Pure Mathematics 理论数学, 2013, 3, 68-71
http://dx.doi.org/10.12677/pm.2013.31011 Published Online January 2013 (http://www.hanspub.org/journal/pm.html)

Inner Product Space and the Least Square Method

Fei Li

Department of Mathematics and System Science, College of Science, National University of Defense Technology, Changsha
Email: ping1guo@163.com.cn

Received: Oct. 13th, 2012; revised: Oct. 27th, 2012; accepted: Nov. 16th, 2012

Abstract: The least square method has a standard method in most of books about linear algebra. In this paper we introduce the element definitions of inner product space and the standard least square method. And in this document we also give a fast calculation for the least square method in inner product space. In the least part we give a method for calculating the smallest distance of the sum of finite points to a subspace. This is an extension for the least square method.

Keywords: Least Square Method; Inner Product Space; Linear Approximation

内积空间与最小二乘法

李　非

国防科学技术大学理学院数学与系统科学系，长沙
Email: ping1guo@163.com.cn

收稿日期：2012 年 10 月 13 日；修回日期：2012 年 10 月 27 日；录用日期：2012 年 11 月 16 日

摘　要：最小二乘法在一般的线性代数中存在标准的推导过程。本文从内积空间的结构出发去发现最小二乘法的本质问题，并且给出了最小二乘法取值的快速算法。同时我们给出了在内积空间中多个点到子空间的最小距离的求法，这是对一般最小二乘的延托。

关键词：最小二乘法；内积空间；线性逼近

1. 引言

在一般有限维线性空间中，我们定义距离为传统的欧氏距离，由此我们可以求出一个向量到子空间的距离，即最小二乘法原理[1]。最小二乘法的理论在广义逆矩阵[2]计算中起重要作用，本文从内积空间出发去研究这个问题，得到更快的距离的算法[3-6]。

2. 最小二乘法的一般推导

令 K 为实数或者复数域，$V=K^n$ 是 n 维向量空间，我们可以在 V 上定义距离为一般的欧氏距离，即假定 $x=(x_1,x_2,\cdots,x_n), y=(y_1,y_2,\cdots,y_n)\in V$。定义

$$d(x,y)\left(\sum_{i=1}^{n}(x_i-y_i)^2\right)^{1/2}$$

为 V 上距离。显然如果我们首先在 V 上定义内积

$$(x,y)=\sum_{i=1}^{n}x_iy_i$$

则我们可以将 V 看作一完备内积空间，距离可以由内积导出，与我们原先定义的欧氏距离相协调。

考虑线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s = b_2 \\ \qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{ns}x_s = b_s \end{cases} \tag{1}$$

可能无解，即任何一组数 $(x_1, x_2, \cdots, x_n)$ 都可能使

$$\sum_{i=1}^{n}\left(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{is}x_s - b_i\right)^2 \tag{2}$$

等于零。于是我们设法找到这样的 $\left(x_1^0, x_2^0, \cdots, x_s^0\right)$ 使上述(2)值最小，这样的 $\left(x_1^0, x_2^0, \cdots, x_s^0\right)$ 称为方程组的最小二乘解。这种问题称为最小二乘问题。

最小二乘法取值 $\left(x_1^0, x_2^0, \cdots, x_s^0\right)$ 的一般推导方法可以在标准的线性代数教课书中看到，我们在这里不再进行具体推导。

我们利用欧氏空间的概念来表达最小二乘法，并给出最小二乘解所满足的代数条件。令

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ns} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix}, \quad Y = AX$$

应用距离的概念，(2)就是 $|Y - B|^2$。最小二乘法就是找 $\left(x_1^0, x_2^0, \cdots, x_s^0\right)$ 使 Y 与 B 的距离最短。把 A 的各列记为 $\alpha_1, \alpha_2, \cdots, \alpha_s$，记他们生成的子空间为 L。最小二乘法的实质就是寻找 X 使(2)的值最小，就是在 L 中找一向量 Y 使它到 B 的距离比 L 中其他向量到 B 的距离都小。

3. 完备内积空间上点到子空间的距离

我们考虑完备内积空间上点到子空间的距离。

3.1 定理 假定 K 为实数或复数域，E 为 K 上之希尔伯特空间，$x, y_i \in E \quad (1 \le i \le s)$ 有

$$\min\left\|x - \sum_{i=1}^{s} c_i y_i\right\| = \frac{G(x, y_1, y_2, \cdots, y_s)}{G(y_1, y_2, \cdots, y_s)} \tag{3}$$

其中，

$$G(y_1, y_2, \cdots, y_s) = \begin{vmatrix} (y_1, y_1) & \cdots & \cdots & (y_1, y_s) \\ (y_2, y_1) & \cdots & \cdots & (y_2, y_s) \\ \vdots & & & \vdots \\ (y_s, y_1) & \cdots & \cdots & (y_s, y_s) \end{vmatrix}$$

证明：令 $M = \left\{y, y = \sum_{i=1}^{s} c_i y_i, c_i \in K, i = 1, 2, \cdots, s\right\}$ 为有限为空间，则 M 为 E 的闭子空间，且 $E = M \oplus M^*$ (M^* 为 M 的正交补空间)。当 $x \in M$ 时，上述公式显然成立。下面假定 $x \in E - M$，欲使 x 与 y 有最短距离的充分必要条件是 $x - y$ 与 M 垂直。

故由正交分解定理。我们不妨假定 $x = y + z$，其中 $y \in M$， $z = x - y \perp M$。此时，有

$$\|x - y\| = \min\left\|x - \sum_{i=1}^{s} c_i y_i\right\| \tag{4}$$

由条件 $y \in M$，我们假定 $y = c_1 y_1 + c_2 y_2 + \cdots + c_s y_s$。由 $x - y \perp M$ 得到

$$\begin{cases} c_1(y_1, y_1) + c_2(y_1, y_2) + \cdots + c_s(y_1, y_s) = 0 \\ c_1(y_2, y_1) + c_2(y_2, y_2) + \cdots + c_s(y_2, y_s) = 0 \\ \vdots \\ c_1(y_s, y_1) + c_2(y_s, y_2) + \cdots + c_s(y_s, y_s) = 0 \end{cases} \tag{5}$$

令 $\delta^2 = \|x - y\|^2$，则

$$\delta^2 = (x - y, x - y) = (x, x) - (y, x)$$

即

$$c_1(y_1, x) + c_2(y_2, x) + \cdots c_s(y_s, x) = (x, x) - \delta^2 \tag{6}$$

联立方程(5)和(6)，看作(s+1)个系数的方程，有非零解 $(c_1, c_2, \cdots, c_s, 1)$，故必有

$$\begin{vmatrix} (y_1, y_1) & (y_2, y_1) & \cdots & (y_s, y_1) - (x, y_1) \\ \vdots & \vdots & \cdots & \vdots \\ (y_1, y_s) & (y_2, y_s) & \cdots & (y_s, y_s) - (x - y_s) \\ (y_1, x) & (y_2, x) & \cdots & (-x, x) \end{vmatrix} = 0$$

即

$$\delta^2 G(y_1, y_2, \cdots, y_s) = G(y_1, y_2, \cdots, y_s, x) = G(x, y_1, y_2, \cdots, y_s)$$

我们考虑希尔伯特空间上的最小二乘法，即考虑 $(Y - B)^2$ 之最小值，由上述定理可知，如果最小值存在，其必定为

$$\left(\frac{G(B, \alpha_1, \alpha_2, \cdots, \alpha_s)}{G(\alpha_1, \alpha_2, \cdots, \alpha_s)} \right)^{1/2} \tag{7}$$

在可分希尔伯特空间中，我们利用此最小距离反推 x 所满足的条件。

$$\frac{G(B, \alpha_1, \alpha_2, \cdots, \alpha_s)}{G(\alpha_1, \alpha_2, \cdots, \alpha_s)} = \frac{1}{|A'A|} \left((B, B)|A'A| + \sum_{i=1}^{s} (-1)^i (\alpha_i, B)|\boldsymbol{C}_i| \right)$$

其中 $\boldsymbol{C}_i$ 为一 $s \times s$ 矩阵，如果 $k, j \neq i$，k 行 j 列元素为 (α_k, α_j)，i 行元素为 (B, α_j)，i 列元素为 (α_k, B) 故我们可以得到

$$\frac{G(B, \alpha_1, \alpha_2, \cdots, \alpha_s)}{G(\alpha_1, \alpha_2, \cdots, \alpha_s)} = \frac{1}{|A'A|} \left((B, B)|A'A| - (A, B)(A'A)^{-}(A, B)' \right) = B'B - B'A(A'A)^{-} A'B$$

由于 $G(\alpha_1, \alpha_2, \cdots, \alpha_s) = A'A$ 在 K 中可逆，假定 $|Y - B|^2 = |AX - B|^2$ 在 x 处取得极小值，故有

$$|Ax - B|^2 = (Ax - B, Ax - B) = (B, B) - (Ax, B) = B'B - B'A(A'A)^{-} A'B$$

此时我们显然有

$$x = A(A'A)^{-} B \tag{8}$$

4. 希尔伯特空间上有限点到子空间的距离

考虑希尔伯特空间上 m 个点 $B_1, B_2, \cdots, B_m$ 到由点 $y_1, y_2, \cdots, y_s$ 生成的子空间 M 的距离的最小和，即

$$\min\sum_{i=1}^{m}\left|x-B_i\right|, x\in M \tag{9}$$

我们将 $y_1,y_2,\cdots,y_s$ 变为一组标准正交基，即

$$\left(y_1,y_2,\cdots,y_s\right)=\left(y_1',y_2',\cdots,y_s'\right)C$$

令 $x=c_1y_1'+c_2y_2'+\cdots+c_sy_s'$，对任意 $B_i=B_{i1}+B_{i2}\ \left(1\le i\le m\right)$，$B_{i1}\in M, B_{i2}\in M^*$ 故(11)式的值为

$$\sum_{i=1}^{m}\left|B_{i2}\right|+\min\sum_{i=1}^{m}\left|x-B_{i1}\right|$$

我们需求

$$\min\sum_{i=1}^{m}\left|x-B_{i1}\right| \tag{10}$$

由于

$$\begin{pmatrix}B_{11}\\B_{21}\\\vdots\\B_{m1}\end{pmatrix}=\begin{pmatrix}d_{11}&d_{12}&\cdots&d_{1s}\\d_{21}&d_{22}&\cdots&d_{2s}\\\vdots&\vdots&&\vdots\\d_{m1}&d_{m2}&\cdots&d_{ms}\end{pmatrix}\begin{pmatrix}y_1'\\y'\\\vdots\\y_s'\end{pmatrix}$$

$$\sum_{i=1}^{m}\left|x-B_{i1}\right|=\sum_{i=1}^{m}\sum_{j=1}^{s}\left|\left(c_j-d_{ij}\right)y_j\right|=\sum_{i=1}^{m}\sum_{j=1}^{s}\left|\left(c_j-d_{ij}\right)\right|$$

此时我们不妨假定 $C_j=\min\sum_{i=1}^{m}\left|c_j-d_{ij}\right|, c_j\in K$，由于 C_j 之间彼此没有影响，故我们求得(9)式的最小值为

$$\sum_{i=1}^{m}G\left(B_i,y_1',y_2',\cdots,y_s'\right)+\sum_{j=1}^{s}C_j \tag{11}$$

同时，假定 c_{ij} 按照绝对值的顺序递增，当 m 为基数时，$C_j=\sum_{i=1}^{(m-1)/2}\left|c_{(m-i-1)j}-c_{ij}\right|$；当 m 为偶数时，$C_j=\sum_{i=1}^{m/2-1}\left|c_{(m-i-1)j}-c_{ij}\right|$。

5. 结论

本文在内积空间中看待最小二乘法，并把有限点到子空间距离的一般算法推广到一般希尔伯特空间上。

6. 致谢

感谢教研室的全体同志在讨论班上对该文题的贡献，尤其感谢冯良贵老师及杨涌老师。

参考文献 (References)

[1] 黄有度, 狄成恩, 朱士信. 矩阵论及其应用[M]. 合肥: 中国科学技术大学出版社, 1995.
[2] 张贤达. 矩阵分析与应用[M]. 北京: 清华大学出版社, 2004.
[3] 王萼芳, 石生明. 高等代数(第三版)[M]. 北京: 高等教育出版社, 2010.
[4] 王声望, 郑维行. 实变函数与泛函分析概要(第三版)[M]. 北京: 高等教育出版社, 2006.

