NOAA Technical Report NOS 79 NGS 12

Optimization of Horizontal Control Networks by Nonlinear Programing

Rockville, Md.
August 1979

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Ocean Survey

NOAA Technical Publications

National Ocean Survey/National Geodetic Survey subseries

The National Geodetic Survey (NGS) of the National Ocean Survey (NOS), NOAA, establishes and maintains the basic National horizontal and vertical networks of geodetic control and provides governmentwide leadership in the improvement of geodetic surveying methods and instrumentation, coordinates operations to assure network development, and provides specifications and criteria for survey operations by Federal, State, and other agencies.

NGS engages in research and development for the improvement of knowledge of the figure of the Earth and its gravity field, and has the responsibility to procure geodetic data from all sources, process these data, and make them generally available to users through a central data base.

NOAA Technical Memorandums and some special NOAA publications are sold by the National Technical Information Service (NTIS) in paper copy and microfiche. Orders should be directed to NTIS, 5285 Port Royal Road, Springfield, VA 22161 (telephone: 703-557-4650). NTIS customer charge accounts are invited; some commercial charge accounts are accepted. When ordering, give the NTIS accession number (which begins with PB) shown in parentheses in the following citations.

Paper copies of NOAA Technical Reports, which are of general interest to the public, are sold by the Superintendent of Documents, U.S. Government Printing Office (GPO), Washington, DC 20402 (telephone: 202-783-3238). For prompt service, please furnish the GPO stock number with your order. If a citation does not carry this number, then the publication is not sold by GPO. All NOAA Technical Reports may be purchased from NTIS in hard copy and microform. Prices for the same publication may vary between the two Government sales agents. Although both are nonprofit, GPO relies on some Federal support whereas NTIS is self-sustained.

An excellent reference source for Government publications is the National Depository Library program, a network of about 1,300 designated libraries. Requests for borrowing Depository Library material may be made through your local library. A free listing of libraries currently in this system is available from the Library Division, U.S. Government Printing Office, 5236 Eisenhower Ave., Alexandria, VA 22304 (telephone: 703-557-9013).

NOAA geodetic publications

Classification, Standards of Accuracy, and General Specifications of Geodetic Control Surveys. Federal Geodetic Control Committee, John O. Phillips (Chairman), Department of Commerce, NOAA, NOS, 1974 reprinted annually, 12 pp (PB265442). National specifications and tables show the closures required and tolerances permitted for first-, second-, and third-order geodetic control surveys.

Specifications To Support Classification, Standards of Accuracy, and General Specifications of Geodetic Control Surveys. Federal Geodetic Control Committee, John O. Phillips (Chairman), Department of Commerce, NOAA, NOS, 1975, reprinted annually, 30 pp (PB261037). This publication provides the rationale behind the original publication, "Classification, Standards of Accuracy, ..." cited above.

NOAA Technical Memorandums, NOS/NGS subseries

NOS NGS-1 Use of climatological and meteorological data in the planning and execution of National Geodetic Survey field operations. Robert J. Leffler, December 1975, 30 pp (PB249677). Availability, pertinence, uses, and procedures for using climatological and meteorological data are discussed as applicable to NGS field operations.

NOS NGS-2 Final report on responses to geodetic data questionnaire. John F. Spencer, Jr., March 1976, 39 pp (PB254641). Responses (20%) to a geodetic data questionnaire, mailed to 36,000 U.S. land surveyors, are analyzed for projecting future geodetic data needs.

NOS NGS-3 Adjustment of geodetic field data using a sequential method. Marvin C. Whiting and Allen J. Pope, March 1976, 11 pp (PB253967). A sequential adjustment is adopted for use by NGS field parties.

NOS NGS-4 Reducing the profile of sparse symmetric matrices. Richard A. Snay, June 1976, 24 pp (PB-258476). An algorithm for improving the profile of a sparse symmetric matrix is introduced and tested against the widely used reverse Cuthill-McKee algorithm.

NOS NGS-5 National Geodetic Survey data: availability, explanation, and application. Joseph F. Dracup, June 1976, 45 pp (PB258475). The summary gives data and services available from from NGS, accuracy of surveys, and uses of specific data.

(Continued at end of publication)

NOAA Technical Report NOS 79 NGS 12

Optimization of Horizontal Control Networks by Nonlinear Programing

Dennis G. Milbert

National Geodetic Survey
Rockville, Md.
August 1979

U.S. DEPARTMENT OF COMMERCE
Juanita M. Kreps, Secretary

National Oceanic and Atmospheric Administration
Richard A. Frank, Administrator

National Ocean Survey
Herbert R. Lippold, Jr., Director

PREFACE

This publication has the same title as the author's thesis which was prepared in partial fulfillment of a Master of Science degree from the Ohio State University (Department of Geodetic Science) in 1978. Two minor changes appear in this version. The text and format conform to U.S. Government style and the original computer program listing is omitted.

Mention of a commercial company or product does not constitute an endorsement by NOAA National Ocean Survey. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

CONTENTS

Preface ii

Abstract 1

Introduction 1

Horizontal control network design 1

Second-order network design 3

Practical considerations for network design 4

Mathematical programing 6

Linear programing 6

Nonlinear programing 15

Implementation 18

Variance constraints 19

Cost function 24

Results 25

Quadrilateral deformation 26

Center point quadrilateral 32

Circular traverse 32

Improvements 35

Computer storage allocation 35

Distance observation correlations 37

Relative accuracy constraints 37

Station occupation costs 39

Acknowledgments 40

References 41

OPTIMIZATION OF HORIZONTAL CONTROL NETWORKS BY NONLINEAR PROGRAMING

Dennis G. Milbert
National Geodetic Survey
National Ocean Survey, NOAA
Rockville, Md. 20852

ABSTRACT. Some practical aspects of horizontal control network design are considered, and the techniques of linear and nonlinear programing are briefly reviewed. Rotationally invariant constraints are written for the coordinate variance sum at each station. These constraints are also quasi-homogeneous in the sense that a Moore-Penrose generalized inverse is used in computation. The objective function to be minimized is a cost proportional to the number of observations. Results are displayed for several test networks. Methods of improvement of the design algorithm are then discussed.

INTRODUCTION

Horizontal Control Network Design

For those with the responsibility to supply horizontal control, the design of geodetic networks providing such control is an important problem. Developing a network design to meet particular accuracy standards can be done before any observations are made because the accuracy a particular design attains does not depend upon values of the observations. All that is required is knowledge of the accuracy of the proposed observations. Traditionally, general guidelines developed over many years were the only criteria for network design. Deficiencies discovered after completion of the field work could only be remedied by expensive reobservation.

The invention of the electronic digital computer provided the technical capability of detailed analysis of a design before work commenced. Such detailed analysis ensures that a given network design will meet user requirements. Further, the computational resources provided by modern computers allow modeling of the economic implications of a design. A more unified view of a geodetic network design is the result of these considerations.

Accuracy Requirements

In an ideal situation, accuracy requirements for a horizontal control network are established in response to user needs. The requirements, in turn, dictate the design of the network and the field procedures to be used. One can see that a network design is a reflection of user requirements.

Current horizontal control standards are based upon the distance relative accuracy between directly connected adjacent points (Federal Geodetic Control Committee 1974, p. 3). Such a standard is attractive since users generally wish to survey from one or more control points to one or more other control points. An absolute positional accuracy is not always required, eliminating the need for a variance-covariance matrix of adjusted parameters. This elimination simplifies the network design task, since the variance-covariance matrix of adjusted parameters depends upon the choice of a constrained position. Relative accuracy criteria are invariant with respect to the selection of a single constrained point. Relative accuracies provide a measure of network accuracy without imposing a coordinate "bias" upon the user.

Network Costs

The cost of execution of a particular horizontal network design is an extremely complex problem. Equipment cost and depreciation, field procedures, labor, transportation, communication, support, and administrative costs all contribute to the final cost of a particular network design. Because of the complexity of the economics, examination of some crude approximations to this problem is instructive.

Consider a level theodolite or electronic distance-measurement instrument centered over the monument, operated by experienced personnel. One can roughly approximate the cost of the observations by the time it takes to perform the observations. One can see that the costs are proportional to the number of repetitions.

The next level of approximation can be added by considering the time necessary to recover the mark, erect and dismantle a portable observation tower (if necessary), and center and level the instrument. This cost is essentially a one-time cost proportional to the number of occupied stations. This cost would be in addition to the cost of observations discussed above.

Further sophistication can be introduced by transportation costs. These costs are roughly proportional to the distance between stations. Terrain and roads, however, dominate the estimate of time spent in transport. An adequate terrain model vastly increases the cost model complexity.

In addition to the above factors, the assignment of labor within the field party will have a major effect upon the total cost (Gergen 1978). Such an assignment would efficiently coordinate reconnaissance, transport, tower construction and dismantlement, light-keeping, and observation. The most efficient task allocation does not seem to vary in any simple fashion with respect to the variation of the elements of a network. Moreover, an optimal task allocation would require a detailed model of field methods in which a subtask would be to perform the observations necessary to fulfill some desired specifications.

In closing, this topic illustrates the complexity inherent in a detailed cost model and serves as a guide to the cost model selected for investigation of horizontal network design.

Second-Order Network Design

In second-order network design, the variance-covariance matrix of the observations is unknown, and is solved using a known network configuration and a desired variance-covariance matrix of the station coordinates (Grafarend 1974, p. 720). Second-order design is the converse of first-order design, where observation variances are known and the network configuration is unknown.

Second-order design merits attention for two reasons. First, a large number of practical difficulties plague the implementation of a first-order design. The geodesist rarely has complete flexibility in the placement of stations. Terrain, property rights, station intervisibility, presence of old control stations, and specific user requirements all impose constraints upon the placement of new stations (Hoyle 1977). Second, the simplest approximation to the cost of a horizontal control network is proportional to the number of observations. Because second-order design solves for the observation variance-covariance matrix, it is possible to compute the cost of a given design since one may compute the number of repetitions of each observation given the variance of each single observation.

The second-order design problem does not yield a unique solution (Bossler et al. 1973). An infinite number of observation covariance matrices will satisfy a required coordinate covariance matrix for a given network configuration. Given this multiplicity, additional constraint is needed to determine a unique solution.

A desirable quality of a second-order network design would be that it would have a minimum cost expressed in some approximate cost model. The solution of such a problem lies in the realm of optimization under constraints, also known as mathematical programing. In such a solution, the costs of a design are minimized while satisfying specified accuracy requirements.

The parametrization of the costs and accuracy requirements in terms of observation repetitions allow such a solution to easily proceed. The dual problem, maximizing the design accuracy while meeting a desired cost requirement, is also solvable by mathematical programing. The dual problem is not pursued further in this report.

Practical Considerations for Network Design

Many practical considerations enter into the network-design problem. The model, types of observations, procedures used in the field, and auxiliary design requirements all need to be considered before a network-design algorithm can be constructed. Allowing for such requirements at an early state ensures a useful design tool for the geodesist.

Model

Assume that all control points lie on a plane. Even for large control networks, this approximation will be quite good. In addition, only direction and distance measurements are considered in the design. These are generally the only observation types at the economic disposal of the design agency.

Using a standard, two-dimension x-y Cartesian coordinate system with azimuth measured clockwise from the north (y axis), one easily derives the observation equations (Schwarz 1974-/5) For a direction from station i to station j,

$$\delta d_{ij} = -\frac{\cos\alpha_{ij}}{S_{ij}}\,\delta x_i + \frac{\sin\alpha_{ij}}{S_{ij}}\,\delta y_i + \frac{\cos\alpha_{ij}}{S_{ij}}\,\delta x_j - \frac{\sin\alpha_{ij}}{S_{ij}}\,\delta y_j + \delta z_i$$

where z_i is the rotation unknown for the direction set at station i; α_{ij}, the azimuth from station i to station j; and S_{ij}, the distance from station i to station j. For distance from station i to station j,

$$\delta S_{ij} = -\sin\alpha_{ij}\delta x_i - \cos\alpha_{ij}\delta y_i + \sin\alpha_{ij}\delta x_j + \cos\alpha_{ij}\delta y_j.$$

The matrix symbol for the coefficients of the differential changes in the observations with respect to the parameters (station coordinates) is A, the design matrix. Since a second-order network design approach is being used, the coordinates of the stations are assumed to be known so that the design matrix is completely determined.

Observation Variance-Covariance Matrix

In second-order network design, the observation variance-covariance matrix, Σ, is solved with respect to a given station configuration and a required coordinate variance-covariance matrix, Σ_x. To determine Σ, one must develop some functional relation in terms of observation repetitions. In this report, we assume that the observations are uncorrelated and the repetitions of each observation are also uncorrelated.

Now, directions are not estimable quantities; but McKay (1973, pp. 9-15) has shown that, with a constraint of the sum of the residuals equaling zero, the adjusted directions are uncorrelated. However, it is not the case that repetitions of distance observations are uncorrelated. One approach to this problem is considered at the end of this report.

Another technique that can use uncorrelated observations is the weight-recovery method of Sprinsky (1974, pp. 203-204). This method computes a unique set of observation weights that will satisfy a required coordinate variance-covariance matrix. The user then selects the type and number of observations from this set. Note that this method may or may not lead to a minimum cost solution.

Field Procedures

When designing a network, special consideration should be made for field operations. It is quite possible for an observation scheme developed in the office to be difficult, if not impossible, to execute in the field. Direction observations are particularly sensitive to this problem. It is extremely cumbersome to develop a scheme of theodolite pointings that would produce direction observations of desired weights. Field procedures require an equal number of pointings to all sighted stations (Dracup and Fronczek 1977). A network-design algorithm should reflect this problem.

Because of the size and distance of a target, different variances may be obtained on direction observations from the same station with the same instrument. Additionally, one may have to perform observations with instruments of different accuracies. A network-design program should be flexible enough to accommodate such cases.

Auxiliary Design Requirements

A geodetic network is not characterized only by some ideal variance-covariance matrix. The design geodesist may also wish to satisfy some auxiliary design criteria. One such criterion is a requirement that each point be redundantly determined to some degree. Such a redundancy requirement is extremely important

since it provides a mechanism for blunder detection and a margin of safety in the event some observation must be deleted.

When using second-order network design and solving for observation repetitions, redundancy requirements can easily be met. This is done by never allowing a network design to have zero repetitions of an observation. This ensures that each observation will be made at least once and the design algorithm will never delete an observation automatically. Still, the design geodesist must initially select a set of observations that will satisfy those redundancy requirements in effect.

MATHEMATICAL PROGRAMING

When maximizing or minimizing some function subject to certain constraints, the problem is said to fall in the realm of optimization under constraints, also known as mathematical programing. The general form for such a problem is

$$\text{max or min} \quad z = f(y_1, \ldots, y_n) \tag{1}$$

where

$$g_i(y_1, \ldots, y_n) \lesseqqgtr b_i \tag{2}$$

$$y_j \geq 0 \quad (j = 1 \text{ to } n) \tag{3}$$

and where $\lesseqqgtr$ represents $\leq$, =, or $\geq$. Eq (1) is the objective function, which is the function to be optimized. Eq (2) is the constraint. Eq (3) covers the nonnegativity restrictions that are assumed to be present in a programing problem although they may not always be explicitly written.

Linear Programing

When f and g_i are linear functions of y_j in eq (1) and (2), the mathematical programing problem is said to be a linear programing problem. This then has the form

$$\text{max or min } z = c_1y_1 + \ldots + c_ny_n \tag{4}$$

where

$$a_{i1}y_1 + \ldots + a_{in}y_n \lesseqqgtr b_i. \tag{5}$$

Note that the nonnegativity restrictions are still in effect although they are not formally expressed. Such a linear problem has a solution computed through a procedure named the Simplex algorithm (Gillett 1976). The remainder of this section will discuss this algorithm.

To best understand the Simplex algorithm, consider

$$\max z = 12y_1 + 18y_2 \tag{6}$$

$$2y_1 + y_2 \leq 4 \tag{7}$$

$$y_1 + 2y_2 \leq 4. \tag{8}$$

A graphical representation of this problem is presented in figure 1.

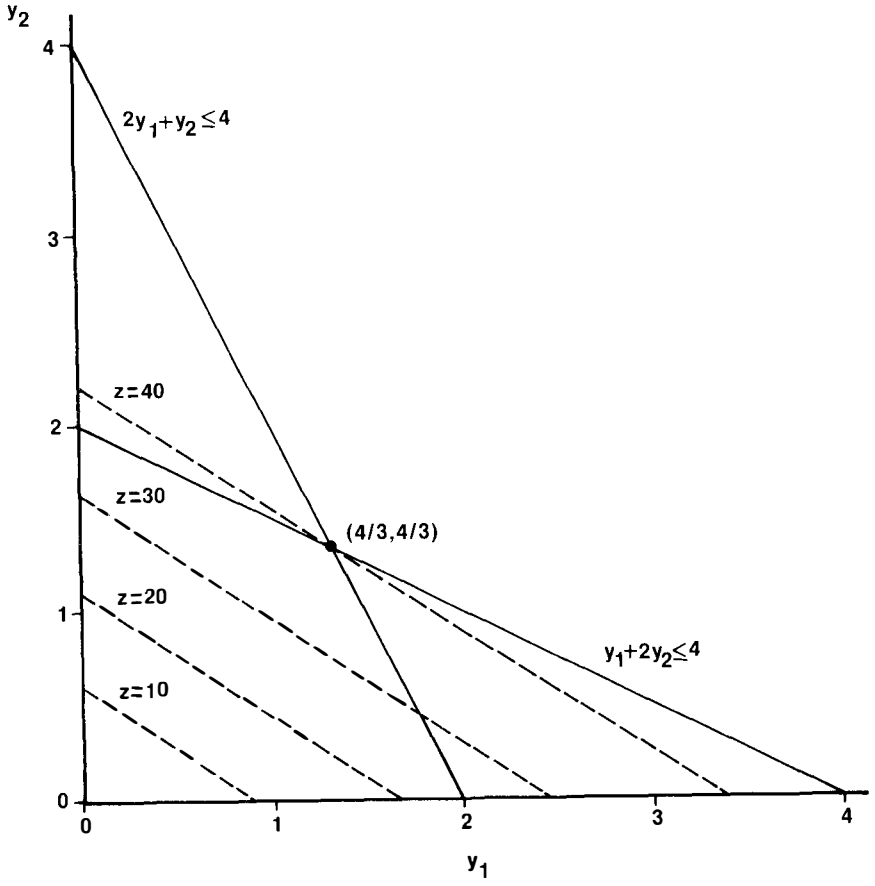

Figure 1.--Linear programing sample.

The nonnegativity requirements ensure that only positive values of y_1 and y_2 are considered so that any possible solution to the problem lies in the first quadrant. Plotting the constraint of eq (7), one sees that the range of possible solutions is now bounded by the lines which intersect at (0,4), (2,0), and (0,0). Similarly, the possible solutions described by eq (8) are bounded by the lines that intersect at (0,2), (4,0), and (0,0). The intersection of these two regions represents the possible solutions to both eq (7) and (8) together. These are called the feasible solutions and are represented by the region bounded by the lines that intersect at (0,2), (4/3,4/3), (2,0), and (0,0).

Now, the region of feasible solutions contains an infinite number of solutions that satisfy the constraints; however, one must find a feasible solution which optimizes the objective function. In figure 1, the objective function is plotted as a dotted line for different values of z; one can see that z achieves a maximum feasible solution at (4/3,4/3) where z = 40. This is the optional feasible solution to this particular linear programing problem.

One can prove, although not done so here, that the optional solution will occur either at one corner or along one edge of the region of feasible solutions in any linear programing problem. If the optional solution falls along an edge of the region, then the solution possesses multiple optimal solutions. The Simplex algorithm starts at one corner of the region and proceeds to another corner such that the objective function is nondecreasing (or nonincreasing in a minimization problem). This cycle is repeated until a corner is reached where all other possible corners have an equal or smaller (or an equal or larger for minimization) objective function. This point is then the optimal feasible solution to the problem.

To apply the Simplex algorithm, one must place the linear programing problem into an equivalent form. This form must satisfy three conditions.

1. All constraints must be represented by equalities.

2. Each constraint must contain a variable possessing a coefficient of 1 in that constraint and 0 in all other constraints.

3. The right-hand side of each constraint, b_i, must be greater than or equal to zero.

Consider, once again, the example. Condition 3 is already satisfied in that the constraints (7) and (8) have positive right-hand sides. Conditions 1 and 2 can be satisfied together by creating a slack variable for each constraint. For less than or equal constraints, this is done by simply adding a unique variable to each less than or equal constraint, and representing the variable with a zero coefficient in the objective function. The example then becomes

$$\max z \quad z - 12y_1 - 18y_2 + 0y_3 + 0y_4 = 0$$

$$2y_1 + y_2 + y_3 = 4$$

$$y_1 + 2y_2 + y_4 = 4$$

where y_3 and y_4 are the slack variables. They are so named because they "take up the slack" in the inequalities. Note that the nonnegativity restrictions apply to the slack variables as well as the original variables. This, then, is the equivalent form of the example.

Suppose we desire to minimize the objective function rather than maximize it. This is expressed in the equivalent form by an objective function in which the coefficients are the negative of the original. Therefore, if it is desired to optimize

$$\min z \quad z - 12y_1 - 18y_2 + 0y_3 + 0y_4 = 0,$$

then the proper equivalent form is expressed as

$$\max z \quad z + 12y_1 + 18y_2 + 0y_3 + 0y_4 = 0.$$

When this is done, remember that the optimal value of the objective function will be -z when the solution is obtained.

Consider the example where eq (8) is now an equality rather than an inequality; then

$$\max z = 12y_1 + 18y_2 \tag{6}$$

$$2y_1 + y_2 \leq 4 \tag{7}$$

$$y_1 + 2y_2 = 4. \tag{9}$$

Eq (9) already satisfies conditions 1 and 3 as it stands. To satisfy condition 2, add another variable called an artificial variable. If this variable, however, possesses any value other than 0 in the final solution, then eq (9) will be incorrect and the solution will be wrong. To ensure that the artificial variables will be 0 in the final solution, we include them in the objective function with a small negative coefficient, -T. Then the Simplex algorithm will "drive out" the artificial variables in the process of finding a solution. The equivalent form of the example now becomes

$$\max z \quad z - 12y_1 - 18y_2 + 0y_3 - Ty_4 = 0$$

$$2y_1 + y_2 + y_3 = 4$$

$$y_1 + 2y_2 + y_4 = 4$$

where y_3 is a slack variable and y_4 is an artificial variable.

Consider the example when eq (8) is now a greater than or equal inequality rather than a less than or equal to inequality:

$$\max z = 12y_1 + 18y_2 \tag{6}$$

$$2y_1 + y_2 \leq 4 \tag{7}$$

$$y_1 + 2y_2 \geq 4. \tag{10}$$

To satisfy condition 1, add a slack variable to eq (10). Since such a variable must be nonnegative because of the nonnegativity restrictions, it must be entered into eq (10) with a negative coefficient:

$$y_1 + 2y_2 - y_4 = 4. \tag{11}$$

Notice that eq (11) still does not satisfy condition 2 since it has a coefficient of -1 rather than +1. Therefore, an artificial variable must also be added to eq (11). Once this is done, the proper equivalent form is found:

$$\max z \quad z - 12y_1 - 18y_2 + 0y_3 + 0y_4 - Ty_5 = 0$$

$$2y_1 + y_2 + y_3 = 4$$

$$y_1 + 2y_2 - y_4 + y_5 = 4$$

where y_3 and y_4 are slack variables and y_5 is an artificial variable.

Now that the technique for converting any linear programing problem into an equivalent form for Simplex algorithm solution has been explored, all that remains in this discussion is to examine how the Simplex algorithm computes the optimal solution from an equivalent form. Consider the equivalent form of the original example:

$$\max z \quad z - 12y_1 - 18y_2 + 0y_3 + 0y_4 = 0 \tag{12}$$

$$2y_1 + y_2 + y_3 = 4 \tag{13}$$

$$y_1 + 2y_2 + y_4 = 4. \tag{14}$$

The constraints (13) and (14) contain a total of four variables, giving two equations in four unknowns. If any two of the variables are set to 0, then a unique solution exists for the remaining two variables. These two remaining variables are called the basic variables and there are $\binom{4}{2} = 6$ ways of choosing them from the set of four variables. When this is done, (y_1, y_2) achieves the values of (0,2), (0,4), (4/3,4/3), (4,0), (2,0), and (0,0) in figure 1. These are called the basic solutions to the linear programing problem.

Notice that two of the basic solutions, (0,4) and (4,0), do not lie in the region of feasible solutions. The Simplex algorithm only works upon the basic feasible solutions (i.e., the subset of basic solutions contained in the region of feasible solutions). This subset is the set of corners (0,2), (4/3,4/3), (2,0), and (0,0). The optimal solution, (4/3,4/3), is the optimal basic feasible solution.

To ensure the Simplex algorithm can always start with a basic feasible solution, condition 2 was imposed upon all equivalent forms. Requiring each constraint in the equivalent form to possess a variable with a coefficient of 1 in that constraint and 0 in all others enables the Simplex algorithm to always begin with a basic feasible solution. This is done by making each artificial variable and each slack variable with a coefficient of 1 a basic variable. All the nonbasic variables are initially set to 0. In the example, y_3 and y_4 are the basic variables. Furthermore, the following are immediately seen:

$$y_1 = 0 \qquad y_2 = 0$$

$$y_3 = 4 \qquad y_4 = 4$$

and

$$z = 0.$$

The Simplex algorithm "moves" from one basic feasible solution to the next by making one basic variable a nonbasic variable and by making a nonbasic variable a basic variable. The basic variable that becomes nonbasic is said to "leave" the basis, and the nonbasic variable which becomes basic is said to "enter" the basis.

The Simplex algorithm selects the variables to leave and enter the basis in such a fashion that the objective function is non-decreasing. The variable with the most negative coefficient in the objective function is the variable to next enter the basis. If all the coefficients are greater than or equal to zero, then the optimal solution has been achieved. Examination of eq (12) shows that y_2 will next enter the basis.

Next, the ratios of the right-hand side of each constraint to the corresponding coefficient of the newly entering variable are computed. For example,

$$4/1 = 4 \quad \text{from eq (13);}$$

and

$$4/2 = 2 \quad \text{from eq (14).}$$

If any of these ratios is not positive, then a finite solution does not exist; that is, the objective function is not bounded by the constraints. The example is bounded, and the ratios are positive. The variable with a coefficient of 1 in the equation that corresponds to the minimum positive ratio will be the variable to next leave the basis. The minimum positive ratio corresponds to eq (14), and y_4 is the variable with a coefficient of 1 in that equation. Therefore, y_4 will next leave the basis.

Once the variables to next leave and enter the basis have been identified, the Simplex algorithm performs a change of basis. The change of basis is a series of elementary transformations executed so that:

1. The new basic variable will have a coefficient of 1 in the constraint equation that currently holds the variable which is leaving the basis.

2. The new basic variable has a coefficient of 0 in all the other constraints and objective equations.

Performing a change of basis from y_4 to y_2, one executes step 1 by dividing eq (14) by 2 to yield

$$1/2y_1 + y_2 + 1/2y_4 = 2. \tag{15}$$

Eq (15) is then multiplied by 18 and added to eq (12); then eq (15) is multiplied by -1 and added to eq (13). This yields a new set of equations:

$$z - 3y_1 + 0y_2 + 0y_3 + 9y_4 = 36 \tag{16}$$

$$1\tfrac{1}{2}y_1 + y_3 - 1/2y_4 = 2 \tag{17}$$

$$1/2y_1 + y_2 + 1/2y_4 = 2. \tag{15}$$

Now y_2 and y_3 are the basic variables. One immediately sees the following:

$$y_1 = 0 \qquad y_2 = 2$$

$$y_3 = 2 \qquad y_4 = 0$$

and

$$z = 36.$$

This completes one cycle of the algorithm. Since a negative coefficient still exists in the objective function, one knows that the solution is not yet optimal.

To complete the solution, execute a second cycle. The most negative coefficient in the objective function belongs to the variable y_1. It shall next enter the basis. When computing the ratios, the values obtained are

$$(2/1\tfrac{1}{2}) = 4/3 \qquad \text{for eq (17)}$$

and

$$(2/\tfrac{1}{2}) = 4 \qquad \text{for eq (15).}$$

The smaller positive ratio belongs to eq (17), indicating that the variable y_3 will next leave the basis. Multiplying eq (17) by 2/3 becomes

$$1y_1 + 2/3y_3 - 1/3y_4 = 4/3. \tag{18}$$

Eq (18) is then multiplied by 3 and added to eq (16); then eq (18) is multiplied by -1/2 and added to eq (15). This yields a new set of equations:

$$z + 0y_1 + 0y_2 + 2y_3 + 8y_4 = 40$$

$$y_1 + 2/3y_3 - 1/3y_4 = 4/3$$

and

$$y_2 - 1/3y_3 + 2/3y_4 = 4/3.$$

Now y_1 and y_2 are the basic variables. Thus one immediately sees:

$$y_1 = 4/3 \qquad y_2 = 4/3$$

$$y_3 = 0 \qquad y_4 = 0$$

and

$$z = 40.$$

The objective function no longer contains any negative coefficients. The solution is optimal and agrees with the solution obtained through graphical considerations.

In general terms, the equivalent form of the linear programing problem is

$$\max z \quad z - \sum_{j=1}^{n} c_j y_j = b_o \tag{19}$$

$$\sum_{j=1}^{n} a_{ij} y_j = b_i \tag{20}$$

where i is 1 to m over the basic variables.

Let x_k be the entry variable with the most negative coefficient in eq (19). If all coefficients are nonnegative, the solution

$$y_i = b_i \qquad i = 1,m \text{ (basic)}$$

is optimal. If not, the ratios are formed:

$$\frac{b_i}{a_{ik}} \qquad i = 1,m \text{ (basic)};$$

and the smallest positive ratio is chosen and represented by b_r/a_{rn} where r represents the leaving variable, y_r, and the rth pivotal equation with the a_{rk} pivotal element. If any ratio is negative, a bounded solution does not exist. If not, a change of basis is performed from y_r to y_k. The change of basis is nothing more than the transformations used in the Gauss-Jordan method of solving linear equations. The pivotal equation becomes

$$\sum_{j=1}^{n} \frac{a_{rj}}{a_{rk}} y_i = \frac{b_r}{a_{rk}} \tag{21}$$

where the pivotal element becomes 1. Now eq (21) is multiplied by c_k and added to eq (19). The new objective equation becomes

$$z - \sum_{j=1}^{n} \left(c_j - c_k \frac{a_{rj}}{a_{rk}}\right) y_j = b_o - \frac{c_r b_r}{a_{rn}}$$

where the j=k coefficient becomes 0.

The remaining equations are computed by multiplying the pivotal eq (21) by a_{ik} and subtracting the result from equations i where i equals 1 to m (basic) and i is not equal to r. Then eq (20) becomes

$$\sum_{j=1}^{n} \left(a_{ij} - \frac{a_{rj} a_{ik}}{a_{rk}}\right) y_j = b_i - \frac{a_{ik} b_r}{a_{rk}} \quad (i \neq r)$$

and

$$\sum_{j=1}^{n} \frac{a_{ri}}{a_{rk}} y_j = \frac{b_r}{a_{rk}} \quad (i = r).$$

This completes one cycle of the Simplex algorithm. The procedure repeats until all the coefficients of eq (22) become negative.

Nonlinear Programing

When f or any g_i is not a linear function of y_j in eq (1) or (2), the mathematical programing problem is said to be a nonlinear programing problem. No algorithm currently exists to solve an arbitrary nonlinear programing problem with the ease and reliability that the Simplex algorithm displays in the solution of the general linear programing problem.

Although no simple general algorithm exists to exactly solve an arbitrary nonlinear problem, a large number of approaches exist for the nonlinear programing problem solution. Some methods use a variety of approximations--some are iterative, and some require derivatives of the functions. An approximation method is now considered.

Separable Functions

A function is said to be separable if it can be written in the form

$$g_1(y_1) + \ldots + g_n(y_n) \lesseqgtr b.$$

This quality is important since, if a function is separable, an approximation can be made to potentially linearize such a function. An approximate linearization would allow the application of the powerful Simplex algorithm to solve for the approximate solution.

Note that a function which appears to be nonseparable may, in fact, be separable. Such a function may be transformed by a change of variables into a clearly separable function. Consider

$$y_1 y_2 \leq b \tag{23}$$

and create two new variables:

$$y_1 = y_3 + y_4$$

and

$$y_2 = y_3 - y_4.$$

The eq (23) can be replaced by

$$y_3^2 - y_4^2 \leq b$$

$$y_1 - y_3 - y_4 = 0$$

and

$$y_2 - y_3 + y_4 = 0.$$

Therefore, eq (23) was indeed a separable function although it did not appear so at first.

Piecewise Linear Approximation

Consider the separable function

$$g_1(y_1) + \ldots + g_n(y_n) \lesseqgtr b.$$

One can approximate each element of the function, g_j, over a range of values U_{oj} to U_{mj} for the y_j variable. The values U_{ij} are known as the break points of the piecewise linear approximation. A set of m new variables is created for each y_j so that

$$y_j = y_{1j} + \ldots + y_{mj} \tag{23}$$

where

$$0 \leq y_{ij} \leq U_{i,j} - U_{i-1,j} \; .$$

For eq (23) to give a satisfactory approximation, impose the condition that,

$$\text{if } y_{ij} > 0, \text{ then } y_{i-1,j} = U_{i-1,j} \; . \tag{24}$$

Exactly how this condition is imposed is discussed later.

The function g_j is then approximated as a linear function over each adjacent pair of break points. This gives an approximation

$$g_j(y_j) \doteq g(U_o) + c_{1j}y_{1j} + \ldots + c_{mj}y_{mj} \tag{25}$$

over the range U_{oj} to U_{mj}. Here, the piecewise slope, c_{ij}, is defined as

$$c_{ij} = \frac{g_j(U_{ij}) - g_j(U_{i-1,j})}{U_{ij} - U_{i-1,j}} \; . \tag{26}$$

At this point, computing an example is constructive. Consider the function $g(y) = 1/y$ over the region $1 \leq y \leq 6$. The function is to be approximated with six equally spaced break points at the integers. In this case, $m = 5$. The results are displayed in table 1.

Table 1.--Piecewise approximation sample

i	U_i	$g(U_i)$	c_i
0	1	1	
1	2	1/2	-1/2
2	3	1/3	-1/6
3	4	1/4	-1/12
4	5	1/5	-1/20
5	6	1/6	-1/30

The expression for y is now

$$y = 1 + y_1 + y_2 + y_3 + y_4 + y_5 \ .$$

Now, the piecewise linear approximation of g(y) is

$$g(y) \doteq 1 - 1/2y_1 - 1/6y_2 - 1/12y_3 - 1/20y_4 - 1/30y_5,$$

and is graphed in figure 2. One can see that this approximation becomes much better for large values of y; and, of course, the approximation is exact at the break points.

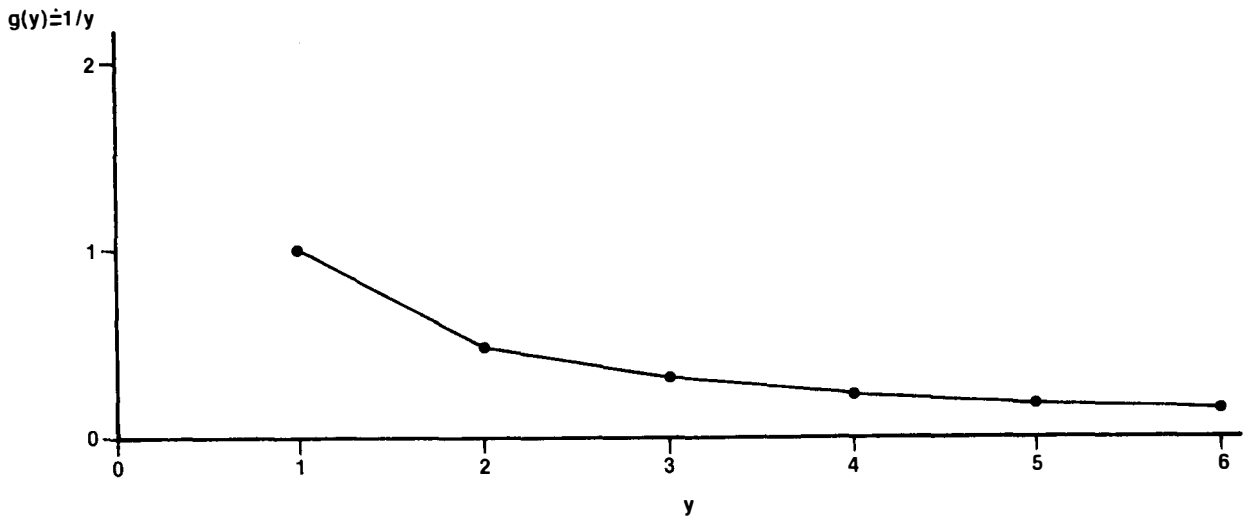

Figure 2.--Piecewise linear approximation of 1/y .

Suppose a solution is desired for the nonlinear programing problem

$$\min z = 2y$$

where

$$1/y \leq 0.2.$$

The objective function is linear, and the constraint is nonlinear and separable. Using a piecewise linear approximation at integer break points from 1 to 6, the problem may be reformulated as

$$\min z = 2 + 2y_1 + 2y_2 + 2y_3 + 2y_4 + 2y_5$$

$$1 - 1/2y_1 - 1/6y_2 - 1/12y_3 - 1/20y_4 - 1/30y_5 \leq 0.2$$

where $y_i \leq 1$, and $i = 1,5$.

Manipulating the first constraint algebraically, the problem becomes

$$\min z = 2 + 2y_1 + 2y_2 + 2y_3 + 2y_4 + 2y_5$$

and

$$1/2y_1 + 1/6y_2 + 1/12y_3 + 1/20y_4 + 1/30y_5 \geq 0.8$$

where $y_i \leq 1$, and $i = 1,5$. From this formulation and by figure 2, one can see that the optimal means of satisfying the constraint is by using the largest possible value of y_i before using y_{i+1}. Thus the condition expressed by eq (24) is automatically satisfied by the form of the piecewise linear approximation. One can show that this will always be the case for an arbitrary function which is said to be convex. This, however, leads into the theory of convex sets and is beyond the intended scope of this report. The interested reader should refer to Hadley (1961).

IMPLEMENTATION

A number of correspondences must be made to implement second-order network design in the form of a mathematical programing problem:

$$\min z = f(y_1, \ldots, y_n) \tag{27}$$

and

$$g_i(y_1, \ldots, y_n) \leq b_i \tag{28}$$

where $i = 1,m$. The variable, y_j will represent the repetitions of the jth observation. Eq (28) will be constraints upon the design coordinate variances. This will ensure that the desired coordinate variances, b_i, are always satisfied. The objective function (27) is in this particular problem the cost function. This will model the costs incurred by a particular combination of observations. The remainder of this section is devoted to

the formulation of the nonlinear programing problem to optimize a second-order, horizontal control geodetic network design problem.

Variance Constraints

From the theory of least-squares adjustment (Uotila 1967),

$$\Sigma_x = (A^t \Sigma^{-1} A)^{-1} \qquad (29).$$

The design matrix, A, was developed in the first section and is completely known. The variance-covariance matrix of the observations, Σ, is unknown and will be a function of the observation repetitions. The variance-covariance of the coordinates, Σ_x, will contain the required coordinate variances and is partially known.

We assume that the repetitions of an observation are uncorrelated so that

$$\sigma^2_{m_i} = \frac{\sigma^2_i}{n_i}$$

where σ^2_i is the variance of a single observation i, n_i is the repetitions of the ith observation, and $\sigma^2_{m_i}$ is the variance of the mean of the repetitions of the ith observation. We also assume that the observations are uncorrelated, yielding the variance-covariance matrix

$$\Sigma = \begin{bmatrix} \frac{\sigma_1^2}{n_1} & & & 0 \\ & \frac{\sigma_2^2}{n_2} & & \\ & & \ddots & \\ 0 & & & \frac{\sigma_n^2}{n_n} \end{bmatrix}$$

This matrix has the inverse

$$\Sigma^{-1} = \begin{bmatrix} \frac{n_1}{\sigma_1^2} & & & 0 \\ & \frac{n_2}{\sigma_2^2} & & \\ & & \ddots & \\ 0 & & & \frac{n_n}{\sigma_2^2} \end{bmatrix}$$

For an arbitrary design matrix, $A^t\Sigma^{-1}A$ is a nondiagonal matrix of variables, n_i. No simple form exists for the inverse of a matrix of variables--so an approximation used by Greve (1972, pp. 39-40) is applied to the problem. Then

$$\Sigma_x \doteq G\Sigma_{i+1}G^t \tag{30}$$

where,

$$G = (A^t\Sigma_i^{-1}A)^{-1}A^t\Sigma_i^{-1} .$$

This is an iterative equation that approximates the variance-covariance matrix of the coordinates in terms of a current estimate of the repetitions of each observation and of prior estimates of the optimal repetitions of the observations. Therefore, the design algorithm will be iterative and require an initial estimate of the optimal repetitions.

Rotationally Invariant Constraints

Having developed a relation for the variance-covariance matrix at coordinates in terms of observation repetitions, one can write $1/2u^2$ constraint inequalities. These inequalities would act as constraints on the variances and covariances of the coordinates. Practice has been to use m pairs of constraints on the variances of the x and y coordinates:

$$f_{x_i}(y_1, \ldots, y_n) \leq \sigma^2_{x_i}$$

and

$$f_{y_i}(y_1, \ldots, y_n) \leq \sigma^2_{y_i}$$

where i = 1,m stations.

Using the approximation

$$\Sigma_x = G\Sigma G^t$$

and the fact that Σ is diagonal with elements σ_i^2/n_i, one sees that

$$\begin{bmatrix} g^2_{x_{11}} & \cdots\cdots\cdots & g^2_{x_{1n}} \\ g^2_{y_{11}} & & g^2_{y_{1n}} \\ \vdots & \ddots & \vdots \\ g^2_{x_{m1}} & \cdots\cdots\cdots & g^2_{x_{mn}} \\ g^2_{y_{m1}} & \cdots\cdots\cdots & g^2_{y_{mn}} \end{bmatrix} \begin{bmatrix} \sigma^2_1/n_1 \\ \vdots \\ \sigma^2_n/n_n \end{bmatrix} = \begin{bmatrix} \sigma^2_{x_1} \\ \sigma^2_{y_1} \\ \vdots \\ \sigma^2_{x_m} \\ \sigma^2_{y_m} \end{bmatrix}$$

for the diagonal elements of the variance-covariance matrix of the coordinates, $\sigma^2_{x_i}$ and $\sigma^2_{y_i}$. Therefore, the constraints for the ith station have the form

$$\sum_{j=1}^{n} g^2_{x_{ji}} \quad \sigma^2_j/n_j \quad \leq \quad \sigma^2_{x_i} \tag{31}$$

and

$$\sum_{j=1}^{n} g^2_{y_{ji}} \quad \sigma^2_j/n_j \quad \leq \quad \sigma^2_{y_i} \tag{32}$$

where $\sigma^2_{x_i}$ and $\sigma^2_{y_i}$ are the maximum allowable variances in x and y at the ith station.

The drawback to these constraints is that different optimal solutions are computed for different orientations of a given configuration. As discussed in the beginning of this report, control networks are based upon a standard independent of network orientation. One desires the design algorithm to reflect this property. It happens that the sum of the variances of the x and y coordinates is invariant with respect to rotation of the network or coordinate system. Therefore, a set of constraints written in terms of the sum of the coordinate variances should give a consistent optimal solution independent of the network or coordinate system orientation. The constraints for $1/2u = m$ stations would have the form

$$f_i'(n_1, \ldots, n_n) \leq \sigma^2_{x_i} + \sigma^2_{y_i} \quad (i = 1,m)$$

where f_i' is the sum of f_{x_i} and f_{y_i}. In terms of elements of the G matrix, this gives

$$\sum_{j=1}^{n} \left(g^2_{x_{ji}} + g^2_{y_{ji}}\right) \frac{\sigma^2_j}{n_j} \leq \sigma^2_{x_i} + \sigma^2_{y_i} \quad . \tag{33}$$

These new constraints are merely the pairwise sum of eq (31) and (32). Note that these constraints are functions of the reciprocal of the observation repetitions and are therefore nonlinear.

Piecewise Linear Approximation

Since the variance constraints are nonlinear in the variable n_j, the Simplex algorithm cannot be used. However, the constraints (33) are immediately seen to be separable. Applying the technique of piecewise linear approximation, discussed previously, the constraints are approximated over ℓ break points at integer spacing from 1 to ℓ. Then for the ith station,

$$\sum_{j=1}^{n} \sigma^2_j \left(g^2_{x_{ji}} + g^2_{y_{ji}}\right)\left(1 - \sum_{k=1}^{\ell-1} \frac{n_{jk}}{k(k+1)}\right) \leq \sigma^2_{x_i} + \sigma^2_{y_i} \tag{34}$$

where $0 \leq n_{jk} \leq i$ for all j and k. Now each variable, n_{jk}, becomes the kth repetition of the ith observation. When rearranging terms, eq (34) becomes

$$\sum_{j=1}^{n} \sigma_j \left(g^2_{x_{ji}} + g^2_{y_{ji}}\right) \sum_{k=1}^{\ell-1} \frac{-n_{jk}}{k(k+1)} \leq \sigma^2_{x_i} + \sigma^2_{y_i} - \sum_{j=1}^{n} \sigma^2_j \left(g^2_{x_{ji}} + g^2_{y_{ji}}\right)$$

and the number of repetitions of the jth observation is

$$n_j = 1 + \sum_{k=1}^{\ell-1} n_{jk}.$$

The piecewise linear approximation is made from 1 rather than from a number close to 0 for two reasons. Primarily, it prevents the elimination (by assigning 0 repetitions) of an observation. This ensures that any redundant observations for error checking are preserved. In addition, a poor approximation is obtained for the reciprocal function over the region 0 to 1 using a single variable. A better approximation is produced by using several variables in the region; however, in the field, only an integral number of observations may be

performed. Rather than do more work to compute a solution that is not physically realizable and may hurt error detection, the simple route is chosen by using an interval from 1.

Direction Constraints

One desires to satisfy the condition that all direction observations in a given set at a station be observed with the same number of repetitions. This condition can be met in either of two ways. First, a series of equality constraints can be constructed requiring that the set of direction repetitions be equal. Second, let n_i represent the number of repetitions of the ith direction observation to all sighted stations. This second approach is used for the design algorithm.

The second approach is particularly attractive because it avoids increasing the number of constraints and reduces the number of variables. Instead of ℓ-1 variables per direction, the method uses ℓ-1 variables per direction set; and in a piecewise linear approximation, an additional constraint is applied to each variable. So a reduction in the number of variables also reduces the number of constraints. Using this approach, for the ith station with a_j direction observations in the jth direction set,

$$\sum_{j=1}^{n} \sum_{h=1}^{a_j} \sigma_{jh}^2 \left(g_{x_{jih}}^2 + g_{y_{jih}}^2\right) \sum_{k-1}^{\ell-1} \frac{-n_{jk}}{k(k+1)}$$

$$\leq \sigma_{x_i}^2 + \sigma_{y_i}^2 - \sum_{j=1}^{n} \sum_{n=1}^{a_j} \sigma_{jh}^2 \left(g_{x_{jih}}^2 + g_{y_{jih}}^2\right)$$

and the number of repetitions of the jth direction set is

$$n_j = 1 + \sum_{k=1}^{\ell-1} n_{jk}$$

Moore-Penrose Generalized Inverse

The coordinate variances at a station assume different values when selecting a different constrained point, even in a minimally constrained network. This is due to the fact that coordinates are not estimable quantities. Because of this, an optimal design will vary according to the selection of the constrained point. In Young (1974, p. 357), one sees that the error ellipses steadily increase in size further from the constrained point. Since a simple relation exists between coordinate variances and error ellipse axes, the example also illustrates the increase in coordinate variances. A network design would possess an increasing number of observation repetitions further from the

constrained point. A control network, which minimally satisfies its relative accuracy specifications, should not vary with respect to the constrained point.

To formulate the constraints such that this increase would not occur is desirable. One approach is to use the Moore-Penrose generalized inverse, $(A^t\Sigma^{-1}A)^+$, in place of the Cayley inverse. The generalized inverse generates coordinate variances that are not distance invariant from the centroid of the network (Pope 1971); but the increase is in a much slower fashion than that of the Cayley inverse. This quality, in combination with rotational invariance, provides a quasi-homogeneous measure of a geodetic network. Such a measure has been proposed by Bjerhammar (1973, pp. 293-295).

Cost Function

The cost function, as discussed earlier, can be exceedingly complex. For this design algorithm, the cost is approximated by a linear function of the repetitions of each observation. The function to be minimized is

$$\sum_{i-1}^{n} c_i n_i$$

where c_i is the cost for each repetition of the ith observation.

Using a piecewise linear approximation increases the number of variables. Each of these must be accounted for by a cost coefficient. Since the cost function is linear and the approximation is made over unit intervals, the function to be minimized becomes

$$\sum_{i=1}^{n} c_i \sum_{j-1}^{\ell-1} n_{ij}$$

for an ℓth order piecewise approximation.

Since n_i represents the repetitions of a set of directions at a station, the cost function should model the larger cost of an occupation that sights a larger number of stations. This condition is satisfied by minimizing

$$\sum_{i=1}^{n} \sum_{k=1}^{a_i} c_{ik} \sum_{j=1}^{\ell=1} n_{ij}$$

for a_i directions in the ith direction set. For this design algorithm, the c_{ik}'s are assumed equal for a specific i^{th} direction set. This gives

$$\sum_{i=1}^{n} a_i c_i \sum_{j=1}^{\ell-1} n_{ij}.$$

RESULTS

A program was written in FORTRAN IV (FORmula TRANslator) to solve the second-order network design problem using the constraints and cost function described earlier. The solution proceeds in an iterative manner, where the most recent estimate of the repetitions of each observation is used to form a new variance-covariance matrix for the next cycle. The output displays the repetitions of each observation required to meet or better the design specifications.

The input consists of three sets of cards--position cards, distance observation cards, and direction observation cards. Each position card consists of a station number, the x and y coordinates in meters, and the required variance sum

$$\sigma_x^2 + \sigma_y^2$$

in square meters. Each distance card contains the occupied and sighted station numbers, the variance and the cost of the single observation, and an initial estimate of the optimal repetitions of that observation. Each direction set consists of an occupation card and any number of direction observation cards. The direction occupation card identifies the occupied station number, the cost of observing one repetition of one of the directions (all are assumed equal), and an initial estimate of the optimal repetitions of the set. Each direction observation card contains the number of the sighted station and the variance of that particular observation in square seconds.

The program inputs the data set, constructing the design matrix as each observation is read. The invariant quantities in the Simplex tableau are entered, and the iteration commences. In each iteration, the G matrix is computed using a library pseudoinverse routine, the constraints are entered into the tableau, and the linearized programing problem is solved by a library Simplex routine. This cycle then repeats for a desired number of iterations. The final solution consists merely of the repetitions of each observation, a resultant cost for each observation, and a total cost of the entire observation scheme.

A serendipitous result of using a piecewise linear approximation at integer points is that a significant number of the observations are optimized at integer values. As seen in figure 2, the piecewise approximation has extreme points at the integers; and it is often there that the solution achieves optimal values. This has two advantages. Since the approximation is exact at the integer points, exact observation variances are being reflected in the optimal solution. In addition, since only an integer number of repetitions can actually be observed, a more physically realizable observation scheme results. All this occurs without incorporating the increased complexity of integer programing techniques.

Quadrilateral Deformation

This series of data sets was run to investigate the behavior of the optimal solution subject to the deformation of a triangulation quadrilateral. Throughout these sets, each direction is assumed to have a variance of $9s^2$ (seconds squared) and the same cost (1 unit) no matter where it is observed. For each data set, the same accuracy constraint is used for each station:

$$\sigma_x^2 + \sigma_y^2 \leq 0.0004 \text{ m}^2 \text{ (meter squared).}$$

By using these same values, the variation in the optional solution is due solely to the variation in the geometry of the quadrilateral.

The first test (fig. 3) was a symmetric quadrilateral, 5 km (kilometers) on a side.

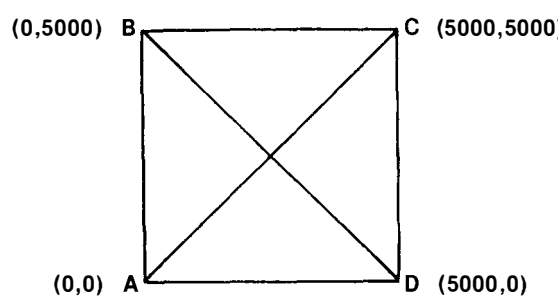

Figure 3.--Symmetric triangulation quadrilateral.

The optimal solution designed by the methods previously described is summarized in table 2.

Table 2.--Symmetric quadrilateral design

Station	Repetitions
A	5.2
B	6.0
C	6.0
D	6.0
Total cost = 69.6	

Notice that no particular reason exists for station A to have less than six observations. This could have just as easily happened at any station. The piecewise linear approximation does tend to keep the solution at integer values as discussed earlier. The total cost is obtained by summing the product of the number of direction observations at a station by the indicated number of repetitions over all the stations.

The next test uses the same data set, except that station C now has the coordinates (4000,5000). The solution recovered for stations A to D was (6.8,5.0,5.0,6.0) for a total cost of 68.4 units. The surprising result of this design is that it has an optimal total cost smaller than the optimal total cost of the symmetric quadrilateral. To investigate this behavior, a number of data sets were run; the results are summarized in table 3.

Table 3.--Deformed quadrilateral design, I

C coordinates	Repetitions	Total cost
	5.2	
	6.0	
(5000,5000):	6.0	
	6.0	69.6
	6.8	
	5.0	
(4000,5000):	5.0	
	6.0	68.4
	7.3	
	4.0	
(3000,5000):	5.0	
	5.3	64.8
	6.0	
	4.0	
(2500,5000):	5.5	
	6.0	64.6
	9.4	
	4.0	
(4000,4000):	5.0	
	3.0	64.3

Table 3.--Deformed quadrilateral design, I (continued)

C coordinates	Repetitions	Total cost
	10.0	
	3.6	
(3000,3000):	4.0	
	3.0	61.6

From the results summarized in table 3, it would appear that the greater the deformation, the cheaper the optimal design. The mechanism by which this occurs is best described by considering the network in figure 4.

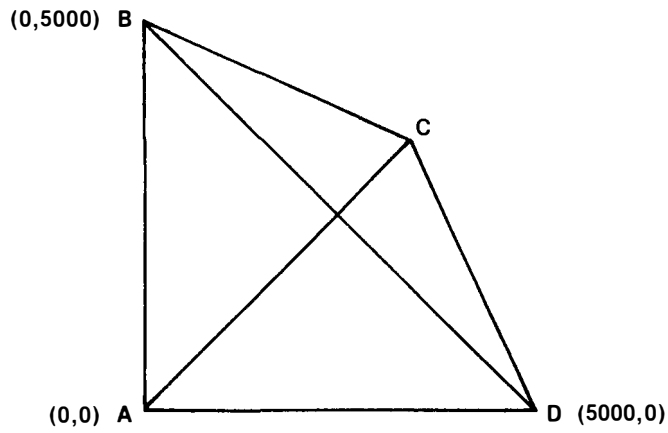

Figure 4.--Deformed triangulation quadrilateral.

As the position of station C comes closer to station A, the angle BCD tends to increase. The direction observations from stations B and D intersect an an increasingly poorer geometry. This makes the observations from A become important. It becomes more effective to make two additional direction sets at A, rather than one additional set at B and at D. This is reflected by a large number of observations at station A in table 3.

Now, since an entire direction set is observed at each station, each repetition of direction AC also means a repetition of AB and AD. The increased number of observations to B and D better determines those positions, causing a reduction of the repetitions needed at B and D. This is also reflected in table 3.

Having identified the mechanism that describes the optimal distribution of observations, one would expect an even lower optimal cost if two positions are brought toward the center of the network. An example of such a network is displayed in figure 5.

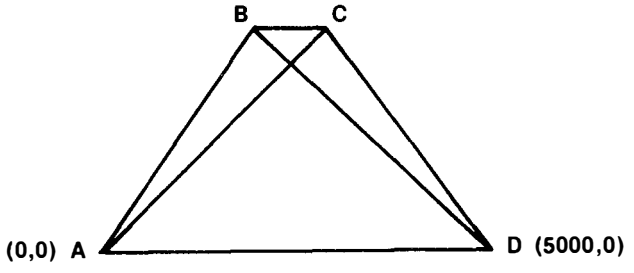

Figure 5.--Symmetric deformed quadrilateral .

The next set of solutions (table 4) was computed with the positions of both B and C altered in a symmetric fashion.

Table 4.--Deformed quadrilateral design, II

B and C coordinates	Repetitions	Total cost
(0,5000) - (5000,5000):	5.2	
	6.0	
	6.0	
	6.0	69.6
(1000,5000) - (4000,5000):	6.0	
	4.0	
	4.0	
	6.0	60.0
(2000,5000) - (3000,5000):	5.5	
	3.0	
	2.6	
	6.0	51.3
(1000,4000) - (4000,4000):	5.0	
	3.4	
	3.4	
	5.0	50.4
(2000,3000) - (3000,3000):	4.5	
	2.0	
	2.0	
	4.5	39.0

Center Point Quadrilateral

The next tests investigate the behavior of the optimal solution of a center point quadrilateral of triangulation. In these tests, the geometry of the quadrilateral remains constant, but the effect of occupation of the central point is examined. The observation variances, costs, and required positional accuracy are identical to those in the preceding topic. The geometry of the first test is displayed in figure 6.

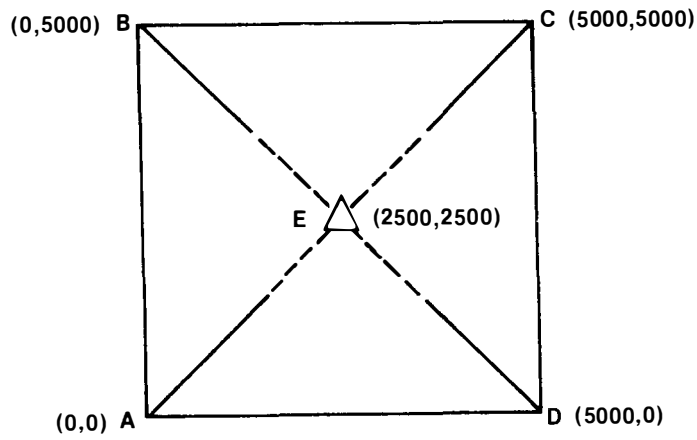

Figure 6.--Center point quadrilateral.

The optimal solution computed for this configuration is (7.0,7.0, 7.0,7.0) with a total cost of 84.0. This should be compared to the total cost of 69.6 for the symmetric quadrilateral with no center point.

Station E is occupied in the next test, with four direction observations originating to the other stations. The location of the stations remains the same, however. In this test the optimal solution is (5.0,5.0,5.0,5.0,3.9) for a total of 75.6. From this one sees that, if it would cost less than 84.0 - 75.6 = 8.4 to occupy station E with a theodolite, net savings result by performing such an occupation.

Circular Traverse

In this final series of tests, the optimal solution of a circular traverse with a 20-km diameter is investigated. Both distance and direction observations are made for three different

design schemes, which will be discussed shortly. The geometry of the first test is displayed in figure 7; and the station coordinates in meters are summarized in table 5.

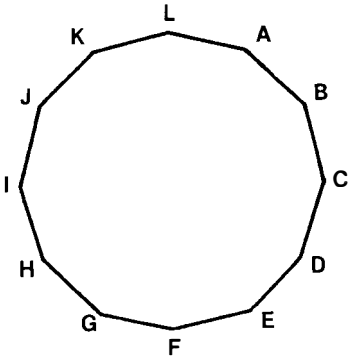

Figure 7.--Circular traverse .

Table 5.--Circular traverse coordinates

Station	x (m)	y (m)
A	15,000	18,500
B	18,500	15,000
C	20,000	10,000
D	18,500	5,000
E	15,000	1,500
F	10,000	0
G	5,000	1,500
H	1,500	5,000
I	0	10,000
J	1,500	15,000
K	5,000	18,500
L	10,000	20,000

For all 12 stations, the required accuracy is

$$\sigma_x^2 + \sigma_y^2 \leq 0.0009m^2 .$$

The variance of a direction is assumed to be $1.4s^2$, (seconds squares) and for a distance is $0.001175m^2$. Finally, assume that the relative costs of a distance and a direction observation are equivalent. Even if this is not the case, the program allows any desired costs to be used for a network design. By using this combination of observation variances, costs, and desired accuracies, the optimal solution computed by the algorithm requires each distance to be measured 3.0 times and each direction set be measured 1.9 times for a total cost of 81.6.

The next test investigates the presence of a direction target in the center of the traverse (10000,10000). No observations of any type are made from this target. Since the target is being observed to strengthen the remainder of the network, the target position is not accurately required, and an extremely large variance sum is assigned to it:

$$\sigma_x^2 + \sigma_y^2 \leq 1.0m^2$$

The optimal solution for this data set states that the distances should be measured from 2.0 to 2.7 times and each direction set should be measured for a total cost of 64.0. This is a dramatic reduction, showing the beneficial effect of observing directions to a central point.

Note the possibility that the solution for the directions, which specifies each direction set to be observed 1.0 times, could be smaller. The constraint of the piecewise linear approximation over the region 1 to ℓ prevents this from happening. If stricter requirements were made on the coordinate variance sum, an even larger reduction in the total cost might have been observed.

The final test investigates the effect upon the optimal solution if both distance and direction observations are made to the central point. In this test, the variance of the distance observations made to the central point is assumed to be $0.0018m^2$. This corresponds to a variance of (3 cm plus 3 ppm)^2 for the distance. The optimal solution indicates that the rim distances be measured 1.0 to 2.0 times, that the spoke distances be measured once, and that the directions be measured once for a total cost of 67.4. This optimal solution is slightly higher than that of the design where distances to the central target were not measured. Remember: no observation is allowed to be

measured less than once; the cost of a direction is assumed to be equivalent to the cost of a distance; and the distances and directions are assumed to be uncorrelated.

IMPROVEMENTS

This final section examines a range of possible improvements that could be made to the design algorithm embodied in the program. As discussed earlier, the program was implemented using a number of approximations. By developing more accurate representations of observation variances and costs, a better geodetic network design tool will result.

Computer Storage Allocation

Although this problem was not mentioned earlier, the design program uses extremely large amounts of computer storage with consequent increases in computer time requirements. These increases are of such magnitude that this problem should be tackled before any operational use could be made of the program.

First, however, one must detail the problem. For all tests, the program used double precision arithmetic on an International Business Machine, IBM 370, with virtual storage. The quadrilateral of triangulation allowing up to 11 repetitions ran in the 192k partition. However, the circular traverse with all the distance and direction observations, allowing up to four repetitions, ran in the 512k partition. This is a large amount of computer storage for such a modest size network.

Tableau Sparsity

The reason behind the dramatic increase of computer storage requirements is primarily due to the piecewise linear approximation. For such an approximation over integer values from 1 to $\ell + 1$, there is an ℓ-fold increase in the number of variables. Further, the piecewise linear approximation produces an additional ℓ-fold increase in slack variables and extra $n \times \ell$ constraints. (See eq 34.) This results in a $2(n \times \ell)^2$ increase in storage for those constraints alone. One approach to the reduction is to consider revised forms of the Simplex algorithm.

The Simplex linear programing subroutine called in the design program already uses a revised Simplex method called the product form of the inverse (Hadley 1961). In this form, the section of the tableau that holds the simple variables is invariant while the section which holds the slack and artificial variables is updated during the change of basis. Now, the invariant section happens to be an augmented identity matrix for the piecewise linear approximation over unit values. Because of this, the identify matrix portion need not be stored so long

as the Simplex subroutine is specially modified to reproduce the effect of the identity portion of the matrix. This approach would reduce storage requirements by almost half.

Variable Piecewise Linear Approximation

A second approach provides for an even more dramatic decrease in the computer storage requirements. Consider the piecewise linear approximation. There is no computational reason for it to be made over integer values or even over equally spaced break points. Moreover, if the approximation is viewed from an iterative standpoint, then it is unnecessary to use the same break points from iteration to iteration. One can use a crude approximation over a large region and select subsequent break points that will always bracket the solution. When done iteratively, one can solve the nonlinear programing problem to an arbitrary accuracy using only a small number of break points.

As an example, suppose a nonlinear function may vary over 1 to 51. Rather than create 50 variables and 50 new constraint equations, only 5 are created with break points at (1, 11, 21, 31, 41,51). Suppose the solution for that approximation is 14. Then, the next set of break points could become (9, 11, 13, 15, 17, 19). The next break point could bring the break points into integer spacing. Using this approach, of course, does not mean abandonment of the advantages of break points with integer spacing. Using this iterative approximation, one can replace each original variable with only two approximation variables. This would lead to a tableau storage requirement of only $8n^2$ for the approximation constraints.

In addition to the greatly reduced storage requirements, three other reasons point to the variable piecewise linear approximation. First, the resultant linear programing problem will be much smaller, so that a solution may be computed much quicker allowing the iterations to proceed at a rapid pace. Second, since the problem is reformulated at each iteration, roundoff error becomes much smaller. Finally, since the approximation to the inverse is iterative (eq 30), the algorithm must repeat anyway, so a smaller linear programing problem might as well be solved instead of a large one.

One slight drawback to the iterative approach is that, if the break points are made to "close in" on the approximate solution too quickly, additional iterations may have to be performed until the break points again properly straddle the solution. One must develop some rules for the optimal rate of contraction of the break points.

Distance Observation Correlations

One of the major assumptions made in the design algorithm was that the distance observations are uncorrelated. This is, in fact, not the case for modern electromagnetic distance-measurement instrumentation (Malla 1978). Because of this correlation, the assumed uncorrelation makes distances appear more effective than they are in reality.

Two approaches can be used to introduce the effect of correlation. The first would be to develop some function of n that adequately models the correlation. A least-squares fit of the function could be made for data that represent the variance of the mean over the number of observations. The least-squares estimate of the variance of the mean could then be used in the design algorithm.

A second approach, however, is much simpler and uses no approximations. The piecewise linear approximation requires the variance of the mean at each break point and nowhere else. Therefore, if the break points are always at integer values, then the mean of the variance is only required at these integer values. These data can be stored in a table. There is no need to compute some approximate least-squares function. The second approach seems most effective for incorporating the correlations of distance observations.

Relative Accuracy Constraints

As discussed earlier, the specifications for horizontal control networks are written in terms of length relative accuracy requirements. It is natural, therefore, to wish to determine the length relative accuracies associated with the optimal solutions computed using the quasi-coordinate invariant constraints.

The length relative accuracy is expressed as 1:r where

$$\frac{1}{r} = \frac{\sigma_{d_{ij}}}{d_{ij}}$$

and

$$d_{ij}^2 = (x_j - x_i)^2 + (y_j - y_i)^2 .$$

Taking the optimal solution for the symmetric triangulation quadrilateral (5.2,6.0,6.0,6.0) and holding points A and D fixed, one sees that between B and C

$$\sigma_{d_{BC}} = 3.7 \text{ cm}$$

$$d_{BC} = 5000 \text{ m}$$

and

$$\frac{1}{r} = 1:135,097.$$

For the deformed quadrilateral with C at (3000,3000), between B and C one computes

$$\sigma_{d_{BC}} = 4.3 \text{ cm}$$

$$d_{BC} = 3605.6 \text{ m}$$

and

$$\frac{1}{r} = 1:83,546;$$

and for the symmetric deformed quadrilateral with B at (2000, 3000) and C at (3000,3000), between B and C one computes

$$\sigma_{d_{BC}} = 3.2 \text{ cm}$$

$$d_{BC} = 1000 \text{ m}$$

and

$$\frac{1}{r} = 1:31,220.$$

Clearly, the relative accuracy between B and C degrades as the quadrilateral is deformed. Primarily, this is due to the change in the distance between B and C. The standard error of the length between B and C exhibits no special pattern. One cannot be surprised at the degradation since the original accuracy constraints were not in terms of length relative accuracies.

It appears that the only way to ensure an optimal design satisfying a particular design requirement is to write the specific requirement into the constraints. Further, before one can develop a network design algorithm, one must decide which accuracy measures are applicable to the problem at hand. For horizontal control networks, one would write the accuracy constraints in terms of length relative accuracies. By linear error propagation, the constraints have the form

$$\Delta x_{ij}^2 \left(\sigma_{x_i}^2 - 2\sigma_{x_i x_j} + \sigma_{x_j}^2\right) + 2\Delta x_{ij}\Delta y_{ij}\left(\sigma_{x_i y_i} - \sigma_{x_i y_j} - \sigma_{y_i x_j} + \sigma_{x_j y_j}\right) + \Delta y_{ij}^2 \left(\sigma_{y_i}^2 - 2\sigma_{y_i y_j} + \sigma_{y_j}^2\right) \leq d_{ij}^4/r^2$$

for all ij over every observed line.

By being able to design networks directly satisfying the specifications, one can scrutinize these designs with other tests, generating ideas for specification revisions. The cycle would repeat itself, resulting in a stronger set of specifications.

Station Occupation Costs

The next level of sophistication in modeling the field work cost is estimating the cost of mark recovery, of tower construction and dismantlement, and of instrument setup and leveling. These costs are incurred at each occupied station regardless of the number of observations made at that station. Such costs seem to be substantial when compared to the costs of making the observations. For this reason, the modeling of station occupation costs is an important problem.

Mixed Integer Methods

To model the cost of station occupation, one can use a fixed charge cost function (Gaver and Thompson 1973, pp. 230-231). The form of this cost function is

$$\min z = \sum_{i=1}^{n} (c_i y_i) + \sum_{j=1}^{k} b_j \delta_j$$

where $\delta_j \leq 1$ and δ_j must be integer for all j. Since δ_j is restricted to integer values, δ_j may only be equal to 0 or 1. A station may only be unoccupied or occupied. Thus, b_j refers to the cost of occupation of the jth station. However, y_i is not restricted to integer values, so the problem becomes a mixed integer problem. One may desire to also restrict y_i to integer values since only an integer number of observations can be made. Then the problem becomes an integer programing problem.

Keep in mind that other constraints will be necessary to enforce the fact that, if no observations from station j are made, all observations y_i originating from that station must be 0. These constraints are not formulated in this report.

Preservation of Network Redundancy

When modeling station occupation costs by a fixed charge cost function, one must allow each observation to achieve a value of zero (i.e., allow each observation not to be made). Unfortunately, any geodetic network designed in this way is likely to have a low degree of redundancy, a most undesirable feature.

One solution that immediately springs to mind is to develop some constraint which would preserve the redundancy in a network. This might require the number of different observations to a

given station meet or exceed a specified amount. Although the network would possess a margin of safety against observation blunders, the set of observations without the blunders may not satisfy the accuracy constraints. It would be self defeating to design a network that would meet accuracy requirements against the worst conceivable set of blunders, but this is edging into an entire problem by itself.

Other Approaches

Perhaps the most efficient approach to network design is to examine a subset of all possible combinations of observations by computing some "value" of each observation. In this strategy, the network would be "built" with the effect of each new observation being added to the foundation of the previous observations. This approach is listed last because its efficiency is bought at the price of examining a subset of all designs, potentially overlooking a more optimal design. When such efficient algorithms are devised, they should be tested against some thorough (if decidedly slower) exhaustive algorithm.

ACKNOWLEDGMENTS

The author acknowledges the assistance of his adviser, Urho A. Uotila, whose insight and perspective were important assets in this work. The author is also indebted to the National Oceanic and Atmospheric Administration for sponsoring his study at the Ohio State University (OSU) and to OSU's Instruction and Research Computer Center for providing funds for computer time.

REFERENCES

Bjerhammar, Arne, 1973: Theory of Errors and Generalized Matrix Inverses. Elsevier Scientific Publishing Co., New York, 420 pp.

Bossler, John D., Grafarend, Erik W., and Kelm, Rainer, 1973: Optimal design of geodetic nets, 2, Journal of Geophysical Research, 78 (26), 5887-5897.

Dracup, Joseph F. and Fronczek, Charles J., 1979: An immediate solution to the optimization problem. Proceedings of the International Association of Geodesy Symposium on Optimization of Design and Computation of Control Networks, Hungarian Academy of Sciences, Geodetic and Geophysical Institute, Sopron, Hungary, July 4-10, 1977, Akademiai Kiado, Budapest (in press).

Federal Geodetic Control Committee, 1974: Classification, Standards of Accuracy, and General Specifications of Geodetic Control Surveys. National Ocean Survey, National Oceanic and Atmospheric Administration, U.S. Dept. Commerce, Rockville, Md., 12 pp.

Gaver, D. P. and Thompson, G. L., 1973: Programming and Probability Models in Operations Research. Brooks/Cole Publishing Co., Monterey, Calif.

Gergen, John G. (National Geodetic Survey, National Ocean Survey, National Oceanic and Atmospheric Administration, Rockville, Md.), 1978 (personal communication).

Gillett, Billy E., 1976: Introduction to Operations Research: A Computer Oriented Algorithmic Approach. McGraw-Hill Book Co., New York, 617 pp.

Grafarend, Erik W., 1974: Optimization of geodetic networks, Canadian Surveyor, 28 (5), 716-723.

Greve, Clifford W., 1972: Design of survey systems using non-linear programming methods, Bulletin Géodésique, 103, Mar. 1, 31-45.

Hadley, G., 1961: Linear Programming. Addison-Wesley Publishing Co., Reading, Pa., 520 pp.

Hoyle, Dixon (National Geodetic Survey, National Ocean Survey, National Oceanic and Atmospheric Administration, Rockville, Md.), 1977 (personal communicaton).

Malla, Rajendra Prakash (Ohio State University, Columbus) 1978: Effects of correlated distance observations on the strength of a horizontal geodetic network. M.S. Thesis, 141 pp.

McKay, Edward J. (Ohio State University, Columbus) 1973: Statistical analysis of angles and directions. M.S. Thesis, 123 pp.

Pope, Allen J. (National Geodetic Survey, National Ocean Survey, National Oceanic and Atmospheric Administration, Rockville, Md.) 1971: Transformation of covariance matrices due to changes in minimal control, presented to the American Geophysical Union annual fall meeting, San Francisco, Dec. 9.

Schwarz, Charles, 1974-75: Adjustment computations (lecture notes), Department of Geography and Regional Science, George Washington University, Washington, D.C.

Sprinsky, William H. (Ohio State University, Columbus) 1974: The design of special purpose horizontal geodetic control networks. Ph.D. Dissertation.

Uotila, Urho A., 1967: Introduction ot adjustment computations with matrices (lecture notes), Department of Geodetic Science, Ohio State University, Columbus, Ohio.

Young, Gary M., 1974: A mixed observational survey method, Bulletin Géodésique, 114, Dec. 1, 349-363.

(Continued from inside front cover)

NOS NGS-6 Determination of North American Datum 1983 coordinates of map corners. T. Vincenty, October 1976, 8 pp (PB262442). Predictions of changes in coordinates of map corners are detailed.

NOS NGS-7 Recent elevation change in Southern California. S.R. Holdahl, February 1977, 19 pp (PB265-940). Velocities of elevation change were determined from Southern Calif. leveling data for 1906-62 and 1959-76 epochs.

NOS NGS-8 Establishment of calibration base lines. Joseph F. Dracup, Charles J. Fronczek, and Raymond W. Tomlinson, August 1977, 22 pp (PB277130). Specifications are given for establishing calibration base lines.

NOS NGS-9 National Geodetic Survey publications on surveying and geodesy 1976. September 1977, 17 pp (PB275181). Compilation lists publications authored by NGS staff in 1976, source availability for out-of-print Coast and Geodetic Survey publications, and subscription information on the Geodetic Control Data Automatic Mailing List.

NOS NGS-10 Use of calibration base lines. Charles J. Fronczek, December 1977, 38 pp (PB279574). Detailed explanation allows the user to evaluate electromagnetic distance measuring instruments.

NOS NGS-11 Applicability of array algebra. Richard A. Snay, February 1978, 22 pp (PB281196). Conditions required for the transformation from matrix equations into computationally more efficient array equations are considered.

NOS NGS-12 The TRAV-10 horizontal network adjustment program. Charles R. Schwarz, April 1978, 52 pp (PB283087). The design, objectives, and specifications of the horizontal control adjustment program are presented.

NOS NGS-13 Application of three-dimensional geodesy to adjustments of horizontal networks. T. Vincenty and B. R. Bowring, June 1978, 7 pp (PB286672). A method is given for adjusting measurements in three-dimensional space without reducing them to any computational surface.

NOS NGS-14 Solvability analysis of geodetic networks using logical geometry. Richard A. Snay, October 1978, 29 pp (PB291286). No algorithm based solely on logical geometry has been found that can unerringly distinguish between solvable and unsolvable horizontal networks. For leveling networks such an algorithm is well known.

NOS NGS-15 Goldstone validation survey - phase 1. William E. Carter and James E. Pettey, November 1978, 44 pp (PB292310). Results are given for a space system validation study conducted at the Goldstone, Calif., Deep Space Communication Complex.

NOS NGS-16 Determination of North American Datum 1983 coordinates of map corners (Second Prediction). T. Vincenty, April 1979, 6 pp. New predictions of changes in coordinates of map corners are given.

NOS NGS-17 The HAVAGO three-dimensional adjustment program. T. Vincenty, May 1979, 18 pp. The HAVAGO computer program adjusts numerous kinds of geodetic observations for high precision special surveys and ordinary surveys.

NOS NGS-18 Determination of astronomic positions for California-Nevada boundary monuments near Lake Tahoe. James E. Pettey, March, 1979. Astronomic observations of the 120th meridian were made at the request of the Calif. State Lands Commission.

NOS NGS-19 HOACOS: A program for adjusting horizontal networks in three dimensions. T. Vincenty, July 1979, 13 pp. Horizontal networks are adjusted simply and efficiently in the height-controlled spatial system without reducing observations to the ellipsoid.

(Continued on inside back cover)

(Continued)

NOAA Technical Reports, NOS/NGS subseries

NOS 65 NGS 1 The statistics of residuals and the detection of outliers. Allen J. Pope, May 1976, 133 pp (PB258428). A criterion for rejection of bad geodetic data is derived on the basis of residuals from a simultaneous least-squares adjustment. Subroutine TAURE is included.

NOS 66 NGS 2 Effect of Geoceiver observations upon the classical triangulation network. R. E. Moose and S. W. Henriksen, June 1976, 65 pp (PB260921). The use of Geoceiver observations is investigated as a means of improving triangulation network adjustment results.

NOS 67 NGS 3 Algorithms for computing the geopotential using a simple-layer density model. Foster Morrison, March 1977, 41 pp (PB266967). Several algorithms are developed for computing with high accuracy the gravitational attraction of a simple-density layer at arbitrary altitudes. Computer program is included.

NOS 68 NGS 4 Test results of first-order class III leveling. Charles T. Whalen and Emery Balazs, November 1976, 30 pp (GPO# 003-017-00393-1) (PB265421). Specifications for releveling the National vertical control net were tested and the results published.

NOS 70 NGS 5 Selenocentric geodetic reference system. Frederick J. Doyle, Atef A. Elassal, and James R. Lucas, February 1977, 53 pp (PB266046). Reference system was established by simultaneous adjustment of 1,233 metric-camera photographs of the lunar surface from which 2,662 terrain points were positioned.

NOS 71 NGS 6 Application of digital filtering to satellite geodesy. C. C. Goad, May 1977, 73 pp (PB-270192). Variations in the orbit of GEOS-3 were analyzed for M_2 tidal harmonic coefficient values which perturb the orbits of artificial satellites and the Moon.

NOS 72 NGS 7 Systems for the determination of polar motion. Soren W. Henriksen, May 1977, 55 pp (PB274698). Methods for determining polar motion are described and their advantages and disadvantages compared.

NOS 73 NGS 8 Control leveling. Charles T. Whalen, May 1978, 23 pp (GPO# 003-017-00422-8) (PB286838). The history of the National network of geodetic control, from its origin in 1878, is presented in addition to the latest observational and computational procedures.

NOS 74 NGS 9 Survey of the McDonald Observatory radial line scheme by relative lateration techniques. William E. Carter and T. Vincenty, June 1978, 33 pp (PB287427). Results of experimental application of the "ratio method" of electromagnetic distance measurements are given for high resolution crustal deformation studies in the vicinity of the McDonald Lunar Laser Ranging and Harvard Radio Astronomy Stations.

NOS 75 NGS 10 An algorithm to compute the eigenvectors of a symmetric matrix. E. Schmid, August 1978, 5 pp (PB287923). Method describes computations for eigenvalues and eigenvectors of a symmetric matrix.

NOS 76 NGS 11 The application of multiquadric equations and point mass anomaly models to crustal movement studies. Rolland L. Hardy, November 1978, 63 pp (PB293544). Multiquadric equations, both harmonic and non-harmonic, are suitable as geometric prediction functions for surface deformation and have potentiality for usage in analysis of subsurface mass redistribution associated with crustal movements.

NOAA Manuals, NOS/NGS subseries

NOS NGS 1 Geodetic bench marks. Lt. Richard P. Floyd, September 1978, 56 pp (GPO# 003-017-00442-2) (PB296427). Reference guide provides specifications for highly stable bench marks, including chapters on installation procedures, vertical instability, and site selection considerations.

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Geodetic Survey, C13x4
Rockville, Maryland 20852

OFFICIAL BUSINESS

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF COMMERCE
COM-210
THIRD CLASS MAIL

U.S.MAIL

NOAA--S/T 79-114