<table>
<thead>
<tr>
<th>Title</th>
<th>Graphical Expression of the Linear Least-Square Theory, and Its Application to the VLBI Observational Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Nakajima, Koichi</td>
</tr>
<tr>
<td>Citation</td>
<td>Hitotsubashi journal of arts and sciences, 27(1): 61-77</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1986-12</td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Text Version</td>
<td>publisher</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://doi.org/10.15057/2524">http://doi.org/10.15057/2524</a></td>
</tr>
</tbody>
</table>
GRAPHICAL EXPRESSION OF THE LINEAR LEAST-SQUARE THEORY, AND ITS APPLICATION TO THE VLBI OBSERVATIONAL EQUATIONS

KOICHI NAKAJIMA

I. Introduction

In the geodetic observations such as Very Long Baseline Interferometry (VLBI) the method of least squares is generally used for determining geodetic parameters from observed data. In most cases the observational equations are linearized, and its theoretical treatment is given as the theory of linear estimation in which the Gauss-Markoff model is applied (Rao 1973).

The linear least-square theory (we abbreviate LLST hereafter) is described in terms of vectors and matrices. It seems that some parts should be expressed graphically to help visual understandings. However, systematic explanations utilizing graphical expressions are hardly seen.

Here we try to give some simplified graphical explanations of a part of LLST, to make clear the characteristics of the normal equation. Then we apply this analysis to an example, a VLBI observation schedule.

The summary of the method of analysis is as follows. The coefficient matrix of the normal equation determines a (hyper)ellipsoid in an m-dimensional space formed by m unknown parameters, and if the composition of the observational equations is improper, some axes are elongated abnormally, causing larger errors in many parameter estimations. This is because the standard deviation of each parameter is given as the projection of the ellipsoid onto each axis of coordinate corresponding to the parameter. The case like this is often called "bad separation of parameters."

In this case, however, we can find a linear combination of unknown parameters whose s.d. is rather small, choosing the estimation-axis toward a thinner part of the ellipsoid. To obtain the lengths and directions of principal axes of the ellipsoid, we solve an eigen value problem of the coefficient matrix of the normal equation.

Applying this analysis to actual observations, we can judge, quantitatively, the efficiency of the observation schedule, or we can find an optimum estimation of parameters in a form of their linear combination in case that the condition of observation is restricted.

II. Equations

In this section equations describing LLST are summarized in terms of vectors and
matrices. Notations are used following Rao's textbook (Rao 1973).

Let $A'$ be the transpose of a matrix (including a vector) $A$. Then we use following notations:

- $Y'=(y_1, \ldots, y_n)$: observed data such as delay times in VLBI. $n$ denotes the number of observations.
- $X=(x_{ij} | i=1, \ldots, n; j=1, \ldots, m)$: $n \times m$ known matrix of coefficients of observational equations. $m$ denotes the number of unknown parameters.
- $P'=(p_1, \ldots, p_m)$: unknown parameters.
- $e'=(e_1, \ldots, e_n) = (Y-X\beta)'$: residuals.

Note that each one $\beta$ determines one $e$. We expect that a true $\beta$ determines $e$ such as $E(e)=0$.

where $E(e)$ means the expectation of $e$. For simplicity we assume that the observations have equal weight and that $y_i$'s are independent, i.e.,

$$D(e) = \sigma^2 I, \quad (\sigma > 0),$$

which we depict in Figure 1. In the case of $n \geq m$, the normal equation is

$$X'X\beta = X'Y.$$
which is derived from equation (3) by minimizing $e'e$, i.e. $(Y - X\beta)'(Y - X\beta)$. Let $\hat{\beta}$ be a solution of (4), and let

$$\hat{\varepsilon} = Y - X\hat{\beta}.$$  

Then we have

$$X'\hat{\varepsilon} = X'(Y - X\hat{\beta}) = X'Y - X'X\hat{\beta} = 0.$$  

If rank $XX' = m$ (i.e. $XX'$ is regular) the solution $\hat{\beta}$ is determined uniquely and is given by

$$\hat{\beta} = (XX')^{-1}X'Y.$$  

If it is not the case (i.e. rank $XX' < m$) we can obtain a $\hat{\beta}$ by using the generalized inverse of matrix (Rao 1973, chapter 4; Bjerhammer 1973). However, we deal with only the former case because in this work we are investigating the efficiency of given observational equations, assuming that $XX'$ is regular.

Note that if $n = m$ and $X$ is regular, we can find a solution $\hat{\beta} = X^{-1}Y$, where $e'e = 0$. This is a special case of equation (6) because if $X$ is an $m \times m$ matrix and is regular, then $X^{-1}X = (XX')^{-1}$.

Here we will show that $(XX')^{-1}$ gives the variance-covariance of estimated $\hat{\beta}$ (i.e. variances and covariances among $\hat{\beta}_1, \ldots, \hat{\beta}_m$). From equations (3) and (1) we have

$$D(Y) = D(e) = \sigma^2 I.$$  

$$\therefore D(XX') = E(XX'YX') = E(YY'X) = \sigma^2 XX'.$$

$$\therefore D(\hat{\beta}) = D((XX')^{-1}YX') = (XX')^{-1}D(XX')(XX')^{-1} = \sigma^2 (XX')^{-1},$$  

because $XX'$ and $(XX')^{-1}$ are symmetric.

### III. Graphical Expressions of LLST and the Error Ellipsoid

In this section we display graphically the relations among the observational equations, the estimated $\hat{\beta}$, and its error ellipsoid.

#### III.1. The relation between observational equations and $\hat{\beta}$

In the $m$-dimensional space of $\beta$, an equation of $\beta$

$$X_i' \beta = y_i,$$  

determines a hyperplane, where $X_i = (x_{i1}, \ldots, x_{im}), \quad i = 1, \ldots, n.$

is a row vector of $X$ (see Figure 1). Let

$$\|X_i\| = (X_i'X_i)^{1/2},$$  

and

$$\epsilon_i = X_i/\|X_i\| \quad ($i.e. a unit vector toward $X_i$).$$

The hyperplane (9) is perpendicular to $\epsilon_i$ (i.e. $X_i$), and its distance from the origin (i.e. $\beta = 0$)
FIG. 2. **RELATIONS AMONG THE HYPERPLANE** \(X_t' \beta = y_t\), \(\varepsilon_t\), \(\delta_t\), AND \(\epsilon_t\)

The projection of every vector \(\beta\) onto the direction \(e_t\) is given by \(e_t' \beta = X_t' \beta / ||X_t||\) (see Figure 2).

If a point \(\beta\) is not on the plane \(9\), there is an \(\epsilon_t\) which is given by equation (3). In this case the distance between \(\beta\) and the plane is

\[
\text{distance} = \frac{X_t' \beta - y_t}{||X_t||}
\]
\[ \delta_t = \text{abs}(e_t)/||X_t||, \]

as shown in Figure 2.

As seen in Figure 3, the least-square solution \( \hat{\beta} \) is in the midst of hyperplanes \( X\beta = Y \), and is given as a point which makes the square-sum of \( \epsilon_t = \delta_t ||X_t|| \) (and not that of distances \( \delta_t \)) be minimum.

III.2. Distributions of the true \( \beta \) around the estimated \( \hat{\beta} \), and the error ellipsoid

Next we investigate the distribution of the true \( \beta \) around the estimated \( \hat{\beta} \). The distribution can be regarded the distribution of \( \hat{\beta} \) itself. From this distribution we will derive the error ellipsoid.

Adopting the Gauss-Markoff model, we have that the distribution of the true \( \beta \) determined by an observational equation

\[ X_t'\beta = y_t + e_t \text{ (in which } E(e_t^2) = \sigma^2 \),

is given by a one-dimensional normal distribution along \( e_t \), such as

\[ N(e_t/||X_t|| | 0, (\sigma/||X_t||)^2) = N((X_t'\beta - y_t)/||X_t|| | 0, (\sigma/||X_t||)^2) \]

\[ = N(e_t'\beta - \hat{\beta} | 0, (\sigma^2)/||X_t||^2) \]

\[ = N(e_t'\beta | e_t'\hat{\beta}, (\sigma^2)/||X_t||^2) \]

where we put \( y_t = X_t'\hat{\beta} \) (see Figure 4).

Now let us define the \( m \)-dimensional normal distribution by

\[ N_m(\beta | \hat{\beta}, v) = N_m(\hat{\beta} | \beta, v) \exp[-(1/2)(\beta - \hat{\beta})'v^{-1}(\beta - \hat{\beta})], \]

\[ \text{ .........................................................(11)} \]

**Fig. 4. One Dimensional Normal Distribution Along } e_t**
where \( v \) denotes the variance-covariance matrix, and \( N_m(\beta | \hat{\beta}, v) \) is determined by the normalization, so that \( \int N_m(\beta | \hat{\beta}, v) \, d\beta = 1 \). Using these notations, we test whether or not the distribution of the true \( \beta \) around the estimated \( \hat{\beta} \) is an \( m \)-dimensional normal distribution.

This distribution is calculated by multiplying \( n \) one-dimensional normal distributions of equation (10), in the \( m \)-dimensional space of \( \beta \). Rewriting equation (10) into \( m \)-dimensional form, as

\[
N_m(\beta | \hat{\beta}_i, v_i) = N_m(\hat{\beta}_i | \hat{\beta}_i, v_i) \exp\left[ -\frac{1}{2} \frac{e_i' (\beta - \hat{\beta}_i)^2}{(\sigma^2 ||X_i||)^2} \right]
\]

we have

\[
\prod_{i=1}^{n} N_m(\beta | \hat{\beta}_i, v_i) = \prod_{i=1}^{n} N_m(\hat{\beta}_i | \hat{\beta}_i, v_i) \exp\left[ -\frac{1}{2} \sum_{i=1}^{n} \{X_i' (\beta - \hat{\beta}_i) \}^2 \right].
\]

Remembering that \( X_i' \hat{\beta}_i = y_i = X_i' \hat{\beta} - \hat{\epsilon}_i \) from equation (5), we have

\[
\sum_{i=1}^{n} \{X_i' (\beta - \hat{\beta}_i) \}^2 = \sum_{i=1}^{n} \{X_i' (\beta - \hat{\beta}_i) + \hat{\epsilon}_i \}^2
\]

\[
= \sum_{i=1}^{n} \{(\beta - \hat{\beta})' X_i' X_i (\beta - \hat{\beta}) + 2 \hat{\epsilon}_i X_i' (\beta - \hat{\beta}) + \hat{\epsilon}_i^2 \}
\]

\[
= (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) + 2 \hat{\epsilon}' X (\beta - \hat{\beta}) + \hat{\epsilon}' \hat{\epsilon}. \quad \text{.................... (14)}
\]

From equation (6) the second term drops, and we have

\[
\prod_{i=1}^{n} N_m(\beta | \hat{\beta}_i, v_i)
\]

\[
= \prod_{i=1}^{n} N_m(\hat{\beta}_i | \hat{\beta}_i, v_i) \exp\left[ -\frac{1}{2} \sigma^2 (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right] \exp\left[ -\frac{1}{2} \sigma^2 \hat{\epsilon}' \hat{\epsilon} \right]. \quad \text{.......... (15)}
\]

**FIG. 5. RELATIONS AMONG EACH HYPERPLANE AND THE SDE**
Thus we have proved that this distribution is an \( m \)-dimensional normal distribution whose variance-covariance matrix is \( \sigma^2(X'X)^{-1} \). Also we have shown equation (8) again. It can be seen in equation (15) that the peak of distribution is the highest when \( \hat{\beta} \) is so determined that \( e'e \) is minimum.

From these analyses we find that those \( \beta \) which give the probability of \( \exp(-1/2) \) of the peak value form a hyperellipsoid in the \( m \)-dimensional space, such as

\[
\sigma^2 = (\beta - \hat{\beta})'X'X(\beta - \hat{\beta}).
\]

The right hand side of equation (16) shows a quadratic form of \( \beta \) which is positive except \( \beta = \hat{\beta} \), and then the symmetric matrix \( X'X \) is positive definit (see Rao 1973, Ic). Thus this ellipsoid is everywhere convex. Let us call it a “standard-deviation ellipsoid” (SDE) because such a \( \beta - \hat{\beta} \) is called a standard-deviation. In Figure 5 we show graphically the relations among each hyperplane for each observational equation and the SDE.

### III.3. The properties of the SDE

In this subsection we investigate the properties of the SDE and their relations to the parameters which appear in the LLST. The properties of an ellipsoid are mainly the lengths and directions of its principal axes, and to find them is an eigen-value problem. For simplicity we put \( \hat{\beta} = 0 \) hereafter.

In this case equation (16) becomes

\[
\sigma^2 = \beta'X'X\beta.
\]

Since \( X'X \) is a positive-definit \( m \times m \) matrix as described above, \( X'X \) has \( m \) eigen values \( \lambda_j(j=1, \ldots, m) \) which are all positive. In this case we have an orthogonal \( m \times m \) matrix

**Fig. 6. Relations among \( \beta_j, \beta_j, \gamma_j, \lambda_j, \) and the SDE**
\( P \), such as
\[ |P| = (\text{determinant of } P) = 1, \]
and
\[ r' = P' \beta \text{ or } \beta = P r, \]

\[ \beta'X'X \beta = r'P'X'X P r \]
\[ = r'/\lambda_1 \ldots 0 \] \[ = \sum_{j=1}^{m} \lambda_j f_j^2. \]

The columnal vectors \( P_j \) \((j=1, \ldots, m)\) of \( P \) are the eigen vectors which are shown in Figure 6. In this case, \( r'=(r_1, \ldots, r_m) \) expresses a set of coordinates of a point \( \beta \) with respect to the reference framework formed by \( P_1, \ldots, P_m \).

From equations (17) and (19) we have
\[ \sigma^2 = \sum_{j=1}^{m} \lambda_j f_j^2, \]

which determines a hyperellipsoid whose principal axis along \( P_j \) has length of \( \sigma \sqrt{\lambda_j} \) (see Figure 6).

Let us examine the relation between the SDE and each s.d. of the estimated value of an unknown parameter \( \beta_j \). Calculating the distribution of the true \( \beta_j \), we will find its s.d. as a function of parameters which determine the SDE. Note that from equation (8) we already know that the variance of a true \( \beta_j \) (or \( \hat{\beta}_j \)) is given by the \( jj \) component of \( \sigma^2(X'X)^{-1} \).

The distribution of \( \beta_j \) is given by
\[ \int \ldots \int N_m(\beta|0, v) d\beta_1 \ldots \gamma \ldots, d\beta_m. \]

To calculate one integral \( \int N_m(\beta|0, v) d\beta_k \) \((k \neq j)\), let us see the \( \beta_k \)-dependence of \( N_m(\beta|0, v) \) which we denote \( f_m(\beta_k) \). Since the exponent in \( N_m(\beta|0, v) \) is a quadratic form of \( \beta_k \), we can write
\[ f_m(\beta_k) \propto \exp\left[-(1/2\sigma^2) \{a^2(\beta_k - b)^2 + c\} \right], \]

where \( a^2 = \sum_{i=1}^{m} x_i^2 \), and \( b \) is a linear combination of \( \beta_1, \ldots, \gamma, \beta_m \), further \( c \) is a quadratic form of \( \beta_1, \gamma, \ldots, \beta_m \).

Since \( a \) is independent of \( \beta \), the integral \( \int N_m(\beta|0, v) d\beta_k = \int f_m(\beta_k) d\beta_k \) is proportional to the peak value of \( f_m(\beta_k) \). (The coefficient of the proportion is \( \sigma \sqrt{\pi/\alpha} \).) Because of the same reason, the integral (21) is proportional to the peak value of \( N_m(\beta|0, v) \) as a function of \( \beta_1, \ldots, \gamma, \beta_m \), that is, (21) is proportional to the peak value of \( N_m(\beta|0, v) \) on the hyperplane given by \( \beta_j=\text{constant} \) (see Figure 7). Since the coefficient of proportion is independent of \( \beta_j \), the distribution of \( \hat{\beta}_j \) is "normal" about \( \beta_j=0 \).

Then it is easy to find the s.d. of \( \beta_j \), i.e. a \( \beta_j \) at which the probability is \( \exp(-1/2) \) of the peak value that occurs at \( \beta_j=0 \), as follows. Such a \( \beta_j \) is given as the coordinates \( \beta_{1+} \).
and $\beta_{2-}$ of the two extreme points on the SDE toward the direction of $\beta_2$ axis (see Figure 7), because at these points the distribution $N_m(\beta,0, \nu)$ confined in the hyperplane $\beta_2=\beta_{2+}$ or $\beta_2=\beta_{2-}$ is maximum, and is $\exp(-1/2)$ of the peak value of $N_m(\beta,0, \nu)$ which occurs at $\beta=0$.

Finally let us see that the $jj$ component of $\sigma^2(X'X)^{-1}$ (i.e. the variance of $\hat{\beta}_j$) surely corresponds to $(\beta_{2+})^2$ or $(\beta_{2-})^2$. To begin with, let us calculate the $m$-dimensional “volume” of the ellipsoid $\sigma^2=\beta'X'X\beta$. From equation (20) and from Figure 6 we see that the volume $V_p$ of the parallelotope which surrounds the SDE is given by

$$V_p=\prod_{j=1}^{m} \sigma_j^2 \sqrt{\lambda_j}. \hspace{1cm} \text{(23)}$$

In other words the determinant of $\sigma^2(X'X)^{-1}$ is, from equation (19),

$$\sigma^{2m}|(X'X)^{-1}|=\sigma^{2m}|X'X|^{-1}$$

$$=\sigma^{2m}|P'X'XP|^{-1}$$

$$=\sigma^{2m}(\prod_{j=1}^{m} \lambda_j)^{-1}$$

$$=V_p^2, \hspace{1cm} \text{...(24)}$$

because we defined $|P|=1$. The ratio of the volume of the SDE (denoted by $V_s$) to that of the parallelotope ($V_p$) is equal to the ratio of the volume of an $m$-dimensional hypersphere to that of a hypercube, and is equal to

$$K_m=\pi^{m/2}/\Gamma(m/2+1), \hspace{1cm} \text{...(24)}$$

that is,

$$V_s=K_m\sigma^m \sqrt{\text{abs} |X'X|^{-1}}. \hspace{1cm} \text{...(25)}$$

Next we calculate the “area” of a cross-section of the SDE, cut by a hyperplane $\beta_j=0$. The

**FIG. 7. RELATION BETWEEN THE STANDARD DEVIATION OF $\hat{\beta}_j$ AND THE SDE**
cross-section is an \((m-1)\)-dimensional hyperellipsoid which is given from equation (17) by putting \(\beta_j=0\), and is

\[
s^2=\begin{pmatrix} \beta_1, \ldots, 0, \ldots, \beta_m \end{pmatrix} X'X \begin{pmatrix} \beta_1 \\ \vdots \\ 0 \\ \vdots \\ \beta_m \end{pmatrix},
\]

that is equal to

\[
s^2=\begin{pmatrix} \beta_1, \ldots, \gamma, \ldots, \beta_m \end{pmatrix} S \begin{pmatrix} \beta_1 \\ \vdots \\ \gamma \\ \vdots \\ \beta_m \end{pmatrix},
\]

where \(S\) denotes an \((m-1)\) \times \((m-1)\) matrix which is made from \(X'X\) by omitting \(j\)-th row and \(j\)-th column. Therefore, the area of the cross-section (denoted by \(S_e\)) is given by

\[
S_e=K_{m-1} a^{m-1} \sqrt{\text{abs } |S|^{-1}}.
\]

On the other hand, we know, from the theorems of matrices, that \(jj\) component of the inverse matrix of \(X'X\) is given by

\[
(X'X)^{-1}_{jj}=(-1)^j S_{jj} / |X'X|.
\]

Thus from equations (25), (27) and (28), we have

\[
s \sqrt{(X'X)^{-1}}=K_{m} V_e / K_{m-1} S_e,
\]

which, we see in Figure 8, is the height of the top of the SDE above the hyperplane \(\beta_j=0\).

**FIG. 8. RELATIONS AMONG THE JJ COMPONENT OF VARIANCE-COVARIANCE MATRIX, AREA OF THE CROSS-SECTION, AND THE VOLUME OF SDE**
IV. Application of the SDE Analysis to the Geodetic Observations

Applying the SDE analysis to actual observations and investigating the shape of the SDE, we can judge the efficiency of observations quantitatively, or we can find an optimum estimation of parameters in a form of their linear combination if the condition of observation is restricted. In this section we explain how to apply this analysis and then see how useful it is, by using an observation schedule of a geodetic VLBI experiment.

IV.1. Observational equations of geodetic VLBI

Here the “geodetic VLBI observation” means such an observation as to determine three components of a baseline vector and a clock parameter in the reference framework formed by some radio sources whose coordinates are assumed to be well determined. Note that the “baseline vector” means a vector connecting two antennas, and the “clock parameter” means the clock difference between the two sites. By the geodetic VLBI observation we can determine these parameters with an unprecedentedly high accuracy, e.g. with a few centi-meters or with 0.1 nano-second for as long a baseline as intercontinental.

As we know rough values of these parameters, we can choose the unknowns to be their small corrections, thus the observational equations can be linearized. In the case of our simple model, however, the observational equations are originally linear. Let us take a reference frame which rotates with a fixed speed with respect to the frame formed by the radio sources. Choosing the rotation speed as close to that of the Earth as possible, we can treat the baseline vector not changing with time in a relatively short time interval (e.g. a few hours). Although vectors pointing radio sources are not constant with time, they can be treated as known parameters at a given moment of time, because the rotation speed is fixed.

Let us introduce the observational equations for the geodetic VLBI. What we observe by the VLBI are the “delay” and the “delay-rate,” in which the “delay” means the difference

FIG. 9. RELATIONS AMONG THE PARAMETERS WHICH APPEAR IN THE OBSERVATIONAL EQUATIONS
of arrival time of a radio signal between the two antennas. Here we consider only the case of delay measurement. Let \( \mathbf{s} \) be a unit vector pointing to a radio source, and \( \mathbf{b} \) be the baseline vector. Let \( \tau \) be the observed delay time. As seen in Figure 9, we have, as an \( i \)-th observational equation,

\[
c(r_0) = \mathbf{b}'s_t + c\tau_e,
\]

where \( c \) denotes the light-velocity, and \( \tau_e \) be the clock advance of the site B against the site A. Since the unknown parameters are \( \mathbf{b} \) and \( \tau_e \), we can find the correspondences

\[
\mathbf{X}' = \begin{pmatrix} c\tau_e, & b_x, & b_y, & b_z \end{pmatrix}, \quad (31)
\]

\[
y_i = (cT_g)_i
\]

where we take into account the dimensions of parameters.

IV.2. An example of the VLBI observation schedule used here

Since we investigate only the SDE (i.e. \( X'X \)), we use only the data of the observation schedule (i.e. a time-table of \( s_t \)) and do not use the data of \( \tau \). By courtesy of the Radio Research Laboratory (RRL) of Japan, we can utilize the VLBI observation schedule which was actually used in an early system-level experiment performed between Kashima branch of RRL and Mojave station of NASA, USA, on January 23, 1984 (N. Kawano and F. Takahashi 1984, private communication).

In order to display the distribution of \( s_i \) (\( i = 1, \ldots, 176 \)) in the celestial sphere, let us define the coordinate system as shown in Figure 10, where the direction of \( x \) axis is taken to be from Kashima to Mojave, \( z \) axis toward the center of the common-view zone, and \( y \) axis to form the right-hand system. Since the distance between the two sites is about 8000 km, and since their latitudes are almost the same, the angle of the common-view zone is about 102° in the \( x-z \) plane, and the celestial north pole lies almost in the \( y-z \) plane (see Figure 10). The number of radio sources used in this experiment is 13, and their 176 observed \( s_t \) are depicted in Figure 11, as the projection onto the \( x-y \) plane.

The whole observations were carried out in 28 hours (about 10 minutes per one observation), and brought a great success into the pioneering VLBI works of RRL in Japan.

IV.3. Methods of analysis

In a general geodetic VLBI experiment all the observations in one or two days' experi-
ment are used to make one normal equation of LLST, producing one estimation of \( b \) in which the theoretical time variations of \( b \) (such as caused by the Earth tide) are corrected. In this work, however, we try to divide the experiment into several subsets of observations each of which gives an estimation of \( b \) and \( \tau_e \) in a short time-span (e.g. 2 hours). By such a method of reduction we can see the short-time variation of \( b \), and we hope to see, for instance, whether or not an abrupt change is happened just before a great earthquake. In the former case (i.e. the case of one normal equation) the distribution of \( s_t \) is uniform and the error ellipsoid is not elongated so much. In the latter case (i.e. dividing into many normal equations), however, some \( s_t \)'s in short time-intervals are not uniform while some are uniform, and then we can see various SDE's and their dependences on \( s_t \) distributions.

IV.4. Results

We divide the whole observations into 14 subsets of 2-hours' observations. In Figure 12 we show some examples of the \( s_t \) distributions. In Table I the eigen values and eigen vectors of the first subset is shown.

It is seen in the Table that two principal axes almost lie in the direction of \( x \) and \( y \) axes, while the others almost lie in the t-z plane and not directed toward \( t \) nor \( z \) axes. As also seen in the Table, one principal axis of the SDE has a very small eigen value, that is, the principal axis is very long (i.e. as long as \( a/\sqrt{0.035} \) where \( a \) denotes the expected error of single observation), and another one principal axis is very short (i.e. as short as \( a/\sqrt{11.56} \)).

Therefore, we depict the SDE in Figure 13 as two cross-sections in the t-z and x-y planes, neglecting those components in eigen vectors which are not arrowed in Table I. The length of each principal axis in the Figure is taken to be \( a/\sqrt{\lambda_t} \), where \( \lambda_t \) denotes each eigen value.

We also show the SDE of the whole observations in Figure 14.
FIG. 12. **Examples of 2-hours' Schedules Depicted by the Same Way**

**a:** The First 2-hours, **b:** The Second one, **c:** The Worst in These Fourteen 2-hours' Schedule, **d:** The Best Case

![Diagram](image)

**Table 1—The Eigen Vectors and Eigen Values of an SDE.**

*This is Calculated with the First 2-hours Schedule Which is Shown in Figures 12a and 13a*

<table>
<thead>
<tr>
<th>EIGEN VECTORS</th>
<th>x: 0.93← 0.04 0.33 0.15</th>
<th>y: 0.95← 0.04 0.28 0.58</th>
<th>z: -0.35 -0.12 0.68 -0.29</th>
<th>t: 0.10 0.14 -0.64 0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>EIGEN VALUES</td>
<td>0.144 1.26 0.035 11.56</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### IV.5. Discussions

Note that the shorter are the lengths of principal axes the better is the composition of
Fig. 13. Cross-sections of the SDEs of Observation Schedules shown in Fig. 12.
observations. In Figure 13 we see that the worst case is exemplified by Figure 13c (and Figure 12c), while the best is by Figure 13d (and Figure 12d). It seems that the reason is because in the former case the distribution of $s_t$ does not fulfill the criterion given by Komaki (1985), i.e. it lacks such $s_t$ as (0, 0, 1), while the latter fulfills the criterion (compare Figure 12c and 12d). It is noteworthy that the efficiency of the observations can be judged "quantitatively" by the SDE analysis. However, it should be also noted that the actual criteria for better observations are derived by different investigations such as given by Komaki (1985).

Next, let us show the method to find an optimum estimation of a linear combination of unknown parameters from a given schedule of observations. As shown in III.3. the s.d.'s of estimated $r_e$ and $b_e$ are given as the projections of the SDE onto y or z axes, respectively. Therefore they have large errors in the case of Figure 13c. However, the errors are correlated with each other, that is, if $r_e$ has a large positive error then $b_e$ has a (absolutely) large negative error. Thus the error of a combination $cr_e + b_e$ becomes rather small. Such a combination can be find as the direction of the shortest axis of SDE, and its s.d. is calculated from the largest eigen value, $\lambda_{max}$, as $\sqrt{\lambda_{max}}$. It is seen in Figure 14 that the geodetic VLBI observations generally provide a higher-accuracy estimation of a combination $cr_e + b_e$. If this combination had some geophysical meanings, the VLBI would have been a very sensitive detector of those geophysical phenomena.

It is noteworthy that if the clocks are synchronized so precisely that $\tau_e$ may be treated as a known parameter, the accuracy of the geodetic VLBI can be improved largely, as seen in Figure 14. This is a reason why the "Connected-element interferometer" at Green Bank (Matsakis et al. 1986) or the "VLA" of the NRAO (Florkowski et al. 1985) in USA produce observations with relatively high accuracy though their baselines are small.

V. Conclusions

First we explained graphically the method of SDE analysis of the observational equations, in terms of LLST. We showed relations of some error parameters to the SDE. It was shown that the s.d. of each unknown parameter is given as the projection of the SDE, then we showed that if the SDE has an elongated axis (which is caused when the combination
of the observational equations is improper), large errors occur in many parameter estima-
tions. The SDE analysis provides a way to judge quantitatively the efficiency of the ob-
servational equations.

Looking at the eigen values and eigen vectors calculated in the SDE analysis, we can
find an optimum estimation of a linear-combinations of unknown parameters even in the
case that the combination of the observational equations is improper.

The SDE analysis was applied to the actual observational equations of a geodetic VLBI
experiment, and its usefulness was explained. Some of the characteristics of the geodetic
VLBI observations were pointed out.

HITOTSUBASHI UNIVERSITY

ACKNOWLEDGEMENTS

The author expresses his hearty thanks to the VLBI staff members of Kashima branch
of the RRL, especially to Drs. N. Kawano and F. Takahashi for forgiving the author to use
their unpublished data of the observation schedule. This work was supported in part by
the Scientific Research Fund of the Ministry of Education, Science, and Culture (60460009).
The computation was carried out with the FACOM-M360 at Computer Center of Hitotsu-
bashi University.

REFERENCES

Bjerhammer, A. 1973, Theory of Errors and Generalized Matrix Inverses (Elsevier Scientific
Sons, New York).