第二讲 特殊矩阵

李宇峰
liyf@nju.edu.cn
人工智能学院

特殊矩阵

在工程应用或者工程问题的求解中，经常会遇到某些具有特殊结构的矩阵。而特殊矩阵是矩阵论中内容非常丰富的一个题目，在这里仅是讲一些我们常用的特殊矩阵。特殊矩阵包括下列三种情形：

（1）元素的分布结构特殊：对角矩阵、上（下）三角矩阵、对称矩阵等；

（2）矩阵的性质特殊：正定矩阵、不定矩阵等；

（3）矩阵的作用特殊：置换矩阵，排列矩阵等。

1. Hermite 矩阵

若$A \in R^{n\times n}$矩阵且满足$A = A^T$, 即$a_{ij} = a_{ji}$, 则成为对称矩阵。

若$A \in C^{n\times n}$, 且$A = A^H$, 即A为复共轭对称矩阵，就是$a_{ij} = a_{ji}^*$,

则称A为$Hermite$矩阵。

譬如观测信号向量$x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$的自协方差矩

阵$C_{xx} = E\{x(t)x(t)^H\}$就是一个典型的$Hermite$矩阵。

$Hermite$矩阵有以下重要性质：

（1）对所有$n\times n$复矩阵A，矩阵$A + A^H$，AA^H和A^HA均是

$Hermite$矩阵。

1. Hermite 矩阵

(2) 若A是$Hermite$矩阵，则矩阵幂A^k对所有$k = 1,2,\ldots$都是$Hermite$矩阵。

(3) 若$Hermite$矩阵A非奇异，则其逆矩阵A^{-1}也是$Hermite$矩阵。

(4) 若A和B都是$Hermite$矩阵，则$\alpha A + \beta B$ (α和β也是任意非零实数)是$Hermite$矩阵。

(5) 若A和B为$Hermite$矩阵，则$AB + BA$和$j(AB - BA)$也是$Hermite$矩阵。

1. Hermite 矩阵

*Hermite*矩阵的正定性判据：一个$n \times n$ *Hermite*矩阵A是正定的，当且仅当它满足以下任何一个条件：

(1) 二次型函数$x^H Ax > 0, \forall x \neq 0$；

(2) 矩阵A的所有特征值都大于零；

(3) 存在一个非奇异的$n \times n$矩阵R，使得$A = R^H R$；

(4) 对任意非奇异的$n \times n$矩阵P，使得$P^H AP$是正定阵。

最简单的*Hermite*矩阵莫过于单位矩阵$I = [e_1, e_2, \dots, e_n]$, 其中 e_i为n维基本向量，其第i个元素为1，其余元素皆等于0。

1. Hermite 矩阵

与单位矩阵紧密相关的置换矩阵、互换矩阵、移位矩阵和选择矩阵。这四种矩阵的共同之点是其元素只由0和1组成，并且每行和每列都只有一个非零元素1，只是非零元素1所处的位置不一样而已。

2. 置换矩阵与互换矩阵

一个正方矩阵称为置换矩阵(*permutation matrix*), 若它的每一行和每一列有一个且仅有一个非零元素1.

例2. 1 给定一个5×4矩阵

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{bmatrix}_{5\times 4}$$

2. 置换矩阵与互换矩阵

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{bmatrix}_{5\times 4}$$

并令置换矩阵

$$P_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{4\times 4}, \qquad P_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}_{5\times 5}$$

则有

$AP_4 = ?\quad P_5A = ?$

2. 置换矩阵与互换矩阵

并令置换矩阵

$$P_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{4\times 4}, \qquad P_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}_{5\times 5}$$

则有

$$AP_4 = \begin{bmatrix} a_{12} & a_{13} & a_{14} & a_{11} \\ a_{22} & a_{23} & a_{24} & a_{21} \\ a_{32} & a_{33} & a_{34} & a_{31} \\ a_{42} & a_{43} & a_{44} & a_{41} \\ a_{52} & a_{53} & a_{54} & a_{51} \end{bmatrix} \qquad P_5 A = \begin{bmatrix} a_{51} & a_{52} & a_{53} & a_{54} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix}$$

2. 置换矩阵与互换矩阵

注意：

（1）用置换矩阵P_4右乘矩阵A，相当于对A的列重新排列；

（2）用置换矩阵P_5左乘矩阵A，相当于将A的行重新排列；

（3）置换矩阵是正交矩阵：$P^TP = PP^T = I$, 也就是$P^{-1} = P^T$；

下列形式的矩阵是一种特殊的置换矩阵，又称反射矩阵

（*reflection matrix*）

$$J = \begin{bmatrix} 0 & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & 0 \end{bmatrix},$$

2. 置换矩阵与互换矩阵

通过左乘或者右乘反射矩阵，可以将矩阵A的行或列的顺序反转，设$A=[a_{ij}]_{m\times n}$为一$m\times n$矩阵矩阵, 则：

$$J_mA=\begin{bmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}_{m\times n}$$

$$AJ_n=\begin{bmatrix} a_{1n} & \cdots & a_{12} & a_{mn} \\ a_{2n} & \cdots & a_{22} & a_{21} \\ \vdots & \vdots & \vdots & \vdots \\ a_{mn} & \cdots & a_{m2} & a_{m1} \end{bmatrix}_{m\times n}$$

J_mA将A的行的次序反转；AJ_n将A的列的次序反转。

3. 广义置换矩阵与选择矩阵

一个正方矩阵称为广义置换矩阵(*generalized permutation matrix*)简称g–矩阵，若每行每列有且仅有一个非零元素。

一个广义置换矩阵可以分解为一个置换矩阵和一个非奇异的对角矩阵的乘积，即有$G = PD$，其中D为非奇异对角矩阵，例如

$$G = \begin{bmatrix} 0 & 0 & 0 & 0 & \alpha \\ 0 & 0 & \beta & 0 & 0 \\ 0 & \gamma & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ \rho & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \rho & & & & \\ & \gamma & & & \\ & & \beta & & \\ & & & \lambda & \\ & & & & \alpha \end{bmatrix}$$

3. 广义置换矩阵与选择矩阵

顾名思义，选择矩阵（*selective matrix*）是一种可以对某个给定矩阵的某些行或者某些列进行选择的矩阵。以$m\times N$矩阵

$$X=\begin{bmatrix} x_1(1) & x_1(2) & \cdots & x_1(N) \\ x_2(1) & x_2(2) & \cdots & x_2(N) \\ \vdots & \vdots & & \vdots \\ x_m(1) & x_m(2) & \cdots & x_m(N) \end{bmatrix}$$

为例。令

$$J_1=[I_{m-1},0_{m-1}],\quad J_2=[0_{m-1},I_{m-1}]$$

是两个$(m-1)\times m$矩阵，式中$I_{m-1}, 0_{m-1}$分别是$(m-1)\times(m-1)$单位矩阵和$(m-1)\times 1$零向量。直接计算得

3. 广义置换矩阵与选择矩阵

$$J_1X=\begin{bmatrix} x_1(1) & x_1(2) & \cdots & x_1(N) \\ x_2(1) & x_2(2) & \cdots & x_2(N) \\ \vdots & \vdots & & \vdots \\ x_{m-1}(1) & x_{m-1}(2) & \cdots & x_{m-1}(N) \end{bmatrix}$$

就是说矩阵J_1X选择的是原矩阵X的前$m-1$行；

$$J_2X=\begin{bmatrix} x_2(1) & x_2(2) & \cdots & x_2(N) \\ x_3(1) & x_3(2) & \cdots & x_3(N) \\ \vdots & \vdots & & \vdots \\ x_m(1) & x_m(2) & \cdots & x_m(N) \end{bmatrix}$$

矩阵J_2X选择的是原矩阵X的后$m-1$行。

3. 广义置换矩阵与选择矩阵

类似的若令

$$J_1 = \begin{bmatrix} I_{N-1} \\ 0_{N-1} \end{bmatrix}, \qquad J_2 = \begin{bmatrix} 0_{N-1} \\ I_{N-1} \end{bmatrix}$$

是两个$N\times(N-1)$矩阵，则

$$XJ_1 = \begin{bmatrix} x_1(1) & x_1(2) & \cdots & x_1(N-1) \\ x_2(1) & x_2(2) & \cdots & x_2(N-1) \\ \vdots & \vdots & & \vdots \\ x_m(1) & x_m(2) & \cdots & x_m(N-1) \end{bmatrix},$$

$$XJ_2 = \begin{bmatrix} x_1(2) & x_1(3) & \cdots & x_1(N) \\ x_2(2) & x_2(3) & \cdots & x_2(N) \\ \vdots & \vdots & & \vdots \\ x_m(2) & x_m(3) & \cdots & x_m(N) \end{bmatrix}$$

就是分别选择原矩阵X的前$N-1$列和后$N-1$列；

4. 广义置换矩阵的应用案例

（1）工程问题的提法

鸡尾酒会问题：n个人在酒会上交谈，由m个传感器观测得到m维的数据向量$x(t)=(x_1(t),x_2(t),\dots,x_m(t))^T$，其中$x_i(t)$表示第$i$个传感器在$t$时刻的观测数据。能否通过观测数据向量$x(t)$将$n$个人的语音信号分离出来？

（2）数学建模

令$x(t)=(x_1(t),x_2(t),\dots,x_m(t))^T$表示$n$维的源信号向量，其中$s_j(t)$表示第$j$个人的语音信号。若$n$维源信号向量$s(t)$经过无线信道的线性混合，被传感器观测，则$m$维观测数据向量

4. 广义置换矩阵的应用案例

$x(t)$可以用线性方程组

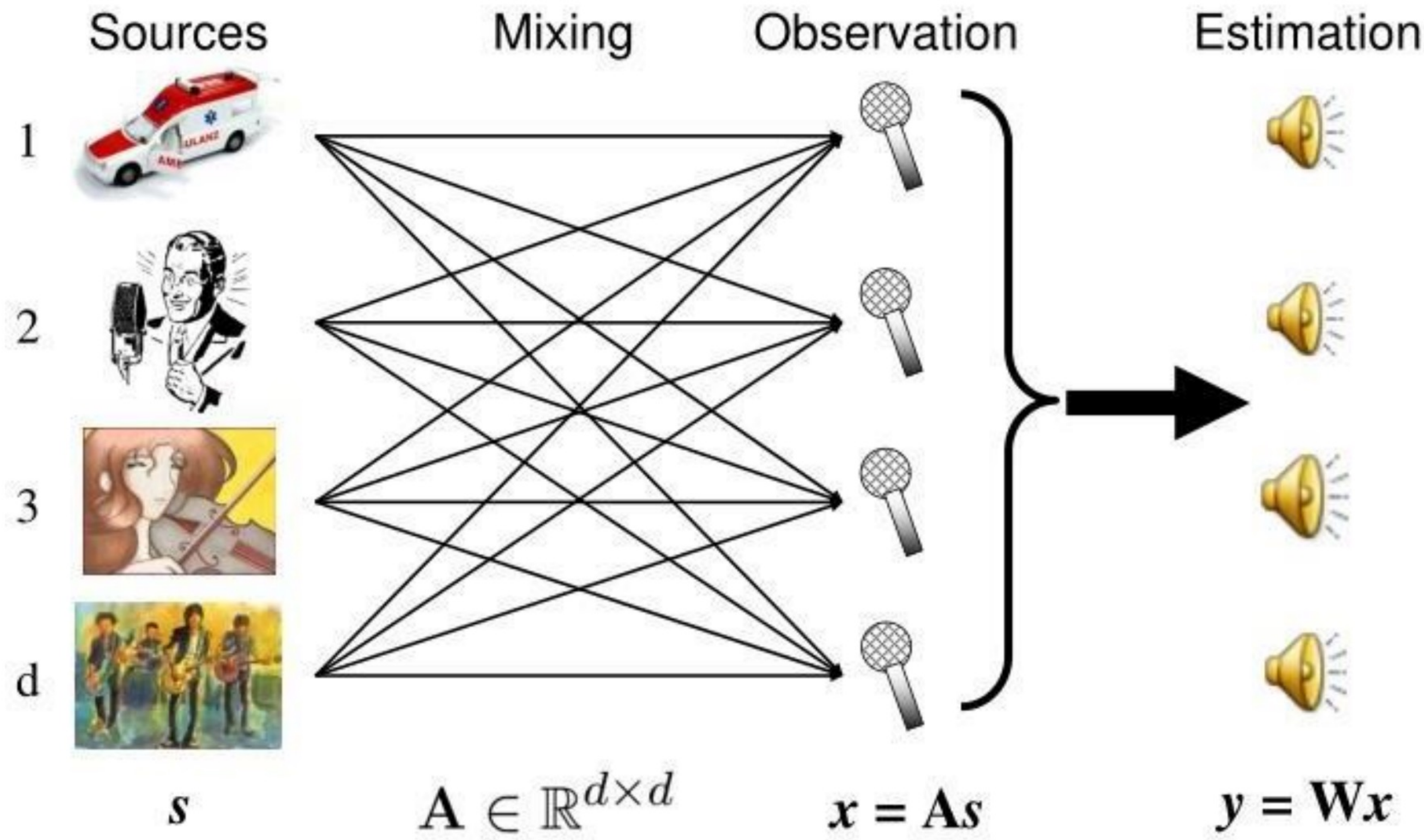
$$x(t) = As(t) \quad 或者 \quad x_i(t) = \sum_{j=1}^{n} a_{ij} s_j(t),\ i = 1,2,\ldots,m$$

来进行数学建模。其中混合矩阵A的元素a_{ij}表示第j个人的语音信号到达第i个传感器所经历信道的混合系数。

（3） 数学问题

现在只有观测数据向量$x(t)$已知，而混合矩阵A未知，能否通过求解盲线性方程组$x(t) = As(t)$得到混合矩阵A与/或源信号向量$s(t)$？这类问题在工程中称为盲信号分离(*blind signal separation*)问题。

4. 广义置换矩阵的应用案例



定理（Pierre Comon, 1994）假设随机信号$\mathbf{z}$服从模型$\mathbf{z} = \mathbf{Bs}$，其中$\mathbf{s}$的分量**相互独立**，且其中至多可以有一个为高斯；$\mathbf{B}$为满秩方阵。那么若$\mathbf{z}$的分量相互独立当且仅当$\mathbf{B} = \mathbf{PD}$，其中$\mathbf{P}$为排列矩阵(permutation matrix)，$\mathbf{D}$为对角矩阵。

盲信号分离的内在不确定性（inherent indeterminacy）

4. 广义置换矩阵的应用案例

（4） 问题求解思路

令混合矩阵$A=[a_1,a_2,\dots,a_n]$, 则盲信号混合的线性方程组可以等价写成

$$x(t)=As(t)=\sum_{j=1}^{n} a_j s_j(t)=\sum_{j=1}^{n} \frac{a_j}{\beta_j}\cdot \beta_j s_j(t) \quad i=1,2,\dots,m$$

由此可以看出盲信号混合方程存在两种不确定性或模糊性：

1. 混合矩阵A的列向量和源信号的排序的不确定性，因为同时对换列向量a_j ↔️ a_k和源信号$s_j(t)$ ↔️ $s_k(t)$, 观测数据向量保存不变

2. 列向量和源信号的测度的不确定性，因为列向量a_j除以一个

4. 广义置换矩阵的应用案例

非零的因子β_j，同时源信号$s_j(t)$乘以一个相同的因子，观测数据向量也保持不变。

这说明盲信号混合方程可以使用广义置换矩阵G进一步等价的写成

$$x(t) = (AG)G^{-1}S(t) = (AG)\bar{s}(t)$$

其中AG和$\bar{s}(t) = G^{-1}S(t)$分别表示混合矩阵和源信号向量的排序不确定性以及尺度不确定性。

4. 广义置换矩阵的应用案例

（5）可行性分析

前面讲到，盲线性方程组$x(t)=(AG)\bar{s}(t)$的解$\bar{s}(t)=(AG)^{+}x(t)$存在两种不确定性：(1)分离信号的排列次序与源信号的排列次序可能不一致；（2)分离的某个信号与对应的源信号相差一个固定的复值因子。然而这两种不确定性在工程应用中是允许的，重要的是已经达到了信号分离的目的，并不影响分离信号与源信号之间的“高保真度”。

5. 正交矩阵与酉矩阵

在n维实线性空间R^n中，我们考虑由n个线性无关的向量组成的向量系$\{x_1, x_2, \dots, x_n\}$, 如果这一组向量系满足

$$x_i^T x_j = \begin{cases} \neq 0 & i = j \\ 0 & i \neq j \end{cases} \qquad i, j = 1,2, \dots, n$$

则称向量系$\{x_1, x_2, \dots, x_n\}$是一组正交向量系，若对$i = 1,2, \dots, n$又有$\|x_i\|_2 = 1$，就称向量系$\{x_1, x_2, \dots, x_n\}$是一组标准正交向量系。正交向量系一定是线性无关的，同样线性无关向量系可以变换为正交向量系。

5. 正交矩阵与酉矩阵

若正方矩阵$Q \in R^{n\times n}$, 且满足

$$QQ^T = Q^TQ = I_n$$

则称矩阵Q是正交矩阵。若Q是一个正交矩阵, 则它的n个列向量或者n个行向量都是标准正交向量系。

正交矩阵有许多重要性质，譬如对任何向量$x \in R^n$, 则

$$\|Qx\|_2 = \|x\|_2,$$

也就是说正交矩阵保持向量的欧氏长度不变。因此正交矩阵是矩阵论和矩阵计算中具有重要应用意义的一类特殊矩阵。

5. 正交矩阵与酉矩阵

若$Q \in R^{m\times n}\,(m \geq n)$不是正方矩阵而满足$Q^TQ = I_n$, 就称矩阵$Q$是列正交矩阵；

若$Q \in R^{m\times n}\,(m \leq n)$不是正方矩阵而满足$QQ^T = I_m$, 则称$Q$是一个行正交矩阵；

若正方矩阵$Q \in C^{n\times n}$, 且满足

$$QQ^H = Q^HQ = I_n$$

则称Q是n阶酉矩阵(复正交矩阵）。

5. 正交矩阵与酉矩阵

设U为$n\times n$酉矩阵，若矩阵$B = U^H AU$，则称矩阵$B \in C^{n\times n}$与$A \in C^{n\times n}$酉等价。如果在实数域讨论，矩阵B与A就是正交等价。

若$A \in R^{n\times n}$是正交矩阵，则有下列重要性质：

（1）A的行向量和列向量都是标准正交向量系；

（2）A的逆矩阵是A^T, 即$A^{-1} = A^T$；

（3）A的行列式的绝对值$|\det(A)| = 1$；

（4）A的特征值的绝对值$|\lambda| = 1$.

复数域上的酉矩阵有类似的性质，正交（酉）矩阵是一类有重要的理论和应用价值的特殊矩阵。

6. 三角矩阵

矩阵计算中一个基本策略就是将一个复杂的问题化解为若干个形式简单的问题来进行求解。其中上（下）三角矩阵就是常用的矩阵。

上三角矩阵(*upper triangular matrix*) $U = [u_{ij}]_{n\times n}$就是主对角线以下元素都为零的矩阵，即$u_{ij} = 0, i > j$：

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{mn} \end{bmatrix}$$

6. 三角矩阵

下三角矩阵（*lower triangular matrix*）$L = \left[l_{ij}\right]_{n\times n}$就是主对角线以上元素都为零的矩阵，即$l_{ij} = 0, i < j$：

$$L = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}$$

上三角矩阵有以下性质：

（1）上三角矩阵的乘积为上三角矩阵，即若$U_1, U_2, \dots, U_k$为上三角矩阵，则$U_1 U_2 \dots U_k$为上三角矩阵；

6. 三角矩阵

(2) 上三角矩阵$U=\left[u_{ij}\right]$的行列式是对角线元素的乘积，即

$$\det(\mathrm{U})=u_{11}u_{12}\ldots u_{nn}=\prod_{i=1}^{n}u_{ii}$$

(3) 非奇异上三角矩阵的逆矩阵为上三角矩阵；

(4) 正定$Hermite$矩阵A可以分解为$A=T^{H}DT$,其中T为上三角矩阵，D为实对角矩阵；

下三角矩阵有以下性质：

(1) 下三角矩阵的乘积为下三角矩阵，即若$L_1,L_2,\ldots,L_k$为下三角矩阵，则$L_1L_2\ldots L_k$为下三角矩阵；

6. 三角矩阵

(2) 下三角矩阵$L=[l_{ij}]$的行列式是对角线元素的乘积，即

$$\det(L)=l_{11}l_{12}\ldots l_{nn}=\prod_{i=1}^{n}l_{ii}$$

(3) 非奇异下三角矩阵的逆矩阵为下三角矩阵；

(4) 实对称正定矩阵$A_{n\times n}$可以分解为$A=LL^T$,其中L为下三角矩阵，这通常称为实对称矩阵A的*Cholesky*分解。

特殊矩阵

- Hermitian matrix
- 置换与互换矩阵
- 广义置换与选择矩阵
- 正交矩阵与酉矩阵
- 三角矩阵