# DETECTION, IDENTIFICATION AND MITIGATION OF OUTLIERS BY SOLVING OBSERVATION EQUATIONS WITH OUTLIERS AS PART OF UNKNOWNS 

Hiroshi Isshiki<br>School of Naval Architecture and Ocean Engineering<br>University of Ulsan<br>Ulsan 680-749, Republic of Korea<br>Tel: 82-52-259-2166<br>Fax: 82-52-259-2836<br>Email: isshiki@dab.hi-ho.ne.jp


#### Abstract

The conventional RAIM based on the least square algorithm of the residuals of the snapshot data can detect the single outlier with high probability, but it can miss the multiple outliers frequently. So, it invites a serious problem of the missed detection. We must clarify this problem, and develop a method to prevent the missed detection of the multiple outliers. In the present paper, a method to detect multiple outliers more correctly is discussed, and the numerical results seem much improved and appropriate.


Keywords: Multiple Outliers, Detection, Observation Equation, Residual.

## 1. INTRODUCTION

The conventional RAIM based on the least square algorithm of the residuals of the snapshot data can detect the single outlier with high probability, but it can miss the multiple outliers frequently. This can be proved mathematically. For example, if the outliers are any linear combinations of the columns of the design matrix, the residuals due to these outliers are equal to zero. Namely, these outliers may not be detected correctly by analyzing the residuals. So, it invites a serious problem of the missed and/or wrong detection. Many researchers have challenged to solve this problem. However, none of them has succeeded in obtaining a reliable method for detecting multiple outliers.

In the present paper, the problem is reexamined from the very basic level, and a new method is proposed for outlier detection. The problem is transformed into a minimum value problem, where the number of the outliers and the list of the satellites with the outliers are the parts of the unknowns. However, several solutions to the problem could exist when the redundancy is not enough, and we must choose a realistic solution among them. Introduction of a constraint to the solution could be effective to choose the solution from the candidates. For example, if the outliers are assumed positive, the numerical results seem improved and appropriate in a case when the outliers are really positive. However, the outliers due to multi-path could be positive or negative. A complete solution without imposing any constraints may require more redundancy or more information other than residuals.

## 2. MATHEMATICAL BACKGROUND

### 2.1. Observation Equations

The pseudo range observation equation is given as

$$
\begin{equation*}
P_{i}=\sqrt{\left(X_{i}-x\right)^{2}+\left(Y_{i}-y\right)^{2}+\left(Z_{i}-z\right)^{2}}+c \tau+e_{i}+f_{i} \tag{1}
\end{equation*}
$$

where
$P_{i}$ : pseudo range measurement of the $i$-th satellite,
$X_{i}, Y_{i}, Z_{i}$ : coordinates of the satellite antenna,
$x, y, z$ : coordinates of the receiver antenna,
$c$ : speed of light,
$\tau$ : receiver clock bias,
$e_{i}$ : zero mean Gaussian white noise $\left(e_{i} \sim N\left(0, \sigma_{e_{i}}^{2}\right)\right)$,
$f_{i}$ : magnitude of outlier related to the $i$-th satellite,
Satellite clock bias, ionosphere and troposphere delays etc. are neglected or included in $e_{i}$.
A linearized observation equation is obtained as follows. The state vector is defined as

$$
\mathbf{x}=\left[\begin{array}{llll}
x & y & z & c \tau \tag{2}
\end{array}\right]^{T} .
$$

Let

$$
\begin{equation*}
x=x_{0}+\Delta x, y=y_{0}+\Delta y, \quad z=z_{0}+\Delta z, c \tau=c \tau_{0}+c \Delta \tau, \tag{3}
\end{equation*}
$$

where $x_{0}, y_{0}, z_{0}$ and $c \tau_{0}$ is a nominal state sufficiently close to the actual state. The perturbation state is given as

$$
\Delta \mathbf{x}=\left[\begin{array}{llll}
\Delta x & \Delta y & \Delta z & c \Delta \tau \tag{4}
\end{array}\right]^{T} .
$$

The observation equation is linearized as

$$
P_{i}=\rho_{0 i}+\left[\begin{array}{llll}
-\frac{X_{i}-x_{0}}{\rho_{0 i}} & -\frac{Y_{i}-y_{0}}{\rho_{0 i}} & -\frac{Z_{i}-z_{0}}{\rho_{0 i}} & 1 \tag{5}
\end{array}\right] \Delta \mathbf{x}+e_{i}+f_{i},
$$

where $\rho_{0 i}$ is defined as

$$
\begin{equation*}
\rho_{0 i}=\sqrt{\left(X_{i}-x_{0}\right)^{2}+\left(Y_{i}-y_{0}\right)^{2}+\left(Z_{i}-z_{0}\right)^{2}} . \tag{6}
\end{equation*}
$$

Let the number of the satellites be $N$. Then, the observation equations can be written in a matrix form as

$$
\begin{equation*}
\mathbf{y}=\mathbf{H} \mathbf{x}+\mathbf{e}+\mathbf{f} \tag{7}
\end{equation*}
$$

where $\Delta \mathbf{x}$ in equation (5) is replaced by $\mathbf{x}$ for simplicity, and

$$
\begin{align*}
& \mathbf{y}=\mathbf{P}-\boldsymbol{\rho}_{0},  \tag{8a}\\
& \mathbf{P}=\left[\begin{array}{llll}
P_{1} & P_{2} & \cdots & P_{N}
\end{array}\right]^{T},  \tag{8b}\\
&=\left[\begin{array}{lllll}
\rho_{01} & \rho_{02} & \cdots & \rho_{0 N}
\end{array}\right]^{T} \\
& \mathbf{H}=\left[\begin{array}{cccc}
-\frac{X_{1}-x_{0}}{\rho_{01}} & -\frac{Y_{1}-y_{0}}{\rho_{01}} & -\frac{Z_{1}-z_{0}}{\rho_{01}} & 1 \\
-\frac{X_{2}-x_{0}}{\rho_{02}} & -\frac{Y_{2}-y_{0}}{\rho_{02}} & -\frac{Z_{2}-z_{0}}{\rho_{02}} & 1 \\
\vdots & \vdots & \vdots & \vdots \\
-\frac{X_{N}-x_{0}}{\rho_{0 N}} & -\frac{Y_{N}-y_{0}}{\rho_{0 N}} & -\frac{Z_{N}-z_{0}}{\rho_{0 N}} & 1
\end{array}\right]  \tag{8d}\\
& \mathbf{e}=\left[\begin{array}{llll}
e_{1} & e_{2} & \cdots & e_{N}
\end{array}\right]^{T}, \tag{8e}
\end{align*}
$$

$$
\mathbf{f}=\left[\begin{array}{llll}
f_{1} & f_{2} & \cdots & f_{N} \tag{8f}
\end{array}\right]^{T} .
$$

The outlier $\mathbf{f}$ may be considered as a part of the solution or a part of errors. If $\mathbf{f}$ is a part of the solution, the observation equation (7) may be written as

$$
\mathbf{y}=\left[\begin{array}{ll}
\mathbf{H} & \mathbf{I}
\end{array}\right]\left[\begin{array}{l}
\mathbf{x}  \tag{9}\\
\mathbf{f}
\end{array}\right]+\mathbf{e} .
$$

If we assume $\mathbf{f}$ is constant through the whole observation, this equation has a same structure as that for the carrier phase positioning. However, $\mathbf{f}$ changes with time in general, and this makes the problem rather difficult.

### 2.2. Property of Residuals and Outliers

In this section, $\mathbf{f}$ is considered as a part of errors (Teunissen (1998), Macabiau et al. (2005), Martini et al. (2006a, b)). Then, the weighted least squares:

$$
\hat{R}(\hat{\mathbf{x}})=\min _{\mathbf{x}} \hat{R}(\mathbf{x})
$$

(10a)

$$
\begin{equation*}
\hat{R}(\mathbf{x}) \equiv\|\mathbf{y}-\mathbf{H x}\|_{\mathbf{Q}_{\mathrm{e}}^{-1}}^{2} \equiv(\mathbf{y}-\mathbf{H} \mathbf{x})^{T} \mathbf{Q}_{\mathbf{e}}^{-1}(\mathbf{y}-\mathbf{H} \mathbf{x}) \tag{10b}
\end{equation*}
$$

is applied, where $\hat{\mathbf{x}}$ is the estimation of $\mathbf{x}$ and $\mathbf{Q}_{\mathrm{e}}$ is the variance-covariance matrix of $\mathbf{y}$ or $\mathbf{e}$ :

$$
\begin{equation*}
\mathbf{Q}_{\mathbf{e}} \equiv D(\mathbf{y})=D(\mathbf{e}) \tag{11}
\end{equation*}
$$

where $D(\mathbf{y})=E\left[(\mathbf{y}-E(\mathbf{y}))(\mathbf{y}-E(\mathbf{y}))^{T}\right]$. The solution of equation (10) is given as

$$
\begin{equation*}
\hat{\mathbf{x}}=\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{y} . \tag{12}
\end{equation*}
$$

The variance and Covariance Matrix of $\hat{\mathbf{x}}$ is given by

$$
\begin{equation*}
\mathbf{Q}_{\hat{\mathbf{x}}} \equiv D(\hat{\mathbf{x}})=\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{H}\right)^{-1} \tag{13}
\end{equation*}
$$

Definition of residual $\hat{\mathbf{r}}=\left[\begin{array}{llll}\hat{r}_{1} & \hat{r}_{2} & \cdots & \hat{r}_{N}\end{array}\right]^{T}$ is given as

$$
\begin{align*}
\hat{\mathbf{r}} & \equiv \mathbf{y}-\mathbf{H} \hat{\mathbf{x}} \\
& =\mathbf{y}-\mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{y}=\mathbf{P} \mathbf{y} \tag{14}
\end{align*}
$$

where

$$
\begin{equation*}
\mathbf{P} \equiv \mathbf{I}-\mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} . \tag{15}
\end{equation*}
$$

The variance and Covariance Matrix of $\hat{\mathbf{r}}$ is given as

$$
\begin{equation*}
\mathbf{Q}_{\hat{\mathbf{r}}} \equiv D(\hat{\mathbf{r}})=\mathbf{P} \mathbf{Q}_{\mathbf{e}} \mathbf{P}^{T} . \tag{16}
\end{equation*}
$$

As easily confirmed, $\mathbf{P}$ is an idempotent matrix, that is,

$$
\begin{equation*}
\mathbf{P}^{n}=\mathbf{P}, \quad n=1,2, \cdots, \tag{17}
\end{equation*}
$$

and satisfies

$$
\begin{align*}
& \mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{P}=0,  \tag{18a}\\
& \mathbf{P H}=0 . \tag{18b}
\end{align*}
$$

Hence

$$
\begin{equation*}
\hat{\mathbf{r}}=\mathbf{P} \mathbf{y}=\mathbf{P}(\mathbf{H} \mathbf{x}+\mathbf{e}+\mathbf{f})=\mathbf{P}(\mathbf{e}+\mathbf{f}) \tag{19}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \hat{\mathbf{r}}=\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{P}(\mathbf{e}+\mathbf{f})=0 \tag{20}
\end{equation*}
$$

Equation (18b) means that the matrix $\mathbf{P}$ is singular. And if the vector $\mathbf{e}+\mathbf{f}$ is a linear combination of columns of the design matrix $\mathbf{H}$, it makes the residual zero because of equation (18b)

Detection of the outliers from the residual $\hat{\mathbf{r}}$ is discussed in Appendix A.

## 3. ALGORITHM FOR OUTLIER IDENTIFICATION

### 3.1. Solution of Observation Equation for Determination of Magnitude of Known Outliers by the conventional method

When $i_{1}, i_{2}, \cdots, i_{q}$-th measurements have outliers $f_{1}, f_{2}, \cdots, f_{q}$, the observation equation may be written as

$$
\begin{equation*}
\mathbf{y}=\mathbf{H x}+\mathbf{C f}+\mathbf{e} \tag{21}
\end{equation*}
$$

where $\mathbf{C}$ is a matrix of order $N \times q$ whose $\left(i_{k}, k\right), k=1,2, \cdots, q$ elements are equal to one and the other ones are zero, and $\mathbf{f}=\left[\begin{array}{llll}f_{1} & f_{2} & \cdots & f_{q}\end{array}\right]^{T}$ is an unknown vector of order $q \times 1$ (Teunissen (1998)).

The least square solution $\overline{\mathbf{x}}$ and $\overline{\mathbf{f}}$ of equation (21) may be obtained by solving $\bar{R}(\overline{\mathbf{x}}, \overline{\mathbf{f}})=\min _{\mathbf{x}, \mathbf{f}} \mathrm{n}(\mathbf{x}, \mathbf{f})$.

$$
\begin{equation*}
\bar{R}(\mathbf{x}, \mathbf{f}) \equiv\|\mathbf{y}-\mathbf{H x}-\mathbf{C} \mathbf{f}\|_{\mathbf{Q}_{\mathrm{e}}^{-1}}^{2} \tag{22a}
\end{equation*}
$$

$$
\begin{equation*}
\equiv(\mathbf{y}-\mathbf{H x}-\mathbf{C f})^{T} \mathbf{Q}_{\mathrm{e}}^{-1}(\mathbf{y}-\mathbf{H x}-\mathbf{C f}) \tag{22b}
\end{equation*}
$$

Equation (22) gives

$$
\begin{equation*}
0=-2\left(\delta \mathbf{x}^{T} \mathbf{H}^{T}+\delta \mathbf{f}^{T} \mathbf{C}^{T}\right) \mathbf{Q}_{\mathbf{e}}^{-1}(\mathbf{y}-\mathbf{H} \overline{\mathbf{x}}-\mathbf{C} \overline{\mathbf{f}}), \tag{23}
\end{equation*}
$$

where $\delta \mathbf{x}$ and $\delta \mathbf{f}$ are the variations of $\mathbf{x}$ and $\mathbf{f}$. Now, we have

$$
\begin{align*}
& \mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{y}=\mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{H} \overline{\mathbf{x}}+\mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{C} \overline{\mathbf{f}}  \tag{24a}\\
& \mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{y}=\mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{H} \overline{\mathbf{x}}+\mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{C} \overline{\mathbf{f}} \tag{24b}
\end{align*}
$$

Solution $\overline{\mathbf{x}}$ and $\overline{\mathbf{f}}$ are now obtained by solving equation (24) as

$$
\begin{align*}
\overline{\mathbf{x}} & =\left(\mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1}\left(\mathbf{I}-\mathbf{C}\left(\mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P C}\right)^{-1} \mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P}\right) \mathbf{y},  \tag{25a}\\
\overline{\mathbf{f}} & =\left(\mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P C}\right)^{-1} \mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P} \mathbf{y} \tag{25b}
\end{align*}
$$

If matrix $\mathbf{A}$ of order $q \times N$ and $\mathbf{B}$ of order $4 \times N$ are defined as

$$
\begin{align*}
& \mathbf{A}=\left(\mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P} \mathbf{C}\right)^{-1} \mathbf{C}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P}  \tag{26a}\\
& \mathbf{B}=\left(\mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \tag{26b}
\end{align*}
$$

$\overline{\mathbf{x}}$ and $\overline{\mathbf{f}}$ may be expressed as

$$
\begin{align*}
& \overline{\mathbf{x}}=\mathbf{B}(\mathbf{I}-\mathbf{C A}) \mathbf{y},  \tag{27a}\\
& \overline{\mathbf{f}}=\mathbf{A y}, \tag{27b}
\end{align*}
$$

and the variance matrices $\mathbf{Q}_{\overline{\mathbf{x}}}$ and $\mathbf{Q}_{\overline{\mathbf{f}}}$ of $\overline{\mathbf{x}}$ and $\overline{\mathbf{f}}$ are derived as

$$
\begin{align*}
& \mathbf{Q}_{\overline{\mathbf{x}}} \equiv D(\overline{\mathbf{x}})=\mathbf{B}(\mathbf{I}-\mathbf{C A}) \mathbf{Q}_{\mathbf{e}}\left(\mathbf{I}-\mathbf{A}^{T} \mathbf{C}^{T}\right) \mathbf{B}^{T},  \tag{28a}\\
& \mathbf{Q}_{\overline{\mathbf{f}}} \equiv D(\overline{\mathbf{f}})=\mathbf{A} \mathbf{Q}_{\mathbf{e}} \mathbf{A}^{T} . \tag{28b}
\end{align*}
$$

The residual in this case is derived as

$$
\begin{align*}
\overline{\mathbf{r}} & =\mathbf{y}-\mathbf{H} \overline{\mathbf{x}}-\mathbf{C} \overline{\mathbf{f}} \\
& =[\mathbf{I}-\mathbf{H B}(\mathbf{I}-\mathbf{C A})-\mathbf{C A}] \mathbf{y} \\
& =\mathbf{P}(\mathbf{I}-\mathbf{C A}) \mathbf{y}, \tag{29}
\end{align*}
$$

and the variance matrix $\mathbf{Q}_{\overline{\mathbf{r}}}$ is obtained as

$$
\begin{equation*}
\mathbf{Q}_{\overline{\mathbf{r}}}=\mathbf{P}(\mathbf{I}-\mathbf{C A}) \mathbf{Q}_{\mathbf{e}}(\mathbf{I}-\mathbf{C A})^{T} \mathbf{P}^{T} . \tag{30}
\end{equation*}
$$

If the order of matrix $\mathbf{C}$ is $N \times 1$ and $\mathbf{C}$ and $\overline{\mathbf{f}}$ become a vector $\mathbf{c}$ and a scalar $\bar{f}$, The row vector $\mathbf{A}$ can be written as

$$
\begin{equation*}
\mathbf{A}=\frac{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P}}{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P} \mathbf{c}} \tag{31}
\end{equation*}
$$

Then, we have

$$
\begin{align*}
& \bar{f}=\frac{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P} \mathbf{y}}{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P} \mathbf{c}}  \tag{32}\\
& Q_{\bar{f}}=\frac{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P} \mathbf{Q}_{\mathbf{e}} \mathbf{P}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{c}}{\left(\mathbf{c}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{P} \mathbf{c}\right)^{2}} \tag{33}
\end{align*}
$$

$\bar{f}$ and $Q_{\bar{f}}$ may be rewritten as

$$
\begin{align*}
& \bar{f}=\frac{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \hat{\mathbf{r}}}{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{P} \mathbf{c}} \frac{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \hat{\mathbf{r}}}{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1}\left(\mathbf{Q}_{\mathbf{e}}-\mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T}\right) \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{c}}=\frac{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \hat{\mathbf{r}}}{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{Q}_{\hat{\mathbf{r}}} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{c}},  \tag{34}\\
& Q_{\bar{f}}=\frac{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{Q}_{\hat{r}} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{c}}{\left(\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{Q}_{\mathbf{r}} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{c}\right)^{2}}=\frac{1}{\mathbf{c}^{T} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{Q}_{\mathbf{r}} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{c}}, \tag{35}
\end{align*}
$$

where the following relation is used to derive equation (34):

$$
\begin{align*}
\mathbf{Q}_{\mathbf{r}} & =\mathbf{P Q}_{\mathbf{e}} \mathbf{P}^{T} \\
& =\left[\mathbf{I}-\mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}{ }^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}{ }^{-1}\right] \mathbf{Q}_{\mathrm{e}}\left[\mathbf{I}-\mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathrm{e}}{ }^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}{ }^{-1}\right]^{T} \\
& =\left[\mathbf{Q}_{\mathbf{e}}-\mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T}\right]\left[\mathbf{I}-\mathbf{Q}_{\mathbf{e}}{ }^{-1} \mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T}\right] \\
& =\mathbf{Q}_{\mathbf{e}}-\mathbf{H}\left(\mathbf{H}^{T} \mathbf{Q}_{\mathbf{e}}{ }^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{T} . \tag{36}
\end{align*}
$$

Hence, we have

$$
\begin{equation*}
w \equiv \frac{\bar{f}}{\sqrt{Q_{\bar{f}}}}=\frac{\mathbf{c}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \hat{\mathbf{r}}}{\sqrt{\mathbf{c}^{T} \mathbf{Q}_{\mathbf{e}}^{-1} \mathbf{Q}_{\hat{\mathbf{r}}} \mathbf{Q}_{\mathrm{e}}^{-1} \mathbf{c}}} \tag{37}
\end{equation*}
$$

In appendix A , the outliers may be obtained from a method based on the residuals $\hat{\mathbf{r}}$ or equation (19). Equations (32) and (33) correspond to equations (A.5) and (A.7) in Appendix A respectively. It has been confirmed numerically by the present authors that the same results are obtained by the method discussed in Appendix A.

### 3.2. Solution of Observation Equation for Identification and Magnitude Determination of Unknown Outliers by the New Method

Let $N$ and $q$ be the numbers of the satellites and the outliers. The Observation equation is given by

$$
\begin{equation*}
\mathbf{y}=\mathbf{H x}+\mathbf{C f}+\mathbf{e} \tag{21}
\end{equation*}
$$

where $\mathbf{C}$ is $N \times q$ matrix, and $\left(i_{k}, k\right), k=0,1, \cdots, q-1$ components are one and the other zero. However, $\mathbf{C}$ is also one of the unknowns in this case. Namely, $q$ and $i_{k}, k=0,1, \cdots, q-1$ are unknowns.

Now, C, $\overline{\mathbf{x}}$ and $\overline{\mathbf{f}}$ are the solution of $\bar{R}(\overline{\mathbf{x}}, \overline{\mathbf{f}})=\min _{\mathbf{x}, \mathbf{f}} \bar{R}(\mathbf{x}, \mathbf{f})$,
under the constraints if necessary, for example:

$$
\begin{equation*}
\bar{f}_{k}>0, k=0,1, \cdots, q-1 \tag{38b}
\end{equation*}
$$

where

$$
\begin{equation*}
\bar{R}(\mathbf{x}, \mathbf{f}) \equiv(\mathbf{y}-\mathbf{H} \mathbf{x}-\mathbf{C} \mathbf{f})^{T} \mathbf{Q}_{\mathbf{e}}^{-1}(\mathbf{y}-\mathbf{H x}-\mathbf{C f}) . \tag{38c}
\end{equation*}
$$

The constraint condition (38b) is not necessarily required, but an appropriate constraint condition may be effective if there is any. In the present paper, we assume that the delay of the signal corresponds to the positive outlier. If any abnormality in the ionosphere or troposphere is the source of outliers, it will be detected as a positive outlier, since the code signal is delayed by the abnormality in the ionosphere or troposphere. However, this does not apply when the outliers are generated by multi-path. In that case, the outliers could be either positive or negative.

The residual vector $\overline{\mathbf{r}}$ is obtained as

$$
\begin{equation*}
\overline{\mathbf{r}}=\mathbf{y}-\mathbf{H} \overline{\mathbf{x}}-\mathbf{C} \overline{\mathbf{f}}=\mathbf{P}(\mathbf{I}-\mathbf{C A}) \mathbf{y} . \tag{39}
\end{equation*}
$$

If the outliers are big enough and the observation errors are small enough, the magnitude or norm of the residual vectors for the observation equations including the correct outliers as a part of the solution becomes the minimum. This could be proven as follows.

If the observation error $\mathbf{e}$ in equation (21) is zero for simplicity, the solution $\overline{\mathbf{x}}$ and $\overline{\mathbf{f}}$ of the least square problem (38) would be the correct ones, that is, $\mathbf{x}$ and $\mathbf{f}$. They make $\bar{R}$ equal to zero, that is, the minimum of $\bar{R}$.

The minimum problem defined by equation (38) is solved numerically by following the flow as shown in Figure 1. In the figure, $\mid$ res_vector $\mid$ is the magnitude of the residual $\overline{\mathbf{r}}$, that is, $\sqrt{\bar{R}(\overline{\mathbf{x}}, \overline{\mathbf{f}})}$.


Fig. 1. Algorithm for determining outliers
(The constraints $f_{0}>0, f_{1}>0, \cdots, f_{q-1}>0$ can be neglected, if not necessary.)

## 4. NUMERICAL RESULTS

The data were downloaded from the homepage of ARGN (Australian Regional GPS Network, http://www.ga.gov.au/bin/data_server/). Data obtained on Nov. 6, 2006 (0:00:00 GPS time) for station Tow2 (Townsville) are analyzed. Nine satellites, that is, $3,6,7,14,15,16,18,21$ and 22 are used.

### 4.1. Seriousness of Wrong Detection

### 4.1.1. A Very Interesting Case

In some cases, the conventional method can't detect correctly even the single outlier. In an example shown in Table 1, six satellites were used and PRN 16 include an outlier with magnitude 100 m . However, the magnitude of residual of PRN 16 is the smallest among the satellites. Furthermore, the root of sum of squares of residuals without outlier is bigger than that with outlier.

$$
\text { Table 1. Residuals } \hat{r}_{i}
$$

(No. of satellites: 6; Sat. used: 7, 15, 16(+100m), 18, 21, 22)

| Sat. No. | Without outlier | With outliers | Note |
| :---: | :---: | :---: | :---: |
| PRN 7 | -2.30 | -0.53 |  |
| PRN 15 | -4.17 | 1.25 |  |
| PRN 16 | -0.59 | 0.03 | Include outlier |
| PRN 18 | -1.43 | -1.76 |  |
| PRN 21 | 3.72 | 0.43 |  |
| PRN 22 | 4.78 | 0.59 |  |
| Root of sum <br> of squares | 7.86 | 2.33 | $7.86>2.33!$ |

Matrix $\mathbf{P}$ for this case is given in Table 2. The magnitude of the component corresponding to PRN 16 in the column corresponding to PRN 16 is 0.01 and smaller than those corresponding to other than PRN 18 in the column. So, the residual of PRN 16 can't be the biggest, and the outlier included in PRN 16 can't be detected by the conventional method.

Table 2. An example of unfavorable matrix $\mathbf{P}$

|  | PRN 07 | PRN 15 | PRN 16 | PRN 18 | PRN 21 | PRN 22 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| PRN 07 | 0.12 | 0.04 | 0.02 | 0.19 | -0.16 | -0.21 |
| PRN 15 | 0.04 | 0.67 | 0.05 | -0.36 | -0.18 | -0.22 |
| PRN 16 | 0.02 | 0.05 | 0.01 | 0 | -0.03 | -0.04 |
| PRN 18 | 0.19 | -0.36 | 0 | 0.57 | -0.17 | -0.23 |
| PRN 21 | -0.16 | -0.18 | -0.03 | -0.17 | 0.24 | 0.3 |
| PRN 22 | -0.21 | -0.22 | -0.04 | -0.23 | 0.3 | 0.39 |

In the following, we consider the reason for wrong detection of a single outlier in general. Since we have

$$
\begin{equation*}
\hat{\mathbf{r}}=\mathbf{P e}+\mathbf{P f} \tag{19}
\end{equation*}
$$

there are two possibilities of wrong detection:
(1) Outlier $\mathbf{f}$ is small and masked by noise $\mathbf{e}$.
(2) Corresponding component in matrix $\mathbf{P}$ is small in the corresponding column.

Generally speaking, case (2) is very serious. The outlier could never be detected correctly.

### 4.1.2. Dependency of Residuals on Magnitude of Outliers

In case of multiple outliers, the correct detection by the conventional method is usually very difficult. In next examples, the number of satellites are nine, and the outliers are in SVs. 0,3 and 5 or PRN 03, 14 and 16. Dependency of residuals on magnitude of outliers is shown in Table 3 and Figure 2. The magnitudes of the outliers are 0, 10, 100, 1000 and 10000. When the magnitudes of the outliers are big, the residuals of SVs 1,3 and 8 are big, and these satellite would be judged to have outliers instead of SVs. 0,3 and 5 , if we consider the number of the outliers is three. However, there is no reason to consider the number of the outliers is three. In general, we can't decide the number of the outliers by the conventional method.

Table 3. Dependency of residuals on magnitude of outliers
(Outliers are in SVs. 0, 3 and 5)

| SV | PRN | Magnitude of outliers |  |  |  |  |  |
| :---: | :---: | ---: | ---: | ---: | ---: | ---: | :---: |
|  |  | +0 | +10 | +100 | +1000 | +10000 |  |
| 0 | 3 | 3.36 | 5.90 | 28.78 | 257.57 | 2545.89 |  |
| 1 | 6 | 0.75 | -2.89 | -35.67 | -363.51 | -3641.37 |  |
| 2 | 7 | -7.27 | -8.71 | -21.68 | -151.38 | -1448.70 |  |
| 3 | 14 | 5.92 | 11.96 | 66.29 | 609.69 | 6043.78 |  |
| 4 | 15 | -3.11 | -3.02 | -2.21 | 5.92 | 87.00 |  |
| 5 | 16 | -5.76 | -7.31 | -19.24 | -140.53 | -1353.79 |  |
| 6 | 18 | 0.29 | -0.11 | -3.73 | -39.94 | -401.90 |  |
| 7 | 21 | -1.24 | 0.12 | 12.37 | 134.84 | 1359.25 |  |
| 8 | 22 | 7.06 | 3.87 | -24.91 | -312.67 | -3190.15 |  |



Fig. 2. Residuals for various outliers
(Outliers are in SVs. 0, 3 and 5)
In Figure 3, components of residual due to noise $\mathbf{e}$ and outlier $\mathbf{f}$ are estimated. The residual component due to noise $(\mathbf{P e})_{i}$ is equal to $\hat{\mathbf{r}}$ obtained by assuming $\mathbf{f}=0$. The residual component $(\mathbf{P f})_{i}$ is estimated as the difference of $\hat{\mathbf{r}}$ between with and without outliers. Figure 3 shows clearly that even if outliers become very big, they are not correctly detected by the conventional method.

(a) $(\mathbf{P e})_{i}$ without outlier

(b) $(\mathbf{P f})_{i} /$ Outlier

Fig. 3. Estimation of components of residual due to noise $\mathbf{e}$ and outlier $\mathbf{f}$
The reasons for wrong detection of multiple outliers by the conventional method are summarized below, where we use

$$
\begin{align*}
& \hat{\mathbf{r}}=\mathbf{P}(\mathbf{e}+\mathbf{f}),  \tag{19}\\
& \mathbf{P H}=0 .
\end{align*}
$$

(1) Pf gives wrong result.
(2) Even if Pf gives a correct result, $\mathbf{f}$ may be small and masked by noise $\mathbf{e}$.
(3) $\mathbf{e}+\mathbf{f}$ is an engen vector of $\mathbf{P}$ or a linear combination of columns of $\mathbf{H}$.
(4) $\mathbf{f}$ is an engen vector of $\mathbf{P}$ or a linear combination of columns of $\mathbf{H}$.
(1) and (4) are most serious case, the outliers could never be detected.

### 4.2. Details of Outlier Detection by the Conventional Method

In the following examples, nine satellites are used.
Observation equation (21) is now given by

$$
\begin{align*}
& \mathbf{H}=\left[\begin{array}{llll}
+0.17 & +0.65 & -0.74 & 1 \\
-0.81 & +0.05 & +0.58 & 1 \\
-0.70 & +0.69 & -0.20 & 1 \\
-0.71 & +0.57 & +0.40 & 1 \\
-0.54 & +0.10 & -0.83 & 1 \\
+0.17 & +0.98 & +0.09 & 1 \\
-0.47 & +0.07 & -0.88 & 1 \\
-0.77 & -0.21 & -0.60 & 1 \\
-0.33 & +0.72 & -0.61 & 1
\end{array}\right]  \tag{40}\\
& \mathbf{y}=[+3.36+0.75-7.27+5.92-3.11-5.76+0.29-1.24+7.06]^{T}, \tag{41}
\end{align*}
$$

where $\mathbf{y}$ is the value when outliers don't exist.
The observation error is assumed to satisfy for simplicity:

$$
\begin{equation*}
\left(\mathbf{Q}_{\mathbf{e}}\right)_{i j}=\delta_{i j} \tag{42}
\end{equation*}
$$

When outliers don't exist, |res_vector $\mid$ is equal to 13.93. The definition of $\mid$ res_vector $\mid$ is given as

$$
\begin{equation*}
\mid \text { res_vector } \mid=\sqrt{\sum_{i=0}^{N-1} r_{i}^{2}} \tag{43}
\end{equation*}
$$

where $r_{i}=\hat{r}_{i}$ or $\bar{r}_{i}$ is the (i+1) -th component of the residual vector $\mathbf{r}=\hat{\mathbf{r}}$ in the conventional method or $\mathbf{r}=\overline{\mathbf{r}}$ in the new method.

Three cases are discussed below:
Case 1. One outlier exists: $+100 m$ for $S V=5$.

$$
\mathbf{y}=\left[\begin{array}{lllllllll}
-35.49 & -14.09 & -2.26 & -6.85 & 6.33 & 32.37 & 5.89 & 14.88 & -0.76 \tag{44a}
\end{array}\right]^{T} .
$$

In this case, $\mathbf{y}$ is equal to $\hat{\mathbf{r}}$ through the repeated linearlization. $\left|y_{0}\right|=\left|\hat{r}_{0}\right|=35.49$ is the biggest, and |res_vector $\mid$ is 87.17 . So, the outlier is not detected.

Case 2. Two outlier exist: +100 m for $S V=3,5$.

$$
\mathbf{y}=\left[\begin{array}{llllllllll}
-23.48 & -43.89 & -33.52 & 54.29 & 9.95 & 19.61 & 13.95 & 15.12 & -12.02 \tag{44b}
\end{array}\right]^{T} .
$$

$\left|y_{3}\right|=\left|\hat{r}_{3}\right|=54.29$ is the biggest, and the absolute value of $\left|y_{1}\right|=\left|\hat{r}_{1}\right|=43.89$ is the second biggest. |res_vector is 53.43 . So, the outliers are not detected.

Case 3. Three outliers exist: +100 m for $S V=0,3,5$.

$$
\mathbf{y}=\left[\begin{array}{lllllllll}
28.78 & -35.67 & -21.68 & 66.29 & -2.21 & -19.24 & -3.73 & 12.37 & -24.91 \tag{44c}
\end{array}\right]^{T} .
$$

$\left|y_{3}\right|=\left|\hat{r}_{3}\right|=66.29$ is the biggest, and the absolute value of $\left|y_{1}\right|=\left|\hat{r}_{1}\right|=35.67$ is the second biggest. $\left|y_{0}\right|=\left|\hat{r}_{0}\right|=28.78$ is the third biggest. $\mid$ res_vector $\mid$ is 90.16 . So, the outliers are not detected.

Hence, the correct answers are not obtained in Case 1 through 3. Namely, the residual can't give the correct identifications of the outliers, even in case of the single outlier in the above example.

Solutions for the three cases by the new method are now obtained below by using algorithm discussed in section 3.2.

### 4.3. Details of Outlier Detection by the New Method

In the following examples, we introduce a constraint that the outliers are positive.
4.3.1 Solution of Case 1 ( One outlier: +100 m for $S V=5$ )
(1) If $q=1$ is assumed, the results are shown in Table 4, and
$\min \mid$ res_vector $\mid=10.34$, when $i_{0}=5$, where $f_{0}=84.89$.
Hence, the outliers are correctly identified without constraints.
Table 4. Search for minimum of $\mid$ res_vector $\mid,(q=1)$.

| $i_{0}$ | $f_{0}$ | $\mid$ res_vector $\mid$ |
| :---: | :---: | :---: |
| 5 | 84.89 | 10.34 |
| 0 | -67.91 | 21.09 |
| 1 | -43.43 | 47.36 |
| 7 | 24.8 | 49.86 |
| 3 | -11.2 | 52.71 |
| 4 | 8.62 | 52.92 |
| 6 | 8.35 | 52.97 |
| 2 | -4.95 | 53.33 |
| 8 | -1.19 | 53.42 |

Coordinates without outliers:

$$
\begin{aligned}
& x_{0}=-5054605.95, y_{0}=3275483.61, z_{0}=-2091543.35, \\
& d\left(c t_{0}\right)=-0.31
\end{aligned}
$$

Coordinates with outliers:

$$
\begin{aligned}
& x_{1}=-505453650, y_{1}=327547648, z_{1}=-2091513.12, \\
& d\left(c t_{1}\right)=53.92
\end{aligned}
$$

Coordinates with outliers detected:

$$
\begin{aligned}
& x_{2}=-5054595.46, y_{2}=327548253, z_{2}=-209153878 \\
& d\left(c t_{2}\right)=7.89
\end{aligned}
$$

Differences of coordinates:

$$
\begin{aligned}
& d x_{1}-d x_{0}=69.45, d y_{1}-d y_{0}=-7.12, d z_{1}-d z_{0}=30.23 \\
& d\left(c t_{1}\right)-d\left(c t_{0}\right)=54.23 \\
& d x_{2}-d x_{0}=10.49, d y_{2}-d y_{0}=-1.08, d z_{2}-d z_{0}=4.57 \\
& d\left(c t_{2}\right)-d\left(c t_{0}\right)=8.20
\end{aligned}
$$

The difference of coordinates between with outliers detected and without outliers is much smaller than the difference of coordinates between with and without outliers, and the precision of the positioning increases by detecting outliers. Since $f_{0}$ is obtained as 84.89 very close to 100 , the improvement of the precision is rather big.

Since min $\mid$ res_vector $\mid$ in Table 4 is smaller than 13.93 which is the value of $\mid$ res_vector $\mid$ without outliers, we consider that the number of outlier $q$ is 1 and the satellite number $i_{0}$ is 5 .
(2) If $q=2$ is assumed, the results are shown in Table 5.

Change of min $\mid$ res_vector $\mid$ by $q$ is shown in Table 6 . If the number bigger than the real outliers is assumed, the residual becomes very small as shown in Table 6.

Table 5. Search for minimum of $\mid$ res_vector $\mid,(q=2)$.

| $i_{0}$ | $i_{1}$ | $f_{0}$ | $f_{1}$ | $\mid$ res_vector $\mid$ |
| ---: | ---: | ---: | ---: | ---: |
| 2 | 5 | -14.44 | 86.78 | 3.59 |
| 5 | 8 | 86.76 | 9.08 | 7.32 |
| 0 | 5 | -19.81 | 64.7 | 7.56 |
| 3 | 5 | 7.01 | 87.23 | 8.89 |
| 1 | 5 | -5.59 | 82.71 | 9.93 |
| 4 | 5 | -2.37 | 85.47 | 10.15 |
| 5 | 7 | 83.93 | 2.26 | 10.21 |
| 5 | 6 | 84.65 | 1.63 | 10.25 |
| 0 | 7 | -66.77 | 21.75 | 12.72 |
| 0 | 1 | -63.62 | -27.33 | 14.57 |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ |

Table 6. Change of min $\mid$ res_vector $\mid$ by $q$.

|  | $q=1$ | $q=2$ | $q=3$ | $q=4$ |
| :--- | :---: | :---: | :---: | :---: |
| min <br> $\mid$ res_vector $\mid$ | $\mathbf{1 0 . 3 4}$ | 7.32 | 1.58 | 0.85 |

4.3.2 Solution of Case 2 (Two outliers: +100 m for $S V=3,5$ )
(1) If $q=1$ is assumed, the results are shown in Table 7, and $\min \mid$ res_vector $\mid=52.71$, when $i_{0}=3$, where $f_{0}=88.80$. $\mid$ res_vector $\mid=40.76$ is smaller than 52.71 , but $f_{0}=-135.3<0$. So not adapted.

Table 7. Search for minimum of $\mid$ res_vector $\mid,(q=1)$.

| $i_{0}$ | $f_{0}$ | $\mid$ res_vector $\mid$ |
| ---: | ---: | ---: |
| 1 | -135.3 | 40.76 |
| 3 | 88.8 | 52.71 |
| 2 | -73.28 | 71.72 |
| 0 | -44.94 | 80.89 |
| 5 | 51.41 | 81.18 |
| 7 | 25.2 | 84.96 |
| 6 | 19.77 | 85.58 |
| 8 | -18.13 | 85.91 |
| 4 | 13.55 | 86.40 |

(2) If $q=2$ is assumed, the results are shown in Table 8, and
$\min \mid$ res_vector $\mid=8.89$, when $i_{0}=3, i_{1}=5$, where $f_{0}=107.01, f_{1}=87.23$.
Hence, the outliers are correctly identified.
Coordinates without outliers:

$$
\begin{aligned}
& x_{0}=-5054605.95, y_{0}=3275483.61, z_{0}=-2091543.35, \\
& d c t_{0}=-0.31
\end{aligned}
$$

Coordinates with outliers:
$x_{1}=-5054573.84, y_{1}=327550682, z_{1}=-2091495.50$,
$d\left(c t_{1}\right)=41.66$

Coordinates with outliers detected:

$$
\begin{aligned}
& x_{2}=-5054594.47, y_{2}=3275480.57, z_{2}=-209154073, \\
& d\left(c t_{2}\right)=7.47
\end{aligned}
$$

Table 8. Search for minimum of $\mid$ res_vector $\mid$.

| $i_{0}$ | $i_{1}$ | $f_{0}$ | $f_{1}$ | $\mid$ res_vector $\mid$ |
| ---: | ---: | ---: | ---: | ---: |
| 3 | 5 | 107.01 | 87.23 | 8.89 |
| 0 | 3 | -68.42 | 102.24 | 21.02 |
| 1 | 2 | -122.84 | -52.67 | 21.05 |
| 1 | 7 | -162.85 | -38.24 | 32.10 |
| 1 | 3 | -97.28 | 41.38 | 32.91 |
| 1 | 8 | -153.14 | 28.59 | 34.96 |
| 0 | 1 | -24.66 | -129.06 | 36.83 |
| 1 | 6 | -134.71 | 17.70 | 37.95 |
| 1 | 4 | -135.54 | 14.53 | 38.81 |
| 2 | 8 | -144.95 | -94.19 | 40.11 |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ |

Differences of coordinates:

$$
\begin{aligned}
& d x_{1}-d x_{0}=32.12, d y_{1}-d y_{0}=23.22, d z_{1}-d z_{0}=47.85 \\
& d\left(c t_{1}\right)-d\left(c t_{0}\right)=41.97 \\
& d x_{2}-d x_{0}=11.49, d y_{2}-d y_{0}=-3.04, d z_{2}-d z_{0}=2.62 \\
& d\left(c t_{2}\right)-d\left(c t_{0}\right)=7.78
\end{aligned}
$$

The difference of coordinates between with outliers detected and without outliers is much smaller than the difference of coordinates between with and without outliers, and the precision of the positioning increases by detecting outliers. Since $f_{0}$ and $f_{1}$ are obtained as 107.01 and 87.23 very close to 100 , the improvement of the precision is rather big.

Since min $\mid$ res_vector $\mid$ in Table 8 is smaller than 13.93 which is the value of $\mid$ res_vector $\mid$ without outliers, we consider that the number of outlier $q$ is 2 and the satellite number $i_{0}$ is 3 and 5 .
(3) If the number bigger than the real outliers is assumed, the residual becomes very small as shown in Table 9.

Table 9. Change of min $\mid$ res_vector $\mid$ by $q$.

|  | $q=1$ | $q=2$ | $q=3$ | $q=4$ |
| :--- | :---: | :---: | :---: | :---: |
| min <br> $\mid$ res_vector $\mid$ | 52.71 | $\mathbf{8 . 8 9}$ | 1.58 | 0.85 |

4.3.3 Solution of Case 3 (Three outliers: +100 m for $S V=0,3,5$ )
(1) If $q=1$ is assumed, the results are shown in Table 10.

Table 10. Search for minimum of $\mid$ res_vector $\mid,(q=1)$.

| $i_{0}$ | $f_{0}$ | $\mid$ res_vector $\mid$ |
| ---: | ---: | ---: |
| 3 | 108.44 | 30.65 |
| 1 | -109.98 | 64.84 |
| 0 | 55.06 | 80.89 |
| 2 | -47.4 | 84.27 |
| 5 | -50.45 | 84.60 |
| 8 | -37.56 | 84.81 |
| 7 | 20.62 | 88.73 |
| 6 | -5.29 | 90.05 |
| 4 | -3.01 | 90.12 |

And
min $\mid$ res_vector $\mid=30.65$, when $i_{0}=3$, where $f_{0}=108.44$.
(2) If $q=2$ is assumed, the results are shown in Table 11, and $\min \mid$ res_vector $\mid=20.94$, when $i_{0}=2, i_{1}=3$, where $f_{0}=41.02, f_{1}=129.41$.
$\mid$ res_vector $\mid=13.06$ is smaller than 20.94, but $f_{0}, f_{1}<0$. So not adapted.
Table 11. Search for minimum of $\mid$ res_vector $\mid,(q=2)$.

| $i_{0}$ | $i_{1}$ | $f_{0}$ | $f_{1}$ | $\mid$ res_vector $\mid$ |
| ---: | ---: | ---: | ---: | ---: |
| 1 | 5 | -161.89 | -113.44 | 13.06 |
| 2 | 3 | 41.02 | 129.41 | 20.94 |
| 0 | 3 | 31.58 | 102.24 | 21.02 |
| 3 | 8 | 104.8 | -19.8 | 26.22 |
| 3 | 7 | 108.36 | 20.19 | 26.36 |
| 3 | 6 | 110.8 | -17.94 | 26.75 |
| 3 | 5 | 105.26 | -15.21 | 29.28 |
| 1 | 3 | -18.77 | 99.29 | 29.6 |
| 3 | 4 | 108.93 | -8.37 | 29.8 |
| 0 | 1 | 75.34 | -129.06 | 36.83 |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ |

(3) If $q=3$ is assumed, the results are shown in Table 12, and
$\min \mid$ res_vector $\mid=5.64$, when $i_{0}=0, i_{1}=3, i_{5}=5$,
where $f_{0}=80.71, f_{1}=106.68, f_{2}=67.47$.
Hence, the outliers are correctly identified.
$\mid$ res_vector $\mid=2.33$ is smaller than 5.64 , but $f_{1}<0$. So not adapted.

Table 12. Search for minimum of $\mid$ res_vector $\mid$.

| $i_{0}$ | $i_{1}$ | $i_{2}$ | $f_{0}$ | $f_{1}$ | $f_{2}$ | $\mid$ res_vector $\mid$ |
| ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| 0 | 1 | 3 | 47.34 | -55.94 | 71.88 | 2.33 |
| 0 | 3 | 5 | 80.71 | 106.68 | 67.47 | 5.64 |
| 1 | 4 | 5 | -163.98 | 13.3 | -117.55 | 6.79 |
| 1 | 5 | 7 | -172.79 | -110.5 | -17 | 6.97 |
| 1 | 3 | 5 | -121.46 | 31 | -87.33 | 7.74 |
| 0 | 1 | 5 | -18.33 | -166.63 | -133.95 | 11.58 |
| 1 | 2 | 5 | -158.78 | -8.6 | -111.1 | 11.83 |
| 2 | 3 | 6 | 43.96 | 133.63 | -20.62 | 12.07 |
| 0 | 3 | 7 | 32.79 | 101.91 | 21.72 | 12.64 |
| 1 | 5 | 8 | -160.62 | -113.35 | -1.97 | 12.98 |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ |

Coordinates without outliers:

$$
\begin{aligned}
& x_{0}=-5054605.95, y_{0}=3275483.61, z_{0}=-2091543.35, \\
& d\left(c t_{0}\right)=-0.31
\end{aligned}
$$

Coordinates with outliers:

$$
\begin{aligned}
& x_{1}=-5054507.61, y_{1}=327548607, z_{1}=-209149800, \\
& d\left(c t_{1}\right)=89.82
\end{aligned}
$$

Coordinates with outliers detected:

$$
\begin{aligned}
& x_{2}=-505456810, y_{2}=3275475.26, z_{2}=-2091535.18, \\
& d\left(c t_{2}\right)=27.44
\end{aligned}
$$

Differences of coordinates:

$$
\begin{aligned}
& d x_{1}-d x_{0}=98.34, d y_{1}-d y_{0}=2.46, d z_{1}-d z_{0}=45.35 \\
& d\left(c t_{1}\right)-d\left(c t_{0}\right)=90.13 \\
& d x_{2}-d x_{0}=37.85, d y_{2}-d y_{0}=-8.35, d z_{2}-d z_{0}=8.17, \\
& d\left(c t_{2}\right)-d\left(c t_{0}\right)=27.75
\end{aligned}
$$

The difference of coordinates between with outliers detected and without outliers is smaller than the difference of coordinates between with and without outliers, and the precision of the positioning increases by detecting outliers. However, the improvement of the precision is not big. This may come from the rather poor estimation of $f_{2}$. The further improvement may be expected for Case 3.

Since min $\mid$ res_vector $\mid$ in Table 12 is smaller than 13.93 which is the value of $\mid$ res_vector $\mid$ without outliers, we consider that the number of outlier $q$ is 3 and the satellite number $i_{0}$ is 0,3 and 5 .
(4) If the number bigger than the real outliers is assumed, the residual becomes very small as shown in Table 13.

Table 13. Change of $\min \mid$ res_vector $\mid$ by $q$.

|  | $q=1$ | $q=2$ | $q=3$ | $q=4$ |
| :--- | :---: | :---: | :---: | :---: |
| min <br> $\mid$ res_vector $\mid$ | 30.65 | 20.94 | $\mathbf{5 . 6 4}$ | 0.08 |

### 4.4. Effectiveness of the New Method

Effectiveness of the New Method is shown below through the detection of 2 outliers in 7~9 satellites given in Table 14. The conventional method considers the outliers as a part of the observation errors. The new method developed in the present paper treats them as unknowns in the observation equations. In this example, no constraints are introduced. When the number of the satellites or the redundancy of the observation equations is increased, the both method show the higher detection rate. However, the correct detection rate by the new method is much higher than those by the conventional method as shown in Figure 4.

Next example shows result of three outliers (PRN 03, 14, 16) among nine satellites (PRN $03,06,07,14,15,16,18,21,22)$. As shown in Figure 5, when the magnitude of the outliers is 100 m , the correct combination of outliers (blue line: i0_3_5) is masked by a wrong combination (magenta line: i0_1_3). However, as the magnitude of the outliers become bigger, the magnitude of the residual of the correct combination becomes the smallest, and the outliers are correctly detected.

Table 14. Satellites used.

| Number of satellites | Satellites used |
| :---: | :---: |
| 7 | PRN $07,14,15,16,18,21,22$ |
| 8 | PRN 06, 07, 14, 15, 16, 18, 21, 22 |
| 9 | PRN 03, 06, 07, 14, 15, 16, 18, 21, 22 |



Fig. 4. Correct detection rates of outliers
The reasons for wrong detection of multiple outliers by the new method are summarized below, where we use

$$
\begin{align*}
& \hat{\mathbf{r}}_{N}=\mathbf{P}_{N} \mathbf{e},  \tag{19}\\
& \mathbf{P}_{N} \mathbf{H}_{N}=0, \text { where } \mathbf{H}_{N}=\left[\begin{array}{l}
\mathbf{H} C]
\end{array} .\right. \tag{18b}
\end{align*}
$$

(1) If magnitude of outliers $\mathbf{f}$ is small, the outliers are masked by noise $\mathbf{e}$.
(2) Noise $\mathbf{e}$ is an eigen-vector of $\mathbf{P}_{N}$ or a linear combination of columns of $\mathbf{H}_{N}$.

However, correct specification of satellites with outliers is possible when item (1) applies, if we can lower the magnitude of noise. For the purpose, we need a low noise receiver.


Fig. 5. Masking of correct detection by noise
If the number of satellites is big enough and we have sufficient redundancy in LSM, we can avoid the wrong detection as shown in Figure 4. When the number of satellites is not enough, introduction of a constraint is very effective. The assumption that the outlier is positive is not always correct, since the multi-path error could be positive or negative. If the receiver is on the ground, map matching would give a reasonable constraint.

## 5. CONCLUSIONS

The properties of the residuals obtained by the least square procedures for solving the observation equations are discussed. Mathematically, the outliers can't be determined by the conventional method or by checking the residuals alone. For example, some kind of outliers make residual zero. So, this kind of outliers may not be detected by the conventional method or by checking the residuals alone.

The following results are obtained through the present discussion:
(1) In the present paper, the number of the outliers and the magnitude of outliers are also unknowns in the observation equation. This strengthens the correct detection of the outliers and also makes it possible to estimate more precise positioning by removing the outliers.
(2) The satellites including outliers are identified correctly even when the redundancy of the observation is not enough, if a constraint such that the outliers are positive is introduced.
(3) However, the outliers due to multi-path could be positive or negative. So, when the redundancy of the observation is not enough, there is a possibility that these outliers can't be solved on the basis of the observation equation alone.
(4) The estimated magnitude of the outliers itself is a little bit poor.
(5) If the outliers are identified correctly, the coordinates of the receiver are estimated much more precisely than those estimated without identifying the outliers.
(6) In order to reduce the masking effects by noise, a low-noise receiver is required.
(7) The present approach alone may not be sufficient for a highly reliable RAIM. The detection of outliers from the sequence of the observation signals should also be investigated. A combination with INS would also increase the reliability.

## ACKNOWLEDGEMENTS

The authors thank Mr. R. Hatch of NavCom and Dr. J. Wang of University of New South Wales for valuable discussions with them.

## REFERENCES

(1) Younes, A., Bakhache, B., Nikiforov, I. and Benhallem A., (1998), Sequential RAIM: Theory and application to Civil Aviation Needs, ION GPS 98.
(2) Macabiau, C., Gerfault, B., Nikiforov, I., Fillatre, L., Roturier, B., Chatre, E., Raimondi, M. and Escher, A. C., (2005), RAIM Performance in presence of multiple range failures, ION NTM 2005.
(3) Martini, I. and Hein, G. W., (2006), An integrity monitoring technique for multiple failures detection, IEEE/ION PLANS 2006.
(4) Martini, I., Wolf, R. and Hein, G. W., (2006), Receiver integrity monitoring in case of multiple failures, ION GNSS 2006.
(5) Hewitson, S., (2003), GNSS receiver autonomous integrity monitoring: A separability analysis, ION GPS/GNSS 2003, monitoring (RAIM) for multiple outliers,
(6) Hewitson, S. and Wang, J., (2004), GNSS receiver autonomous integrity monitoring (RAIM) for multiple outliers,
(7) Teunissen, P. J. G., (1998), Quality control and GPS, Chapter 7 in GPS for Geodesy, Edited by P. J. G. Teunissen and A. Kleusberg, Springer Verlag, 2nd edition.

## APPENDIX A. DETECTION OF THE OUTLIERS FROM THE RESIDUAL $\hat{\mathbf{r}}$

If the number of the outliers is small and the redundancy $N-4$ is big enough, the possibility that $\hat{\mathbf{r}}$ is zero may be small.

From equation (19), the expectation and variance of $\hat{\mathbf{r}}$ are given as

$$
\begin{equation*}
E(\hat{\mathbf{r}})=\mathbf{P} \mathbf{f}, \quad \mathbf{Q}_{\hat{\mathbf{r}}} \equiv D(\hat{\mathbf{r}})=\mathbf{P} \mathbf{Q}_{\mathbf{e}} \mathbf{P}^{T} \tag{A.1}
\end{equation*}
$$

and $\mathbf{Q}_{\hat{\mathbf{r}}}$ is singular, since $\mathbf{P}$ is singular.
If we assume the outlier occurs at the i-th observation and express $\mathbf{f}$ as $\mathbf{f}=\mathbf{c} f$,
where the j -th element $c_{j}, j=1,2, \cdots, N$ of vector $\mathbf{c}$ is equal to 1 when $j=i$ and 0 otherwise, and $f$ is the scale of the outlier. Equation (19) is rewritten as

$$
\begin{equation*}
\hat{\mathbf{r}}=\mathbf{P c} f+\mathbf{P e} \tag{A.3}
\end{equation*}
$$

The least square solution $\tilde{f}$ of $f$ is obtained by solving

$$
\widetilde{R}(\widetilde{f})=\min _{f} \widetilde{R}(f)
$$

(A.4a)

$$
\begin{equation*}
\widetilde{R}(f) \equiv(\hat{\mathbf{r}}-\mathbf{P c} f)^{T}(\hat{\mathbf{r}}-\mathbf{P c} f) \tag{A.4b}
\end{equation*}
$$

since the variance-covariance matrix of $\hat{\mathbf{r}}$ is singular. $\tilde{f}$ may be solved as

$$
\begin{equation*}
\tilde{f}=\frac{\mathbf{c}^{T} \mathbf{P}^{T} \hat{\mathbf{r}}}{\mathbf{c}^{T} \mathbf{P}^{T} \mathbf{P c}} \tag{A.5}
\end{equation*}
$$

Hence, we have

$$
\begin{equation*}
\frac{\tilde{f}}{\sqrt{Q_{\tilde{f}}}}=\frac{\mathbf{c}^{T} \mathbf{P}^{T} \hat{\mathbf{r}}}{\sqrt{\mathbf{c}^{T} \mathbf{P}^{T}\left(\mathbf{P Q}_{\mathbf{e}} \mathbf{P}^{T}\right) \mathbf{P c}}} \tag{A.6}
\end{equation*}
$$

where

$$
\begin{equation*}
\sigma_{\tilde{f}}^{2} \equiv Q_{\tilde{f}} \equiv D(\widetilde{f})=\frac{\mathbf{c}^{T} \mathbf{P}^{T}\left(\mathbf{P} \mathbf{Q}_{\mathbf{e}} \mathbf{P}^{T}\right) \mathbf{P c}}{\left(\mathbf{c}^{T} \mathbf{P}^{T} \mathbf{P} \mathbf{c}\right)^{2}} \tag{A.7}
\end{equation*}
$$

Equation (A.6) may be used for detection of the outlier, since $\tilde{f} / \sigma_{\tilde{f}}$ follows the standard normal distribution $N(0,1)$ when the outlier does not exists. If the outlier exists, $E\left(\tilde{f} / \sigma_{\tilde{f}}\right) \neq 0$.

When $i_{1}, i_{2}, \cdots, i_{q}$-th measurements have outliers $f_{1}, f_{2}, \cdots, f_{q}$, equation (19) may be written as

$$
\begin{equation*}
\hat{\mathbf{r}}=\mathbf{P C f}+\mathbf{P e} \tag{A.8}
\end{equation*}
$$

where $\mathbf{C}$ is a matrix of order $N \times q$ whose $\left(i_{k}, k\right), k=1,2, \cdots, q$ elements are equal to 1 and the other ones are 0 , and $\mathbf{f}=\left[\begin{array}{llll}f_{1} & f_{2} & \cdots & f_{q}\end{array}\right]^{T}$ is an unknown vector of order $q \times 1$.

The least square solution $\widetilde{\mathbf{f}}$ of $\mathbf{f}$ is obtained by solving

$$
\begin{equation*}
\widetilde{R}(\widetilde{\mathbf{f}})=\min _{f} \widetilde{R}(\mathbf{f}), \tag{A.9a}
\end{equation*}
$$

$$
\begin{equation*}
\widetilde{R}(\mathbf{f}) \equiv(\hat{\mathbf{r}}-\mathbf{P C} \mathbf{f})^{T}(\hat{\mathbf{r}}-\mathbf{P C} \widetilde{\mathbf{f}}), \tag{A.9b}
\end{equation*}
$$

since the variance-covariance matrix of $\hat{\mathbf{r}}$ is singular. $\widetilde{\mathbf{f}}$ may be solved as

$$
\begin{equation*}
\widetilde{\mathbf{f}}=\left[\mathbf{C}^{T} \mathbf{P}^{T} \mathbf{P C}\right]^{-1} \mathbf{C}^{T} \mathbf{P}^{T} \hat{\mathbf{r}} . \tag{A.10}
\end{equation*}
$$

The variance may be written as

$$
\begin{equation*}
\mathbf{Q}_{\tilde{\mathbf{f}}} \equiv D(\widetilde{\mathbf{f}})=\left[\mathbf{C}^{T} \mathbf{P}^{T} \mathbf{P C}\right]^{-1} \mathbf{C}^{T} \mathbf{P}^{T} \mathbf{P} \mathbf{Q}_{\mathbf{e}} \mathbf{P}^{T} \mathbf{P C}\left[\mathbf{C}^{T} \mathbf{P}^{T} \mathbf{P C}\right]^{-1} \tag{A.11}
\end{equation*}
$$

Received: 2009-06-29,
Reviewed: 2009-08-27,
Accepted: 2009-09-16.

