



第五讲 子空间分析(1)

李宇峰

liyf@nju.edu.cn

人工智能学院



子空间分析

子空间，顾名思义，就是选择整个空间的一部分加以处理。

- 如何抽取子空间逼近原空间？
- 如何论证两个子空间是等价？
- 有哪些特殊子空间？
- 如何合成子空间？
- 如何对齐子空间？
- ...



子空间分析

- 子空间分析的一般理论
- 子空间的和空间、直接和空间、正交补空间
- 列空间、行空间、零空间
- 基于SVD的子空间旋转

5.1子空间分析的一般理论

所谓 $S \subset V$ 是线性空间 V 的一个子空间，当且仅当对 $\forall x, y \in S, \alpha, \beta \in F$ ，都有 $\alpha x + \beta y \in S$ ，则称 S 是 V 的一个子空间 (*subspace*)。【换言之，空间元素的任何线性合成仍属于空间，则为子空间】

子空间的例子很多，例如

(1) $S = \{(0, \alpha_2, \dots, \alpha_{n-2}, 0), \alpha_i \in R\} \subset R^n$, 则 S 是 R^n 子空间;

(2) S 是所有实系数偶次多项式构成的全体，则其是实系数多项式空间的子空间。

5.1 子空间分析的一般理论

子空间的基

假设 $V \in C^n$ 为 n 维的向量空间，考虑 m 个 n 维向量的子集合，其中 $m < n$ 。

定义 5.1 设 u_1, u_2, \dots, u_m 是向量空间 V 的一组向量系，则 u_1, u_2, \dots, u_m 的所有线性组合 W 称为由 u_1, u_2, \dots, u_m 张成的子空间，定义为

$$W = \text{span}\{u_1, u_2, \dots, u_m\} = \{u \mid u = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_m u_m\}$$

张成子空间 W 的每个向量称为 W 的生成元 (generator)，而所有生成元组成的集合 $\{u_1, u_2, \dots, u_m\}$ 称为子空间的张成集 (spanning set)。一个只包含了零向量的向量子空间称为平凡子空间 (trivial subspace)，通常不予考虑。

5.1 子空间分析的一般理论

定理 5.1.1 (张成集定理 (spanning set theorem)^[65, p.234]) 令 $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ 是向量空间 V 的一个子集, 并且 $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ 是由 S 的 m 个列向量张成的一个子空间。

(1) 如果 S 内有某个向量 (例如 \mathbf{u}_k) 是其他向量的线性组合, 则从 S 中删去向量 \mathbf{u}_k 后, 其他向量仍然张成子空间 W 。

(2) 若 $W \neq \{\mathbf{0}\}$, 即 W 为非平凡子空间, 则在 S 内一定存在某个由线性无关的向量组成的子集合, 它张成子空间 W 。

定理 5.1.2^[65, p.240] 令 $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ 是 n 维向量空间 W 的一组基, 则对于 W 中的任何一个向量 \mathbf{x} , 都存在一组唯一的标量 c_1, c_2, \dots, c_n , 使得 \mathbf{x} 可以表示为

$$\mathbf{x} = c_1 \mathbf{b}_1 + c_2 \mathbf{b}_2 + \dots + c_n \mathbf{b}_n \quad (5.1.2)$$

上述定理称为子空间向量的唯一表示定理。系数 c_1, c_2, \dots, c_n 的唯一性, 使得可以利用它们构成子空间 W 表示的 n 个坐标, 从而组成子空间的坐标系。

5.1 子空间分析的一般理论

如何论证两个子空间是等价？基等价

定义 5.1.2 令 W 是一向量空间。向量集合 $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p\}$ 称为 W 的一组基，若下列两个条件满足：

(1) 子空间 W 由向量 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$ 张成，即

$$W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p\}$$

(2) 向量集合 $B = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p\}$ 是一线性无关的集合。
关于子空间的基，有以下两点重要的事实：

(1) 当使用张成集定理从向量集合 S 中删去某个向量时，一旦 S 变成线性无关向量的集合，则必须立即停止从 S 内再删除向量。如果删去不是其他剩余向量的线性组合的额外向量，则较小的向量集合将不再张成原子空间 W 。因此，子空间的一组基是一个尽可能小的张成集。换句话说，张成子空间 W 的基向量一个也不能少。

(2) 一组基也是线性无关向量的尽可能大的集合。令 S 是子空间 W 的一组基，如果从子空间 W 内，给 S 再扩大一个向量 (例如 \mathbf{w})，则新的向量集合不可能是线性无关的，因为 S 张成子空间 W ，并且 W 内的向量 \mathbf{w} 本身就是 S 内各个基向量的线性组合。因此，张成子空间 W 的基向量一个也不能多。

所谓基，就是最小线性无关集合

5.1 子空间分析的一般理论

举个例子：矩阵张成空间

设矩阵 $A = [a_{ij}]_{m \times n} \in R^{m \times n}$ ， n 个列向量记为

$$a_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, a_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, a_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \in R^{m \times 1}$$

将 A 列向量的所有线性组合的集合构成的一个子空间，称为矩阵 A 的像空间 (记为 $R(A)$) 或者列空间 (column space)，也就是

$$R(A) = \text{span}\{a_1, a_2, \dots, a_n\} = \{y \in R^m \mid y = \sum_{j=1}^n k_j a_j, k_j \in R\}$$

5.1 子空间分析的一般理论

由向量系 u_1, u_2, \dots, u_m 张成的子空间就是该向量系的所有线性组合，而其中一组极大线性无关组构成张成子空间的一组**基底**，这一组基底假定了这个**张成空间 W 的维数**，也就是

$$\dim(W) = \dim(\text{span}\{u_1, u_2, \dots, u_m\}) = \text{rank}([u_1, u_2, \dots, u_m])$$

定义5.3 设 W 是一个向量子空间，向量系 u_1, u_2, \dots, u_p 称为 W 的**一组基底**，其满足下列两个条件：

- (1) 子空间 W 由向量 u_1, u_2, \dots, u_p 张成，即 $W = \text{span}\{u_1, u_2, \dots, u_p\}$ ；
- (2) 向量系 u_1, u_2, \dots, u_p 线性无关。

5.1 子空间分析的一般理论

子空间 W 的基底向量不是唯一的，但是该基底向量系所含向量的个数是唯一的，这个数就是子空间 W 的维数，记成 $\dim(W)$ 。

譬如 u_1, u_2, \dots, u_p 是张成空间 W 的一组基底，即

$$W = \text{span}\{u_1, u_2, \dots, u_p\}$$

则

$$\dim(W) = \text{rank}([u_1, u_2, \dots, u_p]) = p$$

5.1 子空间分析的一般理论

向量空间的基底向量是子空间的重要指标

因为这个子空间中有无穷多个向量，而“最具代表性”的向量系就是基底向量系，因为这个子空间中的任何向量都可以由这一组向量系唯一地表示出来。

定理 5.1.2 ^[65, p.240] 令 $B = \{b_1, b_2, \dots, b_n\}$ 是 n 维向量空间 W 的一组基，则对于 W 中的任何一个向量 x ，都存在一组唯一的标量 c_1, c_2, \dots, c_n ，使得 x 可以表示为

$$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n \quad (5.1.2)$$

上述定理称为子空间向量的唯一表示定理。系数 c_1, c_2, \dots, c_n 的唯一性，使得可以利用它们构成子空间 W 表示的 n 个坐标，从而组成子空间的坐标系。

向量空间 = 基底 + 唯一坐标表示

例题

例5.1 试求向量系

$$a_1 = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 4 \\ 9 \\ 5 \\ 4 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 3 \\ 7 \\ 4 \\ 3 \end{bmatrix}$$

所张成的 R^4 的子空间的基和维数。

解 设

$$k_1 a_1 + k_2 a_2 + k_3 a_3 = 0$$

即

$$\begin{cases} k_1 + 4k_2 + 3k_3 = 0 \\ 3k_1 + 9k_2 + 7k_3 = 0 \\ 2k_1 + 5k_2 + 4k_3 = 0 \\ k_1 + 4k_2 + 3k_3 = 0 \end{cases}$$

例题

解此线性方程组得

$$k_2 = 2k_1, \quad k_3 = -3k_1$$

于是有

$$a_1 + 2a_2 - 3a_3 = 0,$$

故 a_1, a_2, a_3 线性相关, 又显见 a_1 与 a_2 (或 a_2 与 a_3 或 a_1 与 a_3) 线性无关, 因此所论子空间维数是2, 即

$$\dim(\text{Span}\{a_1, a_2, a_3\}) = 2$$

基底由 a_1 与 a_2 (或 a_2 与 a_3 或 a_1 与 a_3) 所组成。

子空间分析

子空间，顾名思义，就是选择整个空间的一部分加以处理。

- 如何合成子空间？
- ...

子空间的和空间、直接和空间、正交补空间

和空间 \Rightarrow 直接和空间 \Rightarrow 正交补空间

都是合成子空间，性质越单纯，计算越便利



子空间的和空间、直接和空间

- 在子空间分析中，两个子空间之间的关系由这两个子空间之间的元素之间的关系来刻画。
- 如何通过两个子空间合成新子空间呢？一个直接做法就是和空间
- 子空间的和空间：设 S_1 和 S_2 是空间 V 的两个子空间，则 S_1 与 S_2 的和空间定义为

$$S = \{s \mid s = s_1 + s_2, \forall s_1 \in S_1, \forall s_2 \in S_2\}$$

记为 $S = S_1 + S_2$.

子空间的和空间、直接和空间

和空间是子空间，但是 元素表示 一般不唯一

譬如设 $S_1 = \text{span}\{e_1, e_2\}$, $S_2 = \text{span}\{e_2, e_3\}$, 其中 e_1, e_2, e_3 是自然基底，则向量

$$s = \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -5 \\ 3 \end{pmatrix} = a_1 + a_2,$$

$$\text{其中 } a_1 = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix} \in S_1, a_2 = \begin{pmatrix} 0 \\ -5 \\ 3 \end{pmatrix} \in S_2$$

$$\text{而 } s \text{ 也可以表示成 } s = \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix} \in S_1, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \in S_2.$$

子空间的和空间、直接和空间

问题：元素唯一表示需要什么条件？

直接和空间 如果 $S_1 + S_2$ 中的任一向量只能唯一地表示为子空间 S_1 的一个向量与子空间 S_2 的一个向量的和，则称 $S_1 + S_2$ 为直接和（或直和），记为 $S_1 \oplus S_2$.

思考： $S_1 + S_2$ 为直接和空间，那么 S_1 与 S_2 应该满足哪些关系？

要回答这个问题，先引入一个定义：交空间

子空间的和空间、直接和空间

交空间：设 S_1 和 S_2 都是线性空间 V 的子空间，所有既属于子空间 S_1 ，又属于 S_2 的向量集合称为 S_1 和 S_2 的交空间，记为

$$S = S_1 \cap S_2 = \{s \mid s \in S_1 \text{ and } s \in S_2\}$$

$S_1 + S_2$ 的交空间就是 S_1 与 S_2 的公共部分，容易证明交空间 $S_1 \cap S_2$ 也是 V 的子空间。

定理5.7 $V_1 + V_2$ 为直接和的充要条件是， V_1 与 V_2 之交 $V_1 \cap V_2$ 为零空间，即

$$V_1 \cap V_2 = \{0\}.$$

子空间的和空间、直接和空间

证明 充分性. 用反证法. 设 $V_1 \cap V_2$ 为零子空间, 若 $z \in V_1 + V_2$ 不能唯一地表示成 V_1 与 V_2 的向量的和.

则必有 $x_1, x_2 \in V_1, y_1, y_2 \in V_2$ 且 $x_1 \neq x_2, y_1 \neq y_2$, 使得

$$z \in x_1 + y_1 \quad \text{及} \quad z \in x_2 + y_2$$

两式相减得 $(x_1 - x_2) + (y_1 - y_2) = 0$,

则取向量 $w = (x_1 - x_2) = -(y_1 - y_2) \neq 0$, 而 $(x_1 - x_2) \in V_1$,
 $-(y_1 - y_2) \in V_2$, 故 $w \in V_1 \cap V_2$, 从而 $V_1 \cap V_2 \neq \{0\}$, 与假设矛盾.

必要性仍可用反证法证明.

子空间的和空间、直接和空间

S_1 与 S_2 的直接和空间具有下列性质：

(1) 直接和空间 $W = S_1 \oplus S_2$ 也是 V 的一个子空间.

(2) 从本质上讲, $S_1 \cap S_2 = \{0\}$ 的条件意味着子空间 S_1 与 S_2 中的任何非零向量是线性无关, 也就是设任意非零向量 $x, y, x \in S_1, y \in S_2$, 则 x 与 y 一定线性无关.

(3) 子空间 W 中的向量表示是唯一的, 即若 $x \in W$, 则 x 可以唯一地表示成 $x = x_1 + x_2$, 其中 $x_1 \in S_1, x_2 \in S_2$.

直接和 = 线性无关 + 唯一表示



如果我们想更加单纯一点, 从线性无关升级为正交呢?

正交子空间、正交补空间

正交子空间： 设 S_1 与 S_2 是线性空间 V 的子空间，设对 $\forall x \in S_1, y \in S_2$ ，都有 x 与 y 正交，则就称子空间 S_1 与 S_2 是互为正交的子空间，简称正交子空间。

正交补(orthogonal complement)空间： 设 S 是线性空间 V 的任意子空间，则 V 中**所有**与子空间 S 正交的向量组成的子空间称为 S 的正交补空间，记成 S^\perp ，也就是

$$S^\perp = \{y \in V \mid x \perp y, \forall x, y \in V\}.$$

正交子空间 不等于 正交补空间，区别在于是否包含 所有

正交子空间、正交补空间

我们称 $V = S \oplus S^\perp$ 为空间 V 的正交分解.

例如

$$R^3 = \text{span}\{e_1, e_2\} \oplus \text{span}\{e_3\} = \text{span}\{e_1, e_2\} \oplus \text{span}\{e_1, e_2\}^\perp$$

子空间的正交补空间

总起来说，对于线性子空间 S_1 与 S_2 ，**四类合成空间**的关系

1. $W = S_1 + S_2$ 称为和空间， S_1 和 S_2 之间没有关系，其维数满足 $\dim(W) \leq \dim(S_1) + \dim(S_2)$;
2. $W = S_1 \oplus S_2$ 称为直接和空间，这时 W 是满足 $S_1 \cap S_2 = \{0\}$ 的特殊和空间，必定有 $\dim(W) = \dim(S_1) + \dim(S_2)$;
3. $W = S_1 \oplus S_2$ 且满足 $S_1 \perp S_2$ ，这时 W 是正交子空间的直接和
4. $W = S_1 \oplus S_1^\perp$ ， S_1^\perp 为 S_1 的正交补空间，这时 S_1 、 S_2 是 W 的正交分解。这种分解在信号处理、模式识别、自动控制等方面都有应用

在以上四种和空间中，条件越来越强，性质越来越纯粹

子空间的和空间、直和空间

和空间

- **全面描述向量空间结构**：和空间是由两个或多个子空间的向量相加所生成的新空间。通过研究和空间，可以了解不同子空间之间的相互关系以及它们如何共同构成更大的向量空间，有助于全面把握向量空间的整体结构。例如，在三维欧几里得空间中，两个不同的二维子空间（如两个不同的平面）的和空间可能是整个三维空间，这体现了子空间组合对空间维度的扩展作用。
- **解决线性方程组问题**：在求解线性方程组时，系数矩阵的列空间和行空间的和空间与方程组的解空间密切相关。通过分析和空间的性质，可以判断线性方程组是否有解，以及解的结构和性质。例如，若两个子空间分别对应线性方程组的系数矩阵的列空间和零空间，它们的和空间的性质能帮助确定方程组解的存在性和唯一性。

直和空间

- **简化空间分解与分析**：直和是和空间的一种特殊且重要的情况，当子空间的和满足特定条件（即每个向量的表示唯一）时，称为直和空间。直和分解将一个复杂的向量空间分解为若干个简单的子空间的直和，使得对整个空间的研究可以转化为对这些子空间的研究，大大简化了问题。例如，一个有限维向量空间可以分解为它的特征子空间的直和，这样就可以通过分别研究各个特征子空间的性质来了解整个向量空间的性质，在矩阵对角化等问题中有着关键应用。
- **物理系统中的应用**：在物理学中，例如量子力学的希尔伯特空间中，系统的状态空间常常可以表示为不同子空间的直和。每个子空间对应着系统的不同性质或自由度，通过直和分解可以更清晰地分析系统的各种状态和相互作用。

正交补空间

- **完备向量空间结构：**正交补空间是对向量空间结构的进一步完善和细化。对于一个给定的子空间，其正交补空间与它共同构成了整个向量空间的一个完备分解。例如，在三维欧几里得空间中，一个平面子空间的正交补空间是一条与该平面垂直的直线，平面与直线共同构成了整个三维空间，这种分解有助于更全面、深入地理解向量空间的内部结构和性质。
- **简化问题求解：**在许多数学问题中，利用正交补空间可以将问题分解为在相互正交的子空间上分别求解，从而简化问题的复杂度。例如，在求解线性方程组时，如果能够将系数矩阵的列空间分解为一个子空间和它的正交补空间，那么可以分别在这两个子空间上考虑方程组的解，利用正交性的性质可以使计算更加简便，并且有助于分析解的存在性和唯一性。
- **信号处理与数据分析：**在信号处理和数据分析中，正交补空间有着广泛的应用。例如，在数据降维中，主成分分析（PCA）实际上就是利用了数据空间的正交分解，将数据投影到主成分子空间和其正交补空间上，通过保留主成分子空间中的信息来实现数据的降维，同时可以利用正交补空间来分析数据中的噪声和异常信息。在信号滤波中，也可以利用信号空间和其正交补空间的关系，设计滤波器来滤除特定频率或方向的噪声信号，保留有用信号。
- **优化与投影算法：**在优化问题和投影算法中，正交补空间起着关键作用。例如，在最小二乘问题中，投影到一个子空间上的向量与该子空间的正交补空间中的向量是正交的，利用这一性质可以通过求解在子空间和其正交补空间上的投影问题来找到最小二乘解。许多迭代算法，如共轭梯度法等，也利用了向量空间的正交分解来加速收敛，提高算法的效率和稳定性。

子空间分析

子空间，顾名思义，就是选择整个空间的一部分加以处理。

- 有哪些特殊子空间？
- ...

列空间、行空间、零空间



列空间、行空间、零空间

定义 5.2.2 若 \mathbf{A} 是一个 $m \times n$ 复矩阵, 则 \mathbf{A} 的值域 (range) 定义为

$$\text{Range}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{C}^m \mid \mathbf{A}\mathbf{x} = \mathbf{y}, \quad \mathbf{x} \in \mathbb{C}^n\} \quad (5.2.5)$$

矩阵 \mathbf{A} 的零空间 (null space) 也称 \mathbf{A} 的核 (kernel), 定义为满足齐次线性方程 $\mathbf{A}\mathbf{x} = \mathbf{0}$ 的所有解向量的集合, 即

$$\text{Null}(\mathbf{A}) = \text{Ker}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{C}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\} \quad (5.2.6)$$

类似地, 复矩阵 $\mathbf{A}_{m \times n}$ 的共轭转置 \mathbf{A}^H 的零空间定义为

$$\text{Null}(\mathbf{A}^H) = \text{Ker}(\mathbf{A}^H) = \{\mathbf{x} \in \mathbb{C}^m \mid \mathbf{A}^H\mathbf{x} = \mathbf{0}\} \quad (5.2.7)$$

零空间的维数称为 \mathbf{A} 的零化维 (nullity), 即有

$$\text{nullity}(\mathbf{A}) = \dim[\text{Null}(\mathbf{A})] \quad (5.2.8)$$

列空间、行空间、零空间

定理 5.2.1 若 \mathbf{A} 是 $m \times n$ 复矩阵, 则 \mathbf{A} 的行空间的正交补 $(\text{Row}(\mathbf{A}))^\perp$ 是 \mathbf{A} 的零空间, 并且 \mathbf{A} 的列空间的正交补 $(\text{Col}(\mathbf{A}))^\perp$ 是 \mathbf{A}^H 的零空间, 即有

$$(\text{Row}(\mathbf{A}))^\perp = \text{Null}(\mathbf{A}), \quad (\text{Col}(\mathbf{A}))^\perp = \text{Null}(\mathbf{A}^H) \quad (5.2.11)$$

总结以上讨论, 即可得到与矩阵 \mathbf{A} 的向量空间之间的关系如下:

(1) 矩阵 \mathbf{A} 的值域与列空间相等, 即

$$\text{Range}(\mathbf{A}) = \text{Col}(\mathbf{A}) = \text{Span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$$

(2) 矩阵 \mathbf{A} 的行空间与 \mathbf{A}^H 的列空间相等, 即

$$\text{Row}(\mathbf{A}) = \text{Col}(\mathbf{A}^H) = \text{Range}(\mathbf{A}^H)$$

(3) 矩阵 \mathbf{A} 的行空间的正交补等于 \mathbf{A} 的零空间, 即

$$(\text{Row}(\mathbf{A}))^\perp = \text{Null}(\mathbf{A})$$

(4) 矩阵 \mathbf{A} 的列空间的正交补就是 \mathbf{A}^H 的零空间, 即

$$(\text{Col}(\mathbf{A}))^\perp = \text{Null}(\mathbf{A}^H)$$

列空间、行空间、零空间

表 5.2.1 $m \times n$ 矩阵 A 的零空间与列空间的对比 [65, p.226]

零空间 $\text{Null}(A)$	列空间 $\text{Col}(A)$
$\text{Null}(A)$ 是 \mathbb{C}^n 的子空间	$\text{Col}(A)$ 是 \mathbb{C}^m 的子空间
$\text{Null}(A)$ 为隐含定义, 与 A 的列向量无直接关系	$\text{Col}(A)$ 为显式定义, 直接由 A 的所有列向量张成
$\text{Null}(A)$ 的基应满足 $Ax = 0$	$\text{Col}(A)$ 的基是 A 的主元列
$\text{Null}(A)$ 与矩阵 A 的元素无任何明显关系	矩阵 A 的每一列都在 $\text{Col}(A)$ 内
$\text{Null}(A)$ 的典型向量 v 满足 $Av = 0$	$\text{Col}(A)$ 的典型向量满足 $Ax = v$ 为一致方程
$v \in \text{Null}(A)$ 的条件: $Av = 0$	$v \in \text{Col}(A)$ 的条件: $[A, v]$ 与 A 具有相同的秩
$\text{Null}(A) = \{0\}$ 当且仅当 $Ax = 0$ 只有零解	$\text{Col}(A) = \{0\}$ 当且仅当 $Ax = b$ 有解
$\text{Null}(A) = \{0\}$ 当且仅当 Ax 为一对一映射	$\text{Col}(A) = \{0\}$ 当且仅当 Ax 为 \mathbb{C}^n 到 \mathbb{C}^m 的映射

子空间分析

子空间，顾名思义，就是选择整个空间的一部分加以处理。

- 如何 对齐 两个子空间？
- ...



基于奇异值分解的子空间旋转对齐

5.2 基于SVD的子空间旋转对齐

现实观察： 在工程中经常会遇到对同一对象进行多次测量，并且每一次的测量数据并不完全相同。

问题形式化： 令 A 和 B 分别是两次测量得到的 $m \times n$ 数据矩阵，希望求一个 $n \times n$ 实正交矩阵 Q ，即 $Q^T Q = I$ ，使得

$$\begin{aligned} \min_Q & \|A - BQ\|_F \\ \text{s.t.} & \quad Q^T Q = I \end{aligned}$$

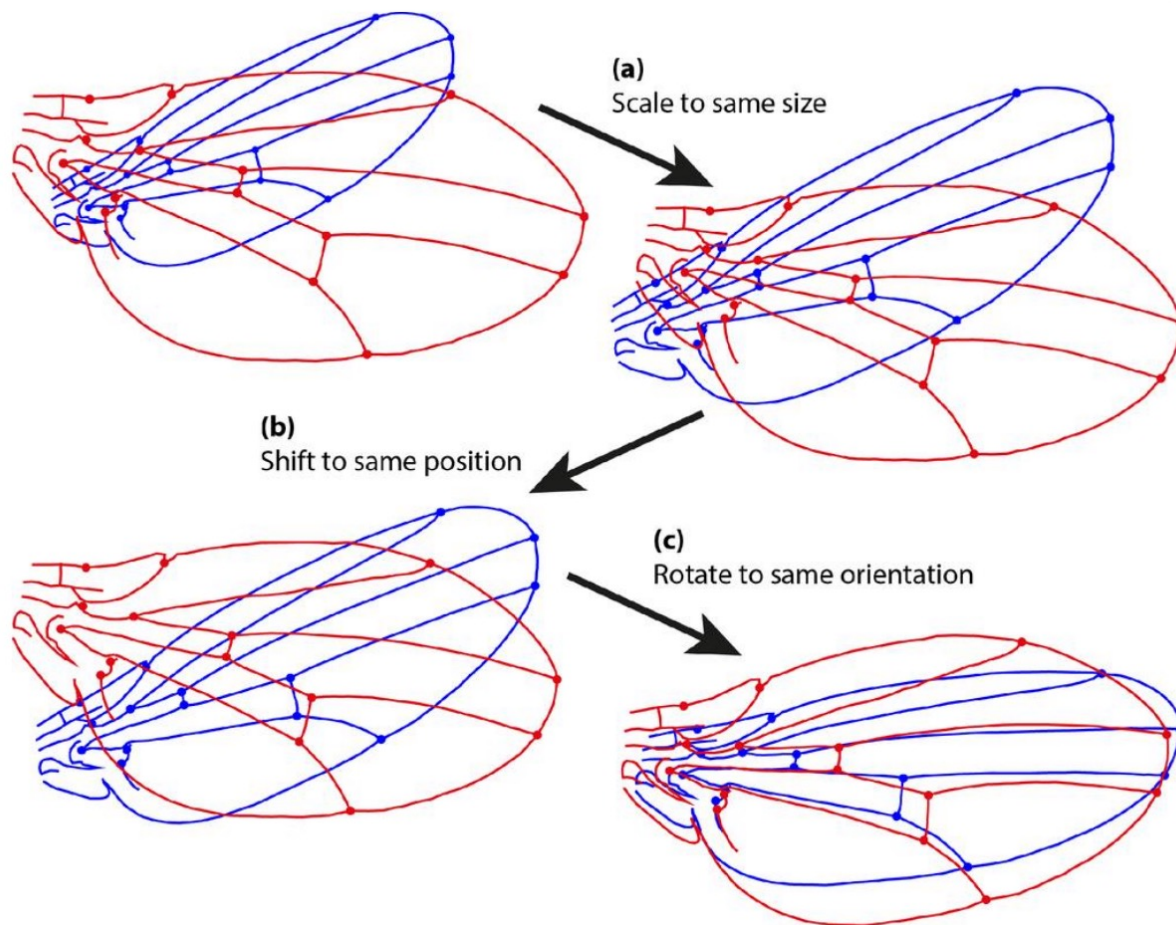
也就是通过正交矩阵 Q ，使得 BQ 与 A 尽可能对齐。

这个问题称为**正交强迫对齐** (Orthogonal Procrustes Problem)，

最早由Green于1952年在计量心理学杂志上提出。

5.2 基于SVD的子空间旋转对齐

建立投影，对齐两个空间



5.2 基于SVD的子空间旋转对齐

思路:

1. 由于 Q 是正交矩阵，矩阵乘积 BQ 并不改变 B 的列向量之间的线性无关性，所以列空间 $R(BQ) = R(B)$ 。
2. 矩阵乘积 BQ 可以看成是对矩阵 B 旋转。因此从子空间的角度看，正交强迫对齐问题相当于使列子空间 $R(B)$ 旋转进入列子空间 $R(A)$ 内。
3. 矩阵的Frobenius范数 $\|A - BQ\|_F$ 起着度量正交强迫对齐解的质量的作用。

矩阵范数平方和 $\|A - BQ\|_F^2$ 可以写成迹函数的形式

$$\|A - BQ\|_F^2 = \text{tr}(A^T A) + \text{tr}(B^T B) - 2\text{tr}(Q^T B^T A)$$

于是原问题就等价于使矩阵的迹 $\text{tr}(Q^T B^T A)$ 最大化。

5.2 基于SVD的子空间旋转对齐

求解思路：迹函数 $tr(Q^T B^T A)$ 的最大化可以通过矩阵乘积 $B^T A$ 的奇异值分解来实现。

令矩阵 $B^T A$ 的奇异值分解 $B^T A = U \Sigma V^T$, 其中

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n).$$

若取正交矩阵 $Z = V^T Q^T U \in R^{n \times n}$, 则有

定理 因 $Z = [z_{ij}]_{n \times n}$ 为正交矩阵, 则 Z 的所有主对角元

$$z_{ii} \leq 1.$$

证明 由于 $Z^T Z = I_n$, $\Rightarrow \sum_{j=1}^n (z_{ij})^2 = 1$, 故对

$$z_{ii} \leq 1 \quad i = 1, 2, \dots, n. \quad \text{定理得证}$$

5.2 基于SVD的子空间旋转对齐

这样就有

$$\begin{aligned} \text{tr}(Q^T B^T A) &= \text{tr}(Q^T U \Sigma V^T) = \text{tr}(V^T Q^T U \Sigma) \\ &= \text{tr}(Z \Sigma) = \sum_{i=1}^n z_{ii} \sigma_i \leq \sum_{i=1}^n \sigma_i \end{aligned}$$

当且仅当 $Z = I$ 即 $Q = UV^T$ 时等号成立。

换言之，当 $Q = UV^T$ ， $\text{tr}(Q^T B^T A)$ 得到最大值，从而实现

$\|A - BQ\|_F$ 最小化

5.2 基于SVD的子空间旋转对齐

子空间旋转对齐小结：

1. 以上分析表明， $B^T A = U \Sigma V^T$ 是矩阵乘积 $B^T A$ 的奇异值分解，则正交矩阵 $Q = UV^T$ 为正交强迫对齐问题的解。
2. 事实上，正交强迫对齐问题也可以看成是将矩阵 A 分解为 BQ ，就是将矩阵分解为 $m \times n$ 矩阵 B 和 $n \times n$ 正交矩阵 Q 的乘积。这种矩阵分解有时也称为极分解 (polar decomposition)。

5.2 基于SVD的子空间旋转对齐

有意思的是，若 $B = I$ ，则正交强迫对齐问题变为

$$Q = \min_{Q^T Q = I} \|A - Q\|_F$$

这一问题的数学描述是求一个与已知 $n \times n$ 矩阵 A 最接近的正交矩阵。根据前面的分析，若 $A = U\Sigma V^T$ 是奇异值分解，则 $Q = UV^T$ 是与矩阵 A 最接近的正交矩阵。

子空间分析-潜在应用

子空间，顾名思义，就是选择整个空间的一部分加以处理。

- 机器学习与模式识别

- 提高模型性能：在机器学习中，子空间分析可以作为预处理步骤，对数据进行降维和特征提取，从而提高机器学习模型的训练速度和泛化能力。例如，在支持向量机（SVM）分类中，使用子空间分析后的特征作为输入，可以减少模型的过拟合风险，提高分类准确率。
- 发现数据中的模式：子空间分析有助于发现数据中的隐藏模式和结构。例如，在聚类分析中，通过子空间分析可以将数据点划分到不同的子空间中，每个子空间对应一个聚类，从而实现了对数据的有效聚类。

- 系统分析与控制

- 系统建模：在控制系统中，子空间分析可以用于建立系统的状态空间模型。通过对系统输入输出数据的分析，确定系统的状态子空间，从而为系统的建模、分析和控制提供基础。
- 故障诊断：子空间分析可以用于检测系统中的故障。通过监测系统状态在子空间中的变化，及时发现系统的异常行为，从而实现故障的早期诊断和预警。

子空间分析-潜在应用

子空间，顾名思义，就是选择整个空间的一部分加以处理。

- 数据降维与特征提取

- 降低数据复杂度：实际数据往往具有高维度，这会增加计算成本和模型复杂度。通过子空间分析，如主成分分析（PCA），可以将数据投影到低维子空间
- 提取关键特征：子空间分析能够找出数据中最具代表性的特征，将原始数据中的噪声和无关信息过滤掉。例如在图像识别中，通过子空间分析可以提取图像的关键特征，如边缘、纹理等，有助于提高图像分类和识别的准确率。

- 信号处理与恢复

- 信号增强：在信号处理中，子空间分析可以用于增强有用信号，抑制噪声。例如，在语音信号处理中，通过将语音信号投影到特定的子空间，可以去除背景噪声，提高语音的清晰度和可懂度。
- 信号恢复：对于受损或丢失的信号，子空间分析可以利用信号在子空间中的特性进行恢复。例如，在压缩感知中，通过对子空间的理解，可以从少量的观测数据中恢复出原始的高维信号。

子空间分析

- 子空间分析的一般理论
- 子空间的和空间、直接和空间、正交补空间
- 列空间、行空间、零空间
- 基于SVD的子空间旋转