§ 20 矩阵的对角化

20.1 矩阵可对角化的条件

设 $n \times n$ 矩阵 A 有 n 个线性无关的特征向量$\mathbf{x}_1, \cdots, \mathbf{x}_n$, 令 $S = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$, 则 $S^{-1}AS$是一个对角矩阵 Λ, 其对角元素是 A 的特征值：

$$S^{-1}AS = \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

20.1 矩阵可对角化的条件

事实上，$AS = A(\mathbf{x}_1, \cdots, \mathbf{x}_n) = (\lambda_1\mathbf{x}_1, \cdots, \lambda_n\mathbf{x}_n)$

$$= (\mathbf{x}_1, \cdots, \mathbf{x}_n)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

于是 $AS = S\Lambda$.

因 S 可逆，故 $S^{-1}AS = \Lambda$.

20.1 矩阵可对角化的条件

若存在可逆矩阵 S, 使 $S^{-1}AS$ 为对角矩阵，则称矩阵 A 是可对角化的(diagonalized).

由上面的分析知，反之也成立. 故有

定理：$n \times n$ 矩阵 A 可对角化的充要条件是 A 有 n 个线性无关的特征向量.

20.1 矩阵可对角化的条件

例：$A=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$ 的特征值为 $\lambda_1=\lambda_2=0.$

$$A\mathbf{x}=\mathbf{0}\Longrightarrow \mathbf{x}=c\begin{pmatrix}1\\0\end{pmatrix},c\in\mathbb{R}.$$

故 A 只有 1 个线性无关的特征向量，因此 A 不能对角化.

20.1 矩阵可对角化的条件

定理：设$\lambda_1, \cdots, \lambda_k$ 是A 的互异特征值，$\mathbf{x}_1, \cdots, \mathbf{x}_k$ 是相应特征向量. 则 $\mathbf{x}_1, \cdots, \mathbf{x}_k$ 线性无关.

证明：设 $c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k = \mathbf{0}$.

两边左乘A, 得 $c_1\lambda_1\mathbf{x}_1 + \cdots + c_k\lambda_k\mathbf{x}_k = \mathbf{0}$.

再左乘 A, 得 $c_1\lambda_1^2\mathbf{x}_1 + \cdots + c_k\lambda_k^2\mathbf{x}_k = \mathbf{0}$.

不断左乘 A, 直到得 $c_1\lambda_1^{k-1}\mathbf{x}_1 + \cdots + c_k\lambda_k^{k-1}\mathbf{x}_k = \mathbf{0}$.

故有

$$(c_1\mathbf{x}_1, \cdots, c_k\mathbf{x}_k)\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{k-1} \end{pmatrix} = 0.$$

20.1 矩阵可对角化的条件

左边第二个矩阵的行列式 $= Vandermonde$ 行列式

$$= \prod_{1\le j<i\le k} (\lambda_i - \lambda_j) \neq 0.$$

因此该矩阵可逆，故 $(c_1\mathbf{x}_1, \cdots, c_k\mathbf{x}_k) = 0.$

由于特征向量 $\mathbf{x}_1, \cdots, \mathbf{x}_k$ 均为非零向量，故 $c_1 = \cdots = c_k = 0.$

所以 $\mathbf{x}_1, \cdots, \mathbf{x}_k$ 线性无关.

20.1 矩阵可对角化的条件

推论：具有 n 个两两互异特征值的 $n\times n$ 矩阵可以对角化.

但若矩阵有相同特征值，其也可能对角化.

例: $I=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$有重特征值 $\lambda_1=\lambda_2=1$. 任何可逆矩阵 S 都使 $S^{-1}IS$ 是对角阵. 这反映了所有非零向量都是单位矩阵的特征向量.

20.2 特征值的代数重数和几何重数

定义：设 $det(A-\lambda I)=(\lambda_1-\lambda)^{n_1}\cdots(\lambda_k-\lambda)^{n_k}$，其中 $\lambda_i\neq\lambda_j(i\neq j)$. 称 n_i 为特征值 λ_i 的代数重数(algebraic multiplicity)，记作 $AM(\lambda_i)=n_i$. 称 $dimN(A-\lambda_i I)$ 为特征值 λ_i 的几何重数(geometric multiplicity)，记作 $GM(\lambda_i)=dimN(A-\lambda_i I)$.

例：$A=\begin{pmatrix}0&1\\0&0\end{pmatrix}$. $\lambda_1=\lambda_2=0.$

$$GM=1<2=AM.$$

20.2 特征值的代数重数和几何重数

例：$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. $\lambda_1 = \lambda_2 = 1$. $GM = 2 = AM$.

例：$A = \begin{pmatrix} 6 & -1 \\ 1 & 4 \end{pmatrix}$. $det(A - \lambda I) = \begin{vmatrix} 6-\lambda & -1 \\ 1 & 4-\lambda \end{vmatrix} = (\lambda - 5)^2$.

$$\Longrightarrow \lambda_1 = \lambda_2 = 5.$$

$$A - 5I = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

$$\Longrightarrow N(A - 5I) = \{c \begin{pmatrix} 1 \\ 1 \end{pmatrix} | c \in \mathbb{R}\} \quad \Longrightarrow GM = 1 < 2 = AM$$

20.2 特征值的代数重数和几何重数

一般地，

命题：$GM(\lambda) \le AM(\lambda)$.

引理1：相似矩阵具有相同的特征多项式.

事实上，设 P 可逆，则我们有

$$det(A-\lambda I) = det(P^{-1}(A-\lambda I)P) = det(P^{-1}AP-\lambda I).$$

20.2 特征值的代数重数和几何重数

引理2：任意复方阵相似于上三角阵，且其对角元为矩阵的特征值.

证明：对方阵的阶数 n 用数学归纳法.

$n=1$ 时结论成立. 假设对 $n-1$ 阶复方阵结论成立.

对任意 n 阶复方阵 A, 设其有特征值 λ_1 及相应特征向量 $\mathbf{x}_1 \neq \mathbf{0}$.

则可将其扩充得 $\mathbb{C}^n$ 的一组基 $\mathbf{x}_1, \cdots, \mathbf{x}_n$, 有

$$A(\mathbf{x}_1, \cdots, \mathbf{x}_n) = (\mathbf{x}_1, \cdots, \mathbf{x}_n) \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}.$$

记 $P_1 = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$, 则有 $P_1^{-1}AP_1 = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}.$

20.2 特征值的代数重数和几何重数

对$n-1$ 阶复方阵A_1, 由归纳假设, 存在可逆阵 Q, 使得$Q^{-1}A_1Q=T_1$ 为上三角阵.

令 $P_2=\begin{pmatrix}1 & \\ & Q\end{pmatrix}, P=P_1P_2.$

$$\Longrightarrow P^{-1}AP=P_2^{-1}P_1^{-1}AP_1P_2=P_2^{-1}\begin{pmatrix}\lambda_1 & * \\ 0 & A_1\end{pmatrix}P_2=\begin{pmatrix}\lambda_1 & * \\ 0 & T_1\end{pmatrix}=:T$$

为上三角阵.

则结论第一部分得证.

由引理1知 $det(A-\lambda I)=det(T-\lambda I)=(t_{11}-\lambda)\cdots(t_{nn}-\lambda).$

上三角阵 T 的对角元 $t_{11},\cdots,t_{nn}$为 A 的特征值.

20.2 特征值的代数重数和几何重数

命题的证明：

由引理2，A 相似于上三角阵 T，则 A 和 T 有相同特征值，且对任意特征值 $\lambda_i, GM_A(\lambda_i) = GM_T(\lambda_i)$.

因此，不妨设 A 是上三角阵，即 $A = \begin{pmatrix} a_{11} & * & * \\ & \ddots & * \\ & & a_{nn} \end{pmatrix}$.

于是 $r(A - \lambda_i I) \geq n - AM(\lambda_i)$.

故 $GM(\lambda_i) = n - r(A - \lambda_i I) \leq AM(\lambda_i)$.

20.2 特征值的代数重数和几何重数

定理：复方阵 A 可对角化$\Longleftrightarrow$ 对任意特征值 $\lambda_i, GM(\lambda_i) = AM(\lambda_i)$.

事实上，$\sum_{i=1}^{k} AM(\lambda_i) = n.$

若 $\forall i, GM(\lambda_i) = AM(\lambda_i)$, 则 $GM(\lambda_1) + \cdots + GM(\lambda_k) = n.$

故 A 有 n 个线性无关的特征向量.

从而 A 可对角化.

20.2 特征值的代数重数和几何重数

例：判断 $A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 5 & -2 \\ -2 & 4 & -1 \end{pmatrix}$ 是否可对角化，若可以求 S 使 $S^{-1}AS$ 为对角阵.

解：$det(A-\lambda I) = -(\lambda-1)^2(\lambda-3) \Longrightarrow \lambda_1 = \lambda_2 = 1, \lambda_3 = 3.$

$$A-\lambda_1 I = \begin{pmatrix} 0 & 0 & 0 \\ -2 & 4 & -2 \\ -2 & 4 & -2 \end{pmatrix} \Longrightarrow dimN(A-\lambda_1 I) = 2.$$

于是 $AM(\lambda_1) = 2 = GM(\lambda_1).$

又 $GM(\lambda_3) = AM(\lambda_3) = 1.$

因此，A 可对角化.

20.2 特征值的代数重数和几何重数

对 $\lambda_1 = \lambda_2 = 1, A - I \longrightarrow \begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$

$(A - I)\mathbf{x} = \mathbf{0}$ 的基础解系为 $\mathbf{x}_1 = (2, 1, 0)^T, \mathbf{x}_2 = (-1, 0, 1)^T.$

对 $\lambda_3 = 3, A - 3I = \begin{pmatrix} -2 & 0 & 0 \\ -2 & 2 & -2 \\ -2 & 4 & -4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$

$(A - 3I)\mathbf{x} = \mathbf{0}$ 的基础解系为 $\mathbf{x}_3 = (0, 1, 1)^T.$

20.2 特征值的代数重数和几何重数

令 $S = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$

则 $S^{-1}AS = \Lambda = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 3 \end{pmatrix}.$

20.2 特征值的代数重数和几何重数

注：可以看到，使 A 对角化的矩阵 S 不是唯一的. 一个特征向量乘以非零常数后仍是属于同一特征值的特征向量，所以若用任意非零常数乘以 S 的各列，则得一个新的使 A 对角化的矩阵. 而对于重特征值则有更大自由度. 上例中由 $\mathbf{x}_1, \mathbf{x}_2$ 的任意线性组合得到的两个线性无关的向量都可充当 S 的前两列.

20.2 特征值的代数重数和几何重数

例：设 $A=\begin{pmatrix} I_r & B \\ 0 & -I_{n-r} \end{pmatrix}$，其中 B 为 $r\times(n-r)$ 矩阵.

$$\Longrightarrow \lambda_1=\cdots=\lambda_r=1, \lambda_{r+1}=\cdots=\lambda_n=-1.$$

$A-\lambda_1 I=\begin{pmatrix} 0 & B \\ 0 & -2I_{n-r} \end{pmatrix}$的秩为$n-r \Longrightarrow GM(\lambda_1)=r=AM(\lambda_1).$

$A-\lambda_{r+1} I=\begin{pmatrix} 2I_r & B \\ 0 & 0 \end{pmatrix}$的秩为$r \Longrightarrow GM(\lambda_{r+1})=n-r=AM(\lambda_{r+1}).$

故 A 可对角化.

20.3 矩阵可对角化的应用

若矩阵 A 可对角化，则可快速计算 A^k.

例：设 $A=\begin{pmatrix}1 & 0\\ 1 & -2\end{pmatrix}$, 求 A^{10}.

解： A 的特征值 $\lambda_1=1, \lambda_2=-2$.

$\Longrightarrow A$ 可对角化.

20.3 矩阵可对角化的应用

对 $\lambda_1 = 1, A - \lambda_1 I = \begin{pmatrix} 0 & 0 \\ 1 & -3 \end{pmatrix}.$

$(A - \lambda_1 I)\mathbf{x} = \mathbf{0}$ 的基础解系为 $\mathbf{x}_1 = (3, 1)^T.$

对 $\lambda_2 = -2, A - \lambda_2 I = \begin{pmatrix} 3 & 0 \\ 1 & 0 \end{pmatrix}.$

$(A - \lambda_2 I)\mathbf{x} = \mathbf{0}$ 的基础解系为 $\mathbf{x}_2 = (0, 1)^T.$

20.3 矩阵可对角化的应用

令 $S=\begin{pmatrix}3 & 0\\ 1 & 1\end{pmatrix}$, 则 $S^{-1}AS=\begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}=\Lambda.$

故 $A=S\Lambda S^{-1}.$

$$\Longrightarrow A^{10}=S\Lambda^{10}S^{-1}=S\begin{pmatrix}1 & 0\\ 0 & (-2)^{10}\end{pmatrix}S^{-1}=\begin{pmatrix}1 & 0\\ \frac{1-2^{10}}{3} & 2^{10}\end{pmatrix}.$$

20.3 矩阵可对角化的应用

例（Markov过程）：

每年海淀区以外人口的 10% 迁入海淀区，而海淀区人口的20%迁出. 这给出一个差分方程:

设最初外部人口为 p_0, 内部人口为 q_0, 则一年以后

$$\text{外部人口}\quad p_1 = 0.9p_0 + 0.2q_0,$$

$$\text{内部人口}\quad q_1 = 0.1p_0 + 0.8q_0.$$

即

$$\begin{pmatrix} p_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} p_0 \\ q_0 \end{pmatrix}.$$

20.3 矩阵可对角化的应用

这个虚构的人口迁移过程有两个特点：(1)人口总数保持不变；(2)海淀区外部和内部的人口数不是负的. 我们称之为Markov(马尔科夫)过程.

由性质(1)，矩阵每一列元素之和为 1; 由性质(2)，矩阵元素非负. 同样 p_0, q_0, p_1, q_1 等也非负.

20.3 矩阵可对角化的应用

记 $A = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$.

$$det(A - \lambda I) = \lambda^2 - 1.7\lambda + 0.7 \implies \lambda_1 = 1, \lambda_2 = 0.7.$$

$$\lambda_1 = 1, A - \lambda_1 I = \begin{pmatrix} -0.1 & 0.2 \\ 0.1 & -0.2 \end{pmatrix} \implies \mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 0.7, A - \lambda_2 I = \begin{pmatrix} 0.2 & 0.2 \\ 0.1 & 0.1 \end{pmatrix} \implies \mathbf{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

取 $S = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$，则 $A = S\Lambda S^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.7 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{pmatrix}$.

20.3 矩阵可对角化的应用

于是我们可求 A^k 和 k 年之后的人口分布：

$$\begin{aligned}
\begin{pmatrix} p_k \\ q_k \end{pmatrix} &= A^k \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = S\Lambda^k S^{-1} \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (0.7)^k \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{pmatrix} \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \\
&= \frac{p_0 + q_0}{3} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \frac{p_0 - 2q_0}{3} (0.7)^k \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\end{aligned}$$

20.3 矩阵可对角化的应用

可以看出，经过很多年之后，$(0.7)^k$会变得非常小，从而这个解达到一个极限状态：

$$\begin{pmatrix} p_\infty \\ q_\infty \end{pmatrix} = \frac{p_0 + q_0}{3} \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

此时，总人口仍为 $p_0 + q_0$，与初始状态相同. 但在此极限状态下，总人口的 $\frac{2}{3}$ 在外部，$\frac{1}{3}$ 在内部，并且这个数据无论初始分布 p_0, q_0 怎样总成立.

20.3 矩阵可对角化的应用

注意到 $A\mathbf{u}_\infty = \mathbf{u}_\infty, \mathbf{u}_\infty = \begin{pmatrix} p_\infty \\ q_\infty \end{pmatrix}.$

即这个稳定状态是Markov矩阵 A 关于 $\lambda = 1$的特征向量.

20.3 矩阵可对角化的应用

例（Fibonacci数列）：

数列 $F_n: 0, 1, 1, 2, 3, 5, 8, 13, \cdots$ 满足规律

$$F_{k+2} = F_{k+1} + F_k.$$

这是一个差分方程.

怎样由 $F_0 = 0, F_1 = 1$ 出发，求出Fibonacci数列的通项公式呢？

20.3 矩阵可对角化的应用

令 $\mathbf{u}_k = \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix}$. 则 $\begin{cases} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{cases}$

即 $\mathbf{u}_{k+1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{u}_k = A\mathbf{u}_k.$

于是 $\mathbf{u}_k = A^k \mathbf{u}_0$. 只需求 A^k.

20.3 矩阵可对角化的应用

$$det(A-\lambda I)=\begin{vmatrix}1-\lambda & 1\\ 1 & -\lambda\end{vmatrix}=\lambda^2-\lambda-1$$

$$\Longrightarrow \lambda_1=\frac{1+\sqrt{5}}{2},\lambda_2=\frac{1-\sqrt{5}}{2}.$$

故 $A=S\Lambda S^{-1}=\begin{pmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{pmatrix}\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}\begin{pmatrix}1 & -\lambda_2\\ -1 & \lambda_1\end{pmatrix}\frac{1}{\lambda_1-\lambda_2}.$

20.3 矩阵可对角化的应用

初始值 $F_0 = 0, F_1 = 1$ 给出 $\mathbf{u}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

于是

$$\begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix} = \mathbf{u}_k = A^k \mathbf{u}_0 = S \Lambda^k S^{-1} \mathbf{u}_0$$
$$= \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{1}{\lambda_1 - \lambda_2}.$$

Fibonacci数 F_k 是这个乘积的第二个分量

$$F_k = \frac{\lambda_1^k}{\lambda_1 - \lambda_2} - \frac{\lambda_2^k}{\lambda_1 - \lambda_2} = \frac{1}{\sqrt{5}} [(\frac{1+\sqrt{5}}{2})^k - (\frac{1-\sqrt{5}}{2})^k].$$

20.3 矩阵可对角化的应用

我们希望研究由差分方程 $\mathbf{u}_{k+1} = A\mathbf{u}_k$ 描述的离散动力系统的长期行为，即 $k \to \infty$ 时解的性质.

设 A 可对角化，即存在可逆矩阵 $S = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$, 其中 $A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \mathbf{x}_i \neq \mathbf{0}$, 使 $S^{-1}AS = \Lambda$ 为对角阵.

则
$$\begin{aligned}\mathbf{u}_k &= A^k \mathbf{u}_0 = S\Lambda^k S^{-1}\mathbf{u}_0 \\ &= c_1\lambda_1^k \mathbf{x}_1 + \cdots + c_n \lambda_n^k \mathbf{x}_n,\end{aligned}$$

其中 $S^{-1}\mathbf{u}_0 = (c_1, \cdots, c_n)^T$, 即 $\mathbf{u}_0 = c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n$.

可以看出，$\mathbf{u}_k$ 的增长由因子 $\lambda_i^k (1 \leq i \leq n)$ 支配. 因此系统的稳定性依赖于 A 的特征值.

20.3 矩阵可对角化的应用

对由一个差分方程 $\mathbf{u}_{k+1} = A\mathbf{u}_k$ 定义的离散动力系统，当 A 的所有特征值 $|\lambda_i| < 1$ 时，它是稳定的(stable)，且 $\mathbf{u}_k \to \mathbf{0}$ ；当所有 $|\lambda_i| \leq 1$时，它是中性稳定的(neutrally stable)，且$\mathbf{u}_k$ 有界；而当至少有一个特征值 $|\lambda_i| > 1$ 时，它是不稳定的(unstable)，且$\mathbf{u}_k$ 是无界的.

Markov过程是中性稳定的，Fibonacci数列是不稳定的.

20.3 矩阵可对角化的应用

例：考虑差分方程 $\mathbf{u}_{k+1}=A\mathbf{u}_k$，其中 $A=\begin{pmatrix}0 & 8\\ 0 & \frac{1}{4}\end{pmatrix}$.

A 的特征值为其对角元 0 和 $\frac{1}{4}$. 故该系统是稳定的.
由任何一个初始向量 $\mathbf{u}_0$ 出发，$\mathbf{u}_{k+1}=A\mathbf{u}_k$ 的解必定最终趋向于 $\mathbf{0}$.
如：

$$\mathbf{u}_0=\begin{pmatrix}0\\ 1\end{pmatrix},\mathbf{u}_1=\begin{pmatrix}8\\ \frac{1}{4}\end{pmatrix},\mathbf{u}_2=\begin{pmatrix}2\\ \frac{1}{16}\end{pmatrix},\mathbf{u}_3=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{64}\end{pmatrix},\cdots$$

20.3 矩阵可对角化的应用

可以看到从 $\mathbf{u}_2$ 开始，$\mathbf{u}_2 = A\mathbf{u}_1 = \frac{1}{4}\mathbf{u}_1, \mathbf{u}_3 = A\mathbf{u}_2 = \frac{1}{4}\mathbf{u}_2, \cdots$

而 $\mathbf{u}_1 = A\mathbf{u}_0$ 的实际作用是，若把 $\mathbf{u}_0$ 分解成 A 的两个特征向量的和：

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 32 \\ 1 \end{pmatrix} + \begin{pmatrix} -32 \\ 0 \end{pmatrix},$$

则 $A\mathbf{u}_0$ 把属于 $\lambda = 0$ 的特征向量 $\begin{pmatrix} -32 \\ 0 \end{pmatrix}$ 化为零，而把属于 $\lambda = \frac{1}{4}$ 的特征向量 $\begin{pmatrix} 32 \\ 1 \end{pmatrix}$ 乘以 $\lambda = \frac{1}{4}$.

20.4 同时对角化

问题：给定两个 n 阶矩阵 A, B, 是否存在可逆矩阵 P, 使得 $P^{-1}AP = \Lambda_1, P^{-1}BP = \Lambda_2$ 同时为对角阵，也即 A, B 同时对角化？

命题：若 A, B 有相同特征向量矩阵 P, 使得 $P^{-1}AP = \Lambda_1$, $P^{-1}BP = \Lambda_2$ 为对角阵，则 $AB = BA$.

事实上，

$$AB = P\Lambda_1 P^{-1} P\Lambda_2 P^{-1} = P\Lambda_1\Lambda_2 P^{-1} = P\Lambda_2\Lambda_1 P^{-1} = BA.$$

20.4 同时对角化

重要的是，“逆”命题也成立. 我们不加证明地给出：

定理：若 A, B 均可对角化，且 $AB = BA$, 则 A, B 可同时对角化.

注意到，若 $A\mathbf{x} = \lambda\mathbf{x}$, 则 $AB\mathbf{x} = BA\mathbf{x} = B\lambda\mathbf{x} = \lambda B\mathbf{x}$,
故 $B\mathbf{x}$ 和 $\mathbf{x}$ 是 A 的属于同一特征值 λ 的特征向量. 看简单的情况.
假设 A 的特征值两两互异，则其所有特征子空间都是一维的. 于是
$B\mathbf{x}$ 必是 $\mathbf{x}$ 的倍数，也即 $\mathbf{x}$ 是 B 的特征向量. 从而 A, B 有公共特征向量矩阵，可同时对角化.

20.4 同时对角化

定理：对 n 阶复矩阵 A, B，若矩阵 A 的特征值两两互异，则

$$AB = BA \Longleftrightarrow A, B \text{ 可同时对角化.}$$

20.4 同时对角化

小结：

1. 矩阵 A 可对角化，指存在可逆矩阵 S, 使 $S^{-1}AS=\Lambda$为对角阵.
2. $n\times n$ 矩阵 A 可对角化$\Longleftrightarrow$ A 有 n 个线性无关的特征向量.
3. 若 $n\times n$ 复矩阵 A 有 n 个互异特征值，则 A 可对角化.
4. 复矩阵 A 可对角化 $\Longleftrightarrow$ 任意特征值的几何重数等于代数重数.
5. 设 A 可对角化, 即存在可逆阵使 $A=S\Lambda S^{-1}$, 则 $A^k=S\Lambda^k S^{-1}$.
6. 差分方程 $\mathbf{u}_{k+1}=A\mathbf{u}_k$ 的解为

$$\mathbf{u}_k=A^k\mathbf{u}_0=S\Lambda^k S^{-1}\mathbf{u}_0=c_1\lambda_1^k\mathbf{x}_1+\cdots+c_n\lambda_n^k\mathbf{x}_n,$$

其中 $\mathbf{u}_0=c_1\mathbf{x}_1+\cdots+c_n\mathbf{x}_n$, $A\mathbf{x}_i=\lambda_i\mathbf{x}_i(1\le i\le n)$.