# § 9. Markov矩阵和正矩阵

# 一 问题引入

设某小镇有3000人，选用两种牙膏A-1和A-2, 本年度有1000人选用牙膏A-1, 2000人选用牙膏A-2. 据调查，下一年度选用牙膏A-1的居民，60%将继续选用A-1，40%将改选A-2. 选用牙膏A-2的居民，70%将继续选用A-2，30%将改选A-1.

可以总结如下表

| | 明年选用A-1的居民 | 明年选用A-2的居民 |
|---|---|---|
| 今年选用A-1的居民 | 60% | 40% |
| 今年选用A-2的居民 | 30% | 70% |

# 一 问题引入

设下一年度选用A-1为$x_1$人，选用A-2为$y_1$人.

(本年度选用A-1为$x_0 = 1000$人,选用A-2为$y_0 = 2000$人). 我们有

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

$\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}$ 被称为转移(概率)矩阵(transition matrix). 假设这个规律不变，

后年度选用A-1为$x_2$人，选用A-2为$y_2$人，则有 $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}^2 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

# 一 问题引入

一般地，在上述规律不变的情况下，我们得到了向量序列

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \quad \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \quad \ldots, \begin{pmatrix} x_k \\ y_k \end{pmatrix}, \ldots$$

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}^k \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

其中，$x_k$是第$k$年度选用A-1的人数，$y_k$是第$k$年度选用A-2的人数. (今年是第0年度)

这个序列给出一个马尔科夫链(Markov chain),可以预测随时间变化序列未来的趋势。

$\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}$ 是一个Markov 矩阵， 当$k \to +\infty$, $\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}^k$ 的极限状态给出了未来的趋势.

# 一 问题引入

参考文献：百度文库>专业资料>自然科学>数学 >马尔科夫链，网址：
http://wenku.baidu.com/link?url=V2b5jZHvANQSeo0zXXSMKd29fEuKLT8-rPkleyC0yzWHp9qRU9ReX1pAolBWeAeCIRUtKdsw55HEFgxqL$_{Y}I9yJJbN8tCr8c0E5KFkYlBga$

# 二 Markov矩阵

一个$n \times n$的实矩阵$\mathbf{A}$被称为Markov矩阵，如果它的元素均为非负，且每个列向量分量之和等于1. 例如：

$$\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \text{等}.$$

Markov矩阵也被称为随机矩阵(Stochastic matrix).一个向量满足分量非负，分量之和等于1 称为随机向量.

# 二 Markov矩阵

我们使用如下技巧： 设$\mathbf{A}=(a_{ij})$, 若$a_{ij} > 0, \forall i, j$, 则写$\mathbf{A}> 0$.

设$\alpha$和$\beta$是两个实向量，$\alpha \geq \beta$意味着$\alpha$的每个分量均大于等于$\beta$的相应分量，例如:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \geq \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

我们写$\alpha > \beta$若$\alpha$的每个分量严格大于$\beta$的相应分量.

**性质：** 设$\mathbf{A}> 0$,$\alpha \geq \beta$且$\alpha \neq \beta$, $\exists \varepsilon > 0$, $\mathbf{A}\alpha>(1+\varepsilon)\mathbf{A}\beta$.

# 二 Markov矩阵

由定义，Markov矩阵的列向量是随机向量.

基本性质：

(1) 设$\mathbf{A}$是$n$阶Markov矩阵，$\mathbf{p}$是$n$维随机向量，则$\mathbf{Ap}$是一个$n$维随机向量.

证明：设$\mathbf{A} = (\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n})$,$\mathbf{v_i}$,是$\mathbf{A}$的第$i$列

$$\mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}, p_1 + ... + p_n = 1, p_i \geq 0.$$

$\mathbf{Ap} = p_1\mathbf{v_1} + ... + p_n\mathbf{v_n}$.

$(1, 1, ..., 1)\mathbf{A} = (1, 1, ..., 1)$, $(1, 1, ..., 1)\mathbf{p} = 1$

$\Rightarrow (1, 1, ..., 1)\mathbf{Ap} = 1$.

# 二 Markov矩阵

(2) 一个Markov矩阵$\mathbf{A}$总有特征值1，其余特征值长度小于等于1.

证明： 因为$|\lambda\mathbf{I}-\mathbf{A}|=|\lambda\mathbf{I}-\mathbf{A}^{\mathbf{T}}|$, 且$\mathbf{A}$和$\mathbf{A}^{\mathbf{T}}$有相同的特征值.由定义

$(1,1,...,1)\mathbf{A}=(1,1,...,1)$, 即$\mathbf{A}^{\mathbf{T}}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$. 所以$\mathbf{A}^{\mathbf{T}}$有特征值1，$\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$是特征向量.

假设$\mathbf{A}$存在特征值$\lambda_0\in\mathbb{C}$, $|\lambda_0|>1$,$\mathbf{A}\alpha=\lambda_{\mathbf{0}}\alpha$.

设 $\alpha=\begin{pmatrix}z_1\\ \vdots\\ z_n\end{pmatrix}$, $z_i\in\mathbb{C}$, $\mathbf{A}=(a_{ij})_{n\times n}$ 令$\alpha^+=\begin{pmatrix}|z_1|\\ \vdots\\ |z_n|\end{pmatrix}$, $\mathbf{A}\alpha^+=\beta=\begin{pmatrix}\omega_1\\ \vdots\\ \omega_n\end{pmatrix}$, $\omega_i\in\mathbb{R}$, $\omega_i\geq 0$.

# 二 Markov矩阵

$$\omega_i = (a_{i1}, ..., a_{in}) \begin{pmatrix} |z_1| \\ \vdots \\ |z_n| \end{pmatrix} = a_{i1}|z_1| + ... + a_{in}|z_n| \geq |a_{i1}z_1 + ... + a_{in}z_n| = |\lambda_0 z_i| = |\lambda_0||z_i|$$

因此,$\mathbf{A}\alpha^+ \geq |\lambda_\mathbf{0}|\alpha^+$.

因为$\mathbf{A}$非负，$\mathbf{A}(\mathbf{A}\alpha^+) \geq \mathbf{A}(|\lambda_0|\alpha^+) = |\lambda_0|\mathbf{A}\alpha^+ \geq |\lambda_0|^2\alpha^+$

$\Rightarrow \mathbf{A}^\mathbf{k}\alpha^+ \geq |\lambda_\mathbf{0}|^\mathbf{k}\alpha^+$. 因为$(1, ..., 1)\mathbf{A} = (1, ..., 1)$, 不等式两端均从左边乘上(1,...,1)则有，

$(1, ..., 1)\mathbf{A}^\mathbf{k}\alpha^+ = (1, 1, ..., 1)\alpha^+ = |z_1| + ... + |z_n|$.

$(1, ..., 1)\lambda_0^k\alpha^+ = |\lambda_0|^k(|z_1| + ... + |z_n|)$.

因此$|\lambda_0| \leq 1$.

# 二 Markov矩阵

由此性质:存在关于$\lambda = 1$特征向量$\alpha_0$, $\mathbf{A}\alpha_0 = \alpha_0$.

性质(2) 中“其余特征值的长度满足$|\lambda| \leq 1$的等号不能去掉”.

例如:

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$ 是一个置换阵，它的特征值是 $\lambda_1 = 1, \lambda_2 = e^{i\frac{2\pi}{3}}, \lambda_3 = e^{i\frac{4\pi}{3}}$

$|\lambda_1| = |\lambda_2| = |\lambda_3|$.可以看到，$\mathbf{P^3} = \mathbf{I_3}$，因此当$k \to +\infty$, $\mathbf{P^k}$无极限.

我们考虑一类特殊的Markov矩阵，使得极限存在.

# 三 正Markov矩阵

**定义：** 一个正Markov矩阵是一个Markov矩阵，其元素均大于0.

例如： $\begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}$是正Markov矩阵， $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$不是正Markov矩阵.

我们来看二阶正Markov矩阵，即$\mathbf{P} = \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix}$. $0 < a < 1, 0 < b < 1$.使用归纳法得到

$\mathbf{P^n} = \frac{\mathbf{1}}{\mathbf{a+b}} \begin{pmatrix} b & b \\ a & a \end{pmatrix} + \frac{\mathbf{(1-a-b)^n}}{\mathbf{a+b}} \begin{pmatrix} a & -b \\ -a & b \end{pmatrix}$, $0 < a < 1, 0 < b < 1, \Rightarrow |1-a-b| < 1 \Rightarrow$ $(1-a-b)^n \to 0, n \to +\infty$.

因此, $\lim\limits_{m\to\infty} \mathbf{P^n} = \begin{pmatrix} \frac{b}{a+b} & \frac{b}{a+b} \\ \frac{a}{a+b} & \frac{a}{a+b} \end{pmatrix}$ 这个极限矩阵的列向量是特征向量，满足分量均大于0，分量之和等于1(随机向量).

这个极限矩阵的列向量是特征向量，满足分量均大于0，分量之和等于1(随机向量).

# 四 正矩阵

$$\mathbf{P}\begin{pmatrix} \frac{b}{a+b} \\ \frac{a}{a+b} \end{pmatrix} = \begin{pmatrix} \frac{b}{a+b} \\ \frac{a}{a+b} \end{pmatrix}.$$

$\lambda = 1$是一个特征值，另一个特征值是$[(1-a)+(1-b)]-1 = 1-a-b.$

$1-a-b \neq 1 \Rightarrow \lambda = 1$的代数重数=几何重数=1.

若$a$或$b$等于$0, 1$,$\mathbf{P}$不再是正Markov矩阵，以上讨论不一定成立.例如$a = b = 0$, $\lambda = 1$是2重特征值.

$a = 1$,$b = 1$，$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$，$\mathbf{P^2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{P^3} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,即$\lim\limits_{m\to\infty} \mathbf{P^n}$不存在.

# 三 正Markov矩阵

**定理：** 设$\mathbf{A}$是正Markov阵，则$\lambda=1$是唯一的长度为1的特征值，且 $\lambda=1$代数重数和几何重数均为1，存在唯一随机向量$\alpha_0$使得$\mathbf{A}\alpha_0=\alpha_0$,$\alpha_0$称为稳定状态向量.
这个定理是$\mathbf{Perron-Frobenius}$定理的特殊情形.

例如：

$$\mathbf{A}=\begin{pmatrix}0.80 & 0.05\\ 0.2 & 0.95\end{pmatrix},\ \alpha_0=\begin{pmatrix}0.2\\ 0.8\end{pmatrix},\mathbf{A}\alpha_{\mathbf{0}}=\alpha_{\mathbf{0}}.$$

**定理：** 设$\mathbf{A}$是一个$n$阶正Markov阵，$\alpha_0$是稳定状态向量，则

$$\lim_{m\to\infty}\mathbf{A^m}=(\alpha_{\mathbf{0}},...,\alpha_{\mathbf{0}})_{\mathbf{n\times n}}$$

# 三 正Markov矩阵

**证明：** 存在可逆矩阵$\mathbf{P}$,$\mathbf{P^{-1}AP} = \mathbf{J} = \begin{pmatrix} \mathbf{J_1} & & \\ & \ddots & \\ & & \mathbf{J_s} \end{pmatrix}$, $\mathbf{J_i}$是若当块(Jordan block).

$\mathbf{J_i} = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$ 对于特征值$\lambda = 1$, 只有一个一阶若当块，不妨记为$\mathbf{J_1}= (1)$，则$\mathbf{P^{-1}AP}=$

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \mathbf{J_2} & & \\ \vdots & & \ddots & \\ 0 & & & \mathbf{J_S} \end{pmatrix} \qquad (\mathbf{P^{-1}AP})^{\mathbf{m}} = \mathbf{P^{-1}A^mP} = \begin{pmatrix} \mathbf{J_1}^{\mathbf{m}} & & \\ & \ddots & \\ & & \mathbf{J_s}^{\mathbf{m}} \end{pmatrix}, \ \mathbf{J_i}^{\mathbf{m}} = \begin{pmatrix} \lambda_i^m & & \\ & \ddots & \\ & & \lambda_i^m \end{pmatrix}_{\mathbf{n_i \times n_i}} \qquad i \geq 2$$

# 三 正Markov矩阵

因为$|\lambda_i|<1,\ i\geq 2$， 当$m\to+\infty,\ \lambda_i^m\to 0.$ 若$n_i=1$,则 $\lim\limits_{m\to\infty}\mathbf{J_i}^{\mathbf{m}}=\mathbf{0}.$

若 $\mathrm{n}_i>1,\ \mathbf{J_i}=\lambda_\mathbf{i}\mathbf{I}+\mathbf{J_0}$,$\mathbf{J_0}=\begin{pmatrix}0 & 1 & & \\ & & \ddots & \\ & & \ddots & 1\\ & & & 0\end{pmatrix}_{\mathbf{n_i}\times\mathbf{n_i}}$, $\mathbf{J_0}^{\mathbf{n_i}}=\mathbf{0}.$ 当$m\to+\infty,\ \lim\limits_{m\to\infty}\mathbf{J_i}^{\mathbf{m}}=\mathbf{0}.$

$$\Rightarrow\lim_{m\to\infty}(\mathbf{P^{-1}AP})^{\mathbf{m}}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & & & \\ \vdots & & \mathbf{0} & \\ 0 & & & \end{pmatrix}=\begin{pmatrix}1\\ \vdots\\ 0\end{pmatrix}\begin{pmatrix}1 & \cdots & 0\end{pmatrix}=\mathbf{B}$$

# 三 正Markov矩阵

$$\Rightarrow \lim_{m\to\infty} \mathbf{A}^{\mathbf{m}} = \mathbf{PBP}^{-1} = \mathbf{C}$$极限是秩等于1的矩阵.

$\mathbf{A}$是正Markov矩阵，则$\mathbf{A}^m$也是正Markov矩阵.

$(1,\dots,1)\mathbf{A}=(1,\dots,1)\quad(1,\dots,1)\mathbf{A}^m=(1,\dots,1)$.

即$\lim\limits_{m\to\infty}\mathbf{A}^{\mathbf{m}}$是一个Markov阵.

设$\mathbf{C}=(t_1\alpha,\dots t_n\alpha)$ (注rank $\mathbf{C}=1$)

$$\mathbf{A}\mathbf{A}^{\mathbf{m}} = \mathbf{A}^{\mathbf{m+1}} \Rightarrow \mathbf{A}\lim_{\mathbf{m}\to\infty}\mathbf{A}^{\mathbf{m}} = \lim_{\mathbf{m}\to\infty}\mathbf{A}^{\mathbf{m+1}} \Rightarrow \mathbf{AC}=\mathbf{C}.$$

即$\mathbf{C}$的每一列是$\mathbf{A}$关于特征值为1的特征值，

且$\mathbf{C}$的每一列均为随机向量,因此$t_1 = t_2 = \dots = t_n = 1, \alpha = \alpha_0$

# 三 正Markov矩阵

例如：

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{4} & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \mathbf{A} = \begin{pmatrix} \frac{8}{9} \\ \frac{2}{3} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{8}{9} \\ \frac{2}{3} \\ 1 \end{pmatrix} N(\mathbf{A} - \mathbf{I_3}) = \{c \begin{pmatrix} \frac{8}{23} \\ \frac{6}{23} \\ \frac{9}{23} \end{pmatrix} \mid c \in \mathbb{R}\} \Rightarrow \alpha_0 = \frac{1}{23} \begin{pmatrix} 8 \\ 6 \\ 9 \end{pmatrix} \Rightarrow$$

$$\lim_{m\to\infty} \mathbf{A^m} = \mathbf{\frac{1}{23}} \begin{pmatrix} 8 & 8 & 8 \\ 6 & 6 & 6 \\ 9 & 9 & 9 \end{pmatrix}.$$

# 三 正Markov矩阵

例如, 对于城市和郊区的人口流动模型, $\begin{pmatrix} u_i \\ v_i \end{pmatrix}$ 表示第$i$年 城市和农村的人口.

人口流动规律为：每年80%的城市人口不动，20%的城市人口流动到郊区；每年95%的郊区人口不动，5% 的郊区人口流动到城市，则

$$u_{i+1} = 0.8u_i + 0.05v_i$$
$$v_{i+1} = 0.2u_i + 0.95v_i$$

即 $\begin{pmatrix} u_{i+1} \\ v_{i+1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.05 \\ 0.2 & 0.95 \end{pmatrix} \begin{pmatrix} u_i \\ v_i \end{pmatrix}$

令$\mathbf{A} = \begin{pmatrix} 0.8 & 0.05 \\ 0.2 & 0.95 \end{pmatrix}$,则 $\mathbf{A}\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$， $\mathbf{A}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = 0.75\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

# 三 正Markov矩阵

设城市和郊区的总人口为S，且$u_1 = 0.9S, v_1 = 0.1S$. 设 $\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = [a\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} + b\begin{pmatrix} -1 \\ 1 \end{pmatrix}]S$

$\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \mathbf{A}\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = a\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} + 0.75b\begin{pmatrix} -1 \\ 1 \end{pmatrix} \Rightarrow \mathbf{f_k}=\mathbf{A^{k+1}}\begin{pmatrix} u_1 \\ v_1 \end{pmatrix}=[a\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}+0.75^k b\begin{pmatrix} -1 \\ 1 \end{pmatrix}]\mathrm{S}.$

$\Rightarrow \mathbf{f_k} \to a\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}S = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}S.$

# 三 正Markov矩阵

一般地，设$\mathbf{A}$是一个正Markov矩阵，设$\lambda_1=1$对应的特征向量(各分量之和为1)$\mathbf{x_1}$称为稳定状态(Steady state). 设$\mathbf{f_1}=a_1\mathbf{x_1}+\ldots+a_n\mathbf{x_n}$,设$\mathbf{f_{k+1}}=\mathbf{A^k f_1}$,则$\lim\limits_{k\to\infty}\mathbf{f_{k+1}}=a_1\mathbf{x_1}$.

若$\mathbf{f_1}$满足各分量之和为1，则$\mathbf{f_k}$也满足这个性质.显然，$a_1\mathbf{x_1}$也满足分量之和为1 $\Rightarrow a_1=1$. 即$\mathbf{f}_\infty=\mathbf{x_1}$.

注：若$\mathbf{A}$是非正Markov矩阵，则可能存在不止一个特征值的模长为1，从而可能没有稳定状态.例如：

$\mathbf{A}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$，有$\lambda_1=1,\lambda_2=-1$. Markov矩阵特征值的性质能部分地推广到非负矩阵.我们考虑正矩阵的情形.

# 四正矩阵

设$\mathbf{A}=(\mathbf{a_{ij}})_{\mathbf{n\times n}}$，若$a_{ij}>0,\forall i,j=1,...,n$，则称$\mathbf{A}$为正矩阵.正Markov矩阵是正矩阵的特殊情形.

**定义：** 一个方阵$\mathbf{A}$的谱半径(记作$\rho(\mathbf{A})$)是$\mathbf{A}$的特征值长度的最大值.例如：$\mathbf{A}=\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$，$\rho(\mathbf{A})=\mathbf{1}$.

# 四正矩阵

**Perron-Frobenius定理：** 设$\mathbf{A} = (\mathbf{a_{ij}})_{\mathbf{n}\times\mathbf{n}}$是一个正矩阵，则

(a)$\rho(\mathbf{A})$是$\mathbf{A}$的一个特征值，它有正特征向量;

(b)$\rho(\mathbf{A})$是$\mathbf{A}$的长度等于$\rho(\mathbf{A})$的唯一特征值;

(c)$\rho(\mathbf{A})$的几何重数等于1;

(d)$\rho(\mathbf{A})$的代数重数等于1。

# 四正矩阵

证明:

(a)设$\lambda$是$\mathbf{A}$的一个特征值，$|\lambda| = \rho(\mathbf{A})$.

设$\alpha = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$, $z_i \in \mathbb{C}$, $\alpha \neq 0$, $\mathbf{A}\alpha = \lambda\alpha$.考虑$\alpha^+ = \begin{pmatrix} |z_1| \\ \vdots \\ |z_n| \end{pmatrix}$,

检查$\mathbf{A}\alpha^+ \geq \rho(\mathbf{A})\alpha^+$,若$\mathbf{A}\alpha^+ \neq \rho(\mathbf{A})\alpha^+$.

因为$\mathbf{A}$是正矩阵，$\mathbf{A}(\mathbf{A}\alpha^+) \geq \rho(\mathbf{A})\alpha^+$, 即$\exists\varepsilon > 0$, $\mathbf{A^2}\alpha^+ > (\mathbf{1}+\varepsilon)\rho(\mathbf{A})^{\mathbf{2}}\alpha^+$.

$\mathbf{B} = \frac{1}{\rho(\mathbf{A})^{\mathbf{2}}}\mathbf{A^2}$, 则$\rho(\mathbf{B}) = 1$，$\mathbf{B}\alpha^+ > (1+\varepsilon)\alpha^+$, $\mathbf{B}^k\alpha^+ > (1+\varepsilon)^k\alpha^+$，即$(\frac{\mathbf{B}}{1+\varepsilon})^k\alpha^+ > \alpha^+$.
但$\rho(\frac{\mathbf{B}}{1+\varepsilon}) < 1 \Rightarrow (\frac{\mathbf{B}}{1+\varepsilon})^k \to 0$，矛盾. 因此$\mathbf{A}\alpha^+ = \rho(\mathbf{A})\alpha^+$.

# 四正矩阵

(b) 在（a）中我们得到，若$\mathbf{A}\alpha = \lambda\alpha, |\lambda| = \rho(\mathbf{A})$,则$\mathbf{A}\alpha^{+} = \rho(\mathbf{A})\alpha^{+}$.
这说明 $\sum_{j=1}^{n} a_{ij}|z_{ij}| = |\sum_{j=1}^{n} a_{ij}z_j|, i = 1, ..., n$,则存在$c \neq 0 \in \mathbb{C}$,使得$c\alpha$是一个非负向量.

我们以$n = 2$来解释:

$\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \alpha = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \mathbf{A}\begin{pmatrix} |z_1| \\ |z_2| \end{pmatrix} = \rho(\mathbf{A})\begin{pmatrix} |z_1| \\ |z_2| \end{pmatrix}$. 由 $\mathbf{A}\alpha = \lambda\alpha$,

$$\begin{cases} \rho(\mathbf{A})|z_1| = |a_{11}z_1 + a_{12}z_2| \\ \rho(\mathbf{A})|z_2| = |a_{21}z_1 + a_{22}z_2| \end{cases} \Rightarrow \begin{cases} |a_{11}z_1 + a_{12}z_2| = a_{11}|z_1| + a_{12}|z_2| \\ |a_{21}z_1 + a_{22}z_2| = a_{21}|z_1| + a_{22}|z_2| \end{cases}$$

$\Rightarrow |a_{11}z_1 + a_{12}z_2|^2 = (a_{11}z_1 + a_{12}z_2)(a_{11}\bar{z}_1 + a_{12}\bar{z}_2)$.

$(a_{11}|z_1| + a_{12}|z_2|)^2 = a_{11}^2|z_1|^2 + 2a_{11}a_{12}|z_1||z_2| + a_{12}^2|z_2|^2$.

比较上述两个式子右边得：

$a_{11}a_{12}(z_1\bar{z}_2 + \bar{z}_1 z_2) = 2a_{11}a_{12}|z_1||z_2|$.

# 四正矩阵

即$2|z_1||z_2| = (z_1\bar{z}_2 + \bar{z}_1 z_2) = 2Re(z_1\bar{z}_2)$. $|z_1 z_2|^2 = z_1 z_2 \bar{z}_1 \bar{z}_2 = (z_1\bar{z}_2)(\bar{z}_1 z_2) = ||z_1\bar{z}_2||^2$.

即$||z_1\bar{z}_2|| = Re(z_1\bar{z}_2)$，取$c = \bar{z}_2$, $c\alpha = \begin{pmatrix} z_1\bar{z}_2 \\ z_2\bar{z}_2 \end{pmatrix} \geq 0$.

我们得到 $\mathbf{A}\alpha = \lambda\alpha$, $|\lambda| = \rho(\mathbf{A})$，且$\exists c \in \mathbb{C}$, $c\alpha \geq 0$.

$\mathbf{A}(c\alpha) = \lambda c\alpha \geq 0 \Rightarrow \lambda > 0$,即$\lambda = \rho(\mathbf{A})$.

# 四正矩阵

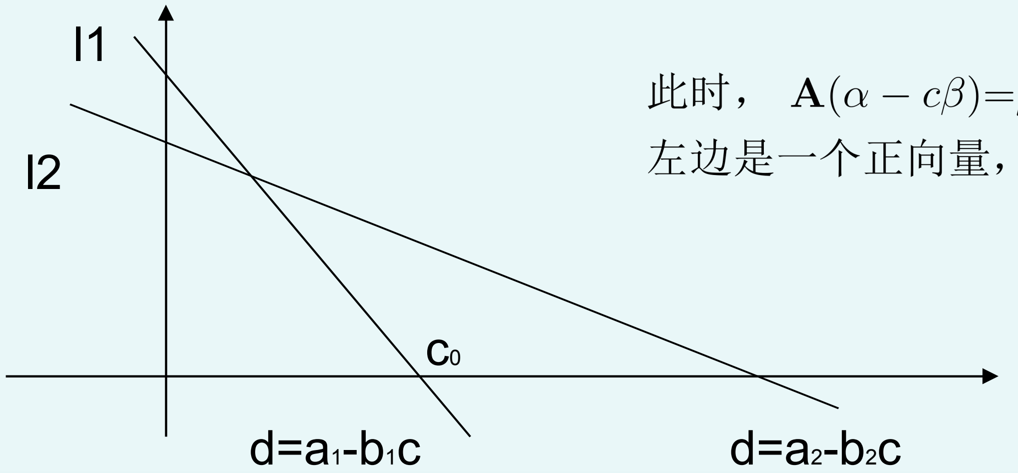
**证明：**

(c)由(a)存在$\alpha > 0$, $\mathbf{A}\alpha > \mathbf{0} = \rho(\mathbf{A})\alpha$,设存在$\beta \neq 0$, $\mathbf{A}\beta = \rho(\mathbf{A})\beta$,可以假设$\beta$是实向量.

$\exists c > 0$, $\alpha - c\beta \geq 0$,且存在某一分量为0，因此$\alpha - c\beta$不严格大于0.

我们接下来介绍如何找c,例如$n = 2$, $\alpha = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $a_i > 0$, $\beta = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$,$b_i > 0$.

考虑函数$d = a_1 - b_1 c$和$d = a_2 - b_2 c$，当$c = c_0$时，$a_1 - b_1 c_0 = 0$,$a_2 - b_2 c_0 > 0$，如下图所示.




此时，$\mathbf{A}(\alpha - c\beta)$=$\rho(\mathbf{A})(\alpha - c\beta)$，若$\alpha - c\beta \neq 0$,

左边是一个正向量，右边有零分量，矛盾.因此$\alpha - c\beta = 0$.

# 四 正矩阵

**证明：**

(d)因为$\mathbf{A}$和$\mathbf{A^T}$有相同特征值.关于$\mathbf{A^T}$,存在一正向量$\beta > 0$,$\mathbf{A^T}\beta = \rho(\mathbf{A})\beta$.考虑 $\beta^{\perp} = \{x \in$

$\mathrm{R}$, $\beta^T x = 0\}$，它是$n-1$维的.任取一组基$\alpha_2, ..., \alpha_n$，$\mathbf{A}\alpha = \rho(\mathbf{A})\alpha$, $\alpha > 0$, $\beta^{\perp}\alpha > 0 \Rightarrow \alpha \notin \beta^{\perp}$.

令 $\alpha_1 = \alpha$, $\mathbf{P} = (\alpha_\mathbf{1}, ..., \alpha_\mathbf{n})$ 可逆，$\mathbf{A}\alpha_i \in \beta^{\perp}, i \geq 2$

$$\mathbf{A}(\alpha_1, ..., \alpha_n) = (\alpha_1, ..., \alpha_n)\begin{pmatrix} \rho(\mathbf{A}) & 0 \\ 0 & \mathbf{A_1} \end{pmatrix}.$$

$\mathbf{A_1}$不包含属于$\rho(\mathbf{A})$的特征向量，否则$dimN(\mathbf{A}-\rho(\mathbf{A})\mathbf{I}) \geq 2$.

$\Rightarrow$ $\mathbf{A_1}$没有特征值$\rho(\mathbf{A})$.因此，$\rho(\mathbf{A})$的代数重数为1.

# 四 正矩阵

例子： 1973年，Leontief(里昂惕夫)获得了诺贝尔经济学奖(Nobel prize in economics)，我们使用一个简单的例子说明Leontief输入-输出模型.假设有两个产业：木材和电力. 假设生产单位电需要0.4单位电和0.2单位木材.假设生产单位木材需要0.2单位电和0.1单位木材.如下图所示：

| | 电力 | 木材 |
|---|---|---|
| 电力 | 0.4 | 0.2 |
| 木材 | 0.2 | 0.1 |

$\Rightarrow \mathbf{A} = \begin{pmatrix} 0.4 & 0.2 \\ 0.2 & 0.1 \end{pmatrix}$称为消耗矩阵(Consumption matrix).

# 四 正矩阵

令$\mathbf{P} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$是产量向量(Production vector)，即生产出$x_1$单位电和$x_2$单位木材. $\mathbf{AP} = \begin{pmatrix} 0.4x_1 + 0.2x_2 \\ 0.2x_1 + 0.1x_2 \end{pmatrix}$.

第一行表示生产出$x_1$单位电需要的成本，第二行表示生产出$x_2$单位木材需要的成本.

$\mathbf{AP}$是成本向量, $\mathbf{P} - \mathbf{AP} = (\mathbf{I} - \mathbf{A})\mathbf{P}$

设$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$表示市场对电的需要为$d_1$单位，对木材的需求为$d_2$单位.

若$(\mathbf{I} - \mathbf{A})\mathbf{P} = \mathbf{d}$,表示结余=需求(理想状态).

目标：给定消耗矩阵$\mathbf{A}$，需求向量$\mathbf{d}$,求$\mathbf{P}$,使得$(\mathbf{I} - \mathbf{A})\mathbf{P} = \mathbf{d}$.

若$\mathbf{I} - \mathbf{A}$可逆，则$\mathbf{P} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{d}$.还需要$\mathbf{P} \geq \mathbf{0}$.

定义：我们说$\mathbf{A}$是有效率的(Productive)，若$\mathbf{I} - \mathbf{A}$可逆，且$\mathbf{I} - \mathbf{A}$的逆是正矩阵.

# 四 正矩阵

**定理:** 若 $\mathbf{A}$是$n$阶正矩阵，且$\mathbf{A}$的每行元素和小于1，则$\mathbf{A}$是有效率的.

证明： 考虑$\mathbf{x}=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}_{n\times 1}$.由已知，$\mathbf{Ax} < \mathbf{x}$,则存在$0 < \varepsilon < 1$, $\mathbf{Ax} < (\mathbf{1}-\varepsilon)\mathbf{x}$.递归地，$\mathbf{A^n x} < (\mathbf{1}-\varepsilon)^{\mathbf{n}}\mathbf{x}$. $\lim\limits_{n\to\infty} \mathbf{A^n x} = \mathbf{0} \Rightarrow \lim\limits_{n\to\infty} \mathbf{A^n} = \mathbf{0}$.

另一方面，$\mathbf{A}$是正矩阵，存在$\alpha > 0$,$\mathbf{A}\alpha$=$\rho(\mathbf{A})\alpha$, $\mathbf{A^n}\alpha$=$\rho(\mathbf{A})^n\alpha$, 因为$\lim\limits_{n\to\infty} \mathbf{A^n} = \mathbf{0}$，因此$0 <$ $\rho(\mathbf{A})$$< 1$. 则$\mathbf{I}-\mathbf{A}$的特征值均不等于0，故$\mathbf{I}-\mathbf{A}$可逆，且
$(\mathbf{I}-\mathbf{A})^{-1}$=$\mathbf{I}+\mathbf{A}+\mathbf{A^2}+\ldots$
上式右边除第一项外每项均为正矩阵，因此$(\mathbf{I}-\mathbf{A})^{-1}$是正矩阵.
我们的例子满足定理条件.