回顾上节课内容：

1. n阶方阵A可对角化：$A = X\Lambda X^{-1}$ ， Λ为对角矩阵

2. A可对角化⟺有n个线性无关的特征向量

3. 特殊情形：如果方阵有n个互异的特征值, 则可对角化

   原因: 属于不同特征值的特征向量线性无关

4. $A \in M_n(\mathbb{C}), p_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$, 其中$\lambda_i \in \mathbb{C}$互不相同，$n_1, \ldots, n_k \geq 1, n_1 + \cdots + n_k = n$.

$AM(\lambda_i) = n_i$ 称为特征值λ_i的**代数重数** (<u>A</u>lgebraic <u>M</u>ultiplicity)

$GM(\lambda_i) = \dim \mathcal{N}(\lambda_i I_n - A)$称为$\lambda_i$的**几何重数**(<u>G</u>eometric <u>M</u>ultiplicity)

总有 $GM(\lambda_i) \leq AM(\lambda_i)$

A可对角化当且仅当对所有特征值$\lambda_i, GM(\lambda_i) = AM(\lambda_i)$.

4 给出了对角化矩阵的方法

例:

形如$J_n(\lambda_0)=\begin{bmatrix}\lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0\end{bmatrix}$的$n$阶方阵称为若尔当块（Jordan block).

讨论$J_n(\lambda_0)$是否可对角化。

$p_{J_n(\lambda_0)}(\lambda)=(\lambda-\lambda_0)^n$因此，$J_n(\lambda_0)$的特征值为$\lambda_0, AM(\lambda_0)=n.$

$$\lambda_0 I_n - J_n(\lambda_0)=\begin{bmatrix}0 & -1 & & \\ & 0 & \ddots & \\ & & \ddots & -1 \\ & & & 0\end{bmatrix}$$

$$\dim\mathcal{N}\big(\lambda_0 I_n - J_n(\lambda_0)\big)=n-rank\left(\begin{bmatrix}0 & -1 & & \\ & 0 & \ddots & \\ & & \ddots & -1 \\ & & & 0\end{bmatrix}\right)=1.$$

因此，如果$n\geq 2, J_n(\lambda_0)$不可对角化。

5.3 矩阵的相似

主要内容:

(a) 相似等价与相似标准形（Jordan标准形）

(b) Hamilton-Cayley定理

(c) 两矩阵可同时对角化

(a) 相似等价与相似标准形

定义:

对方阵A, B，如果可逆矩阵X使得$B = X^{-1}AX$，则称A**相似** (similar)于B.

A可对角化就是A相似于对角阵

例:

$\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}$相似于$\begin{bmatrix} \lambda_2 & \\ & \lambda_1 \end{bmatrix}$

$$\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}\begin{bmatrix} \lambda_2 & \\ & \lambda_1 \end{bmatrix}\begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$$

一般的，两个对角矩阵的对角元素如果只差一个排列，则它们相似

命题:

方阵的相似是一个等价关系，即：

(1)自反性：A与自身相似

(2)对称性：若A与B相似，则B与A相似

(3)传递性：若A与B相似，B与C相似, 则A与C相似

证明:

$A = I_n^{-1} A I_n$

$A = X^{-1} B X, B = X A X^{-1}$

$A = X_1^{-1} B X_1, B = X_2^{-1} C X_2$,则$A = (X_2 X_1)^{-1} C (X_2 X_1)$

问题:

1. 矩阵的哪些量是在相似关系(或相似变换)下保持的?
   即相似关系下的**不变量**
2. 相似关系下的标准形?

命题:

相似关系有如下不变量:

1. 秩

2. 特征多项式

3. 特征值，迹，行列式

4. 特征值的代数重数

5. 特征值的几何重数

证明:

X可逆，$rank\left(X^{-1}AX\right) = rank(A)$

设$B = X^{-1}AX.\ |\lambda I - B| = \left|\lambda I - X^{-1}AX\right| = \left|X^{-1}(\lambda I - A)X\right| = |\lambda I - A|$.

于是$p_B(\lambda) = p_A(\lambda)$.

由特征多项式相同得到特征值、迹、行列式、特征值的代数重数均相同

证明（续）：

特征值的几何重数

设$B = X^{-1}AX$。设λ_0为A, B的特征值。

$\lambda_0 I - B = \lambda_0 I - X^{-1}AX = X^{-1}(\lambda_0 I - A)X.$

于是，$rank(\lambda_0 I - B) = rank(\lambda_0 I - A).$

因此，$\dim\mathcal{N}(\lambda_0 I - B) = \dim\mathcal{N}(\lambda_0 I - A).$

实际上，$\boldsymbol{v} \in \mathcal{N}(\lambda_0 I - B)$当且仅当$X\boldsymbol{v} \in \mathcal{N}(\lambda_0 I - A).$

$(\lambda_0 I - B)\boldsymbol{v} = \mathbf{0} \Rightarrow \left(\lambda_0 I - X^{-1}AX\right)\boldsymbol{v} = \mathbf{0} \Rightarrow X^{-1}(\lambda_0 I - A)X\boldsymbol{v} = \mathbf{0}$

$\Rightarrow (\lambda_0 I - A)X\boldsymbol{v} = \mathbf{0}$

例:

$$\begin{bmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}\text{与}\begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}\text{特征多项式均为}(\lambda-2)^4.$$

$GM = 2$ $GM = 1$

两者不相似，因为两者对于特征值2的几何重数不同。

相似标准形（Jordan标准形）：

回顾Jordan块 $J_k(\lambda)=\begin{bmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{bmatrix}_{k\times k}$

定理:

(1)对于n阶复方阵A，存在可逆矩阵$X\in M_n(\mathbb{C})$ 使得

$X^{-1}AX=\begin{bmatrix}J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_s}(\lambda_s)\end{bmatrix}$，这里$\lambda_i$可能相同。

$\begin{bmatrix}J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_s}(\lambda_s)\end{bmatrix}$称为$A$的**Jordan标准形**（Jordan canonical form).

(2)两个矩阵相似当且仅当它们具有相同的若尔当标准形。

Jordan标准形的证明超出了课程范围，我们证明较弱版本

命题: (每个相似等价类里有一个上三角矩阵.)

A为n阶复方阵，存在可逆矩阵X使得$X^{-1}AX$为上三角阵。

证明:

取A的一个特征值λ_1。令$\boldsymbol{x}_1$为λ_1对应的一个特征向量。

将$\boldsymbol{x}_1$扩充为$\mathbb{C}^n$一组基$\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$. 记矩阵$X = [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n]$，那么

$$AX = [A\boldsymbol{x}_1, A\boldsymbol{x}_2, \dots, A\boldsymbol{x}_n] = [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n]\begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \cdots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

$$= [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n]\begin{bmatrix} \lambda_1 & \boldsymbol{a}^T \\ & A_1 \end{bmatrix} = X\begin{bmatrix} \lambda_1 & \boldsymbol{a}^T \\ & A_1 \end{bmatrix}$$

因此， $X^{-1}AX = \begin{bmatrix} \lambda_1 & \boldsymbol{a}^T \\ & A_1 \end{bmatrix}$

A_1为$n-1$阶方阵，由归纳假设A_1相似于上三角阵，即有B_1使得$B_1^{-1}A_1B_1 = U_1$.

令$B = X\begin{bmatrix}1 & \\ & B_1\end{bmatrix}$,

$$B^{-1}AB = \begin{bmatrix}1 & \\ & B_1\end{bmatrix}^{-1} X^{-1}AX\begin{bmatrix}1 & \\ & B_1\end{bmatrix} = \begin{bmatrix}1 & \\ & B_1\end{bmatrix}^{-1}\begin{bmatrix}\lambda_1 & \boldsymbol{a}^T \\ & A_1\end{bmatrix}\begin{bmatrix}1 & \\ & B_1\end{bmatrix}$$

$$= \begin{bmatrix}\lambda_1 & \boldsymbol{a}^T B_1 \\ & B_1^{-1}A_1B_1\end{bmatrix} = \begin{bmatrix}\lambda_1 & \boldsymbol{a}^T B_1 \\ & U_1\end{bmatrix}$$

(b) Hamilton-Cayley定理:

设$p_A(\lambda)$为A的特征多项式，则$p_A(A) = O_{n\times n}$.

证明思路：（转化为可对角化情形）

(1) A为对角矩阵时，可直接验证。

(2) A为可对角化时，利用相似矩阵有相同的特征多项式证明。

(3) 对一般的矩阵，想法是转化成可对角化情形。

下面简略说明（3）的证明思路

首先A可上三角化，即$A = XUX^{-1}$.

如果U对角线上的元素均不同，那么U可对角化，A也可对角化，已解决

如果U对角线上可能有相同的元素，取一列上三角矩阵U_k满足

(1) U_k对角线元素均不相同，对角线以外元素与U元素相同

(2) $U_k \to U, k \to \infty$（每个ij位置的元素构成的数列趋于U的ij位置元素）

由于上三角矩阵的特征值为对角线元素，

$p_{U_k}(\lambda) \to p_U\ (\lambda) = p_A(\lambda),\ k \to \infty$

令$A_k = XU_kX^{-1}$. $p_{A_k}(\lambda) = p_{U_k}(\lambda)$且$A_k \to A, k \to \infty$.

因此，$p_{A_k}(A_k) \to p_A(A)$.

由于A_k可对角化，$p_{A_k}(A_k) = O_{n\times n}$, 于是$p_A(A)= O_{n\times n}$.

(c) 可同时对角化:

问题:

假定A, B都是可对角化矩阵，一般说来，AB的特征值不等于A的特征值乘以B的特征值

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

$$\lambda_A=1,2, \lambda_B=1,-1; \quad \lambda_{AB}=\sqrt{2},-\sqrt{2}$$

如果有$\boldsymbol{x}$同时为A, B特征值分别为λ_A, λ_B的特征向量，则

$$AB\boldsymbol{x}=A(\lambda_B\boldsymbol{x})=\lambda_B A\boldsymbol{x}=\lambda_B\lambda_A\boldsymbol{x}.$$

因此，$\boldsymbol{x}$也为AB的特征向量，特征值为 $\lambda_A\lambda_B$.

可同时对角化的定义:

设A, B是n阶方阵，如果存在可逆矩阵X使得

$X^{-1}AX = \Lambda(A), X^{-1}BX = \Lambda(B)$

$\Lambda(A), \Lambda(B)$都是对角矩阵，则称A, B**可同时对角化。**

将X写为列向量形式$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]$。

A, B由X同时对角化等价于$\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$为$A, B$公共的线性无关的特征向量。

π

定理：（可同时对角化的等价条件）

对可对角化的n阶方阵A, B，以下叙述等价：

(1) A, B可同时对角化

(2) 存在n个线性无关的向量，同时是A, B的特征向量

(3) A, B可交换，即$AB = BA$.

证明：

(1)(2)等价前面已经说明

(1) $\Rightarrow$ (3): 设有X使得$X^{-1}AX = \Lambda(A), X^{-1}BX = \Lambda(B)$.

于是$X^{-1}AXX^{-1}BX = \Lambda(A)\Lambda(B) = \Lambda(B)\Lambda(A) = X^{-1}BXX^{-1}AX$.

因此，$X^{-1}ABX \ = X^{-1}BAX$,得到$AB = BA$.

(3)⇒(1)较为复杂。我们看一个简单情况

假定B有n个互不相同的特征值。

根据假设$X^{-1}BX=\Lambda(B)$。令$A'=X^{-1}AX=\left(a'_{ij}\right)_{1\le i,j\le n}$.

断言A'是对角矩阵，由此得到，A可由X对角化。

首先，由于A,B可交换，A'与$\Lambda(B)$可交换：

$A'\Lambda(B)=X^{-1}AX\Lambda(B)=X^{-1}ABX=X^{-1}BAX=\left(X^{-1}BX\right)\left(X^{-1}AX\right)=\Lambda(B)A'$.

比较ij位置元素得到：$\lambda_j(B)a'_{ij}=\lambda_i(B)a'_{ij}$

当$i\neq j$时，$\lambda_j(B)\neq\lambda_i(B)$，于是 $a'_{ij}=0$。

因此，A'为对角矩阵。

一般情形的证明见教材5.4.11

例: 人口模型与马尔可夫矩阵

回到本章开始的例子：

每年有97%的人口选择留在城市，3%的人口选择由城镇迁移至乡村；

每年有95%的人口选择留在农村，5%的人口选择由乡村迁移至城镇；

$A = \begin{bmatrix} 0.97 & 0.05 \\ 0.03 & 0.95 \end{bmatrix}$

A称为马尔可夫矩阵或者转移概率矩阵。分析A的性质

求A的特征值：首先$\lambda_1 = 1$是A的一个特征值，

$I_2 - A = \begin{bmatrix} 0.03 & -0.05 \\ -0.03 & 0.05 \end{bmatrix}, (I_2 - A)\boldsymbol{x} = \mathbf{0}$有解$\begin{bmatrix} 5 \\ 3 \end{bmatrix}$。

做单位化，令$\boldsymbol{x}_1 = \begin{bmatrix} 5/8 \\ 3/8 \end{bmatrix}$.

由$trace$, A的另一个特征值$\lambda_2 = 0.97 + 0.92 - 1 = 0.92 < 1, \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$为其特征向量。

$A[\boldsymbol{x}_1, \boldsymbol{x}_2] = [\boldsymbol{x}_1, \boldsymbol{x}_2] \begin{bmatrix} 1 & \\ & \lambda_2 \end{bmatrix}$.

π

我们假设总人口数保持不变，单位化为1.

令$\boldsymbol{u}_0=\begin{bmatrix}u_1\\u_2\end{bmatrix}$为任一初始状态满足$u_1+u_2=1$.

$\boldsymbol{u}_k=A^k\boldsymbol{u}_0$ 为k年后的人口分布。

结论：(1)不论初始值$\boldsymbol{u}_0$如何选取，$\boldsymbol{u}_k\to\boldsymbol{x}_1,k\to\infty$

(2)$A^\infty=[\boldsymbol{x}_1,\boldsymbol{x}_1]$

由于$\boldsymbol{x}_1,\boldsymbol{x}_2$构成$\mathbb{R}^2$的一组基，有$c_1,c_2$满足$\boldsymbol{u}_0=c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2$.

容易得到$c_1=1$。这是因为$1=u_1+u_2=c_1\big((\boldsymbol{x}_1)_1+(\boldsymbol{x}_1)_2\big)+c_2\big(1+(-1)\big)=c_1$.

因此，$\boldsymbol{u}_0=\boldsymbol{x}_1+c_2\boldsymbol{x}_2$.

于是，$\boldsymbol{u}_k=A^k\boldsymbol{u}_0=A^k\boldsymbol{x}_1+c_2A^k\boldsymbol{x}_2=\boldsymbol{x}_1+c_2\lambda_2^k\boldsymbol{x}_2$

由于$0<\lambda_2<1$,我们得到$k\to\infty,\boldsymbol{u}_\infty=\boldsymbol{x}_1$.

计算A^∞: $A^k=A^k[\boldsymbol{e}_1,\boldsymbol{e}_2]=\left[\boldsymbol{x}_1+c\lambda_2^k\boldsymbol{x}_2,\boldsymbol{x}_1+d\lambda_2^k\boldsymbol{x}_2\right],c,d\in\mathbb{R}$;

取$k\to\infty$，得到$A^\infty=[\boldsymbol{x}_1,\boldsymbol{x}_1]$. $\boldsymbol{x}_1$也称为**稳定状态向量**(steady-state vector).

本章小结:

1. 特征对的求法: 特征多项式→特征值, 解方程→特征向量
2. 特征值的代数重数和几何重数
3. 判断方阵是否可对角化:

   足够多的线性无关的特征向量

4. 矩阵的相似, 相似等价下的不变量、若尔当(Jordan)标准形
5. Hamilton-Cayley定理、可同时对角化问题