XU PEILIANG
Department of Geodesy
Wuhan Technical University of Surveying & Mapping
Wuhan, Hubei Province (The People's Republic of China)

# ON ROBUST ESTIMATION WITH CORRELATED OBSERVATIONS

**Abstract**

*Robust estimation methods have been widely used in Photogrammetry and Geodesy. They have an implicit basic premise that all observations are uncorrelated. However, we have to deal with various types of correlated observations in surveying practice. This paper therefore extends robust estimation methods to the case of correlated observations with the help of bivariate functions. The properties of the solution have been investigated and the proper confidence region problems have been discussed. A simulated example shows the applications of the proposed methods.*

## 1. Introduction

Since Huber presented an M-estimator for the median parameter, robust estimation methods have been rapdily developed, with various methods appearing for different purposes from different points of view (Andrews, 1974, Chen and Wang, 1984, El-Hakim, 1982, Kubik, 1982, Kubik et al., 1984). For instance, when estimating parameters in a linear regression model, Andrews (1974) used a sine function for robust estimation functions. All these methods are concerned with some parameters in these functions, which have been selected on the basis of researcher experience. They not only have an effect on the solution but also determine to some extent the confidence regions of the solution and the observations. But the relationship between the currently used significance level and its corresponding confidence region is frequently not true (Xu, 1989). Kubik et al., (1984a) pointed out that, for all methods, we had lacked proper confidence regions. Fortunately, a general procedure for selecting proper confidence regions has been developed (Xu, 1987, 1989). Robust estimation methods have been widely used in practice (Caspary et al., 1983, Caspary and Borutta, 1987, Chen, 1984, Ebong, 1985, Fuchs, 1982, 1983, Juhl, 1984, Kelm, 1986). They are sometimes named robust adjustment methods in surveying. The present robust adjustment methods, however, have an implicit basic premise that all observations are uncorrelated : This is by no means always true. For example, in deformation analysis, we get coordinate differences between any two measurement epochs by means of a free net adjustment, and then use them to solve for the deformation parameters by the L1 norm (the least sum method) or other robust methods (Caspary, et al., 1983, Chen, 1984). These coordinate differences are certainly correlated. It is noted that, in deformation applications of robust adjustment methods, displacements can be referred to as "gross errors" and outliers between observations are often supposed to have been detected at an earlier stage. Height differences from levelling measurements are correlated observations (Chen and Chrzanowski, 1986, Lucht, 1983). In the satellite Doppler positioning systems, two adjacent Doppler counter observations are correlated. To make measurement results un-influenced by outliers as far as possible, robust adjustment methods using e.g. the L1 norm are used to obtain parameters in adjustment models (Meissl, 1980).

When considering the case in which observations are correlated, one possible choice is to change all correlated observations into uncorrelated ones. Suppose that observations $l$ contain outliers $\Delta l$, with the variance-covariance matrix $D$ which is supposed to be positive definite and symmetric. Therefore an orthogonal matrix $T$ exists so that we have

$$D = T^T \lambda \, T \tag{1}$$

where $\lambda$ is a diagonal matrix with positive elements.

Thus we make a linear transformation

$$Z = T \cdot l \tag{2}$$

then

$$D(Z) = \lambda \tag{3}$$

Expression (3) means that $Z$ is uncorrelated. It is clear from (2) that the relationship between outliers of observations $l$ and $Z$ is given by

$$\Delta Z = T \cdot \Delta l \tag{4}$$

In general, the observations with outliers are only a small portion of all observations. We can therefore partition $l$ into two parts $l_1$ and $l_2$. $l_1$ is a random observational vector and $l_2$ is an observational vector with outliers. That is to say $\Delta l_1 = 0$, $\Delta l_2 \neq 0$.

Expression (4) therefore becomes

$$\Delta Z = T \begin{pmatrix} 0 \\ \Delta \ell_2 \end{pmatrix} = \begin{pmatrix} T_{12} \Delta \ell_2 \\ T_{22} \Delta \ell_2 \end{pmatrix}$$

where

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$$

It can be seen that all new observations $Z$ are contaminated by outliers $\Delta l_2$, for $T_{12} \neq 0$, $T_{22} \neq 0$. Therefore robust adjustment with observations $Z$ would become more difficult than before. The reason for this may be explained by saying that the essence of robust adjustments is to make the adjustment results depend mainly on the non-outlier observations. We may therefore come to a preliminary conclusion that robust adjustment methods should not be used after changing all correlated observations into uncorrelated ones if the observations have outliers.

This paper presents an extension of the robust adjustment methods with correlated observations using multivariate functions which serve as estimation functions for robust methods (or so called robust functions).

**2. Robust Functions**

Let us begin with the model

$$l = AX + \epsilon \;, \; P \tag{5}$$

or

$$l_i = \sum_{j=1}^{t} a_{ij} x_j + \epsilon_i \qquad i = 1, 2, \ldots, n \tag{6}$$

where $l$ = the observational vector

$P$ = the observational weight matrix

$A$ = the design matrix with full column rank

$X$ = the $t$ unknown parameter vector

$\epsilon$ = the normally distributed error vector.

The corresponding error equations are in :

$$V = A \cdot \hat{X} - l \tag{7}$$

$$v_i = \sum_{j=1}^{t} a_{ij} \hat{x}_j - l_i \qquad i = 1, 2, \ldots, n \tag{8}$$

where $V$ is the residual vector.

When observations $l$ are uncorrelated, their cofactor matrix $Q = P^{-1}$ is diagonal. We therefore obtain solutions by $\sum_{i=1}^{n} p_i v_i^2 = min$ (the least squares adjustment). However, if there are outliers between the observations, robust adjustment methods are adopted to weaken or avoid their influence on adjustment results. Thus we choose an alternative function $\rho(v)$ whose speed of increase is lower than that of the function $v^2$. In other words, we have

$$\Omega(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_t) = \sum_{i=1}^{n} \rho(l_i - \sum_{i=1}^{t} a_{ij} \hat{x}_j) = \sum_{i=1}^{n} \rho(v_i) = min \tag{9}$$

We obtain solutions by minimizing $\Omega$.

The function $\rho(y)$ should satisfy the following conditions (Chen and Wang, 1984) :

1) its definition domain is $\mathbb{R}'$

2) $\rho(y)$ is continuous

3) $\rho(0) = 0$, $\rho(y) = \rho(-y)$; and if $|v| \leqslant |u|$ then $\rho(v) \leqslant \rho(u)$

4) $\lim\limits_{y \to \infty} \rho(y) \cdot y^{-2} = 0$

Furthermore, if $\rho(y)$ has a second order differential and $\rho(y)$, $\rho'(y) \to \infty$ as $y \to \infty$, condition 4) can be equivalently rewritten as

4') $\lim\limits_{y \to \infty} \rho''(y) = 0$

The reason for this is $\lim\limits_{y \to \infty} \rho(y) \cdot y^{-2} = \lim\limits_{y \to \infty} \rho'(y) \cdot y^{-1}/2 = \frac{1}{2} \lim \rho''(y) = 0$.

However, if the observations l are correlated, the weight matrix $P = Q^{-1}$ is non-diagonal. The least squares adjustment results can easily be obtained by minimizing $V^T P \cdot V$, that is

$$V^T P \cdot V = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} v_i v_j = min \tag{10}$$

But, if there are outliers between the observations, the above estimation criterion may not be acceptable, and robust adjustment methods are used. Here we select an alternative function $f(p_{ij}, v_i, v_j)$ instead of $p_{ij} v_i v_j$. Thus the estimation criterion will take the form

$$\Omega(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_t) = \sum_{i=1}^{n} \sum_{j=1}^{n} f(p_{ij}, v_i, v_j) \tag{11}$$

By minimizing $\Omega$, we can get parameter estimates $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_t$. The function $f(p_{xy}, x, y)$ is also called a robust function.

A robust function $f(p_{xy}, x, y)$ should satisfy the following conditions :

1) its definition domain is $\mathbb{R}^2$

2) the function $f(\cdot)$ is continuous

3) $f(p_{xy}, x, 0) = 0$ and $f(p_{xy}, 0, y) = 0$

4) if $|x| \leqslant |\alpha|$ and $|y| \leqslant |\beta|$ then $|f(p_{xy}, x, y)| \leqslant |f(p_{xy}, \alpha, \beta)|$

5) $\lim\limits_{\substack{x \to \infty \\ \text{or } y \to \infty}} f(p_{xy}, x, y) \cdot x^{-2} = 0$, $\quad \lim\limits_{\substack{x \to \infty \\ \text{or } y \to \infty}} f(p_{xy}, x, y) \cdot y^{-2} = 0$

$\lim\limits_{\substack{x \to \infty \\ \text{or } y \to \infty}} f(p_{xy}, x, y) \cdot x^{-1} y^{-1} = 0$, $\quad \lim\limits_{x \to \infty} f(p_x, x, x) \cdot x^{-2} = 0$

In particular, if the function $f(p_{xy}, x, y)$ has differentials, and the values of the function and its first order partial differentials tend to $\infty$ as $x$ or $y \to \infty$, condition 5) is equivalent to :

5′) $$\lim_{\substack{x\to\infty \\ \text{or } y\to\infty}} f''_1(p_{xy}, x, y) = 0 \ ; \qquad \lim_{\substack{x\to\infty \\ \text{or } y\to\infty}} f''_2(p_{xy}, x, y) = 0$$

$$\lim_{\substack{x\to\infty \\ \text{or } y\to\infty}} f''_{12}(p_{xy}, x, y) = 0 \ , \qquad \lim_{x\to\infty} f''(p_x, x, x) = 0$$

We can even suggest extra conditions such as :

6) symmetric condition : $f(p_{xy}, x, y) = f(p_{xy}, y, x)$ (12)

The conditions are basically necessary to the robustness of solutions (Chen and Wang, 1984). But which robust functions are the best is a problem for further research.

### 3. Computation procedures and some special cases

Suppose that the function $f(p, x, y)$ is differentiable. Differentiating expression (11) with respect to $x_m$ $(m = 1, 2, \ldots, t)$ respectively, we obtain the following equation :

$$\frac{\partial \Omega}{\partial x_m} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial}{\partial x_m} f(p_{ij}, v_i, v_j) = 0 \qquad m = 1, 2, \ldots, t \tag{13}$$

or

$$\frac{\partial \Omega}{\partial x_m} = \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{ \frac{\partial}{\partial v_i} f(p_{ij}, v_i, v_j) \cdot a_{im} + \frac{\partial}{\partial v_j} f(p_{ij}, v_i, v_j) \cdot a_{jm} \right\} = 0 \tag{14}$$

where $a_{im}$, $a_{jm}$ are the elements of the matrix $A$.

Denoting

$$\frac{\partial}{\partial v_i} f(p_{ij}, v_i, v_j) = f'_1(i, j), \quad \frac{\partial}{\partial v_j} f(p_{ij}, v_i, v_j) = f'_2(i, j)$$

$$\frac{\partial^2}{\partial v_i^2} f(p_{ij}, v_i, v_j) = f''_1(i, j), \quad \frac{\partial^2}{\partial v_j^2} f(p_{ij}, v_i, v_j) = f''_2(i, j)$$

$$\frac{\partial^2}{\partial v_i \, \partial v_j} f(p_{ij}, v_i, v_j) = f''_{12}(i, j)$$

Using (12) we have

$$f'_1(i, j) = f'_2(j, i) \tag{15a}$$

$$f_1''(i,j) = f_2''(j,i) \tag{15b}$$

$$f_{12}''(i,j) = f_{12}''(j,i) \tag{15c}$$

Inserting (15a) into (14) we obtain

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial}{\partial v_i} f(p_{ij}, v_i, v_j) \cdot a_{im} = 0 \qquad m = 1, 2, \ldots, t \tag{16}$$

Clearly, eq. (16) is nonlinear. It is therefore difficult to obtain the solution directly. We have to solve eq. (16) by means of iterative techniques.

Let $\hat{X}^k$ be the $k$'th iterative solution of $X$, then the corresponding residual vector is below

$$V^k = A \cdot \hat{X}^k - l \tag{17}$$

We have to calculate, in the $(k+1)$ th iteration, neglecting the terms with and above order II in the Taylor series expansions,

$$\frac{\partial}{\partial v_i} f(p_{ij}, v_i^{k+1}, v_j^{k+1})$$

$$= \frac{\partial}{\partial v_i} f(p_{ij}, v_i^k, v_j^{k+1}) + \frac{\partial^2}{\partial v_i^2} f(p_{ij}, v_i^k, v_j^{k+1}) (v_i^{k+1} - v_i^k)$$

$$= \frac{\partial}{\partial v_i} f(p_{ij}, v_i^k, v_j^k) + \frac{\partial^2}{\partial v_i \partial v_j} f(p_{ij}, v_i^k, v_j^k)(v_j^{k+1} - v_j^k) + \frac{\partial^2}{\partial v_i^2} f(p_{ij}, v_i^k, v_j^k) \cdot \sum_{h=1}^{t} a_{ih} \Delta\hat{X}_h$$

$$= \frac{\partial}{\partial v_i} f(p_{ij}, v_i^k, v_j^k) + \frac{\partial^2}{\partial v_i \partial v_j} f(p_{ij}, v_i^k, v_j^k) \cdot \sum_{h=1}^{t} a_{jh} \Delta\hat{X}_h + \frac{\partial^2}{\partial v_i^2} f(p_{ij}, v_i^k, v_j^k) \cdot \sum_{h=1}^{t} a_{ih} \Delta\hat{X}_h$$

$$= f_1'(i,j) + f_{12}''(i,j) \cdot \sum_{h=1}^{t} a_{jh} \Delta\hat{X}_h + f_1''(i,j) \cdot \sum_{h=1}^{t} a_{ih} \Delta\hat{X}_h \tag{18}$$

where

$$\Delta\hat{X}_h = \hat{X}_h^{k+1} - \hat{X}_h^k \quad , \quad h = 1, 2, \ldots, t$$

Therefore eq. (16) becomes

$$\sum_{i=1}^{n} \sum_{j=1}^{n} f_1'(i,j) a_{im} + \sum_{i=1}^{n} \sum_{j=1}^{n} f_{12}''(i,j) \cdot \sum_{h=1}^{t} a_{jh} \Delta\hat{X}_h a_{im} + \sum_{i=1}^{n} \sum_{j=1}^{n} f_1''(i,j) \cdot \sum_{h=1}^{t} a_{ih} \Delta\hat{X}_h a_{im} = 0$$

or

$$u_m + \sum_{h=1}^{t} A_m^T D_1 A_h \Delta\hat{X}_h + \sum_{h=1}^{t} A_m^T D_2 A_h \Delta\hat{X}_h = 0\,, \quad m = 1, 2, \ldots, t \tag{19}$$

where

$$u_m = \sum_{i=1}^{n} \sum_{j=1}^{n} f_1'(i,j) \cdot a_{im}$$

$A_m$ = the $m'$th column of the matrix $A$

$$D_1 = (f_{12}''(i,j))$$

$$= \begin{pmatrix} f_{12}''(1,1) & f_{12}''(1,2) \ldots & f_{12}''(1,n) \\ f_{12}''(2,1) & f_{12}''(2,2) \ldots & f_{12}''(2,n) \\ . & . \quad \ldots & . \\ f_{12}''(n,1) & f_{12}''(n,2) \ldots & f_{12}''(n,n) \end{pmatrix}$$

$$D_2 = \mathit{diag}\,(\sum_{j=1}^{n} f_1''(i,j))$$

Eq. (19) can be rewritten as a normal equation system :

$$N \cdot \Delta X = U \tag{20}$$

where

$$N = A^T D_\Omega A$$

$$D_\Omega = D_1 + D_2$$

$$U = -(u_1, u_2, \ldots, u_t)^T$$

$$\Delta X = (\Delta\hat{x}_1, \Delta\hat{x}_2, \ldots, \Delta\hat{x}_t)^T \;.$$

Therefore, we can obtain the parameter estimates $\hat{x}_1$, $\hat{x}_2, \ldots, \hat{x}_t$ through (20) by means of an iterative technique if the matrix N is of full rank. Properties of the solution depend mainly on the robust function and the convergence of the solution is related to the initial values of the parameters at the beginning of the iterative computation. If they are too approximate, the iteration may fail. For more details see (Chen and Wang, 1984, Xu, 1987a).

Here we give some special cases.

If $p_{ij}$ in (11) is equal to zero, we give an extra condition to the function $f(p_{ij}, v_i, v_j)$,

$$f(0, v_i, v_j) = 0 \tag{21}$$

Therefore, if observations $l$ are uncorrelated, the estimate criterion (11) will become

$$F = \sum_{i=1}^{n} f(p_i, v_i) = min \tag{22}$$

It is clear that (22) is the starting point of robust estimation methods, which is naturally a special case of the robust adjustment methods for correlated observations given here. We therefore can readily obtain robust adjustment methods such as the ones presented by Huber, Andrews and so on by selecting an adequate bivariate function $f(p_{xy}, x, y)$. For instance, take the bivariate function below :

$$f(p_{ij}, v_i, v_j) = \begin{cases} p_{ij}\, sin(v_i/b)\, sin(v_j/b), & |v_i| \leqslant c_i, |v_j| \leqslant c_j \\ k & \text{elsewhere} \end{cases} \tag{23}$$

where $b$, $c_i$, $c_j$ and $k$ are positive constants.

If the observations are uncorrelated, or if the information on the observational correlations is neglected, (23) will become

$$F = f(p_i, v_i) = \begin{cases} p_i\, sin^2(v_i/b) & |v_i| \leqslant c_i \\ k & \text{elsewhere} \end{cases} \tag{24}$$

Differentiating (24) with respect to $v_i$, we have

$$\varphi(v_i) = \begin{cases} p_i\, sin(2v_i/b)/b & |v_i| \leqslant c_i \\ 0 & \text{elsewhere} \end{cases} \tag{25}$$

(25) is Andrew's sine estimate (Andrews, 1974, Chen and Wang, 1984). It is a special case of (23).

Setting

$$f(p_{ij}, v_i, v_j) = \begin{cases} p_{ij}\, v_i\, v_j & |v_i| \leqslant c_i, |v_j| \leqslant c_j \\ k & \text{elsewhere} \end{cases} \tag{26}$$

where $c_i$, $c_j$ and $k$ are positive constants.

If we take the terms with information on $p_{ij}$ from (26), it is possible to get the method given by Huber (Andrews, 1974, Chen and Wang, 1984).

Alternatively, we can realise the robustness of the parameter estimates by using iterative methods with chosen correlation functions. They begin with :

$$V^T \bar{Q}^{-1} V = min \tag{27}$$

where $\bar{Q}$ is the iterative cofactor matrix whose elements $\bar{q}_{ij}$ are computed from an iterative correlation function.

Here we recommend the iterative correlation function below :

$$\bar{q}_{ij}(q_{ij}, v_i, v_j) = \begin{cases} q_{ii} \exp(a\,|v_i|) & i=j\,,\ |v_i| > c_i \\ q_{ij} & \text{elsewhere} \end{cases} \tag{28}$$

where $a$, $c_i$ are positive constants, $q_{ij}$ is an element of the a priori cofactor matrix $Q = D/\sigma^2$, $\sigma^2$ is the unit weight variance factor. The iterative cofactor matrix $\bar{Q}$ generated by using (28) remains positive definite. The proof is easy and omitted here.

Clearly, different functions $\bar{q}_{ij}(\cdot)$ will lead to different iterative methods. Properties of the solution is related to some extent to the iterative correlation function. But, if its parameters $a$ and $c_i$ are properly selected, the form of the functions $\bar{q}_{ij}(\cdot)$ has less effect on the results in the case of uncorrelated observations (Xu, 1989).

## 4. Properties of solutions and confidence region problems

We investigate properties of solutions and the proper confidence region problems, based on equations (26) and (28). Any one successful robust method should make the iterated correlations of the observations with outliers become less and less, until even zero. Therefore it is possible to divide the observational vector $l$ into two parts $l_1$ and $l_2$, which again respectively denote a random observational vector and an observational vector with outliers. Furthermore, let $\hat{X}_1$ be the estimate obtained from $l_1$ by either iterative or robust methods, and $\hat{X}$ the estimate from all observations $l$. Without losing generality, we suppose that $l_1$ and $l_2$ respectively have the iterative weight matrices $P_1$, $P_2 \approx 0$ and $P_{12} \approx 0$. Therefore we have the normal equation :

$$A^T \bar{P} A \cdot \hat{X} = A^T \bar{P} \cdot l \tag{29}$$

When the iterative method such as (26) is used, we get

$$\bar{P}_1 = P_1 \tag{30}$$

Inserting (30) into (29) gives :

$$(A_1^T\ A_2^T) \begin{pmatrix} P_1 & \bar{P}_{12} \\ \bar{P}_{21} & \bar{P}_2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \hat{X} = (A_1^T\ A_2^T) \begin{pmatrix} P_1 & \bar{P}_{12} \\ \bar{P}_{21} & \bar{P}_2 \end{pmatrix} \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} \tag{31}$$

where $A_1$, $A_2$ are the design matrices corresponding respectively to the observations $l_1$ and $l_2$. $A_1$ is of full rank.

Omitting the terms with $\bar{P}_{12}$, $\bar{P}_{21}$ and $\bar{P}_2$ in (31), we have

$$A_1^T P_1 A_1 \hat{X} \approx A_1^T P_1 l_1$$

or

$$\hat{X} \approx (A_1^T P_1 A_1)^{-1} A_1^T P_1 l_1 = \hat{X}_1 \qquad (32)$$

Applying the mathematical expectation operator $E$ to (32), we have

$$E(\hat{X}) \approx (A_1^T P_1 A_1)^{-1} A_1^T P_1 E(l_1) = X \qquad (33)$$

It is clear that $\hat{X}$ is asymptotely unbiased. In particular, in the limiting case when $\bar{P}_{12} = 0$, $\bar{P}_2 = 0$ and $\bar{P}_1^{-1} = Q_{11}$ (the prior cofactor matrix of $l_1$), $\hat{X}$ is the minimum variance unbiased estimate. It should be pointed out that the solution obtained from (26) is asymptotically unbiased but not the minimum variance estimate, for $P_1 = Q_{11}^{-1} + Q_{11}^{-1} Q_{12} \tilde{Q}_{22}^{-1} Q_{21} Q_{11}^{-1} \neq Q_{11}^{-1}$ where

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \quad \tilde{Q}_{22} = Q_{22} - Q_{21} Q_{11}^{-1} Q_{12}$$

On the other hand, the solution obtained by using (27) and (28) is the asymptotically minimum variance unbiased estimate, because the iterative cofactor matrix used is, without loss of generality, denoted by

$$\bar{Q} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & \bar{Q}_{22} \end{pmatrix}$$

It is obvious from (28) that if the parameter a is properly selected, each diagonal element of $\bar{Q}_{22}$ is an especially large positive value $\bar{q}_i$ so that $1/\bar{q}_i \approx 0$. We therefore have

$$\bar{Q}_{22}^{-1} \approx 0$$
$$\bar{P}_{12} \approx 0$$
$$\bar{P}_2 \approx 0$$
$$\bar{P}_1^{-1} \approx Q_{11}$$

The proof is simple and omitted here.

We now make some changes to the observation equation before going on to the proper confidence region problems. Taking $Q_{21} Q_{11}^{-1} l_1$ from $l_2$, we have

$$l_2 - Q_{21} Q_{11}^{-1} l_1 = A_2 X - Q_{21} Q_{11}^{-1} A_1 X + \epsilon_2 - Q_{21} Q_{11}^{-1} \epsilon_1 \qquad (34)$$

Denoting

$$l_3 = l_2 - Q_{21} Q_{11}^{-1} l_1$$
$$A_3 = A_2 - Q_{21} Q_{11}^{-1} A_1$$
$$\epsilon_3 = \epsilon_2 - Q_{21} Q_{11}^{-1} \epsilon_1$$

then (34) becomes

$$l_3 = A_3 X + \epsilon_3 \tag{35}$$

and the cofactor matrix of $l_3$ is

$$Q_3 = Q_{22} - Q_{21} Q_{11}^{-1} Q_{12} \tag{36}$$

$$\mathrm{COV}(l_1, l_3) = 0 \tag{37}$$

Therefore $l_1$ and $l_3$ are uncorrelated. Following Xu (1989), we have in the limiting case, if the observational vector of (5) is normally distributed,

$$U = V_3^T \left\{ A_3 (A_1^T Q_{11}^{-1} A_1)^{-1} A_3^T + Q_3 \right\}^{-1} V_3 / \sigma^2 \sim \chi^2 (m) \tag{38}$$

$$F = \frac{U \cdot \sigma^2}{m \hat{\sigma}^2} \sim F(m, n-m-t) \tag{39}$$

and

$$T = v_i / \left\{ A_i (A_1^T Q_{11}^{-1} A_1)^{-1} A_i^T + q_i \right\}^{\frac{1}{2}} \hat{\sigma} \sim t(n-m-t) \tag{40}$$

where

$$V_3 = A_3 \hat{X}_1 - l_3$$

and $v_i$ is an element of $V_3$, $q_i$ a main diagonal element of $Q_3$, $m$ is the dimensional number of $l_2$ and $A_i$ is a row of $A_3$, $\hat{\sigma}^2$ is the estimate of unit weight variance obtained from $l_1$.

Based on the statistic $T$, we can determine the proper confidence region for different geodetic or photogrammetric problems under a certain significance level, which have received less attention. When the significance level is $\alpha_T$, the critical value from (40) satisfies :

$$\Pr \left\{ \frac{|v_i|}{[A_i (A_1^T Q_{11}^{-1} A_1)^{-1} A_i^T + q_i]^{1/2} \hat{\sigma}} \leqslant t_{\alpha_{T/2}} (n-m-t) \right\} = 1 - \alpha_T \tag{41}$$

Therefore the proper confidence region is

$$|v_i| \leqslant \left\{ A_i (A_1^T Q_{11}^{-1} A_1)^{-1} A_i^T + q_i \right\}^{\frac{1}{2}} \cdot t_{\alpha_{T/2}} (n-m-t) \cdot \hat{\sigma} \tag{42}$$

## 5. An example

To demonstrate the applications of the proposed methods, we have simulated the example which is taken from (Andrews, 1974) but some changes have been made. Though the original data are collected from the operation of a plant for the oxidation of ammonia to nitric acid, the methods proposed can find applications in most of the linear or linearized adjustment models of geodesy. It consists of 21 observations and 4 unknown parameters. The initial value of the unit weight variance is 1 . If the 21 observations are given different coordinates, these uncorrelated observations can be correlated by using correlation functions (e.g. Tscherning and Rapp, 1974, Schwarz and Lachapelle, 1980) :

$$C(\psi) = A_o \sum_{n=n_o}^{\infty} \frac{(n-1)}{(n-2)(n+B)} s_o^{n+2} \cdot P_n(\cos\psi)$$

where $\psi$ is the spherical distance between any two points, $A_o = 181.2$ , $B = 24$ , $s_o = 0.99895$ , $n_o = 100$ .

Hence each element of the cofactor matrix $Q$ can be obtained by using

$$q_{ij} = C(\psi)/C(0)$$

We then put 4 outliers 4.5 , 6.5 , 8.0 and 9.5 into the observations 1 , 3 , 9 and 15 , respectively. These adjustments have been computed on the Siemens 7570–C computer system of the computing center of Wuhan Technical University of Surveying & Mapping, using the following programs :

- ERRCO : a least squares adjustment with correlated observations
- $SIN_0$ : a robust adjustment without consideration of the correlation information by using (23) , $b \geqslant 3.0$
- $SIN_1$ : a robust adjustment with consideration of the correlation information by using (23), $b \geqslant 3.0$
- $HURE_0$ : a robust adjustment without consideration of the correlation information by using (26)
- $HURE_1$ : a robust adjustment with consideration of the correlation information by using (26)
- $KUB_1$ : a robust adjustment with consideration of the correlation information by using (28) , $a \geqslant 3.5$

The proper confidence regions are determined using (42) and $\alpha = 0.001$.

***Table 1*** gives the residuals of all these methods. It is obvious after looking at the column ERRCO that 6 maximum residuals are observations 2 , 9 , 13 , 15 , 17 and 18 . On the other hand, the residuals of observations 1 and 3 with the outliers are relatively small. Therefore the results of the least squares adjustment are not acceptable, for they are very sensitive to the outliers, but robust methods correctly detect the outliers. The differences of the residuals in the comparison of the two cases, with and without consideration of the correlation information for each robust method, do not seem large. Parameter estimates of the methods are listed in ***Table 2*** where column TRUE gives the true values of 4 parameters. It is clear that the least squares adjustment

**Table 1**

**Residuals of the Methods**

| OBS. | ERRCO | $SIN_0$ | $SIN_1$ | $HURE_0$ | $HURE_1$ | $KUB_1$ |
|---|---|---|---|---|---|---|
| 1 | – 3.333 | – 4.815 | – 4.811 | – 4.817 | – 4.812 | – 5.031 |
| 2 | 4.585 | – 0.480 | 0.149 | – 0.479 | 0.163 | 0.039 |
| 3 | – 2.511 | – 7.308 | – 6.795 | – 7.309 | – 6.787 | – 6.973 |
| 4 | 1.701 | – 0.401 | – 0.498 | – 0.405 | – 0.504 | – 0.664 |
| 5 | 3.275 | – 0.149 | 0.179 | – 0.150 | 0.183 | – 0.048 |
| 6 | 2.979 | 0.524 | 0.284 | 0.517 | 0.273 | – 0.022 |
| 7 | 2.087 | – 1.194 | – 1.118 | – 1.195 | – 1.112 | – 1.250 |
| 8 | 0.601 | – 0.505 | – 0.749 | – 0.510 | – 0.762 | – 1.149 |
| 9 | – 5.597 | – 7.700 | – 7.838 | – 7.704 | – 7.842 | – 7.861 |
| 10 | 2.740 | 0.166 | 0.422 | 0.166 | 0.429 | 0.411 |
| 11 | 3.453 | 0.286 | 0.542 | 0.283 | 0.541 | 0.243 |
| 12 | 2.646 | 0.939 | 0.737 | 0.937 | 0.738 | 0.721 |
| 13 | 4.959 | 0.351 | 0.997 | 0.352 | 1.010 | 1.058 |
| 14 | 2.613 | – 1.226 | – 1.418 | – 1.232 | – 1.419 | – 1.229 |
| 15 | – 5.976 | – 11.275 | – 10.619 | – 11.274 | – 10.601 | – 10.382 |
| 16 | 0.712 | – 1.293 | – 1.281 | – 1.297 | – 1.287 | – 1.394 |
| 17 | 5.007 | 1.042 | 1.413 | 1.041 | 1.422 | 1.523 |
| 18 | 4.344 | 1.809 | 1.676 | 1.805 | 1.670 | 1.574 |
| 19 | 2.709 | – 0.393 | – 0.111 | – 0.392 | – 0.100 | 0.063 |
| 20 | 1.005 | 0.237 | – 0.081 | 0.232 | – 0.092 | – 0.253 |
| 21 | 3.446 | 0.806 | 0.890 | 0.806 | 0.895 | 0.871 |

**Table 2**

**Parameter Estimations of the Methods**

| X (I) | ERRCO | $SIN_0$ | $SIN_1$ | $HURE_0$ | $HURE_1$ | $KUB_1$ | TRUE |
|---|---|---|---|---|---|---|---|
| 1 | – 2.824 | – 3.143 | – 3.169 | – 3.145 | – 3.172 | – 3.186 | – 3.173 |
| 2 | 2.385 | 1.545 | 1.750 | 1.547 | 1.756 | 1.806 | 1.816 |
| 3 | 1.025 | 0.668 | 0.547 | 0.667 | 0.546 | 0.593 | 0.524 |
| 4 | – 1.435 | – 0.953 | – 1.015 | – 0.953 | – 1.017 | – 1.082 | – 1.072 |

leads to the worst estimation of the parameters which are greatly different from their true values. The robust estimations considering the correlations are better than those neglecting them, for the accuracies of the best estimates of the four parameters are 0.082, 0.062, 0.057 and 0.043 respectively, which are nearly completely the same as those of $KUB_1$ . It is clear from *Table 2* that the three methods considering correlation information are equivalent. Furthermore, because $KUB_1$ has good asymptote properties of the solution and proper confidence regions, it is suggested that it should be used in practice. If (23) is used, the results will more and more approach those of $HURE_1$ as its parameter b increases. If $p_{ij}$ in (26) is determined according to the initial variance matrix of those observations whose $|v_i| \leqslant c_i$ , the solution is the same as that of $KUB_1$ and, of course, the asymptotically minimum variance unbiased estimate. Finally, it should be pointed out that a relatively small critical value has been used in these tests, otherwise the solutions do not converge. Further details are given in (Chen and Wang, 1984, Xu, 1987a).

## 6. Conclusion

Robust adjustment methods have been widely used in deformation analysis (Caspary et al., 1983, Caspary and Borutta, 1987, Chen, 1984), in the adjustment of levelling networks (Ebong, 1985, Fuchs, 1982, 1983, Meissl, 1980) and in the detection of outliers (Kelm, 1986, Werner, 1984, etc.). They are more accurate than the least squares adjustment when outliers are present (Krarup et al., 1980), but they however have an implicit basic premise that observations are uncorrelated. Various correlated observations must be dealt with in surveying practice and an extended robust adjustment technique has therefore been proposed. As special cases we can obtain almost all robust methods used at present. The convergence of the solution is related to initial critical values or initial values of parameters (Chen and Wang, 1984). Therefore a relatively small critical value is helpful at the beginning of iterative computations, as also is a set of good initial values of the parameters. Parameter estimation has the property of asymptotically minimum variance if the robust function is properly selected. The proper confidence region problems have also been solved here. It can be seen from (42) that the region is related to the design matrix A , the cofactor matrix Q , the number of the observations and the significance level and so on. Therefore different problems will have different confidence regions. Finally, a simulated example is computed by using the three proposed robust estimation methods with correlated observations. In actual cases the results will depend on the number of observations and the character of the problem.

## Acknowledgements

The author wishes to thank the two reviewers for making valuable comments which have been considered in improving the presentation. He is also grateful to Editor in Chief, Prof. Dr. Tscherning for the valuable suggestions and the improvement of the English of the text.

o

o o

## REFERENCES

D.F. ANDREWS : A robust method for multiple linear regression. Technometrics, Vol. 16, No. 4, pp. 523–531, 1974.

W. CASPARY, Y.Q. CHEN and R. KONIG : Kongruenzuntersuchungen in Deformationsnetzen durch Minimierung der Summe der Klaffungsbetrage, Deformationsanalysen'83, 1983.

W. CASPARY and H. BORUTTA : Robust estimation in deformation models. Survey Reviews, Vol. 29, No. 223, pp. 29–46, 1987.

Y.Q. CHEN : The least sum method and its application in deformation analysis. Journal of Wuhan Technical University of Surveying & Mapping, No. 1 (in Chinese with English abstract), 1984.

Y.Q. CHEN and A. CHRZANOWSKI : A method to assess levelling measurements. Journal of Wuhan Technical University of Surveying & Mapping, No. 1 (in Chinese with English abstract). pp. 1–9, 1986.

H.R. CHEN and S.G. WANG : Modern applied regression methods. Guangxi People Press (in Chinese), 1984.

M.B. EBONG : The least sum adjustment of a geodetic levelling network. Manuscript Geodetica, Vol. 10, pp. 32–36, 1985.

S.F. EL–HAKIM : Data snooping with weighted observations. Symposium of ISP Comm. 3, Helsinki, pp. 126–133, 1982.

H. FUCHS : Adjustment by minimizing the sum of absolute residuals, Heft Nr. 258/V, DGK, pp. 29–48, 1982.

H. FUCHS : Contribution to the adjustment by minimizing the sum of absolute residuals. Manuscript Geodetica, Vol. 7, pp. 151–207, 1983.

J. JUHL : The Danish method of weight reduction for gross error detection. Symposium of ISPRS, Comm. 3, Rio de Janeiro, pp. 468–472, 1984.

R. KELM : Lectures on European geodetic network adjustment . Presented at Wuhan Technical University of Surveying & Mapping, 1986.

T. KRARUP, J. JUHL and K. KUBIK : Gotterdammerung over least squares adjustment. Symposium of ISP, Comm. 3, Hamburg, pp. 369–378, 1980.

K. KUBIK : An error theory for the Danish method. Symposium of ISP, Comm. 3, Helsinki, pp. 299–309, 1982.

K. KUBIK, C. JORGENSEN, P. FREDERIKEN and W. WENG : Ah, robust estimation. Symposium of ISPRS, Comm. 3, Rio de Janeiro, pp. 268–277, 1984.

K. KUBIK, W. WENG, P. FREDERIKEN : Oh, grosserrors ! Symposium of ISPRS, Comm. 3, Rio de Janeiro, pp. 278–289, 1984a.

H. LUCHT : Neighbourhood correlations among observations in leveling networks. In : H. Pelzer and W. Niemeier (Editors) : Precise Levelling, Dummler Verlag, Bonn, pp. 315–326, 1983.

P. MEISSL : Adjustment of levelling networks by minimizing the sum of the absolute residuals. Proc. Second Intern. Symp. on Problems related to the redefinition of North American Vertical Geodetic Networks, Ottawa, pp. 393–415, 1980.

K.P. SCHWARZ and G. LACHAPELLE : Local characteristics of the gravity anomaly covariance function, Bulletin Géodésique, No. 1, pp. 21–36, 1980.

C.C. TSCHERNING and R.H. RAPP : Closed covariance expression for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models. Rep. 208, Dep. of Geod. Sci., Ohio State University, 1974.

H. WERNER : Automatic gross error detection by robust estimators. Symposium of ISPRS, Comm. 3, Rio de Janeiro, pp. 1101–1108, 1984.

P.L. XU : A F–T method for outliers. Journal of Wuhan Technical University of Surveying & Mapping, Vol. 12, No. 2 (in Chinese with English abstract), pp. 41–47, 1987.

P.L. XU : A test method for many outliers. ITC Journal, No. 4, pp. 314–317. 1987a.

P.L. XU : Statistical criteria for robust estimation methods. ITC Journal, No. 1, 1989.

*Received : 02.03.1988*
*Accepted : 03.04.1989*