文章编号:1671-8860(2013)05-0505-08 文献标志码:A

整体最小二乘估计的研究进展

刘经南[1,2] 曾文宪[2] 徐培亮[3]

(1 武汉大学卫星导航定位技术研究中心,武汉市珞喻路 129 号,430079)
(2 武汉大学测绘学院,武汉市珞喻路 129 号,430079)
(3 京都大学防灾研究所,京都府宇治市,611-0011)

摘 要:整体最小二乘估计方法作为经典最小二乘估计方法的扩展,近 20 年来被广泛地应用于信号处理、计算机视觉、图像处理、通信工程以及大地测量与摄影测量等测绘相关领域,成为各专业领域进行数据处理的基本方法。概述了整体最小二乘估计的发展历史,从整体最小二乘估计的算法、统计特性和可靠性研究三方面综述了整体最小二乘估计方法的研究进展情况,侧重强调各种算法的本质特征,并对整体最小二乘估计的研究方向进行了展望。

关键词:整体最小二乘平差; EIV 模型; 参数估计; 统计分析; 可靠性
中图法分类号:P207

经典 Gauss-Markov 模型假定函数模型已知、非随机,仅假定观测值向量包含随机误差。在许多实际问题中如数字地面模型拟合、大地测量反演、GIS 空间数据分析、滑坡监测、坐标变换等数学模型中,观测向量和描述函数模型的系数矩阵均由观测数据组成,两者都包含随机误差,这类平差模型称为 EIV(errors-in-variables)模型:

$$\boldsymbol{y} = (\boldsymbol{A} - \boldsymbol{E}_A)\boldsymbol{\beta} + \boldsymbol{e}_y \tag{1}$$

式中,$\boldsymbol{y}$ 表示 $n\times 1$ 的观测向量,$\boldsymbol{e}_y$ 是 $\boldsymbol{y}$ 中包含的 $n\times 1$ 随机误差向量;$\boldsymbol{A}$ 表示 $n\times m$ 系数矩阵,含 $n\times m$ 随机误差 $\boldsymbol{E}_A$,$\boldsymbol{e}_A = \mathrm{vec}\,\boldsymbol{E}_A$,即 $\boldsymbol{e}_A$ 是将 $\boldsymbol{E}_A$ 按列向量化后得到的 $nm\times 1$ 向量;$\boldsymbol{\beta}$ 为 $m\times 1$ 的待估参数向量。记 $\boldsymbol{e} = [\boldsymbol{e}_y, \boldsymbol{e}_A]^{\mathrm{T}}$,并假定 $\boldsymbol{e}$ 的均值为零,方差协方差阵为 $\sigma_0^2\boldsymbol{Q}$,其中 σ_0^2 表示单位权方差,$\boldsymbol{Q}$ 表示观测数据的协因数阵,相应权阵为 $\boldsymbol{W}$。

由于 EIV 模型的系数矩阵包含误差,经典最小二乘估计方法不再具有无偏性和方差最小等统计特性。为寻求 EIV 模型的最优估计,Adcock[1] 通过将最小二乘估计准则进行一般化的扩展,首次提出了所有观测数据(包括观测向量和系数矩阵)的残差的平方和最小化的整体最小二乘平差准则。Adcock[2] 和 Kummell[3] 分别推导了直线拟合模型的整体最小二乘解。Pearson[4] 导出了 EIV 模型的严密的整体最小二乘解。上述文献中,整体最小二乘估计被直接作为一个具有一般性的最小二乘问题进行研究。Deming[5-7] 称这种方法为广义最小二乘(general least squares),整体最小二乘估计从而作为区别于最小二乘估计被独立提出,同时推导了真正统计意义上的整体最小二乘解。Gerhold[8] 研究了整体最小二乘解的统计特性,导出了整体最小二乘解的一阶近似精度评定公式,并且提出了整体最小二乘的另一专业术语 complete least squares。作为对 Pearson[4] 工作的扩展,Golub 和 Van Loan[9] 提出了著名的 SVD(singular value decomposition)算法,即奇异值分解算法,并采用 total least squares(TLS)命名整体最小二乘估计方法,该术语自此成为整体最小二乘估计方法的通用术语,从而统一了不同领域及不同文献中的名称。由于采用奇异值分解相对容易得到 EIV 模型的整体最小二乘解,因此,自 SVD 算法提出以后,整体最小二乘估计引起了各领域的广泛关注。Van Huffel 和 Vandewalle[10] 出版了整体最小二乘估计方法的专著,进一步促进了各领域对整体最小二乘估计方法的研究工作。

尽管大地测量领域最早开展整体最小二乘估计的研究,但是,自 20 世纪 80 年代至 21 世纪初,

收稿日期:2013-03-18。
项目来源:国家自然科学基金资助项目(41231174,41174012,41274022)。

主要在大地坐标系统转换等应用中引入了整体最小二乘方法，如 Wolf[11]、刘经南等[12-14]、Soler[15]、Grafarend[16]等。真正研究适用于大地测量领域的整体最小二乘估计方法始于 2005 年 Schaffrin 和 Felus[17]发表的带有等式约束的整体最小二乘估计方法（equality-constrained TLS，ECTLS）研究成果，之后 Schaffrin[18-21]陆续发表了加权整体最小二乘估计（weighted TLS，WTLS）、多元整体最小二乘估计（multivariate TLS，MTLS）等研究成果。Shen[22]等提出了基于 Newton-Gauss 迭代法的加权整体最小二乘算法。Mahboub[23]研究了系数矩阵呈现规律性的结构的加权整体最小二乘估计（structuxed TLS，STLS）。以上算法中假定的权矩阵通常为某种特殊形式，Fang[24]进一步研究了权矩阵为一般情况下的普遍适用的加权整体最小二乘算法，并讨论了一般性权矩阵下的 STLS、系数矩阵部分元素非随机情况下的 TLS 算法、MTLS、非线性 TLS（nonlinear TLS）等算法。随着 EIV 模型在应用中的扩展，Zhang[25]等研究了附有不等式约束的整体最小二乘算法（inequality-constrained TLS，ICTLS）；Choi[26]等研究了稳健整体最小二乘算法（robust TLS）；Schaffrin[27]和 Snow[28]提出了整体最小二乘配置算法（TLS-collocation）等。针对系数矩阵中随机和固定元素同时存在的一般情况，Xu[29]等提出采用 partial-EIV 模型作为 EIV 模型更为一般的表示形式并推导了基于 partial-EIV 模型的整体最小二乘解，并从理论上推导了有限样本情况下的整体最小二乘解的精度评定严密计算公式。EIV 模型的可靠性方面仅 Schaffrin 和 Uzun[30-31]进行了研究工作。可以看到，目前整体最小二乘算法的研究成果较为丰富，但对整体最小二乘估计统计特性的研究成果相对来说还十分有限。

本文从整体最小二乘算法、统计特性和可靠性三个方面系统地综述了整体最小二乘估计方法的研究进展，分析了各种算法的本质特征，并提出了整体最小二乘估计的若干研究方向。

1 整体最小二乘算法

1.1 经典整体最小二乘算法

1980 年，Golub 和 van Loan[9]采用奇异值分解方法导出了 EIV 模型的整体最小二乘解，随后该方法在各专业领域得到了广泛应用。实际上早在 1901 年，Pearson[4]就提出了基于正交回归的估计思想，并证明整体最小二乘解等价于最小特征向量，与 Golub 和 van Loan[9]的 SVD 解本质上完全一致，EIV 模型这类基于数值逼近理论的求解算法被称为经典整体最小二乘算法（Markovsky 和 van Huffel[32]）。

1.1.1 Pearson 算法基本思想

从任意维空间的超平面 $\Omega(\boldsymbol{x},\boldsymbol{y})$ 的最佳拟合问题出发，Pearson[4]分析了由于第 i 个数据点中的因变量 y_i 和自变量 $(x_1,x_2\cdots x_m)_i$ 均为包含随机误差的观测值，因此，$\Omega(\boldsymbol{x},\boldsymbol{y})$ 一个好的估计值（最优估计值）不再仅仅满足数据点在 y 方向上至 $\Omega(\boldsymbol{x},\boldsymbol{y})$ 的距离的平方和最小的经典最小二乘准则，应采用 n 个数据点各分量到 $\Omega(\boldsymbol{x},\boldsymbol{y})$ 的距离的平方和最小即整体最小二乘准则：

$$\min:S(d)=\sum_{i=1}^{n}s_i^2 \tag{2}$$

以二维平面中的直线拟合为例，整体最小二乘平差准则的几何意义如图 1 所示。

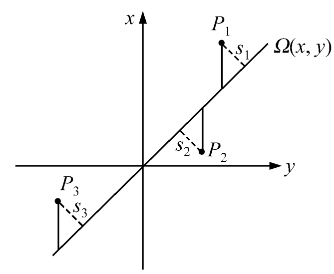

图 1 整体最小二乘平差准则的几何意义（根据 Pearson[4]修改）

Fig. 1 Geometrical Illustration of the Principle of Total Least Squares

图 1 中，s_i 表示数据点到拟合直线 $\Omega(\boldsymbol{x},\boldsymbol{y})$ 的垂直距离。Pearson[4]证明了整体最小二乘解 $\Omega(\boldsymbol{x},\boldsymbol{y})$ 必定通过数据点的重心，并且准则(2)下的整体最小二乘解 $\Omega(\boldsymbol{x},\boldsymbol{y})$ 为模型最小特征值对应的特征向量。

1.1.2 SVD 算法

SVD 算法的平差准则为：

$$\min:S=\|\boldsymbol{D}[\boldsymbol{E}_A,\boldsymbol{e}_y]\boldsymbol{T}\|_F \tag{3}$$

式中，$\boldsymbol{D}=\mathrm{diag}(d_1,\cdots d_n)$ 和 $\boldsymbol{T}=\mathrm{diag}(t_1,\cdots t_{m+1})$ 为权对角阵，且 d_i 和 t_i 均为正常数；$\|\cdot\|_F$ 表示 Frobenius 范数。

根据平差准则(3)，Golub 和 Van Loan[9]导出的 SVD 算法如下。

1）对 EIV 模型的增广矩阵 $[\boldsymbol{A},\boldsymbol{y}]$ 进行奇异值分解：

$$\boldsymbol{C}=\boldsymbol{D}[\boldsymbol{A},\boldsymbol{y}]\boldsymbol{T}=\boldsymbol{U\Sigma V} \tag{4}$$

式中，$\underset{n,n}{\boldsymbol{U}}=[\boldsymbol{U}_1,\cdots\boldsymbol{U}_n]$；$\underset{m+1,m+1}{\boldsymbol{V}}=[\boldsymbol{V}_1,\cdots\boldsymbol{V}_{m+1}]$；

$\underset{n,m+1}{\boldsymbol{\Sigma}}=\mathrm{diag}(\sigma_1,\cdots\sigma_{m+1})$ 且 $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_{m+1}$。

2) 若 $\sigma_m>\sigma_{m+1}$,存在唯一的整体最小二乘解:

$$\hat{\boldsymbol{\beta}}_{\mathrm{TLS}}=-\frac{\boldsymbol{V}_{m+1}}{\boldsymbol{V}_{m+1,m+1}} \tag{5}$$

式(5)表明,整体最小二乘的 SVD 解对应于最小特征值相应的特征向量,与 Pearson[4] 算法得到的整体最小二乘解一致。

Golub 和 van Loan[9] 分析了 SVD 解的扰动性,定义$(\sigma_m-\sigma_{m+1})^{-1}$作为衡量算法敏感性的指标,得出了整体最小二乘估计的条件数通常大于相应的最小二乘估计的条件数的结论,并且对 EIV 模型的 TLS 解和 LS 解进行了比较。van Huffel 和 Vandewalle[33] 采用对解空间附加约束的方法得到了无解情况下($\sigma_m=\sigma_{m+1}$)满足式(4)准则的更普遍适用的 SVD 算法。Gleser[34] 研究了多元 EIV 模型的整体最小二乘算法。van Huffel 和 Vandewalle[10] 采用先做 QR 分解再进行 SVD 分解的方法解决了广义整体最小二乘问题。

1.1.3　经典整体最小二乘算法评述

Xu[29] 等分析了 Pearson[4] 算法和 SVD 算法的本质特性,指出式(4)中的 $\boldsymbol{T}$ 仅对参数和观测向量起到线性变换的作用,即经过线性变换后 $\boldsymbol{T}$ 完全可以从式中移除,可表示为:

$$\min:S=\|\boldsymbol{D}[\boldsymbol{E}_A,\boldsymbol{e}_y]\boldsymbol{T}\|_F=\sum_{i=1}^{n}d_i^2 s_i^2 \tag{6}$$

如果把 d_i 视为权,式(6)说明 SVD 算法实质是将 Pearson 算法中的数据点至拟合超平面的距离的平方和最小的准则扩展到了距离的加权平方和最小,同样说明了 SVD 算法与 Pearson 算法具有一致性。

式(2)表明,Pearson 算法并没有顾及数据点的精度信息;同样,尽管 SVD 算法中式(4)的 $\boldsymbol{D}$ 和 $\boldsymbol{T}$ 被称为权阵,但其对角线元素的值并不包含观测数据的精度信息,并非真正的权阵,除非在特殊情况下刚好等于观测数据的权[29]。即 SVD 算法与 Pearson 算法都没有顾及 EIV 的随机模型,仅仅是基于数值逼近理论的整体最小二乘解,并非真正统计意义上的整体最小二乘解。因此,尽管 SVD 算法简单,但显然无法获得大地测量领域普遍存在的观测数据不等精度以及相关情况下的平差模型的统计意义上的最佳估计值,这也是自 SVD 算法提出以来在大地测量领域的应用受到限制的主要原因。

1.2　加权整体最小二乘算法(WTLS)

现实世界中由于数据来源、获取手段和获取方法等存在差异性,观测数据存在等精度、独立和不等精度、相关等各种情况,因此,在考虑 EIV 随机模型的情况下,加权整体最小二乘算法 WTLS 的研究由最初特殊的仅考虑不等精度独立观测值、不等精度相关观测值、观测向量与系数矩阵不相关直到研究最一般的观测向量与系数矩阵相关情况下的 WTLS 算法。

1.2.1　WTLS 算法基本进展

对于 EIV 模型(1),加权整体最小二乘估计的平差准则为:

$$\min:S=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{e}=\begin{bmatrix}\boldsymbol{e}_y\\ \boldsymbol{e}_A\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q}_y & \boldsymbol{Q}_{yA}\\ \boldsymbol{Q}_{Ay} & \boldsymbol{Q}_A\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{e}_y\\ \boldsymbol{e}_A\end{bmatrix} \tag{7}$$

如果观测向量 $\boldsymbol{y}$ 和系数矩阵 $\boldsymbol{A}$ 的协因数阵 $\boldsymbol{Q}$ 为对称正定阵,WTLS 算法的基本思想是在给定的平差准则下,根据拉格朗日乘数法原理构建 EIV 模型的目标函数:

$$\Phi=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{e}+2\lambda^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{A}\boldsymbol{\beta}+\boldsymbol{E}_A\boldsymbol{\beta}-\boldsymbol{e}_y) \tag{8}$$

目标函数分别对待估量求偏导并令其为 0,从而可以求得 EIV 模型的统计意义上的最优估计值。由于方程为非线性,必须采用迭代解法得到最终的参数解。

WTLS 算法的发展可视为协因数阵 $\boldsymbol{Q}$(权阵 $\boldsymbol{W}$)从特殊到一般的发展过程。1931 年 Deming[5] 针对所有观测数据不相关情况下的 EIV 模型($\boldsymbol{Q}$ 为对角阵)首次导出了 WTLS,这也是首次得到了真正统计意义上的整体最小二乘解。在 Deming[5] 的假定条件下,式(7)简化为:

$$\min:S=\sum_{j=1}^{mn}\omega_j e_{Aj}^2+\sum_{i=1}^{n}w_i e_{yi}^2 \tag{9}$$

对于非线性的 EIV 模型,Deming 运用泰勒级数将其在$(\boldsymbol{A},\boldsymbol{y},\boldsymbol{\beta}_0)$进行了线性化,从而将非线性模型的条件极值问题变成了线性模型的条件极值,再采用拉格朗日乘数法求得了 EIV 模型的 WTLS 解。

Gerhold[8] 对 Deming 的线性化解法进行了详细的分析,认为当近似值$(\boldsymbol{A},\boldsymbol{y},\boldsymbol{\beta}_0)$不够好的情况下,整体最小二乘解会有较大的偏差,并运用拉格朗日乘数法得到了协因数阵为对角阵情况下的 WTLS 解。

Schaffrin 和 Wieser[20] 研究了基于元素相关的加权 WTLS 算法,假定 $\boldsymbol{Q}$ 形式如下:

$$\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{Q}_y & 0\\ 0 & \boldsymbol{Q}_A=\underset{m,m}{\boldsymbol{Q}_0}\otimes\underset{n,n}{\boldsymbol{Q}_X}\end{bmatrix} \tag{10}$$

式(10)表明系数矩阵 $\boldsymbol{A}$ 的协因数阵的任意两列

仅相差一个比例系数，$\boldsymbol{Q}_A$实际上由 $m\times m$ 个对称方阵$\boldsymbol{Q}_X(n\times n)$倍数组成。笔者进一步讨论了当系数矩阵 $\boldsymbol{A}$ 中某列为不含随机误差的固定值(如坐标转换模型中平移参数对应的系数矩阵的列)的情况，这时可将$\boldsymbol{Q}_0$ 中对应的行和列的值取0，尽管$\boldsymbol{Q}_0$ 为奇异阵，但只要$[\boldsymbol{Q}_y+(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{Q}_0\boldsymbol{\beta})\boldsymbol{Q}_X]^{-1}$存在就可得到 WTLS 解。显然 Schaffrin 和 Wieser[20]为便于公式推导而提出的协因数阵不具一般性。

Xu[29]等在仅假定系数矩阵和观测值不相关并且 $\boldsymbol{Q}$ 为非奇异阵的情况下导出了 WTLS 解，即式(7)中的协因素阵 $\boldsymbol{Q}$ 形式如下：

$$\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{Q}_y & 0\\ 0 & \boldsymbol{Q}_A\end{bmatrix} \tag{11}$$

式(11)对系数矩阵 $\boldsymbol{A}$ 中元素的精度和相关性等无限制，但对 $\boldsymbol{Q}$ 非奇异性的要求限制了当 $\boldsymbol{A}$ 中出现固定值或者结构性特征情况下算法的应用，同一文献中，Xu[29]等推导的基于 partial-EIV 模型算法得到了更一般的 WTLS 解。

Fang[24]在对系数矩阵和观测向量的精度和相关性没有限制的一般情况下系统地研究了 WTLS 算法，给定协因素阵 $\boldsymbol{Q}$ 一般形式：

$$\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{Q}_y & \boldsymbol{Q}_{yA}\\ \boldsymbol{Q}_{Ay} & \boldsymbol{Q}_A\end{bmatrix} \tag{12}$$

Fang[24]导出了 WTLS 算法3种不同形式的解并列出了相应的算法。为避免 $\boldsymbol{Q}$ 矩阵奇异，Fang[24]提出了当系数矩阵 $\boldsymbol{A}$ 中某些列为固定值或者当某些元素为固定值时一般 WTLS 算法的改进形式，基本思想是将 $\boldsymbol{A}$ 表示成分别包含固定列(固定元素)和包含随机元素的矩阵维数相同的两部分($\boldsymbol{A}=\boldsymbol{A}_1+\boldsymbol{A}_2$)再进行求解。

Snow[28]通过在解方程中引入正定或者半正定的对称方阵得到了当 $\boldsymbol{BQB}^{\mathrm{T}}$($\boldsymbol{B}=[\boldsymbol{\beta}^{T}\otimes\boldsymbol{I}_n,-I_n]$)为奇异矩阵时的一般情况下的 WTLS 算法。对于系数矩阵呈现结构性特征的情况，Mahboub[23]通过对$\boldsymbol{Q}_A$进行处理将 STIS 问题归入到 WTLS 方法解算。

以上得到了一般情况下的 WTLS 算法，但是算法中均包含了 m 个参数 $\boldsymbol{\beta}$、n 个观测值误差改正量$\boldsymbol{e}_y$以及 mn 个系数矩阵误差改正量$\boldsymbol{E}_A$。相对于经典的最小二乘估计而言，WTLS 算法属于含大量待估量的非线性估计，因此，WTLS 在实际应用中尤其是在海量数据处理中的一个关键问题是如何简化算法以提高算法的有效性。

1.2.2 基于 Partial-EIV 模型的 WTLS 算法

当系数矩阵 $\boldsymbol{A}$ 中出现非随机性的固定元素的情况下，协因素阵 $\boldsymbol{Q}$ 奇异，必须对 $\boldsymbol{Q}$ 进行特殊处理才能求得唯一的 WTLS 解。针对这类情况，Xu[29]等将 EIV 模型改写成 partial-EIV 模型形式：

$$\begin{aligned}\boldsymbol{y}&=(\boldsymbol{\beta}^{\mathrm{T}}\otimes\boldsymbol{I}_n)(\boldsymbol{h}+\boldsymbol{B}\bar{\boldsymbol{a}})+\boldsymbol{e}_y\\ \boldsymbol{a}&=\bar{\boldsymbol{a}}+\boldsymbol{e}_A\end{aligned} \tag{13}$$

式中，$(\boldsymbol{h}+\boldsymbol{B}\bar{\boldsymbol{a}}):=\mathrm{vec}(\boldsymbol{A}-\boldsymbol{E}_A)$。

Partial-EIV 模型的基本思想是将随机的系数矩阵的向量化形式 $\mathrm{vec}(\boldsymbol{A}-\boldsymbol{E}_A)$分成两部分：第一部分 h 是由 $\mathrm{vec}(\boldsymbol{A}-\boldsymbol{E}_A)$中非随机项构成的向量；第二部分 $\boldsymbol{B}\bar{\boldsymbol{a}}$ 是由 $\mathrm{vec}(\boldsymbol{A}-\boldsymbol{E}_A)$中随机项构成的向量，其中 $\bar{\boldsymbol{a}}$ 仅仅为待估计的系数矩阵中的随机元素，$\boldsymbol{B}$ 是由常数构成的 $\bar{\boldsymbol{a}}$ 的系数矩阵。Xu[29]等采用自由极值的拉格朗日乘数法导出基于 partial-EIV 模型的 WTLS 计算公式。

基于 partial-EIV 模型的 WTLS 算法对一般 WTLS 算法的改进主要体现在以下方面：① partial-EIV 模型是 EIV 模型更为一般的表达形式，涵括了一般 WTLS 算法需要特殊处理的各种情况，如部分元素为非随机量导致的 $\boldsymbol{Q}$ 奇异，$\boldsymbol{A}$ 呈现结构性特征的 STLS 等；② 对待估计量采用了分离的估计方法，算法每次迭代的效率由一般 WTLS 算法的 $O([t+m]^3)$提高到了 $O(\max[t,m]^3)$，其中 t 表示 $\boldsymbol{A}$ 中随机元素个数；③ partial-EIV 模型的系数矩阵误差改正估计量个数 t 要小于等于 EIV 模型相应的估计量个数 $m\times n$，尤其当系数矩阵非随机量的个数较多时，模型的待估量大大减少；④ 模型算法的简单形式大大便利了后续估计值精度的评定。

1.3 其他整体最小二乘扩展算法

在(加权)整体最小二乘算法的基础上，针对应用中的需求发展了附有约束的整体最小二乘、非线性整体最小二乘、多元整体最小二乘、整体最小二乘配置、正则化整体最小二乘、稳健整体最小二乘等扩展算法。

1) 附有约束的整体最小二乘算法(ECTLS 和 ICTLS)。当模型存在某些先验信息或者需要设定一定的限制条件时，我们必须对 EIV 模型附加约束方程。约束方程分为等式和不等式两类。Dowling[35]等在线性约束子空间的预测频率估计技术中研究了含等式 EIV 模型的 ECTLS 算法，Schaffrin[18]推导了含随机约束值但约束方程被认为是非随机的等权条件下的 ECTLS 解，Schaffrin 和 Felus[21]研究了附有线性约束等式和二次型等式约束的等权条件下 EIV 模型的 ECTLS 解。Fang[24]得到了一般情况下的附有线性等式

约束的 WTLS。ECTLS 实质是在整体最小二乘准则下求解下列模型的最优估计值：

$$\begin{cases} \boldsymbol{y} = (\boldsymbol{A} - \boldsymbol{E}_A)\boldsymbol{\beta} + \boldsymbol{e}_y \\ \boldsymbol{K\beta} = \boldsymbol{Z} \end{cases} \tag{14}$$

与 TLS 求解思路相同，采用拉格朗日乘数法即可导出 ECTLS 解。

Zhang[25] 等研究了等权条件下附有不等式约束的 EIV 模型的 ICTLS 算法，基本方法是采用穷举搜索法将不等式约束方程转变为等式约束进行求解，该算法随着不等式约束方程数目的增大计算量显著增长，适用性受到一定的限制。

2）非线性 EIV 模型整体最小二乘算法。当 EIV 模型为非线性形式时的估计问题称为非线性 TLS 估计。Schwetlick 和 Tiller[36] 讨论了非线性 TLS 的数值解法，Fang[24] 推导了一般情况下基于拉格朗日乘数法和 Gauss-Newton 法的附有约束的加权 non-linear TLS 算法。

3）多元整体最小二乘算法（MTLS）。将 EIV 模型中的观测向量 $\boldsymbol{y}$ 和参数向量 $\boldsymbol{\beta}$ 扩展为观测矩阵 $\boldsymbol{Y}$ 和参数矩阵 $\boldsymbol{\Xi}$：

$$\boldsymbol{Y} - \boldsymbol{E}_Y = (\boldsymbol{A} - \boldsymbol{E}_A)\boldsymbol{\Xi} \tag{15}$$

以上模型估计称为多元整体最小二乘估计。van Huffel 和 Vandewalle[10] 研究了基于 SVD 分解的 MTLS 算法。Schaffrin 和 Felus[37] 讨论了基于 SVD 和拉格朗日乘数的等权条件下 MTLS 的 3 种具体算法并进行了比较。Fang[24] 推导了一般情况下的加权 MTLS 算法。

4）整体最小二乘配置算法。Schaffrin[27] 提出当考虑 EIV 模型待估参数 $\boldsymbol{\beta}$ 的先验统计信息时（$\boldsymbol{\beta}$ 视为随机量），相应模型称为 EIV-REM 模型（errors-in-variables with random effects model）。Snow[28]] 采用拉格朗日乘数法推导了一般权矩阵条件下的 EIV-REM 模型的 TLS 解。

5）正则化整体最小二乘算法（regularized TLS）。当 EIV 模型的系数矩阵病态时会导致 TLS 解的不稳定，Golub[38] 等提出了 Tykhonov 正则化方法获得 TLS 的稳定解。Shaffrin 和 Snow[39] 指出 Tykhonov 正则化方法通过在 TLS 解中引入一个二次型约束尽管从数值角度上分析具有等价性，但并不具备统计上的等价性。Beck 和 Ben-Tal[40]、Pruessner 和 O'Leary[41] 研究了改进的 Tykhonov 正则化 TLS 算法。Schaffrin 和 Felus[21] 研究了基于拉格朗日乘数法的 Tykhonov 正则化 TLS 改进算法。

6）稳健整体最小二乘算法。稳健整体最小二乘算法基本沿用了最小二乘的稳健估计理论和方法，如 Choi[26] 等设计了基于权的影响函数，Waston[42] 研究了变量较少情况下将模型转化为凸最优化问题进行求解。

2　整体最小二乘统计特性

EIV 模型的整体最小二乘估计方法目前在算法研究上取得了较为丰富的成果，但对整体最小二乘估计统计特性的研究尚未引起足够的重视，研究成果相当有限。Gleser[34]、van Huffel 和 Vandewalle[10] 以及 Markovsky 和 van Huffel[32] 等证明了当观测数据个数趋于无穷大时 EIV 模型的整体最小二乘估计具有渐进无偏性。Golub 和 van Loan[9]、Markovsky 和 van Huffel[32]、刘永辉和魏木生[43] 等从数值分析的角度对 EIV 模型的 TLS 解和 LS 解进行了概略的比较。Fang[24] 导出了 EIV 模型 WTLS 和 LS 参数估计值的差异的计算公式。

尽管 TLS 解被证明具有渐进无偏性，但是在现实的有限样本条件下，模型的非线性会导致 TLS 解产生偏差，偏差的大小取决于模型的非线性程度（Xu[29]），因此有限样本情况下整体最小二乘解的统计特性研究更具实际意义。Gerhold[8] 采用 Deming[5] 线性化的思路首次得到了权矩阵为对角阵的 WTLS 解的一阶近似精度估计公式。Schaffrin[20] 对 TLS 的非线性解运用误差传播公式导出了等权 TLS 解的精度近似公式。以上精度评定没有考虑估计值偏差的影响，Schaffrin[18] 直接对非线性的 TLS 解运用误差传播技术理论上并不严密。

Xu[29] 等系统地研究了基于有限样本的一般情况下的 WTLS 解的统计特性，对非线性 EIV 模型运用线性化技术推导了 WTLS 估计值的精度、偏差、置信区间以及单位权方差估计的严密公式。TLS 参数估计值的一阶近似精度公式为：

$$D(\hat{\boldsymbol{\beta}}) = (\boldsymbol{N}_{11} - \boldsymbol{N}_{11}\ \boldsymbol{N}_{22}\ \boldsymbol{N}_{21})^{-1}\sigma^2 \tag{16}$$

式中，$\boldsymbol{N}_{ij}$ 是根据观测数据和权阵计算的对称方阵；$\boldsymbol{N}_{11}^{-1}$ 表示对应的最小二乘解的协因数阵。

Xu[29] 指出由于有限样本的 WTLS 为有偏估计，显然直接采用残差求得的单位权方差必然有偏，因此，Xu[29] 通过偏差改正技术求出估计残差 $\boldsymbol{r}_A$ 和 $\boldsymbol{r}_y$ 及改正值 $\boldsymbol{r}_{AC}$ 和 $\boldsymbol{r}_{yC}$，进而推出了 WTLS 单位权方差的一个改善估计值：

$$\hat{\sigma}^2 = \frac{\boldsymbol{r}_{AC}^{\mathrm{T}}\boldsymbol{\omega}_A\boldsymbol{r}_{AC} + \boldsymbol{r}_{yC}^{\mathrm{T}}\boldsymbol{W}_y\boldsymbol{r}_{yC}}{n - m} \tag{17}$$

以上精度等公式是基于一般情况下的 Partial-EIV 模型,因此具有普遍的适用性。

3 EIV 模型的可靠性

Baarda[44]首次提出了 Gauss-Markov 模型的基于假设检验的可靠性理论,之后在应用中得到了进一步扩展(如 Pelzer[45]、李德仁[46]、Koch[47]等)。Schaffrin 和 Uzun[30-31]在假定仅存在单个粗差且观测向量和系数矩阵中各列的协因数阵相同和不相关的特殊条件下初步探讨了 EIV 模型的可靠性理论。

基于 Schaffrin[48]提出的 Gauss-Markov 模型的内部和外部可靠性的思想,Schaffrin 和 Uzun[30-31]在忽略 TLS 解的迭代扰动项的情况下,分别推导了当观测向量和系数矩阵存在单个粗差时的 EIV 模型的内部可靠性指标 $\boldsymbol{r}_E$和外部可靠性计算公式 $\bar{\delta}_E$。如当系数矩阵存在单个粗差时:

$$r_{\boldsymbol{EA}} = [\boldsymbol{\eta}_j^{\mathrm{T}}(\boldsymbol{W}_n\boldsymbol{Q}_e\boldsymbol{W}_n)\boldsymbol{\eta}_j][\boldsymbol{\eta}_j^{\mathrm{T}}\boldsymbol{W}_n(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\otimes\boldsymbol{I}_n)(\boldsymbol{\eta}_k\otimes\boldsymbol{\eta}_j)]^{-1}(1+\hat{\boldsymbol{\beta}}^{\mathrm{T}}\hat{\boldsymbol{\beta}})^{-1}[(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\hat{\boldsymbol{\beta}})(\boldsymbol{\eta}_k^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{(jk)})] \tag{18}$$

$$\bar{\delta}_{\boldsymbol{EA}} = (\boldsymbol{\eta}_j^{\mathrm{T}}\boldsymbol{W}_n\boldsymbol{\eta}_j)\{1-[\boldsymbol{\eta}_j^{\mathrm{T}}(\boldsymbol{W}_n\boldsymbol{Q}_e\boldsymbol{W}_n)\boldsymbol{\eta}_j][\boldsymbol{\eta}_j^{\mathrm{T}}\boldsymbol{W}_n(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\otimes\boldsymbol{I}_n)(\boldsymbol{\eta}_k\otimes\boldsymbol{\eta}_j)]^{-1})\boldsymbol{\eta}_k^{\mathrm{T}}\hat{\boldsymbol{\beta}}\}[\boldsymbol{\eta}_k^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{(jk)}]^2 \tag{19}$$

式中,$\boldsymbol{Q}_e$表示观测值改正数的协因数阵;$\boldsymbol{W}_n$ 表示观测向量(系数矩阵中的各列)的协因数阵;$\boldsymbol{\eta}_j$ 和 $\boldsymbol{\eta}_k$ 表示单位向量;$\hat{\boldsymbol{\beta}}$ 表示不含粗差时的参数估计值;$\hat{\boldsymbol{\beta}}^{(jk)}$ 表示当系数矩阵第 j 行第 k 列元素包含粗差时的参数估计值。

经典的平等模型的可靠性理论是根据系数矩阵和观测值的先验信息(协因数阵或权阵)计算出模型可发现粗差的下界值以及不可发现粗差对平差结果的影响。由于系数矩阵仅与模型的几何或物理结构有关(如控制网的结构),因而在确定平差模型后不需要进行实际观测即可根据系数矩阵和观测值先验信息计算出模型的可靠性指标,从而在设计阶段就可以对平差模型设计方案进行评价和改进。Schaffrin 和 Uzun[30-31]推导的可靠性公式中包含了参数估计值,但并没有阐明可靠性公式为何与待估参数解有关的原因,我们认为这个问题涉及如何实际上预估 EIV 模型的可靠性,还有待进一步研究与探索。

4 结 语

整体最小二乘方法本质上是对经典最小二乘方法的扩展。对于 EIV 模型,在基于同时顾及观测向量和系数矩阵误差的平差准则下,整体最小二乘解比最小二乘解更真实。

整体最小二乘算法研究经历了从基于数值逼近理论的 SVD 等经典算法等到统计意义下的加权整体最小二乘算法。同时,通过对 EIV 模型的不断扩展和精化,由假定系数矩阵元素全部包含误差的 EIV 模型的 TLS 算法扩展到基于一般形式的 partial-EIV 模型的 TLS 算法,由等权条件下的 TLS 算法发展到一般权矩阵下的 WTLS 算法,由线性 EIV 模型扩展到非线性 EIV 模型,由仅考虑随机误差的 TLS 估计扩展到同时考虑粗差的稳健 TLS 估计,由待估参数为固定值的 TLS 估计扩展到考虑待估参数先验信息的 TLS 配置算法等等。整体最小二乘的算法研究在理论上已经取得了较为丰富的成果,但由于整体最小二乘属于非线性估计,模型和算法的复杂性要远高于最小二乘估计,因此,在应用上尚受到一定的限制,如何进一步简化算法和提高算法的效率是今后整体最小二乘估计算法研究的重要目标。

Xu[29]等推导的有限样本情况下的加权整体最小二乘解的精度评定公式具有普遍适用性。相对于 EIV 模型的已经取得的 TLS 各种扩展算法研究成果,如非线性 TLS 解、附有约束的 TLS 解等的精度分析尚待研究。另外,对于 EIV 模型可靠性指标的定义、公式的推导以及一般情况下(如含多个粗差)可靠性评价的方法等有待下一步开展系统的研究工作。

参 考 文 献

[1] Adcock R J. Note on the Method of Least Squares [J]. Analyst, 1877,4:183-184

[2] Adcock R J. A Problem in Least Squares[J]. Analyst, 1878, 5: 53-54

[3] Kummell C H. Reduction of Observation Equations Which Contain more than One Observed Quantity [J]. Analyst, 1879, 6: 97-105

[4] Pearson K. On Lines and Planes of Closest Fit to Systems of Points in Space[J]. Phil Mag, 1901, 2: 559-572

[5] Deming W E. The Application of Least Squares[J]. Phil Mag, 1931, 11: 146-158

[6] Deming W E. On the Application of Least Squares-II[J]. Phil Mag, 1934, 17: 804-829

[7] Deming W E. Statistical Adjustment of Data[M]. New York:Dover Publications, 1964

[8] Gerhold G A. Least-squares Adjustment of Weighted Data to a General Linear Equation[J]. Am J Phys, 1969,37:156-161

[9] Golub G H, Van Loan C F. An Analysis of the Total Least Squares Problem[J]. SIAM J Numer, Anal, 1980, 17: 883-893
[10] Van Huffel S, Vandewalle J. The Total Least Squares Problem: Computational Aspects and Analysis[M]. Philadelphia:SIAM, 1991
[11] Wolf H. Scale and Orientation in Combined Doppler and Triangulation Nets[J]. Bull, Geod, 1980, 54: 45-53
[12] 刘经南. 卫星网与地面网联合平差坐标转换模型的等价性[J]. 武汉测绘学院学报, 1983,8(1):37-50
[13] 刘经南. 刘大杰,大地坐标和地心坐标精度对联合平差的精度影响[J]. 测绘学报, 1985, 14(2):133-144
[14] 刘经南,刘大杰,崔希璋. 卫星网与地面网联合平差的理论和应用[J]. 武汉测绘科技大学学报,1987, 12(4):1-9
[15] Soler T. A Compendium of Transformation Formulas Useful in GPS Work[J]. J Geod, 1998, 72: 482-490
[16] Grafarend E W. Nonlinear Analysis of the Three-dimensional Datum Transformation [J]. J Geod, 2003, 77: 66-76
[17] Schaffrin B, Felus Y A. On Total Least-squares Adjustment with Constraints [J]. IAG-Symp, 2005,128: 417-421
[18] Schaffrin B. A Note on Constrained Total Least-squares Estimation[J]. Linear Algebra Appl, 2006, 417: 245-258
[19] Schaffrin B, Lee I, Choi Y, et al. Total Least-squares (TLS) for Geodetic Straight-line and Plane Adjustment[J]. Boll Geod Sc Aff, 2006,65: 141-168
[20] Schaffrin B, Wieser A. On Weighted Total Least-squares Adjustment for Linear Regression[J]. J Geod, 2008, 82: 415-421
[21] Schaffrin B, Felus Y A. An Algorithmic Approach to the Total Least-squares Problem with Linear and Quadratic Constraints[J]. Stud Geo-phys Geod, 2009, 53: 1-16
[22] Shen Y Z, Li B F, Chen Y. An Iterative Solution of Weighted Total Least Squares Adjustment[J]. J Geod, 2010, 85: 229-238
[23] Mahboub V. On Weighted Total Least-squares for Geodetic Transformation[J]. J Geod, 2012, 86: 359-367
[24] Fang X. Weighted Total Least Squares Solutions for Applications in Geodesy[D]. Germany:Leibniz University Hannover, 2011
[25] Zhang S L, Tong X H, Zhang K l. A Solution to EIV Model with Inequality Constraints and Its Geodetic Applications[J]. J Geod, 2013, 87: 89-99
[26] Choi Y J, Kim J Y, Sumg K M. A Robust Algorithm of Total Least Squares Method[J]. IEICE TRANS, Fundamentals, 1997, 7(E80-A): 1 336-1 339
[27] Schaffrin B. TLS-collocation: The Total-least Squares Approach to EIV-models with Stochastic Prior Information[C]. The 18th Intl Workshop Matr Stat Smolenice Castle, Slovakia, 2009
[28] Snow K. Topics in Total Least-squares Adjustment Within the Errors-in-variables Model: Singular Cofactor Matrices and Prior Information[D]. American:The Ohio State University, 2012
[29] Xu P L, Liu J N, Shi C. Total Least Squares Adjustment in Partial Errors-in-variables Models: Algorithm and Statistical Analysis[J]. J Geod, 2012, 86: 661-675
[30] Schaffrin B, Uzun S. Errors-in-variables for Mobile Mapping Algorithms in the Presence of Outliers[J]. Archives of Photogra, Cartog Rem Sen, 2011, 22: 377-387
[31] Schaffrin B, Uzun S. On the Reliability of Errors-in-variables Models[J]. Acta et Commentationes Universitatis Tartuensis de Mathematica, 2012, 16: 69-81
[32] Markovsky I, Van Huffel S. Overview of Total Least-squares Methods[J]. Signal Proc, 2008, 87: 2 283-2 302
[33] Van Huffel S, Vandewalle J. Analysis and Solution of the Nongeneric Total Least Squares Problem[J]. SIAM J Matrix Anal Appl, 1988, 9(3): 360-372
[34] Gleser L J. Estimation in a Multivariate "Errors in Variables" Regression Model: Large Sample Results[J]. Ann Statist,1981, 9(1):24-44
[35] Dowling E M, DeGroat R D, Linebarger D A. Total Least Squares with Linear Constrains[J]. ICASSP-92,1992, 5: 341-344
[36] Schwetlick H, Tiller V. Numerical Methods for Estimating Parameters in Nonlinear Models with Errors in the Variables[J]. Technometrics, 1985, 27 (1): 17-24
[37] Schaffrin B, Felus Y A. On the Multivariate Total Least-squares Approach to Empirical Coordinate Transformations: Three Algorithms[J]. J Geod, 2008, 82(6):373-383
[38] Golub G H, Hanse P C, O'Leary D P. Tikhonov Regularization and Total Least Squares[J]. SIAMJ Matrix Anal Appl, 1991, 21(1): 185-194
[39] Schaffrin B, Snow K. Total Least-Squares Regularization of Tykhonov Type and an Ancient Racetrack in Corinth[J]. Linear Algebra Appl, 2010, 432

(8)：2 061-2 076

[40] Beck A，Ben-Tal A. On the Solution of the Tikhonov Regularization of the Total Least Squares[J]. SIAM J Optimization，2006，17(1)：98-118

[41] Pruessner A，O'Leary D P. Blind Deconvolution Using a Regularized Structured Total Least Norm Algorithm[J]. SIAM J Matrix Anal Appl，2003，24(4)：1 018-1 037

[42] Waston G A. Robust Counterparts of Errors-in-variables Problems[J]. Comp Stat Data Analy，2007，52(2)：1 080-1 089

[43] 刘永辉，魏木生. TLS 和 LS 问题的比较[J]. 计算数学，2003，25(4)：479-492

[44] Baarda W. A Testing Procedure for Use in Geodetic Networks[R]. Neth Geodetic Comm，Netherlands，1968

[45] Pelzer H. Some Criteria for the Accuracy and Reliability of Network[R]. DGK Reihe B Nr 252，Germany，1980

[46] 李德仁. 误差处理和可靠性理论[J]. 北京：测绘出版社，1988

[47] Koch K R. Parameter Estimation and Hypothesis Testing in Linear Models[M]. Berlin：Springer-Verlag，1999

[48] Schaffrin B. Reliability Measures for Correlated Observations[J]. J Surv Eng，1997，123(3)：126-137

第一作者简介：刘经南，中国工程院院士，教授，博士生导师，长期从事大地测量理论及应用研究，包括现代测量数据处理理论与应用、卫星定位理论与应用、软件开发和重大工程应用等方面。

E-mail：jnliu@whu.edu.cn

Overview of Total Least Squares Methods

LIU Jingnan[1,2] *ZENG Wenxian*[2] *XU Peiliang*[3]

(1 Research Center of GNSS，Wuhan University，129 Luoyu Road，Wuhan 430079，China)

(2 School of Geodesy and Geomatics，Wuhan University，129 Luoyu Road，Wuhan 430079，China)

(3 Disaster Prevention Research Institute，Kyoto University，Uji，Kyoto 611-0011，Japan)

Abstract：Total least squares (TLS) is a basic estimation method to account for random errors in functional models and has found a wide variety of applications in different areas of science and engineering，including signal and image processing，computer vision，communication engineering and our own subject area (geodesy，photogrammetry，geomatics and GIS). The purpose of this paper is to briefly review TLS methods and algorithms，including a discussion of the accuracy of TLS estimates. Since reliability is of interest in our subject area，we will also briefly touch the reliability issue of TLS. Finally，we will outline some topics for further investigations in the future.

Key words：total least squares；EIV model；parameter estimate；statistic analysis；reliability

About the first author：LIU Jingnan，academy of Chinese Academy of Engineering，professor，Ph. D supervisor. He is an expert in geodesy and surveying engineering with the specialty of GNSS technology and applications. He has been engaged in the research of data processing，GNSS technology and software development，as well as large project implementation.

E-mail：jnliu@whu.edu.cn