作者	**生姜 DrGinger**
脚本	**生姜 DrGinger**
视频	**崔崔 CuiCui**
开源学习资源	https://github.com/Visualize-ML
平台	https://www.youtube.com/@DrGinger_Jiang https://space.bilibili.com/3546865719052873 https://space.bilibili.com/513194466

9.3 正交矩阵性质

本节你将掌握的核心技能：

- ▶ 正交矩阵：方阵的逆等价于其转置。
- ▶ 正交矩阵的列向量、行向量都构成规范正交基。
- ▶ 矩阵乘法不同视角展开正交矩阵的格拉姆矩阵。
- ▶ 正交矩阵的规范正交基保持向量长度、夹角不变。
- ▶ 正交矩阵使求逆和投影计算变得简单且稳定。
- ▶ 半正交矩阵、正交矩阵关系。
- ▶ 伪逆是对非方阵或不可逆矩阵的一种“广义逆”。

本节介绍正交矩阵的各种性质，内容过于“硬核”；这也是，为什么我们先用前两节做铺垫的原因。请大家务必掌握前两节内容，特别是几何视角理解正交矩阵之后，再学习本节。本节用到了前文各种线性代数工具，比如矩阵乘法第一、第二视角，格拉姆矩阵，逆矩阵，矩阵转置，正交投影，行列式，规范正交基，分块矩阵乘法等。

定义

如果实数 $n \times n$ 方阵 $\boldsymbol{A}$ 是正交矩阵，当且仅当

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I} \tag{1}$$

图 1 所示为 $\boldsymbol{A}^{\mathrm{T}} @ \boldsymbol{A} = \boldsymbol{I}$ 对应的热图，$\boldsymbol{A}^{\mathrm{T}} @ \boldsymbol{A}$ 是矩阵 $\boldsymbol{A}$ 的格拉姆矩阵。

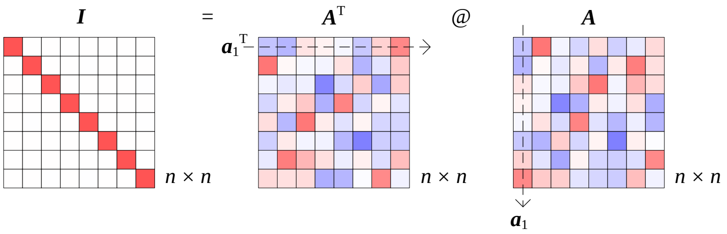

图 1. $\boldsymbol{A}^{\mathrm{T}} @ \boldsymbol{A} = \boldsymbol{I}$

(1) 还告诉我们 $\boldsymbol{A}$ 的逆矩阵等于它的转置矩阵

$$\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}} \tag{2}$$

正交矩阵这个特性让求逆变得特别方便。

此外，(1) 说明 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 都是正交矩阵，具体如图 2 所示。$\boldsymbol{A} @ \boldsymbol{A}^{\mathrm{T}}$ 是矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的格拉姆矩阵。

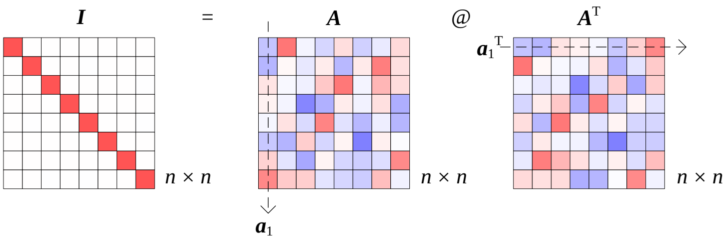

图 2. $\boldsymbol{A} @ \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$

下面让我们用本书前文介绍的矩阵乘法不同展开视角来分析图 1、图 2 对应的矩阵乘法。

矩阵乘法第一视角展开 $\boldsymbol{A}^{\mathrm{T}} @ \boldsymbol{A} = \boldsymbol{I}$

让我们先用矩阵乘法第一视角——内积视角——展开 $\boldsymbol{A}^{\mathrm{T}} @ \boldsymbol{A} = \boldsymbol{I}$，即

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{a}_n^{\mathrm{T}} \end{bmatrix} @ \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} @ \boldsymbol{a}_1 & \boldsymbol{a}_1^{\mathrm{T}} @ \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1^{\mathrm{T}} @ \boldsymbol{a}_n \\ \boldsymbol{a}_2^{\mathrm{T}} @ \boldsymbol{a}_1 & \boldsymbol{a}_2^{\mathrm{T}} @ \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_2^{\mathrm{T}} @ \boldsymbol{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_n^{\mathrm{T}} @ \boldsymbol{a}_1 & \boldsymbol{a}_n^{\mathrm{T}} @ \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n^{\mathrm{T}} @ \boldsymbol{a}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{3}$$

我们也可以把 (3) 写成内积形式 (即内积视角)

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_1 \cdot \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1 \cdot \boldsymbol{a}_n \\ \boldsymbol{a}_2 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_2 \cdot \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_2 \cdot \boldsymbol{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_n \cdot \boldsymbol{a}_1 & \boldsymbol{a}_n \cdot \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \cdot \boldsymbol{a}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{4}$$

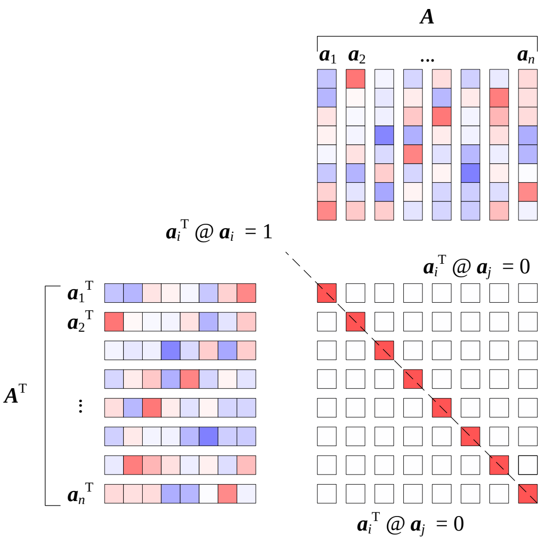

图 3. $A^{T} @ A = I$，矩阵乘法第一视角 (内积视角)

图 3 中主对角线元素为 1，这说明正交矩阵的**列向量**均为单位向量 (方向向量)，向量长度 (大小、模、L^2 范数、欧几里得范数) 为 1，即

$$a_i^{T} @ a_i = a_i \cdot a_i = \langle a_i, a_i \rangle = \|a_i\|_2^2 = 1 \tag{5}$$

图 3 中非主对角线元素为 0，这说明正交矩阵的**列向量**两两正交，即

$$a_i^{T} @ a_j = a_j^{T} @ a_i = a_i \cdot a_j = a_j \cdot a_i = \langle a_i, a_j \rangle = \langle a_j, a_i \rangle = 0 \tag{6}$$

其中，i 不等于 j。

综上，正交矩阵 A **列向量**均为单位向量，且两两正交。

矩阵乘法第二视角展开 $A^{T} @ A = I$

再用矩阵乘法第二视角——外积视角——展开 $A^{T} @ A = I$，即

$$A^{T}A = \begin{bmatrix} a^{(1)T} & a^{(2)T} & \cdots & a^{(n)T} \end{bmatrix} @ \begin{bmatrix} a^{(1)} \\ a^{(2)} \\ \vdots \\ a^{(n)} \end{bmatrix} = a^{(1)T} @ a^{(1)} + a^{(2)T} @ a^{(2)} + \cdots + a^{(n)T} @ a^{(n)} = I \tag{7}$$

显然，A 的行向量都是非零向量，因此，$a^{(i)T} @ a^{(i)}$ 也是秩一矩阵。

此外，$a^{(i)T} @ a^{(i)}$为对称矩阵，即满足

$$\left(a^{(i)} @ a^{(i)T}\right)^{T} = a^{(i)T} @ a^{(i)} \tag{8}$$

如图 4 所示，n 个秩一矩阵之和为单位矩阵。

根据本章前两节内容，图 4 中每个秩一矩阵，即 $\boldsymbol{a}^{(i)\mathrm{T}} @ \boldsymbol{a}^{(i)}$，都是投影矩阵。$(\boldsymbol{a}^{(i)\mathrm{T}} @ \boldsymbol{a}^{(i)}) @ \boldsymbol{x}$ 就是向量 $\boldsymbol{x}$ 朝 $\boldsymbol{a}^{(i)\mathrm{T}}$ 的正交投影。

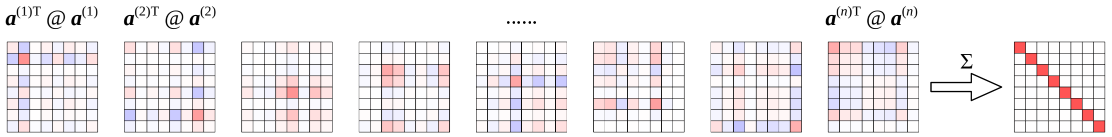

图 4. $\boldsymbol{A}^{\mathrm{T}} @ \boldsymbol{A} = \boldsymbol{I}$，矩阵乘法第二视角 (外积视角)

矩阵乘法第一视角展开 $\boldsymbol{A} @ \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$

对于矩阵乘法 $\boldsymbol{A} @ \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$，让我们也先用矩阵乘法第一视角——内积视角——展开，即

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{a}^{(1)} \\ \boldsymbol{a}^{(2)} \\ \vdots \\ \boldsymbol{a}^{(n)} \end{bmatrix} @ \begin{bmatrix} \boldsymbol{a}^{(1)\mathrm{T}} & \boldsymbol{a}^{(2)\mathrm{T}} & \cdots & \boldsymbol{a}^{(n)\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}^{(1)} @ \boldsymbol{a}^{(1)\mathrm{T}} & \boldsymbol{a}^{(1)} @ \boldsymbol{a}^{(2)\mathrm{T}} & \cdots & \boldsymbol{a}^{(1)} @ \boldsymbol{a}^{(n)\mathrm{T}} \\ \boldsymbol{a}^{(2)} @ \boldsymbol{a}^{(1)\mathrm{T}} & \boldsymbol{a}^{(2)} @ \boldsymbol{a}^{(2)\mathrm{T}} & \cdots & \boldsymbol{a}^{(2)} @ \boldsymbol{a}^{(n)\mathrm{T}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}^{(n)} @ \boldsymbol{a}^{(1)\mathrm{T}} & \boldsymbol{a}^{(n)} @ \boldsymbol{a}^{(2)\mathrm{T}} & \cdots & \boldsymbol{a}^{(n)} @ \boldsymbol{a}^{(n)\mathrm{T}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{9}$$

把 (9) 写成内积形式 (即内积视角)

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{a}^{(1)} \cdot \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(1)} \cdot \boldsymbol{a}^{(2)} & \cdots & \boldsymbol{a}^{(1)} \cdot \boldsymbol{a}^{(n)} \\ \boldsymbol{a}^{(2)} \cdot \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(2)} \cdot \boldsymbol{a}^{(2)} & \cdots & \boldsymbol{a}^{(2)} \cdot \boldsymbol{a}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}^{(n)} \cdot \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(n)} \cdot \boldsymbol{a}^{(2)} & \cdots & \boldsymbol{a}^{(n)} \cdot \boldsymbol{a}^{(n)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{10}$$

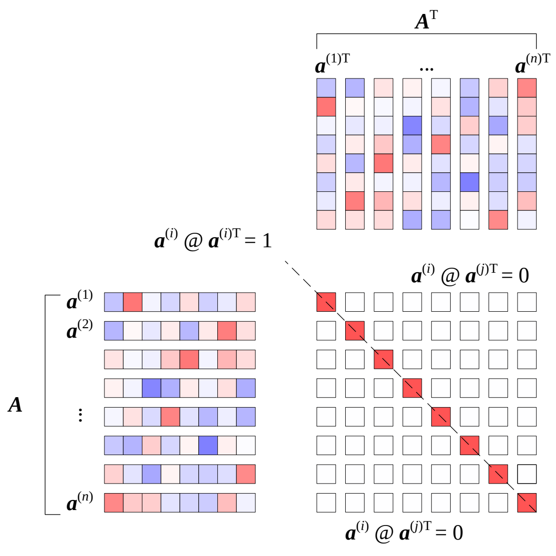

图 5. $\boldsymbol{A}^{\mathrm{T}} @ \boldsymbol{A} = \boldsymbol{I}$，矩阵乘法第一视角 (内积视角)

图 5 中主对角线元素为 1，这说明正交矩阵 $\boldsymbol{A}$ 的**行向量**均为单位向量 (方向向量)，向量长度 (大小、模、L^2 范数、欧几里得范数) 为 1，即

$$\boldsymbol{a}^{(i)} @ \boldsymbol{a}^{(i)\mathrm{T}} = \boldsymbol{a}^{(i)} \cdot \boldsymbol{a}^{(i)} = \left\langle \boldsymbol{a}^{(i)}, \boldsymbol{a}^{(i)} \right\rangle = \left\| \boldsymbol{a}^{(i)} \right\|_2^2 = 1 \tag{11}$$

图 5 中非主对角线元素为 0，这说明正交矩阵的**行向量**两两正交，即

$$\boldsymbol{a}^{(i)} @ \boldsymbol{a}^{(j)\mathrm{T}} = \boldsymbol{a}^{(j)} @ \boldsymbol{a}^{(i)\mathrm{T}} = \boldsymbol{a}^{(i)} \cdot \boldsymbol{a}^{(j)} = \boldsymbol{a}^{(j)} \cdot \boldsymbol{a}^{(i)} = \left\langle \boldsymbol{a}^{(i)}, \boldsymbol{a}^{(j)} \right\rangle = \left\langle \boldsymbol{a}^{(j)}, \boldsymbol{a}^{(i)} \right\rangle = 0 \tag{12}$$

其中，i 不等于 j。

综上，正交矩阵 $\boldsymbol{A}$ **行向量**均为单位向量，且两两正交。

矩阵乘法第二视角展开 $A @ A^{\mathrm{T}} = I$

矩阵乘法第二视角——外积视角——展开 $\boldsymbol{A} @ \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$，即

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix} @ \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{a}_n^{\mathrm{T}} \end{bmatrix} = \boldsymbol{a}_1 @ \boldsymbol{a}_1^{\mathrm{T}} + \boldsymbol{a}_2 @ \boldsymbol{a}_2^{\mathrm{T}} + \cdots + \boldsymbol{a}_n @ \boldsymbol{a}_n^{\mathrm{T}} = \boldsymbol{I} \tag{13}$$

$\boldsymbol{A}$ 列向量都是非零向量，因此，$\boldsymbol{a}_i @ \boldsymbol{a}_i^{\mathrm{T}}$ 为秩一矩阵。

此外，$\boldsymbol{a}_i @ \boldsymbol{a}_i^{\mathrm{T}}$ 为对称矩阵，满足

$$\left(\boldsymbol{a}_i @ \boldsymbol{a}_i^{\mathrm{T}}\right)^{\mathrm{T}} = \boldsymbol{a}_i @ \boldsymbol{a}_i^{\mathrm{T}} \tag{14}$$

如图 6 所示，n 个秩一矩阵之和为单位矩阵。每个秩一矩阵也都是投影矩阵。

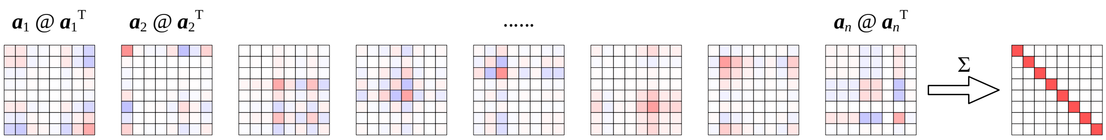

图 6．$\boldsymbol{A} @ \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$，矩阵乘法第二视角 (外积视角)

行列式绝对值为 1：面积保持不变

如果矩阵 $\boldsymbol{A}$ 为正交矩阵，它的行列式为±1。证明也很简单，具体如下

$$\det(\boldsymbol{A})^2 = \det\left(\boldsymbol{A}^{\mathrm{T}}\right)\det(\boldsymbol{A}) = \det\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right) = \det(\boldsymbol{I}) = 1 \tag{15}$$

本书前文提过，单位矩阵、旋转矩阵、置换矩阵、镜像矩阵，以及它们的组合都是正交矩阵。

显然单位矩阵满足 (1)，即

$$\boldsymbol{I}^{\mathrm{T}}\boldsymbol{I} = \boldsymbol{I}\boldsymbol{I}^{\mathrm{T}} = \boldsymbol{I} \tag{16}$$

比如，

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{\mathrm{T}} @ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} @ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{17}$$

单位矩阵的行列式为 1。

平面 (绕原点) 逆时针旋转矩阵也满足 (1)

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{\mathrm{T}} @ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} @ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{18}$$

上式旋转矩阵的行列式为 1。

可以这样理解，交换单位矩阵的行或列，我们便得到置换矩阵。矩阵乘法中，置换矩阵的作用是调换行或列。置换矩阵也是正交矩阵，比如下例

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{\mathrm{T}} @ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} @ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{19}$$

上式置换矩阵的行列式为−1。

⚠ 注意，置换矩阵的行列式可以是−1，也可以是 1；取决于置换次数的奇偶数。这也告诉我们正交矩阵的行列式可以是 1，也可以是−1。

镜像矩阵也是正交矩阵，比如

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^{\mathrm{T}} @ \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} @ \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{20}$$

⚠ 注意，镜像矩阵的行列式也是−1。

规范正交基

既然 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 都是正交矩阵，$\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 也都是规范正交基。

首先，$\boldsymbol{A}$ 的列向量线性组合 $\mathrm{span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ 张成 $\mathbb{R}^n$ 。

用前两节的思路，给定 n 维列向量 $\boldsymbol{x}$，如下乘法代表 $\boldsymbol{x}$ 向 $\boldsymbol{A}$ 的列向量分别投影，再求和

$$\boldsymbol{I}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} = \left(\boldsymbol{a}_1 @ \boldsymbol{a}_1^{\mathrm{T}} + \boldsymbol{a}_2 @ \boldsymbol{a}_2^{\mathrm{T}} + \cdots + \boldsymbol{a}_n @ \boldsymbol{a}_n^{\mathrm{T}}\right)\boldsymbol{x} = \left(\boldsymbol{a}_1 @ \boldsymbol{a}_1^{\mathrm{T}}\right) @ \boldsymbol{x} + \left(\boldsymbol{a}_2 @ \boldsymbol{a}_2^{\mathrm{T}}\right) @ \boldsymbol{x} + \cdots + \left(\boldsymbol{a}_n @ \boldsymbol{a}_n^{\mathrm{T}}\right) @ \boldsymbol{x} \tag{21}$$

图 6 中每个秩一矩阵，即 $\boldsymbol{a}_i @ \boldsymbol{a}_i^{\mathrm{T}}$，都是正交投影矩阵。

$(\boldsymbol{a}_i @ \boldsymbol{a}_i^{\mathrm{T}}) @ \boldsymbol{x}$ 就是向量 $\boldsymbol{x}$ 朝 $\boldsymbol{a}_i$ 的正交投影。

类似地，$\boldsymbol{A}^{\mathrm{T}}$ 的列向量线性组合 $\mathrm{span}(\boldsymbol{a}^{(1)\mathrm{T}}, \boldsymbol{a}^{(2)\mathrm{T}}, \ldots, \boldsymbol{a}^{(n)\mathrm{T}})$ 也张成 $\mathbb{R}^n$ 。请大家试着解释如下等式

$$\boldsymbol{I}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \left(\boldsymbol{a}^{(1)\mathrm{T}} @ \boldsymbol{a}^{(1)}\right) @ \boldsymbol{x} + \left(\boldsymbol{a}^{(2)\mathrm{T}} @ \boldsymbol{a}^{(2)}\right) @ \boldsymbol{x} + \cdots + \left(\boldsymbol{a}^{(n)\mathrm{T}} @ \boldsymbol{a}^{(n)}\right) @ \boldsymbol{x} \tag{22}$$

图 4 中每个秩一矩阵，即 $\boldsymbol{a}^{(i)\mathrm{T}} @ \boldsymbol{a}^{(i)}$，也都是投影矩阵。$(\boldsymbol{a}^{(i)\mathrm{T}} @ \boldsymbol{a}^{(i)}) @ \boldsymbol{x}$ 就是向量 $\boldsymbol{x}$ 朝 $\boldsymbol{a}^{(i)\mathrm{T}}$ 的正交投影。

保持向量长度不变

对任意 n 维向量 $\boldsymbol{x}$，$\boldsymbol{A}\boldsymbol{x}$ 的模等于 $\boldsymbol{x}$ 的模，即

$$\|\boldsymbol{A}\boldsymbol{x}\| = \|\boldsymbol{x}\| \tag{23}$$

证明很简单，具体如下

$$\|Ax\|^2 = (Ax)^{\mathrm{T}}(Ax) = x^{\mathrm{T}} \underbrace{A^{\mathrm{T}}A}_{I} x = x^{\mathrm{T}}x = \|x\|^2 \tag{24}$$

以旋转为例，如图 7 所示，几何角度来看，旋转不会改变向量的长度。

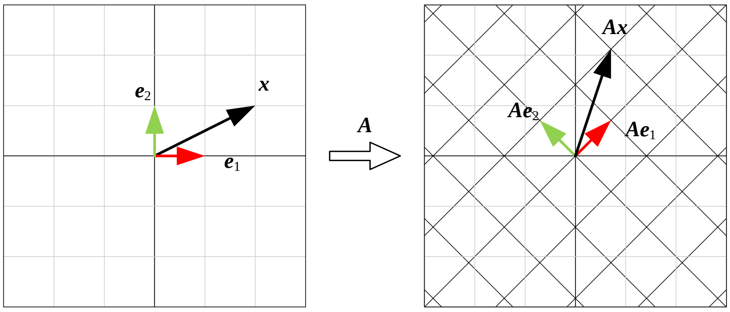

图 7. 正交变换不改变向量长度

对任意 n 维向量，$A^{\mathrm{T}}x$ 的模等于 x 的模，即

$$\|A^{\mathrm{T}}x\| = \|x\| \tag{25}$$

请大家自行证明 (25)，并解释几何角度意味着什么。

保持角度与内积不变

如果 A 为正交矩阵，对两个 n 维向量 x、y，Ax、Ay 的夹角等于 x、y 夹角。

证明也很简单，我们已经知道 Ax 的模等于 x 的模，Ay 的模等于 y 的模。只需要证明 Ax、Ay 的内积等于 x、y 的内积，即

$$(Ax)\cdot(Ay) = x \cdot y \tag{26}$$

用矩阵乘法计算上述内积

$$(Ax)^{\mathrm{T}} @ (Ay) = x^{\mathrm{T}} @ y \tag{27}$$

整理便得证

$$(Ax)^{\mathrm{T}} @ (Ay) = x^{\mathrm{T}} \underbrace{A^{\mathrm{T}}A}_{I} y = x^{\mathrm{T}}y \tag{28}$$

如图 8 所示，几何角度来看，正交变换不会改变向量之间夹角。

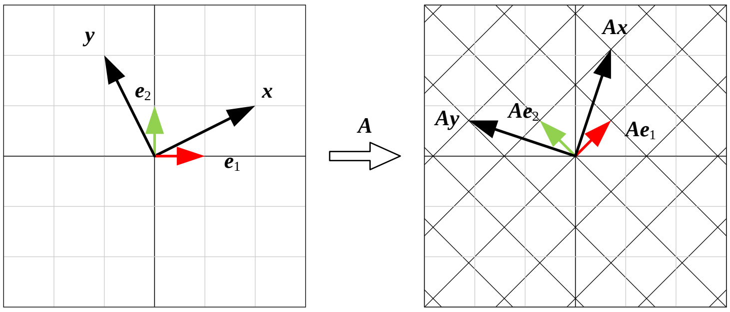

图 8. 正交变换不改变向量之间夹角

正交矩阵的连乘仍然是正交矩阵

如果 A_1、A_2 都是正交矩阵，则

$$\left(A_1A_2\right)^{\mathrm{T}} @ \left(A_1A_2\right) = A_2^{\mathrm{T}} \underbrace{A_1^{\mathrm{T}} A_1}_{I} A_2 = A_2^{\mathrm{T}} A_2 = I \tag{29}$$

同样

$$\left(A_1A_2\right) @ \left(A_1A_2\right)^{\mathrm{T}} = A_1 \underbrace{A_2 A_2^{\mathrm{T}}}_{I} A_1^{\mathrm{T}} = A_1 A_1^{\mathrm{T}} = I \tag{30}$$

这意味着“旋转 → 镜像”、“置换 → 镜像”、“旋转 → 置换 → 镜像”等等复合几何变换矩阵也都是正交矩阵。

半正交矩阵

一个非方阵 A_1，满足

$$A_1^{\mathrm{T}} A_1 = I \tag{31}$$

但不满足

$$A_1 A_1^{\mathrm{T}} = I \tag{32}$$

我们称其为半正交 (semi-orthogonal matrix)。这个矩阵的列向量也构成规范正交基。

实际上，图 1 中的矩阵 A 任意去掉若干列剩下的矩阵就是半正交矩阵 A_1，具体如图 9 所示。

⚠ 注意，删除的列向量不必连续。

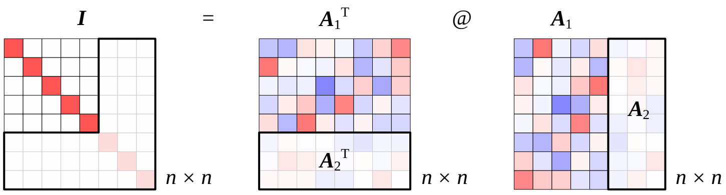

图 9. 正交矩阵删去若干列得到半正交矩阵

如图 10 所示，$\boldsymbol{A}$ 矩阵取消的部分 ($\boldsymbol{A}_2$) 和保留的子矩阵 ($\boldsymbol{A}_1$) 张成的空间互为正交补。

把 $\boldsymbol{A}$ 写成 $[\boldsymbol{A}_1, \boldsymbol{A}_2]$，根据本书前文介绍的分块矩阵乘法，矩阵乘法 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可以展开

$$\begin{aligned}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} &= \begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2\end{bmatrix} \\ &= \begin{bmatrix}\boldsymbol{A}_1^{\mathrm{T}} \\ \boldsymbol{A}_2^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_1 & \boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_2 \\ \boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{A}_1 & \boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{A}_2\end{bmatrix} = \begin{bmatrix}\boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I}\end{bmatrix}\end{aligned} \tag{33}$$

上式中，$\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_1$、$\boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{A}_2$ 均为半正交矩阵的格拉姆矩阵。

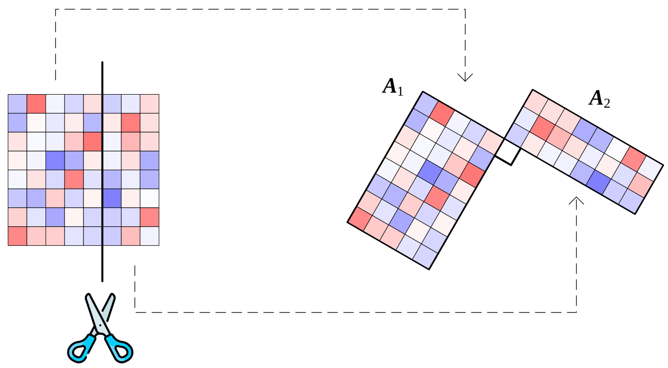

图 10. 张成的空间互为正交补

举个例子，给定 3×2 矩阵 $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{bmatrix}1 & 0 \\ 0 & 0 \\ 0 & 1\end{bmatrix} \tag{34}$$

再计算 (34) 的 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{bmatrix}1 & 0 & 0 \\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & 0 \\ 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} \tag{35}$$

结果为单位矩阵。这意味着 $\boldsymbol{A}$ 的列向量均为单位向量，并且两两正交。

再计算 (34) 的 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix}1 & 0 \\ 0 & 0 \\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 \\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1\end{bmatrix} \tag{36}$$

这个结果不是单位矩阵。

因此，$\boldsymbol{A}$ 是半正定矩阵，$\boldsymbol{A}$ 的列向量为规范正交基；但是，$\boldsymbol{A}$ 不是正定矩阵。

也请大家注意，如果矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$，但是不满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$，这意味着矩阵 $\boldsymbol{A}$ 的行向量构成规范正交基。有些时候，我们管这个矩阵 $\boldsymbol{A}$ 也叫做半正定矩阵。

举个例子，比如 2×3 矩阵 $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{bmatrix} 4/5 & 0 & 3/5 \\ -3/5 & 0 & 4/5 \end{bmatrix} \tag{37}$$

再计算 (37) 的 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{bmatrix} 4/5 & -3/5 \\ 0 & 0 \\ 3/5 & 4/5 \end{bmatrix} \begin{bmatrix} 4/5 & 0 & 3/5 \\ -3/5 & 0 & 4/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{38}$$

结果不是单位矩阵。

再计算 (37) 中 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} 4/5 & 0 & 3/5 \\ -3/5 & 0 & 4/5 \end{bmatrix} \begin{bmatrix} 4/5 & -3/5 \\ 0 & 0 \\ 3/5 & 4/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{39}$$

这个结果是单位矩阵。

这意味着 $\boldsymbol{A}$ 的行向量均为单位向量，并且两两正交。

因此，$\boldsymbol{A}$ 也可以叫做半正定矩阵，$\boldsymbol{A}$ 的行向量为规范正交基；但是，$\boldsymbol{A}$ 不是正定矩阵。

一般矩阵的正交投影

我们这么“渴望”正交矩阵，是因为在处理正交投影问题时，一般矩阵的运算很麻烦！

下面让我们举个例子看一下。

给定如下两个三维列向量

$$\boldsymbol{x}_1 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \tag{40}$$

如图 11 所示，显然 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$ 不共线，非正交基 $[\boldsymbol{x}_1, \boldsymbol{x}_2]$ 在三维空间张成平面 $\mathrm{span}(\boldsymbol{x}_1, \boldsymbol{x}_2)$。

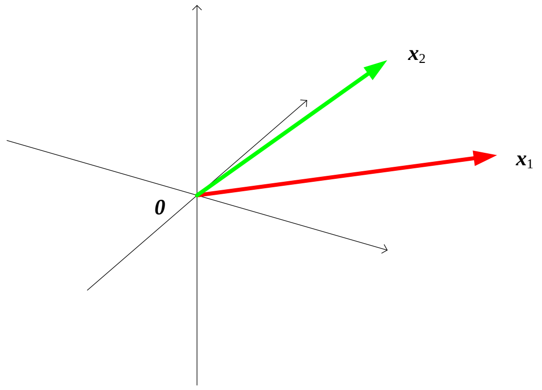

图 11. $\boldsymbol{x}_1$、$\boldsymbol{x}_2$具体位置

再给定三维列向量 $\boldsymbol{y}$

$$\boldsymbol{y} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \tag{41}$$

如图 12 所示，向量 $\boldsymbol{y}$ 不在平面 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 上。

向量 $\boldsymbol{y}$ 向 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 正交投影，结果为 $\hat{\boldsymbol{y}}$。$\hat{\boldsymbol{y}}$ 显然在 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 平面上！

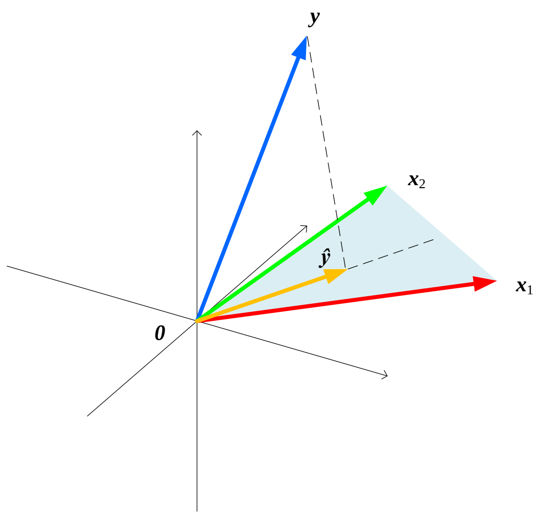

图 12. $\boldsymbol{y}$ 向 span($\boldsymbol{x}_1$、$\boldsymbol{x}_2$) 平面投影

由于 [$\boldsymbol{x}_1$, $\boldsymbol{x}_2$] 是 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 的非正交基；也就是说，$\hat{\boldsymbol{y}}$ 可以写成 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$ 的线性组合

$$\hat{\boldsymbol{y}} = b_1\boldsymbol{x}_1 + b_2\boldsymbol{x}_2 = b_1\begin{bmatrix}2\\2\\0\end{bmatrix} + b_2\begin{bmatrix}2\\0\\2\end{bmatrix} \tag{42}$$

想要求得 $\hat{\boldsymbol{y}}$，我们就需要计算 b_1、b_2。

把 (42) 写成矩阵乘法形式

$$\hat{\boldsymbol{y}} = \underbrace{\begin{bmatrix}\boldsymbol{x}_1 & \boldsymbol{x}_2\end{bmatrix}}_{\boldsymbol{X}}\underbrace{\begin{bmatrix}b_1\\b_2\end{bmatrix}}_{\boldsymbol{b}} = \begin{bmatrix}\begin{bmatrix}2\\2\\0\end{bmatrix} & \begin{bmatrix}2\\0\\2\end{bmatrix}\end{bmatrix}\boldsymbol{b} = \begin{bmatrix}2 & 2\\2 & 0\\0 & 2\end{bmatrix}\boldsymbol{b} = \boldsymbol{X}\boldsymbol{b} \tag{43}$$

矩阵 $\boldsymbol{X}$ 为非方阵，显然不可逆！

让我们想想其他办法！

如图 13 所示，向量 $\boldsymbol{y}$ 可以正交分解成两部分：

a) $\hat{\boldsymbol{y}}$，在 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 平面上；

b) $\boldsymbol{\varepsilon}$，垂直于 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$)。

$\boldsymbol{\varepsilon}$ 为 $\boldsymbol{y}$、$\hat{\boldsymbol{y}}$ 两者之差

$$\boldsymbol{\varepsilon} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{b} \tag{44}$$

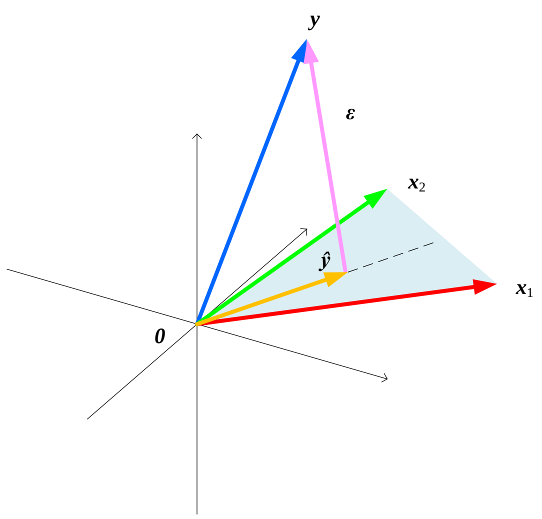

图 13. $\boldsymbol{y}$ 正交分解成两部分

让我们从 $\boldsymbol{\varepsilon}$ 入手！

既然，$\boldsymbol{\varepsilon}$ 垂直于 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$)，$\boldsymbol{\varepsilon}$ 必然分别垂直于 $\boldsymbol{x}_1$、$\boldsymbol{x}_1$，于是有

$$\begin{aligned}\boldsymbol{x}_1^{\mathrm{T}} @ \boldsymbol{\varepsilon} = 0\\ \boldsymbol{x}_2^{\mathrm{T}} @ \boldsymbol{\varepsilon} = 0\end{aligned} \tag{45}$$

把这两个式子合并

$$\begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}} \\ \boldsymbol{x}_2^{\mathrm{T}} \end{bmatrix} @ \boldsymbol{\varepsilon} = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{X}^{\mathrm{T}} \boldsymbol{\varepsilon} = \boldsymbol{0} \tag{46}$$

将 (44) 带入 (46)

$$\boldsymbol{X}^{\mathrm{T}} \left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b} \right) = \boldsymbol{0} \tag{47}$$

整理上式得到

$$\boldsymbol{X}^{\mathrm{T}} \boldsymbol{y} = \boldsymbol{X}^{\mathrm{T}} \boldsymbol{X} \boldsymbol{b} \tag{48}$$

上式实际上就是在 $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b}$ 等式两侧左边分别乘 $\boldsymbol{X}^{\mathrm{T}}$。

由于 $\boldsymbol{X}$ 列满秩 (列向量不共线)，$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 这个格拉姆矩阵可逆，因此

$$\boldsymbol{b} = \left(\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X} \right)^{-1} \boldsymbol{X}^{\mathrm{T}} \boldsymbol{y} \tag{49}$$

实际上，如图 14 所似乎，$\left(\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X} \right)^{-1} \boldsymbol{X}^{\mathrm{T}}$ 叫**伪逆** (pseudo-inverse)，也叫**广义逆** (generalized inverse)。

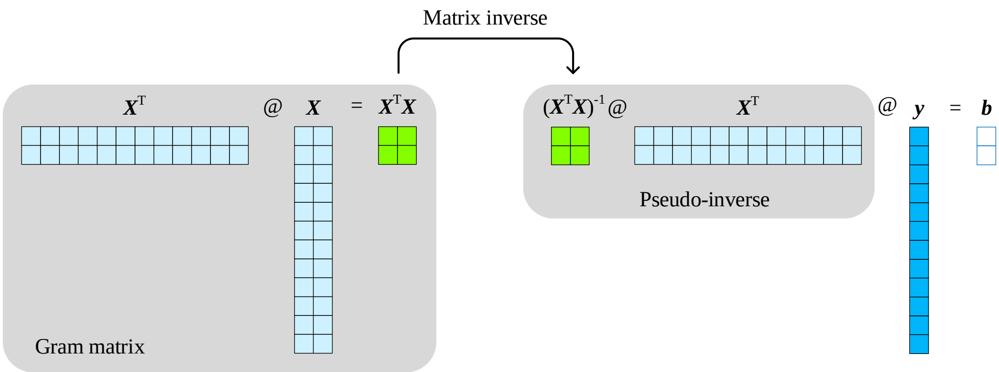

图 14. 计算 $\boldsymbol{b}$

简单来说，伪逆是对非方阵 (比如 $\boldsymbol{X}$) 或不可逆矩阵的一种“广义逆”，用于在无法直接求逆时近似求解。

⚠ 必须强调一点，只有 $\boldsymbol{X}$ 为列满秩时，$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 才存在逆。

这样就求得 $\boldsymbol{y}$ 在 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 平面上的投影结果 $\hat{\boldsymbol{y}}$ 可以写成

$$\hat{\boldsymbol{y}} = \boldsymbol{X} \left(\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X} \right)^{-1} \boldsymbol{X}^{\mathrm{T}} \boldsymbol{y} \tag{50}$$

其中，$\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$ 就是 $\hat{\boldsymbol{y}}$ 向 $\boldsymbol{X}$ 列项构造的向量空间正交投影的投影矩阵。

上式中，$\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$ 常被称作**帽子矩阵** (hat matrix)。

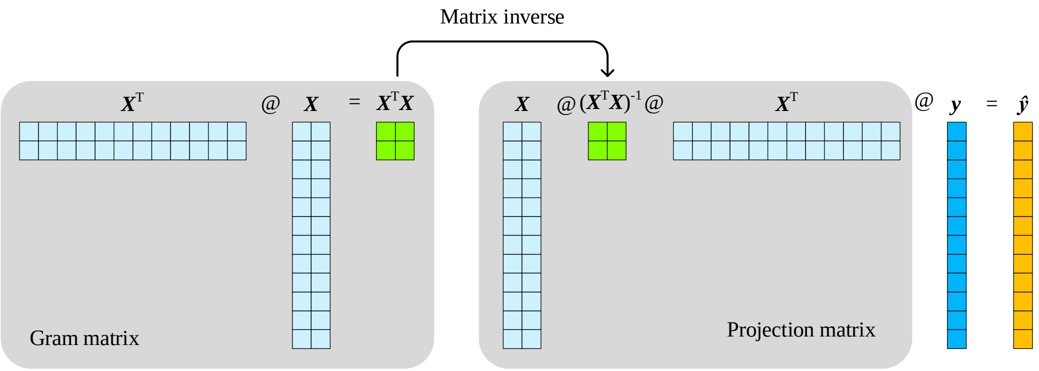

图 15. 计算 $\hat{\boldsymbol{y}}$

$\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$ 是**幂等矩阵** (idempotent matrix)。

本书前文提过，投影矩阵也是幂等矩阵；幂等矩阵与自身相乘仍等于自身，即

$$\left(\boldsymbol{X}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\right)^{2}=\boldsymbol{X}\underbrace{\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}_{\boldsymbol{I}}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}=\boldsymbol{X}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}} \tag{51}$$

半正交矩阵的正交投影

如果我们在 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 平面找到一个规范正交基，它们同样张成 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$) 平面的话，完成上述投影运算便“易如反掌”！

通过某个特殊的方法，我们还真找到了一个规范正交基

$$\boldsymbol{Q}_1=\begin{bmatrix}\boldsymbol{q}_1 & \boldsymbol{q}_2\end{bmatrix}=\begin{bmatrix}\sqrt{2}/2 & \sqrt{6}/6\\ \sqrt{2}/2 & -\sqrt{6}/6\\ 0 & \sqrt{6}/3\end{bmatrix} \tag{52}$$

如图 16 所示，列向量 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 张成的平面 span($\boldsymbol{q}_1$, $\boldsymbol{q}_2$) 等价于 span($\boldsymbol{x}_1$, $\boldsymbol{x}_2$)。

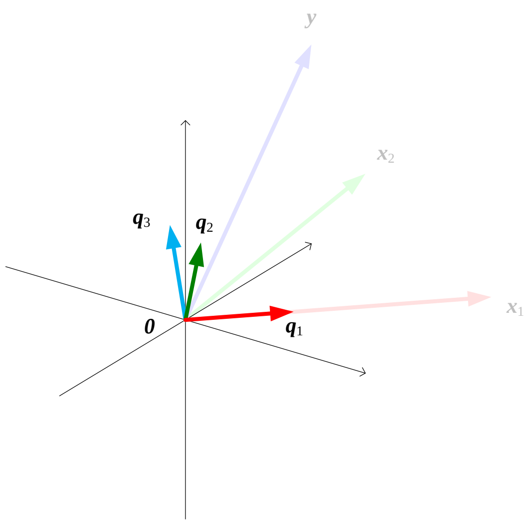

图 16. 规范正交基 $[\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3]$

$\mathrm{span}(\boldsymbol{q}_1, \boldsymbol{q}_2)$ 为什么等价 $\mathrm{span}(\boldsymbol{x}_1, \boldsymbol{x}_2)$？我们如何计算得到 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$？这些问题将在下一节得到解答。

列向量 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 都在 $\mathrm{span}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ 平面上；$\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 均为单位向量，且正交。因此，$[\boldsymbol{q}_1, \boldsymbol{q}_2]$ 为规范正交基。

这也告诉我们，(52) 中 $\boldsymbol{Q}_1$ 是半正交矩阵，即满足 $\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{Q}_1 = \boldsymbol{I}$。

用 $\boldsymbol{Q}_1$ 替换 (50) 中的 $\boldsymbol{X}$，我们可以得到

$$\begin{aligned}\hat{\boldsymbol{y}} &= \boldsymbol{Q}_1 \underbrace{\left(\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{Q}_1\right)}_{\boldsymbol{I}}{}^{-1} \boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{y} \\ &= \boldsymbol{Q}_1\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{y}\end{aligned} \tag{53}$$

把 $\boldsymbol{Q}_1$ 写成 $[\boldsymbol{q}_1, \boldsymbol{q}_2]$，展开上式矩阵乘法

$$\begin{aligned}\hat{\boldsymbol{y}} &= \boldsymbol{Q}_1\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{y} \\ &= \begin{bmatrix}\boldsymbol{q}_1 & \boldsymbol{q}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{q}_1^{\mathrm{T}} \\ \boldsymbol{q}_2^{\mathrm{T}}\end{bmatrix}\boldsymbol{y} \\ &= \left(\boldsymbol{q}_1 @ \boldsymbol{q}_1^{\mathrm{T}} + \boldsymbol{q}_2 @ \boldsymbol{q}_2^{\mathrm{T}}\right)\boldsymbol{y} \\ &= \boldsymbol{q}_1 @ \boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{y} + \boldsymbol{q}_2 @ \boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{y}\end{aligned} \tag{54}$$

这相当于 $\boldsymbol{y}$ 向 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 分别正交投影，然后再合成！

半正交矩阵 $\boldsymbol{Q}_1$ 的 $\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{Q}_1 = \boldsymbol{I}$ 这个性质起到了至关重要的作用！

细心的读者可能已经发现，撑起三维空间需要三根线性无关的“棍子“。

而撑起三维空间的规范正交基要求三个“棍子”的长度为 1，且两两正交。

现在有了正交单位向量 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$，还缺一根 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 和均正交的单位向量，我们管它叫 $\boldsymbol{q}_3$。

而 $\boldsymbol{q}_3$ 可以用下式求解

$$\boldsymbol{q}_3 @ \boldsymbol{q}_3^{\mathrm{T}} = \boldsymbol{I} - \boldsymbol{q}_1 @ \boldsymbol{q}_1^{\mathrm{T}} + \boldsymbol{q}_2 @ \boldsymbol{q}_2^{\mathrm{T}} \tag{55}$$

显然，$\boldsymbol{q}_3$ 为 $\mathrm{span}(\boldsymbol{q}_1, \boldsymbol{q}_2)$ 平面法向量；这也说明 $\boldsymbol{q}_3$ 平行图 13 中的 $\boldsymbol{\varepsilon}$。

大家很快就会发现，以上这个算例，实际上给格拉姆-施密特正交化、QR 分解、线性回归这三个话题做铺垫！格拉姆-施密特正交化，便是获得本例中 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、$\boldsymbol{q}_3$ 的方法！

怎么获得正交矩阵？

既然正交矩阵 (甚至半正交矩阵) 有这么多“优异”的特性，该如何获得正交矩阵呢？

本书后续介绍 QR 分解、谱分解、奇异值分解都会得到正交矩阵。

注意，此处大家不需要掌握这几个矩阵分解，本书后续会展开讲解这些矩阵分解。

用 QR 分解，任意实数矩阵 $\boldsymbol{A}$ 可以分解为

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R} \tag{56}$$

其中，$\boldsymbol{Q}$ 为正交矩阵。

⚠ 注意，有些场合 $\boldsymbol{Q}$ 为半正交矩阵。

下一节先介绍的格拉姆–施密特正交化本质上就是把 $\boldsymbol{A}$ 的列向量转换成 $\boldsymbol{Q}$ 的列向量的过程。

用谱分解，对称矩阵 $\boldsymbol{A}$ 可以分解为

$$\boldsymbol{A}=\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}} \tag{57}$$

其中，$\boldsymbol{V}$ 为正交矩阵。

用奇异值分解，任意实数矩阵 $\boldsymbol{A}$ 可以分解为

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}} \tag{58}$$

其中，$\boldsymbol{U}$、$\boldsymbol{V}$ 均为正交矩阵。

⚠ 注意，有些场合 $\boldsymbol{U}$、$\boldsymbol{V}$ 也可以是半正交矩阵。

这三个矩阵分解都是本书后续要重点介绍的话题；显然，正交矩阵在其中扮演重要角色。这就是我们要拿出来三节内容介绍正交矩阵的原因。

请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 如下矩阵 $\boldsymbol{A}$ 均为正交矩阵，请分别用矩阵乘法第一、第二视角展开 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{I}$。

- $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
- $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
- $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$
- $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
- $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
- $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
- $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

Q2. 如下矩阵 $\boldsymbol{A}$ 均为半正交矩阵，请分别用矩阵乘法第一、第二视角展开 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$、$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{I}$。

- $\begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}$

▶ $\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

Q3. 给定如下三个列向量，请大家计算 $\boldsymbol{a}_3$ 向 $\text{span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$ 正交投影的坐标。

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},\quad \boldsymbol{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},\quad \boldsymbol{a}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Q4. 请用伪逆求解如下线性方程组。

▶ $\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

▶ $\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

▶ $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$