

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

9.2 3×3 正交矩阵



本节你将掌握的核心技能：

- ▶ 单位矩阵、旋转矩阵、镜像矩阵、置换矩阵，以及它们的复合矩阵都是正交矩阵。
- ▶ 矩阵乘法展第一视角展开正交矩阵的格拉姆矩阵。
- ▶ 投影矩阵合成单位矩阵。
- ▶ 两次投影：先标量投影，再向量投影。
- ▶ 正交补空间包含了垂直于原空间的所有向量。

上一节从单位矩阵、旋转矩阵两个实例入手，用平面几何视角讲解 2×2 正交矩阵。本节“升维”，从三维几何视角讲解 3×3 正交矩阵。

3×3 单位矩阵

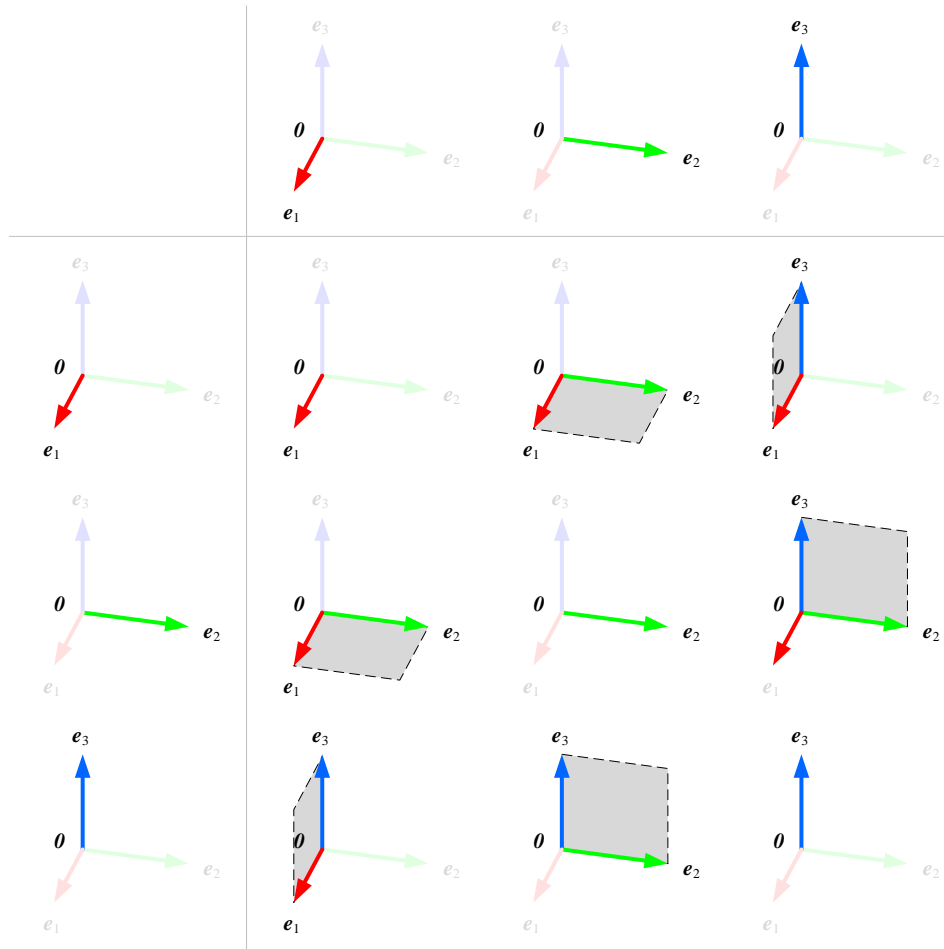
3×3 单位矩阵 I 也是一个正交矩阵，即满足

$$I_{3 \times 3} @ I_{3 \times 3}^T = I_{3 \times 3}^T @ I_{3 \times 3} = I_{3 \times 3} \quad (1)$$

先看格拉姆矩阵 $I_{3 \times 3}^T @ I_{3 \times 3} = I_{3 \times 3}$ ，用矩阵乘法第一视角展开

$$\begin{bmatrix} e_1^T \\ e_2^T \\ e_3^T \end{bmatrix} @ \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} = \begin{bmatrix} e_1^T @ e_1 & e_1^T @ e_2 & e_1^T @ e_3 \\ e_2^T @ e_1 & e_2^T @ e_2 & e_2^T @ e_3 \\ e_3^T @ e_1 & e_3^T @ e_2 & e_3^T @ e_3 \end{bmatrix} = \begin{bmatrix} e_1 \cdot e_1 & e_1 \cdot e_2 & e_1 \cdot e_3 \\ e_2 \cdot e_1 & e_2 \cdot e_2 & e_2 \cdot e_3 \\ e_3 \cdot e_1 & e_3 \cdot e_2 & e_3 \cdot e_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_{3 \times 3} \quad (2)$$

请大家根据图 1 自行分析 (2) 主对角线元素、非主对角线元素的特点。

图 1. 用矩阵乘法第一视角展开矩阵乘法 $I^T I$

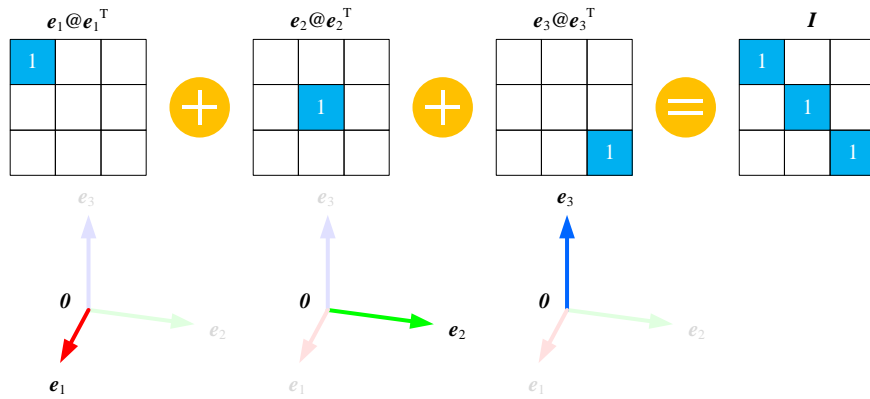
合成单位矩阵

再看 $I_{3 \times 3} @ I_{3 \times 3}^T = I_{3 \times 3}$ ，用矩阵乘法第二视角展开

$$\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix} @ \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \mathbf{e}_3^T \end{bmatrix} = \mathbf{e}_1 @ \mathbf{e}_1^T + \mathbf{e}_2 @ \mathbf{e}_2^T + \mathbf{e}_3 @ \mathbf{e}_3^T = I_{3 \times 3} \quad (3)$$

如图 2 所示，我们看到的是三个正交投影矩阵的叠加，结果是单位矩阵。

比如， $\mathbf{e}_1 @ \mathbf{e}_1^T @ \mathbf{x}$ 代表将向量 \mathbf{x} 投影到 x_1 轴； $\mathbf{e}_2 @ \mathbf{e}_2^T @ \mathbf{x}$ 代表将向量 \mathbf{x} 投影到 x_2 轴，以此类推。下面让我们逐个讲解。

图 2. 三个正交投影矩阵叠加得到单位矩阵 I

朝 e_1 方向正交投影

三维列向量 x 朝方向向量 (单位向量) e_1 方向投影，对应的投影矩阵 P_1 为

$$P_1 = e_1 @ e_1^T = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} @ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (4)$$

投影矩阵 P_1 是秩一矩阵，也是对称矩阵，即满足

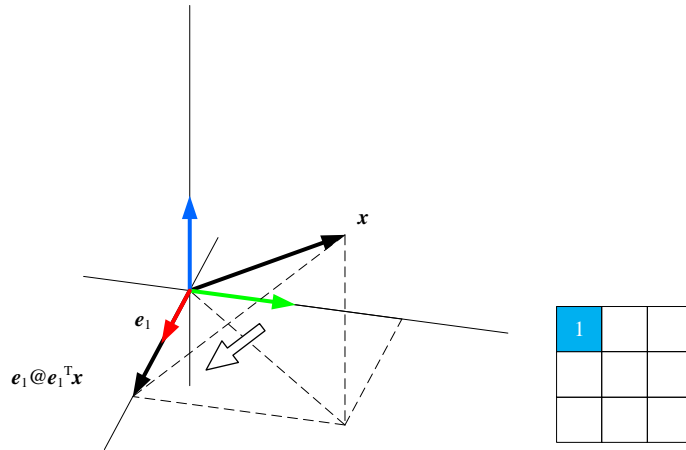
$$(e_1 @ e_1^T)^T = e_1 @ e_1^T \quad (5)$$

如图 3 所示，投影结果对应如下矩阵乘法

$$P_1 x = e_1 @ e_1^T @ x = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{e_1 @ e_1^T} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} \quad (6)$$

列向量 x 在 $\text{span}(e_1)$ 的坐标，就是 x 在 e_1 方向上标量投影结果，即

$$e_1^T @ x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \quad (7)$$

图 3. 三维列向量 x 在 e_1 方向正交投影

比较 (6)、(7)，我们可以把正交投影写成两部分

$$P_1 x = \underbrace{e_1}_{\text{Direction}} @ \underbrace{e_1^T @ x}_{\text{Scalar}} \quad (8)$$

这实际上就是“二次投影”——先把 x 投影到 $\text{span}(e_1)$ ，这部分是标量投影；然后在三维空间中描述投影结果，这一步相当于向量投影。

等式 (8) 右侧转置得到

$$\underbrace{x^T @ e_1}_{\text{Scalar}} @ \underbrace{e_1^T}_{\text{Direction}} \quad (9)$$

上式，相当于用行向量 x^T 代表三维空间中的一点。

朝 e_2 方向正交投影

三维列向量 x 朝 e_2 方向投影，投影矩阵 P_2 为

$$P_2 = e_2 @ e_2^T = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} @ \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (10)$$

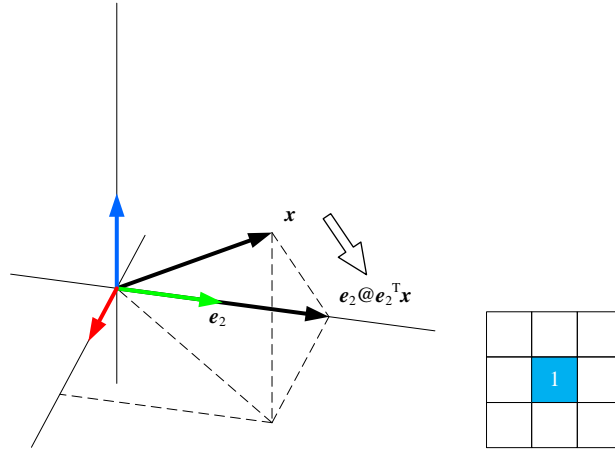
投影矩阵 P_2 同样是秩一矩阵，也是对称矩阵。

如图 4 所示，投影结果对应如下矩阵乘法

$$P_2 x = \underbrace{e_2 @ e_2^T}_{e_2 @ e_2^T} @ x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} \quad (11)$$

列向量 x 在 $\text{span}(e_2)$ 的坐标，就是标量投影结果，即

$$\mathbf{e}_2^T @ \mathbf{x} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \quad (12)$$

图 4. 三维列向量 \mathbf{x} 在 \mathbf{e}_2 方向正交投影

朝 \mathbf{e}_3 方向正交投影

三维列向量 \mathbf{x} 朝 \mathbf{e}_3 方向投影，对应的投影矩阵 \mathbf{P}_3 为

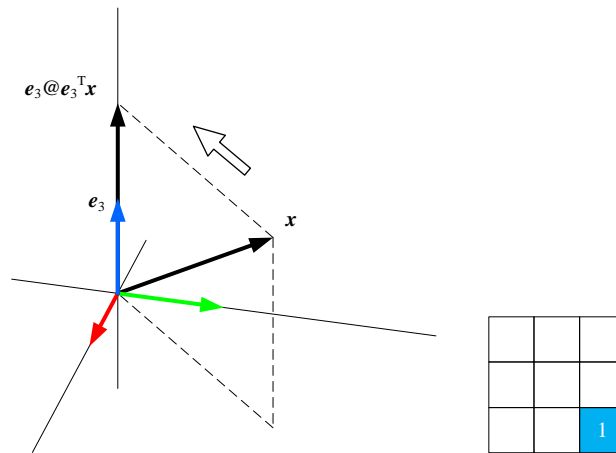
$$\mathbf{P}_3 = \mathbf{e}_3 @ \mathbf{e}_3^T = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} @ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (13)$$

如图 5 所示，投影结果对应如下矩阵乘法

$$\mathbf{P}_3 \mathbf{x} = \mathbf{e}_3 @ \underbrace{\mathbf{e}_3^T}_{\mathbf{e}_3 @ \mathbf{e}_3^T} @ \mathbf{x} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} \quad (14)$$

列向量 \mathbf{x} 在 $\text{span}(\mathbf{e}_3)$ 的坐标，就是标量投影结果，即

$$\mathbf{e}_3^T @ \mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \quad (15)$$

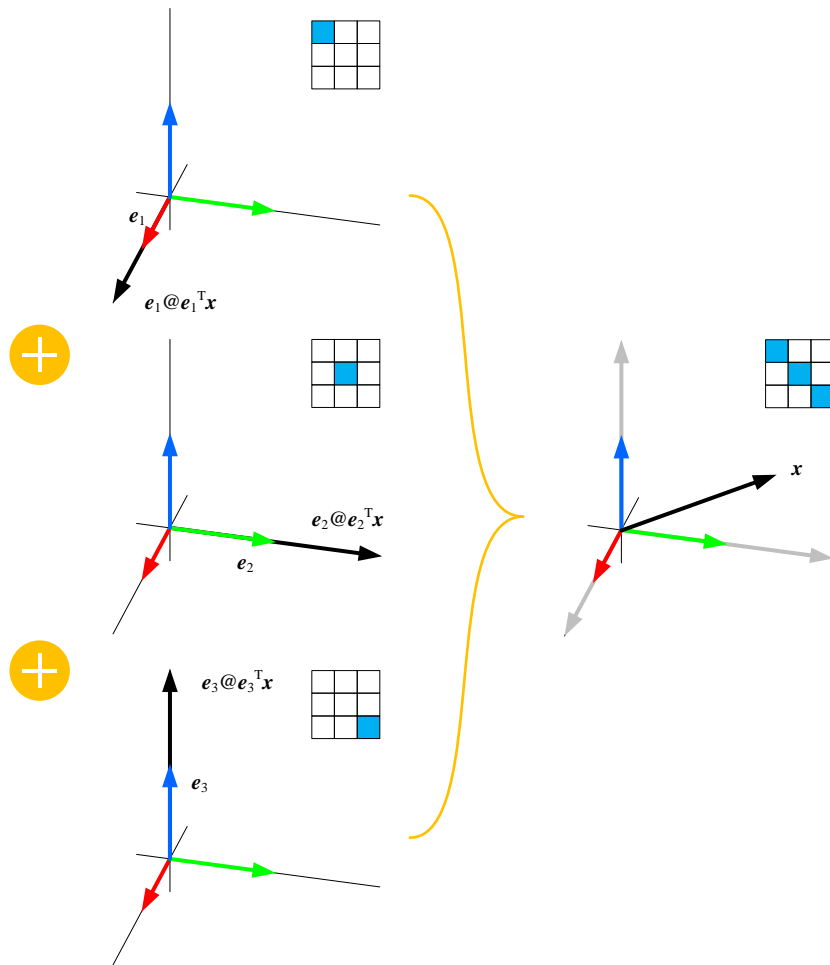
图 5. 三维列向量 x 在 e_3 方向正交投影

向量合成

P_1x 、 P_2x 、 P_3x 均含有 x 的部分信息。但是，如图 6 所示，把三个正交投影分量叠加

$$P_1x + P_2x + P_3x = (P_1 + P_2 + P_3)x = (e_1 @ e_1^T + e_2 @ e_2^T + e_3 @ e_3^T)x = I @ x \quad (16)$$

我们便在 $[e_1, e_2, e_3]$ 中还原 x 所有信息！

图 6. $[e_1, e_2, e_3]$ 中还原三维列向量 x

正交补

(3) 可以整理得到三个有趣的等式

$$\begin{aligned} I - e_1 @ e_1^T &= e_2 @ e_2^T + e_3 @ e_3^T \\ I - e_2 @ e_2^T &= e_1 @ e_1^T + e_3 @ e_3^T \\ I - e_3 @ e_3^T &= e_1 @ e_1^T + e_2 @ e_2^T \end{aligned} \quad (17)$$

这三个矩阵又代表什么呢？大家是否想到了用法向量的正交矩阵。

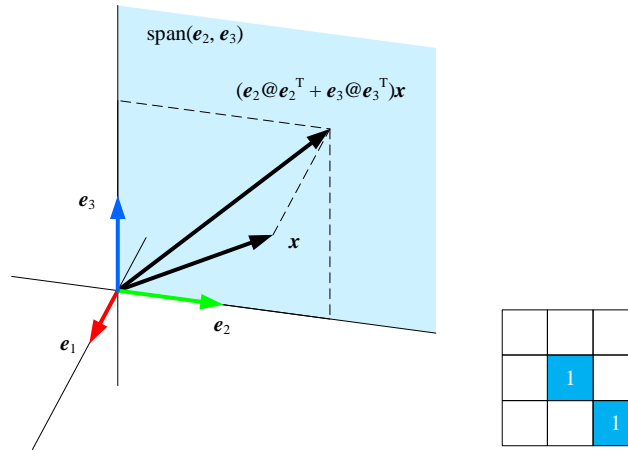
举个例子，如下等式

$$(I - e_1 @ e_1^T)x = (e_2 @ e_2^T + e_3 @ e_3^T)x = (e_2 @ e_2^T)x + (e_3 @ e_3^T)x \quad (18)$$

左侧代表向 x_2x_3 平面正交投影。

x_2x_3 平面的法向量为 e_1 ，向这个平面投影对应的投影矩阵为 $I - e_1 @ e_1^T$ 。具体如图 7 所示。

等式右侧代表分别向 x_2 轴、 x_3 轴投影，再叠加。值得注意的是， $\text{span}(e_1)$ 、 $\text{span}(e_2, e_3)$ 互为正交补。

图 7. 三维列向量朝 x_2x_3 平面投影

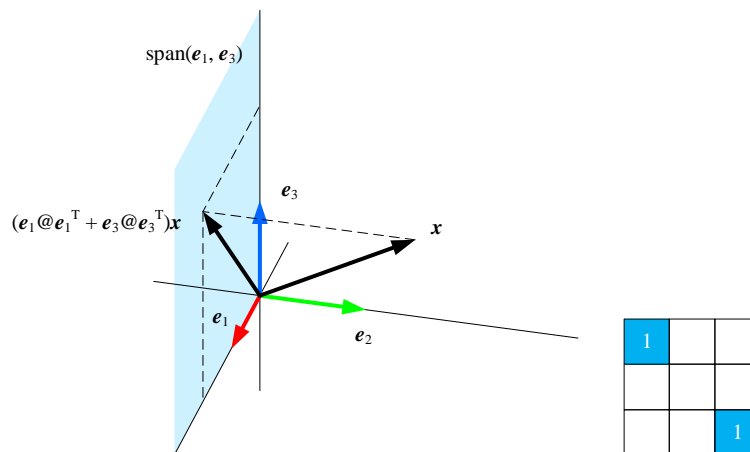
再看一个例子，请大家看下式

$$(I - e_2 @ e_2^T)x = (e_1 @ e_1^T + e_3 @ e_3^T)x = (e_1 @ e_1^T)x + (e_3 @ e_3^T)x \quad (19)$$

左侧代表向 x_1x_3 平面正交投影， e_2 是 x_1x_3 平面的法向量。图 8 所示为三维列向量朝 x_1x_3 平面投影。

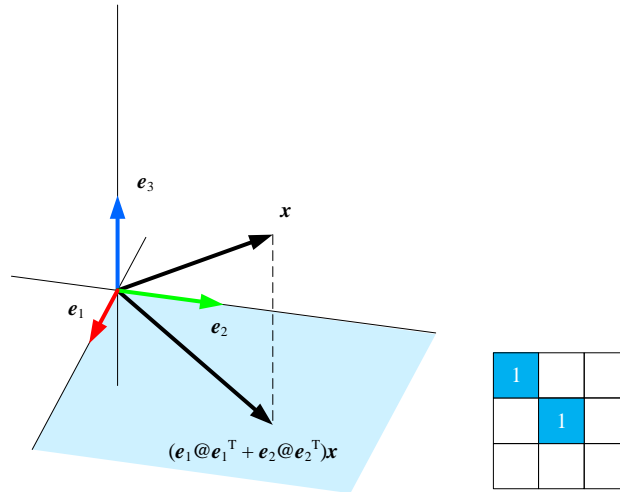
等式右侧代表分别向 x_1 轴、 x_3 轴投影，再叠加。

也就是说， $\text{span}(e_2)$ 、 $\text{span}(e_1, e_3)$ 互为正交补。向两个空间的正交投影得到的都是部分信息；两者拼凑得到全部信息。

图 8. 三维列向量朝 x_1x_3 平面投影

请大家自行分析如下等式和图 9

$$(I - e_3 @ e_3^T)x = (e_1 @ e_1^T + e_2 @ e_2^T)x = (e_1 @ e_1^T)x + (e_2 @ e_2^T)x \quad (20)$$

图 9. 三维列向量朝 x_1x_2 平面投影

3 × 3 旋转矩阵

给定如下三维 (绕原点) 旋转矩阵 V

$$V = \begin{bmatrix} 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4/5 \\ 3/5 \\ 0 \end{bmatrix}_{v_1} \underbrace{\begin{bmatrix} -3/5 \\ 4/5 \\ 0 \end{bmatrix}}_{v_2} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{v_3} \quad (21)$$

矩阵 V 也是正交矩阵，满足

$$VV^T = V^TV = I \quad (22)$$

先看格拉姆矩阵 V^TV ，把 V 写成 $[v_1, v_2, v_3]$ ，用矩阵乘法第一视角展开

$$\begin{bmatrix} v_1^T \\ v_2^T \\ v_3^T \end{bmatrix} @ [v_1 \ v_2 \ v_3] = \begin{bmatrix} v_1^T @ v_1 & v_1^T @ v_2 & v_1^T @ v_3 \\ v_2^T @ v_1 & v_2^T @ v_2 & v_2^T @ v_3 \\ v_3^T @ v_1 & v_3^T @ v_2 & v_3^T @ v_3 \end{bmatrix} = I_{3 \times 3} \quad (23)$$

V^TV 的主对角线元素为 V 的列向量和自身的内积， L^2 范数平方； V^TV 非主对角线元素为 V 的列向量的成对内积，均为 0，这是因为 V 的列向量两两正交。

合成单位矩阵

再看 VV^T ，用矩阵乘法第二视角展开

$$[v_1 \ v_2 \ v_3] @ \begin{bmatrix} v_1^T \\ v_2^T \\ v_3^T \end{bmatrix} = v_1 @ v_1^T + v_2 @ v_2^T + v_3 @ v_3^T = I_{3 \times 3} \quad (24)$$

我们看到的是三个正交投影矩阵的叠加，结果是单位矩阵。

朝 v_1 方向正交投影

三维列向量 x 朝方向向量 (单位向量) v_1 方向投影，对应的投影矩阵 P_1 为

$$P_1 = v_1 @ v_1^T = \begin{bmatrix} 4/5 \\ 3/5 \\ 0 \end{bmatrix} @ \begin{bmatrix} 4/5 \\ 3/5 \\ 0 \end{bmatrix}^T = \begin{bmatrix} 16/25 & 12/25 & 0 \\ 12/25 & 9/25 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (25)$$

上述投影矩阵是秩一矩阵，也是对称矩阵。

投影结果对应如下矩阵乘法

$$P_1 x = v_1 @ v_1^T x = \underbrace{\begin{bmatrix} 16/25 & 12/25 & 0 \\ 12/25 & 9/25 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{v_1 @ v_1^T} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 16x_1/25 + 12x_2/25 \\ 12x_1/25 + 9x_2/25 \\ 0 \end{bmatrix} \quad (26)$$

列向量 x 在 $\text{span}(v_1)$ 的坐标，即标量投影为

$$v_1^T @ x = \begin{bmatrix} 4/5 \\ 3/5 \\ 0 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4x_1/5 + 3x_2/5 \quad (27)$$

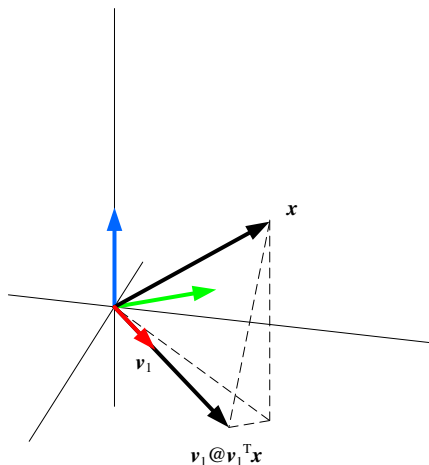


图 10. 三维列向量 x 在 v_1 方向正交投影

朝 v_2 方向正交投影

三维列向量 x 朝方向向量 (单位向量) v_2 方向投影，对应的投影矩阵 P_2 为

$$P_2 = v_2 @ v_2^T = \begin{bmatrix} -3/5 \\ 4/5 \\ 0 \end{bmatrix} @ \begin{bmatrix} -3/5 \\ 4/5 \\ 0 \end{bmatrix}^T = \begin{bmatrix} 9/25 & -12/25 & 0 \\ -12/25 & 16/25 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (28)$$

投影结果对应如下矩阵乘法

$$\mathbf{P}_2 \mathbf{x} = \mathbf{v}_2 @ \mathbf{v}_2^T \mathbf{x} = \underbrace{\begin{bmatrix} 9/25 & -12/25 & 0 \\ -12/25 & 16/25 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{v}_2 @ \mathbf{v}_2^T} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9x_1/25 - 12x_2/25 \\ -12x_1/25 + 16x_2/25 \\ 0 \end{bmatrix} \quad (29)$$

列向量 \mathbf{x} 在 $\text{span}(\mathbf{v}_2)$ 的坐标，就是标量投影结果，即

$$\mathbf{v}_2^T @ \mathbf{x} = \begin{bmatrix} -3/5 \\ 4/5 \\ 0 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -3x_1/5 + 4x_2/5 \quad (30)$$

朝 \mathbf{v}_3 方向正交投影

三维列向量 \mathbf{x} 朝方向向量 (单位向量) \mathbf{v}_3 方向投影，对应的投影矩阵 \mathbf{P}_3 为

$$\mathbf{P}_3 = \mathbf{v}_3 @ \mathbf{v}_3^T = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} @ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (31)$$

投影结果对应如下矩阵乘法

$$\mathbf{P}_3 \mathbf{x} = \mathbf{v}_3 @ \mathbf{v}_3^T \mathbf{x} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{v}_3 @ \mathbf{v}_3^T} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} \quad (32)$$

列向量 \mathbf{x} 在 $\text{span}(\mathbf{v}_3)$ 的坐标，就是标量投影结果，即

$$\mathbf{v}_3^T @ \mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \quad (33)$$

向量合成

把 (26)、(29)、(32) 叠加还原 \mathbf{x} 所有信息

$$(\mathbf{v}_1 @ \mathbf{v}_1^T + \mathbf{v}_2 @ \mathbf{v}_2^T + \mathbf{v}_3 @ \mathbf{v}_3^T) \mathbf{x} = \begin{bmatrix} 16x_1/25 + 12x_2/25 \\ 12x_1/25 + 9x_2/25 \\ 0 \end{bmatrix} + \begin{bmatrix} 9x_1/25 - 12x_2/25 \\ -12x_1/25 + 16x_2/25 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (34)$$

正交补

(3) 也可以整理得到三个有趣的等式

$$\begin{aligned} \mathbf{I} - \mathbf{v}_1 @ \mathbf{v}_1^T &= \mathbf{v}_2 @ \mathbf{v}_2^T + \mathbf{v}_3 @ \mathbf{v}_3^T \\ \mathbf{I} - \mathbf{v}_2 @ \mathbf{v}_2^T &= \mathbf{v}_1 @ \mathbf{v}_1^T + \mathbf{v}_3 @ \mathbf{v}_3^T \\ \mathbf{I} - \mathbf{v}_3 @ \mathbf{v}_3^T &= \mathbf{v}_1 @ \mathbf{v}_1^T + \mathbf{v}_2 @ \mathbf{v}_2^T \end{aligned} \quad (35)$$

先看第一个等式,

$$(I - v_1 @ v_1^T)x = (v_2 @ v_2^T)x + (v_3 @ v_3^T)x \quad (36)$$

左侧代表向 $\text{span}(v_2, v_3)$ 平面正交投影, 这个平面的法向量为 v_1 。

等式右侧代表分别向 $\text{span}(v_2)$ 、 $\text{span}(v_3)$ 轴投影, 再叠加。 $\text{span}(v_1)$ 、 $\text{span}(v_2, v_3)$ 互为正交补。

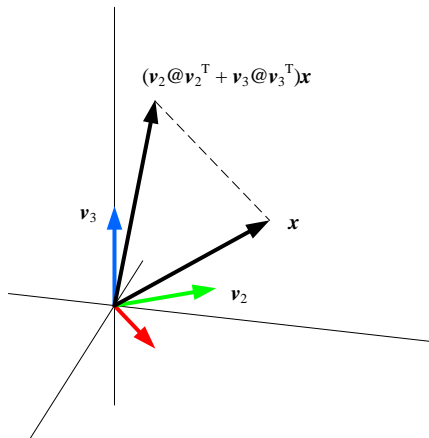


图 11. 三维列向量 x 朝 $\text{span}(v_2, v_3)$ 投影

再看第二个等式

$$(I - v_2 @ v_2^T)x = (v_1 @ v_1^T)x + (v_3 @ v_3^T)x \quad (37)$$

等式 (37) 左侧代表向 $\text{span}(v_1, v_3)$ 正交投影, 这个平面的法向量为 v_2 。

等式 (37) 右侧代表分别向 $\text{span}(v_1)$ 、 $\text{span}(v_3)$ 轴投影, 再叠加。 $\text{span}(v_2)$ 、 $\text{span}(v_1, v_3)$ 互为正交补。

最后看第三个等式

$$(I - v_3 @ v_3^T)x = (v_1 @ v_1^T)x + (v_2 @ v_2^T)x \quad (38)$$

请大家自定分析等式 (38) 和图 12。

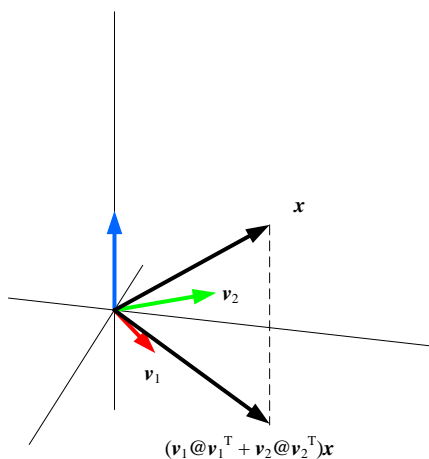


图 12. 三维列向量 x 朝 $\text{span}(v_1, v_2)$ 投影

Q1. 请判断哪些矩阵为正交矩阵。

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

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

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

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

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

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

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

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

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

▶ $\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Q2. 几何角度来看，**Q1** 中的正交矩阵对应怎样的几何变换？