对角化与相似 BY 范中平

I 对角化

1. 对角化的定义

设A为n阶方阵，若存在可逆阵P和对角阵D，使得$\mathrm{P}^{-1}\mathrm{AP}=\mathrm{D}=\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_{\mathrm{n}} \end{pmatrix}=\mathrm{diag}(d_1\cdots d_{\mathrm{n}})$，则称A可**对角化**，称P为A的一个**对角变换阵**.

举例: (1) 对角阵是可对角化的; (2) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$;

目前研究对角化问题的意义:

(1) 从学习的过程来说，是特征值与特征向量向矩阵方向的自然延续. 对角阵是最好研究的矩阵之一. 给出一个对角阵，可以马上读出它的特征值，特征向量，秩，逆矩阵(若存在)，行列式，迹等等重要的性质. 自然会想办法把不好研究的矩阵转化成好研究的矩阵.

(2) 从矩阵分类比较的角度来说，将对角化到相同对角阵的矩阵归为一类，会发现有非常多共有的性质，有助于一次性理解一大类矩阵. 这种分类的想法也是矩阵相似的意义所在.

(3) 从二次型的研究可知，化简一个二次型成为标准型，就是将这个二次型对应的实对称矩阵进行对角化.

以后会看到几何上确定一个非退化二次曲面的退化型，力学里确定一个惯量矩阵的主惯量等具体问题都可以转化为一个矩阵的对角化问题.

2. 对角化的条件

设A为n阶方阵，讨论A可对角化的充要条件.

- 必要条件:

若A可对角化，设P的列向量为$X_1,\cdots,X_n$，也就是

$$\mathrm{P}=[X_1,\cdots,X_n].$$

由$\mathrm{P}^{-1}\mathrm{AP}=\mathrm{D}$，有$\mathrm{AP}=\mathrm{PD}$，也就是

$$[AX_1,\cdots,AX_n]=[X_1,\cdots,X_n]\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_{\mathrm{n}} \end{pmatrix}=[d_1X_1,\cdots,d_{\mathrm{n}}X_n],$$

从而有

$$AX_i=d_iX_i,\qquad i=1,2,\cdots,n.$$

由此可得下列命题的(1)(2)，由P可逆知$X_1,\cdots,X_n$线性无关知(3),

命题 1: 若A可对角化，则

(1) $d_1,\cdots,d_n$为A的特征值;

(2) X_i为d_{i}的特征向量，$\mathrm{i}=1,2,\cdots,\mathrm{n}$.

(3) A有n个线性无关的特征向量.

- 充分条件:

若A有n个无关的特征向量$X_1,\cdots,X_n$，对应的特征值为$\lambda_1,\cdots,\lambda_{\mathrm{n}}$，令

$$\mathrm{P}=[X_1,\cdots,X_n],\qquad \Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_{\mathrm{n}}).$$

显然P可逆，且

$$AP=[AX_1,\cdots,AX_n]=[\lambda_1X_1,\cdots,\lambda_nX_n]=[X_1,\cdots,X_n]\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}=P\Lambda,$$

从而有

$$\mathrm{P}^{-1}\mathrm{AP}=\Lambda.$$

命题 2: 若A有n个无关的特征向量，则A可对角化.

推论 2.1: 若A有n个互异的特征值，则A可对角化.

由命题 1(3)与命题 2 可知A可对角化的充要条件.

- 充要条件:

定理 3: 设 A为n阶方阵，则A可对角化等价于A有n个线性无关的特征向量.

- 小结:

(1) 定理 3 给出了A可对角化的充要条件，不过通过定理 1 的判别A是否可对角化比较麻烦.

(2) 若A可对角化，也就是

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

则计算Λ就等价于求A的特征值，计算$\mathrm{P}=[x_1,\cdots,x_n]$就等价于求A的特征向量.

3.可对角化的判定:

对n阶方阵A的对角化的讨论主要有以下两方面:

(1) 如何判定方阵A是否可对角化?

(2) 若A可以对角化，如何将A对角化?

设A的全部互异特征值为$\lambda_1,\cdots,\lambda_t$ ，其代数重数分别为$n_1,\cdots,n_t$. 对$\mathrm{j}\in\{1,\cdots,t\}$，求$(\lambda_jI-\mathrm{A})\mathrm{X}=0$的基础解系得到

$$\mathrm{X}_1^{(j)},\cdots,\mathrm{X}_{m_j}^{(j)},$$

其中m_j为λ_j的几何重数. 因此关于λ_j的无关特征向量最多有m_j个，则A最多有$\sum_{j=1}^{t}m_j$个线性无关的特征向量.

由定理 3 知，A可对角化等价于

$$\sum_{j=1}^{t}m_j=n.$$

已知$\sum_{j=1}^{t}n_j=n$，且$m_j\le n_j$，若有$\sum_{j=1}^{t}m_j=n$，则对任意的$\mathrm{j}\in\{1,\cdots,t\}$，有

$$m_j=n_j.$$

因此，有

定理 4: 设A为n阶方阵. A可对角化的等价条件是A的任意特征值的几何重数都等于其代数重数.

接下来回答本节开始的两个问题，举例

(1) 判别A是否可以对角化: 若对所有特征值λ_j，都有$m_j=n_j$，则可对角化.

(2) 若A可对角化，要知道如何将A对角化，也就是要确定可逆阵P和对角阵Λ.

Step1: 解$|\lambda I-\mathrm{A}|=0$，求出A的全部不同特征值$\lambda_1,\cdots,\lambda_t$ ，与代数重数$n_1,\cdots,n_t$.

Step2: 对$\mathrm{j}\in\{1,\cdots,t\}$，求$(\lambda_jI-\mathrm{A})\mathrm{X}=0$的基础解系得到$\mathrm{X}_1^{(j)},\cdots,\mathrm{X}_{m_j}^{(j)}$.

Step3: 令

$$\Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_1,\lambda_2,\cdots,\lambda_2,\cdots,\lambda_t,\cdots,\lambda_t)=\operatorname{diag}(\cdots,\lambda_j,\cdots,\lambda_j,\cdots),$$

$$P=\left(X_1^{(1)},\cdots,X_{m_1}^{(1)},\cdots,X_1^{(j)},\cdots,X_{m_j}^{(j)},\cdots,X_1^{(t)},\cdots,X_{m_t}^{(t)}\right).$$

需要注意，并非所有方阵都可对角化.

例如，对$A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$，容易验证A的特征值$\lambda_1=\lambda_2=1$. 若假设A可以对角化，既存在 P，使得$P^{-1}AP=\Lambda$，其中$\Lambda$为对角阵. 由命题 1 知，$\Lambda=\begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}$. 则$A=P\Lambda P^{-1}=\begin{pmatrix}1&0\\0&1\end{pmatrix}$，矛盾.

4. 性质

对n阶可对角化方阵A, B，设$P^{-1}AP=\Lambda$，$Q^{-1}BQ=\Lambda'$，其中Λ,Λ'为对角阵:

- 与加法: 若$P=Q$，则$P^{-1}(A+B)P=P^{-1}AP+Q^{-1}BQ=\Lambda+\Lambda'$.

  可对角化方阵的和不一定可对角化.

  举例，$\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}1&0\\0&0\end{pmatrix}+\begin{pmatrix}0&1\\0&1\end{pmatrix}$不可对角化，其中$\begin{pmatrix}1&0\\0&0\end{pmatrix}$与$\begin{pmatrix}0&1\\0&1\end{pmatrix}$可对角化，后者有

$$\begin{pmatrix}1&-1\\0&1\end{pmatrix}\begin{pmatrix}0&1\\0&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}1&-1\\0&0\end{pmatrix}.$$

- 与数乘: $P^{-1}kAP=k\Lambda$.
- 与乘法: 若$P=Q$，则$P^{-1}(AB)P=P^{-1}APP^{-1}BP=\Lambda\Lambda'$.

  可对角化方阵的积不一定可对角化.

  举例，

$$\begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}-2&-1\\\frac{3}{2}&1\end{pmatrix}=\begin{pmatrix}1&1\\0&1\end{pmatrix},$$

  但左边两个矩阵可对角化.

- 与幂:

$$P^{-1}A^mP=PA(P^{-1}P)A(P^{-1}P)\cdots(P^{-1}P)AP^{-1}=\Lambda^m,$$

  由此可计算A^m，有

$$A^m=\left(P\Lambda P^{-1}\right)^m=P\Lambda\left(P^{-1}P\right)\Lambda\left(P^{-1}P\right)\cdots\left(P^{-1}P\right)\Lambda P^{-1}=P\Lambda^mP^{-1}.$$

5. 实对称矩阵的对角化

称A是**实对称矩阵**，若A为实方阵且

$$A^T=A.$$

定理 5: 对实对称矩阵A，有以下结论成立:

(1) 特征值都是实数;

(2) 不同特征值对应的特征向量彼此正交;

(3) 总可以对角化，进一步，总存在n阶正交矩阵Q，使得$Q^{-1}AQ=\Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$.

证: (1) 要证$\lambda=\bar{\lambda}$. 由

$$AX=\lambda X,$$

有

$$(\overline{AX})^T=(\overline{\lambda X})^T,$$

由$\overline{A}^T = A$有,

$$(\overline{AX})^T X = \overline{X^T}\,\overline{A^T} X = \overline{X^T} A X = \lambda \overline{X^T} X,$$

同时有

$$(\overline{AX})^T X = (\overline{\lambda X})^T X = \overline{\lambda}\,\overline{X^T} X.$$

因此有

$$\overline{\lambda}\,\overline{X^T} X = \lambda \overline{X^T} X.$$

(2) 设$X_1 \sim \lambda_1, X_2 \sim \lambda_2, \lambda_1 \neq \lambda_2$. 有

$$\lambda_1 X_2^T X_1 = X_2^T A X_1 = X_2^T A^T X_1 = (AX_2)^T X_1 = \lambda_2 X_2^T X_1$$

由$\lambda_1 \neq \lambda_2$知,

$$X_2^T X_1 = 0.$$

(3) P243 定理 5.12

实对称矩阵的对角化方法: P245 例 1

设n阶实对称阵A的全部互异特征值为$\lambda_1, \cdots, \lambda_t$ ，已知A可以对角化，且不同特征值下的特征向量正交.

Step1 对$j = 1 \cdots t$，由$(\lambda_j \mathrm{I} - A)X = 0$，求基础解系 $\lambda_j \sim X_1^{(j)}, \cdots, X_{m_j}^{(j)}$;

Step2 对$j = 1 \cdots t$，将同一特征值下的特征向量进行 Schmidt 正交化，把$X_1^{(j)}, \cdots, X_{m_j}^{(j)}$转化为$\mathrm{Y}_1^{(j)}, \cdots, \mathrm{Y}_{m_j}^{(j)}$;

Step3 令$\mathrm{Q} = \left(\cdots, \mathrm{Y}_1^{(j)}, \cdots, \mathrm{Y}_{m_j}^{(j)}, \cdots\right)$ ，有$\mathrm{Q}^{\mathrm{T}}\mathrm{Q} = \mathrm{I}$且$\mathrm{Q}^{-1}\mathrm{AQ} = \Lambda = \mathrm{diag}(\cdots, \lambda_j, \cdots, \lambda_j, \cdots)$.

III 相似

1.定义:

设A, B为n阶方阵，若存在可逆阵P，使得$\mathrm{P}^{-1}\mathrm{AP} = \mathrm{B}$，则称A**相似于**B，记为A~ B；称P为A到B的**相似变换阵**.

显然，A可对角化等价于A相似于一个对角阵.

2.性质:

- 等价关系：自反性 A~ A；对称性 若A~B，则B~ A；传递性：若A~ B，B~C，则A~C.
- 与数乘：若A~ B，则

$$\mathrm{kA} \sim \mathrm{kB}.$$

- 与加法：若$\mathrm{P}^{-1}\mathrm{AP} = \mathrm{B}$，则$(\mathrm{A} + \mathrm{C}) \sim (\mathrm{B} + \mathrm{P}^{-1}\mathrm{CP})$;
- 与乘法：若$\mathrm{P}^{-1}\mathrm{AP} = \mathrm{B}$，则$\mathrm{AC} \sim \mathrm{B}(\mathrm{P}^{-1}CP)$;
- 与幂：若A~B，则

$$\mathrm{A}^m \sim \mathbf{B}^m.$$

- 与特征值:

定理 1: 相似矩阵的特征值相同.

证：设A~ B，按定义存在可逆阵P，使得

$$\mathrm{P}^{-1}\mathrm{AP} = \mathrm{B}.$$

由

$$|\lambda I-\mathrm{B}|=|\mathrm{P}^{-1}(\lambda I-\mathrm{A})P|=\left|\mathrm{P}^{-1}\right||\lambda I-\mathrm{A}|\left|P\right|=|\lambda I-\mathrm{A}|,$$

A与B特征值相同.

定理 1 逆命题不成立，即有相同特征值的矩阵不一定相似. 例如$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$与$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$的特征值都为 1，但不相似.

- 与多项式：若A~B，则

$$\mathrm{f(A)}\sim\mathrm{f(B)}.$$

证：设$\mathrm{P}^{-1}\mathrm{AP}=\mathrm{B}$. 则$\mathrm{P}^{-1}\mathrm{f(A)P}=\mathrm{f(B)}$.

定理 2: 若$\lambda_1,\cdots,\lambda_n$为A的所有特征值，则$\mathrm{f}(\lambda_1),\cdots,f(\lambda_n)$为f(A)的所有特征值.

证：一般情形都对，这里只对A为可对角化的矩阵的情形给出证明. $\mathrm{A}\sim\Lambda$，$\mathrm{f(A)}\sim\mathrm{f}(\Lambda)$，由定理 1 知，f(A)与$\mathrm{f}(\Lambda)$特征值相同.