X-565-67-601

NASA TM X-63127

THE THEORY AND APPLICATION OF LEAST SQUARES

GENE A. SMITH

MAR 1968
RECEIVED
NASA STI FACILITY
INPUT BRANCH

GPO PRICE $ ______

CFSTI PRICE(S) $ ______

Hard copy (HC) 3.00

Microfiche (MF) .65

ff 653 July 65

DECEMBER 1967

NASA

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

N68-19109

(ACCESSION NUMBER) (THRU)

45 (PAGES) 1 (CODE)

TMX-63127 (NASA CR OR TMX OR AD NUMBER) 19 (CATEGORY)

FACILITY FORM 602

X-565-67-601

THE THEORY AND APPLICATION OF LEAST SQUARES

An Analytical Approach to Curvefitting, Data Smoothing, and Solution of Many Overdetermined Sets of Equations

Gene A. Smith

December 1967

Goddard Space Flight Center
Greenbelt, Maryland

CONTENTS

PRECEDING PAGE BLANK NOT FILMED.

PREFACE

This paper is intended for those unfamiliar with the method of least squares. It was written to present completely a few examples of easily programmed equations which serve a variety of purposes and to introduce without extensive development the more complex general non-linear equation treatment.

PRECEDING PAGE BLANK NOT FILMED.

THE THEORY AND APPLICATION OF LEAST SQUARES

Simple Linear Equation

A problem area in science is that of determining the proper formula for relating observations. Even when an equation which correlates experimental data is known, the best scaling and transformation coefficients are initially uncertain.

For example, assume a simple set of observations of X and a corresponding set of measurements of Y are plotted:

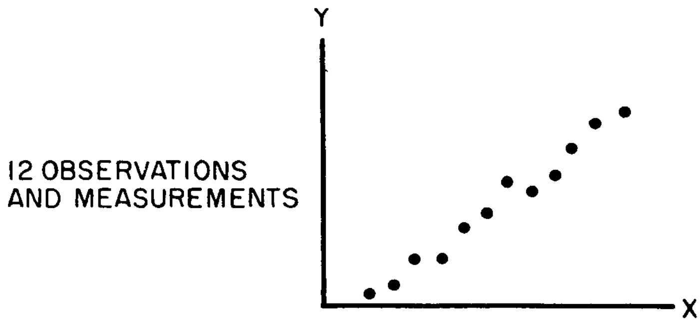


Figure 1

A reasonable assumption would be that there is a linear relationship between X and Y. That is,

$$Y = AX + B \tag{1}$$

This is an equation for a straight line. But what choice of A and B best represents the data? For any two choices of data points, a solution for A and B can be obtained from (1). This supplies a straight line through those two points. For fig. 1, this is $12!/10!2! = 66$ possibilities. Certainly, it would be too time consuming examining each possible line with respect to the others to see which one appears best. Often a straight edge is aligned visually on the plot in an approximation of a close fit.

For many purposes either of the above rough techniques may be satisfactory; however, a more analytic treatment is normally preferred. This treatment should be an objective, general approach. That is, the problem is stated, equations proposed, observations or measurements made, and solutions derived from the equations.

For equation (1) and fig. 1, the problem is to determine by a mathematical algorithm the best straight line for the given points. Twelve observations relating X and Y are given. The coefficients A and B of (1) are, therefore, overdetermined.

$$\begin{aligned} Y_1 &\approx AX_1 + B \\ Y_2 &\approx AX_2 + B \\ &\cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \\ Y_n &\approx AX_n + B \end{aligned} \tag{2a}$$

The general set of n equations is represented in (2a). Our example from fig. 1 has $n = 12$. The approximate nature of the equations is eliminated by explicitly including the residuals Δ_i, which are the differences from observations Y_i and the straight line computed values $(AX_i + B)$,

$$Y_i - (AX_i + B) = \Delta_i \tag{3}$$

for $i = 1, 2, \ldots, n$. The equations (3) are known as <u>condition equations</u>. We should like to find values for A and B such that a minimum error in Y could be expected for additional observations of X.

A method which achieves this aim is that of least squares, so named because the object is to minimize the sum of the squares of the residuals or observation vs. computation errors. That is,

$$\sum_{i=1}^{n} [Y_i - (AX_i + B)]^2 = \sum_{i=1}^{n} \Delta_i^2 = \text{minimum}. \tag{4}$$

Expanding (4) gives,

$$\sum_{i=1}^{n} Y_i^2 + A^2 \sum_{i=1}^{n} X_i^2 + nB^2 - 2A \sum_{i=1}^{n} Y_i X_i - 2B \sum_{i=1}^{n} Y_i + 2AB \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \Delta_i^2 . \quad (4a)$$

To minimize $\sum_{i=1}^{n} \Delta_i^2$ with respect to particular coefficients A and B, the partial derivatives of $\Sigma\, \Delta_i^2$ with respect to A and B must be zero.

$$\frac{\partial \sum_{i=1}^{n} \Delta_i^2}{\partial A} = 0 = 2A \sum_{i=1}^{n} X_i^2 - 2 \sum_{i=1}^{n} Y_i X_i + 2B \sum_{i=1}^{n} X_i \quad (5)$$

$$\frac{\partial \sum_{i=1}^{n} \Delta_i^2}{\partial B} = 0 = 2nB - 2 \sum_{i=1}^{n} Y_i + 2A \sum_{i=1}^{n} X_i \quad (6)$$

where (5) and (6) are <u>normal equations</u>. Thus, with two equations in two unknowns, a solution is given by:

$$B = \frac{\sum_{i=1}^{n} Y_i - A \sum_{i=1}^{n} X_i}{n}$$

$$A \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} Y_i X_i + \frac{\sum_{i=1}^{n} Y_i - A \sum_{i=1}^{n} X_i}{n} \sum_{i=1}^{n} X_i = 0$$

or

$$nA \sum_{i=1}^{n} X_i^2 - n \sum_{i=1}^{n} Y_i X_i + \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i - A \left(\sum_{i=1}^{n} X_i \right)^2 = 0$$

whence,

$$A = \frac{n \sum_{i=1}^{n} Y_i X_i - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i \right)^2} . \tag{7}$$

Then,

$$B = \frac{-n \sum_{i=1}^{n} X_i Y_i \left(\sum_{i=1}^{n} X_i \right) + \left(\sum_{i=1}^{n} X_i \right)^2 \sum_{i=1}^{n} Y_i + n \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \left(\sum_{i=1}^{n} X_i \right)^2 \sum_{i=1}^{n} Y_i}{n \left[n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i \right)^2 \right]}$$

whence,

$$B = \frac{\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i Y_i \left(\sum_{i=1}^{n} X_i \right)}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i \right)^2} \tag{8}$$

Equation (2) can be expressed in matrix notation.

$$\underset{(n \times 2)}{\begin{bmatrix} X_1 & 1 \\ X_2 & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ X_n & 1 \end{bmatrix}} \underset{(2 \times 1)}{\begin{bmatrix} A \\ B \end{bmatrix}} \approx \underset{(n \times 1)}{\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{bmatrix}}$$

But this is just $A\Sigma X_i^2 + B\Sigma X_i = \Sigma X_i Y_i$ and $A\Sigma X_i + nB = \Sigma Y_i$, equation (5) and (6).

Q.E.D.

Singularity of a matrix is indicated if the determinant of it is zero, so that if $X^T X$ is nonsingular, then an inverse exists and P can be solved for by pre-multiplying (10) by $(X^T X)^{-1}$, the inverse of $X^T X$.

$$(X^T X)^{-1} (X^T X) P = (X^T X)^{-1} X^T Y \tag{11a}$$

$$P = (X^T X)^{-1} X^T Y \tag{11b}$$

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \dfrac{n}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} & \dfrac{-\sum_{i=1}^{n} X_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \\ \dfrac{-\sum_{i=1}^{n} X_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_j\right)^2} & \dfrac{\sum_{i=1}^{n} X_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} X_i Y_i \\ \sum_{i=1}^{n} Y_i \end{bmatrix} \tag{11c}$$

However, this is equivalent to (7) and (8).

As was indicated before equation (9), an analytic justification is necessary to prove that the matrix representation of (9) is equivalent to a least squares treatment. Observe that the condition equations (3) were not given an exact expression in a matrix form. This is simply done by:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} - \begin{bmatrix} X_1 & 1 \\ X_2 & 1 \\ \vdots & \vdots \\ X_n & 1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_n \end{bmatrix} \tag{12a}$$

A simplification of the solution is achieved by replacing the approximate sign by an equal sign (the justification for this step is given on page #6). Matrices may then be expressed by

$$\underset{(n\times 2)}{X}\ \underset{(2\times 1)}{P} = \underset{(n\times 1)}{Y} \tag{9}$$

Because the matrix X is not a square one ($n\times n$), inversion directly is not possible. This dilemma is resolved by premultiplying (9) by the transpose of X.

$$\underbrace{\underset{(2\times n)}{X^T}\ \underset{(n\times 2)}{X}}_{(2\times 2)}\ \underset{(2\times 1)}{P} = \underset{(2\times n)}{X^T}\ \underset{(n\times 1)}{Y} \tag{10}$$

This is the matrix form of the normal equations (5) and (6). Because the transpose matrix product of a matrix is a square matrix, the product matrix (X^TX) is a 2×2 square matrix.

We are justified in calling both (10) and the set (5)-(6) normal equations because they are equivalent. This is shown by expanding (10).

$$X^T X \begin{bmatrix} A \\ B \end{bmatrix} = X^T Y$$

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} X_1 & 1 \\ X_2 & 1 \\ \vdots & \vdots \\ X_n & 1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

$$\begin{bmatrix} \sum_{i=1}^{n} X_i^2 & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & n \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} X_i Y_i \\ \sum_{i=1}^{n} Y_i \end{bmatrix}$$

or

$$Y - XP = V. \tag{12b}$$

Equations (12) are the matrix form of the condition equations (3). For matrices, the sum of the squares of the residuals is represented by:

$$(Y - XP)^T (Y - XP) = V^T V. \tag{13a}$$

Manipulating (12) by matrix rules yields,

$$Y^T Y - Y^T XP - P^T X^T Y + P^T X^T XP = V^T V \tag{13b}$$

but for differentiable matrices, $V^T V$ may be minimized with respect to P by:

$$\frac{\partial V^T V}{\partial P} = 0 = -(Y^T X)^T - X^T Y - 2X^T XP \tag{14}$$

$$0 = 2X^T XP - 2X^T Y \tag{14a}$$

which may be solved for P to give

$$P = (X^T X)^{-1} X^T Y. \tag{15}$$

This is just our result (10), justifying our simplification made at that time.

Three Dimensional Orthogonal Transformation

A second example which is more complex, more general, and more interesting is the three-dimensional orthogonal vector transformation (Orthogonality implies mutually perpendicular axes). This transformation serves to rotate and translate a vector $\vec{R}$ into another vector $\vec{R}'$ (fig. 2). Another way of viewing the transformation is as a rotation and translation of a coordinate system into another reference frame (fig. 3).

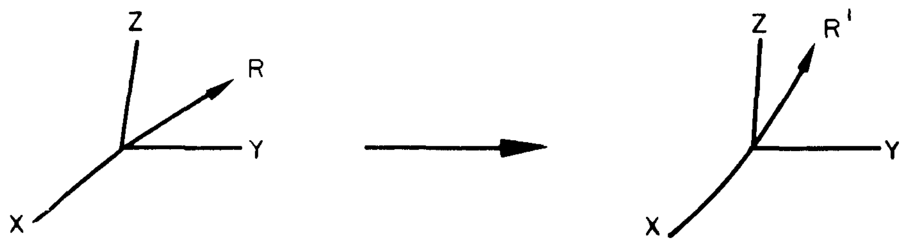


Figure 2

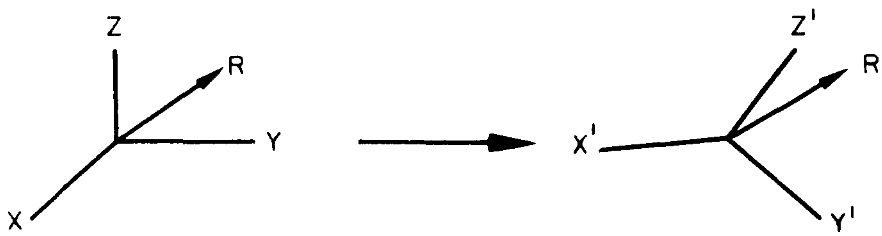


Figure 3

An axial rotation is a rotation of a frame about one of the axes X, Y, or Z. At most, three axial rotations, each representable as a direction cosine matrix in form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}, \text{ or } \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

are necessary to perform any transformation. We shall specifically examine a set of three axial rotations defined by:

(1) rotation around X axis by roll angle ω

(2) rotation around new Y axis by pitch angle ϕ

(3) rotation around newest Z axis by yaw angle κ.

A large number of possible choices exist however, e.g.

pitch (ϕ), roll (ω), yaw (κ) about Y - X - Z

heading (H), roll (ω), pitch (ϕ) about Z - X - Y

heading (H), pitch (ϕ), roll (ω) about Z - Y - X

azimuth (α), tilt (t), swing (s) about Z - X - Z

azimuth (α), elevation (h), swing (s) about Z - Y - Z etc.

The direction cosine matrices are represented by:

$$\left.\begin{aligned} A_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{pmatrix} \\ A_2 &= \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix} \\ A_3 &= \begin{pmatrix} \cos\kappa & \sin\kappa & 0 \\ -\sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}\right\} \tag{16}$$

A vector R can now be transformed into R′ by a series of three axial rotations:

$$R_1 = A_1 R; \quad R_2 = A_2 R_1; \quad R' = A_3 R_2. \tag{17}$$

Instead of making three separate steps, a single rotation represented by direction cosine matrix $A = A_3 A_2 A_1$ is sufficient. "A" is thereby a product matrix (3×3) of three orthogonal matrices and is therefore orthogonal itself.

$$A = \begin{pmatrix} \cos\phi\cos\kappa & \cos\omega\sin\kappa + \sin\omega\sin\phi\cos\kappa & \sin\omega\sin\kappa - \cos\omega\sin\phi\cos\kappa \\ -\cos\phi\sin\kappa & \cos\omega\cos\kappa - \sin\omega\sin\phi\sin\kappa & \sin\omega\cos\kappa + \cos\omega\sin\phi\sin\kappa \\ \sin\phi & -\sin\omega\cos\phi & \cos\omega\cos\phi \end{pmatrix} \quad (18)$$

Note: All rotations considered in this section are counter clock-wise in accord with current scientific convention.

A review of a few properties of orthogonal direction cosine matrices may be useful.

If A transforms vector

$$\vec{R} <=> \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \text{ into } \vec{R}' <=> \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix},$$

then

$$R' = AR \quad (19)$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{pmatrix} \cos\widehat{X'X} & \cos\widehat{X'Y} & \cos\widehat{X'Z} \\ \cos\widehat{Y'X} & \cos\widehat{Y'Y} & \cos\widehat{Y'Z} \\ \cos\widehat{Z'X} & \cos\widehat{Z'Y} & \cos\widehat{Z'Z} \end{pmatrix}. \quad (18a)$$

That is, the elements of a direction cosine matrix are the cosines of the angles between two axes, one from each reference frame.

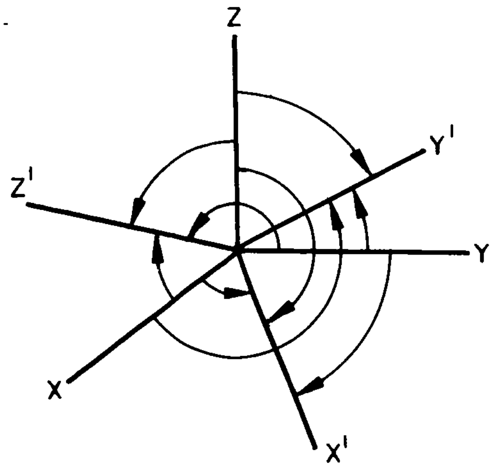


Figure 4

Thus the relations $a_{11} = \cos \widehat{X'X} = \cos\phi \cos\kappa$, $a_{12} = \cos \widehat{X'Y'} = \cos\omega \sin\kappa + \sin\omega \sin\phi \cos\kappa$, . . . $a_{33} = \cos \widehat{Z'Z} = \cos\omega \cos\phi$ supply the dependence of the angles between the X - Y - Z axes and the X' - Y' - Z' axes as functions of the rotation angles ω, ϕ, κ. Explicitly, "A" is produced by:

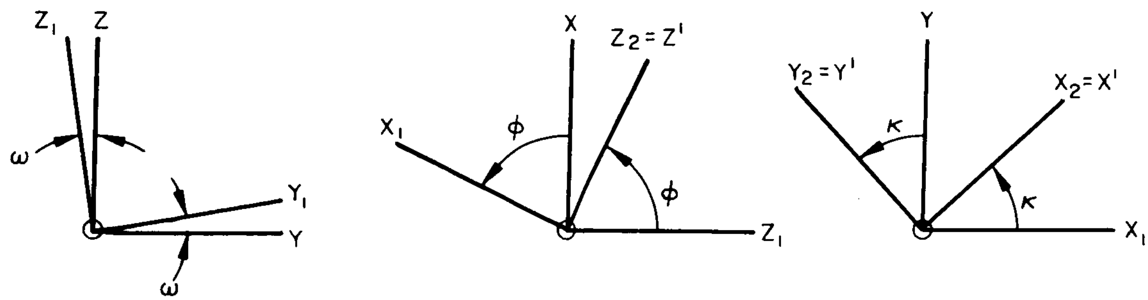


Figure 5

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \longrightarrow \begin{pmatrix} X \\ Y_1 \\ Z_1 \end{pmatrix}, \quad \begin{pmatrix} X \\ Y_1 \\ Z_1 \end{pmatrix} \longrightarrow \begin{pmatrix} X_1 \\ Y_1 \\ Z_2 \end{pmatrix}, \quad \begin{pmatrix} X_1 \\ Y_1 \\ Z_2 \end{pmatrix} \longrightarrow \begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} = \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix}$$

via $A_1 (\omega)$ via $A_2 (\phi)$ via $A_3 (\kappa)$

For any orthogonal matrix A,

$$\sum_{j=1}^{3} a_{ij}\, a_{kj} = \delta_{ik} \quad \text{holds for rows:} \quad i = 1, 2, 3;\ k = 1, 2, 3$$

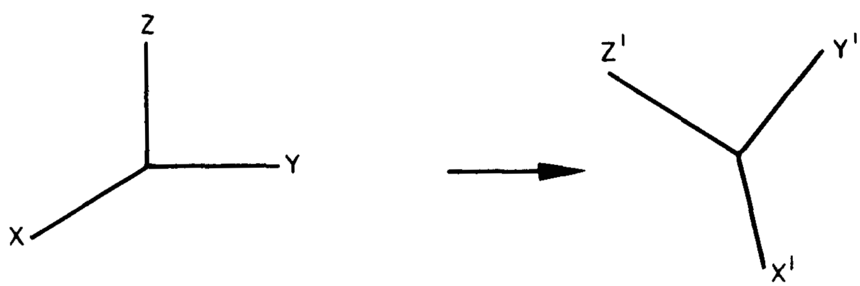
$$\sum_{i=1}^{3} a_{ij}\, a_{ik} = \delta_{jk} \quad \text{holds for columns:} \quad j = 1, 2, 3;\ k = 1, 2, 3$$

where

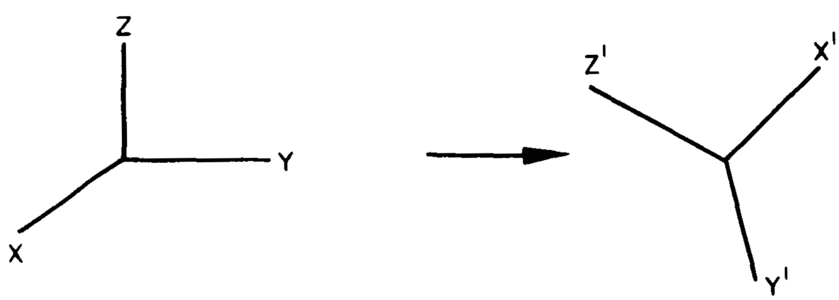
$$\delta_{ik} = 1 \text{ for } i = k; \quad \delta_{ik} = 0 \text{ for } i \neq k$$

$$\delta_{jk} = 1 \text{ for } j = k; \quad \delta_{jk} = 0 \text{ for } i \neq k$$

det A $= |A| = \pm 1$, $+1$ when a right (left)-handed coordinate system is transformed into another right (left)-handed system



and -1 when a right (left)-handed system is transformed into a left (right)-handed one.



A number of graphic techniques have been devised to aid in visualization of the transformation. Two are illustrated below.

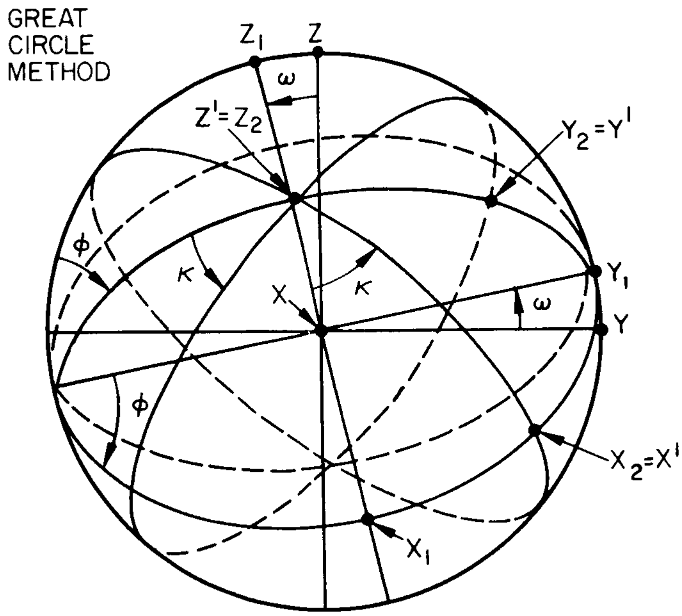


Figure 6(a)

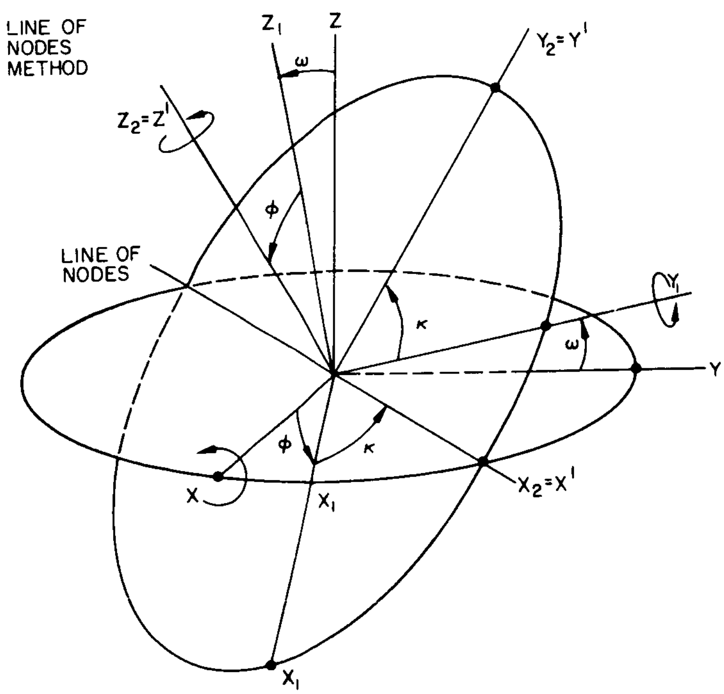


Figure 6(b)

For our purposes, we wish to choose a transformation

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} A & B & C \\ F & G & H \\ D & E & I \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{19a}$$

or

$$R' = A\,R. \tag{19b}$$

The peculiar choice of elements for the direction cosine matrix will be useful in a later development. Now we turn to the solution of (19) by least squares.

If n vectors are measured in both of two coordinate systems (where $n \geq 3$, since each vector supplies 3 points),

$$\underset{(3 \times n)}{\begin{pmatrix} X_1' & X_2' & \cdots & X_n' \\ Y_1' & Y_2' & \cdots & Y_n' \\ Z_1' & Z_2' & \cdots & Z_n' \end{pmatrix}} = \underset{(3\times3)}{A} \underset{(3 \times n)}{\begin{pmatrix} X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \\ Z_1 & Z_2 & \cdots & Z_n \end{pmatrix}} \tag{20a}$$

$$R_n' = A\,R_n\,. \tag{20}$$

The subscript n denotes the number of vectors in the matrix. To make the right-hand side invertible, post multiply by R_n^T.

$$R_n'\,R_n^T = A\,R_n\,R_n^T \tag{21}$$

We may solve for A as we did for P in the previous example.

$$R_n'\,R_n^T\,(R_n\,R_n^T)^{-1} = A\,R_n\,R_n^T\,(R_n\,R_n^T)^{-1} \tag{22a}$$

$$A = R_n'\,R_n^T\,(R_n\,R_n^T)^{-1} \tag{22b}$$

Explicitly,

$$R_n R_n^T = \begin{pmatrix} X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \\ Z_1 & Z_2 & \cdots & Z_n \end{pmatrix} \begin{pmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ \vdots & \vdots & \vdots \\ X_n & Y_n & Z_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n X_i & \sum_{i=1}^n X_i Y_i & \sum_{i=1}^n X_i Z_i \\ \sum_{i=1}^n Y_i X_i & \sum_{i=1}^n Y_i^2 & \sum_{i=1}^n Y_i Z_i \\ \sum_{i=1}^n Z_i X_i & \sum_{i=1}^n Z_i Y_i & \sum_{i=1}^n Z_i^2 \end{pmatrix}$$

$$R_n' R_n^T = \begin{pmatrix} X_1' & X_2' & \cdots & X_n' \\ Y_1' & Y_2' & \cdots & Y_n' \\ Z_1' & Z_2' & \cdots & Z_n' \end{pmatrix} \begin{pmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ \vdots & \vdots & \vdots \\ X_n & Y_n & Z_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n X_i' X_i & \sum_{i=1}^n X_i' Y_i & \sum_{i=1}^n X_i' Z_i \\ \sum_{i=1}^n Y_i' X_i & \sum_{i=1}^n Y_i' Y_i & \sum_{i=1}^n Y_i' Z_i \\ \sum_{i=1}^n Z_i' X_i & \sum_{i=1}^n Z_i' Y_i & \sum_{i=1}^n Z_i' Z_i \end{pmatrix}$$

Projective Transformation

Consider a scaled form of equation (18),

$$\begin{pmatrix} \frac{X}{Z} \\ \\ \frac{Y}{Z} \\ \\ \frac{Z}{Z} \end{pmatrix} = \begin{pmatrix} A & B & C \\ \\ F & G & H \\ \\ D & E & I \end{pmatrix} \begin{pmatrix} \frac{x}{z} \\ \\ \frac{y}{z} \\ \\ \frac{z}{z} \end{pmatrix} . \tag{23a}$$

We wish to project the points in an xy plane onto points in an XY plane. That is, a projection with $Z = z =$ constant. We thus have:

$$\left.\begin{aligned} \frac{X}{Z} &= A\frac{x}{z} + B\frac{y}{z} + C\frac{z}{z} \\ \frac{Y}{Z} &= F\frac{x}{z} + G\frac{y}{z} + H\frac{z}{z} \\ \frac{Z}{Z} &= D\frac{x}{z} + E\frac{y}{z} + I\frac{z}{z} \end{aligned}\right\} \tag{23b}$$

Moreover, if the constant $Z = z$ is chosen to be one,

$$\left.\begin{aligned} X &= \frac{Ax + By + C}{Dx + Ey + I} \\ Y &= \frac{Fx + Gy + H}{Dx + Ey + I} \end{aligned}\right\} \tag{24}$$

For an orthogonal-conformal transformation, as I approaches the value of one, $D = E = C = H = 0$. If the requirement for conformality is now relaxed such that D, E, C, and H may not be zero for $I = 1$, then

$$X = \frac{Ax + By + C}{Dx + Ey + 1} \tag{25}$$

$$Y = \frac{Fx + Gy + H}{Dx + Ey + 1} \tag{26}$$

This is an eight parameter nonconformal projective transformation. It is used for mapping planar surfaces into other planar surfaces.

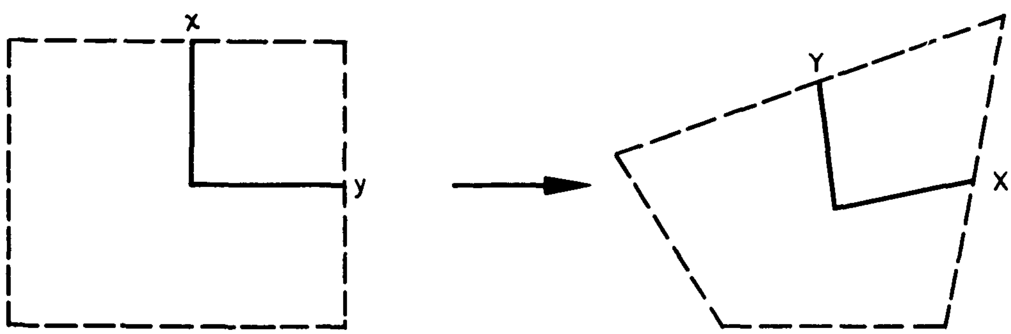
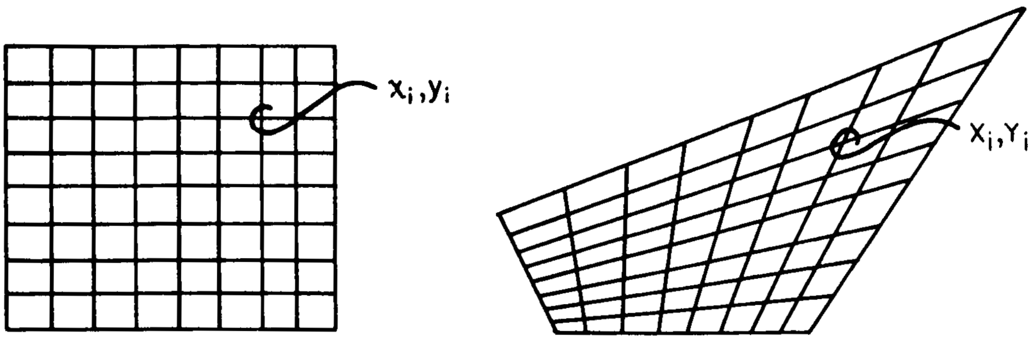


Figure 7

Specifically, an oblique photograph from the air may be transformed point by point into a map of the same area. Also, a linearly distorted and scaled network or grid may be correlated to a reference grid.

We shall be concerned with how the best choice of the coefficients A through H are made. Assume a network with n grid intersections which are measured. Let this be a reference grid. Similarly measure the n intersections of a distorted grid. The ith reference grid intersection is denoted by x_i, y_i and the corresponding ith distorted grid intersection is denoted by X_i, Y_i. Since there are eight parameters to be determined, there must be at least four measured intersections (each intersection supplies two equations). A "good" pattern of points is recommended. That is, not all points from the same area, in a line, or in any way likely to create a bias in the solution should be chosen. After a solution for the coefficients is made, any arbitrary point within the point pattern on the distorted grid may be measured and transformed into the reference grid.



REFERENCE GRID

DISTORTED GRID

Figure 8

In many cases of interest, orthogonality and conformality are almost preserved; that is, $A \approx G$, $B \approx -F$, $D \approx E \approx 0$.

The following 2n equations are formed when n intersections from reference and distorted grids are measured.

$$
\begin{aligned}
X_1 &\approx \frac{Ax_1 + By_1 + C}{Dx_1 + Ey_1 + 1} & Y_1 &\approx \frac{Fx_1 + Gy_1 + H}{Dx_1 + Ey_1 + 1} \\
X_2 &\approx \frac{Ax_2 + By_2 + C}{Dx_2 + Ey_2 + 1} & Y_2 &\approx \frac{Fx_2 + Gy_2 + H}{Dx_2 + Ey_2 + 1} \\
&\vdots \quad \vdots & &\vdots \quad \vdots \\
X_n &\approx \frac{Ax_n + By_n + C}{Dx_n + Ey_n + 1} & Y_n &\approx \frac{Fx_n + Gy_n + H}{Dx_n + Ey_n + 1}
\end{aligned}
\tag{27}
$$

Expressed as condition equations,

$$
\begin{aligned}
X_1 - (Ax_1 + By_1 + C - Dx_1X_1 - Ey_1X_1) &= \Delta_{11} \\
X_2 - (Ax_2 + By_2 + C - Dx_2X_2 - Ey_2X_2) &= \Delta_{12} \\
\vdots \qquad\qquad & \qquad \vdots \\
X_n - (Ax_n + By_n + C - Dx_nX_n - Ey_nX_n) &= \Delta_{1n} \\
Y_1 - (Fx_1 + Gy_1 + H - Dx_1Y_1 - Ey_1Y_1) &= \Delta_{21} \\
Y_2 - (Fx_2 + Gy_2 + H - Dx_2Y_2 - Ey_2Y_2) &= \Delta_{22} \\
\vdots \qquad\qquad & \qquad \vdots \\
Y_n - (Fx_n + Gy_n + H - Dx_nY_n - Ey_nY_n) &= \Delta_{2n} \quad .
\end{aligned}
\tag{28}
$$

Square the residuals Δ_{ji}, sum over j and i, and minimize with respect to the coefficients.

$$
\frac{\partial \sum_{j,i}^{2,n} \Delta_{ji}^2}{\partial A} = 0, \quad \frac{\partial \sum_{j,i}^{2,n} \Delta_{ji}^2}{\partial B} = 0, \quad \cdots, \quad \frac{\partial \sum_{j,i}^{2,n} \Delta_{ji}^2}{\partial H} = 0
$$

Once again by resorting to matrix notation, the solution is simplified.

$$
\begin{array}{ccc}
(2n \times 8) & (8\times 1) & (2n\times 1) \\
\begin{bmatrix}
x_1 & y_1 & 1 & -x_1 X_1 & -y_1 X_1 & 0 & 0 & 0 \\
x_2 & y_2 & 1 & -x_2 X_2 & -y_2 X_2 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_n & y_n & 1 & -x_n X_n & -y_n X_n & 0 & 0 & 0 \\
0 & 0 & 0 & -x_1 Y_1 & -y_1 Y_1 & x_1 & y_1 & 1 \\
0 & 0 & 0 & -x_2 Y_2 & -y_2 Y_2 & x_2 & y_2 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & -x_n Y_n & -y_n Y_n & x_n & y_n & 1
\end{bmatrix}
&
\begin{bmatrix}
A \\ B \\ C \\ D \\ E \\ F \\ G \\ H
\end{bmatrix}
=
&
\begin{bmatrix}
X_1 \\ X_2 \\ \vdots \\ X_n \\ Y_1 \\ Y_2 \\ \vdots \\ Y_n
\end{bmatrix}
\end{array}
\qquad (29a)
$$

$$
X \ P = Y \qquad (29b)
$$

Again, square matrix X by premultiplying by the transpose of X,

$$
X^T \, X \, P = X^T \, Y \qquad (30)
$$

where $X^T X =$

$$\begin{bmatrix}
\sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} x_i & -\sum_{i=1}^{n} x_i^2 X_i & -\sum_{i=1}^{n} x_i y_i X_i & 0 & 0 & 0 \\
\sum x_i y_i & \sum_{i=1}^{n} y_i^2 & \sum_{i=1}^{n} y_i & -\sum_{i=1}^{n} x_i y_i X_i & -\sum_{i=1}^{n} y_i^2 X_i & 0 & 0 & 0 \\
\sum_{i=1}^{n} x_i & \sum_{i=1}^{n} y_i & n & -\sum_{i=1}^{n} x_i X_i & -\sum_{i=1}^{n} y_i X_i & 0 & 0 & 0 \\
-\sum_{i=1}^{n} x_i^2 X_i & -\sum_{i=1}^{n} x_i y_i X_i & -\sum_{i=1}^{n} x_i X_i & \begin{pmatrix} \sum x_i^2 X_i^2 + \\ \sum x_i^2 Y_i^2 \end{pmatrix} & \begin{pmatrix} \sum y_i x_i X_i^2 + \\ \sum y_i x_i Y_i^2 \end{pmatrix} & -\sum_{i=1}^{n} x_i^2 Y_i & -\sum_{i=1}^{n} x_i y_i Y_i & -\sum_{i=1}^{n} x_i Y_i \\
-\sum_{i=1}^{n} x_i y_i X_i & -\sum_{i=1}^{n} y_i^2 X_i & -\sum_{i=1}^{n} y_i X_i & \begin{pmatrix} \sum x_i y_i X_i^2 + \\ \sum x_i y_i Y_i^2 \end{pmatrix} & \begin{pmatrix} \sum y_i^2 X_i^2 + \\ \sum y_i^2 Y_i^2 \end{pmatrix} & -\sum_{i=1}^{n} x_i y_i Y_i & -\sum_{i=1}^{n} y_i^2 Y_i & -\sum_{i=1}^{n} y_i Y_i \\
0 & 0 & 0 & -\sum_{i=1}^{n} x_i^2 Y_i & -\sum_{i=1}^{n} x_i y_i Y_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} x_i \\
0 & 0 & 0 & -\sum_{i=1}^{n} x_i y_i Y_i & -\sum_{i=1}^{n} y_i^2 Y_i & \sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} y_i^2 & \sum_{i=1}^{n} y_i \\
0 & 0 & 0 & -\sum_{i=1}^{n} x_i Y_i & -\sum_{i=1}^{n} y_i Y_i & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} y_i & n
\end{bmatrix}$$

and

$$X^T Y = \begin{bmatrix} \sum x_i X_i \\ \sum y_i X_i \\ \sum X_i \\ \sum x_i X_i^2 + \sum x_i Y_i^2 \\ \sum y_i X_i^2 + \sum y_i Y_i^2 \\ \sum x_i Y_i \\ \sum y_i Y_i \\ \sum Y_i \end{bmatrix}$$

Inverting $(X^T X)$ and premultiplying into (30),

$$(X^T X)^{-1} X^T X \; P = (X^T X)^{-1} X^T Y \tag{31a}$$

$$P = (X^T X)^{-1} X^T Y \tag{31b}$$

To reemphasize a statement previously made, points computed from the coefficient parameters should lie within or very near the spread of points used for calculating the coefficients.

The projective transformation serves to analytically project a plane (x , y) onto another (X , Y) where one plane is tilted with respect to the other and the rays of the projection may be viewed as parallel, diverging, or converging.

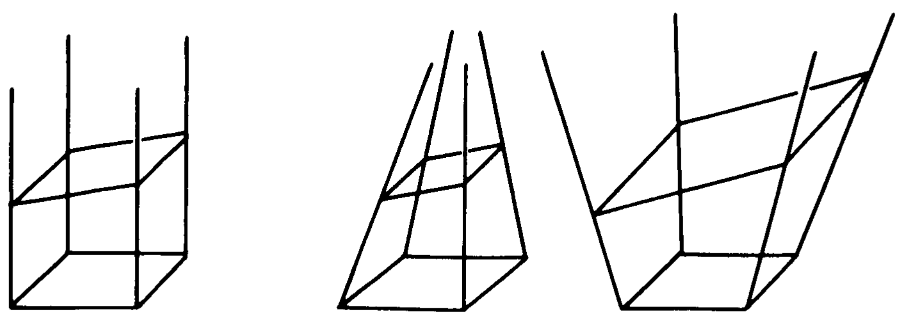


Figure 9

N Degree Polynomial

As a third example, the general n degree polynomial transformation is illuminating. In many situations, data points can be plotted such that a simple curve can adequately represent the observations.

$$Y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \tag{32}$$

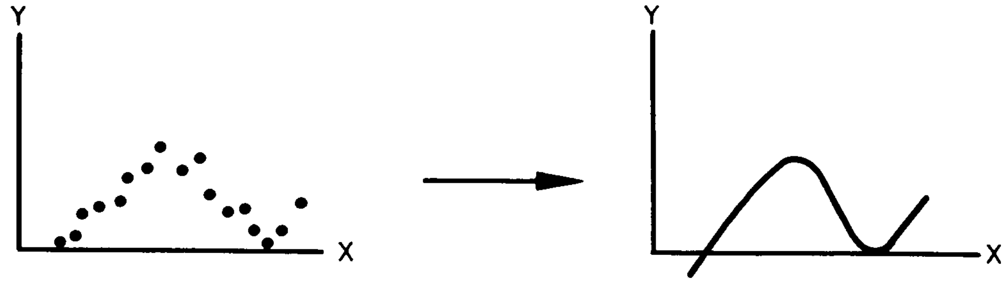


Figure 10

A least squares solution for coefficients a_i can be formed whenever $m \geq n$ observations in X are made.

$$\begin{aligned}
y_1 &\approx a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n \\
y_2 &\approx a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_n x_2^n \\
&\vdots \\
y_m &\approx a_0 + a_1 x_m + a_2 x_m^2 + \cdots + a_n x_m^n
\end{aligned} \tag{33}$$

for $m \geq n$.

The condition equations in matrix notation are:

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} - \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ 1 & x_3 & x_3^2 & \cdots & x_3^n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{bmatrix} = \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta_m \end{bmatrix} \tag{34}$$

$$Y - X \ A = V. \tag{34a}$$

Normal equations are formed via,

$$- A^T X^T Y + A^T X^T X A + Y^T Y - Y^T X A = V^T V. \tag{35}$$

Minimizing $V^T V$ as before (eq. (14)) gives normal equations,

$$X^T Y - X^T X A = 0. \tag{36}$$

Whence,

$$A = (X^T X)^{-1} X^T Y,$$

where

$$X^T X = \begin{bmatrix} m & \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \cdots & \sum_{i=1}^m x_i^n \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 & \cdots & \sum_{i=1}^m x_i^{n+1} \\ \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 & \sum_{i=1}^m x_i^4 & \cdots & \sum_{i=1}^m x_i^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m x_i^n & \sum_{i=1}^m x_i^{n+1} & \sum_{i=1}^m x_i^{n+2} & \cdots & \sum_{i=1}^m x_i^{2n} \end{bmatrix}$$

For demonstrative purposes, the above will be hand computed.

$$
\begin{aligned}
y_1 - (a_0 + a_1 x_1 + \cdots + a_n x_1^n) &= \Delta_1 \\
y_2 - (a_0 + a_1 x_2 + \cdots + a_n x_2^n) &= \Delta_2 \\
\vdots \qquad\qquad & \quad \vdots \\
y_m - (a_0 + a_1 x_m + \cdots + a_n x_m^n) &= \Delta_m
\end{aligned}
\tag{34b}
$$

Squaring and summing the residuals,

$$
\begin{aligned}
\sum_{i=1}^{m} \Delta_i^2 = \sum_{i=1}^{m} y_i^2 &+ m\, a_0 + 2\, a_0 a_1 \sum_{i=1}^{m} x_i + 2\, a_0 a_2 \sum_{i=1}^{m} x_i^2 + \cdots \\
&+ 2\, a_0 a_n \sum_{i=1}^{m} x_i^n + a_1^2 \sum x_i^2 + 2\, a_1 a_2 \sum x_i^3 + \cdots \\
&+ 2\, a_1 a_n \sum x_i^{n+1} + a_2^2 \sum x_i^4 + 2\, a_2 a_3 \sum x_i^5 + \cdots \\
&+ 2\, a_2 a_n \sum x_i^{n+2} + \cdots + a_n^2 \sum x^{2n} - 2\, a_0 \sum y_i \\
&- 2\, a_1 \sum x_i y_i - 2\, a_2 \sum x_i^2 y_i - \cdots - 2\, a_n \sum x_i^n y_i \,.
\end{aligned}
\tag{37}
$$

To minimize the sum of the squares, find the first partials with respect to the coefficients a_i and equate each to zero.

$$\frac{\partial \sum_{i=1}^{m} \Delta_i^2}{\partial a_0} = 2 m a_0 + 2 a_1 \sum_{i=1}^{m} x_i + 2 a_2 \sum_{i=1}^{m} x_i^2 + \cdots + 2 a_n \sum_{i=1}^{m} x_i^n - 2 \sum_{i=1}^{m} y_i = 0$$

$$\frac{\partial \sum_{i=1}^{m} \Delta_i^2}{\partial a_1} = 2 a_0 \sum_{i=1}^{m} x_i + 2 a_1 \sum_{i=1}^{m} x_i^2 + 2 a_2 \sum_{i=1}^{m} x_i^3 + \cdots + 2 a_n \sum_{i=1}^{m} x_i^{n+1} - 2 \sum_{i=1}^{m} x_i y_i = 0 \qquad (38a)$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$\frac{\partial \sum_{i=1}^{m} \Delta_i^2}{\partial a_n} = 2 a_0 \sum_{i=1}^{m} x_i^n + 2 a_1 \sum_{i=1}^{m} x_i^{n+1} + 2 a_2 \sum_{i=1}^{m} x_i^{n+2} + \cdots + 2 a_n \sum_{i=1}^{m} x_n^{2n} - 2 \sum_{i=1}^{m} x_i^n y_i = 0$$

This is $n+1$ equations in $n+1$ unknowns a_j $(j = 0, 1, \ldots, n)$.

The multiplicative factor 2 may be removed, allowing (38a) to be expressed in matrix form,

$$\underset{(n+1)\times(n+1)}{\begin{bmatrix} m & \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 & \cdots & \sum_{i=1}^{m} x_i^n \\ \sum_{i=1}^{m} x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{n+1} \\ \sum_{i=1}^{m} x_i^2 & \sum x_i^3 & \sum x_i^4 & \cdots & \sum x_i^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} x_i^n & \sum x_i^{n+1} & \sum x_i^{n+2} & \cdots & \sum x_i^{2n} \end{bmatrix}} \underset{(n+1)\times 1}{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}} = \underset{(n+1)\times m}{\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_m \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_m^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & x_3^n & \cdots & x_m^n \end{bmatrix}} \underset{m\times 1}{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix}} \qquad (38b)$$

Which can be expressed as

$$(X^T X) \quad A = X^T Y . \tag{38c}$$

Continue solution as for eq. (36).

General Non-Linear Equation Approach

For our last development, we wish to derive a general approach whereby non-linear equations can be solved by least squares. A Taylor's expansion linearization will be used leading to approximate solutions. By an iterative process, the approximate solutions will be improved until changes in the solution become less than a predetermined limit. The generality is further increased by also allowing a limited number of coordinates x_k to be better determined than from observation. For j_{max} = maximum number of coefficients to be found and k_{max} = maximum number of coordinates to be re-evaluated, $m \geq j_{max} + k_{max}$ is necessary for a least squares method to work (m = number of observations).

Assume a general function,

$$X_i = X_i\ (a_j,\ x_k) \tag{39}$$

This is expressible as a polynomial in the coefficient parameters a_j and coordinates x_k.

$$X_i \begin{cases} X_1 X_2 \cdots \\ Y_1 Y_2 \cdots \\ Z_1 Z_2 \cdots \end{cases} \qquad a_j \left\{ A\,B\,C \cdots \right. \qquad x_k \begin{cases} x_1 x_2 \cdots \\ y_1 y_2 \cdots \\ z_1 z_2 \cdots \end{cases}$$

Form condition equations,

$$f_\ell\,(X_i, a_j, x_k) = 0 \tag{40}$$

As implied, linearization can be achieved by expanding the condition equations in a Taylor's Series and dropping second and higher order terms.

$$0 = f_\ell^\circ + \sum_{k=1}^{k_{max}} \left(\frac{\partial f_\ell}{\partial x_k}\right)_0 dx_k + \sum_{j=1}^{j_{max}} \left(\frac{\partial f_\ell}{\partial a_j}\right)_0 da_j = f_\ell \tag{40a}$$

a°_ℓ = initial or interated value of a_j , x°_k = initial or iterated value of x_k , such that $f^\circ_\ell = f_\ell(X_i, a^\circ_j, x^\circ_k)$.

For notational simplicity let

$$\left(\frac{\partial f_\ell}{\partial x_k}\right)_0 = \frac{\partial f_\ell}{\partial x^\circ_k}, \quad \left(\frac{\partial f_\ell}{\partial a_j}\right)_0 = \frac{\partial f_\ell}{\partial a^\circ_j}$$

and

$$0 = f_\ell = f_\ell(X_i, a^\circ_j + \delta_j, x^\circ_k + v_k) = f^\circ_\ell + A_k v_k + B_j \delta_j \tag{40b}$$

In matrix notation the ℓ condition equations become:

$$0 = \underset{\ell_{max} X 1}{\begin{pmatrix} f^\circ_1 \\ f^\circ_2 \\ \cdot \\ \cdot \\ \cdot \\ f^\circ_{\ell_{max}} \end{pmatrix}} + \underset{\ell_{max} X k_{max}}{\begin{pmatrix} \frac{\partial f_1}{\partial x^\circ_j} & \frac{\partial f_1}{\partial x^\circ_2} & \cdots & \frac{\partial f_1}{\partial x^\circ_{k_{max}}} \\ \frac{\partial f_2}{\partial x^\circ_1} & \frac{\partial f_2}{\partial x^\circ_2} & \cdots & \frac{\partial f_2}{\partial x^\circ_{k_{max}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{\ell_{max}}}{\partial x^\circ_1} & \frac{\partial f_{\ell_{max}}}{\partial x^\circ_2} & \cdots & \frac{\partial f_{\ell_{max}}}{\partial x^\circ_{k_{max}}} \end{pmatrix}} \underset{k_{max} X 1}{\begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_{k_{max}} \end{pmatrix}} + \underset{\ell_{max} X j_{max}}{\begin{pmatrix} \frac{\partial f_1}{\partial a^\circ_1} & \frac{\partial f_1}{\partial a^\circ_2} & \cdots & \frac{\partial f_1}{\partial a^\circ_{j_{max}}} \\ \frac{\partial f_2}{\partial a^\circ_1} & \frac{\partial f_2}{\partial a_2} & \cdots & \frac{\partial f_2}{\partial a^\circ_{j_{max}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{\ell_{max}}}{\partial a^\circ_1} & \frac{\partial f_{\ell_{max}}}{\partial a^\circ_2} & \cdots & \frac{\partial f_{\ell_{max}}}{\partial a^\circ_{j_{max}}} \end{pmatrix}} \underset{j_{max} X 1}{\begin{pmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta_{j_{max}} \end{pmatrix}} \tag{40c}$$

$$0 = F + AV + B\Delta \tag{40d}$$

The matrices of (40d) relate directly to the ones of (40c) and are specifically:

A , the matrix of partials with respect to coordinates

B , the matrix of partials with respect to coefficients

V , the matrix of residuals of measurements

Δ , the matrix of parameter corrections

F , the matrix of initial or iterated estimates of function.

In this case, the sum of the squares of the residuals must be minimized with respect to the coordinate and coefficient parameters. The parameter corrections Δ are needed to do this.

Consider,

$$S = V^T P V, \tag{41}$$

the matrix form of the <u>weighted</u> sum of the squares of the residuals. Thus P is a weighting matrix which is the identity matrix when all measurements are considered to be equally good.

To introduce an explicit dependence of S on the condition equations, equations of constraint

$$2\lambda_\ell \left[f_\ell^\circ + \sum_{k=1}^{k_{max}} \left(\frac{\partial f_\ell}{\partial x_k^\circ}\right) dx_k + \sum_{j=1}^{j_{max}} \left(\frac{\partial f_\ell}{\partial a_j^\circ}\right) da_j \right]$$

are inserted in

$$S = V^T P V - 2 \Lambda^T (\underbrace{F + A V + B \Delta}_{\equiv 0}) \tag{42}$$

where

$$\Lambda^T = \left[\lambda_1 \; \lambda_2 \; \cdots \; \lambda_{\ell_{max}} \right],$$

and the individual λ_ℓ are Lagrangian undetermined multipliers. S is now a function of the residuals and of the parameter corrections, $S = S(V, \Delta)$. Thus S can be differentiated with respect to V and Δ and set to zero for minimizing

$$P\,V - A^T \Lambda = 0 \tag{43}$$

$$- B^T \Lambda = 0. \tag{44}$$

(43) and (44) are matrix forms of the normal equations. The weighting matrix P is always square and nonsingular such that

$$V = P^{-1} A^T \Lambda. \tag{45a}$$

This leads to

$$F + A\,P^{-1} A^T \Lambda + B\,\Delta = 0, \tag{40e}$$

whence,

$$\Lambda = -\,(A\,P^{-1} A^T)^{-1} (B\,\Delta + F). \tag{46}$$

Therefore,

$$- B^T \Lambda = -B^T [-(A\,P^{-1} A^T)^{-1} (B\,\Delta + F)] = 0 \tag{44a}$$

or

$$[B^T (A\,P^{-1} A^T)^{-1} B]\,\Delta + B^T (A\,P^{-1} A^T)^{-1} F = 0 \tag{44b}$$

yielding,

$$\Delta = [B^T (A P^{-1} A^T)^{-1} B]^{-1}\ [B^T (A\,P^{-1} A^T)^{-1}]\,F \tag{47}$$

and

$$V = P^{-1} A^T [-(A\,P^{-1} A^T)^{-1} B\{[B^T (A P^{-1} A^T)^{-1} B]^{-1} [B^T (A\,P^{-1} A^T)^{-1}] F\} + F]. \tag{45b}$$

Since

$$\Delta = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta_{j_{max}} \end{pmatrix} \quad \text{and} \quad \mathbf{V} = \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_{k_{max}} \end{pmatrix}$$

the values of the coefficient parameters become for the next iteration

$$(a_j^\circ)_{new} = a_j^\circ + \delta_j \tag{48}$$

and for the coordinates

$$(x_k^\circ)_{new} = x_k^\circ + v_k \ . \tag{49}$$

The δ_j are compared with a small finite limit ϵ_1 and v_k are compared with another limit ϵ_2. If the δ_j and v_k standard deviations are less than ϵ_1 and ϵ_2 repsectively, where an error analysis is performed on the inverted normal matrix to produce the standard deviations then the solution is complete; otherwise, (48) and (49) replace a_j° and x_k° respectively in (40b) and the algorithm proceeds as before. Iterations continue until ϵ_1 and ϵ_2 are reached. If, after a given number of iterations, ϵ_2 is not achieved by the standard deviation of the residuals, those still greater than ϵ_2 are dropped from the equations (40) and a last iteration is performed.

In setting up the equations, if approximations for the a_j and x_k are initially made, acceptable answers may be achieved after only a few iterations; however, complete uncertainty may exist at first such that zero values are the approximations. This means the iterative process will be longer, but residuals and corrections should converge.

Other methods, differing in specific detail but following this general outline, exist in the literature. One increase in complexity occurs when the need for partitioning of all of the matrices is necessitated by use of different point types among the coordinates.

Obviously a computer routine is essential to make the above approach feasible. Indeed, for the examples given in this paper, the establishment of computer programs to perform the least squares algorithms is desirable and necessary to achieve the best utilization of this powerful concept.

Bibliography

H. H. Schmid, "An Analytical Treatment of the Orientation of a Photogrammetric Camera," Ballistic Research Laboratories Report #880, Aberdeen Proving Ground, Md. (1954)

H. H. Schmid, "A General Analytical Solution to the Problem of Photogrammetry," B.R.L. Report #1065, Aberdeen P.G., Md. (1959)

D. C. Brown, "A Matrix Treatment of the General Problem of Least Squares Considering Correlated Observations," B.R.L. Report #937, Aberdeen P.G., Md. (1955)

D. C. Brown, "A Treatment of Analytical Photogrammetry," R.C.A. Data Reduction Technical Report #39, Patrick Air Force Base, Florida (1957)

F. R. Helmert, "Die Ausgleichungsrechnung nach der Methode der Kleinsten Quadrate," Leipzig (1872)

Jordan-Eggert, "Handbuck der Vermessungskunde (Handbook of Geodesy)," Vol. I, Stuttgart (1939) (Transl.: M. Carta, Army Map Service (1962))

J. B. Scarborough, NUMERICAL METHEMATICAL ANALYSIS, The Johns Hopkins Press, 5th Edition, Baltimore, Md. (1962)

APPENDIX

NASA Oriented Application: Satellite Position and Range Determination

To illustrate least squares applications in NASA oriented problems, consider the procedures involved in determining the position, in geocentric coordinates, of an orbiting satellite. If the satellite is observed simultaneously from several ground stations, its position can thence be evaluated. Additionally, the range distance from each station to the satellite can be found. Many such determinations of satellite position within the same orbit enable the orbital parameters to be computed.

Notational definitions:

Let X_j , Y_j , and Z_j be the Cartesian geocentric coordinates of the J^{th} station.

Let x'_{ji} , y'_{ji} , and z'_{ji} be the geocentric coordinates of the satellite as observed at position i and as determined from station j .

Let x_{ji}, y_{ji} , and z_{ji} be the local Cartesian coordinates of the satellite at position i with the local system origined at station j .

Let x_i , y_i , and z_i be the "true" (that is, best fit) position of the satellite in the geocentric coordinate system.

In spherical coordinates the satellite at position i with respect to the station j local system is:

$$\left.\begin{aligned} x_{ji} &= \rho_{ji} \sin \varphi_{ji} \cos \alpha_{ji} \\ y_{ji} &= \rho_{ji} \sin \varphi_{ji} \sin \alpha_{ji} \\ z_{ji} &= \rho_{ji} \cos \varphi_{ji} \end{aligned}\right\} \tag{50}$$

where α is right ascension, φ is codeclination ($90° - \delta$),δ is declination, and ρ is range.

Assume the geocentric coordinates X_j, Y_j, and Z_j for each station j are known and the celestial angles φ_{ji} and α_{ji} are measured. By requiring the axes of each local system to be parallel with the geocentric one, the following relations hold:

$$\left.\begin{aligned} x'_{ji} &= x_{ji} + X_j \\ y'_{ji} &= y_{ji} + Y_j \\ z'_{ji} &= z_{ji} + Z_j \end{aligned}\right\} \quad (51)$$

Furthermore, the differences expressed as residuals in X, Y, Z between "true" satellite positions in geocentric coordinates and station-relative satellite positions in geocentric coordinates are:

$$\left.\begin{aligned} \Delta x_{ji} &= x'_{ji} - x_i \\ \Delta y_{ji} &= y'_{ji} - y_i \\ \Delta z_{ji} &= z'_{ji} - z_i \end{aligned}\right\} \quad (52)$$

$$\Delta \rho_{ji} = \sqrt{\Delta x_{ji}^2 + \Delta y_{ji}^2 + \Delta z_{ji}^2} \quad (53)$$

Therefore, if we assume n ground stations simultaneously observe the satellite position i,

$$\begin{aligned}
\rho_{1i} \sin \varphi_{1i} \cos \alpha_{1i} + X_1 &\approx x_i \\
\rho_{2i} \sin \varphi_{2i} \cos \alpha_{2i} + X_2 &\approx x_i \\
&\vdots \\
\rho_{ni} \sin \varphi_{ni} \cos \alpha_{ni} + X_n &\approx x_i \\
\rho_{1i} \sin \varphi_{1i} \sin \alpha_{1i} + Y_1 &\approx y_i \\
\rho_{2i} \sin \varphi_{2i} \sin \alpha_{2i} + Y_2 &\approx y_i \\
&\vdots \\
\rho_{ni} \sin \varphi_{ni} \sin \alpha_{ni} + Y_n &\approx y_i \\
\rho_{1i} \cos \varphi_{1i} + Z_1 &\approx z_i \\
\rho_{2i} \cos \varphi_{2i} + Z_2 &\approx z_i \\
&\vdots \\
\rho_{ni} \cos \varphi_{ni} + Z_n &\approx z_i
\end{aligned} \quad (54)$$

From these 3n equations per satellite position i we wish to determine the 3 + n unknowns per satellite position i. As before, expressing the approximate equations (as for (56)) in matrix form produces the condition equations. For our purposes, we wish to separate the unknowns ρ_{1i}, ρ_{2i}, . . . , ρ_{ni}, x_i, y_i, and z_i from the known factors. The resulting condition equation is:

$$
\underset{(3n)\times(n+3)}{\begin{bmatrix}
1 & 0 & 0 & \sin\varphi_{1i}\cos\alpha_{1i} & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \sin\varphi_{1i}\sin\alpha_{1i} & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cos\varphi_{1i} & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & 0 & \sin\varphi_{2i}\cos\alpha_{2i} & 0 & \cdots & 0 \\
0 & 1 & 0 & 0 & \sin\varphi_{2i}\sin\alpha_{2i} & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & \cos\varphi_{2i} & 0 & \cdots & 0 \\
1 & 0 & 0 & 0 & 0 & \sin\varphi_{3i}\cos\alpha_{3i} & \cdots & 0 \\
0 & 1 & 0 & 0 & 0 & \sin\varphi_{3i}\sin\alpha_{3i} & \cdots & 0 \\
0 & 0 & 1 & 0 & 0 & \cos\varphi_{3i} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 0 & 0 & 0 & 0 & 0 & \cdots & \sin\varphi_{ni}\cos\alpha_{ni} \\
0 & 1 & 0 & 0 & 0 & 0 & \cdots & \sin\varphi_{ni}\sin\alpha_{ni} \\
0 & 0 & 1 & 0 & 0 & 0 & \cdots & \cos\varphi_{ni}
\end{bmatrix}}
\underset{(n+3)\times 1}{\begin{bmatrix}
x_i \\ y_i \\ z_i \\ -\rho_{1i} \\ -\rho_{2i} \\ \vdots \\ -\rho_{ni}
\end{bmatrix}}
=
\underset{(3n)\times 1}{\begin{bmatrix}
X_1 \\ Y_1 \\ Z_1 \\ X_2 \\ Y_2 \\ Z_2 \\ \vdots \\ X_n \\ Y_n \\ Z_n
\end{bmatrix}}
\qquad (55a)
$$

For convenience express (55a) as:

$$S_i \; R_i = X_i \qquad (55b)$$

The solution of 55b) for R_i is equivalent to a least squares evaluation of

$$Q_i = \sum_{j=1}^{n} \Delta\rho_{ji} = \sum_{j=1}^{n} \Delta x_{ji}^2 + \Delta y_{ji}^2 + \Delta z_{ji}^2 =$$

(56)

$$\sum_{j=1}^{n} [(\rho_{ji}\sin\varphi_{ji}\cos\alpha_{ji} + X_j - x_i)^2 + (\rho_{ji}\sin\varphi_{ji}\sin\alpha_{ji} + Y_j - y_i)^2$$

$$+ (\rho_{ji}\cos\varphi_{ji} + Z_j - z_i)^2] .$$

That is, the sum of the squares of the residuals Q_i is minimized by finding the partials of Q_i with respect to the unknowns evaluated as zero. e.g.

$$\frac{\partial Q_i}{\partial \rho_{1i}} = \frac{\partial Q_i}{\partial \rho_{2i}} = \cdots = \frac{\partial Q_i}{\partial \rho_{ni}} = \frac{\partial Q_i}{\partial X_i} = \frac{\partial Q_i}{\partial Y_i} = \frac{\partial Q_i}{\partial Z_i} = 0 \tag{57}$$

Considering (55b) first,

$$S_i^T S_i R_i = S_i^T X_i \tag{58a}$$

which is, explicitly,

$$
\begin{array}{cc}
(n+3)\times(n+3) & (n+3)\times 1 \\
\begin{bmatrix}
n & 0 & 0 & \sin\varphi_{1i}\cos\alpha_{1i} & \sin\varphi_{2i}\cos\alpha_{2i} & \cdots & \sin\varphi_{ni}\cos\alpha_{ni} \\
0 & n & 0 & \sin\varphi_{1i}\sin\alpha_{1i} & \sin\varphi_{2i}\sin\alpha_{2i} & \cdots & \sin\varphi_{ni}\sin\alpha_{ni} \\
0 & 0 & n & \cos\varphi_{1i} & \cos\varphi_{2i} & \cdots & \cos\varphi_{ni} \\
\sin\varphi_{1i}\cos\alpha_{1i} & \sin\varphi_{1i}\sin\alpha_{1i} & \cos\varphi_{1i} & 1 & 0 & \cdots & 0 \\
\sin\varphi_{2i}\cos\alpha_{2i} & \sin\varphi_{2i}\sin\alpha_{2i} & \cos\varphi_{2i} & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\sin\varphi_{ni}\cos\alpha_{ni} & \sin\varphi_{ni}\sin\alpha_{ni} & \cos\varphi_{ni} & 0 & 0 & \cdots & 1
\end{bmatrix}
&
\begin{bmatrix}
X_i \\ Y_i \\ Z_i \\ -\rho_{1i} \\ -\rho_{2i} \\ \vdots \\ -\rho_{ni}
\end{bmatrix}
\end{array} =
$$

$$
\underset{(n+3)\times(3n)}{\begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 0 & \cdots & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & \cdots & 0 & 0 & 1 \\
\sin\varphi_{1i}\cos\alpha_{1i} & \sin\varphi_{1i}\sin\alpha_{1i} & \cos\varphi_{1i} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & 0 & \sin\varphi_{2i}\cos\alpha_{2i} & \sin\varphi_{2i}\sin\alpha_{2i} & \cos\varphi_{2i} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & \sin\varphi_{ni}\cos\alpha_{ni} & \sin\varphi_{ni}\sin\alpha_{ni} & \cos\varphi_{ni}
\end{bmatrix}}
\underset{(3n)\times 1}{\begin{bmatrix}
X_1 \\ Y_1 \\ Z_1 \\ X_2 \\ Y_2 \\ Z_2 \\ \vdots \\ X_n \\ Y_n \\ Z_n
\end{bmatrix}}
\qquad (58b)
$$

Premultiplying (58) by the inverse of $(S_i^T\ S_i)$ produces the desired solution for R_i .

$$(S_i^T S_i)^{-1} S_i^T S_i R_i = (S_i^T S_i)^{-1} S_i^T X_i \tag{59a}$$

$$R_i = (S_i^T S_i)^{-1} S_i^T X_i \tag{59b}$$

The equation (58) may also be produced by straight forward evaluation of the partials in equation (57).

$$\left.\begin{aligned}
&\frac{\partial Q_i}{\partial x_i} = -2 \sum_{j=1}^{n} (\rho_{ji} \sin \varphi_{ji} \cos \alpha_{ji} + X_j - x_i) = \sum_{j=1}^{n} (\rho_{ji} \sin \varphi_{ji} \cos \alpha_{ji}) + \sum_{j=1}^{n} X_j - n x_i = 0 \\
&\frac{\partial Q_i}{\partial y_i} = -2 \sum_{j=1}^{n} (\rho_{ji} \sin \varphi_{ji} \sin \alpha_{ji} + Y_j - y_i) = \sum_{j=1}^{n} (\rho_{ji} \sin \varphi_{ji} \sin \alpha_{ji}) + \sum_{j=1}^{n} Y_j - n y_i = 0 \\
&\frac{\partial Q_i}{\partial z_i} = -2 \sum_{j=1}^{n} (\rho_{ji} \cos \varphi_{ji} + Z_j - z_i) = \sum_{j=1}^{n} (\rho_{ji} \cos \varphi_{ji}) + \sum_{j=1}^{n} Z_j - n z_i = 0 \\
&\frac{\partial Q_i}{\partial \rho_{ji}} = 2 (\rho_{ji} \sin \varphi_{ji} \cos \alpha_{ji} + X_j - x_i) \sin \varphi_{ji} \cos \alpha_{ji} + 2 (\rho_{ji} \sin \varphi_{ji} \sin \alpha_{ji} + \\
&(j = 1, 2, \cdots, n) \qquad Y_j - y_i) \sin \varphi_{ji} \sin \alpha_{ji} + 2 (\rho_{ji} \cos \varphi_{ji} + Z_j - z_i) \cos \varphi_{ji} = 0
\end{aligned}\right\} \tag{60}$$

By rearranging the terms of (60), i.e.

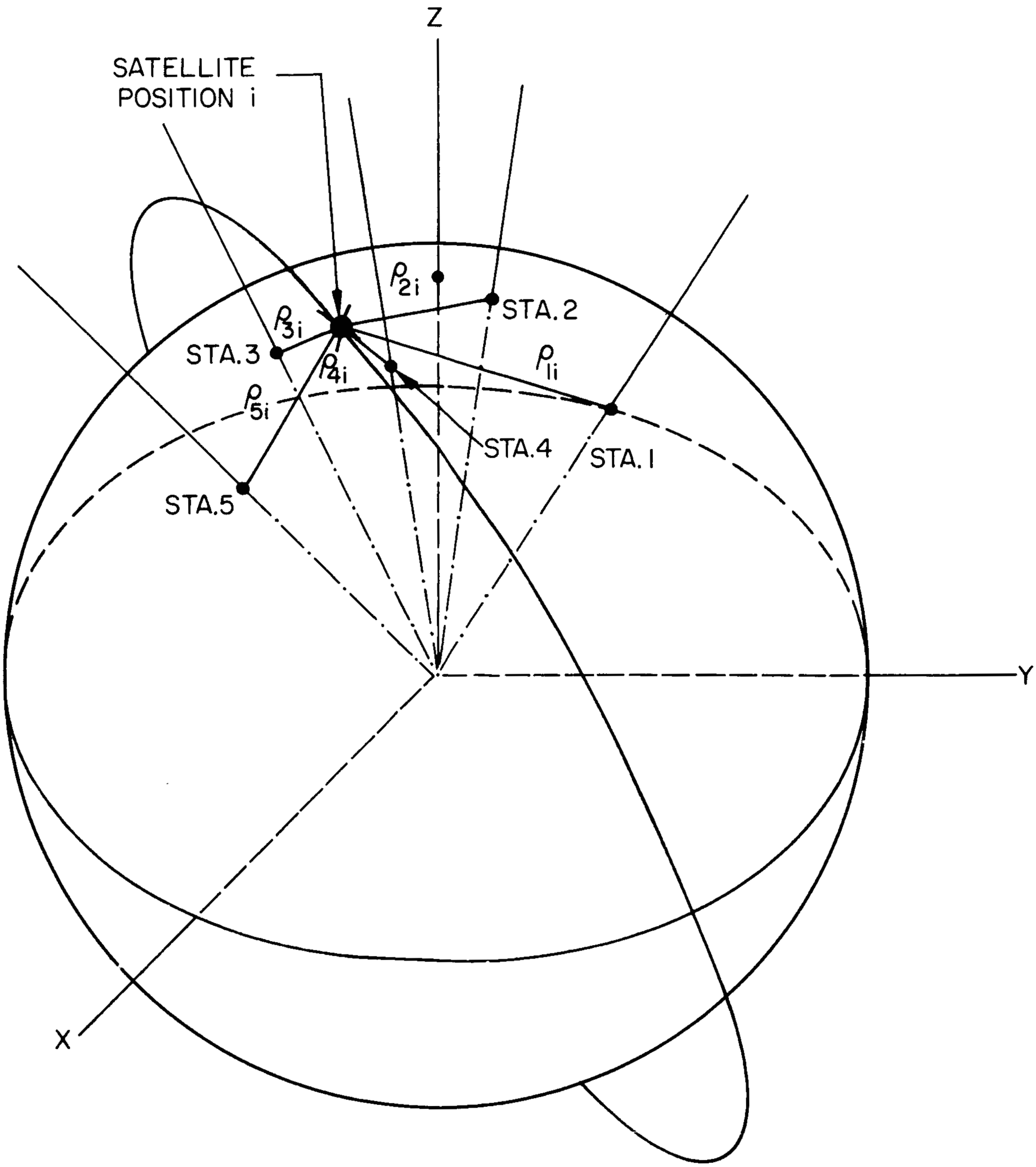
$$\rho_{ji} + (Z_j - z_i) \cos \varphi_{ji} + (X_j - x_i) \sin \varphi_j \cos \alpha_{ji} + (Y_j - y_i) \sin \varphi_{ji} \sin \alpha_{ji} = 0 \left\{ \begin{array}{l} \text{n equations} \\ j = 1, 2, \cdots, n \end{array} \right.$$

$$n x_i - \sum_{j=1}^{n} X_j = \sum_{j=1}^{n} \rho_{ji} \sin \varphi_{ji} \cos \alpha_{ji}$$

$$n y_i - \sum_{j=1}^{n} Y_j = \sum_{j=1}^{n} \rho_{ji} \sin \varphi_{ij} \sin \alpha_{ji}$$

$$n z_i - \sum_{j=1}^{n} Z_j = \sum_{j=1}^{n} \rho_{ji} \cos \varphi_{ji} \, ,$$

the relations represented in matrix notation in (58) can be readily obtained, once more justifying the matrix approach to least squares.



Z
SATELLITE POSITION i
ρ_{2i}
STA. 2
ρ_{3i}
STA. 3
ρ_{4i}
ρ_{1i}
ρ_{5i}
STA. 4
STA. 1
STA. 5
Y
X

Figure 11