《线性代数》

第七章 线性运算与线性变换

§7.3 相似矩阵与对角化问题

2019秋

杨晶 主讲

内容提要

- 对角化与相似矩阵的回顾
- 相似矩阵的基本性质
- 矩阵可对角化的条件
- 矩阵对角化的方法

(一) 问题回顾

(1) 对于给定的线性变换σ, 如果存在一组基 $\vec{\beta}_1, \vec{\beta}_2, \ldots, \vec{\beta}_n$ 使得σ在这组基下的矩阵D是一个对角矩阵, 则称线性变换σ可对角化. 中心问题: 线性变换σ可对角化的条件?

(2) 对于给定方阵A, 如果存在可逆矩阵P使得$P^{-1}AP$为对角矩阵, 则称方阵A可对角化. 中心问题: 方阵A可对角化的条件?

分析: 仅讨论n阶矩阵$A=(a_{ij})$可对角化的条件. 设存在n阶可逆矩阵P使得

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

$$\Rightarrow\ AP = P\,\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

进一步, 令 $P = (\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n)$, 代入得

$$A(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n) = (\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n)\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

$$\Rightarrow\ (A\vec{X}_1, \ldots, A\vec{X}_n) = (\lambda_1\vec{X}_1, \ldots, \lambda_n\vec{X}_n)$$

$$\Rightarrow\ A\vec{X}_i = \lambda_i\vec{X}_i,\ (i=1,2,\ldots,n).$$

- 说明我们需要专门研究满足下列条件的$\vec{X}$和λ:

$$\boxed{A\vec{X} = \lambda\vec{X}.}$$

本节用特征值理论, 来解决矩阵对角化的问题. 首先, 考虑一种特殊情况.

例1 设n阶方阵$A=(a_{ij})$有n个互异的实特征值$\lambda_1, \lambda_2, \ldots, \lambda_n$, 而 $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n$为相对应的特征向量, 于是 $A\vec{X}_i = \lambda_i \vec{X}_i$, $i = 1, 2, \ldots, n$.

将上述n个列向量排起来, 得到如下分块矩阵的形式:

$$(A\vec{X}_1, A\vec{X}_2, \ldots, A\vec{X}_n) = (\lambda_1\vec{X}_1, \lambda_2\vec{X}_2, \ldots, \lambda_n\vec{X}_n),$$

$$\Rightarrow A(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n) = (\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

令$P=(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n)$, $\Lambda=\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, 则由上一讲的定理2知,

$\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n$线性无关, 从而$P$为可逆阵. 故上式可表为 $AP = P\Lambda \quad \Rightarrow \quad P^{-1}AP = \Lambda.$

这说明, 通过可逆矩阵P, 我们可把原方阵A变化为对角阵Λ. 其中Λ包含了A的特征值的信息, 而P包含了A的特征向量的信息.

相似矩阵的概念回顾

定义1 设 A, B 是两个 n 阶方阵, 如果存在一个 n 阶可逆矩阵P, 使得 $P^{-1}AP = B$, 则称方阵 B **相似**于方阵 A, 记作 $A \backsim B$.

- 相似作为 n 阶方阵之间的一种关系, 满足以下三条性质:

(1) 自反性: $A \backsim A$; $\quad A \xrightarrow[\sim]{I} A$

(2) 对称性: 若$A \backsim B$, 则$B \backsim A$; $\quad A \xrightarrow[\sim]{P} B \Rightarrow B \xrightarrow[\sim]{P^{-1}} A$

(3) 传递性: 若$A \backsim B$, $B \backsim C$, 则$A \backsim C$. $\quad A \xrightarrow[\sim]{P_1} B, B \xrightarrow[\sim]{P_2} C \Rightarrow A \xrightarrow[\sim]{P_1P_2} C$

- 相似是一种等价关系, 自然问题是在相似形成的等价类中, 什么矩阵是最简代表元呢?
- 在例1中, 具有互异特征值的矩阵相似于对角阵, 那么矩阵可对角化充要条件是什么呢?

(二) 相似矩阵的基本性质

- 相似的两个矩阵有很多共同的特性.

性质1 若$A\backsim B$, 则$kA\backsim kB$, $A^m\backsim B^m$, 进而有$g(A)\backsim g(B)$, 对任意多项式$g(x)$成立.

证明 易证相似保持加法和数乘. 下只证方幂和多项式的结果. 由$A\backsim B$, 则存在可逆矩阵 P 使得 $P^{-1}AP=B$, 故

$$B^m=(P^{-1}AP)^m=(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)$$
$$=P^{-1}A(PP^{-1})A\cdots A(PP^{-1})AP=P^{-1}A^mP \qquad \Rightarrow A^m\backsim B^m.$$

设多项式 $g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0$, 则

$$P^{-1}g(A)P=P^{-1}(b_mA^m+b_{m-1}A^{m-1}+\cdots+b_1A+b_0I)P$$
$$=b_mP^{-1}A^mP+b_{m-1}P^{-1}A^{m-1}P+\cdots+b_1P^{-1}AP+b_0I$$
$$=b_mB^m+b_{m-1}B^{m-1}+\cdots+b_1B+b_0I=g(B) \qquad \Rightarrow g(A)\backsim g(B). \qquad \blacksquare$$

性质2 相似的矩阵有相同的秩, 即若$A\backsim B$, 则$\mathrm{r}(A)=\mathrm{r}(B)$;

进而, 相似的矩阵有相同的可逆性, 且若$A\backsim B$, 则$A^{-1}\backsim B^{-1}$.

证明 由于乘以可逆阵不改变矩阵的秩, 于是 $P^{-1}AP=B\Rightarrow \mathrm{r}(A)=\mathrm{r}(B)$.

故若A满秩, 则B也满秩, 即若A可逆, 则B也可逆, 并且此有,

$$(P^{-1}AP)^{-1}=B^{-1} \Rightarrow P^{-1}A^{-1}P=B^{-1}, \text{即 } A^{-1}\backsim B^{-1}.$$

■

性质3 相似矩阵有相同的特征多项式.

证明 设$B=P^{-1}AP$, 其中P为可逆阵, 则

$$f_B(\lambda)=|\lambda I-B|=|\lambda(P^{-1}P)-P^{-1}AP|=|P^{-1}(\lambda I-A)P|$$
$$=|P^{-1}||\lambda I-A||P|=|\lambda I-A|=f_A(\lambda). \qquad \blacksquare$$

性质4 相似矩阵具有相同的特征值.

逆命题不成立!

(证明略, 就是性质3的推论.)

性质5 相似矩阵具有相同的行列式(det)与迹(tr).

(证明略, 就是性质4的推论.)

反例: 特征值相同, 但不相似

$$A=\begin{pmatrix}1&0\\0&1\end{pmatrix},\ B=\begin{pmatrix}1&1\\0&1\end{pmatrix}$$

$$(\because\ P^{-1}AP=P^{-1}P=A=I_2)$$

下面，来讨论相似矩阵的特征向量的关系.

性质6 设$P^{-1}AP=B$，且λ_0为A的特征值，$\vec{X}$为相对应的特征向量，则 $P^{-1}\vec{X}$为B的属于特征值λ_0的特征向量.

相似不保持特征向量!

性质6 设$P^{-1}AP=B$，且λ_0为A的特征值，$\vec{X}$为相对应的特征向量，则 $P^{-1}\vec{X}$为B的属于特征值λ_0的特征向量.

证明 $A\vec{X}=\lambda_0\vec{X} \Rightarrow P^{-1}A\vec{X}=\lambda_0 P^{-1}\vec{X}$

$$\Rightarrow \boxed{P^{-1}AP}P^{-1}\vec{X}=\lambda_0 P^{-1}\vec{X}$$

$$\Rightarrow B(P^{-1}\vec{X})=\lambda_0(P^{-1}\vec{X}).$$

所以，$\underline{P^{-1}\vec{X}}$为$B$的属于特征值$\lambda_0$的特征向量. ■

多选题 1.5分

设置

相似矩阵保持下列哪些性质不变?

A 秩

B 特征值

C 特征向量

D 特征多项式

E 行列式 det

F 迹函数 tr

G 可逆性

提交

相似矩阵的简单应用

例2 设 $A=\begin{pmatrix}0.95 & 0.15\\ 0.05 & 0.85\end{pmatrix}, \Lambda=\begin{pmatrix}1 & 0\\ 0 & 0.8\end{pmatrix}, P=\begin{pmatrix}3 & -1\\ 1 & 1\end{pmatrix}.$

验证: $P^{-1}AP=\Lambda$, 并且求A^k (k为正整数).

解 为验证 $P^{-1}AP=\Lambda$, 只需要分别计算 AP 与 $P\Lambda$, 并验证它们是否相等(此步免去计算P^{-1}) .

$$AP=\begin{pmatrix}0.95 & 0.15\\ 0.05 & 0.85\end{pmatrix}\begin{pmatrix}3 & -1\\ 1 & 1\end{pmatrix}=\begin{pmatrix}3 & -0.8\\ 1 & 0.8\end{pmatrix}=\begin{pmatrix}3 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 0.8\end{pmatrix}=P\Lambda.$$

$$A^k=(P\Lambda P^{-1})^k=(P\Lambda P^{-1})(P\Lambda P^{-1})\dots(P\Lambda P^{-1})=P\Lambda^k P^{-1}$$

$$=\begin{pmatrix}3 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1^k & \\ & (0.8)^k\end{pmatrix}\frac{1}{4}\begin{pmatrix}1 & 1\\ -1 & 3\end{pmatrix}=\frac{1}{4}\begin{pmatrix}3+0.8^k & 3-3\cdot 0.8^k\\ 1-0.8^k & 1+3\cdot 0.8^k\end{pmatrix}.$$

注: 本例实际上就是上一讲例1的人口迁移问题的一种求解方式. 将 $k=20$ 代入上页结果可得:

$$A^{20}=\frac{1}{4}\begin{pmatrix}3+0.8^{20} & 3-3\cdot 0.8^{20}\\ 1-0.8^{20} & 1+3\cdot 0.8^{20}\end{pmatrix}\approx\begin{pmatrix}0.7525 & 0.7425\\ 0.2475 & 0.2575\end{pmatrix}.$$

而当$k\to\infty$时, 有

$$\lim_{k\to\infty}A^{k}=\lim_{k\to\infty}\frac{1}{4}\begin{pmatrix}3+0.8^{k} & 3-3\cdot 0.8^{k}\\ 1-0.8^{k} & 1+3\cdot 0.8^{k}\end{pmatrix}=\begin{pmatrix}3/4 & 3/4\\ 1/4 & 1/4\end{pmatrix}.$$

从而, 无需借助计算机, 我们也可算出,

$$\lim_{k\to\infty}\begin{pmatrix}x_k\\ y_k\end{pmatrix}=\lim_{k\to\infty}A^{k}\begin{pmatrix}x_0\\ y_0\end{pmatrix}=\begin{pmatrix}3/4 & 3/4\\ 1/4 & 1/4\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix}=(a+b)\begin{pmatrix}3/4\\ 1/4\end{pmatrix}.$$

上述讨论的关键在于, 验证了相似关系: $P^{-1}AP = \Lambda$ (对角阵).

问题: 任意给定n阶方阵A, 是否存在可逆阵P与对角阵Λ, 满足上述相似关系? 若存在, 如何确定P与Λ?

分析: 由本节开始的讨论知, 当A有n个互异的实特征值时, 满足上述关系的可逆阵P与对角阵Λ是存在的, 并且

- 对角阵Λ即包含了A的特征值的信息;
- 可逆阵P包含了A的特征向量的信息.

而这, 正是我们引入特征值与特征向量的概念的原因之一. 我们将讨论一般方阵可对角化的充分必要条件. 它与矩阵的特征值与特征向量有密切关系, 具体有如下的结论.

(三) 矩阵可对角化的条件

定理1 n阶方阵A的可对角化的充要条件是A有n个线性无关的特征向量.

证明 “⇒ (必要性)” 设n阶方阵A可对角化, 则有可逆阵P, 使得

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

令$P=(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n)$, $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, 则上式可写为 $AP = P\Lambda$, 故

$$左边 = AP = (A\vec{X}_1, A\vec{X}_2, \ldots, A\vec{X}_n);$$

左行右列做倍乘

$$右边 = P\Lambda = (\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = (\lambda_1\vec{X}_1, \lambda_2\vec{X}_2, \ldots, \lambda_n\vec{X}_n).$$

故得 $\quad A\vec{X}_i = \lambda_i\vec{X}_i,\ i=1, 2, \ldots, n.$

由P可逆知 $\vec{X}_i \neq \vec{0}$, 故$\vec{X}_i$为A的特征向量, 且线性无关.

"⇐ (充分性)" 设A有n个线性无关的特征向量$\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$, 设它们对应的特征值分别为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则

$$A\vec{X}_i = \lambda_i \vec{X}_i, \quad i = 1, 2, \dots, n.$$

令$P = (\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n)$, 则$P$为可逆阵, 并且有

$$AP = (A\vec{X}_1, A\vec{X}_2, \dots, A\vec{X}_n) = (\lambda_1\vec{X}_1, \lambda_2\vec{X}_2, \dots, \lambda_n\vec{X}_n)$$

$$= (\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n)\underbrace{\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}}_{\Lambda} = P\Lambda.$$

因而, $P^{-1}AP = \Lambda$, 即A可 (相似) 对角化. ■

由于属于不同特征值的特征向量是线性无关的，故当A有n个互异的特征值时，每个特征值至少有一个特征向量，这n个特征向量必线性无关，从而，我们有如下推论.

推论1 若n阶方阵A有n个互异的特征值，则A可对角化.

此外，从定理1的证明过程知，A可否对角化与它的特征值与特征向量有密切关系，具体来说，有如下结论.

推论2 若n阶方阵A可对角化，即存在可逆阵P与对角阵Λ，满足$P^{-1}AP=\Lambda$，则Λ对角线上的n个元即为A的n个特征值 $\lambda_i\ (1\le i\le n)$；而可逆阵P的n个列向量$\vec{X}_i$即为属于 $\lambda_i\ (1\le i\le n)$ 的特征向量.

实际上，推论2给出了判断和求解矩阵对角化问题的方法.

(四) 矩阵对角化的方法

先从具体例子着手.

例3 试判断下列矩阵能否对角化, 如能对角化, 求出可逆阵P与对角阵Λ, 使得 $P^{-1}AP=\Lambda$.

$$(1)\quad A=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}\qquad\qquad (2)\quad A=\begin{bmatrix}2&2&-1\\-1&0&1\\-1&0&2\end{bmatrix}$$

解: (1) 先求A的特征值. 特征多项式为

$$f_A(\lambda)=\begin{vmatrix}\lambda-2&-1&-1\\-1&\lambda-2&-1\\-1&-1&\lambda-2\end{vmatrix}=\begin{vmatrix}0&-\lambda+1&\lambda^2-4\lambda+3\\0&\lambda-2&-\lambda+1\\-1&-1&\lambda-2\end{vmatrix}$$

$$\Rightarrow \quad f_A(\lambda) = -(\lambda-1)^2 \begin{vmatrix} -1 & \lambda-3 \\ 1 & -1 \end{vmatrix} = (\lambda-1)^2(\lambda-4)$$

所以, 得到A的特征值为：$\lambda_1 = 1$ (重数为2)；$\lambda_2 = 4$.

再求A的特征向量.

对于特征值$\lambda_1 = 1$, 解齐次线性方程组 $(\lambda_1 I - A)\vec{X} = \vec{0}$,

$$(\lambda_1 I - A) = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{1} & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

得属于特征值 $\lambda_1 = 1$的两个线性无关的特征向量为:

$$\vec{X}_{11} = (-1, 1, 0)^{\mathrm{T}};\ \vec{X}_{12} = (-1, 0, 1)^{\mathrm{T}}.$$

对于特征值 $\lambda_2=4$, 解齐次线性方程组 $(\lambda_2 I-A)\vec{X}=\vec{0}$,

$$(\lambda_2 I-A)=\begin{pmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{pmatrix}\to\begin{pmatrix}1&1&-2\\0&3&-3\\0&-3&3\end{pmatrix}\to\begin{pmatrix}\boxed{1}&0&-1\\0&\boxed{1}&-1\\0&0&0\end{pmatrix}.$$

得属于特征值 $\lambda_2=4$的一个线性无关的特征向量为:

$$\vec{X}_{21}=(1,1,1)^{\mathrm{T}}.$$

因此, A存在3个线性无关的特征向量$\vec{X}_{11},\vec{X}_{12},\vec{X}_{21}$, 故$A$可对角化, 且

$$P=(\vec{X}_{11},\vec{X}_{12},\vec{X}_{21})=\begin{pmatrix}-1&-1&1\\1&0&1\\0&1&1\end{pmatrix},\quad \Lambda=P^{-1}AP=\begin{pmatrix}1&0&0\\0&1&0\\0&0&4\end{pmatrix}.$$

解: (2) 对于3阶方阵 $A=\begin{bmatrix} 2 & 2 & -1 \\ -1 & 0 & 1 \\ -1 & 0 & 2 \end{bmatrix}$, 计算其特征多项式, 得

$$f_A(\lambda)=|\lambda I_3 - A| = \begin{vmatrix} \lambda-2 & -2 & 1 \\ 1 & \lambda & -1 \\ 1 & 0 & \lambda-2 \end{vmatrix} = (\lambda-1)^2(\lambda-2).$$

得A的特征值为: $\lambda_1 = 1$ (2重根); $\lambda_2 = 2$.

对于特征值 $\lambda_1 = 1$, 解齐次线性方程组 $(\lambda_1 I - A)\vec{X} = \vec{0}$.

$$(\lambda_1 I - A) = \begin{pmatrix} -1 & -2 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & -1 \end{pmatrix} \to \begin{pmatrix} -1 & -2 & 1 \\ 0 & -1 & 0 \\ 0 & -2 & 0 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \vec{X}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

对于特征值 $\lambda_2 = 2$, 解齐次线性方程组 $(\lambda_2 I - A)\vec{X} = \vec{0}$.

$$(\lambda_2 I - A) = \begin{pmatrix} 0 & -2 & 1 \\ 1 & 2 & -1 \\ 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -2 & 1 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & -1/2 \\ 0 & 0 & 0 \end{pmatrix},$$

得属于特征值$\lambda_2 = 2$的一个线性无关的特征向量为: $\vec{X}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

对于A的任意特征向量, 要么是$\vec{X}_1$倍数, 要么是$\vec{X}_2$倍数, 故A的线性无关的特征向量最多是2个, 由定理1知, A不可对角化.

总结一下, 得到对给定n阶方阵A的对角化的具体步骤如下:

Step 1. 分解特征多项式, 求解A的特征值.

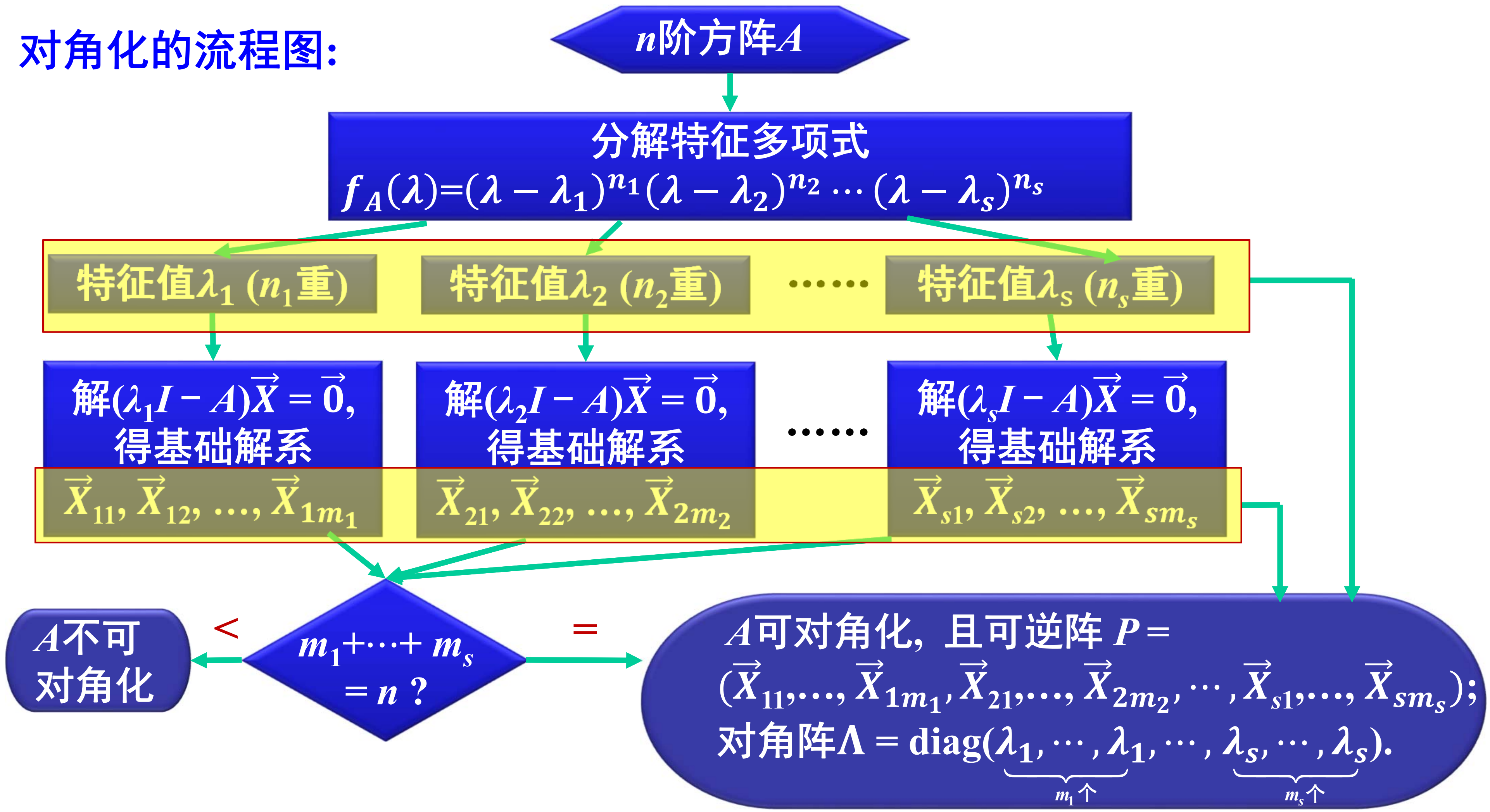
设$f_A(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$, 则$\lambda_1,\lambda_2,\cdots,\lambda_s$为$A$的互异的特征值, 其代数重数分别为$n_1,n_2,\cdots,n_s$.

Step 2. 对每个特征值λ_i ($1\leq i\leq s$) 计算特征向量. 即求解齐次线性方程组$(\lambda_i I - A)\vec{X} = \vec{0}$, 得到一组基础解系$\vec{X}_{i1}, \vec{X}_{i2}, \ldots, \vec{X}_{im_i}$, 其中$m_i$为$\lambda_i$的几何重数.

Step 3. 判断是否可对角化. 若$m_1+\cdots+m_s< n$, 则A不可对角化, 计算停止; 否则, 若$m_1+\cdots+m_s= n$, 则进行下一步.

Step 4. 得到可逆阵P与对角阵Λ, 其中$P = (\vec{X}_{11}, \vec{X}_{12}, \ldots, \vec{X}_{1m_1},\cdots,\vec{X}_{s1}, \vec{X}_{s2}, \ldots, \vec{X}_{sm_s})$, 而$\Lambda=P^{-1}AP = \mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{m_1\text{个}},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{m_s\text{个}})$.

对角化的流程图:


n阶方阵A
分解特征多项式
$f_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$
特征值λ_1 (n_1重)
特征值λ_2 (n_2重)
……
特征值λ_s (n_s重)
解$(\lambda_1 I-A)\vec{X}=\vec{0}$, 得基础解系
解$(\lambda_2 I-A)\vec{X}=\vec{0}$, 得基础解系
……
解$(\lambda_s I-A)\vec{X}=\vec{0}$, 得基础解系
$\vec{X}_{11}, \vec{X}_{12}, \dots, \vec{X}_{1m_1}$
$\vec{X}_{21}, \vec{X}_{22}, \dots, \vec{X}_{2m_2}$
$\vec{X}_{s1}, \vec{X}_{s2}, \dots, \vec{X}_{sm_s}$
$m_1+\cdots+m_s=n$?
<
=
A不可对角化
A可对角化，且可逆阵 $P=(\vec{X}_{11},\dots,\vec{X}_{1m_1},\vec{X}_{21},\dots,\vec{X}_{2m_2},\cdots,\vec{X}_{s1},\dots,\vec{X}_{sm_s})$;
对角阵$\Lambda=\mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{m_1个},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{m_s个})$.

注: (i) 简而言之, 对每个特征值λ_i, 找出属于λ_i的个数最多的线性无关的特征向量组(m_i个, **局部上**), 合并起来就得到A的个数最多的线性无关的特征向量组 (**整体上**).

(ii) 对每个特征值 λ_i, 其几何重数m_i可取到多大呢?

可以证明, m_i的上界是 n_i, 即

定理2 对n阶方阵A的每个特征值λ_i, 其几何重数不超过代数重数, 即$m_i \le n_i$ $(\forall i)$.

又因为 $n_1 + n_2 + \cdots + n_s = n$, 可得到以下矩阵可对角化的另一个充要条件:

定理3 n阶方阵A的可对角化 $\Longleftrightarrow$ $m_i = n_i$ $(\forall i)$.

例如, 在本节之前的例3(2)中, 对特征值 $\lambda_1 = 1$, 它是特征多项式的二重根, 故 $n_1 = 2$; 而用Gauss消去法得到 $\mathrm{r}(\lambda_1 I - A) = 2$, 因此 $m_1 = \dim \mathrm{N}(\lambda_1 I - A) = n - \mathrm{r}(\lambda_1 I - A) = 1$. 从而, 由$m_1 < n_1$ 得, A不可对角化.

定理2的证明:

设$\vec{x}_{i1},\vec{x}_{i2},\dots,\vec{x}_{im_i}$为特征值$\lambda_i$的特征子空间$V_{\lambda_i}(\subseteq \mathbb{C}^n)$的一组基, 将其扩充为$\mathbb{C}^n$的一组基$\{\vec{x}_{i1},\vec{x}_{i2},\dots,\vec{x}_{im_i},\vec{y}_1,\vec{y}_2,\dots,\vec{y}_{n-m_i}\}$, 则

$$A\left(\vec{x}_{i1},\vec{x}_{i2},\dots,\vec{x}_{im_i},\vec{y}_1,\vec{y}_2,\dots,\vec{y}_{n-m_i}\right)$$
$$=\left(\vec{x}_{i1},\vec{x}_{i2},\dots,\vec{x}_{im_i},\vec{y}_1,\vec{y}_2,\dots,\vec{y}_{n-m_i}\right)\begin{pmatrix}\lambda_i I_{m_i} & * \\ O & A_1\end{pmatrix}$$

其中, A_1是 $n-m_i$阶方阵, 令$P=\left(\vec{x}_{i1},\vec{x}_{i2},\dots,\vec{x}_{im_i},\vec{y}_1,\vec{y}_2,\dots,\vec{y}_{n-m_i}\right)$, 则

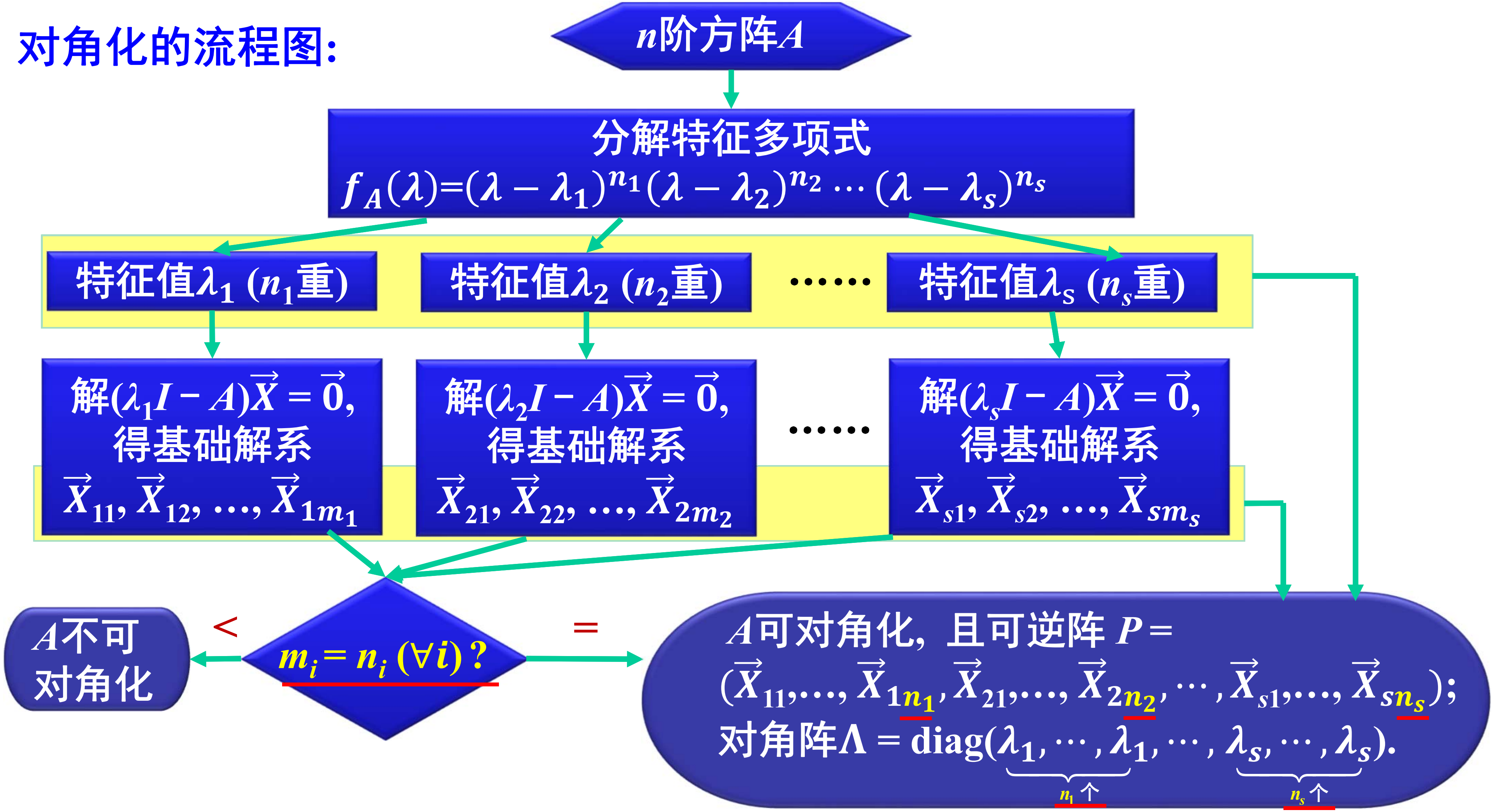
$$AP=P\begin{pmatrix}\lambda_i I_{m_i} & * \\ O & A_1\end{pmatrix} \Rightarrow P^{-1}AP=\begin{pmatrix}\lambda_i I_{m_i} & * \\ O & A_1\end{pmatrix} \Rightarrow A\sim\begin{pmatrix}\lambda_i I_{m_i} & * \\ O & A_1\end{pmatrix}.$$

由相似的矩阵有相同的特征多项式, 以及准三角阵特征多项式的结论知:

$$(\lambda-\lambda_i)^{m_i} \mid f_A(\lambda) \ =(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$$

故得: $m_i\le n_i\ (\forall i=1,2,\dots,s)$. ■

对角化的流程图:


n阶方阵A
分解特征多项式
$f_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$
特征值λ_1 (n_1重)
特征值λ_2 (n_2重)
……
特征值λ_s (n_s重)
解$(\lambda_1 I-A)\vec{X}=\vec{0}$, 得基础解系 $\vec{X}_{11}, \vec{X}_{12}, \ldots, \vec{X}_{1m_1}$
解$(\lambda_2 I-A)\vec{X}=\vec{0}$, 得基础解系 $\vec{X}_{21}, \vec{X}_{22}, \ldots, \vec{X}_{2m_2}$
……
解$(\lambda_s I-A)\vec{X}=\vec{0}$, 得基础解系 $\vec{X}_{s1}, \vec{X}_{s2}, \ldots, \vec{X}_{sm_s}$
$m_i=n_i\ (\forall i)$?
$<$
A不可对角化
$=$
A可对角化, 且可逆阵 $P=$
$(\vec{X}_{11},\ldots,\vec{X}_{1n_1},\vec{X}_{21},\ldots,\vec{X}_{2n_2},\cdots,\vec{X}_{s1},\ldots,\vec{X}_{sn_s})$;
对角阵$\Lambda=\mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{n_1个},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{n_s个})$.

注: (iii) 再次, 来讨论对角化后所得可逆阵P与对角阵Λ的关系和唯一性.

若A可对角化, 则 $m_i = n_i\ (\forall i)$, 且

$$\text{对角阵 } \Lambda = \mathrm{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}_{n_1\text{个}}, \underbrace{\lambda_2, \cdots, \lambda_2}_{n_2\text{个}}, \cdots\cdots, \underbrace{\lambda_s, \cdots, \lambda_s}_{n_s\text{个}}).$$

$$\text{可逆阵 } P = (\vec{X}_{11}, \cdots, \vec{X}_{1n_1}, \vec{X}_{21}, \cdots, \vec{X}_{2n_2}, \cdots, \vec{X}_{s1}, \cdots, \vec{X}_{sn_s});$$

① P与Λ 的列之间有对应原则

② Λ 主对角线上的元素为全体特征值, 可以任意交换次序, 故Λ在不计主对角线上元素次序的意义下是唯一的.

$n!/(n_1! \cdots n_s!)$种

③ 当Λ 取定以后, P的列按①的对应原则来取. 但在同一个特征值下, 基础解系的次序可交换; 不仅如此, 每组基础解系也有不同取法, 即V_{λ_i}可有不同的基. 因此P 的结果有约束但并不唯一.

$n!$种

矩阵对角化的简单应用

矩阵对角化的一个重要应用是用来简化矩阵方幂的计算: 若A可对角化, 即存在可逆阵P与对角阵Λ, 使得

$$P^{-1}AP = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

两边作k次方, 得

$$(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP) = P^{-1}A^kP = \mathrm{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k).$$

得到计算矩阵的方幂的公式:

$$\boxed{A^k = P\,\mathrm{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k)\,P^{-1}.}$$

例3(1)续 设 $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$, 求$A^k$.

解: 之前已经将A对角化, 得 $P^{-1}AP=\Lambda$, 其中

$$P=\begin{pmatrix}-1&-1&1\\1&0&1\\0&1&1\end{pmatrix},\quad \Lambda=\begin{pmatrix}1&0&0\\0&1&0\\0&0&4\end{pmatrix}.$$

$$(P\,|\,I_3)=\left(\begin{array}{ccc|ccc}-1&-1&1&1&0&0\\1&0&1&0&1&0\\0&1&1&0&0&1\end{array}\right)\to\left(\begin{array}{c|c}\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}&\overbrace{\begin{matrix}-1/3&2/3&-1/3\\-1/3&-1/3&2/3\\1/3&1/3&1/3\end{matrix}}^{P^{-1}}\end{array}\right),$$

$$\Rightarrow A^k=P\begin{pmatrix}1^k&0&0\\0&1^k&0\\0&0&4^k\end{pmatrix}P^{-1}=\frac{1}{3}\begin{pmatrix}2+4^k&-1+4^k&-1+4^k\\-1+4^k&2+4^k&-1+4^k\\-1+4^k&-1+4^k&2+4^k\end{pmatrix}.$$

与A同为对称阵.

例4 (1) 若A为n阶方阵, 且A可对角化, 则A^k (k为正整数), $g(A)$ (g为实系数多项式) 均可对角化.

(2) 进一步, 若A还可逆, 则A^{-1}, A^*也可对角化.

证明: (1) 由A可对角化, 则存在可逆阵P与对角阵Λ, 使得

$$P^{-1}AP = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

故 $$A^k = P\,\mathrm{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k)\,P^{-1}, \Rightarrow A^k \text{ 可对角化.}$$

再由矩阵乘法的分配律有:

$$g(A) = P\,\mathrm{diag}(g(\lambda_1), g(\lambda_2), \cdots, g(\lambda_n))\,P^{-1}, \Rightarrow g(A) \text{ 可对角化.}$$

(2) 若A可逆, 则λ_i均不为0, 两边求逆可得

$$P^{-1}A^{-1}P = \mathrm{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}), \qquad \Rightarrow A^{-1}\text{可对角化.}$$

由 $A^* = |A|A^{-1}$ 得: $P^{-1}A^*P = |A|\,\mathrm{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}), \Rightarrow A^*$ 可对角化.

■ (留作习题)

例4 设A为n阶方阵，满足 $A^2=I$，证明: A可对角化.

证明: 设λ是A的任一特征值, $\vec{X}$是λ所属的特征向量, 则

$$A^2\vec{X}=\lambda^2\vec{X}=\vec{X}, \ \Rightarrow \ (\lambda^2-1)\vec{X}=\vec{0} \ \ (\vec{X}\neq\vec{0}) \ \ \Rightarrow \lambda=\pm 1.$$

设$\lambda_1=1$, $\lambda_2=-1$, 且m_1, m_2为对应的几何重数, 则

$$m_1+m_2=(n-\mathrm{r}(A-I))+(n-\mathrm{r}(A+I)) \ =2n-[\mathrm{r}(I-A)+\mathrm{r}(I+A)] \ .$$

由矩阵秩的相关结论知

$$n=\mathrm{r}(2I)=\mathrm{r}\big[(I-A)+(I+A)\big]\leq \mathrm{r}(I-A)+\mathrm{r}(I+A)\leq n,$$

其中, 第二个小于等于号是因为 $(I-A)(I+A)=O$. 故, 有

$$\mathrm{r}(A-I)+\mathrm{r}(A+I)=n, \ \Rightarrow \ \ m_1+m_2=n.$$

从而，A可对角化. ■

例5 简单的种群增长问题

案例问题： 经过统计，某地区猫头鹰和森林鼠的数量具有如下规律：每个月只有一半的猫头鹰可以存活；老鼠的数量每个月会增加10%．如果老鼠充足(数量为R)，则下个月猫头鹰的数量将会增加$0.4R$．平均每个月每只猫头鹰的捕食会导致104只老鼠死亡. 试确定该系统的长期演化情况.

- 上述例3中我们用了两种方法，但方法二相对比方法一简单，所以对于本例，我们只用方法二来求解.

模型准备： 不考虑其他因素对猫头鹰和森林鼠的数量的影响.

模型建立： 设猫头鹰和森林鼠在时刻k的数量为$\vec{x}_k=\begin{pmatrix}O_k\\R_k\end{pmatrix}$，其中$k$是以月份为单位的时间，$O_k$为研究区域中猫头鹰的数量 (只)，$R_k$为老鼠的数量 (千只)，则

$$\begin{cases}O_{k+1}=0.5O_k+0.4R_k,\\R_{k+1}=-0.104O_k+1.1R_k,\end{cases}\Rightarrow\begin{pmatrix}O_{k+1}\\R_{k+1}\end{pmatrix}=\begin{pmatrix}0.5&0.4\\-0.104&1.1\end{pmatrix}\begin{pmatrix}O_k\\R_k\end{pmatrix}.$$

目的就是分析 $\vec{x}_k=(O_k,R_k)^{\mathrm{T}}$的变化趋势.

模型求解： 我们直接用方法二，令 $A=\begin{pmatrix}0.5&0.4\\-0.104&1.1\end{pmatrix}$

经计算得,矩阵A的两个特征值为: $\lambda_1=1.02$，$\lambda_2=0.58$，对应的特征向量分别为: $\vec{v}_1=(10,\ 13)^T$，$\vec{v}_2=(5,\ 1)^T$.

设初始向量$\vec{x}_0$可以表为: $\vec{x}_0=c_1\vec{v}_1+c_2\vec{v}_2$，于是对于$k\geq 0$, 有

$$\begin{pmatrix} O_k \\ R_k \end{pmatrix} = \vec{x}_k = A^k(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1\lambda_1^k\vec{v}_1 + c_2\lambda_2^k\vec{v}_2$$

$$= c_1(1.02)^k\binom{10}{13} + c_2(0.58)^k\binom{5}{1}.$$

当$k \to \infty$时, $(0.58)^k \to 0$，假定$c_1 > 0$，则对于充分大的k，$\vec{x}_k$近似地等于$(1.02)^k c_1\vec{v}_1$，即

$$\vec{x}_k \approx (1.02)^k c_1\binom{10}{13} \approx 1.02\vec{x}_{k-1}.$$

问题： 经过上述分析, 你能得到是什么结论? 这是一个稳定的生态系统吗?

多选题 1分

设置

当k充分大时，由算式$\begin{pmatrix} O_k \\ R_k \end{pmatrix} = \vec{x}_k \approx (1.02)^k c_1 \binom{10}{13} \approx 1.02\vec{x}_{k-1}$，说明经过充分长的时间后，该生态系统有什么情况?

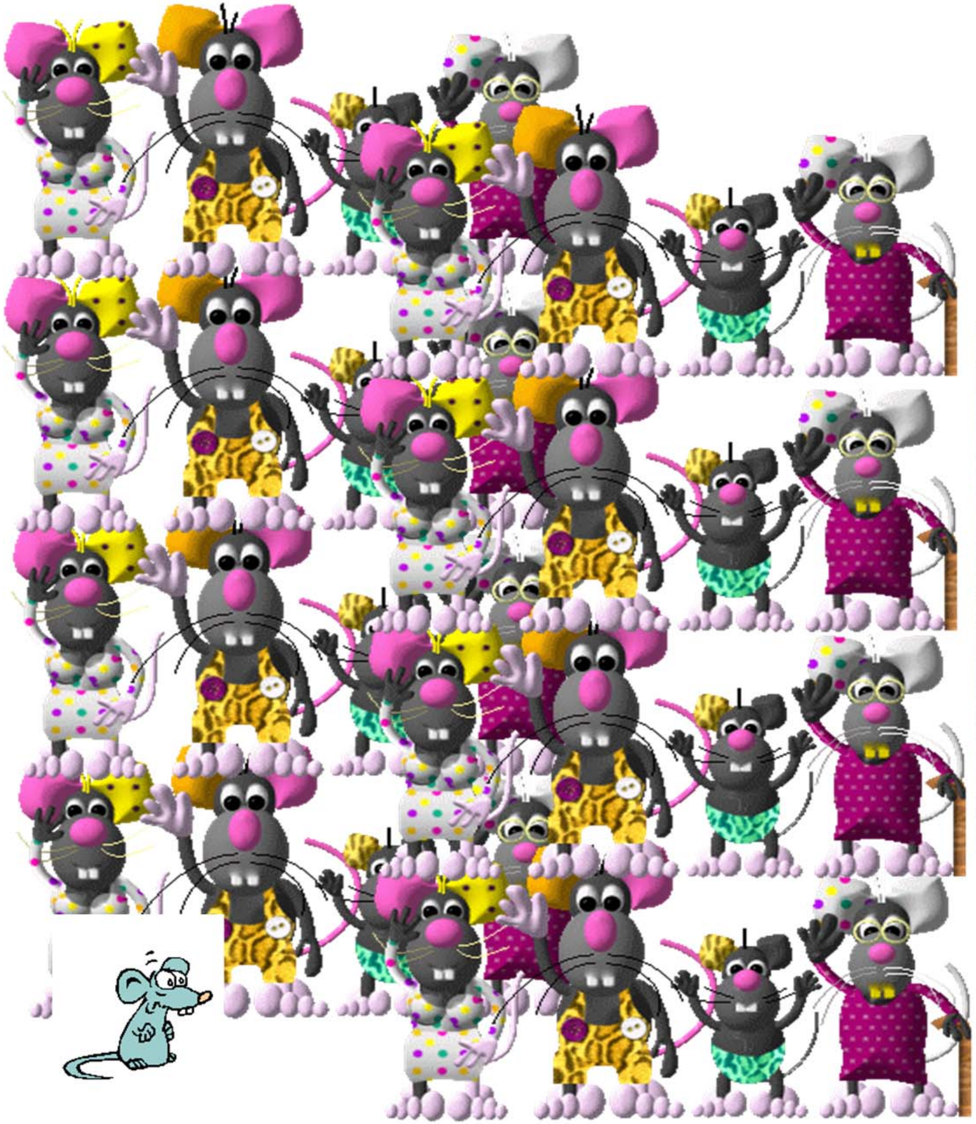
A 猫头鹰和老鼠的数量都有2%的月增长率

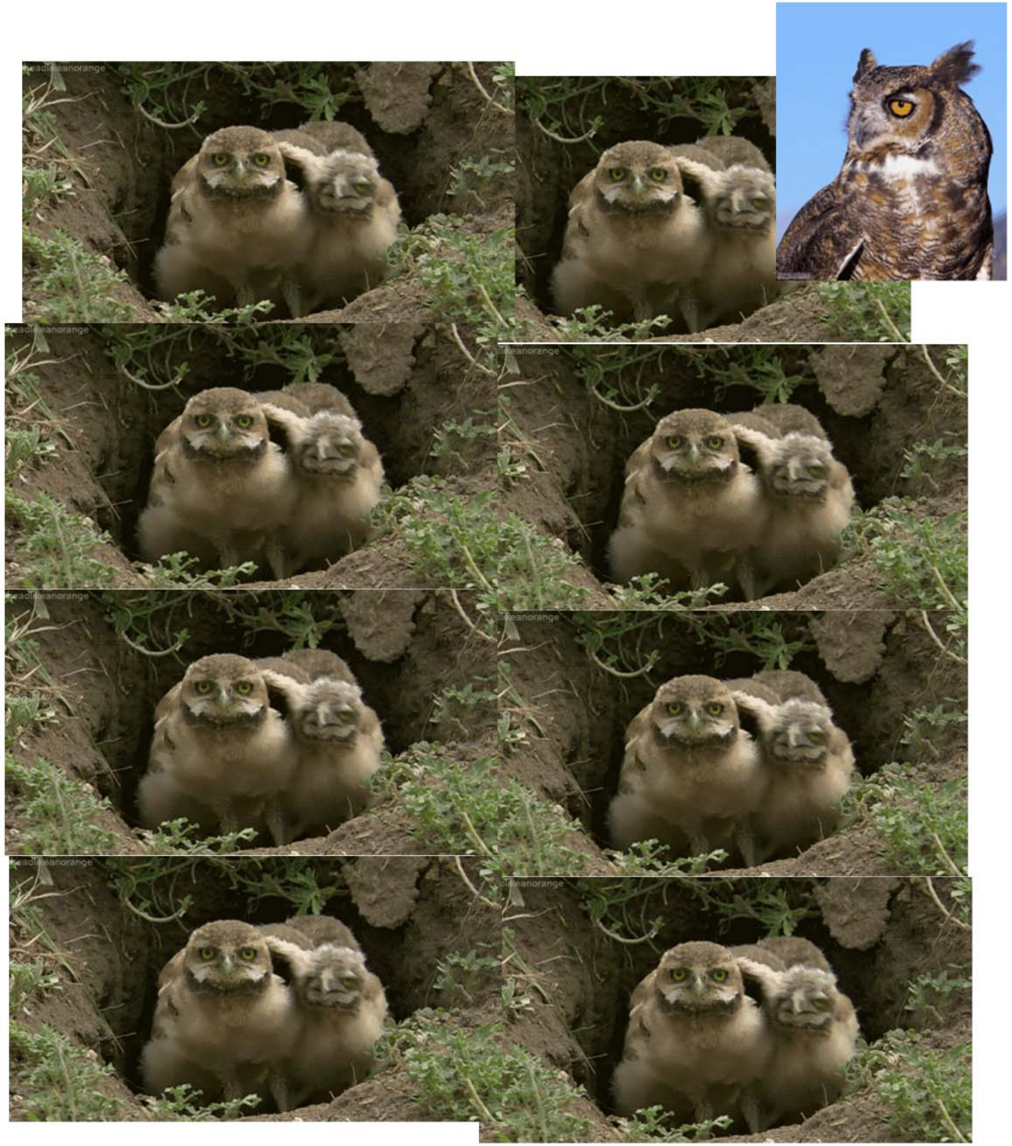
B O_k与R_k的比值约为10:13，即每10只猫头鹰对应着约13000只老鼠

C O_k与R_k的比值就是绝对值较大的特征值对应的特征向量$\vec{v}_1$的斜率.

D 这不是一个稳定的系统，长期发展下去，猫头鹰和老鼠的数量会膨胀溢出

提交





$$\begin{pmatrix} O_k \\ R_k \end{pmatrix} = \vec{x}_k = A^k(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1\lambda_1^k\vec{v}_1 + c_2\lambda_2^k\vec{v}_2$$

$$= c_1(1.02)^k\binom{10}{13} + c_2(0.58)^k\binom{5}{1}.$$

当$k \to \infty$时, $(0.58)^k \to 0$, 假定$c_1 > 0$，则对于充分大的k, $\vec{x}_k$近似地等于$(1.02)^k c_1\vec{v}_1$, 即

$$\vec{x}_k \approx (1.02)^k c_1\binom{10}{13} \approx 1.02\vec{x}_{k-1}.$$

注1: 相较计算A^k的方法, 本例中使用了将初始向量用特征向量线性表示的方法, 后者明显更简单, 原因在于避开了计算A^k和P^{-1}, 而是直奔主题, 直接计算所需的结果 $(x_k, y_k)^T$.

注2: 矩阵左乘在特征向量上, 仅为简单的倍乘变换 (伸缩变换)——而这就是特征向量的实际意义. 此外, 左乘A多次迭代后, 绝对值较大的特征值逐渐起到主导作用, 属于该特征值的特征向量反映了系统长期发展的比例.

本讲小结1

- **相似矩阵的概念**

  — ∃可逆阵P, s.t. $P^{-1}AP=B$; $A\backsim B$为等价关系.

- **相似矩阵的性质**

  — 矩阵相似 $\Rightarrow$ 矩阵的多项式相似

  — 相似保持矩阵的秩及可逆性不变

  — 相似保持矩阵的特征多项式, 特征值, 行列式, 迹不变

  — 相似矩阵的特征向量的关系: $\vec{X} \rightarrow P^{-1}\vec{X}$

本讲小结2

- 矩阵可对角化的条件 —— (i) $m_1+\cdots+m_s=n$, (ii) $m_i=n_i\ (\forall i)$
- 矩阵对角化的方法 —— (i) 对应原则, (ii) 非唯一性

对角阵 $\Lambda=\operatorname{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{n_1\text{个}},\underbrace{\lambda_2,\cdots,\lambda_2}_{n_2\text{个}},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{n_s\text{个}})$

可逆阵 $P=(\vec{X}_{11},\ldots,\vec{X}_{1n_1},\vec{X}_{21},\ldots,\vec{X}_{2n_2},\cdots,\vec{X}_{s1},\ldots,\vec{X}_{sn_s})$

- 矩阵对角化的简单应用 —— $A^k=P\operatorname{diag}(\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k)\,P^{-1}$.

—— 离散Markov过程: 长期演化的系统.