第九讲 特征值分析方法的应用

贺丽君
信息与通信工程学院
Email: lijunhe@mail.xjtu.edu.cn
2023-04

内容提要

- Markov矩阵及其性质
- 应用举例：Google 网页排名算法

内容提要

- Markov矩阵及其性质
- 应用举例：Google 网页排名算法

Markov矩阵

稳态与收敛速度

- 人口迁移问题

$$\mathbf{A}=\begin{bmatrix}0.9 & 0.2\\ 0.1 & 0.8\end{bmatrix}\qquad \mathbf{u}_0=\begin{bmatrix}a\\ b\end{bmatrix}\qquad \mathbf{u}_\infty=?$$

$$\mathbf{A}^k\mathbf{u}_0\to\mathbf{u}_\infty,\quad k>?$$

Markov矩阵：

- A的每一个元素均为非负实数
- A的每一列元素之和等于1

⇒ 每个列向量称为概率向量

Markov矩阵的性质

- 若$\mathbf{A}$为Markov矩阵，且$\mathbf{u}_0$为一非负向量，则$\mathbf{u}_1=\mathbf{A}\mathbf{u}_0$为非负向量
- 若$\mathbf{A}$为Markov矩阵且非负向量$\mathbf{u}_0$的各元素之和等于1，则$\mathbf{u}_1=\mathbf{A}\mathbf{u}_0$，$\mathbf{u}_2=\mathbf{A}^2\mathbf{u}_0$，…，$\mathbf{u}_k=\mathbf{A}^k\mathbf{u}_0$的各元素之和都等于1

$$[1,1,\cdots]\mathbf{A}=[1,1,\cdots]\Rightarrow[1,1,\cdots]\mathbf{A}\mathbf{u}_0=[1,1,\cdots]\mathbf{u}_0=1$$

$$[1,1,\cdots]\mathbf{u}_1=1$$

Markov矩阵的性质

- 若$\mathbf{A}$为Markov矩阵，且$\mathbf{u}_0$为一非负向量，则$\mathbf{u}_1 = \mathbf{A}\mathbf{u}_0$为非负向量
- 若$\mathbf{A}$为Markov矩阵且非负向量$\mathbf{u}_0$的各元素之和等于1，则$\mathbf{u}_1=\mathbf{A}\mathbf{u}_0$，$\mathbf{u}_2=\mathbf{A}^2\mathbf{u}_0$，…，$\mathbf{u}_k= \mathbf{A}^k\mathbf{u}_0$的各元素之和都等于1
- 如果$\mathbf{A}$是正的Markov矩阵(各元素大于0，且每一列元素之和等于1)，则$\lambda=1$是唯一的模为1的特征值，其代数重数和几何重数均为1，与该特征值对应的特征向量各分量均大于0；其余所有特征值的模均小于1

$\mathbf{A}-\mathbf{I}$各行向量相加为零$\Rightarrow$行向量线性相关

$\det(\mathbf{A}-\mathbf{I})=0 \Rightarrow$（$\mathbf{A}-\lambda\mathbf{I}$）$\mathbf{x}=\mathbf{0}$在$\lambda=1$时，有非零解

正Markov矩阵

$$\mathbf{u}_k = \mathbf{A}^k \mathbf{u}_0$$

$$= c_1(\lambda_1)^k \mathbf{x}_1 + c_2(\lambda_2)^k \mathbf{x}_2 + \ldots + c_n(\lambda_n)^k \mathbf{x}_n$$

假设上式中的各项满足：

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

如果A是正Markov矩阵，则：

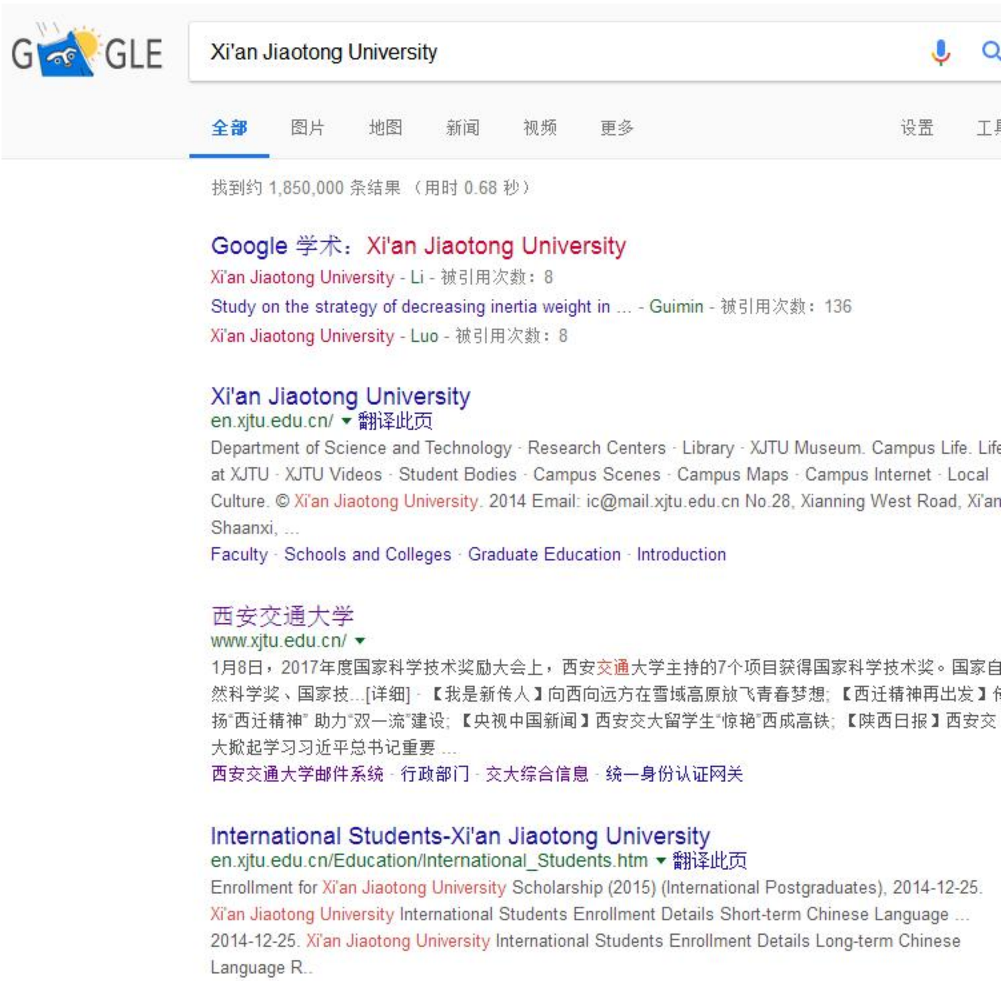
$$\mathbf{u}_k \to c_1\mathbf{x}_1, \quad k \to \infty$$

最大特征值所对应的特征向量决定稳态，稳态与初值无关；第二大的特征值控制收敛速度

内容提要

- Markov矩阵及其性质
- 应用举例：Google 网页排名算法

网页排名

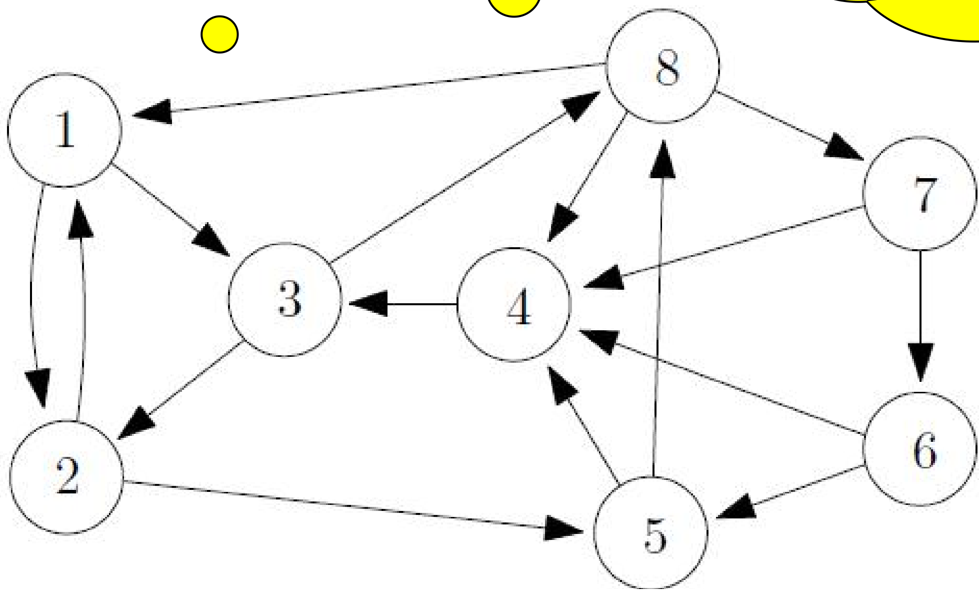


- 网页内容与搜索关键词的相关性 (Relevance Score)
- 网页本身的重要性 (Importance Score)

Google PageRank算法

链接到该网页的
其他网页越重要，
该网页页越重要

π_1



$\frac{1}{O_2}$

$$\pi_1 = \sum_{j \to 1} \frac{\pi_j}{O_j}$$

$$\pi_1 = 0\pi_1 + \frac{1}{2}\pi_2 + 0\pi_3 + 0\pi_4 + 0\pi_5 + 0\pi_6 + 0\pi_7 + \frac{1}{3}\pi_8$$

$$\pi_2 = \frac{1}{2}\pi_1 + 0\pi_2 + \frac{1}{2}\pi_3 + 0\pi_4 + 0\pi_5 + 0\pi_6 + 0\pi_7 + 0\pi_8$$

$$\vdots$$

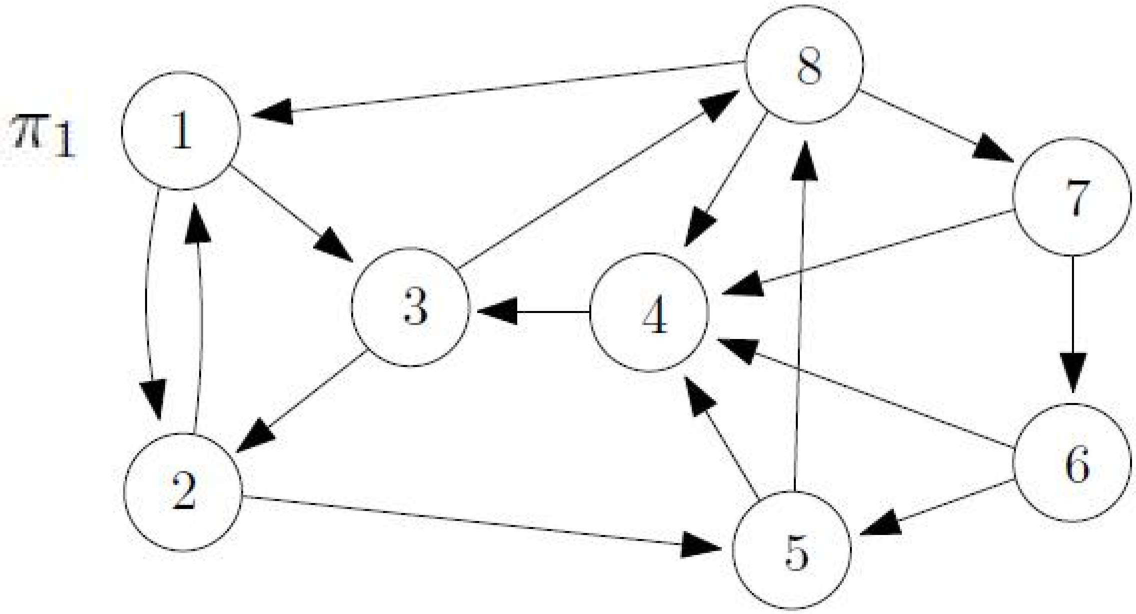
$\frac{1}{O_5}$

$$\pi_8 = 0\pi_1 + 0\pi_2 + \frac{1}{2}\pi_3 + 0\pi_4 + \frac{1}{2}\pi_5 + 0\pi_6 + 0\pi_7 + 0\pi_8$$

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{H}$$

$$\mathbf{H}_{ij} = \begin{cases} \frac{1}{O_i}, & i \to j \\ 0, & \text{otherwise} \end{cases}$$

Google PageRank算法



$$\pi_1 = \sum_{j \to 1} \frac{\pi_j}{O_j}$$

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{H}$$

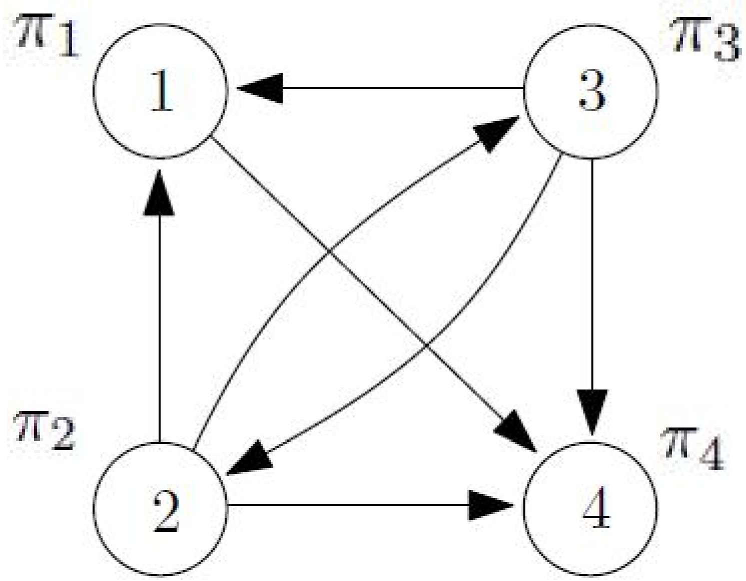
$$\mathbf{H}_{ij} = \begin{cases} \dfrac{1}{O_i}, & i \to j \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{H} = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 1/3 & 0 \end{bmatrix}$$

Google PageRank算法

一定存在这样的向量π么？

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{H}$$



$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{cases} \frac{1}{3}(\pi_2 + \pi_3) = \pi_1 \\ \frac{1}{3}\pi_3 = \pi_2 \\ \frac{1}{3}\pi_2 = \pi_3 \\ \pi_1 + \frac{1}{3}(\pi_2 + \pi_3) = \pi_4 \end{cases} \Rightarrow \begin{matrix} \pi_1 = 0 \\ \pi_2 = 0 \\ \pi_3 = 0 \\ \pi_4 = 0 \end{matrix}$$

Google PageRank算法

$\mathbf{H}^T$应为正Markov矩阵

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{H} \Rightarrow \boldsymbol{\pi} = \mathbf{H}^T \boldsymbol{\pi}$$

1、$\mathbf{H}^T$具有特征值1

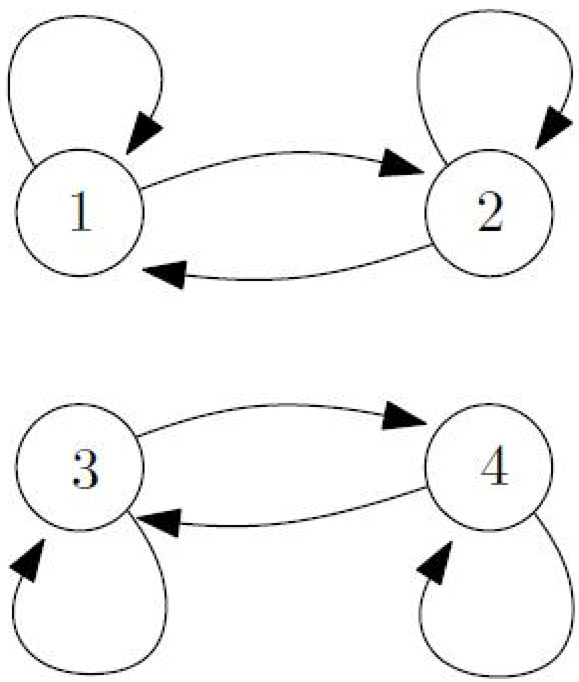
2、π是与1对应的唯一特征向量 $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \hat{\mathbf{H}} = \mathbf{H} + \frac{1}{N}(\mathbf{w}\underline{\mathbf{1}}^T)$$

$$w_i = \begin{cases} 1, & i \text{ is a dangling node} \\ 0, & \text{otherwise} \end{cases}$$

Google PageRank算法

H^T不是正Markov矩阵



$$\mathbf{H} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$$

解不唯一：

$$\pi^* = [0.5\ 0.5\ 0\ 0]^T$$

$$\pi^* = [0.15\ 0.15\ 0.35\ 0.35]^T$$

$$\hat{\mathbf{H}} = \mathbf{H} + \frac{1}{N}(\mathbf{w}\mathbf{1}^T) \Longrightarrow \mathbf{G} = \theta\hat{\mathbf{H}} + (1-\theta)\frac{1}{N}\mathbf{1}\mathbf{1}^T$$

Google PageRank算法

正Markov矩阵

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{H} \quad \Rightarrow \quad \boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{G} \quad \Leftrightarrow \quad \boldsymbol{\pi} = \mathbf{G}^T \boldsymbol{\pi}$$

如何求解π？

$$\left(\mathbf{G}^T - \mathbf{I}\right)\mathbf{x} = \mathbf{0}$$

如果A是正Markov矩阵，则：

$$\mathbf{u}_k = \mathbf{A}^k \mathbf{u}_0 \to c_1 \mathbf{x}_1, \quad k \to \infty$$

迭代算法：

$$\boldsymbol{\pi}[k] = \mathbf{G}^T \boldsymbol{\pi}[k-1]$$

谢谢大家！