2. Dillinger, W. H., and Hanson, R. H., "Status of the Computer System for the New Adjustment of the North American Datum," presented to the American Geophysical Union Annual Fall Meeting, Dec., 1976. Preprint, National Ocean Survey, National Geodetic Survey, Rockville, Md.
3. Dracup, J. F., "The New Adjustment of the North American Datum," *Article No. 9: Plane Coordinate Systems, ACSM Bulletin,* Vol. 59, 1977, p. 27.
4. "Geodesy, Dealing with an Enormous Computer Task," *Science,* Vol. 200, Apr., 1978, p. 421.
5. Meissl, P., "New Adjustment of the North American Datum," *Article No. 10: Prediction of Roundoff Errors, ACSM Bulletin,* Vol. 60, 1978, pp. 17, 21.
6. "Policy on Publication of Plane Coordinates," *Federal Register,* Vol. 42, No. 57, 1977, pp. 15913–15914.
7. Vincenty, T., "Determination of North American Datum 1983 Coordinates of Map Corners (Second Prediction)," *NOAA Technical Memorandum NOS NGS-16,* Apr., 1979, pp. 3–4.

17289 AUGUST 1982 SU2

Minimal Constraints in Two-Dimensional Networks

By Alfred Leick,[1] A. M. ASCE

Abstract: The use of coordinates in the adjustment of two-dimensional networks leads to inherently singular solutions. Various minimal constraints are applied to eliminate the singularity. The variant and invariant quantities, with respect to a choice of minimal constraints, are identified. Inner constraints, which constitute a subset of all sets of minimal constraints, are useful in displaying the internal network geometry through error ellipses whose size, shape and orientation are independent of the definition of the coordinate system. Inner constraint solutions are helpful in simulation studies where the optimum station locations and observing program are investigated. Any minimal constraint solution can be used for checking the quality of observations and the detection of blunders.

Introduction

The mathematical model for two-dimensional network adjustments uses either auxiliary parameters (station coordinates) or conditions between the observations. Since the use of auxiliary parameters leads to a formulation which itself can be programmed on computers, even on microcomputers, the method of variation of parameters is rapidly gaining popularity. Any two-dimensional network adjustment can be formulated in a unified manner.

The use of coordinates leads to inherently singular adjustments. Various sets of minimal constraints (conditions) can be applied to solve this singularity. Inner constraints, which constitute a subset of all sets of minimal constraints, are particularly useful in displaying the internal network geometry through error ellipses whose size, shape, and orientation are independent of the definition of the coordinate system.

A general formulation of minimal constraints is given. Sample adjustments demonstrate the relationship between minimal constraints and the coordinate system definition. Variant and invariant elements with respect to the specific choice of minimal constraints are identified. Not all sets of minimal constraints are equivalent from the numerical viewpoint. In the extreme cases, other singularities of the normal equation arises.

[1]Asst. Prof., Dept. of Civ. Engrg., Univ. of Maine, Surveying Engrg. Program, Orono, Me. 04469.

Note.—Discussion open until January 1, 1983. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on May 7, 1981. This paper is part of the Journal of the Surveying and Mapping Division, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 108, No. SU2, August, 1982. ISSN 0569-8073/82/0002-0053/$01.00.

Minimal constraint solutions are used in checking the quality of the observations, detection of blunders, etc. Inner constraint solutions are helpful in simulation studies where the optimum station locations and observing program are investigated.

Dual Formulation

There are two methods used to adjust horizontal networks. One method uses a nonlinear mathematical model in which the observations are related explicitly to a set of unknown quantities (coordinates) called the parameters. Eq. 1 expresses the general form of this model. The symbols $\mathbf{L_a}$ and $\mathbf{X_a}$ = the adjusted observations and parameters, respectively:

$$\mathbf{L_a} = \mathbf{F}(\mathbf{X_a}) \quad (1)$$

Because the parameters are determined such that some well-defined quadratic form is minimized, the method is also commonly referred to as the "method of variation of parameters." The other method, commonly referred to as the "method of condition equations," uses a nonlinear model in which the observations are related implicitly:

$$\mathbf{G}(\mathbf{L_a}) = 0 \quad (2)$$

The complete formulation of the adjustment also requires information about the quality of the observations. This information is introduced through the so-called stochastic model:

$$\mathbf{P} = \sigma_0^2 \, \Sigma_{\mathbf{L_b}}^{-1} \quad (3)$$

in which $\Sigma_{\mathbf{L_b}}$ = the variance-covariance matrix of the observation, $\mathbf{L_b}$; σ_0^2 = the *a priori* variance of unit weight; and $\mathbf{P}$ = the weight matrix.

Both mathematical models are equivalent. They yield the same set of adjusted observations, $\mathbf{L_a}$, and variance-covariance matrix of adjusted observations, $\Sigma_{\mathbf{L_a}}$. Without elaborating any further on the extent of this equivalence, it is simply stated that a network adjustment can be formulated with either model 1 or 2.

Finding the mathematical model for the method of variation of parameters is made easy by the following rule: "Each observation—one equation." If there are n observations (observed distances and angles) in the network then Eq. 1 represents n nonlinear equations, which, in turn, are related to, say, u parameters. $\mathbf{L_a}$ is an n-dimensional vector which is related through the n-dimensional function, F, to the u-dimensional vector of parameters. The difference, $n - u$, is called the degree-of-freedom and equals the number of redundant observations, i.e. there are $n - u$ more equations than needed for the direct solutions of the system of Eq. 1. Degenerate networks such as the triangle, quadrilateral, central system, and traverse do not require any special considerations. It is easy to program the computer for setting up one equation for each observation.

The method of condition (Eq. 2) requires exactly $n - u$ independent conditions. There are difficulties involved: First, independent conditions are required; for a complex network it is generally not easy to identify such conditions. Secondly, it is very difficult to program the computer as to find the $n - u$ conditions automatically. Traditionally, the various conditions were found manually.

Preference is occasionally given to the method of Eq. 2 because there are only $n - u$ normal equations to be reduced as opposed to u in the case of method 1. In the precomputer era, this difference in computational effort was significant, especially if $n - u \ll u$. Today, when microcomputers readily perform all computations in the surveyor's office, it is more important that the adjustment be performed completely automatically, i.e. the computer not only performs the numerical work but should also "set-up" the adjustment.

Method of Variation of Parameters

The least squares algorithm for the method of observation equations can be found in any textbook on estimation. For the sake of subsequent studies, the pertinent expressions are

$$\mathbf{V} = \mathbf{AX} + \mathbf{L} \quad (4)$$

$$\mathbf{X} = -(\mathbf{A}^T\mathbf{PA})^{-1}\,\mathbf{A}^T\,\mathbf{PL} \quad (5)$$

$$\hat{\sigma}_0^2 = \frac{\mathbf{V}^T\mathbf{PV}}{n - u} \quad (6)$$

$$\Sigma_{\mathbf{X}} = \hat{\sigma}_0^2\,(\mathbf{A}^T\mathbf{PA})^{-1} \quad (7)$$

$$\Sigma_{\mathbf{L_a}} = \mathbf{A}\Sigma_{\mathbf{X}}\mathbf{A}^T \quad (8)$$

$$\Sigma_{\mathbf{L_a}} = \Sigma_{\mathbf{L_b}} - \Sigma_V \quad (9)$$

in which $\mathbf{V} = \mathbf{L_a} - \mathbf{L_b}$ (10)

$$\mathbf{X} = \mathbf{X_a} - \mathbf{X_0} \quad (11)$$

$$\mathbf{L} = \mathbf{L_0} - \mathbf{L_b} \quad (12)$$

$$\mathbf{L_0} = \mathbf{F}(\mathbf{X_0}) \quad (13)$$

$$\mathbf{A} = \left.\frac{\partial \mathbf{F}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X_0}} \quad (14)$$

in which $\mathbf{L_a}$, $\mathbf{L_b}$, and $\mathbf{V}$ = the n-dimensional vector of adjusted observations, observed values, and residuals; $\mathbf{X_a}$, $\mathbf{X_0}$, and $\mathbf{X}$ = the u-dimensional vector of the adjusted parameters, approximate parameters, and the parameter corrections, respectively; and $\mathbf{L_0}$ = an n-dimensional vector containing the values of the function $\mathbf{F}$, evaluated at $\mathbf{X_0}$. The design matrix, $\mathbf{A}$, has the size $(n \times u)$. $\hat{\sigma}_0$ = *a posteriori* variance of unit weight, and finally, $\Sigma_{\mathbf{X}}$ and $\Sigma_{\mathbf{L_a}}$ = the variance-covariance matrices of the adjusted parameters and observations.

A linear function of the adjusted parameters and the respective variance-covariances is computed from the well-known law of variance-covariance propagation:

$$\mathbf{Y} = \mathbf{HX} \quad (15)$$

$$\Sigma_{\mathbf{Y}} = \mathbf{H}\Sigma_{\mathbf{X}}\mathbf{H}^T \quad (16)$$

in which $\mathbf{H}$ = an $(m \times u)$ matrix for m linear functions. Eqs. 15 and 16 can be used to compute the adjusted distances and angles between points including their variance-covariances.

The variance-covariance matrix, Σ_X, reflects the geometry of the network. Traditionally, this geometry is made visible by ellipses of standard deviations. Expressions are given in Appendix I.

Rank Defect in Horizontal Network Adjustments

Consider the case of a trilateration network in which only distances are observed. Such a network can be translated as a whole and rotated in the plane without changes in distances between points. The distances determine the relative location of points and not the absolute position of the points in the plane. An alternative way of expressing this property is to say that distances are invariant with respect to translation and rotation. The coordinate system serves only as auxiliary means to express distances.

Consider the case of a triangulation network in which only angles are observed. Such a network can be translated, rotated, and even magnified without any angles undergoing changes. The size of the network cannot be determined from angle measurements only. Thus, angles are said to be invariant with respect to translation, rotation, and scale.

A trilateration/triangulation network in which both distances and angles are measured is invariant with respect to rigid translation and rotation. One distance determines the size of the network.

The invariance of angles and distances, with respect to the definition of the coordinate system (translation plus rotation) and the scale, is reflected by rank deficiencies in the design matrix. Assume that a network consists of m stations. Since each station contributes two parameters, one for each coordinate, there are $u = 2m$ columns in the design matrix **A**. Inspecting the coefficients of the **A**-matrix yields the respective linear combinations of the columns.

The row of matrix **A** which pertains to a distance observation between stations P_i and P_j contains the following elements:

$$\begin{bmatrix} & dx_i & dy_i & & dx_j & dy_j & \\ \cdots & -\frac{x_j - x_i}{D_{ij}} & -\frac{y_j - y_i}{D_{ij}} & \cdots & \frac{x_j - x_i}{D_{ij}} & \frac{y_j - y_i}{D_{ij}} & \cdots \end{bmatrix} \quad (17)$$

The zero elements are represented by dots. The symbols written immediately above the coefficients of Eq. 17 relate the columns to the individual parameters. The vector of the unknown parameter corrections is

$$\mathbf{X}^T = (dx_1, dy_1 \ldots dx_i, dy_i \ldots dx_j, dy_j \ldots dx_m, dy_m) \quad (18)$$

in which D_{ij} = the distance between the points P_i and P_j as computed from the approximate coordinates. The y-axis and the x-axis denote the ordinate and the abscissa, respectively. The coefficients for angle observations, $P_i - P_j - P_k$, as measured at P_j from P_i to P_k in a clock-wise sense, are

$$\begin{bmatrix} & dx_i & dy_i & & dx_j & dy_j & & dx_k & dy_k \\ \cdots & -\frac{y_i - y_j}{D_{ij}^2} & \frac{x_i - x_j}{D_{ij}^2} & \cdots & -\frac{y_k - y_j}{D_{jk}^2} + \frac{y_i - y_j}{D_{ij}^2} & \frac{x_k - x_j}{D_{kj}^2} - \frac{x_i - x_j}{D_{ij}^2} & \cdots & \frac{y_k - y_j}{D_{kj}^2} & -\frac{x_k - x_j}{D_{kj}^2} \end{bmatrix} \quad (19)$$

In trilateration networks, the **A**-matrix consists of n rows of the type shown in

Eq. 17. The following three linear relations can be readily verified:

$$\mathbf{A}\mathbf{E}_D^T = 0 \quad (20)$$

$$\text{with } \mathbf{E}_D = \begin{pmatrix} 1 & 0 & 1 & 0 \ldots 1 & 0 \\ 0 & 1 & 0 & 1 \ldots 0 & 1 \\ y_1 & -x_1 & y_2 & -x_2 \ldots y_m & -x_m \end{pmatrix} \quad (21)$$

In triangulation networks, the **A**-matrix consists of rows of the type shown in Eq. 19. The following four linear relations can be readily verified:

$$\mathbf{A}\mathbf{E}_A^T = 0 \quad (22)$$

$$\text{with } \mathbf{E}_A = \begin{pmatrix} 1 & 0 & 1 & 0 \ldots 1 & 0 \\ 0 & 1 & 0 & 1 \ldots 0 & 1 \\ y_1 & -x_1 & y_2 & -x_2 \ldots y_m & -x_m \\ x_1 & y_1 & x_2 & y_2 \ldots x_m & y_m \end{pmatrix} \quad (23)$$

In trilateration/triangulation networks, the **A**-matrix consists of rows of the type shown in Eqs. 17 and 19. It can be readily verified that Eqs. 20 and 21 apply to this case.

Free Network Adjustment

Because the columns of the design matrix are linearly dependent, the normal matrix

$$\mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A} \quad (24)$$

is singular. The singularity will disappear if information about the definition of the coordinate system, and possibly even the scale is introduced. This is accomplished by imposing as many constraints (conditions) upon the adjustment as there is rank deficiency. Adjustments which incorporate no more and no less conditions than are necessary to define the coordinte system and the scale lead to so-called minimal constraint solutions, or "free network adjustments."

Each specific choice of constraints results in a different adjustment. Some quantities remain invariant with respect to such a choice; other vary. As shown in Appendix II, the following quantities remain invariant: adjusted observation, $\mathbf{L}_a$; variance-covariance matrix of the adjusted observations, Σ_{L_a}; estimated distances and angles and their variance-covariances, **V**, $\mathbf{V}^T\mathbf{P}\mathbf{V}$, and $\hat{\sigma}_0$ (by implication). The variant quantities are those which carry information about the specific choice of the coordinate system. In Appendix II, the following quantities are identified as variants: parameters, **X**; variance-covariance matrix of adjusted parameters, Σ_X; and correlation matrix of the parameters (by implication). The auxiliary nature of the coordinates becomes apparent when realizing that angles, distances, and their respective variance-covariances, as computed from the adjusted results, are invariant. Angles and distances are sometimes referred to as estimable quantities in this context.

Constraining Subset of Coordinates.—The most simple method for introducing the minimal number of constraints is to assign numerical values to a

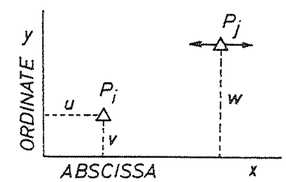

FIG. 1.—Defining Coordinate System

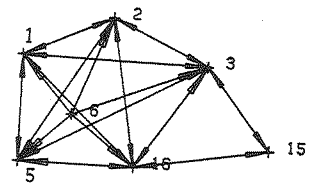

FIG. 2.—Sample Network Configuration

subset of coordinates. These coordinates, then, are considered constants and are deleted from the parameter vector (Eq. 18). The respective columns of the design matrix are deleted also. The origin of the coordinate system is defined by assuming arbitrary numerical values for the coordinates of some arbitrary station. For example, the coordinates of P_i are $x_i = u$ and $y_i = v$. The values u and v must not necessarily be zero. The coordinate system is oriented by assigning a numerical value to another coordinate.

Fig. 1 shows the case where y_j is assigned the value w. The adjustment can "move" the station P_j only in the x-direction, as indicated by the arrows. The coordinate w must be chosen such that the absolute value, $|w - v|$, does not exceed the distance between stations P_i and P_j. In triangulation networks, the scale of the network is defined by fixing the second coordinate of station P_j. In that case, the values assigned to the coordinates of P_j are arbitrary. The distance between P_i and P_j, as is implied by the assigned coordinate values, determines the size of the network.

The network shown in Fig. 2 serves as a sample to demonstrate the effects of various sets of minimal constraints. The network consists of uncorrelated angle and distance observations. There is no need to list the actual observations since these computations only serve to demonstrate the variant quantities, in particular the ellipses of standard deviation. No attempt is made to report the adjustment of real field observations. All adjustments use the same observations and variances.

Figs. 3–5 show the error ellipses for the minimal constraints listed in Table 1. The figures show an obvious dependency of size, shape, and orientation of the ellipses upon the choice of minimal constraints. The ellipses tend to increase in size in proportion to their distance from the fixed station. The adjusted parameters are also different for each case. Their actual values are of no importance to the present analysis and therefore are not rendered. The same is true for the numerical values of the invariant quantities $\mathbf{L}_a$, $\Sigma_{\mathbf{L}_a}$, and $\mathbf{V}$, $\mathbf{V}^T\mathbf{P}\mathbf{V}$, $\hat{\sigma}_0$, and any distance or angle computed from the adjusted coordinates.

Implementation of minimal constraints by fixing three or four coordinates is accomplished simply by deleting the respective columns from the design matrix. However, there is one pitfall: The coordinate system may be ill-defined by some particular set of minimal constraints. In such cases, even after the respective columns have been deleted in the A-matrix, it is found that the remaining columns are "numerically nearly" linear dependent, the normal matrix is "numerically nearly" singular, and the correlation matrix of the parameters exhibits high correlation. In the extreme case, there is a "total" singularity. Such a sit-

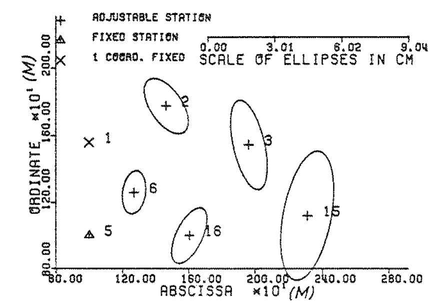

FIG. 3.—Minimal Constraints on Stations 1 and 5

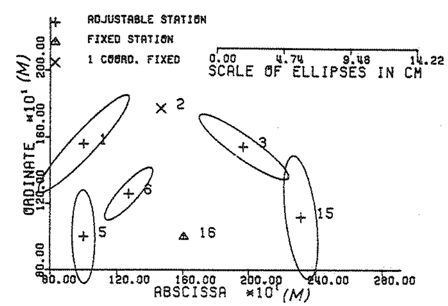

FIG. 4.—Minimal Constraints on Stations 16 and 2 (Ordinate)

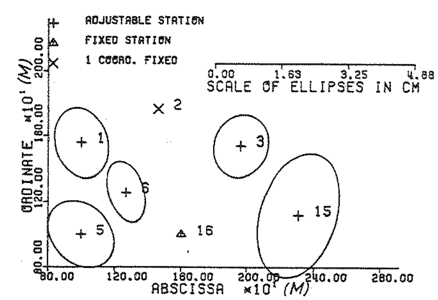

FIG. 5.—Minimal Constraints on Stations 16 and 2 (Abscissa)

uation arises for the following minimal constraints: $y_5 = 1{,}000.$, $x_5 = 1{,}000.$ $y_1 =$ distance $P_1P_5 + y_5$.

This set of constraints implies that the y-axis is parallel to the station P_1 and P_5. The station P_1 is allowed to move in the x-direction. The additional rank defect, which is now solely due to this particular choice of coordinate system, can be explained geometrically as follows: If the coordinate system is rotated by a differentially small amount, then the change in the coordinate difference, $y_1 - y_5$, is of second order, i.e. the change is a function of the square of differential elements, whereas the change in $x_1 - x_5$ is of the first order. Thus, the

TABLE 1.—Sample Sets of Minimal Constraints

Figure (1)	Constraint coordinates, in meters (2)		
3	$y_5 = 1{,}000.00$	$x_5 = 1{,}000.00$	$x_1 = 1{,}000.00$
4	$y_{16} = 1{,}000.00$	$x_{16} = 1{,}604.82$	$y_2 = 1{,}775.09$
5	$y_{16} = 1{,}000.00$	$x_{16} = 1{,}604.82$	$x_2 = 1{,}464.90$

coordinate difference, $y_1 - y_5$, is invariant with respect to a differential rotation, given this particular choice of coordinate systems, and neglecting higher order small terms. The mathematical model used in the adjustment is the linearized part of the true nonlinear mathematical relation, as is well known. Thus, imposing a condition on $y_1 - y_5$ does not define the coordinate system; the coordinate system still can be rotated by a differentially small amount without violating the conditions imposed.

For the limiting case where one of the axes is parallel to the two stations whose coordinate difference of that axis is constrained, it is easy to verify analytically the linear dependency of the columns of the design matrix. If the axis is nearly parallel to the respective two stations, the columns are nearly linearly dependent. Inversion of the normal matrix may be impossible on the computer because of the limited number of significant digits. The adjustment is said to be ill conditioned. The following three examples demonstrate this kind of ill-conditioning. In each case, the angle between the line $\overline{P_1P_5}$ and the ordinate is altered, as shown in Table 2. Fig. 6 shows the ellipses for all three cases. The ill-conditioning decreases as the angular separation increases. The semimajor axes of the error ellipses change dramatically by a factor of approximately four. The ellipses become systematically narrower as the ill-conditioning increases (correlations between the parameters increase). The orientation of the error ellipses changes also. In the limiting case, where the y-axis is parallel to $\overline{P_1P_5}$, the error ellipses degenerate into straight lines. These lines would be tangent to circles whose center are at P_5 and which go through the station under consideration. The degenerated error ellipses would indicate the direction of possible motion of the station due to the lack of coordinate system definition.

It has been demonstrated that each adjustment must be considered together with its specific set of minimal constraints. Pope (5) presents expressions for the transformation of the parameter vector and the variance-covariance matrix due to changes in the minimal constraints. One solution can be transformed into another without redoing the adjustment. It is emphasized that the expressions and properties listed in Appendix I apply only to a straightforward rotation of

TABLE 2.—Rotation of Coordinate System

Adjustment (1)	Approximate angle between ordinate and $\overline{P_1P_5}$ (2)
1	5°
2	20°
3	45°

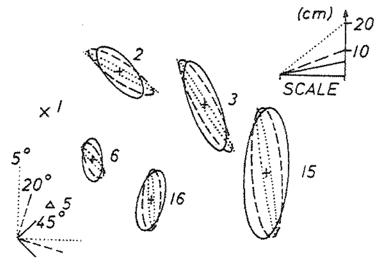

FIG. 6.—Ill Conditioned Adjustment Due to Coordinate System Definition

the coordinate system (after the adjustment) without a change in the underlying minimal constraints.

Inner Constraint Solution.—Thus far, the minimal set of constraints is applied by simply removing three of four columns from the design matrix. It was demonstrated that the error ellipses heavily depend on the specific choice of the minimal set of constraints. This dependency is eliminated by choosing so-called inner constraints

$$\mathbf{EX} = 0 \quad \text{(25)}$$

The **E** matrix is given in Eq. 21 and 23. The number of conditions is again equal to the rank defect of the design matrix. The inner constraints Eq. 25 encompass all parameters; no column is deleted from the full design matrix.

The general formulation of the inner constrain solution can be stated as follows: Find the solution for **X** of $\mathbf{V} = \mathbf{AX} + \mathbf{L}$ such that $\mathbf{V}^T\mathbf{PV}$ is a minimum, and that the condition, $\mathbf{EX} = 0$, is fulfilled. This problem has been thoroughly investigated by Meissl (4) and Pope (5), both of which were primarily concerned with three-dimensional networks. The solution is (5)

$$\mathbf{X} = -\mathbf{N}^+\mathbf{A}^T\mathbf{PL} \quad \text{(26)}$$

$$\text{with} \quad \mathbf{N}^+ = (\mathbf{A}^T\mathbf{PA} + \mathbf{E}^T\mathbf{E})^{-1} - \mathbf{E}^T(\mathbf{EE}^T\mathbf{EE}^T)^{-1}\mathbf{E} \quad \text{(27)}$$

$$\text{and} \quad \Sigma_X = \hat{\sigma}_0\mathbf{N}^+ \quad \text{(28)}$$

The $(u \times u)$ matrix, $\mathbf{N}^+$, is called the pseudoinverse of the normal matrix, **N**.

There are several alternative ways to compute the pseudoinverse. The attraction of Eq. 27 is that **E** can be found readily without any extra computations. However, Eq. 27 involves the subtraction of two matrices which can cause numerical problems since the computer operates only with a limited number of significant digits. Such numerical difficulties arise when the station coordinates have large numerical values. Usually, it will be possible to avoid these numerical difficulties by increasing the computational accuracy (e.g. double precision to quadruple precision) or by locating the origin of the coordinate system at the center of the network.

There are many interesting properties associated with the pseudoinverse, but most of them are not of direct concern to the contents of this paper. It is, important to note, however, that the inner constraints are just a particular set of minimal constraints. All invariant quantities considered earlier also remain in-

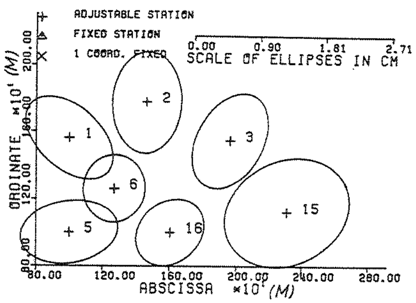

FIG. 7.—Inner Constraint Solution

variant with respect to the inner constraints (Eq. 25). In Appendix III, a geometric interpretation of the inner constraints in two-dimensional networks is considered. The fields of adjusted station positions and the approximate positions are related by a zero translation and rotation (and scale) in the least squares sense.

Fig. 7 shows the ellipses of standard deviation for the case of inner constraints. There is one ellipse for each station. The size of the ellipses is generally smaller than in other minimal constraint solutions. But, most importantly, the effect of the coordinate system definition on the ellipses has disappeared altogether. The shapes and sizes of the ellipses truely represent the geometry of the network.

Conclusions

The concepts of minimal constraint solutions have been dealt with theoretically and demonstrated by numerical examples. Two cases were considered: (1) Conditions on a subset of the parameters; and (2) conditions encompassing all parameters. Case (2) was termed "inner constraint solution." It was shown that numerical instabilities can still arise, even though the minimal number of conditions is applied. Those difficulties rise for an ill-selected set of minimal constraints (case 1) and for large coordinate values (case 2). It was further demonstrated that the adjusted observations, $\mathbf{L}_a$, their variance-covariance matrix, Σ_{L_a}, and consequently the quadratic form, $\mathbf{V}^T\mathbf{P}\mathbf{V}$, the *a posteriori* variance of unit weight $\hat{\sigma}_0$, and the residuals, $\mathbf{V}$, are independent of the specific choice of minimal constraints. The ellipses of standard deviation, on the contrary, depend on the definition of the coordinate system.

The most simple way of applying minimal constraints is to delete a subset of parameters. Numerical difficulties, due to the coordinate system definition, are avoided by selecting the two stations as far apart as possible. Define the origin by treating both coordinates of one station as constants (delete them from the list of parameters). Orient the coordinate system by fixing one of the coordinate of the other station. Take the x or y-axis to be parallel to both stations and hold the y or x coordinate fixed respectively. If only angles are observed, all four coordinates of both stations are kept constant.

Any minimal constraint solution is adequate for the statistical testing of the quality of the observations, the detection of blunders, and the computation of adjusted distances and angles and their variance-covariances. The inverse of the normal matrix should be computed only if necessary.

Inner constraint solutions are particularly useful in displaying the geometry of the network since the size, shape, and orientation of the ellipses of standard deviation are independent of the definition of the coordinate system. Such a display of the ellipses, in connection with the ability to compute selected variances of distances and angles, provides a powerful tool to discover local geometric weaknesses in a network and to find those additional observations which would strengthen the figure most effectively. Those techniques are especially useful in simulation studies prior to taking observations.

This study did not address the inner constraints required in the presence of direction observations. Examples for this are found in Leick and Tyler (3). Minimal constraints are sometimes applied by holding fixed one station and the azimuth to another station. The corresponding minimal constraint solution, as studied in this paper, would be to let one axis coincide with the direction between the two stations and to hold the appropriate three coordinates fixed. Relative ellipses of standard deviation between two stations are not considered either. Since they are invariant with respect to the definition of the coordinate system, they are another tool for the analysis of networks.

It is hoped that this paper is a help in clarifying questions regarding ellipses of standard deviation. Most important of all, the ellipses of standard deviation, as well as the standard deviations of the coordinates, depend on the definition of the coordinate system. The statement that a point has been determined with plus or minus x centimeters is, therefore, misleading and should give rise to further questions. A safe and unique way is to quote the standard deviations of observables, i.e. angles and distances as computed from adjusted coordinates. One may rightfully ask the question, "Why wasn't the method of condition equations (Eq. 2) used in the first place, since it deals with the angles and distances directly and does not use coordinates?" Recall that the method of observation equations has one big advantage: the adjustment can be completely setup by the computer. The simple rule, "each observation—one equation," can be translated readily into the computer language. Even the surveyor's own microcomputer can do the job.

Appendix I.—Rotational Invariance of Ellipses of Standard Deviation

The size and shape of the ellipses of standard deviation, commonly referred to as error ellipses, are given by the variance-covariance matrix, $\Sigma_{\mathbf{X}}$. Let σ_x^2 and

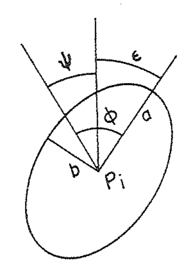

FIG. 8.—Elements of Ellipse of Standard Deviation

σ_y^2 denote the variances of the coordinates for some point P_i, and let σ_{xy} be their covariance. Then the elements of the ellipse of standard deviation are given by the following well-known expressions:

$$\text{semimajor axis} \quad a = \sqrt{\frac{\sigma_y^2 + \sigma_x^2}{2} + \frac{1}{2}w} \quad \text{(29)}$$

$$\text{semiminor axis} \quad b = \sqrt{\frac{\sigma_y^2 + \sigma_x^2}{2} - \frac{1}{2}w} \quad \text{(30)}$$

$$\text{azimuth of semimajor axis} \quad \sin 2\phi = \frac{2\sigma_{xy}}{w} \quad \text{(31)}$$

$$\cos 2\phi = \frac{\sigma_y^2 - \sigma_x^2}{w} \quad \text{(32)}$$

$$w = \sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4\sigma_{xy}^2} \quad \text{(33)}$$

Elements a, b, and ϕ are shown in Fig. 8.

Assume that the coordinate system is rotated by the angle ψ as expressed by the transformation

$$Y = \mathbf{R}\mathbf{X} \quad \text{(34)}$$

$$\Sigma_Y = \mathbf{R}\Sigma_{\mathbf{X}}\mathbf{R}^T \quad \text{(35)}$$

$\mathbf{R}$ is an $(u \times u)$ rotation matrix whose elements are a function of ψ. Using the elements of Σ_Y, Eqs. 29–33 are obtained for the semimajor and semiminor axes of the ellipses, after some algebraic computations. The azimuth of the semimajor axis becomes

$$\epsilon = \phi - \psi \quad \text{(36)}$$

The azimuth, ϵ, differs by the rotation angle, ψ, from the previous azimuth, ϕ.

Thus, it is concluded that the size and the relative orientation of ellipses of standard deviations remain invariant under rotation of the coordinate system.

Appendix II.—Invariant Quantities in Minimal Constraint Solutions

In order to identify the invariant quantities, the least squares solution is derived through appropriate orthogonal transformation. This approach requires knowledge of some theorems from linear algebra which can be found in any textbook on linear algebra, e.g. Graybill (1).

Without loss of generality, it is assumed for the purpose of this derivation that the weight matrix is an identity matrix, i.e. $\mathbf{P} = \mathbf{I}$. There is always an orthogonal transformation of the observation such that the transformed observations have an identity matrix as weight matrix (2). Thus, the least squares solution can be formulated as follows:

$$\mathbf{V} = \mathbf{A}\mathbf{X} + \mathbf{L} \quad \text{(37)}$$

$$\mathbf{P} = \mathbf{I} \quad \text{or} \quad \Sigma_{\mathbf{L}_b} = \sigma_0^2 I \quad \text{(38)}$$

$$\mathbf{V}^T\mathbf{V} = \text{minimum} \quad \text{(39)}$$

Let r denote the rank of the design matrix. Let the matrix $\mathbf{F}$ be an $(n \times r)$ matrix whose columns constitute an orthonormal basis for the column space of $\mathbf{A}$ (one such choice for the columns of $\mathbf{F}$ may be to take the normalized eigenvectors of $\mathbf{A}\mathbf{A}^T$). Let $\mathbf{G}$ be an $(n \times n - r)$ matrix such that $(\mathbf{F}{:}\mathbf{G})$ is orthogonal, i.e. the columns of $\mathbf{G}$ span the orthogonal complement of the column space of $\mathbf{A}$ or, equivalently, $\mathbf{G}$ spans the null space of $\mathbf{A}^T$. From the definition of $\mathbf{F}$ and $\mathbf{G}$, it follows that

$$\begin{pmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{pmatrix} (\mathbf{F}\mathbf{G}) = \begin{pmatrix} \mathbf{F}^T\mathbf{F} & \mathbf{F}^T\mathbf{G} \\ \mathbf{G}^T\mathbf{F} & \mathbf{G}^T\mathbf{G} \end{pmatrix} = \begin{pmatrix} {}_rI_r & 0 \\ 0 & {}_{n-r}I_{n-r} \end{pmatrix} \quad \text{(40)}$$

$$(\mathbf{F}\mathbf{G}) \begin{pmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{pmatrix} = \mathbf{F}\mathbf{F}^T + \mathbf{G}\mathbf{G}^T = I \quad \text{(41)}$$

$$\mathbf{A}^T\mathbf{G} = 0 \quad \text{(42)}$$

$$\mathbf{G}^T\mathbf{A} = 0 \quad \text{(43)}$$

Next, the observation equations Eq. 37 are transformed orthogonally:

$$\begin{pmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{pmatrix} V = \begin{pmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{pmatrix} \mathbf{A}\mathbf{X} + \begin{pmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{pmatrix} \mathbf{L} \quad \text{(44)}$$

Introducing the new variables

$$Z_1 = \mathbf{F}^T\mathbf{L}; \quad Z_2 = \mathbf{G}^T\mathbf{L} \quad \text{(45)}$$

Eq. 44 can be written as

$$\mathbf{V}_Z \equiv \begin{pmatrix} \mathbf{V}_{Z_1} \\ \mathbf{V}_{Z_2} \end{pmatrix} = \begin{pmatrix} \mathbf{F}^T\mathbf{A}\mathbf{X} \\ 0 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \quad \text{(46)}$$

where Eq. 43 was used. The quadratic form (Eq. 39) remains invariant under the orthogonal transformation since

$$\mathbf{V}_Z^T\mathbf{V}_Z = \mathbf{V}_{Z_1}^T\mathbf{V}_{Z_1} + \mathbf{V}_{Z_2}^T\mathbf{V}_{Z_2} = \mathbf{V}^T(\mathbf{F}\mathbf{F}^T + \mathbf{G}\mathbf{G}^T)\mathbf{V} = \mathbf{V}^T\mathbf{V} \quad \text{(47)}$$

The latter equality follows from Eq. 41. The actual quadratic form is obtained from Eq. 46:

$$\mathbf{V}_Z^T\mathbf{V}_Z = (\mathbf{F}^T\mathbf{A}\mathbf{X} + Z_1)^T(\mathbf{F}^T\mathbf{A}\mathbf{X} + Z_1) + Z_2^TZ_2 \quad \text{(48)}$$

According to the principles of least squares, this expression must be minimized through variation of the parameter $\mathbf{X}$. Clearly, $\mathbf{X}$ is the only variable (unknown) in Eq. 48. The matrices $\mathbf{F}$ and $\mathbf{G}$ are determined by the design matrix $\mathbf{A}$, and Z_1 and Z_2 are linear functions of the observations. Inspecting Eq. 48, it is seen that there is no alternative for minimizing $\mathbf{V}_Z^T\mathbf{V}_Z$ but taking $\mathbf{X}$ such that

$$-\mathbf{F}^T\mathbf{A}\mathbf{X} = Z_1 \quad \text{(49)}$$

The solution for $\mathbf{X}$ effectively eliminates the first term in Eq. 48 and yields for the quadratic form

$$\mathbf{V}_Z^T\mathbf{V}_Z \equiv \mathbf{V}^T\mathbf{V} = \mathbf{Z}_2^T\mathbf{Z}_2 = \mathbf{L}\mathbf{G}\mathbf{G}^T\mathbf{L} \quad \text{(50)}$$

Thus, the first invariant quantity has been found. Since $\mathbf{G}$ is only dependent on the design matrix $\mathbf{A}$, and $\mathbf{L} = \mathbf{L}_0 - \mathbf{L}_b$, the minimum of the quadratic form, $\mathbf{V}^T\mathbf{V}$, and, thus, $\mathbf{V}^T\mathbf{P}\mathbf{V}$, is altogether independent of the parameter $\mathbf{X}$. Consequently the *a posteriori* variance of unit weight, $\hat{\sigma}_0$, is also invariant.

Eq. 49 provides r equations for the u parameters. Since $u > r$, and the rank of $\mathbf{F}^T\mathbf{A}$ is r, exactly $u - r$ parameters can be selected arbitrarily; $u - r$ is equal to the rank defect of the normal matrix, which can amount to three or four in two-dimensional networks. There is no limitation on the choice for selecting the $u - r$ parameters in Eq. 49. For example, one may simply equate the last $u - r$ parameters to zero, or to any other constant. This is identical to the procedure considered earlier where a subset of station coordinates was held fixed. The equation $\mathbf{EX} = 0$ constitutes a set of $u - r$ possible conditions which involve all parameters.

Another group of invariant quantities is the residuals. Substituting Eq. 49 into Eq. 46 gives

$$\begin{pmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{pmatrix} \mathbf{V} = \begin{pmatrix} -\mathbf{Z}_1 \\ 0 \end{pmatrix} + \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{G}^T\mathbf{L} \end{pmatrix}; \quad \text{(51)}$$

Thus $$\mathbf{V} = (\mathbf{FG})\begin{pmatrix} 0 \\ \mathbf{G}^T\mathbf{L} \end{pmatrix} = \mathbf{GG}^T\mathbf{L} \quad \text{(52)}$$

which is independent of the specific choice of $\mathbf{X}$. Since $\mathbf{L}_a = \mathbf{L}_b + \mathbf{V}$, the adjusted observations are also invariant with respect to the choice of the minimal constraints. As a consequence of this invariance, it follows from Eq. 37 that the product

$$\mathbf{AX}_i = \mathbf{AX}_j \quad \text{(53)}$$

is invariant with respect to the alternative solution, $\mathbf{X}_i$ and $\mathbf{X}_j$. Each of these solutions has its own variance-covariance matrix, $\Sigma_{\mathbf{X}_i}$ and $\Sigma_{\mathbf{X}_j}$, respectively. Applying the law of variance-covariance propagation to Eq. 49, it follows that

$$\Sigma_{Z_1} = \mathbf{F}^T\mathbf{A}\,\Sigma_{\mathbf{X}_i}\mathbf{A}^T\mathbf{F} = \mathbf{F}^T\mathbf{A}\,\Sigma_{\mathbf{X}_j}\mathbf{A}^T\mathbf{F} \quad \text{(54)}$$

Since $\mathbf{F}$ has full column rank, it must be that

$$\mathbf{A}\Sigma_{\mathbf{X}_i}\mathbf{A}^T = \mathbf{A}\,\Sigma_{\mathbf{X}_j}\mathbf{A}^T \quad \text{(55)}$$

Thus, the variance-covariance matrix of the adjusted observations

$$\Sigma_{\mathbf{L}_a} = \mathbf{A}\,\Sigma_{\mathbf{X}}\mathbf{A}^T \quad \text{(56)}$$

is also invariant with respect to the specific minimal constraints. It follows from Eq. 52 that the covariance matrix of the adjusted residuals

$$\Sigma_V = \hat{\sigma}_0^2\mathbf{GG}^T\mathbf{GG}^T = \hat{\sigma}_0^2\mathbf{GG}^T \quad \text{(57)}$$

is also invariant with respect to the choice of the parameters. Since

$$\Sigma_{\mathbf{L}_a} = \Sigma_{\mathbf{L}_b} - \Sigma_V \quad \text{(58)}$$

an alternative expression of 56 is

$$\Sigma_{\mathbf{L}_a} = \hat{\sigma}_0^2(I - \mathbf{GG}^T) \quad \text{(59)}$$

Finally, the fact that the adjusted observations and their variances and covariances are independent implies that any other distances and angles and their respective variances and covariances, which are computed from the adjusted results, are invariant.

Appendix III.—Geometric Interpretation of Inner Constraints

The minimal constraint solution with the specific constraints

$$\mathbf{EX} = 0 \quad \text{(60)}$$

is called an Inner Constraint Solution. The matrix $\mathbf{E}$ is given in Eqs. 21 and 23 for trilateration and triangulation networks. The conditions of Eq. 60 can be rewritten as follows:

$$\sum_{i=1}^{m} dx_i = 0 \quad \text{(61)}$$

$$\sum_{i=1}^{m} dy_i = 0 \quad \text{(62)}$$

$$\sum_{r=1}^{m} (y_i dx_i - x_i dy_i) = 0 \quad \text{(63)}$$

$$\sum_{r=1}^{m} (x_i dx_i + y_i dy_i) = 0 \quad \text{(64)}$$

The condition (Eq. 64) is used only in triangulation networks where only angles are observed.

In order to arrive at the geometric interpretation, consider a similarity transformation between the adjusted coordinates, $\mathbf{X}_a$, and the approximate coordinates, $\mathbf{X}_0$. The translation rotation, and scale are determined in the least square sense. The mathematical model is

$$\mathbf{X}_a = \Delta + (1 - k)\,\mathbf{R}(\alpha)\,\mathbf{X}_0 \quad \text{(65)}$$

in which Δ = the translation vector; $1 - k$ is the scale factor; and $\mathbf{R}(\alpha)$ = the $(u \times u)$ rotation matrix. Since the approximate coordinates are generally close to the adjusted coordinates, the rotation angle, α, is expected to be small. Each station contributes two equations to the model (65). For a small rotation angle, these equations are for the station, P_i:

$$\begin{pmatrix} x_{ia} \\ y_{ia} \end{pmatrix} = \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix}\begin{pmatrix} x_i \\ y_i \end{pmatrix}(1 - k) \quad \text{(66)}$$

This expression can be simplified as

$$\begin{pmatrix} x_i + dx_i \\ y_i + dy_i \end{pmatrix} = \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} \alpha y_i \\ -\alpha x_i \end{pmatrix} + k\begin{pmatrix} x_i \\ y_i \end{pmatrix} \quad \text{(67)}$$

in which the small terms of the second order were neglected. Thus

$$\begin{pmatrix} dx_i \\ dy_i \end{pmatrix} = \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \begin{pmatrix} \alpha y_i \\ -\alpha x_i \end{pmatrix} + k \begin{pmatrix} x_i \\ y_i \end{pmatrix} \quad \text{(68)}$$

The respective rows in the design matrix are

$$\mathbf{A}: \begin{bmatrix} \Delta x & \Delta y & \alpha & k \\ 1 & 0 & y_i & x_i \\ 0 & 1 & -x_i & y_i \end{bmatrix} \quad \text{(69)}$$

This pattern of the A-matrix repeats for each point. The least squares estimate of the parameters

$$\mathbf{X}^T = (x, y, \alpha, k) \quad \text{(70)}$$

is $\mathbf{X} = -(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{L}$ ..... (71)

in which the identify matrix was used as a weight matrix. The L-vector is

$$\mathbf{L}^T = \mathbf{L}_0^T - \mathbf{L}_b^T = -\mathbf{L}_b^T = (-dx_1, -dy_1, -dx_2, -dy_2, \ldots -dx_m, -dy_m) \quad \text{(72)}$$

Using the full design matrix, which is indicated partially by Eq. 69, the product $\mathbf{A}^T\mathbf{L}$ becomes

$$\mathbf{A}^T\mathbf{L} = \begin{pmatrix} -\sum_{i=m}^{m} dx_i \\ -\sum_{i=1}^{m} dy_i \\ -\sum_{i=1}^{m} (y_i dx_i - x_i dy_i) \\ -\sum_{i=1}^{m} (x_i dx_i + y_i dy_i) \end{pmatrix} \quad \text{(73)}$$

However, the expressions in Eq. 73 are all zero according to the conditions of Eqs. 61–64.

It can be concluded that the inner constraint solutions, the information for the coordinate system definition and the scale, if applicable, is derived from all approximate coordinates. The least squares estimates of the similarity transformation parameters are zero for a transformation between the point fields represented by $\mathbf{X}_a$ and $\mathbf{X}_0$.

Appendix IV.—References

1. Graybill, F. A., *Introduction to Matrices with Applications in Statistics,* Wadsworth Publishing Co., Inc., Belmont, Calif., 1969, Chap. 5.
2. Leick, A., *Adjustment Computations,* University of Maine, Department of Civil Engineering, Orono, Me., 1980.
3. Leick, A., and Tyler, D., "Winter Institute in Surveying Engineering," University of Maine, Department of Civil Engineering, Orono, Me., 1980.
4. Miessl, P., "Zusammenfassung and Ausbau der Inneren Fehlertheorie eines Punkthaufens," *Beiträge zur Theorie der Geodätischen Netze im Raum,* Deutsche Geodätische Kommission, Reihe A, Vol. 61, 1969.
5. Pope, A. J., "Transformation of Covariance Matrices Due to Changes in Minimal Control," Paper presented at American Geophysical Union Fall Meeting, San Francisco, Calif., Dec., 1971.

Community Geographic Data Base[a]

By Henry A. Emery[1]

Abstract: The need for a community geographic and utility management system is reviewed. Information needs are identified, as are potential user groups such as local goverment agencies, utility companies, and engineering and surveying companies. The uses of such a management system are also considered. Potential methods to accomplish a community geographic and utility management information system are also analyzed. The results of several projects in both the public and private sectors are presented.

Introduction

During the next ten years, we will witness the largest computer data entry in the history of mankind. Geographic and utility data will be entered in far greater amounts than what was accomplished during the approximately 40 year life of computers. It is a matter of when, not if, the community geographic information management systems will occur.

Unbelievable?

A major U.S. county government has attempted to count the number of times the information on a subdivision plat was redrawn—they reached 150 redraws and then stopped the count. Unbelievable, isn't it? Another individual attempted to find out how many aerial photography flights occurred in one year over a major metropolitan area, finaling stopping at 50. Again, unbelievable? One utility has more than 100 different map types to maintain costing in excess of $2 million per year. In fact, these maps are not being maintained on a current basis; it would take 80 person years and approximately $2.3 million to eliminate this backlog. Unbelievable? One telephone company had to tell its regulatory agency the number and type of poles in its system. It estimated a $4 million cost for a one time inventory. Unbelievable?

Isolated examples, you contend, but representative of every metropolitan area in the developed world. Every city, every county, every school district, every water district, every sewer district, every recreation district, every electric utility,

[a]Presented at the May 9–15, 1981, ASCE International Convention and Exposition, held at New York, N.Y. (Preprint 81-105).

[1]Div. Mgr., Kellogg Corp., 5601 S. Broadway, Suite 400, Littleton, Colo. 80121.

Note.—Discussion open until January 1, 1983. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on May 26, 1981. This paper is part of the Journal of the Surveying and Mapping Division, Proceedings of the American Society of Civil Engineers, ©ASCE, Vol. 108, No. SU2, August, 1982. ISSN 0569-8073/82/0002-0069/$01.00.