UNIVERSAL FORMULAE OF ROBUST ESTIMATES OF PARAMETER VECTOR AND VARIANCE COMPONENTS[①]

Wang Zhizhong
Institute of Surveying Engineering and Land Information, Central South University of Technology, Changsha 410083

ABSTRACT The results of the robust estimates depend mainly on their weight functions. But nearly all weight functions are empirical and their parameters are determined personally. A new approach of determining the weight functions of the robust estimates has been presented and the universal formulae of the robust estimaters of parameter vector and variance components were established according to statistics.

Key words variance component robust estimate adjustment model

1 INTRODUCTION

Since Krarup and Kubik introduced the robust estimate into surveying data processing in 1967[2], many approaches of the robust estimates have been found and widely used in surveying adjustment. But these robust estimates have not strictly theoretical bases. To solve these problems, Zhu Jianjun[1, 2] gave the robust estimate with minimum mean squared error (MSE) and with minimum mean Cook distance respectively. However, Zhu Jianjun only considered the robust estimate of parameter vector in the adjustment model. The paper generalizes the result of Zhu Jianjun and establishes the universal formulae of the robust estimates of parameter vector and variance components with minimum MSE and mininum mean Cook distance according to the principle of statistics.

2 ROBUST ESTIMATES

Consider the following adjustment model

$$l_{ij} = \boldsymbol{a}_{ij}^{\mathrm{T}}\boldsymbol{X} + e_{ij} \qquad \begin{matrix} i = 1, 2, \cdots, k \\ j = 1, 2, \cdots, n_i \end{matrix} \tag{1a}$$

or

$$E(\boldsymbol{L}) = \boldsymbol{AX} \tag{1b}$$

$$\mathrm{Cov}(\boldsymbol{L}) = \Sigma = \mathrm{diag}[\sigma_1^2 p_{11}, \cdots \sigma_1^2 p_{1n_1}, \cdots, \sigma_k^2 p_{k1}, \cdots, \sigma_k^2 p_{kn_k}] \tag{1c}$$

where $l_{ij} \in R$ is the jth observation of ith group $\boldsymbol{L}_i$, $\boldsymbol{a}_{ij}^{\mathrm{T}} \in R^t$ is the jth row of ith group a_i of the design matrix $\boldsymbol{A}_{n\times t}$ with rank$(\boldsymbol{A}) = t$, $n_1 + n_2 + \cdots + n_k = n$.

$\boldsymbol{X} \in R^t$ is t-vector of unknown parameter, $\sigma_i^2 > 0$ is the ith unknown variance component. The observation error equations are

$$\boldsymbol{V} = \boldsymbol{L} - \boldsymbol{A}\hat{\boldsymbol{X}} \quad \text{or} \quad v_{ij} = l_{ij} - \boldsymbol{a}_{ij}^{\mathrm{T}}\hat{\boldsymbol{X}}$$

where v_{ij} is the jth residual value of the ith group $\boldsymbol{V}_i$ of the residual vector $\boldsymbol{V}$ and $\hat{\boldsymbol{X}}$ is the estimate of $\boldsymbol{X}$.

If the observation $\boldsymbol{L}$ is assumed to be normally distributed, the likelihood function $f(\boldsymbol{L}, \boldsymbol{X}, \sigma_i^2)$ of the observation $\boldsymbol{L}$ with unknown parameters $\boldsymbol{X}$ and σ_i^2 is given by

$$f(\boldsymbol{L}, \boldsymbol{X}, \sigma_i^2) = \frac{1}{(2\pi)^{n/2}(\det\Sigma)^{1/2}} \cdot \exp\left[-\frac{1}{2}(\boldsymbol{L} - \boldsymbol{AX})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{L} - \boldsymbol{AX})\right] \tag{2}$$

Using classical maximum likelihood approach to estimate $\boldsymbol{X}$ and σ_i^2 may often be carried

---

① Supported by the National Doctorate Program Fund of the State Education Committee of China
Received Oct. 15, 1996; accepted Apr. 27, 1997

out by setting the partial derivatives with respect to $\boldsymbol{X}$ and σ_i^2, respectively, of the log-likelihood equations to zero, and solving the resulting likelihood equations. Letting $\psi(z_{ij}) = \rho'(z_{ij})$ and $\rho(z_{ij}) = z_{ij}^2$, these equations may be expressed as follows,

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\psi(z_{ij})\frac{\boldsymbol{a}_{ij}}{\sigma_i\sqrt{p_{ij}}} = \boldsymbol{0} \tag{3}$$

$$\sum_{j=1}^{n_i}\psi(z_{ij})z_{ij} - n_i = 0,$$
$$i = 1, 2, \cdots, k \tag{4}$$

where

$$z_{ij} = \frac{l_{ij} - \boldsymbol{a}_{ij}\boldsymbol{X}}{\sigma_j\sqrt{p_{ij}}} \quad \begin{matrix} i = 1, 2, \cdots, k \\ j = 1, 2, \cdots, n_i \end{matrix}$$

If there are outliers or blunders in the model (1), many estimates are better than the classical maximum likelihood estimates[3]. Now, the universal formulae of all robust estmates of $\boldsymbol{X}$ and σ_i^2, are used in surveying adjustment. Firstly, generalize Eqn. (3) and Eqn. (4), then get

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\psi(z_{ij})\frac{\boldsymbol{a}_{ij}}{\sigma_i\sqrt{p_{ij}}} = \boldsymbol{0} \tag{5}$$

$$\sum_{j=1}^{n_i}\psi_i(z_{ij})z_{ij} - b_i = 0$$
$$i = 1, 2, \cdots, k \tag{6}$$

where b_i is a positive number ($i = 1, 2, \cdots, k$). $\psi(z_{ij}) = \rho'(z_{ij})$ and $\psi_i(z_{ij}) = \rho'_i(z_{ij})$, ($i = 1, 2, \cdots, k$) are suitable functions satisfying:

(i) ρ and ρ_i ($i = 1, 2, \cdots, k$) are symmetric, continuously differential and $\rho(0) = 0$ and $\rho_i(0) = 0$ ($i = 1, 2, \cdots, k$).

(ii) There exist $c > 0$ and $c_i > 0$ ($i = 1, 2, \cdots, k$) so that ρ and ρ_i ($i = 1, 2, \cdots, k$) are strictly increasing on $[0, c]$ and $[0, c_i]$ ($i = 1, 2, \cdots, k$) respectively and that ψ and ψ_i ($i = 1, 2, \cdots, k$) are constant on (c, ∞) and (c_i, ∞) ($i = 1, 2, \cdots, k$) respectively. Eqn. (5) and Eqn. (6) can be rewritten as

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\psi(z_{ij})}{z_{ij}}\frac{\boldsymbol{a}_{ij}v_{ij}}{\sigma_i^2\sqrt{p_{ij}}} = \boldsymbol{0} \tag{7}$$

$$\sum_{j=1}^{n_i}\frac{\psi_i(z_{ij})}{z_{ij}}\frac{v_{ij}^2}{\sigma_i^2\sqrt{p_{ij}}} - b_i = 0$$
$$i = 1, 2, \cdots, k \tag{8}$$

where $b_{ij} = \dfrac{\psi(z_{ij})}{z_{ij}}$ and $c_{ij} = \dfrac{\psi_i(z_{ij})}{z_{ij}}$ are called weight factors[4].

Then the universal formulae of the robust estimates of $\boldsymbol{X}$ and σ_i^2 can be obtained

$$\hat{\boldsymbol{X}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{L} \tag{9}$$

$$\hat{\sigma}_i^2 = \frac{1}{\sigma_i}\sum_{j=1}^{b_i}\frac{c_{ij}}{p_{ij}}v_{ij}^2$$
$$i = 1, 2, \cdots, k \tag{10}$$

where

$$\boldsymbol{P} = \mathrm{diag}\left[\frac{b_{11}}{\sigma_1^2 p_{11}}, \cdots, \frac{b_{1n_1}}{\sigma_1^2 p_{1n_1}}, \cdots, \frac{b_{k_1}}{\sigma_k^2 p_{1k_1}}, \cdots, \frac{b_{kn_k}}{\sigma_k^2 p_{kn_k}}\right]$$

Obviously, the robust estimate $\hat{\boldsymbol{X}}$ has the same form as the least squares estimate of $\boldsymbol{X}$ and $\hat{\sigma}_i^2$ is plus weight squared sums of the residual value v_{ij} of the ith group residual vector $\boldsymbol{V}_i$. Differences among the robust estimates are approaches of the weight functions determined. Therefore the development of the weight functions of the robust estimates is of great importance.

3 ROBUST ESTIMATES WITH MINIMUM MSE

The robust estimates depend on their weight functions. But almost weight functions are empirical, therefore the robust estimates $\hat{\boldsymbol{X}}$ and $\hat{\sigma}^2$ are determined personally, that is, all robust estimates can be considered as empirical estimates. Zhu Jianjun[1, 2] presented that the weight functions of the robust estimates should be determined according to statistics. Similarly to the proof of Refs. [1] and [2], we can give the robust estimate of $\boldsymbol{X}$ with MSE and minimum mean Cook distance.

$$\hat{\boldsymbol{X}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{L}_{\delta} \tag{11}$$

where $\mathrm{Cov}(\boldsymbol{L}_{\delta}) = \sigma_0^2\boldsymbol{P}^{-1} = \sum_{i=1}^{k}\sigma_i^2\bar{\boldsymbol{P}}_i$,

$$\boldsymbol{P}_i = \mathrm{diag}\left[0, \cdots, 0, \frac{1}{\bar{p}_{i1}}, \cdots, \frac{1}{\bar{p}_{in_i}}, 0, \cdots, 0\right],$$

$$\bar{\boldsymbol{P}}_i = \mathrm{diag}\left[0, \cdots, 0, \bar{p}_{i1}, \cdots, \bar{p}_{in_i}, 0, \cdots, 0\right].$$

The definition of $\bar{p}_{ij}$ is described like that:

When the linear model (1) is contaminated and the contaminated model is the stochastic error model, $\bar{p}_{ij}$ will be[5]

$$\bar{p}_{ij} = \begin{cases} p_{ij}, & \dfrac{v_{ij}^2}{r_{ij}\sigma_{i0}^2 p_{ij}} \leqslant \chi_\alpha \\ \dfrac{v_{ij}^2}{r_{ij}\sigma_{i0}^2}, \quad \sigma & \dfrac{v_{ij}^2}{r_i\sigma_{i0}^2 p_{ij}} > \chi_\alpha \end{cases} \tag{12}$$

where σ_{i0}^2 is initial value of σ_i^2, r_{ij} is the jth redundancy observation number of the ith group. When the contaminated model is the mean shift error model, $\bar{p}_{ij}$ will be[5]

$$\bar{p}_{ij} = \begin{cases} p_{ij}, & \left|\dfrac{v_{ij}}{\sigma_{v_{ij}}}\right| < w_\alpha \\ p_{ij} + \dfrac{v_{ij}^2}{r_i^2\sigma_{i0}^2}, & \left|\dfrac{v_{ij}}{\sigma_{v_{ij}}}\right| \geqslant w_\alpha \end{cases} \tag{13}$$

where the definitions of w_α and $\sigma_{v_{ij}}$, etc, are found in Ref. [5].

So, we obtain the estimate $\boldsymbol{X}$. Now, we study the robust estimates σ_i^2 with minimum MSE.

To estimate the linear function of σ_i^2

$$\Omega = C_1\sigma_1^2 + C_2\sigma_2^2 + \cdots + C_k\sigma_k^2 \tag{14}$$

Choose the estimate from Eqn. (10), then

$$\hat{\Omega} = \sum_{j=1}^{n_j}\sum_{i=1}^{k} d_{ij} v_{ij}^2 = \boldsymbol{V}^{\mathrm{T}} \Lambda \boldsymbol{V} \tag{15}$$

Naturally, demand that $\hat{\Omega}$ is unbiased estimate of Ω, that is

$$\begin{aligned} E(\hat{\Omega}) &= E(\boldsymbol{V}^{\mathrm{T}} \Lambda \boldsymbol{V}) \\ &= E(\boldsymbol{L}_\delta^{\mathrm{T}} \boldsymbol{PQ}\Lambda\boldsymbol{QPL}_\delta) \\ &= \sum_{i=1}^{k} \sigma_i^2 \mathrm{tr}(\boldsymbol{QP}_i\boldsymbol{Q}\Lambda) \\ &= \sum_{i=1}^{k} C_i\sigma_i^2 \end{aligned} \tag{16}$$

Therefore

$$\mathrm{tr}(\boldsymbol{QP}_i\boldsymbol{Q}\Lambda) = C_i \quad i = 1, 2, \cdots, k \tag{17a}$$

where

$$\boldsymbol{Q} = \boldsymbol{P}^{-1} - \boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA})^{-1}\boldsymbol{A}^{\mathrm{T}} \tag{17b}$$

$$\boldsymbol{V} = \boldsymbol{QPL}_\delta \tag{17c}$$

According to Ref. [1], the mean squared error (MSE) is

$$\begin{aligned} \mathrm{MSE}(\hat{\Omega}) &= E(\hat{\Omega} - \Omega)^2 \\ &= E(\hat{\Omega})^2 - \Omega^2 \\ &= E(\boldsymbol{L}_\delta^{\mathrm{T}}\boldsymbol{PQ}\Lambda\boldsymbol{QPL}_\delta)^2 - \Omega^2 \\ &= 2\sigma_0^4\mathrm{tr}(\boldsymbol{Q}\Lambda)^2 - \Omega^2 \end{aligned} \tag{18}$$

Now, take the unbiased estimate of Ω as

$$\hat{\Omega}_0 = \sum_{i=1}^{k}\lambda_i\boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}_i\boldsymbol{V} = \boldsymbol{V}^{\mathrm{T}}\bar{\boldsymbol{P}}\boldsymbol{V} \tag{19}$$

where

$\bar{\boldsymbol{P}} = \lambda_1\boldsymbol{P}_1 + \lambda_2\boldsymbol{P}_2 + \cdots + \lambda_k\boldsymbol{P}_k$.

Let $\boldsymbol{H} = \Lambda - \bar{\boldsymbol{P}}$

and from Eqn. (17), one can have

$$\mathrm{tr}(\boldsymbol{QP}_i\boldsymbol{Q}\Lambda) = \mathrm{tr}(\boldsymbol{QP}_i\boldsymbol{QH}) + \mathrm{tr}(\boldsymbol{QP}_i\boldsymbol{Q}\bar{\boldsymbol{P}})$$

therefore,

$$\mathrm{tr}(\boldsymbol{QP}_i\boldsymbol{QH}) = 0 \qquad i = 1, 2, \cdots, k \tag{20}$$

From Eqns. (16) and (18), one can obtain

$$\begin{aligned} \mathrm{MSE}(\hat{\Omega}) &= 2\sigma_0^4\mathrm{tr}[\boldsymbol{Q}(\boldsymbol{H} + \bar{\boldsymbol{P}})]^2 - \Omega^2 \\ &= 2\sigma_0^4\mathrm{tr}(\boldsymbol{QH})^2 + 2\sigma_0^4\mathrm{tr}(\boldsymbol{Q}\bar{\boldsymbol{P}})^2 + 4\sigma_0^4\mathrm{tr}(\boldsymbol{QHQ}\bar{\boldsymbol{P}}) - \Omega^2 \\ &= 2\sigma_0^4\mathrm{tr}(\boldsymbol{QH})^2 + 2\sigma_0^4\mathrm{tr}(\boldsymbol{Q}\bar{\boldsymbol{P}})^2 + 4\sigma_0^4\sum_{i=1}^{k}\mathrm{tr}(\boldsymbol{QP}_i\boldsymbol{QH}) - \Omega^2 \\ &\geqslant \mathrm{MSE}(\hat{\Omega}_0) \end{aligned} \tag{21}$$

Therefore $\hat{\Omega}_0$ is the robust estimete of Ω with minimum MSE.

Because

$$\begin{aligned} E(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}_i\boldsymbol{V}) &= E(\boldsymbol{L}_\delta^{\mathrm{T}}\boldsymbol{PQP}_i\boldsymbol{QPL}_\delta) \\ &= \sigma_i^2\mathrm{tr}(\boldsymbol{QP}_i) = r_i\sigma_i^2 \end{aligned}$$

$$i = 1, 2, \cdots, k \tag{22}$$

so, the robust estimate of σ^2 with minimum MSE is

$$\hat{\sigma}_i^2 = \frac{\boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}_i\boldsymbol{V}}{r_i} = \sum_{j=1}^{n_i}\frac{1}{r_i\bar{p}_{ij}}v_{ij}^2$$

$$i = 1, 2, \cdots, k \tag{23}$$

where r_i is the ith group redundancy observation number.

Therefore, one can obtain the conclusion: the robust estimate $\hat{\sigma}_i^2$ of the ith variance component with minimum MSE is plus weight squared sums of the residual value jth v_{ij} of the ith group $\boldsymbol{V}_i$. The weight function is the weight of the ith group observation value which is revised because there are blunders.

Now the robust estimates of σ_i^2 with minimum mean Cook distance[2] are proved to be Eqn. (23).

According to Ref. [6], maximum model space likelihood and orthogonal complement likelihood estimates of $\boldsymbol{X}$ and σ_i^2 in the model (1) are

$$X = (A^{T}P_{0}A)^{-1}A^{T}P_{0}L \tag{24}$$

$$\hat{\sigma}_i^2 = \frac{1}{r_i}\sum_{j=1}^{n_i}\frac{1}{P_{ij}}v_{ij}^2 \tag{25}$$

where $P_0 = \Sigma^{-1}$.

The residual error equation is

$$V = QPL_\delta \tag{26}$$

Assume that Q in Eqns. (17c) and (26) are approximately invariable, in Eqn. (15) the robust estimate $\hat{\Omega}$ of Ω in Eqn. (14) satisfies the unbiased conditions in Eqn. (17). Now, consider the mean Cook distance[2].

$$\begin{aligned} D_n &= E(\hat{\Omega} - \hat{\Omega}_0)^2 \\ &= E(L_\delta^{T}PQ\Lambda PQL_\delta - LP_0Q\Lambda_0QP_0L)^2 \\ &= 2\sigma_0^4\mathrm{tr}(Q\Lambda)^2 + 2\sigma_0^4\mathrm{tr}(Q\Lambda_0)^2 \\ &\quad - 4\sigma_0^4\mathrm{tr}(Q\Lambda Q\Lambda_0) \end{aligned} \tag{27}$$

where $\hat{\Omega}_0^2 = V^{T}\Lambda_0 V$

$$\Lambda_0 = \mathrm{diag}\left[\frac{C_1}{r_1P_{11}}, \cdots, \frac{C_1}{r_1P_{1n_1}}, \cdots, \frac{C_k}{r_kP_{k1}}, \cdots, \frac{C_k}{r_kP_{kn_k}}\right]$$

It is evident that $\hat{\Omega}_0$ is maximum orthogonal complement estimate of Ω_0.

Because

$$\begin{aligned} \mathrm{tr}(Q\Lambda Q\Lambda_0) &= \sum_{i=1}^{k}\frac{C_i}{r_i}\mathrm{tr}(QP_iQ\Lambda) \\ &= \sum_{i=1}^{k}\frac{C_i^2}{r_i} \end{aligned} \tag{28}$$

Therefore, the robust estimates $\hat{\sigma}_i^2$ with minimum mean Cook distance satisfy

$$\mathrm{tr}(Q\Lambda)^2 = \min \tag{29a}$$

$$\mathrm{tr}(QP_iQ\Lambda) = C_i, \quad i = 1, 2, \cdots, k \tag{29b}$$

So, it is easy to prove that the robust estimates $\hat{\sigma}_i^2 (i = 1, 2, \cdots, k)$ with minimum Cook distance are Eqn. (23).

In above derivation, Q in Eqns. (17c) and (26) is assumed invariable. But Q in fact is relative to P_0. In practical computation one must iterate the computation.

4 CONCLUSION

The paper considers the variance component model and gives the robust estimates X and $\hat{\sigma}_i^2$ with minimum MSE and with minimum mean Cook distance. These formulae can be found in Eqns. (11) and (23). The weight factors can be found in Eqns. (12) and (13).

The formulae are simple and easy in use, and have widely application and theorecal values.

REFERENCES

1 Zhu Jianjun. The Australia Surveyor, 1991, 36(2): 111- 115.
2 Zhu Jianjun. Journal of Geodesy, (in Chinese), 1996, 70: 586- 590.
3 Huber P J. Robust Statistic. New York: Wiley, 1981: 1.
4 Zhou Jiangwen. Acta Geodetia et Cartographica Sinica, 1986, 18: 115- 120.
5 Zhu Jianjun. Journal of Central South University of Technology, (in Chinese), 1996, 27(3): 273- 277.
6 Koch K R. Bulletin Géodésique, 1986, 60: 329- 338.

(Edited by He Xuefeng)