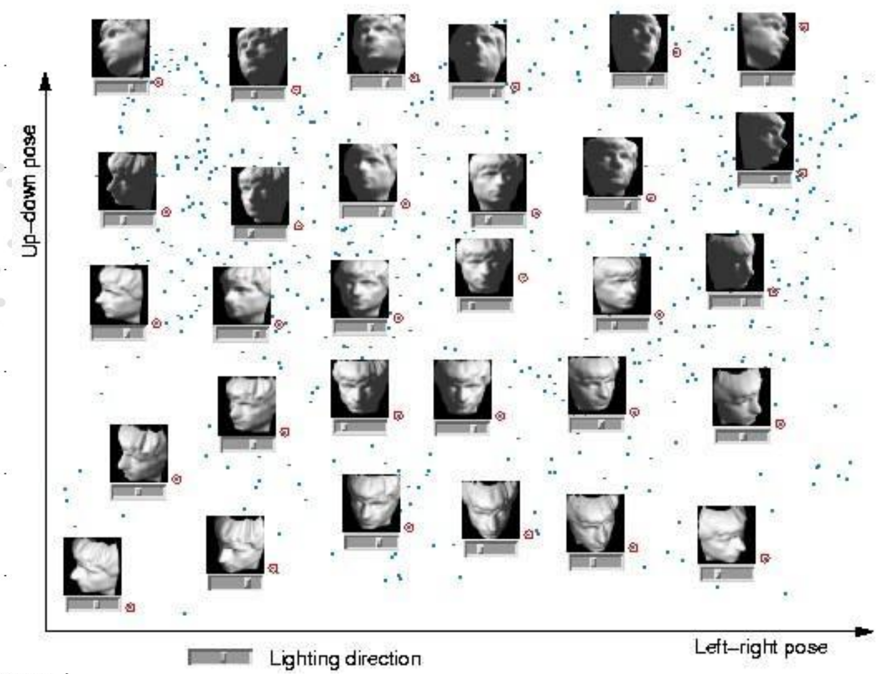
- 大数据：数据信息爆炸时代

这么多的数据，我们应该怎么办！

课程内容背景

- **大数据：数据爆炸时代**





提取主要信息

这是离人工智能最近的一章！

矩阵特征值

特征值性质：

特征值与特征向量

$$\boxed{Ax = \lambda x} \quad (\lambda \in C,\ x \neq 0)$$

- **性质**

(1) $Ax = \lambda x \Rightarrow (A - \mu I)x = (\lambda - \mu)x$

(2) $Ax = \lambda x \Rightarrow A^k x = \lambda^k x$

(3) $B = P^{-1}AP, Ax = \lambda x \Rightarrow By = \lambda y,\ y = P^{-1}x$

(4) 若 A 对称，则存在正交矩阵 Q，使得

$$Q^T AQ = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

矩阵特征值

特征值性质：

定义 12.1 设 A 是 $m\times m$ 矩阵, A 的**主特征值**(dominant eigenvalue) 是这样一个特征值 λ, 其大于 A 的所有其他特征值. 如果它存在, 那么与 λ 对应的特征向量称为**主特征向量**(dominant eigenvector).

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

$$x_1=Ax_0=\begin{bmatrix}1&3\\2&2\end{bmatrix}\begin{bmatrix}-5\\5\end{bmatrix}=\begin{bmatrix}10\\0\end{bmatrix},$$

$$x_2=A^2x_0=\begin{bmatrix}1&3\\2&2\end{bmatrix}\begin{bmatrix}10\\0\end{bmatrix}=\begin{bmatrix}10\\20\end{bmatrix},$$

$$x_3=A^3x_0=\begin{bmatrix}1&3\\2&2\end{bmatrix}\begin{bmatrix}10\\20\end{bmatrix}=\begin{bmatrix}70\\60\end{bmatrix},$$

$$x_4=A^4x_0=\begin{bmatrix}1&3\\2&2\end{bmatrix}\begin{bmatrix}70\\60\end{bmatrix}=\begin{bmatrix}250\\260\end{bmatrix}=260\begin{bmatrix}\frac{25}{26}\\1\end{bmatrix}$$

矩阵特征值

- **瑞利商：**

定理： 设 A 是 n 阶实对称矩阵，其特征值为

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$$

则对任意非零向量 x，有

$$\boxed{\lambda_n \le \frac{(Ax,x)}{(x,x)} \le \lambda_1}$$

且 $$\boxed{\lambda_1 = \max_{x\neq 0}\frac{(Ax,x)}{(x,x)},\quad \lambda_n = \min_{x\neq 0}\frac{(Ax,x)}{(x,x)}}$$

- $R(x) = \dfrac{(Ax,x)}{(x,x)}$ 称为矩阵 A 关于 x 的 Rayleigh 商。

矩阵特征值

- **幂法：**
  - **幂法（乘幂法，幂迭代）**
    - 计算矩阵的主特征值（按模最大）及其特征向量

假设：(1) $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n| \geq 0$

(2) 对应的 n 个线性无关特征向量为：$x_1, x_2, \ldots, x_n$

计算过程：

(1) 任取一个非零向量 v_0，要求满足 $(x_1, v_0) \neq 0$

(2) 对 $k = 1, 2, \ldots$，直到收敛，计算

$$v_k = Av_{k-1}$$

矩阵特征值

- **幂法的收敛性：**

定理： 设 A 有 n 个线性无关的特征向量，其特征值满足

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

则由幂法生成的向量满足

$$\lim_{k\to\infty}\frac{v_k}{\lambda_1^k} = \alpha_1 x_1, \quad \lim_{k\to\infty}\frac{(v_{k+1})_j}{(v_k)_j} = \lambda_1$$

- 注：幂法的收敛速度取决于 $\left|\dfrac{\lambda_2}{\lambda_1}\right|$ 的大小

矩阵特征值

- **幂法的收敛性：**

设 $v_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \quad (\alpha_1 \neq 0)$

$$\Rightarrow v_1 = A v_0 = \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \cdots + \alpha_n \lambda_n x_n$$

$$\vdots$$

$$v_k = A v_{k-1} = \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \cdots + \alpha_n \lambda_n^k x_n$$

$$= \lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right]$$

$$\longrightarrow \lambda_1^k \alpha_1 x_1$$

$\boxed{\left| \frac{\lambda_2}{\lambda_1} \right| \text{越小，收敛越快}}$

矩阵特征值

- **幂法的收敛性：**

当 k 充分大时，有

$$\left.\begin{aligned} v_k &\approx \lambda_1^k \alpha_1 x_1 \\ v_{k+1} &\approx \lambda_1^{k+1} \alpha_1 x_1 \end{aligned}\right\} \Rightarrow v_{k+1} \approx \lambda_1 v_k \Rightarrow \frac{(v_{k+1})_j}{(v_k)_j} \approx \lambda_1 \ (j = 1, 2, \ldots, n)$$

又 $v_{k+1} = A v_k$

$$\Rightarrow A v_k \approx \lambda_1 v_k$$

$$\Downarrow$$

v_k 为 λ_1 的近似特征向量

矩阵特征值

- 幂法:
  - 幂法中存在的问题

$$v_k \approx \lambda_1^k \alpha_1 x_1 \rightarrow \begin{cases} \infty, & |\lambda_1|>1 \\ 0, & |\lambda_1|<1 \end{cases}$$

改进方法：规范化

$$v_{k+1} = Av_k \Rightarrow u_k = \frac{v_k}{\|v_k\|_\infty}, \quad v_{k+1} = Au_k$$

$$\Rightarrow \lim_{k\to\infty} u_k = \frac{x_1}{\|x_1\|_\infty}$$

矩阵特征值

- 幂法:
  - λ_1 的计算

$$u_k = \frac{v_k}{\|v_k\|_\infty} = \frac{A^k v_0}{\|A^k v_0\|_\infty}$$

$$\Rightarrow\quad v_{k+1} = A u_k = \frac{A^{k+1} v_0}{\|A^k v_0\|_\infty} = \frac{\lambda_1^{k+1}\left[\alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k+1} x_i\right]}{\left\|\lambda_1^{k}\left[\alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k} x_i\right]\right\|_\infty}$$

$$\Rightarrow\quad \lim_{k\to\infty} \|v_k\|_\infty = \lambda_1$$

矩阵特征值

改进的幂法：

(1) 任取一个非零向量 v_0，要求满足 $(x_1,v_0)\neq 0$

(2) 对 $k=1, 2, \ldots$ ，直到收敛，计算

$$u_k=\frac{v_k}{\|v_k\|_\infty},\quad v_{k+1}=Au_k$$

定理： 设 A 有 n 个线性无关的特征向量，其特征值满足

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

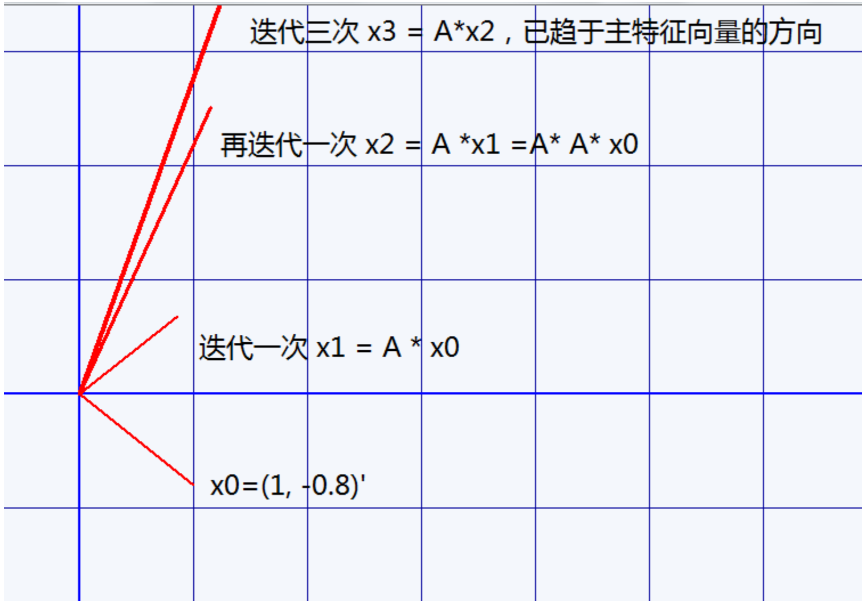
则由改进的幂法生成的向量满足

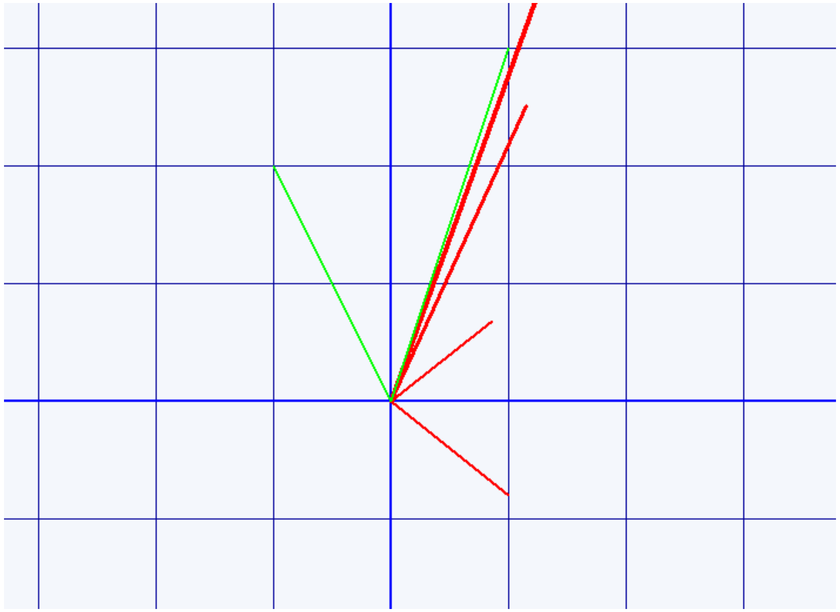
$$\lim_{k\to\infty}u_k=\frac{x_1}{\|x_1\|_\infty},\quad \lim_{k\to\infty}\|v_k\|_\infty=\lambda_1$$

矩阵特征值

- **举例：**

A=[1.1 , 0.3； 1.8, 1.4]， x=[1,-0.8]

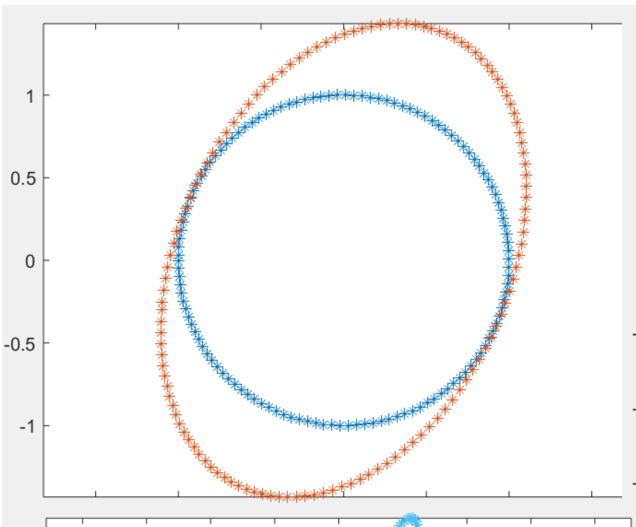


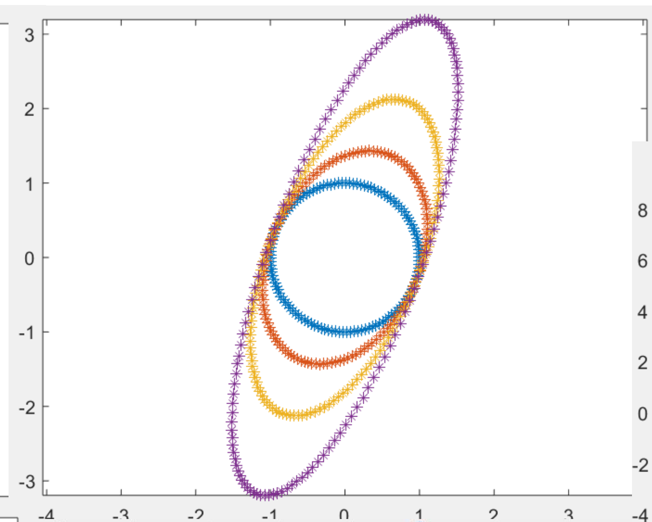


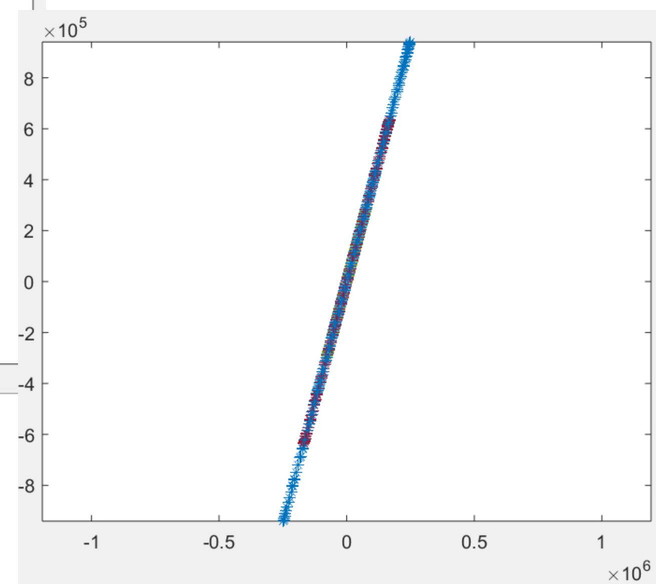
矩阵特征值

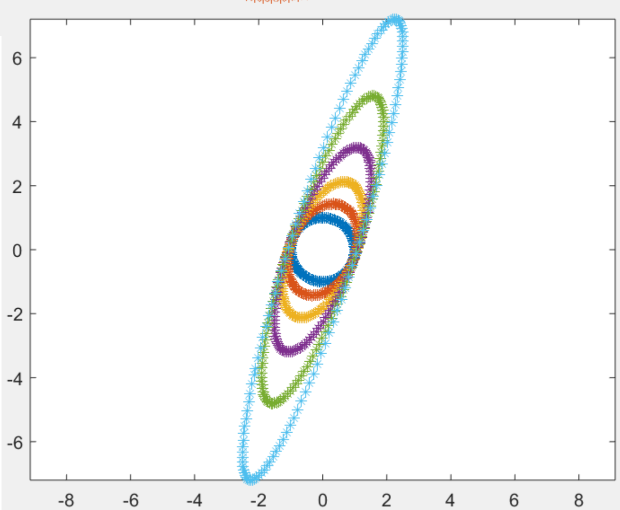
- **举例：**

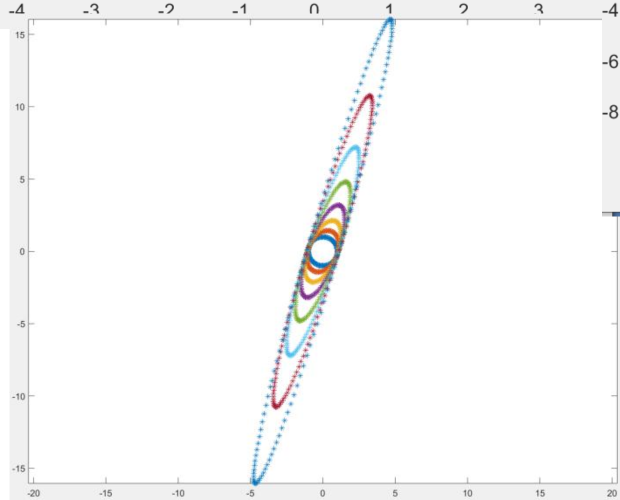
A=[1.1 , 0.1； 0.3, 1.4]，主方向[1.0209 1.4791]，A对单位圆作用。











25次迭代后，单位圆几乎变成一条直线，与主方向吻合

矩阵特征值

- **幂法的加速：**

幂法的收敛速度取决于 $r=\left|\frac{\lambda_2}{\lambda_1}\right|$ 的大小

$\Rightarrow$ 当 r 接近于 1 时，乘幂法收敛会很慢！

❑ **幂法的加速：原点平移法** $\longrightarrow$ $\boxed{\text{带位移的幂法}}$

令 $B=A-pI$，则 B 的特征值为：λ_i-p

选择适当的 p 满足：

(1) $|\lambda_1-p|>|\lambda_j-p|$ $(j=2,\ldots,n)$ $\longrightarrow$ $\boxed{\text{保持主特征值}}$

(2) $\max\limits_{2\le j\le n}\left|\frac{\lambda_j-p}{\lambda_1-p}\right|<\left|\frac{\lambda_2}{\lambda_1}\right|$ $\longrightarrow$ $\boxed{\text{加快收敛速度}}$

用幂法计算矩阵 B 的主特征值：λ_1-p

矩阵特征值

- **幂法：**

幂法可以计算最大特征值和特征向量，怎么求最小特征值?

矩阵特征值

- **反幂法：**
  - 计算矩阵的按模最小的特征值及其特征向量

假设：(1) $|\lambda_1| \ge |\lambda_2| \ge \ldots \ge |\lambda_{n-1}| > |\lambda_n| > 0$

(2) 对应的 n 个线性无关特征向量为：$x_1, x_2, \ldots, x_n$

A^{-1} 的特征值为：$\left|\dfrac{1}{\lambda_1}\right| \le \left|\dfrac{1}{\lambda_2}\right| \le \cdots \le \left|\dfrac{1}{\lambda_{n-1}}\right| < \left|\dfrac{1}{\lambda_n}\right|$

对应的特征向量仍然为 $x_1, x_2, \ldots, x_n$

- **反幂法**：对矩阵 A^{-1} 使用幂法

矩阵特征值

- **反幂法：**

(1) 任取一个非零向量 v_0，要求满足 $(x_1, v_0) \neq 0$

(2) 对 $k = 1, 2, \ldots,$ 直到收敛，计算

$$u_k = \frac{v_k}{\|v_k\|_\infty}, \quad v_{k+1} = A^{-1} u_k$$

定理： 设 A 有 n 个线性无关的特征向量，其特征值满足

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{n-1}| > |\lambda_n| > 0$$

则由反幂法生成的向量满足

$$\lim_{k\to\infty} u_k = \frac{x_n}{\|x_n\|_\infty}, \quad \lim_{k\to\infty} \|v_k\|_\infty = \frac{1}{\lambda_n}$$

矩阵特征值

- **反幂法的收敛：**

反幂法的收敛速度取决于 $r = \left|\dfrac{\lambda_n}{\lambda_{n-1}}\right|$ 的大小

⇒ 当 r 接近于 1 时，反乘幂法收敛会很慢！

可以使用原点平移法对反幂法进行加速

问题： 如何选择参数 p ？

⇒ 离 λ_n 越近越好（但不能相等）

矩阵特征值

- **反幂法的瑞利商加速：**

定理 设 A 是 n 阶实对称矩阵，其特征值为

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

对应的特征向量 $x_1, x_2, \ldots, x_n$ 满足：$(x_i, x_j) = \delta_{ij}$ ，

使用改进的乘幂法计算 A 的按模最大特征值 λ_1 时， u_k 的 Rayleigh商给出了 λ_1 的较好的近似，即

$$\frac{(Au_k, u_k)}{(u_k, u_k)} = \lambda_1 + O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right)$$

证： $u_k = \dfrac{A^k u_0}{\max(A^k u_0)}$， $v_{k+1} = \dfrac{A^{k+1} u_0}{\max(A^k u_0)}$

$$\Rightarrow \frac{(Au_k, u_k)}{(u_k, u_k)} = \frac{(A^{k+1}u_0, A^k u_0)}{(A^k u_0, A^k u_0)} = \frac{\sum_{i=1}^{n} \alpha_i^2 \lambda_i^{2k+1}}{\sum_{i=1}^{n} \alpha_i^2 \lambda_i^{2k}} = \lambda_1 + O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right)$$

20

矩阵特征值

- **反幂法的瑞利商加速：**

❑ Rayleigh 商加速

$$\lim_{k\to\infty} u_k = \frac{x_n}{\|x_n\|_\infty} \Rightarrow \lim_{k\to\infty} \frac{(u_k, Au_k)}{(u_k, u_k)} = \frac{(x_n, Ax_n)}{(x_n, x_n)} = \lambda_n$$

(1) 任取一个非零向量 v_0，要求满足 $(x_1, v_0) \neq 0$

(2) 对 $k = 1, 2, \ldots,$ 直到收敛，计算

$$u_k = \frac{v_k}{\|v_k\|_\infty}, \quad p_k = \frac{(u_k, Au_k)}{(u_k, u_k)}$$

$$v_{k+1} = (A - p_k I)^{-1} u_k$$

矩阵特征值

■ **反幂法的瑞利商加速：**

- 带位移的反幂法中需要计算 $v_{k+1} = (A - p_k I)^{-1} u_k$

$$\Leftrightarrow \quad (A - p_k I) v_{k+1} = u_k$$

解线性方程组

- 带位移的反幂法可以用于计算任何一个特征值 λ_k

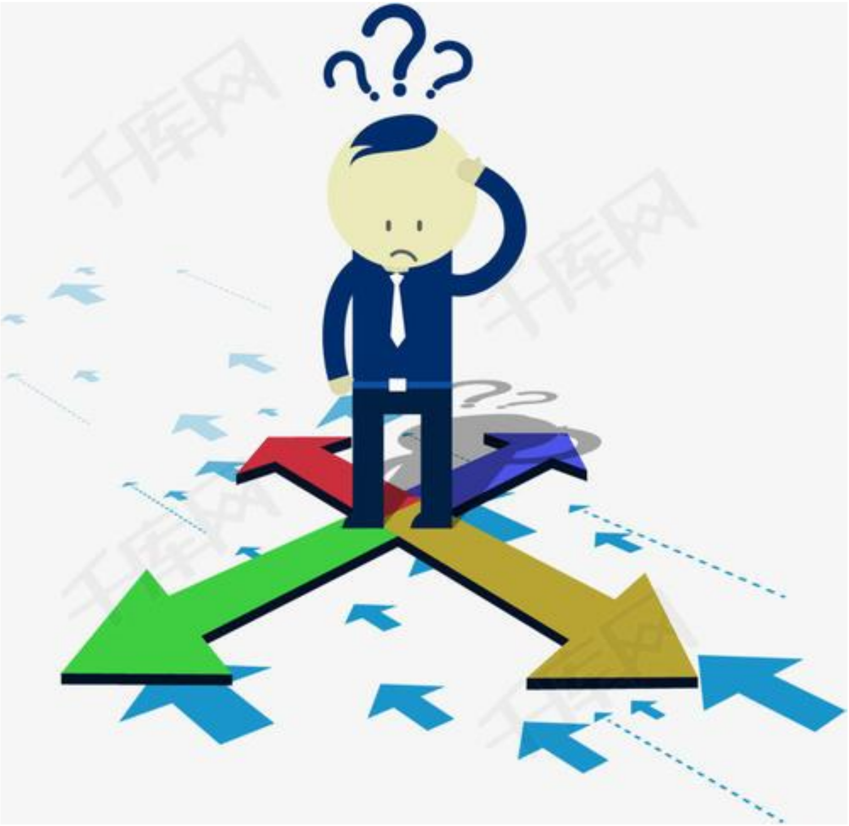
$\Rightarrow$ 将参数 p 取为 λ_k 附近

- 若已知特征值，计算特征向量时，可使用带位移的反幂法

$\Rightarrow$ 令 p 足够靠近 λ_k

矩阵特征值

- **思考：**
  怎样基于幂法迭代，算出矩阵A的所有特征值和特征向量？



矩阵特征值

- **作业：**

  无

矩阵特征值

THE END