



第二讲 特殊矩阵 (1)

李宇峰

liyf@nju.edu.cn

人工智能学院



特殊矩阵

在矩阵计算中，特殊矩阵具有**简化计算、揭示内在结构、提供变换工具**等作用，在各个领域都有广泛而重要的应用：

简化计算

- **对角矩阵**：其非对角线元素全为零，在矩阵乘法、求逆等运算中具有极大的便利性。例如，两个同阶对角矩阵相乘，只需将对应角线上的元素相乘即可得到结果。计算对角矩阵的幂次时，只需将对角线上的元素进行相应的幂运算，大大简化了计算过程。
- **三角矩阵**：无论是上三角矩阵还是下三角矩阵，在求解线性方程组时都具有优势。对于上三角线性方程组（为上三角矩阵），可以采用回代法求解，从最后一个方程开始，逐步向前求解未知数，计算量相对较小。

揭示结构与性质

- **对称矩阵**：其关于主对角线对称的元素相等，这种结构特点反映了所描述问题中的某种对称性。在二次型（为对称矩阵）中，对称矩阵的特征值和特征向量可以揭示二次型的几何性质，如椭球的形状、方向等。
- **正定矩阵**：作为一种特殊的对称矩阵，正定矩阵具有所有特征值都大于零的性质。这一性质在优化问题中非常重要，例如在最小二乘法问题中，如果目标函数的海森矩阵是正定的，那么该函数存在唯一的全局最小值，为优化算法的收敛性和稳定性提供了理论依据。

特殊矩阵

作为变换工具

- **正交矩阵**：它的列向量（或行向量）是两两正交的单位向量，其转置等于逆。在几何变换中，正交矩阵常用于表示旋转、反射等变换，因为它能够保持向量的长度和夹角不变，保证了变换后的图形与原图形在几何性质上的一致性。
- **置换矩阵**：可以实现对矩阵的行或列进行置换操作，在矩阵的重排、分块等操作中具有重要作用。例如在矩阵的 LU 分解中，有时需要使用置换矩阵来确保分解的稳定性和可行性，通过对矩阵进行适当的行置换，使得分解过程能够顺利进行。

用于矩阵分解

- **奇异值分解 (SVD)**：对于任意矩阵，都可以分解为 $A = U \Sigma V^T$ 的形式，其中 U 和 V 是正交矩阵， Σ 是对角矩阵。SVD 在数据压缩、图像识别、信号处理等领域有广泛应用。例如在图像压缩中，通过对图像的像素矩阵进行 SVD 分解，可以根据奇异值的大小保留主要信息，去除次要信息，从而实现图像的压缩。
- **QR分解**：将矩阵分解为一个正交矩阵和一个上三角矩阵的乘积，即 $A=QR$ 。QR分解在求解线性方程组、特征值问题等方面有重要应用，例如在求解最小二乘问题时，通过QR分解可以将问题转化为更易于求解的形式，提高计算效率和精度。

构建数学模型

- **图论**：邻接矩阵用于表示图中节点之间的连接关系，是一种特殊矩阵。通过对邻接矩阵的运算和分析，可以研究图的各种性质，如连通性、最短路径等。拉普拉斯矩阵也是图论中的重要特殊矩阵，它与图的谱性质密切相关，可用于图的聚类、社区发现等任务。

特殊矩阵

特殊矩阵是矩阵论中内容非常丰富的题目，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。

特殊矩阵

- Hermitian 矩阵
- 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

1. Hermite 矩阵

- 简化特征值问题

- **对称矩阵具有实特征值**，这是其非常重要的特性。根据线性代数理论，对于实对称矩阵 A ，其所有特征值都是实数。这一性质使得在求解特征值问题时，避免了处理复数的复杂性，简化了计算过程。
- **对称矩阵不同特征值对应的特征向量相互正交**。设 a 和 b 是对称矩阵 A 的两个不同特征值， u 和 v 分别是对应的特征向量， u 和 v 正交。这一正交性为特征向量的求解和矩阵的对角化提供了便利。

- 计算机应用

- 机器学习：在机器学习算法中，许多模型的目标函数涉及到对称矩阵的运算。例如，**在支持向量机（SVM）中，核矩阵是对称矩阵**，它在计算样本之间的相似度时起着关键作用。通过对核矩阵的分析和处理，可以优化 SVM 模型的性能。
- 计算机图形学：在图形变换中，对称矩阵可以用于表示一些具有对称性的变换，如反射变换。**对称矩阵的性质可以保证变换的稳定性和准确性，使得图形在变换过程中保持良好的几何性质。**

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 $A^H A$ 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 $k = 1, 2, \dots$ 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

例如, 设 $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, A 和 B 都是对称矩阵。

计算 $AB = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 7 & 11 \end{bmatrix}$, 结果是对称矩阵。

再设 $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, 同样 A 和 B 是对称矩阵。

计算 $AB = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 6 & 10 \end{bmatrix}$, 而 $(AB)^T = \begin{bmatrix} 3 & 6 \\ 5 & 10 \end{bmatrix}$, 此时 $AB \neq (AB)^T$, 即 AB 不是对称矩阵。

所以只有当 A 和 B 可交换, 即 $AB = BA$ 时, AB 才是对称矩阵。

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 $x^H A x > 0, \forall x \neq 0$;
- (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 $P^H A P$ 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

*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 $x^H A x > 0, \forall x \neq 0$;
- (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 $P^H A P$ 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。

2. 置换矩阵与互换矩阵

实现矩阵的行或列变换

- **行或列的置换**：置换矩阵可以对其他矩阵的行或列进行重新排列。比如，用一个置换矩阵左乘矩阵，即对行进行置换；用右乘，即对列进行置换。在求解线性方程组时，如果系数矩阵的某些行或列需要交换顺序以方便计算，如在高斯消元法中为了避免出现除数为零的情况，就可以使用置换矩阵来实现。
- **调整矩阵结构**：在矩阵的分块运算中，通过置换矩阵可以将矩阵的行或列进行重新分组，使分块后的矩阵结构更便于分析和计算。例如，将一个大型矩阵按照某种规则进行行或列的置换后，可能会得到一个块对角矩阵或块三角矩阵，从而简化矩阵的运算。

数据处理和分析中的应用

- **数据重排**：在数据分析和处理中，经常需要对数据矩阵进行重新排列。例如，在数据挖掘中，可能需要根据某些特征对数据样本进行重新排序，以便更好地进行特征提取、分类等操作。置换矩阵可以方便地实现数据矩阵的行或列重排，为后续的数据处理提供便利。
- **加密与解密**：在密码学中，置换矩阵可以用于构建一些简单的加密算法。通过对明文数据矩阵进行特定的行或列置换操作，实现数据的加密。在解密过程中，使用相应的逆置换矩阵将密文恢复为明文。这种基于置换矩阵的加密方法虽然相对简单，但在一些特定的场景下具有一定的应用价值。

2. 置换矩阵与互换矩阵

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

一个正方形矩阵称为置换矩阵 (*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

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

$$P_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}_{5 \times 5}$$

则有

$$AP_4 = ? \quad P_5A = ?$$

2. 置换矩阵与互换矩阵

并令置换矩阵

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

则有

$$AP_4 = \begin{bmatrix} a_{12} & a_{13} & a_{14} & a_{11} \\ a_{22} & a_{23} & a_{24} & a_{21} \\ a_{32} & a_{33} & a_{34} & a_{31} \\ a_{42} & a_{43} & a_{44} & a_{41} \\ a_{52} & a_{53} & a_{54} & a_{51} \end{bmatrix} \quad P_5A = \begin{bmatrix} a_{51} & a_{52} & a_{53} & a_{54} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix}$$

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： $P^T P = P P^T = I$ ，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*)

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

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，则：

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

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

$J_m A$ 将 A 的行的次序反转； $A J_n$ 将 A 的列的次序反转。

3. 广义置换矩阵与选择矩阵

一个方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矩阵。以 $m \times N$ 矩阵

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

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

$$J_1 X = \begin{bmatrix} x_1(1) & x_1(2) & \dots & x_1(N) \\ x_2(1) & x_2(2) & \dots & x_2(N) \\ \vdots & \vdots & & \vdots \\ x_{m-1}(1) & x_{m-1}(2) & \dots & x_{m-1}(N) \end{bmatrix}$$

就是说矩阵 $J_1 X$ 选择的是原矩阵 X 的前 $m-1$ 行；

$$J_2 X = \begin{bmatrix} x_2(1) & x_2(2) & \dots & x_2(N) \\ x_3(1) & x_3(2) & \dots & x_3(N) \\ \vdots & \vdots & & \vdots \\ x_m(1) & x_m(2) & \dots & x_m(N) \end{bmatrix}$$

矩阵 $J_2 X$ 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

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

是两个 $N \times (N-1)$ 矩阵, 则

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

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

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

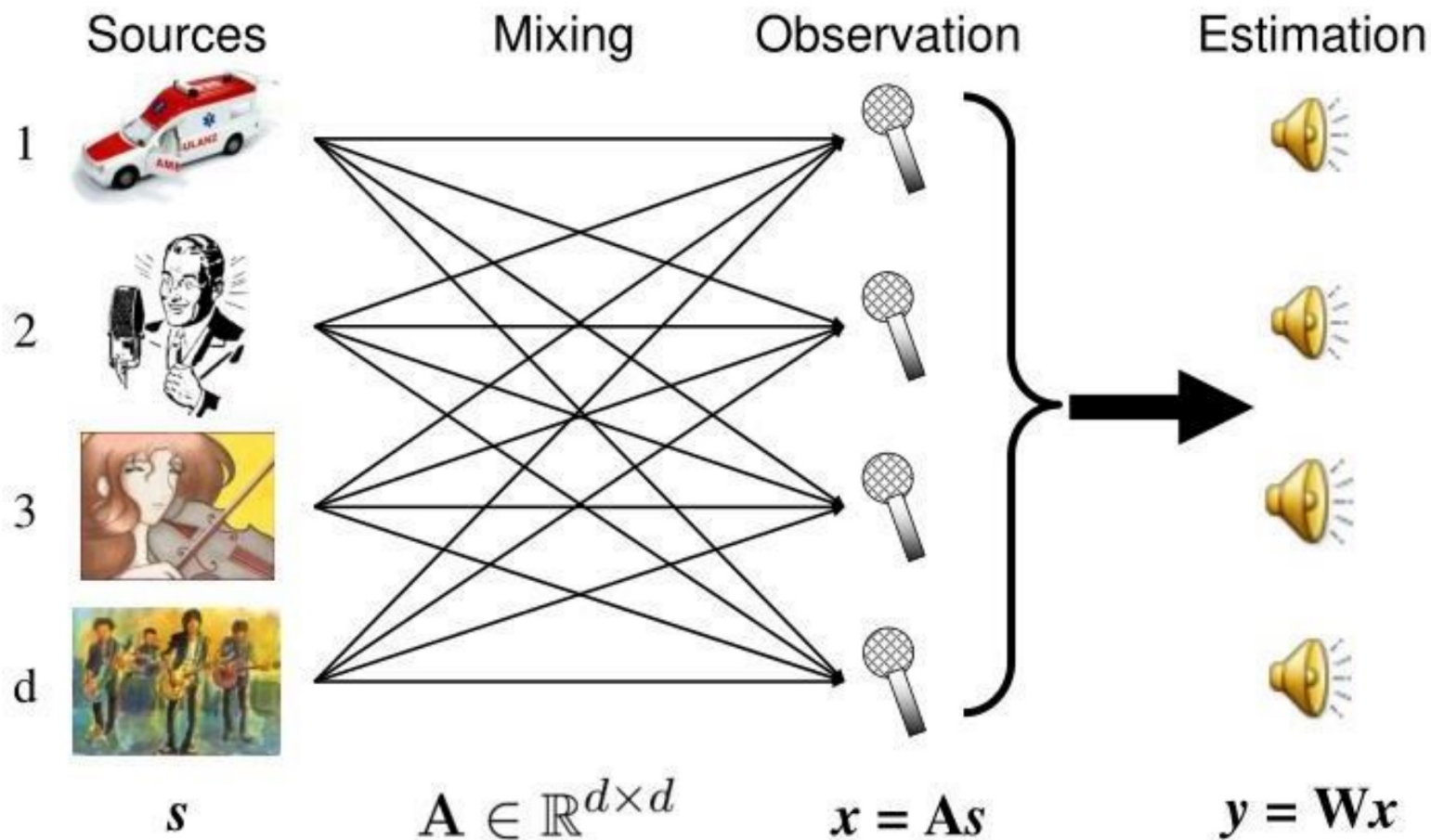
$$x(t) = As(t) \quad \text{或者} \quad x_i(t) = \sum_{j=1}^n a_{ij}s_j(t), \quad i = 1, 2, \dots, m$$

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



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 $\mathbf{z} = \mathbf{B}\mathbf{s}$ ，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 $\mathbf{B} = \mathbf{P}\mathbf{D}$ ，其中 \mathbf{P} 为排列矩阵 (permutation matrix)， \mathbf{D} 为对角矩阵。

盲信号分离问题经典算法之一

- **算法示例**：经典的 PCA 算法通过对混合信号矩阵 X 计算协方差矩阵 $C = XX^T$ ，然后对 C 进行特征分解，得到特征值和特征向量，选择主要的特征向量构成投影矩阵 P ，分离后的信号 $Y = P^T X$ 。

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 $a_j \leftrightarrow a_k$ 和源信号 $s_j(t) \leftrightarrow s_k(t)$, 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

正交矩阵在简化计算、提高计算精度和稳定性等方面发挥着重要作用：

简化矩阵运算

- **乘法计算简化**：若 Q 为正交矩阵，对于矩阵乘法 QA 或 AQ ，由于正交矩阵的特殊性质 $Q^T Q = Q Q^T = I$ ，在后续计算中可能会出现 $Q^T Q A = A$ 或 $A Q Q^T = A$ 等简化形式，从而减少计算量。
- **求逆计算简化**：一般矩阵求逆运算较为复杂，但正交矩阵的逆矩阵等于其转置矩阵，即 $Q^{-1} = Q^T$ ，大大简化了求逆的计算过程。例如在求解线性方程组 $Qx = b$ 时，其解为 $x = Q^{-1}b = Q^T b$ ，无需进行复杂的求逆运算。

提高计算精度与稳定性

- **减少误差积累**：在数值计算中，误差的积累可能会导致结果的严重偏差。正交矩阵具有保范性和保角性，在矩阵变换过程中能保持向量的长度和夹角不变，使得计算过程中的误差传播相对稳定，减少了误差积累的可能性，提高了计算精度。
- **改善条件数**：矩阵的条件数反映了矩阵计算对误差的敏感度。对于正交矩阵 Q ，其条件数 $\kappa(Q) = 1$ ，是最小的条件数。当用正交矩阵对其他矩阵进行变换或分解时，能够改善原矩阵的条件数，使计算结果更稳定、可靠。比如在最小二乘法问题中，使用正交矩阵进行QR分解可以有效改善问题的条件数，提高求解的精度和稳定性。

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

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

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

若正方形矩阵 $Q \in R^{n \times n}$, 且满足

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

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 $Q^T Q = I_n$, 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 $Q Q^T = I_m$, 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

$$Q Q^H = Q^H Q = I_n$$

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 $B = U^H A U$ ，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

在矩阵计算中，三角矩阵（包括上三角矩阵和下三角矩阵）具有特殊的结构和性质，在简化计算、求解线性方程组、求矩阵的行列式和特征值等方面都有重要作用，具体如下：

简化矩阵乘法运算

- **上三角矩阵相乘：**两个上三角矩阵相乘的结果仍然是上三角矩阵，且其主对角线及上方元素的计算只需考虑参与运算矩阵的对应位置及上方的元素，下方元素均为零，无需计算，大大减少了乘法和加法的运算量。
- **下三角矩阵相乘：**同理，两个下三角矩阵相乘结果为下三角矩阵，计算时只需关注主对角线及下方元素，简化了计算过程。

6. 三角矩阵

利于求矩阵的行列式

- **上三角矩阵**: 上三角矩阵的行列式等于其主对角线元素之积, 即 $|A| = \prod_{i=1}^n a_{ii}$, 无需进行复杂的行列式展开计算。
- **下三角矩阵**: 下三角矩阵的行列式同样等于主对角线元素的乘积, 计算非常简便, 这在计算高阶矩阵行列式时优势尤为明显。

便于求矩阵的特征值

- **上三角矩阵**: 根据特征方程 $|\lambda I - A| = 0$, 对于上三角矩阵 A , $\lambda I - A$ 仍然是上三角矩阵, 其行列式为 $|\lambda I - A| = \prod_{i=1}^n (\lambda - a_{ii}) = 0$, 所以上三角矩阵的特征值就是其主对角线元素 $a_{11}, a_{22}, \dots, a_{nn}$ 。
- **下三角矩阵**: 下三角矩阵的特征值求解与上三角矩阵类似, 也是其主对角线元素, 这种特性使得三角矩阵在特征值计算中具有很大的便利性, 无需通过复杂的计算来求解特征多项式的根。

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 $u_{ij} = 0, i > j$:

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

6. 三角矩阵

下三角矩阵 (*lower triangular matrix*) $L = [l_{ij}]_{n \times n}$ 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 $U = [u_{ij}]$ 的行列式是对角线元素的乘积，即

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

(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 $A = T^H D T$, 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, \dots, L_k 为下三角矩阵，则 $L_1 L_2 \dots L_k$ 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

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

(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