

MIT 18.06 线性代数

第十四讲：正交向量与子空间

整理自 MIT 18.06 课程及相关公开笔记

引言

本讲重点关注向量之间、空间之间的**正交** (Orthogonal) 关系，这也是线性代数基本定理的第二部分。此前讨论的四个基本子空间，实际存在着——对应的正交补关系：行空间与零空间构成正交对，列空间与左零空间构成正交对。理解这种正交性，是后续学习投影、最小二乘法以及 $A^T A$ 矩阵性质的关键基础。

1 正交向量与勾股定理

正交 (Orthogonal) 指的是“垂直”的关系。如果两个向量是正交的，那么它们之间的夹角为 90° 。根据勾股定理，如果向量 $\mathbf{x}$ 与 $\mathbf{y}$ 正交，那么斜边 $\mathbf{x} + \mathbf{y}$ 的平方长度等于两条直角边的平方和：

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} + \mathbf{y}\|^2$$

利用向量内积的定义： $\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}$ ，将上述等式展开：

$$\mathbf{x}^T \mathbf{x} + \mathbf{y}^T \mathbf{y} = (\mathbf{x} + \mathbf{y})^T (\mathbf{x} + \mathbf{y})$$

展开右侧得到： $\mathbf{x}^T \mathbf{x} + \mathbf{y}^T \mathbf{y} + \mathbf{x}^T \mathbf{y} + \mathbf{y}^T \mathbf{x}$ 。进一步简化即可得到 $2\mathbf{x}^T \mathbf{y} = 0$ 。因此，两个向量正交的充要条件是其点积为 0：

$$\mathbf{x}^T \mathbf{y} = 0$$

说明：零向量 $\mathbf{0}$ 与任意向量都正交，因为 $\mathbf{0}^T \mathbf{y} = 0$ 恒成立。

2 子空间的正交

两个子空间 S 与 T **正交** 是指： S 中的任意一个向量都与 T 中的任意一个向量正交。

注意：两个子空间如果有交集，且交集包含非零向量，则这两个子空间一定不正交。例如，教室内的黑板所在平面与地板所在平面（两平面交于一条线）就不是正交子空间。因为取交线上的某个非零向量，该向量同时属于两个平面，但它无法满足“与自身正交”，此时点积为该向量长度的平方，必然不为 0。

在二维平面 $\mathbb{R}^2$ 中讨论子空间的正交：

- 整个平面 $\mathbb{R}^2$ 与任何子空间都不正交（除了仅含零向量的原点空间）。
- 过原点的直线 L 与仅包含原点的空间 $\mathbf{0}$ 总是正交的。
- 两条过原点的直线正交，当且仅当它们在原点处的夹角为 90° 。

3 零空间与行空间的正交关系

对于一个 $m \times n$ 的矩阵 A ，可以将其写成行向量的形式： $A = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_m^T \end{bmatrix}$ 。考察方程 $A\mathbf{x} = \mathbf{0}$ ，向量 $\mathbf{x}$ 属于零空间 $N(A)$ ，等价于：

$$\mathbf{r}_1^T \mathbf{x} = 0, \quad \mathbf{r}_2^T \mathbf{x} = 0, \quad \dots, \quad \mathbf{r}_m^T \mathbf{x} = 0$$

由于 $\mathbf{x}$ 与 A 的每一行都正交，而 A 的行空间是这些行向量的所有线性组合。根据线性组合的分配律， $\mathbf{x}$ 必然与行空间内的任意向量正交。因此，零空间 $N(A)$ 与行空间 $C(A^T)$ 是正交的。

正交补的概念：行空间和零空间不仅仅是正交，它们的维数相加恰好等于 $\mathbb{R}^n$ 的维数 n 。因此，行空间和零空间被称为 $\mathbb{R}^n$ 空间中的**正交补** (Orthogonal Complements)，这意味着零空间包含了所有与行空间正交的向量，反之亦然。

示例：设 $A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix}$ ，矩阵 A 的第二行是第一行的 2 倍，秩 $r = 1$ 。

- **行空间**维数为 1，即由向量 $\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$ 张成的过原点的直线。
- **零空间**维数为 $n - r = 3 - 1 = 2$ 。零空间就是垂直于行向量 $\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$ 且过原点的一个二维平面。
- 行空间与零空间构成 $\mathbb{R}^3$ 空间的一对正交补。

4 列空间与左零空间的正交关系

将矩阵 A 写成列向量的形式： $A = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \dots \quad \mathbf{c}_n]$ 。考察方程 $A^T \mathbf{y} = \mathbf{0}$ ，其中向量 $\mathbf{y}$ 属于左零空间 $N(A^T)$ ，由于 A^T 的行向量正是 A 的列向量的转置，该方程等价于：

$$\mathbf{c}_1^T \mathbf{y} = 0, \quad \mathbf{c}_2^T \mathbf{y} = 0, \quad \dots, \quad \mathbf{c}_n^T \mathbf{y} = 0$$

由于 $\mathbf{y}$ 与 A 的每一列都正交，而 A 的列空间是这些列向量的所有线性组合。根据线性运算的性质， $\mathbf{y}$ 必然与列空间内的任意向量正交。因此，左零空间 $N(A^T)$ 与列空间 $C(A)$ 是正交的，它们的维数加起来恰好等于 $\mathbb{R}^m$ 的维数 m 。

至此，线性代数基本定理的第二部分（正交性）得到了完整的阐述。

5 正规方程与 $A^T A$

在实际应用中，矩阵 A 的行数 m 通常远大于列数 n （样本量多于待估参数），方程 $A\mathbf{x} = \mathbf{b}$ 往往无精确解。此时要求一个“最优解”，即找到使误差最小的 $\hat{\mathbf{x}}$ 。通过将 $A\mathbf{x} = \mathbf{b}$ 两边同时左乘 A^T ，可以得到著名的**正规方程** (Normal Equations)：

$$A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$$

求解 $\hat{\mathbf{x}}$ 即可得到原方程的最优近似解。

5.1 $A^T A$ 的基本性质

1. **方阵与对称性**: 若 A 为 $m \times n$ 矩阵, 则 $A^T A$ 为 $n \times n$ 方阵, 且总是对称的:

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

2. **零空间相同**: $A^T A$ 的零空间与 A 的零空间完全相同。

$$N(A^T A) = N(A)$$

证明: 若 $A\mathbf{x} = \mathbf{0}$, 则 $A^T A\mathbf{x} = \mathbf{0}$ 显然成立。反之, 若 $A^T A\mathbf{x} = \mathbf{0}$, 则两边左乘 $\mathbf{x}^T$ 得 $\mathbf{x}^T A^T A\mathbf{x} = \mathbf{0}$, 即 $(A\mathbf{x})^T (A\mathbf{x}) = \mathbf{0}$ 。向量与自身的点积为 0, 意味着该向量为 0, 故 $A\mathbf{x} = \mathbf{0}$ 。

3. **秩相同**: 因为 $A^T A$ 与 A 的零空间相同, 所以它们的秩也相同。

5.2 $A^T A$ 的可逆性判断

由于 $r(A^T A) = r(A)$, 因此可以推导出 $A^T A$ 是否可逆的判别条件:

- **可逆**: 当且仅当 A 的列向量线性无关时, A 的零空间仅含零向量, 此时 $A^T A$ 的零空间也仅含零向量, 故 $A^T A$ **可逆**。
- **不可逆**: 当 A 的列向量线性相关时, $A^T A$ **不可逆**。

示例:

1. $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 5 \end{bmatrix}$ 的列向量线性无关, 计算得 $A^T A = \begin{bmatrix} 3 & 8 \\ 8 & 30 \end{bmatrix}$, 是可逆矩阵。

2. $A = \begin{bmatrix} 1 & 3 \\ 1 & 3 \\ 1 & 3 \end{bmatrix}$ 的两列成比例, 线性相关, 计算得 $A^T A = \begin{bmatrix} 3 & 9 \\ 9 & 27 \end{bmatrix}$, 是不可逆矩阵。

因此, 在实际使用正规方程求解最优解之前, 必须先检查原矩阵 A 的列向量是否线性无关。

总结

本讲是线性代数基本定理的重要章节, 主要结论如下:

- 向量正交的判定依据是点积为零: $\mathbf{x}^T \mathbf{y} = 0$ 。
- **零空间** $N(A)$ 与**行空间** $C(A^T)$ 是 $\mathbb{R}^n$ 中的正交补; **列空间** $C(A)$ 与**左零空间** $N(A^T)$ 是 $\mathbb{R}^m$ 中的正交补。
- 当 $A\mathbf{x} = \mathbf{b}$ 无精确解时, 可求正规方程 $A^T A\hat{\mathbf{x}} = A^T \mathbf{b}$ 的最优解。
- 矩阵 $A^T A$ 总是对称方阵, 且 $r(A^T A) = r(A)$ 。当且仅当 A 的列向量线性无关时, $A^T A$ 可逆。