

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

9.4 格拉姆-施密特正交化



本节你将掌握的核心技能：

- ▶ 正交化：将一组线性无关向量转化为两两正交的单位向量
- ▶ 格拉姆-施密特正交化：通过逐步投影与单位化实现向量正交化。
- ▶ 格拉姆-施密特正交化的每一步都是将一个向量投影到已有正交基的正交补中。
- ▶ 每个向量都可以正交分解为“已知方向 + 独特方向”。
- ▶ $A = QR$ ：最终获得的 Q 为正交矩阵， R 为上三角坐标矩阵。

本章前三节，我们看到了正交矩阵的各种“优质”特性，本节开始介绍如何将线性无关向量正交化、并获得正交矩阵的一种方法。我们用到的算法叫做格拉姆-施密特正交化，这是获得正交矩阵的几种常用方法之一。

几何视角

如图 1 所示，从几何角度来看，**格拉姆-施密特正交化** (Gram-Schmidt orthogonalization) 将一组线性无关的向量 (矩阵 A 列向量) 转换成两两正交的单位向量 (矩阵 Q 列向量)。

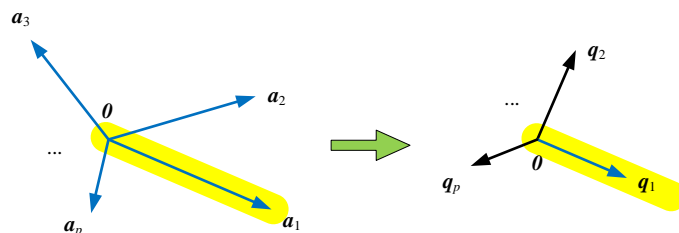


图 1. 格拉姆-施密特正交化背后的几何意义

⚠ 注意，图 1 特别用黄色高亮强调了列向量 a_1 和 q_1 平行，这是格拉姆-施密特正交化重要特征之一。

如图 2 所示，从矩阵转换角度来看，通过格拉姆-施密特正交化，一个列满秩矩阵 A 转化为半正交矩阵 Q 。

从基底角度来看，格拉姆-施密特正交化把非正交基 $[a_1, a_2, \dots, a_p]$ 转化成规范正交基 $[q_1, q_2, \dots, q_p]$ 。

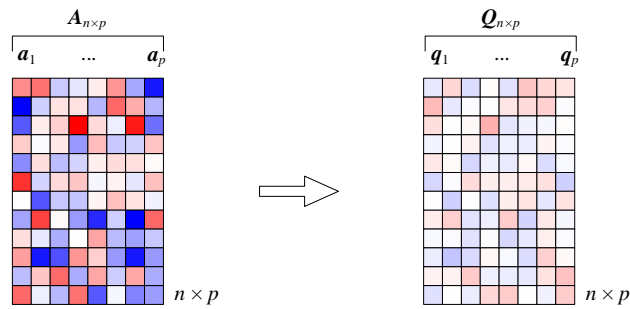


图 2. 列满秩矩阵 A 转化为半正交矩阵 Q

正交化过程

利用向量投影，格拉姆-施密特正交化过程为

$$\begin{aligned}
 \eta_1 &= a_1 \\
 \eta_2 &= a_2 - \text{proj}_{\eta_1}(a_2) \\
 \eta_3 &= a_3 - \text{proj}_{\eta_1}(a_3) - \text{proj}_{\eta_2}(a_3) \\
 &\vdots \\
 \eta_p &= a_p - \sum_{j=1}^{p-1} \text{proj}_{\eta_j}(a_p)
 \end{aligned} \tag{1}$$

最后单位化，获得单位正交向量

$$q_1 = \frac{\eta_1}{\|\eta_1\|}, \quad q_2 = \frac{\eta_2}{\|\eta_2\|}, \quad q_3 = \frac{\eta_3}{\|\eta_3\|}, \quad \dots, \quad q_p = \frac{\eta_p}{\|\eta_p\|} \tag{2}$$

结果中， q_i 为单位向量，且 q_i 、 q_j 两两正交。

值得再次强调的是，规范正交基 $[q_1, q_2, q_3, \dots, q_p]$ 的特别之处在于 q_1 平行 a_1 。本书后续还会介绍其他获得规范正交基的算法，请大家注意比对。

⚠ 注意，(1) 这个格拉姆-施密特正交化过程大家千万不要死记硬背。我们马上会用一个例子让大家看到这个正交化过程的几何视角，这会极大地帮助大家理解上述过程。

正交投影

根据本书前文介绍的正交投影，(1) 中第二个等式用到的正交投影可以写成

$$\text{proj}_{\eta_1}(a_2) = \frac{\eta_1 @ \eta_1^T}{\eta_1^T @ \eta_1} @ a_2 = q_1 @ q_1^T @ a_2 \quad (3)$$

得到 η_1 后，先单位化获得 q_1 ，这样计算 η_2 时，可以用

$$\eta_2 = a_2 - q_1 @ q_1^T @ a_2 = (I - q_1 @ q_1^T) @ a_2 \quad (4)$$

上式的几何含义是，把向量 a_2 投影到 $\text{span}(q_1)$ 正交补空间中；得到的投影结果 η_2 必定垂直 q_1 。

得到 η_2 后，立刻将其单位化获得 q_2 ；计算 η_3 也用正交投影

$$\eta_3 = a_3 - q_1 @ q_1^T @ a_3 - q_2 @ q_2^T @ a_3 = (I - q_1 @ q_1^T - q_2 @ q_2^T) @ a_3 \quad (5)$$

几何角度来看，上式相当于把向量 a_3 投影到 $\text{span}(q_1, q_2)$ 正交补空间中；得到的投影结果 η_3 分别垂直 q_1 、 q_2 。

依次类推，用同样的办法我们可以将 (列满秩的，即线性无关的) $[a_1, a_2, a_3, \dots, a_p]$ 正交化得到 $[q_1, q_2, q_3, \dots, q_p]$ 。

举个实例

下面，让我们一步步用格拉姆-施密特正交化如下方阵列向量

$$A = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 0 & 2 \\ 0 & 2 & 2 \end{bmatrix} \quad (6)$$

A 的列向量

$$a_1 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \quad (7)$$

如图 3 所示， a_1 、 a_2 、 a_3 都不是单位向量，两两非正交；但是 a_1 、 a_2 、 a_3 线性无关，所以它们可以构成非正交基，也同样张成三维空间。

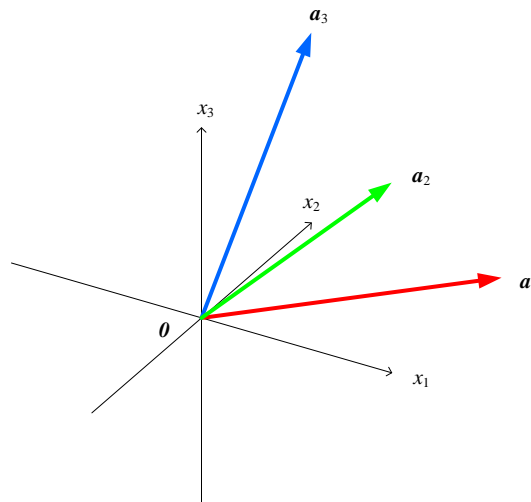


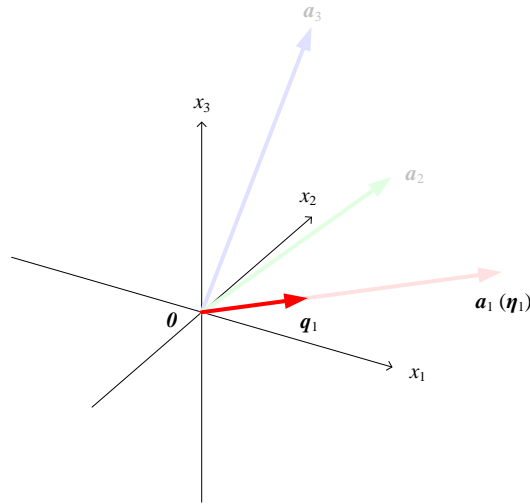
图 3. a_1 、 a_2 、 a_3 具体位置**第一步：** $a_1 \rightarrow q_1$

第一步最简单，获得 η_1 ，只需要 $\eta_1 = a_1$ 。

然后，对非零向量 η_1 单位化，便得到 q_1

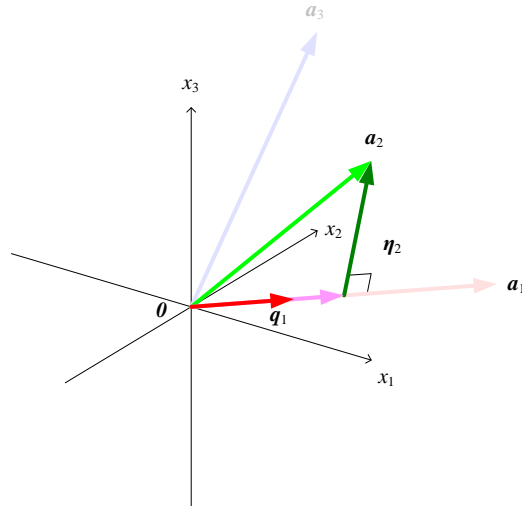
$$q_1 = \frac{\eta_1}{\|\eta_1\|} = \frac{a_1}{\|a_1\|} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} \quad (8)$$

具体如图 4 所示；很明显， q_1 、 a_1 (η_1) 同向。

图 4. a_1 (η_1) 单位化**第二步：** $a_2 \rightarrow q_2$

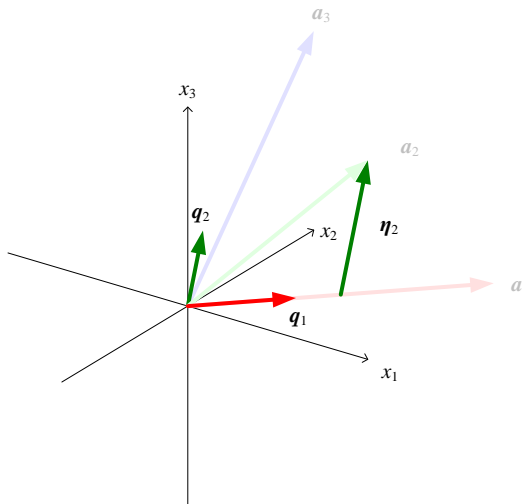
获得 q_1 后，用正交投影计算 η_2

$$\begin{aligned} \eta_2 &= (I - q_1 @ q_1^T) @ a_2 \\ &= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} @ \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 & 0 \end{bmatrix}^T \right) @ \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} @ \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \end{aligned} \quad (9)$$

图 5. 正交投影计算 η_2

然后，如图 6 所示， η_2 单位化得到 q_2

$$q_2 = \frac{\eta_2}{\|\eta_2\|} = \frac{\sqrt{6}}{6} \times \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} \quad (10)$$

图 6. η_2 单位化

q_2 算完了，下面做一些“额外”的分析。

如图 7 所示，这一步相当于把 a_2 正交分解成两部分 $q_1 @ q_1^T @ a_2$ 、 η_2 ，对应

$$a_2 = q_1 @ q_1^T @ a_2 + \eta_2 \quad (11)$$

也就是说， a_2 正交分解为平行于 q_1 (a_1, η_1)、垂直于 q_1 两个分量。

而刨除 a_2 中和 a_1 同向分量，剩余垂直于 a_1 的分量 (η_2) 正是 a_2 贡献的“独特”方向！

有了 q_2 之后， η_2 可以写成

$$\eta_2 = q_2 @ q_2^T @ a_2 \quad (12)$$

如图 7 所示， a_2 向量正交分解就可以写成

$$\begin{aligned} a_2 &= q_1 @ q_1^T @ a_2 + q_2 @ q_2^T @ a_2 \\ &= (q_1 @ q_1^T + q_2 @ q_2^T) @ a_2 \end{aligned} \quad (13)$$

⚠ 注意， $q_1 @ q_1^T + q_2 @ q_2^T$ 不满秩，所以方阵不可逆！

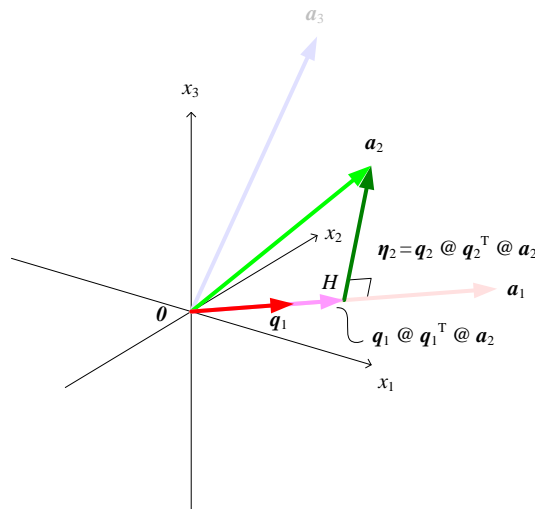


图 7. a_2 正交分解

计算 $q_1 @ q_1^T @ a_2$,

$$\begin{aligned} q_1 @ q_1^T @ a_2 &= \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} @ \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix}^T @ \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} @ \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \end{aligned} \quad (14)$$

算式 $q_1 @ q_1^T @ a_2$ 中， $q_1^T @ a_2$ 是标量投影，即向量 a_2 在方向向量 (单位向量) q_1 上的坐标；再乘上 q_1 ，便得到向量投影。这就是前文提到的“二次投影”！

$q_1 @ q_1^T @ a_2$ 就是图 7 中点 H 的坐标值！

显然， $q_1 @ q_1^T @ a_2$ 平行 a_1

$$q_1 @ q_1^T @ a_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{a_1}{2} = \frac{1}{2} \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} \quad (15)$$

验证 η_2 、 $q_1 @ q_1^T @ a_2$ 正交

$$\eta_2 \cdot (q_1 @ q_1^T @ a_2) = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 0 \quad (16)$$

再验证 η_2 、 $q_1 @ q_1^T @ a_2$ 合成得到 a_2 ，即

$$\eta_2 + q_1 @ q_1^T @ a_2 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = a_2 \quad (17)$$

第三步： $a_3 \rightarrow q_3$

获得 q_1 、 q_2 后，用正交投影计算 η_3

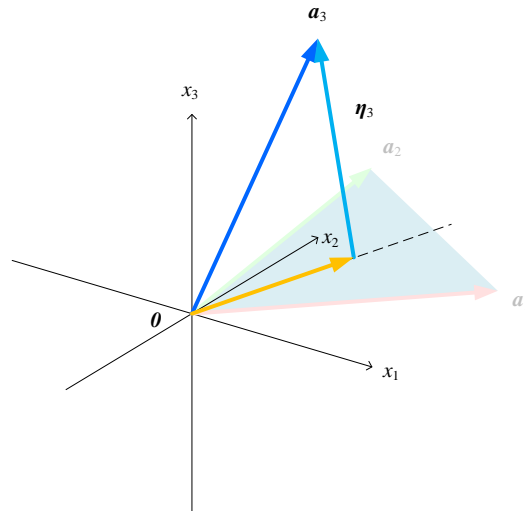
$$\begin{aligned} \eta_3 &= (I - q_1 @ q_1^T - q_2 @ q_2^T) @ a_3 \\ &= \begin{pmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} @ \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix}^T - \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} @ \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}^T \end{pmatrix} @ \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{pmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1/6 & -1/6 & 1/3 \\ -1/6 & 1/6 & -1/3 \\ 1/3 & -1/3 & 2/3 \end{bmatrix} \end{pmatrix} @ \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1/3 & -1/3 & -1/3 \\ -1/3 & 1/3 & 1/3 \\ -1/3 & 1/3 & 1/3 \end{bmatrix} @ \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -4/3 \\ 4/3 \\ 4/3 \end{bmatrix} \end{aligned} \quad (18)$$

从几何角度来看，第三步是 a_3 向 $[\eta_1, \eta_2]$ 张成的平面投影。而 η_3 为 a_3 中不在 $[\eta_1, \eta_2]$ 平面上的向量。

⚠ 注意， $\text{span}(\eta_1, \eta_2)$ 等价于 $\text{span}(q_1, q_2)$ 。

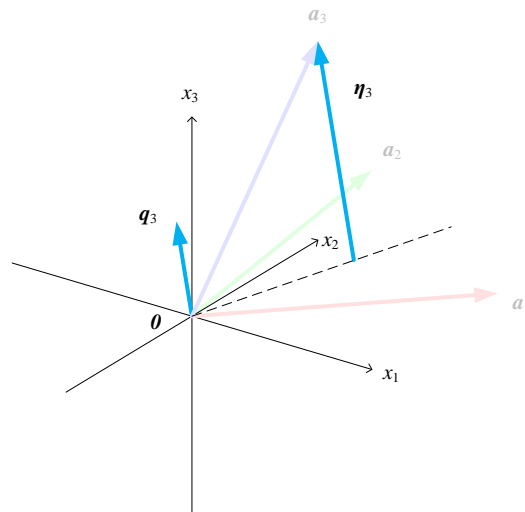
显然， η_3 垂直 $\text{span}(\eta_1, \eta_2)$ ，也就是说 η_3 分别垂直 η_1 和 η_2 。这也就是说， $\text{span}(\eta_3)$ 和 $\text{span}(\eta_1, \eta_2)$ 互为正交补。

如图 8 所示，剔除 a_3 中和 a_1 、 a_2 平行的分量，余下的 η_3 是 a_3 的“独特”贡献。

图 8. 正交投影计算 η_3

然后，如图 9 所示， η_3 单位化得到 q_3

$$q_3 = \frac{\eta_3}{\|\eta_3\|} = \begin{bmatrix} -\sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \end{bmatrix} \quad (19)$$

图 9. η_3 单位化

LA_09_04_01.ipynb 上述运算，请大家对照本节内容逐行分析代码。

至此，如图 10 所示，我们已经完成 $[a_1, a_2, a_3]$ 正交化，得到了 $[q_1, q_2, q_3]$ 。

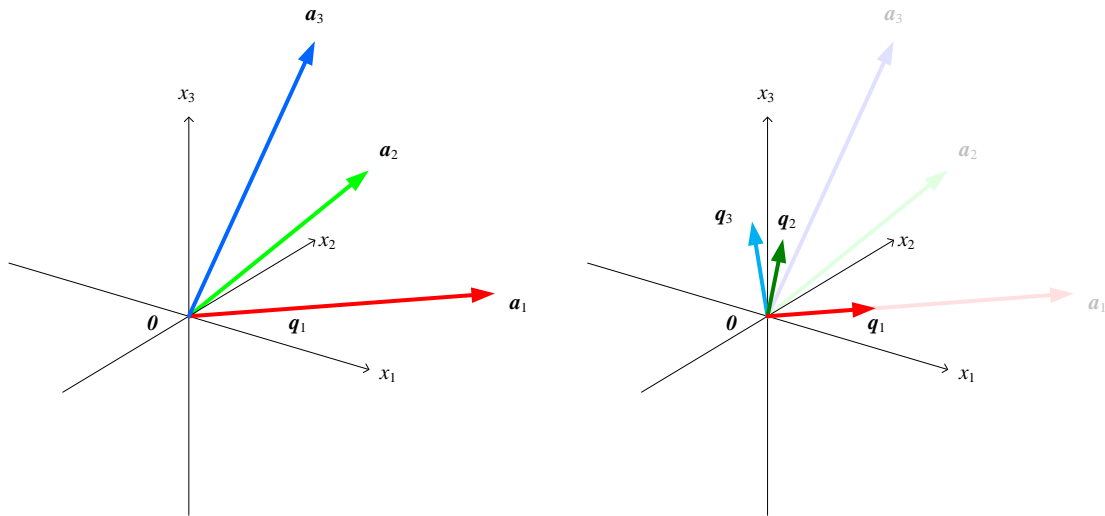


图 10. 将 $[a_1, a_2, a_3]$ 正交化得到 $[q_1, q_2, q_3]$

但是，我们的分析还没有结束！

如图 11 所示，这一步相当于把 a_3 正交分解成两部分 $(q_1 @ q_1^T + q_2 @ q_2^T) @ a_3$ 、 η_3 ，对应

$$a_3 = (q_1 @ q_1^T + q_2 @ q_2^T) @ a_3 + \eta_3 \quad (20)$$

也就是说， a_3 正交分解为 $\text{span}(q_1, q_2)$ 平面上、垂直于 $\text{span}(q_1, q_2)$ 两个分量。

而刨除 a_3 中 $\text{span}(q_1, q_2)$ 平面上分量，剩余垂直于 $\text{span}(q_1, q_2)$ 的分量 (η_3) 便是 a_3 贡献的“独特”方向！

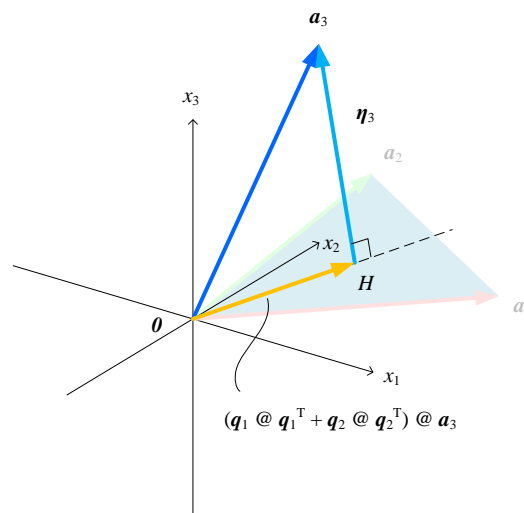


图 11. a_3 正交分解

如图 12 所示， $(q_1 @ q_1^T + q_2 @ q_2^T) @ a_3$ 还可以进一步正交分解成 $(q_1 @ q_1^T) @ a_3$ 、 $(q_2 @ q_2^T) @ a_3$ 。

计算 $q_1 @ q_1^T @ a_3$,

$$\begin{aligned} q_1 @ q_1^T @ a_3 &= \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} @ \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix}^T @ \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} @ \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \end{aligned} \quad (21)$$

如图 12 所示，几何角度，算式 $q_1 @ q_1^T @ a_3$ 代表 a_3 向方向向量 q_1 的正交投影。

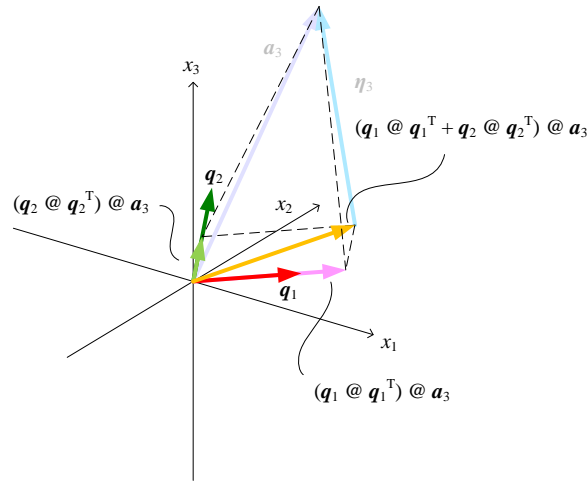


图 12. $(q_1 @ q_1^T + q_2 @ q_2^T) @ a_3$ 正交分解

计算 $q_2 @ q_2^T @ a_3$,

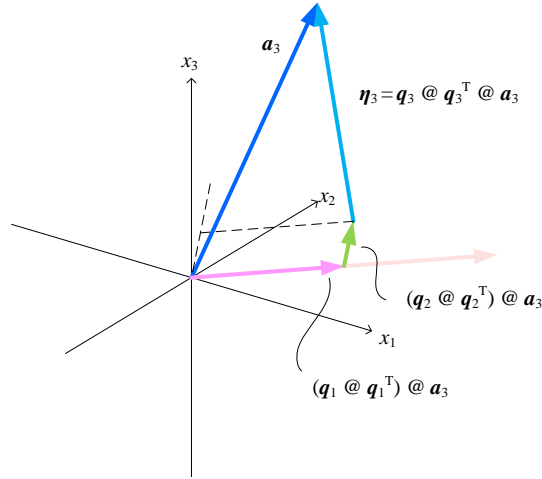
$$\begin{aligned} q_2 @ q_2^T @ a_3 &= \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} @ \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}^T @ \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1/6 & -1/6 & 1/3 \\ -1/6 & 1/6 & -1/3 \\ 1/3 & -1/3 & 2/3 \end{bmatrix} @ \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -1/3 \\ 2/3 \end{bmatrix} \end{aligned} \quad (22)$$

如图 12 所示，几何角度，算式 $q_2 @ q_2^T @ a_3$ 代表 a_3 向方向向量 q_2 的正交投影。

如图 13 所示， η_3 、 $q_1 @ q_1^T @ a_3$ 、 $q_2 @ q_2^T @ a_3$ 合成得到 a_3 ，即

$$\eta_3 + q_1 @ q_1^T @ a_3 + q_2 @ q_2^T @ a_3 = \begin{bmatrix} -4/3 \\ 4/3 \\ 4/3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = a_3 \quad (23)$$

请大家自行验证 η_3 、 $q_1 @ q_1^T @ a_3$ 、 $q_2 @ q_2^T @ a_3$ 两两正交。

图 13. 三个向量首尾相接还原 a_3

有了 q_3 之后，而 η_3 实际上可以写成

$$\eta_3 = q_3 @ q_3^T @ a_3 \quad (24)$$

这样，(23) 就是

$$a_3 = (q_1 @ q_1^T + q_2 @ q_2^T + q_3 @ q_3^T) @ a_3 \quad (25)$$

其中包含着我们熟悉的

$$q_1 @ q_1^T + q_2 @ q_2^T + q_3 @ q_3^T = I \quad (26)$$

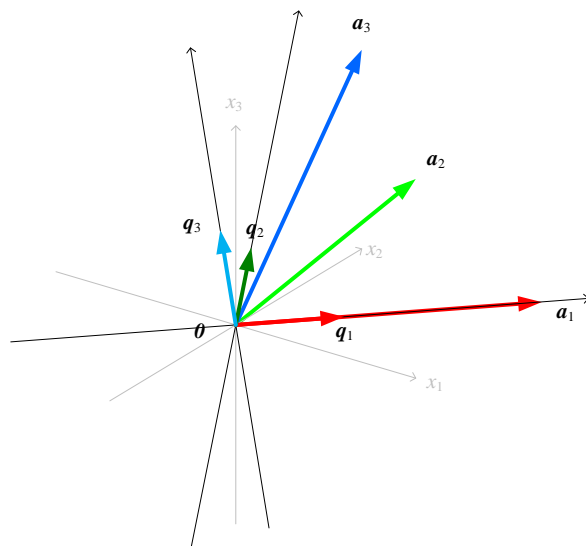
$[q_1, q_2, q_3]$ 坐标

在 $[e_1, e_2, e_3]$ 坐标系中，我们知道 a_1 、 a_2 、 a_3 的坐标，即

$$a_1 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \quad (27)$$

而 $[q_1, q_2, q_3]$ 是规范正交基，撑起的平行六面体是旋转单位正方体。

如图 14 所示， a_1 、 a_2 、 a_3 在 $[q_1, q_2, q_3]$ 当然也有自己的坐标！

图 14. a_1 、 a_2 、 a_3 在规范正交基 $[q_1, q_2, q_3]$

下面让我们逐一计算这些坐标值。

a_1 在 $[q_1, q_2, q_3]$ 坐标

设 a_1 在 $[q_1, q_2, q_3]$ 坐标为 $(r_{1,1}, r_{2,1}, r_{3,1})$ ，也就是说 a_1 可以写成如下线性组合

$$a_1 = r_{1,1}q_1 + r_{2,1}q_2 + r_{3,1}q_3 \quad (28)$$

写成矩阵乘法

$$a_1 = [q_1 \quad q_2 \quad q_3] \begin{bmatrix} r_{1,1} \\ r_{2,1} \\ r_{3,1} \end{bmatrix} \quad (29)$$

根据标量乘法，向量 a_1 在方向向量（单位向量） q_1 的坐标就是两者的内积，即

$$r_{1,1} = a_1 \cdot q_1 = q_1 \cdot a_1 = \langle a_1, q_1 \rangle = a_1^T @ q_1 = q_1^T @ a_1 \quad (30)$$

带入具体数值

$$r_{1,1} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} = 2\sqrt{2} \quad (31)$$

这是 a_1 在 q_1 轴上的坐标（标量投影）。

类似地，我们可以得到

$$\begin{aligned} r_{2,1} &= a_1 \cdot q_2 = q_2 \cdot a_1 = \langle a_1, q_2 \rangle = a_1^T @ q_2 = q_2^T @ a_1 \\ r_{3,1} &= a_1 \cdot q_3 = q_3 \cdot a_1 = \langle a_1, q_3 \rangle = a_1^T @ q_3 = q_3^T @ a_1 \end{aligned} \quad (32)$$

由于 a_1 平行 q_1 ，而 q_1 、 q_2 、 q_3 两两正交；显然， a_1 在 q_2 、 q_3 标量投影为 0。

这意味着 $r_{2,1} = 0$, $r_{3,1} = 0$, (29) 可以写成

$$\mathbf{a}_1 = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] \begin{bmatrix} r_{1,1} \\ 0 \\ 0 \end{bmatrix} \quad (33)$$

请大家用向量内积，带入具体值计算 $r_{2,1}$ 、 $r_{3,1}$ ，看是否均为 0。

图 15 所示为 \mathbf{a}_1 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 中的坐标。

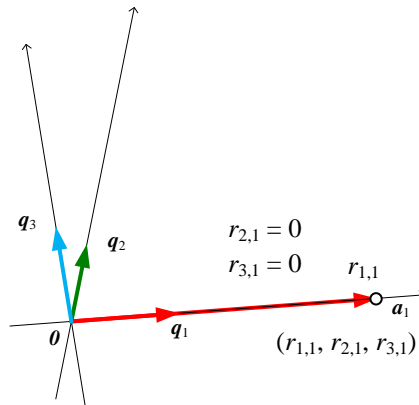


图 15. \mathbf{a}_1 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 中的坐标

\mathbf{a}_2 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 坐标

设 \mathbf{a}_2 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 坐标为 $(r_{1,2}, r_{2,2}, r_{3,2})$ ，也就是说 \mathbf{a}_2 可以写成如下线性组合

$$\mathbf{a}_2 = r_{1,2}\mathbf{q}_1 + r_{2,2}\mathbf{q}_2 + r_{3,2}\mathbf{q}_3 \quad (34)$$

写成矩阵乘法

$$\mathbf{a}_2 = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] \begin{bmatrix} r_{1,2} \\ r_{2,2} \\ r_{3,2} \end{bmatrix} \quad (35)$$

同样根据标量乘法，向量 \mathbf{a}_2 在方向向量 \mathbf{q}_1 的坐标就是两者的内积，即

$$r_{1,2} = \mathbf{a}_2 \cdot \mathbf{q}_1 = \mathbf{q}_1 \cdot \mathbf{a}_2 = \langle \mathbf{a}_2, \mathbf{q}_1 \rangle = \mathbf{a}_2^T @ \mathbf{q}_1 = \mathbf{q}_1^T @ \mathbf{a}_2 \quad (36)$$

带入具体数值

$$r_{1,2} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} = \sqrt{2} \quad (37)$$

这是 \mathbf{a}_2 在 \mathbf{q}_1 轴上的坐标 (标量投影)。

类似地，我们可以得到

$$r_{2,2} = \mathbf{a}_2 \cdot \mathbf{q}_2 = \mathbf{q}_2 \cdot \mathbf{a}_2 = \langle \mathbf{a}_2, \mathbf{q}_2 \rangle = \mathbf{a}_2^T @ \mathbf{q}_2 = \mathbf{q}_2^T @ \mathbf{a}_2 \quad (38)$$

带入具体数值

$$r_{2,2} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} = \sqrt{6} \quad (39)$$

这是 \mathbf{a}_2 在 \mathbf{q}_2 轴上的坐标 (标量投影)。

由于 \mathbf{a}_2 在 $\text{span}(\mathbf{q}_1, \mathbf{q}_2)$ 中, \mathbf{a}_2 在 \mathbf{q}_3 标量投影为 0, 这意味着 $r_{3,2} = 0$ 。(35) 可以写成

$$\mathbf{a}_2 = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] \begin{bmatrix} r_{1,2} \\ r_{2,2} \\ 0 \end{bmatrix} \quad (40)$$

请大家用向量内积, 带入具体值计算 $r_{3,2}$, 看是否为 0。

图 16 所示为 \mathbf{a}_2 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 中的坐标。

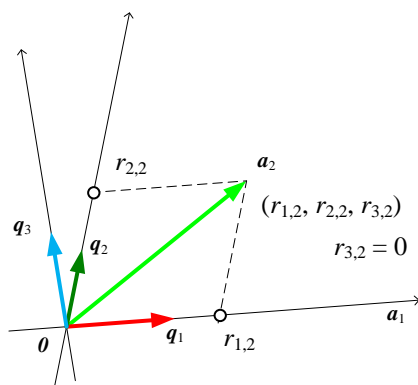


图 16. \mathbf{a}_2 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 中的坐标

\mathbf{a}_3 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 坐标

设 \mathbf{a}_3 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 坐标为 $(r_{1,3}, r_{2,3}, r_{3,3})$, 也就是说 \mathbf{a}_3 可以写成如下线性组合

$$\mathbf{a}_3 = r_{1,3}\mathbf{q}_1 + r_{2,3}\mathbf{q}_2 + r_{3,3}\mathbf{q}_3 \quad (41)$$

根据标量乘法, 向量 \mathbf{a}_3 在方向向量 \mathbf{q}_1 的坐标就是两者的内积, 即

$$r_{1,3} = \mathbf{a}_3 \cdot \mathbf{q}_1 = \mathbf{q}_1 \cdot \mathbf{a}_3 = \langle \mathbf{a}_3, \mathbf{q}_1 \rangle = \mathbf{a}_3^T @ \mathbf{q}_1 = \mathbf{q}_1^T @ \mathbf{a}_3 \quad (42)$$

带入具体数值

$$r_{1,3} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} = \sqrt{2} \quad (43)$$

这是 \mathbf{a}_3 在 \mathbf{q}_1 轴上的坐标 (标量投影)。

类似地，我们可以得到

$$r_{2,3} = \mathbf{a}_3 \cdot \mathbf{q}_2 = \mathbf{q}_2 \cdot \mathbf{a}_3 = \langle \mathbf{a}_3, \mathbf{q}_2 \rangle = \mathbf{a}_3^\top @ \mathbf{q}_2 = \mathbf{q}_2^\top @ \mathbf{a}_3 \quad (44)$$

带入具体数值

$$r_{2,3} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} = \sqrt{6}/3 \quad (45)$$

这是 \mathbf{a}_3 在 \mathbf{q}_2 轴上的坐标 (标量投影)。

类似地，我们可以得到

$$r_{3,3} = \mathbf{a}_3 \cdot \mathbf{q}_3 = \mathbf{q}_3 \cdot \mathbf{a}_3 = \langle \mathbf{a}_3, \mathbf{q}_3 \rangle = \mathbf{a}_3^\top @ \mathbf{q}_3 = \mathbf{q}_3^\top @ \mathbf{a}_3 \quad (46)$$

带入具体数值

$$r_{3,3} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -\sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \end{bmatrix} = 4\sqrt{3}/3 \quad (47)$$

写成矩阵乘法

$$\mathbf{a}_3 = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] \begin{bmatrix} r_{1,3} \\ r_{2,3} \\ r_{3,3} \end{bmatrix} \quad (48)$$

图 17 所示为 \mathbf{a}_3 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 中的坐标。

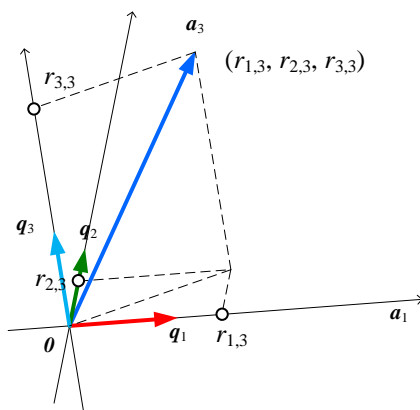


图 17. \mathbf{a}_3 在 $[\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$ 中的坐标

一组正交分解

这样我们便得到一组向量分解：

$$\begin{aligned}
a_1 &= q_1 @ q_1^T @ a_1 \\
a_2 &= q_1 @ q_1^T @ a_2 + q_2 @ q_2^T @ a_2 \\
a_3 &= q_1 @ q_1^T @ a_3 + q_2 @ q_2^T @ a_3 + q_3 @ q_3^T @ a_3
\end{aligned} \tag{49}$$

上述等式中，我们再一次看到了正交投影和向量正交分解！

这里面还有一个有趣的细节

$$a_1 = (q_1 @ q_1^T) @ a_1 \tag{50}$$

这实际上，引出了特征值分解 $Av = \lambda v$ 。

简单来说，对于方阵 A ，特征值分解是寻找一组非零向量 v 和对应标量 λ ，使得 $Av = \lambda v$ ，即方阵 A 在这些方向上只拉伸不改变方向（方向同向、反向）。

▲ 特征值分解是本书后续要着重展开讲解的话题之一。

矩阵乘法第三视角：列向量组合

把 (33)、(40)、(48) 三个矩阵乘法顺序排列得到

$$\underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}}_A = \underbrace{\begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix}}_Q \underbrace{\begin{bmatrix} \begin{bmatrix} r_{1,1} \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} r_{1,2} \\ r_{2,2} \\ 0 \end{bmatrix} & \begin{bmatrix} r_{1,3} \\ r_{2,3} \\ r_{3,3} \end{bmatrix} \end{bmatrix}}_R \tag{51}$$

R 为上三角矩阵。

根据 (8)、(10)、(19)， Q 为

$$Q = [q_1 \quad q_2 \quad q_3] = \begin{bmatrix} \sqrt{2}/2 & \sqrt{6}/6 & -\sqrt{3}/3 \\ \sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ 0 & \sqrt{6}/3 & \sqrt{3}/3 \end{bmatrix} \tag{52}$$

我们发现本例中 Q 为正交矩阵，因为列向量为单位向量，且两两正交。

根据本节前文涉及到的正交投影关系，上三角矩阵 R 还可以写成

$$R = \begin{bmatrix} a_1 \cdot q_1 & a_2 \cdot q_1 & a_3 \cdot q_1 \\ 0 & a_2 \cdot q_2 & a_3 \cdot q_2 \\ 0 & 0 & a_3 \cdot q_3 \end{bmatrix} = \begin{bmatrix} q_1^T @ a_1 & q_1^T @ a_2 & q_1^T @ a_3 \\ 0 & q_2^T @ a_2 & q_2^T @ a_3 \\ 0 & 0 & q_3^T @ a_3 \end{bmatrix} \tag{53}$$

R 的每一列 r_1 、 r_2 、 r_3 对应 A 的 a_1 、 a_2 、 a_3 在 $[q_1, q_2, q_3]$ 中坐标。

最后，把 (51) 写成

$$A = QR \tag{54}$$

这便是下一节的 QR 分解！



LA_09_04_02.ipynb 自定义函数完成矩阵列向量的格拉姆-施密特正交化，请大家自学。此外，请大家根据 (53)、(54) 修改代码重新计算矩阵 R 。



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 本节只有一道题目，我们将 (7) 中前两个列向量数值调换，具体如下；请大家根据前文内容对如下向量正交化。

$$\mathbf{a}_1 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$