DOI:10.13718/j.cnki.xsxb.2017.09.002

基于组合测量的一种新的最小二乘算法①

孙艳波

南京航空航天大学 金城学院，南京 211156

摘要： 利用线性互补问题，建立了一种基于组合测量的新的最小二乘算法，并在适当的条件下证明了算法的收敛性和唯一性. 最后，计算机仿真验证了算法的有效性.

关　键　词： 非负线性最小二乘问题；线性互补问题；组合测量

中图分类号： O151.21　　**文献标志码：** A　　**文章编号：** 1000-5471(2017)09-0006-05

最小二乘法(Least Squares Method，LSE)，最早是由德国数学家高斯于 18 世纪末在预测行星运行轨道时提出的. 基本思想是通过误差的平方和最小化，找出相关数据之间的联系，建立最佳匹配的函数关系. 其中线性最小二乘问题作为数据处理的有力工具，已被广泛应用于系统工程、控制理论、经济学等领域[1]. 尤其在测绘学科中，成为研究的热点问题之一[2]. 文献[3－4]阐述了 LSE 在误差分析和数据处理应用中的基本原理及算法. 文献[5－6]将 LSE 应用于小波测量及纳米颗粒粒径测量中. 文献[7]较详细地研究和总结了 LSE 在测绘学中的发展过程.

然而，在算法方面的研究还停留在一些经典的算法上，如正交化方法、SOR 方法、共轭梯度法[8]、并行算法、可行内点算法等[9－12]. 本文与传统的最小二乘算法构造思路不同，利用线性互补问题[13]与非负线性最小二乘问题的等价性，建立了一种不动点方程，在此基础上建立了算法，并对算法的收敛性进行了分析，最后实例表明新的算法应用于测绘学科中的组合测量相关问题是可行的.

1　基于组合测量的最小二乘问题

在组合测量中，为了确定 t 个不可直接测量的未知数 $x_1, x_2, \cdots, x_t$ 的估计值，可以建立与这 t 个未知数有函数关系的线性模型：

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1t}x_t \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2t}x_t \\ \cdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nt}x_t \end{cases}$$

简记为 $\boldsymbol{Y} = \boldsymbol{AX}$，其中 $\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{pmatrix}$ 为待求未知量，$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nt} \end{pmatrix}$ 为系数矩阵，$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ 为最小二乘估计量.

设 $\boldsymbol{L} = (l_1 l_2 \cdots l_t)^{\mathrm{T}}$ 为 $\boldsymbol{Y}$ 的测量值，含有误差 $\boldsymbol{V}$，则

---

① 收稿日期：2015-12-31
基金项目：江苏省自然科学基金项目(BK2011719)；国家自然科学基金项目(11101211).
作者简介：孙艳波(1979-)，女，内蒙古通辽人，副教授，主要从事变分不等式和互补问题研究.

$$\boldsymbol{V} = \boldsymbol{L} - \boldsymbol{Y} = \boldsymbol{L} - \boldsymbol{AX} \tag{1}$$

称为残余误差. 测量结果的最可信赖值应在残余误差平方和为最小的条件下求得，且在实际测量问题中多数测量为非负的，所以转化为非负线性最小二乘问题的模型，即

$$\min_{X \geqslant 0} \frac{1}{2} \| \boldsymbol{L} - \boldsymbol{AX} \|^2 \tag{2}$$

以上所介绍的问题就是本文要解决的问题.

2　非负线性最小二乘问题与线性互补问题的等价性

问题(2)按范数展开得：

$$\min_{\boldsymbol{X} \geqslant 0} q(\boldsymbol{X}) = \frac{1}{2} \boldsymbol{X}^{\mathrm{T}} (\boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}) \boldsymbol{X} - \boldsymbol{L}^{\mathrm{T}} \boldsymbol{AX} + \frac{1}{2} \boldsymbol{L}^{\mathrm{T}} \boldsymbol{L} \tag{3}$$

这是一个以对称矩阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为二阶 Hesse 矩阵的极小化问题，由于 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 至少半正定，所以(3)式的任何最优解也是全局最优解.

由 K-K-T 条件，问题(3)等价于：求 $\boldsymbol{X} \in \mathbb{R}^n$，使得

$$\boldsymbol{X} \geqslant 0,\ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{L} \geqslant 0,\ \boldsymbol{X}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{L}) = 0 \tag{4}$$

令 $\boldsymbol{M} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$，$\boldsymbol{q} = -\boldsymbol{A}^{\mathrm{T}}\boldsymbol{L}$，$F(\boldsymbol{X}) = \boldsymbol{MX} + \boldsymbol{q}$，则(4)式即为线性互补问题[13]，

$$\boldsymbol{X} \geqslant 0,\ F(\boldsymbol{X}) \geqslant 0,\ \boldsymbol{X}^{\mathrm{T}} F(\boldsymbol{X}) = 0 \tag{5}$$

从而问题(2)的求解可以转化为线性互补问题(5)的求解.

在 $\mathbb{R}^n_+$ 中，由投影映射的性质[14]，建立不动点方程：

$$\boldsymbol{X} = (\boldsymbol{X} - \tau F(\boldsymbol{X}))_+ \tag{6}$$

其中：τ 为任意正数，$\boldsymbol{X}_+ = \max\{\boldsymbol{X}, \boldsymbol{0}\} = (\max(x_1, 0), \max(x_2, 0), \cdots, \max(x_n, 0))^{\mathrm{T}}$. 使得方程(6)成立的点称为不动点.

下面的定理给出不动点方程产生的不动点 $\boldsymbol{X}^*$ 为线性互补问题(5)的解，也是问题(2)的解.

定理 1　$\boldsymbol{X}^*$ 为不动点方程(6)的解，则$(\boldsymbol{X}^*)_+$ 为非负线性最小二乘问题(2)的解.

证　设 $\boldsymbol{X}^*$ 为(6)式的解，即$(\boldsymbol{X}^* - \tau F(\boldsymbol{X}^*))_+ = \boldsymbol{X}^*$，由于

$$((\boldsymbol{X}^* - \tau F(\boldsymbol{X}^*))_+ - (\boldsymbol{X}^* - \tau F(\boldsymbol{X}^*)), \boldsymbol{Y}) \geqslant 0,\ \forall \boldsymbol{Y} \geqslant \boldsymbol{0}$$

$$((\boldsymbol{X}^* - \tau F(\boldsymbol{X}^*))_+ - (\boldsymbol{X}^* - \tau F(\boldsymbol{X}^*)), (\boldsymbol{X}^* - \tau F(\boldsymbol{X}^*))_+) = 0$$

从而

$$(\tau F(\boldsymbol{X}^*), \boldsymbol{Y}) \geqslant 0,\ \forall \boldsymbol{Y} \geqslant \boldsymbol{0} \text{ 且} (\tau F(\boldsymbol{X}^*), \boldsymbol{X}^*) = 0$$

即

$$F(\boldsymbol{X}^*) \geqslant \boldsymbol{0},\ (\boldsymbol{X}^*, F(\boldsymbol{X}^*)) = 0$$

又因为 $\boldsymbol{X}^* \geqslant \boldsymbol{0}$，所以 $\boldsymbol{X}^* = (\boldsymbol{X}^*)_+$ 为线性互补问题(5)的解，也是问题(2)的解.

3　算法及收敛性分析

以线性互补问题作为桥梁，利用非负线性最小二乘问题和不动点方程之间求解的关系，可以建立以下算法.

算法 1

给定控制误差 $\varepsilon > 0$.

Step1　取初始点 $\boldsymbol{X}^0 \in \mathbb{R}^n$，取一个常数 $\tau > 0$，令 $k := 0$.

Step2　令 $i = 1, 2, \cdots, n$. 计算 $y_i^k = x_i^k - \tau F_i(\boldsymbol{X}^k)$，得到 $\boldsymbol{Y}^k = (y_1^k, y_2^k, \cdots, y_n^k)^{\mathrm{T}}$.

Step3　令 $i = 1, 2, \cdots, n$. 计算 $x_i^{k+1} = \max\{0, y_i^k\}$，得到 $\boldsymbol{X}^{k+1} = (x_1^{k+1}, x_2^{k+1}, \cdots, x_n^{k+1})^{\mathrm{T}}$.

Step4　若 $\| \boldsymbol{X}^{k+1} - \boldsymbol{X}^k \| < \varepsilon$，停止. 否则令 $k := k + 1$，转 Step2.

下面我们对算法进行收敛性分析. 首先再明确几个定义.

定义 1[15]　若存在 $\alpha > 0$，使得映射 $F: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ 满足

$$(\boldsymbol{X}-\boldsymbol{Y},\ F(\boldsymbol{X})-F(\boldsymbol{Y}))\geqslant\alpha\|\boldsymbol{X}-\boldsymbol{Y}\|^2,\ \forall\ \boldsymbol{X},\boldsymbol{Y}\in\mathbb{R}^n,\ \boldsymbol{X}\neq\boldsymbol{Y}$$

则称 F 是 α -强单调映射.

定义 2[15] 若 $\beta>0$，使得映射 $F:\mathbb{R}^n\longrightarrow\mathbb{R}^n$ 满足

$$\|F(\boldsymbol{X})-F(\boldsymbol{Y})\|\leqslant\beta\|\boldsymbol{X}-\boldsymbol{Y}\|,\ \forall\ \boldsymbol{X},\boldsymbol{Y}\in\mathbb{R}^n,\ \boldsymbol{X}\neq\boldsymbol{Y}$$

则称 F 是 β - Lipschitz 连续的.

引理 1 函数 $F(\boldsymbol{X})=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{L}$，若 $\boldsymbol{A}$ 是列满秩的，那么函数 $F(\boldsymbol{X})$ 是 α -强单调映射，并且是 β - Lipschitz 连续的，其中 $\alpha=\lambda_{\min}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$，$\beta=\lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$，并且 $\lambda_{\min}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$，$\lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$ 分别表示 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的最小和最大特征值.

证 若 $\boldsymbol{A}$ 是列满秩的，那么 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是正定的. 由正定矩阵性质可知，矩阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的特征值都大于 0. 于是 $\forall\ \boldsymbol{X},\boldsymbol{Y}\in\mathbb{R}^n$，$\boldsymbol{X}\neq\boldsymbol{Y}$，有

$$\begin{aligned}(\boldsymbol{X}-\boldsymbol{Y},\ F(\boldsymbol{X})-F(\boldsymbol{Y}))=(\boldsymbol{X}-\boldsymbol{Y})^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{Y})\geqslant\\ \lambda_{\min}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\|\boldsymbol{X}-\boldsymbol{Y}\|^2=\alpha\|\boldsymbol{X}-\boldsymbol{Y}\|^2\end{aligned}$$

又因为

$$\begin{aligned}\|F(\boldsymbol{X})-F(\boldsymbol{Y})\|=\|\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{Y})\|\leqslant\\ \|\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\|\ \|\boldsymbol{X}-\boldsymbol{Y}\|=\|\boldsymbol{A}\|^2\|\boldsymbol{X}-\boldsymbol{Y}\|\leqslant\\ \lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\|\boldsymbol{X}-\boldsymbol{Y}\|=\beta\|\boldsymbol{X}-\boldsymbol{Y}\|\end{aligned}$$

从而 $F(\boldsymbol{X})$ 是 α -强单调映射，并且是 β - Lipschitz 连续的.

定理 2 假设 $F(\boldsymbol{X})=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 中 $\boldsymbol{A}$ 列满秩，若式(6)中 τ 满足 $\tau\in\left(0,\ \dfrac{2}{\beta}\right)$，其中 $\alpha=\lambda_{\min}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$，$\beta=\lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$，则存在 $\boldsymbol{X}^*\in\mathbb{R}^n$，满足等式 $\boldsymbol{X}=(\boldsymbol{X}-\tau F(\boldsymbol{X}))_+$，并且由算法产生的序列 $\{\boldsymbol{X}^k\}$ 在 $\mathbb{R}^n$ 中收敛到 $\boldsymbol{X}^*$.

证 因为 $\boldsymbol{A}$ 是列满秩的，所以矩阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是对称正定的，所有特征值都是正实数，记为 λ_i，$i=1,2,\cdots,n$，由引理 1 可知：

$$0<\alpha=\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n=\beta$$

由算法 1 可知：

$$\begin{aligned}\|\boldsymbol{X}^{k+1}-\boldsymbol{X}^k\|=\|(\boldsymbol{X}^k-\tau F(\boldsymbol{X}^k))_+-(\boldsymbol{X}^{k-1}-\tau F(\boldsymbol{X}^{k-1}))_+\|\leqslant\\ \|\boldsymbol{X}^k-\boldsymbol{X}^{k-1}-\tau[F(\boldsymbol{X}^k)-F(\boldsymbol{X}^{k-1})]\|=\\ \|\boldsymbol{X}^k-\boldsymbol{X}^{k-1}-\tau\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{X}^k-\boldsymbol{X}^{k-1})\|=\\ \|(I-\tau\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})(\boldsymbol{X}^k-\boldsymbol{X}^{k-1})\|\leqslant\\ \|\boldsymbol{I}-\tau\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\|\ \|\boldsymbol{X}^k-\boldsymbol{X}^{k-1}\|=\\ \max_{i=1,2,\cdots,n}(|1-\tau\lambda_i|)\|\boldsymbol{X}^k-\boldsymbol{X}^{k-1}\|\end{aligned}$$

令 $\theta=\max\limits_{i=1,2,\cdots,n}(|1-\tau\lambda_i|)$，则有

$$\|\boldsymbol{X}^{k+1}-\boldsymbol{X}^k\|\leqslant\theta\|\boldsymbol{X}^k-\boldsymbol{X}^{k-1}\|\leqslant\cdots\leqslant\theta^k\|\boldsymbol{X}^1-\boldsymbol{X}^0\|$$

因为 $\tau\in\left(0,\ \dfrac{2}{\beta}\right)$，所以 $1-\dfrac{2\lambda_i}{\beta}<1-\tau\lambda_i<1$，又 $\dfrac{\lambda_i}{\beta}<1$，所以 $-1<1-\tau\lambda_i<1$，即 $\theta=\max\limits_{i=1,2,\cdots,n}(|1-\tau\lambda_i|)<1$，由此推得级数 $\sum\limits_{k=0}^{\infty}(\boldsymbol{X}^{k+1}-\boldsymbol{X}^k)$ 收敛，即 $\{\boldsymbol{X}^k\}$ 是一个基本点列. 设 $\{\boldsymbol{X}^k\}$ 的极限为 $\boldsymbol{X}^*$，由点列 $\{\boldsymbol{X}^k\}$ 在算法 1 中产生的过程知：

$$\boldsymbol{X}^{k+1}=(\boldsymbol{X}^k-\tau F(\boldsymbol{X}^k))_+$$

两边同时令 $k\rightarrow\infty$ 取极限，可得 $\boldsymbol{X}^*=(\boldsymbol{X}^*-\tau F(\boldsymbol{X}^*))_+$，即 $\boldsymbol{X}^*$ 是方程(5)的解，也是问题(2)的解.

4 实例仿真

取文献[3]中的一个例子验证算法的有效性.

例 要检定刻线 A,B,C,D 间的距离 x_1,x_2,x_3. 为此，直接测量刻线间距的各种组合量(如图1)，得到

如下数据：

$l_1 = 1.015$ mm，$l_2 = 0.985$ mm，$l_3 = 0.985$ mm，

$l_4 = 2.016$ mm，$l_5 = 1.981$ mm，$l_6 = 3.032$ mm.

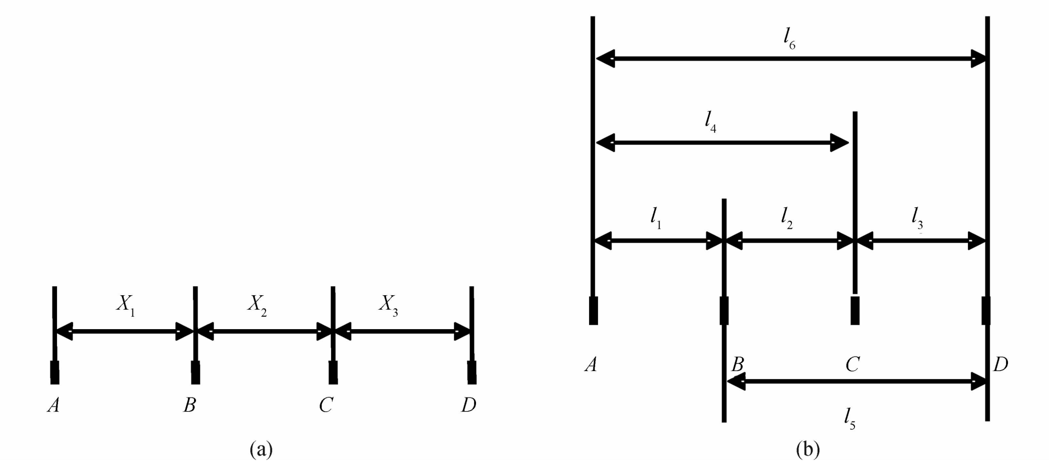

图 1　刻线图

按(1)式，残余误差方程为

$$\begin{cases} v_1 = l_1 - x_1 \\ v_2 = l_2 - x_2 \\ v_3 = l_3 - x_4 \\ v_4 = l_4 - (x_1 + x_2) \\ v_5 = l_5 - (x_2 + x_3) \\ v_6 = l_6 - (x_1 + x_2 + x_3) \end{cases}$$

则

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad \boldsymbol{L} = \begin{pmatrix} 1.015 \\ 0.985 \\ 1.020 \\ 2.016 \\ 1.981 \\ 3.032 \end{pmatrix} \qquad \boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

利用算法算出 AB，BC，CD 间距的最佳估计值，

$$\boldsymbol{X}^* = \begin{pmatrix} 1.028 \\ 0.983 \\ 1.013 \end{pmatrix}$$

符合参考文献[3]中的计算结果，表明此算法是有效的.

5　结束语

本文利用线性互补问题与非负线性最小二乘问题的等价性，建立了 LSE 的一个新的算法，证实了此算法在测绘学科中有一定的实用性. 新的算法计算简单、保稀疏、存储量小.

在算法的收敛性方面，本文只证明了对于 $\tau \in \left(0, \dfrac{2}{\beta}\right)$ 时，算法是收敛的，而对于更一般的情况算法能否收敛，将是下一阶段的研究工作.

参考文献：

[1]　陈希孺，王松桂. 线性模型中的最小二乘法 [M]. 上海：上海科学技术出版社，2003.

[2] 沈云中，陶本藻. 实用测量数据处理方法 [M]. 北京：测绘出版社，2000.
[3] 李金海. 误差理论与测量不确定度评定 [M]. 北京：中国计量出版社，2003.
[4] 吴石林，张 玘. 误差分析与数据处理 [M]. 北京：清华大学出版社，2010.
[5] 车金锐，梁 刚，陈敏茹. 基于小波测量预处理的最小二乘估计 [J]. 数学的实践与认知，2006，36(4)：98－101.
[6] 周述苍，杨燕婷，周誉昌. 纳米颗粒粒径测量的非负最小二乘算法分析与改进 [J]. 中国粉体技术，2009，15(5)：72－75.
[7] 王玉玮. 误差论 [M]. 北京：国防工业出版社，2014.
[8] 刘钦圣. 最小二乘问题计算方法 [M]. 北京：北京工业大学出版社，1988.
[9] 孙文瑜，徐成贤，朱德通. 最优化方法 [M]. 北京：高等教育出版社，2004.
[10] 王德人，白中治. 大型线性最小二乘问题的并行矩阵多分裂松弛算法 [J]. 高校计算数学学报，1993，3(1)：72－85.
[11] 杨志霞，张知难. 解线性最小二乘问题的一种新并行算法 [J]. 新疆大学学报(自然科学版)，2004，21(4)：370－376.
[12] 雍龙泉. 非负线性最小二乘问题的一个严格可行内点算法 [J]. 陕西理工学院学报，2010，26(4)：84－89.
[13] 韩继业，修乃华，戚厚铎. 非线性互补理论与算法 [M]. 上海：上海科技出版社，2006.
[14] 程其襄，张奠宇. 实变函数与泛函分析基础 [M]. 北京：北京工业大学出版社，2003.
[15] 殷洪友，徐成贤，张忠秀. F-互补问题及其与极小元问题的等价性 [J]. 数学学报，2001，4(4)：679－686.

A New Least Squares Algorithm Based On Combined Measurement

SUN Yan-bo

College of Jincheng, Nanjing University of Aeronautics and Astronautics, Nanjing 211156, China

Abstract: In this article, a new least squares algorithm based on combined measurement has been established by linear complementarity. And the convergence of the algorithm and the only under appropriate conditions have been proved. Finally, the computer simulation illustrated the feasibility and effectiveness.

Key words: the linear nonnegative least-squares problem; the complementarity problem; combined measurement

责任编辑 张 枸