第五讲 子空间分析

李宇峰

liyf@nju.edu.cn

人工智能学院

子空间分析

在涉及拟合逼近、最优化、微分方程、通信、信号处理、模式识别、自动控制等问题中，向量子空间的概念起着重要的作用。向量子空间也为解决大量工程应用问题提供了一类有效的方法，称为子空间方法。

这里我们主要讨论子空间的分析理论，介绍子空间方法的一些典型工程应用。

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，是指对$\forall x, y \in S, \alpha, \beta \in F$, 都有$\alpha x + \beta y \in S$, 则称$S$是$V$的一个子空间（*subspace*）.

子空间的例子很多，例如

（1）$S = \{(0, \alpha_2, \cdots, \alpha_{n-2}, 0), \alpha_i \in R\} \subset R^n$, 则$S$是$R^n$的一个子空间；

（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。

Subspace, Basis and Spanning Set

子空间的基

在具体讨论各种子空间之前，有必要先介绍子空间的一些基本概念、子空间之间的代数关系和几何关系。

假设$V \in C^n$为n维的复向量空间，考虑m个n维复向量的子集合，其中设$m < n$.

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

$$W = span\{u_1, u_2, \cdots, u_m\} = \{u | u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m\}$$

Subspace, Basis and Spanning Set

下面我们来给出一些常用的张成空间。

设矩阵$A=[a_{ij}]_{m\times n}\in R^{m\times n}$, 其$m$个行向量表示为

$$r_1=(a_{11},a_{12},\cdots,a_{1n})\in R^{1\times n},\qquad \cdots\ ,$$

$$r_m=(a_{m1},a_{m2},\cdots,a_{mn})\in R^{1\times n}$$

矩阵A的n个列向量记为

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

Subspace, Basis and Spanning Set

这时我们将A的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

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

Subspace, Basis and Spanning Set

定义5.2（张成集定理 *spanning set theorem*）设$u_1, u_2, \cdots, u_m$是向量空间V的一组向量系，并且$W = span\{u_1, u_2, \cdots, u_m\}$是由$m$个列向量$u_1, u_2, \cdots, u_m$张成的一个子空间，那么

（1）如果向量系$u_1, u_2, \cdots, u_m$内某个向量（譬如u_k）是其他向量的线性组合，则从这m个向量中删去u_k后，其他向量仍然张成子空间W，即$span\{u_1, u_2, \cdots, u_m\} = span\{u_1, \cdots, u_{k-1}, u_{k+1}, \cdots, u_m\}$；

（2）若$W \neq 0$，即W为非平凡子空间，则在$u_1, u_2, \cdots, u_m$中一定存在某个由线性无关向量组成的子集，它张成子空间W。

Subspace, Basis and Spanning Set

由定理5.1我们可以看出，由向量系$u_1, u_2, \cdots, u_m$张成的子空间就是该向量系的所有线性组合，而这一组向量系的一组极大线性无关组可以构成这个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

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

定义5.3 设W是一个向量子空间，向量系$u_1, u_2, \cdots, u_p$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$u_1, u_2, \cdots, u_p$张成，即$W = span\{u_1, u_2, \cdots, u_p\}$；

(2) 向量系$u_1, u_2, \cdots, u_p$线性无关。

Subspace, Basis and Spanning Set

前面已经讲到，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$u_1, u_2, \cdots, u_p$是张成空间W的一组基底，即

$$W = span\{u_1, u_2, \cdots, u_p\}$$

则

$$\dim(W) = rank([u_1, u_2, \cdots, u_p]) = p$$

Subspace, Basis and Spanning Set

向量子空间的基底向量是这个子空间中的一个重要指标，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。

Subspace, Basis and Spanning Set

定理5.4 设$u_1, u_2,\cdots, u_n$是n维向量子空间W的一组基底，则对于W中的任何一个向量x，都存在一组唯一的标量 $c_1, c_2,\cdots, c_n$ ，使得x可以表示为

$$x = c_1u_1 + c_2u_2 +\cdots +c_nu_n$$

这就是向量x在这组基底下的坐标表示。

以上定理就是子空间向量的唯一表示定理，系数$c_1, c_2,\cdots, c_n$就是x在$u_1, u_2,\cdots, u_n$基底下的坐标，同一个向量x在不同的基底下的坐标是不同的。

例题

例5.1 试求向量系

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

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

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

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（即向量）之间的关系来刻画。

- 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$S = S_1 + S_2$.

Sum Space and Direct Sum Space

和空间是V的一个子空间，但和空间的元素表示一般不唯一。

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

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

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

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

Sum Space and Direct Sum Space

问题：能否找到一种特殊的和子空间，使得这个和子空间中的任一元素能唯一地表示成两个子空间中元素的和。

定义5.6 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：如果要求和空间$S_1 + S_2$为直接和空间，那么其中S_1与S_2应该有哪些关系？

Sum Space and Direct Sum Space

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

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

$S_1 + S_2$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。

Sum Space and Direct Sum Space

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

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

Sum Space and Direct Sum Space

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

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

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.

Sum Space and Direct Sum Space

S_1与S_2的直接和空间是一个特殊的和空间。

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的，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$x_1 \in S_1, x_2 \in S_2$.

Orthogonal Subspaces

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$x \perp y$, 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的子空间，而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，记成$S^{\perp}$, 也就是

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

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

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

Orthogonal Subspaces

总起来说，对于V的线性子空间S_1与S_2：

（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$S_1 \cap S_2 = \{0\}$的特殊的和空间，必定有$\dim(W) = \dim(S_1) + \dim(S_2)$；

（3）$W = S_1 \oplus S_2$, 且满足$S_1 \perp S_2$, 这时称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.

（4）如果$V = S_1 \oplus {S_1}^{\perp}$，则$S_1^{\perp}$称为$S_1$的正交补空间.

在以上四种和空间中，条件一个比一个强。

5.2 基于SVD的子空间的旋转

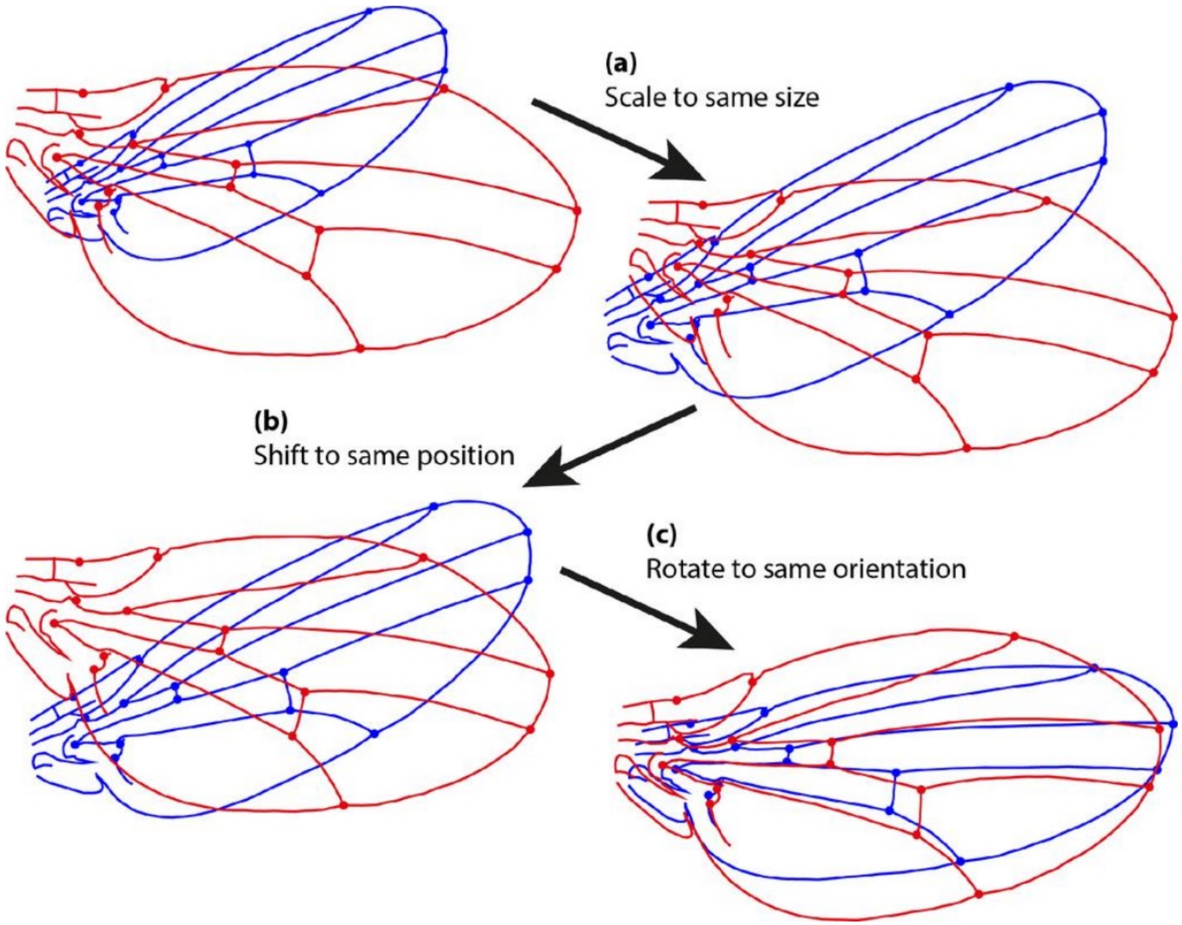
在工程中经常会遇到对同一对象进行多次测量，并且每一次的测量数据并不完全相同。令A和B分别是两次测量得到的$m \times n$数据矩阵，现在我们希望求一个$n \times n$实正交矩阵Q, 在$Q^TQ = I$的约束下，使得

$$\min_{Q} \|A - BQ\|_F$$
$$s.t. \quad Q^TQ = I$$

也就是通过正交矩阵Q，强迫BQ与A尽可能一致。

上述问题称为正交强迫一致问题(Orthogonal Procrustes Problem), 最早由Green于1952年在计量心理学杂志上提出。

Orthogonal Procrustes Problem



Orthogonal Procrustes Problem

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$R(BQ) = R(B)$。另外矩阵乘积BQ可以看成是对矩阵B旋转。因此从子空间的角度看，正交强迫一致的问题相当于使列子空间$R(B)$旋转进入列子空间$R(A)$内。显然矩阵的$Frobenius$范数$\|A - BQ\|_F$起着度量正交强迫一致问题的解的质量的作用。

矩阵范数平方和$\|A - BQ\|_F^2$可以写成迹函数的形式

$$\|A - BQ\|_F^2 = tr(A^TA) + tr(B^TB) - 2tr(Q^TB^TA)$$

于是原问题就等价于使矩阵的迹$tr(Q^TB^TA)$最大化。

Orthogonal Procrustes Problem

迹函数$tr(Q^TB^TA)$的最大化可以通过矩阵乘积B^TA的奇异值分解来实现。

令矩阵B^TA的奇异值分解$B^TA = U\Sigma V^T$, 其中

$$\Sigma = diag(\sigma_1, \sigma_2, \cdots, \sigma_n).$$

若取正交矩阵$Z = V^TQ^TU \in R^{n\times n}$，则

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

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

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

$$z_{ii} \le 1 \quad i = 1,2,\cdots,n.$$ 定理得证

Orthogonal Procrustes Problem

这样就有

$$tr(Q^T B^T A) = tr(Q^T U\Sigma V^T) = tr(V^T Q^T U\Sigma)$$

$$= tr(Z\Sigma) = \sum_{i=1}^{n} z_{ii}\sigma_i \le \sum_{i=1}^{n} \sigma_i$$

当且仅当$Z = I$即$Q = UV^T$时等号成立。换言之若选择$Q = UV^T$, 则$tr(Q^T B^T A)$取最大值.

从而使$\|A - BQ\|_F$取最小值.

Orthogonal Procrustes Problem

以上分析表明，若$B^TA = U\Sigma V^T$是矩阵乘积B^TA的奇异值分解,

则只要取正交矩阵$Q = UV^T$就是正交强迫一致问题的解.

事实上正交强迫一致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*)。

Orthogonal Procrustes Problem

有意思的是，若$B = I$, 则正交强迫一致问题变为

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

这一问题的数学描述是求与已知$n \times n$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。