线性代数

-15 课　子空间投影

一、知识概要

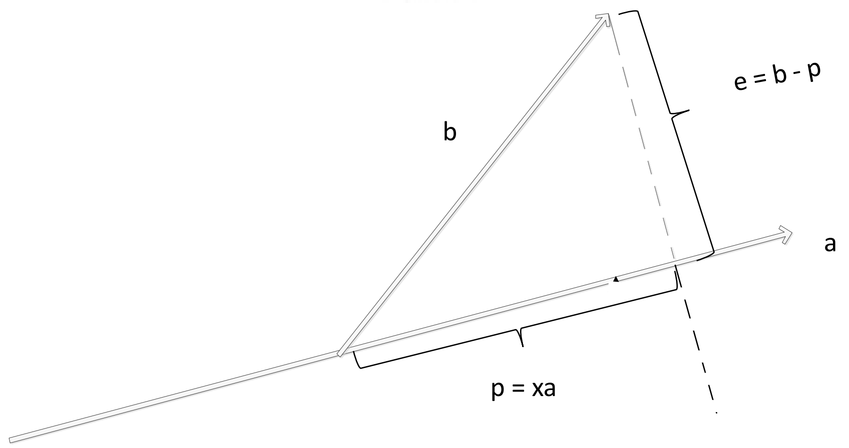
这一节主要在讲解投影，从向量的投影入手，延伸到高维投影，并将投影用矩阵形式给出，上一节中我们介绍了正交的基本概念，这一节中我们需要使用这个概念，做投影即是向另一个向量上做垂线。通过本节的学习，我们可以更好的理解正交概念以及空间之间的投影概念。

二. 投影

2.1 简单的投影

介绍一下简单的投影，如图（一）所示，p 即为 b 在 a 上的投影，写做 p = xa（x 为倍数）。而这时，它们之间的差值是 b-p。那么现在的问题就是这个投影有什么特点呢？很明显，联系我们上节所学习的正交的概念，a，b 都看作向量，由 a 与 e 的垂直关系，容易得到：

$$a^T e = a^T(b-p) = a^T(b-ax) = 0 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots（1）$$



图（一）

解式（1）得到 x，再将 x 代入 p = xa，容易得到：

$$p = a\frac{a^T b}{a^T a}\ 。$$

很明显，p 的形式中含有 b，那么就说明投影是通过前面系数(矩阵形式)来完成的，我们把它称为：**投影矩阵**。即：

p = Pb（P 是投影矩阵，作用于 b 向量上）

对比一下，不难得到： $P = a\frac{a^T}{a^Ta}$，也就是说，P 矩阵生成了投影 p。

注：这里的 aa^T 与 a^Ta 是不同的，当 a 是列向量时，前者是一个矩阵，后者是一个具体数字。而图（一）这个情况中，矩阵 P 的秩为 1，而且对称。这点我们通过具体计算或上一节末尾的结论就可以看出来，在此不赘述。

我们更需要注意的是 P 的两条性质，首先我们很容易看出：

$$P^T = P$$

因为 $P = a\frac{a^T}{a^Ta}$， 下面分母是数字，上面分子转置与之形式一样。

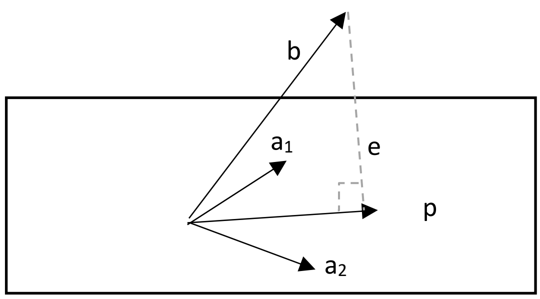
其次，我们来考虑这样一个问题，如果将用 P 投影两次会怎样？答案是矩阵不变，因为在图（一）中，我们如果投影两次，那么其表现出的结果与只投影一次并没有区别。所以我们还可以得到：

$$P^2 = P$$

以上两条性质是我们继续扩展投影概念的基础。

2.2 平面上的投影

介绍完了向量之间的投影，我们来看一看向量与平面之间的投影。如图(二)



图（二）

其中a_1，a_2 为构成平面的一组基，p 在平面上。所以有：p = $\widehat{x_1}a_1 + \widehat{x_2}a_2$。或者写做：p = A$\hat{x}$，A = $[a_1 \quad a_2]$(其中的a_n都是列向量)，$\hat{x} = \begin{bmatrix}\widehat{x_1}\\ \widehat{x_2}\end{bmatrix}$。

那么现在我们可以将关于平面投影的问题与关于向量之间投影的问题联系起来，用之前的思路，分别研究a_1与a_2，因为它们都与 b-p 向量垂直。将 p = Ax 代入这种垂直关系。可以得到下面两个式子：

$$a_1{}^T(\text{b-A}\hat{x}) = 0 \qquad\qquad a_2{}^T(\text{b-A}\hat{x}) = 0$$

合在一起，写成矩阵形式：

$$\begin{bmatrix}a_1{}^T\\ a_2{}^T\end{bmatrix}(\text{b-A}\hat{x}) = \begin{bmatrix}0\\0\end{bmatrix} \longleftrightarrow A^T(\text{b-A}\hat{x}) = 0$$

得到：

$$A^T(\text{b-A}\hat{x}) = 0..........（2）$$

（2）式看上去和（1）式形式差不多，也就是说（1）式其实是（2）的一种情况，为 A 只有一列时的情况。而在（2）式说明：b-A$\hat{x}$（垂直的 e 向量）在 A^T 的零空间里。

上节课学过，零空间与行空间正交，所以有 e 向量与 A^T的行向量正交，也就与 A 的列向量正交，也就是 e 向量也垂直于a_1，a_2向量。联系图像，显然成立。

研究完了这个方程，我们继续看一看面上的的投影矩阵 **P** 是什么。我们根据（2）式，先计算 p = A$\hat{x}$来找到 p 与 b 之间投影关系。

首先化简式(2)，得到 A^Tb = A^TA$\hat{x}$。然后我们要注意一点：不能直接在两边左乘$(A^T)^{-1}$。因为 A^T不一定是方阵，不一定有逆矩阵。但是这里 A 为两个基向量构成的矩阵，两个基向量线性无关，根据上节的知识，A^TA是可逆的。所以我们应该在两边同时左乘$(A^TA)^{-1}$解这个方程。

得到：

$$\hat{x} = (A^TA)^{-1}A^Tb$$

代入 p = A$\hat{x}$：

$$\text{p} = \text{A}\hat{x} = \boxed{\text{A}(A^TA)^{-1}A^T}\,b$$

这样就找到了 p 与向量 b 之间的投影关系，进而得到投影矩阵：

$$\mathbf{P} = \text{A}(A^TA)^{-1}A^T..................（3）$$

这就意味着，(3)式是投影矩阵 **P** 的一般情况，而我们上面计算的 $P = a\frac{a^T}{a^Ta}$，是投影矩阵的一维特殊情况。同样，**P** = $A(A^TA)^{-1}A^T$也具有两个性质：

$$P^T = P \quad , \quad P^2 = P$$

利用同上面一维的变换就可以得到了，这里不详述。大家可以自己验证一下。

三、最小二乘法初涉

那么我们之前学习的投影有什么用呢？我们不难发现，上面求得投影中的 e 是可以看做 b 与 a 之间偏移量的大小。这样就为我们使用最小二乘法拟合直线提供了方便。

【例】

求解：三个点（1，1），（2,2），（3，2）拟合的直线方程

• 我们假设最优直线方程：b = C + Dt，代入三个点列出方程。

$$C + D = 1$$
$$C + 2D = 2$$
$$C + 3D = 2$$

• 列成矩阵方程 Ax = b 发现无解。这时就要采取我们在投影中讲到的方法来拟合这条曲线。上一节讲过，方程化为：求解 $A^T b = A^T A\hat{x}$。

△ 这个方法中的一个关键部分就是 Ax = b 本无解，但是经过我们处理得到 $A^T A\hat{x} = A^T b$ 却可以求出最优解。我们将一个无解的方程转换为了可以求解的最优方程。

最小二乘法下节还会详述。

四、学习感悟

这部分内容是上一节的延伸。通过正交我们开始计算投影，并利用投影与向量之间的偏差来运算最小二乘法，解决无解时候方程 AX = b 的最优解问题，借此拟合直线。这部分的实际应用性较强，主要目的就是解决最优解问题，计算 $A^T A\hat{x} = A^T b$。