Weighted total least squares problems with inequality constraints solved by standard least squares theory

Jian Xie[1], Sichun Long[1]

1. Hunan Province Key Laboratory of Coal Resources Clean-utilization and Mine Environment Protection, Hunan University of Science and Technology, Taoyuan Road, Xiangtan 411201, China

1. Research contents:

Since the weighted total least squares problem with inequality constraints (ICWTLS) is nonlinear, it is solved either by linearizing the function model or by nonlinear programming. The existing algorithms are computationally expensive and not familiar to geodesists. A new iterative method is proposed based on standard least squares (SLS) adjustment. Compared with linear approximation (LA) method and sequential quadratic programming (SQP), the SLS method is simple and easy to implement.

2. The ICPEIV model and SLS algorithm

The inequality constrained partial errors-in-variables model is,

$$\boldsymbol{y} = \left(\boldsymbol{\beta}^{\mathrm{T}} \otimes \boldsymbol{I}_n\right)\left(\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}\right) + \boldsymbol{e}_y \tag{1a}$$

$$\boldsymbol{a} = \bar{\boldsymbol{a}} + \boldsymbol{e}_a \tag{1b}$$

$$\boldsymbol{G\beta} - \boldsymbol{d} \geq 0 \tag{1c}$$

$$D\left(\boldsymbol{e}_y\right) = \sigma_0^2 \boldsymbol{W}^{-1}, \quad D\left(\boldsymbol{e}_a\right) = \sigma_0^2 \boldsymbol{\omega}^{-1} \tag{1d}$$

The WTLS solution is the minimum of the nonlinear programming,

$$\left.\begin{aligned} &\min: \Phi\left(\bar{\boldsymbol{a}}, \boldsymbol{\beta}\right) = \left(\bar{\boldsymbol{a}} - \boldsymbol{a}\right)^{\mathrm{T}} \boldsymbol{\omega}\left(\bar{\boldsymbol{a}} - \boldsymbol{a}\right) + \\ &\left(\left(\boldsymbol{\beta}^{\mathrm{T}} \otimes \boldsymbol{I}_n\right)\left(\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}\right) - \boldsymbol{y}\right)^{\mathrm{T}} \boldsymbol{W}\left(\left(\boldsymbol{\beta}^{\mathrm{T}} \otimes \boldsymbol{I}_n\right)\left(\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}\right) - \boldsymbol{y}\right) \\ &s.t. \quad \boldsymbol{G\beta} - \boldsymbol{d} \geq 0 \end{aligned}\right\} \tag{2}$$

Both the model parameters and random coefficients are estimated simultaneously with the LA and SQP methods. We separately estimate two kinds of parameters. Based on the Kuhn-Tucker condition, if an estimate of $\boldsymbol{\beta}$ is given (it is assumed to have been estimated), say $\hat{\boldsymbol{\beta}}$, then we can estimate $\bar{\boldsymbol{a}}$ by either of the two formulae,

$$\hat{\bar{\boldsymbol{a}}} = \left(\boldsymbol{\omega} + \boldsymbol{S}_\beta^{\mathrm{T}} \boldsymbol{W} \boldsymbol{S}_\beta\right)^{-1} \times \left(\boldsymbol{\omega a} - \boldsymbol{S}_\beta^{\mathrm{T}} \boldsymbol{W} \boldsymbol{h}_\beta + \boldsymbol{S}_\beta^{\mathrm{T}} \boldsymbol{W y}\right) \tag{3a}$$

$$\hat{\bar{\boldsymbol{a}}} = \boldsymbol{a} + \boldsymbol{\omega}^{-1} \boldsymbol{S}_\beta^{\mathrm{T}} \boldsymbol{E}^{-1}\left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}\right) \tag{3b}$$

Where $\boldsymbol{h}_\beta = \sum_{i=1}^{m} \boldsymbol{h}_i \hat{\boldsymbol{\beta}}_i$, $\boldsymbol{S}_\beta = \sum_{i=1}^{m} \boldsymbol{B}_i \hat{\boldsymbol{\beta}}_i = \left(\hat{\boldsymbol{\beta}}^{\mathrm{T}} \otimes \boldsymbol{I}_n\right)\boldsymbol{B}$ and $\boldsymbol{E} = \boldsymbol{W}^{-1} + \boldsymbol{S}_\beta \boldsymbol{\omega}^{-1} \boldsymbol{S}_\beta^{\mathrm{T}}$. When the number of $\bar{\boldsymbol{a}}$ is less than that of observables, (3a) is recommended as it is computationally efficient. Otherwise, (3b) is a better choice. The equation $\left(\hat{\boldsymbol{\beta}}^{\mathrm{T}} \otimes \boldsymbol{I}_n\right)\left(\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}\right) = \bar{\boldsymbol{A}}\hat{\boldsymbol{\beta}}$ proves true (where $\bar{\boldsymbol{A}} = Ivec\left(\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}\right)$), so the target function (2) can be rewritten as,

$$\begin{cases} \min \phi\left(\boldsymbol{\beta}\right) = \left(\boldsymbol{a} - \bar{\boldsymbol{a}}\right)^{\mathrm{T}} \boldsymbol{\omega}\left(\boldsymbol{a} - \bar{\boldsymbol{a}}\right) + \left(\bar{\boldsymbol{A}}\boldsymbol{\beta} - \boldsymbol{y}\right)^{\mathrm{T}} \boldsymbol{W}\left(\bar{\boldsymbol{A}}\boldsymbol{\beta} - \boldsymbol{y}\right) \\ s.t. \quad \boldsymbol{G\beta} - \boldsymbol{d} \geq \boldsymbol{0} \end{cases} \tag{4}$$

When $\hat{\bar{\boldsymbol{a}}}$ is obtained from (3), (4) is equivalent to the QP problem,

$$\begin{cases} \min \phi\left(\boldsymbol{\beta}\right) = \left(\bar{\boldsymbol{A}}\boldsymbol{\beta} - \boldsymbol{y}\right)^{\mathrm{T}} \boldsymbol{W}\left(\bar{\boldsymbol{A}}\boldsymbol{\beta} - \boldsymbol{y}\right) \\ s.t. \quad \boldsymbol{G\beta} - \boldsymbol{d} \geq \boldsymbol{0} \end{cases} \tag{5}$$

The Kuhn-Tucker condition for (5) is as follows,

$$\begin{cases} \boldsymbol{\lambda} \geq 0 \text{ and } \boldsymbol{D\lambda} - \boldsymbol{l} \geq 0 \\ \boldsymbol{\lambda}^T\left(\boldsymbol{D\lambda} - \boldsymbol{l}\right) = 0 \end{cases} \tag{6}$$

Where $\boldsymbol{D} = \boldsymbol{G}\boldsymbol{N}^{-1}\boldsymbol{G}^{\mathrm{T}}$, $\boldsymbol{N} = \bar{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{W}\bar{\boldsymbol{A}}$, $\boldsymbol{l} = \boldsymbol{d} - \boldsymbol{G}\boldsymbol{N}^{-1}\bar{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{W y}$ and $\boldsymbol{\lambda}$ is the Lagrange multipliers. We designed a modified Jacobi iterative algorithm to (6). If the k-th iteration of $\boldsymbol{\lambda}$ is $\boldsymbol{\lambda}^{(k)} = \left(\lambda_1^{(k)}, \lambda_2^{(k)}, \cdots \lambda_s^{(k)}\right)^{\mathrm{T}}$, we solve $\boldsymbol{D\lambda} = \boldsymbol{l}$ by Jacobi iterative method. If the components of some $\boldsymbol{\lambda}$ are nonnegative, keep it unchanged. If some other $\boldsymbol{\lambda}$ are negative and we assign them to 0. Repeat until the last two solutions are close enough and $\boldsymbol{\lambda}$ is obtained. Thus $\hat{\boldsymbol{\beta}}$ can be calculated from $\boldsymbol{\lambda}$. Afterwards, the estimator $\hat{\boldsymbol{\beta}}$ is used as a new approximation to start the next iteration. Stop until the norm of the difference of last two solutions is within a given threshold.

3. Numerical examples

The examples are cited from "Fang X, On Non-combinatorial Weighted Total Least Squares with Inequality Constraints. J. Geod, 2014, 88(8):805-816". The inequality constrained EIV model is transformed into model (1).

Tab.1 Comparison of computing efficiency

	11 constraints			3 constraints		
	A	B	C	A	B	C
$\hat{\boldsymbol{\beta}}_{LA}$	23	2	0.038	7	2	0.021
$\hat{\boldsymbol{\beta}}_{SQP}$	5	2	0.020	4	2	0.018
$\hat{\boldsymbol{\beta}}_{SLS}$	22	132	0.025	14	27	0.009

A-Ouer iterations, B-Average inner iterations, C-Computing time (seconds)

For brevity we don't display the parameter estimates and they are almost the same. In Tab.1, when the number of constraints is large, the computation efficiency is not better than the SQP method. Although the number of iterations is much more than the former two, the computing time is short since each iteration is much simpler. When all the box constraints are abandoned and only 3 constraints remained, the computing efficiency improves as the number of constraints reduces. So the SLS method is feasible and efficient especially when the number of constraints is small.